A COMMON TRENDS MODEL: IDENTIFICATION, ESTIMATION AND INFERENCE

ANDERS WARNE

ABSTRACT. Common trends models provide a useful tool for studying growth and business cycle phenomena in a joint framework (see King, Plosser, Stock and Watson (1991)). In this paper we study the problem of how to estimate and analyse a common stochastic trends model for an n dimensional time series which is cointegrated of order (1,1) with $r < n$ cointegration vectors. Identification of $k = n - r$ permanent (trend) and r transitory innovations is discussed in terms of impulse responses and variance decompositions. Finally, we derive analytical expressions of the asymptotic distributions for estimates of these functions, thereby making formal hypothesis testing and inference possible within this framework.

KEYWORDS: Cointegration, common trends, impulse response function, permanent and transitory shocks, variance decomposition.

JEL CLASSIFICATION NUMBERS: C32, C51.

1. INTRODUCTION

In many models on macroeconomic fluctuations the dichotomy between growth and cycles has played an important role. Traditionally, growth has often been treated as independent of factors that result in cyclical fluctuations (see King, Plosser and Rebelo (1988a)). In contrast, stochastic growth models (see e.g. King, Plosser and Rebelo (1988b), and King, Plosser, Stock and Watson (1987)) allow growth shocks to influence the short run fluctuations. A common feature of these models is that the number of growth disturbances is rather low relative to the number of variables.

The prevailing view in the theoretical literature seems to be that macroeconomic fluctuations arise from shocks to fundamental variables such as economic policy, preferences, and technology. These shocks are then propagated through the economy and result in systematic patterns of persistence and comovements among macroeconomic aggregates. Consequently, it should be of interest to analyse a simple time series model which makes

Date: October 1993.
I have benefitted from discussions with Michael Bergman, Nils Gottfries, Niels Haldrup, Dennis Hoffman, Svend Hylleberg, Tor Jacobson, Søren Johansen, Katarina Juselius, Sune Karlsson, Erik Mellander, Lars–Erik Öller, and Anders Vredin. This paper is based on a chapter of my PhD thesis. Responsibility for errors and obscurities rests with me alone. Financial support from *Bankforskningsinstitutet* and *Humanistisk–Samhällsvetenskapliga Forskningsrådet* is gratefully acknowledged.

it possible to examine connections between growth related shocks and transient fluctuations. Such a model will then by necessity incorporate stochastic rather than deterministic trends. Furthermore, to consider the notion of a few important growth disturbances, there will in general be fewer stochastic trends than time series.

In papers by King, Plosser, Stock, and Watson (1987,1991) and Stock and Watson (1988) the connection between cointegration and common stochastic trends was first examined in some detail. The basic idea is that there is a *reduced* number of linear stochastic trends feeding the system. This implies that there exists certain linear combinations of the levels series which ensure that the trends average out, i.e. the residuals from the linear combinations are wide sense stationary stochastic processes. King, Plosser, Stock, and Watson (1987) investigate a common trends model for five U.S. macroeconomic time series (output, consumption, investments, the price level, and the money stock) and model growth by two stochastic trends, a nominal and a real trend. With five time series and two independent stochastic trends, common sense (or algebra) suggests that we can construct three independent vectors which eliminate the trends, i.e. there are three cointegrating vectors which describe a steady state in such a system. A shortcoming of their paper is that the description of the estimation and computation strategy they make use of is somewhat limited. For example, an inversion algorithm needed to obtain estimates of, e.g. impulse response functions and forecast error variance decompositions is only mentioned. More importantly, asymptotic properties of these functions are not considered.

A purpose of this paper is to mathematically establish how one may estimate the parameters in a common stochastic trends model when the time series of interest are cointegrated of order (1,1) (see Blanchard and Quah (1989), Park (1990), and Shapiro and Watson (1988) for approaches which are related to the one I shall examine here; or Gonzalo and Granger (1992) for a factor model approach to common trends). Furthermore, I shall show how one may perform dynamic analysis within this framework when the innovations to the system are either permanent or transitory, i.e. when the responses in at least one variable to an innovation are *or* are not persistent. In particular, the calculation of impulse response functions and forecast error variance decompositions will be looked into in some detail. Finally, I shall derive asymptotic distributions of estimates of these functions in the present setting. Here, the theory is based on Baillie

(1981,1987), Lütkepohl (1988,1989,1990), Lütkepohl and Poskitt (1990), Lütkepohl and Reimers (1992), and Schmidt (1973,1974), although the particular innovations I examine complicate the analysis somewhat.

The paper is organized as follows. In section 2, I discuss some representations which are equivalent for cointegrated time series. There it is shown that a restricted vector autoregressive representation for cointegrated time series exists under familiar circumstances. Since this representation is invertible, it is well suited for calculating all other parameters of interest (see also Warne (1990)). Section 3 is concerned with the Wold moving average parameters and with identification of permanent and transitory innovations. In section 4, I analyse the asymptotic properties of impulse response functions and forecast error variance decompositions under the assumption that the lag order has a known upper bound. Section 5 summarizes the main results. Mathematical proofs of the theorems are provided in the Appendix.

2. Common Trends and Cointegration

Linear time series models are generally specified in terms of variables which can be observed and a purely nondeterministic and serially uncorrelated error. Accordingly, they can be estimated with standard tools. In contrast, a common trends model consists of a vector of trends and a vector of stationary variables, where neither component can be observed as an individual factor. Without loss of generality, let $\{x_t\}$ be a vector time series such that

$$x_t = x_t^{\mathrm{p}} + x_t^{\mathrm{s}}. \tag{1}$$

Here, x_t^{p} represents a vector of trends of x_t, while x_t^{s} is a stationary residual.

King, Plosser, Stock and Watson (KPSW) (1987) and Stock and Watson (1988) show that there is a simple duality between the concepts of cointegration and common trends. In particular, the cointegrating restrictions determine the number of independent trends and how a vector of observed variables is related to all the independent trends. That is, if α is a cointegrating vector, then $\alpha' x_t^{\mathrm{p}} = 0$ for $\alpha' x_t = \alpha' x_t^{\mathrm{s}}$ to be stationary. These restrictions, however, *neither* specify *nor* suggest whether a certain trend is related to, e.g. technology shocks or economic policy. To be able to make such interpretations it is necessary to consider further identifying assumptions. In this section I shall devote the first part to the mathematical structure of cointegrated time series and the second part to estimation and identification of the common trends parameters. As an illustration, the third part contains an example of a three dimensional system for output, the price level, and the money stock with one cointegration vector and two common trends.

2.1. **The Model.** Let $\{x_t\}$ denote an n dimensional real valued vector (discrete) time series which is driven by $k \leq n$ common stochastic trends. Specifically,

$$x_t = x_0 + \Upsilon\tau_t + \Phi(L)\nu_t. \tag{2}$$

Here, L is the lag operator, i.e. $L^j x_t = x_{t-j}$ for any integer j. The n dimensional vector sequence $\{\nu_t\}$ is assumed to be white noise with $\mathrm{E}[\nu_t] = 0$ and $\mathrm{E}[\nu_t\nu_t'] = I_n$, the $n \times n$ identity matrix. Furthermore, the $n \times n$ matrix polynomial $\Phi(\lambda) = \sum_{j=0}^{\infty} \Phi_j \lambda^j$ is finite for all λ on and inside the unit circle and, without loss of generality, we assume that x_0 is stationary. In other words, the component $\Phi(L)\nu_t$ is jointly (wide sense) stationary.

The trend or growth component of x_t is described by $\Upsilon\tau_t$. The loading matrix Υ is of dimension $n \times k$ with rank k whereas

$$\tau_t = \mu + \tau_{t-1} + \varphi_t. \tag{3}$$

Hence, τ_t is a k dimensional vector of random walks with drift μ and innovation φ_t. Let us assume that the trend disturbance sequence $\{\varphi_t\}$ is white noise with $\mathrm{E}[\varphi_t] = 0$ and $\mathrm{E}[\varphi_t\varphi_t'] = I_k$.

In relation to the decomposition in (1) we find that the common trends model in (2) and (3) specifies that

$$\begin{aligned} x_t^{\mathrm{s}} &= x_0 + \Phi(L)\nu_t, \\ x_t^{\mathrm{p}} &= \Upsilon\left[\tau_0 + \mu t + \textstyle\sum_{j=1}^{t} \varphi_j\right]. \end{aligned} \tag{4}$$

Furthermore, whenever the number of common trends, k, is less than the number of variables, n, there are exactly $r = n-k$ linearly independent vectors which are orthogonal to the columns of the loading matrix Υ. In other words, there exists and $n \times r$ matrix α such that $\alpha'\Upsilon = 0$. Accordingly, $\alpha' x_t^{\mathrm{p}} = 0$ for all t so that $z_t := \alpha' x_t$ ($:=$ denotes a definition) is jointly stationary.

The common trends model in (2) and (3), originally due to Stock and Watson (1988), has some appealing properties. First, the trends include a stochastic element which is consistent with the notion that some shocks to an economy are persistent. Second, there may be fewer trends than variables so that the model allows for steady state relationships between the variables. In this framework, these steady states are described by the matrix α. Furthermore, if φ_t and ν_t are correlated it is possible for the trend disturbances to influence not only growth but also fluctuations about the trends. In fact, the approach we shall take in this paper implies that the first k elements of ν_t are given by φ_t.

To determine how we can estimate the common trends model let us assume that $\{x_t\}$ is generated by the unrestricted vector autoregression (VAR) of order p:

$$A(L)x_t = \rho + \varepsilon_t. \tag{5}$$

The n dimensional sequence of (reduced form) disturbances $\{\varepsilon_t\}$ is white noise with $\mathrm{E}[\varepsilon_t] = 0$ and $\mathrm{E}[\varepsilon_t \varepsilon_t'] = \Sigma$, a positive definite matrix. The $n \times n$ matrix polynomial $A(\lambda) = I_n - \sum_{j=1}^{p} A_j \lambda^j$ satisfies $\det[A(\lambda)] = 0$ if and only if $|\lambda| > 1$ or $\lambda = 1$ so that explosive $\{x_t\}$ processes are ruled out. Moreover, the only form of nonstationarity which is possible is due to unit roots. In other words, if $\{x_t\}$ is generated by (5), then the process is integrated of order d, where d is a nonnegative integer (for a definition of integration, see Johansen (1991)).

If $\{x_t\}$ in (5) is cointegrated of order (1,1) with r cointegration vectors we know from Granger's Representation Theorem (GRT) that (i) $\mathrm{rank}[A(1)] = r$, and (ii) $A(1) = \gamma\alpha'$ (see Engle and Granger (1987), Hylleberg and Mizon (1989), and Johansen (1988a,1989, 1991)). The matrices γ and α are of dimension $n \times r$ and the columns of α are called the cointegration vectors. Under the assumption of cointegration it follows by GRT that an alternative form of (5) is

$$A^*(L)\Delta x_t = \rho - \gamma z_{t-1} + \varepsilon_t, \tag{6}$$

where $\Delta := 1 - L$ is the first difference operator and the matrix polynomial $A^*(\lambda) = I_n - \sum_{i=1}^{p-1} A_i^* \lambda^i$ is related to $A(\lambda)$ through $A_i^* = -\sum_{j=i+1}^{p} A_j$ for $i = 1, \dots, p-1$.

The representation in (6) is widely known as the vector error correction (VEC) model (see e.g. Hylleberg and Mizon (1989)). Cointegration implies that the r dimensional process $\{z_t\}$ is jointly stationary. If we regard the cointegration vectors as describing a steady state or a long run equilibrium for x the term γz_{t-1} represents the correction of the change in x_t due to last periods long run equilibrium error. Note that the major difference between equations (5) and (6) is that the latter representation is conditioned on cointegration while the former is merely consistent with unit roots.

Engle and Granger's (1987) version of GRT is based on the existence of a Wold vector moving average (VMA) representation of the form

$$\Delta x_t = \delta + C(L)\varepsilon_t. \tag{7}$$

The matrix polynomial $C(\lambda) = I_n + \sum_{j=1}^{\infty} C_j \lambda^j$ is assumed to be 1–summable in the sense of Brillinger (1981), i.e. $\sum_{j=1}^{\infty} j|C_j|$ is finite. In other words, the time series $\{\Delta x_t\}$ is jointly stationary. In addition, if $C(1) \neq 0$ it follows that $\{x_t\}$ is nonstationary. Specifically, Engle and Granger find that if $\{x_t\}$ is cointegrated of order (1,1), then $C(1)$ has rank $n - r$ and $\alpha' C(1) = 0$. That is, a VMA representation of the form in (7) and cointegration jointly imply the existence of the unrestricted VAR and of the VEC representations in (5) and (6), respectively (with p, the lag order, possibly infinite).

Johansen (1991), on the other hand, shows that if $\{x_t\}$ is generated by (5), $A(1) = \gamma\alpha'$, and the $(n - r) \times (n - r)$ matrix

$$\gamma_\perp' \left(\sum_{j=1}^{p} j A_j \right) \alpha_\perp$$

is nonsingular, then $\{x_t\}$ is cointegrated of order (1,1) with r cointegration vectors such that there exists a Wold representation of the form in (7). The nonsingularity requirement for the above matrix rules out the possibility that $\{x_t\}$ is integrated of an order greater than 1.

