Discussion Papers

Department of Economics
University of Copenhagen

No. 07-24

Some Identification Problems in the Cointegrated Vector Autoregressive Model

Søren Johansen

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel.: +45 35 32 30 82 – Fax: +45 35 32 30 00
http://www.econ.ku.dk

ISSN: 1601-2461 (online)

Some identification problems in the cointegrated vector autoregressive model

Søren Johansen*
Department of Economics, University of Copenhagen
and
CREATES, University of Aarhus

October 25, 2007

Abstract

An analysis of some identification problems in the cointegrated VAR is given. We give a new criteria for identification by linear restrictions on individual relations which is equivalent to the rank condition. We compare the asymptotic distribution of the estimators of α and β, when they are identified by linear restrictions on β, and when they are identified by linear restrictions on α, in which case a component of $\hat{\beta}$ is asymptotically Gaussian. Finally we discuss identification of shocks by introducing the contemporaneous and permanent effect of a shock and the distinction between permanent and transitory shocks, which allows one to identify permanent shocks from the long-run variance and transitory shocks from the short-run variance.

Keywords: Identification, cointegration, common trends

JEL Classification: C32

---

*The author wishes to thank the Danish Social Sciences Research Council for continuing support.
Address: Department of Economics, University of Copenhagen, Studiestræde 6, DK-1455 Copenhagen K, Denmark. Email: sjo@math.ku.dk

1 Introduction

We analyze some identification problems in the cointegrated vector autoregressive model, where it is well known that the adjustment and cointegration parameters enter through the impact matrix $\Pi = \alpha\beta'$. To get meaningful estimates of these parameters we therefore need to impose identifying restrictions. The long-run impact matrix C is the coefficient matrix of the cumulated shocks to the model and $C\sum_{i=1}^{t}\varepsilon_i$ defines the common stochastic trends generated by the permanent shocks $\alpha_\perp'\varepsilon_t$. We discuss identification of permanent shocks based upon the rows of C. Finally the transitory shocks $\alpha'\Omega^{-1}\varepsilon_t$ enter the conditional model of ΔX_t given $\alpha_\perp'\Delta X_t$ and the past via the matrix $B = \alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}$and we propose to discuss identification of transitory shocks via the rows of B.

As a result of the lack of identification, the asymptotic information matrix for the parameters (α, β) is singular, see Theorem 1. To find meaningful estimates we need to impose identifying restrictions on β or α, and this is usually done by restricting the individual cointegrating vectors by linear restrictions. We give a new criterion for identification, which is equivalent to the classical rank condition, see Lemma 5, and give some applications. In Theorem 7 we give the asymptotic distribution of the identified coefficients when they are identified by linear restrictions on β. If instead we impose linear restrictions on the adjustment coefficients we find that the asymptotic distribution of the cointegrating coefficient have a Gaussian component, see Theorem 9.

Finally we define permanent and transitory shocks and the contemporary and permanent effects of these. If we analyze the permanent shocks by a Cholesky decomposition of the long-run variance $C\Omega C'$ and the transitory shocks by a Cholesky decomposition of the short-run variance $B\Omega B'$, we need the asymptotic distribution of the linear combinations thus derived. These are found in Theorems 11 and 12.

2 The unrestricted cointegration model

We define the cointegrated VAR model, the cointegrating relations, and the commons stochastic trends. We then discuss estimation of β by an eigenvalue problem and find the asymptotic score and information under local alternatives.

2.1 Cointegration and common trends

We consider the $p-$dimensional process $X_t,\ t = 1, \ldots, T$, given by the cointegration vector autoregressive model

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k}\Gamma_i \Delta X_{t-i} + \varepsilon_t, \tag{1}$$

where α and β are $p \times r$ and ε_t are i.i.d. $N_p(0, \Omega)$. We have left out deterministic terms for simplicity. Under the assumption that the roots of the characteristic polynomial are either outside the unit circle or equal to one, and that $\det(\alpha_\perp'(I_p - \sum_{i=1}^k \Gamma_i)\beta_\perp) \neq 0$, the solution can be represented as

$$X_t = C\sum_{i=1}^t \varepsilon_i + Y_t + A, \tag{2}$$

where Y_t is stationary and A depends on initial values, so that $\beta' A = 0$. The long-run impact matrix matrix C is given by

$$C = \beta_\perp(\alpha_\perp'(I_p - \sum_{i=1}^k \Gamma_i)\beta_\perp)^{-1}\alpha_\perp'. \tag{3}$$

It follows from (2) and (3) that the cointegrating linear combinations, $\beta' X_t$, are stationary, and the nonstationarity of X_t is generated by the common stochastic trends

$$S_t = \alpha_\perp' \sum_{i=1}^t \varepsilon_i.$$

We also need the conditional model for ΔX_t given the past and $\alpha_\perp'\Delta X_t$, as given by

$$\Delta X_t = \alpha\beta' X_{t-1} + \omega\alpha_\perp'\Delta X_t + \sum_{i=1}^{k-1}(\Gamma_i - \omega\alpha_\perp'\Gamma_i)\Delta X_{t-i} + \varepsilon_t - \omega\alpha_\perp'\varepsilon_t \tag{4}$$

with $\omega = \Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}$, where $\varepsilon_t - \omega\alpha_\perp'\varepsilon_t = B\varepsilon_t$, and

$$B = I_p - \Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp' = \alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}. \tag{5}$$

Asymptotic inference for model (1) has been worked out, see for instance Johansen (1996). In the following we give some asymptotic results without detailed proofs, but appeal to the general idea that the asymptotic distribution of the maximum likelihood estimator can be found as the ratio of the limit under local alternatives of the score function and the information. As usual the estimators are derived from the Gaussian likelihood function and their properties given under general i.i.d. errors with mean zero and finite variance.

2.2 Estimation of the unrestricted parameters by reduced rank regression

If the parameters $(\alpha, \beta, \Gamma_1, \ldots, \Gamma_k, \Omega)$ are unrestricted, it is well known that the parameters can be estimated by reduced rank regression, see Anderson (1951). We use the Frisch-Waugh theorem to eliminate the short run dynamics $\Gamma_1, \ldots, \Gamma_k$ and define the residuals

$$R_{0t} = (\Delta X_t | \Delta X_{t-1}, \ldots, \Delta X_{t-k}) \text{ and } R_{1t} = (X_{t-1} | \Delta X_{t-1}, \ldots, \Delta X_{t-k})$$

and the product moments $S_{ij} = T^{-1}\sum_{t=1}^T R_{it}R_{jt}'$ and $S_{\varepsilon 1} = T^{-1}\sum_{t=1}^T \varepsilon_t R_{1t}$. The rest of the analysis is conducted in the profile likelihood function

$$\begin{aligned} -\log L_T(\alpha,\beta,\Omega) &= -\max_{\Gamma_1,\dots,\Gamma_1} \log L_T(\alpha,\beta,\Gamma_1,\dots,\Gamma_k,\Omega) \\ &= \frac{1}{2} tr[\Omega^{-1}\sum_{t=1}^T (R_{0t}-\alpha\beta' R_{1t})(R_{0t}-\alpha\beta' R_{1t})']. \end{aligned}$$

The parameter β is now estimated by solving an eigenvalue problem, and the maximized likelihood function is used to construct tests for rank and hypotheses on β, see Johansen (1996).

2.3 Asymptotic distribution of score and information under local alternatives

The local experiment, see van der Vaart (1988), is constructed by finding the limit under local alternatives of the parameters of the score and information. We introduce some notation. Let $f(x,y)$ be a matrix valued function of matrix arguments x and y, and let dx and dy denote directions of the same dimensions as x and y respectively. We define the partial derivative of f with respect to x in the direction dx by

$$D_x f(x,y)(dx) = \lim_{s\to 0} s^{-1}\{f(x+sdx),y) - f(x,y)\}.$$

We use the notation $\{A_{ij}\}$ for a block matrix with blocks A_{ij}, and $diag\{A_i\}$ for a block diagonal matrix with diagonal blocks A_i. Finally we write $A\otimes B = \{a_{ij}B\}$, and use throughout that if Z is a stochastic matrix with variance $Var(Z) = A\otimes B$, then for vectors μ,ν we have

$$Var(\mu' Z\nu) = \nu' A\nu\mu' B\mu.$$

We let W be the limit in distribution on $D^p[0,1]$ of $X_{[Tu]} = T^{-1/2}\sum_{i=1}^{[Tu]}\varepsilon_i$, and write

$$X_{[Tu]} \Longrightarrow CW(u). \tag{6}$$

Theorem 1 *The limit in distribution of the score with respect to β in the direction $T^{-1}\beta_\perp\bar\beta_\perp'(d\beta)$ is given by*

$$tr[\alpha'\Omega^{-1}\int_0^1 (dW)W'C'\beta_\perp(\bar\beta_\perp' d\beta)], \tag{7}$$

which is mixed Gaussian with asymptotic conditional variance

$$vec(\bar\beta_\perp' d\beta)'\left(\alpha'\Omega^{-1}\alpha\otimes\beta_\perp C\int_0^1 WW'duC'\beta_\perp\right)vec(\bar\beta_\perp' d\beta). \tag{8}$$

The score with respect to β, α in the directions $T^{-1/2}(\beta\bar{\beta}'(d\beta), (d\alpha)')$ is asymptotically Gaussian with mean zero and variance

$$vec(\bar{\beta}'(d\beta), (d\alpha)')' \left(\begin{pmatrix} \alpha'\Omega^{-1}\alpha & \alpha'\Omega^{-1} \\ \Omega^{-1}\alpha & \Omega^{-1} \end{pmatrix} \otimes \Sigma_{\beta\beta} \right) vec(\bar{\beta}'(d\beta), (d\alpha)'), \tag{9}$$

where $\Sigma_{\beta\beta} = Var(\beta' X_t | \Delta X_t, \ldots, \Delta X_{t-k+1})$.

The asymptotic distribution of the information is given by a block diagonal matrix with a block corresponding to $T^{-1}\beta_\perp \bar{\beta}_\perp' d\beta$ given by (8) and a block corresponding to $T^{-1/2}(\beta\bar{\beta}'(d\beta), (d\alpha)')$ given by the singular matrix (9).

