

**Guía básica de temas del 2º parcial: Econometría. 4º Curso de Economía. USC.  
Curso 2009-2010. Profesora María-Carmen Guisán**

**Tema 6. Econometría Aplicada. Introducción.**

**Lectura recomendada:** [Documento 95: Causalidad y desarrollo económico: Análisis econométrico de los países de la OCDE, 1965-2005.](#)

**6.1. Clasificación de las relaciones causales en los modelos econométricos.**

**Cuadro 1.** Relaciones de causalidad en modelos econométricos

1. Causalidad directa unidireccional de sentido fijo: a) X causa a Y ó b) Y causa a X
2. Causalidad directa unidireccional de sentido variable.  
Existe en modelos de desequilibrio: X a veces causa a Y, e Y a veces causa a X
3. Causalidad directa bidireccional (X causa a Y e Y causa a X):  
a) sin relación contemporánea (modelo recursivo):  $Y_t = f(X_t)$ ,  $X_t = f(Y_{t-1})$   
b) con relación contemporánea (interdependencia):  $Y_t = f(X_t, Z_{1t})$ ,  $X_t = f(Y_t, Z_{2t})$
4. Causalidad indirecta: X e Y no están relacionadas directamente pero influyen en, o están influenciadas por, otra variable Z.  
Ejemplos: a)  $Z = F(X, \text{otras variables})$   $Y = F(Z)$ , b)  $Y = F(Z)$ ,  $X = F(Z)$
5. Ausencia de causalidad: a) incorrelación, b) correlación casual

**6.2. Estimación de modelos temporales, atemporales y mixtos.**

Muchos modelos econométricos relacionan variables en términos reales y por lo tanto en los modelos temporales y mixtos tenemos que expresar las variables a precios constantes, para lo cual debemos hacer uso de las fuentes estadísticas correspondientes. En general las variables de diversos países deben ir expresadas en una moneda común y en ese sentido podemos utilizar tipos de cambio (TC) o las paridades de poder de compra (PC), dependiendo de cual sea el método que proporcione comparaciones más realistas según las variables utilizadas. Desde el punto de vista metodológico en los modelos temporales en general se contrasta la posible presencia de autocorrelación, en los modelos atemporales la de heterocedasticidad y en los modelos mixtos se tienen en cuenta ambas cuestiones. En cuanto a la estabilidad en general se contrasta la homogeneidad de coeficientes entre distintas unidades atemporales (países, regiones u otras) en modelos atemporales y mixtos. También puede contrastarse la estabilidad temporal en modelos temporales y mixtos.

**6.3. Introducción a los modelos de Consumo, Producción, Inversión, Empleo y Desarrollo**

*Función de producción:* La función de producción es una relación importante que a veces se utiliza, a nivel macroeconómico, para explicar el PIB real de un país. Presenta una elevada bondad del ajuste pero no siempre sirve para explicar el nivel de esta variable, sino que en ocasiones explica el nivel utilizado del capital disponible. Algunos autores han utilizado la función para explicar el nivel de empleo deseado ( $L^*$ ), y otros la consideran útil para explicar el nivel utilizado del capital físico (K). Existen otras relaciones, tanto por el lado de la oferta como de la demanda que también contribuyen a explicar el crecimiento del PIB real como se analiza en 10.4.

*Función de consumo:* El consumo privado de un país depende fundamentalmente de la renta disponible. La renta familiar disponible depende a su vez fundamentalmente del PIB, de forma que la función a veces se simplifica haciendo depender el Consumo directamente del PIB.

*Modelos de inversión:* La inversión es en general una condición necesaria, aunque no siempre suficiente, para garantizar el crecimiento del PIB real. Depende en gran medida de la evolución del PIB en el período anterior y de diversas variables que proporcionan incentivos a la inversión.

*Modelos de empleo:* El empleo de un país, o de una región, depende fundamentalmente de la evolución del Valor Añadido real no agrario (QNA). En general la Industria influye positivamente, junto a otras variables, en el desarrollo de los sectores no agrarios, de forma que el desarrollo industrial a veces genera más empleos indirectos, en los sectores de construcción y servicios, que empleos directos en el propio sector industrial. El aumento de QNA depende en gran medida de los incentivos a la inversión productiva.

*Modelos de desarrollo socio-económico:* Explican la evolución de la renta real per cápita y de otras variables que tienen una importancia destacada en la calidad de vida y en el desarrollo económico y social como el nivel educativo, los valores sociales, la calidad del gobierno, el medio ambiente y otras.

## Tema 7.- Modelos dinámicos causales

Lectura recomendada: [Documento 96: Modelos econométricos dinámicos y desarrollo económico: Análisis del salario real, la productividad y el empleo en los países de la OCDE, 1965-2005.](#)

La clasificación de modelos dinámicos en sentido estricto y en sentido amplio que aparecen en el estudio “Crecimiento económico en los países de la OCDE” y disponibles en el Documento nº 96 es importante.

El concepto de modelo dinámico en sentido estricto va unido al “efecto propagación” entendiendo que dicho efecto existe cuando el incremento o disminución de una variable explicativa, X, en un momento del tiempo, afecta a Y no sólo en el momento t sino que el efecto se transmite a varios períodos futuros (t+1, t+2, ....). Este efecto se propaga en el tiempo a veces de forma amortiguada hasta desaparecer pasados varios períodos, pero otras veces tiene un efecto duradero estable o incluso creciente en el tiempo. Por ejemplo las ventajas de industrialización que tuvieron países como USA o Gran Bretaña en el siglo XIX han tenido efectos de propagación que se han transmitido hasta la actualidad, ya que su mayor nivel de renta per cápita en comparación con otros países les ha permitido invertir más en capital físico y humano y eso ha dado lugar a nuevos incrementos a lo largo de todo este tiempo.

Es especialmente importante el Cuadro 2 de esta guía, que corresponde al apartado 4 del Cuadro 2 del Documento 96, e indica la *forma de expresar la relación* en un modelo dinámico en sentido estricto, distinguiendo 4 formas: a) en niveles b) en incrementos simple c) en incrementos con efecto “corrección del error”, o efecto CE, y d) dinámico mixto.

Cuadro 2. Clasificación por la forma de expresar la relación funcional entre las variables:

- a) *dinámico en niveles*:  $Y=f(X, Y(-1))$
- b) *dinámico en incrementos simple, sin efecto CE*:  $Y=f(DY(-1), DY(-2), DX, DX(-1), DX(-2))$
- c) *dinámico en incrementos con efecto CE, sin y con inclusión de DX*:  
 c1)  $DY=f(DY(-1), DY(-2), DX(-1), DX(-2), \hat{u}(-1))$ ; c2)  $DY=f(DY(-1), DY(-2), DX, DX(-1), DX(-2), \hat{u}(-1))$
- d) *dinámico mixto, en nivel y en incremento*:  $Y=f(DX, Y(-1))$

La letra D indica *incremento* de forma que por ejemplo el modelo dinámico mixto en el que  $Y= f(DX, Y(-1))$ , significa que el valor de Y en el momento t depende del valor de Y en (t-1) y del incremento que haya tenido la variable X al pasar del período (t-1) al período t. Es decir  $DX= X_t - X_{t-1}$ . Si el modelo tiene varias variables explicativas (k) aparecerán los incrementos de cada una de ellas:  $DX_1, DX_2, \dots, DX_k$ .

