Identifying a Simplifying Structure in Time Series

Daniel Pena, George E. P. Box

Journal of the American Statistical Association, Volume 82, Issue 399 (Sep., 1987), 836-843.

Your use of the JSTOR database indicates your acceptance of JSTOR's Terms and Conditions of Use. A copy of JSTOR's Terms and Conditions of Use is available at http://www.jstor.org/about/terms.html, by contacting JSTOR at jstor-info@umich.edu, or by calling JSTOR at (888)388-3574, (734)998-9101 or (FAX) (734)998-9113. No part of a JSTOR transmission may be copied, downloaded, stored, further transmitted, transferred, distributed, altered, or otherwise used, in any form or by any means, except: (1) one stored electronic and one paper copy of any article solely for your personal, non-commercial use, or (2) with prior written permission of JSTOR and the publisher of the article or other text.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Journal of the American Statistical Association is published by American Statistical Association. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/asa.html.

Journal of the American Statistical Association
©1987 American Statistical Association

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2001 JSTOR

http://www.jstor.org/
Sat Apr 28 17:06:25 2001

Identifying a Simplifying Structure in Time Series

DANIEL PEÑA and GEORGE E. P. BOX*

This article studies how to identify hidden factors in multivariate time series process. This problem is important because, when the series are driven by a set of common factors, (a) a large number of parameters may be needed to obtain an adequate representation of the system and (b) the estimated parameters will be highly correlated. Therefore, a complex and badly defined relationship can appear when, in fact, a simpler and parsimonious model in terms of a few common factors can be operating. This article develops a methodology to identify the number of factors and to build a simplifying transformation to represent the series. It is proved that the number of factors is equal to the rank of the covariance matrices and the parameter matrices of the infinite moving average representation of the process. The eigenvectors of these matrices will provide the canonical transformation. The method is illustrated with one example, using series of the price of wheat in five provinces of Spain in the 19th century. The standard approach to build a vector autoregressive integrated moving average model showed a complex relationship with all kinds of feedback operating. When the methodology developed in the article was applied, however, two factors were identified and a clearer and simpler representation of the system was achieved.

KEY WORDS: Multiple time series; Canonical transformation; Dynamic factor analysis.

1. INTRODUCTION

A central problem in the study of multivariate observations is the reduction of dimensionality. In the time-domain study of vector time series, Quenouille (1957) suggested how to use the eigenvectors of the parameter matrices of an autoregressive process to obtain a simpler interpretation of the system, Box and Tiao (1977) showed how to build a canonical transformation of an autoregressive vector process that orders the components from least to most predictable, and Reinsel (1983) suggested the use of index variables to simplify the representation of a vector autoregressive time series, using a formulation presented by Sargent and Sims (1977).

Brillinger (1981) extended the standard principal-component approach to the frequency domain. Related work was done by Priestley (1981) and Rao (1975). Geweke (1977) and Geweke and Singleton (1981) discussed a frequency-domain version of factor analysis. Finally, Hannan (1981) presented methods for estimating the dimension of a linear system.

In this article a method to identify common factors in a vector of time series is proposed. It is assumed that an observed k vector of time series $\mathbf{z}_t$ can be written as

$$\mathbf{z}_t = \mathbf{P}\mathbf{y}_t + \boldsymbol{\epsilon}_t, \tag{1.1}$$

where $\mathbf{P}$ is a $k \times r$ matrix of unknown parameters, $\mathbf{y}_t$ is an unobservable r-dimensional vector process, and $\boldsymbol{\epsilon}_t$ is a k-dimensional white-noise sequence with full-rank covariance matrix $\boldsymbol{\Sigma}_\varepsilon$. This model is, of course, relevant only if $r < k$ or if $r = k$ and $\boldsymbol{\Sigma}_\varepsilon = 0$. In the first case, a reduction of dimensionality can be achieved without loss of information; in the second case, $r = k$, an interesting problem is to find a linear transformation of the series that allows a simpler representation of the system.

In Section 2 the general model (1.1) is analyzed, some of its properties are studied, and a canonical transformation is introduced for structural simplification. In Section 3 a model with dependent factors is presented, and in Section 4 an example of the methodology is given.

2. THE GENERAL UNCOUPLED-FACTORS MODEL

2.1 Formulation

Let Z_t be a vector of k time series, and let $z_t = Z_t - \mu_z$ be the vector of deviations from some origin μ_z that will be the mean if the series are stationary. We assume that these series are generated by r ($r \leq k$) factors, $\mathbf{y}_t$, plus a measurement error $\boldsymbol{\epsilon}_t$ as

$$\mathbf{z}_t = \mathbf{P}\mathbf{y}_t + \boldsymbol{\epsilon}_t, \tag{2.1}$$

where $\mathbf{P}$ is a $k \times r$ matrix of parameters of rank r and $\boldsymbol{\epsilon}_t$ is a white-noise sequence with full-rank covariance matrix $\boldsymbol{\Sigma}_\varepsilon$. The vector $\mathbf{y}_t$ follows an r-dimensional (autoregressive moving average) ARMA(p_y, q_y) process of the form

$$\boldsymbol{\phi}_y(B)\mathbf{y}_t = \boldsymbol{\theta}_y(B)\mathbf{a}_t, \tag{2.2}$$

where

$$\boldsymbol{\phi}_y(B) = \mathbf{I} - \boldsymbol{\phi}_y(1)B - \cdots - \boldsymbol{\phi}_y(p)B^p$$

and

$$\boldsymbol{\theta}_y(B) = \mathbf{I} - \boldsymbol{\theta}_y(1)B - \cdots - \boldsymbol{\theta}_y(q)B^q$$

are matrix polynomials in the backshift operator B, the $\boldsymbol{\phi}$'s and the $\boldsymbol{\theta}$'s are $r \times r$ matrices, the roots of the determinantal polynomial $|\boldsymbol{\phi}_y(B)|$ are all outside the unit circle, and those of $|\boldsymbol{\theta}_y(B)|$ are all outside the unit circle. In addition, $\{\mathbf{a}_t\}$ is a sequence of vector Gaussian white noise with zero mean and covariance matrix $\boldsymbol{\Sigma}_a$. We shall assume in this section that the r factors are independent and that all of the $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ matrices will be diagonal. We will allow contemporaneous dependency in the noise matrix $\boldsymbol{\Sigma}_a$, however, and assume only that it is positive definite. In Section 3 this model will be generalized to the case in which the factors have dynamic dependence and the parameter matrices $\boldsymbol{\phi}_y$ and $\boldsymbol{\theta}_y$ are no longer diagonal.

The matrix $\mathbf{P}$ will be called the factor-loadings or factor-weights matrix, and its elements, p_{ij}, represent the weight

* Daniel Peña is Professor, Department of Statistics, Universidad Politécnica de Madrid, 28006 Madrid, Spain. George E. P. Box is Vilas Research Professor, Department of Statistics, University of Wisconsin, Madison, WI 53705. This research was carried out while Peña was Visiting Professor at the Mathematics Research Center, University of Wisconsin, and was sponsored by U.S. Army Contract DAAG29-80-C-0041. Professor Peña acknowledges support from the U.S.-Spanish Joint Committee for Educational and Cultural Affairs.

