

ASEPELT 99

Título: “Un Modelo Econométrico Predictivo en el Marco de la Teoría de la Cointegración para la Economía Extremeña”.

Resumen:

El objetivo de este trabajo es la especificación, estimación y validación de un modelo econométrico con fines fundamentalmente predictivos para el crecimiento de las cifras de VABpm de las ramas de la economía extremeña. El modelo obtenido es uniregional, descendente y de orientación de base económica, dentro del marco de la teoría de la cointegración. Metodológicamente, ocupa un lugar muy importante el tratamiento del cambio estructural, dado que la práctica econométrica estándar no es aplicable a las economías regionales que, como la extremeña, están sujetas a una gran volatilidad.

Autores: Miguel Ángel Márquez Paniagua y Julián Ramajo Hernández

Área temática: Crecimiento, empleo y distribución de la renta.

Palabras clave: *Modelización econométrica regional, Integrabilidad, Cointegración, Cambio estructural.*

Universidad o centro de investigación: Universidad de Extremadura

Dirección postal:

Universidad de Extremadura
Facultad de Ciencias Económicas y Empresariales
Departamento de Economía Aplicada y Organización de Empresas
Avda. de Elvas, s/n
06071-Badajoz

Teléfono: 924289300-Extensión 9178 (Miguel Ángel Márquez)
924289547 (Julián Ramajo)

FAX: 924-272509

E-mail: mmárquez@unex.es (Miguel Ángel Márquez) y
ramajo@unex.es (Julián Ramajo).

“Un Modelo Econométrico Predictivo en el Marco de la Teoría de la Cointegración para la Economía Extremeña.”

Miguel Ángel Márquez Paniagua.
Universidad de Extremadura
Julián Ramajo Hernández
Universidad de Extremadura

Resumen

El objetivo de este trabajo es la especificación, estimación y validación de un modelo econométrico con fines fundamentalmente predictivos para el crecimiento de las cifras de VABpm de las ramas de la economía extremeña. El modelo obtenido es uniregional, descendente y de orientación de base económica, dentro del marco de la teoría de la cointegración. Metodológicamente, ocupa un lugar muy importante el tratamiento del cambio estructural, dado que la práctica econométrica estándar no es aplicable a las economías regionales que, como la extremeña, están sujetas a una gran volatilidad.

1. INTRODUCCIÓN

En este trabajo se expone (de manera resumida) los principales resultados obtenidos en el proceso de especificación, estimación y validación de un modelo econométrico con fines fundamentalmente predictivos para el crecimiento de las cifras de VABpm de las ramas de la economía extremeña. Así, atendiendo a la clasificación de Weber (1986), se presenta el bloque de producción que constituye la estructura de un modelo econométrico uniregional satélite para Extremadura. El hecho de caracterizarlo de esta forma se justifica porque las variables endógenas están en función de variables endógenas regionales, exógenas regionales, y exógenas nacionales (que ocupan un lugar muy importante en la dirección del sistema en general). Además, se ignora la posible causalidad de los cambios en la región sobre las variables nacionales. En definitiva, se trata de un modelo uniregional (porque no consideramos las interrelaciones con otras regiones), descendente (“top-down”, en el sentido de que las variables nacionales tienen un papel director importante), y en el que no se tienen en cuenta las posibles retroalimentaciones (“feed-back”) de la economía regional extremeña en la nacional (por lo que en términos de la tipología descrita en Courbis (1994) se trataría de un modelo abierto). Se ha considerado necesario adoptar un enfoque de tipo descendente (“top-down”), en los que se trabaja tanto con variables nacionales (que se consideran exógenas) como con variables regionales, ya sean endógenas y/o exógenas.

Las herramientas tradicionales de análisis econométrico se fundamentan generalmente en el supuesto de que la estructura del sistema económico bajo estudio es estable. Sin embargo, esta hipótesis es muy restrictiva en el caso de modelizar economías regionales, puesto que es difícil asumir que el comportamiento futuro será similar al observado en el pasado. Así, Extremadura ha sufrido grandes cambios a lo largo de las últimas décadas, apareciendo el problema de la no existencia de una estructura estable para todo el período en estudio.

En consecuencia, y desde la perspectiva metodológica, ocupa un lugar muy importante la investigación de los efectos que produce la presencia de cambio estructural sobre el proceso de elaboración de modelos econométricos regionales, ya que ésta ha sido la dificultad principal a la que hemos tenido que hacer frente. La especificación y

estimación de un modelo econométrico para Extremadura ha supuesto proponer un procedimiento (que se aparta de la práctica econométrica estándar) para aplicar cuando las relaciones estructurales no son constantes.

La construcción de un modelo econométrico a escala regional tiene como objetivo típico la obtención de inferencias acerca del presente o futuro de la región analizada a partir de los datos históricos sobre la misma. Para que las inferencias sean válidas se requiere que los modelos construidos sean “estables”, en el sentido de que se supone que el futuro será similar al pasado, implicando este hecho que los parámetros del modelo econométrico han de ser constantes. Si la hipótesis de constancia de los parámetros no se verifica en la práctica, cualquier inferencia que se obtenga a partir de los mismos estará sesgada. En el caso de las economías regionales, el problema de la inestabilidad se ve aún más agudizado que en el caso de economías nacionales, puesto que a ese nivel el impacto de “*shocks*” externos o internos es mucho mayor que para el total del país.

El marco estadístico de partida es la teoría de la cointegración (Engle y Granger, 1987), la cual combina en su especificación econométrica básica las relaciones de equilibrio a largo plazo sugeridas por la teoría económica con el proceso de ajuste al equilibrio (en el corto plazo) de dichas relaciones a través de los mecanismos de corrección de error. En el largo plazo, y dada la interpretación de las relaciones de cointegración, es previsible la presencia de un número muy reducido de puntos de ruptura, por lo que la hipótesis que se utiliza es que pueden modelizarse dichos cambios a través de la introducción de variables ficticias¹. En el corto plazo esta hipótesis se reemplaza por otra más general, en la que los parámetros del modelo pueden variar de forma continua a lo largo del tiempo. La hipótesis alternativa a la constancia es que los parámetros son estocásticos² y varían según un modelo tipo >paseo aleatorio= (multivariante), por lo que el modelo resultante puede acomodar cambios estructurales de todo tipo (bruscos o suaves) que hayan tenido lugar durante el período muestral. La organización del trabajo es como sigue. En la sección dos se detallan los resultados básicos concernientes a la especificación estructural que sirve de soporte para el análisis empírico posterior. En la sección tres se discute la metodología utilizada, así como el modelo estadístico común para todas las ecuaciones del modelo econométrico, el cual se construye en la sección cuatro. En esta última sección se procede en primer lugar al análisis de las propiedades estocásticas de las series utilizadas. Posteriormente, se pasa a la especificación de las ecuaciones a largo plazo de los siete sectores económicos considerados, analizando en cada caso la estabilidad de la relación estimada. Una vez reestimada ésta (si se hace necesario, a la vista de los resultados de los contrastes de estabilidad), se modelizan los mecanismos de corrección de error para el corto plazo, investigando de nuevo la verificación de la hipótesis de estabilidad paramétrica y reespecificando cada ecuación, si ello es preciso, siguiendo los dos enfoques metodológicos considerados en la sección tres. En la última sección se ofrecen algunas conclusiones que pueden obtenerse de nuestro trabajo.

¹ En Canarella *et al.* (1990) la presencia de inestabilidad estructural en el largo plazo se modeliza usando un enfoque de parámetros variables en el tiempo. Sin embargo, tal como señala Hall (1994), sería difícil dar una interpretación económica al modelo de largo plazo resultante de tal enfoque.

² La modelización mediante variables ficticias implica que la alternativa a la hipótesis de constancia es un cambio determinista de los parámetros en el tiempo. Aparte de esta diferencia, el enfoque estocástico propuesto asume que el cambio en los parámetros es >suave=, frente a la hipótesis de cambio discreto (y por tanto brusco) del enfoque de variables ficticias.

2. EL MODELO ESTRUCTURAL BÁSICO

El modelo econométrico que se ha construido (el cual se detalla en la sección cuatro) está basado en las ideas fundamentales de los modelos de “base económica” (ver, por ejemplo, Treyz, 1993). Por razones estadísticas (ausencia en España de series temporales a nivel regional suficientemente largas³ desagregadas desde la perspectiva de la demanda), el modelo económico propuesto realiza una desagregación sectorial por el lado de la oferta. En concreto, se divide la producción regional en siete sectores básicos: agricultura, energía, industria manufacturera, construcción, servicios destinados a la venta (excepto transportes y comunicaciones), transportes y comunicaciones, y servicios no destinados a la venta. Las variables endógenas a explicar (y predecir) vienen dadas por la producción de cada uno de esos sectores, medida a través del valor añadido bruto a precios de mercado en pesetas constantes de 1986 (que representaremos como W). Siguiendo la línea argumental de los modelos de base económica, pueden distinguirse los sectores básicos, cuya producción sirve para abastecer el mercado nacional o supranacional y los sectores no básicos o locales, los cuales venden la mayor parte de su producción en el mercado regional. Para los primeros, el nivel de actividad viene determinado fundamentalmente por factores externos, por lo que la especificación tipo es de la forma:

$$E[W^b/IE^b, IR_1^b, IR_2^b, \dots] = a_b + b_{b_0} IE^b + \sum_{i=1}^{n_b} b_{b_i} IR_i^b \quad (1)$$

donde $E[.]$ representa el valor esperado condicional, el superíndice hace referencia al sector básico b , IE es un indicador externo que mide la evolución del ciclo del mercado nacional del sector b ⁴, e IR son indicadores regionales que complementan la especificación básica (entre los que pueden encontrarse variables que midan las ventajas de localización del sector b en la región, así como otras que expliquen las relaciones intersectoriales dentro del entorno regional). Para los sectores locales, la ecuación tipo viene dada por

$$E[W^l/IR, IR_1^l, IR_2^l, \dots] = a_l + b_{l_0} IR + \sum_{i=1}^{n_l} b_{l_i} IR_i^l \quad (2)$$

donde ahora IR es un indicador del nivel de demanda interna total en la región⁵ e IR^l son indicadores regionales que complementan la relación básica (entre ellos pueden estar algunas variables que capten la relación del sector local con actividades básicas de la región).

