

***ELABORACIÓN DE UN MODELO PARA EL VAB SUBSECTORIAL DE LA ECONOMÍA
CATALANA A PARTIR DE LA TEORÍA DE LA COINTEGRACIÓN.***

Miquel CLAR
Departament d'Econometria, Estadística i Economia Espanyola
Universitat de Barcelona

De un tiempo a esta parte en España el interés por el estudio de temas relacionados con las economías regionales ha ido en aumento. En este sentido, en el seno del Departament d'Econometria, Estadística i Economia Espanyola de la Universitat de Barcelona se viene siguiendo en los últimos años una línea de investigación en economía regional. Así, el modelo econométrico para el bloque de Producción que presentamos en este trabajo cabe enmarcarlo en esta línea de investigación. En concreto, modelizamos basándonos en la teoría econométrica de la cointegración las variables de VAB de nuestra economía a un nivel de desagregación sectorial de ocho ramas: Agricultura, Energía, Bienes Intermedios, Bienes de Equipo, Construcción, Transportes y Comunicaciones y Resto de Servicios destinados a la Venta más Servicios no destinados a la Venta. Así, en primer lugar estudiamos mediante los tests de Sargan y Bhargava (1983), Dickey y Fuller (1979) y Dickey y Fuller ampliado (1981) el orden de integrabilidad de las variables del modelo realizando, para aquellas variables que resulten $I(2)$, un análisis de tendencias segmentadas en la media. A continuación estudiamos la existencia de relaciones de cointegración entre ellas. Por último, presentamos el modelo de VAB subsectorial catalán, en el que las ecuaciones de comportamiento están especificadas, cuando ha sido posible, en términos de *MCE* para así contemplar la dinámica a corto y a largo plazo y realizamos un análisis de la capacidad predictiva del modelo.

Palabras clave: Modelización macroeconométrica regional, tendencias comunes, modelos satélite.
e-mail: mclar@riscd2.eco.ub.es

1. INFORMACIÓN ESTADÍSTICA.

En el estudio realizado se ha puesto de manifiesto la falta de información estadística a nivel regional que hoy en día vivimos en nuestro país tanto en el reducido número de observaciones, como de variables regionales disponibles. Además, el problema queda agravado por el nivel de desagregación sectorial considerado: cuanto mayor es éste mayor es la dificultad para modelizar los subsectores económicos considerados ya que su evolución dependerá en mayor medida de la propia dinámica del sector en la región y, en consecuencia, la falta de información estadística se hace más evidente. A pesar de ello preferimos asumir los inconvenientes de trabajar con un mayor grado de detalle subsectorial dado que así disponemos de un mayor conocimiento de la realidad económica de nuestra Comunidad.

Todo ello hace que las especificaciones que aquí presentamos no deben considerarse definitivas sino que están sujetas a revisión a medida que dispongamos de información adicional dado que las pocas observaciones disponibles hacen que las estimaciones obtenidas sean poco estables. Además, la utilidad de cualquier modelo depende directamente del proceso de actualización al que sea sometido, tanto en lo que hace referencia a las especificaciones de las ecuaciones como a la consideración de nuevas técnicas econométricas que se puedan aplicar en el ámbito regional.

Las principales fuentes de información que hemos utilizado han sido para las variables regionales la *CRE* del INE y la base de datos *Hispadat* mientras que para las nacionales ha sido la *CNE* del INE. Por consiguiente, las series con las que trabajamos son de periodicidad anual. Así, para las variables regionales el período muestral del que disponemos abarca desde el año 1970, 1971 o 1972, según la variable en cuestión, hasta el 1992. Esto es debido al retraso temporal de la *CRE* que no permite disponer de datos oficiales para los años 1993-95; por tanto, para estos años tendríamos que trabajar con estimaciones. En consecuencia, a pesar que dispondríamos de tres grados de libertad más, teniendo en cuenta que las conclusiones a las que lleguemos no deberían verse alteradas por el hecho de no considerar estos años dado que, en último término, estudiamos la existencia de relaciones a largo plazo entre las variables, no los consideramos en el análisis de integrabilidad de las variables regionales para así no introducir ningún sesgo en la información muestral.

2. ANÁLISIS DEL ORDEN DE INTEGRABILIDAD DE LAS VARIABLES.

Como señalan Dickey y Pantula (1987) la estrategia más adecuada para determinar el orden de integrabilidad de una serie temporal consiste en partir del orden más elevado posible que pueda presentar (d) y en caso de rechazar la hipótesis nula contrastar si la serie es integrable de orden inmediatamente inferior ($d-1$). En nuestro estudio teniendo en cuenta que, como demuestran trabajos empíricos anteriores¹, las variables serie-temporales de carácter macroeconómico presentan como máximo dos raíces unitarias, en primer lugar contrastamos la hipótesis $d=2$. Por tanto, la regresión auxiliar del test *DFA* es:

$$\Delta^d x_t = g \Delta^{d-1} x_{t-1} + \sum_{j=1}^k d_j \Delta^d x_{t-j} + m + bt + e_t \quad d = 1, 2$$

¹ Por ejemplo en Nelson y Plosser (1982), en Molinas *et al.* (1990), en Andrés *et al.* (1990a y 1990b) así como en diferentes modelos econométricos regionales españoles actualmente en vigor.

En el cuadro 1 recogemos los principales resultados obtenidos al analizar el orden de integrabilidad de las variables del modelo². Analizando dichos resultados podemos observar que el estadístico DW toma un valor cercano a cero en todos los casos. Dados los valores críticos tabulados por Sargan y Bhargava (1983) para ninguna de las variables podemos rechazar la hipótesis nula, lo cual significa que como mínimo son $I(1)$. Un comentario aparte merece el VAB del sector agrícola. Para dicha variable hemos obtenido un valor del estadístico DW que cae en la zona de rechazo de la hipótesis nula lo cual nos lleva a considerar que no presenta tendencia estocástica.

