TWO PAPERS ON ARIMA SIGNAL EXTRACTION

Agustín Maravall

Banco de España. Servicio de Estudios
Documento de Trabajo n.º 8801

El Banco de España al publicar esta serie pretende facilitar la difusión de estudios de interés que contribuyan al mejor conocimiento de la economía española.

Los análisis, opiniones y conclusiones de estas investigaciones representan las ideas de los autores, con las que no necesariamente coincide el Banco de España.

ISBN: 84-7793-002-3
Depósito legal: M. 9811 - 1988

Imprenta del Banco de España

INDICE

Introducción

La descomposición de series temporales en componentes (y en concreto, la desestacionalización y la estimación de tendencias) por medio de un método basado en modelos consiste, fundamentalmente, en dos partes:

(a) La especificación de los modelos y la estimación de los parámetros de los mismos.

(b) La estimación de los componentes (o señales).

El método hoy en día más desarrollado en este campo es el que utiliza modelos ARIMA en la parte (a), y técnicas de extracción de señales (el llamado filtro de Wiener-Kolmogorov, abreviadamente W-K) en la parte (b).

En lo que concierne a la especificación de los modelos, el problema de la especificación del componente estacional plantea problemas importantes y sucede, con cierta frecuencia, que los métodos basados en modelos parten de una especificación incorrecta. En el primero de los dos trabajos que adjunto presento una forma de abordar el problema de descomponer una serie que elimina varias de las fuentes de error que surgen en la especificación, y que se extiende de forma natural a la especificación de cualquier componente cíclico. El enfoque consiste en utilizar componentes elementales, fáciles de modelizar e interpretar. Cada uno de estos componentes se asocia con una frecuencia específica (por ejemplo, con cada una de las frecuencias estacionales); a continuación, los componentes pueden agregarse fácilmente (y obtenerse, por ejemplo, el componente estacional total).

En cuanto a las técnicas de extracción de señales en procesos estocásticos, estas últimas estaban disponibles desde hace algún tiempo para series estacionarias (el filtro de W–K). Puesto que la inmensa mayoría de las series económicas son no–estacionarias, las técnicas tenían poco interés aplicado. La generalización a series no–estacionarias se ha producido en los últimos años, gracias fundamentalmente a trabajos de Cleveland y Tiao y de Bell, ambos densos y díficiles. Por otra parte, puesto que la solución utiliza unas funciones que aparentemente no están definidas en el caso no–estacionario, la solución no resulta de ningún modo intuitiva.

En el segundo trabajo que incluyo, presento una derivación original alternativa de la solución óptima para la extracción de cualquier señal en series no–estacionarias. La derivación es intuitiva y sorprendentemente sencilla. La racionalización práctica del trabajo puede realizarse de la siguiente manera: La extensión del filtro de W–K a series no–estacionarias es lo que ha permitido el desarrollo de programas como el de Burman o el del SCA, que se utilizan constantemente en el Servicio de Estudios (también ha permitido un análisis más riguroso de X11). Puesto que las herramientas analíticas se utilizan mejor cuando se comprende porqué funcionan, si se facilita y simplifica esta comprensión, mejorará el uso de esas herramientas.

ON THE DYNAMIC STRUCTURE
OF A SEASONAL COMPONENT

Summary

A model often used to characterize the dynamic structure of a seasonal component in unobserved component models and, in particular, in seasonal adjustment is analysed. It is found that the characterization is inadequate, since the seasonal component itself would then accept a perfectly sensible decomposition into trend, seasonal and irregular components. In the analysis, an alternative approach to the problem of specifying component models is suggested.

Key words: Time Series models; ARIMA models; Seasonal Adjustment; Unobserved-Components model; Signal Extraction; Dynamic Seasonality.

1. Introduction

Over the past years, model-based estimation of unobserved components has been increasingly developed, with the main purpose of providing a flexible alternative to "ad hoc" methods such as X11 or X11 ARIMA. Given that the most important application of unobserved components estimation is seasonal adjustment, a crucial part of a model-based seasonal adjustment method is the specification of the particular model assumed to generate the seasonal component.

Let s denote the number of observations in one year and B the lag operator. A model often used to characterize the seasonal component (z_t) is an ARIMA process of the type

$$(1-\Phi B^s)z_t = \theta(B)a_t \ , \qquad (1)$$

where $0<\Phi\leq 1$, $\theta(B)$ represents a moving average (MA) polynomial in the lag operator B, and a_t is a white-noise series. This type of model for the seasonal component has been used by Nerlove, Grether and Carvalho (1979), Pierce (1978), Pagan (1975), Engle (1978), Cleveland and Tiao (1976), Granger (1978), Harvey (1981), Ansley (1983), Gourieroux and Monfort (1983), Pierce, Cleveland and Grupe (1984), Hausman and Watson (1985), and Hylleberg (1986) among others. Although, in later work, some of these authors have changed the specification of the seasonal component, models that fit into (1) are nevertheless heavily used.

In the next section using, for simplicity of exposition, a particular case of (1), it is argued that such a specification is inappropriate to characterize the dynamic structure of a seasonal component. The discussion suggests a natural approach to the problem of specifying models for the components where, in a first stage, elementary components are assigned to each individual spectral peak and, in a second stage, the components can be aggregated as desired.

2. An Example

Consider a quarterly seasonal series that follows the model

$$\nabla_4 z_t = a_t \ , \qquad (2)$$

where $\nabla_4 = 1-B^4$. The pseudospectrum of z_t, equal to

$$g_z(\omega) = \frac{1}{2(1-\cos 4\omega)} \sigma_a^2, \quad (0 \leq \omega \leq \pi), \qquad (3)$$

is displayed in Figure 1 (the prefix "pseudo" will be removed in later references.) It is symmetric around $\omega=\pi/2$ and presents three peaks, associated with the frequencies $\omega=0$, $\omega=\pi/2$ and $\omega=\pi$. One may try to decompose the series into orthogonal components, each one capturing a spectral peak. Thus we seek to express z_t as:

$$z_t = z_{1t}+z_{2t}+z_{3t}+u_t \ , \qquad (4)$$

where u_t is a white-noise residual. We shall only concern ourselves with the specification of the models for the components; once these are specified, estimation, diagnosis and inference can be carried out as in, for example, Bell and Hillmer (1984) and Maravall (1988).