Sin embargo, como en la práctica no existen mercados exclusivamente nacionales o regionales, los sectores productivos son habitualmente mixtos, en el sentido de que

³ Las bases de datos HISPALINK e HISPADAT, de las cuales se han extraído las principales macromagnitudes usadas en el trabajo, sólo contienen series de demanda final regionales para el período 1986-1993.

⁴ El coeficiente Ξ_{b0} medirá entonces la sensibilidad de la producción regional del sector ante variaciones en el valor añadido del mismo a escala nacional.

⁵ En este caso, el coeficiente Ξ_{l0} medirá la sensibilidad del sector ante cambios en la renta regional.

parte de su actividad se determina por factores exógenos a la región y parte por circunstancias endógenas⁶. Este hecho implica que las relaciones que se especificarán para los distintos sectores pueden ser una mezcla de las ecuaciones [1] y [2].

3. METODOLOGÍA

3.1. Especificación econométrica

La estructura funcional de las ecuaciones que componen el modelo econométrico regional está basada en la moderna teoría de la cointegración (Engle y Granger, 1987). En concreto, se proponen ecuaciones “tipo” en forma de mecanismo de corrección de error (MCE), en el que se plantea una relación de equilibrio a largo plazo entre las variables explicativas (no necesariamente exógenas) y la explicada (endógena), a la vez que se permite la existencia de desajustes en el corto plazo respecto a esa situación de equilibrio a través de la introducción de términos dinámicos. La estructura básica que proponemos para las ecuaciones del modelo es una variante de los mecanismos de corrección de error tradicionales, viniendo dada por la expresión

$$\Delta W_t^s = g_0 + \sum_{i=1}^p g_{1i} \Delta W_{t-i}^s + \sum_{i=0}^q g_{2i} \Delta Z_{t-i}^s - a (W_{t-1}^s - b' X_{t-1}^s) + u_t^s \quad (3)$$

donde W^s representa el valor añadido bruto del sector s (en logaritmos), X^s es un vector de variables explicativas (en general también en logaritmos) que cointegran con la variable dependiente W^s ⁷ y Z^s es un vector de variables que explica (junto con los valores retrasados de la variable dependiente) las desviaciones en el corto plazo de la situación de equilibrio (entre las componentes de Z^s se pueden encontrar algunas de las variables del vector X^s ⁸).

3.2. El modelo estadístico

En la especificación [3] hemos supuesto que los parámetros son fijos, es decir, que las relaciones estructurales a largo y corto plazo son estables. Para relajar esta hipótesis distinguiremos dos casos, según que se presente inestabilidad estructural en el largo o en el corto plazo. La ecuación [3] supone que la combinación lineal $W_t^s - \exists X_t^s$ de las variables integradas tiene una distribución estacionaria. Sin embargo, existe la posibilidad de un tipo más general de cointegración permitiendo que el vector de

⁶ Lo cual no es óbice para que en cada sector en particular prevalezca la característica básica o local. En este trabajo, la determinación de los sectores básicos y no básicos (hipótesis que posteriormente se contrasta mediante la teoría de la cointegración) y de las relaciones intersectoriales que se dan dentro de la economía extremeña, se ha realizado mediante la explotación de las tablas *input-output* que existen para Extremadura.

⁷ $W^s = \exists X^s$ será entonces la relación de equilibrio a largo plazo, y $W_t^s - \exists X_t^s$ medirá las desviaciones del equilibrio en cada instante t .

⁸ La especificación propuesta supone que todas las variables que aparecen (en niveles) en [3] son I(1).

cointegración \exists cambie en algún punto del período muestral⁹. La hipótesis nula de cointegración estándar implica el modelo

$$W_t^s = \mathbf{b}_0 + \mathbf{b}_1' \mathbf{X}_t^s + e_t^s \quad (4)$$

donde W_t^s y \mathbf{X}_t^s son I(1) y e_t^s es I(0). Si la relación [4] es estable, los parámetros Ξ_0 y Ξ_1 deben ser constantes (invariantes en el tiempo), pero si existe inestabilidad estructural dichos parámetros permanecerán constantes durante algún período de tiempo, para posteriormente cambiar (Ξ_0 o alguna componente del vector Ξ_1) hacia un nuevo nivel, obteniéndose otra relación de equilibrio con distintos valores para las pendientes o la ordenada en el origen; el cambio producido puede ser definitivo, pero también puede ocurrir que se vuelva a la situación original transcurrido un cierto período de tiempo o que se pase a otro estado de equilibrio caracterizado por un nuevo conjunto de coeficientes. Si suponemos que el cambio en los parámetros es discreto, entonces se puede modelizar el cambio estructural introduciendo una variable ficticia del tipo

$$\mathbf{j}_{[t_0, t_1]}(t) = \begin{cases} 0 & \text{si } t < t_0 \\ 1 & \text{si } t_0 \leq t \leq t_1 \\ 0 & \text{si } t > t_1 \end{cases} \quad (5)$$

donde t_0 denota el punto de ruptura de la relación de cointegración y t_1 el punto de vuelta a la situación inicial, con $1 < t_0 \neq t_1 \neq T$. En el caso más general donde el cambio estructural implica una modificación tanto de la ordenada en el origen como en las pendientes, la relación de cointegración con cambio estructural viene dada por

$$W_t^s = \mathbf{b}_{01} + \mathbf{b}_{02} \mathbf{j}_{[t_0, t_1]}(t) + \mathbf{b}_{11}' \mathbf{X}_t^s + \mathbf{b}_{12}' \mathbf{X}_t^s \mathbf{j}_{[t_0, t_1]}(t) + e_t^s \quad (6)$$

donde Ξ_{01} y Ξ_{02} representan, respectivamente, las ordenadas en el origen antes y después del cambio estructural, y Ξ_{11} y Ξ_{12} son los coeficientes de las pendientes en la relación de cointegración antes y después del cambio de régimen.

El modelo anterior se puede generalizar para permitir más de un punto de ruptura simplemente introduciendo variables ficticias adicionales. En cualquier caso, el cambio de régimen que se está suponiendo es completamente discreto. Podría relajarse esta hipótesis a través de dos mecanismos distintos. Uno de ellos (que será utilizado en este trabajo para el caso de la modelización de las relaciones de corto plazo) consiste en permitir que los coeficientes de la ecuación [4] sigan un proceso estocástico del tipo $\Xi_t = \Theta \Xi_{t-1} + \eta_t$, de forma tal que los parámetros cambien de forma continua a lo largo de todo el período muestral¹⁰. La otra alternativa consiste en modelizar el cambio estructural de una forma paramétrica, permitiendo que los parámetros cambien de forma gradual de un sistema estable a otro a través de alguna función Λ ; este el modelo propuesto por Lin y Teräsvirta (1994), quienes consideran una especificación del tipo¹¹

⁹ Por simplicidad en la exposición, suponemos que existe un único punto de ruptura.

¹⁰ Si $\Theta = \mathbf{I}$ tendremos un modelo de “paseo aleatorio” y si $\Theta = \mathbf{B} = \text{diag}\{b_1, b_2, \dots, b_K\}$ se tendrá el modelo de “vuelta a la normalidad”. Al menos bajo la perspectiva del largo plazo, ambos modelos - considerados globalmente- son “excesivamente” generales, en el sentido de que suponen un cambio estructural continuo (de período a período cambian todos los coeficientes); en este sentido, al menos en el modelo de vuelta a la normalidad los coeficientes convergen (si los parámetros b_i cumplen la restricción de ser menores que uno en valor absoluto) a un estado estacionario, Ξ^* , que sería el que tendría la interpretación como vector de cointegración.