Por lo que se refiere a los resultados obtenidos con los tests DF y DFA en el caso de la hipótesis nula $d=2$ para las únicas variables que no es posible rechazarla son $\ln VLGCT86_t$, $\ln VTOTCT86_t$ y $\ln PPRIEN_t$. Sin embargo, los gráficos de dichas series muestran que existe un cambio estructural en su evolución temporal. Además, si observamos el valor estimado para g vemos que no es concluyente: si las variables realmente fuesen $I(2)$ cabría esperar un valor para dicho parámetro muy próximo a cero lo cual no sucede en este caso. Lo anterior, es suficiente como para que no podamos establecer una conclusión definitiva respecto al orden de integrabilidad de estas variables hasta que no estudiemos la posible existencia de tendencias deterministas segmentadas en la media del PGD puesto que, tal y como demostraron Rappoport y Reichlin (1989) y Perron (1990), la presencia de tal tipo de tendencias (si no son tratadas adecuadamente) puede generar una raíz unitaria espúrea.

En cuanto al contraste de la hipótesis nula $d=1$ vemos que para ninguna variable el estadístico t_g obtenido permite rechazarla. Sin embargo hay dos excepciones: el VAB agrícola catalán y el del subsector de Bienes Intermedios de la economía española ($\ln VABI86_t$) para los que el resultado obtenido pone de manifiesto que no presentan tendencias estocásticas.

3. ANÁLISIS DE TENDENCIAS SEGMENTADAS EN LA MEDIA.

En el cuadro 2 presentamos la evolución de la desviación estándar de las cuatro series de interés ante sucesivas diferenciaciones y en el 3 los resultados obtenidos al aplicar el contraste de Rappoport y Reichlin (1989). Dichos resultados nos llevan a la conclusión que son $I(1)$ con tendencias deterministas segmentadas (menos $\ln VABI86_t$ que simplemente es $I(1)$). Por lo tanto, en base al análisis realizado podemos afirmar que las variables del modelo son $I(1)$ excepto el VAB agrícola catalán que es $I(0)$.

4. RELACIONES ESTABLES A LARGO PLAZO.

El concepto de cointegración (Engle y Granger, 1987) se basa en la idea que si en el largo plazo dos (o más) variables presentan un mismo comportamiento, incluso aunque presenten tendencias estocásticas y en consecuencia sean no estacionarias, puede existir una (o más) combinación lineal entre ellas en la que se cancelen dichas tendencias siendo

² En el anexo I presentamos la definición de las variables utilizadas.

por tanto estacionaria. En definitiva se trata de buscar relaciones estacionarias entre variables que no lo son.

$(1-L)^2 x_t = g(1-L)x_{t-1} + a(1) \left[b_1 + \sum_{i=2}^n (b_i - b_{i-1}) E_{i-1,t-2} \right] +$ $+ a^*(L) \left[\sum_{i=2}^n (b_i - b_{i-1})(1-L) E_{i-1,t-1} \right] + a^{**}(L)(1-L)^2 x_{t-1} + e_t$							
d=2							
<i>Variable</i>	<i>Modelo</i>	\hat{g}	t_g^2	DW^1	$Q(4)^1$	<i>A.C.T.</i>	<i>T</i>
<i>lnVLGCT86</i>	$a^{**}(L)=0$	-1.06	-2.63	2.21	6.71	80, 87	18
<i>lnVTOTCT86</i>	$a^{**}(L)=0$	-0.61	-3.92	1.77	3.29	80, 86	19
<i>lnVABI86</i>	$a^{**}(L)=0$	-1.05	-5.09	1.95	1.68	74	24
<i>lnPPRIEN</i>	$a^{**}(L)=a_1$	-0.99	-4.01	2.26	2.30	85	24
$(1-L)x_t = a(1) \left[(c_1 - b_1) + b_1 t + \sum_{i=2}^n (b_i - b_{i-1}) E_{i-1,t-1} t - \sum_{i=2}^n (b_i - b_{i-1}) E_{i-1,t-1} (1 + t_{i-1}^*) \right] +$ $+ \gamma x_{t-1} + a^*(L) \left[b_1 + \sum_{i=2}^n (b_i - b_{i-1})(1-L) E_{i-1,t-1} \right] + a^{**}(L)(1-L)^2 x_{t-1} + \varepsilon_t$							
d=1							
<i>Variable</i>	<i>Modelo</i>	\hat{g}	t_g^3	DW^1	$Q(4)^1$	<i>A.C.T.</i>	<i>T</i>
<i>lnVLGCT86</i>	$a^{**}(L)=a_1+a_2L$	-0.22	-0.06	2.34	3.04	80, 87	16
<i>lnVTOTCT86</i>	$a^{**}(L)=a_1+a_2L$	-0.41	-4.38	2.03	5.38	80, 86	17
<i>lnVABI86</i>	$a^{**}(L)=a_1$	-0.62	-3.62	1.98	0.66	74	23
<i>lnPPRIEN</i>	$a^{**}(L)=0$	-0.51	-3.38	1.63	8.71	85	24

Cuadro 3. Resultados del contraste tendencias segmentadas en la media. Test de Rappoport y Reichlin (1989).

¹ Estadísticos obtenidos a partir de los residuos estimados de la regresión del test.

² No existen valores críticos tabulados con los que comparar el estadístico t_g para $d=2$.

³ El valor crítico tabulado por Rappoport i Reichlin (1989) es -4.76 para $d=1$.

En cualquier caso, los valores críticos tabulados por Rappoport y Reichlin adolecen de ciertos inconvenientes que en nuestro caso se ponen de manifiesto. En concreto, el tamaño muestral con el que trabajamos es inferior al considerado por los autores (los valores críticos los obtuvieron a partir de simulaciones de Monte Carlo con series de 100 observaciones) y el punto de cambio estructural no está situado sobre el primer cuartil ni sobre la mediana que fueron los supuestos considerados por los autores. En consecuencia debemos tomar dicho valor crítico como orientativo y las conclusiones que se deriven con prudencia.

