



W
h9
(9721)



Documento de trabajo

Previsión y seguimiento de ocupados por sectores, población activa y parados en España: una comparación de modelos alternativos

Silvia Relloso Pereda

No. 9721

Diciembre 1997

ICAE

Instituto Complutense de Análisis Económico

UNIVERSIDAD COMPLUTENSE

FACULTAD DE ECONOMICAS

Campus de Somosaguas

28223 MADRID

Teléfono 394 26 11 - FAX 294 26 13

ICAE

Instituto Complutense de Análisis Económico

UNIVERSIDAD COMPLUTENSE

PREVISIÓN Y SEGUIMIENTO DE OCUPADOS POR SECTORES,

POBLACIÓN ACTIVA Y PARADOS EN ESPAÑA.

UNA COMPARACIÓN DE MODELOS ALTERNATIVOS*

Silvia Relloso Pereda
Departamento de Economía Cuantitativa
Universidad Complutense de Madrid
Campus de Somosaguas, 28223, Madrid
Tfno: 394 23 84, e-mail: eccua27@sis.ucm.es



ABSTRACT

It is shown that both the disaggregation of employment into productive sectors and the study of relationships among the resulting components are relevant tools for obtaining important improvements in forecasting employment and unemployment in Spain. For this purpose, the predictive capacity of a variety of univariate and transfer function models is investigated and compared with the predictive capacity of a multivariate model in which disaggregation and relationships are incorporated.

RESUMEN

Se demuestra que la desagregación del número de ocupados en sectores productivos y el análisis de las relaciones entre los componentes resultantes es relevante para obtener mejoras importantes en la previsión de los parados y los ocupados en España. Para ello, se investiga la capacidad predictiva de un conjunto de modelos univariantes y de transferencia, comparándola con la capacidad predictiva de un modelo multivariante que incorpora la desagregación del número de ocupados y sus relaciones.

N.C.: X-53-797058-5

N.E.: 531077576

* Este trabajo se deriva de la tesis doctoral realizada por la autora bajo la dirección del profesor Arthur B. Treadway. La autora desea agradecer el apoyo recibido de los miembros del *Servicio de Previsión y Seguimiento de la Economía Española* y la ayuda financiera facilitada por Caja de Madrid. Todos los posibles errores del trabajo son responsabilidad única de la autora.

1. INTRODUCCIÓN

El elevado nivel de desempleo es el problema más destacado al que se enfrenta hoy día la economía española. Esto motiva un intento de elaborar modelos útiles que proporcionen mejoras en las previsiones del número de Parados. Este trabajo muestra la importancia de estudiar la evolución del número de Ocupados (E_t) desagregando el total en sectores productivos y analizando las relaciones entre los componentes, para obtener mejoras relevantes en las operaciones de previsión y seguimiento de E_t , de la Población Activa (A_t) y de los Parados ($D_t \equiv A_t - E_t$). Se compara la capacidad predictiva de un modelo multivariante que incorpora la desagregación de E_t y las relaciones de sus componentes, tanto internas como con A_t (y D_t), con la de otros modelos que no incorporan estas dos características.

La investigación parte del trabajo realizado para el boletín mensual *Previsión y Seguimiento de la Economía Española* (SPS:EE), publicado por la Caja de Madrid hasta diciembre de 1996, donde se ha ofrecido información sobre previsiones de E_t , A_t , D_t y la Población Total restada de la población contada aparte (P_t), usando los datos estadísticos (trimestrales) proporcionados por la Encuesta de Población Activa (EPA) elaborada por el Instituto Nacional de Estadística (INE).

Los sectores de E_t considerados son: Agricultura, Ganadería y Pesca (EA_t), Industria (EI_t), Construcción (EC_t), Servicios Privados (EP_t) y Servicios Públicos (EU_t), que comprende la Administración Pública, la Educación y la Sanidad.

Este trabajo consta de 4 secciones. En la Sección 2 se presentan los modelos alternativos para generar las previsiones de cada variable. La Sección 3 recoge una comparación de los resultados en previsión generados por cada tipo de modelo. La Sección 4 resume las principales conclusiones.

2. MODELOS USADOS PARA LA PREVISIÓN

En la elaboración de los modelos se emplea la metodología de Análisis de Series Estadísticas Temporales del enfoque Box y Jenkins (Box *et al.* (1994)) extendida con el análisis de la

estacionalidad utilizando el Modelo Estacional Generalizado (MEG), propuesto en Gallego (1995) y en Gallego y Treadway (1996), y se usa el criterio de Máxima Verosimilitud Exacta en la estimación según Mauricio (1995, 1996, 1997). El Apéndice A recoge unas nociones del análisis de la estacionalidad usando el MEG. La muestra empleada en la elaboración de los modelos, abarca desde el tercer trimestre de 1976 hasta el tercer trimestre de 1993. Se emplea la muestra comprendida entre el cuarto trimestre de 1993 y el segundo de 1996 para comparar los modelos en operaciones de previsión y seguimiento.

