Journal of Economic Dynamics and Control 16 (1992) 53–78. North-Holland

Impulse response analysis of cointegrated systems

Helmut Lütkepohl and Hans-Eggert Reimers*

Christian-Albrechts-Universität Kiel, D-2300 Kiel, Germany

Received March 1990, final version received April 1991

Impulse response or dynamic multiplier analysis of vector autoregressive systems with cointegrated variables is considered. The asymptotic distribution of the responses estimated with Johansen's (1988) maximum likelihood procedure is derived. The results are illustrated with an analysis of a West German money demand system. The investigation shows that a direct interpretation of the cointegration relations may be difficult or misleading. Thereby the virtue of impulse response analysis for applied work is illustrated.

1. Introduction

Impulse response or dynamic multiplier analysis is a common tool for investigating the interrelationships among the variables in dynamic models. In the following we will argue that this tool is also valuable in cointegrated systems. In such systems it is assumed that although the individual variables are nonstationary, there are linear combinations of them which are stationary. These linear combinations are often interpreted as long-run equilibrium relations. In other words, it is assumed that the deviations from the equilibrium relations are stationary. Consequently, assuming that the variables are in an equilibrium at some time t, say $t = 0$, any shock to one of the variables will result in time paths of the system that will eventually settle down in a new equilibrium provided no further shocks occur. These time paths of the variables may give interesting insights into the short-term and long-run

*The authors are grateful to Søren Johansen, Katarina Juselius, and two anonymous referees for helpful comments on earlier versions of the paper which have led to substantial improvements. An earlier version of the paper entitled 'Impulse Response Analysis of Cointegrated Systems with an Application to German Money Demand' was presented at the 6th World Congress of the Econometric Society in Barcelona, 1990.

relations among the variables that are of interest in many applied analyses especially since it is often difficult to interpret the cointegration relations directly. This is especially true if there are two or more such relations. Dickey (1988) gives an example demonstrating the power of these devices for interpreting cointegrated systems.

In the next section we will lay out the formal framework and we will formally define the quantities of interest. In section 3 the estimation of cointegrated systems is reviewed and the resulting estimators of the impulse responses and of forecast error variance components are investigated. In section 4 a West German money demand system is discussed and thereby the virtue of impulse response analysis of cointegrated systems is illustrated. Conclusions are contained in section 5. The proof of the theorems stated in section 3 is given in the appendix.

2. A vector autoregressive model with cointegrated variables

We consider a K-dimensional vector autoregressive (VAR) model of the form

$$y_t = A_1 y_{t-1} + \cdots + A_p y_{t-p} + u_t, \tag{1}$$

where $y_t = (y_{1t}, \ldots, y_{Kt})'$, the A_i are $(K \times K)$ coefficient matrices, and u_t is *Gaussian* white noise, that is, u_t and u_s are independent for $s \neq t$ and $u_t \sim \mathrm{N}(0, \Sigma_u)$ for all t. The covariance matrix Σ_u is assumed to be positive definite. Furthermore, we assume that the first differences $\Delta y_t = y_t - y_{t-1}$ are stationary and that

$$\det(I_K - A_1 z - \cdots - A_p z^p) \tag{2}$$

has all its roots outside the complex unit circle except for possibly some roots that are unity. In other words,

$$\Pi = I_K - A_1 - \cdots - A_p \tag{3}$$

may be singular, say of rank $r \leq K$.

The $(K \times K)$ matrix Π can be expressed as a product of a $(K \times r)$ matrix B and an $(r \times K)$ matrix C, which have both rank r, that is, $\Pi = BC$. Here C is a matrix representing the cointegration relations such that Cy_t is stationary. Commonly Cy_t is interpreted as the long-run equilibrium relation between the y variables. Since this relation is often of interest, the model is usually reparameterized in one of several equivalent forms. In the following we will use the representation

$$\Delta y_t = \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} - \Pi y_{t-p} + u_t, \tag{4}$$

where

$$\Gamma_i = -(I_K - A_1 - \cdots - A_i), \qquad i = 1, \ldots, p-1.$$

In this representation it becomes quite obvious that the deviations from the equilibrium relations Cy_{t-p} form a stationary process.

The quantities of interest in the following are the impulse responses or dynamic multipliers that represent the effects of shocks in the variables of the system. As in the stationary case, they are most easily obtained from the representation (1) and may be defined as

$$\Phi_n = (\varphi_{ik,n}) = \sum_{j=1}^{n} \Phi_{n-j} A_j, \qquad n = 1, 2, \ldots, \tag{5}$$

where $\Phi_0 = I_K$, $A_j = 0$ for $j > p$ and $\varphi_{ik,n}$ is the ikth element of Φ_n and represents the response of variable y_i to a unit shock in variable k, n periods ago. In many econometric studies responses to orthogonalized impulses are preferred. They are defined as $\Theta_n = (\theta_{ik,n}) = \Phi_n P$, where P is the lower triangular Choleski decomposition of Σ_u, that is, $PP' = \Sigma_u$ [see, e.g., Lütkepohl (1990)]. Again $\theta_{ik,n}$ is interpreted as response of variable i to an impulse in variable k, n periods ago. These impulses can be thought of as transformed residuals of the form $w_t = P^{-1} u_t$ which have identity covariance matrix, $\mathrm{E}(w_t w_t') = I_K$. Thus, a unit impulse has size one standard deviation in this case. For both types of impulse responses the difference to the stationary case is that the effect of a shock in one of the variables will in general not die out in the long run, that is, the variables may not return to their initial values even if no further shocks occur. In other words, a one-time impulse may have a permanent effect in the sense that it shifts the system to a new equilibrium. Therefore the Φ_n and Θ_n cannot be interpreted as MA coefficient matrices and their sums will in general not be finite.

Other quantities of potential interest are accumulated impulse responses,

$$\Psi_m = I_K + \sum_{n=1}^{m} \Phi_n \quad \text{and} \quad \Xi_m = \sum_{n=0}^{m} \Theta_n$$

[see Lütkepohl (1990)]. Given the previous remarks it is clear that these quantities will usually diverge to infinity for $m \to \infty$. Again this contrasts with the stationary case where Ψ_∞ and Ξ_∞ are finite matrices and are often interpreted as accumulated long-run effects. In addition to these quantities forecast error variance decompositions are often considered in VAR analyses. They are also available for cointegrated systems and are computed using the same formulas as in the stationary case. Specifically,

$$\omega_{kj,h} = \sum_{n=0}^{h-1} \theta_{kj,n}^2 / MSE_k(h), \qquad h = 1,2,\dots, \tag{6}$$

where $\theta_{kj,n}$ is the kjth element of Θ_n and $MSE_k(h)$ is the kth diagonal element of

$$MSE(h) = \Sigma_u + \sum_{n=1}^{h-1} \Phi_n \Sigma_u \Phi_n',$$

the mean squared error matrix of the optimal h-step-ahead forecast of the y_t process.

It may also be worth noting the relation between the impulse responses and Granger-causation. In a bivariate system with variables y_{1t} and y_{2t} the latter variable is not Granger-causal for y_{1t} if and only if $\varphi_{12,n} = 0$ for $n = 1,2,\dots$. Denoting the upper right-hand element of A_i in (1) by $\alpha_{12,i}$, $i = 1,\dots,p$, this condition for noncausality is equivalent to $\alpha_{12,i} = 0$, $i = 1,\dots,p$, just as in the stationary case.

In the next section maximum likelihood (ML) estimation of the system (4) will be considered and the resulting asymptotic distributions of the A_i coefficients in (1) and the impulse responses and forecast error variance decompositions will be given.