PROOF. We let Ω be known to simplify the calculations, and define the concentrated likelihood function $l_T(\alpha, \beta) = -2 \log L_T(\alpha, \beta, \Omega)$. We first find the score with respect to β in the direction $T^{-1}\beta_\perp \bar{\beta}_\perp'(d\beta)$,

$$T^{-1}\mathsf{D}_\beta l_T(\alpha, \beta)(\beta_\perp \beta_\perp'(d\beta)) = -tr[\Omega^{-1} S_{\varepsilon 1} \beta_\perp \bar{\beta}_\perp'(d\beta)\alpha'],$$

and the limits in (7) and (8) follow from (6) and

$$S_{1\varepsilon} \xrightarrow{d} C \int_0^1 W(dW)',$$
$$T^{-1} S_{11} \xrightarrow{d} C \int_0^1 WW' du C'.$$

Next, the derivatives with respect to (β, α) in the directions $T^{-1/2}(\beta\bar{\beta}'(d\beta), (d\alpha)')$ are

$$T^{-1/2}\mathsf{D}_\beta l_T(\alpha, \beta)(\beta\bar{\beta}'(d\beta)) = -tr[\Omega^{-1} T^{1/2} S_{\varepsilon 1} \beta\bar{\beta}'(d\beta)\alpha'], \tag{10}$$
$$T^{-1/2}\mathsf{D}_\alpha l_T(\alpha, \beta)(d\alpha) = -tr[\Omega^{-1} T^{1/2} S_{\varepsilon 1} \beta (d\alpha)']. \tag{11}$$

It follows from

$$T^{1/2}\beta' S_{1\varepsilon} \xrightarrow{d} N_{r\times p}(0, \Omega \otimes \Sigma_{\beta\beta}),$$

that the scores are (jointly) asymptotically Gaussian with the mean zero and variance (9). The calculation of the information and its limit is given in the Appendix. ∎

It is seen that the asymptotic distribution of the maximum likelihood estimators cannot be derived from the results in Theorem 1 because the parameters are not identified or equivalently the asymptotic information (9) is singular. By imposing restrictions on α or β we restrict the variation of $d\alpha$ and $d\beta$ so that α and β become identified and the asymptotic information has full rank. This has the implication that there are cases in which $\hat{\beta} - \beta$ has a component in the direction β, so that $T^{1/2}(\bar{\beta}'(\hat{\beta} - \beta), (\hat{\alpha} - \alpha))$ is asymptotically Gaussian and in general correlated. This is discussed in sections 4 and 5, but first we give some criteria for identification by linear restrictions.

3 Definitions and criteria for identification

We give here the definition of identification and discuss some criteria for identification by linear restrictions on individual relations.

Definition 2 *Let $\{P_\theta, \theta \in \Theta\}$ be a family of probability measures, that is, a statistical model. We say that a parameter function $g(\theta)$ is identified if $g(\theta_1) \neq g(\theta_2)$ implies that $P_{\theta_1} \neq P_{\theta_2}$, or equivalently if $P_{\theta_1} = P_{\theta_2}$ implies that $g(\theta_1) = g(\theta_2)$.*

In the cointegration model (1), the parameter function $\Pi = \alpha\beta'$ and the parameters $\Gamma_1, \ldots, \Gamma_k, \Omega$ are identified, because if $\theta^i = (\alpha^i, \beta^i, \Gamma_1^i, \ldots, \Gamma_k^i, \Omega^i),\ i = 1, 2,$ and the (conditional) densities are equal

$$p(X_1, \ldots, X_T, \theta^1 | X_0, \ldots, X_{-k}) = p(X_1, \ldots, X_T, \theta^2 | X_0, \ldots, X_{-k}),$$

then $\Pi^1 = \alpha^1\beta^{1\prime} = \alpha^2\beta^{2\prime} = \Pi^2$, $\Gamma_i^1 = \Gamma_i^2$, $i = 1, \ldots, k$, and $\Omega^1 = \Omega^2$. On the other hand, given any choices of (α^1, β^1) and (α^2, β^2) for which $\Pi = \alpha^1\beta^{1\prime} = \alpha^2\beta^{2\prime}$, one can find a full rank $r \times r$ matrix for which $\alpha^1 = \alpha^2\xi'$ and $\beta^1 = \beta^2\xi^{-1}$. Thus the identification of α and β reduces to giving such restrictions on α or β, that if α and β satisfy the restrictions and $\beta\xi$ and $\alpha\xi'^{-1}$ satisfy the same restrictions, then $\xi = I_r$.

A general formulation of linear restrictions on individual cointegrating relations allows s_i linear restrictions the i'th vector: $R_i'\beta_i = 0,\ i = 1, \ldots, r$, where R_i is $p \times s_i$ of rank s_i. If we define $H_i = R_{i\perp}$, then the restrictions can be formulated as

$$R_i'\beta_i = 0 \text{ or } \beta_i = H_i\phi_i,\ \phi_i \in R^{m_i},\ i = 1, \ldots, r. \tag{12}$$

Thus β_i, or ϕ_i, has $m_i = p - s_i$ free parameters. The general definition of identification implies that β_i is identified (up to a constant factor) by (12) if it is the only linear combination of β which satisfies the same restriction. That is, if $R_i'\beta v = 0$, $v \in R^r$, implies that v is proportional to the i'th unit vector $v = \lambda e_i$.

Often we normalize the vectors and in this case the restrictions $R_i'\beta_i = 0$ can be formulated as

$$\beta_i = h_i + H^i\psi_i,\ \psi \in R^{m_i - 1}, \tag{13}$$

where $H_i = (h_i, H^i) = R_{i\perp}$, and H^i is $p \times (m_i - 1)$.

We give a simple example:

Example 1 If $r = 2$, a set of restrictions that is sometimes useful is

$$\beta' = \begin{pmatrix} 1 & 0 & \beta_{31} & \beta_{41} \\ 0 & 1 & \beta_{32} & \beta_{42} \end{pmatrix} = (I_2, -B), \tag{14}$$

which corresponds to solving the cointegrating relations for the first two variables, that is, if $X_t = (X_{1t}, X_{2t})$, each of dimension 2, we can write the cointegrating relations as

$$X_{1t} = BX_{2t} + u_t.$$

In the general formulation, (13), h_1 and h_2 are the first two unit vectors and $H^{i\prime} = (0_{2\times 2}, I_2)$ and ψ_i is 2×1. Clearly if ξ is 2×2 and $\beta\xi$ satisfies the same restrictions, than $\xi = I_2$, so that the restrictions in (14) identify β.

If instead

$$\beta' = \begin{pmatrix} 1 & 0 & \beta_{31} & \beta_{41} \\ \beta_{12} & 1 & 0 & \beta_{42} \end{pmatrix}, \tag{15}$$

then if $\beta\xi$ satisfy the same restrictions, one can see that $\xi_{11} = \xi_{22} = 1,\ \xi_{12} = 0$, and $\xi_{21}\beta_{31} = 0$. Thus if $\beta_{31} \neq 0$, then $\xi_{21} = 0$, and $\xi = I_2$, so that the restrictions (15) identify β, whereas if $\beta_{31} = 0$, then only the first relation is identified. In this case we say that β_2 is generically identified because the set of β for which β_2 is not identified is a small set which has Lebegue measure zero.■

The next lemma is the classical rank condition for identification, see for instance Fisher (1966).

Lemma 3 *(Rank condition) The vector β_i is identified up to a normalization by the s_i restrictions $R_i'\beta_i = 0$, if and only if the $s_i \times r$ matrix $R_i'\beta$ has rank $r - 1$.*

PROOF. We let $i = 1$. If rank$(R_1'\beta)$ =rank$(R_1'(\beta_2, \ldots, \beta_r)) < r - 1$, there would be an $(r-1)-$vector $v = (v_2, \ldots v_r) \neq 0$ for which $R_1'(\beta_2, \ldots, \beta_r)v = 0$. In this case the vector $\beta_1^* = \beta_1 + \sum_{i=2}^r \beta_i v_i \neq \lambda\beta_1$ would satisfy $R_1'\beta_1^* = 0$, so that β_1 is not identified.

If on the other hand β_1 is not identified identified, we can find an $r - 1$ vector $v \neq 0$, for which $\beta_1^* = (\beta_2, \ldots, \beta_r)v$ satisfies $R_1'\beta_1^* = 0$, but then rank$(R_1'\beta) = rank(R_1'(\beta_2, \ldots, \beta_r)) < r - 1$. ■

What is usually done in practice in order to implement this criterion, is to consider the determinant

$$P_1(\beta_2, \ldots, \beta_r) = |(\beta_2, \ldots, \beta_r)' R_1 R_1'(\beta_2, \ldots, \beta_r)|$$

which, if β_1 is identified, has rank $r - 1$. The polynomial $P_1(\beta_2, \ldots, \beta_r)$ is either identically zero, in which case no vector is identified by R_1 or there is only a finite number of roots, so that except for finitely many values of $(\beta_2, \ldots, \beta_r)$ the rank condition is satisfied and β_1 is identified by R_1. In this case we say that the restrictions $R_1, \ldots, R_r$ are generically identifying.

Example 2 Let

$$\beta' = \begin{pmatrix} 0 & \beta_{21} & \beta_{31} & \beta_{41} \\ \beta_{12} & 0 & \beta_{32} & \beta_{42} \end{pmatrix},$$

which satisfies the restrictions $R_i'\beta_i = 0$, where

$$R_1' = (1, 0, 0, 0), \quad R_2' = (0, 1, 0, 0).$$

In this case $P_1(\beta_2) = (R_1'\beta_2)^2 = \beta_{12}^2$, has only one zero, $\beta_{12} = 0$, so that R_1 is generically identifying, and β_1 is identified if $\beta_{12} \neq 0$. Similarly β_2 is identified if $\beta_{21} \neq 0$.■

We next give an elementary result from matrix algebra and apply this to reformulate the rank condition.