En general los modelos que mejores resultados proporcionan son el dinámico mixto y el modelo CE en su modalidad “*con inclusión de DX*”, es decir cuando considera el efecto de un cambio “contemporáneo” en X. El modelo CE es utilizado por algunos autores sin ese efecto contemporáneo, incluyendo sólo retardos de DX, pero en general el modelo CE sólo funciona bien cuando incluye el incremento de X en el momento t, es decir  $DX = X_t - X_{t-1}$ .

Un ejemplo de modelo dinámico mixto interesante es el de la página 114 del libro “OCDE 1” (página 14 del [Documento 43](#)). Se trata de una ecuación de empleo mediante un modelo dinámico mixto en el que se ha incluido una variable ficticia para cada país y además se ha corregido la autocorrelación.

Es muy importante saber comparar la bondad del ajuste de los distintos modelos dinámicos, ya que en algunos casos la variable explicada está en niveles (Y) como ocurre en los apartados a) y d), pero en otros están en incrementos (DY) como ocurre en los apartados b) y c). Para compararlos hay que observar los valores de la SCE de cada uno de los modelos y no el  $R^2$  que resulta automáticamente de la regresión pues la Suma de Cuadrados Totales que se utiliza para calcular el coeficiente de determinación es muy distinta para las variables en niveles y para las variables en incrementos.

La SCE sí se puede comparar, porque es fácil de comprobar que la Suma de Cuadrados de Errores que un modelo cualquiera presenta para Y es la misma que ese modelo implica para DY, ya que cada residuo para la variable en niveles es igual al residuo para la variable en incrementos. En efecto, si denominados  $y_t$  al valor verdadero de Y en el momento t, e  $yf_t$  al valor estimado por el modelo tenemos que:

Residuo del modelo para y:  $e_t(y) = y_t - yf_t$

Residuo del modelo para Dy:  $e_t(Dy) = Dy - Dyf = (y_t - y_{t-1}) - (yf_t - yf_{t-1}) = y_t - yf_t$

Por lo tanto si en un ejercicio nos dicen que se ha estimado un modelo en niveles con una SCE= 2.5 y un  $R^2=0.98$ , y un modelo CE con una SCE=1.8 y un  $R^2=0.65$ , y nos preguntan cual de ellos proporciona un mejor ajuste, en principio hay que responder que el mejor ajuste corresponde al modelo CE ya que, en ese caso, su SCE es mucho menor. Para tener en cuenta el tamaño muestral y el número de parámetros, conviene efectuar la comparación de estos modelos mediante el % S.E. (el porcentaje que el Error Standard supone respecto a la media de  $y_t$ ), criterio que depende en gran medida de cual de ellos tenga el menor valor de la SCE. La bondad del ajuste, teniendo en cuenta el menor valor del % S.E, es un criterio importante pero no el único a tener en cuenta a la hora de seleccionar un modelo, como comentaremos en el ejemplo práctico.

### **Tema 8.- Regresores ficticios, datos de panel y estabilidad muestral**

Lecturas recomendadas: [Capítulo 6 del libro de Econometría, Guisán\(1997\)](#) o [contenidos similares](#).

Los conceptos más importantes que es preciso saber son: 1) inclusión de variables ficticias para expresar diferencias en la ordenada en el origen, 2) contrastes F de homogeneidad de parámetros, 3) inclusión de variables ficticias multiplicativas. 4) Estimación consistente de White en caso de heterocedasticidad.

1) Las fórmulas (6.2) y (6.4) de las páginas 152 y 153 del libro de Econometría (p.152 y 153) expresan dos formas de plantear el uso de ficticias ambas situaciones. Podemos elegir cualquiera de ellas: En la situación a) tenemos que tener en cuenta que el parámetro que acompaña a cada variable ficticia ( $\alpha_j$  para  $j=1,2,...,p$ ) es la “ordenada en el origen” de la modalidad correspondiente, mientras que en la opción b) la ordenada en el origen del grupo 1 es  $\beta_0$ , la del grupo 2 es  $(\beta_0 + \alpha_2)$ , y en general la del grupo j-ésimo, para j distinto de 1, es  $(\beta_0 + \alpha_j)$ . La opción a) es la que utiliza el programa econométrico cuando seleccionamos la opción “efectos fijos”. La opción b) es mejor para contrastar de forma rápida si existe diferencia significativa o no en la ordenada en el origen de los distintos grupos, ya que basta con ver si las  $\alpha_j$  son significativamente diferentes de cero.

RECORDAR: Conceptos de “significatividad y precisión”, y de “aceptación, rechazo e incertidumbre”.

2) Contrastes de homogeneidad de parámetros (también llamados de estabilidad muestral). Son tres contrastes que denominados  $F_1$ ,  $F_2$  y  $F_3$ , cuyo cálculo se basa en la SCE de diferentes modelos. Dichos contrastes se resumen en el cuadro 6.2 de la página 163 del libro de Econometría y, las SCE que son necesarias para calcularlos se resumen en el cuadro 6.1 de la página 160:

$S_1$  = SCE suma de las  $SCE_i$  obtenidas mediante una regresión para cada grupo

$S_2$  = SCE de la regresión conjunta con efectos fijos (ordenada diferente)

$S_3$  = SCE de la regresión conjunta sin efectos fijos (ordenada común o sin ordenadas).

3) Inclusión de variables ficticias multiplicativas: Se incorporan para tener en cuenta que hay un cambio en un determinado parámetro, bien a través del tiempo o del espacio. Ejemplo:

$$W_{it} = \beta_0 + \beta_1 PM_{it} + \beta_2 W_{i,t-1} + \alpha_1 D1_{it} PM_{it} + \varepsilon_{it}$$

donde D1 es una ficticia que vale uno en las observaciones de Alemania y cero en los otros dos países. Podemos contrastar, con la t de Student, si el parámetro  $\alpha_1$  es significativamente distinto de cero. Si se rechaza la nulidad de ese parámetro podemos afirmar que no hay diferencia significativa entre Alemania y los otros dos países en este sentido.

E-Views: POOL

### **Tema 8. Contrastes de causalidad y cointegración**

El análisis de causalidad se puede abordar desde los test de causalidad Granger(1969), el análisis de cointegración y otros contrastes de causalidad y especificación propios de niveles más avanzados de Econometría. A nivel introductorio el test de Granger(1969) es uno de los más conocidos y se basa en analizar la relación causal bilateral entre dos variables Y y X en función de un modelo VAR con algunos retardos (por ejemplo un VAR(2), el cual incluye como variables explicativas  $Y(-1)$ ,  $Y(-2)$ ,  $X(-1)$  y  $X(-2)$ . El test es interesante pero en ocasiones presenta resultados poco concluyentes debido a que no tiene en cuenta la relación contemporánea entre ambas variables y a que la introducción de varios retardos provoca mucha multicolinealidad. El test de Hausman, que no se estudia en este curso por ser propio de niveles más avanzados, se basa en un sistema multiecuacional con interdependencia, tiene en cuenta la relación contemporánea entre las variables y permite también analizar si la relación causal es unilateral o bilateral. Un estudio más detallado de estos contrastes se realiza en la asignatura Econometría Aplicada al Sector Público optativa de 5º curso de Economía.