3. Estimation of the impulse responses

Suppose we have estimators $\hat{A}_1,\dots,\hat{A}_p$ of $A_1,\dots,A_p$, that have an asymptotic normal distribution with $(K^2p \times K^2p)$ covariance matrix Σ_A. More precisely,

$$\sqrt{T}\,\mathrm{vec}\Big[\big(\hat{A}_1,\dots,\hat{A}_p\big) - (A_1,\dots,A_p)\Big] \xrightarrow{d} \mathrm{N}(0,\Sigma_A). \tag{7}$$

Then it follows, e.g., from Baillie (1987) or Lütkepohl (1990, proposition 1) that

$$\hat{\Phi}_n = \sum_{j=1}^{n} \hat{\Phi}_{n-j}\hat{A}_j, \qquad n = 1,2,\dots,$$

and

$$\hat{\Psi}_m = I_K + \sum_{n=1}^{m} \hat{\Phi}_n, \qquad m = 1, 2, \ldots,$$

also have asymptotic normal distributions,

$$\sqrt{T}\,\mathrm{vec}\big(\hat{\Phi}_n - \Phi_n\big) \xrightarrow{\mathrm{d}} \mathrm{N}(0, G_n \Sigma_A G_n'), \qquad n = 1, 2, \ldots, \tag{8}$$

and

$$\sqrt{T}\,\mathrm{vec}\big(\hat{\Psi}_m - \Psi_m\big) \xrightarrow{\mathrm{d}} \mathrm{N}(0, F_m \Sigma_A F_m'), \qquad m = 1, 2, \ldots, \tag{9}$$

where $G_n = \partial\,\mathrm{vec}\,\Phi_n / \partial\,\mathrm{vec}(A_1, \ldots, A_p)'$ and $F_m = G_1 + \cdots + G_m$. Here 'vec' is the column stacking operator. An explicit expression for G_n is given, for example, in Lütkepohl (1990). Given the remarks at the end of the previous section the asymptotic distribution in (7) can also be used in the usual way to perform tests for Granger-causality.

Suppose now that $\hat{\Sigma}_u$ is an estimator of Σ_u which is asymptotically independent of the $\hat{A}_i$ and satisfies

$$\sqrt{T}\,\mathrm{vech}\big(\hat{\Sigma}_u - \Sigma_u\big) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Sigma_\sigma), \tag{10}$$

where 'vech' is the column stacking operator that stacks the elements on and below the diagonal only. Assume furthermore that $\hat{P}$ is the lower triangular Choleski decomposition of $\hat{\Sigma}_u$ and Θ_n is estimated as $\hat{\Theta}_n = \hat{\Phi}_n \hat{P}$ for $n = 0, 1, \ldots,$ with $\hat{\Phi}_0 = I_K$. Then the $\hat{\Theta}_n$ and $\hat{\Xi}_m = \sum_{n=0}^{m} \hat{\Theta}_n$ have asymptotic normal distributions

$$\sqrt{T}\,\mathrm{vec}\big(\hat{\Theta}_n - \Theta_n\big) \xrightarrow{\mathrm{d}} \mathrm{N}\big(0, C_n \Sigma_A C_n' + \bar{C}_n \Sigma_\sigma \bar{C}_n'\big), \qquad n = 0, 1, \ldots, \tag{11}$$

and

$$\sqrt{T}\,\mathrm{vec}\big(\hat{\Xi}_m - \Xi_m\big) \xrightarrow{\mathrm{d}} \mathrm{N}\big(0, B_m \Sigma_A B_m' + \bar{B}_m \Sigma_\sigma \bar{B}_m'\big), \qquad m = 0, 1, \ldots, \tag{12}$$

where the matrices C_n, $\bar{C}_n$, B_m, and $\bar{B}_m$ are also given in Lütkepohl (1990, proposition 1). Finally, the estimated forecast error variance components, obtained by replacing the unknown quantities in (6) by the above estimators, have the following asymptotic distributions:

$$\sqrt{T}\left(\hat{\omega}_{kj,h}-\omega_{kj,h}\right) \xrightarrow{\mathrm{d}} \mathrm{N}\left(0, d_{kj,h}\Sigma_A d'_{kj,h}+\bar{d}_{kj,h}\Sigma_\sigma \bar{d}'_{kj,h}\right), \tag{13}$$

where the $d_{kj,h}$ and $\bar{d}_{kj,h}$ are again given in Lütkepohl (1990, proposition 1). For all the above asymptotic distributions the remarks of the latter article regarding possible singularities apply here too.

In order to actually obtain the asymptotic covariance matrices of the asymptotic distributions in (8)–(13), we need Σ_A and Σ_σ. These covariance matrices can be derived using the approach of Johansen (1988, 1989) and Johansen and Juselius (1990a) who consider ML estimation of the system (4) subject to the restriction that Π has rank r. We will review these results here to the extent necessary for deriving the matrices Σ_A and Σ_σ. For the moment we assume that the data generation process does not include a constant term just as in (1). Later we will comment on processes with constant terms and seasonal dummy variables.

In order to state the required results we assume that a time series of length T and p presample values are available and use the following notation:

$$\begin{aligned}
&\Delta Y=[\Delta y_1,\ldots,\Delta y_T] && (K\times T),\\
&X_t=\begin{bmatrix}\Delta y_{t-1}\\ \vdots\\ \Delta y_{t-p+1}\end{bmatrix} && (K(p-1)\times 1),\\
&X=[X_1,\ldots,X_T] && (K(p-1)\times T),\\
&Y_{-p}=[y_{1-p},\ldots,y_{T-p}] && (K\times T),\\
&U=[u_1,\ldots,u_T] && (K\times T),\\
&\Gamma=[\Gamma_1,\ldots,\Gamma_{p-1}] && (K\times K(p-1)).
\end{aligned}$$

With this notation the model (4) can be written compactly as

$$\Delta Y=\Gamma X-\Pi Y_{-p}+U. \tag{14}$$

Assuming fixed presample values[1] the ML estimator of Γ, for given Π, is

[1]The assumption of fixed presample values is a common one in VAR analyses. We do not know of any studies investigating the small-sample consequences of that assumption in the present context. As a referee points out, an ML estimator for the full sample could be considered alternatively.

known to be

$$\hat{\Gamma}(\Pi) = (\Delta Y + \Pi Y_{-p})X'(XX')^{-1}. \tag{15}$$

This result may be used to concentrate on $\Pi = BC$ and determine the ML estimator of B for given C by substituting (15) for Γ in (14),

$$\Delta Y = (\Delta Y + \Pi Y_{-p})X'(XX')^{-1}X - \Pi Y_{-p} + \hat{U}$$

or

$$\Delta YM = -BCY_{-p}M + \hat{U},$$

where $M = I - X'(XX')^{-1}X$. Hence

$$-\hat{B}(C) = \Delta YMY'_{-p}C'\left[CY_{-p}MY'_{-p}C'\right]^{-1}, \tag{16}$$

where the idempotency of M has been used.

Now we can concentrate on C and obtain an estimator of this matrix. Note, however, that the matrices B and C are not identified because for any nonsingular $(r \times r)$ matrix G, $BGG^{-1}C = \Pi$. Thus, $B^* = BG$ and $C^* = G^{-1}C$ also qualify to decompose Π. Therefore identifying constraints must be imposed for consistent estimation of the cointegration matrix C. Johansen (1988) suggests a procedure that amounts to restricting the length of the cointegration vectors. While this does not identify the true cointegration vectors, it permits to estimate a basis of the space spanned by the rows of the cointegration matrix. Specifically the following procedure is proposed: Define

$$R_0 = \Delta YM, \qquad R_1 = Y_{-p}M,$$

$$S_{ij} = \frac{1}{T}R_iR'_j, \qquad i,j = 0,1,$$

and L is such that $LS_{11}L' = I$. The eigenvalues of $LS_{10}S_{00}^{-1}S_{01}L'$ are

$$\lambda_1 \geq \cdots \geq \lambda_K, \tag{17}$$

and the corresponding eigenvectors

$$V = [v_1, \ldots, v_K]$$

are normalized such that $V'V = I$. These normalization restrictions are the constraints which are imposed to obtain unique estimates. It can be shown that an ML estimator of a basis of the space spanned by C is given by

$$\hat{C} = [v_1, \ldots, v_r]' L'. \tag{18}$$

Note that this is not a consistent estimator of C because identifying restrictions are not imposed. However, the resulting estimator $\hat{\Pi} = \hat{B}(\hat{C})\hat{C}$ is consistent, and so is the ML estimator $\hat{\Gamma} = \hat{\Gamma}(\hat{\Pi})$ of Γ. The asymptotic distribution of these estimators is given in the following theorem the proof of which is a consequence of Johansen's results and is given in the appendix. In the theorem the matrix $D_K^+ = (D_K' D_K)^{-1} D_K'$ is used, where D_K is the $(K^2 \times K(K+1)/2)$ duplication matrix [see Magnus (1988)].