© 1987 American Statistical Association

of the factor j in the observed ith component. If $\mathbf{C}$ is any $r \times r$ nonsingular matrix, the generating equation could also be written as

$$\mathbf{z}_t = \mathbf{P}^*\mathbf{y}_t^* + \boldsymbol{\epsilon}_t,$$

where $\mathbf{P}^* = \mathbf{P}\mathbf{C}^{-1}$ is the new rectangular matrix of coefficients and $\mathbf{y}_t^* = \mathbf{C}\mathbf{y}_t$ is a linear transformation of the factors. Multiplying (2.2) by $\mathbf{C}$ gives

$$\mathbf{C}\boldsymbol{\phi}(B)\mathbf{C}^{-1}\mathbf{C}\mathbf{y}_t = \mathbf{C}\boldsymbol{\theta}(B)\mathbf{C}^{-1}\mathbf{C}\mathbf{a}_t,$$

$$\boldsymbol{\phi}^*(B)\mathbf{y}_t^* = \boldsymbol{\theta}^*(B)\mathbf{a}_t^*,$$

so the model for the new set of factors is again an r-dimensional ARMA(p_y, q_y) model with parameters

$$\boldsymbol{\phi}^*(B) = \mathbf{C}\boldsymbol{\phi}(B)\mathbf{C}^{-1},$$

$$\boldsymbol{\theta}^*(B) = \mathbf{C}\boldsymbol{\theta}(B)\mathbf{C}^{-1},$$

$$\boldsymbol{\Sigma}_a^* = \mathbf{C}\boldsymbol{\Sigma}_a\mathbf{C}'.$$

To remove this source of indeterminacy we shall assume that $\mathbf{P}'\mathbf{P} = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

2.2 The Covariance and Correlation Matrices of the Observed Process

Let us call $\boldsymbol{\Gamma}_z(k) = E[\mathbf{z}_{t-k}\mathbf{z}_t']$ the covariance matrices of the process $\mathbf{z}_t$ and $\boldsymbol{\Gamma}_y(k)$ the covariance matrices for the generating vector $\mathbf{y}_t$. Then

$$\boldsymbol{\Gamma}_z(0) = \mathbf{P}\boldsymbol{\Gamma}_y(0)\mathbf{P}' + \boldsymbol{\Sigma}_\epsilon, \tag{2.3}$$

$$\boldsymbol{\Gamma}_z(k) = \mathbf{P}\boldsymbol{\Gamma}_y(k)\mathbf{P}', \qquad k \geq 1, \tag{2.4}$$

and the rank of $\boldsymbol{\Gamma}_z(k)$ for $k \geq 1$ will be equal to the number r of common factors. If the factors are independent for all lags and the matrix $\boldsymbol{\Sigma}_a$ is diagonal, then all of the covariance matrices $\boldsymbol{\Gamma}_y(k)$ will be diagonal and hence (a) the matrices $\boldsymbol{\Gamma}_z(k)$ will be symmetric for $k \geq 1$ and (b) the columns of $\mathbf{P}$ will be eigenvectors of $\boldsymbol{\Gamma}_z(k)$ with eigenvalues $\gamma_i(k)$, where $\gamma_i(k)$ are the diagonal elements of $\boldsymbol{\Gamma}_y(k)$.

The partial autocorrelation matrices will be given by

$$\begin{aligned}\mathcal{P}'(l) &= \boldsymbol{\Gamma}_z^{-1}(0)\boldsymbol{\Gamma}_z(1), \qquad l = 1,\\ &= [\boldsymbol{\Gamma}_z(0) - \mathbf{b}'(l)\mathbf{A}^{-1}(l)\mathbf{b}(l)]^{-1},\\ &= [\boldsymbol{\Gamma}_z(l) - b'(l)A^{-1}(l)c(l)], \qquad l > 1,\end{aligned}$$

where

$$\mathbf{b}'(l) = [\boldsymbol{\Gamma}_z'(l-1)\boldsymbol{\Gamma}_z'(l-2)\cdots\boldsymbol{\Gamma}_z'(1)],$$

$$\mathbf{c}'(l) = [\boldsymbol{\Gamma}_z'(1)\boldsymbol{\Gamma}_z'(2)\cdots\boldsymbol{\Gamma}_z'(l-1)],$$

and

$$\mathbf{A}(l) = \begin{bmatrix} \boldsymbol{\Gamma}_z(0) & \boldsymbol{\Gamma}_z'(1)\cdots & \boldsymbol{\Gamma}_z'(l-2) \\ \boldsymbol{\Gamma}_z(1) & \boldsymbol{\Gamma}_z(0) & \\ & \ddots & \\ & & \ddots \\ & & \ddots \\ \boldsymbol{\Gamma}_z(l-2) & & \boldsymbol{\Gamma}_z(0) \end{bmatrix}$$

(see Tiao and Box 1981), so the partial autocorrelation matrices can be written as

$$\mathcal{P}'(l) = \mathbf{B}(l)\left[\boldsymbol{\Gamma}_z(l) - \sum_{i=1}^{l-1}\sum_{j=1}^{l-1}\boldsymbol{\Gamma}_z'(l-i)\mathbf{A}^{-1}(i,j,l)\boldsymbol{\Gamma}_z(j)\right],$$

where $\mathbf{B}(l) = [\boldsymbol{\Gamma}_z(0) - \mathbf{b}'(l)\mathbf{A}^{-1}(l)\mathbf{b}(l)]^{-1}$ and $\mathbf{A}^{-1}(i, j, l)$ is the (i, j) element of $\mathbf{A}^{-1}(l)$. Then

$$\mathcal{P}'(l) = \mathbf{B}(l)\mathbf{P}\left[\boldsymbol{\Gamma}_y(l) - \sum_{i=1}^{l-1}\sum_{j=1}^{l-1}\boldsymbol{\Gamma}_y'(l-1)\mathbf{P}'\mathbf{A}^{-1}(i,j,l)\mathbf{P}\boldsymbol{\Gamma}_y(j)\right]\mathbf{P}'. \tag{2.5}$$

Now, if we call v the null space of $\mathbf{P}$ defined by all of the vectors that verify $\mathbf{P}'v = 0$, and if $\mathbf{V}$ is a $k \times (k - r)$ matrix the columns of which are a basis for this space v, then

$$\mathcal{P}'(l)\mathbf{V} = \mathbf{H} \cdot \mathbf{P}'\mathbf{V} = 0. \tag{2.6}$$

This shows that the rank of $\mathcal{P}'(l)$ is at most r and that the eigenvectors linked to zero eigenvalues of $\mathcal{P}'(l)$ belong to the v space.

2.3 The Model for the Observed Process and Its Parameters

The properties that we have studied for the covariance and correlation matrices impose strong restriction on the type of model that the observed $\mathbf{z}_t$ vector of series can follow and on the characteristics of its parameter values.