El análisis de cointegración consiste en examinar si la perturbación de la ecuación que relaciona a Y con X es estacionaria. Si la perturbación está autocorrelacionada se precisa que el coeficiente  $\rho$  sea en valor absoluto menor que 1. Dicho análisis puede realizarse mediante el test ADF, aumentado de Dickey y Fuller, o mediante otros tests, pero presenta muchas limitaciones en todas sus versiones:

1) cointegración no es un concepto idéntico al de causalidad, y por lo tanto puede haber relaciones causales que no cointegran por un pequeño problema de especificación (por ejemplo la relación entre el empleo de un sector y el valor añadido de ese sector puede no cointegrar si la relación está expresada mediante un modelo en niveles y puede cointegrar si la expresamos mediante un modelo dinámico mixto) y también puede haber relaciones no casuales que cointegran por casualidad.

2) Los tests de causalidad presentan muchas situaciones de incertidumbre, especialmente cuando  $0.90 < p < 0.99$  y por lo tanto son con frecuencia poco prácticos para determinar si hay cointegración o no.

3) el análisis de cointegración no permite analizar el sentido de la causalidad. Este test se realiza mediante la instrucción UROOT del programa E-Views y su resultado es favorable a la existencia de cointegración cuando el valor del estadístico t está a la izquierda de los valores críticos de MacKinnon que proporciona el programa en la cola izquierda de la distribución.

E-Views: Causalidad: CAUSE. Cointegración: UROOT

### **Tema 10: Modelos especiales: ARIMA. No lineales. Multiecuacionales. Otros.**

Lecturas recomendadas: [Capítulos 8 y 9 del libro de Econometría de Guisán\(1997\)](#) o [contenidos similares](#).

#### **10.1. Modelos dinámicos no causales: Modelos ARIMA**

Los modelos dinámicos no causales, no son dinámicos en sentido estricto y generalmente sus resultados son peores (peor capacidad predictiva por ejemplo). Por lo que respecta al modelo ARIMA el capítulo 9 del libro de Econometría presenta una síntesis muy completa, de la cual es interesante conocer las situaciones que se pueden presentar según el Cuadro 9.1 de la página 216, y la sección 9.4.2 que figura a continuación de dicho cuadro, en la que se describe la fase de estimación y chequeo. En un examen se puede preguntar que se escriba un modelo ARIMA(2,1,0) por ejemplo. Entonces hay que saber aplicar la fórmula general y poner dos parámetros autorregresivos ( $p=2$ ), diferenciar una vez la variable ( $d=1$ ) y ningún parámetro en la parte llamada de “medias móviles” correspondiente a retardos de la perturbación ( $q=0$ ).

Las fases de elaboración de un modelos ARIMA son las siguientes: 1) Identificación. b) Estimación. 3) Chequeo. 4) Predicción.

E-Views: IDENT, LS MODEL+M(q), BOX-LJUNG

#### **10.2. Modelos no lineales**

Según se indica en la bibliografía recomendada es frecuente que los modelos econométricos no sean lineales pero sean fácilmente linealizables, como es el caso de las funciones potenciales y exponenciales. Así por ejemplo en el caso de la función de producción Cobb-Douglas con parámetro de eficiencia, linealizamos mediante logaritmos y podemos aplicar mínimos cuadrados ordinarios:

LS  $\log(Q)$  c  $\log(KD)$   $\log(L)$  TI

Donde Q es la producción, KD el capital disponible, L es el empleo y TI es el tiempo. También podemos generar las variables  $LQ=\log(Q)$ ,  $LKD=\log(KD)$  y  $LL=\log(L)$  y aplicar:

LS LQ c LKD LL TI

Además podemos estimar la función no lineal utilizando mínimos cuadrados no lineales, lo cual es especialmente útil en el caso de funciones que no sean fácilmente linealizables. En el caso de la función de producción Cobb-Douglas expresaríamos:

NLS  $Q=c(1)*KD^{c(2)}*L^{c(3)}*\exp(c(4)*TI)$

Donde el símbolo \* indicado multiplicación, ^ indicada potencia, y exp es el número e, y los parámetros son  $c(1)$ ,  $c(2)$ ,  $c(3)$ ,  $c(4)$ .

La relación potencial se aplica con frecuencia a los modelos de demanda de un producto en los que, entre las variables explicativas, figura el precio del producto si la elasticidad cantidad/precio es estable. Es importante saber comparar la bondad del ajuste entre distintas funciones teniendo en cuenta el  $R^2$  respecto a la misma variable “y”, así como el % S.E. de dicha variable.

### 10.3. Modelos multiecuacionales

Un modelo multiecuacional está constituido por 2 o más ecuaciones que se relacionan entre sí de alguna manera, y que explican a 2 o más variables endógenas. Cuando existen relaciones bilaterales contemporáneas entre 2 o más variables del modelo multiecuacional se dice que el modelo presenta interdependencia y en ese caso el método más utilizado de estimación es el de Mínimos Cuadrados en dos etapas, MC2E, el cual se denomina en inglés TSLS o 2SLS (Two Stages Least Squares). Para estimar con el programa E-Views un modelo en niveles en el que  $Y_1$  depende de  $Y_2$  y de  $X_1$  e  $Y_2$  depende de  $Y_1$ , de  $X_2$  y de  $Y_2(-1)$  expresamos las siguientes órdenes:

TSLS  $Y_1 \ Y_2 \ X_1 \ @ \ X_1 \ X_2 \ Y_2(-1)$       TSLS  $Y_2 \ Y_1 \ X_2 \ @ \ X_1 \ X_2 \ Y_2(-1)$

Las variables que figuran antes del signo “arroba”, @, son las incluidas en cada ecuación del modelo y después del signo @ son todas las predeterminadas del sistema (exógenas y endógenas retardadas que aparezcan como explicativas en cualquier ecuación).