From mathematical and statistical perspectives, Johansen's approach to the GRT is the more natural. The VAR in (5) is a system of stochastic linear difference equations whose solution is given by (7) and where $\{x_t\}$ is cointegrated under the *parametric* conditions stated by Johansen. In contrast, Engle and Granger assume cointegration and show that there exists a VEC representation.

Using these results it is now possible to rewrite equation (7) as a common trends model. In particular, let

$$C(\lambda) = C(1) + (1 - \lambda) C^*(\lambda), \tag{8}$$

where $C^*(\lambda) = \sum_{i=0}^{\infty} C_i^* \lambda^i$ is absolutely summable and $C_i^* = -\sum_{j=i+1}^{\infty} C_j$ for $i \geq 0$ as shown by Stock (1987). Substituting equation (8) into (7) for $C(\lambda)$, recursively substituting for $x_{t-1}, \dots, x_1$, and letting $\varepsilon_s = 0$ for $s = 0$, we obtain

$$x_t = x_0 + C(1)\xi_t + C^*(L)\varepsilon_t. \tag{9}$$

For this (reduced form common) stochastic trends representation we have that $\xi_t = \rho + \xi_{t-1} + \varepsilon_t$ and $\delta = C(1)\rho$. In terms of equation (1) this means that

$$\begin{aligned} x_t^{\mathrm{s}} &= x_0 + C^*(L)\varepsilon_t, \\ x_t^{\mathrm{p}} &= C(1)\left[\xi_0 + \rho t + \textstyle\sum_{j=1}^{t} \varepsilon_j\right]. \end{aligned} \tag{10}$$

Stock and Watson (1988) derive the common trends model in (2) from (9). Leaving the algebra aside for now, the basic idea is simply to use the fact that $C(1)$ has a reduced rank under the assumption of cointegration. Accordingly, only $k = n - r$ elements of $C(1)\varepsilon_t$ result in (linearly) independent permanent effects on x_t. In fact, from (4) and (10) we find that the equality of the trend components imply that

$$\Upsilon\varphi_t = C(1)\varepsilon_t, \quad \Upsilon\Upsilon' = C(1)\Sigma C(1)', \quad \text{and} \quad \Upsilon\mu = C(1)\rho. \tag{11}$$

When one is concerned with estimating the loading matrix, Υ, of the common trends model in (2), it is clear that we need to have information about the parameters of $C(1)$ and Σ. While Σ can be estimated directly from (5) or (6), to obtain an estimate of $C(1)$ we must know how to invert the VEC representation.

Campbell and Shiller (1988) show that it is straightforward to rewrite the VEC representation as a restricted VAR system when $n = 2$ and $r = 1$. To generalize their result, let M be an $n \times n$ *nonsingular* matrix given by $[S_k' \ \ \alpha]'$, where the rows of the $k \times n$ selection matrix S_k satisfy $S_{i,k}C(1) \neq 0$ for all $i \in \{1, \dots, k\}$. Also, let γ^* be an $n \times n$ matrix equal to $[0 \ \ \gamma]$, while the $n \times n$ matrix polynomials $D(\lambda)$ and $D_\perp(\lambda)$ are

$$D(\lambda) := \begin{bmatrix} I_k & 0 \\ 0 & (1-\lambda)I_r \end{bmatrix}, \qquad D_\perp(\lambda) := \begin{bmatrix} (1-\lambda)I_k & 0 \\ 0 & I_r \end{bmatrix},$$

Next, let $\theta := M\rho$ and $\eta_t := M\varepsilon_t$. We can now derive a VAR representation for x_t which is conditioned on the cointegration vectors. We shall call this representation a restricted VAR.

Premultiplying both sides of equation (6) by M we get

$$MA^*(L)\Delta x_t = \theta - M\gamma z_{t-1} + \eta_t.$$

Define the n dimensional random variable y_t from $y_t := D_\perp(L)Mx_t$. Noting that $(1-\lambda)I_n = D(\lambda)D_\perp(\lambda)$ and $\gamma z_t = \gamma^* y_t$, we can express this system as

$$B(L)y_t = \theta + \eta_t, \tag{12}$$

where

$$B(\lambda) := M\left[A^*(\lambda)M^{-1}D(\lambda) + \gamma^*\lambda\right].$$

Note that $B(0) = I_n$ and that the matrix polynomial $B(\lambda)$ is (at most) of order p. The following version of GRT turns out to be very useful in the coming analysis of common trends:

Theorem 1. (Granger's Representation Theorem) *Suppose $\{x_t\}$ is generated according to* (5) *with* $\text{rank}[A(1)] = r < n$ *and* $\det[B(\lambda)] = 0$ *if and only if* $|\lambda| > 1$*, then $\{y_t\}$, $\{z_t\}$, and $\{\Delta x_t\}$ are integrated of order zero. In addition,*

$$A(\lambda) = M^{-1}B(\lambda)D_\perp(\lambda)M, \tag{13}$$

and

$$C(\lambda) = M^{-1}D(\lambda)B(\lambda)^{-1}M. \tag{14}$$

Note that the rank condition ensures that $\{x_t\}$ is *not* integrated of order zero. The determinant condition, on the other hand, means that y_t in (12) has an invertible Wold moving average representation and, accordingly, $\{y_t\}$ (and thus $\{z_t\}$) is integrated of order zero. The rank condition then implies that $\{x_t\}$ is integrated of order one. Premultiplication by M^{-1} in (12) and using the definitions of y_t, θ and η_t gives us the expression in (13). Similarly, $C(\lambda)$ is obtained by premultiplying

$$y_t = B(1)^{-1}\theta + B(L)^{-1}\eta_t,$$

by $M^{-1}D(\lambda)$ and using the same definitions and the property that $(1-\lambda)I_n = D(\lambda)D_\perp(\lambda)$.

In a sense, Theorem 1 summarizes all we need to know about the (reduced form) mathematical properties of a vector time series which is cointegrated of order (1,1) with cointegrating rank r. The matrix polynomial $B(\lambda)$ captures the general 'short run' dynamics, whereas $(D_{\perp}(\lambda), D(\lambda))$ and M represent integration and cointegration, respectively. Furthermore, the restricted VAR may be used as a convenient data generating process for testing linear rational expectations models when the time series of interest are cointegrated (cf. Baillie (1989), Campbell and Shiller (1987), and Warne (1990)). However, for my purpose here, a more important result is that we have found a simple mathematical connection to the Wold VMA representation.[1] Hence, the restricted VAR in (12) is very well suited for estimating a common trends model.[2]

2.2. **Estimation of the Common Trends Parameters.** From Theorem 1 it follows that the lag order of the restricted VAR in (12) is never greater than that of the unrestricted VAR in (5). In fact, unless all elements in the final r columns of the matrix A_p are zero the restricted VAR is also of order p. Hence, let us consider $B(\lambda) = I_n - \sum_{k=1}^{p} B_k \lambda^k$. Furthermore, the Theorem establishes that the matrix $C(1)$ is equal to $M^{-1}D(1)F(1)M$, where $F(1)$ is the inverse of $B(1)$. It then follows that if M, $\Omega := M\Sigma M' (= \mathrm{E}[\eta_t \eta_t'])$ and $B(1)$ were known, so would Σ and $C(1)$ be.

The space spanned by the rows of α' may be estimated and analysed by applying the maximum likelihood based methods developed in Johansen (1988b,1989,1991) and Johansen and Juselius (1990). Another possibility is to let these parameters be determined by the steady state of an appropriate economic theory (cf. KPSW (1987) and Mellander, Vredin and Warne (1992)). In both cases, knowledge of these parameters suffices for the purpose of determining the matrices M and $D_{\perp}(\lambda)$, needed to construct the vector time series $\{y_t\}$. Furthermore, consistent and asymptotically efficient estimates of the parameters in (12) may be obtained from e.g. Gaussian maximum likelihood estimation of y_t on a constant and p lags.[3]

[1] KPSW make use of the VEC representation in their study. Based on the results in Theorem 1 it can be shown that $C(\lambda) = M^{-1}D(\lambda)[A^*(\lambda)M^{-1}D(\lambda) + \gamma^*\lambda]^{-1}$; see also Johansen (1991) and Lütkepohl and Reimers (1992).

[2] It may be noted that from a purely mathematical point of view it is *always* possible to consider a selection matrix S_k of the form $S_k = [I_k \;\; 0]$ when the cointegrating vectors are known. The reason for this is that $\mathrm{rank}[\alpha] = r$ so that the components of x_t can be ordered to ensure that the last r columns of α' is an invertible matrix. In fact, we can let $\alpha' = [\alpha_k' \;\; I_r]$, where α_k' is an $r \times k$ matrix. It is now easily established that $\alpha' C(1) = 0$.

[3] Note that the asymptotic properties of, e.g. $C(1)$ are *independent* of how α has been determined. A consistently estimated and a known α matrix will always be associated with identical asymptotic covariance matrices for the estimate of $C(1)$ as long as the corresponding M matrix converges in probability to a nonsingular (and true) matrix at the rate $\sqrt{T}$ (where T is the sample size). For example, Johansen's ML estimator of the cointegration vectors converges in probability at all rates less than T. Moreover,

The next step is to calculate the matrix of common trends parameters. To identify these parameters one may proceed along the route suggested by KPSW (see also the proof of Theorem 4.1 in Johansen (1991) for an alternative approach). That is, when $\{x_t\}$ has k common stochastic trends, we may write the matrix Υ as

$$\Upsilon = \Upsilon_0 \pi, \tag{15}$$

where Υ_0 is an $n \times k$ matrix with known parameters, chosen so that $\alpha'\Upsilon_0 = 0$ and so that the innovations to the trends have an economic interpretation (we shall return to this issue below). The "free" parameters of Υ are lumped into the $k \times k$ matrix π. One problem is thus how we can determine π.

Using the relationship $\Upsilon\Upsilon' = C(1)\Sigma C(1)'$ and equation (15) we have that

$$\pi\pi' = (\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0' C(1)\Sigma C(1)'\Upsilon_0(\Upsilon_0'\Upsilon_0)^{-1}. \tag{16}$$

The right hand side of equation (16) is a $k \times k$ positive definite and symmetric matrix with $k(k+1)/2$ unique parameters. We cannot, however, solve for π *uniquely* without making some additional assumptions. For the above system of equations exactly $k(k+1)/2$ parameters can be uniquely determined, e.g. from a Choleski decomposition. Other procedures, such as a method of moments decomposition, may also be considered (cf. Bernanke (1986)).

It should be noted that although the Choleski decomposition of π indicates a recursive structure for the influence of τ_t on x_t, the choice of Υ_0 actually determines how the trends affect x_t. Thus, Υ need *not* represent any recursiveness for the common trends model. To summarize this discussion, to identify the nk parameters of Υ we first use the rk restrictions $\alpha'\Upsilon = 0$. Hence, there remains to determine k^2 parameters. Second, we can solve the $k(k+1)/2$ independent equations in $\pi\pi'$ if Υ_0 is known. Accordingly, in addition to the requirement $\alpha'\Upsilon_0 = 0$, $k(k-1)/2$ further restrictions must be imposed on Υ to achieve exact identification. These additional constraints should be motivated by economic theory since they cannot be tested.

the matrix S_k must be appropriately chosen. A suitable choice is to let $S_k = \alpha_\perp'$, where $\alpha'\alpha_\perp = 0$ and $\alpha_\perp'\alpha_\perp = I_k$. This choice of a selection matrix guarantees that the estimate of M converges to a non-singular matrix in probability (under suitable regularity conditions for the process $\{\varepsilon_t\}$). Note that the Johansen ML estimation procedure "automatically" provides an estimate of $\alpha_\perp$, i.e. those eigenvectors which are associated with the smallest $(n-r)$ eigenvalues in the cointegration rank testing scheme.

At this stage, it should be emphasized that this procedure for identifying the common trends parameters implies that the innovations to the common trends influence transient fluctuations in x_t as well as the growth path. To see this, note that $\varphi_t = (\Upsilon'\Upsilon)^{-1}\Upsilon' C(1)\varepsilon_t$. Consequently, the covariance matrix between φ_t and ε_t is

$$\mathrm{E}[\varphi_t \varepsilon_t'] = (\Upsilon'\Upsilon)^{-1}\Upsilon' C(1)\Sigma.$$

Obviously, this matrix is nonzero since the columns of Υ cannot be orthogonal to the columns of $C(1)$. It is precisely this fact which allows us to study connections between growth and transitory fluctuations.

2.3. **An Example.** Suppose we are interested in examining the interactions between the logarithms of real output ($\ln Y_t$), the price level ($\ln P_t$), and the money stock ($\ln M_t$). Furthermore, suppose we model $x_t = [\ln Y_t \;\, \ln P_t \;\, \ln M_t]'$ as being cointegrated of order (1,1) with one cointegrating vector. Let $\alpha' = [1 \;\, 1 \;\, -1]$, so that the logarithm of the velocity of money is integrated of order zero. Accordingly, each series is nonstationary in levels.[4]

In this case, we can estimate the restricted VAR in equation (12) with y_t given by $y_t = [\Delta \ln Y_t \;\, \Delta \ln P_t \;\, \ln V_t]'$, where $\ln V_t = (\ln Y_t + \ln P_t - \ln M_t)$. From estimates of the B_k matrices and of Ω we can then determine estimates of $C(1)$ and Σ. Furthermore, to estimate the 3×2 matrix of common trends parameters, Υ, we need to specify Υ_0. A choice, suggested by economic theory, is to let τ_t include a real (technology) and a nominal (monetary policy) trend, and assume that the nominal trend does not influence the long run growth path of $\ln Y_t$. Naturally, with $k = 2$ this is one conceivable additional restriction which Υ must satisfy.