2.1 Modelos usados en el boletín *Previsión y Seguimiento de la Economía Española*

Los modelos elaborados para las operaciones de previsión y seguimiento publicadas en el SPS:EE, tienen en cuenta las siguientes identidades que relacionan las variables:

$$A_t \equiv E_t + D_t \quad (1)$$

$$P_t \equiv A_t + I_t \quad (2)$$

donde I_t denota la Población Inactiva. Se modelizan D_t e I_t mediante la aplicación de estas dos identidades. Se emplean dos supuestos no contrastables de identificación para abordar el estudio de las relaciones entre estas cinco variables: (1) ausencia de efectos contemporáneos de E_t y A_t sobre P_t y (2) ausencia de efectos contemporáneos de A_t sobre E_t . Estos dos supuestos facilitan la identificación exacta del sistema en (A_t, E_t, P_t) . En relación a la primera hipótesis no contrastable de identificación, P_t podría teóricamente reaccionar frente a A_t y E_t (como factores que reflejan las condiciones económicas) por dos tipos de razones: (1) natalidad y defunciones y (2) migración. Si la natalidad reacciona a las condiciones económicas, la realimentación sería muy lenta, ya que deben pasar 16 años para que opere sobre P_t ; por eso se ignora este aspecto de posible realimentación. Tampoco parece plausible que las defunciones reaccionen en el mismo trimestre a las condiciones económicas. La migración supone una reacción más rápida y difícil de ignorar, pero parece plausible que no reaccione

a las condiciones económicas en el mismo trimestre. La hipótesis de que A_t no influye contemporáneamente en el empleo en ningún sector es otra hipótesis no contrastable de identificación. Se considera que la decisión sobre E_t se toma por agentes empleadores que son distintos de los agentes potencialmente empleados. El autoempleo es una excepción evidente que puede ser investigada mediante la desagregación entre asalariados y no asalariados, tema no abordado en este trabajo.

Para la previsión de E_t en el SPS:EE se ha usado un modelo univariante con intervención (Box y Tiao (1975)), recogido en la Ecuación (3) donde, $\xi_t^{S,t*}$ denota una variable determinista de intervención de tipo escalón con $\xi_t^{S,t*} = 0$ ($t < t^*$) ó 1 ($t \geq t^*$), g denota la ganancia a largo plazo de una función de transferencia, \bar{a} denota la media de la serie residual, $\hat{\sigma}_a$ denota la desviación típica estimada de la serie residual y $Q(GL)$ es el estadístico Q de Ljung y Box (1978) donde GL denota los grados de libertad con los que se calcula. Se presentan entre paréntesis las desviaciones típicas asociadas a los parámetros estimados, debajo de los mismos.

$$E_t = 190.6 \xi_t^{S,II/87} - 103.6 \xi_t^{S,IV/89} - (150.4 + 156.1B) \xi_t^{S,IV/92} + NE_t \quad (3)$$

$(30.5) \quad (24.2) \quad (45.9) \quad (42.5)$
 $g = -306.5$
 (73.8)

$$\nabla \nabla_4 NE_t = (1 - .40B) (1 + .86B^2) (1 + 1.00B) \hat{a}_t$$

$(.10) \quad (.06) \quad (.04)$

$$\bar{a} = 0.0 \quad \hat{\sigma}_a = 46.5 \quad Q(12) = 18.7$$

(5.7)

El término de intervención en II/87 se debe a un cambio metodológico en la EPA. El resto de los términos de intervención son debidos a observaciones anómalas que resultan influyentes en la estructura estocástica de la variable y para los que no existe información extramuestral.

La variable P_t se describe mediante el modelo univariante estocástico:

$$\nabla^2 P_t = (1 - .82B) \hat{a}_t \quad (4)$$

$$\bar{a} = -0.2(2.2) \quad \hat{\sigma}_a = 17.9 \quad Q(14) = 8.4$$

Para la previsión y seguimiento de A_t , se usaba un modelo de transferencia de un sólo output (UT) con E_t y P_t como inputs independientes, donde $E_t^c \equiv NE_t$ corresponde a la variable E_t corregida de cuatro términos escalón en II/87, IV/89, IV/92 y I/93, con los coeficientes estimados presentados en (3). Sólo el término escalón en IV/89 influye en la evolución de A_t y la magnitud del efecto es la misma que la de E_t , aunque de signo contrario. Los términos de intervención en IV/88, III/91, I/92,

$$A_t = 207.2 \xi_t^{S,II/87} - 150.9 \xi_t^{S,IV/88} + 61.9 \xi_t^{S,III/91} - 43.8 \xi_t^{S,I/92} + (70.8 + 76.0B) \xi_t^{S,II/93} \quad (21.3) \quad (20.4) \quad (20.1) \quad (20.7) \quad (24.2) \quad (25.8)$$

$$g = \frac{146.8}{(41.3)}$$

$$+ .42 (E_t^* - \xi_t) + .46 P_t + NA_t \quad (15)$$

$$\nabla \nabla_4 NA_t = (1 - .75B) (1 + .57B^2) (1 + 1.00B) \hat{a}_t \quad (14) \quad (13) \quad (05)$$

$$\bar{a} = 2.5 \quad \hat{\sigma}_a = 23.8 \quad Q(12) = 6.5 \quad (2.9)$$

$$\xi_t = \begin{cases} -103.6, & t < IV/89 \\ 0.0, & t \geq IV/89 \end{cases} \quad (5)$$

Los modelos no invertibles (3) y (5) son equivalentes en operaciones de previsión a los modelos invertibles con componentes de estacionalidad determinista (Gallego (1995) y Bell (1987)). Los términos de estacionalidad determinista no aparecen en la representación no invertible, lo que supone una simplificación en las operaciones de previsión y seguimiento.

La previsión y seguimiento de D_t se realiza a través de la identidad (1). En la subsección 2.2 se recogen los modelos elaborados al considerar la desagregación de E_t y el estudio de sus relaciones.

2.2 Modelos que incorporan la desagregación de Ocupados en sectores productivos

Para ampliar el conocimiento sobre la evolución de E_t y mejorar las previsiones generadas por estos modelos se plantea la desagregación de E_t en cinco sectores productivos, de forma que la previsión del agregado se obtiene mediante la agregación de las previsiones de cada uno de sus componentes, a través de la identidad:

$$E_t \equiv EA_t + EP_t + EI_t + EU_t + EC_t \quad (6)$$

Esta desagregación no es sólo interesante en términos de previsión, sino también para conocer y comprender mejor la muy distinta evolución de cada componente (véase Figuras 1 y 2).