Theorem 1. If y_t is a Gaussian process as in (4) with $rk(\Pi) = r$, then

$$\sqrt{T}\, \mathrm{vec}\left(\left[\hat{\Gamma}, -\hat{\Pi}\right] - [\Gamma, -\Pi]\right) \xrightarrow{d} \mathrm{N}(0, \Sigma_{c0}),$$

where

$$\Sigma_{c0} = \begin{pmatrix} I_{K(p-1)} & 0 \\ 0 & C' \end{pmatrix} \Omega^{-1} \begin{pmatrix} I_{K(p-1)} & 0 \\ 0 & C \end{pmatrix} \otimes \Sigma_u \tag{19}$$

and

$$\Omega = \mathrm{plim}\, \frac{1}{T} \begin{pmatrix} XX' & XY_{-p}'C' \\ CY_{-p}X' & CY_{-p}Y_{-p}'C' \end{pmatrix}.$$

Σ_{c0} is consistently estimated by

$$\hat{\Sigma}_{c0} = \begin{pmatrix} I & 0 \\ 0 & \hat{C}' \end{pmatrix} \hat{\Omega}^{-1} \begin{pmatrix} I & 0 \\ 0 & \hat{C} \end{pmatrix} \otimes \hat{\Sigma}_u, \tag{20}$$

where

$$\hat{\Omega} = \frac{1}{T} \begin{pmatrix} XX' & XY_{-p}'\hat{C}' \\ \hat{C}Y_{-p}X' & \hat{C}Y_{-p}Y_{-p}'\hat{C}' \end{pmatrix}$$

and $\hat{\Sigma}_u = S_{00} - \hat{B}\hat{B}'$. Moreover, $\hat{\Sigma}_u$ is asymptotically independent of the other parameters and

$$\sqrt{T}\, \mathrm{vech}\left(\hat{\Sigma}_u - \Sigma_u\right) \xrightarrow{d} \mathrm{N}\left(0, \Sigma_\sigma = 2D_K^+ (\Sigma_u \otimes \Sigma_u) D_K^{+\prime}\right). \tag{21}$$

The last result stated in the theorem means that $\hat{\Sigma}_u$ has the same asymptotic distribution as a white noise covariance matrix estimator based on the true residuals u_t rather than estimation residuals [see Magnus (1988)].

The existence of Ω is guaranteed since CY_{-p} and X all involve stationary variables only. Note that Σ_{c0} is singular for $r<K$. It is perhaps worth pointing out that the estimators $\hat{\Gamma}$ and $\hat{\Pi}=\hat{B}\hat{C}$ converge at rate $\sqrt{T}$ although $\hat{C}$ converges at rate T (see the appendix). The $\sqrt{T}$-convergence of $\hat{\Pi}$ is a consequence of the $\sqrt{T}$-convergence of B which implies that the product $\hat{B}\hat{C}$ converges at the slower rate $\sqrt{T}$ and not at rate T. In the proof in the appendix it is essentially shown that, in deriving the asymptotic distributions of the parameters of interest, C may be treated as known due to the rapid convergence of $\hat{C}$. This result is obtained despite the aforementioned fact that $\hat{C}$ is not a consistent estimator of C. A consequence of this result is that Σ_{c0} is exactly the covariance matrix one would obtain by assuming that C is known.

If the process y_t contains a constant term, say ν, that is,

$$y_t=\nu+A_1y_{t-1}+\cdots+A_py_{t-p}+u_t$$

or

$$\Delta y_t=\nu+\Gamma_1\Delta y_{t-1}+\cdots+\Gamma_{p-1}\Delta y_{t-p+1}-BCy_{t-p}+u_t, \tag{22}$$

two cases may be distinguished. The first case assumes that ν can be absorbed into the cointegration relation which is possible if there exists an $(r\times 1)$ vector η such that $-B\eta=\nu$ and hence (22) may be rewritten as

$$\Delta y_t=\Gamma_1\Delta y_{t-1}+\cdots+\Gamma_{p-1}\Delta y_{t-p+1}-B[C,\eta]\begin{bmatrix} y_{t-p}\\ 1\end{bmatrix}+u_t. \tag{23}$$

The ML estimators for the parameters of interest may then be obtained using the formulas (15)–(18) with Y_{-p} replaced by

$$Y_{-p}^*=\begin{pmatrix} y_{1-p},\dots,y_{T-p}\\ 1,\dots,1\end{pmatrix},$$

and C is replaced by $[C,\eta]$. The asymptotic distribution of the resulting estimators $[\hat{\Gamma},-\hat{\Pi}]$ may be obtained from Theorem 1 by removing the last row and column from the inverse of

$$\Omega^*=\operatorname{plim}\frac{1}{T}\begin{pmatrix} XX' & XY_{-p}^{*\prime}[C,\eta]' \\ [C,\eta]Y_{-p}^*X' & [C,\eta]Y_{-p}^*Y_{-p}^{*\prime}[C,\eta]'\end{pmatrix}$$

and using that matrix in place of Ω^{-1} [see Johansen (1989)].

If ν cannot be absorbed into the cointegration relation, the ML estimators of the parameters

$$[\nu, \Gamma] = [\nu, \Gamma_1, \ldots, \Gamma_{p-1}]$$

can again be obtained from the formulas (15)–(18) by redefining X_t as

$$X_t = (1, \Delta y'_{t-1}, \ldots, \Delta y'_{t-p+1})'.$$

The resulting estimators $\hat{\Gamma}_i$ and $\hat{\Pi}$ again have an asymptotic normal distribution as in Theorem 1. The covariance matrix of that distribution can be obtained as in (19) by deleting the first row and column from Ω^{-1}. Note, however, that in general the resulting estimator $\hat{\nu}$ of the constant term does not have an asymptotic normal distribution. This is no problem in the following because $\hat{\nu}$ does not enter the estimators for the impulse responses and forecast error variance components.

Seasonal dummy variables may also be added into the model. For instance, for quarterly data seasonal terms may be included by replacing Γ with

$$[\nu, \delta_2, \delta_3, \delta_4, \Gamma]$$

and defining X_t as

$$X_t = [1, d_{2t}, d_{3t}, d_{4t}, \Delta y'_{t-1}, \ldots, \Delta y'_{t-p+1}]',$$

where, e.g.,

$$d_{it} = \begin{cases} \frac{3}{4} & \text{if } t \text{ belongs to the } i\text{th quarter}, \\ -\frac{1}{4} & \text{otherwise}. \end{cases}$$

With these modifications the same formulas as in (15)–(18) may be used in computing the estimators of the parameters of interest. Again the resulting $\hat{\Gamma}_i$ and $\hat{\Pi}$ have asymptotic normal distributions as in Theorem 1 [Johansen (1989)]. This approach is used in the empirical example in section 4.

From the asymptotic normal distribution of the $\hat{\Gamma}_i$ and $\hat{\Pi}$ it is easy to obtain the desired asymptotic distribution of the corresponding estimators $\hat{A}_i$. Defining $J = [I_K \; 0 \; \ldots \; 0]$ and

$$D = \begin{pmatrix} I_K & -I_K & 0 & \ldots & 0 & 0 \\ 0 & I_K & -I_K & & 0 & 0 \\ 0 & 0 & I_K & \ddots & 0 & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & 0 & 0 & \ddots & -I_K & 0 \\ 0 & 0 & 0 & \ldots & I_K & -I_K \\ 0 & 0 & 0 & \ldots & 0 & I_K \end{pmatrix} \quad (Kp \times Kp), \tag{24}$$

we get

$$\left[\hat{A}_1,\dots,\hat{A}_p\right]=\left[\hat{\Gamma},-\hat{B}\hat{C}\right]D+J, \tag{25}$$

and the asymptotic distribution of this estimator is given in the following corollary.