¹¹ A la que Lin y Teräsvirta denominan modelo de transición suave (STR). El modelo [7] también supone que existe únicamente un punto de ruptura en la relación a largo plazo, pero fácilmente

$$W_t^s = \mathbf{b}_1' \mathbf{X}_t^s + \mathbf{b}_2' \mathbf{X}_t^s \Lambda(Z_t, \mathbf{q}) + e_t^s \quad (7)$$

donde $\Lambda(Z_t, \mathbf{q})$ es una función de transición que permite al modelo cambiar del estado $E[W_t^s | \mathbf{X}_t^s] = \exists_1' \mathbf{X}_t^s$ a $E[W_t^s | \mathbf{X}_t^s] = (\exists_1 + \exists_2)' \mathbf{X}_t^s$ a través de las variables Z_t . Si $Z_t = t$ y $\Lambda = \mathbf{j}_{[t_0, t_1]}$ obtenemos el caso de cambio discreto, pero otro tipo de funciones flexibilizará el modelo de cambio de un régimen a otro (ver Lin y Teräsvirta (1994) para un análisis de diferentes elecciones). En lo que respecta a los parámetros del mecanismo de corrección del error¹², y en previsión de que en el corto plazo puedan existir inestabilidades importantes, formulamos un modelo que permite adaptarse a cualquier tipo de cambio (brusco o suave) que pueda producirse. En este sentido, el modelo determinista de variables ficticias resulta *a priori* demasiado rígido, siendo más conveniente un modelo estocástico¹³ que permita mayor flexibilidad en la evolución temporal de los parámetros. Reespecificando el modelo [3] de tal forma que en el vector \forall aparezcan todos los parámetros $((\alpha_0, (\alpha_{11}, \dots, \alpha_{1p}), (\alpha_{20}, \dots, \alpha_{2q}), \dots, \forall)$ y en el vector \mathbf{H}^s todas las variables explicativas¹⁴ del mismo, la estructura de las ecuaciones del modelo que se propone son del tipo

$$\begin{aligned} \Delta W_t^s &= \mathbf{a}_t' \mathbf{H}_t^s + u_t^s \\ \mathbf{a}_t &= \mathbf{a}_{t-1} + \mathbf{h}_t \end{aligned} \quad (8)$$

donde suponemos que los errores u_t^s se distribuyen normalmente con media cero y varianza constante Φ^2 y son independientes entre sí, y que \mathbf{h}_t es un vector de variables aleatorias normales con media cero y matriz de covarianzas $\Phi^2 \mathbf{P} = \mathbf{Q}$ cuya distribución es independiente de la de los errores u_t^s y de la correspondiente al vector \forall_0 . La primera ecuación del sistema [8] se conoce como ecuación de medida y la segunda como ecuación de transición, la cual describe la evolución temporal del vector parámetros de interés, \forall_t , que ahora se denomina vector de estados (y a sus componentes variables de estado). En nuestra aplicación supondremos que la matriz \mathbf{Q} (conocida como matriz de dispersión) es diagonal, es decir, no se permite a las variables de estado interactuar entre sí, para lo cual deberían aparecer elementos no nulos fuera de la diagonal. Lógicamente, en el caso en que $\mathbf{Q} = \mathbf{0}$ se obtiene el modelo [3] de parámetros constantes.

La especificación [8] supone que el vector de parámetros \forall_t sigue una distribución multivariante tipo paseo aleatorio (sin deriva) y, al ser no estacionario, evoluciona en el tiempo de forma tal que puede acomodar todos los cambios estructurales que hayan tenido lugar durante el período muestral. Obviamente, pueden proponerse otros modelos estocásticos para \forall_t dependiendo del nivel de información a

puede generalizarse al caso de más de un cambio estructural.

¹² En el trabajo seguiremos una estrategia de estimación bietápica (Engle y Granger, 1987) en la que en la primera etapa se estima la relación de cointegración, y posteriormente se estima el mecanismo de corrección de error introduciendo los residuos (retrasados) de la relación a largo plazo estimada en el primer paso.

¹³ También podría utilizarse un modelo del tipo STR tal como se hace en Wolters *et al.* (1996).

¹⁴ Substituyendo el vector de cointegración $W_{t-1}^s - \exists = X_{t-1}^s$ por el valor de los residuos estimados.

priori que se posea sobre la forma, fechado y velocidad del cambio estructural¹⁵. En nuestro caso, dada la carencia de dicha información¹⁶, hemos preferido utilizar como hipótesis alternativa (por otro lado habitual) un modelo de paseo aleatorio.

4. RESULTADOS EMPÍRICOS

4.1. Fuentes estadísticas y propiedades estocásticas de los datos

La mayor parte de las variables utilizadas en el trabajo se han obtenido de las bases de datos HISPALINK¹⁷ y Cordero y Gayoso (1997). La primera es una base histórica que recoge datos para el período 1970-1985 (ver Hispalink, 1993) mientras que la segunda abarca desde 1980 hasta 1995 con datos oficiales del Instituto Nacional de Estadística y utilización de deflatores propios. En concreto, todas las variables referentes al valor añadido bruto (por sectores y a nivel regional y nacional) se han obtenido de dichas bases. El resto se ha construido a partir de distintas fuentes de información estadística regionales o nacionales¹⁸. Los datos utilizados comprenden, en general, desde 1970 hasta 1995, aunque en algunos sectores la longitud es más reducida debido a la inexistencia de series desagregadas para varios años al principio del período considerado. Uno de los objetivos del trabajo consiste en analizar el funcionamiento predictivo post-muestral del modelo, por lo que hemos reservado los dos últimos años de las series (los años 1994-1995) para hacer predicciones *ex-post* para dichos años. Por lo tanto, el período muestral que se utiliza en la estimación de los epígrafes siguientes es el comprendido, en general, entre 1970 y 1993.

Como paso previo a la estimación de la ecuaciones del modelo, necesitamos analizar el orden de integrabilidad (*d*) de todas las variables endógenas y exógenas que aparecerán en el modelo. Para tal fin se ha utilizado el test de Dickey-Fuller aumentado (*ADF*) (Dickey y Fuller, 1981). No obstante, y teniendo en cuenta experiencias previas sobre el estudio de las propiedades estocásticas de las series macroeconómicas españolas (Andrés et al. 1990; Molinas *et al.*, 1991), así como la propia naturaleza de las series regionales (con frecuentes puntos de ruptura¹⁹), se ha considerado también la versión más general del test *ADF* para tener en cuenta la posibilidad de existencia de tendencias determinísticas segmentadas en la media (Rappoport y Reichlin, 1989). Los resultados de aplicar los contrastes antes citados a las variables objeto de estudio se muestran en la

¹⁵ Ver al respecto el apartado 4 de Hall (1994).

¹⁶ Y el reducido número de observaciones, que limita cualquier tipo de generalización.

¹⁷ En Otero *et al.* (1996) se discuten el origen, contenido y metodología utilizada en la construcción de esta base de datos.

¹⁸ En Ramajo y Márquez (1996) se hace un análisis detallado de las distintas fuentes de información que han sido utilizadas en mayor o menor medida en el trabajo.

¹⁹ Debidos no sólo a cambios estructurales como los analizados en este trabajo, sino también a problemas “en origen” tales como cambios de base, redefiniciones de variables, errores de medida causados por la aplicación de métodos de reparto, utilización de deflatores aproximados, etc.

Tabla 1, donde también se presenta un breve resumen estadístico de las mismas²⁰. La conclusión a la que se llega es que todas las series pueden considerarse como variables $I(1)$, algunas de ellas con una única tendencia determinística y el resto de variables con varias tendencias segmentadas en la media.

4.2. Estimación del modelo

Siguiendo la filosofía del método bietápico propuesto por Engle y Granger (1987), en primer lugar se han estimado las relaciones a largo plazo del tipo [4] utilizando los argumentos teóricos esbozados en la sección dos. En todos los casos se efectuó el contraste de cointegración de Engle y Granger (que emplea el test de raíces unitarias de Dickey-Fuller, 1979) para determinar si las variables implicadas en la regresión estaban cointegradas y el resultado fue que en la mayor parte de los tests no se rechazaba la hipótesis nula de existencia de una raíz unitaria en los residuos estimados (es decir, de ausencia de cointegración).

Ante esta evidencia, y teniendo en cuenta los resultados de Campos *et al.* (1996) (los cuales ponen de manifiesto no sólo que la presencia de cambio estructural en series estacionarias puede inducir raíces unitarias espúreas²¹, sino que tales rupturas afectan de forma considerable a la potencia de los contrastes de cointegración en general, y en particular al procedimiento bietápico de Engle y Granger²²) se procedió a aplicar una serie de contrastes de estabilidad paramétrica para detectar la presencia de puntos de ruptura en cada una de las ecuaciones de comportamiento a largo plazo²³. En concreto, se han aplicado varios tests basados en el cálculo de estadísticos de Wald secuenciales, $F_T(\delta)$, los cuales contrastan la hipótesis nula de estabilidad frente a la alternativa de existencia de algún punto de ruptura en la observación t_0 (más concretamente, en la fracción $\delta_0=t_0/T$ de la muestra). La localización exacta de los posibles puntos de corte no es conocida *a priori*, por lo que se calculan los estadísticos $F_T(t/T)$ para todos los puntos de la muestra²⁴ y a continuación se construye algún funcional de los mismos. Los tres funcionales considerados²⁵ en nuestro trabajo son el estadístico de razón de verosimilitud

²⁰ Por razones de espacio no se presenta el detalle de cada una de las regresiones efectuadas.

²¹ Resultado que ya fue demostrado analíticamente por Perron (1989) y evidenciado empíricamente por Hendry y Neale (1991).

²² Hay que tener en cuenta que los contrastes estándar de cointegración asumen que el vector de cointegración es invariante en el tiempo bajo la hipótesis alternativa.