Lemma 4 *For any two matrices $\beta(p \times r)$ and $R(p \times s)$ of rank r and s respectively it holds that*

$$\mathit{rank}(R'\beta) = r - q \text{ if and only if } \mathit{rank}(\beta'_{\perp} R_{\perp}) = p - s - q.$$

PROOF. If if $\mathrm{rank}(R'\beta) \leq r - q$, there exists a matrix $\xi(r \times q)$ of rank q for which $R'\beta\xi = 0$, but then there exists a matrix $\phi((p - s) \times q)$ of rank q for which $\beta\xi = R_{\perp}\phi$, so that $0 = \beta'_{\perp}\beta\xi = \beta'_{\perp}R_{\perp}\phi$, which implies that $\mathrm{rank}(\beta'_{\perp}R_{\perp}) \leq p - s - q$. Thus

$$\mathrm{rank}(R'\beta) \leq r - q \text{ implies } \mathrm{rank}(\beta'_{\perp}R_{\perp}) \leq p - s - q.$$

By interchanging β with $R_{\perp}$ and R with $\beta_{\perp}$ and $p - s$ with r, we see that also the inverse implication holds and therefore the ranks have to be equal. ■

Lemma 5 *The relation β_i is identified up to a normalization by the s_i restrictions $R_i'\beta_i = 0$, if and only if the $(p - r) \times (p - s_i)$ matrix $\beta'_{\perp}H_i$ has rank $p - s_i - 1$.*

The lemma is a Corollary of Lemma 4 for $q = 1$.

Another algebraic criterion that does not depend on the parameters but only on the restrictions is given in

Lemma 6 *A necessary and sufficient condition that β_i is generically identified by $R_i'\beta_i = 0$ is that for any $k = 1, \ldots, r-1$ and any k indices $1 \leq i_1 < i_2 < \cdots < i_k \leq r$ for which $i_j \neq i$, we have*

$$\mathit{rank}(R_i'(R_{i_1\perp}, \ldots, R_{i_k\perp})) \geq k.$$

PROOF. To show that the condition is necessary we take $\beta_{i_j} = R_{i_j\perp}\phi_j$, $j = 1, \ldots, k$, for some ϕ_j. Then

$$\begin{aligned} \mathit{rank}(R_1'(R_{i_1\perp}, \ldots, R_{i_k\perp})) &\geq \mathit{rank}(R_1'(R_{i_1\perp}\phi_1, \ldots, R_{i_k\perp}\phi_k)) \\ &= \mathit{rank}(R_i'(\beta_{i_1}, \ldots, \beta_{i_k})) = k. \end{aligned}$$

That this criterion is also sufficient follows from a theorem of Rado (1942), see Johansen (1995). ■

It shows that based upon the restrictions alone it is possible to check for generic identification without knowing the parameter values.

4 Identification by linear restrictions on β

We give the well known asymptotic distribution of the estimated cointegrated relations when they are identified by individual linear restrictions.

4.1 Asymptotic distribution of $\hat{\beta}$ and $\hat{\alpha}$ with identifying restrictions on β

Theorem 7 *If β is identified by the restrictions $\beta_i = h_i + H^i\psi_i, i = 1, \ldots, r$, it holds that the limit of the information matrix for the parameters ψ_i, $i = 1, \ldots, r$, is the full rank matrix*

$$\{\alpha_i'\Omega^{-1}\alpha_j H^{i\prime} C \int_0^1 WW'duC'H^j\}.$$

The asymptotic distribution of $Tvec(\hat{\beta} - \beta)$ is mixed Gaussian with asymptotic conditional variance

$$diag\{H^i\}\{\alpha_i'\Omega^{-1}\alpha_j H^{i\prime} C \int_0^1 WW'duC'H^j\}^{-1} diag\{H^j\}',$$

which can be estimated by

$$Tdiag\{H^i\}\{\hat{\alpha}_i'\hat{\Omega}^{-1}\hat{\alpha}_j H^{i\prime} S_{11} H^j\}^{-1} diag\{H^j\}'. \tag{16}$$

The asymptotic distribution of $T^{1/2}vec(\hat{\alpha} - \alpha)$ is Gaussian with mean zero and variance matrix

$$\Sigma_{\beta\beta}^{-1} \otimes \Omega. \tag{17}$$

The estimators $\hat{\beta}$ and $\hat{\alpha}$ are asymptotically independent.

PROOF. The score with respect to β in the direction $T^{-1}H^i d\psi_i$ is

$$\begin{aligned} -tr[\Omega^{-1}S_{\varepsilon 1}H^i(d\psi_i)\alpha_i'] &= -tr[\Omega^{-1}S_{\varepsilon 1}(\beta\bar{\beta}' + \beta_\perp\bar{\beta}_\perp')H^i(d\psi_i)\alpha_i'] \\ &\xrightarrow{d} -tr[\alpha_i'\Omega^{-1}\int_0^1 (dW)W'C'H^i(d\psi_i)], \end{aligned}$$

so that in the limit the directions $\beta\bar{\beta}'H^i\psi_i$ play no role. The limit of the vector of scores with respect to $T^{-1}(H^1(d\psi_1), \ldots, H^r(d\psi_r))$ is therefore mixed Gaussian with asymptotic conditional variance matrix

$$vec(\psi)'\{\alpha_i'\Omega^{-1}\alpha_j H^{i\prime} C \int_0^1 WW'duC'H^j\}vec(\psi),$$

where $vec(\psi) = (\psi_1', \ldots, \psi_r')'$. This is also the limit of the information matrix. It follows from Lemma (5) that this matrix has full rank because if v_i, $i = 1, \ldots, r$ are vectors so that

$$\begin{aligned} &\sum_{i=1}^r \sum_{j=1}^r \alpha_i'\Omega^{-1}\alpha_j v_i' H^{i\prime} C \int_0^1 WW'duC'H^j v_j \\ =\ & tr[\alpha'\Omega^{-1}\alpha\{v_i'H^{i\prime}C\int_0^1 WW'duC'H^j v_j\}] = 0, \end{aligned}$$

then $v_i' H^{i\prime} C \int_0^1 WW' du C' H^j v_j = 0$, and hence $\beta_\perp' H^j v_j = 0$, but this implies by Lemma 5 that $v_j = 0$. We can therefore find the asymptotic distribution of $T\hat{\psi}_i$ by dividing the score by the information.

The asymptotic distribution of the score with respect to α in the direction $T^{-1/2}(d\alpha), T^{-1/2}\mathsf{D}_\alpha l_T(\alpha,\beta)(d\alpha) = -tr[\Omega^{-1}T^{1/2}S_{\varepsilon 1}\beta(d\alpha)']$ see (11), is Gaussian with mean zero and variance

$$vec((d\alpha)')'(\Sigma_{\beta\beta} \otimes \Omega^{-1})vec((d\alpha)'),$$

which is also the limit of the information, and that proves (17). Finally the estimates are asymptotically independent because

$$\begin{aligned} & T^{-3/2}D^2_{\alpha\beta}l_T(\alpha,\beta)(d\alpha_i, H^i(d\psi_i)) \\ = \ & -T^{-1/2}tr\{\Omega^{-1}S_{\varepsilon 1}H^i(d\psi_i)(d\alpha_i)') + T^{-1/2}tr\{\Omega^{-1}(d\alpha_i)\beta_i' S_{11}H^i d\psi_i)\alpha_i') \xrightarrow{P} 0. \end{aligned}$$

■

The result above is derived under the assumption that all vectors in β are identified. If only some are identified, we can impose just identifying restrictions on the remaining ones and apply the result.

It may therefore appear that the asymptotic distribution of an identified vector, $T(\hat{\beta}_1 - \beta_1)$ say, could depend on how some of the other columns of β are being just identified. It is, however, a consequence of Lemma 5 that this is not the case.

Theorem 8 *Let $\beta = (\beta_1, \ldots, \beta_r)$ be identified by $R_1, \ldots, R_r$, and assume that β_r is just identified. Then the asymptotic conditional variance for the vector $Tvec(\hat{\beta}_1 - \beta_1, \ldots, \hat{\beta}_{r-1} - \beta_{r-1})$ does not depend on how β_r is just identified, and it is estimated by the relevant block of (16).*

PROOF. Let $\rho_{ij} = \alpha_i'\Omega^{-1}\alpha_j$, and $M = \bar{\beta}_\perp' C \int_0^1 WW' du C' \bar{\beta}_\perp$. The asymptotic conditional variance of $Tvec(\hat{\psi}_1 - \psi_1, \ldots, \hat{\psi}_{r-1} - \psi_{r-1})$ is given by the inverse of a matrix with i, j'th block

$$\begin{aligned} & \rho_{ij}H^{i\prime}\beta_\perp M \beta_\perp' H^j - \rho_{ir}H^{i\prime}\beta_\perp M\beta_\perp' H^r(\rho_{rr}H^{r\prime}\beta_\perp M\beta_\perp' H^r)^{-1}\rho_{rj}H^{r\prime}\beta_\perp M\beta_\perp' H^j \\ = \ & (\rho_{ij} - \frac{\rho_{ir}\rho_{rj}}{\rho_{rr}})H^{i\prime}\beta_\perp M\beta_\perp' H^j, \end{aligned}$$

where we use that $H^{r\prime}\beta_\perp$ is $(p-r) \times (p-r)$ of full rank, see Lemma 5, for any choice of a just identifying H^r. Next we show that the coefficient $\rho_{ij} - \frac{\rho_{ir}\rho_{rj}}{\rho_{rr}}$ does not depend on the choice of H^r. We find

$$\begin{aligned} \rho_{ij} - \frac{\rho_{ir}\rho_{rj}}{\rho_{rr}} &= \alpha_i'\Omega^{-1}\alpha_j - \alpha_i'\Omega^{-1}\alpha_r(\alpha_r'\Omega^{-1}\alpha_r)^{-1}\alpha_r'\Omega^{-1}\alpha_j \\ &= \alpha_i'\alpha_{r\perp}(\alpha_{r\perp}'\Omega\alpha_{r\perp})^{-1}\alpha_{r\perp}'\alpha_j, \end{aligned}$$

so that the adjustment coefficient α_r only enters through its orthogonal complement. If $\beta_r^* = \sum_{i=1}^r \beta_i v_i$ ($v_r \neq 0$) is chosen so that β_r^* is just identified by some other restrictions $R_r^{*\prime}\beta_r^* = 0$, then

$$\sum_{i=1}^{r} \alpha_i \beta_i' = \sum_{i=1}^{r-1} (\alpha_i - \alpha_r \frac{v_i}{v_r})\beta_i' + \alpha_r \frac{1}{v_r}\beta_r^* = \sum_{i=1}^{r-1} \alpha_i^* \beta_i' + \alpha_r^* \beta_r^{*\prime},$$

which shows that the adjustment coefficient to β_i^* is $\alpha_i^* = \alpha_i - \alpha_r \frac{v_i}{v_r}$, and the coefficient to β_r^* is $\alpha_r^* = \alpha_r \frac{1}{v_r}$, so that $\alpha_{r\perp}^* = \alpha_{r\perp}$. Therefore

$$\begin{aligned} \rho_{ij}^* - \frac{\rho_{ir}^*\rho_{rj}^*}{\rho_{rr}^*} &= \alpha_i^{*\prime}\alpha_{r\perp}^*(\alpha_{r\perp}^{*\prime}\Omega\alpha_{r\perp}^*)^{-1}\alpha_{r\perp}^{*\prime}\alpha_j^* = \alpha_i^{*\prime}\alpha_{r\perp}(\alpha_{r\perp}'\Omega\alpha_{r\perp})^{-1}\alpha_{r\perp}'\alpha_j^* \\ &= \alpha_i'\alpha_{r\perp}(\alpha_{r\perp}'\Omega\alpha_{r\perp})^{-1}\alpha_{r\perp}'\alpha_j = \rho_{ij} - \frac{\rho_{ir}\rho_{rj}}{\rho_{rr}}. \end{aligned}$$

