

XXIX REUNIÓN DE ESTUDIOS REGIONALES

Santander 27 y 28 de noviembre de 2003

Universidad de Cantabria

Una estrategia para la especificación de un modelo econométrico estructural: Aplicación al VAB agrícola de Andalucía

Fernando Isla Castillo

Profesor titular de Economía Aplicada
Departamento de Estadística y Econometría (68)
Universidad de Málaga
Tfno: 952137189
e-mail: isla@uma.es

Francisco Trujillo Aranda

Catedrático de Economía Aplicada
Departamento de Estadística y Econometría (68)
Universidad de Málaga
Tfno: 952131207
e-mail: trujillo@uma.es

RESUMEN

En este trabajo describimos y aplicamos una estrategia metodológica para llevar a cabo la especificación de un modelo econométrico trimestral que permite predecir el VAB agrícola de la economía andaluza mediante un *Modelo Estructural Vectorial de Corrección por el Error*. En muchas ocasiones el investigador realiza la especificación del modelo econométrico mediante un procedimiento no secuencial de prueba y error, con los riesgos de “*data mining*” que ello conlleva. La principal aportación de este trabajo es el planteamiento y aplicación de una estrategia para aproximar de forma adecuada el verdadero *Proceso Generador de los Datos*. En lugar de proceder de una forma no ordenada se trata de ir “interrogando a los datos” a través de la metodología “de lo general a lo simple”, en la que se parte de una especificación VAR irrestricta (UVAR) y se hace uso de un *software* específico (*Pc-Give*). Esta estrategia permite analizar la idoneidad de las variables que integrarán el modelo, determinar el orden del UVAR e, incluso, identificar la opción del contraste de cointegración de *Johansen* más adecuada dadas las características de dichas variables. Al final se ilustra la validez de esa estrategia mediante la evaluación de la capacidad predictiva de los modelos estructurales estimados.

Palabras clave: modelización de lo general a lo simple, modelización econométrica regional, cointegración, modelos de corrección por el error, predicción.

1. INTRODUCCIÓN

En este artículo mostramos la especificación y estimación de un Modelo Estructural Vectorial de Corrección por el Error (SECM)¹ del VAB del sector agrícola de Andalucía, modelo estimado con fines de predicción a corto plazo.

Se denomina modelo estructural a un *Modelo de Ecuaciones Simultáneas (SEM)*² en el que se hacen explícitas las relaciones contemporáneas entre las variables endógenas y en el que también pueden existir correlaciones contemporáneas entre los términos de perturbación de las ecuaciones del modelo. Por modelo vectorial de corrección por el error (VECM)³ se denominan los modelos multivariantes y multiecuacionales que recogen explícitamente las relaciones de corto y largo plazo entre variables económicas haciendo uso del mecanismo de corrección por el error. La unión de ambos tipos de modelos se denomina *Modelo Estructural Vectorial de Corrección por el Error (SECM)*⁴.

Un paso previo en la especificación de un modelo estructural es la identificación de la estructura, a largo y corto plazo, a partir de restricciones extramuestrales impuestas por la Teoría Económica. Sin embargo, uno de los principales problemas que el analista se encuentra a la hora de especificar su modelo es que la Teoría Económica no le proporciona suficiente información ni de la dinámica a corto plazo entre las variables ni, incluso, respecto a las relaciones de equilibrio a largo plazo. Es más, el orden de la causalidad queda sujeto en algunos casos al arbitrio del investigador.

El enfoque inicial de la *Cowles Commission* sobre identificación y estimación de modelos econométricos estructurales fue duramente criticado en los años 70 del siglo pasado, cuando diversos autores mostraron que esos modelos "...no representaban la realidad reflejada en los datos, no representaban la teoría...eran ineficientes en el campo aplicado de la predicción y simulación de políticas alternativas..." (Pesaran y Smith, 1995).

En el proceso de explicar los fracasos del enfoque de la *Cowles Commission* surgieron nuevos desarrollos metodológicos en la investigación econométrica, por ejemplo, como apunta Favero (2001), dentro de la línea del problema de la identificación se puede distinguir el enfoque de la

¹ *Structural Error Correction Model.*

² *Simultaneous Equation Models.*

³ *Vector Error Correction Model.*

⁴ Sobre modelos estructurales, su identificación y estimación pueden consultarse diversos trabajos recientes como los de Johansen y Juselius (1994), Ericsson, N.R. (1995), Clements, M. P. y Mizon, G.E. (1991) y Urbain, J.-P. (1995).

London School of Economics (LSE), el enfoque de los modelos VAR⁵ y el enfoque de la optimización intertemporal del Ciclo Económico (RBC).

En este trabajo hemos aplicado la metodología de la *LSE* para tratar de aproximar el proceso generador de los datos (PGD) que explica el comportamiento de la serie del VAB de la Agricultura en Andalucía, haciendo uso además del VAB de la Agricultura en España y de la Relación Asalariados-Ocupados Agrícolas en Andalucía. En líneas generales este enfoque parte de la idea de que el proceso generador de los datos es inobservable, en nuestro caso porque el modelo estructural que subyace bajo una forma reducida irrestricta (UVAR)⁶ que tomamos como modelo base no está identificado. Sin embargo, en la práctica los datos pueden soportar ciertas restricciones que permiten una aproximación del PGD mediante una forma reducida restringida. Es más, si se disponen de suficientes restricciones es posible contrastar dicho proceso de reducción e, inclusive, contrastar ciertas restricciones estructurales a priori que establece la Teoría Económica.

En los modelos estimados todas las variables se expresan en logaritmos neperianos y se refieren al período 1980.1-2002.4 (datos trimestrales). Las cifras de los VAB (regional y nacional) y del indicador de empleo corresponden a la señal tendencia-ciclo de dichas variables, extraídas con el programa TRAMO-SEATS, esto es, son series depuradas de estacionalidad, variaciones irregulares y valores atípicos. Los VAB se expresan en millones de euros constantes de 1995. En la sección 2 abordamos la especificación de un modelo VAR no restringido (UVAR) como punto de arranque para llevar a cabo de forma adecuada el análisis de cointegración y los posteriores contrastes de hipótesis. Dicha especificación inicial se entiende que está libre de errores de especificación y constituye una buena aproximación de los datos. En la sección 3 procedemos a la simplificación del modelo UVAR obteniendo un VAR parsimonioso (PVAR) y su representación en forma de *Modelo Vectorial de Corrección por el Error (VECM)*, así mismo analizamos la adecuación de la simplificación. En la sección 4 llevamos a cabo la especificación y estimación de un modelo estructural de corrección por el error (SECM) partiendo del VECM parsimonioso obtenido en la sección anterior y se contrasta también la idoneidad de este nuevo paso, así como su simplificación mediante adecuados contrastes de exogeneidad. Finalmente, en la última sección hacemos una evaluación de la capacidad predictiva de las especificaciones finales estimadas y extraemos algunas conclusiones⁷.

⁵ *Vector Autoregressive Model.*

⁶ *Unrestricted Vector Autoregressive Model.*

⁷ Todo el análisis se ha llevado a cabo haciendo uso de los programas de econometría *PC-Give 10.0* y *EViews 4.0*.

