

# COINTEGRATION AND COMMON FACTORS

BY ALVARO ESCRIBANO AND DANIEL PEÑA

*Universidad Carlos III de Madrid*

*First version received April 1993*

**Abstract.** Alternative common factor representations for cointegrated vectors are studied. This is done by embedding them into the dynamic factor model proposed by Peña and Box (Identifying a simplifying structure in time series. *J. Am. Statist. Assoc.* 82 (1987), 836–43). It is shown that dynamic factor models produce as a particular case the alternative common trend representations for cointegrated variables available in the literature. Furthermore a new normalization is proposed which has the advantage of producing common trend representations with moving-average polynomials and under certain circumstances with uncorrelated shocks.

**Keywords.** Dynamic factor models; cointegration; common factors; common trends; uncorrelated shocks.

## 1. INTRODUCTION

One of the most important contributions of the autoregressive integrated moving-average (ARIMA) methodology advocated by Box and Jenkins (1976) is to show the advantages of differencing univariate time series to obtain stationary processes. However, in the multivariate case differencing should be made with great care as shown by Box and Tiao (1977). Granger (1981, 1986) and Engle and Granger (1987) proposed the concept of cointegration when the number of unit roots in the multivariate system is less than the number of variables in the system. Granger's representation theorem established the relationship between the vector autoregressive (VAR), the vector moving-average (VMA) and the vector error correction (VEC) representations. Peña and Box (1984, 1987) showed how to identify common factors in a vector of time series and how to build a simplifying transformation to recover the factors as linear combinations of the original series. Stock and Watson (1988) showed that cointegrated multiple time series must have, at least, one common trend or factor. This work was a first step towards clarifying the connection between cointegration and common factor analysis, further studied by Peña (1990), Johansen (1991) and Gonzalo and Granger (1992).

The importance of the concept of cointegration in the economic literature is due to the possibility of putting together the information about the long-run equilibrium, coming from economic theory, and the statistical evidence about

short-run dynamics in the observed series. In particular, the presence of common factors implies that the autoregressive (AR) and moving-average (MA) matrices of the vector autoregressive integrated moving-average (VARIMA) representation are not unique (Peña and Box, 1987). Then if a vector autoregressive moving-average (VARMA) model is fitted a large number of parameters may be needed and these parameters will be highly correlated. Also, at the scalar level, the estimation and formulation of single equation models is strongly influenced by the fact that the variables are cointegrated, see Granger and Newbold (1974) and Phillips (1991). Surveys on some of these topics are given in Escribano (1990), Campbell and Perron (1991) and Johansen (1994).

The paper is organized as follows. Section 2 presents briefly the notation and the time series representations used in the paper. Section 3 introduces a new normalization of the common trend representation and proves the relationships between cointegration and the alternative common factor models. Section 4 shows that most of the common trend representations proposed in the literature are 'equivalent' in the long run and that the dynamic factor models nest all common trend (factor) representations. Finally Section 5 incorporates some concluding remarks.

## 2. MULTIPLE TIME SERIES REPRESENTATIONS AND COINTEGRATION

Let us consider an  $n \times 1$  time series vector  $X_t^*$  and, for simplicity, we shall measure all variables in deviations from the mean  $X_t = X_t^* - \mu_t$ . Given the initial conditions,  $X_j^* = 0$  for  $j \leq 0$ , the mean  $\mu_t = E(X_t^*)$  is a column vector of  $n$  components which can include constants, deterministic time trends, dummy variables etc. Let the vector  $X_t$  follow a VARMA representation

$$\phi(B)X_t = \theta(B)\varepsilon_t \quad (2.1)$$

where  $\phi(B) = I - \phi_1 B - \dots - \phi_p B^p$ ,  $\theta(B) = I - \theta_1 B - \dots - \theta_q B^q$ ,  $\phi_i$  ( $i = 1, \dots, p$ ) and  $\theta_i$  ( $i = 1, \dots, q$ ) are  $n \times n$  square matrices,  $B$  is the backshift operator and  $\varepsilon_t$  is a sequence of uncorrelated variables with zero mean and positive definite covariance matrix  $\Sigma$ . We assume that  $\phi(B)$  and  $\theta(B)$  are coprime, i.e. they do not have roots in common, and that the zeros of the determinantal polynomials are as follows:

- (i)  $|\phi(B)|$  has zeros *on or outside* the unit circle;
- (ii)  $|\theta(B)|$  has all zeros *outside* the unit circle.