Se elaboran, en primer lugar, los MEG para cada uno de los componentes de E_t recogidos en la Tabla 1. Los símbolos $\hat{\theta}$, $\hat{\lambda}_1$, $\hat{\lambda}_2$ representan los parámetros estimados de los operadores media móvil de frecuencias cero, uno y dos respectivamente, $\hat{\sigma}_a$ representa la desviación típica estimada de la serie residual, S indica que el término determinista de intervención es de tipo escalón, I de tipo impulso, IC de tipo impulso compensado y $\hat{\omega}_0$ y $\hat{\omega}_1$ representan los valores estimados de los coeficientes de los términos deterministas de intervención. Todos los términos de intervención se deben a valores anómalos influyentes en la estructura estocástica de cada variable, salvo el correspondiente a II/87 para el que existe información extramuestral. Se presenta entre paréntesis la desviación típica estimada asociada al valor estimado de cada parámetro. Los MEG son el punto de partida para la elaboración de un modelo multivariante estocástico (MS) que recoge las relaciones entre los componentes de E_t , P_t y A_t . Las implicaciones para D_t se derivan de la identidad (1).

Este artículo no recoge de forma detallada el proceso seguido en la elaboración de los modelos univariantes ni del modelo MS, ya que el objetivo central aquí es mostrar las mejoras obtenidas en la

previsión con la desagregación de E_i , el análisis de las relaciones entre sus componentes y la evaluación del distinto efecto de cada sector de E_i sobre A_i (y por tanto sobre D_i). La descripción exhaustiva de la elaboración de los MEG y del MS se encuentra en Reloso (1997) y (1998).

El modelo MS utilizado para generar las previsiones se presenta en la Ecuación (7). La Tabla 2 recoge los términos deterministas de intervención asociados a cada variable. En la Ecuación (7), se denota por: (1) NEI_i , la desviación de la variable original Ei_i de sus términos de intervención deterministas, para $i = A, P, I, U, C$, (2) NA_i la desviación típica de la variable original A_i de sus términos de intervención deterministas, (3) $\hat{a}_{A_t}, \hat{a}_{P_t}, \hat{a}_{I_t}, \hat{a}_{U_t}, \hat{a}_{C_t}$ y \hat{a}_{ct} las series residuales, y (4) \hat{P} la matriz de correlaciones contemporánea estimada, proporcionando debajo de \hat{P} el error estándar estimado.

$Z(B)$ es una matriz tal que:

$$[Z_{ij}(B)] = \begin{cases} u(B) & i = j \neq 2 \\ 1 & i = j = 2 \\ 0 & i \neq j \end{cases} \quad \text{con:} \\ u(B) = (1+B)(1+B^2)$$

El vector $\hat{v}(B)^T$ es:

$$\hat{v}(B)^T = [\hat{\omega}_1, \hat{\omega}_2(1+B), \hat{\omega}_3, \hat{\omega}_4, \hat{\omega}_5, 0]$$

y el vector ξ se define como:

$$\xi_t = [0, \xi_t^A, \xi_t^P, \xi_t^I, \xi_t^U, \xi_t^C]^T$$

siendo ξ_t^i : (1) la suma de los efectos deterministas asociados a la variable Ei_i , para $i = A, C, P$ y

(2) la suma de los efectos deterministas asociados a la variable Ei_i , menos la suma de los efectos deterministas asociados a dicha variable correspondientes a II/87, IV/92 y I/93, para $i = I, U$.

(7)

$$\begin{pmatrix} \nabla \nabla_4 [NA_i - \psi(B)^T \xi_t] \\ \nabla \nabla_4 P_t \\ \nabla \nabla_4 NEA_t \\ \nabla \nabla_4 NEP_t \\ \nabla \nabla_4 NEI_t \\ \nabla \nabla_4 NEU_t \\ \nabla \nabla_4 NEC_t \end{pmatrix} = \begin{pmatrix} 1 & -44 & -43(1+B) & -45 & -51 & -30 & 0 \\ (.15) & (.02) & (.03) & (.05) & (.05) & (.05) & \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \hat{a}_{At} \\ \hat{a}_{Pt} \\ \hat{a}_{It} \\ \hat{a}_{Ut} \\ \hat{a}_{Ct} \end{pmatrix}$$

$$\nabla \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & (1-.80B) & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & (1-.41B)(1+.65B^2)(1+.70B) & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .67B(1-.45B) & 0 & 0 & 0 \\ 0 & 0 & 0 & (1-.45B) & (22B+.30B^2+.26B^3) & 0 & 0 \\ 0 & 0 & 0 & 0 & (1-.88B) & .30B^2(1-.88B) & 0 \\ 0 & 0 & 0 & 0 & 0 & (1-.49B) & 0 \end{pmatrix} \begin{pmatrix} \hat{a}_{At} \\ \hat{a}_{Pt} \\ \hat{a}_{It} \\ \hat{a}_{Ut} \\ \hat{a}_{Ct} \end{pmatrix}$$