[4]Note, however, it is *not* necessary that each time series in a common trends model is nonstationary. For example, the time series model studied by Blanchard and Quah (1989) can easily be fitted into a common trends framework. Since they examine real output and unemployment and model the former as first difference stationary and the latter as stationary, there is one cointegrating vector which assigns a nonzero coefficient to unemployment and a zero coefficient to real output. Then, the time series y_t is given by the first difference of real output and the level of unemployment. Furthermore, the matrix Υ is 2×1 with a nonzero coefficient in the output equation and a zero coefficient in the unemployment equation.

Given α and a lower triangular π matrix, the above discussion implies that

$$\Upsilon_0 = \begin{bmatrix} \Upsilon_{0,11} & 0 \\ \Upsilon_{0,21} & \Upsilon_{0,22} \\ \Upsilon_{0,11} + \Upsilon_{0,21} & \Upsilon_{0,22} \end{bmatrix}.$$

It can be verified that the choice of $\Upsilon_{0,21}$ does *not* influence the matrix Υ, whereas $\Upsilon_{0,11}$ and $\Upsilon_{0,22}$ must be nonzero. Therefore, let $\Upsilon_{0,21} = 0$ and $\Upsilon_{0,11} = \Upsilon_{0,22} = 1$.[5] With π being lower triangular, the general form of Υ is then

$$\Upsilon = \begin{bmatrix} \pi_{11} & 0 \\ \pi_{21} & \pi_{22} \\ \pi_{11} + \pi_{21} & \pi_{22} \end{bmatrix},$$

where π_{ij} denotes the (i,j):th element of π. From Υ it can be seen that the first element of τ_t is the real trend, while the second element is the nominal trend. Furthermore, the matrix Υ satisfies three restrictions and thus has three free parameters. That is, $\Upsilon_{12} = 0$ is a restriction suggested by economic theory while $\Upsilon_{11} + \Upsilon_{21} = \Upsilon_{31}$, and $\Upsilon_{22} = \Upsilon_{32}$ are implied by the two restrictions from cointegration. The common trends parameters are then exactly identified.

3. Inversion and Identification

3.1. **The VMA Parameters.** The VMA representation in equation (7) is a natural starting point for analysing some dynamic properties of a vector time series $\{x_t\}$ with k common trends. The central issues for performing impulse response analysis and forecast error variance decompositions are those of (a) calculating the sequence of matrices $\{C_j\}_{j=1}^{\infty}$, (b) identifying the innovations to the system, and (c) computing standard errors for the estimated impulse responses and variance decompositions. Here, I shall focus on (a), whereas (b) and (c) are studied in sections 3.2 and 4, respectively. It can be

[5]To verify this claim, let ϕ be a $k \times k$ matrix such that $\Upsilon = \Upsilon_0 \phi^{-1} \phi \pi = \tilde{\Upsilon}_0 \tilde{\pi}$. Since π is lower triangular it follows that ϕ must also be lower triangular for $\tilde{\pi} := \phi\pi$ to be lower triangular. Standard matrix theory tells us that the inverse of a lower triangular matrix is lower triangular. Thus, the matrix $\tilde{\Upsilon}_0 := \Upsilon_0 \phi^{-1}$ can be constructed in any way we desire as long as $\alpha' \Upsilon_0 = 0$ and ϕ is lower triangular. In the example, we may consider a lower triangular ϕ matrix whose diagonal elements are given by $\phi_{ii} = \Upsilon_{0,ii}$ for $i \in \{1,2\}$, thus the nonzero requirement. Also, $\phi_{21} = \Upsilon_{0,21}$ is permissible. If π is not lower triangular it becomes more difficult to determine what kind of ϕ matrices that may be considered.

noted that we will abstract from non–Wold moving average representations (see Lippi and Reichlin (1990)).

From Theorem 1 we find that the $C(\lambda)$ polynomial is equal to $M^{-1}D(\lambda)F(\lambda)M$, and $F(\lambda)$ is the inverse of $B(\lambda)$. Letting $F(\lambda) = I_n + \sum_{j=1}^{\infty} F_j\lambda^j$, and using the above relationship, we find that

$$C(\lambda) = I_n + \sum_{j=1}^{\infty} M^{-1}(F_j - DF_{j-1})M\lambda^j, \tag{17}$$

where $F_0 = I_n$ and the $n \times n$ matrix D is defined from $D(\lambda) = I_n - D \cdot \lambda$. It may be noted that $D = D_\perp(1)$ is idempotent, a property we shall use below.

One algorithm for performing the inversion is obtained by stacking equation (12) into a first order system of the form

$$\begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix} = \begin{bmatrix} \theta \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} B_1 & B_2 & \cdots & B_{p-1} & B_p \\ I_n & 0 & \cdots & 0 & 0 \\ \vdots & \ddots & & \vdots & \vdots \\ 0 & 0 & & I_n & 0 \end{bmatrix} \cdot \begin{bmatrix} y_{t-1} \\ y_{t-2} \\ \vdots \\ y_{t-p} \end{bmatrix} + \begin{bmatrix} \eta_t \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

or

$$Y_t = \Theta + BY_{t-1} + N_t. \tag{18}$$

Since $\det[B(\lambda)] = 0$ has all solutions outside the unit circle, it follows that the eigenvalues of B are inside the unit circle. Accordingly, $\lim_{s\to\infty} B^s = 0$ so that the solution to the system of stochastic difference equations in (18) is

$$Y_t = \sum_{j=0}^{\infty} B^j\Theta + \sum_{j=0}^{\infty} B^j N_{t-j}. \tag{19}$$

Defining the $n \times np$ matrix J_p as $[I_n \;\; 0 \cdots 0]$, we find that $y_t = J_pY_t$, $\Theta = J_p'\theta$, and $N_t = J_p'\eta_t$. Hence, the solution to equation (12) in terms of current and past realizations of η_t can be written as

$$y_t = \sum_{j=0}^{\infty} J_pB^jJ_p'\theta + \sum_{j=0}^{\infty} J_pB^jJ_p'\eta_{t-j}, \tag{20}$$

where $F(1) = \sum_{j=0}^{\infty} J_pB^jJ_p'$. Finally, by equation (17) we get

$$C_j = M^{-1}J_pB^jJ_p'M - M^{-1}DJ_pB^{j-1}J_p'M, \quad j = 1, 2, \ldots \tag{21}$$

where $C_0 = I_n$. Given estimates of the B_k matrices and of M, it is thus straightforward to estimate the Wold VMA parameters in C_j.

3.2. **Identification of Permanent and Transitory Innovations.** In section 2.2 we identified the long run coefficients on the k common trends innovations. My objective is now to be more specific about identification of all parameters in the common trends model. In particular, it is important to be precise about identification in the sense that implications from impulse response functions and forecast error variance decompositions are fully consistent with the common trends model. Before we come to that, however, two definitions and some new notation is introduced to minimize ambiguities.

Let Γ be any $n \times n$ nonsingular matrix such that $\Gamma\Sigma\Gamma'$ is diagonal. The matrix $R(1) = C(1)\Gamma^{-1}$ is called the total impact matrix. Also, let ν_{it} be the i:th component of the vector $\Gamma\varepsilon_t$.

Definition 1. *An $n \times n$ matrix Γ is said to identify a common trends model if* (i) *it is uniquely determined from the parameters of the model in equation* (6)*,* (ii) *the covariance matrix of $\Gamma\varepsilon_t$ is diagonal with nonzero diagonal elements, and* (iii) *the total impact matrix is given by $R(1) = [\Upsilon \;\; 0]$.*

Definition 2. *An innovation ν_{it} is said to be permanent (transitory) if the i:th column of the total impact matrix is nonzero (zero).*

From these two definitions it follows that if an $n \times n$ matrix Γ identifies a common trends model, then the permanent innovations are those which are associated with the common trends.

Let the $n \times n$ nonsingular matrix Γ be chosen so that (i) the permanent innovations are equal to φ_t, (ii) the permanent and the transitory innovations, ψ_t, are independent, and (iii) the transitory innovations are mutually independent. We then have that

$$\Delta x_t = \delta + C(L)\varepsilon_t = \delta + R(L)\nu_t, \tag{22}$$

where $R(\lambda) = C(\lambda)\Gamma^{-1}$, $\nu_t = \Gamma\varepsilon_t$, and $\mathrm{E}[\nu_t\nu_t'] = I_n$. The component $R(L)\nu_t$ in equation (22) is called the impulse response function of Δx_t.

In order to derive a suitable matrix Γ, it may first be noted that

$$\nu_t = \begin{bmatrix} \varphi_t \\ \psi_t \end{bmatrix} = \begin{bmatrix} \Gamma_k \\ \Gamma_r \end{bmatrix} \varepsilon_t = \Gamma\varepsilon_t, \tag{23}$$

where Γ_k and Γ_r are $k \times n$ and $r \times n$ matrices, respectively. It has already been established that $\Upsilon\varphi_t = C(1)\varepsilon_t$ and that Υ as well as $C(1)$ had rank equal to k. Hence, it follows that the permanent innovations may be described by

$$\varphi_t = (\Upsilon'\Upsilon)^{-1}\Upsilon' C(1)\varepsilon_t, \tag{24}$$

and, accordingly, the top $k \times n$ matrix Γ_k in (23) is $(\Upsilon'\Upsilon)^{-1}\Upsilon' C(1)$.

To find a matrix Γ_r which satisfies the conditions (ii) φ_t and ψ_t are independent, and (iii) the components of ψ_t are mutually independent, we may either make use of a Jordan decomposition of some suitable matrix or mathematically related schemes (cf. Stock and Watson (1988)). I shall first consider condition (ii). Evaluating the covariance between the permanent and transitory innovations, we find that

$$\mathrm{E}[\varphi_t\psi_t'] = (\Upsilon'\Upsilon)^{-1}\Upsilon' C(1)\Sigma\Gamma_r'. \tag{25}$$

For this $k \times r$ matrix to be zero, it seems natural to let Γ_r include Σ^{-1}. That allows us to focus on the matrix $C(1)$, which is known to have reduced rank. From linear algebra it is well known that there exists exactly r linearly independent vectors which are orthogonal to the rows of $C(1)$. Letting $\Gamma_r = H_r\Sigma^{-1}$, we are therefore seeking an $r \times n$ matrix H_r such that $C(1)H_r' = 0$.

One possibility is to consider the space spanned by the columns of γ. From the properties of the $A(1)$ and $C(1)$ matrices, we have that $C(1)\gamma = 0$. In fact, the following relationship may be established

$$\gamma = M^{-1}B(1)D_{\perp}(1)M\alpha(\alpha'\alpha)^{-1} = M^{-1}B(1)P_r, \tag{26}$$

where P_r is the $n \times r$ matrix determined from $D = [0 \;\; P_r]$, i.e. $P_r = [0 \;\; I_r]'$.[6] Premultiplying γ in equation (26) by $C(1)$, we find that

$$
\begin{aligned}
C(1)\gamma &= (M^{-1}D(1)F(1)M) \cdot (M^{-1}B(1)D_{\perp}(1)M\alpha(\alpha'\alpha)^{-1}) \\
&= 0,
\end{aligned}
$$

since D is idempotent, i.e. $D(1)D_{\perp}(1) = (I_n - D)D = 0$.

Let $H_r = Q^{-1}\zeta'$, where Q is an $r \times r$ matrix, $\zeta = \gamma(U\gamma)^{-1}$, and U is an $r \times n$ matrix chosen so that $U\gamma$ is invertible (the specific use of U will be discussed below). The covariance matrix for the transitory innovations is then given by

$$
\mathrm{E}[\psi_t\psi_t'] = Q^{-1}\zeta'\Sigma^{-1}\zeta(Q')^{-1}. \tag{27}
$$

In order to ensure that this matrix is compatible with the assumption of ψ_t being mutually independent, Q must be chosen such that $\zeta'\Sigma^{-1}\zeta$ is diagonalized. A convenient normalization is then to let $\mathrm{E}[\psi_t\psi_t'] = I_r$. The transitory innovations are now determined from

$$
\psi_t = Q^{-1}\zeta'\Sigma^{-1}\varepsilon_t. \tag{28}
$$

Accordingly, the matrix Γ_r is given by $Q^{-1}\zeta'\Sigma^{-1}$ so that the matrix Γ becomes

$$
\Gamma = \begin{bmatrix} (\Upsilon'\Upsilon)^{-1}\Upsilon'C(1) \\ Q^{-1}\zeta'\Sigma^{-1} \end{bmatrix}. \tag{29}
$$

It may be noted that the k linearly independent rows of Γ_k are linearly independent to the r linearly independent rows of Γ_r. These properties imply that Γ is of full rank (cf. Theorem 3.19 in Magnus and Neudecker (1988)).

We are now in a position to state the following result concerning the properties of the matrix Γ.