### Ejemplos de modelos multiecuacionales lineales o no lineales que se explican en clase.

Modelo 1. Modelo del PIB basado en la demanda agregada

$$(1.1) \text{ PIB}^d = C + G + IS + EXP - IMP$$

$$(1.2) C = f(D(\text{PIB}), C(-1))$$

$$(1.3) IS = f(D(\text{PIB}), IS(-1))$$

$$(1.4) D(\text{PIB}) = \text{PIB} - \text{PIB}(-1)$$

Modelo 2. Modelo del PIB basado en la oferta de Inputs primarios:

$$(2.1) \text{ PIB}^{s1} = f(KD, L, t)$$

$$(2.2) L = f(Q/W, L(-1))$$

$$(2.3) KD = KD(-1) + IN(-1)$$

$$(2.4) IN(-1) = f(\text{PIB}(-1), \text{Factores externos (préstamos e inversión)})$$

Modelo 3. Modelo del PIB basado en la oferta de inputs intermedios y relaciones intersectoriales

$$(3.1) \text{ PIB}^{s2} = QA + QI + QB + QS$$

$$(3.2) QBS = f(D(QAI), D(IMPB), D(EXPB), QBS(-1))$$

$$(3.3) IMPB = F(EXPB + EXPS - IMPS, \text{otras variables})$$

$$(3.4) EXPB = F(D(QI), EXPB(-1), \text{otras variables})$$

### 10.3. Otros modelos especiales.

*Modelos con variable dependiente cualitativa:* La variable cualitativa puede ser binaria (toma los valores 1 y 0, según se de o no una determinada característica) o un conjunto de variables binarias. En el caso de una variable binaria, si disponemos de una muestra grande podemos agrupar a los individuos, por ejemplo por provincias, y relacionar la *frecuencia relativa* con las variables explicativas. En estos modelos hay que tener en cuenta que se produce un problema de heterocedasticidad, como se indica en la página 117 del libro de Econometría. Si no disponemos de una muestra amplia que nos permita calcular la frecuencia relativa podemos relacionar la variable binaria y (con valores 1 o 0 según el individuo adopte o no la decisión) con una *variable latente*  $y^*$  que es la frecuencia esperada en función de unas variables explicativas. Esa función puede adoptar la forma probit, logit u otras. Por lo que respecta a la bondad del ajuste es interesante comparar el número de casos de estimación correcta con el número de casos totales, de forma que la variable binaria estimada,  $y_f$ , toma el valor 0 cuando la variable latente estimada,  $y_f^* < 0.5$  y el valor 1 si  $y_f^* > 0.5$ .

*Modelos anidados:* La selección de modelos en los que uno es un caso particular de otro, bajo determinadas restricciones, se pueden realizar mediante los contrastes clásicos W, LR y LM, como se analiza en la sección 8.3.1 del libro de Econometría. En muestras pequeñas  $W > LR > LM$ . Los tests son asintóticamente equivalentes por lo que en muestras grandes proporcionan el mismo resultado.

### Ejercicio para interpretar en el examen

En el examen se incluirá un ejercicio práctico para ser explicado por los alumnos que no hayan podido asistir al número suficiente de prácticas del 2º parcial para obtener una evaluación positiva.

La interpretación de los resultados incluye indicar si los parámetros son “significativos” o no (importante conocer el concepto de “significatividad” y no confundirlo con “precisión”, ya que son dos conceptos diferentes: un estimador puede ser preciso o poco preciso (intervalo de confianza amplio, desviación típica relativamente grande en relación con el valor de  $b_i$ ), y en ambos casos su parámetro ( $\beta_i$ ) puede ser significativamente distinto de cero o no (depende de que el estadístico  $t_i = b_i / (\text{desviación típica de } b_i)$  sea mayor que el nivel crítico de la t de student ( $t$  de  $\alpha/2$ ) o no. Las situaciones posibles respecto al carácter significativo de un parámetro son: a) rechazo de la nulidad (el estadístico t es mayor que el nivel crítico, b) aceptación de la nulidad (el estadístico t es pequeño y además el intervalo de confianza tiene sus extremos próximos a cero, por ej. Entre  $-0.02$  y  $0.015$ ) c) incertidumbre (el estadístico t es pequeño pero el intervalo de confianza es muy amplio porque el estimador es poco preciso, por ejemplo sus extremos están entre  $-1.2$  y  $0.4$ ). La incertidumbre se debe generalmente a un problema de multicolinealidad entre los regresores, que al estar muy correlacionados no nos permiten distinguir con claridad el impacto de cada uno de ellos sobre la variable explicada. En una situación de incertidumbre las variables que consideramos importantes desde un punto de vista económico generalmente no se excluyen del modelo aunque sus parámetros no sean significativos. Además de comentar otros resultados (bondad del ajuste, etc.) se debe relacionar el resultado del ejercicio con las conclusiones de los modelos analizados en clase y en los documentos de lectura recomendada.

#### Ejercicio 1. Comparación de modelos dinámicos del empleo no agrario en España.

Presentamos ahora un ejercicio de modelos dinámicos del empleo no agrario de España, LNAE. Se pide comparar los resultados de varios modelos dinámicos, teniendo en cuenta además que en otros estudios las mejores predicciones se obtuvieron con el modelo dinámico mixto y las peores con un modelo puramente autorregresivo. El gráfico 1 muestra la evolución del empleo y del valor añadido no agrario de España expresados, respectivamente, en miles de personas y en millones de Euros, a precios de 1995.

Gráfico 1. LNAE y QNAE

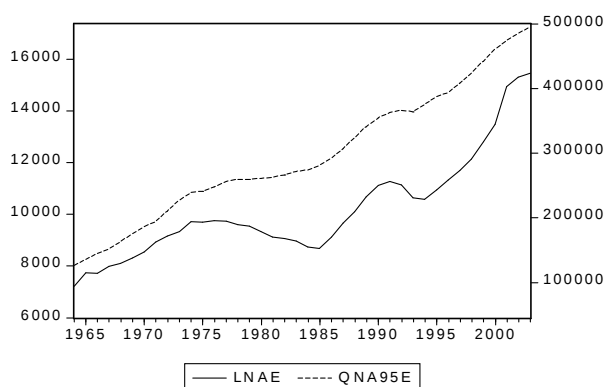


Gráfico 2. Empleo no agrario por cada 100 habitantes

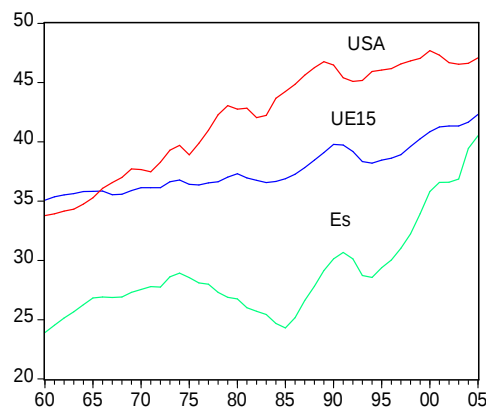
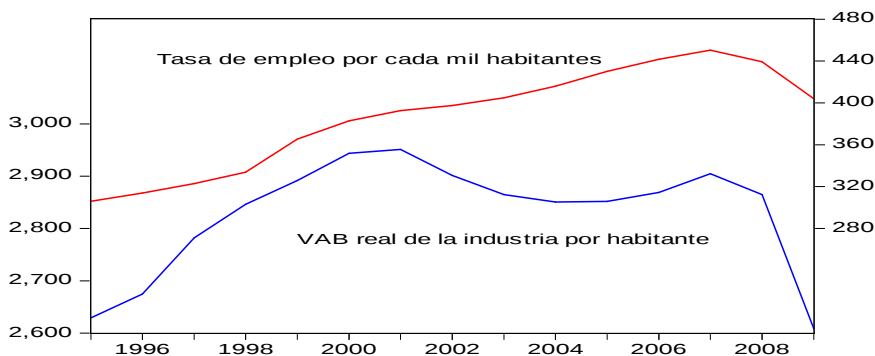


Gráfico 3. VAB real de la industria por habitante (miles € de 2000) y tasa de empleo



El gráfico 2 muestra una comparación de la tasa de empleo no agrario de España con la UE15 y con USA. El gráfico 3 presenta la relación entre VAB real de la industria y tasa de empleo no agrario en España, con estimaciones provisionales del año 2009.