This shows that the asymptotic conditional variance does not depend on how the last vector is just identified. ■

5 Identification of β using linear restrictions on α

Another possibility for identifying β is to identify the r cointegrating relations as the first r rows of the matrix Π. This corresponds to identifying β by choosing $\alpha' = (I_r, \alpha_2')$, see (14). A general formulation of this is

$$\alpha_i = a_i + A^i\psi_i,\ \ \psi_i \in R^{m_i - 1},\ \ i = 1, \ldots, r,$$

or $(a_i, A^i)_\perp'\alpha_i = 0$, see (13). We give the asymptotics of $\hat{\beta}$ and $\hat{\alpha}$ in this case.

5.1 Asymptotic distribution of $\hat{\alpha}$ and $\hat{\beta}$ with identifying restrictions on α

Theorem 9 *Under the identification $\alpha_i = a_i + A^i\psi_i$ the limit of the information matrix in the directions $\bar{\beta}'(d\beta_1, \ldots, d\beta_r), d\psi_1, \ldots, d\psi_r$ is*

$$\mathcal{I} = \begin{pmatrix} \{e_i'\alpha'\Omega^{-1}\alpha e_j \Sigma_{\beta\beta}\} & \{\Sigma_{\beta\beta} e_l e_i' \alpha'\Omega^{-1}A_l\} \\ \{A_k'\Omega^{-1}\alpha e_j e_k' \Sigma_{\beta\beta}\} & \{e_k'\Sigma_{\beta\beta} e_l A_k'\Omega^{-1}A_l\} \end{pmatrix}, \tag{18}$$

which has full rank. The asymptotic distribution of $T(vec(\bar{\beta}_\perp'(\hat{\beta} - \beta)))$ is mixed Gaussian with asymptotically conditional variance

$$(\alpha'\Omega^{-1}\alpha \otimes \beta_\perp' C \int_0^1 WW'duC'\beta_\perp)^{-1} \tag{19}$$

and $T^{1/2}(vec(\bar{\beta}'(\hat{\beta} - \beta))', \hat{\psi}_1 - \psi_1, \ldots, \hat{\psi}_r - \psi_r)$ is asymptotically Gaussian with mean zero and variance matrix $\mathcal{I}^{-1}$, see (18).

In the special case where $A_i = A,\ i = 1, \dots, r,$ the covariance matrix $\mathcal{I}^{-1}$ becomes

$$\begin{pmatrix} \{e_i'(\alpha' A_\perp (A_\perp' \Omega A_\perp)^{-1} A_\perp' \alpha)^{-1} e_j \Sigma_{\beta\beta}^{-1}\} & -\{\Sigma_{\beta\beta}^{-1} e_l e_i' (\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1} A M\} \\ -\{M A' \Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1} e_j e_k' \Sigma_{\beta\beta}^{-1}\} & \{e_k' \Sigma_{\beta\beta}^{-1} e_l M\} \end{pmatrix}, \quad (20)$$

where $M = (A'\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp' A)^{-1}$. When we identify β by $\alpha = (I_r, \alpha_2')'$ we have $A = (0, I_{p-r})$, and find the asymptotic variance matrix for $T^{1/2}(vec(\bar{\beta}'(\hat{\beta} - \beta)))$

$$\Sigma_{\beta\beta}^{-1} \otimes \Omega_{1:r,1:r}.$$

Proof. Proof see the appendix ■

The results of Theorem 9 show that the asymptotic distribution of $\hat{\beta}$ has two components: one in the direction of β which is $T^{1/2}$ consistent and one in the direction of $\beta_\perp$ which is T consistent. It follows that

$$T^{1/2}(\hat{\beta} - \beta) = T^{1/2}\beta\bar{\beta}'(\hat{\beta} - \beta) + T^{1/2}\beta_\perp\bar{\beta}_\perp'(\hat{\beta} - \beta) = T^{1/2}\beta\bar{\beta}'(\hat{\beta} - \beta) + o_P(1).$$

Thus, the distribution of $T^{1/2}(\hat{\beta} - \beta)$ is asymptotically Gaussian, but with a singular covariance matrix, so there are some hypotheses that one cannot test using the Gaussian distribution. We illustrate by an example.

Example 3 Suppose we have four variables and two cointegrating relations, and we identify using $\alpha' = (I_2, \alpha_2')$ so that the first two rows of Π are the identified cointegrating parameters. Now assume that the true value β has $\beta_{11} = 0$. We can test this hypothesis by applying the asymptotic Gaussian distribution of $\hat{\beta}_{11}$:

$$\frac{T^{1/2}\hat{\beta}_{11}}{\sqrt{e_1'\hat{\Omega}e_1 e_1'\hat{\beta}\Sigma_{\hat{\beta}\hat{\beta}}^{-1}\hat{\beta}' e_1}} \xrightarrow{d} N(0,1), \quad (21)$$

where

$$e_1'\beta\Sigma_{\beta\beta}^{-1}\beta' e_1 = (\beta_{11}, \beta_{12})\Sigma_{\beta\beta}^{-1}(\beta_{11}, \beta_{12})' = (0, \beta_{12})\Sigma_{\beta\beta}^{-1}(0, \beta_{12})' = (\beta_{12})^2(\Sigma_{\beta\beta}^{-1})_{22}.$$

Thus we see that (21) only holds if $\beta_{12} \neq 0$. In this case we let $B_T = (\beta, T^{-1/2}\beta_\perp)$ and find a consistent estimator of $e_1'\beta\Sigma^{-1}\beta' e_1$ from

$$\begin{aligned} e_1' S_{11}^{-1} e_1 &= e_1' B_T (B_T' S_{11} B_T)^{-1} B_T' e_1 & (22) \\ &= e_1'\beta\Sigma_{\beta\beta}^{-1}\beta' e_1 + T^{-1} e_1'\beta_\perp (T^{-1}\beta_\perp' S_{11}\beta_\perp)^{-1}\beta_\perp' e_1 + o_P(1) \xrightarrow{P} e_1'\beta\Sigma_{\beta\beta}^{-1}\beta' e_1. \end{aligned}$$

Thus a simulation experiment will show that $\hat{\beta}_{11}$ is approximately Gaussian in case $\beta_{12} \neq 0$. If, however, $\beta_{12} = 0$, that is, β has the first row equal to zero, $\beta' e_1 = 0$, we cannot use the asymptotic result (21), and find instead that

$$\begin{aligned} T\hat{\beta}_{11} &= Te_1'(\hat{\beta} - \beta)e_1 = Te_1'\beta\bar{\beta}'(\hat{\beta} - \beta)e_1 + Te_1'\beta_\perp\bar{\beta}_\perp'(\hat{\beta} - \beta)e_1 = Te_1'\beta_\perp\bar{\beta}_\perp'(\hat{\beta} - \beta)e_1 \\ &\xrightarrow{d} e_1'\beta_\perp(\beta_\perp' C \int_0^1 WW' du C'\beta_\perp)^{-1}\beta_\perp' C \int_0^1 W(dW)'\Omega^{-1}\alpha_1, \end{aligned}$$

and because $e_1'\beta = 0$, we have $e_1'\beta_\perp \neq 0$. Thus in case $\beta_{12} = 0$, the simulation above will show that $\hat{\beta}_{11}$ is not approximately Gaussian, but mixed Gaussian, which means that we no longer want to normalize by the standard error, but by an estimate of the asymptotic conditional standard error. Thus in this case inference is not based on the asymptotic distribution of $\hat{\beta}_{11}$ but on the joint asymptotic distribution of $\hat{\beta}_{11}$ and the observed information $e_1'\hat{\Omega}e_1e_1'S_{11}^{-1}e_1$.

Thus, the mixed Gaussian limit has to be used to make inference on β_{11}. Note, however, that one can use the same statistic:

$$\frac{T^{1/2}\hat{\beta}_{11}}{\sqrt{e_1'\hat{\Omega}e_1e_1'S_{11}^{-1}e_1}} = \frac{T\hat{\beta}_{11}}{\sqrt{e_1'\hat{\Omega}e_1e_1'(T^{-1}S_{11})e_1}} \xrightarrow{d} N(0,1),$$

because in this case, it follows from (22) that when $e_1'\beta = 0$,

$$\begin{aligned} e_1'(T^{-1}S_{11})^{-1}e_1 &= e_1'\beta_\perp(T^{-1}\beta_\perp'S_{11}\beta_\perp)^{-1}\beta_\perp'e_1 + o_P(1) \\ &\xrightarrow{d} e_1'\beta_\perp(\beta_\perp'C\int_0^1 WW'duC'\beta_\perp)^{-1}\beta_\perp'e_1, \end{aligned}$$

so that $T\hat{\beta}_{11}$ is normalized by an estimator of the square root of its asymptotic conditional variance, which is the appropriate scaling factor, and the limit is again $N(0,1)$.