### 2.1. Jointly $I(1)$ variables

**DEFINITION 1.** The univariate time series  $x_t$  is *integrated of order 1*,  $I(1)$ , if  $(1 - B)x_t$  follows an invertible  $MA(\infty)$  process.

DEFINITION 2. The time series vector  $X_t$  is *jointly integrated of order 1*,  $JI(1)$ , if (i) all the individual components are  $I(1)$ ; and (ii)  $(1 - B)X_t$  follows an invertible  $VMA(\infty)$  with full rank matrices.

We have added the term *jointly* to distinguish the multivariate  $I(1)$  concept from the univariate one.

If  $X_t$  is  $JI(1)$ , the determinantal equation of  $\phi(B)$  has  $n$  roots on the unit circle,  $\phi(1)$  is equal to 0, we can write<sup>1</sup>  $\phi(B) = (1 - B)\phi^*(B)$  and calling  $\Pi^*(B) = \theta^{-1}(B)\phi^*(B)$  model (2.1) can be written as a  $VAR(\infty)$  in first differences,

$$\Pi^*(B)(1 - B)X_t = \varepsilon_t \quad (2.2)$$

where  $\Pi^*(B) = I - \sum_{i=1}^{\infty} \Pi_i^* B^i$  has all its roots outside the unit circle. The  $VMA(\infty)$  is obtained by inverting the matrix  $\Pi^*(B)$  and calling  $C(B) = \Pi^*(B)^{-1}$ ,

$$(1 - B)X_t = C(B)\varepsilon_t \quad (2.3)$$

where  $C(B) = \sum_{j=0}^{\infty} C_j B^j$  is an  $n \times n$  matrix of polynomials in  $B$  with  $C_0 = I$ . The  $JI(1)$  condition of  $X_t$  implies that the roots of the equation  $|C(B)| = 0$  are outside the unit circle, and therefore  $C(1)$  has full rank. In the next section the multiple unobserved components model, where  $X_t$  is driven by  $n \times 1$  linearly independent random walks, is needed and it is obtained by decomposing the matrix  $C(B)$  as the sum of the zero frequency ( $B = 1$ ) components and the rest of the frequency terms,

$$C(B) = C(1) + (1 - B)C^*(B) \quad (2.4)$$

where  $C^*(B) = (1 - B)^{-1}\{C(B) - C(1)\}$  and  $C_j^* = -\sum_{i=j+1}^{\infty} C_i$ , under the condition that  $\sum_{j=0}^{\infty} j^{1/2} |C_j| < \infty$ . Substituting (2.4) into (2.3) under the initial condition that  $\varepsilon_{t-j} = 0$  for  $j \geq t$ , we get

$$X_t = C(1)T_t + C^*(B)\varepsilon_t \quad (2.5)$$

$$T_t = T_{t-1} + \varepsilon_t \quad (2.6)$$

where  $T_t$  is an  $n \times 1$  vector of *random walks* and  $C(1)$  is an  $n \times n$  matrix of *full rank*,  $\text{rank}\{C(1)\} = n$ , and so the  $n \times 1$  system of equations is driven by linear combinations of  $n \times 1$  different random walks.

## 2.2. Cointegrated variables, $JI(0)$

*Cointegration*, a term coined by Granger (1981), occurs when all individual series are  $I(1)$  but some linear combinations are  $I(0)$ .

DEFINITION 3. The components of the vector  $X_t$  are *cointegrated of order 1*,  $CI(1, 1)$ , with rank  $r$  if (i) all individual components are univariate  $I(1)$ ; and (ii) there are  $r$  linearly independent combinations  $\beta' X_t$  which are  $JI(0)$  and therefore  $\text{rank}(\beta) = r < n$ .