The roots of the autoregressive (AR) polynomial $(1-B^4)$ are

$$\nabla_4 = (1-B)(1+B)(1+B^2) \ . \tag{5}$$

and it is easily seen that the spectral peak for $\omega=0$ is induced by the AR factor $(1-B)$ in (5) and, similarly, the peaks for $\omega=\pi$ and $\omega=\pi/2$ are induced by the AR factors $(1+B)$ and $(1+B^2)$, respectively. Thus, from (2), (4) and (5), in order to capture the individual spectral peaks, the z_i-components will have to be of the type:

$$\begin{aligned} (1-B)\ z_{1t} &= \alpha_1(B)\ b_{1t} \\ (1+B)\ z_{2t} &= \alpha_2(B)\ b_{2t} \\ (1+B^2)\ z_{3t} &= \alpha_3(B)\ b_{3t} \end{aligned} \tag{6}$$

where $\alpha_i(B)$ represents some polynomial in B, and b_{1t}, b_{2t} and b_{3t} are mutually independent white noises, also independent of u_t.

Replacing in (4) the components by their expressions in (6) and considering (2), the following identity is obtained:

$$a_t = (1+B)(1+B^2)\ \alpha_1(B)\ b_{1t} + (1-B)(1+B^2)\ \alpha_2(B)\ b_{2t} +$$

$$+ (1-B)(1+B)\ \alpha_3(B)\ b_{3t} + \nabla_4\ u_t . \tag{7}$$

Since the r.h.s. of (7) has to be white noise, the lag-4 autocorrelation of the term $\nabla_4\ u_t$ should cancel out with that of the other terms. Hence $\alpha_1(B)$ and/or $\alpha_2(B)$ have to be at least of order one, and/or $\alpha_3(B)$ has to be at least of order two. Therefore, we can assume that

$$\begin{aligned} \alpha_1(B) &= 1-\alpha_{11}\ B \\ \alpha_2(B) &= 1-\alpha_{21}\ B \\ \alpha_3(B) &= 1-\alpha_{31}\ B-\alpha_{32}\ B^2 \end{aligned} \tag{8}$$

Considering (6) and (8), the models for the four components are seen to depend on eight parameters: —the four α–parameters and the four variances (σ_i^2 of b_{it}, i=1,2,3, and σ_u^2)-- which have to satisfy the constraints implied by equating the autocovariances of the l.h.s. and the r.h.s. of (7). These constraints yield a system of five equations with eight unknowns, and hence there will be an infinite number of parameter values in the component models which satisfy the system. In econometrics terminology, if equations (4), (6) and (8) describe the "structure" of the model, an infinite number of structures are compatible with the "reduced" form model, given by equation (2), and therefore the models for the components are not identified.

In order to achieve identification, additional information is required. Following Box, Hillmer and Tiao (1978) and Pierce (1978), the following requirement will be imposed: Let z_{it} denote any component (not u_t). Then, z_{it} should not accept a decomposition of the type:

$$z_{it} = z^*_{it} + n_t \quad ,$$

where z^*_{it} and n_t are independent and the latter is white noise. (If such a decomposition were feasible, then the component z_{it} should be replaced with z^*_{it}, and n_t should be added to u_t.) We refer to this requirement as the "canonical" requirement.

Let $g_i(\omega)$ be the spectrum of z_{it}, for $0 \le \omega \le \pi$. The canonical requirement implies that, for some ω in that range, $g_1(\omega)$ should be zero. From (6) and (8) it is found that $g_1(\omega)$ is monotonically decreasing in ω, hence no noise will contaminate z_{1t} when

$g_1(\pi)=0$. Since this condition implies the presence of the factor $(1+\cos\omega)$ in the numerator of $g_1(\omega)$, in the time domain, it is equivalent to the presence of the factor $(1+B)$ in $\alpha_1(B)$ and, therefore, the model for z_{1t} is given by

$$(1-B)z_{1t} = (1+B)b_{1t}. \tag{9a}$$

Similarly, from (6) and (8), it is seen that $g_2(\omega)$ is monotonically increasing in the range $0<\omega<\pi$. Hence the canonical requirement implies $g_2(0)=0$, or equivalently, the presence of the factor $(1-\cos\omega)$ in the numerator of $g_2(\omega)$. Thus $\alpha_{21}=1$ and the model for z_{2t} is equal to

$$(1+B)z_{2t} = (1-B)b_{2t} \quad . \tag{9b}$$

Concerning the third component z_{3t}, because of symmetry, its spectrum will reach a minimum of zero, for $\omega=0$ and $\omega=\pi$. Hence the two factors $(1-B)$ and $(1+B)$ have to be present in $\alpha_3(B)$, so that $\alpha_3(B)=(1-B)(1+B)=1-B^2$. The model for z_{3t} is therefore given by

$$(1+B^2)z_{3t} = (1-B^2)b_{3t} \quad . \tag{9c}$$

Considering (9a), (9b) and (9c), it is seen that the canonical requirement has allowed us to identify the α-parameters of the component models. This is sufficient to identify fully these models since, plugging the α-values in the system of autocovariance equations implied by (7), a unique solution for the four unknown variances is obtained. This solution is given by

$$\sigma_1^2 = \sigma_2^2 = \sigma_a^2/64 \quad ,$$

$$\sigma_3^2 = \sigma_a^2/16 \quad , \tag{10}$$

$$\sigma_u^2 = 3\ \sigma_a^2/32 \quad .$$

Expressions (9) and (10) completely specify the models for the components; Figure 2 exhibits the z_{it} –component spectra. The first component, z_{1t}, obviously represents a trend component, and the white-noise u_t an irregular component. The other two components, z_{2t} and z_{3t}, contain the series variation for the frequencies $\omega=\pi$ and $\omega=\pi/2$, the twice-a-year and once-a-year seasonal frequencies in quarterly data.

The approach we have outlined provides elementary components, each one (except the irregular) unambiguously assigned to a peak in the series spectrum. In a second stage, the basic components can be aggregated as desired. Thus, in our example, the total seasonal component, s_t, would be equal to

$$s_t = z_{2t} + z_{3t} \quad ,$$

and, replacing z_{2t} and z_{3t} with their expressions (9b) and (9c), it is obtained that

$$(1+B)(1+B^2)s_t = (1+B^2)(1-B)b_{2t}+(1+B)(1-B^2)b_{3t} \tag{11}$$

The r.h.s. of (11) is an MA(3), say $\beta(B)$, for which $\beta(1)=0$. The total seasonal componen' will follow then the model,

$$(1+B+B^2+B^3)s_t = (1-B)(1-\beta_1 B-\beta_2 B^2)c_t \tag{12}$$

where c_t is white-noise. Equating the autocovariances of the r.h.s. of (11) and of (12), it is found that β_1=–.819, β_2 = –.344, and σ_c^2 = .227 σ_a^2; the spectrum of s_t is given in Figure 3. The model obtained for s_t is, in this case, the same that would result from the seasonal adjustment method developed in Burman (1980) –who first made the point that the spectral peak at ω=0, implied by a seasonal difference, should not be part of the seasonal component– and Hillmer and Tiao (1982); this equivalence, however, will not be true in general. Alternatively, the model for the seasonally adjusted series, z_t^a, can be obtained by summing the other two components z_{1t} and u_t, which eventually yields

$$(1-B)z_t^a = (1-.42B)d_t, \tag{13}$$

with d_t white noise and σ_d^2=.186 σ_a^2.