2. ANÁLISIS DE COINTEGRACIÓN

2.1 Búsqueda de una especificación inicial general congruente con los datos

En la práctica es frecuente desconocer el número de relaciones de cointegración existentes entre N variables, así como la especificación de dichas relaciones. No obstante, la Teoría Económica puede proporcionar cierta información a priori sobre dichas relaciones económicas de largo plazo. Desafortunadamente, en nuestro caso la Teoría Económica poco nos puede decir sobre la relación entre las variables analizadas, esto es, los VAB de la Agricultura en Andalucía y España y el cociente Asalariado-Ocupados Agrícolas. A priori cabe esperar que exista una relación de largo plazo entre el VAB y el empleo regionales, mientras que la relación entre el VAB regional y nacional no parece tan evidente si admitimos que la agricultura presenta en Andalucía peculiaridades propias, ligadas con la climatología, que no se dan en el resto del país.

Existen múltiples contrastes de cointegración⁸, el que aplicamos es el contraste multiecuacional de Johansen (1988, 1992a), quizás el más popular e implementado en numerosos programas de econometría. El primer paso de este contraste es la especificación y estimación de un modelo *UVAR* de N variables que sea congruente, irrestricto, cerrado y de orden p . En nuestro caso $N=3$ y se supone que p no es conocida a priori, con lo que el modelo vendría dado por:

$$y_t = \Pi_1 y_{t-1} + \Pi_2 y_{t-2} + \dots + \Pi_p y_{t-p} + \Psi D_t + \varepsilon_t; \quad t = 1, 2, \dots, T \quad (1)$$

donde $y_t' = (vaan95_t, vaes95_t, asaocut_t)'$ es el vector de variables modelizadas (**vaan95** = ln del VAB de la Agricultura en Andalucía, **vaes95** = ln del VAB de la Agricultura en España y **asaocut** = ln del cociente Asalariados/Ocupados). $D_t' = (1, t, w_t)'$, donde w_t denota un vector de variables ficticias $w8303, w8304, w8502, w8602, w8604, w8802, w8804, w8902, w9001, w9002, w9404, w9502, w9601, w0003$ ⁹. En tanto que ε_t denota un vector de perturbaciones estacionarias¹⁰ y Π_i y Ψ denotan las matrices de coeficientes.

Mediante manipulaciones algebraicas la ecuación (1) se puede escribir también en la forma:

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + \Psi D_t + \Pi y_{t-1} + \varepsilon_t; \quad t = 1, 2, \dots, T \quad (2)$$

⁸ Véanse, por ejemplo Watson (1994) y Maddala y Kim (1998).

⁹ En general, $wAAQQ$ denota una variable que toma el valor 1 en el trimestre QQ del año AA y cero en el resto. Estas variables ficticias se incluyen para garantizar la hipótesis previa de que el *UVAR* constituye un modelo correctamente especificado, esto es, no presenta problemas de mala especificación. Con ello se garantiza también la validez de los contrastes posteriores.

¹⁰ Véase Banerjee *et al.* (1993, pág. 87-88) para las condiciones mínimas que deben satisfacer las perturbaciones. Además, se suponen fijos los p valores iniciales de las variables.

donde Π es una matriz cuadrada de orden N y rango r ($0 < r \leq N-1$) si las variables son $CI(1,1)^{11}$, siendo también r el rango de cointegración, que se puede expresar como el producto de dos matrices de orden $N \times r$ y rango r , esto es:

$$\Pi = \alpha\beta', \quad (3)$$

descomposición en la que β es la matriz de cointegración y α es la matriz de ajuste. Los términos de corrección por el error, que figuran como regresores en el modelo, vienen dados por $\beta'y_{t-1}$. Justamente por eso las expresiones (2) y (3) constituyen la representación del *UVAR* en forma de modelo vectorial de corrección por el error (*VECM*). Un paso clave previo al análisis de cointegración es determinar la longitud del retardo del modelo *UVAR*, esto es, el valor del parámetro p . La tabla 1 muestra los criterios empleados al efecto en este trabajo.

Tabla 1. Elección del orden p del modelo *UVAR*

Criterios de información							
<i>UVAR(p)</i>	<i>T</i>	<i>k</i>		log-likelihood	SC	HQ	AIC
UVAR(1)	74	57	OLS	625.41754	-13.588	-14.655	-15.363
UVAR(2)	74	66	OLS	770.20404	-16.978	-18.213	-19.033
UVAR(3)	74	75	OLS	902.54769	-20.031	-21.435	-22.366
UVAR(4)	74	84	OLS	1015.0916	-22.549	-24.121	-25.165
UVAR(5)	74	93	OLS	1043.4267	-22.792	-24.532	-25.687
UVAR(6)	74	102	OLS	1058.3784	-22.672	-24.581	-25.848

T = Número de observaciones de cada una de las tres ecuaciones

k = Número total de parámetros del sistema

Contrastes de mala especificación			
	AR (1-5) test ~ F ^a	Normality test ~ Chi ² ^b	Heteros test ~ Chi ² ^c
UVAR(6)	0.79672 [0.7906]	5.6120 [0.4680]	288.42 [0.8269]
UVAR(5)	0.90405 [0.6383]	7.3712 [0.2879]	250.21 [0.8654]
UVAR(4)	0.88748 [0.6652]	18.860 [0.0044]**	0.046592 [1.0000]
UVAR(3)	1.9008 [0.0045]**	12.092 [0.0599]	0.40928 [1.0000]
UVAR(2)	5.9946 [0.0000]**	15.558 [0.0163]*	0.63642 [0.9962]
UVAR(1)	6.7125 [0.0000]**	11.821 [0.0661]	0.58122 [0.9994]

a Contraste multivariante de autocorrelación de orden 5 (Godfrey, 1988), que se distribuye según una $\chi^2 ((q-h+1)N^2)$, en el que se especifica un $AR(h-q)$ con $1 \leq h \leq q$, tal que en este caso $q \leq 5$.

b Contraste multivariante de normalidad (Doornik y Hansen, 1994) que se distribuye según una $\chi^2 (6)$

c Contraste multivariante de homoscedasticidad (Doornik, 1996) que se distribuye según una $\chi^2 (sN(N+1)/2)$ donde s es el número de variables no redundantes.

Contrastes de especificación ^d	
UVAR(6) --> UVAR(5):	$F(9, 92) = 1.8590 [0.0681]$
UVAR(6) --> UVAR(4):	$F(9, 99) = 4.1062 [0.0002]**$
UVAR(5) --> UVAR(4):	$F(18, 107) = 3.0728 [0.0002]**$

d Contraste *F* basado en una aproximación del test de la razón de verosimilitud (véase Anderson, 1984)

¹¹ Cointegradas de manera que las perturbaciones de las ecuaciones de largo plazo entre las variables son estacionarias. Previamente se han aplicado los contrastes de raíces unitarias pertinentes, regulares y estacionales, y se ha admitido que todas las variables estocásticas sólo tienen una raíz unitaria regular, esto es, son integradas de orden 1 o $I(1)$.