Existen diversos métodos para estimar relaciones de cointegración. Así, de entre los propuestos en la literatura podemos destacar los siguientes: estimación bietápica por MCO (Engle y Granger, 1987), estimación máximo verosímil (Johansen, 1988) y el método de Kremers, *et al.* (1992).

En la búsqueda de estas relaciones estacionarias entre las variables seguimos una estrategia de lo particular a lo general en el sentido que partimos de la relación más sencilla posible (aquella en la que consideramos un menor número de variables) y, mientras no rechazamos la hipótesis nula de no cointegración, incorporamos variables hasta encontrar una relación estable³. Como hemos visto anteriormente las variables del modelo son $I(1)$; en consecuencia,

³ Esta estrategia es seguida en numerosos trabajos. Véanse por ejemplo los trabajos de Andrés *et al.* (1990a y 1990b).

cumplen la condición necesaria para estar cointegradas (la única excepción es el VAB agrícola catalán). Las relaciones de cointegración que hemos encontrado las presentamos en el cuadro 4.

A la hora de interpretar los resultados hay que tener en cuenta que la potencia de los tests de integrabilidad se reduce de manera significativa para muestras pequeñas, por lo que los valores críticos pueden estar ligeramente sesgados al alza (en valor absoluto) con lo cual la probabilidad de no rechazar la hipótesis nula de no cointegración aumenta.

5. PRESENTACIÓN Y RESULTADOS DE LA ESTIMACIÓN DEL MODELO.

La variable endógena que hemos modelizado es el VAB a precios de mercado y a pesetas constantes de 1986. El nivel de desagregación sectorial con el que hemos trabajado es de ocho ramas de actividad⁴. En consecuencia, el modelo que presentamos constituye por sí mismo un modelo de ecuaciones simultáneas formado por ocho ecuaciones de comportamiento (cada una de ellas corresponde a cada uno de los ocho subsectores de actividad analizados) y tres identidades contables.

El objetivo perseguido con el modelo es predecir la evolución a corto plazo del VAB subsectorial catalán⁵. Para ello hemos optado, condicionados en parte por la información estadística disponible, por un enfoque *Top Down*.

Como podemos ver en el cuadro 5, las ecuaciones de VAB subsectoriales están especificadas en términos de *MCE* excepto para el sector primario, que es lineal dado que la variable endógena ($\ln VACT86_i$) es $I(0)$.

Para estimar el modelo hemos optado por aplicar métodos de estimación directa (uniecuacional). Así, estimamos las ecuaciones de comportamiento por *MCNL* excepto la ecuación correspondiente al sector agrícola que al ser lineal hemos aplicado *MCO*. La elección de los métodos de estimación directa frente a los métodos de estimación de información limitada y completa a pesar que son estos métodos los que garantizan la consistencia (en el caso de los primeros) y la consistencia y la eficiencia (los segundos) de los estimadores ha estado motivada por distintas causas: *a)* por el reducido número de observaciones disponibles (en este sentido cabe señalar que la evidencia empírica ha puesto de manifiesto que con series cortas no están tan claras las ventajas de los métodos de estimación conjunta respecto de los uniecuacionales); *b)* por el hecho que los métodos de estimación conjunta sean más sensibles que los directos ante problemas de multicolinealidad y de errores de medida en las variables; y *c)* porque al estimar de forma conjunta se trasladan los errores que se puedan cometer en una ecuación a las otras ecuaciones del modelo.

En cuanto a los criterios de selección entre ecuaciones alternativas (sobre la inclusión de determinadas variables

⁴ No hemos podido realizar el estudio a nueve ramas de actividad diferenciando entre Resto de Servicios destinados a la Venta y Servicios no destinados a la Venta porque las series correspondientes a estas dos variables sólo están disponibles desde 1980, por lo que el número de observaciones es insuficiente para realizar un análisis como el que presentamos en este trabajo.

⁵ En este sentido, el modelo que presentamos ha sido utilizado para elaborar predicciones sobre el comportamiento de dicha macromagnitud de la economía catalana para 1995 y 1996 a partir de diferentes escenarios nacionales de referencia.

exógenas o no) han sido los habituales en todo trabajo econométrico: análisis de la significación individual y conjunta de los parámetros, estudio de los coeficientes de determinación corregidos por los grados de libertad, estudio de la varianza estimada del término de perturbación, tests de cumplimiento de las hipótesis básicas sobre el término de perturbación, forma funcional, etc. Además también hemos valorado de la capacidad de ajuste de las diferentes ecuaciones especificadas, etc. En cualquier caso, dado el objetivo del modelo, el criterio que en última instancia hemos adoptado para seleccionar entre las posibles especificaciones contempladas en condiciones de igualdad en otros estadísticos, ha sido el de menor *EPAM* de predicción *ex-post* a dos períodos. Es decir, para cada sector hemos optado por la especificación que garantizaba una mejor capacidad predictiva.