Theorem 1. Suppose $\mathbf{z}_t = \mathbf{P}\mathbf{y}_t + \boldsymbol{\epsilon}_t$, where $\mathbf{y}_t$ is an r-dimensional ARMA(p_y, q_y) process, $\mathbf{P}$ is a $k \times r$ matrix ($k \geq r$) of rank r, and $\boldsymbol{\epsilon}_t$ is a k-dimensional white-noise sequence with covariance matrix $\boldsymbol{\Sigma}_\epsilon$. Then $\mathbf{z}$ follows a k-dimensional ARMA(p_z, q_z) with $p_z = p_y$ and $q_z = \max(p_y, q_y)$.

The representation obtained for $\mathbf{z}_t$ is not unique, as shown in the following theorem.

Theorem 2. If $\mathbf{z}_t$ is an observed-vector time series generated by r common factors as in (2.1) and follows model (2.2), then $\mathbf{z}_t$ could be equally represented as $\boldsymbol{\phi}^*(B)\mathbf{z}_t = \boldsymbol{\theta}^*(B)\mathbf{b}_t$ with new parameters $\boldsymbol{\phi}_l^* = \boldsymbol{\phi}_l + \mathbf{A}$ and $\boldsymbol{\theta}_l^* = \boldsymbol{\theta}_l + \mathbf{A}$, where $\mathbf{A}$ is a nonnull matrix of rank $= k - r$.

The proofs of these theorems are in the Appendix.

An important conclusion from Theorem 1 is that the autoregressive matrices of the multivariate observed process $\mathbf{z}_t$ must satisfy

$$\boldsymbol{\phi}_z(k)\mathbf{P} = \mathbf{P}\boldsymbol{\phi}_y(k), \tag{2.7}$$

which has the general solution

$$\boldsymbol{\phi}_z(k) = \mathbf{P}\boldsymbol{\phi}_y(k)\mathbf{P}^- + \mathbf{A}(\mathbf{I} - \mathbf{P}\mathbf{P}^-), \tag{2.8}$$

where $\mathbf{P}^-$ is any generalized inverse of $\mathbf{P}$ that satisfies

$\mathbf{PP}^{-}\mathbf{P} = \mathbf{P}$ and $\mathbf{A}$ is any arbitrary matrix, with the only restriction being that the roots of $|\boldsymbol{\phi}_z(B)|$ are on or outside the unit circle. As the matrix $\boldsymbol{\phi}_y(k)$ is diagonal, Equation (2.7) shows that the columns of $\mathbf{P}$ are eigenvectors of $\boldsymbol{\phi}_z(k)$, with eigenvalues as the diagonal elements of $\boldsymbol{\phi}_y(k)$. The matrix $\boldsymbol{\phi}_z(k)$ can have any rank, however, because of the presence of the arbitrary matrix $\mathbf{A}$. On the other hand, according to Theorem 2, for every possible solution $\boldsymbol{\phi}_z(k)$ there will be a set of $\boldsymbol{\theta}_z(k)$ matrices that must verify some restrictions [see (A.6) in the Appendix]. For example, if $p_y = 1$, $q_y = 0$, and the vector of components follows a multivariate (autoregressive) AR(1) process, the system of equations to determine the parameters is

$$\boldsymbol{\Sigma}_\varepsilon + \mathbf{P}\boldsymbol{\Sigma}_a\mathbf{P}' = \boldsymbol{\Sigma}_u - \boldsymbol{\phi}_z\boldsymbol{\Sigma}_u\boldsymbol{\theta}_z' + \boldsymbol{\theta}_z\boldsymbol{\Sigma}_u\boldsymbol{\theta}_z', \quad (2.9)$$

$$\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\phi}_z' = \boldsymbol{\Sigma}_u\boldsymbol{\theta}_z', \quad (2.10)$$

and

$$\boldsymbol{\phi}_z\mathbf{P} = \mathbf{P}\boldsymbol{\phi}_y,$$

where $\boldsymbol{\Sigma}_u$ is the noise covariance matrix and $\boldsymbol{\phi}_z$, $\boldsymbol{\theta}_z$ are the parameter matrices in the model for $\mathbf{Z}_t$. These matrices can be characterized as follows: $\boldsymbol{\phi}_z$ is given by (2.8); then we have $k(k+1)/2$ equations in (2.9) and k^2 in (2.10) that are equal to the number of unknown parameters. Furthermore, by using result (A.6),

$$\mathbf{V}'\boldsymbol{\psi}_1 = \mathbf{V}_1'(\boldsymbol{\phi}_z - \boldsymbol{\theta}_z) = \mathbf{0}, \quad (2.11)$$

which implies that, although $\boldsymbol{\phi}_z$ and $\boldsymbol{\theta}_z$ can have any rank, their difference must have a rank no greater than the number of components in the system. Moreover, suppose that $\mathbf{h}$ is an eigenvector of the matrix $\boldsymbol{\phi}_z - \boldsymbol{\theta}_z$ and that $(\boldsymbol{\phi}_z - \boldsymbol{\theta}_z)\mathbf{h} = \lambda\mathbf{h}$; then assume that λ is nonzero. If $\mathbf{v}_i'$ is any row of $\mathbf{V}$ as

$$\mathbf{v}_i'(\boldsymbol{\phi}_z - \boldsymbol{\theta}_z)\mathbf{h} = \lambda\mathbf{v}_i'\mathbf{h} = 0,$$

then $\mathbf{h}$ must be orthogonal to the subspace generated by the matrix $\mathbf{V}$ and, therefore, it belongs to the subspace generated by $\mathbf{P}$.

In summary, when building a multivariate time series model for a vector of series $\mathbf{z}_t$ that is driven by a number of common factors, as in (2.1), it may appear that a complicated model with all kinds of dependency is needed. In particular we may finish with a model showing feedback among all of the components of the vector $\mathbf{z}_t$. But the parameter matrices in this case must satisfy many restrictions. These may be summarized as follows: (a) All of the autoregressive matrices will have r common eigenvectors that define the subspace $\mathbf{P}$ of the factors. (b) All of the $\boldsymbol{\psi}$ matrices must have rank equal to r, and their columns must belong to the subspace generated by $\mathbf{P}$.

2.4 A Canonical Transformation

An important problem is to find a transformation of the series $\mathbf{z}_t$ that allows us to simplify the structure and to gain understanding. Our interest here is in simplifying the underlying structure of the observed series. Efficient estimation can be achieved, for example, by formulating the model in state space and estimating by maximum likelihood as suggested by Engle and Watson (1981). To explain how to find such a transformation, first, note that when $\mathbf{P}^-$ is any generalized inverse of $\mathbf{P}$, if $\mathbf{z}_t = \mathbf{P}\mathbf{y}_t + \boldsymbol{\epsilon}_t$ then $\mathbf{y}_t = \mathbf{P}^-\mathbf{z}_t - \mathbf{P}^-\boldsymbol{\epsilon}_t$. Second, suppose that we apply a linear transformation $\mathbf{M}$ to $\mathbf{z}_t$, where $\mathbf{M} = \begin{bmatrix} P^- \\ B \end{bmatrix}$. Then

$$\mathbf{x}_t = \mathbf{M}\mathbf{z}_t = \begin{bmatrix} \mathbf{P}^-\mathbf{z}_t \\ \mathbf{B}\mathbf{z}_t \end{bmatrix} = \begin{bmatrix} \mathbf{y} + \mathbf{P}^-\boldsymbol{\epsilon}_t \\ \mathbf{B}\mathbf{P}\mathbf{y}_t + \mathbf{B}\boldsymbol{\epsilon}_t \end{bmatrix} = \begin{bmatrix} \mathbf{x}_{1t} \\ \mathbf{x}_{2t} \end{bmatrix}.$$