Observamos que el esfuerzo por crear empleo en el período 1996-2008 ha sido destacado, pero lamentablemente no ha ido acompañado de forma suficiente por otras medidas importantes para garantizar que el empleo sea duradero ni para garantizar un incremento suficiente del salario real y de la renta real per cápita, como veremos al final de este documento.

#### 1.- Modelo en niveles

| Dependent Variable: LNAE                            |             |                       |             |          |
|-----------------------------------------------------|-------------|-----------------------|-------------|----------|
| Method: Least Squares. Sample(adjusted): 1965 2002  |             |                       |             |          |
| Included observations: 38 after adjusting endpoints |             |                       |             |          |
| Variable                                            | Coefficient | Std. Error            | t-Statistic | Prob.    |
| C                                                   | -208.5589   | 512.5009              | -0.406943   | 0.6865   |
| QNA95E                                              | 1.402424    | 1.597206              | 0.878049    | 0.3859   |
| LNAE(-1)                                            | 1.000560    | 0.093266              | 10.72802    | 0.0000   |
| R-squared                                           | 0.968121    | Mean dependent var    |             | 10120.99 |
| Adjusted R-squared                                  | 0.966299    | S.D. dependent var    |             | 1804.858 |
| S.E. of regression                                  | 331.3309    | Akaike info criterion |             | 14.51977 |
| Sum squared resid                                   | 3842307.    | Schwarz criterion     |             | 14.64905 |
| Log likelihood                                      | -272.8756   | F-statistic           |             | 531.4511 |
| Durbin-Watson stat                                  | 0.903105    | Prob(F-statistic)     |             | 0.000000 |

#### 2.- Modelo en incrementos simple

| Dependent Variable: D(LNAE)                         |             |                       |             |          |
|-----------------------------------------------------|-------------|-----------------------|-------------|----------|
| Method: Least Squares. Sample(adjusted): 1965 2002  |             |                       |             |          |
| Included observations: 38 after adjusting endpoints |             |                       |             |          |
| Variable                                            | Coefficient | Std. Error            | t-Statistic | Prob.    |
| D(QNA95E)                                           | 28.55807    | 3.877101              | 7.365831    | 0.0000   |
| R-squared                                           | 0.440342    | Mean dependent var    |             | 212.8085 |
| Adjusted R-squared                                  | 0.440342    | S.D. dependent var    |             | 349.7075 |
| S.E. of regression                                  | 261.6172    | Akaike info criterion |             | 13.99761 |
| Sum squared resid                                   | 2532412.    | Schwarz criterion     |             | 14.04070 |
| Log likelihood                                      | -264.9545   | Durbin-Watson stat    |             | 1.265215 |

#### 3.- Modelos en incrementos tipo CE

Relación a largo plazo

| Dependent Variable: LNAE                     |             |                       |             |          |
|----------------------------------------------|-------------|-----------------------|-------------|----------|
| Method: Least Squares                        |             |                       |             |          |
| Sample: 1964 2002. Included observations: 39 |             |                       |             |          |
| Variable                                     | Coefficient | Std. Error            | t-Statistic | Prob.    |
| C                                            | 4962.912    | 337.4963              | 14.70508    | 0.0000   |
| QNA95E                                       | 17.40076    | 1.095818              | 15.87924    | 0.0000   |
| R-squared                                    | 0.872039    | Mean dependent var    |             | 10046.34 |
| Adjusted R-squared                           | 0.868580    | S.D. dependent var    |             | 1840.957 |
| S.E. of regression                           | 667.3805    | Akaike info criterion |             | 15.89452 |
| Sum squared resid                            | 16479678    | Schwarz criterion     |             | 15.97983 |
| Log likelihood                               | -307.9431   | F-statistic           |             | 252.1503 |
| Durbin-Watson stat                           | 0.188063    | Prob(F-statistic)     |             | 0.000000 |

Una vez estimada la relación a largo plazo por MCO (aunque exista autocorrelación, como puede observarse en el valor del estadístico  $dw=0.18$ ), obtenemos los residuos de esta regresión con: FIT YF y generando los residuos: GENR UF=RESID. A continuación se estima la relación a corto plazo, para lo que tenemos dos opciones: sin y con relación contemporánea entre  $\Delta Y_t$  e  $\Delta X_t$ .

### 3.1. Relación a corto plazo del modelo CE, no contemporánea

| Dependent Variable: D(LNAE). Method: Least Squares.    |             |                       |             |          |
|--------------------------------------------------------|-------------|-----------------------|-------------|----------|
| Sample(adjusted): 1966 2002. Included observations: 37 |             |                       |             |          |
| Variable                                               | Coefficient | Std. Error            | t-Statistic | Prob.    |
| D(LNAE(-1))                                            | 0.449769    | 0.180577              | 2.490727    | 0.0178   |
| D(QNA95E(-1))                                          | 0.015540    | 0.006262              | 2.481422    | 0.0182   |
| UF(-1)                                                 | -0.133520   | 0.080040              | -1.668162   | 0.1045   |
| R-squared                                              | 0.529002    | Mean dependent var    |             | 204.3739 |
| Adjusted R-squared                                     | 0.501296    | S.D. dependent var    |             | 350.5908 |
| S.E. of regression                                     | 247.5836    | Akaike info criterion |             | 13.93898 |
| Sum squared resid                                      | 2084119.    | Schwarz criterion     |             | 14.06959 |
| Log likelihood                                         | -254.8711   | Durbin-Watson stat    |             | 1.775962 |

### 3.2. Relación a corto plazo del modelo CE, contemporánea

| Dependent Variable: D(LNAE). Method: Least Squares.    |             |                       |             |          |
|--------------------------------------------------------|-------------|-----------------------|-------------|----------|
| Sample(adjusted): 1966 2002. Included observations: 37 |             |                       |             |          |
| Variable                                               | Coefficient | Std. Error            | t-Statistic | Prob.    |
| D(LNAE(-1))                                            | 0.460285    | 0.156038              | 2.949835    | 0.0057   |
| D(QNA95E)                                              | 16.24813    | 5.416358              | 2.999826    | 0.0050   |
| UF(-1)                                                 | -0.096728   | 0.080763              | -1.197679   | 0.2393   |
| R-squared                                              | 0.560127    | Mean dependent var    |             | 204.3739 |
| Adjusted R-squared                                     | 0.534252    | S.D. dependent var    |             | 350.5908 |
| S.E. of regression                                     | 239.2632    | Akaike info criterion |             | 13.87061 |
| Sum squared resid                                      | 1946394.    | Schwarz criterion     |             | 14.00123 |
| Log likelihood                                         | -253.6063   | Durbin-Watson stat    |             | 1.862978 |