Since there is no generally accepted definition of what is a seasonal component, the choice of a model for said component is, to some degree, arbitrary. Be that as it may, in the additive decomposition of z_t as in (4), the components z_{2t} and z_{3t} are clearly associated with the seasonal variation, but in which way can $z_{1t}+u_t$ –or, equivalently, expression (13)– also be considered to represent seasonal variation?

Notice that estimation criteria, however, cannot be used to decide whether the seasonal component should be characterized by, say, model (2) or (12). Having estimated model (2), it is always possible to decompose it as in (12) plus (13), and both representations will be observationally equivalent. The decision of which representation should be used for the seasonal component depends on its implicit or explicit definition, and it is difficult to accept a definition that includes as part of the seasonal a spectral peak for the (nonseasonal) frequency $\omega=0$, just as a spectral peak for the seasonal frequency $\omega=\pi$ would not be assigned to the trend.

3. Some Extensions

If, in (2), ∇_4 is replaced by $1-\Phi B^4$, with $0<\Phi\leq 1$, and a_t is replaced by an invertible moving average $\theta_q(B)\, a_t$, with $q\leq 4$, the discussion remains basically unchanged. The variable z_t will, in general, accept a perfectly sensible decomposition into trend plus seasonal plus irregular, although the numerical values of some of the parameters in the component models will change. The discussion can also be easily extended to the case s=12 by introducing new components for the additional seasonal frequencies.

In the final analysis, although a model such as (1) seems inadequate to characterize a seasonal component, it is of interest to know the practical effect of this inadequacy. As a first example, the model

$$(1-.8B^4)z_t = a_t \quad , \qquad (14)$$

very close to the ones used in Hylleberg (1986, Ch. 7) to characterize the stochastic seasonal component of several series, can be decomposed –following a reasoning similar to that in Section 2– into a purely seasonal component and a nonseasonal one consisting of the sum of a trend, given by

$$(1-.95\ B)p_t = (1+B)c_t \quad ,$$

with $\sigma_c^2 = .0366\ \sigma_a^2$, and a white–noise irregular, with $\sigma_u^2 = .1161\ \sigma_a^2$. It is easily seen that more than half of the variance of z_t in (14) is explained by its nonseasonal component. To have an idea of the relative size of this component, its standard deviation equals 72.6% of the standard deviation of z_t.

Another example, with monthly data this time, is provided by the model

$$(1-.5\ B^{12})z_t = (1-.6\ B)a_t\ ,$$

very similar to the ones used by Hausman and Watson (1985) for their stochastic seasonals. Since $(.944)^{12}=.5$, using the factorization $1-.5B^{12}=(1-.944B)(1+.944B+(944\ B)^2+...+(.944B)^{11})$, as before, the component z_t accepts a seasonal plus nonseasonal decomposition. The trend is given by

$$(1-.944\ B)p_t = (1+B)c_t\ ,$$

with $\sigma_c^2=.00095\ \sigma_a^2$, and the variance of the noise is $\sigma_u^2=.0325\ \sigma_a^2$. Although in this case most of the variation of z_t is "purely" seasonal, the standard deviation of the nonseasonal component it contains is found to be equal to 19.2% of the standard deviation of z_t, a relatively small (though not negligible) percentage.

4. A Final Comment

The approach followed is easily extended to the case of series with different frequencies of observations (see Maravall and Pierce, 1987). It can also be extended to estimate components associated with nonzero and nonseasonal frequencies. Consider, for example, a series that presents a nonstationary cycle of period T (larger than a year), revealed perhaps by the factorization of the autoregressive polynomial in the model for the observed series, or perhaps by a priori knowledge that such a cycle exists. One of the components of the series would follow then the model

$$(1-\Phi B+B^2)z_{it} = (1+B)(1-\beta B)\ a_{it}, \tag{15}$$

where $\Phi=2\cos\omega$ and $\omega=2\pi/T$. The model depends on two parameters, β and σ_i^2, which have a rather natural interpretation. The variance of a_{it} (expressed as a percentage of the variance in the series innovation) reflects the relative importance of the cyclical component in the series, while the β-parameter will be associated with its random nature. Larger values of β imply more stable components, as seen in Figure 4, which displays the spectra of a cyclical component with a period of 2 1/2 years in quaterly data, for $\beta=\pm 5$ and $\sigma_i^2=1$.

The ARIMA representation of the cycle, given by (15), can of course be translated into a state-space representation following, for example, Aoki (1983, pp. 29–31). In fact, component models somewhat similar to (15) is a state space framework have been used by Harvey (1985) to estimate cycles.

REFERENCES

Ansley, C.F. (1983), "Comment to 'Forecasting Economic Time Series with Structural and Box–Jenkins Models'," Journal of Business and Economic Statistics, 1, 4, 307–309.

Aoki, M. (1983), Notes on Economic Time Series Analysis: System Theoretic Perspectives, Berlin, Springer-Verlag.

Bell, W.R. and Hillmer, S.C. (1984), "Issues Involved with the Seasonal Adjustment of Economic Time Series", Journal of Business and Economic Statistics, 2, 291–320.

Box, G.E.P., Hillmer, S.C. and Tiao, G.C. (1978), "Analysis and Modeling of Seasonal Time Series" in Seasonal Analysis of Economic Time Series, ed. A. Zellner, Washington, D.C.: U.S. Dept. of Commerce – Bureau of the Census, 309–334.

Burman, J.P. (1980), "Seasonal Adjustment by Signal Extraction", Journal of the Royal Statistical Society, A, 143, 321–337.

Cleveland, W.P. and Tiao, G.C. (1976), "Decomposition of Seasonal Time Series: A Model for the X–11 Program", Journal of the American Statistical Association, 71, 581–587.

Engle, R.F. (1978), "Estimating Structural Models of Seasonality", in Seasonal Analysis of Economic Time Series, ed. A. Zellner, Washington, D.C.: U.S. Dept. of Commerce–Bureau of the Census, 281–297.

Granger, C.W.J. (1978),"Seasonality:Causation, Interpretation and Implications" in Seasonal Analysis of Economic Time Series, ed. A. Zellner, Washington, D.C.: U.S. Dept. of Commerce - Bureau of the Census, 33–46.