Theorem 2. *If the n dimensional vector time series $\{x_t\}$ satisfies the assumptions in Theorem 1, then the $n \times n$ nonsingular matrix Γ in equation* (29) *identifies a common*

[6] This is easily established by noting that $M\alpha(\alpha'\alpha)^{-1} = [(\alpha'\alpha)^{-1}\alpha' S_k' \;\; I_r]'$. Premultiplying by D we obtain P_r. Alternatively, from Theorem 1 we know that $B(1) = M[A^*(1)M^{-1}D(1)+\gamma^*]$. Premultiplying by M^{-1} and postmultiplying by P_r, we find that $M^{-1}B(1)P_r = \gamma^* P_r$. Since $\gamma^* = [0 \;\; \gamma]$, it follows that $\gamma^* P_r = \gamma$.

trends model, i.e.

$$R(1) = C(1)\Gamma^{-1} = [\Upsilon \;\; 0], \tag{30}$$

and $\Gamma\Sigma\Gamma' = I_n$. *Furthermore,*

$$\Gamma^{-1} = \begin{bmatrix} \Gamma_k^+ & \Gamma_r^+ \end{bmatrix} = \begin{bmatrix} \Sigma C(1)'\Upsilon(\Upsilon'\Upsilon)^{-1} & \zeta(Q')^{-1} \end{bmatrix}. \tag{31}$$

From the conclusions in this Theorem it is straightforward to derive the common trends model in (2) from the reduced form representation in (9). For the trend components we have that

$$C(1)\xi_t = R(1)\left[\Gamma\xi_0 + \Gamma\rho t + \sum_{j=1}^{t} \nu_j\right] = \Upsilon\tau_t,$$

since $\mu = \Gamma_k\rho$. Also, from the stationary components we get $C^*(L)\varepsilon_t = C^*(L)\Gamma^{-1}\nu_t$ so that $\Phi(\lambda) = C^*(\lambda)\Gamma^{-1}$.

The matrix U can be used to give the transitory disturbances an economic interpretation. Suppose we want to identify the transitory innovations based on their contemporaneous relation to Δx (or to x). In that case, with $R(0) = \Gamma^{-1}$ it follows that we should impose restrictions on Γ_r^+. For example, given γ and Q the matrix U can always be chosen so that $r(r-1)/2$ elements in $\gamma(U\gamma)^{-1}(Q')^{-1}$ are zero. Now suppose Q is lower triangular and that $n = 4$ and $r = 3$. Letting q_{ij}^+ and ζ_{ij} denote the (i, j):th elements of Q^{-1} and ζ, respectively, we have that

$$\Gamma_r^+ = \begin{bmatrix} \zeta_{11}q_{11}^+ & \sum_{j=1}^{2}\zeta_{1j}q_{2j}^+ & \sum_{j=1}^{3}\zeta_{1j}q_{3j}^+ \\ \zeta_{21}q_{11}^+ & \sum_{j=1}^{2}\zeta_{2j}q_{2j}^+ & \sum_{j=1}^{3}\zeta_{2j}q_{3j}^+ \\ \zeta_{31}q_{11}^+ & \sum_{j=1}^{2}\zeta_{3j}q_{2j}^+ & \sum_{j=1}^{3}\zeta_{3j}q_{3j}^+ \\ \zeta_{41}q_{11}^+ & \sum_{j=1}^{2}\zeta_{4j}q_{2j}^+ & \sum_{j=1}^{3}\zeta_{4j}q_{3j}^+ \end{bmatrix}.$$

To exactly identify the transitory innovations we need to consider three restrictions on this matrix. A simple procedure is to let certain elements of ζ be equal to zero, say, ζ_{11}, ζ_{12}, and ζ_{21}. This, however, requires that the first and second and either the third or fourth row of γ are linearly independent. If this is true we can let the 3×4 matrix U be given by a zero–one matrix which appropriately normalizes γ. Accordingly, the first transitory innovation has a zero contemporaneous effect on the first two elements of Δx (and x), while the second transitory innovation has a zero contemporaneous effect on the first element of Δx. On the other hand, if the first and second rows of γ are linearly dependent another identification procedure has to be considered.

It should be emphasized that a lower triangular Q matrix is neither necessary nor sufficient for identification of the transitory disturbances. Rather, it is the relationship between the space spanned by the columns of γ and the mathematical structure of Q that jointly determine which class of identification procedures that are allowed. This is analogous to the identification of the permanent shocks in that Υ is a normalisation of the space spanned by the columns of $\alpha_\perp$ and the choice of Υ_0 and structure of π specify the implications of the identifying assumptions in terms of Υ. Furthermore, it is straightforward to test whether the desired identification scheme of the transitory disturbances is consistent with data. By equation (26) we know that $\gamma = M^{-1}B(1)P_r$. Thus, restriction on γ are immediately transformed into restrictions on the B_k matrices and standard classical tests can thus be applied (see Englund, Vredin, and Warne (1992) for an application).

4. Inference in the Common Trends Model

In this section we shall analyse the asymptotic properties of estimated impulse response functions and forecast error variance decompositions. The maintained hypothesis is that a finite upper bound for the lag order, $\bar{p}$, of the VAR system in (5) is known. Furthermore, we assume that the VAR representation is not misspecified. For a recent study of the consequences of misspecification in VAR models, see Braun and Mittnik (1993).

4.1. **Impulse Response Functions.** If the lag order of a covariance stationary VAR is finite and known or if an upper bound for the order is known, results from Baillie (1981,1987), Schmidt (1973,1974), and others can be applied to obtain the asymptotic distribution of the estimated Wold VMA and of the impulse response parameters. Also, the case of unknown and possibly infinite lag order is examined by Lütkepohl (1988,1990) and Lütkepohl and Poskitt (1990). Furthermore, the asymptotic distribution for the impulse responses in a cointegrated framework with Gaussian innovations is analysed in Lütkepohl and Reimers (1992). Unfortunately, neither of these approaches can *fully* be applied in the present setting. In particular, while the results in Lütkepohl and Reimers[7] imply that some innovations will have a permanent effect on some components of x_t,

[7] They consider a Choleski decomposition of Σ, the covariance matrix of ε_t.

permanent and transitory innovations in the sense implied by the common trends model have generally *not* been identified.

Without loss of generality, suppose Q is lower triangular, i.e. the transitory innovations are based on a Choleski decomposition of $\zeta'\Sigma^{-1}\zeta$.[8] Let vec denote the column stacking operator for any matrix, vech the corresponding operator (primarily used for symmetric and lower triangular matrices) that only stacks elements on and below the diagonal. The Kronecker product is denoted by $\otimes$, the $mn \times mn$ commutation matrix K_{mn} is defined such that for any $m \times n$ matrix G, $K_{mn}\text{vec}(G) = \text{vec}(G')$, and the $m^2 \times m^2$ matrix $N_m = \frac{1}{2}(I_{m^2} + K_{mm})$. The $m^2 \times m(m+1)/2$ duplication matrix D_m is defined such that $D_m\text{vech}(G) = \text{vec}(G)$ for any symmetric $m \times m$ matrix G. The Moore–Penrose inverse of D_m is given by $D_m^+ = (D_m'D_m)^{-1}D_m'$, while the $m(m+1)/2 \times m^2$ elimination matrix L_m is defined such that for any $m \times m$ matrix G, $\text{vech}(G) = L_m\text{vec}(G)$. In addition, if G is lower triangular, then $\text{vec}(G) = L_m'\text{vech}(G)$; see Henderson and Searle (1979), Magnus and Neudecker (1985,1986,1988), and Neudecker (1983). Finally, let $\xrightarrow{\text{p}}$ and $\xrightarrow{\text{d}}$ denote convergence in probability and in distribution, respectively, T the sample size for the estimated parameters, which are denoted by a hat, while $\mathbb{N}$ denotes the (multivariate) normal distribution.

Suppose that we shock Δx_t at $t = t^*$ by a one standard deviation change in ν_{t^*}. The dynamic responses in Δx_{t^*+s} are then given by

$$\text{resp}(\Delta x_{t^*+s}) = R_s, \tag{32}$$

where $\text{resp}(\Delta x_{\text{inf}}) = \lim_{s\to\infty} \text{resp}(\Delta x_{t^*+s}) = 0$. Similarly, the responses in the levels, x_{t^*+s}, are given by

$$\text{resp}(x_{t^*+s}) = \sum_{j=0}^{s} R_j, \tag{33}$$

where $\text{resp}(x_{\text{inf}}) = \lim_{s\to\infty} \text{resp}(x_{t^*+s}) = R(1) = [\Upsilon\ 0]$. To estimate these impulse response functions we replace R_s with $\hat{R}_s = \hat{C}_s\hat{\Gamma}^{-1}$.

For the purpose of deriving asymptotic distributions of functions of estimated parameters a result from Serfling (1980, Theorem 3.3.A) is employed. Let $\phi \in \mathbb{R}^m$ be a vector

[8] Results similar to those obtained below hold with other Q and π matrices satisfying $QQ' = \zeta'\Sigma^{-1}\zeta$ and $\pi\pi' = (\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'C(1)\Sigma C(1)'\Upsilon_0(\Upsilon_0'\Upsilon_0)^{-1}$, respectively.

of parameters and $\hat{\phi}$ a consistent estimator of ϕ_0, the true value of ϕ, such that

$$\sqrt{T}\left(\hat{\phi}-\phi_0\right) \xrightarrow{\mathrm{d}} \mathbb{N}(0, V_\phi).$$

Suppose $f(\phi)$ is a continuously differentiable function which maps ϕ into $\mathbb{R}^n$ and $\partial f_i/\partial\phi' \neq 0$ at ϕ_0 for all $i \in \{1,\dots,n\}$. Then,

$$\sqrt{T}\left(f(\hat{\phi})-f(\phi_0)\right) \xrightarrow{\mathrm{d}} \mathbb{N}(0, V_f),$$

where

$$V_f = \frac{\partial f}{\partial \phi'} V_\phi \frac{\partial f}{\partial \phi},$$

and $(\partial f/\partial\phi)$ is the transpose of $(\partial f/\partial\phi')$. This well known result is vital in the following analysis. Hence, if $\hat{\phi}$ is a consistent estimator of ϕ and converges at the rate $\sqrt{T}$ to a joint asymptotic normal distribution, then $f(\hat{\phi})$ has similar properties. Accordingly, only $(\partial f/\partial\phi')$ has to be derived.

We are now ready to state the main results regarding analytical expressions for the asymptotic distributions of the impulse response functions.

Theorem 3. (i) *If the lag order in equation* (12) *has a finite upper bound* $\bar{p} \leq p$, (ii) $\{x_t\}$ *is cointegrated of order (1,1) with* r *cointegration vectors,* (iii) *the estimated cointegration vectors satisfy* $\sqrt{T}(\hat{\alpha}-\alpha) \xrightarrow{\mathrm{p}} 0$, (iv)

$$\sqrt{T}\begin{bmatrix} \hat{\beta}-\beta \\ \hat{\omega}-\omega \end{bmatrix} \xrightarrow{\mathrm{d}} \mathbb{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} V_\beta & 0 \\ 0 & V_\omega \end{bmatrix}\right),$$

where $\beta = \mathrm{vec}(J_p B)$ *and* $\omega = \mathrm{vech}(\Omega)$, *and* (v) *the matrix* Γ *is given by equation* (29), *with* π *and* Q *being lower triangular matrices, then*

$$\sqrt{T}\left(\mathrm{vec}(\hat{R}_j)-\mathrm{vec}(R_j)\right) \xrightarrow{\mathrm{d}} \mathbb{N}(0, V_{R_j}), \tag{34}$$

for $j = 0,1,2,\dots,$ *where*

$$V_{R_j} = \frac{\partial \mathrm{vec}(R_j)}{\partial \beta'} V_\beta \frac{\partial \mathrm{vec}(R_j)}{\partial \beta} + \frac{\partial \mathrm{vec}(R_j)}{\partial \omega'} V_\omega \frac{\partial \mathrm{vec}(R_j)}{\partial \omega},$$

and

$$\frac{\partial \mathrm{vec}(R_j)}{\partial \beta'} = \left[(\Gamma^{-1})' \otimes I_n\right] \frac{\partial \mathrm{vec}(C_j)}{\partial \beta'} + [I_n \otimes C_j] \begin{bmatrix} \partial \mathrm{vec}(\Gamma_k^{+})/\partial \beta' \\ \partial \mathrm{vec}(\Gamma_r^{+})/\partial \beta' \end{bmatrix},$$

$$\frac{\partial \text{vec}(C_j)}{\partial \beta'} = \sum_{k=0}^{j-1} \left[M' J_p (B')^{j-1-k} \otimes M^{-1} F_k\right] - \sum_{k=0}^{j-2} \left[M' J_p (B')^{j-2-k} \otimes M^{-1} D F_k\right],$$

where the second term on the right hand side is zero for $j \leq 1$ and the first term is zero for $j = 0$, while

$$\frac{\partial \text{vec}(R_j)}{\partial \omega'} = [I_n \otimes C_j] \begin{bmatrix} \partial \text{vec}(\Gamma_k^+)/\partial \omega' \\ \partial \text{vec}(\Gamma_r^+)/\partial \omega' \end{bmatrix},$$