Notice that it is possible to define the concept of cointegration allowing some of the elements of the vector  $X_t$  to be I(0) as long as some of the others are I(1), but this complicates the algebra (see Davidson, 1991) and for simplicity we shall not consider this possibility here.

It is important to stress that if  $\beta' X_t$  is I(0),  $R\beta' X_t$  will also be I(0) for any non-singular  $R$  matrix. Calling  $T' = R\beta'$  the cointegration matrix we can always choose  $R$  to make  $T'T = I$ . To see this, note that given any  $r \times n$  matrix  $\beta'$  of full rank  $r$ , we can always take  $R = A'(\beta'\beta)^{-1}$ , where  $A$  is the square root of the positive definite matrix  $\beta'\beta$ , i.e. it verifies that  $(\beta'\beta) = AA'$ . Then  $T'T = A'(\beta'\beta)^{-1}\beta'\beta(\beta'\beta)^{-1}A = I$ . Therefore, without loss of generality, we can always assume that the cointegration matrix  $\beta$  is normalized in such a way that  $\beta'\beta = I$ .

From Granger's representation theorem (see Engle and Granger, 1987) we know that cointegration implies certain restrictions on the matrices  $\Pi(1)$  and  $C(1)$ . The autoregressive matrix  $\Pi(1)$  has rank  $\{\Pi(1)\} = r$  and can be written as  $\pi(1) = \alpha\beta'$ , where  $\beta$  is the cointegration matrix introduced in Definition 3 and  $\alpha$  and  $\beta$  are  $n \times r$  matrices. Calling  $\beta_\perp$  the  $n \times (n-r)$  matrix that spans the null space of  $\beta$  ( $\beta'\beta_\perp = 0$ ), we obtain  $\Pi(1)\beta_\perp = 0$  and the rows of the matrix  $\Pi(1)$  span the same space as the rows of the matrix  $\beta'$ , which will be called the cointegration space. Also, the columns of the matrix  $\Pi(1)$  span the same space as  $\alpha$ , and the matrix  $\alpha$  verifies  $C(1)\alpha = 0$ . Cointegration implies that  $\Pi(1) \neq 0$  and so Equation (2.2) is misspecified.

In terms of the VMA( $\infty$ ), Equation (2.3), cointegration implies that the matrix  $C(1)$ , which measures the long-run impact of shocks in impulse response analysis, has rank  $n-r$ . The reason for the reduced rank of  $C(1)$  is clear, since premultiplying Equation (2.5) by  $\beta'$  we see that for  $\beta' X_t$  to be I(0) the condition  $\beta' C(1) = 0$  must be satisfied.

After having reviewed the main known implications of cointegration we shall study in the next section the relationship with dynamic common factor representations.

### 3. COINTEGRATION AND COMMON FACTOR MODELS

Following Peña and Box (1987) we shall say that  $X_t$  follows a dynamic factor model when

$$X_t = Af_t + u_t \quad (3.1)$$

$$\phi(B)f_t = \theta(B)a_t \quad (3.2)$$

where  $A$  is a  $n \times k$  factor loading matrix, that without loss of generality can be taken such that  $A'A = I$ ,  $f_t$  is a vector of common factors, which follows the vector ARIMA representation (3.2), and  $a_t$  and  $u_t$  are uncorrelated zero mean white noise processes with covariance matrices  $\Sigma_u$  and  $\Sigma_a$ , where  $\Sigma_a$  is diagonal.

If  $k = n$ , for the model to be identified  $\Sigma_u = 0$ . Otherwise, if  $k < n$ , we

assume that  $\Sigma_u$  is positive definite. When all of the matrices  $\phi_i$ ,  $\theta_i$  and  $\Sigma_a$  are diagonal the model is called the uncorrelated factors model. In the dynamic case the model is identified even if the factors are correlated and, then, the matrices  $\phi_i$ ,  $\theta_i$  and  $\Sigma_a$  need not be diagonal. However, one of the matrices  $\Sigma_u$  or  $\Sigma_a$  must be chosen to be diagonal in order for the model to be identified. We shall assume here that  $\Sigma_a$  is diagonal but the factors may be correlated with non-diagonal parameter matrices.