Gourieroux, A. and Monfort, A. (1983), Cours de Series Temporelles. Paris: Economica.

Harvey, A.C. (1981), Time Series Models, Deddington: Philip Allan.

Harvey, A.C. (1985), "Trends and Cycles in Macroeconomic Time Series", Journal of Business and Economic Statistics, 3, 216–227.

Hausman, J.A. and Watson, M.W. (1985), "Errors-in-Variables and Seasonal Adjustment Procedures", Journal of the American Statistical Association, 80, 531–540.

Hillmer, S.C. and Tiao, G.C. (1982), "An ARIMA-Model Based Approach to Seasonal Adjustment", Journal of the American Statistical Association, 77, 63–70.

Hylleberg, S. (1986), Seasonality in Regression, New York: Academic Press.

Maravall, A. and Pierce, D.A. (1987), "A Prototypical Seasonal Adjustment Model", Journal of Time Series Analysis, 8, 177–193.

Maravall, A. (1988), "The Use of ARIMA Models in Unobserved Components Estimation", in Dynamic Econometric Modelling, eds. W. Barnett, E. Berndt y H. White, Cambridge University Press, 171–196.

Nerlove, M., Grether, D.M. and Carvalho, J.L. (1979), Analysis of Economic Time Series: A Synthesis, New York: Academic Press.

Pagan, A. (1975), "A Note on the Extraction of Components from Time Series", Econometrica, 43, 1, 163–168.

Pierce, D.A., Grupe, M.R. and Cleveland, W.P. (1984), "Seasonal Adjustment of the Weekly Monetary Aggregates: A Model–Based Approach", Journal of Business and Economic Statistics, 2, 3, 260–271.

Pierce, D.A. (1978), "Seasonal Adjustment When Both Deterministic and Stochastic Seasonality are Present", in Seasonal Analysis of Economic Time Series, ed. A. Zellner, Washington, D.C.: U.S. Dept. of Commerce – Bureau of the Census, 242–269.

FIGURE 1.

SPECTRUM OF $\nabla_4 \; z_t = a_t$

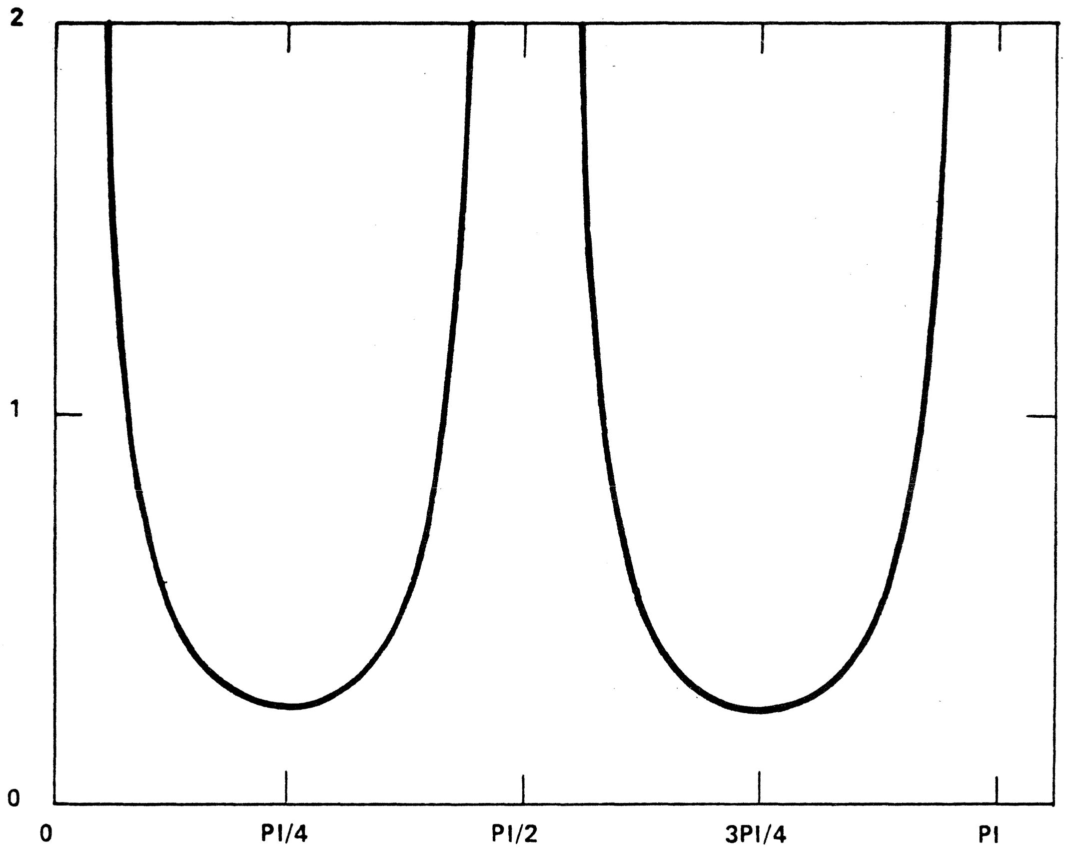


FIGURE 2.