En el cuadro 5 presentamos las ecuaciones de comportamiento estimadas así como diferentes estadísticos calculados para cada una de ellas con el fin de valorar el ajuste obtenido⁶. Si analizamos dichos resultados podemos afirmar, en líneas generales, que son satisfactorios. En este sentido, los ajustes obtenidos son buenos tal y como se deduce de los coeficientes de determinación que en todos los casos son superiores a 0.8 excepto en la ecuación correspondiente al VAB del subsector de Bienes de Consumo que es aproximadamente 0.79; los parámetros estimados son en todos los casos significativos; las relaciones de comportamiento son adecuadas en el sentido que el contraste de significación global de los parámetros nos lleva a concluir que la ecuación especificada es correcta; no hay problemas de autocorrelación tal y como pone de manifiesto el análisis de los estadísticos *DW*, *Q* de Box-Pierce y *Q* de Ljung-Box (estos últimos calculados a partir de los primeros cuatro coeficientes de autocorrelación simple de los residuos estimados de cada ecuación); el estadístico de *B-J* pone de relieve que no existen problemas de no normalidad en los residuos estimados de las ecuaciones; los parámetros asociados a las relaciones de largo plazo (parámetros que miden la velocidad de ajuste) están comprendidos tal y como cabía esperar entre cero y uno (en valor absoluto) siendo todos ellos claramente significativos puesto que los *t*-estadísticos superan los valores críticos tabulados por Banerjee *et al.* (1993b). Ello permite en base al Teorema de la Representación de Granger (Granger (1981) y Engle y Granger (1987)) concluir, como demostraron Kremers *et al.* (1992), que las variables incluidas en la relación a largo plazo están cointegradas.

6. ANÁLISIS DE LA CAPACIDAD PREDICTIVA DEL MODELO.

Para evaluar la capacidad predictiva del modelo hemos procedido como sigue: una vez seleccionadas las ecuaciones para todo el período muestral, las hemos reestimado para el período 1973-92. Una vez reestimadas con las primeras veinte observaciones realizamos una predicción *ex-post* para los años 1993 y 1994. De esta forma, dado que los valores correspondientes a estos dos últimos años son conocidos, podemos analizar la bondad de las predicciones obtenidas

⁶ El período muestral utilizado para estimar las ecuaciones del modelo ha sido desde 1973 hasta 1994 (ambos incluidos). En consecuencia disponemos de veintidós observaciones para la estimación.

para los años 1993 y 1994⁷.

En el cuadro 6 recogemos los valores predichos por el modelo y los valores proporcionados por Hispadat para 1993 y 1994. A partir de dichos datos calculamos diferentes estadísticos para valorar la capacidad predictiva del modelo. Así, en primer lugar, presentamos el *EPAM* de predicción (*ex-post*) para los dos años evaluados. Teniendo en cuenta que existe la opinión consensuada que valores del *EPAM* inferiores a tres son indicativos de una buena capacidad de predicción, entre tres y cinco no del todo buena y que si es superior a cinco podemos considerar que el modelo presenta una mala capacidad predictiva, los resultados obtenidos ponen de manifiesto que la capacidad predictiva del modelo es, en términos generales, aceptable excepto para el sector de la Construcción, aunque en este caso el *EPAM* obtenido no supera la cota del 5%.

Adicionalmente hemos calculado una batería de estadísticos a partir de los errores de predicción. En concreto, recogemos el *EMA*, el *RECM*, el *ECMP* y la U^2 de Theil. El análisis de los resultados obtenidos muestra que la capacidad predictiva es buena siendo el sector de la Construcción el que presenta peor capacidad predictiva.

Por último, recogemos las tasas de variación reales y predichas para el VAB subsectorial de la economía catalana para los años 1993 y 1994. Los resultados obtenidos muestran que para el año 1994 siempre queda recogida la trayectoria de crecimiento de los VAB subsectoriales (los signos coinciden con los reales) excepto para el subsector de los Bienes de Consumo. Sin embargo, para 1993 hay una mayor divergencia entre las tasas reales y las predichas. Además, tanto para el año 1993 como para el 1994 el sector de la Construcción es, una vez más el que presenta un mayor desajuste entre los valores reales y los predichos.

A modo de conclusión y, a la vista de todo lo anterior, podemos considerar que el modelo presenta una buena capacidad predictiva. En cualquier caso, hay que tener en cuenta que el año 1993 supuso el principio de una fase de recuperación económica tras los años de crisis anteriores. Ello produce un efecto negativo en la capacidad predictiva dado que es difícil que el modelo recoja a la perfección tal cambio.

Respecto al sector de la Construcción, que es el sector para el cual se han obtenido peores resultados en términos de capacidad predictiva, cabe señalar como descargo que se trata de un sector muy dinámico y que además, dadas las características singulares que nuestra Comunidad vivió durante los primeros años de la década de los noventa (nos estamos refiriendo a las Olimpiadas de 1992), hicieron que este sector en dichos años experimentara un auge muy importante que, pasado dicho evento, no tuvo continuidad produciéndose en consecuencia un cambio difícilmente predecible por cualquier modelo econométrico.

⁷ En cualquier caso la metodología utilizada para evaluar la capacidad predictiva del modelo presenta varios inconvenientes. En este sentido hay que señalar que para los años 1993 y 1994 no se disponen de cifras oficiales sobre las variables de VAB subsectorial estudiadas. Por ello hemos tomado los datos proporcionados por Hispadat con lo cual en realidad no estamos validando la capacidad predictiva del modelo respecto de las cifras oficiales. Ello nos debe poner en alerta en el sentido que los resultados que a continuación presentamos hay que tomarlos con prudencia.