**THEOREM.** *The three following propositions are equivalent:*

- (i)  $X_t$  is  $CI(1, 1)$  with rank  $r$ ;
- (ii)  $X_t$  can be written as generated by  $n - r$  common trends;
- (iii)  $X_t$  has  $n - r$  common factors that are  $II(1)$ , and  $r$  factors that are  $II(0)$ .

**PROOF:** First, we show that cointegration implies a common trend representation, i.e. (i)  $\rightarrow$  (ii). From Section 2.2, we can decompose the matrix  $C(1)$  by using the Jordan canonical form

$$C(1) = HJH^{-1} \quad (3.3)$$

where  $H$  is partitioned as  $H = (H_1, H_2)$  with  $H_1$  and  $H_2$  of order  $n \times (n - r)$  and  $n \times r$  respectively, containing the right eigenvectors of  $C(1)$ . The  $J$  matrix is block diagonal, contains all eigenvalues of  $C(1)$  and, since under cointegration the rank of  $C(1)$  is  $n - r$ , will have  $n - r$  eigenvalues not equal to zero in the first block of the diagonal,  $J_{11}$ , and zeros in all the other blocks. If  $H^{-1}$  is partitioned conformably as  $H^{-1} = (H_1^*, H_2^*)'$ , where  $H_2^*$  contains the left eigenvectors linked to zero eigenvalues, under cointegration the decomposition (3.3) can be reduced to

$$C(1) = H_1 J_{11} H_1^* = H_1 H_3 \quad (3.4)$$

where  $H_3 = J_{11} H_1^*$ . Note that as  $\beta' C(1) = 0$ ,  $\beta'$  contains the left eigenvectors linked to the zero eigenvalues of  $C(1)$  and, therefore,  $H_2^* = \beta'$  and  $H_1 = \beta_\perp$ , where  $\beta_\perp$  is the  $(n - r) \times n$  matrix such that  $\beta_\perp \beta = 0$ . Substituting (2.4) with  $C(1) = \beta_\perp H_3$  in (2.3) and calling the stationary process  $C^*(B)\varepsilon_t = e_t$  and  $H_3 \varepsilon_t = v_t$ , we obtain

$$(1 - B)X_t = \beta_\perp v_t + (1 - B)e_t \quad (3.5)$$

which can be written as

$$X_t = \beta_\perp \tau_t + e_t \quad (3.6)$$

$$\tau_t = \tau_{t-1} + v_t \quad (3.7)$$

where the common stochastic trend or common factor  $\tau_t$  is a vector of order  $(n - r) \times 1$  instead of the  $n \times 1$  vector  $T_t$  of (2.5). This representation was obtained by Stock and Watson (1988). Note that (a) both equations are driven by the same white noise process  $\varepsilon_t$ ; and (b)  $C^*(B)$  is not an MA polynomial because  $C^*(0) = C_0^* \neq 0$ .

However, we can write (3.6) and (3.7) in an alternative way that avoids both of these problems. To do so, assume that  $\beta$  and  $\beta_\perp$  are normalized so that  $\beta'\beta = I$ ,  $\beta'_\perp\beta_\perp = I$ . Then  $\beta\beta'$  and  $\beta_\perp\beta'_\perp$  are idempotent projection matrices on the spaces defined by  $\beta$  and  $\beta_\perp$  and they verify

$$\beta_\perp\beta'_\perp + \beta\beta' = I. \quad (3.8)$$

Using (3.8), we can write (3.6) as

$$X_t = \beta_\perp(\tau_t + \beta'_\perp e_t) + \beta\beta' C^*(B)\varepsilon_t \quad (3.9)$$

calling  $\tau_t^* = \tau_t + \beta'_\perp e_t$  and  $\eta_t = \beta' C^*(B)\varepsilon_t$ , we have