COMPONENTS SPECTRA FOR THE MODEL $\nabla_4 z_t = a_t$

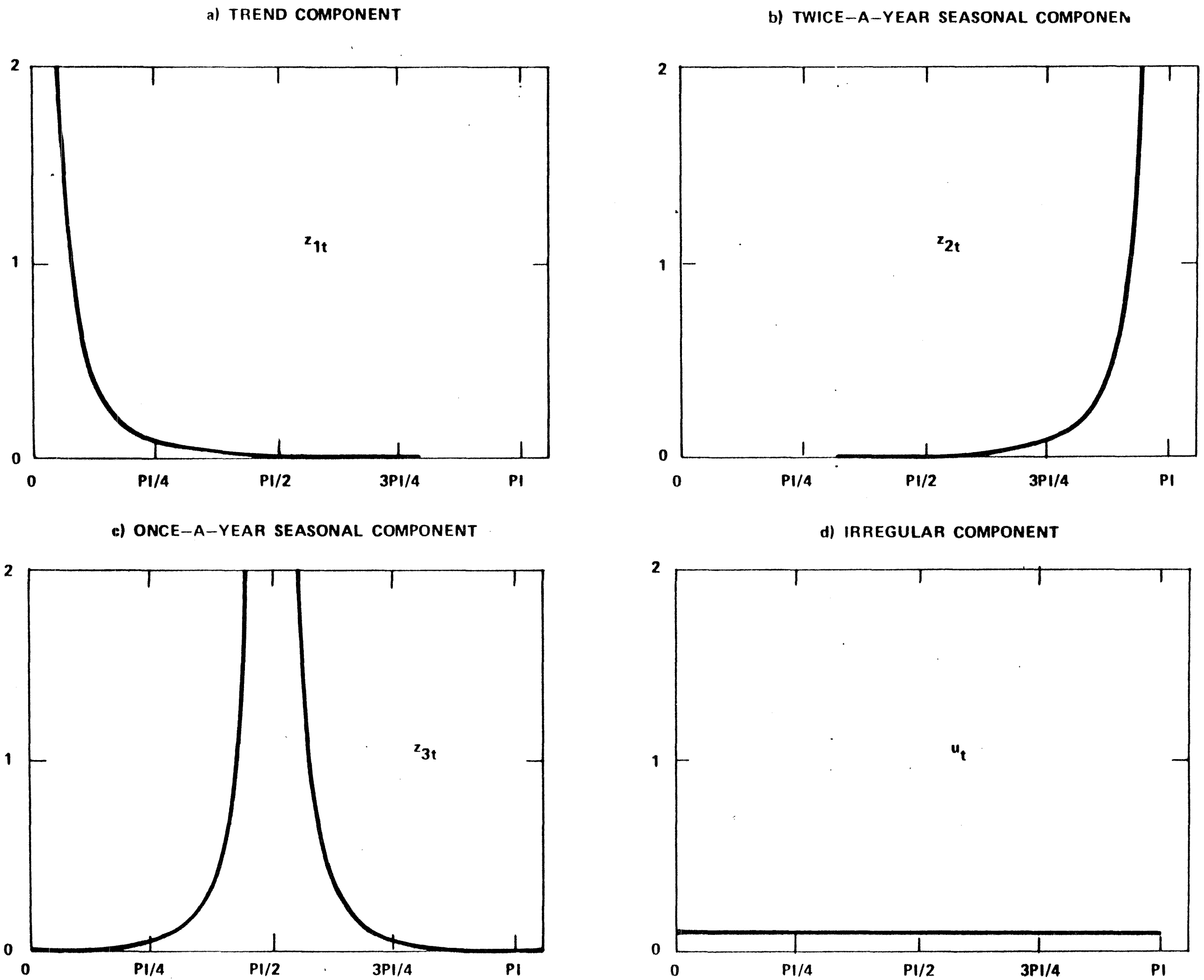


FIGURE 3.

SPECTRUM OF THE AGGREGATED SEASONAL COMPONENT

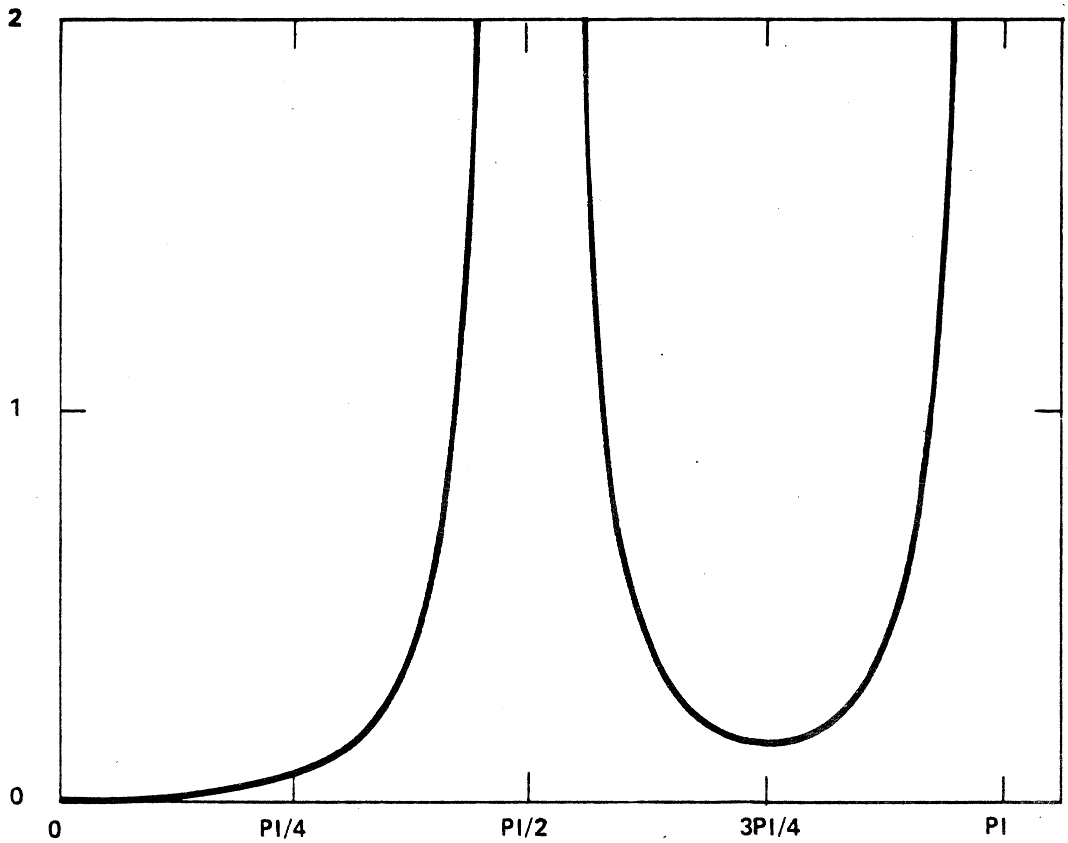
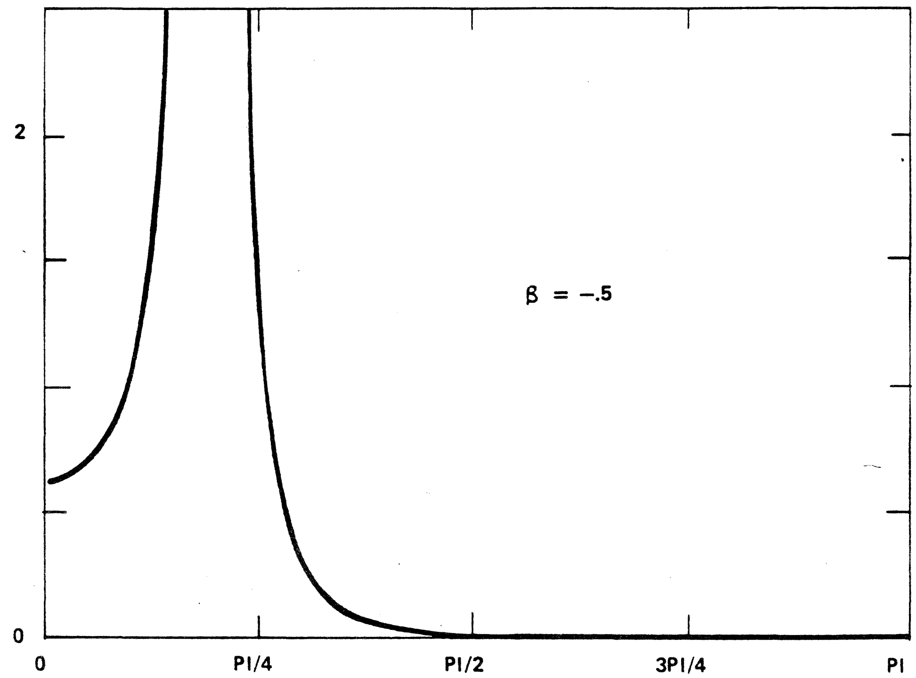
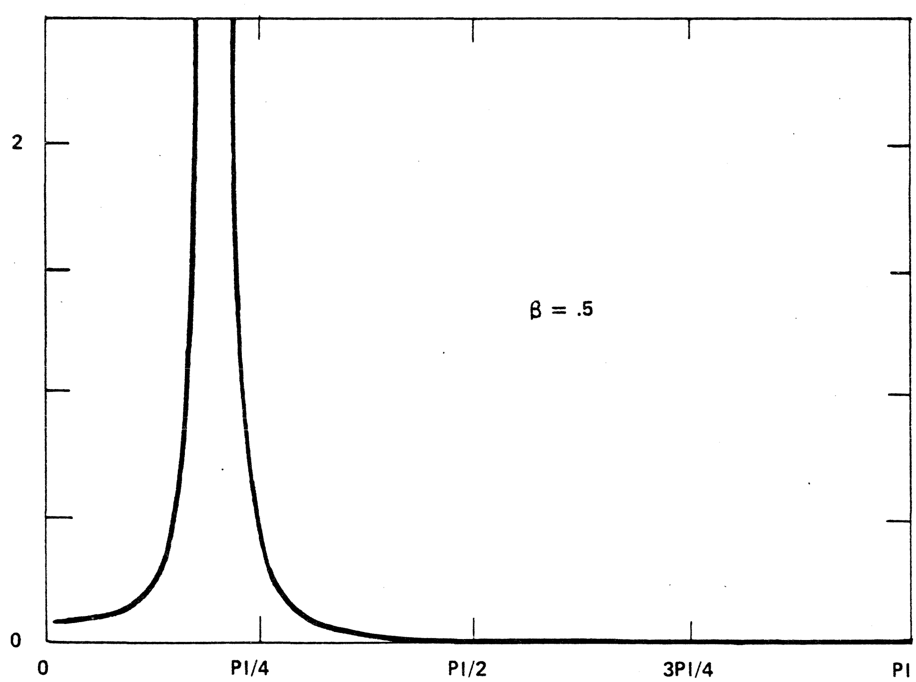