²³ De hecho, al aplicar los contrastes de cointegración propuestos por Gregory y Hansen (1996), los cuales permiten la posibilidad de cambios de régimen, el resultado fue que en todos los casos se rechazó la hipótesis nula de no cointegración (en este caso frente a la alternativa de cointegración en presencia de un posible punto de ruptura). No obstante, puesto que los contrastes de Gregory y Hansen sólo permiten un punto de ruptura (el procedimiento que utilizan permite además identificar tal punto), y ante la posibilidad de que exista un número mayor en nuestro caso, se ha continuado el proceso de modelización del cambio estructural.

²⁴ Tal como señala Andrews (1993), no se pueden utilizar todos los puntos $t/T \in [0,1]$, puesto que en ese caso los contrastes divergerán hacia infinito, por lo que propone utilizar la región $T=[*_1,*_2]=[0.15,0.85]$.

²⁵ Las distribuciones asintóticas de los mismos se discuten en Andrews (1993) y Andrews y

propuesto por Quandt (1960)

$$SupF = \sup_{\mathbf{d} \in [\mathbf{d}_1, \mathbf{d}_2]} F_T(\mathbf{d}) \quad (13)$$

el estadístico promedio propuesto por Andrews y Ploberger (1994) y Hansen (1992)

$$MeanF = \int_{\mathbf{d}_2}^{\mathbf{d}_1} F_t(\mathbf{d}) d\mathbf{d} \quad (14)$$

y el estadístico promedio exponencial propuesto por Andrews y Ploberger (1994)

$$ExpF = \ln \left\{ \int_{\mathbf{d}_2}^{\mathbf{d}_1} \exp \left(\frac{1}{2} F_t(\mathbf{d}) \right) d\mathbf{d} \right\} \quad (15)$$

Al aplicar los tres contrastes señalados a las ecuaciones a largo plazo del modelo, en todos los casos, salvo en uno los valores de dichos tests, sobrepasaron los valores críticos correspondientes a un nivel de significación del 1% (el único caso donde no se superaron dichos valores se rechazó la hipótesis nula al nivel del 5%).

El siguiente paso consistió en introducir las variables ficticias necesarias para aproximar los cambios estructurales detectados a través de la aplicación secuencial de los estadísticos de Wald²⁶, es decir, se estimaron relaciones de cointegración del tipo [6], con tantas funciones $v_{[t_0, t_1]}$ como puntos de ruptura detectados. Los resultados obtenidos para cada uno de los sectores considerados se presentan en la Tabla 2. Haremos a continuación algunos comentarios al respecto de dichos resultados. En primer lugar, para cada una de las ecuaciones estimadas se aplicaron de nuevo los contrastes de estabilidad *SupF*, *MeanF* y *ExpF* y el resultado fue que en ningún caso se rechazó la hipótesis nula de estabilidad de los parámetros estimados. En segundo lugar, tal como puede comprobarse en la tabla, el estadístico *DF* rechaza en todos los casos la presencia de una raíz unitaria en los errores estimados de cada ecuación, es decir, las combinaciones lineales de las variables de cada modelo son estacionarias y, por tanto, se pueden interpretar dichas relaciones como ecuaciones de cointegración o equilibrio a largo plazo con cambios de régimen. En tercer lugar, se observa que el número de puntos de ruptura es reducido²⁷, con a lo sumo tres cambios estructurales por ecuación (en el caso de los sectores industrial, de transportes y comunicaciones y de servicios no destinados a la venta). Por otro lado, en varios casos se presentan simultáneamente cambios en el nivel y en las pendientes del modelo, con dos sectores (los de servicios destinados y no destinados a la venta) en los que se detectaron regímenes distintos, con dos puntos de corte con cambio tanto en la pendiente como en la ordenada en el origen.

Una vez estimadas las relaciones a largo plazo, el segundo paso del procedimiento de Engle y Granger (1987) consiste en estimar las ecuaciones a corto

Ploberger (1994), siendo en todos los casos no estándar (funcionales de movimientos Brownianos multidimensionales).

²⁶ Somos conscientes de algunos de los problemas que este enfoque puede plantear. En primer lugar, la estimación por MCO no es eficiente y los contrastes de significación no tienen las distribuciones asintóticas estándar bajo la hipótesis de cointegración con cambios de régimen. En segundo lugar, bajo la hipótesis nula de estabilidad en los parámetros, y puesto que los puntos de ruptura son desconocidos *a priori*, los estadísticos de Wald construidos tienen también distribuciones no estándar. Los resultados de Hansen (1992) y Quintos y Phillips (1993) podrían ser de utilidad en la resolución de estos problemas.

²⁷ Lo que en cierto modo justifica el enfoque de variables ficticias, por su fácil implementación frente a otras alternativas más elaboradas.

plazo, que vienen dadas por [3], con la expresión entre paréntesis (que ahora sería del tipo [6] con varias funciones $v_{[t_0, t_1]}$) substituida por lo errores estimados que se derivan de la Tabla 2. Igual que para el modelo a largo plazo, en primer lugar se estimaron las ecuaciones individuales sin incluir variables ficticias y, a continuación, se llevaron a cabo los contrastes de estabilidad *SupF*, *MeanF* y *ExpF*. Excepto en el caso del sector de transportes y comunicaciones, en el resto de sectores los tres estadísticos utilizados rechazaron de forma simultánea la hipótesis nula de estabilidad de los parámetros de los mecanismos de corrección de error. Para tener en cuenta la presencia de cambio estructural se consideraron dos opciones. Una de ellas fue utilizar, igual que en largo plazo, variables ficticias para recoger el efecto de dichos cambios estructurales. La otra consistió en modelizar la ruptura estructural proponiendo como alternativa un modelo adaptativo como el representado por las ecuaciones [8]. Ambas opciones son desarrolladas a continuación. La información proporcionada por los estadísticos de Wald utilizados en los contrastes de estabilidad (complementada con un análisis gráfico de cada variable dependiente y de los residuos estimados del modelo inicial) ayudó a identificar las variables ficticias a introducir en cada una de las ecuaciones de comportamiento a corto plazo. El resultado final son las regresiones que se presentan en la Tabla 3. En lo referente al número de variables ficticias introducidas, cabe hacer dos anotaciones. En primer lugar, el número de variables ficticias de cada ecuación viene determinado por los valores de los contrastes de estabilidad; con las variables introducidas (y sólo con esas) se consigue que los mecanismos de corrección sean estables, en el sentido de que ningún estadístico de los tres propuestos supera el nivel crítico correspondiente. En segundo lugar, y como cabía esperar, el número de puntos de ruptura que aparece es mucho más elevado que en las relaciones a largo plazo, señalando la presencia de mayor inestabilidad en las relaciones de corto plazo que en la ecuaciones de equilibrio.

Por lo que respecta a la estimación de las ecuaciones [8] para cada uno de los siete sectores considerados²⁸, la técnica de estimación que se ha utilizado está basada en la aplicación recursiva del filtro de Kalman (Kalman, 1960). En cada instante t de tiempo, y dadas las observaciones (W_1^s, \dots, W_t^s) , el interés se centra en estimar el vector ∇_t usando las ecuaciones del modelo [8]. Bajo las hipótesis de normalidad establecidas para los errores del modelo (y para el vector de estado inicial ∇_0), y suponiéndose conocidos Φ^2 , \mathbf{P} (o \mathbf{Q}) y la media y matriz de covarianzas de ∇_0 ²⁹, el estimador óptimo de ∇_t usando la información I_s disponible hasta el instante s , viene dado por la esperanza condicional de ∇_t supuesto conocido I_s , a la que denotaremos por $E[\nabla_t | I_s] = \mathbf{a}_{t|s}$ ³⁰; y el estimador óptimo para la matriz de covarianzas de ∇_t usando la información disponible I_s vendrá dado por $Cov[\nabla_t | I_s] = \mathbf{E}_{t|s}$. Las recursiones del filtro de Kalman, para $t=0,1,2,\dots$ vienen dadas por las siguientes ecuaciones (ver, por ejemplo, Lühtkepohl, 1993):

²⁸ Aunque en principio no era necesario reestimar el modelo correspondiente al sector de transportes y comunicaciones, por ser estable en el corto plazo, también se incluyó en esta fase al objeto de comparar los resultados entre el modelo con parámetros fijos y el de parámetros variables.

²⁹ A las que denotaremos por $\mathbf{a}_{0|0}$ y $\mathbf{E}_{0|0}$, respectivamente.