ANEXO. DEFINICIÓN DE LAS VARIABLES¹

Variable	Definición
<i>VACT86</i>	Valor Añadido Bruto sector Agrícola Cataluña
<i>VECT86</i>	Valor Añadido Bruto subsector Bienes Energéticos Cataluña
<i>VQCT86</i>	Valor Añadido Bruto subsector Bienes Intermedios Cataluña
<i>VKCT86</i>	Valor Añadido Bruto subsector Bienes de Equipo Cataluña
<i>VCCT86</i>	Valor Añadido Bruto subsector Bienes de Consumo Cataluña
<i>VBCT86</i>	Valor Añadido Bruto subsector de la Construcción Cataluña
<i>VZCT86</i>	Valor Añadido Bruto subsector Transportes y Comunicaciones Cataluña
<i>VLGCT86</i>	Valor Añadido Bruto subsector Resto Servicios Cataluña (incluye los subsectores de Resto de Servicios destinados a la Venta y Servicios no destinados a la Venta)
<i>VTOTCT86</i>	Valor Añadido Bruto total Cataluña
<i>VABI86</i>	Valor Añadido Bruto subsector Bienes Intermedios España
<i>VABE86</i>	Valor Añadido Bruto subsector Bienes de Equipo España
<i>VATC86</i>	Valor Añadido Bruto subsector Transportes y Comunicaciones España
<i>PROAG</i>	Productividad aparente del sector Agrícola (se define como el cociente entre el VAB Agrícola nacional y el Empleo Agrícola nacional: $VAAG86/EEAG$)
<i>PPRIAG</i>	Precios Compuestos del sector Agrícola
<i>PPRIEN</i>	Precios Compuestos del subsector de Bienes Energéticos
<i>EEC</i>	Población Ocupada en el subsector de Bienes de Consumo en España
<i>EEIND</i>	Población Ocupada en el sector Industrial en España
<i>IVCST86</i>	Formación Bruta de Capital Fijo en Construcción a precios constantes de 1986
<i>EECST</i>	Población Ocupada en el sector de la Construcción en España
<i>YDH86</i>	Renta Disponible a precios constantes de 1986

¹ Todas las variables están expresadas a precios de mercado y a pesetas constantes de 1986.

Las letras *ln* delante del nombre de la variable indica que la variable está expresada en logaritmos (naturales).

La letra Δ delante del nombre de la variable indica que la variable está en primeras diferencias.

El primer año del que disponemos de información muestral es 1970 excepto para *VACT86* y *VLGCT86* que es 1971 y 1972.

Variables	$S-G^1$	DF y DFA d=2				DF y DFA d=1				T
	DW	Modelo ²	\mathbf{g}	t_g	$t \ t_m \ t_i$ ³	Modelo ²	\mathbf{g}	t_g	$t \ t_m \ t_i$ ³	
lnVACT86	1.515	T,0	-1.39	-6.227	-3.6746	C,0	-0.79	-3.238	-3.0199	1971-92
lnVECT86	0.076	T,1	-1.27	-4.102	-3.6591	C,1	-0.15	-2.807	-3.0114	1970-92
lnVQCT86	0.078	T,0	-0.95	-3.916	-3.6591	C,0	-0.11	-2.186	-3.0114	1970-92
lnVKCT86	0.237	N,0	-0.51	-2.913	-1.9592	C,1	-0.27	-2.988	-3.0114	1970-92
lnVCCT86	0.468	T,0	-0.99	-4.023	-3.6591	N,0	0.000	0.760	-1.9583	1970-92
lnVBCT86	0.299	C,0	-0.73	-3.269	-3.0199	C,1	-0.20	-1.669	-3.0114	1970-92
lnVZCT86	0.038	T,0	-0.95	-3.953	-3.6591	C,0	-0.06	-2.593	-3.0114	1970-92
lnVLGCT86	0.097	T,1	-0.63	-2.158	-3.6920	T,0	-0.07	-0.580	-3.6746	1972-92
lnVTOTCT86	0.105	T,0	-0.53	-2.415	-3.6920	T,1	-0.13	-1.491	-3.6746	1970-92
lnVABI86	0.097	C,1	-0.80	-3.717	-3.0038	C,1	-0.23	-4.098	-2.9969	1970-94
lnVABE86	0.152	C,0	-0.57	-3.521	-3.0038	C,1	-0.21	-2.952	-2.9969	1970-94
lnVATC86	0.030	N,0	-0.12	-2.080	-1.9574	C,2	-0.01	-0.763	-2.9969	1970-94
lnPROAG ⁴	0.035	T,0	-1.10	-4.700	-3.6330	C,0	-0.01	-0.626	-2.9969	1970-94
lnPPRIAG	0.024	T,0	-1.05	-4.656	-3.6330	C,1	-0.06	-2.967	-2.9969	1970-94
lnPPRIEN	0.029	N,0	-0.23	-1.644	-1.9574	C,1	-0.04	-1.850	-2.9969	1970-94
lnEEC	0.090	N,0	-0.30	-1.917	-1.9574	C,1	-0.06	-1.277	-2.9969	1970-94
lnEEIND	0.124	N,0	-0.51	-2.766	-1.9574	C,1	-0.06	-0.901	-2.9969	1970-94
lnEECST	0.227	N,1	-0.48	-2.809	-1.9574	N,1	-0.00	-0.114	-1.9566	1970-94
lnIVCST86	0.112	N,1	-0.41	-2.963	-1.9574	C,2	-0.02	-0.375	-3.0038	1970-94
lnYDH86	0.066	N,1	-0.27	-2.609	-1.9574	C,1	-0.05	-1.877	-2.9969	1970-94

Cuadro 1. Resultados de los tests de Sargan y Bhargava, Dickey y Fuller y Dickey y Fuller ampliado.

¹ Sargan y Bhargava (1983) únicamente tabularon los valores críticos para determinados tamaños muestrales. Los valores críticos para 21 y 31 observaciones son (para $\alpha=0.05$) 1.069 y 0.770.

² Las letras T, C y N hacen referencia a la consideración de constante y tendencia, sólo constante y ningún componente determinista en la regresión del test. La cifra que les acompaña indica el número de retardos de la endógena considerados en la regresión.

³ Valores críticos tabulados por Mackinnon (1991) para contrastar la significación del parámetro \mathbf{g} en la regresión correspondiente del test a un nivel de significación del 5%.

⁴ El resultado obtenido respecto al orden de integrabilidad de esta variable ($I(1)$) coincide con el esperado *a priori* puesto que dicha variable se obtiene a partir de dos variables que son integrables de orden uno. Ello constituye, por tanto, una prueba de la propiedad dominante del orden de integrabilidad.