Corollary 1. Under the conditions of Theorem 1, the estimators $\hat{A}_i$ have the asymptotic normal distribution given in (7) with covariance matrix

$$\Sigma_A=D'\begin{pmatrix} I & 0\\ 0 & C'\end{pmatrix}\Omega^{-1}\begin{pmatrix} I & 0\\ 0 & C\end{pmatrix}D\otimes\Sigma_u, \tag{26}$$

and

$$\hat{\Sigma}_A=D'\begin{pmatrix} I & 0\\ 0 & \hat{C}'\end{pmatrix}\hat{\Omega}^{-1}\begin{pmatrix} I & 0\\ 0 & \hat{C}\end{pmatrix}D\otimes\left(S_{00}-\hat{B}\hat{B}'\right) \tag{27}$$

is a consistent estimator of Σ_A.

Proof. The result is trivial since

$$\operatorname{vec}[(\Gamma,-\Pi)D]=(D'\otimes I_K)\operatorname{vec}[\Gamma,-\Pi]. \quad \blacksquare$$

Instead of imposing the constraint $rk(\Pi)=r<K$, one may estimate the model (1) in unrestricted form by multivariate least squares. Park and Phillips (1989) and Sims, Stock, and Watson (1990) argue that the resulting estimator $[\tilde{A}_1,\dots,\tilde{A}_p]$ has an asymptotic normal distribution whose covariance matrix may be estimated by the usual formula for stationary processes [e.g., Judge et al. (1988, ch. 18)]. For our purposes it may be advantageous to impose the rank constraint in order to isolate the equilibrium relations.

In some applications restrictions have been imposed on the matrices B and C. If the restrictions are merely identifying constraints, that is, if they ensure uniqueness of B and C and do not constrain the parameter space, the results in Theorem 1 and Corollary 1 are maintained because the estimator of C converges more rapidly than the estimators of the other parameters [e.g., Engle and Granger (1987) or Stock (1987)]. For instance, if C can be written in the form $C=[I_r\ C_1]$, where C_1 is a $(r\times(K-r))$ matrix, these restrictions just identify C. This type of identifying restriction has been used in particular for the case of just one cointegration relation ($r=1$), where it amounts to normalizing the coefficient of the first variable [e.g., Engle and Granger (1987)].

In some studies partially identifying or overidentifying restrictions are imposed either on the cointegration matrix C or on the loading matrix B.

For instance, if one wants to exclude the cointegration relations (the level variables) from some of the equations of the system zero constraints may be placed on B. Alternatively, if one of the variables is not cointegrated with any of the other variables one may want to exclude it from the cointegration relations by constraining some elements of C to zero. Johansen (1988, 1989) and Johansen and Juselius (1990a) consider restrictions for the loading and cointegration matrices that constrain all columns of B or all rows of C in the same manner. Formally the constraints can be written as

$$B = HB_1, \tag{28}$$

where H is $(K \times s)$ and B_1 is $(s \times r)$, $r \leq s \leq K$, and

$$C = C_1 R, \tag{29}$$

where C_1 is $(r \times m)$ and R is $(m \times K)$, $r \leq m \leq K$. In this case the ML estimators of Γ, B_1, C_1, and Σ_u can be obtained as follows [see Johansen and Juselius (1990a)]: Let $\bar{H}$ be a $(K \times (K-s))$ matrix such that $\bar{H}'H = 0$ and $\bar{H}'\bar{H} = I_{K-s}$ and define

$$\tilde{R}_0 = \left[H' - H'S_{00}\bar{H}\left(\bar{H}'S_{00}\bar{H}\right)^{-1}\bar{H}'\right]R_0,$$

$$\tilde{R}_1 = R\left[R_1 - S_{10}\bar{H}\left(\bar{H}'S_{00}\bar{H}\right)^{-1}\bar{H}'R_0\right],$$

$$\tilde{S}_{ij} = \frac{1}{T}\tilde{R}_i\tilde{R}_j', \qquad i,j = 0,1,$$

$\tilde{L}$ is defined such that $\tilde{L}\tilde{S}_{11}\tilde{L}' = I_m$,

$\tilde{\lambda}_1 \geq \cdots \geq \tilde{\lambda}_m$ are the eigenvalues of $\tilde{L}\tilde{S}_{10}\tilde{S}_{00}^{-1}\tilde{S}_{01}\tilde{L}'$,

$\tilde{V} = [\tilde{v}_1, \ldots, \tilde{v}_m]$ are the corresponding eigenvectors satisfying $\tilde{V}'\tilde{V} = I_m$.

Then we get the ML estimators

$$\tilde{C}_1 = [\tilde{v}_1, \dots, \tilde{v}_r]'\tilde{L}', \qquad \tilde{C} = \tilde{C}_1 R,$$
$$\tilde{B}_1 = -(H'H)^{-1}\tilde{S}_{01}\tilde{C}_1', \qquad \tilde{B} = H\tilde{B}_1,$$

$$\tilde{\Pi} = \tilde{B}\tilde{C},$$

$$\tilde{\Gamma} = \left[\Delta Y + H\tilde{B}_1\tilde{C}_1 R Y_{-p}\right] X'(XX')^{-1},$$

$$\tilde{\Sigma}_u = \left(\Delta Y - \tilde{\Gamma}X + H\tilde{B}_1\tilde{C}_1 R Y_{-p}\right)\left(\Delta Y - \tilde{\Gamma}X + H\tilde{B}_1\tilde{C}_1 R Y_{-p}\right)'/T.$$

The following theorem is also proven in the appendix.

Theorem 2. If y_t is a Gaussian process as in (4) with $rk(\Pi) = r$ and B and C as in (28) and (29), respectively, then

$$\sqrt{T}\,\mathrm{vec}\left(\left[\tilde{\Gamma}, -\tilde{\Pi}\right] - [\Gamma, -\Pi]\right) \xrightarrow{\mathrm{d}} \mathrm{N}(0, \Sigma_{c0}^r),$$

where

$$\Sigma_{c0}^r = \begin{pmatrix} I_{K^2(p-1)} & 0 \\ 0 & C' \otimes H \end{pmatrix} \Omega_r^{-1} \begin{pmatrix} I_{K^2(p-1)} & 0 \\ 0 & C \otimes H' \end{pmatrix} \tag{30}$$

and

$$\Omega_r = \mathrm{plim}\frac{1}{T}\begin{pmatrix} XX' \otimes \Sigma_u^{-1} & XY_{-p}'C' \otimes \Sigma_u^{-1}H \\ CY_{-p}X' \otimes H'\Sigma_u^{-1} & CY_{-p}Y_{-p}'C' \otimes H'\Sigma_u^{-1}H \end{pmatrix}.$$

A consistent estimator $\tilde{\Sigma}_{c0}^r$ of Σ_{c0}^r is obtained by replacing C and Σ_u in (30) and Ω_r by $\tilde{C}$ and $\tilde{\Sigma}_u$, respectively. Furthermore, $\tilde{\Sigma}_u$ is asymptotically independent of the other parameters and

$$\sqrt{T}\,\mathrm{vech}\left(\tilde{\Sigma}_u - \Sigma_u\right) \xrightarrow{\mathrm{d}} \mathrm{N}\left(0, \Sigma_\sigma = 2D_K^+(\Sigma_u \otimes \Sigma_u)D_K^{+\prime}\right).$$

Although Theorem 1 is a special case of this theorem with $H = I$ and $R = I$, we have chosen to state the former separately because the covariance matrix Σ_{c0} has a particularly simple form and is clearly of interest from a practical point of view. In practice, in many situations there may be no restrictions for B and C. Obviously, the asymptotic distribution of $\tilde{\Sigma}_u$ in Theorem 2 is the same as that of $\hat{\Sigma}_u$ in Theorem 1. The comments following Theorem 1 regarding the inclusion of intercept terms and seasonal dummy variables apply here too [see Johansen (1989)]. In the next corollary we give the Σ_A matrix obtained from Theorem 2.

Corollary 2. Under the conditions of Theorem 2, the covariance matrix in (7) is

$$\Sigma_A = (D' \otimes I_K)\Sigma^r_{c0}(D \otimes I_K),$$

where D is as defined in (24). $\tilde{\Sigma}_A = (D' \otimes I_K)\tilde{\Sigma}^r_{c0}(D \otimes I_K)$ is a consistent estimator of Σ_A.