³⁰ Cuando $s=t$ la evaluación de $\mathbf{a}_{t|t}$ se conoce como filtrado, y cuando $s>t$ se conoce como suavizado.

$$\begin{aligned}
\mathbf{a}_{t+1/t} &= \mathbf{a}_{t/t} \\
\Sigma_{t+1/t} &= \Sigma_{t/t} + \mathbf{Q} \\
\mathbf{a}_{t+1/t+1} &= \mathbf{a}_{t+1/t} + k_{t+1} [\Delta W_{t+1}^s - \mathbf{a}_{t+1/t}^{\prime} \mathbf{H}_{t+1}^s]
\end{aligned} \tag{16}$$

$$\begin{aligned}
\Sigma_{t+1/t+1} &= \Sigma_{t+1/t} - k_{t+1} \mathbf{H}_{t+1}^s \Sigma_{t+1/t} \\
\text{donde } k_{t+1} &= [\mathbf{H}_{t+1}^s \Sigma_{t+1/t} \mathbf{H}_{t+1}^{\prime} + \mathbf{s}^2]^{-1} \Sigma_{t+1/t} \mathbf{H}_{t+1}^{\prime}
\end{aligned} \tag{17}$$

Y entonces la predicción n -períodos hacia adelante de W_{t+n}^s , $W_{t+n|t}^s$ vendrá dada por

$$\Delta W_{t+n|t}^s = \mathbf{a}_{t'/t}^{\prime} \mathbf{H}_{t+n}^s \tag{18}$$

En la aplicación de las fórmulas anteriores existen varios problemas que es necesario explicitar y que hacen referencia al conjunto de parámetros que se asume son conocidos *a priori*. En concreto, puesto que se ha especificado un modelo de variación paramétrica tipo “paseo aleatorio” para \forall_t , no existen valores automáticos (como la media o matriz de covarianzas incondicionales) para los valores $\mathbf{a}_{0|0}$ y $\mathbf{E}_{0|0}$. Por otro lado, tampoco se conocen ni los elementos de la matriz \mathbf{Q} ni el parámetro Φ^2 . Este último es el menos problemático puesto que no existen inconvenientes para estimarlo por máxima verosimilitud aislándolo del resto de parámetros (Chow, 1984, pág. 1222). Con respecto a los valores de inicialización del filtro de Kalman se pueden emplear las primeras K observaciones -siendo K la dimensión del vector de estado- (Harvey, 1981, 1989), un valor *a priori* “difuso” (Ansley y Kohn, 1983) o se pueden estimar por máxima verosimilitud junto con el resto de parámetros del modelo (Chow, 1984). En nuestro caso hemos utilizado otra alternativa, fijando [tal como se hace en Hackl y Westlund (1996)] los elementos $\mathbf{a}_{0|0}$ y $\mathbf{E}_{0|0}$ en los valores MCO obtenidos estimando el modelo con parámetros constantes (y sin variables ficticias) sobre el período muestral completo³¹.

En lo referente a los elementos de la matriz \mathbf{Q} (conocidos como hiperparámetros) se pueden seguir dos vías. Estimarlos por el método de máxima verosimilitud (Chow, 1984), o fijarlos de antemano tal como proponen, de nuevo, Hackl y Westlund (1996), para evitar problemas de falta de identificación y grandes oscilaciones en las estimaciones de los parámetros de las variables de estado. En nuestro caso, hemos seguido las dos opciones: estimar los elementos q_{ii} de la matriz \mathbf{Q} por máxima verosimilitud; y fijar de antemano su valores, en nuestro caso tomando $\mathbf{Q}=\mathbf{I}$ ³².

4.3. Análisis predictivo ex-post

En este epígrafe comparamos el funcionamiento predictivo post-muestral de los

³¹ Reconocemos que, supuesto que se decide fijar los valores iniciales en lugar de estimarlos, cualquiera de los otros dos métodos de inicialización considerados es más ortodoxo y correcto que el utilizado por nosotros. En uno de los casos, el reducido tamaño de la muestra nos ha impedido prescindir de algunas observaciones al inicio del período muestral. En el segundo, creímos conveniente dar un valor inicial incorporando información *a priori* debido a que con tan pocas observaciones un valor de inicialización difuso podría originar trayectorias con grandes fluctuaciones de período a período originadas por una mala elección del punto inicial.

³² Wolff (1987) considera un rango de matrices del tipo $\mathbf{Q}=(\mathbf{E}_{0|0})$ para $(=0.0, 0.01, 0.05, 0.10$ y 0.25 permitiendo mayor ($=0.25$) o menor ($=0.0$) variabilidad en las variables de estado e interacción entre las mismas.

modelos a corto plazo estimados en el apartado anterior. Tal como se apuntó en el epígrafe 4.1, todos los modelos fueron estimados (en general) para el período 1970-1993, dejando los años 1994 y 1995 como período de prueba para la realización de un experimento de predicción ex-post. Por tanto, las predicciones realizadas con los modelos estructurales están basadas en valores reales de las variables explicativas, y se comparan dichas predicciones con los valores observados de las variables endógenas para los años considerados.

Para medir el grado de bondad de las predicciones utilizamos cuatro conocidos estadísticos, todos ellos basados en funciones de pérdida simétricas. En concreto, se calculan el error medio (ME), el error absoluto medio (MAE), el error cuadrático medio (RMSE) y el coeficiente de desigualdad (U) de Theil (1966)³³. Los resultados de la simulación ex-post realizada con los tres tipos de modelos utilizados ($\mathbf{Q}=\mathbf{0}$, es decir, el modelo con parámetros fijos y variables ficticias; $\mathbf{Q}=\mathbf{I}$, es decir, fijando de antemano los hiperparámetros; y, $\mathbf{Q}=\text{diag}\{q_{11}, q_{22}, \dots, q_{KK}\}$, es decir, estimando -por máxima verosimilitud- los elementos de la diagonal pero restringiendo el resto a cero) se presentan en la Tabla 4. Merece la pena destacar algunos comentarios acerca de las cifras que aparecen en la misma. En primer lugar, de la comparación de los valores de los estadísticos se observa que no existe un claro dominio de los modelos con parámetros variables sobre el modelo con parámetros fijos, mejorando los resultados respecto a este último en tres de los siete sectores considerados y produciendo resultados similares en otros dos casos. En segundo lugar, el modelo en que se estiman los hiperparámetros (\mathbf{Q} diagonal) produce en general peores resultados que el modelo donde éstos se fijan de antemano. Ello puede deberse a los problemas de identificación de tales parámetros ocasionado por el reducido tamaño de la muestra. Por último, con respecto al estadístico U de Theil (1966), sólo en el caso del sector de la energía los tres modelos predicen sistemáticamente peor que el modelo “ingenuo” de paseo aleatorio. En el resto de sectores los modelos estimados funcionan en general mejor que dicho modelo ingenuo aunque, de nuevo, se observa que el modelo de parámetros variables donde se estiman los elementos q_{ii} de la diagonal de la matriz de dispersión \mathbf{Q} produce peores resultados que los otros dos modelos propuestos.

5. CONCLUSIONES

En este trabajo se ha presentado el bloque de producción de un modelo econométrico uniregional, descendente, abierto, de base económica dinámica (porque la especificación dinámica parte de la consideración de la existencia de sectores básicos y no básicos en la economía extremeña) y con fines predictivos para la economía extremeña.

La estabilidad de un modelo econométrico es un requisito básico para que éste pueda utilizarse con fines predictivos o inferenciales. A escala regional, sin embargo, la inestabilidad paramétrica de los modelos estimados con datos históricos suele ser bastante común, debido a que a ese nivel son frecuentes los “cambios de régimen”. Por este motivo, los experimentos predictivos que se realicen tendrán sentido únicamente si se cumple la condición (necesaria) de estabilidad. Existe una literatura bastante amplia

³³ El estadístico U utilizado es el ratio de los RMSE de las predicciones obtenidas con el modelo estimado y con el modelo (ingenuo) de paseo aleatorio. Este estadístico tiene una fácil interpretación: si $U < 1$ las predicciones del modelo son mejores que las predicciones “ingenuas” y si $U > 1$ el modelo de paseo aleatorio funciona mejor que el propuesto.

sobre el tema de la contrastación de la hipótesis de estabilidad paramétrica (ver, por ejemplo, Stock y Watson (1996) y las referencias allí contenidas), pero las aportaciones son mucho más escasas sobre la cuestión de cómo modelizar el cambio estructural una vez que éste es detectado. Con frecuencia, un contraste de estabilidad significativo indica algún tipo de mala especificación, por lo que se procede a probar con especificaciones alternativas. En otros casos, se permite a los parámetros variar a lo largo de la muestra, siendo el método utilizado más común la introducción de variables ficticias que interaccionan con las variables originales para permitir cambios en las pendientes y/o en la ordenada en el origen. Como alternativa, también se puede dejar variar aleatoriamente a los coeficientes de regresión, dotando al modelo resultante de un mayor nivel de flexibilidad frente al caso de las variables ficticias. En este trabajo hemos utilizado ambos enfoques para construir un modelo econométrico para una región española, Extremadura, la cual ha experimentado profundos cambios en su estructura económica desde comienzos de los años ochenta. El marco econométrico de partida ha sido la teoría de la cointegración, que distingue las relaciones económicas en el largo plazo de la dinámica a corto plazo mediante la introducción de modelos de corrección de error, que son la base analítica del modelo econométrico construido en este trabajo.

Los resultados obtenidos muestran, en primer lugar, que en el caso de Extremadura existe inestabilidad estructural tanto en el largo como en el corto plazo, por lo que la metodología econométrica estándar resulta inaplicable en nuestro caso. En el largo plazo, la introducción de variables ficticias es suficiente para recoger los cambios de régimen habidos durante el período muestral, pero en el corto plazo la inestabilidad es mucho mayor por lo que parece más razonable flexibilizar el modelo de cambio permitiendo a los parámetros variar de forma aleatoria. En el análisis realizado, el modelo de parámetros variables proporciona resultados similares a los del modelo con parámetros fijos y variables ficticias, pero consideramos que el primero posee la ventaja de incorporar en las predicciones la incertidumbre futura sobre los valores de los parámetros, cosa que no ocurre con el modelo con variables ficticias.

TABLA 1. Resultados de los contrastes de raíces unitarias.