REGRESIÓN DE COINTEGRACIÓN			RESULTADOS DE LOS TESTS DE DF Y DFA SOBRE LOS RESIDUOS DE LA REGRESIÓN DE COINTEGRACIÓN							
<i>Ecuación</i>	<i>R²</i>	<i>DW</i>	<i>k¹</i>	<i>g</i>	<i>t_g</i>	<i>VC 5%²</i>	<i>DW³</i>	<i>Q(4)³</i>	<i>d</i>	<i>t_d</i>
<i>lnVECT86_t</i> =-1.66+0.86 <i>lnVTOTCT86_t</i> +0.22 <i>lnPPRIEN_t</i> + <i>e_t</i> (-1.21) (9.37) (12.96)	0.97	2.12	0	-1.08	-5.01	-4.1710	1.88	2.09	---	---
<i>lnVQCT86_t</i> =-0.33+0.43 <i>lnVTOTCT86_t</i> +0.88 <i>lnVABI86_t</i> + <i>e_t</i> (-0.32) (3.97) (7.00)	0.95	1.87	1	-1.24	-5.32	-4.1710	2.04	3.45	0.42	2.57
<i>lnVKCT86_t</i> =7.12+0.79 <i>lnVABE86_t</i> + <i>e_t</i> (15.67) (13.28)	0.89	0.64	1	-0.39	-2.21	-3.6141	1.99	0.91	0.20	0.95
<i>lnVCCT86_t</i> =2.19+0.55 <i>lnVTOTCT86_t</i> +0.42 <i>lnEEC_t</i> + <i>e_t</i> (1.55) (8.23) (5.32)	0.77	1.73	0	-0.88	-4.01	-4.1710	1.90	1.97	---	---
<i>lnVBCT86_t</i> =3.78+1.07 <i>lnEECST_t</i> +0.20 <i>lnIVCST86_t</i> + <i>e_t</i> (6.28) (11.21) (2.93)	0.91	1.30	0	-0.65	-3.27	-4.1314	1.85	0.41	---	-----
<i>lnVZCT86_t</i> =2.69+1.33 <i>lnVATC86_t</i> + <i>e_t</i> (17.94) (65.22)	0.99	1.23	0	-0.68	-3.71	-3.6141	1.91	1.57	---	---
<i>lnVLGCT86_t</i> =-1.64+0.64 <i>lnVTOTCT86_t</i> +0.63 <i>lnYDH86_t</i> + <i>e_t</i> (-2.39) (5.71) (5.41)	0.99	1.84	0	-0.98	-4.45	-4.1710	1.85	3.10	---	---

Cuadro 4. Resultados del análisis de relaciones de cointegración. Método de Engle y Granger (1987). Este método consiste en estimar por MCO una regresión entre las variables (en niveles) sobre las cuales se quiere analizar si están cointegradas. En una segunda etapa se contrasta si los residuos estimados de la regresión anterior son $I(1)$ frente a la alternativa que son estacionarios. Para llevar a cabo tal contraste los autores propusieron siete tests entre los que se encuentran los de Sargan y Bhargava, DF y DFA. Bajo la hipótesis nula que las variables no están cointegradas los residuos presentarán una raíz unitaria y la regresión en niveles no es más que una relación espúrea; por el contrario, si los residuos son estacionarios las variables están cointegradas.

¹ Número de retardos de la primera diferencia de los residuos de la regresión de cointegración incluidos en la regresión del test.

² Valores críticos de Mackinnon (1991).

³ Estadísticos calculados a partir de los residuos de la regresión del test de DF y DFA.