Los criterios de información de Akaike (*AIC*), Schwarz (*SC*) y Hannan-Quinn (*HQ*) utilizados en selección de modelos indican que debe elegirse un *UVAR* con 5 o 6 retardos. Los contrastes de mala especificación permiten concluir que son los modelos *UVAR*(5) y *UVAR*(6) los que superan todos los contrastes. Efectivamente, el *UVAR*(4) presenta problemas de normalidad, mientras que con un número inferior de retardos el problema se agrava al detectarse autocorrelación. El contraste *F* de especificación, una aproximación al contraste de la razón de verosimilitud, no rechaza la especificación *UVAR*(5), por lo que finalmente seleccionamos dicho modelo.

2.2 Contraste del rango de cointegración

El orden de cointegración r es desconocido a priori, para determinarlo se formula un contraste de cointegración, del tipo razón de verosimilitudes, en el que se compara el valor de la verosimilitud del modelo estimado sin restricciones (rango de Π igual a N) y con restricciones (rango igual a r con $r > 0$)¹². El estadístico de la *traza*¹³ sirve para contrastar que, como máximo, hay r vectores de cointegración ($0 \leq r \leq N - 1$) y $N - r$ raíces unitarias. La expresión del estadístico es:

$$\eta_r = -T \sum_{i=r+1}^N \log(1 - \lambda_i), \quad r = 0, 1, 2, \dots, N-1. \quad (4)$$

Siendo λ_i el valor propio i -ésimo en tamaño de la matriz a partir de la que se estima β ¹⁴. La distribución de este estadístico se obtiene bajo la hipótesis nula de que hay r vectores de cointegración, frente a la alternativa de que hay N . El contraste se aplica secuencialmente, calculando desde η_0 hasta η_{N-1} , y seleccionando como orden de integración el valor r si el último estadístico significativo (que rechaza la hipótesis nula) es η_{r-1} . Este contraste admite distintas especificaciones que se recogen en la tabla 2.

La distribución del estadístico no es la habitual, depende de la dimensión N del vector de variables analizado y de la naturaleza de los elementos deterministas (constante, tendencia y ficticias estacionales) incluidos¹⁵. Los valores críticos para las distintas especificaciones han sido calculados por los autores del contraste, si bien las tablas de Osterwald-Lenum (1992) son más completas.

¹² Como es sabido, si $r=0$ no existiría cointegración entre las variables y debería estimarse un *UVAR* en niveles. Si $r=N$ las variables serían estacionarias, en contra de lo que se supone inicialmente y de los resultados de los contrastes de raíces unitarias obtenidos en este caso.

¹³ También se puede aplicar el estadístico del *valor máximo*, definido como: $\lambda_{max} = -T \log(1 - \lambda_{r+1})$, $r = 0, 1, 2, \dots, N-1$. Con el que contrasta la hipótesis nula de que hay r vectores de cointegración frente a la alternativa de que hay $r+1$.

¹⁴ Véase, entre otros, Banerjee *et al.* (1993, pág. 261-268).

¹⁵ Véanse Banerjee *et al.*, 1993, pág. 271-276 y Maddala y Kim, 1988, pág. 213-214.

Tabla 2. Conjunto de posibles modelos a contrastar

Especificaciones	Características
Modelo 1	Ni el <i>VECM</i> ni las ecuaciones de cointegración tienen constante.
Modelo 2	Las ecuaciones de cointegración tienen constante, pero el <i>VECM</i> no.
Modelo 3	El <i>VECM</i> y las ecuaciones de cointegración tienen constante.
Modelo 4	El <i>VECM</i> tiene constante, las ecuaciones de cointegración tienen constante y tendencia lineal.
Modelo 5	El <i>VECM</i> y las ecuaciones de cointegración tienen constante y tendencia lineal.

Antes de aplicar el contraste de Johansen se pueden calcular las raíces del polinomio característico del *UVAR*(5), seleccionado y estimado previamente, con objeto de explorar el número de raíces unitarias. Las 15 raíces estimadas se muestran en la tabla 3, en la que se puede apreciar la presencia de dos raíces con modulo igual a la unidad, lo que constituye un indicador de la posible presencia de un vector de cointegración al estar operando con tres variables.

Tabla 3. Raíces del polinomio característico del *UVAR*(5)

<i>Roots of Characteristic Polynomial</i>			
Raíces	Módulo	Raíces	Módulo
1.001548 - 0.004996i	1.001560	0.508611 + 0.655145i	0.829398
1.001548 + 0.004996i	1.001560	0.029830 - 0.759461i	0.760046
0.919692	0.919692	0.029830 + 0.759461i	0.760046
0.800956 + 0.444860i	0.916205	0.656088 - 0.301632i	0.722104
0.800956 - 0.444860i	0.916205	0.656088 + 0.301632i	0.722104
0.707996 + 0.519976i	0.878427	0.294907	0.294907
0.707996 - 0.519976i	0.878427	0.025478	0.025478
0.508611 - 0.655145i	0.829398		

La tabla 4 muestra el resumen de los contrastes de cointegración de la Traza y del valor Máximo para el conjunto de modelos recogidos en la tabla 2. Con ambos contrastes no se descarta, inicialmente, la posibilidad de que existan uno o dos vectores de cointegración¹⁶. Nótese que aunque con el modelo 1 se concluiría que no existe cointegración, esta opción se descarta porque la ausencia de constante en los vectores de cointegración es una restricción muy fuerte e innecesaria en nuestro caso.

La tabla 5 resume los valores del criterio *SC* de Schwarz que se utiliza para determinar cuál es el modelo de contraste más congruente con los datos. Según este criterio, se podría seleccionar el modelo 2 (las ecuaciones de cointegración tienen constante, el *VECM* no) con un vector de

¹⁶ Los valores críticos utilizados en los contrastes de cointegración no tienen en cuenta la existencia de variables ficticias en el *UVAR*. Sin embargo, hemos comprobado que excluyendo dichas variables llegamos al mismo resultado desde el punto de vista de la cointegración.

cointegración ($r=1$), o el modelo 4 (el *VECM* tiene constante, las ecuaciones de cointegración tienen constante y tendencia lineal) con dos vectores de cointegración ($r=2$).

Tabla 4. Resumen de los contrastes de cointegración

Número de vectores de cointegración (r)					
Contraste	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
Traza	0	1	1	2	2
Valor-Max.	0	1	1	2	2

Tabla 5. Criterio SC para seleccionar la especificación del *VECM* y el número de vectores de cointegración

Criterio de Schwarz por Rango (filas) y Modelo (columnas)					
R	Modelo 1	Modelo 2	Modelo 3	Modelo 4	Modelo 5
0	-24.51338	-24.51338	-24.42697	-24.42697	-24.40274
1	-24.31249	-25.01897	-24.96610	-24.90819	-24.84799
2	-24.00338	-24.72126	-24.68477	-25.05473	-25.02608
3	-23.66445	-24.34012	-24.34012	-24.71094	-24.71094

Atendiendo al principio de parametrización escueta, teniendo en cuenta que previamente se han estimado dos raíces unitarias en el polinomio característico del *UVAR*, que los valores del criterio SC son muy similares en ambos casos (-25,02 y -25,05 para los modelos 2 y 3 respectivamente) y que la serie a predecir, el VAB agrícola en Andalucía, no presenta una clara tendencia lineal, nos decantamos por aceptar el modelo 2 con un vector de cointegración como la especificación correcta en este caso. Al considerar sólo un vector de cointegración éste se encuentra identificado.