To isolate the effect of the factors, $\mathbf{B}$ should be chosen such that $\mathbf{BP} = 0$. Then the first r components of x_t will be equal to the fector $\mathbf{y}_t$ of common factors plus some added noise, and the second $k - r$ components will be just white noise. A matrix $\mathbf{B}$ that satisfies this condition may be obtained by putting the $k - r$ eigenvectors linked to zero eigenvalues of $\mathbf{PP}'$ as rows of $\mathbf{B}$. Calling v_i the eigenvectors

$$\mathbf{PP}'[v_1, \ldots, v_{k-r}] = \mathbf{0} = \mathbf{PP}'\mathbf{V}$$

implies that $\mathbf{V}'\mathbf{P} = \mathbf{0}$, so $\mathbf{B} = \mathbf{V}'$ has the desired property. Although the selection of $\mathbf{P}^-$ could be arbitrary, it seems sensible to choose a generalized inverse that leads to canonical variables that are as simple as possible. A convenient property is that the $\mathbf{x}_1$ and $\mathbf{x}_2$ components of the new vector are independent, but in general this is not possible unless $\boldsymbol{\Sigma}_\varepsilon = \mathbf{I}$. If we assume this equality, then

$$E[\mathbf{x}_{1t}\mathbf{x}_{2t}'] = \mathbf{P}^-\boldsymbol{\Sigma}_\varepsilon\mathbf{B}' = \mathbf{P}^-\mathbf{B}';$$

and if we choose $\mathbf{P}^-$ as the Moore–Penroe generalized inverse of $\mathbf{P}$, given by $(\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'$, the covariance matrix will be 0 and the components will be independent.

To illustrate this transformation, let us suppose that $\mathbf{y}_t$ follows an r-dimensional AR(1) process. Then $\mathbf{z}_t$ is ARMA(1, 1), and so is any linear transformation of $\mathbf{z}_t$. But the parameters of the k-dimensional $\mathbf{x}_t$ process given by $\mathbf{x}_t = \mathbf{M}\mathbf{z}_t$, where

$$\mathbf{M} = \begin{bmatrix} (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}' \\ \mathbf{V}' \end{bmatrix},$$

will be

$$\boldsymbol{\phi}_x = \mathbf{M}\boldsymbol{\phi}_z\mathbf{M}^{-1}, \qquad \boldsymbol{\theta}_x = \mathbf{M}\boldsymbol{\theta}_z\mathbf{M}^{-1},$$

where $\mathbf{M}^{-1} = [\mathbf{P} \mid \mathbf{V}]$, as can be verified by direct multiplication, $\mathbf{M}\mathbf{M}^{-1} = \mathbf{M}^{-1}\mathbf{M} = \mathbf{I}$. So

$$\boldsymbol{\phi}_x = \begin{vmatrix} (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\boldsymbol{\phi}_z\mathbf{P} & (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\boldsymbol{\phi}_z\mathbf{V} \\ \mathbf{V}'\boldsymbol{\phi}_z\mathbf{P} & \mathbf{V}'\boldsymbol{\phi}_z\mathbf{V} \end{vmatrix},$$

and, by using the general expression (2.8) for $\boldsymbol{\phi}_z$,

$$\boldsymbol{\phi}_x = \begin{vmatrix} \boldsymbol{\phi}_y & \mathbf{0} \\ \mathbf{0} & \mathbf{V}'\mathbf{A}\mathbf{V} \end{vmatrix}$$

in the same way, calling $\boldsymbol{\theta}_0$ the value of $\boldsymbol{\theta}_z$ associated with $\boldsymbol{\phi}_z = \mathbf{P}\boldsymbol{\phi}_y\mathbf{P}^-$ when $\mathbf{A} = \mathbf{0}$. Then for Theorem 2 the general expression of $\boldsymbol{\theta}_z$ is

$$\boldsymbol{\theta}_z = \boldsymbol{\theta}_0 + \mathbf{A}(\mathbf{I} - \mathbf{PP}^-),$$

so

$$\boldsymbol{\theta}_x = \begin{vmatrix} (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\boldsymbol{\theta}_0\mathbf{P} & (\mathbf{P}'\mathbf{P})^{-1}\mathbf{P}'\boldsymbol{\theta}_z\mathbf{V} \\ \mathbf{0} & \mathbf{V}'\mathbf{A}\mathbf{V} \end{vmatrix},$$

where we have used $\mathbf{V}'\boldsymbol{\phi}_z = \mathbf{V}'\boldsymbol{\theta}_z$ for (2.11). Therefore, by calling $\mathbf{x}_t' = [\mathbf{x}_{1t}'; \mathbf{x}_{2t}']$ and by taking into account that the addition of moving average (MA) processes is a new MA process, we have decomposed the $\mathbf{z}_t$ series into two components—(a) an ARMA(1, 1) process of dimension r, x_{1t}, with the factors-AR parameter matrix, which will contain the factors mixed up with noise, and (b) a white-noise process, x_{2t}, of dimension $k - r$.

3. THE FACTOR MODEL WITH DEPENDENT FACTORS

Table 1 summarizes the differences between the basic properties of the uncoupled-factors model and the coupled (or dependent)-factors model. The main difference is related to the eigenvectors of the covariance and autoregressive parameter matrices that are not columns of the factor-loading matrix in the second case. Equations (2.4) and (2.7) hold in this case, but the conclusions that we have to draw from them are different. For instance, starting with (2.4),

$$\boldsymbol{\Gamma}_z(k) = \mathbf{P}\boldsymbol{\Gamma}_y(k)\mathbf{P}', \qquad k \geqslant 1.$$

Assuming that $\boldsymbol{\Gamma}_y(k) = \mathbf{U}_k\mathbf{D}_k\mathbf{U}_k^{-1}$, where $\mathbf{D}_k$ is diagonal, it is clear that $\boldsymbol{\Gamma}_z(k)$ has the same eigenvalues as $\boldsymbol{\Gamma}_y(k)$ but with eigenvectors $\mathbf{PU}_k$ and $\mathbf{V}$, where the columns of $\mathbf{V}$ belong to the null space of $\mathbf{PP}'$. To see this, write

$$\boldsymbol{\Gamma}_z(k) = [\mathbf{PU}_k : \mathbf{V}] \begin{vmatrix} \mathbf{D}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{vmatrix} \begin{vmatrix} \mathbf{U}_k^{-1}\mathbf{P}' \\ \mathbf{V}' \end{vmatrix},$$

which shows the eigenstructure of $\boldsymbol{\Gamma}_z(k)$ clearly. Of course, if $\boldsymbol{\Gamma}_y(k)$ is diagonal, then $\mathbf{U}_k = \mathbf{I}$ and we obtain the uncoupled case.