4. Analysis of a German money demand system – An illustration

To illustrate the theoretical results we consider West German money demand. Our point of departure is a standard money demand function of the form

$$M/P = f(Y, i),$$

where M is a money variable, P is the price level, so M/P is real money demand, Y is the transactions volume, and i is an interest rate. We use the broader definition *M2* as money variable because its demand function is regarded to be more stable than that of *M1* [see Judd and Scadding (1982)]. The opportunity cost for holding *M2* is reflected in the difference between short- and long-term interest rates. Therefore, we use two interest rates in the model, a short-term rate i^s and a bond rate i^b. The transactions volume is approximated by real domestic demand (DD) to reduce the impact of external effects such as changes in the terms of trade. For Germany there is some evidence that domestic demand and money are more closely related than the gross national product and money [Scheide (1984)]. Thus, we get a four-dimensional system consisting of the variables:

$y_1 = 1 +$ yields outstanding for total fixed interest securities (i^b),
$y_2 = 1 +$ interest rate on three-month loans in the money market, Frankfurt am Main (i^s),
$y_3 =$ real domestic demand (DD) (in billions of German marks),
$y_4 =$ real money demand ($M2$) (in billions of German marks).

We assume a log-linear relationship. Therefore, all variables are in logarithms. Unadjusted quarterly data from 1960.I–1987.IV were used. DD, i^s, and i^b are from the data base of the Deutsche Bundesbank. The latter variable is based on a wide range of longer-term securities. It may be thought of as a 5- to 7-year bond rate and, hence, it represents a long-term interest

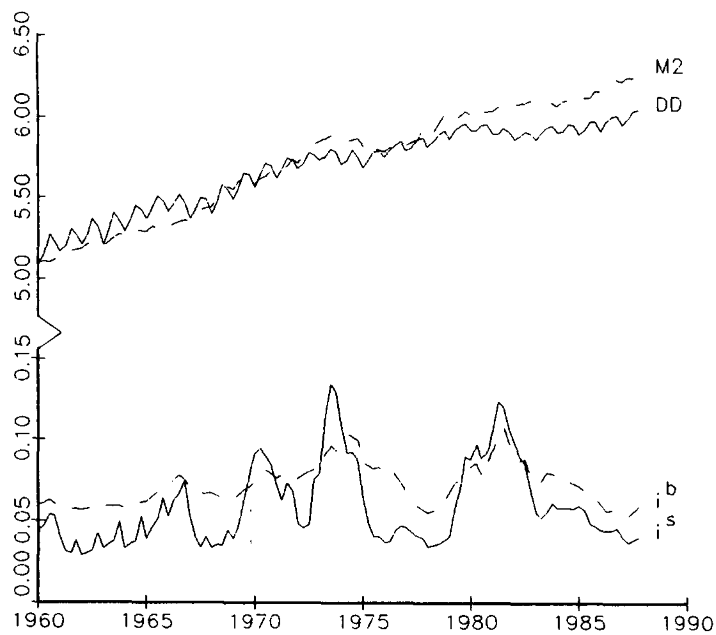

Fig. 1. Plot of ln $M2$, ln DD, ln i^s, and ln i^b.

rate. Nominal $M2$ is constructed by the Institute of World Economics, Kiel. Real $M2$ is obtained by deflating the nominal figures with the implicit deflator of domestic demand. The quarterly interest rates are averages of monthly rates. A plot of the four series is given in fig. 1, where it is seen that the series are trending together. This behavior indicates potential cointegration among the variables.

Using a maximum order of five, the model selection criteria HQ and AIC [see Judge et al. (1985, ch. 16)] suggested the orders $p = 2$ and $p = 3$, respectively. We have chosen the less restricted model and work with a VAR(3) model with intercept terms and seasonal dummies in the following. The first three observations of each series are treated as presample values leaving a sample size of $T = 109$. In table 1 the results of Johansen's (1989) likelihood ratio tests for the rank of cointegration are documented. At the common significance level of 5% both tests indicate a rank of $r = 2$. The corresponding estimated cointegration matrix is

$$\begin{array}{cc} & \begin{array}{cccc} i^b & i^s & DD & M2 \end{array} \\ \hat{C} = & \begin{pmatrix} 127.3 & -92.8 & -4.41 & 3.64 \\ -104.0 & 14.3 & 6.23 & -4.96 \end{pmatrix}. \end{array}$$

Table 1

Tests for cointegration.

Hypotheses	Trace test				Max. eigenvalue test		
H_0	$r=3$	$r=2$	$r=1$	$r=0$	$r=2$	$r=1$	$r=0$
H_1	$r=4$	$r\geq 3$	$r\geq 2$	$r\geq 1$	$r=3$	$r=2$	$r=1$
Values of test statistics	0.88	7.97	31.53[a]	74.38[b]	7.09	23.56[b]	42.84[b]

[a]Significant at 5% level [see Johansen and Juselius (1990a)].
[b]Significant at 1% level [see Johansen and Juselius (1990a)].

In studying the term structure of U.S. interest rates, Campbell and Shiller (1987) find that the interest rate differential (the spread) $i^{\mathrm{b}}-i^{\mathrm{s}}$ may be regarded as a stationary variable. Thus, one of the two cointegration relations may relate the two interest rates only and may not involve *M2* and *DD*, while the other cointegration relation may represent a money demand function. To check whether a similar interpretation is possible for our German system, we have tested the hypotheses that i^{b} and i^{s} enter the first cointegration relation with coefficients of identical size but opposite sign (i.e., $i^{\mathrm{b}}-i^{\mathrm{s}}$ is the relevant variable) and that *DD* and *M2* do not appear in that relation (the attached coefficients are zero). A possible test for this hypothesis is described by Johansen and Juselius (1990b). For the present system it clearly rejects the null hypothesis even at a 1% level of significance. Thus, the above interpretation seems not possible for the example system.

It is still possible that the spread $i^{\mathrm{b}}-i^{\mathrm{s}}$ is the variable of importance in the system if the coefficients of the two interest rates are identical with opposite sign in the cointegration relations. Thus, we have tested this hypothesis with a χ^2 test due to Johansen (1989). The test value is 17.3, which is highly significant compared with critical values from a $\chi^2(2)$ distribution. Therefore we reject the null hypothesis and treat i^{b} and i^{s} as separate variables.

The strict version of the quantity theory of money implies that money and income are proportional [see also Juselius (1989)]. Focussing just on the cointegration relations this means that in our log-linear system *DD* and *M2* may have equal coefficients with opposite sign. For that hypothesis,

$$R=\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{pmatrix},$$

in the framework of section 3. The test statistic of a $\chi^2(2)$ test of that hypothesis assumes the value 0.64 and is therefore not significant. Hence, we impose the restriction in the following analysis in order to highlight a potential problem that results from focussing attention on the cointegration vectors only and neglecting the remaining parts of the dynamic system.

The estimated restricted cointegration matrix becomes

$$\begin{array}{cccccc} & & i^{\mathrm{b}} & i^{\mathrm{s}} & DD & M2 \\ \hat{C}_1 R = & & \left(\begin{matrix} 122.5 \\ 101.3 \end{matrix}\right. & \begin{matrix} -91.9 \\ -14.8 \end{matrix} & \begin{matrix} -2.19 \\ -2.69 \end{matrix} & \left.\begin{matrix} 2.19 \\ 2.69 \end{matrix}\right). \end{array}$$

Normalizing the money coefficient implies relations

$$M2 = \nu_1 + DD + 42.0 i^{\mathrm{s}} - 55.9 i^{\mathrm{b}},$$

$$M2 = \nu_2 + DD + 5.50 i^{\mathrm{s}} - 37.7 i^{\mathrm{b}}, \tag{31}$$

where ν_1 and ν_2 are constants. If these cointegration relations were interpreted as long-run relations, it might be tempting to argue that a permanent increase in DD leads to an increase of $M2$ of equal size. We will return to this interpretation shortly in the light of an impulse response analysis. An assessment of the impact of the interest rates is much more difficult because the two equations (31) seem to imply different effects. This interpretation ignores, however, that the two relations in (31) cannot be interpreted individually. If the two equations really describe the equilibrium relations between the four variables and if the equilibrium is to be maintained, a change in one interest rate will not only induce a shift in $M2$ but is necessarily accompanied by a change in the other interest rate.