CONTRASTES DE RAÍCES UNITARIAS													
	$H_0: \{I(2)\}$ frente a $H_1: \{I(1)\}$						$H_0: \{I(1)\}$ frente a $H_1: \{I(0)\}$						
	DFA			PP			DFA			PP			
VARIABLE	PGD	t	VC	PGD	t	VC	PGD	t	VC	PGD	t	VC	Resultado
LVAES	CT,0	-5,49	-3,61	CT,2	-5,70	-3,61	C,0	-2,46	-2,98	C,2	-2,43	-2,98	I(1)
LVAEX	N,1	-5,58	-1,95	N,2	-12,59	-1,95	C,0	-3,08	-3,73*	C,2	-2,93	-2,99	I(1)
LVOLAG	N,0	-5,95	-1,95	N,2	-5,95	-1,95	CT,0	-2,52	-3,60	CT,2	-2,63	-3,60	I(1)
LVEES	CT,0	-5,21	-3,61	CT,2	-5,37	-3,61	C,0	-2,69	-2,98	C,2	-3,46	-3,72*	I(1)
LVEEX	N,0	-6,17	-1,95	N,2	-6,20	-1,95	CT,0	-2,44	-3,60	CT,2	-2,47	-3,60	I(1)
LPRECIP	N,2	-4,58	-1,95	N,2	-8,58	-1,95	C,0	-4,20	-2,98	C,2	-4,27	-2,98	I(0)
LENERG	RR,2	-6,58	-4,08	-	-	-	RR,2	-2,48	-4,08	-	-	-	I(1)
LVBEX	C,0	-5,59	-2,99	C,2	-6,83	-2,99	CT,0	-2,72	-3,60	CT,2	-2,48	-3,60	I(1)
LVABES	C,1	-3,09	-2,99	C,2	-2,09	-2,99	CT,1	-3,93	-4,39*	CT,2	-2,79	-3,61	I(1)
LVABEX	C,0	-5,17	-3,00	C,2	-8,96	-3,00	CT,0	-2,40	-3,62	CT,2	-2,09	-3,62	I(1)
LVIEX	N,0	-5,21	-1,95	N,2	-5,27	-1,95	C,0	-2,42	-2,98	C,2	-3,56	-3,72*	I(1)
LVIES	N,0	-3,20	-1,95	N,2	-3,42	-1,95	CT,1	-2,77	-3,62	CT,2	-2,66	-3,60	I(1)
LVNOAE	C,0	-5,07	-3,00	C,2	-5,02	-3,00	CT,0	-1,95	-3,62	CT,2	-2,39	-3,62	I(1)
LVLEX	N,1	-2,35	-1,95	N,2	-4,07	-1,95	RR,3	-4,26	-4,76	-	-	-	I(1)
LVLES	RR,2	-4,14	-4,08	-	-	-	CT,1	-2,85	-3,62	CT,2	-2,21	-3,62	I(1)
LVZEX	N,0	-3,11	-1,95	N,2	-3,10	-1,95	CT,1	-2,28	-3,62	CT,2	-2,19	-3,61	I(1)
LVZES	RR,2	-4,39	-4,08	-	-	-	CT,1	-3,33	-3,62	CT,2	-3,16	-3,61	I(1)
LVGEX	RR,2	-7,55	-4,08	-	-	-	RR,2	-3,72	-4,08	-	-	-	I(1)
LVGES	RR,2	-4,16	-4,08	-	-	-	CT,1	-2,86	-3,61	CT,2	-1,46	-3,60	I(1)

Notas:

-La notación utilizada para representar a las variables es:

L indica que la variable está en logaritmos,

V denota el valor añadido bruto en pesetas constantes de 1986. **V** se acompaña de una sola letra para indicar el sector al que hace referencia (**VA**: agrario, **VE**: energía, **VB**: construcción, **VI**: industria manufacturera, **VL**: servicios destinados a la venta (excluido transportes y comunicaciones), **VZ**: transportes y comunicaciones y **VG**: servicios no destinados a la venta); o de varias letras para cada una de las agrupaciones sectoriales consideradas (**VAB**: valor añadido bruto total, **VNOAE**: valor añadido bruto total no agrario y no energético).

La terminación final para cada variable, especifica si se trata de una variable nacional (**ES**) o extremeña (**EX**).

VOLAG es el volumen de agua embalsada al final de cada año en Extremadura.

PRECIP es el volumen medio de precipitaciones recogidas en Extremadura.

ENERG es la producción bruta de energía eléctrica en Extremadura.

-La columna **PGD** especifica el proceso generador de datos que se ha considerado para la variable en estudio; de esta forma, la letra **N** indica que la regresión auxiliar estimada no incorpora ninguna componente determinística, **C** señala la admisión de un término constante y **CT** denota la presencia de una tendencia lineal determinística. El número a continuación de estas letras (separado por coma) señala el número de retardos introducidos en el contraste DFA o el número de períodos de correlación serial a incluir (retardos de truncamiento) en el procedimiento de Newey-West para ajustar los errores estándar (en el contraste PP). En el caso de que aparezcan las letras **RR**, se advierte de la consideración de tendencias determinísticas segmentadas en la media para el contraste DFA (tantos segmentos como se indique en la cifra).

-En la columna siguiente, t es el estadístico de prueba utilizado en cada uno de los contrastes.

-**VC** es el valor crítico tabulado a un nivel de significación del 5% para cada uno de los contrastes obtenidos de MacKinnon (1991) o de Rappoport y Reichlin (1989). En el caso de trabajar al 1% se utiliza * al lado de la cifra correspondiente.

TABLA 2. Estimación de las relaciones de cointegración (MCO con variables ficticias).

AGRICULTURA				
ECUACIÓN ESTIMADA; Método: MCO con variables ficticias. (t = 1971,...,1993)				
LVAEX _t = -2,431 + 0,951 LVAES _t + 0,231 D7175 _t				
(-0,422) (2,388) (2,685)				
CONTRASTES DE VALIDACIÓN				
R ² =0,279; Durbin-Watson=1,457; DFA = -3,191*; SupF = 7,784; MediaF = 2,041; ExpF = 1,797				
ENERGIA				
ECUACIÓN ESTIMADA; Método: MCO con variables ficticias. (t = 1970,...,1993)				
LVEEX _t = -5,070 + 1,050 LVEES _t + 0,561 D7073 _t + 1,087D8493 _t				
(-0,743) (2,193) (2,652) (6,014)				
CONTRASTES DE VALIDACIÓN				
R ² =0,899;Durbin-Watson=2,153;DFA= -4,747***;SupF= 7,090;MediaF= 2,326; ExpF= 1,767				
CONSTRUCCIÓN				
ECUACIÓN ESTIMADA; Método: MCO con variables ficticias. (t = 1972,...,1993)				
LVBEX _t = -2,366 + 1,011 LVABEX _t - 19,077 D7079 _t + 1,446 D7079*LVABEX _t				
(-1,154) (6,547) (-2,545) (2,511)				
CONTRASTES DE VALIDACIÓN				
R ² =0,941;Durbin-Watson=1,829;DFA = -4,046***;SupF= 7,948; MediaF= 1,745; ExpF= 1,618				
INDUSTRIA MANUFACTURERA				
ECUACIÓN ESTIMADA ; Método: MCO con variables ficticias. (t = 1970,...,1993)				
LVIEX _t = -7,994 + 1,182 LVIES _t + 0,206 F78 _t + 0,223 F80 _t + 0,104 D8285 _t				
(-5,158) (12,039) (2,551) (2,767) (2,400)				
CONTRASTES DE VALIDACIÓN				
R ² =0,894; Durbin-Watson=1,995;DFA = -4,476***;SupF= 5,472; MediaF= 3,612; ExpF= 1,937				
SERVICIOS DESTINADOS A LA VENTA				
ECUACIÓN ESTIMADA ; Método: MCO con variables ficticias. (t = 1972,...,1993)				
LVLEX _t = -5,366 + 1,083 LVLES _t + 43,937 D8185 _t - 2,724 D8185*LVLES _t - 11,327 D8693 _t				
(-2,016) (6,556) (5,657) (-5,673) (-3,192)				
+ 0,682 D8695*LVLES _t				
(3,118)				
CONTRASTES DE VALIDACIÓN				
R ² =0,962;Durbin-Watson=2,351;DFA= -5,586***;SupF= 9,726*; MediaF= 2,983; ExpF= 2,464				
TRANSPORTES Y COMUNICACIONES				
ECUACIÓN ESTIMADA; Método: MCO con variables ficticias. (t = 1972,...,1993)				
LVZEX _t = -5,269 + 0,722 LVZES _t + 0,399 LVLEX _t + 0,257 F72 _t + 17,686 D8084 _t				
(-6,491) (16,269) (4,422) (7,038) (2,505)				
- 1,232 D8084*LVZES _t - 0,131 D8889 _t				
(-2,493) (-5,414)				
CONTRASTES DE VALIDACIÓN				
R ² =0,980;Durbin-Watson= 1,988; DFA = -4,335***; SupF=7,559; MediaF= 3,168; ExpF= 2,224				
SERVICIOS NO DESTINADOS A LAVENTA				
ECUACIÓN ESTIMADA; Método: MCO con variables ficticias. (t = 1971,...,1993)				
LVGEX _t = -3,636 + 1,004 LVGES _t + 3,645 D7078 _t - 0,257 D7078*LVGES _t - 10,452 D7985 _t				
(-4,066) (17,219) (3,447) (-3,683) (-7,288)				
+ 0,689 D7985*LVGES _t + 0,024 D9293 _t				
(7,289) (2,225)				
CONTRASTES DE VALIDACIÓN				
R ² =0,998;Durbin-Watson=2,682; DFA = -6,395***; SupF= 2,603; MediaF= 0,513; ExpF= 0,340				

Notas: a) Debajo de los coeficientes estimados aparecen, entre paréntesis (y sólo como medidas descriptivas), los estadísticos *t*. b) * denota el rechazo de la hipótesis nula al 10%, ** al 5% y *** al 1%. c) D7175_t denota a una variable ficticia que toma el valor 1 en los años 1971-1975, y 0 en el resto; y F78_t toma 1 en 1978 y 0 en el resto.