As for the autoregressive parameter matrices, the condition is

$$\boldsymbol{\phi}_z(k)\mathbf{P} = \mathbf{P}\boldsymbol{\phi}_y(k),$$

assuming that $\boldsymbol{\phi}_y(k)$ has linearly independent eigenvectors and writing $\boldsymbol{\phi}_y(k) = \mathbf{W}_k\mathbf{F}_k\mathbf{W}_k^{-1}$, where $\mathbf{F}_k$ is diagonal and contains the eigenvalues of $\boldsymbol{\phi}_y(k)$. Then

$$\boldsymbol{\phi}_z(k)\mathbf{PW}_k = \mathbf{PW}_k\mathbf{F}_k$$

shows that the eigenvalues of $\boldsymbol{\phi}_y(k)$ are eigenvalues of $\boldsymbol{\phi}_z(k)$ as well, with eigenvectors $\mathbf{PW}_k$.

To build the canonical transformation, we need to determine the null space of $\mathbf{PP}'$. Suppose that $\mathbf{H}$ is the matrix of eigenvectors of the first autoregressive parameter matrix of the observed series $\mathbf{z}_t$. Then r columns of $\mathbf{H}$ are $\mathbf{PW}_1$, whereas the other $k - r$ columns are arbitrary. As

$$(\mathbf{PW}_1)(\mathbf{W}_1^{-1}\mathbf{P}') = \mathbf{PP}',$$

the null space of $(\mathbf{PW}_1)(\mathbf{W}_1^{-1}\mathbf{P}')$ is the same as that of $\mathbf{PP}'$.

To select which columns of $\mathbf{H}$ are the set $\mathbf{PW}_1$, we must know that (a) their number, r, is known because it is the rank of $\boldsymbol{\Gamma}_z(k)$, $\mathcal{P}(1)$, and all of the $\boldsymbol{\psi}_i$'s and (b) they will appear approximately in $\mathcal{P}(1)$ as the eigenvectors corresponding to their nonzero eigenvalues. Then, given $(\mathbf{PW}_1)$ from $\mathbf{H}$, we can obtain $\mathbf{W}^{-1}\mathbf{P}'$ from $\mathbf{H}^{-1}$. In addition, by calling $\mathbf{V}$ the matrix of eigenvectors corresponding to nonzero eigenvalues of $(\mathbf{PW}_1)(\mathbf{W}_1^{-1}\mathbf{P}')$, the transformation

$$\mathbf{M} = \begin{vmatrix} \mathbf{W}_1^{-1}\mathbf{P}' \\ \mathbf{V}' \end{vmatrix}$$

will produce, as the uncoupled case, a decomposition of the $\mathbf{z}_t$ vector into two components: The first will contain the factors mixed up with noise, and the second component will be a vector of white noise:

$$\mathbf{x}_t = \mathbf{Mz}_t = \begin{vmatrix} \mathbf{W}_1^{-1}\mathbf{y}_t + \mathbf{W}_1^{-1}\boldsymbol{\epsilon}_t \\ \mathbf{V}'\boldsymbol{\epsilon}_t \end{vmatrix} = \begin{vmatrix} \mathbf{x}_{1t} \\ \mathbf{x}_{2t} \end{vmatrix}.$$

An interesting property of the transformation is that the components obtained in $\mathbf{x}_{1t}$ will be, in practice, weekly related. For example, if $\mathbf{y}_t$ follows an AR(1) process, $\mathbf{W}_1^{-1}\mathbf{y}_t$ will follow an AR(1) process too, but with a diagonal autoregressive parameter matrix.

4. EXAMPLE

Figure 1 shows the series of the price of wheat in five provinces of Castillia, Spain, from July 1880 to December 1890. These provinces are Avila, Palencia, Segovia, Valladolid, and Zamora. The source of the data is Sánchez-Albornoz (1975). Peña and Sánchez-Albornoz (1984) studied some of these series during the second half of the 19th century, pointing out their lack of seasonality.

The analysis that follows is made with the series in logs, although the conclusions are not depending on the particular transformation to be used.

Table 2 shows the univariate model for the five series. These models are broadly consistent with the hypothesis of some common factors following an AR(1) model. Table 3 displays the eigenvalues and eigenvectors of the covariance matrices (see Table 4) up to lag 6. There is a strong eigenvalue that decreases slowly, linked to a fairly stable eigenvector. The second eigenvalue is rather small, and

Table 1. Differences Between Properties of the Uncoupled-Factors and Coupled-Factors Models

Property	*Uncoupled factors*	*Coupled factors*
Rank $(\boldsymbol{\Gamma}_z(k)) = r$ $(k \geqslant 1)$	Yes	Yes
$\boldsymbol{\Gamma}_z(k)$ symmetric	Yes	No
Eigenvectors of $\boldsymbol{\Gamma}_z(k)$ are columns of $\mathbf{P}$	Yes	No
Eigenvalues of $\boldsymbol{\Gamma}_z(k)$ are eigenvalues of $\boldsymbol{\Gamma}_y(k)$	Yes	Yes
Rank $(\mathbf{P}(l)) = r$	Yes	Yes
$\mathbf{z}_t \sim$ ARMA$(p_z = p_y, q_z = \max(p_y, q_y))$	Yes	Yes
Rank $(\boldsymbol{\psi}_i) = r$ $(i \geqslant 1)$	Yes	Yes
Eigenvalues of $\boldsymbol{\phi}_z(k)$ are eigenvalues of $\boldsymbol{\phi}_y(k)$	Yes	Yes
Eigenvectors of $\boldsymbol{\phi}_z(k)$ are eigenvectors of $\boldsymbol{\phi}_y(k)$	Yes	No

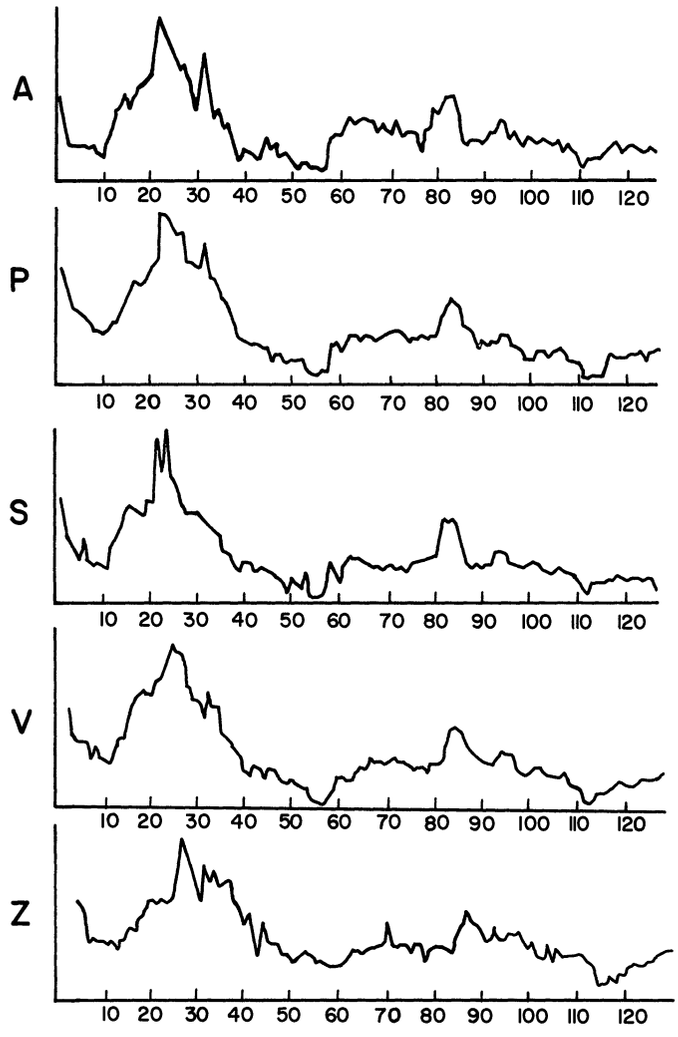

Figure 1. Wheat-Price Series (July 1880–December 1890) of Avila (A), Palencia (P), Segovia (S), Valladolid (V), and Zamora (Z).