### 4. Modelo Dinámico Mixto

| Dependent Variable: LNAE                           |             |                       |             |          |
|----------------------------------------------------|-------------|-----------------------|-------------|----------|
| Method: Least Squares. Sample(adjusted): 1965 2002 |             |                       |             |          |
| Variable                                           | Coefficient | Std. Error            | t-Statistic | Prob.    |
| LNAE(-1)                                           | 0.983452    | 0.008174              | 120.3167    | 0.0000   |
| D(QNA95E)                                          | 41.73182    | 7.497667              | 5.565974    | 0.0000   |
| R-squared                                          | 0.981137    | Mean dependent var    |             | 10120.99 |
| Adjusted R-squared                                 | 0.980613    | S.D. dependent var    |             | 1804.858 |
| S.E. of regression                                 | 251.3063    | Akaike info criterion |             | 13.94242 |
| Sum squared resid                                  | 2273574.    | Schwarz criterion     |             | 14.02861 |
| Log likelihood                                     | -262.9059   | Durbin-Watson stat    |             | 1.535879 |

### 5. Modelo ARIMA

#### 5.1. Fase de identificación. IDENT LNAE: Sample: 1964 2002. Included observations: 39

| Autocorrelation | Partial Correlation |    | AC     | PAC    | Q-Stat | Prob  |
|-----------------|---------------------|----|--------|--------|--------|-------|
| .  *****        | .  *****            | 1  | 0.838  | 0.838  | 29.515 | 0.000 |
| .  *****        | . * .               | 2  | 0.668  | -0.111 | 48.818 | 0.000 |
| .  ****         | .  .                | 3  | 0.535  | 0.019  | 61.513 | 0.000 |
| .  ***          | .  .                | 4  | 0.420  | -0.029 | 69.567 | 0.000 |
| .  **           | .  .                | 5  | 0.327  | -0.003 | 74.582 | 0.000 |
| .  **           | .  .                | 6  | 0.252  | -0.007 | 77.649 | 0.000 |
| .  *            | .  .                | 7  | 0.196  | 0.011  | 79.569 | 0.000 |
| .  *            | .  .                | 8  | 0.167  | 0.046  | 81.007 | 0.000 |
| .  *            | .  .                | 9  | 0.160  | 0.050  | 82.374 | 0.000 |
| .  *            | .  .                | 10 | 0.150  | -0.011 | 83.619 | 0.000 |
| .  *            | .  .                | 11 | 0.121  | -0.056 | 84.460 | 0.000 |
| .  *            | . * .               | 12 | 0.075  | -0.066 | 84.795 | 0.000 |
| .  .            | .  .                | 13 | 0.023  | -0.048 | 84.829 | 0.000 |
| .  .            | .  .                | 14 | -0.022 | -0.024 | 84.861 | 0.000 |
| .  .            | .  .                | 15 | -0.057 | -0.008 | 85.077 | 0.000 |
| . * .           | .  .                | 16 | -0.082 | -0.008 | 85.544 | 0.000 |

#### IDENT D(LNAE)

| Sample: 1964 2002. Included observations: 38 |                     |    |        |        |        |       |  |
|----------------------------------------------|---------------------|----|--------|--------|--------|-------|--|
| Autocorrelation                              | Partial Correlation |    | AC     | PAC    | Q-Stat | Prob  |  |
| .  *****                                     | .  *****            | 1  | 0.599  | 0.599  | 14.762 | 0.000 |  |
| .  ***                                       | .  .                | 2  | 0.369  | 0.015  | 20.516 | 0.000 |  |
| .  *                                         | .  *                | 3  | 0.156  | -0.111 | 21.574 | 0.000 |  |
| .  .                                         | .  *                | 4  | 0.006  | -0.079 | 21.576 | 0.000 |  |
| .  *                                         | .  *                | 5  | -0.091 | -0.059 | 21.960 | 0.001 |  |
| .  *                                         | .  *                | 6  | -0.162 | -0.078 | 23.214 | 0.001 |  |
| .  **                                        | .  *                | 7  | -0.230 | -0.114 | 25.800 | 0.001 |  |
| .  **                                        | .  .                | 8  | -0.223 | -0.021 | 28.324 | 0.000 |  |
| .  .                                         | .  **               | 9  | -0.022 | 0.258  | 28.349 | 0.001 |  |
| .  .                                         | .  .                | 10 | 0.049  | -0.019 | 28.481 | 0.002 |  |
| .  *                                         | .  .                | 11 | 0.125  | 0.024  | 29.360 | 0.002 |  |
| .  **                                        | .  *                | 12 | 0.204  | 0.113  | 31.788 | 0.001 |  |
| .  *                                         | .  *                | 13 | 0.146  | -0.096 | 33.092 | 0.002 |  |
| .  *                                         | .  .                | 14 | 0.106  | -0.029 | 33.801 | 0.002 |  |
| .  .                                         | .  .                | 15 | 0.042  | -0.034 | 33.918 | 0.003 |  |
| .  *                                         | .  *                | 16 | -0.115 | -0.161 | 34.826 | 0.004 |  |

#### 5.2. Fase de estimación

##### Modelo ARIMA(1,0,0)

| Dependent Variable: LNAE                            |             |                       |             |          |
|-----------------------------------------------------|-------------|-----------------------|-------------|----------|
| Method: Least Squares. Sample(adjusted): 1965 2002  |             |                       |             |          |
| Included observations: 38 after adjusting endpoints |             |                       |             |          |
| Variable                                            | Coefficient | Std. Error            | t-Statistic | Prob.    |
| LNAE(-1)                                            | 1.077137    | 0.032945              | 32.69452    | 0.0000   |
| C                                                   | -551.4811   | 330.7976              | -1.667126   | 0.1042   |
| R-squared                                           | 0.967419    | Mean dependent var    |             | 10120.99 |
| Adjusted R-squared                                  | 0.966514    | S.D. dependent var    |             | 1804.858 |
| S.E. of regression                                  | 330.2753    | Akaike info criterion |             | 14.48893 |
| Sum squared resid                                   | 3926944.    | Schwarz criterion     |             | 14.57511 |
| Log likelihood                                      | -273.2896   | F-statistic           |             | 1068.932 |
| Durbin-Watson stat                                  | 0.972782    | Prob(F-statistic)     |             | 0.000000 |

Fit Yf. Genr Uf5=resid

##### Modelo ARIMA(1,1,0)

| Dependent Variable: D(LNAE)                         |             |                       |             |          |
|-----------------------------------------------------|-------------|-----------------------|-------------|----------|
| Method: Least Squares. Sample(adjusted): 1966 2002  |             |                       |             |          |
| Included observations: 37 after adjusting endpoints |             |                       |             |          |
| Variable                                            | Coefficient | Std. Error            | t-Statistic | Prob.    |
| D(LNAE(-1))                                         | 0.701956    | 0.116016              | 6.050510    | 0.0000   |
| R-squared                                           | 0.331025    | Mean dependent var    |             | 204.3739 |
| Adjusted R-squared                                  | 0.331025    | S.D. dependent var    |             | 350.5908 |
| S.E. of regression                                  | 286.7512    | Akaike info criterion |             | 14.18176 |
| Sum squared resid                                   | 2960146.    | Schwarz criterion     |             | 14.22530 |
| Log likelihood                                      | -261.3626   | Durbin-Watson stat    |             | 1.845806 |

Fit YF. Genr Uf6=resid

Generamos los residuos de estos modelos para poder utilizarlos en la fase de chequeo para contrastar mediante el estadístico Q de Box-Pierce si se cumple la condición de que la perturbación  $u_t$  es un ruido blanco en la fórmula de los modelos ARIMA:  $\Phi(B) y_t = \Theta(B) u_t + \delta$ . En la estimación de estos modelos el parámetro  $\delta$  puede suprimirse cuando no es significativamente distinto de cero.