The difference between the situation $\beta_{21} = 0$ and $\beta_{21} \neq 0$, is that if $\beta_{21} \neq 0$, the hypothesis $\beta_{11} = 0$ does not change the cointegration space. When $\beta_{21} \neq 0$, it is possible to change the coordinate system inside sp(β) and obtain $\beta_{11} = 0$, that is find a linear combination of β_1 and β_2 with first coefficient zero. If $\beta_{21} = 0$, however, this is not possible in general and the hypothesis $\beta_{11} = 0$, becomes a hypothesis on the cointegrating space leading to mixed Gaussian inference.■

5.1.1 Discussion.

The mathematical formulation of all this is that it is the cointegration space which is estimated superconsistently and the estimator is mixed Gaussian and asymptotically independent of the limit of the remaining parameters. The problem is how to reformulate this statement into something that is economically useful. This is what is done by the identifying restrictions of the form $R_i'\beta_i = 0$, whereby the cointegrating space is parameterized using economically meaningful parameters. On the other hand, identifying the cointegration space as the first two rows, say, of the matrix Π, that is, as the stationary linear combination to which the first two variables react, is clearly an identification of the cointegration space, but it also specifies which vectors in that space we are interested in. Thus if the first two rows of Π are

$$\beta' = \begin{pmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \beta_{41} \\ \beta_{12} & \beta_{22} & \beta_{32} & \beta_{42} \end{pmatrix},$$

it is clear that by multiplying by the matrix

$$\begin{pmatrix} \beta_{22} & -\beta_{21} \\ -\beta_{12} & \beta_{11} \end{pmatrix},$$

we get

$$\begin{pmatrix} \beta_{11}\beta_{22}-\beta_{12}\beta_{21} & 0 & * & * \\ 0 & \beta_{11}\beta_{22}-\beta_{12}\beta_{21} & * & * \end{pmatrix}.$$

This does not change the cointegrating space, but we can achieve zeros at the positions 12 and 21. Thus the hypothesis that, say, $\beta_{12}=\beta_{21}=0$, is clearly testable, because the parameters are identified, but it is not a hypothesis on the cointegration space and hence does not involve mixed Gaussian inference.

6 Shocks and their effect

The cointegrated vector autoregressive model associates with each variable X_{it} a contemporaneous shock ε_{it}, which is the unanticipated part of X_{it} given the past. The identification of shocks is often done by postulating an ordering of the variables, and therefore the corresponding shocks, and perform a Cholesky decomposition of the variance Ω. If we decompose the shock into a permanent $\alpha_\perp'\varepsilon_t$ and a transitory part, $\alpha'\Omega^{-1}\varepsilon_t$, there is no natural ordering because of the lack of identification of $(\alpha,\alpha_\perp)$. We propose to give names to the permanents shocks by considering their effects on the process, that is, by analyzing the long-run variance

$$C\Omega C',$$

see (3). Similarly we give names to the transitory shocks by considering their effect on the equations given the long-run shocks. This leads to an analysis of the short-run variance

$$B\Omega B'=\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'.$$

If we apply a Cholesky decomposition of $C\Omega C'$ or $B\Omega B'$ we find linear combinations of shocks. We end this section by finding the asymptotic variances of these estimated linear combinations.

6.1 Permanent and transitory shocks and their contemporaneous and permanent effects

Model (1) can be written as

$$X_t=(I_p+\alpha\beta')X_{t-1}+\sum_{i=1}^{k}\Gamma_i\Delta X_{t-i}+\varepsilon_t.$$

This shows that a change in ε_t ($\varepsilon_t\mapsto\varepsilon_t+c$) is equivalent to a change in X_t ($X_t\mapsto X_t+c$). We now discuss the effect of such a change on later values X_{t+h}, using the Granger representation theorem (2). We find

$$X_{t+h}=C\sum_{i=1}^{t+h}\varepsilon_i+\sum_{i=0}^{\infty}C_i\varepsilon_{t+h-i}+A,$$

which shows that the effect at time $t+h$ of a change $c \in R^p$ to ε_t (or X_t) is

$$\frac{\partial X_{t+h}}{\partial \varepsilon_t} c = (C + C_h)c \to Cc, h \to \infty.$$

A change c to the system at time t propagates through the system and becomes Cc in the long run, that is, the permanent effect of a change is Cc. Moreover from the expression for C, see (3), we see that the part of the shock ε_t, which has an effect in the long run is $\alpha_\perp' \varepsilon_t$.

We note the decomposition of ε_t into independent components

$$\varepsilon_t = \alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}\varepsilon_t + \Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp'\varepsilon_t. \tag{23}$$

Based on this we use the definitions

Definition 10 *We call ε_t the shock at time t, and define the permanent shock as $\alpha_\perp'\varepsilon_t$ and the transitory shock as $\alpha'\Omega^{-1}\varepsilon_t$. We define the contemporaneous effect of ε_t as I_p.*

From the decomposition (23) of the shock ε_t, we find the contemporaneous effect of the permanent shock, $\alpha_\perp'\varepsilon_t$, is $\Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}$, and that of a transitory shock, $\alpha'\Omega^{-1}\varepsilon_t$, is $\alpha(\alpha'\Omega^{-1}\alpha)^{-1}$.

Finally the long-run effect of a shock ε_t is defined to be C, of a permanent shock it is $C\Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1} = \beta_\perp(\alpha_\perp'\Gamma\beta_\perp)^{-1}$, and of a transitory shock it is $C\alpha(\alpha'\Omega^{-1}\alpha)^{-1} = 0$.

7 Identification of permanent shocks

We start by an example.

Example 4 Let us assume we have among others the real variable y_t^r and the nominal variable p_t, and that we have two permanent shocks, which we want to call a real shock and a nominal shock. The Granger representation of the variables is

$$\begin{aligned}
y_t^r &= e_1'C\sum_{i=1}^t \varepsilon_i + z_{1t} = c_1'\sum_{i=1}^t \varepsilon_i + z_{1t}, \\
p_t &= e_2'C\sum_{i=1}^t \varepsilon_i + z_{2t} = c_2'\sum_{i=1}^t \varepsilon_i + z_{2t}, \\
&\vdots
\end{aligned}$$

It seems clear that we want the permanent shocks to be defined by $\alpha_\perp'\varepsilon_t$. The problem is to identify which shock can influence which variable.

Let us assume that

A nominal shock cannot have a permanent influence on a real variable

From this it follows that the only permanent shock which appears in a real variable is a permanent real shock.

Thus we define (or identify) the permanent real shock as that which generates the random walk in the real variable, that is, the linear combination given by the row of C which corresponds to the real variable.

It seems natural then to define (or identify) the permanent nominal shock as the linear combination given by that row of C which corresponds to the nominal variable but orthogonalized to the permanent real shock.■

If we want to announce that the cumulated (permanent) shocks which enter into the 'real' variable X_{1t} is generated by the 'permanent real' shocks to the economy, we can define the 'permanent real' shock as

$$v_{1t} = c_1'\varepsilon_t = e_1'C\varepsilon_t, \tag{24}$$

where c_1' is the first row of the matrix C corresponding to X_{1t}. If we next want to define the 'permanent nominal' shock as that part of the cumulated shock in the 'nominal' variable X_{2t} which is independent of $c_1'\varepsilon_t$ we define

$$v_{2t} = c_2'\varepsilon_t - c_2'\Omega c_1(c_1'\Omega c_1)^{-1}c_1'\varepsilon_t. \tag{25}$$

If more orthogonalized shocks are needed the Cholesky decomposition is continued by choosing further variables, that is, rows of C until we have identified $p-r$ shocks. We have to choose the variables so that the rows c_i are linearly independent, that is, so that the next vector chosen is linearly independent of those previously chosen, or so that the next variable is not cointegrating with those already chosen. In general when we have chosen $c_{1:m} = (c_1, \dots, c_m)$, we define

$$v_{m+1.t} = c_{m+1.1:m}'\varepsilon_t = c_{m+1}'\varepsilon_t - c_{m+1}'\Omega c_{1:m}(c_{1:m}'\Omega c_{1:m})^{-1}c_{1:m}'\varepsilon_t.$$

Note that we can estimate $C\varepsilon_t$ from

$$\hat{C}R_{0t} = \hat{C}(\hat{\alpha}\hat{\beta}'R_{1t} + \hat{\varepsilon}_t) = \hat{C}\hat{\varepsilon}_t.$$

The above discussion of defining the permanent shocks by their long-run properties, is of course not new, see Blanchard and Quah (1989), but here the different terms are defined in the cointegrated VAR. A discussion of long-run propositions as restrictions on $\beta_\perp$ is found in the paper by Lin (2007).

Thus, instead of the Cholesky decomposition of Ω, based upon an ordering of the variables, one may choose a Cholesky decomposition of the long-run variance $C\Omega C'$, in order to identify the permanent shocks.

7.1 The variance of the weights of the permanent shocks

We find here the asymptotic Gaussian distribution of the estimators of the weights found by the Cholesky decomposition of the long-run variance $C\Omega C'$.

Theorem 11 *The asymptotic variance of the coefficients*

$$\hat{c}_1' = e_1'\hat{C}$$

is given by

$$asVar(\hat{c}_1') = \xi'\Sigma^{-1}\xi c_1'\Omega c_1 \tag{26}$$

whereas the asymptotic variance of $\hat{c}_{2.1}$ *is given by*

$$\begin{aligned} asVar(\hat{c}_{2.1}) &= c_{2.1}'(\Omega + \Omega\xi'\Sigma^{-1}\xi\Omega)c_{2.1}C_1 \\ &\quad +(c_{2.1}'\Omega c_{2.1})\Omega^{-1}C_{1\perp}\xi'\Sigma^{-1}\xi C_{1\perp}\Omega^{-1}, \end{aligned} \tag{27}$$

where we have used the notation

$$\begin{aligned} \Sigma &= EZ_tZ_t', \\ Z_t' &= (X_{t-1}'\beta, \Delta X_{t-1}', \ldots, \Delta X_{t-k+1}'), \\ \xi' &= ((C'\Gamma' - I_p)\bar{\alpha}, C', \ldots, C') \\ C_1 &= c_1(c_1'\Omega c_1)^{-1}c_1' \\ C_{1\perp} &= c_{1\perp}(c_{1\perp}'\Omega^{-1}c_{1\perp})^{-1}c_{1\perp}' \end{aligned}$$

The general case is found by replacing c_1' *by* $c_{1:m}'$ *and* $c_{2.1}'$ *by* $c_{m+1.1:m}' = c_{m+1}' - c_{m+1}'\Omega c_{1:m}(c_{1:m}'\Omega c_{1:m})^{-1}c_{1:m}'$.