the eigenvector is not completely stable. Finally, the other three eigenvectors are practically 0. Then it seems that there are at least one important common component and perhaps a second one, although the second is much less important than the first. The two main eigenvalues of the covariance matrices appear again linked to the largest and third eigenvalues of the first partial autocorrelation matrix. The pattern of significant coefficients in the exact maximum likelihood estimation of the parameters of an ARMA(1, 1) is

$$\begin{vmatrix} + & \cdot & - & + & \cdot \\ \cdot & + & \cdot & + & \cdot \\ \cdot & - & + & + & \cdot \\ \cdot & \cdot & \cdot & + & \cdot \\ \cdot & \cdot & \cdot & + & + \end{vmatrix}_{\boldsymbol{\phi}} \begin{vmatrix} + & - & - & \cdot & + \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & - & + & \cdot & \cdot \\ \cdot & \cdot & - & \cdot & \cdot \\ \cdot & - & - & \cdot & + \end{vmatrix}_{\boldsymbol{\theta}} . \quad (4.1)$$

It can be concluded that there is a very complex relationship among the series with all kinds of feedback present. Moreover, the five eigenvalues of $\boldsymbol{\phi}$ are close to the unit circle, which can incorrectly lead to the conclusion that there are five nonstationary factors. The rank of $\boldsymbol{\phi} - \boldsymbol{\theta}$ is nearly 2, however, leading to the conclusion that there are two factors in the system. The eigenvectors linked to these main eigenvalues are again very close to the previous eigenvectors of the covariance and partial correlation matrices, and these two eigenvectors appear approximately in the $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ matrices as well.

In summary, there is strong evidence of at least one and probably two common factors driving the system. The first

Table 3. Eigenvalues (λ) and Eigenvectors (v) of the Covariance Matrices

	Lag					
	1	2	3	4	5	6
Largest eigenvector and eigenvalue						
λ	777	735	693	647	604	560
v	.35	.35	.34	.33	.32	.32
	.44	.44	.45	.45	.45	.45
	.49	.49	.48	.47	.47	.46
	.44	.44	.44	.44	.43	.43
	.49	.50	.51	.52	.53	.55
Second-largest eigenvector and eigenvalue						
λ	12	13	11	8	5	7
v	.65	.40	.08	.21	.38	.18
	−.18	−.17	.01	−.09	.26	.23
	.28	.46	.18	.31	.05	.34
	.06	.16	.04	.16	.10	.43
	−.68	−.77	−.31	−.57	−.68	−.78
Leftover eigenvalues						
3rd	0	1	4	3	5	1
4th	5	4	4	3	4	2
5th	5	4	4	1	0	2

NOTE: All of the eigenvalues have been multiplied by 10^4.

Table 2. Univariate Model for the Series

$\nabla \ln \text{Avila} = \underset{(1.30)}{(1 + .11B)a_t}$	$\hat{\sigma}_a = .039$	$Q(23) = 32.8$
$\nabla \ln \text{Palencia} = \underset{(2.02)}{(1 + .18B)a_t}$	$\hat{\sigma}_a = .030$	$Q(23) = 24$
$\nabla \ln \text{Segovia} = \underset{(1.77)}{(1 - .16B)a_t}$	$\hat{\sigma}_a = .057$	$Q(23) = 16$
$\nabla \ln \text{Valladolid} = \underset{(1.70)}{(1 + .15B)a_t}$	$\hat{\sigma}_a = .031$	$Q(23) = 36.9$
$\underset{(1.70)}{(1 - .15B} - \underset{(2.30)}{.20B^2)} \nabla \ln \text{Valladolid} = a_t$	$\hat{\sigma}_a = .029$	$Q(23) = 34.6$
$\nabla \ln \text{Zamora} = \underset{(1.84)}{(1 - .16B)}$	$\hat{\sigma}_a = .049$	$Q(23) = 23.9$

NOTE: The values under the estimated parameters are t values. $Q(l)$ is the statistic of Ljung and Box with l df. $X^2_{23}(.05) = 35.2$.

Table 4. Eigenvalues and Eigenvectors of the First Covariance Matrix

Γ_0				
810	*20*	*10*	*10*	*10*
.36	.27	.80	.33	.24
.43	−.14	.06	−.80	.40
.50	.57	−.58	.22	.19
.44	.10	.12	−.21	−.86
.49	−.76	−.15	.41	.05

is linked to an eigenvalue that is, broadly speaking, a mean of the series. The second is roughly the ratio among Avila and Zamora. This result makes sense because this second factor affects the two more western provinces that are closer to the Portuguese border.

We have chosen the eigenvectors (.37, .49, .53, .48, .57) and (.69, 0, 0, 0, −.74) to build the transformation. Any other selection does not seem to change the results. With these two generators, the matrix $\mathbf{M}$ is

$$\mathbf{M} = \begin{vmatrix} .52 & .07 & -.51 & -.48 & .49 \\ .26 & -.88 & .09 & .31 & .24 \\ .04 & -.15 & .70 & -.70 & .03 \\ .41 & .41 & .45 & .41 & .38 \\ .74 & .07 & .07 & .07 & -.66 \end{vmatrix}.$$

The new ϕ_x and θ_x for the transformed $\mathbf{x}$ series obtained transforming the parameter matrices of $\mathbf{z}$ turn out to be

$$\phi_x = \mathbf{M}\phi\mathbf{M}^{-1} = \begin{vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -.71 & .94 & .18 \\ \cdot & \cdot & \cdot & \cdot & .92 \end{vmatrix} + \begin{vmatrix} .96 & \cdot & \cdot & \cdot & \cdot \\ \cdot & .98 & -.19 & \cdot & \cdot \\ \cdot & .16 & .84 & \cdot & \cdot \\ -.30 & .82 & -.24 & \cdot & \cdot \\ \cdot & .25 & -.27 & \cdot & \cdot \end{vmatrix}$$

and

$$\theta_x = \mathbf{M}\theta\mathbf{M}^{-1} = \begin{vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & -.55 & \cdot & \cdot \\ \cdot & \cdot & .20 & .50 & \cdot \end{vmatrix} + \begin{vmatrix} .90 & .19 & .15 & \cdot & \cdot \\ \cdot & .93 & -.23 & \cdot & \cdot \\ \cdot & .16 & 1.0 & \cdot & \cdot \\ .12 & .83 & -.12 & \cdot & \cdot \\ \cdot & .33 & -.19 & \cdot & \cdot \end{vmatrix},$$

where we have used a dot to indicate a parameter value below .1. This decomposition shows that there is a matrix A of rank equal to 3 that can be subtracted from both sides. All of the rows of this matrix are orthogonal to the vectors (00010) and (00001), so, according to Theorem 2, they could be subtracted to obtain three independent sequences of white noise plus two factors. The first factor is, according to the fourth row of M, a mean of all of the series and describes the general trend of the market. The second factor is stationary and affects the most westerly provinces, Avila and Zamora, and it seems to have an ARMA(1, 1) structure. The factors are not completely uncorrelated, and a weak relationship seems to exist between them. The canonical transformation obtained leads to a representation of the system for the transformed series $\mathbf{x}$ that needs fewer parameters and shows clearly the dynamic structure that is operating.