$Z(B)$

$$\begin{array}{lll} \bar{a}_A = -2.0 (3.7) & \hat{\sigma}_A^* = 22.6 & Q(15) = 15.8 \\ \bar{a}_p = 0.6 (2.2) & \hat{\sigma}_p = 19.0 & Q(14) = 9.1 \\ \bar{a}_a = 1.5 (3.4) & \hat{\sigma}_a = 27.4 & Q(12) = 9.6 \\ \bar{a}_p = 0.5 (2.7) & \hat{\sigma}_p = 24.1 & Q(15) = 12.0 \\ \bar{a}_i = -0.8 (2.0) & \hat{\sigma}_i = 17.3 & Q(14) = 12.9 \\ \bar{a}_u = 3.1 (1.8) & \hat{\sigma}_u = 15.6 & Q(14) = 18.5 \\ \bar{a}_c = -1.4 (1.9) & \hat{\sigma}_c = 16.2 & Q(14) = 11.4 \end{array}$$

$$(.13) \quad \hat{P} = \begin{pmatrix} 1 & & & & & & \\ 0 & 1 & & & & & \\ 0 & 0 & 1 & & & & \\ 0 & 0 & 0 & 1 & & & \\ 0 & 0 & -0.22 & 0 & 1 & & \\ 0 & 0 & 0 & 0 & .31 & 1 & \\ 0 & 0 & -0.43 & 0 & 0 & 0 & 1 \end{pmatrix}$$

La elaboración del modelo MS permite investigar las relaciones dinámicas entre los componentes de ocupados, un enfoque nuevo en el estudio de este tipo de datos, hasta donde la autora conoce. Se observa que: (1) EC_t no recibe influencias retardadas del resto de los sectores, (2) se encuentra una relación de cointegración $CI(2,1)$ en el vector (NEP_t, NEI_t) , lo que implica que las variaciones en los dos sectores operan en un equilibrio estacionario, (3) EU_t recibe influencias contemporáneas y retardadas de EI_t (debido a que la correlación contemporánea positiva entre \hat{a}_t y \hat{a}_{ut} se interpreta dando un papel dependiente a EU_t) e influencias retardadas de EC_t , (4) EI_t recibe influencias retardadas de las variaciones de EC_t y de EU_t , (5) EP_t recibe influencias retardadas de las variaciones EC_t y de EU_t , y (6) EA_t no influye en la evolución del resto de los sectores. Los valores elevados de las correlaciones contemporáneas entre los residuos de EA_t con EI_t y EC_t se interpretan suponiendo que EI_t y EC_t influyen en EA_t sin realimentación.

Además, el análisis MS considerando los cinco sectores de ocupados revela que: (1) se logra una disminución de las varianzas residuales al aplicar la metodología de análisis MS en comparación con los valores obtenidos en los análisis univariantes, (2) la varianza de la serie residual de E_t obtenida como la suma de los residuos de los cinco componentes de ocupados que se derivan del modelo MS,

es ligeramente menor que la obtenida a partir del modelo univariante del agregado, (3) P_t no recibe efectos contemporáneos de los componentes de E_t ni de A_t (hipótesis de identificación de partida), ni efectos retardados, (4) la desagregación de E_t permite evaluar el distinto efecto de cada uno de sus componentes sobre A_t y, por tanto, sobre D_t dada la identidad que relaciona estas variables y (5) se consigue reducir la estructura estocástica del ruido a un paseo aleatorio, en relación a lo obtenido con el modelo UT de A_t usado en las operaciones del SPS:EE que utilizaba E_t agregado.

3. RESULTADOS EN PREVISIÓN Y SEGUIMIENTO

En esta sección se presenta una comparación de los resultados en previsión generados por:

(1) los modelos univariantes, (2) los modelos usados en el SPS:EE y (3) el modelo MS.

La comparación de la capacidad predictiva se hace en términos de: (1) la desviación típica estimada (DTE) de los errores de previsión generada por cada modelo en la muestra comprendida entre el III/76 y el III/93 y (2) el comportamiento en pruebas postmuestreales. La varianza estimada de los errores es un elemento importante en la comparación de distintas representaciones de los datos, ya que tal comparación no debe basarse únicamente en la previsión puntual, sino también en una medida del riesgo asociado a dicha previsión puntual. Las previsiones se realizan para once observaciones fuera de la muestra empleada en la elaboración de los modelos, por lo que es posible calcular los errores de previsión cometidos para los distintos horizontes y comprobar si dichos errores pertenecen o no a los intervalos de previsión generados por el modelo. Los errores de previsión a horizonte uno dentro de la muestra parecen seguir procesos de ruido blanco.

Es conocido que no existe un criterio único y objetivo para la comparación de la capacidad predictiva de un conjunto de modelos (Brian y Diamantopoulos (1994) y García-Ferrer *et al.* (1994)). Lo mismo ocurre con los criterios de estimación. (Zellner (1983). En este trabajo la estimación se realiza usando el criterio de Máxima Verosimilitud Exacta, lo que implica asumir los supuestos de

distribución a priori no informativa sobre los valores de los parámetros y función de pérdida cuadrática. Bajo estos supuestos, se considera apropiado usar en las comparaciones de la capacidad predictiva el valor de la DTE generada por cada modelo y los resultados obtenidos en pruebas postmuestreales de los valores absolutos de los errores de previsión y los valores de la Raíz del Error Cuadrático Medio (RECM). Se calculan también los valores de la Media del Error Absoluto, pero no se presentan estos datos por no diferir los resultados de la comparación de lo obtenido con la RECM. Este es un tipo de procedimiento común en los estudios de comparación de previsiones, pero no el único posible o relevante en la práctica (Diebold y Mariano (1995)).

Las opciones para la previsión y el seguimiento de las variables son distintas para cada una. No se consideran modelos alternativos para: (1) EC_t y P_t , variables exógenas en el modelo MS, por lo que los resultados obtenidos con este modelo no deben diferir de los obtenidos con los modelos univariantes y (2) EA_t , variable que sólo tiene relaciones contemporáneas débiles en el modelo MS.

El tratamiento dado a EU_t , EI_t y EP_t en los modelos univariantes no coincide con el del modelo MS, ya que este último incluye relaciones retardadas para estas variables y además una relación de cointegración para EP_t . Por tanto, existen dos opciones para realizar previsión y seguimiento de estas tres variables.