PROOF. See the Appendix. ■

8 Identification of transitory shocks

Having now exploited the matrix C for $p-r$ linearly independent rows to define $p-r$ permanent shocks, we next want to identify the r transitory shocks. These can be found in the conditional model for ΔX_t given the past and $\alpha_\perp'\Delta X_t$, or equivalently given the permanent shocks $\alpha_\perp'\varepsilon_t$, that is, in model (4).

Thus to each equation is associated a transitory shock through the corresponding row of the matrix

$$B = \alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}.$$

With this notation, the formulae for identifying the transitory shocks by a Cholesky decomposition of $B\Omega B'$ are the same as for the permanent shocks, only C is replaced by B.

Thus, whereas the permanent shocks are ordered according to *variable,* the transitory shocks are ordered according to *equation.*

A Cholesky decomposition is then found as before

$$u_{1t} = e_1'B\varepsilon_t = b_1'\varepsilon_t, \tag{28}$$

where e_1 is the unit vector corresponding to the first equation chosen. Next we choose another equation and define

$$u_{2t} = (b_2' - b_2'\Omega b_1(b_1'\Omega b_1)^{-1}b_1')\varepsilon_t, \tag{29}$$

and continue the Cholesky decomposition up to r shocks.

The transitory shocks, $B\varepsilon_t$, can be found from

$$R_{0t} = \hat{\alpha}\hat{\beta}' R_{1t} + \hat{\omega}\hat{\alpha}'_{\perp} R_{0t} + \hat{B}\hat{\varepsilon}_t.$$

8.1 The variance of the weights of the transitory shocks

We find here the asymptotic Gaussian distribution of the estimators of the weights found by the Cholesky decomposition of the short-run variance $B\Omega B'$.

Theorem 12 *The asymptotic variance of $\hat{b}_1'$ is*

$$asVar(\hat{b}_1) = (\Omega A_{\perp}\Omega)_{11}\Omega^{-1}\Sigma_{\alpha\alpha}\Omega^{-1} + \Omega^{-1}\Sigma_{\alpha\alpha}e_1e_1'\Omega A_{\perp} + A_{\perp}\Omega e_1 e_1'\Sigma_{\alpha\alpha}\Omega^{-1} + A_{\perp}(\Sigma_{\alpha\alpha} + A)_{11} \tag{30}$$

where

$$A = \alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha' \text{ and } A_{\perp} = \alpha_{\perp}(\alpha'_{\perp}\Omega\alpha_{\perp})^{-1}\alpha'_{\perp} \tag{31}$$

which satisfy $A\Omega^{-1} + \Omega A_{\perp} = I_p$, and

$$\Sigma_{\alpha\alpha} = \alpha(\alpha'\Omega^{-1}\alpha)^{-1}\Sigma_{\beta\beta}^{-1}(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'. \tag{32}$$

The asymptotic variance of $\hat{b}_{2.1}'$ is

$$\begin{aligned} & asVar(\hat{b}_{2.1}') \\ = \; & (b_{2.1}'\Omega b_{2.1})b_1(b_1'\Omega b_1)^{-1}b_1' \\ & +(b_{2.1}'\Sigma_{\alpha\alpha}b_{2.1})b_1(b_1'\Omega b_1)^{-1}e_1'\Omega A_{\perp}\Omega e_1(b_1'\Omega b_1)^{-1}b_1' \\ & +(e_{2.1}'\Omega A_{\perp}\Omega e_{2.1})\Omega^{-1}B_{1\perp}\Omega^{-1}\Sigma_{\alpha\alpha}\Omega^{-1}B_{1\perp}\Omega^{-1} + (\Sigma_{\alpha\alpha} + A)_{22}A_{\perp} \\ & +\Omega^{-1}B_{1\perp}\Omega^{-1}\Sigma_{\alpha\alpha}e_{2.1}e_{2.1}'\Omega A_{\perp} + A_{\perp}\Omega e_{2.1}e_{2.1}'\Sigma_{\alpha\alpha}\Omega^{-1}B_{1\perp}\Omega^{-1} \\ & +(e_{2.1}'\Omega A_{\perp}\Omega e_1)(b_1'\Omega b_1)^{-1}\Omega^{-1}B_{1\perp}\Omega^{-1}\Sigma_{\alpha\alpha}b_{2.1}b_1' \\ & +(e_{2.1}'\Sigma_{\alpha\alpha}b_{2.1})\Omega^{-1}B_{1\perp}A_{\perp}\Omega e_1(b_1'\Omega b_1)^{-1}b_1', \end{aligned} \tag{33}$$

where

$$B_1 = b_1(b_1'\Omega b_1)^{-1}b_1' \text{ and } B_{1\perp} = b_{1\perp}(b_{1\perp}'\Omega^{-1}b_{1\perp})^{-1}b_{1\perp}',$$

which satisfy $I_p = B_1\Omega + \Omega^{-1}B_{1\perp}$, and $e_{2.1}' = e_2' - b_2'\Omega b_1(b_1'\Omega b_1)^{-1}e_1'$.

The result for $\hat{b}_{m+1.1:m}$ is given by replacing $b_1,\ b_{2.1},$ by $b_{1:m},\ b_{m+1.1:m}$.

PROOF. See the Appendix. ■

9 Appendix

9.1 Proof of Theorem 1

We give here a calculation of the information with respect to the parameters (β, α) and its limit. We find the components of the information matrix and their limits in the directions

$$(T^{-1}\beta_{\perp}\bar{\beta}'_{\perp}(d\beta), T^{-1/2}\beta\bar{\beta}'(d\beta), T^{-1/2}d\alpha).$$

1. $\mathsf{D}^2_{\alpha\alpha}$:

$$T^{-1}\mathsf{D}^2_{\alpha\alpha}l_T(\alpha,\beta)(d\alpha,d\alpha) = tr[\Omega^{-1}(d\alpha)\beta' S_{11}\beta(d\alpha)'] \xrightarrow{P} tr[\Omega^{-1}(d\alpha)\Sigma_{\beta\beta}(d\alpha)'],$$

2. $\mathsf{D}^2_{\alpha\beta}$:

$$\begin{aligned} T^{-1}\mathsf{D}^2_{\alpha\beta}l_T(\alpha,\beta)(d\alpha,\beta\bar{\beta}'d\beta)) &= -tr[\Omega^{-1}S_{\varepsilon 1}\beta\bar{\beta}'(d\beta)(d\alpha)'] \\ &\quad +tr[\Omega^{-1}\alpha(d\beta)'\bar{\beta}\beta' S_{11}\beta(d\alpha)'], \\ T^{-3/2}\mathsf{D}^2_{\alpha\beta}l_T(\alpha,\beta)(d\alpha,\beta_\perp\bar{\beta}_\perp' d\beta)) &= -T^{-1/2}tr[\Omega^{-1}S_{\varepsilon 1}\beta_\perp\bar{\beta}_\perp'(d\beta(d\alpha)')] \\ &\quad +T^{-1/2}tr[\Omega^{-1}\alpha'(d\beta)'\bar{\beta}_\perp\beta_\perp' S_{11}\beta(d\alpha)']. \end{aligned}$$

We apply the convergence $S_{\varepsilon 1}\beta \xrightarrow{P} 0$, and $\beta' S_{11}\beta \xrightarrow{P} \Sigma_{\beta\beta}$, and find

$$\begin{aligned} &T^{-1}\mathsf{D}^2_{\alpha\beta}l_T(\alpha,\beta)(d\alpha,\beta\bar{\beta}'d\beta)) \xrightarrow{P} tr[\Omega^{-1}\alpha(d\beta)'\bar{\beta}\Sigma_{\beta\beta}(d\alpha)'], \\ &T^{-3/2}\mathsf{D}^2_{\alpha\beta}l_T(\alpha,\beta)(d\alpha,\beta_\perp\bar{\beta}_\perp' d\beta)) \xrightarrow{P} 0. \end{aligned}$$

3. $\mathsf{D}^2_{\beta\beta}$:

$$\begin{aligned} T^{-1}\mathsf{D}^2_{\beta\beta}l_T(\alpha,\beta)(\beta\bar{\beta}'d\beta,\beta\bar{\beta}'d\beta) &= tr[\alpha'\Omega^{-1}\alpha(d\beta)'\bar{\beta}\beta' S_{11}\beta\bar{\beta}'(d\beta)] \\ T^{-3/2}\mathsf{D}^2_{\beta\beta}l_T(\alpha,\beta)(\beta\bar{\beta}'d\beta,\beta_\perp\bar{\beta}_\perp' d\beta) &= T^{-1/2}tr[\alpha'\Omega^{-1}\alpha(d\beta)'\bar{\beta}\beta' S_{11}\beta_\perp\bar{\beta}_\perp'(d\beta) \\ T^{-2}\mathsf{D}^2_{\beta\beta}l_T(\alpha,\beta)(\beta_\perp\bar{\beta}_\perp' d\beta,\beta_\perp\bar{\beta}_\perp' d\beta) &= tr[\alpha'\Omega^{-1}\alpha(d\beta)'\bar{\beta}_\perp T^{-1}\beta_\perp' S_{11}\beta_\perp\bar{\beta}_\perp'(d\beta)] \end{aligned}$$

Here we apply the convergence $\beta' S_{11}\beta \xrightarrow{P} \Sigma_{\beta\beta}$, $\beta' S_{11}\beta_\perp \in O_P(1)$, and $T^{-1}\beta_\perp' S_{11}\beta_\perp \xrightarrow{d} \beta_\perp' C\int_0^1 WW'duC'\beta_\perp$ to see that

$$\begin{aligned} &T^{-1}\mathsf{D}^2_{\beta\beta}l_T(\alpha,\beta)(\beta\bar{\beta}'d\beta,\beta\bar{\beta}'d\beta) \xrightarrow{P} tr[\alpha'\Omega^{-1}\alpha(d\beta)'\bar{\beta}\Sigma_{\beta\beta}\bar{\beta}'(d\beta)], \\ &T^{-3/2}\mathsf{D}^2_{\beta\beta}l_T(\alpha,\beta)(\beta\bar{\beta}'d\beta,\beta_\perp\bar{\beta}_\perp' d\beta) \xrightarrow{P} 0, \\ &T^{-2}\mathsf{D}^2_{\beta\beta}l_T(\alpha,\beta)(\beta_\perp\bar{\beta}_\perp' d\beta,\beta_\perp\bar{\beta}_\perp' d\beta) \xrightarrow{d} tr[\alpha'\Omega^{-1}\alpha(d\beta)'C\int_0^1 WW'duC'(d\beta)]. \end{aligned}$$

Collecting these terms we find the results.