TABLA 3. Estimación de los modelos de corrección de error (MCO con variables ficticias).

AGRICULTURA
ECUACIÓN ESTIMADA (t = 1972,...,1993)
$DLVAEX_t = 0,053 - 0,923 \hat{u}_{t-1} + 1,005 DLVAES_t + 0,171 DLVOLAG_{t-1} + 0,214 DD7175_t$ <p>(4,358) (-8,426) (4,908) (3,337) (4,498)</p> $- 0,130 D7375_t - 0,400 F83_t - 0,197 D9293_t$ <p>(-4,386) (-8,353) (-5,057)</p>
CONTRASTES DE VALIDACIÓN
$R^2 = 0,998$; R^2 ajustado = 0,922; Durbin-Watson: DW = 2,214; Jarque-Bera: P-val. = 0,361; Breusch-Godfrey: a) [AR(1)] P-val: 0,221; b) [AR(2)] P-val: 0,356; Ljung-Box (p=6): P-val.= 0,829; ARCH: a) [ARCH(1)] P-val: 0,201; a) [ARCH(2)] P-val: 0,600; White: P-val. = 0,761; <i>SupF</i> = 7,279; <i>MeanF</i> = 3,918; <i>ExpF</i> = 2,350.
ENERGIA
ECUACIÓN ESTIMADA (t = 1972,...,1993)
$DLVEEX_t = -0,142 - 0,771 \hat{u}_{t-1} + 2,574 DLVEES_t + 0,407 DLPRECIP_t + 1,089 DLENERG_t$ <p>(-2,623)(-6,024) (2,876) (5,751) (10,128)</p> $- 0,307 D7476_t + 0,590 F77_t + 0,526 F82_t + 0,171 D8892_t$ <p>(-3,954) (4,343) (3,814) (2,697)</p>
CONTRASTES DE VALIDACIÓN
$R^2 = 0,951$; R^2 ajustado = 0,922; Durbin-Watson: DW = 2,384; Jarque-Bera: P-val. = 0,874; Breusch-Godfrey: a) [AR(1)] P-val: 0,083; b) [AR(2)] P-val: 0,073; Ljung-Box (p=6): P-val.= 0,065; ARCH: a) [ARCH(1)] P-val: 0,118; a) [ARCH(2)] P-val: 0,264; White: P-val. = 0,304; <i>SupF</i> = 10,621; <i>MeanF</i> = 5,270; <i>ExpF</i> = 3,757.
CONSTRUCCIÓN
ECUACIÓN ESTIMADA (t = 1973,...,1993)
$DLVBEX_t = 0,012 - 0,448 \hat{u}_{t-1} + 0,892 DLVABEX_t + 0,661 DLVABES_t + 0,143 DLVBEX_{t-1}$ <p>(2,559)(-7,792) (9,230) (4,604) (3,398)</p> $- 0,083 D7475_t + 0,066 D7677_t + 0,033 F81_t + 0,068 F83_t - 0,183 F85_t + 0,113 F86_t$ <p>(-9,965) (8,060) (2,333) (5,744) (-16,459) (8,532)</p> $- 0,250 F88_t - 0,050 F89_t$ <p>(-21,128) (-3,298)</p>
CONTRASTES DE VALIDACIÓN
$R^2 = 0,994$; R^2 ajustado = 0,985; Durbin-Watson: DW = 2,305; Jarque-Bera: P-val. = 0,746; Breusch-Godfrey: a) [AR(1)] P-val: 0,304; b) [AR(2)] P-val: 0,159; Ljung-Box (p=6): P-val.= 0,331; ARCH: a) [ARCH(1)] P-val: 0,629; a) [ARCH(2)] P-val: 0,253; White: P-val. = 0,521; <i>SupF</i> = 14,859; <i>MeanF</i> = 7,278; <i>ExpF</i> = 5,183.
INDUSTRIA MANUFACTURERA
ECUACIÓN ESTIMADA (t = 1973,...,1993)
$DLVIEX_t = -0,061 - 0,555 \hat{u}_{t-1} + 2,262 DLVNOAEEX_t + 0,231 DLVIEX_{t-1} + 0,171 DLVAEX_t$ <p>(-2,830)(-2,337) (5,103) (1,968) (2,107)</p> $+ 0,111 DF80_t + 0,128 DD8285_t + 0,142 F77_t + 0,152 F85_t$ <p>(2,352) (2,857) (2,084) (2,452)</p>
CONTRASTES DE VALIDACIÓN
$R^2 = 0,873$; R^2 ajustado = 0,788; Durbin-Watson: DW = 2,051; Jarque-Bera: P-val. = 0,564; Breusch-Godfrey: a) [AR(1)] P-val: 0,506; b) [AR(2)] P-val: 0,799; Ljung-Box (p=6): P-val.= 0,942; ARCH: a) [ARCH(1)] P-val: 0,308; a) [ARCH(2)] P-val: 0,446; White: P-val. = 0,611; <i>SupF</i> = 14,247; <i>MeanF</i> = 8,141*; <i>ExpF</i> = 5,723*.

SERVICIOS DESTINADOS A LA VENTA
ECUACIÓN ESTIMADA (t = 1973,...,1993)
$DLVLEX_t = 0,006 - 0,836 \hat{u}_{t-1} + 0,420 DLVNOAE_t + 0,924 DLVLES_t - 0,059 F77_t$ <p style="text-align: center;">(0,815)(-3,801) (2,253) (3,617) (-3,440)</p> $- 0,054 D8085_t - 0,036 F90_t$ <p style="text-align: center;">(-5,665) (-2,134)</p>
CONTRASTES DE VALIDACIÓN
$R^2 = 0,938$; R^2 ajustado = 0,911; Durbin-Watson: DW = 2,365; Jarque-Bera: P-val. = 0,542; Breusch-Godfrey: a) [AR(1)] P-val: 0,166; b) [AR(2)] P-val: 0,015; Ljung-Box (p=6): P-val.= 0,324; ARCH: a) [ARCH(1)] P-val: 0,957; a) [ARCH(2)] P-val: 0,166; White: P-val. = 0,157; <i>SupF</i> = 3,350; <i>MeanF</i> = 2,059; <i>ExpF</i> = 1,132.
TRANSPORTES Y COMUNICACIONES
ECUACIÓN ESTIMADA (t = 1974,...,1993)
$DLVZEX_t = 0,008 - 0,690 \hat{u}_{t-1} + 0,222 DLVLEX_t + 0,348 DLVZES_t + 0,245 DLVZEX_{t-1}$ <p style="text-align: center;">(1,123)(-4,123) (3,054) (1,967) (3,142)</p> $+ 20,551 DD8084_t - 0,082 DD8889_t - 1,435 D(D8084LVZES)_t$ <p style="text-align: center;">(7,676) (-7,614) (-7,658)</p>
CONTRASTES DE VALIDACIÓN
$R^2 = 0,944$; R^2 ajustado = 0,912; Durbin-Watson: DW = 1,756; Jarque-Bera: P-val. = 0,865; Breusch-Godfrey: a) [AR(1)] P-val: 0,741; b) [AR(2)] P-val: 0,184; Ljung-Box (p=6): P-val.= 0,781; ARCH: a) [ARCH(1)] P-val: 0,147; a) [ARCH(2)] P-val: 0,159; White: P-val. = 0,540; <i>SupF</i> = 5,486; <i>MeanF</i> = 2,539; <i>ExpF</i> = 1,580.
SERVICIOS NO DESTINADOS A LA VENTA
ECUACIÓN ESTIMADA (t = 1972,...,1993)
$DLVGEX_t = -0,001 - 0,701 \hat{u}_{t-1} + 0,974 DLVGES_t + 2,796 DD7078_t - 0,205 D(D7078LVGES)_t$ <p style="text-align: center;">(-0,194)(-7,689) (13,952) (4,335) (-4,797)</p> $- 17,353 DD7985_t + 1,141 D(D7985LVGES)_t + 0,027 DD9293_t + 0,015 F80_t - 0,040 F82_t$ <p style="text-align: center;">(-17,291) (14,593) (6,191) (3,112) (-7,621)</p> $- 0,041 F84_t + 0,021 F88_t$ <p style="text-align: center;">(-8,694) (4,808)</p>
CONTRASTES DE VALIDACIÓN
$R^2 = 0,991$; R^2 ajustado = 0,982; Durbin-Watson: DW = 2,345; Jarque-Bera: P-val. = 0,645; Breusch-Godfrey: a) [AR(1)] P-val: 0,191; b) [AR(2)] P-val: 0,012; Ljung-Box (p=6): P-val.= 0,209; ARCH: a) [ARCH(1)] P-val: 0,076; a) [ARCH(2)] P-val: 0,167; White: P-val. = 0,154; <i>SupF</i> = 6,164; <i>MeanF</i> = 3,407; <i>ExpF</i> = 2,025.