Seguidamente se ha contrastado la restricción de que el VAB de España no interviene en el vector de cointegración, esto es, en la relación de largo plazo que "explica" el VAB agrícola en Andalucía. Los resultados del contraste figuran en la tabla 6, donde se puede ver que el estadístico de la Razón de Verosimilitud deja una probabilidad (0,76) muy superior a los niveles de significación habituales y, por tanto, permite no rechazar dicha restricción¹⁷. De acuerdo con la estimación del vector de cointegración restringido, la elasticidad a largo plazo entre el VAB agrícola en Andalucía y el cociente Asalariados/Ocupados agrícolas en la región es significativa y supera ligeramente el 1,5%. La constante del vector de cointegración también es significativa. Nótese que el coeficiente de ajuste (coeficiente del término de corrección de error) en la ecuación de $\Delta asaocut$ no es significativamente distinto de cero, ello estaría indicando que dicha variable puede ser débilmente exógena.

¹⁷ Este estadístico sigue una distribución $\chi^2(1)$. Para un resumen de los contrastes de restricciones sobre el espacio de cointegración puede consultarse Doornik y Hendry (2001), páginas 178-183.

Tabla 6. Contraste de restricciones sobre el vector de cointegración en el VECM

Restricted cointegrating coefficients (std.err. in parentheses)			
VAAN95	VAES95	ASAOUCU	C
1.000000	0.000000	-1.565486	-7.941964
(0.00000)	(0.00000)	(0.15528)	(0.07746)
Adjustment coefficients (std.err. in parentheses)			
DVAAN95	-0.036134		
	(0.00458)		
DVAES95	-0.025443		
	(0.01155)		
DASAOUCU	-0.002954		
	(0.01007)		
log-likelihood	1018.19431		
-T/2log Omega	1333.19866		
no. of observations	74		
no. of parameters	83		
LR test of restrictions: Chi^2(1) = 0.089748 [0.7645]			

$$DVAAN95 = (1-L)vaan95; DVAES95 = (1-L)vaes95; DASAOUCU = (1-L)asaocu$$

3. Simplificación del UVAR en forma de VECM : el modelo PVAR

El VECM especificado en el epígrafe anterior, cuya expresión general viene dada por las ecuaciones (2) y (3), se puede considerar la forma reducida del siguiente modelo estructural de corrección por el error (SECM):

$$\Phi \Delta y_t = a\beta' y_{t-1} + \Gamma_1^+ \Delta y_{t-1} + \dots + \Gamma_{p-1}^+ \Delta y_{t-p+1} + \Psi^+ D_t + u_t, \quad (5)$$

en el que la matriz a recoge los coeficientes de ajuste, esto es, los coeficientes de los términos de corrección por el error en el modelo estructural. Nótese que los términos de corrección por el error, dados por $\beta' y_{t-1}$, son los mismos que en el VECM.

Efectivamente, despejando las variables endógenas en el SECM se tendrá su forma reducida:

$$\Delta y_t = \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \dots + \Gamma_{p-1} \Delta y_{t-p+1} + \Psi D_t + \varepsilon_t, \quad (6)$$

que coincidirá formalmente con el VECM, ecuaciones (2) y (3), ya que: $\alpha = \Phi^{-1}a$, $\Gamma_i = \Phi^{-1}\Gamma_i^+$, $\Psi = \Phi^{-1}\Psi^+$ y $\varepsilon_t = \Phi^{-1}u_t$.

El *VECM* en (6) no tiene incorporada ninguna restricción, no obstante, al especificarse el *VECM* a partir de un *UVAR* (*VAR* irrestricto), muy probablemente tendrá en la práctica un exceso de parámetros, esto es, algunos de sus regresores no serán relevantes y podrán eliminarse.

En este trabajo se ha especificado inicialmente un *UVAR*(5) en niveles y se ha contrastado que se puede expresar en forma de *VECM*¹⁸ no restringido con $r=1$. Ello supone estimar un sistema de tres ecuaciones en el que cada una tiene 27 parámetros, lo que supone un total de 81 parámetros en el sistema¹⁹. Es muy probable que este sistema pueda simplificarse, eliminando regresores redundantes, hasta llegar a un *PVAR* (*VAR* parsimonioso) expresado también en forma de *VECM*. De hecho, es muy posible que no todas las variables ficticias sean relevantes en todas las ecuaciones del modelo, si bien es preciso tener en cuenta que, para poder llevar a cabo la simplificación, es necesario que la variable en cuestión no sea relevante en ninguna de las tres ecuaciones.

Tabla 7. Contraste F para retener regresores en el *UVAR* en forma de *VECM*

F-tests para retener regresores, $F(3,45) =$

DVAAN95_1	531.150	[0.000]**	DVAAN95_2	72.9723	[0.000]**
DVAAN95_3	7.11908	[0.001]**	DVAAN95_4	1.06220	[0.375]
DVAES95_1	108.404	[0.000]**	DVAES95_2	27.8814	[0.000]**
DVAES95_3	7.36357	[0.000]**	DVAES95_4	1.31586	[0.281]
DASAOUCU_1	94.7969	[0.000]**	DASAOUCU_2	31.1213	[0.000]**
DASAOUCU_3	14.8495	[0.000]**	DASAOUCU_4	13.0535	[0.000]**

W8304	1.76137	[0.168]	W8502	7.66011	[0.000]**
W8602	6.64462	[0.001]**	W8604	4.62821	[0.007]**
W8802	9.81541	[0.000]**	W8804	4.48673	[0.008]**
W8902	15.1099	[0.000]**	W9001	3.35633	[0.027]*
W9002	14.0040	[0.000]**	W9502	10.4831	[0.000]**
W9601	2.09220	[0.115]	W0003	1.54509	[0.216]
W9404	3.09019	[0.036]*	W8303	2.37425	[0.083]

La tabla 7 muestra el resumen del contraste de la F para retener un regresor cualquiera en las tres ecuaciones del sistema. Dicho contraste confirma que es posible suprimir determinados regresores entre los que se encuentran algunas variables ficticias (**w8303**, **w8304**, **w9601** y **w0003**) y los retardos de orden cuatro (**vaan95** y **vaes95** en diferencias). Los resultados de la simplificación se muestran en las tablas 8 y 9 donde:

- **VECM_1**: representa el *UVAR* en forma de *VECM*.

¹⁸ Nótese que al expresar el *UVAR* como *VECM* el número de retardos se reduce a 4.

¹⁹ Se incluye en cada ecuación el término de corrección de error restringido que figura en la tabla 2.6.