9.2 Proof of Theorem 9

The result (19) follows as in Theorem 1. We find the score functions for α and β in the direction $T^{-1/2}(\bar{\beta}'(d\beta_i), A_i(d\psi_i))$ to be

$$\begin{aligned} T^{-1/2}\mathsf{D}_\beta l_T(\alpha,\beta)(\beta\bar{\beta}'(d\beta_i)) &= -tr[\Omega^{-1}T^{1/2}S_{\varepsilon 1}\beta\bar{\beta}'(d\beta_i)\alpha_i'] \\ T^{-1/2}\mathsf{D}_\alpha l_T(\alpha,\beta)(A_i(d\psi_i)) &= -tr[(d\psi_i)'A_i'\Omega^{-1}T^{1/2}S_{\varepsilon 1}\beta_i] \end{aligned}$$

which are asymptotically jointly Gaussian with mean zero with a variance matrix as given in (20).

In the special case where $A_i = A$ we can invert the information matrix and find

$$\begin{pmatrix} \{e_i'\alpha'\Omega^{-1}\alpha e_j\Sigma_{\beta\beta}\} & \{\Sigma_{\beta\beta}e_l e_i'\alpha'\Omega^{-1}A\} \\ \{A'\Omega^{-1}\alpha e_j e_k'\Sigma\} & \{e_k'\Sigma_{\beta\beta}e_l A'\Omega^{-1}A\} \end{pmatrix}^{-1} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}.$$

We first find that $\{e_k'\Sigma_{\beta\beta}e_l A'\Omega^{-1}A\}^{-1} = \{e_k'\Sigma_{\beta\beta}^{-1}e_l(A'\Omega^{-1}A)^{-1}\}$, and therefore, for $\rho_{ij} = \alpha_i'\Omega^{-1}\alpha_j$,

$$\begin{aligned}(M_{11}^{-1})_{ij} &= \rho_{ij}\Sigma_{\beta\beta} - \sum_{l,k}\Sigma_{\beta\beta}e_l e_i'\alpha'\Omega^{-1}Ae_k'\Sigma_{\beta\beta}^{-1}e_l(A'\Omega^{-1}A)^{-1}A'\Omega^{-1}\alpha e_j e_k'\Sigma_{\beta\beta} \\ &= \rho_{ij}\Sigma_{\beta\beta} - \sum_{l,k}\Sigma_{\beta\beta}e_l e_k'\Sigma_{\beta\beta}^{-1}e_l e_k'\Sigma_{\beta\beta}[e_i'\alpha'\Omega^{-1}A(A'\Omega^{-1}A)^{-1}A'\Omega^{-1}\alpha e_j]\end{aligned}$$

Now $\sum_{l,k}\Sigma_{\beta\beta}e_l e_l'\Sigma_{\beta\beta}^{-1}e_k e_k'\Sigma_{\beta\beta} = \Sigma_{\beta\beta}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta\beta} = \Sigma_{\beta\beta}$, and hence

$$\begin{aligned}(M_{11}^{-1})_{ij} &= e_i'[\alpha'\Omega^{-1}\alpha - \alpha'\Omega^{-1}A(A'\Omega^{-1}A)^{-1}A'\Omega^{-1}\alpha]e_j\Sigma_{\beta\beta} \\ &= e_i'[\alpha'A_\perp(A_\perp'\Omega A_\perp)^{-1}A_\perp\alpha]e_j\Sigma_{\beta\beta}\end{aligned}$$

and therefore

$$M_{11} = \{e_i'(\alpha'A_\perp(A_\perp'\Omega A_\perp)^{-1}A_\perp\alpha)^{-1}e_j\Sigma_{\beta\beta}^{-1}\}.$$

Similarly we find

$$\begin{aligned}(M_{22}^{-1})_{kl} &= \sigma_{kl}A'\Omega^{-1}A - \sum_{i,j}(A'\Omega^{-1}\alpha e_j e_k'\Sigma_{\beta\beta}e_i'(\alpha'\Omega^{-1}\alpha)^{-1}e_j\Sigma_{\beta\beta}^{-1}\Sigma_{\beta\beta}e_l e_i'\alpha'\Omega^{-1}A \\ &= \sigma_{kl}A'\Omega^{-1}A - e_k'\Sigma_{\beta\beta}e_l\sum_{i,j}(A'\Omega^{-1}\alpha e_j e_i'(\alpha'\Omega^{-1}\alpha)^{-1}e_j e_i'\alpha'\Omega^{-1}A \\ &= \sigma_{kl}[A'\Omega^{-1}A - (A'\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}A] \\ &= \sigma_{kl}[A'\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp'A].\end{aligned}$$

Hence

$$M_{22} = \{e_l'\Sigma_{\beta\beta}^{-1}e_k[A'\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp'A]^{-1}\}.$$

The off-diagonal elements are found from

$$\{\rho_{ij}\Sigma_{\beta\beta}\}M_{12} + \{\Sigma_{\beta\beta}e_l e_i'\alpha'\Omega^{-1}A\}M_{22} = 0,$$

which implies that

$$\begin{aligned}(M_{12})_{il} &= -\sum_{k,m}\rho^{ik}\Sigma_{\beta\beta}^{-1}\Sigma_{\beta\beta}e_m e_k'\alpha'\Omega^{-1}A\sigma^{ml}M \\ &= -\sum_m e_m e_m'\Sigma_{\beta\beta}^{-1}e_l\sum_k e_i'(\alpha'\Omega^{-1}\alpha)^{-1}e_k e_k'\alpha'\Omega^{-1}AM \\ &= -\Sigma_{\beta\beta}^{-1}e_l e_i'(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}AM.\end{aligned}$$

In particular for $A = (0, I_{p-r})'$ we find, using $\alpha'A_\perp = I_r$, and $A_\perp = (I_r, 0)'$

$$\begin{aligned}M_{11} &= \{e_i'(\alpha'A_\perp(A_\perp'\Omega A_\perp)^{-1}A_\perp\alpha)^{-1}e_j\Sigma_{\beta\beta}^{-1}\} \\ &= \{e_i'\Omega^{-1}e_j\Sigma^{-1}\} = \Omega^{11}\otimes\Sigma_{\beta\beta}^{-1}.\end{aligned}$$

9.3 Proof of Theorem 11

We want to derive the asymptotic distribution in Theorem 11 using the $\delta-$method. For this we need the asymptotic distribution of $(\hat{C}, \hat{\Omega})$, which is asymptotically Gaussian with variance matrix given by the next lemma and an expansion of $c_{2.1}$ as function of C and Ω.

Lemma 13 *The asymptotic variance and covariance matrix of $\hat{\Omega}$ and $\hat{C}$ are given by*

$$\begin{aligned} asVar(\kappa'\hat{\Omega}\eta) &= (\kappa'\Omega\kappa)(\eta'\Omega\eta) + (\eta'\Omega\kappa)^2, \\ asVar(\kappa'\hat{C}\eta) &= (\eta'\xi'\Sigma^{-1}\xi\eta)(\kappa' C\Omega C'\kappa), \\ asCov(\hat{\Omega}, \hat{C}) &= 0, \end{aligned}$$

where

$$\begin{aligned} \Sigma &= EZ_tZ_t', \\ Z_t' &= (X_{t-1}'\hat{\beta}, \Delta X_{t-1}', \ldots, \Delta X_{t-k+1}'), \\ \xi' &= ((C'\Gamma' - I_p)\bar{\alpha}, C', \ldots, C'). \end{aligned}$$

see Paruolo (1997). We also need the derivatives of

$$c_{2.1}' = c_2' - c_2'\Omega c_1(c_1'\Omega c_1)^{-1}c_1' = [e_2' - e_2'C\Omega C'e_1(e_1'C\Omega C'e_1)^{-1}e_1']C$$

with respect to C and Ω.

Lemma 14 *We have the expansion*

$$\begin{aligned} & d(c_2' - c_2'\Omega c_1(c_1'\Omega c_1)^{-1}c_1') \\ &= (e_2' - c_2'\Omega c_1(c_1'\Omega c_1)^{-1}e_1')(dC)C_{1\perp}\Omega^{-1} - c_{2.1}'(d\Omega)C_1 - c_{2.1}'\Omega(dC)'e_1(c_1'\Omega c_1)^{-1}c_1', \end{aligned}$$

where $C_1 = c_1(c_1'\Omega c_1)^{-1}c_1'$ and $C_{1\perp} = c_{1\perp}(c_{1\perp}'\Omega^{-1}c_{1\perp})^{-1}c_{1\perp}'$.

PROOF. By Taylor's expansion we find

$$\begin{aligned} & d(c_2' - c_2'\Omega c_1(c_1'\Omega c_1)^{-1}c_1') \\ &= (dc_2)' - (dc_2)'\Omega c_1(c_1'\Omega c_1)^{-1}c_1' - c_2'(d\Omega)c_1(c_1'\Omega c_1)^{-1}c_1' - c_2'\Omega(dc_1)(c_1'\Omega c_1)^{-1}c_1' \\ &+ c_2'\Omega c_1(c_1'\Omega c_1)^{-1}((dc_1)'\Omega c_1 + c_1'(d\Omega)c_1 + c_1'\Omega(dc_1))(c_1'\Omega c_1)^{-1}c_1' - c_2'\Omega c_1(c_1'\Omega c_1)^{-1}(dc_1)' \\ &= e_2'(dC)(I_p - \Omega C_1) - c_{2.1}'(d\Omega)C_1 - c_2'\Omega[I_p - C_1\Omega] \\ & (dC)'e_1(c_1'\Omega c_1)^{-1}c_1' - c_2'\Omega c_1(c_1'\Omega c_1)^{-1}(dC)'e_1C_{1\perp}\Omega^{-1} \\ &= e_2'(dC)C_{1\perp}\Omega^{-1} - c_{2.1}'(d\Omega)C_1 \\ &- c_{2.1}'\Omega(dC)'e_1(c_1'\Omega c_1)^{-1}c_1' - c_2'\Omega c_1(c_1'\Omega c_1)^{-1}e_1'(dC)C_{1\perp}\Omega^{-1}, \end{aligned}$$

which reduces to the result.