Furthermore, letting $E_p = [I_n \ \cdots \ I_n]$ be an $n \times np$ matrix we have

$$\begin{aligned} \partial \text{vec}(\Gamma_k^+)/\partial \beta' \;=\;& K_{nk}[M^{-1}\Omega F(1)' E_p \otimes (\Upsilon'\Upsilon)^{-1}\Upsilon' M^{-1} D(1) F(1)] - \\ & [\pi^{-1} \otimes \Sigma C(1)'\Upsilon(\Upsilon'\Upsilon)^{-1}] K_{kk} L_k' \{L_k N_k (\pi \otimes I_k) L_k'\}^{-1} D_k^+ \times \\ & [(\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0' C(1) M^{-1}\Omega F(1)' E_p \otimes (\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0' M^{-1} D(1) F(1)], \end{aligned}$$

$$\begin{aligned} \partial \text{vec}(\Gamma_r^+)/\partial \beta' \;=\;& \{[Q^{-1} \otimes \zeta(Q')^{-1}] K_{rr} L_r' \{L_r N_r (Q \otimes I_r) L_r'\}^{-1} D_r^+ [I_r \otimes \zeta'\Sigma^{-1}] - \\ & [Q^{-1} \otimes I_n]\}\{[(\gamma' U')^{-1} \otimes I_n] - [(\gamma' U')^{-1} \otimes \zeta U]\}[P_r' E_p \otimes M^{-1}], \end{aligned}$$

whereas

$$\begin{aligned} \partial \text{vec}(\Gamma_k^+)/\partial \omega' \;=\;& [(\Upsilon'\Upsilon)^{-1}\Upsilon' C(1) M^{-1} \otimes M^{-1}] D_n - \tfrac{1}{2}[\pi^{-1} \otimes \Sigma C(1)'\Upsilon(\Upsilon'\Upsilon)^{-1}] K_{kk} \times \\ & L_k' \{L_k N_k (\pi \otimes I_k) L_k'\}^{-1} D_k^+ [(\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0' C(1) M^{-1} \\ & \otimes (\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0' C(1) M^{-1}] D_n, \end{aligned}$$

and

$$\begin{aligned} \partial \text{vec}(\Gamma_r^+)/\partial \omega' \;=\;& \tfrac{1}{2}[Q^{-1} \otimes \zeta(Q')^{-1}] K_{rr} L_r' \{L_r N_r (Q \otimes I_r) L_r'\}^{-1} D_r^+ \times \\ & [\zeta' M' \Omega^{-1} \otimes \zeta' M' \Omega^{-1}] D_n. \end{aligned}$$

The asymptotic distributions for the accumulated response functions are given by

$$\sqrt{T}\left(\sum_{i=0}^{j} \text{vec}(\hat{R}_i) - \sum_{i=0}^{j} \text{vec}(R_i)\right) \xrightarrow{\text{d}} \mathbb{N}(0, V_{\Sigma R_j}), \tag{35}$$

for $j = 0, 1, 2, \ldots,$ where

$$V_{\Sigma R_j} = \left(\sum_{i=0}^{j} \frac{\partial \text{vec}(R_i)}{\partial \beta'}\right) V_\beta \left(\sum_{i=0}^{j} \frac{\partial \text{vec}(R_i)}{\partial \beta}\right) + \left(\sum_{i=0}^{j} \frac{\partial \text{vec}(R_i)}{\partial \omega'}\right) V_\omega \left(\sum_{i=0}^{j} \frac{\partial \text{vec}(R_i)}{\partial \omega}\right),$$

and finally,

$$\sqrt{T}\left(\mathrm{vec}(\hat{\Upsilon})-\mathrm{vec}(\Upsilon)\right)\xrightarrow{\mathrm{d}}\mathbb{N}(0,V_{\Upsilon}),\tag{36}$$

where

$$V_{\Upsilon}=\frac{\partial\mathrm{vec}(\Upsilon)}{\partial\beta'}V_{\beta}\frac{\partial\mathrm{vec}(\Upsilon)}{\partial\beta}+\frac{\partial\mathrm{vec}(\Upsilon)}{\partial\omega'}V_{\omega}\frac{\partial\mathrm{vec}(\Upsilon)}{\partial\omega},$$

while the remaining matrices of partial derivatives are given by

$$\begin{aligned}\partial\mathrm{vec}(\Upsilon)/\partial\beta' &= [I_k\otimes\Upsilon_0]\,L_k'\{L_kN_k(\pi\otimes I_k)L_k'\}^{-1}D_k^{+}\times\\ &\quad [(\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0C(1)M^{-1}\Omega F(1)'E_p\otimes(\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'M^{-1}D(1)F(1)],\end{aligned}$$

and

$$\begin{aligned}\partial\mathrm{vec}(\Upsilon)/\partial\omega' &= \tfrac{1}{2}\,[I_k\otimes\Upsilon_0]\,L_k'\{L_kN_k(\pi\otimes I_k)L_k'\}^{-1}D_k^{+}\times\\ &\quad [(\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'C(1)M^{-1}\otimes(\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'C(1)M^{-1}]D_n.\end{aligned}$$

A few remarks about the asymptotic covariance matrices for consistent and asymptotically normal estimators of β and ω may be of interest. Here, I shall only consider the multivariate least squares and Gaussian maximum likelihood estimators of the parameters. First, it is well known that such estimators of β and ω are asymptotically independent under standard assumptions. Moreover, if $\theta=0$, then the asymptotic covariance matrix for either estimator of β is given by $V_\beta=(\mathbb{I}_Y^{-1}\otimes\Omega)$, where $\mathbb{I}_Y:=p\lim T^{-1}\sum_{t=1}^{T}Y_{t-1}Y_{t-1}'$ is assumed to be nonsingular. If a constant is included in the restricted VAR, the expression for V_β looks similar to that above. In fact, letting $Y_t^*:=[1\ \ Y_t']'$, the $np\times(np+1)$ matrix $G=[0\ \ I_{np}]$, and $\mathbb{I}_{Y^*}:=p\lim T^{-1}\sum_{t=1}^{T}Y_{t-1}^*Y_{t-1}^{*\prime}$, we find that $V_\beta=(G\mathbb{I}_{Y^*}^{-1}G'\otimes\Omega)$. Second, if η_t is i.i.d. Gaussian, from the inverse of the information matrix we find that $V_\omega=2D_n^{+}(\Omega\otimes\Omega)D_n^{+\prime}$ (cf. Magnus and Neudecker (1986)). In practise, the asymptotic covariance matrices of $\hat{\beta}$ and $\hat{\omega}$ are, as usual, estimated by using consistent estimates of $\mathbb{I}_{Y^*}$ and Ω, respectively.

Second, it can be seen that (ζ,Q) does not influence the upper left $nk\times nk$ submatrices of V_{R_j} and $V_{\Sigma R_j}$. The diagonal elements of these matrices are the asymptotic variances of the responses in Δx_t and x_t, respectively, from a one standard deviation impulse to the permanent innovations. Similarly, (Υ_0,π) does not appear in the lower right $nr\times nr$ submatrices of these covariance matrices. Hence, identification of the permanent (transitory) innovations neither influence the estimated impulse responses nor the standard errors for the transitory (permanent) innovations. This is clearly a desirable property, whether we are primarily concerned with the permanent innovations or not.

Third, some of the asymptotic variances may be zero. For example, if an element of the k:th column of Υ_0 is constrained to zero, then the corresponding elements of $\hat{\Upsilon}$ and Υ are also zero. Although this is really not troublesome from a theoretical point of view, in applied work one has to be cautious to this possibility when e.g. t–ratios are calculated (cf. Lütkepohl (1990)).

Finally, it should be emphasized that estimators of the cointegration vectors generally converge in probability at the rate $\sqrt{T}$ (see Johansen (1991), Park (1992), and Stock (1987)). That is, if an estimator of α (and, hence, of M) is consistent, it also converges in probability at the rate $\sqrt{T}$. Accordingly, in terms of asymptotics, an estimator of the matrix M may be treated as if it is known. This is related to the result in Warne (1992) that the multivariate least squares estimator of the parameters in $A(\lambda)$ has identical limiting distributions for the cases when these parameters are estimated freely and under cointegration restrictions. Also, since the cointegration vectors determine the space spanned by the columns of Υ_0 it follows that for estimators of α which converge in probability at the rate $\sqrt{T}$ (to its true value), the corresponding estimator of Υ_0 converges in probability at the rate $\sqrt{T}$ as well.

4.2. **Forecast Error Variance Decompositions.** To my knowledge, the only papers which provide analytical expressions of the asymptotic distributions of estimated forecast error variance decompositions are those by Lütkepohl (1990) and Lütkepohl and Poskitt (1990). Lütkepohl shows that the asymptotic variances are quite easy to calculate when the lag order in a VAR model for covariance stationary time series has a known upper bound. In case the lag order is unknown and possibly infinite, then the asymptotic variances of forecast error variance decompositions can be derived under the assumptions made by Lütkepohl and Poskitt.

Before the main results are stated and proven, some additional notation will be useful. Let $v_{il,s}$ denote the fraction of the s steps ahead forecast error variance of Δx_i which is accounted for by shocks in ν_l, where $i, l \in \{1, \ldots, n\}$. Similarly, $v^*_{il,s}$ is the fraction of the s steps ahead forecast error variance of x_i which is accounted for by shocks in ν_l, whereas $\bar{v}_{il}$ denotes the long run fraction of the forecast error variance in the levels series x_i which is accounted for by shocks in ν_l. Accordingly,

$$v_{il,s} = \frac{\sum_{j=1}^{s}(e_i' R_{j-1} e_l)^2}{\sum_{j=1}^{s} e_i' R_{j-1} R_{j-1}' e_i},$$

for $i,l \in \{1,\ldots,n\}$ and $s = 1,2,\ldots$. Here, e_i is the i:th column of I_n. Furthermore, letting $R_m^* = \sum_{j=1}^{m} R_{j-1}$, we have that

$$v_{il,s}^* = \frac{\sum_{m=1}^{s}(e_i' R_m^* e_l)^2}{\sum_{m=1}^{s} e_i' R_m^* R_m^{*\prime} e_i},$$

for $i,l \in \{1,\ldots,n\}$ and $s = 1,2,\ldots$, and finally,

$$\bar{v}_{il} = \frac{(e_i' \Upsilon e_{l(k)})^2}{e_i' \Upsilon \Upsilon' e_i},$$

for $i \in \{1,\ldots,n\}$ and $l \in \{1,\ldots,k\}$. Here, $e_{l(k)}$ is the l:th column of I_k.

As we shall soon see, it is preferable from a theoretical perspective to examine the forecast error variance decompositions is terms of matrices. In particular, letting $\odot$ denote the Hadamard product (see e.g. Magnus and Neudecker (1988)) we find that

$$v_s = \left[\sum_{j=1}^{s}\left(R_{j-1} R_{j-1}' \odot I_n\right)\right]^{-1}\left[\sum_{j=1}^{s}\left(R_{j-1} \odot R_{j-1}\right)\right], \tag{37}$$

for $s = 1,2,\ldots$. The (i,l):th element of v_s is $v_{il,s}$. Furthermore,

$$v_s^* = \left[\sum_{m=1}^{s}\left(R_m^* R_m^{*\prime} \odot I_n\right)\right]^{-1}\left[\sum_{m=1}^{s}\left(R_m^* \odot R_m^*\right)\right], \tag{38}$$

whereas

$$\bar{v}_{\mathrm{inf}} = \left[\Upsilon \Upsilon' \odot I_n\right]^{-1}\left[\Upsilon \odot \Upsilon\right]. \tag{39}$$

The (i,l):th elements of these two matrices correspond to $v_{il,s}^*$ and $\bar{v}_{il}$, respectively. It should be noted that if all elements in some row of Υ_0 are equal to zero, this row must be deleted from $\hat{\Upsilon}$ when computing $\hat{\bar{v}}_{\mathrm{inf}}$. The reason is, of course, that the corresponding diagonal element of $\hat{\Upsilon}\hat{\Upsilon}'$ is zero and $[\hat{\Upsilon}\hat{\Upsilon}' \odot I_n]$ is consequently singular.