As a further check of the methodology, we have analyzed the wheat-price series of three provinces of different economic behavior in which a unique common factor is not expected. The provinces that we have chosen are La Coruña, which is not a wheat producer, in the North of Spain; Zaragoza, which is a producer of wheat supplied to Aragon and Cataluña; and Valladolid, again as representative of Castillia. Then the eigenvalues of the covariance matrices are fairly stable, the partial autocorrelation coefficients show three strong roots, the estimation of an ARMA(1, 1) model shows that $\phi - \theta$ has rank 3, and it is not possible to reduce the dimension of the system. There are three factors, and the three are nonstationary, as expected.

5. CONCLUSION

It has been shown that the multivariate time series model for an observed vector of time series may appear very complicated, when in fact a simple representation using a small set of common factors is adequate. Besides, failing to notice this structure can be dangerous, because the interpretation of the internal dynamic that we have to draw from the model will be completely different in both cases.

Unlike in the factor analysis of static variables, the restrictions that the presence of common factors puts on the covariance, correlation, partial autocorrelation, and parameter matrices make the identification of the number of factors and the estimation of the factor-loading matrix relatively easy.

APPENDIX: PROOFS

A.1 Proof of Theorem 1

The covariance matrices of an ARMA(p, q) model satisfy

$$\Gamma_y(l) = \sum_{j=l-m}^{l-1} \Gamma_y(j)\phi_y'(l-j) - \sum_{j=0}^{m-l} \psi_y(j)\Sigma_a\theta_y'(j+l), \qquad l = 0, \ldots, m$$

$$= \sum_{j=1}^{m} \Gamma_y(l-j)\phi_y'(j), \qquad l > m,$$

where $m = \max(p_y, q_y)$, the ψ's are obtained from the relationship $\psi(B) = \phi^{-1}(B)\theta(B)$, and $\theta_0 = -\mathbf{I}$ (see Tiao and Box 1981). It is assumed that (a) if $p < q$, then $\phi(p+1) = \cdots = \phi(m) = 0$, and (b) if $q < p$, then $\theta(q+1) = \theta(m) = 0$. Premultiplying the system of equations by $\mathbf{P}$ and postmultiplying by

$\mathbf{P}'$ give

$$\begin{aligned}\boldsymbol{\Gamma}_z(l) &= \sum_{j=-m}^{-1} \mathbf{P}\boldsymbol{\Gamma}_y(j)\boldsymbol{\phi}_y'(l-j)\mathbf{P}' \\ &\quad - \sum_{j=0}^{m} \mathbf{P}\boldsymbol{\psi}_y(j)\boldsymbol{\Sigma}_a\boldsymbol{\theta}_y'(j)\mathbf{P}' + \boldsymbol{\Sigma}_\varepsilon, \qquad l = 0 \\ &= \sum_{j=l-m}^{l-1} \mathbf{P}\boldsymbol{\Gamma}_y(j)\boldsymbol{\phi}_y'(l-j)\mathbf{P}' \\ &\quad - \sum_{j=0}^{m-l} \mathbf{P}\boldsymbol{\psi}_y(j)\boldsymbol{\Sigma}_a\boldsymbol{\theta}_y'(j+l)\mathbf{P}', \qquad l = 1, \ldots, m \\ &= \sum_{j=1}^{m} \mathbf{P}\boldsymbol{\Gamma}_y(l-j)\boldsymbol{\phi}_y'(j)\mathbf{P}', \qquad l > m. \end{aligned} \tag{A.1}$$

For $l > m$, equating (A.1) to the general expression for the covariances of an ARMA model gives

$$\sum_{j=1}^{m} \mathbf{P}\boldsymbol{\Gamma}_y(l-j)\boldsymbol{\phi}_y'(j)\mathbf{P}' = \sum_{j=1}^{p_z} \boldsymbol{\Gamma}_z(l-j)\boldsymbol{\phi}_z'(j), \qquad l = m+1, \ldots.$$

Therefore, using (2.4), $p_z = p_y$,

$$\boldsymbol{\phi}_y'(k)\mathbf{P}' = \mathbf{P}'\boldsymbol{\phi}_z'(k), \qquad k \geq 1, \tag{A.2}$$

and, substituting (A.2) into (A.1),

$$\begin{aligned}\boldsymbol{\Gamma}_z(l) &= \sum_{j=-m}^{-1} \boldsymbol{\Gamma}_z(j)\boldsymbol{\phi}_z'(l-j) \\ &\quad - \sum_{j=0}^{m} \mathbf{P}\boldsymbol{\psi}_y(j)\boldsymbol{\Sigma}_a\boldsymbol{\theta}_y'(j)\mathbf{P}' + \boldsymbol{\Sigma}_\varepsilon, \qquad l = 0 \\ &= \sum_{j=l-m}^{l-1} \boldsymbol{\Gamma}_z(j)\boldsymbol{\phi}_z'(l-j) - \boldsymbol{\Sigma}_\varepsilon\boldsymbol{\phi}_z'(l) \\ &\quad - \sum_{j=0}^{m-l} \mathbf{P}\boldsymbol{\psi}_y(j)\boldsymbol{\Sigma}_a\boldsymbol{\theta}_y'(j+l)\mathbf{P}', \qquad l = 1, \ldots, m \\ &= \sum_{j=1}^{m} \boldsymbol{\Gamma}_z(l-j)\boldsymbol{\phi}_z'(j), \qquad l > m. \end{aligned} \tag{A.3}$$