Las Tablas 3 a 5 recogen los valores de la DTE y de la RECM para las tres variables y los ocho horizontes de previsión calculados. El modelo que genere menores valores de estos criterios (que aparecen en negrilla en las tablas) es preferido al resto. Se denota por (1) los resultados obtenidos por el modelo univariante y por (2) los obtenidos con el modelo MS. En las dos últimas columnas de cada tabla se detalla, para cada horizonte de previsión, el número de veces que cada tipo de modelo proporciona un error de previsión menor en valor absoluto. Así, *p.e.*, en la previsión de EP_t a horizonte uno, para seis de los orígenes de previsión el modelo univariante proporciona errores de previsión menores en valor absoluto a los proporcionados por el modelo MS, el cual proporciona

errores de previsión menores en valor absoluto en los cinco orígenes restantes.

Para EP_t (Tabla 3) el Modelo MS proporciona valores menores de la DTE que los generados por el modelo univariante para todos los horizontes y la mejora es mayor para las previsiones a horizontes mayores. Los valores de la RECM asociados a los errores de previsión a horizontes cortos son similares con los dos tipos de modelos. Las diferencias entre los valores de la RECM aumentan al aumentar el horizonte de previsión, siendo mucho menores los valores obtenidos con el modelo MS. Los valores absolutos de los errores de previsión a horizontes mayores son siempre menores con el modelo MS, aunque las diferencias entre los errores es inferior a dos desviaciones típicas en la mayoría de los casos. El modelo MS genera una mejora notable en la previsión de EP_t mayor a horizontes mayores.

No hay apenas diferencias (Tablas 4 y 5) entre los resultados obtenidos por los dos tipos de modelos en la previsión de EI_t y de EU_t , aunque el modelo MS parece proporcionar mejores resultados para EU_t a medida que aumenta el horizonte de previsión.

En la Tabla 6 se presentan los resultados obtenidos en la previsión de E_t con tres tipos de modelos. Se denota por (1) los resultados obtenidos por la agregación de componentes utilizando sus modelos univariantes, por (2) los resultados obtenidos con el modelo univariante del agregado, usado en las operaciones de previsión y seguimiento publicadas en el SPS:EE y por (3) los resultados obtenidos por agregación de componentes usando el modelo MS.

La previsión de E_t obtenida por la agregación de sus componentes usando sus modelos univariantes genera, a partir del horizonte 2, valores de la DTE y de la RECM menores que los generados por el modelo univariante del agregado, y la mejora es mayor al aumentar el horizonte de previsión. La introducción de las relaciones entre los sectores hace posible obtener: (1) valores menores de la DTE para todos los horizontes de previsión, (2) valores menores de la RECM a partir del horizonte tres de previsión y (3) en la mayoría de los casos, errores de previsión menores.

Se produce, por tanto, una notable mejora en la previsión de E_t al utilizar el modelo MS, siendo esta mejora mayor a medida que aumenta el horizonte de previsión.

La Tabla 7 presenta los valores de la DTE y de la RECM obtenidos en la previsión de A_t con tres tipos de modelos: (1) el modelo univariante, (2) el modelo UT en uso hasta ahora en el SPS:EE y (3) el modelo MS. También se presenta el número de veces que cada tipo de modelo proporciona un error de previsión menor en valor absoluto.

El modelo MS genera los menores valores de la DTE a todos los horizontes, no existiendo diferencias significativas entre los valores de la DTE generados por el modelo univariante de A_t y el modelo de transferencia donde E_t se trata de forma agregada. El modelo MS genera, a partir del horizonte tres, los menores valores de la RECM y de los errores de previsión.

La previsión de D_t se realiza utilizando la identidad (1) de la Sección 2, que relaciona esta variable con E_t y A_t . Se comparan los resultados obtenidos para esta variable por esta vía con tres tipos de modelos: (1) usando los modelos univariantes de E_t y A_t , (2) usando el modelo UT de A_t , en el que E_t se trata de forma agregada (modelo usado en el SPS:EE) y (3) usando el modelo MS, en el que E_t se trata de forma desagregada y se tienen en cuenta las relaciones. La Tabla 8 recoge información similar a la presentada para la comparación de los resultados en la previsión de E_t y A_t .

La comparación de los valores de la DTE y de la RECM obtenidos con los tres tipos de modelos, conduce a conclusiones semejantes: (1) la previsión de D_t a partir del modelo MS proporciona los mejores resultados y (2) no hay diferencias significativas entre los valores de la DTE y de la RECM, obtenidos con los modelos de tipo (1) y (2).

Los valores de la DTE generados en la previsión de D_t con el modelo MS son bastante menores a los generados por el tipo de modelo (2) (aproximadamente una reducción del 68% de la varianza estimada asociada al error de previsión a horizonte uno), siendo la reducción mayor para horizontes mayores. Los errores de previsión generados por el modelo MS, para la mayoría de los

orígenes de previsión y para todos los horizontes, son menores que los generados por los otros dos tipos de modelo. Es en esta variable donde la ganancia en la previsión y el seguimiento con el modelo MS es mayor.

4. CONCLUSIONES

Este trabajo muestra la relevancia de desagregar el total de Ocupados en sectores productivos, analizar sus relaciones y estudiar el distinto efecto de cada sector sobre la Población Activa y los Parados, para obtener mejoras importantes en las operaciones de previsión y seguimiento de un grupo de las variables analizadas. Se compara la capacidad predictiva de un modelo multivariante que incorpora la desagregación de Ocupados y las relaciones entre las variables, con otros modelos que no incorporan esta información. Los modelos considerados son: (1) los usados en las operaciones de previsión en el boletín mensual *Previsión y Seguimiento de la Economía Española*, donde se consideran los agregados de Ocupados, Población Activa, Población Total y Parados, (2) los modelos univariantes de cada uno de los componentes de Ocupados y (3) el modelo multivariante que implica la desagregación de los Ocupados en sectores productivos y el análisis de las relaciones tanto entre los componentes como con la Población Activa y los Parados.