It may also be worth noting that little additional information regarding the long-run behavior of the system is gained from the matrix Π which is sometimes referred to as the matrix of long-run responses [see Hylleberg and Mizon (1989)]. For the present example the estimate of the Π matrix is

$$\hat{\Pi} = \hat{B}\hat{C}_1 R = \begin{bmatrix} \underset{(0.050)}{0.126} & \underset{(0.029)}{-0.061} & \underset{(0.001)}{-0.003} & \underset{(0.001)}{0.003} \\ \underset{(0.102)}{-0.278} & \underset{(0.059)}{0.225} & \underset{(0.002)}{0.005} & \underset{(0.002)}{-0.005} \\ \underset{(0.279)}{0.185} & \underset{(0.164)}{0.288} & \underset{(0.006)}{-0.010} & \underset{(0.006)}{0.010} \\ \underset{(0.250)}{0.404} & \underset{(0.146)}{0.048} & \underset{(0.006)}{-0.012} & \underset{(0.006)}{0.012} \end{bmatrix},$$

where estimated standard errors are given in parentheses. It is not clear what this matrix is actually telling us. Therefore we will perform an impulse response analysis next.

Since we intend to consider orthogonal impulse responses, the order of the variables may be of importance. Using the order $i^b, i^s, DD, M2$, interest rate and income shocks are permitted to have an instantaneous effect on $M2$, while $M2$ can only have a lagged impact on the other variables. It seems reasonable to assume that agents adjust their money demand quickly in response to interest rate shocks. However, the fact that the order of the variables may have an effect on the results has been a reason for substantial criticism of impulse response analyses and we will return to this point later.

In discussing the impulse responses we use the following convention. Suppose the system is in an equilibrium and that equilibrium is placed at the origin of our coordinate system, that is, all variables are zero in the equilibrium. Then an effect of a one-time impulse on a variable is called transitory if the variable returns to its previous equilibrium value of zero after some periods. If it does not return to zero and settles at a different equilibrium value, the effect is called permanent. Although expressions like long-run effect or permanent effect are sometimes used in a different sense in discussions of stationary systems, the aforementioned use of the terms is intuitively appealing in the present context and is therefore used here.

In fig. 2 we show the responses of DD and $M2$ to an orthogonalized impulse in DD of size one standard deviation. Estimated two standard error bounds are depicted as dashed lines. Obviously the impulse leads to a permanent (long-term) increase in DD but has no significant effect on $M2$. Note that after 30 quarters the variables have approximately reached their long-term positions in the sense that they remain constant if no further shocks hit the system. The reaction of the variable $M2$ contrasts sharply with the above interpretation of the cointegration relations and sheds doubt on the meaning of the restrictions we have imposed on the cointegration vectors. Therefore we have also computed impulse responses from a system estimated without constraints on the cointegration vectors. They are practically identical to those in fig. 2 and are therefore not given to save space.

In figs. 3 and 4 the effects of impulses in the short- and long-term interest rates on themselves and $M2$ are displayed. The impulses in both variables are not permanent. However, i^b leads to a long-term decline in $M2$ which is plausible if agents buy long-term bonds in response to the increase in the interest rate. On the other hand, there is no significant permanent effect on $M2$ due to a transitory increase in the short-term interest rate i^s. Again this is a plausible reaction because three-month money is included in $M2$. In all these cases almost the same impulse responses are obtained from a system without restrictions on the cointegration vectors.

In summary this exercise demonstrates that quite plausible reactions of $M2$ are compatible with the data. Since the impulse responses are not unique, it

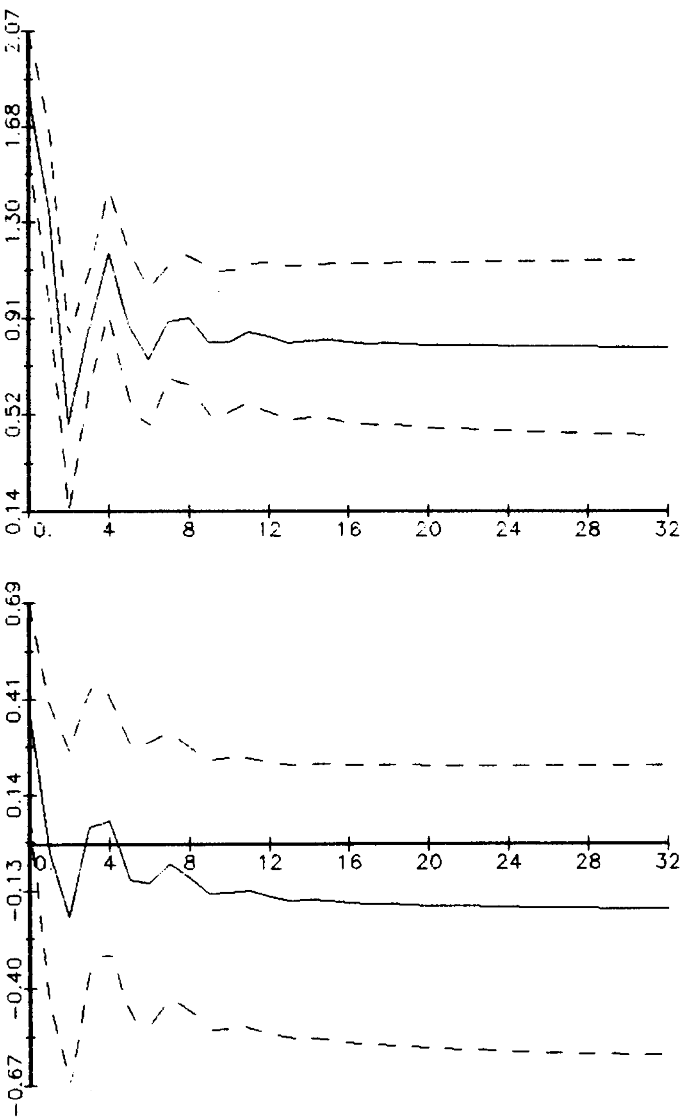

Fig. 2. Responses of ln DD (upper panel) and ln $M2$ (lower panel) to an impulse in ln DD, y-axis values $\times 10^{-2}$.

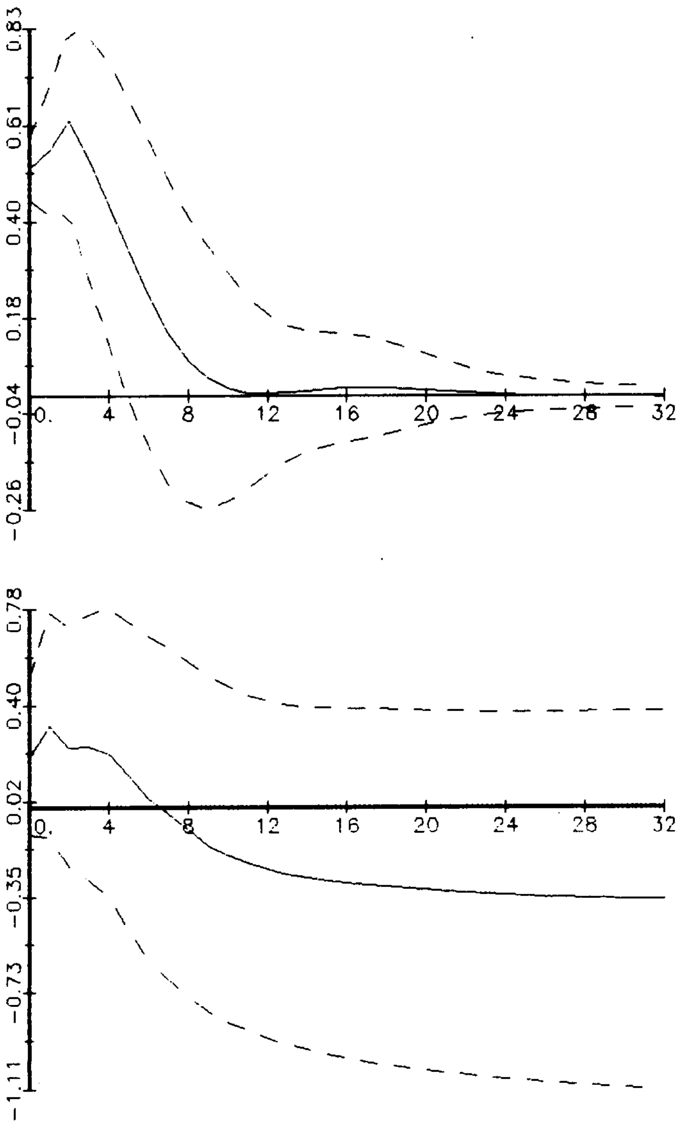

Fig. 3. Responses of i^s (upper panel) and ln *M2* (lower panel) to an impulse in i^s, y-axis values $\times 10^{-2}$.