For any $m \times n$ matrix G, let diag[vec(G)] denote the $mn \times mn$ diagonal matrix whose diagonal elements are given by vec(G). Regarding the asymptotic distributions of the estimated forecast error variance decompositions it can now be stated that:

Theorem 4. *If the assumptions in Theorem 3 are satisfied, then*

$$\sqrt{T}\,(\mathrm{vec}(\hat{v}_s) - \mathrm{vec}(v_s)) \xrightarrow{\mathrm{d}} \mathbb{N}(0, w_s), \tag{40}$$

for $s = 1, 2, \ldots,$ *where*

$$w_s = \frac{\partial \mathrm{vec}(v_s)}{\partial \beta'} V_\beta \frac{\partial \mathrm{vec}(v_s)}{\partial \beta} + \frac{\partial \mathrm{vec}(v_s)}{\partial \omega'} V_\omega \frac{\partial \mathrm{vec}(v_s)}{\partial \omega},$$

while

$$\begin{aligned}
\partial \mathrm{vec}(v_s)/\partial \beta' \;&=\; 2[I_n \otimes [\textstyle\sum_{i=1}^{s}(R_{i-1}R_{i-1}' \odot I_n)]^{-1}] \textstyle\sum_{j=1}^{s}\{\mathrm{diag}[\mathrm{vec}(R_{j-1})] - \\
&\quad [v_s' \otimes I_n]\mathrm{diag}[\mathrm{vec}(I_n)][R_{j-1} \otimes I_n]\}(\partial \mathrm{vec}(R_{j-1})/\partial \beta'),
\end{aligned}$$

$$\begin{aligned}
\partial \mathrm{vec}(v_s)/\partial \omega' \;&=\; 2[I_n \otimes [\textstyle\sum_{i=1}^{s}(R_{i-1}R_{i-1}' \odot I_n)]^{-1}] \textstyle\sum_{j=1}^{s}\{\mathrm{diag}[\mathrm{vec}(R_{j-1})] - \\
&\quad [v_s' \otimes I_n]\mathrm{diag}[\mathrm{vec}(I_n)][R_{j-1} \otimes I_n]\}(\partial \mathrm{vec}(R_{j-1})/\partial \omega').
\end{aligned}$$

Regarding the asymptotic distribution of $\hat{v}_s^*$*, we have that*

$$\sqrt{T}\,(\mathrm{vec}(\hat{v}_s^*) - \mathrm{vec}(v_s^*)) \xrightarrow{\mathrm{d}} \mathbb{N}(0, w_s^*), \tag{41}$$

for $s = 1, 2, \ldots,$ *where*

$$w_s^* = \frac{\partial \mathrm{vec}(v_s^*)}{\partial \beta'} V_\beta \frac{\partial \mathrm{vec}(v_s^*)}{\partial \beta} + \frac{\partial \mathrm{vec}(v_s^*)}{\partial \omega'} V_\omega \frac{\partial \mathrm{vec}(v_s^*)}{\partial \omega},$$

while

$$\begin{aligned}
\partial \mathrm{vec}(v_s^*)/\partial \beta' \;&=\; 2[I_n \otimes [\textstyle\sum_{i=1}^{s}(R_i^* R_i^{*\prime} \odot I_n)]^{-1}] \textstyle\sum_{m=1}^{s}\{\mathrm{diag}[\mathrm{vec}(R_m^*)] - \\
&\quad [v_s^{*\prime} \otimes I_n]\mathrm{diag}[\mathrm{vec}(I_n)][R_m^* \otimes I_n]\}(\partial \mathrm{vec}(R_m^*)/\partial \beta'),
\end{aligned}$$

$$\begin{aligned}
\partial \mathrm{vec}(v_s^*)/\partial \omega' \;&=\; 2[I_n \otimes [\textstyle\sum_{i=1}^{s}(R_i^* R_i^{*\prime} \odot I_n)]^{-1}] \textstyle\sum_{m=1}^{s}\{\mathrm{diag}[\mathrm{vec}(R_m^*)] - \\
&\quad [v_s^{*\prime} \otimes I_n]\mathrm{diag}[\mathrm{vec}(I_n)][R_m^* \otimes I_n]\}(\partial \mathrm{vec}(R_m^*)/\partial \omega').
\end{aligned}$$

Finally, the asymptotic distribution for $\hat{\bar{v}}_{\mathrm{inf}}$ *is given by*

$$\sqrt{T}\left(\mathrm{vec}(\hat{\bar{v}}_{\mathrm{inf}}) - \mathrm{vec}(\bar{v}_{\mathrm{inf}})\right) \xrightarrow{\mathrm{d}} \mathbb{N}(0, \bar{w}_{\mathrm{inf}}), \tag{42}$$

where

$$\bar{w}_{\mathrm{inf}} = \frac{\partial \mathrm{vec}(\bar{v}_{\mathrm{inf}})}{\partial \beta'} V_\beta \frac{\partial \mathrm{vec}(\bar{v}_{\mathrm{inf}})}{\partial \beta} + \frac{\partial \mathrm{vec}(\bar{v}_{\mathrm{inf}})}{\partial \omega'} V_\omega \frac{\partial \mathrm{vec}(\bar{v}_{\mathrm{inf}})}{\partial \omega},$$

while

$$
\begin{aligned}
\partial\mathrm{vec}(\bar{v}_{\mathrm{inf}})/\partial\beta' &= 2[I_k \otimes [\Upsilon\Upsilon' \odot I_n]^{-1}]\{\mathrm{diag}[\mathrm{vec}(\Upsilon)]- \\
&\quad [\bar{v}_{\mathrm{inf}}' \otimes I_n]\mathrm{diag}[\mathrm{vec}(I_n)][\Upsilon \otimes I_n]\}(\partial\mathrm{vec}(\Upsilon)/\partial\beta'),
\end{aligned}
$$

$$
\begin{aligned}
\partial\mathrm{vec}(\bar{v}_{\mathrm{inf}})/\partial\omega' &= 2[I_k \otimes [\Upsilon\Upsilon' \odot I_n]^{-1}]\{\mathrm{diag}[\mathrm{vec}(\Upsilon)]- \\
&\quad [\bar{v}_{\mathrm{inf}}' \otimes I_n]\mathrm{diag}[\mathrm{vec}(I_n)][\Upsilon \otimes I_n]\}(\partial\mathrm{vec}(\Upsilon)/\partial\omega').
\end{aligned}
$$

In a common trends framework it may be of particular interest to study the joint influence of e.g. the permanent versus that of the transitory innovations on the forecast error variance of the time series. Other linear functions of variance decompositions that can be relevant in empirical studies of macroeconomic time series are sums of real versus sums of nominal innovations and sums of domestic versus sums of foreign innovations. To examine such linear functions, let G be an $n \times q$ matrix and $\bar{G}$ be a $k \times q$ matrix whose rows are taken from the first k rows of G. Consider the matrix functions

$$
\begin{aligned}
H_s &= v_s G, \\
H_s^* &= v_s^* G, \qquad (43) \\
\bar{H}_{\mathrm{inf}} &= \bar{v}_{\mathrm{inf}} \bar{G},
\end{aligned}
$$

for $s = 1, 2, \ldots$. The following results can easily be proven:

Corollary 1. *If the assumptions in Theorem 3 are satisfied, then*

$$
\sqrt{T}\left(\mathrm{vec}(\hat{H}_s) - \mathrm{vec}(H_s)\right) \xrightarrow{\mathrm{d}} \mathbb{N}(0, [G' \otimes I_n] w_s [G \otimes I_n]), \qquad (44)
$$

$$
\sqrt{T}\left(\mathrm{vec}(\hat{H}_s^*) - \mathrm{vec}(H_s^*)\right) \xrightarrow{\mathrm{d}} \mathbb{N}(0, [G' \otimes I_n] w_s^* [G \otimes I_n]), \qquad (45)
$$

for $s = 1, 2, \ldots,$ *while*

$$
\sqrt{T}\left(\mathrm{vec}(\hat{\bar{H}}_{\mathrm{inf}}) - \mathrm{vec}(\bar{H}_{\mathrm{inf}})\right) \xrightarrow{\mathrm{d}} \mathbb{N}(0, [\bar{G}' \otimes I_n] \bar{w}_{\mathrm{inf}} [\bar{G} \otimes I_n]). \qquad (46)
$$

The asymptotic covariance matrices in Theorem 4 take into account that the rows of v_s, v_s^*, and $\bar{v}_{\mathrm{inf}}$ sum to one. Accordingly, the rank of w_s and w_s^* is (less than or) equal to $n(n-1)$, while the rank of $\bar{w}_{\mathrm{inf}}$ is (less than or) equal to $n(k-1)$. It then follows that formal significance tests of the hypothesis that an element of, say, v_s is equal to zero or one cannot be conducted since the corresponding standard error is equal to zero.[9] Furthermore, by using the technique which allows us to derive the partial derivatives in Theorem 4 it can be shown that the expressions given by Lütkepohl (1990) indeed also take into account that the sum of a variance decomposition is always equal to one.

The reduced rank property of the covariance matrices can be stated in a very simple yet general form. Let $\imath_m$ be the $m \times 1$ unit vector. We then have that:

Corollary 2. *If $G = \imath_n$ and $\bar{G} = \imath_k$, then the asymptotic covariance matrices in Corollary 1 are equal to zero.*

This result can now be used to determine the asymptotic relationship between two groups of innovations for a particular variance decomposition. In particular, let g be an $n \times 1$ vector with known elements and $g_\perp := \imath_n - g$. Similarly, let $\bar{g}$ be the $k \times 1$ vector obtained from the first k elements of g and $\bar{g}_\perp := \imath_k - \bar{g}$. It now follows that

Corollary 3. *The asymptotic covariance matrices in Corollary 1 are equal for $G = g$ and $G = g_\perp$ for $s = 1, 2, \ldots$ and also for $\bar{G} = \bar{g}$ and $\bar{G} = \bar{g}_\perp$.*

Hence, letting $G = g$ be a vector with ones in the first k elements and zeros elsewhere *or* vice versa, i.e. $G = g_\perp$, will provide us with identical estimates of the asymptotic covariances matrices for $\hat{H}_s$ and $\hat{H}_s^*$. That is, the standard error for an estimate of the joint influence of the permanent innovations in a variance decomposition is equal to the standard error for an estimate of the joint influence of the transitory innovations.

5. Concluding Comments

In this paper I analyse how we can estimate and conduct inference in a common trends model with k permanent and r transitory innovations when the time series are cointegrated of order (1,1) with r cointegrating vectors. Such innovations may be of particular interest when we are interested in studying connections between growth and business cycle fluctuations in macroeconomic time series.[10] Theorem 1, a version of the well known Granger's representation theorem, shows how all reduced form parameters

[9] In fact, if the asymptotic standard errors would not be zero under such null hypotheses we would be alarmed since the variance decompositions are bounded by zero and one. In fact, this suggests that the limiting distributions are generally associated with smaller standard errors the closer the true value of a variance decomposition gets to either of these bounds.

[10] Readers interested in macroeconomic applications are referred to papers by Englund, Vredin and Warne (1992) (who study the behavior of output, money, the price level, the interest rate, public consumption

can be calculated directly from the (estimated) parameters of a restricted VAR. The restrictions this model satisfies are the cointegration constraints which may, e.g., describe the steady state of a stochastic growth model. Furthermore, a simple description of the solution to the unrestricted and restricted VAR as well as to the VEC representation in terms of estimable parameters is provided. The solution is expressed as a Wold VMA representation which describes a reduced form of the propagation mechanism.

Second, an identification matrix for permanent and transitory innovations has been derived. Since this matrix is generally not triangular and is based on certain assumptions regarding the nature of permanent and transitory innovations, it cannot be computed by a Choleski, an eigen, or a method of moments decomposition of Σ. However, it is shown that its parameters are solely determined by parameters which have already been calculated and is therefore easily obtained in practise.

Third, I have analysed the asymptotic properties of impulse response functions and forecast error variance decompositions within a common trends model given an upper bound for the lag order. Such functions are calculated directly from the Wold VMA representation and from the identification matrix. Based on the results in Theorems 3 and 4, we find that the analytical expressions of the asymptotic covariances are somewhat more complex than in the model studied by Lütkepohl and Reimers (1992). The reason for the added complexity is that I have analysed identification of permanent and transitory innovations, whereas Lütkepohl and Reimers study the case when the covariance matrix Σ is orthogonalized via a Choleski decomposition. From a practical point of view, however, the added complexity is not severe. For example, to compute estimates of the asymptotic covariance matrices of impulse response functions is about as time consuming as for ordinary VAR's. Moreover, Corollary 1 provides us with asymptotic distributions of the estimated forecast error variances which are accounted for by linear functions of the innovations. Based on these distributions it is, e.g. possible to analyse how important innovations to growth are, at a business cycle horizon, relative to transitory shocks for the time series of interest. Also, it highlights the fact that the asymptotic covariance

---

and foreign output in a common trends framework), Jacobson, Vredin and Warne (1993) (output, employment, unemployment and real product wages), and Mellander, Vredin and Warne (1992) (output, consumption, investments and terms of trade). Other applied studies on common trends include Blakemore and Hoffman (1993), Blanchard and Quah (1989), KPSW (1987,1991), and Shapiro and Watson (1988).

matrices of the estimated forecast error variance decompositions are conditioned on the property that the sum of a decomposition for any time series is equal to one.

Mathematical Appendix

Proof of Theorem 2: Let us partition the inverse of Γ into

$$\Gamma^{-1} = \begin{bmatrix} \Gamma_k^+ & \Gamma_r^+ \end{bmatrix},$$

where Γ_k^+ and Γ_r^+ are $n \times k$ and $n \times r$ matrices, respectively. Postmultiplying Γ in equation (29) by this expression for Γ^{-1}, we obtain

$$\Gamma\Gamma^{-1} = \begin{bmatrix} (\Upsilon'\Upsilon)^{-1}\Upsilon' C(1)\Gamma_k^+ & (\Upsilon'\Upsilon)^{-1}\Upsilon' C(1)\Gamma_r^+ \\ Q^{-1}\zeta'\Sigma^{-1}\Gamma_k^+ & Q^{-1}\zeta'\Sigma^{-1}\Gamma_r^+ \end{bmatrix} = I_n.$$

Letting $\Gamma_r^+ = \zeta(Q')^{-1}$ and $\Gamma_k^+ = \Sigma C(1)'\Upsilon(\Upsilon'\Upsilon)^{-1}$ we have found the inverse of Γ. Substituting for these relationships in equation (30), we have that

$$R(1) = [R(1)_k \;\; R(1)_r] = \begin{bmatrix} C(1)\Sigma C(1)'\Upsilon(\Upsilon'\Upsilon)^{-1} & C(1)\zeta(Q')^{-1} \end{bmatrix}.$$

Clearly, $R(1)_k = C(1)\Sigma C(1)'\Upsilon(\Upsilon'\Upsilon)^{-1} = \Upsilon$ since $C(1)\Sigma C(1)' = \Upsilon\Upsilon'$, while $R(1)_r = 0$ by virtue of the fact that $C(1)\zeta = 0$. Finally, it is obvious from the above analysis that $\Gamma\Sigma\Gamma' = I_n$. Q.E.D.