- **VECM_2:** representa el modelo *VECM_1* eliminando, a su vez, las siguientes variables ficticias: *w8303*, *w8304*, *w9601* y *w0003*,
- **VECM_3:** representa el *VECM_2* eliminando la variable *dvaan95_4* (retardo de orden cuatro de la diferencia de *vaan95*),
- **VECM_4:** representa el *VECM_3* eliminando posteriormente la variable *dvaes95_4* (retardo de orden cuatro de la diferencia de *vaes95*),

El modelo resultante, *VECM_4*, no presenta problemas de autocorrelación serial, heteroscedasticidad o no normalidad en las perturbaciones (véase tabla 8). Asimismo, observando en la misma tabla, si nos atenemos exclusivamente a criterios de información, tanto el criterio de *Akaike* (*AIC*), *Schwarz* (*SC*) y *Hannan-Quinn* (*HQ*), hay una coincidencia absoluta en elegir el modelo con menor número de parámetros. Finalmente, el contraste *F* para seleccionar modelos anidados, no permite rechazar la última especificación por lo que concluimos que el *PVAR* en forma de *VECM* constituye una reducción adecuada del modelo. El modelo *VECM_4* tiene un total de 63 parámetros, el logaritmo de la función de verosimilitud asciende a 997,2 y el determinante de la matriz de varianzas y covarianzas de los residuos, $|\hat{\Omega}|$, asciende a $3,96 \cdot 10^{-16}$.

Tabla 8. Determinación del PVAR: contrastes de mala especificación

Model	Contrastes de mala especificación ^a		
	AR 1-5 test ~ F	Normality test ~ Chi^2	Heteros test ~ Chi^2
VECM_4	0.90247 [0.6442]	5.7942 [0.4466]	0.27631 [1.0000]
VECM_3	0.85855 [0.7131]	4.5860 [0.5979]	0.25087 [1.0000]
VECM_2	0.75722 [0.8509]	3.0260 [0.8056]	0.20635 [1.0000]
VECM_1	0.92044 [0.6138]	8.2990 [0.2170]	0.05684 [1.0000]

Criterios de información							
Model	T	K	log-likelihood	SC	HQ	AIC	
VECM_4	74	63	OLS 997.20492	-23.287	-24.466	-25.249	
VECM_3	74	66	OLS 1000.1057	-23.191	-24.426	-25.246	
VECM_2	74	69	OLS 1002.9624	-23.094	-24.385	-25.242	
VECM_1	74	81	OLS 1018.1943	-22.808	-24.324	-25.330	

Contrastes de Especificación	
Secuencia de simplificación	
VECM_1 --> VECM_2: F(12,119)=	1.6744 [0.0809]
VECM_2 --> VECM_3: F(3,49) =	1.3110 [0.2814]
VECM_1 --> VECM_3: F(15,124)=	1.6097 [0.0802]
VECM_3 --> VECM_4: F(3,50) =	1.3592 [0.2660]
VECM_2 --> VECM_4: F(6,98) =	1.3215 [0.2548]
VECM_1 --> VECM_4: F(18,127)=	1.5764 [0.0757]

^a Sobre la distribución de estos contrastes véase tabla 2.2.

4. Especificación de un modelo estructural

4.1. ¿Por qué los modelos estructurales?

El enfoque VAR representa un planteamiento ateórico ya que no se cuestiona en ningún momento la exogeneidad o el grado de exogeneidad de las variables, un aspecto relevante si se desea analizar un posible orden de causalidad en las relaciones entre variables económicas. En el planteamiento estructural de la *Cowles Commission* la exogeneidad de las variables venía impuesta a priori, cuando, en realidad, su validez debería ser contrastada a partir de los datos, tal y como insiste el enfoque de la *London School of Economics*²⁰. Asimismo, la importancia de la exogeneidad es fundamental si deseamos reducir la dimensión de un modelo. En el caso de que una variable no sea exógena, esta debería ser modelizada dentro del sistema lo que podría invalidar ciertas restricciones tipo “cero” impuestas a priori bajo el enfoque estructural. En definitiva, la no exogeneidad puede concebirse como una necesidad de incrementar el tamaño o dimensión del modelo (número de ecuaciones) y viceversa.

En un enfoque VAR no se contemplan las relaciones contemporáneas entre las variables endógenas que si se recogen en un modelo estructural, siendo su principal característica diferenciadora frente a otros enfoques de modelización. De acuerdo con Wold (1954), en su crítica a los modelos estructurales, añade que la simultaneidad no se concibe en economía porque siempre existe cierto retardo entre la causa y el efecto, y esta es la razón de ser de los modelos VAR y los modelos recursivos. Sin embargo, dependiendo de la frecuencia de los datos, es posible que tal simultaneidad esté presente, dando sentido a los modelos estructurales.

Como ya apuntamos en la introducción, en un modelo estructural o modelo de ecuaciones simultáneas se hacen explícitas las relaciones contemporáneas entre las variables endógenas, pudiendo al mismo tiempo existir correlaciones contemporáneas entre los términos de perturbación del propio modelo. El segundo tipo de correlaciones se obtienen a partir de la propia muestra, mientras que las correlaciones entre variables conllevan una serie de restricciones adicionales, necesarias para que el modelo esté identificado, y que no pueden ser obtenidas a partir de los datos, sino de forma extramuestral, por ejemplo, a través de la Teoría Económica. En este sentido, la

²⁰ Véase, por ejemplo, el capítulo 5 de Favero (2001) dedicado a la metodología de la LSE.

metodología de la *Cowles Commission* imponía a priori las restricciones necesarias para lograr la identificación del modelo estructural sin contrastar posteriormente su validez²¹.

En relación con la identificación, si partimos del *SECM* en (5) e imponemos exactamente las restricciones necesarias para que dicho modelo esté identificado, la forma reducida restringida y no restringida coinciden²². Pero, si la información extramuestral permite imponer más restricciones que las estrictamente necesarias para identificar los parámetros estructurales, entonces el *SECM* estará sobreidentificado y la forma reducida restringida constituirá un caso particular del *UVAR* en forma de *VECM*. En este sentido, también se dice que el *UVAR* abarca al modelo estructural. Sin embargo, siguiendo a Espasa y Cancelo (1993), de ser correctas estas restricciones, el modelo estructural, podría llevarnos a una mejora de la modelización con respecto al enfoque VAR en términos de aumento de eficiencia de los estimadores y de capacidad de explicar la realidad. Incluso, como añaden estos autores, si el interés del analista es meramente predictivo, la posible ventaja del modelo estructural es la de proporcionar una forma reducida restringida, que de ser correcta, generará predicciones más eficientes que una forma reducida no restringida. En este caso, ¿sería factible que el modelo más simple abarque de forma parsimoniosa al modelo más general? Esto es, ¿podría plantearse la posibilidad de que el *SECM* pudiera recoger idénticos resultados a los del *UVAR* o en su caso al *PVAR*?

4.2. Especificación, estimación y validación del modelo estructural

Para llevar a cabo la especificación del modelo estructural dado por la ecuación (5), es necesario hacer un análisis previo de la identificación de la estructura a corto plazo²³, esto es, necesitamos que exista un número de restricciones tipo “cero” suficientes para que el modelo estructural esté sobreidentificado²⁴. Bajo este criterio, hemos especificado un modelo estructural imponiendo a priori las restricciones necesarias para que se garantice las condiciones necesarias y suficientes de identificación cuya estimación figura en la tabla 9. Asimismo, todos los coeficientes resultan significativamente distintos de cero al 5% haciendo uso del contraste de la *t* de *student*²⁵. Dado, que

²¹ El enfoque de la *London School of Economic* (LSE) explica precisamente el fracaso de esta metodología en que no se contrastaba la validez de la forma reducida restringida derivada de la forma estructural.

²² En este caso, no es factible contrastar la validez de la forma reducida a partir de los datos.

²³ No se plantea la identificación del largo plazo, dado que sólo hay un vector de cointegración.