FIGURE 4.

SPECTRUM OF A CYCLICAL COMPONENT

(Frequency: once every 10 periods)





A NOTE ON MINIMUM MEAN SQUARED ERROR ESTIMATION OF SIGNALS WITH UNIT ROOTS

Abstract

Using an ARIMA parametrization, this note provides a very simple proof of how the Wiener-Kolmogorov-Whittle filter to estimate signals in time series can be extended to the nonstationary case. The proof is valid for any number and type of unit roots (not simply those implied by differencing) in both the signal and the overall model.

Introduction

The use of the Wiener–Kolmogorov filter, as developed in Whittle (1963), to estimate signals or unobserved components in ARIMA models became a powerful tool in applied time series work when the technique was extended to cover nonstationary series and signals. Key references are the pioneer work of Cleveland and Tiao (1976) and the comprehensive paper by Bell (1984). An important practical application, for example, has been the so-called ARIMA model-based seasonal adjustment method; see Burman (1980) and Hillmer and Tiao (1982).

In this note we present an extremely simple alternative proof of how Whittle's result can be extended, in a rather natural way, to provide the minimum mean squared error estimator of the signal in the general nonstationary (unit roots) case.

1. General Background

Consider a series x_t that can be decomposed as the sum of two independent unobserved components,

$$x_t = s_t + n_t \, , \tag{1.1}$$

which follow the ARIMA models

$$\Phi_s(B)\ s_t = \Theta_s(B)\ b_t \, , \tag{1.2a}$$

$$\Phi_n(B)\ n_t = \Theta_n(B)\ c_t \, , \tag{1.2b}$$

where b_t and c_t are independent gaussian white noises with variances V_b and V_c; and Φ_s, Θ_s, Φ_n and Θ_n are finite polynomials in B, the lag operator. Let s_t represent the signal of interest. Since the roots of the autoregressive polynomials Φ_s and

Φ_n imply peaks in the spectrum of x_t for certain frequencies, and since different components are associated with different peaks, it is assumed that Φ_s and Φ_n have no roots in common (for a more complete discussion see Pierce, 1979.) Furthermore, although θ_s and θ_n may be noninvertible, it will be assumed that they do not share the same unit root, so that the overall model for x_t is invertible.

Equations (1.1)–(1.2) imply that the observed series x_t follows the ARIMA model

$$\Phi(B)\; x_t = \theta(B)\; a_t \;, \tag{1.3}$$

where

$$\Phi(B) = \Phi_s(B)\; \Phi_n(B) \;, \tag{1.4}$$

and the polynomial $\theta(B)$ and the variance of a_t, V_a, can be obtained through the identity

$$\theta(B)\; a_t = \Phi_n(B)\; \theta_s(B)\; b_t + \Phi_s(B)\; \theta_n(B)\; c_t \;, \tag{1.5}$$

easily derived from expressions (1.1) to (1.4).

2. Estimation of the Signal

Let all summation signs in this section extend from $-\infty$ to ∞, and consider the linear estimator of s_t

$$\hat{s}_t = \nu(B)\; x_t \;, \tag{2.1}$$

where $\nu(B)=\Sigma\, \nu_j\, B^j$. In order to simplify notation, let a polynomial in B be denoted simply by a greek letter; an upper bar will denote the polynomial with B replaced by $F=B^{-1}$. Thus, for example, $\phi_s=\phi_s(B)$ and $\bar{\phi}_s=\phi_s(F)$.