$\ln VACT \hat{86}_t = 10.9893 - 0.3776 \ln PROAG_t + 0.1987 \ln PPRIAG_t - 0.1197 D7677_t + 0.1813 D80_t - 0.1448 D86_t + 0.1112 D8990_t$ $(69.1948) \quad (-7.2779) \quad (5.3653) \quad (-4.6708) \quad (5.2592) \quad (-4.0142) \quad (4.2369)$ $R^2 = 0.9722 \quad F = 125.6313 \quad DW = 1.952 \quad Q_{B-P}(4) = 1.27 \quad Q_{L-B}(4) = 1.53 \quad BJ = 0.624$
$\Delta \ln VECT \hat{86}_t = 0.7439 \Delta \ln VTOTCT86_t + 0.1092 D7475_t - 0.9771 [\ln VECT86_{t-1} + 0.4475 - 0.7833 \ln VTOTCT86_{t-1} - 0.2047 \ln PPRIEN_{t-1}]$ $(3.708) \quad (5.030) \quad (-6.391) \quad (0.440) \quad (-11.714) \quad (-17.821)$ $R^2 = 0.8559 \quad F = 19.0064 \quad DW = 2.401 \quad Q_{B-P}(4) = 3.86 \quad Q_{L-B}(4) = 4.63 \quad BJ = 0.281$
$\Delta \ln VCCT \hat{86}_t = 0.9078 \Delta \ln VTOTCT86_t - 0.0698 D74_t - 0.8142 D80_t - 0.8723 [\ln VCCT86_{t-1} - 2.8047 - 0.4955 \ln VTOTCT86_{t-1} - 0.4439 \ln EEC_{t-1}]$ $(3.978) \quad (-2.058) \quad (-2.511) \quad (-4.186) \quad (-1.763) \quad (-6.436) \quad (-5.035)$ $R^2 = 0.7895 \quad F = 8.349 \quad DW = 1.739 \quad Q_{B-P}(4) = 3.87 \quad Q_{L-B}(4) = 4.66 \quad BJ = 0.896$
$\Delta \ln VQCT \hat{86}_t = 0.2096 + 0.5301 \Delta \ln VTOTCT86_t + 0.6301 \Delta \ln VABI86_t - 0.0691 D7475_t - 0.8159 [\ln VQCT86_{t-1} - 2.8828 \ln VTOTCT86_{t-1} - 0.4439 \ln EEC_{t-1}]$ $(0.3930) \quad (3.6656) \quad (3.6239) \quad (-4.7381) \quad (-6.5033) \quad (-3.6577) \quad (-9.1342)$ $R^2 = 0.8727 \quad F = 17.137 \quad DW = 1.874 \quad Q_{B-P}(4) = 3.39 \quad Q_{L-B}(4) = 4.46 \quad BJ = 1.03$
$\Delta \ln VKCT \hat{86}_t = 0.5592 \Delta \ln VABE86_t + 0.7253 \Delta \ln EEIND_t + 0.1617 D73_t + 0.0566 D78_t - 0.4577 [\ln VKCT86_{t-1} - 5.9986 - 0.9402 \ln VABE86_{t-1}]$ $(4.0517) \quad (3.3272) \quad (4.8171) \quad (2.0146) \quad (-4.1553) \quad (-7.5856) \quad (-9.1342)$ $R^2 = 0.9145 \quad F = 26.725 \quad DW = 1.947 \quad Q_{B-P}(4) = 3.61 \quad Q_{L-B}(4) = 4.62 \quad BJ = 0.079$
$\Delta \ln VBCT \hat{86}_t = 1.3543 \Delta \ln VTOTCT86_t + 0.6174 \Delta \ln EECST_t + 0.0813 D74_t + 0.1032 D86_t - 0.8328 [\ln VBCT86_{t-1} - 3.2839 - 1.0402 \ln EECST_{t-1}]$ $(3.8373) \quad (3.3231) \quad (1.9462) \quad (2.3185) \quad (-4.7345) \quad (-5.7717) \quad (-11.231)$ $R^2 = 0.9192 \quad F = 22.7693 \quad DW = 1.923 \quad Q_{B-P}(4) = 1.67 \quad Q_{L-B}(4) = 2.07 \quad BJ = 0.079$
$\Delta \ln VZCT \hat{86}_t = 2.0092 \Delta \ln VATC86_t + 0.0706 D86_t - 0.7084 [\ln VZCT86_{t-1} - 1.9991 - 1.4188 \ln VATC_{t-1}]$ $(6.6867) \quad (3.0825) \quad (-3.9590) \quad (-7.4150) \quad (-40.619)$ $R^2 = 0.8325 \quad F = 23.6155 \quad DW = 1.891 \quad Q_{B-P}(4) = 5.19 \quad Q_{L-B}(4) = 6.44 \quad BJ = 2.577$
$\Delta \ln VLGCT \hat{86}_t = -1.4955 + 0.9972 \Delta \ln VTOTCT86_t - 0.0355 D83_t + 0.0281 D85_t - 0.9139 [\ln VLGCT86_{t-1} - 0.6154 \ln VTOTCT86_{t-1} - 0.4439 \ln EEC_{t-1}]$ $(-2.7015) \quad (13.2217) \quad (-2.7248) \quad (2.2584) \quad (-5.6803) \quad (-6.7867) \quad (-7.343)$ $R^2 = 0.9358 \quad F = 36.4779 \quad DW = 2.220 \quad Q_{B-P}(4) = 2.90 \quad Q_{L-B}(4) = 3.63 \quad BJ = 0.798$

Cuadro 5. Resultados de la estimación del modelo de VAB de la economía catalana.

<i>VARIABLE</i>	<i>D.S. niveles</i>	<i>D.S. 1º diferencia</i>	<i>D.S. 2ª diferencia</i>	<i>T</i>
<i>lnVLGCT86</i>	0.1396211	0.0377489	0.0514157	1972-92
<i>lnVTOTCT86</i>	0.1177709	0.0308221	0.0320936	1974-92
<i>lnVABI86</i>	0.1185201	0.0504744	0.0551507	1972-94
<i>lnPPRIEN</i>	0.7593468	0.1148038	0.1030850	1972-94

Cuadro 2. Desviaciones estándar de las series ante sucesivas diferenciaciones.

<i>Sector</i>	<i>Predicción ex-post</i>		<i>Hispatat</i>		<i>EPAM</i> ¹	<i>EMA</i> ²	<i>RECM</i> ³	<i>ECMP</i> ⁴	<i>U</i> ⁵	<i>Tasa variación real</i> ⁶		<i>Tasa variación predicha</i> ⁶	
	1993	1994	1993	1994						$\frac{y_{1993}}{y_{1992}} - 1$	$\frac{y_{1994}}{y_{1993}} - 1$	$\frac{\hat{y}_{1993}}{y_{1992}} - 1$	$\frac{\hat{y}_{1994}}{\hat{y}_{1993}} - 1$
<i>A</i>	124655.3	127431.6	121885.8	124811.0	0.765	2695.05	2696.08	0.0478	0.00017	-4.00	2.40	-1.82	2.22
<i>E</i>	379768.7	388403.3	388254.1	397960.4	2.293	9021.28	9037.18	0.0527	0.00013	1.10	2.50	-1.10	2.27
<i>C</i>	995052.4	991501.1	992301.3	1005201.1	0.820	8225.55	9880.75	0.0096	0.00002	-3.20	1.30	-2.94	-0.35
<i>Q</i>	496672.2	513290.3	514725.9	528108.8	3.156	16436.1	16515.5	0.1008	0.00025	-0.70	2.60	-4.18	3.34
<i>K</i>	575782.9	595136.4	585122.1	593313.7	0.952	5580.93	6728.39	0.0132	0.00003	-7.60	1.40	-9.07	3.35
<i>B</i>	459801.6	450277.8	477050.6	472280.1	4.137	19625.6	19769	0.1738	0.00045	-9.20	-0.10	-12.48	-2.07
<i>Z</i>	431380.1	458014.7	424851.9	442270.7	2.548	11136.1	12051.7	0.0751	0.00018	-1.60	4.10	-0.09	6.17
<i>LG</i>	3429928	3550425	3509695	3593928	1.741	61634.7	64246.6	0.0331	0.00008	0.30	2.40	-1.97	3.51

Cuadro 6. Estadísticos para valorar la capacidad predictiva del modelo.