Notas:

-Debajo de los coeficientes estimados aparecen, entre paréntesis, los estadísticos t correspondientes.

-* denota el rechazo de la hipótesis nula al 10%, ** al 5% y *** al 1%.

- \hat{u}_{t-1} son los residuos (retrasados) de la relación a largo plazo estimada en cada una de las ecuaciones del Cuadro nº 75.

-F83_t es una variable ficticia que toma 1 en 1983 y 0 en el resto.

-D seguido de valores numéricos, por ejemplo D7375_t, denota a una variable ficticia que toma el valor 1 en los años 1973-1975, y 0 en el resto.

-D seguido de una variable, denota la primera diferencia de una variable, ya sea una ficticia (DD7175), una variable en logaritmos (DLVOLAG), o el producto de dos variables (D(D7078LVGES)).

TABLA 4. Estadísticos sobre el funcionamiento predictivo de los modelos de simulación “*ex-post*” 1994-1995.

MODELO	ME	MAE	RMSE	U
Agricultura				
Q=0	-0,065	0,065	0,075	0,587
Q=I	0,063	0,063	0,085	0,667
Q diag. estimada	-0,060	0,066	0,089	0,698
Energía				
Q=0	0,010	0,042	0,043	3,165
Q=I	-0,026	0,045	0,052	3,800
Q diag. estimada	-0,084	0,084	0,109	7,932
Construcción				
Q=0	-0,016	0,016	0,017	0,898
Q=I	-0,021	0,021	0,021	1,128
Q diag. estimada	0,012	0,013	0,014	0,723
Industria Manufacturera				
Q=0	0,026	0,026	0,028	1,075
Q=I	0,012	0,024	0,027	1,012
Q diag. estimada	0,009	0,009	0,013	0,494
Servicios destinados a la venta				
Q=0	0,011	0,011	0,013	0,664
Q=I	0,016	0,016	0,017	0,874
Q diag. estimada	0,041	0,041	0,042	2,152
Transportes y comunicaciones				
Q=0	0,006	0,017	0,018	0,569
Q=I	-0,002	0,017	0,017	0,549
Q diag. estimada	-0,040	0,040	0,048	1,547
Servicios no destinados a la venta				
Q=0	-0,001	0,003	0,003	0,135
Q=I	-0,003	0,003	0,004	0,172
Q diag. estimada	-0,021	0,021	0,026	1,127

Nota: ME = error medio; MAE = error absoluto medio; RMSE = error cuadrático medio; U= coeficiente de desigualdad de Theil (1966); en negrita, los mejores resultados.

REFERENCIAS

Andrés J.A., Escribano C., Molinas C. y Taguas D., 1990, *La inversión en España: Modelización con Restricciones de Equilibrio* (Barcelona-Madrid: Antoni Bosch, ed., e Instituto de Estudios Fiscales).

Andrews D., 1993, Tests for Parameter Instability and Structural Change With Unknown Change Point, *Econometrica*, 61, 821-856.

Andrews D. y Ploberger W., 1994, Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative, *Econometrica*, 62, 1383-1414.

Ansley C. y Kohn R., 1983, State Space Models and Diffuse Initial Conditions, I: Filtering and Likelihood, Technical Report 13, University of Chicago, Statistics Research Center, Graduate School of Business.

Campos J., Ericsson N. y Hendry D.F., 1996, Cointegration test in the presence of structural breaks, *Journal of Econometrics*, 70, 187-220.

Canarella G., Pollard S. y Lai K., 1990, Cointegration between exchange rates and relative price: another view, *European Economic Review*, 34, 1303-1322.

Chow G., 1984, Random and Changing Coefficient Models, en: Z. Griliches y M. Intriligator, eds., *Handbook of Econometrics* (Vol. II) (Amsterdam: North-Holland), 1213-1245.

Cordero, G. y Gayoso, A., 1997, "Evolución de las Economías Regionales en los Primeros 90", Ministerio de Economía y Hacienda. Secretaría de Estado de Presupuestos y Gastos. Dirección General de Análisis y Programación Presupuestaria. Noviembre de 1997.

Courbis, R., 1994, La modelización Multirregional en Europa Occidental: Balance y Perspectivas, en Pulido, A. y Cabrer, B., (Coords.), 1994, 145-181.

Dickey D. and Fuller W., 1979, Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 84, 427-431.

Dickey D. and Fuller W. , 1981, Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, 50, 1057-1072.

Engle R. and Granger C., 1987, Co-integration and error correction: representation, estimation and testing, *Econometrica*, 55, 251-276.

Gregory A. y Hansen B., 1996, Residual-based tests for cointegration in models with regime shifts, *Journal of Econometrics*, 70, 99-126.

Hackl P. and Westlund A., 1996, Demand for international telecommunication. Time-varying price elasticity, *Journal of Econometrics*, 70, 243-260.

Hall S., 1994, Modelling economies in transition, Centre for Economic Forecasting, London Business School.

Hansen B., 1992, Tests for Parameter Instability in regressions with I(1) Processes, *Journal of Business and Economic Statistics*, 10, 321-336.

Harvey A., 1981, *The Econometric Analysis of Time Series* (Oxford: Philip Allan Publishers Limited).

Harvey A., 1989, *Forecasting, Structural Time Series Models, and the Kalman Filter* (Cambridge: Cambridge University Press).

Hendry D.F. y Neale A., 1991, A Monte Carlo study of the effects of structural breaks on tests for unit roots, en: P. Hackl y A. Westlund, eds., *Economic structural change: analysis and forecasting* (Berlin: Springer-Verlag), 95-119.

Hispalink, 1993, Banco de datos multirregional HISPALINK (Madrid: Consejo Superior de Cámaras de Comercio, Industria y Navegación de España).

Kalman R., 1960, A new approach to linear filtering and prediction problems, *ASME Journal of Basic Engineering*, 82, 35-45.

Lin C-F. y Teräsvirta T., 1994, Testing the constancy of regression parameters against continuous structural change, *Journal of Econometrics*, 62, 211-228.

Lucas R.E., 1976, Econometric Policy Evaluation: A Critique, *Carnegie-Rochester Conference Series on Public Policy*, 1, 19-46.

Lühtkepohl H., 1993, *Introduction to Multiple Time Series Analysis* (New York: Springer-Verlag).

McKinnon J., 1991, Critical values for cointegration tests in long-run economic relationships, en: R. Engle and C. Granger, eds., *Readings in Cointegration* (New York: Oxford University Press), 267-276.

Molinas C., Sebastián M. y Zabalza A., 1991, *La economía española: una perspectiva macroeconómica* (Barcelona-Madrid: Antoni Bosch, ed., e Instituto de Estudios Fiscales).

Otero J.M., Isla F., Trujillo F., Fernández A. y López P., 1996, Modelización económica regional: el proyecto Hispalink-Andalucía, *Boletín Económico de Andalucía*, 21, 49-66.

Perron P., 1989, The great crash, the oil price shock, and the unit root hypothesis, *Econometrica*, 57, 1361-1401.

Perryman, M. R. y Schmidt, J. R. (Ed.), 1986, *Regional Econometric Modeling*, Boston: Kluwer•Nijhoff Publishing.

Pulido, A. y Cabrer, B., (Coords.), 1994, *Datos, Técnicas y Resultados del Moderno Análisis Económico Regional*, Ediciones Mundi-Prensa.

Quandt R., 1960, Tests of the Hypothesis That a Linear Regression System Obeys Two Separate Regimes, *Journal of the American Statistical Association*, 55, 324-330.

Quintos C. y Phillips P., 1993, Parameter constancy in cointegrating regressions, *Empirical Economics*, 18, 675-703.

Ramajo J. y Márquez M.A., 1996, Elaboración de indicadores sintéticos para el seguimiento de la coyuntura económica de Extremadura, Monográfico de la Consejería de Economía, Industria y Hacienda, Junta de Extremadura.

Rappoport P. y Reichlin L., 1989, Segmented trends and non-stationary time series, *The Economic Journal*, 99, 168-177.

Stock J. and Watson M., 1996, Evidence on structural instability in macroeconomic time series relations, *Journal of Business and Economic Statistics*, 14, 11-30.

Theil H., 1966, *Applied Economic Forecasting* (Amsterdam: North-Holland).

Treyz G., 1993, *Regional economic modeling. A Systematic Approach to Economic Forecasting and Policy* (Boston: Kluwer Academic Publishers).

Weber, R. E., 1986, Regional Econometric Modelling and the New Jersey State Model; en Perryman, M. R. y Schmidt, J. R. (Ed.), 1986, 11-39.

Wolf C., 1987, Time-Varying Parameters and the Out-of-Sample Forecasting Performance of Structural Exchange Rate Models, *Journal of Business and Economic Statistics*, 5, 87-97.

Wolters J., Teräsvirta T. y Lühtkepohl H., 1996, Modelling the Demand for M3 in the Unified Germany, Discussion Paper 23, Humboldt Universität zu Berlin.