²⁴ Si el número de restricciones es tal que el modelo estructural está identificado, entonces no podemos contrastar la validez de la forma reducida. La condición necesaria de identificación es que el número de restricciones nulas en una ecuación cualquiera debe ser mayor o igual a $G-1$ siendo G el número de variables modelizadas.

²⁵ Por limitaciones de espacio, no hemos podido incluir los errores estándar de las estimaciones.

no se descarta la simultaneidad entre todas las variables, hemos estimado dicho modelo por el método de información completa (*MC3E = Mínimos Cuadrados en Tres Etapas*) procediendo a su reducción o simplificación²⁶ hasta llegar a la especificación que figura en la tabla 9.

Tabla 9. Modelo Estructural estimado por MC3E

$$\begin{bmatrix} 1 & 0,036 & 0 \\ 0 & 1 & 0 \\ 0 & 0,047 & 1 \end{bmatrix} \begin{pmatrix} \Delta \text{vaan95}_t \\ \Delta \text{vaes95}_t \\ \Delta \text{asaocu}_t \end{pmatrix} = \begin{bmatrix} -0,03 \\ -0,02 \\ 0 \end{bmatrix} [\text{ecm}_{t-1}] + \begin{bmatrix} 2,208 & 0 & 0 \\ 0 & 1,936 & 0 \\ 0 & 0 & 1,675 \end{bmatrix} \begin{pmatrix} \Delta \text{vaan95}_{t-1} \\ \Delta \text{vaes95}_{t-1} \\ \Delta \text{asaocu}_{t-1} \end{pmatrix} +$$

$$+ \begin{bmatrix} -1,856 & 0 & 0 \\ 0 & -1,619 & -0,452 \\ 0 & 0 & -1,487 \end{bmatrix} \begin{pmatrix} \Delta \text{vaan95}_{t-2} \\ \Delta \text{vaes95}_{t-2} \\ \Delta \text{asaocu}_{t-2} \end{pmatrix} + \begin{bmatrix} 0,602 & 0 & 0 \\ 0 & 0,597 & 0,640 \\ 0 & 0 & 0,999 \end{bmatrix} \begin{pmatrix} \Delta \text{vaan95}_{t-3} \\ \Delta \text{vaes95}_{t-3} \\ \Delta \text{asaocu}_{t-3} \end{pmatrix}$$

$$+ \begin{bmatrix} 0 & 0 & -0,127 \\ 0 & 0 & -0,384 \\ 0 & 0 & -0,476 \end{bmatrix} \begin{pmatrix} \Delta \text{vaan95}_{t-4} \\ \Delta \text{vaes95}_{t-4} \\ \Delta \text{asaocu}_{t-4} \end{pmatrix} + \begin{bmatrix} \text{(ficticias)} \end{bmatrix} + \begin{bmatrix} \text{e}_{1t} \\ \text{e}_{2t} \\ \text{e}_{3t} \end{bmatrix}$$

Ecuación (1) $\Rightarrow \bar{R}^2 = 0,9578$; $DW = 1,88$ Ecuación (2) $\Rightarrow \bar{R}^2 = 0,9577$; $DW = 1,97$
Ecuación (3) $\Rightarrow \bar{R}^2 = 0,8865$; $DW = 1,67$
 $\text{ecm}_{t-1} = \text{vaan95}_{t-1} - 1,565486 \text{ asaocu}_{t-1} - 7,941964$
 $|\hat{\Omega}| = 6,33 * 10^{-16}$
 $k = 35$ (número de coeficientes incluyendo las variables ficticias)

Coefficientes de corrección de error $\hat{a} = \begin{bmatrix} -0,03 \\ -0,02 \\ 0 \end{bmatrix}$	Coefficientes de ajuste $\hat{\alpha} = \Phi^{-1} \hat{a} = \begin{bmatrix} -0,029276 \\ -0,02 \\ 0,0009331 \end{bmatrix}$
--	--

A continuación, contrastamos la validez de las restricciones impuestas a priori, siguiendo el enfoque de la *London School of Economics (LSE)*, esto es, contrastamos si la forma reducida restringida dada por el modelo estructural que figura en la tabla anterior soporta las correspondientes restricciones que no impone la forma reducida no restringida, esto es, el *PVAR*.

Hemos aplicando un contraste de razón de verosimilitud basado en el logaritmo de las funciones de verosimilitud de ambos modelos. Dicho contraste puede expresarse también en función de los

²⁶ Hemos eliminado aquellas variables que han resultado redundantes mediante los contrastes adecuados.

determinantes de las matrices de varianzas y covarianzas de los residuos, y viene dada por (véase *Patterson, 2000*):

$$\chi^2_{obs.} = T \left[\ln |\hat{\Omega}_0| - \ln |\hat{\Omega}| \right] \sim \chi^2 (k - k_0) \quad (7)$$

Donde T es el tamaño muestral, $|\hat{\Omega}_0|$ y $|\hat{\Omega}|$ son los determinantes de las matrices de varianzas y covarianzas de los residuos del modelo estructural y el *PVAR* en forma de *VECM* respectivamente, k y k_0 son el número de parámetros en el modelo no restringido y restringido respectivamente. Los resultados que figuran en la tabla 10 muestran que tal reducción es viable con un nivel de significación del 5%.

Tabla 10. Contraste de abarcamiento parsimonioso

	PVAR $\Rightarrow$ SECM_1
$\chi^2_{obs.} = T \left[\ln \hat{\Omega}_0 - \ln \hat{\Omega} \right]$	34,74
$\chi^2 (63-35=28) (\alpha=0,05)$	41,34
H₀	No se rechaza

4.3. Contrastes de exogeneidad: ¿se puede reducir la dimensión del *SECM*?

Tal y como se recoge en *Charemza y Deadman (1997, págs. 213-247)* la metodología VAR puede presentar limitaciones importantes si se opera con un número excesivo de variables dado que el elevado número de parámetros puede hacer inviable la estimación por falta de grados de libertad. No es nuestro caso, al trabajar con sólo tres variables, pero si puede llegar a ser relevante si deseamos aplicar el enfoque, por ejemplo, a cinco o más variables. En estas situaciones, puede resultar interesante trabajar con dos subsistemas de ecuaciones. Uno formado por el conjunto de variables endógenas a partir de las cuales se especifica y estima el vector de cointegración, y otro formado por aquellas variables fuertemente, o al menos, débilmente exógenas, con respecto a los parámetros de interés que intervienen en el subsistema anterior de forma condicional pero cuya modelización es independiente. Esto nos permitirá reducir el tamaño inicial del modelo. Pero ¿qué

papel puede jugar el grado de exogeneidad de las variables de un modelo en relación a su tamaño? Dicho de otra manera, dado que nuestro interés es estimar la primera ecuación y predecir con ella el valor añadido bruto de Andalucía en el sector primario, ¿hasta que punto sería posible estimar dicha ecuación prescindiendo del resto del modelo? ¿Podríamos igualmente predecir las variables explicativas de la primera ecuación sin necesidad de predecir previamente el VAB agrícola de Andalucía?