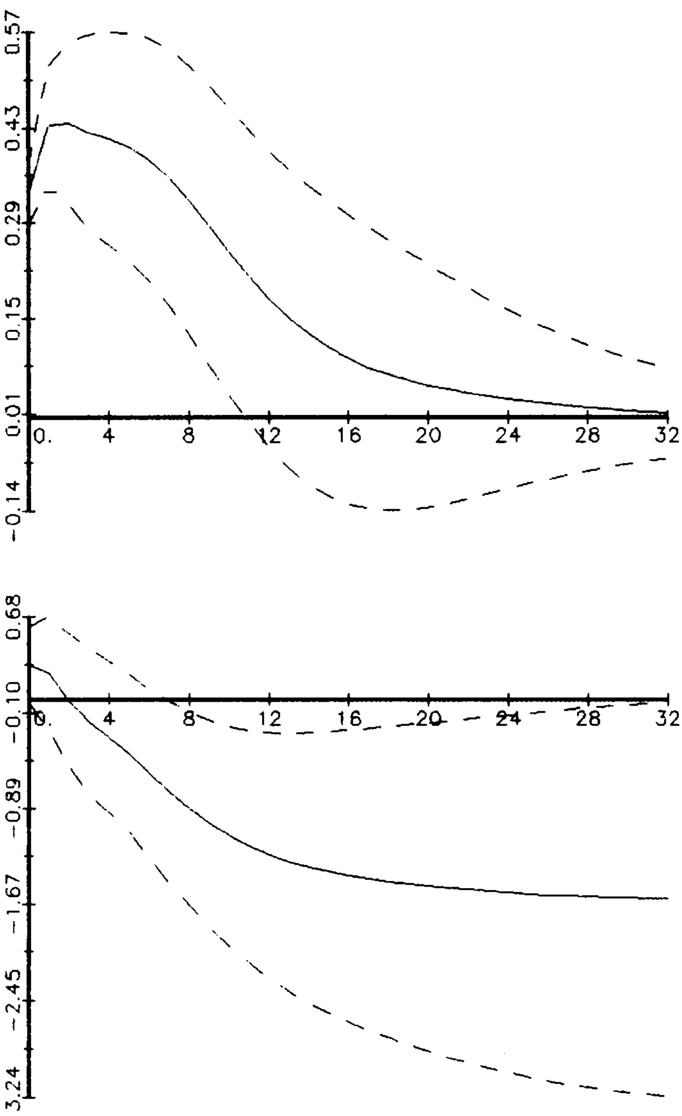

Fig. 4. Responses of i^b (upper panel) and ln *M2* (lower panel) to an impulse in i^b, y-axis values $\times 10^{-2}$.

cannot be concluded that other reactions of the system are impossible. To investigate the impact of the order of the variables, we have repeated the impulse response analysis with a reversed system, that is, the variables have the order $M2, DD, i^s, i^b$. In this case an impulse in DD resulted in a significant permanent decrease in $M2$ which is quite implausible. This result may be an indication that this order does indeed not reflect the actual reactions of the system. Alternatively, it may be an indication that the system is not adequate for studying the demand for money. In any case, this example shows that impulse response analysis can give interesting insights into a system that cannot be gained from the cointegration vectors alone.

5. Conclusions

In the previous sections we have discussed impulse response analysis in cointegrated systems. We have given the asymptotic distribution of the impulse responses and forecast error variance components of a Gaussian VAR process with cointegrated variables. The asymptotic theory is derived from Johansen's (1988, 1989) ML estimation procedure for such systems. We have demonstrated the practical feasibility of such an analysis using a system of West German macro economic data.

Appendix: Proof of Theorems 1 and 2

We give a proof of Theorem 2. Theorem 1 is then obtained by setting $R = I$ and $H = I$. The conditions stated in Theorem 2 are assumed to be satisfied. The notation of section 3 is used. In addition we denote by $\tilde{B}(C)$ and $\tilde{\Gamma}(C)$ the ML estimators of B and Γ given that C is known and we define $\bar{C} = (CC')(\tilde{C}C')^{-1}$. The following results of Johansen (1988, 1989) and Phillips and Durlauf (1986) will be helpful:

$$XX'/T \text{ has a well-defined fixed probability limit,} \tag{A.1}$$

$$Y_{-p}X'/T \text{ converges weakly,} \tag{A.2}$$

$$CY_{-p}Y'_{-p}/T \text{ converges weakly,} \tag{A.3}$$

$$Y_{-p}U'/T \text{ converges weakly,} \tag{A.4}$$

$$\operatorname{plim} CY_{-p}U'/T = 0, \tag{A.5}$$

$$Y_{-p}Y'_{-p}/T^2 \text{ converges weakly,} \tag{A.6}$$

$$T(\bar{C}\tilde{C} - C) \text{ converges weakly,} \tag{A.7}$$

$$\operatorname{plim} \sqrt{T}\left[\tilde{B}\tilde{C} - \tilde{B}(C)C\right] = 0. \tag{A.8}$$

We first demonstrate that the covariance matrix Σ_{c0}^{r} of the asymptotic normal distribution in Theorem 2 is precisely the one that is obtained when C is known.

Lemma 1. If C is known and $\tilde{\Gamma}(C)$ and $\tilde{B}_1(C)$ are the resulting ML estimators of Γ and B_1, respectively, and $\tilde{B}(C) = H\tilde{B}_1(C)$, then

$$\sqrt{T}\,\mathrm{vec}\left(\left[\tilde{\Gamma}(C), -\tilde{B}(C)C\right] - [\Gamma, -\Pi]\right) \xrightarrow{d} \mathrm{N}(0, \Sigma_{c0}^{r}).$$

Proof. With the notation of section 3 we have

$$\Delta Y = \Gamma X - HB_1C_1RY_{-p} + U. \tag{A.9}$$

Applying the vec operator gives

$$\begin{aligned}\mathrm{vec}(\Delta Y) &= (X' \otimes I)\mathrm{vec}(\Gamma) - (Y'_{-p}R'C'_1 \otimes H)\mathrm{vec}(B_1) + \mathrm{vec}(U)\\ &= \left[X' \otimes I, Y'_{-p}C' \otimes H\right]\begin{pmatrix}\mathrm{vec}(\Gamma)\\ -\mathrm{vec}(B_1)\end{pmatrix} + \mathrm{vec}(U).\end{aligned}$$

Since CY_{-p} is stationary, all the 'regressors' are stationary, and the ML estimator of

$$\begin{pmatrix}\mathrm{vec}(\Gamma)\\ -\mathrm{vec}(B_1)\end{pmatrix}$$

is known to have an asymptotic normal distribution with covariance matrix

$$\operatorname{plim} T\left(\begin{bmatrix}X \otimes I\\ CY_{-p} \otimes H'\end{bmatrix}(I \otimes \Sigma_u^{-1})\left[X' \otimes I, Y'_{-p}C' \otimes H\right]\right)^{-1} = \Omega_r^{-1}.$$

Noting that

$$\mathrm{vec}(\Gamma, -\Pi) = \begin{pmatrix}I & 0\\ 0 & C' \otimes H\end{pmatrix}\begin{pmatrix}\mathrm{vec}(\Gamma)\\ -\mathrm{vec}(B_1)\end{pmatrix}$$

gives the desired result. ■

As a consequence of Lemma 1 the asymptotic normal distribution in Theorem 2 is obtained from (A.8) and the following lemma:

Lemma 2. $\operatorname{plim}\sqrt{T}[\tilde{\Gamma} - \tilde{\Gamma}(C)] = 0.$

Proof. It is easily seen from (15) that

$$\tilde{\Gamma} = \left[\Delta Y + \tilde{B}\tilde{C}Y_{-p}\right]X'(XX')^{-1}$$

and

$$\tilde{\Gamma}(C) = \left[\Delta Y + \tilde{B}(C)CY_{-p}\right]X'(XX')^{-1}.$$

Hence,

$$\sqrt{T}\left[\tilde{\Gamma} - \tilde{\Gamma}(C)\right] = \sqrt{T}\left[\tilde{B}\tilde{C} - \tilde{B}(C)C\right]\frac{Y_{-p}X'}{T}\left(\frac{XX'}{T}\right)^{-1}$$

converges to zero in probability by (A.1), (A.2), and (A.8). ■

The next lemma implies that $\tilde{\Sigma}_u$ has the properties claimed in Theorem 2.