$$X_t = \beta_\perp \tau_t^* + \beta \eta_t \quad (3.10)$$

where  $\tau_t^*$  does not follow a random walk process but a general multivariate ARIMA model. Note that (3.10) represents  $X_t$  as driven by  $n - r$   $\text{JI}(1)$  common trend components  $\tau_t^*$ , plus a vector of  $r$   $\text{JI}(0)$  components,  $\eta_t$ . The model for  $\tau_t^*$  is

$$(1 - B)\tau_t^* = H_3\varepsilon_t + \beta'_\perp(1 - B)C^*(B)\varepsilon_t$$

and knowing that  $(1 - B)C^*(B) = C(B) - C(1)$ , and  $\beta'_\perp C(1) = H_3$ , we obtain

$$(1 - B)\tau_t^* = \beta'_\perp C(B)\varepsilon_t. \quad (3.11)$$

Let  $u_{1t}$  and  $u_{2t}$  be defined by  $\beta_\perp u_{1t} = \varepsilon_t$  and  $\beta u_{2t} = \varepsilon_t$ . Then,  $u_{1t} = \beta'_\perp \varepsilon_t$  and  $u_{2t} = \beta' \varepsilon_t$ , and (3.10) can be written as

$$X_t = \beta_\perp \tau_t^* + \beta L(B)u_{2t} \quad (3.12)$$

where  $L(B) = \beta' C^*(B)\beta$ , and (3.11) can be written as

$$(1 - B)\tau_t^* = G(B)u_{1t} \quad (3.13)$$

where  $G(B) = \beta'_\perp C(B)\beta_\perp$ . Equations (3.12) and (3.13) form a new common trend representation, different from the one given by Stock and Watson (1988). This representation has the following advantages:

(i) when  $\Sigma_\varepsilon = \sigma^2 I$ ,  $u_{1t}$  and  $u_{2t}$  are uncorrelated and each equation is driven by noises that are orthogonal. If  $\Sigma_\varepsilon$  is general, then  $E[u_{1t} u_{2t}'] = \beta'_\perp \Sigma_\varepsilon \beta$ ;

(ii) the polynomials  $L(B)$  and  $G(B)$  are MA polynomials because as  $\beta' C(1) = 0$  implies that  $\beta' C_0^* = \beta'$ , then  $L(0) = \beta' C_0^* \beta = \beta' \beta = I$  and  $G(0) = \beta'_\perp \beta_\perp = I$ .

Now, we shall prove (ii)  $\rightarrow$  (iii), i.e. this common trend model implies a particular case of the dynamic factor model. Starting from (3.10), let us call  $A = [\beta_\perp, \beta]$  and  $f_t' = [\tau_t^*, \eta_t']$ . Then, it is obvious that (3.10) is a particular case of the factor model (3.1) with  $k = n$  and, therefore,  $\Sigma_u = 0$ . The vector of factors is formed by  $n - r$   $\text{JI}(1)$  factors  $\tau_t^*$  and  $r$   $\text{JI}(0)$  factors  $\eta_t$ , and it follows the model

$$\begin{bmatrix} (1 - B)I & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \tau_t^* \\ \eta_t \end{bmatrix} = \begin{bmatrix} G(B) & 0 \\ 0 & L(B) \end{bmatrix} \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \quad (3.14)$$

where the covariance matrix of  $u_t$  is block diagonal if  $\Sigma_\varepsilon = \sigma^2 I$ . It is clear that (3.12) and (3.14) are a particular case of (3.1) and (3.2) in which the model for the factors is written as a moving average with a block diagonal structure for the factors.

The proof that (ii)  $\rightarrow$  (i) (common trends implies cointegration) is trivial: we multiply (3.6) by  $\beta'$  and obtain that  $\beta' X_t = \beta' C^*(B) \varepsilon_t$  is a stationary process. The proof that (iii)  $\rightarrow$  (ii) is also straightforward. Starting from the dynamic factor model, let us assume that the  $k$  factors can be decomposed into  $(n - r) < k$  JI(1) factors,  $(1 - B)\tau_t^* = \theta_1(B)a_{1t}$ , and  $r$  JI(0) factors that follow the model  $w_t = \theta_2(B)a_{2t}$ , where  $f'_t = [\tau_t^{*'} w_t']$  and  $\theta_i(B)$  for  $i = 1, 2$  have all roots outside the unit circle. Then, partitioning  $A$  as  $[A_1, A_2]$ , the model (3.1) can be written as

$$X_t = A_1 \tau_t^* + A_2 w_t \quad (3.15)$$

where the first term  $A_1 \tau_t^*$  contains the JI(1) factors and the second term  $A_2 w_t$  is JI(0). ■