### 5.3. Fase de chequeo.

Mediante la instrucción IDENT para los residuos UF5 y UF6 obtenemos los siguientes resultados, de los cuales nos interesa especialmente contrastar la hipótesis de ruido blanco con el estadístico Q de Box-Pierce para un número suficientemente grande de retardos. Por ello presentamos sólo los primeros y últimos retardos de los residuos.

IDENT de Residuos Uf5 del Modelo ARIMA en niveles

| Sample: 1964 2002. Included observations: 38 |                     |     |        |        |        |       |  |
|----------------------------------------------|---------------------|-----|--------|--------|--------|-------|--|
| Autocorrelation                              | Partial Correlation | Lag | AC     | PAC    | Q-Stat | Prob  |  |
| .  ****                                      | .  ****             | 1   | 0.472  | 0.472  | 9.1681 | 0.002 |  |
| .  **                                        | .  *                | 2   | 0.307  | 0.108  | 13.142 | 0.001 |  |
| .  *                                         | .  *                | 3   | 0.074  | -0.138 | 13.382 | 0.004 |  |
| .  *                                         | .  *                | 4   | -0.070 | -0.117 | 13.602 | 0.009 |  |
| .  *                                         | .  *                | 20  | -0.169 | -0.081 | 34.139 | 0.025 |  |
| .  *                                         | .                   | 21  | -0.117 | -0.014 | 35.356 | 0.026 |  |
| .  *                                         | .                   | 22  | -0.092 | -0.004 | 36.152 | 0.029 |  |
| .  *                                         | .  *                | 23  | -0.123 | -0.182 | 37.680 | 0.028 |  |
| .                                            | .                   | 24  | -0.049 | -0.037 | 37.940 | 0.035 |  |

IDENT de Residuos Uf6 del Modelo ARIMA en incrementos.

| Sample: 1964 2002. Included observations: 37 |                     |    |        |        |        |       |
|----------------------------------------------|---------------------|----|--------|--------|--------|-------|
| Autocorrelation                              | Partial Correlation |    | AC     | PAC    | Q-Stat | Prob  |
| .  *                                         | .  *                | 1  | -0.083 | -0.083 | 0.2761 | 0.599 |
| .                                            | .                   | 2  | 0.046  | 0.039  | 0.3627 | 0.834 |
| .                                            | .                   | 3  | -0.029 | -0.023 | 0.3990 | 0.940 |
| .                                            | .  *                | 4  | -0.054 | -0.061 | 0.5286 | 0.971 |
| .  *                                         | .  *                | 20 | -0.132 | -0.069 | 11.378 | 0.936 |
| .  *                                         | .  *                | 21 | -0.095 | -0.136 | 12.189 | 0.934 |
| .                                            | .                   | 22 | -0.007 | -0.003 | 12.194 | 0.953 |
| .  *                                         | .  *                | 23 | -0.109 | -0.099 | 13.408 | 0.943 |
| .                                            | .  **               | 24 | 0.000  | -0.189 | 13.408 | 0.959 |

Para la aceptación de la hipótesis de ruido blanco, el valor del estadístico Q de Box-Pierce, de acuerdo con la exposición del Capítulo 9 del libro de Econometría debe ser menor que el nivel crítico,  $\chi^2_{\alpha}$ . La Q del modelo en niveles está en la cola derecha (rechazo) mientras que la Q del modelo en incrementos está en la cola izquierda (aceptación de la hipótesis), y por lo tanto es preferible en este caso el modelo en incrementos.

### Conclusiones del ejercicio:

a) Bondad del Ajuste: Se puede observar que la SCE es menor en el caso de los modelos dinámico mixto, CE y modelo en incrementos que en los otros casos. Podríamos calcular el valor de  $R^2$ , tanto ajustado como no ajustado, referido a la variable LNAE y no a su incremento y comprobar que, como es habitual, dichos modelos son los que mejor se ajustan. Calculando el % S.E. (porcentaje del Error Standard respecto a la media de la variable LNAE) podemos observar también los modelos que tienen los valores más bajos que son los de mejor ajuste. También los criterios Akaike y Schwartz tienen valores bajos en los modelos con buen ajuste.

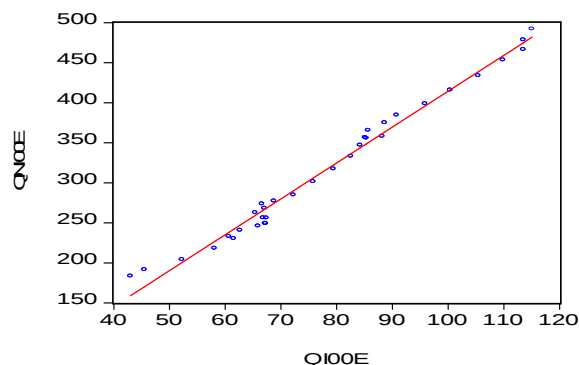
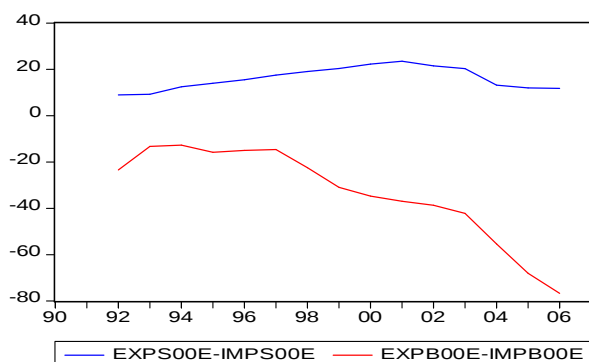
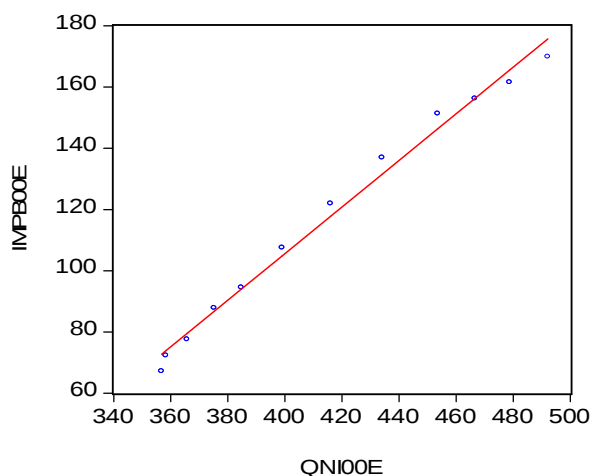
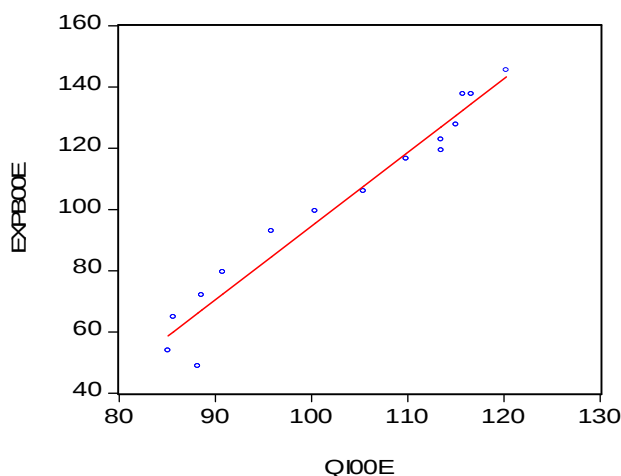
b) Capacidad predictiva. Por lo que respecta a la capacidad predictiva el modelo dinámico mixto y el modelo CE, con relación a corto plazo contemporánea, son los que generalmente se comportan mejor.