La comparación de la capacidad predictiva se realiza en términos de la Desviación Típica Estimada de los errores de previsión generado por cada modelo, de los valores de la Razón del Error Cuadrático Medio y del valor absoluto de los errores de previsión obtenidos con cada tipo de modelo en once observaciones postmuestrales.

Se muestra que la información adicional que incorpora el modelo multivariante es relevante para obtener mejoras considerables en la comprensión, previsión y seguimiento de Ocupados en Servicios Privados, Ocupados (agregado), Población Activa y sobre todo Parados, siendo en esta variable donde la mejora obtenida es mayor.

APÉNDICE A: Representación Univariante con estacionalidad ARIMA generalizada

Se recogen en esta sección unas nociones del análisis de la estacionalidad usando el MEG siguiendo la metodología propuesta por Gallego (1995); véase también Gallego y Treadway (1996). Sólo se expone el caso de $s = 4$, ya que todas las variables utilizadas en este trabajo son trimestrales.

Si en el modelo $ARIMA(p,d,q)(P,D,Q)_4$ se tiene que $D = Q = 1$, entonces el polinomio $MA(1)_4$ invertible se puede descomponer en tres factores simples irreducibles:

$$(1 - \Theta B^4) = (1 - \Theta^{1/2} B)(1 + \Theta^{1/4} B^2)(1 + \Theta^{1/2} B) \quad [A.1]$$

Si se relaja la restricción de que todas las raíces tengan el mismo módulo, pero se mantienen las restricciones de frecuencias, es posible estudiar la existencia de algunos factores simples MA que cancelen con los correspondientes factores simples no estacionarios que componen ∇_4 , haciendo posible la representación mediante procesos estacionales mixtos.

Se considera entonces el proceso:

$$\nabla_4 z_t = (1 - \lambda_0 B)(1 - \lambda_1 B^2)(1 - \lambda_2 B) a_t \quad [A.2]$$

donde $0 < \lambda_0 \leq 1$, $-1 \leq \lambda_1 < 0$, $-1 \leq \lambda_2 < 0$ y donde el subíndice denota la frecuencia del operador MA correspondiente.

Si el operador MA de frecuencia uno resulta no invertible, esto es, si $\lambda_1 = -1$, se cancela con el operador $(1 + B^2)$ y el proceso z_t se puede escribir:

$$z_t = \alpha_1 \cos \frac{\pi}{2} t + \beta_1 \sin \frac{\pi}{2} t + N_t \quad [A.3]$$

$$\nabla(1 + B)N_t = (1 - \lambda_0 B)(1 - \lambda_2 B)a_t$$

Si $\lambda_2 = -1$, se cancela el operador $(1 + B)$ y el proceso z_t se puede escribir:

Cuando alguno de los operadores resulta no invertible, al cancelar con los correspondientes operadores de la diferencia anual, además de introducir los términos de estacionalidad determinista,

$$z_t = \alpha_2 (-1)^t + N_t \quad [A.4]$$

$$\nabla(1 + B^2)N_t = (1 - \lambda_0 B)(1 - \lambda_1 B^2)a_t$$

se puede introducir un operador AR de la misma frecuencia que el operador MA cancelado, para realizar diagnóstico por sobreajuste de la posibilidad de una estructura estacionaria de la misma frecuencia. En este trabajo, en todos los casos en los que se ha procedido a la cancelación de un operador MA no invertible y se ha introducido un operador AR de sobreajuste, el parámetro autorregresivo ha resultado claramente no significativo.

La aplicación de esta metodología lleva a resultados útiles, no sólo en el análisis univariante, sino también en la elaboración de modelos multivariantes, porque: (1) elimina la confusión entre tendencia y estacionalidad estocásticas que surge del empleo de factores $IMA(1,1)_4$ que trata las dos con un sólo parámetro, (2) permite que el carácter de la estacionalidad de una serie sea determinista en una frecuencia y estocástica en otra y (3) facilita una flexibilidad útil en las condiciones iniciales de la función de previsión puntual.

BIBLIOGRAFÍA

Bell, W.R. (1987). A Note on Overdifferencing and the Equivalence of Seasonal Time Series Models with Monthly Means and Models with $(0,1,1)_{12}$ Seasonal Parts when $\theta = 1$. *Journal of Business and Economic Statistics*, 5, pp. 383-387.

Box, G.E.P., G.M. Jenkins y G.C. Reinsel (1994). *Time Series Analysis: Forecasting and Control*. 3ª edición. New Jersey: Prentice Hall.

Box, G.E.P. y G.C. Tiao (1975). Intervention Analysis with Applications to Economic and Environmental Data. *Journal of the American Statistical Association*, 70, pp. 70-79.

Brian, P.M. y A. Diamantopoulos (1994). Towards a Taxonomy of Forecast Error Measures. A Factor-Comparative Investigation of Forecast Error Dimensions. *Journal of Forecasting*, 13, pp. 409-416.

Caja de Madrid (6/1995-12/1996). *Previsión y Seguimiento de la Economía Española*. Madrid.

Diebold, F.X y R.S Mariano (1995). Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, 3, Vol.13, pp.253-263.

Gallego, J.L. (1995). *Una Familia general de procesos estocásticos estacionales*. Tesis Doctoral. Facultad de Ciencias Económicas y Empresariales. Universidad Complutense de Madrid.

Gallego, J.L. y A.B. Treadway (1996). *The General Seasonal ARIMA Family of Stochastic Processes*. DT 96-01. Departamento de Economía, Universidad de Cantabria.