¹ Error Porcentual Absoluto Medio.

² Error Medio Absoluto.

³ Raíz cuadrada del Error Cuadrático Medio.

⁴ Error Cuadrático Medio Porcentual.

⁵ U^2 de Theil. Este estadístico compara los valores reales con los predichos por el modelo valorando la importancia de un posible error sistemático en la predicción. Una de las ventajas de este estadístico es que está acotado entre cero y uno. Si toma el valor cero significa que el coeficiente de correlación entre los valores reales y los predichos es uno y además coinciden año a año, es decir, cuando $\hat{y} = y$.

⁶ En tanto por ciento.

Los tres primeros estadísticos presentan el inconveniente que dependen de las unidades de medida en que están expresadas las variables por lo que hay que interpretarlos con prudencia. En cualquier caso, cuanto más próximos a cero estén mejor capacidad predictiva tiene el modelo (esto último también es válido para los dos últimos).

Los estadísticos calculados se definen de la forma siguiente:

$$e_t = y_t - \hat{y}_t ; \quad EMA = \frac{\sum_{t=1993}^{1994} e_t}{2} ; \quad RECM = \sqrt{\frac{\sum_{t=1993}^{1994} e_t^2}{2}} ; \quad ECMP = \frac{\sum_{t=1993}^{1994} \left(\frac{e_t}{y_t} \right)^2}{2} 100$$

$$U^2 = \frac{\frac{\sum_{t=1993}^{1994} y_t^2}{2} + \frac{\sum_{t=1993}^{1994} \hat{y}_t^2}{2}}{\left[\sqrt{\frac{\sum_{t=1993}^{1994} y_t^2}{2}} + \sqrt{\frac{\sum_{t=1993}^{1994} \hat{y}_t^2}{2}} \right]^2}$$

BIBLIOGRAFÍA

- * ANDRÉS, J., ESCRIBANO, A., MOLINAS, C. y TAGUAS, D. (1990a): *La Inversión en España. Econometría con Restricciones de Equilibrio*. Antoni Bosch editor e IEF.
- * ANDRÉS, J., MOLINAS, C. y TAGUAS, D. (1990b): "Una Función de Consumo Privado para la Economía Española: Aplicación del Análisis de Cointegración", *Cuadernos Económicos de ICE*, 1990/1, 44, pp. 173-212.
- * ARTÍS, M., CLAR, M., Del BARRIO, T., GARCÍA, J.R., SANSÓ, A. y SURINACH, J. (1996): "Model Economètric Regional per Catalunya. Previsions Macromagnituds sobre l'Economia Catalana pel 1995 i 1996", *Document de treball DOCT-96R07*, Departament d'Econometria, Estadística i Economia Espanyola, Universitat de Barcelona.
- * BANERJEE, A., DOLADO, J., GALBRAITH, J. y HENDRY, D.F. (1993a): *Co-Integration, Error-Correction and the Economic Analysis of Non-Stationary Data*. Oxford University Press.
- * BANERJEE, A., DOLADO, J. y MESTRE, R. (1993b): "On Some Tests from Cointegration: the Cost of Simplicity", *Banco de España. Documento de trabajo* 9302.
- * CLAR, M. (1996): Una Aplicación de la Teoría de la Cointegración a la Modelización del VAB Subsectorial de la Economía Catalana, Tesina de Licenciatura, Departament d'Econometria, Estadística i Economia Espanyola, Universitat de Barcelona.
- * DICKEY, D.A. y FULLER, W. (1979): "Distribution of the Estimators for Autoregressive Time Series with a Unit Root", *Journal of American Statistical Association*, 74, pp. 427-431.
- * DICKEY, D.A. y FULLER, W. (1981): "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica*, 49, pp. 1057-1072.
- * DICKEY, D.A. y PANTULA, S.G. (1987): "Determining the Order of Differencing in Autoregressive Processes", *Journal of Business and Economics Statistics*, 5, 4, pp. 455-461.
- * ENGLE, R.F. y GRANGER, C.W.J. (1987): "Cointegration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55, 2, pp. 251-276.
- * JOHANSEN, S. (1988): "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12, pp. 231-254.
- * KREMERS, J., ERICSSON, N., DOLADO, J. (1992): "The Power of Cointegration Tests", *Banco de España. Documento de trabajo* 9218.
- * MACKINNON, J. (1991): "Critical Values for Cointegration Test", en ENGLE, R. y GRANGER, C.W.J. editores (1991), *Long-Run Economic Relationships*, pp. 267-276.
- * MOLINAS, C., BALLABRIGA, F.C., CANADELL, E., ESCRIBANO, A., LÓPEZ, E., MANZANEDO, L., MESTRE, R., SEBASTIÁN, M. y TAGUAS, D. (1990): *MOISEES: Un Modelo de Investigación y Simulación de la Economía Española*. Antoni Bosch editor e IEF.
- * NELSON, C.R., y PLOSSER, C.I. (1982): "Trends and Random Walks in Macroeconomic Time Series. Some Evidence and Implications", *Journal of Monetary Economics*, 10, pp. 139-162.
- * PERRON, P. (1990): "Testing for a Unit Root in Time Series with a Changing Mean", *Journal of Economic Statistics*, 8, pp. 153-162.
- * RAPPOPORT, P. y REICHLIN, L. (1989): "Segmented Trends and Non-Stationary Time Series", *The Economic Journal*, 99, pp. 168-177.
- * SARGAN, J.D. y BHARGAVA, A. (1983): "Testing Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk", *Econometrica*, 51, pp. 153-174.
- * SURINACH, J., ARTÍS, M., LÓPEZ, E. y SANSÓ, A. (1995): *Análisis Económico Regional. Nociones Básicas de la Teoría de la Cointegración*, Antoni Bosch editor.