Before we prove Theorems 3 and 4 the following Lemma is useful:

Lemma 1. *Let A, B and C be $n \times n$ matrices with B nonsingular and $A := B^{-1}CB$. Suppose that*

$$\sqrt{T}\left(\mathrm{vec}(\hat{B}) - \mathrm{vec}(B)\right) \xrightarrow{\mathrm{p}} 0, \tag{A.1}$$

while $\hat{C}$ is a consistent estimator of C with

$$\sqrt{T}\left(\mathrm{vec}(\hat{C}) - \mathrm{vec}(C)\right) \xrightarrow{\mathrm{d}} \mathbb{N}\left(0, \Sigma\right). \tag{A.2}$$

It then follows that $\hat{A} = \hat{B}^{-1}\hat{C}\hat{B}$ is a consistent estimator of A and

$$\sqrt{T}\left(\mathrm{vec}(\hat{A}) - \mathrm{vec}(A)\right) \xrightarrow{\mathrm{d}} \mathbb{N}\left(0, \left[B' \otimes B^{-1}\right]\Sigma\left[B \otimes (B')^{-1}\right]\right). \tag{A.3}$$

Proof: It is straightforward to show that

$$\hat{A} - A = B^{-1}\left(\hat{C} - C\right)B + B^{-1}\hat{C}\left(\hat{B} - B\right) + \left(\hat{B}^{-1} - B^{-1}\right)\hat{C}\hat{B}.$$

It now follows from (A.1) that

$$\sqrt{T}B^{-1}\hat{C}\left(\hat{B}-B\right)\xrightarrow{\text{p}}0.$$

Similarly, since B^{-1} is a continuous function of B by Mann and Wald's theorem (see e.g. Serfling (1980, p. 24) we find that

$$\sqrt{T}\left(\hat{B}^{-1}-B^{-1}\right)\hat{C}\hat{B}\xrightarrow{\text{p}}0.$$

Using this, equation (A.2), and taking vec's of $(\hat{A}-A)$ the conclusion follows. Q.E.D.

This lemma thus illustrates that an estimator which converges in probability (to its true value) at the rate $\sqrt{T}$ can be treated as known when we study a function that depends on estimates that converge in probability as well as estimates that converge in distribution at the rate $\sqrt{T}$. Hence, the lemma is a special case of Slutzky's theorem (see e.g. Serfling (1980, p. 19)). Since the condition $\sqrt{T}(\hat{\alpha}-\alpha)\xrightarrow{\text{p}}0$ implies that $\sqrt{T}(\hat{M}-M)\xrightarrow{\text{p}}0$ and $\sqrt{T}(\hat{\Upsilon}_0-\Upsilon_0)\xrightarrow{\text{p}}$ (and the assumption about the limiting behavior of β and ω in Theorem 3 is consistent with this; cf. Johansen (1991) or Warne (1992)), it follows that $\hat{M}$ and $\hat{\Upsilon}_0$ can be treated as known when we examine the limiting behavior of $\hat{C}_j$, $\hat{R}_j$, etc.

Proof of Theorem 3: In order to derive the matrices in equation (34), note that the following relationships hold: $R_j=C_j\Gamma^{-1}$, $\Gamma^{-1}=[\Sigma C(1)'\Upsilon(\Upsilon'\Upsilon)^{-1}\ \ \zeta(Q')^{-1}]$, $\Sigma=M^{-1}\Omega(M')^{-1}$, $\zeta=\gamma(U\gamma)^{-1}$, $\gamma=M^{-1}(I_n-J_pBE_p')P_r$, $\Upsilon=\Upsilon_0\pi$, $QQ'=\zeta'\Sigma^{-1}\zeta$, and

$$\pi\pi'=(\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'C(1)\Sigma C(1)'\Upsilon_0(\Upsilon_0'\Upsilon_0)^{-1}.$$

Then, taking vec's of the first differential of R_j, we find that

$$d\text{vec}(R_j)=\left[(\Gamma^{-1})'\otimes I_n\right]d\text{vec}(C_j)+[I_n\otimes C_j]\begin{bmatrix}d\text{vec}(\Gamma_k^+)\\ d\text{vec}(\Gamma_r^+)\end{bmatrix}. \tag{A.4}$$

From equation (21) we know that

$$dC_j=\sum_{k=0}^{j-1}M^{-1}J_pB^kdBB^{j-1-k}J_p'M-\sum_{k=0}^{j-2}M^{-1}DJ_pB^kdBB^{j-2-k}J_p'M,$$

where the second term on the right hand side is zero for $j \leq 1$ and the first term is equal to zero for $j = 0$ since $dF_0 = 0$. Noting that $dB = J_p' db$, where $b = [B_1 \ \cdots \ B_p]$, and applying the vec operator we have

$$d\text{vec}(C_j) = \left\{ \sum_{k=0}^{j-1} \left[M' J_p (B')^{j-1-k} \otimes M^{-1} F_k \right] - \sum_{k=0}^{j-2} \left[M' J_p (B')^{j-2-k} \otimes M^{-1} D F_k \right] \right\} d\beta,$$

where $F_k = J_p B^k J_p'$.

Before we consider $d\text{vec}(\Gamma_k^+)$, note that

$$\Upsilon(\Upsilon'\Upsilon)^{-1} = \Upsilon_0 \pi (\pi' \Upsilon_0' \Upsilon_0 \pi)^{-1} = \Upsilon_0 (\Upsilon_0' \Upsilon_0)^{-1} (\pi')^{-1}.$$

Since $\hat{\Upsilon}_0$ can be treated as known we only need to consider π when deriving the first differential of $\Upsilon(\Upsilon'\Upsilon)^{-1}$.

The first differential of Γ_k^+ is equal to

$$\begin{aligned} d\Gamma_k^+ &= (d\Sigma) C(1)' \Upsilon (\Upsilon'\Upsilon)^{-1} + \Sigma (dC(1)') \Upsilon (\Upsilon'\Upsilon)^{-1} \\ &\quad - \Sigma C(1)' \Upsilon_0 (\Upsilon_0' \Upsilon_0)^{-1} (\pi')^{-1} (d\pi') (\pi')^{-1}. \end{aligned}$$

Taking vec's and using the commutation matrix, we have that

$$\begin{aligned} d\text{vec}(\Gamma_k^+) &= [(\Upsilon'\Upsilon)^{-1} \Upsilon' C(1) \otimes I_n] d\text{vec}(\Sigma) + [(\Upsilon'\Upsilon)^{-1} \Upsilon' \otimes \Sigma] K_{nn} d\text{vec}(C(1)) \\ &\quad - [\pi^{-1} \otimes \Sigma C(1)' \Upsilon (\Upsilon'\Upsilon)^{-1}] K_{kk} d\text{vec}(\pi). \end{aligned} \tag{A.5}$$

From the function mapping the parameters of Ω into Σ, we get

$$d\text{vec}(\Sigma) = [M^{-1} \otimes M^{-1}] d\text{vec}(\Omega) = [M^{-1} \otimes M^{-1}] D_n d\omega. \tag{A.6}$$

Furthermore, by Theorem 1 it immediately follows that

$$d\text{vec}(C(1)) = [M' F(1)' E_p \otimes M^{-1} D(1) F(1)] d\beta. \tag{A.7}$$

Next, the matrix π is lower triangular and obtained from a symmetric matrix. Lemma 1 in Lütkepohl (1989) implies that

$$d\text{vech}(\pi) = \{2 L_k N_k (\pi \otimes I_k) L_k'\}^{-1} d\text{vech}(\pi\pi').$$

From the properties of the elimination and duplication matrices, we then find that

$$d\text{vec}(\pi) = L_k' \{2 L_k N_k (\pi \otimes I_k) L_k'\}^{-1} D_k^+ d\text{vec}(\pi\pi'). \tag{A.8}$$

The first differential of $\pi\pi'$ is

$$\begin{aligned} d(\pi\pi') &= (\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'(dC(1))\Sigma C(1)'\Upsilon_0(\Upsilon_0'\Upsilon_0)^{-1}+ \\ &\quad (\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'C(1)\Sigma(dC(1)')\Upsilon_0(\Upsilon_0'\Upsilon_0)^{-1}+ \\ &\quad (\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'C(1)(d\Sigma)C(1)'\Upsilon_0(\Upsilon_0'\Upsilon_0)^{-1}. \end{aligned}$$

Taking vec's and making use of the commutation and N matrices, we find that

$$\begin{aligned} d\text{vec}(\pi\pi') &= 2N_k[(\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'C(1)\Sigma \otimes (\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0']d\text{vec}(C(1))+ \\ &\quad [(\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'C(1) \otimes (\Upsilon_0'\Upsilon_0)^{-1}\Upsilon_0'C(1)]d\text{vec}(\Sigma), \end{aligned} \tag{A.9}$$

by employing Lemma 4 in Magnus and Neudecker (1986). Substituting for the expressions in equation (A.6)–(A.9) into (A.5), using the relationship between Σ and Ω, and the fact that $D_k^+N_k = D_k^+$ (see Theorem 3.12 in Magnus and Neudecker (1988)), the matrices with partial derivatives of $\text{vec}(\Gamma_k^+)$ with respect to β and ω follow directly.

The first differential of Γ_r^+ is

$$d\Gamma_r^+ = (d\zeta)(Q')^{-1} - \zeta(Q')^{-1}(dQ')(Q')^{-1}.$$

Taking vec's we get

$$d\text{vec}(\Gamma_r^+) = [Q^{-1} \otimes I_n]d\text{vec}(\zeta) - [Q^{-1} \otimes \zeta(Q')^{-1}]K_{rr}d\text{vec}(Q). \tag{A.10}$$

Next, the first differential of ζ is

$$d\zeta = (d\gamma)(U\gamma)^{-1} - \zeta U(d\gamma)(U\gamma)^{-1},$$

whereas the first differential of γ is

$$d\gamma = -M^{-1}J_p(dB)E_p'P_r.$$

Combining these two relationships and employing the vec operator we obtain

$$d\text{vec}(\zeta) = \left\{[(\gamma'U')^{-1} \otimes \zeta U] - [(\gamma'U')^{-1} \otimes I_n]\right\}[P_r'E_p \otimes M^{-1}]d\beta, \tag{A.11}$$

since $d\text{vec}(B) = [I_{np} \otimes J_p']d\beta$ and $J_pJ_p' = I_n$.

With Q being lower triangular, Lemma 1 in Lütkepohl (1989) provides us with

$$d\text{vec}(Q) = L_r'\{2L_rN_r(Q \otimes I_r)L_r'\}^{-1}D_r^+d\text{vec}(QQ'). \tag{A.12}$$

The first differential of QQ' is

$$d(QQ') = (d\zeta')\Sigma^{-1}\zeta + \zeta'\Sigma^{-1}(d\zeta) - \zeta'\Sigma^{-1}(d\Sigma)\Sigma^{-1}\zeta.$$

Taking vec's, we get

$$d\text{vec}(QQ') = 2N_r[I_r \otimes \zeta'\Sigma^{-1}]d\text{vec}(\zeta) - [\zeta'\Sigma^{-1} \otimes \zeta'\Sigma^{-1}]d\text{vec}(\Sigma). \tag{A.13}$$

Substituting for the expressions in equations (A.6) and (A.11)–(A.13) into (A.10) and making use of the property $D_r^+ N_r = D_r^+$, the matrices with partial derivatives of $\text{vec}(\Gamma_r^+)$ with respect to β and ω are immediate. Finally, equations (35) and (36) follow directly from (34), the above expressions, and the fact that $d\text{vec}(\Upsilon) = [I_k \otimes \Upsilon_0]d\text{vec}(\pi)$. Q.E.D.

Proof of Theorem 4: Comparing the expressions for v_s, v_s^* and $\bar{v}_{\text{inf}}$ it is obvious that the asymptotic covariance matrices for estimates of these functions have a similar parametric structure. Hence, I shall limit the proof to equation (40).