García-Ferrer A., J. Hoyo, A. Martín-Arroyo y P. Young (1997). On Univariate Forecasting Comparisons: The Case of the Spanish Automobile Industry. *Journal of Forecasting*, 16, pp. 117-125.

Ljung, G.M. y G.E.P. Box (1978). On a Measure of Lack of Fit in Time Series Models. *Biometrika*, 65(2), pp. 297-303.

Mauricio, J.A. (1995). Exact Maximum Likelihood Estimation of Stationary Vector ARMA Models. *Journal of American Statistical Association*, 90, pp. 282-291.

Mauricio, J.A. (1996). Some Computational Aspects of Exact Maximum Likelihood Estimation of Time Series Models. *COMPSTAT 1996 - Proceedings on Computational Statistics*, pp. 361-366. Heidelberg: Physica-Verlag.

Mauricio, J.A. (1997). The Exact Likelihood Function of a Vector Autorregresive Moving Average Model. *Applied Statistics*, 46(1), pp. 157-171.

Relloso, S. (1997). *Un Modelo Multivariante para la Previsión y el Seguimiento del Empleo por Sectores, Activos y Parados*. Tesis Doctoral. Facultad de Ciencias Económicas y Empresariales. Universidad Complutense de Madrid.

Relloso, S. (1998). *Un Modelo Multivariante para el Empleo por Sectores, Activos y Parados*. DT del Instituto Complutense de Análisis Económico. En Impresión.

Zellner, A. (1983). *Handbook of Econometrics*, Volumen I. Z. Griliches y M.D. Intriligator editores. North-Holland Publishing Company.

Tabla 1: Representación con Modelos Estacionales Generalizados

SERIE	FACTOR AR CON RAÍCES DE MÓDULO UNO	TÉRMINOS DE INTERVENCIÓN DETERMINISTAS			PARÁMETROS OPERADOR MA			$\hat{\sigma}_a$					
		FECHA	TIPO	$(\hat{\omega}_0 - \hat{\omega}_1 B)$		$(1 - \hat{\theta}B)(1 - \hat{\lambda}_1 B)(1 - \hat{\lambda}_2 B)$							
				$\hat{\omega}_0$	$\hat{\omega}_1$	$\hat{\theta}$	$\hat{\lambda}_1$		$\hat{\lambda}_2$				
EA_t	$\nabla\nabla_4$	I/85	S	86.5(11.5)		.43(.20)	-.68(.10)	-.72(.14)	27.4				
		II/87	S	63.3(20.1)									
		I/89	S	-84.0(22.8)									
		IV/89	S	-70.9(20.5)									
		I/92	S	42.4(22.3)									
EI_t	$\nabla\nabla_4$	II/80	I	-49.0(8.8)		.46(.14)	-1.0(.10)	-1.0(.04)	18.9				
		II/82	IC	18.3(5.7)									
		II/87	S	63.0(6.9)									
		I/88	S	-52.7(15.0)									
		IV/92	S	-61.8(15.8)	82.0(14.9)								
			$g = -143.8(18.9)$										
EC_t	$\nabla\nabla_4$	I/92	S	-30.9(17.9)		.60(.10)	-.83(.14)	-.94(.05)	19.0				
EP_t	$\nabla\nabla_4$	II/80	S	-24.0(13.8)	40.2(26.5)	.44(.24)	-.71(.11)	-.77(.19)	32.7				
					$g = -64.2(31.3)$								
		II/87	S	196.9(27.0)									
		III/87	S	33.8(17.4)									
		III/88	I	47.7(17.6)									
EU_t	∇^2	I/87	S	-60.8(11.6)	118.7(11.1)	.88(.05)			16.7				
					$g = -179.5(.52)$								
		III/88	I	-60.3(11.1)									
		I/92	S	-44.6(16.2)									
		IV/92	S	-44.5(15.9)	39.2(16.1)								
			$g = -83.7(21.7)$										

Tabla 2: Incidentes influyentes del Modelo (7)

SERIE	TÉRMINOS DE INTERVENCIÓN DETERMINISTAS			
	FECHA	TIPO	$(\hat{\omega}_0 - \hat{\omega}_1 B)$	
			$\hat{\omega}_0$	$\hat{\omega}_1$
A_t	II/87	S	190.5(39.1)	
	IV/88	S	-107.4(41.6)	
	I/92	S	-41.1(22.6)	
	II/93	S	100.6(23.4)	-59.1(22.9)
			g = 159.7(39.8)	
EA_t	I/85	S	91.3(19.1)	
	II/87	S	51.5(18.4)	
	I/89	S	-83.6(19.1)	
	IV/89	S	-73.7(19.1)	
	I/92	S	47.9(19.8)	
EP_t	II/80	S	-58.7(24.5)	73.6(25.3)
			g = -132.3(30.0)	
	II/87	S	227.6(25.4)	
	III/87	S	51.2(25.8)	
	III/88	I	54.2(24.9)	
EI_t	II/80	I	-51.5(8.4)	
	II/82	IC	18.7(4.8)	
	II/87	S	70.4(14.5)	
	I/88	S	-49.3(14.3)	
	IV/92	S	-56.2(13.3)	84.3(13.2)
			g = -140.5(15.7)	
EU_t	I/87	S	-65.1(14.2)	120.9(14.1)
			g = -186.0(25.4)	
	III/88	I	-49.0(10.0)	
	I/92	S	-39.2(14.2)	
	IV/92	S	-46.7(14.9)	35.0(15.1)
		g = -81.7(26.3)		
EC_t	II/83	S	32.0(10.2)	
	II/89	I	-50.2(10.4)	
	I/92	S	-41.0(10.5)	

Tabla 3: Comparación de los resultados obtenidos en la previsión de EP_t con: (1) el modelo univariante y (2) el modelo MS.