c) Factores que explican la evolución del empleo. El valor real de la producción no agraria es la variable que tiene mayor impacto sobre el empleo en España y en la mayoría de los países. Se pueden hacer más comentarios sobre la relación entre producción y empleo, y los factores que más positivamente influyen en ambos, como los que figuran en los documentos recomendados. Respecto a la comparación de distintos enfoques teóricos para la explicación del empleo, el Documento 77 presenta resultados muy interesantes, los cuales se analizan en cursos de Econometría Aplicada avanzada. La explicación del nivel de QNAE es por lo tanto la cuestión principal en este tema.

*Ejercicio 2. Modelo intersectorial explicativo de la producción no agraria de España.*

Teniendo en cuenta las explicaciones que se indican en el Anexo 2 del Documento nº 97 y de los modelos que allí se citan, estimamos un modelo para el Valor Añadido real de los sectores B (Construcción) y S (Servicios) mediante un modelo dinámico mixto que relaciona a la variable QBS con su valor anterior, y con los incrementos de las variables QI, IMPB y EXPB, correspondientes al Valor Añadido real de la Industria, e Importaciones y Exportaciones de bienes. Los datos están en millones de dólares a precios y tipos de cambio del año 2000.

| Dependent Variable: QBSE                           |             |                       |             |          |
|----------------------------------------------------|-------------|-----------------------|-------------|----------|
| Method: Least Squares. Sample(adjusted): 1995 2003 |             |                       |             |          |
| Variable                                           | Coefficient | Std. Error            | t-Statistic | Prob.    |
| QBSE(-1)                                           | 1.019978    | 0.004425              | 230.4826    | 0.0000   |
| D(QIE)                                             | 0.567179    | 0.680758              | 0.833157    | 0.4427   |
| D(IMPBE)                                           | 0.777185    | 0.310552              | 2.502589    | 0.0543   |
| D(EXPBE)                                           | -0.606374   | 0.283255              | -2.140736   | 0.0852   |
| R-squared                                          | 0.998394    | Mean dependent var    |             | 412.1419 |
| Adjusted R-squared                                 | 0.997430    | S.D. dependent var    |             | 40.48321 |
| S.E. of regression                                 | 2.052142    | Akaike info criterion |             | 4.576747 |
| Sum squared resid                                  | 21.05643    | Schwarz criterion     |             | 4.664403 |
| Log likelihood                                     | -16.59536   | Durbin-Watson stat    |             | 1.482393 |

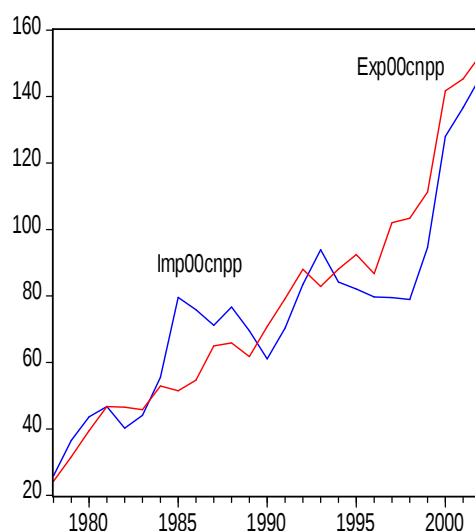


Documento nº 100, a incluir en <http://www.usc.es/economet/>

Ejercicio 3. Examen de Mayo de 2009

1. Se pide explicar brevemente como realizaría con el programa econométrico de prácticas las siguientes aplicaciones para las variables IMP=imp00cnpp (importaciones de China en dólares del año 2000 según Paridades de Poder de Compra) y EXP=exp00cnpp (exportaciones de China en \$2000 según Paridades):

a) Contrastar la cointegración y la causalidad entre ambas variables. b) Estimar un modelo ARIMA para las Exportaciones y comparar dicha estimación con la que daría la estimación de un modelo dinámico mixto que relacione las exportaciones con su valor retardado y con el incremento de la producción industrial de China. c) Comentar los resultados, tanto en sentido econométrico como en sentido económico, de la ecuación estimada, teniendo en cuenta que el coeficiente de la variable D(EXP) es algo bajo en comparación con otros países en los que suele estar más próximo a la unidad.



|                                                     |             |                       |             |          |
|-----------------------------------------------------|-------------|-----------------------|-------------|----------|
| Dependent Variable: IMP00CN                         |             |                       |             |          |
| Method: Least Squares                               |             |                       |             |          |
| Sample(adjusted): 1979 2002                         |             |                       |             |          |
| Included observations: 24 after adjusting endpoints |             |                       |             |          |
| Variable                                            | Coefficient | Std. Error            | t-Statistic | Prob.    |
| D(EXP00CN)                                          | 0.729966    | 0.199980              | 3.650199    | 0.0014   |
| IMP00CN(-1)                                         | 1.011276    | 0.036796              | 27.48350    | 0.0000   |
| R-squared                                           | 0.927752    | Mean dependent var    |             | 146.9876 |
| Adjusted R-squared                                  | 0.924468    | S.D. dependent var    |             | 67.36471 |
| S.E. of regression                                  | 18.51394    | Akaike info criterion |             | 8.754581 |
| Sum squared resid                                   | 7540.856    | Schwarz criterion     |             | 8.852752 |
| Log likelihood                                      | -103.0550   | Durbin-Watson stat    |             | 1.981116 |