If the order of the AR component of $\mathbf{y}_t$, p_y, is 0, then there will be exactly q_y, $\boldsymbol{\Gamma}_z(l)$ covariance matrices different from 0 and the observed process $\mathbf{z}_t$ will follow an MA(q_y) process. On the other hand, if $q_y = 0$, the system of equations reduces to

$$\boldsymbol{\Gamma}_z(0) = \sum_{j=-m}^{-1} \boldsymbol{\Gamma}_z(j)\boldsymbol{\phi}_z'(-j) + \boldsymbol{\Sigma}_\varepsilon + \mathbf{P}\boldsymbol{\Sigma}_a\mathbf{P}',$$

$$\boldsymbol{\Gamma}_z(l) = \sum_{j=l-m}^{l-1} \boldsymbol{\Gamma}_z(j)\boldsymbol{\phi}_z'(l-j) - \boldsymbol{\Sigma}_\varepsilon\boldsymbol{\phi}_z'(l), \qquad l = 1, \ldots, p,$$

$$\boldsymbol{\Gamma}_z(l) = \sum_{j=1}^{m} \boldsymbol{\Gamma}_z(l-j)\boldsymbol{\phi}_z'(j), \qquad l > p,$$

and by writing

$$\boldsymbol{\Sigma}_\varepsilon + \mathbf{P}\boldsymbol{\Sigma}_a\mathbf{P}' = \sum_{j=0}^{m} \boldsymbol{\psi}_z(j)\boldsymbol{\Sigma}_u\boldsymbol{\theta}_z'(j),$$

$$\boldsymbol{\Sigma}_\varepsilon\boldsymbol{\phi}_z'(l) = \sum_{j=0}^{m-l} \boldsymbol{\psi}_z(j)\boldsymbol{\Sigma}_u\boldsymbol{\theta}_z'(j+l), \qquad l = 1, \ldots, p,$$

the process will follow an ARMA(p, p) model. In the general case, if $p_y > q_y$ and $m = p_y$, the first p_y covariance matrices will be given by (A.3), so they will not follow the AR pattern and an MA(p) component will be needed. Thus the required model will be ARMA(p, p). On the other hand, if $q_y > p_y$, the first q_y matrices will follow the nonsystematic structure and the process will be ARMA(p_y, q_y).

A.2 Proof of Theorem 2

Let $m = \max(p, q)$; then, assuming that $\theta(q+1) = \cdots = \theta(m) = 0$ if $m > q$ and $\phi(p+1) = \cdots = \phi(m) = 0$ if $m > p$,

$$\begin{aligned}\boldsymbol{\theta}^*(B) &= (\mathbf{I} - (\boldsymbol{\theta}_1 + \mathbf{A})B - \cdots - (\boldsymbol{\theta}_m + \mathbf{A})B^m) \\ &= \boldsymbol{\theta}(B) + \mathbf{A}\mathbf{I}_m(B)\end{aligned}$$

and

$$\boldsymbol{\phi}^*(B) = \boldsymbol{\phi}(B) + \mathbf{A}\mathbf{I}_m(B).$$

The matrices $\theta^*(B)$ and $\phi^*(B)$ are equally acceptable as parameter matrices of the process if

$$\boldsymbol{\phi}(B)\mathbf{z}_t + \mathbf{A}\mathbf{I}_m(B)\mathbf{z}_t = \boldsymbol{\theta}(B)\mathbf{u}_t + \mathbf{A}\mathbf{I}_m(B)\mathbf{u}_t,$$

so the condition is

$$\mathbf{A}\mathbf{z}_t = \mathbf{A}\mathbf{u}_t \tag{A.4}$$

and there must be $k - r$ linear combination of the series that are random noise. Suppose that $\mathbf{z}_t = \mathbf{P}\mathbf{y}_t + \boldsymbol{\epsilon}_t$ if we multiply for any matrix $\mathbf{V}'$ such that $\mathbf{V}'\mathbf{P} = 0$. Then by Sylvester's law of nullity, rank(V) $\leq k - r$ and

$$\mathbf{V}'\mathbf{z}_t = \mathbf{V}'\boldsymbol{\epsilon}_t. \tag{A.5}$$

We now prove that any matrix V' that verifies (A.5) must verify (A.4), since

$$\mathbf{z}_t = \boldsymbol{\phi}(B)^{-1}\boldsymbol{\theta}(B)\mathbf{u}_t = \boldsymbol{\psi}(B)\mathbf{u}_t$$

and

$$\boldsymbol{\epsilon}_t = \boldsymbol{\psi}(B)\mathbf{u}_t - \mathbf{P}\mathbf{y}_t.$$

If $\mathbf{V}'\mathbf{P} = 0$, then $\mathbf{V}'\boldsymbol{\epsilon}_t = \mathbf{V}'\boldsymbol{\psi}(B)\mathbf{u}_t$, and, as $\mathbf{V}'\boldsymbol{\epsilon}_t$ must be a white-noise sequence,

$$\mathbf{V}'\boldsymbol{\epsilon}_t = \mathbf{V}'\mathbf{u}_t, \qquad \mathbf{V}'\boldsymbol{\psi}_i = \mathbf{0}, \qquad i \geq 1. \tag{A.6}$$

Therefore, if we take $\mathbf{A} = \mathbf{V}'$ the proof is complete. Note that there always exists a matrix $\mathbf{V}'$ such that $\mathbf{V}'\mathbf{P} = 0$. We can take, for example, as the first $k - r$ rows of $\mathbf{V}'$ a set of the $k - r$ linear independent eigenvectors of $\mathbf{PP}'$ that are linked to zero eigenvalues. Then we can put in the leftover r positions any r vectors of the linear subspace generated by the first $k - r$ rows.

[*Received August 1984. Revised November 1986.*]

REFERENCES

Box, G. E. P., and Tiao, G. C. (1977), "A Canonical Analysis of Multiple Time Series," *Biometrika*, 64, 355–365.

Brillinger, D. R. (1981), *Time Series, Data Analysis and Theory* (expanded ed.), San Francisco: Holden-Day.

Engle, R., and Watson, M. (1981), "A One-Factor Multivariate Time Series Model of Metropolitan Wage Rates," *Journal of the American Statistical Association*, 76, 774–781.

Geweke, J. (1977), "The Dynamic Factor Analysis of Economic Time Series Models," in *Latent Variables in Socio-Economic Models*, eds. D. J. Aigner and A. S. Goldberger, Amsterdam: North-Holland, pp. 365–383.

Geweke, J., and Singleton, K. (1981), "Maximum Likelihood Confirmatory Factor Analysis of Economic Time Series," *International Economic Review*, 22, 37–54.

Hannan, E. J. (1981), "Estimating the Dimension of a Linear System," *Journal of Multivariate Analysis*, 11, 459–473.

Peña, D., and Sánchez-Albornoz, N. (1984), "Wheat Prices in Spain, 1857–1890: An Application of the Box–Jenkins Methodology," *Journal of European Economics History*, 13, 353–373.

Priestley, M. B. (1981), *Spectral Analysis and Time Series* (Vol. 2), New York: Academic Press.

Quenouille, M. H. (1957), *The Analysis of Multiple Time Series,* London: Charles W. Griffin.

Rao, T. Subba (1975), "An Innovation Approach to the Reduction of the Dimensions in a Multivariate Stochastic System," *International Journal of Control,* 21, 673–680.

Reinsel, G. (1983), "Some Results on Multivariate Autoregressive Index Models," *Biometrika,* 70, 145–156.

Sánchez-Albornoz, N. (1975), *Los Precios Agricolas Durante la Segunda Mitad del Siglo XIX,* Madrid: Banco de España.

Sargent, T. J., and Sims, G. A. (1977), "Business Cycle Modeling Without Pretending to Have Too Much A Priori Economic Theory," in *New Methods in Business Cycle Research: Proceedings From a Conference,* ed. C. A. Sims, Minneapolis, MN: Federal Reserve Bank of Minneapolis, pp. 45–110.

Tiao, G. C., and Box, G. E. P. (1981), "Modeling Multiple Time Series With Applications," *Journal of the American Statistical Association,* 76, 802–816.