The mean squared error (MSE) of $\hat{s}_t$ is

$$MSE = E(s_t-\hat{s}_t)^2 = E[s_t-\nu(s_t+n_t)]^2 =$$

$$= E\left[(1-\nu)\frac{\theta_s}{\phi_s} b_t - \nu \frac{\theta_n}{\phi_n} c_t\right]^2 \tag{2.2}$$

and it will always be finite, even when the roots of ϕ_s and ϕ_n are nonstationary, when

$$1-\nu = \phi_s \omega_1 \tag{2.3a}$$

$$\nu = \phi_n \omega_2 , \tag{2.3b}$$

where ω_1 and ω_2 are polynomials in B such that $\Sigma_j \omega_{ij}^2<\infty$ for i=1,2. From (2.3), the following identity is obtained

$$1 = \phi_s \omega_1 + \phi_n \omega_2 . \tag{2.4}$$

Equating the autocovariance generating functions on both sides of (1.5) yields

$$\theta \bar{\theta} V_a = \phi_n \bar{\phi}_n \theta_s \bar{\theta}_s V_b + \phi_s \bar{\phi}_s \theta_n \bar{\theta}_n V_c . \tag{2.5}$$

Dividing both sides of (2.5) by $\theta\bar{\theta}V_a$ and then adding and substracting the polynomial $\phi_s\phi_n\omega$ in the right hand side, where ω is any arbitrary polynomial in B, (2.5) can be rewritten

$$1 = \phi_s\left[\frac{\bar{\phi}_s \theta_n \bar{\theta}_n V_c}{\theta \bar{\theta} V_a} + \phi_n \omega\right] + \phi_n\left[\frac{\bar{\phi}_n \theta_s \bar{\theta}_s V_b}{\theta \bar{\theta} V_a} - \phi_s \omega\right], \tag{2.6}$$

Comparing (2.4) and (2.6), it is seen that

$$\omega_1 = \frac{\bar{\Phi}_s \, \theta_n \, \bar{\theta}_n \, V_c}{\theta \, \bar{\theta} \, V_a} + \Phi_n \, \omega \tag{2.7a}$$

$$\omega_2 = \frac{\bar{\Phi}_n \, \theta_s \, \bar{\theta}_s \, V_b}{\theta \, \bar{\theta} \, V_a} - \Phi_s \, \omega \tag{2.7b}$$

and the filter ν, from (2.3b), becomes:

$$\nu = \frac{\Phi_n \bar{\Phi}_n \, \theta_s \, \bar{\theta}_s \, V_b}{\theta \, \bar{\theta} \, V_a} - \Phi_s \, \omega \quad . \tag{2.8}$$

Let W be the set $W=\{\omega \,|\, \Sigma\omega_j^2<\infty\}$, and denote by Ω the set of filters ν defined by (2.8) with $\omega \in W$. The MSE of $\hat{s}_t$ will always be finite if $\nu \in \Omega$. When the roots of Φ are nonstationary, only filters $\nu \in \Omega$ will yield finite MSE estimators.

The Whittle filter, given by

$$\nu_o = \frac{V_b}{V_a} \; \frac{\Psi_s \, \bar{\Psi}_s}{\Psi \, \bar{\Psi}} \quad ,$$

where $\Psi_s=\theta_s/\Phi_s$ and $\Psi=\theta/\Phi$, becomes, in our ARIMA notation,

$$\nu_o = \frac{V_b}{V_a} \; \frac{\Phi_n \, \bar{\Phi}_n \, \theta_s \, \bar{\theta}_s}{\theta \, \bar{\theta}} \tag{2.9}$$

and it obviously belongs to Ω since then ω=o. When the roots of Φ are stationary, ν_o produces the minimum MSE estimator of s_t. Our aim is to show that (2.9) still yields the minimum MSE estimator of s_t in the nonstationary case.

3. The Result

Let the polynomials θ_s, θ_n, Φ_n and the variances V_b and V_c be fixed, assume $\Phi_s=1+\Phi_1B+...+\Phi_pB^p$, and consider the roots of the equation $z^p+\Phi_1z^{p-1}+...+\Phi_p=0$. The coefficient Φ_j (j=1,...,p) is given by $\Phi_j=(-1)^jS_j$, where S_j is the symmetric function consisting of the sum of all combinations of the products of j roots; thus Φ_j is a continuous function of the roots. Assume there are m real roots and 2n complex roots (m+2n=p). Express the complex conjugate roots in terms of the modulus, $r_i (r_i>0)$, and the frequency, $f_i (f_i \in [0,2\pi])$. Let the frequencies be fixed and construct the vector r with elements the real positive roots, the negative of the real negative roots, and the n modulii r_i of the complex roots. The vector r belongs to R_+^{m+n}. Define the vector u of the same dimension as r and with all elements equal to one. Let $F=(\Phi_1,...,\Phi_p)'$; since u is an interior point of R_+^{m+n}, F is a continuous function of r at u.

Expression (2.5) implies a system of covariance equations in the θ and V_a parameters, where all equations are polynomials. The Jacobian of the system evaluated at F(u) will be therefore nonzero (except possibly on a zero measure set on the space of the Φ_n, θ_s, θ_n, V_b and V_c parameters.) Thus the θ and V_a parameters will be (a.e.) continuous in F at F(u). It is then inmediate, from (2.7) and (2.8) that the coefficients of ω_1, ω_2 and ν are also continuous functions of F at F(u). Let v denote the vector with elements the coefficients of ν. For ν given by (2.8), $v=v[F(r),\omega]$ is a continuous function of r at u. Define the polynomials $\lambda=\omega_1(r,\omega)\theta_s$ and $\delta=-\omega_2(r,\omega)\theta_n$, with ω_1 and ω_2 given by (2.7). From (2.2) and (2.3), for $\nu \in \Omega$, the MSE of $\hat{s}_t$ can be expressed as

$$\text{MSE } \{F(r),\ v[F(r),\ \omega]\} = E[\lambda\ b_t + \delta\ c_t]^2 =$$

$$= V_b\ \Sigma\ \lambda_j^2 + V_c\ \Sigma\ \delta_j^2\ ,$$

and hence MSE is a continuous function of r at u.

In our notation, v_o given by (2.9), for the nonstationary case, can be written as $V[F(u),o]$. Its minimum MSE property is established in the following theorem.

Theorem: $MSE\{F(u),v[F(u),o]\} \leq MSE\{F(u),v[F(u),\omega]\}$ for any $\omega \neq 0$.

Proof: Assume the theorem is false. Then $\exists\ \omega^*$, $\omega^* \in W$ and $\omega^* \neq 0$, such that

$$MSE\{F(u),v[F(u),o]\} > MSE\{F(u),v[F(u),\omega^*]\} \tag{3.1}$$

Then, by continuity, for a vector $\varepsilon > 0$, of the same dimension as r and with small enough positive elements, (3.1) implies

$$MSE\{F(u-\varepsilon),v[F(u-\varepsilon),o]\} > MSE\{F(u-\varepsilon),v[F(u-\varepsilon),\omega^*]\} \tag{3.2}$$

Since for $F(u-\varepsilon)$ we are in the starionary case, $v[F(u-\varepsilon),0]$ minimizes the MSE of $\hat{s}_t$, and hence (3.2) cannot be true, proving the Theorem.