One conclusion we can draw from this theorem is the following. Peña and Box (1987) proved that if we build a VARIMA model for a vector of time series generated by a dynamic factor model, the AR and MA matrices of the VARIMA representation are not uniquely determined. Then, a large number of parameters may be needed to obtain an adequate representation of the system and the estimated parameters will be highly correlated. Therefore, a very complex relationship seems to occur when in fact, if the common factors are taken into account, the model could be very simple. This result suggests that when the vector of data may be cointegrated it is safer to build a VAR model because in this case the AR matrices will be identified. As many economic variables are cointegrated, this result might explain the wide use of VAR modelling in the econometric literature.

#### 4. COMMON FACTOR REPRESENTATIONS

Although common factors, common trends and cointegration are three very related ideas, as proved in the previous theorem, several authors have suggested different ways to obtain the factors or the cointegration relationships. Peña and Box (1987) suggested to get the factors as

$$\hat{f}_t = M X_t = \begin{bmatrix} A_\perp \\ \bar{A} \end{bmatrix} X_t \quad (4.1)$$

where  $\bar{A}$  is the Moore–Penrose generalized inverse of  $A$  given by the  $k \times n$  matrix  $\bar{A} = (A' A)^{-1} A' = A'$  and  $A'_\perp$  is the  $(n - k) \times n$  orthogonal complement of  $A$  such that  $\text{rank}(A'_\perp) = n - k$  and  $A'_\perp A = 0_{n-k}$ . When the vector of series is driven by  $n - r$  JI(1) common factors and  $r$  JI(0) factors, the transformation to recover the JI(1) common factors can be easily built as

follows: the orthogonal complement of the factor loading matrix  $A_1$  for the  $JI(1)$  factors is  $\beta'$ , the cointegration matrix, because  $\beta' A_1$  must be zero if  $\beta' X_t$  is  $JI(0)$ . As the matrix  $M$  is partitioned into two orthogonal complements, calling  $\beta'_\perp$  the  $(n-r) \times n$  orthogonal complement of  $\beta$  the transformation will be

$$\hat{f}_t = \begin{bmatrix} \beta'_\perp \\ \beta' \end{bmatrix} X_t \quad (4.2)$$

where  $\beta'_\perp$  is also the Moore–Penrose generalized inverse of  $A_1$  and  $A_1$  must be equal to  $\beta_\perp$ . Transformation (4.2) decomposes  $X_t$  into two components: the first,  $\beta'_\perp X_t$ , is  $JI(1)$  and contains the non-stationary common factors driving the observed vector of time series; the second,  $\beta' X_t$ , is  $JI(0)$  and includes the cointegration relationships. This decomposition can also be expressed in general terms as

$$X_t = P_1 \beta'_\perp X_t + P_2 \beta' X_t \quad (4.3)$$

and it will be called the *observed factors model* decomposition of  $X_t$  for certain  $P_1$  and  $P_2$  matrices of dimension  $n \times (n-r)$  and  $n \times r$  that must satisfy  $P_1 \beta'_\perp + P_2 \beta' = I$ . Premultiplying Equation (4.3) by  $\beta'$  we obtain the conditions  $\beta' P_1 = 0$ ,  $\beta' P_2 = I$  and by premultiplying by  $\beta'_\perp$  we obtain  $\beta'_\perp P_1 = I$ ,  $\beta'_\perp P_2 = 0$ . This set of conditions will be satisfied by choosing, for instance,  $P_1 = B_\perp$ ,  $P_2 = \beta$ , and then a simple case of (4.3) is