To derive the matrices of partial derivatives of $\text{vec}(v_s)$ with respect to β and ω, note that the first differential of v_s is given by

$$\begin{aligned} dv_s &= -[\textstyle\sum_{i=1}^s (R_{i-1}R_{i-1}' \odot I_n)]^{-1}[\textstyle\sum_{j=1}^s d(R_{j-1}R_{j-1}' \odot I_n)]v_s + \\ &\quad [\textstyle\sum_{i=1}^s (R_{i-1}R_{i-1}' \odot I_n)]^{-1}[\textstyle\sum_{j=1}^s d(R_{j-1} \odot R_{j-1})] \\ &= [\textstyle\sum_{i=1}^s (R_{i-1}R_{i-1}' \odot I_n)]^{-1} \textstyle\sum_{j=1}^s \{[(dR_{j-1} \odot R_{j-1}) + (R_{j-1} \odot dR_{j-1})] \\ &\quad -[(d(R_{j-1}R_{j-1}') \odot I_n)]v_s\}. \end{aligned}$$

Taking vec's and using Lemma 2 in Magnus and Neudecker (1985), we obtain

(A.14)

$$\begin{aligned} d\text{vec}(v_s) &= [I_n \otimes [\textstyle\sum_{i=1}^s (R_{i-1}R_{i-1}' \odot I_n)]^{-1}] \textstyle\sum_{j=1}^s \{2\text{diag}[\text{vec}(R_{j-1})]d\text{vec}(R_{j-1}) \\ &\quad -[v_s' \otimes I_n]\text{diag}[\text{vec}(I_n)]d\text{vec}(R_{j-1}R_{j-1}')\}. \end{aligned}$$

The first differential of $R_{j-1}R_{j-1}'$ is

$$d(R_{j-1}R_{j-1}') = (dR_{j-1})R_{j-1}' + R_{j-1}(dR_{j-1}').$$

Applying the vec operator and the commutation matrix, we have that

$$\begin{aligned} d\text{vec}(R_{j-1}R'_{j-1}) &= [R_{j-1} \otimes I_n]d\text{vec}(R_{j-1}) + [I_n \otimes R_{j-1}]K_{nn}d\text{vec}(R_{j-1}) \\ &= 2N_n[R_{j-1} \otimes I_n]d\text{vec}(R_{j-1}), \end{aligned} \tag{A.15}$$

by Lemma 4 in Magnus and Neudecker (1986). Subsituting equation (A.15) for the first differential $d\text{vec}(R_{j-1}R'_{j-1})$ into equation (A.14), it is clear that equation (40) follows if

$$\text{diag}[\text{vec}(I_n)]N_n = \text{diag}[\text{vec}(I_n)]K_{nn} = \text{diag}[\text{vec}(I_n)].$$

If the second equality is valid, the first is an immediate consequence of the definition of N_n. To show the second, note that

$$\text{diag}[\text{vec}(I_n)] = \begin{bmatrix} e_1e'_1 \otimes e'_1 \\ e_2e'_2 \otimes e'_2 \\ \vdots \\ e_ne'_n \otimes e'_n \end{bmatrix}, \tag{A.16}$$

whereas

$$K_{nn} = \begin{bmatrix} I_n \otimes e'_1 \\ I_n \otimes e'_2 \\ \vdots \\ I_n \otimes e'_n \end{bmatrix}. \tag{A.17}$$

Since K_{nn} is symmetric (cf. Magnus and Neudecker (1988), p. 47) the result follows from an inspection of (A.16) and (A.17). Q.E.D.

Proof of Corollary 3: Below I shall show that for any $n \times 1$ vector g with $g_\perp := \imath_n - g$ the following equality holds true:

$$[g' \otimes I_n]w_s[g \otimes I_n] = [g'_\perp \otimes I_n]w_s[g_\perp \otimes I_n].$$

The two remaining equalities then follow from similar arguments.

The first differential of $v_s\imath_n \equiv \imath_n$ is

$$(dv_s)\imath_n = 0.$$

Using the fact that $\imath_n = g + g_\perp$, we obtain

$$(dv_s)g = -(dv_s)g_\perp.$$

Taking vec's provides us with

$$[g' \otimes I_n] d\mathrm{vec}(v_s) = -[g'_\perp \otimes I_n] d\mathrm{vec}(v_s),$$

or in terms of partial derivatives

$$[g' \otimes I_n] \frac{\partial \mathrm{vec}(v_s)}{\partial \beta'} = -[g'_\perp \otimes I_n] \frac{\partial \mathrm{vec}(v_s)}{\partial \beta'}, \tag{A.18}$$

and

$$[g' \otimes I_n] \frac{\partial \mathrm{vec}(v_s)}{\partial \omega'} = -[g'_\perp \otimes I_n] \frac{\partial \mathrm{vec}(v_s)}{\partial \omega'}. \tag{A.19}$$

Postmultiplying both sides of (A.18) by V_β times the transpose of the expression in the equation we get

$$[g' \otimes I_n] \frac{\partial \mathrm{vec}(v_s)}{\partial \beta'} V_\beta \frac{\partial \mathrm{vec}(v_s)}{\partial \beta} [g \otimes I_n] = [g'_\perp \otimes I_n] \frac{\partial \mathrm{vec}(v_s)}{\partial \beta'} V_\beta \frac{\partial \mathrm{vec}(v_s)}{\partial \beta} [g_\perp \otimes I_n]. \tag{A.20}$$

Similarly, for (A.19) we find that

$$[g' \otimes I_n] \frac{\partial \mathrm{vec}(v_s)}{\partial \omega'} V_\omega \frac{\partial \mathrm{vec}(v_s)}{\partial \omega} [g \otimes I_n] = [g'_\perp \otimes I_n] \frac{\partial \mathrm{vec}(v_s)}{\partial \omega'} V_\omega \frac{\partial \mathrm{vec}(v_s)}{\partial \omega} [g_\perp \otimes I_n]. \tag{A.21}$$

Adding the expressions in equations (A.20) and (A.21) the result follows. Q.E.D.

References

[1] Baillie, Richard T. (1981), "Prediction from the Dynamic Simultaneous Equation Model with Vector Autoregressive Errors", *Econometrica*, 49:1331–1337.

[2] Baillie, Richard T. (1987), "Inference in Dynamic Models Containing 'Surprise' Variables", *Journal of Econometrics*, 35:101–117.

[3] Baillie, Richard T. (1989), "Econometric Tests of Rationality and Market Efficiency", *Econometric Reviews*.

[4] Bernanke, Ben S. (1986), "Alternative Explanations of the Money–Income Correlation", *Carnegie–Rochester Conference Series on Public Policy*, 25:49–100.

[5] Blakemore, Arthur E. and Dennis L. Hoffman (1993), "Nonstationary Behavior in the Labor Market and the Intertemporal Substitution Effect", Unpublished Manuscript, Arizona State University.

[6] Blanchard, Olivier J. and Danny Quah (1989), "The Dynamic Effects of Aggregate Demand and Supply Disturbances", *American Economic Review*, 79:655–673.

[7] Braun, Phillip A. and Stefan Mittnik (1993), "Misspecifications in Vector Autoregressions and their Effects on Impulse Responses and Variance Decompositions", *Journal of Econometrics*, 59:319–341.

[8] Brillinger, D. R. (1981), *Time Series Data Analysis and Theory*, Holden–Day, San Francisco.

[9] Campbell, John Y. and Robert J. Shiller (1987), "Cointegration and Tests of Present Value Models", *Journal of Political Economy*, 95:1062–1088.

[10] Campbell, John Y. and Robert J. Shiller (1988), "Interpreting Cointegrated Models", *Journal of Economic Dynamics and Control*, 12:505–522.

[11] Engle, Robert F. and Clive W.J. Granger (1987), "Co–Integration and Error Correction: Representation, Estimation and Testing", *Econometrica*, 55:251–276.

[12] Englund, Peter, Anders Vredin, and Anders Warne (1992), "Macroeconomic Shocks in an Open Economy: A Common Trends Representation of Swedish Data 1871–1990", *FIEF Studies in Labor Markets and Economic Policy*, vol. 5, Clarendon Press, Oxford (forthcoming).

[13] Gonzalo, Jesus and Clive W.J. Granger (1992), "Estimation of Common Long–Memory Components in Cointegrated Systems", Unpublished Manuscript, University of California, San Diego.

[14] Henderson, Harold V. and S.R. Searle (1979), "Vec and Vech Operators for Matrices, with some uses in Jacobians and Multivariate Statistics", *Canadian Journal of Statistics*, 7:65–81.

[15] Hylleberg, Svend and Grayham E. Mizon (1989), "Cointegration and Error Correction Mechanisms", *Economic Journal*, 99:113–125.

[16] Jacobson, Tor, Vredin, Anders and Anders Warne (1993), "Are Real Wages and Unemployment Related?", Unpublished Manuscript, Institute for International Economic Studies, Stockholm University.

[17] Johansen, Søren (1988a), "The Mathematical Structure of Error Correction Models", *Contemporary Mathematics*, 80:359–386.

[18] Johansen, Søren (1988b), "Statistical Analysis of Cointegration Vectors", *Journal of Economic Dynamics and Control*, 12:231–254.

[19] Johansen, Søren (1989), "Likelihood Based Inference on Cointegration: Theory and Applications", Lecture Notes, Institute of Mathematical Statistics, University of Copenhagen, Denmark.

[20] Johansen, Søren (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models", *Econometrica*, 59:1551–1580.

[21] Johansen, Søren and Katarina Juselius (1990), "Maximum Likelihood Estimation and Inference on Cointegration – with Applications to the Demand for Money", *Oxford Bulletin of Economics and Statistics*, 52:169–210.

[22] King, Robert G., Charles I. Plosser, and Sergio T. Rebelo (1988a), "Production, Growth and Business Cycles: I. The Basic Neoclassical Model", *Journal of Monetary Economics*, 21:195–232.

[23] King, Robert G., Charles I. Plosser, and Sergio T. Rebelo (1988b), "Production, Growth and Business Cycles: II. New Directions", *Journal of Monetary Economics*, 21:309–341.

[24] King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson (1987), "Stochastic Trends and Economic Fluctuations", NBER Working Paper No. 2229.

[25] King, Robert G., Charles I. Plosser, James H. Stock, and Mark W. Watson (1991), "Stochastic Trends and Economic Fluctuations", *American Economic Review*, 81:819–840.

[26] Lippi, Marco and Lucrezia Reichlin (1990), "A Note on Measuring the Dynamic Effects of Aggregate Demand and Supply Disturbances", Working Paper No. 90–06, OFCE, Paris.

[27] Lütkepohl, Helmut (1988), "Asymptotic Distribution of the Moving Average Coefficients of an Estimated Vector Autoregressive Process", *Econometric Theory*, 4:77–85.

[28] Lütkepohl, Helmut (1989), "A Note on the Asymptotic Distribution of Impulse Response

Functions of estimated VAR Models with Orthogonal Residuals", *Journal of Econometrics*, 42:371–376.

[29] Lütkepohl, Helmut (1990), "Asymptotic Distributions of Impulse Response Functions and Forecast Error Variance Decompositions of Vector Autoregressive Models", *Review of Economics and Statistics*, 72:116–125.

[30] Lütkepohl, Helmut and Don S. Poskitt (1990), "Estimating Orthogonal Impulse Responses via Vector Autoregressive Models", Manuscript, Institut für Statistik und Ökonometrie der Christian–Albrechts–Universität, Kiel, Germany.

[31] Lütkepohl, Helmut and Hans–Eggert Reimers (1992), "Impulse Response Analysis of Cointegrated Systems", *Journal of Economic Dynamics and Control*, 16:53–78.

[32] Magnus, Jan R. and Heinz Neudecker (1985), "Matrix Differential Calculus with Applications to Simple, Hadamard, and Kronecker Products", *Journal of Mathematical Psychology*, 29:474–492.

[33] Magnus, Jan R. and Heinz Neudecker (1986), "Symmetry, 0–1 Matrices and Jacobians: A Review", *Econometric Theory*, 2:157–190.

[34] Magnus, Jan R. and Heinz Neudecker (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, John Wiley, Chichester.

[35] Mellander, Erik, Anders Vredin, and Anders Warne (1992), "Stochastic Trends and Economic Fluctuations in a Small Open Economy", *Journal of Applied Econometrics*, 7:369–394.

[36] Neudecker, Heinz (1983), "On Jacobians of Transformations with Skew–Symmetric, Strictly (Lower) Triangular or Diagonal Matrix Arguments", *Linear and Multilinear Algebra*, 14:271–295.

[37] Park, Joon Y. (1990), "Disequilibrium Impulse Analysis", Memo 1990–19, Institute of Economics, University of Aarhus, Denmark.

[38] Park, Joon Y. (1992), "Canonical Cointegrating Regressions", *Econometrica*, 60:119–143.

[39] Schmidt, Peter (1973), "The Asymptotic Distribution of Dynamic Multipliers", *Econometrica*, 41:161–164.

[40] Schmidt, Peter (1974), "The Asymptotic Distribution of Forecasts in the Dynamic Simulation of an Econometric Model", *Econometrica*, 42:303–309.

[41] Serfling, Robert J. (1980), *Approximation Theorems of Mathematical Statistics*, John Wiley, New York.

[42] Shapiro, Matthew D. and Mark W. Watson (1988), "Sources of Business Cycle Fluctuations", In Stanley Fisher, editor, *NBER Macroeconomics Annual*, pages 111–148, MIT Press.

[43] Stock, James H. (1987), "Asymptotic Properties of Least Squares Estimators of Cointegrating Vectors", *Econometrica*, 55:1035–1056.

[44] Stock, James H. and Mark W. Watson (1988), "Testing for Common Trends", *Journal of the American Statistical Association*, 83:1097–1107.

[45] Warne, Anders (1990), *Vector Autoregressions and Common Trends in Macro and Financial Economics*, PhD thesis, Stockholm School of Economics, Sweden.

[46] Warne, Anders (1992), "Wald Tests of Nonlinear Restrictions in Vector Autoregressions with Integrated Time Series", Unpublished Manuscript, Institute for International Economic Studies, Stockholm University.

Institute for International Economic Studies, Stockholm University, 10691 Stockholm, Sweden

E-mail address: warnea@iies.su.se