Notice that the Theorem assumes that all roots of Φ_s are nonstationary. The case in which only some of them are nonstationary can be proved in an identical way simply by fixing the stationary roots and removing them from the vector r. The proof is thus general in that it can handdle any number and type of unit roots (not simply those implied by differencing), in both Φ_s and Φ_n.

REFERENCES

Bell, W.R. (1984), "Signal Extraction for Nonstationary Time Series", The Annals of Statistics, 12, 2, 646–664.

Burman, J.P. (1980), "Seasonal Adjustment by Signal Extraction", Journal of the Royal Statistical Society, A. 143, 321–337.

Cleveland, W.P. and Tiao, G.C. (1976), "Decomposition of Seasonal Time Series: A Model for the X–11 Program", Journal of the American Statistical Association, 71, 581–587.

Hillmer, S.C. and G.C. Tiao (1982), "An ARIMA–Model Based Approach to Seasonal Adjustment", Journal of the American Statistical Association, 77, 63–70.

Pierce, D. A. (1979), "Signal Extraction Error in Nonstationary Time Series", Annals of Statistics, 7, 1303–20.

Whittle, P. (1963), Prediction and Regulation by Linear Least–Square Methods, London: English Universities Press.

DOCUMENTOS DE TRABAJO (1):

8501 **Agustín Maravall**: Predicción con modelos de series temporales.

8502 **Agustín Maravall**: On structural time series models and the characterization of components.

8503 **Ignacio Mauleón**: Predicción multivariante de los tipos interbancarios.

8504 **José Viñals**: El déficit público y sus efectos macroeconómicos: algunas reconsideraciones.

8505 **José Luis Malo de Molina y Eloísa Ortega**: Estructuras de ponderación y de precios relativos entre los deflactores de la Contabilidad Nacional.

8506 **José Viñals**: Gasto público, estructura impositiva y actividad macroeconómica en una economía abierta.

8507 **Ignacio Mauleón**: Una función de exportaciones para la economía española.

8508 **J. J. Dolado, J. L. Malo de Molina y A. Zabalza**: Spanish industrial unemployment: some explanatory factors *(versión inglés)*. El desempleo en el sector industrial español: algunos factores explicativos *(versión español)*.

8509 **Ignacio Mauleón**: Stability testing in regression models.

8510 **Ascensión Molina y Ricardo Sanz**: Un indicador mensual del consumo de energía eléctrica para usos industriales, 1976-1984.

8511 **J. J. Dolado y J. L. Malo de Molina**: An expectational model of labour demand in Spanish industry.

8512 **J. Albarracín y A. Yago**: Agregación de la Encuesta Industrial en los 15 sectores de la Contabilidad Nacional de 1970.

8513 **Juan J. Dolado, José Luis Malo de Molina y Eloísa Ortega**: Respuestas en el deflactor del valor añadido en la industria ante variaciones en los costes laborales unitarios.

8514 **Ricardo Sanz**: Trimestralización del PIB por ramas de actividad, 1964-1984.

8515 **Ignacio Mauleón**: La inversión en bienes de equipo: determinantes y estabilidad.

8516 **A. Espasa y R. Galián**: Parsimony and omitted factors: The airline model and the census X-11 assumptions *(versión inglés)*. Parquedad en la parametrización y omisiones de factores: el modelo de las líneas aéreas y las hipótesis del census X-11 *(versión español)*.

8517 **Ignacio Mauleón**: A stability test for simultaneous equation models.

8518 **José Viñals**: ¿Aumenta la apertura financiera exterior las fluctuaciones del tipo de cambio? *(versión español)*. Does financial openness increase exchange rate fluctuations? *(versión inglés)*.

8519 **José Viñals**: Deuda exterior y objetivos de balanza de pagos en España: Un análisis de largo plazo.

8520 **José Marín Arcas**: Algunos índices de progresividad de la imposición estatal sobre la renta en España y otros países de la OCDE.

8520 **José Marín Arcas**: Algunos índices de progresividad de la imposición estatal sobre la renta en España y otros países de la OCDE.

8601 **Agustín Maravall**: Revisions in ARIMA signal extraction.

8602 **Agustín Maravall y David A. Pierce**: A prototypical seasonal adjustment model.

8603 **Agustín Maravall**: On minimum mean squared error estimation of the noise in unobserved component models.

8604 **Ignacio Mauleón**: Testing the rational expectations model.

8605 **Ricardo Sanz**: Efectos de variaciones en los precios energéticos sobre los precios sectoriales y de la demanda final de nuestra economía.

8607 **José Viñals**: La política fiscal y la restricción exterior. (Publicada una edición en inglés con el mismo número).

8608 **José Viñals y John Cuddington**: Fiscal policy and the current account: what do capital controls do?

8609 **Gonzalo Gil**: Política agrícola de la Comunidad Económica Europea y montantes compensatorios monetarios.

8610 **José Viñals**: ¿Hacia una menor flexibilidad de los tipos de cambio en el sistema monetario internacional?

8701 **Agustín Maravall**: The use of ARIMA models in unobserved components estimation: an application to spanish monetary control.

8702 **Agustín Maravall**: Descomposición de series temporales: especificación, estimación e inferencia (Con una aplicación a la oferta monetaria en España).
8703 **José Viñals y Lorenzo Domingo**: La peseta y el sistema monetario europeo: un modelo de tipo de cambio peseta-marco.
8704 **Gonzalo Gil**: The functions of the Bank of Spain.
8705 **Agustín Maravall**: Descomposición de series temporales, con una aplicación a la oferta monetaria en España: Comentarios y contestación.
8706 **P. L'Hotellerie y J. Viñals**: Tendencias del comercio exterior español. Apéndice estadístico.
8707 **Anindya Banerjee y Juan Dolado**: Tests of the Life Cycle-Permanent Income Hypothesis in the Presence of Random Walks: Asymptotic Theory and Small-Sample Interpretations.
8708 **Juan J. Dolado y Tim Jenkinson**: Cointegration: A survey of recent developments.
8709 **Ignacio Mauleón**: La demanda de dinero reconsiderada.
8801 **Agustín Maravall**: Two papers on arima signal extraction.

(1) Los Documentos de Trabajo anteriores a 1985 figuran en el catálogo de publicaciones del Banco de España.

Información: Banco de España
Publicaciones. Negociado de Distribución y Gestión
Teléfono: 446 90 55, ext. 2180
Alcalá, 50. 28014 Madrid