$$X_t = \beta_\perp \beta'_\perp X_t + \beta \beta' X_t. \quad (4.4)$$

The representation obtained by Stock and Watson (1988) is also a particular case of (4.3). They define the observed common factors by  $\beta'_\perp X_t$  and take as the factor loading matrix  $\beta_\perp (\beta'_\perp \beta_\perp)^{-1}$  so that the decomposition (4.3) is

$$X_t = \beta_\perp (\beta'_\perp \beta_\perp)^{-1} \beta'_\perp X_t + \beta (\beta' \beta)^{-1} \beta' X_t \quad (4.5)$$

which is identical to (4.4) if we impose the normalization conditions that  $\beta' \beta = I$  and  $\beta'_\perp \beta_\perp = I$ . Representation (4.5) was used as well by Johansen (1991). Kasa (1992) suggested using decomposition (4.5) but defining the common factors as  $(\beta'_\perp \beta_\perp)^{-1} \beta'_\perp X_t$  with loading matrix  $\beta_\perp$ . Of course, the difference between these two formulations is only due to the normalization chosen. These questions are examples of the well-known identification problem between the loading matrix and the factors that appear in the dynamic as well as in the static factor model, as stressed by Peña and Box (1987).

Gonzalo and Granger (1992) have suggested an alternative observed common factor model for cointegrated variables in which the common factors are given by  $\alpha'_\perp X_t$ , where  $\alpha_\perp$  is the null space of  $\alpha$  and  $C(1)\alpha = 0$ . Then, their decomposition is

$$X_t = \beta_\perp (\alpha'_\perp \beta_\perp)^{-1} \alpha'_\perp X_t + \alpha (\beta' \alpha)^{-1} \beta' X_t. \quad (4.6)$$

We shall now show the relationship between representations (4.5) and (4.6). Subtracting (4.6) from (4.5) we get

$$\beta_{\perp}(\beta'_{\perp}\beta_{\perp})^{-1}\beta'_{\perp}X_t = \beta_{\perp}(\alpha'_{\perp}\beta_{\perp})^{-1}\alpha'_{\perp}X_t + \{\beta(\beta'\beta)^{-1} - \alpha(\beta'\alpha)^{-1}\}\beta'X_t \quad (4.7)$$

which indicates that both  $\text{JI}(1)$  components are cointegrated, with a cointegrating matrix equal to the identity matrix  $I_{n-r}$ . Premultiplying by  $\beta'_{\perp}$  in (4.7) we get the  $\text{JI}(1)$  common factors of (4.5) in terms of those of (4.6),

$$\beta'_{\perp}X_t = \beta'_{\perp}\beta_{\perp}(\alpha'_{\perp}\beta_{\perp})^{-1}\alpha'_{\perp}X_t - \beta'_{\perp}\alpha(\beta'\alpha)^{-1}\beta'X_t. \quad (4.8)$$

Alternatively, from (4.7) we get, by premultiplying by  $\alpha'_{\perp}$ , the  $\text{JI}(1)$  common factors of Gonzalo and Granger (1992) in terms of those of (4.5),

$$\alpha'_{\perp}X_t = \alpha'_{\perp}\beta_{\perp}(\beta'_{\perp}\beta_{\perp})^{-1}\beta'_{\perp}X_t - \alpha'_{\perp}\beta(\beta'\beta)^{-1}\beta'X_t. \quad (4.9)$$

Once again, from (4.8) and (4.9) it is clear that each  $\text{JI}(1)$  common factor is not simply a linear combination of the other but they differ as well by linear combinations of the cointegration vectors  $\beta'X_t$ . The main difference between  $\alpha'_{\perp}X_t$  and  $\beta'_{\perp}X_t$ , as was pointed out by Gonzalo and Granger (1992), is that the former cancels out the error correction terms  $\alpha'_{\perp}\Pi(1)X_{t-1} = \alpha'_{\perp}\alpha\beta'X_{t-1} = 0$  from the vector error correction model, forcing the shocks of the cointegrating factors to have only a temporary effect on  $X_t$ .