Desde el punto de vista de la estimación del modelo estructural, nos interesa estudiar **el grado de exogeneidad débil** y desde el punto de vista de la predicción **el grado de exogeneidad fuerte**²⁷. En nuestro modelo la matriz Π viene dada por:

$$\Pi = \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} (\beta_{11} \quad \beta_{21} \quad \beta_{31}) \quad (7)$$

donde previamente hemos normalizado el vector de cointegración β haciendo $\beta_{11} = 1$ e imponiendo la restricción $\beta_{21} = 0$ (véase tabla 6). Contrastar la exogeneidad débil con respecto a los parámetros de interés de la primera ecuación equivale a contrastar la hipótesis nula $H_0: \alpha_i = 0$ para $i = 2$ y 3 . Bajo la hipótesis nula cierta el siguiente estadístico WETS (*Weak exogeneity test statistic*)²⁸ se distribuye como una χ^2 ($r \times s$) donde $r = 1$ y $s = 2$ (número de variables débilmente exógenas):

$$\text{WETS} = T \sum_{i=1}^r \ln \left[\frac{(1 - \hat{\lambda}_i^*)}{(1 - \hat{\lambda}_i)} \right] \sim \chi^2 (r \times s) \quad (8)$$

donde $\hat{\lambda}_i^*$ son los valores propios estimados bajo el ECM restringido y $\hat{\lambda}_i$ los valores propios estimados bajo el ECM no restringido. Los resultados del contraste figuran en la siguiente 11 donde además hemos incorporado el contraste conjunto de que ambas variables son débilmente exógenas. Los resultados muestran que la variable de España no es débilmente exógena y sí lo es, en cambio, el indicador de empleo. Hemos repetido el contraste imponiendo además que el VAB de España no figura en el vector de cointegración (esto es, $\beta_{21} = 0$). Los resultados coinciden bajo los contrastes individuales, aunque no ocurre lo mismo bajo el contraste conjunto, ya que ahora no se rechaza la hipótesis nula de que ambas variables son débilmente exógenas. En este último caso, podríamos estimar la primera ecuación por MCO manteniendo aún las buenas propiedades de los estimadores. Nótese además, que al hacer el contraste se reestima el vector de cointegración. Los resultados del vector reestimado son prácticamente idénticos a los que figuran en la tabla 2.6 confirmándose que

²⁷ Si una variable no es exógena en sentido débil tampoco lo será en sentido fuerte.

²⁸ Véase Johansen (1992) para una definición e implicación de la exogeneidad débil.

los datos pueden admitir el grado de exogeneidad débil en ambas variables. Confirmado el grado de exogeneidad débil, quedaría por contrastar a efectos de predicción si ambas variables son, además, exógenas en sentido fuerte, lo que supone en este caso, contrastar la *No Causalidad de Granger*²⁹. La tabla 12 muestra que ambas variables son fuertemente exógenas ya que los retardos del VAB de Andalucía no son conjuntamente explicativas ni del VAB de España ni del cociente del empleo al ser la probabilidad superior el nivel de significación del 5%.

Tabla 11. Contraste de exogeneidad débil

$\beta' = (1 \ \beta_{21} \ \beta_{31})'$		
Contrastes individuales	vaes95 es exóg. débil $\alpha' = (\alpha_1 \ 0 \ \alpha_3)'$	asaocu es exóg. débil $\alpha' = (\alpha_1 \ \alpha_2 \ 0)'$
WETS =	6,97	0,16
$\chi^2 (1) (\alpha = 0,05) =$	3,14	3,14
Contraste conjunto	vaes95 y asaocu son exóg. en sentido débil $\alpha' = (\alpha_1 \ 0 \ 0)'$	
WETS =	7,09	
$\chi^2 (2) (\alpha = 0,05) =$	5,99	

$\beta' = (1 \ 0 \ \beta_{31})'$		
Contrastes individuales	vaes95 es exóg. débil $\alpha' = (\alpha_1 \ 0 \ \alpha_3)'$	Asaocu es exóg. débil $\alpha' = (\alpha_1 \ \alpha_2 \ 0)'$
WETS =	7,31	0,22
$\chi^2 (2) (\alpha = 0,05) =$	5,99	5,99
	$\hat{\beta}' = (1 \ 0 \ -1,5765)'$	$\hat{\beta}' = (1 \ 0 \ -1,5687)'$
Contraste conjunto	vaes95 y asaocu son exóg. en sentido débil $\alpha' = (\alpha_1 \ 0 \ 0)'$	
WETS =	7,53	
$\chi^2 (3) (\alpha = 0,05) =$	7,81	
	$\hat{\beta}' = (1 \ 0 \ -1,5721)'$	

Tabla 12 Contraste de No Causalidad de Granger

Dependent variable: D(LVAES95T)				Dependent variable: D(LASAOCUT)			
Exclude	Chi-sq	df	Prob.	Exclude	Chi-sq	df	Prob.
D(LVAAN95)	4.401057	4	0.3544	D(LVAAN95)	4.732809	4	0.3158

²⁹ Véase Granger (1969) para el concepto de causalidad en las relaciones entre variables económicas.

Si comparamos los resultados de los contrastes de exogeneidad con los obtenidos bajo el modelo estructural (*SECM*), podemos observar que $\hat{\alpha} = \hat{\Phi}^{-1}\hat{a}$ tiene un cero en la tercera fila confirmándose igualmente, la exogeneidad débil de *asaocu*. Sin embargo, el coeficiente $\hat{\alpha}_2 = -0,02 \neq 0$ por lo que, en principio, el modelo estructural no confirma la exogeneidad débil del VAB de España (*vaes95*)³⁰. Por otro lado, la variable *asaocu* que resulta ser débilmente exógena, no interviene de forma contemporánea en la ecuación de interés. En cualquier caso, si analizamos los resultados globales de la tabla 11 y comparamos las estimaciones de la primera ecuación por MCO que figuran en la tabla 13 con las del modelo estructural, ¿no sería factible admitir que los datos confirman que ambas variables son exógenas en sentido débil? De hecho, de ser ciertas las hipótesis de exogeneidad, no sería necesario el plantemiento multiecuacional de cointegración de *Johansen*, es decir, bastaría un enfoque uniecuacional para estimar el largo³¹ para después plantear un modelo condicionado a las variables exógenas en la especificación del corto plazo, esto es, el modelo de la tabla 2.18, reduciendo de esta manera la dimensión del modelo.

Tabla 13 Modelo uniecuacional del VAB de Andalucía estimado por MCO

$$\Delta \text{vaan95}_t = -0,03 \text{ ecm}_{t-1} + 0,035 \Delta \text{vaes95}_t + 2,21 \Delta \text{vaan95}_{t-1} - 1,86 \Delta \text{vaan95}_{t-2} + 0,60 \Delta \text{vaan95}_{t-3} - 0,13 \Delta \text{asaocu}_{t-4} + (\text{ficticias}) + e_{1t} \quad \bar{R}^2 = 0,997 ; DW = 1,88 \quad \text{ecm}_{t-1} = \text{vaan95}_{t-1} - 1,565486 \text{ asaocut}_{t-1} - 7,941964$$

(-8,4) (2,3) (64,4) (-36,5) (20,9)

(-5,6)

5. Evaluación de la capacidad predictiva

Por último, hemos llevado a cabo la evaluación de la capacidad predictiva del modelo uniecuacional, obteniendo predicciones para el período 2001-2002 (ocho trimestres). Para ello, necesitamos conocer los valores de las exógenas. A modo de ejercicio, hemos considerado dos escenarios: el “optimista” donde conocemos los valores reales de las exógenas (las predicciones se recogen con el nombre *VAB_ECM_1* y el “pesimista”, que sería el caso real, donde desconocemos el valor de las mismas. Para este escenario, podemos hacer uso del modelo estructural (*SECM*) con el fin de obtener las predicciones del VAB de España y de la variable de empleo. Las predicciones del VAB andaluz bajo este escenario se recogen con el nombre *VAB_ECM_2*. Finalmente, y a efectos de comparar el modelo uniecuacional con el multiecuacional, hemos recogido también las predicciones dinámicas que proporciona el modelo estructural del VAB andaluz bajo el nombre de

³⁰ De hecho, el contraste individual de exogeneidad tampoco confirmaba la exogeneidad débil de dicha variable.