Tipo de modelo	DTE		RECM		Comparación valor absoluto del error	
	(1)	(2)	(1)	(2)	(1)	(2)
Horizonte 1	32.2	21.4	44.2	49.9	6	5
Horizonte 2	53.7	32.7	58.7	67.5	5	5
Horizonte 3	76.8	49.9	88.5	64.1	4	5
Horizonte 4	99.1	64.5	129.3	45.1	2	6
Horizonte 5	131.9	78.3	181.8	66.4	1	6
Horizonte 6	164.2	92.1	247.9	115.8	0	6
Horizonte 7	197.6	106.3	298.5	165	0	5
Horizonte 8	229.9	121	355	214	0	4

Tabla 4: Comparación de los resultados obtenidos en la previsión de EI_t con: (1) el modelo univariante y (2) el modelo MS.

Tipo de modelo	DTE		RECM		Comparación valor absoluto del error	
	(1)	(2)	(1)	(2)	(1)	(2)
Horizonte 1	18.1	16.3	23.4	24.6	5	6
Horizonte 2	33.3	31.1	42.5	46	8	2
Horizonte 3	50.2	48.5	59.9	65.9	5	4
Horizonte 4	69.1	68.1	77.2	83.7	5	3
Horizonte 5	89.6	88.5	104.7	107.9	4	3
Horizonte 6	111.9	110	137.4	144.1	5	1
Horizonte 7	135.6	132.7	183.3	172.4	3	2
Horizonte 8	160.8	156.5	214.4	208.2	1	3

Tabla 5: Comparación de los resultados obtenidos en la previsión de EU_t con: (1) el modelo univariante y (2) el modelo MS.

Tipo de modelo	DTE		RECM		Comparación valor absoluto del error	
	(1)	(2)	(1)	(2)	(1)	(2)
Horizonte 1	16.5	14.1	27.1	28.4	6	5
Horizonte 2	24.8	21.1	46	49.3	6	4
Horizonte 3	32.3	27.8	54.4	57.1	5	4
Horizonte 4	39.4	34.1	58.2	53.5	2	6
Horizonte 5	46.6	40.3	67.1	58	1	6
Horizonte 6	53.8	46.5	79.6	71.2	1	5
Horizonte 7	61.1	52.8	87.7	68.1	0	5
Horizonte 8	68.5	59.2	86.2	67.8	0	4

Tabla 6: Comparación de los resultados obtenidos en la previsión de E_t mediante: (1) la agregación de componentes usando sus modelos univariantes, (2) el modelo US del agregado y (3) la agregación de componentes usando el modelo MS.

Tipo de modelo	DTE			RECM			Comparación valor absoluto del error			
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(3)	(2)	(3)
Horizonte 1	51.1	46	37.6	64	56.6	58.6	6	5	4	7
Horizonte 2	85.5	86.7	66.2	99.6	120.5	87.8	6	4	5	5
Horizonte 3	122.5	128.5	104.4	168.2	190.5	117.8	2	7	2	7
Horizonte 4	159.3	174.9	142.1	257	285.7	163.8	2	6	1	7
Horizonte 5	209.6	229.6	180.8	378.6	413.6	263.1	0	7	0	7
Horizonte 6	260.1	289.4	220.5	522	569.9	381.4	1	5	0	6
Horizonte 7	312.6	350.4	261.9	661.1	717.9	507.7	1	4	0	5
Horizonte 8	364.6	414.5	304.7	804	862.8	652.6	1	3	0	4

Tabla 7: Comparación de los resultados obtenidos en la previsión de A_t mediante: (1) el modelo univariante, (2) el modelo UT de A_t con E_t agregado y (3) el modelo MS con E_t desagregado por sectores.

Tipo de modelo	DTE			RECM			Comparación valor absoluto del error			
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(3)	(2)	(3)
Horizonte 1	34.5	31.7	29.5	44.6	43.1	51.4	7	4	6	5
Horizonte 2	58.1	54.3	50.9	74	69.6	82.3	7	3	6	4
Horizonte 3	76.9	73	70.8	103.7	76.7	76.5	4	5	5	4
Horizonte 4	97.3	94.1	90.9	140.4	89.6	62.2	3	5	2	6
Horizonte 5	124.4	121.7	113.1	165.3	99.3	64.5	3	4	2	5
Horizonte 6	153.4	151.5	137.9	183	125.4	72.1	3	3	1	5
Horizonte 7	179.6	179.5	163.2	191.4	151	86.6	1	4	0	5
Horizonte 8	206.8	208.8	188.9	151.6	165.6	100.8	1	3	0	4

Tabla 8: Comparación de los resultados obtenidos en la previsión de D_t usando la identidad (1) a partir de: (1) los modelos univariantes de E_t y A_t , (2) el modelo UT de A_t con E_t agregado y (3) el modelo MS con E_t desagregado por sectores.

Tipo de modelo	DTE			RECM			Comparación valor absoluto del error			
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(3)	(2)	(3)
Horizonte 1	57.5	55.8	31.6	62.3	68.2	43.8	3	8	2	9
Horizonte 2	104.4	102.3	51.6	132.2	141.5	69.3	1	9	2	8
Horizonte 3	149.8	147.8	78	207	211.1	131.3	2	7	2	7
Horizonte 4	200.2	198.6	104.6	293.7	290	199.4	0	8	0	8
Horizonte 5	261.1	259.8	129.3	375.2	377.1	279.4	0	7	0	7
Horizonte 6	327.6	326.7	156.5	467.4	485.5	363.3	0	6	1	5
Horizonte 7	393.8	393.7	184.7	580.9	597.8	455.2	0	5	0	5
Horizonte 8	463.2	464.1	214.7	712.4	713.9	564.4	0	4	0	4