## 5. CONCLUSION

We have shown that a simple reformulation of the Peña and Box (1987) dynamic factors model can incorporate common trends and cointegration relationships. This is done by specifying general dynamic factors which are divided in two groups of  $\text{JI}(1)$  factors and  $\text{JI}(0)$  factors, in such a way that the total number of factors equals the total number of variables.

On a common trend representation, a new normalization on the shocks of the  $\text{JI}(1)$  and  $\text{JI}(0)$  factors is proposed, which has the advantage of producing MA polynomials and in certain cases generates uncorrelated permanent and transitory shocks. Furthermore, when obtaining the factors as linear combinations of the observed variables it is shown that Stock and Watson's (1988) factors are derived automatically following the suggestion of Peña and Box (1987). Alternatively, one could obtain the factors following Gonzalo and Granger (1992).

Finally, it is shown that both of the alternative representations of the  $\text{JI}(1)$  factors in terms of observable variables are cointegrated with an identity cointegrating matrix.

## NOTE

1 This decomposition of  $\phi(B)$  is explained in Equation (2.4).

#### ACKNOWLEDGEMENTS

We acknowledge financial support from the spanish DGICYT contracts PS-90-0014 and PB-90-0266, Cátedra Argentaria, and the first author also acknowledges financial support from the SPES Program of the EEC, contract SPES-CT91-0059. We are grateful to the referees and the associate editor for helpful comments that have led to a better paper and to Jesús Gonzalo for useful discussions.

#### REFERENCES

- BOX, G. E. P. and JENKINS, G. M. (1976) *Time Series Analysis, Forecasting and Control*, revised edn. San Francisco, CA: Holden-Day.
- and TIAO (1977) A canonical analysis of multiple time series. *Biometrika* 64, 355–65.
- CAMPBELL, J. Y. and PERRON, P. (1991) Pitfalls and opportunities: What macroeconomists should know about unit roots. In *NBER Macroeconomics Annual 1991* (ed. O. J. Blanchard and S. Fisher). Cambridge, MA: MIT Press, pp. 141–201.
- DAVIDSON, J. E. (1991) The cointegration properties of vector autoregressions models. *J. Time Ser. Anal.* 12 (1), 41–62.
- ENGLE, R. F. and GRANGER, C. W. J. (1987) Co-integration and error correction: representation, estimation and testing. *Econometrica* 55, 251–76.
- ESCRIBANO A. (1990) Introducción al tema de cointegración y tendencias. *Cuad. Econ. ICE* 44, 7–42.
- GONZALO, J. and GRANGER, C. W. J. (1992) Estimation of common long-memory components in cointegrated systems. Discussion Paper, University of California, San Diego.
- GRANGER, C. W. J. (1981) Some properties of time series data and their use in econometric model specification. *J. Econ.* 16, 121–30.
- (1986) Developments in the study of cointegrated economic variables. *Oxford Bull. Econ. Statist.* 48 (3), 213–28. In *Log-run Economic Relations. Readings in Cointegration* (ed. C. W. J. Granger and R. F. Engle). Oxford: Oxford University Press, pp. 237–66.
- and NEWBOLD, P. (1974) Spurious regressions in econometrics, *J. Econ.* 2, 111–20
- JOHANSEN, S. (1991) Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. *Econometrica* 59 (6), 1551–80.
- (1994) Estimating systems of trending variables. Forthcoming in *Econ. Rev.*
- KASA, K. (1992) Common stochastic trends in international stock markets. *J. Monetary Econ.* 29, 95–124.
- PEÑA, D. (1990) Cointegración y reducción de dimensionalidad en series temporales multivariantes. *Cuad. Econ. ICE* 44, 109–26.
- and BOX, G. E. P. (1984) Hidden relationships in multivariate time series. *Proc. of Bus. Ec. St. Am. Statist. Assoc.* 494–99.
- and — (1987) Identifying a simplifying structure in time series. *J. Am. Statist. Assoc.* 82, 836–43.
- PHILLIPS, P. C. B. (1991) Optimal inference in co-integrated systems. *Econometrica* 59, 283–306.
- STOCK, J. H. and WATSON, M. W. (1988). Testing for common trends. *J. Am. Statist. Assoc.* 83, 1097–107.