³¹ Por ejemplo, a partir del enfoque bietápico de *Engle-Granger* (1987) o resolviendo un modelo dinámico (véase por ejemplo *Patterson* (2000) o *Doornik and Hendry* (2001) se mantienen las propiedades asintóticas de los estimadores.

VAB_SECM. En el gráfico 1 comparamos las predicciones del VAB regional bajo los dos escenarios “optimista” y “pesimista”. Bajo el supuesto de conocer las exógenas, las predicciones son muy acertadas, pero tampoco son despreciables las que se obtienen prediciendo las exógenas, de hecho, ambas predicciones son casi idénticas durante los primeros cinco trimestres.

Gráfico 1. Predicciones uniecuacionales

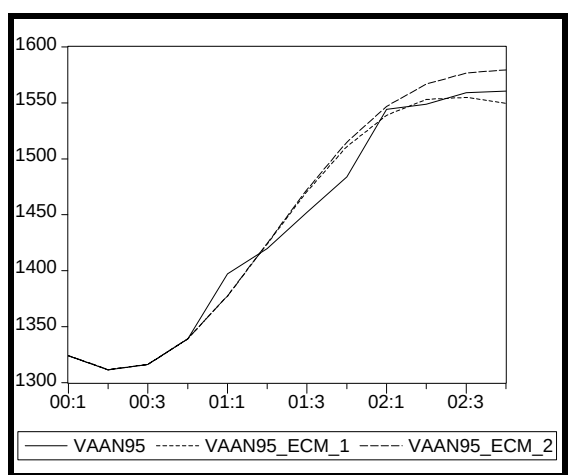
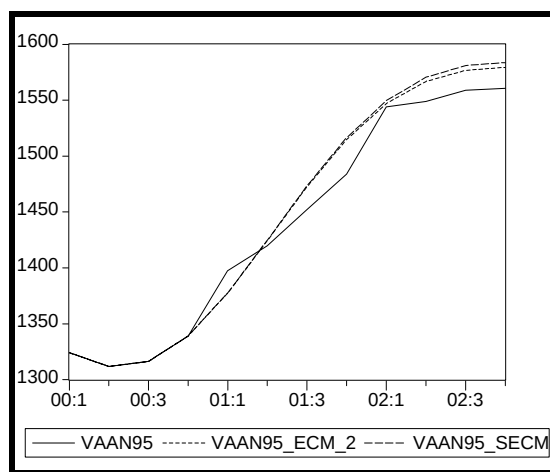


Gráfico 2. Predicciones del modelo uniecuacional y multiecuacional



En el gráfico 2 comparamos las predicciones del modelo uniecuacional y del modelo multiecuacional, prediciendo en ambos casos las exógenas bajo el modelo estructural. La principal conclusión que extraemos de este análisis es que las dos ecuaciones adicionales del modelo estructural no aportan ninguna información relevante a efectos de estimación de la ecuación de interés.

7. BIBLIOGRAFÍA

- Anderson, T.W. (1984). *An introduction to Multivariate Statistical Analysis*, 2nd edition, New York: John Wiley & Sons.
- Banerjee, A., Dolado, J.J., Galbraith, J.W. and Hendry, D.F. (1993). *Co-integration, Error Correction and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press.
- Charemza, W.W. and Deadman, D.F. (1997). *New directions in Econometric Practice: General to Specific Modelling*, Cointegration and Vector Autoregression. Edward Elgar.
- Clements, M. P. and Mizon, G.E. (1991): “Empirical analysis of macroeconomic time series: VAR and structural models”, *European Economic Review* 35, 887-932. North Holland.

- Doornik, J. and Hansen, D.F. (1994). *A practical test of multivariate normality*. Oxford. Nuffield College.
- Doornik, J.A. (1996). Testing vector autocorrelation and heteroscedasticity in dynamics models, Nuffield College.
- Doornik, J.A. and Hendry, D.F. (2001). *Modelling Dinamic Systems using PcGive*. Volume II. Timberlake Consultants LTD. London. United Kingdom.
- Espasa, A. y Cancelo J.R. (1993). *Métodos Cuantitativos para el análisis de la coyuntura económica*. Alianza Editorial, S.A. Madrid.
- Favero, Carlo A. (2001). *Applied Macroeconometrics*. Oxford University Press Inc. New York.
- Engle, R.F.y Granger, C.W.J. (1987): “Cointegration and Error Correction: Representation, Estimation and Testing”, *Econometrica*, 55.
- Goodfrey, L.G. (1988). *Misspecification test in Econometrics*. Cambridge: Cambrige University Press.
- Granger, C.W.J. (1969): “Investigating casual relations by econometric models and cross-spectral methods”, *Econometrica* 37, pp. 24-36.
- Johansen, S. (1988): “Statistical Analysis of Cointegration Vectors”, *Journal of Economic Dinamics and Control*, Vol. 12, pp. 231-254.
- Johansen, S. (1992a): “Determination of Cointegration Rank in the Presence of a Linear Trend”, *Oxford Bulletin of Edonomics and Statistics*, Vol. 54, pp. 383-397.
- Johansen, S. (1992b): “Testing Weak Exogeneity and the Order of Cointegration in UK Money Demand Data”, *Journal of Policy Modelling*, Vol. 14, pp. 313-334.
- Maddala, G.S. and KIM, I.M. (1998): *Unit Roots, Cointegration and Structural Change*, Cambridge University Press, Cambridge.
- Osterwald-Lenum, M. (1992). *A note with quantiles of the asymptotic distribution of the ML cointegration rank test statistics*, Oxford Bulletin of Economics and Statistics, 54, 461-472.
- Pesaran, M.H. and Smith, R. (1994): *Theory and Evidence in Economics*. Journal of Econometrics.
- Patterson, K. (2000). *An Introduction to Applied Econometrics: A time series approach*. MacMillan Press LTD. London.
- Urbain, J.-P. (1995): “ Partial versus full system modelling of cointegrated systems: An empirical illustration”, *Journal of Econometrics*, 69, 177-210.
- Watson, M.W. (1994): “Vector Autoregression and Cointegration” in *The Hadbook of Econometrics*, Vol. IV, North-Holland, Amsterdam, pp. 2844-2915.
- Wold, H. (1954): “Causality and econometrics”, *Econometrica* 22, pp. 162-177.