Lemma 3. $\operatorname{plim}\sqrt{T}(\tilde{\Sigma}_u - UU'/T) = 0.$

Proof.

$$\begin{aligned}
\tilde{\Sigma}_u &= \left(U + \Gamma X - \Pi Y_{-p} - \tilde{\Gamma}X + H\tilde{B}_1\tilde{C}_1RY_{-p}\right) \\
&\quad \times\left(U + \Gamma X - \Pi Y_{-p} - \tilde{\Gamma}X + H\tilde{B}_1\tilde{C}_1RY_{-p}\right)'/T \\
&= \frac{UU'}{T} + \left\{\frac{UX'}{T}(\Gamma - \tilde{\Gamma})' - \frac{UY'_{-p}\left(\Pi - H\tilde{B}_1\tilde{C}_1R\right)'}{T}\right. \\
&\quad + (\Gamma - \tilde{\Gamma})\frac{XU'}{T} - \frac{\left(\Pi - H\tilde{B}_1\tilde{C}_1R\right)Y_{-p}U'}{T} \\
&\quad + (\Gamma - \tilde{\Gamma})\frac{XX'}{T}(\Gamma - \tilde{\Gamma})' - \frac{\left(\Pi - H\tilde{B}_1\tilde{C}_1R\right)Y_{-p}X'}{T}(\Gamma - \tilde{\Gamma})' \\
&\quad - (\Gamma - \tilde{\Gamma})\frac{XY'_{-p}\left(\Pi - H\tilde{B}_1\tilde{C}_1R\right)'}{T} \\
&\quad \left. + \frac{\left(\Pi - H\tilde{B}_1\tilde{C}_1R\right)Y_{-p}Y'_{-p}\left(\Pi - H\tilde{B}_1\tilde{C}_1R\right)'}{T}\right\}.
\end{aligned}$$

Multiplying by $\sqrt{T}$, the term in curly brackets converges to zero in probability by (A.1)–(A.8). ■

It remains to show that

$$\tilde{\Sigma}_{c0}^{r} = T\begin{pmatrix} I & 0 \\ 0 & \tilde{C}' \otimes H \end{pmatrix}$$

$$\times \begin{pmatrix} XX' \otimes \tilde{\Sigma}_u^{-1} & XY_{-p}'\tilde{C}' \otimes \tilde{\Sigma}_u^{-1}H \\ \tilde{C}Y_{-p}X' \otimes H'\tilde{\Sigma}_u^{-1} & \tilde{C}Y_{-p}Y_{-p}'\tilde{C}' \otimes H'\tilde{\Sigma}_u^{-1}H \end{pmatrix}^{-1}$$

$$\times \begin{pmatrix} I & 0 \\ 0 & \tilde{C} \otimes H' \end{pmatrix}$$

is a consistent estimator of Σ_{c0}^{r}. In proving this result it must be kept in mind that $\tilde{C}$ is not a consistent estimator of C. However, we have

$$\tilde{\Sigma}_{c0}^{r} = T\begin{pmatrix} I & 0 \\ 0 & \tilde{C}'\overline{C}' \otimes H \end{pmatrix}$$

$$\times \begin{pmatrix} XX' \otimes \tilde{\Sigma}_u^{-1} & XY_{-p}'\tilde{C}\overline{C}' \otimes \tilde{\Sigma}_u^{-1}H \\ \overline{C}\tilde{C}\tilde{Y}_{-p}X' \otimes H'\tilde{\Sigma}_u^{-1} & \overline{C}\tilde{C}Y_{-p}Y_{-p}'\tilde{C}'\overline{C}' \otimes H'\tilde{\Sigma}_u^{-1}H \end{pmatrix}^{-1}$$

$$\times \begin{pmatrix} I & 0 \\ 0 & \overline{C}\tilde{C} \otimes H' \end{pmatrix},$$

and thus the consistency of $\tilde{\Sigma}_{c0}^{r}$ follows from (A.1)–(A.8).

As a final note we mention that, under the conditions of Theorem 1, $\hat{\Sigma}_u = S_{00} - \hat{B}\hat{B}'$ is equivalent to $\tilde{\Sigma}_u = (Y - \hat{\Gamma}X + \hat{B}\hat{C}Y_{-p})(Y - \hat{\Gamma}X + \hat{B}\hat{C}Y_{-p})'/T$ by a result of Johansen (1989).

References

Baillie, R.T., 1987, Inference in dynamic models containing 'surprise' variables, Journal of Econometrics 35, 101–117.

Campbell, J.Y. and R.J. Shiller, 1987, Cointegration and tests of present value models, Journal of Political Economy 95, 1062–1088.

Dickey, D.A., 1988, Comment, Journal of Business and Economic Statistics 6, 410–412.

Engle, R.F. and C.W.J. Granger, 1987, Co-integration and error correction: Representation, estimation and testing, Econometrica 55, 251–276.

Hylleberg, S. and G.E. Mizon, 1989, Cointegration and error correction mechanisms, Economic Journal 99, 113–125.

Johansen, S., 1988, Statistical analysis of cointegration vectors, Journal of Economic Dynamics and Control 12, 231–254.

Johansen, S., 1989, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Working paper (Institute of Mathematical Statistics, University of Copenhagen, Copenhagen).

Johansen, S. and K. Juselius, 1990a, Maximum likelihood estimation and inference on cointegration – with applications to the demand for money, Oxford Bulletin of Economics and Statistics 52, 169–210.

Johansen, S. and K. Juselius, 1990b, Some structural hypotheses in a multivariate cointegration analysis of the purchasing power parity and the uncovered interest parity for UK, Paper presented at the World Congress of the Econometric Society, Barcelona.

Judd, J.P. and J.L. Scadding, 1982, The search for a stable money demand function: A survey of the post-1973 literature, Journal of Economic Literature 20, 993–1023.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lütkepohl, and T.C. Lee, 1985, The theory and practice of econometrics, 2nd ed. (Wiley, New York, NY).

Judge, G.G., R.C. Hill, W.E. Griffiths, H. Lütkepohl, and T.C. Lee, 1988, Introduction to the theory and practice of econometrics, 2nd ed. (Wiley, New York, NY).

Juselius, K., 1989, Stationary equilibrium error processes in the Danish money market: An application of ML cointegration, Paper presented at the ESEM, Munich.

Lütkepohl, H., 1990, Asymptotic distributions of impulse response functions and forecast error variance decompositions of vector autoregressive models, Review of Economics and Statistics 72, 116–125.

Magnus, J.R., 1988, Linear structures (Charles Griffin, London).

Park, J.Y. and P.C.B. Phillips, 1989, Statistical inference in regressions with integrated processes: Part 2, Econometric Theory 5, 95–131.

Phillips, P.C.B and S.N. Durlauf, 1986, Multiple time series regression with integrated processes, Review of Economic Studies 80, 473–495.

Scheide, J., 1984, Geldpolitik, Konjunktur und rationale Erwartungen, Kieler Studien 188 (J.-C.B. Mohr, Tübingen).

Sims, C.A., J.H. Stock, and M.W. Watson, 1990, Inference in linear time series models with some unit roots, Econometrica 58, 113–144.

Stock, J.H., 1987, Asymptotic properties of least squares estimators of cointegrating vectors, Econometrica 55, 1035–1056.