We start the proof of Theorem 11 by (26), which follows directly from the expression for the asymptotic variance of $\hat{C}$:

$$asVar(\hat{c}_1'\eta) = asVar(e_1'\hat{C}\eta) = (\eta'\xi'\Sigma^{-1}\xi\eta)(e_1'C\Omega C'e_1) = (\eta'\xi'\Sigma^{-1}\xi\eta)(c_1'\Omega c_1).$$

We next prove (27). We find from Lemma 14 that, for $e_{2.1}' = e_2' - c_2'\Omega c_1(c_1'\Omega c_1)^{-1}e_1'$, it holds that

$$(dc_{2.1})'\eta = -c_{2.1}'(d\Omega)C_1\eta + e_{2.1}'(dC)C_{1\perp}\Omega^{-1}\eta - \eta'c_1(c_1'\Omega c_1)^{-1}e_1'(dC)\Omega c_{2.1},$$

so that

$$\begin{aligned} asVar(\hat{c}_{2.1}'\eta) &= asVar(-c_{2.1}'\hat{\Omega}C_1\eta + e_{2.1}'\hat{C}C_{1\perp}\Omega^{-1}\eta - \eta'c_1(c_1'\Omega c_1)^{-1}e_1'\hat{C}\Omega c_{2.1}) \\ &= asVar(K_\Omega + K_1 + K_2). \end{aligned}$$

We first find the variances

$$asVar(K_\Omega) = (c_{2.1}'\Omega c_{2.1})(\eta'C_1\Omega C_1\eta) + (c_{2.1}'\Omega C_1\eta)^2 = (c_{2.1}'\Omega c_{2.1})(\eta'C_1\eta)$$

because $C_1\Omega C_1 = C_1$ and $c_{2.1}'\Omega C_1 = 0$. Next

$$asVar(K_1) = (\eta'\Omega^{-1}C_{1\perp}\xi'\Sigma^{-1}\xi C_{1\perp}\Omega^{-1}\eta)(c_{2.1}'\Omega c_{2.1}).$$

Finally

$$\begin{aligned} asVar(K_2) &= (c_{2.1}'\Omega\xi'\Sigma^{-1}\xi\Omega c_{2.1})(\eta'c_1(c_1'\Omega c_1)^{-1}e_1'C\Omega C'e_1(c_1'\Omega c_1)^{-1}c_1'\eta) \\ &= (c_{2.1}'\Omega\xi'\Sigma^{-1}\xi\Omega c_{2.1})(\eta'C_1\eta). \end{aligned}$$

We have $Cov(K_\Omega; K_1) = Cov(K_\Omega; K_2) = 0$, because $\hat{C}$ and $\hat{\Omega}$ are asymptotically independent, see Lemma 13, and find

$$\begin{aligned} asCov(K_1; K_2) &= -asCov[e_{2.1}'\hat{C}C_{1\perp}\Omega^{-1}\eta; \eta'c_1(c_1'\Omega c_1)^{-1}e_1'\hat{C}\Omega c_{2.1}] \\ &= (e_{2.1}'C\Omega C'e_1(c_1'\Omega c_1)^{-1}c_1'\eta)(c_{2.1}'\Omega\xi'\Sigma^{-1}\xi C_{1\perp}\Omega^{-1}\eta) \\ &= (c_{2.1}'\Omega C_1\eta)(c_{2.1}'\Omega\xi'\Sigma^{-1}\xi C_{1\perp}\Omega^{-1}\eta) = 0, \end{aligned}$$

because $c_{2.1}'\Omega C_1 = 0$. Collecting the result, we have proved (27). The general case is proved the same way by replacing c_1 by $c_{1:m}$. This completes the proof of Theorem 11. ■

9.4 Proof of Theorem 12

We first find an expansion of B as a function of Ω and α, and apply this to find the asymptotic distribution of $(\hat{B}, \hat{\Omega})$.

Lemma 15 *We have the expansions*

$$\begin{aligned} dB &= d(\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}) \\ &= \Omega A_\perp(d\alpha)(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1} + \alpha(\alpha'\Omega^{-1}\alpha)^{-1}(d\alpha)'A_\perp - A\Omega^{-1}(d\Omega)A_\perp, \end{aligned} \tag{34}$$

where $A = \alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'$ *and* $A_\perp = \alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp'$.

PROOF. By Taylor's expansion we find

$$d\Omega^{-1} = -\Omega^{-1}(d\Omega)\Omega^{-1}$$

and

$$d(\alpha'\Omega^{-1}\alpha)^{-1} = -(\alpha'\Omega^{-1}\alpha)^{-1}[(d\alpha)'\Omega^{-1}\alpha - \alpha'\Omega^{-1}(d\Omega)\Omega^{-1}\alpha + \alpha'\Omega^{-1}(d\alpha)](\alpha'\Omega^{-1}\alpha)^{-1}$$

so that

$$\begin{aligned}
&d(\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}) \\
&= (d\alpha)(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1} + \alpha(\alpha'\Omega^{-1}\alpha)^{-1}(d\alpha)'\Omega^{-1} - \alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}(d\Omega)\Omega^{-1} \\
&- \alpha(\alpha'\Omega^{-1}\alpha)^{-1}[(d\alpha)'\Omega^{-1}\alpha - \alpha'\Omega^{-1}(d\Omega)\Omega^{-1}\alpha + \alpha'\Omega^{-1}(d\alpha)](\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}
\end{aligned}$$

This expression is the same as in (34), as can be seen by multiplying both by the matrix $(\Omega\alpha_\perp, \alpha)$. ■

The asymptotic variance of $\hat{\alpha}$, see Johansen (1996), is given by

$$\begin{aligned}
asCov(\kappa'\hat{\alpha}\eta; \xi'\hat{\alpha}\phi) &= (\kappa'\Omega\xi)(\eta'\Sigma_{\beta\beta}^{-1}\phi) \\
\Sigma_{\beta\beta} &= Var(\beta' X_t | \Delta X_t, \dots, \Delta X_{t-k+1}), \\
asCov(\hat{\alpha}; \hat{\Omega}) &= 0,
\end{aligned}$$

and we apply that to find the asymptotic variance of $(\hat{B}, \hat{\Omega})$.

Lemma 16 *The asymptotic variance of $\hat{B} = \hat{\alpha}(\hat{\alpha}'\hat{\Omega}^{-1}\hat{\alpha})^{-1}\hat{\alpha}'\hat{\Omega}^{-1}$ is*

$$\begin{aligned}
&asCov(\kappa'\hat{B}\eta; \xi'\hat{B}\phi) \\
&= (\kappa'\Omega A_\perp\Omega\xi)(\eta'\Omega^{-1}\Sigma_{\alpha\alpha}\Omega^{-1}\phi) + (\eta'\Omega^{-1}\Sigma_{\alpha\alpha}\xi)(\kappa'\Omega A_\perp\phi) \\
&+ (\eta' A_\perp\Omega\xi)(\kappa'\Sigma_{\alpha\alpha}\Omega^{-1}\phi) + (\eta' A_\perp\phi)(\kappa'[\Sigma_{\alpha\alpha} + A]\xi),
\end{aligned} \tag{35}$$

and the covariance with $\hat{\Omega}$ is

$$asCov(\kappa'\hat{B}\eta; \xi'\hat{\Omega}\phi) = -(\kappa' A\xi)(\eta' A_\perp\Omega\phi), \tag{36}$$

where A and $A_\perp$ are given in Lemma 15, and

$$\Sigma_{\alpha\alpha} = \alpha(\alpha'\Omega^{-1}\alpha)^{-1}\Sigma_{\beta\beta}^{-1}(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'.$$

PROOF. We find from the expansion of B in Lemma 15 that $(\alpha_\Omega = \alpha(\alpha\Omega^{-1}\alpha)^{-1})$

$$\begin{aligned}
&asCov(\kappa'\hat{B}\eta; \xi'\hat{B}\phi) \\
&= asCov(\kappa'\Omega A_\perp\hat{\alpha}\alpha_\Omega'\Omega^{-1}\eta + \kappa'\alpha_\Omega\hat{\alpha}' A_\perp\eta - \kappa' A\Omega^{-1}\hat{\Omega} A_\perp\eta; \\
&\quad \xi'\Omega A_\perp\hat{\alpha}\alpha_\Omega'\Omega^{-1}\phi + \xi'\alpha_\Omega\hat{\alpha}' A_\perp\phi - \xi' A\Omega^{-1}\hat{\Omega} A_\perp\phi),
\end{aligned}$$

We first take the terms with $\hat{\Omega}$:

$$
\begin{aligned}
&asCov(\kappa' A\Omega^{-1}\hat{\Omega}A_\perp\eta;\xi' A\Omega^{-1}\hat{\Omega}A_\perp\phi)\\
&=(\kappa' A\Omega^{-1}\Omega\Omega^{-1}A\xi)(\eta' A_\perp\Omega A_\perp\phi)+(\kappa' A\Omega^{-1}\Omega A_\perp\phi)(\xi' A\Omega^{-1}\Omega A_\perp\eta)\\
&=(\kappa' A\xi)(\eta' A_\perp\phi),
\end{aligned}
$$

because

$$
A\Omega^{-1}\Omega A_\perp=0,\quad A\Omega^{-1}A=A,\quad A_\perp\Omega A_\perp=A_\perp.
$$

Next consider the terms involving $\hat{\alpha}$ and $\hat{\alpha}'$

$$
\begin{aligned}
&asCov(\kappa'\Omega A_\perp\hat{\alpha}\alpha_\Omega'\Omega^{-1}\eta+\kappa'\alpha_\Omega\hat{\alpha}'A_\perp\eta;\xi'\Omega A_\perp\hat{\alpha}\alpha_\Omega'\Omega^{-1}\phi+\xi'\alpha_\Omega\hat{\alpha}'A_\perp\phi)\\
&=asCov(\kappa'\Omega A_\perp\hat{\alpha}\alpha_\Omega'\Omega^{-1}\eta+\eta' A_\perp\hat{\alpha}\alpha_\Omega'\kappa;\xi'\Omega A_\perp\hat{\alpha}\alpha_\Omega'\Omega^{-1}\phi+\phi' A_\perp\hat{\alpha}\alpha_\Omega'\xi)\\
&=(\kappa'\Omega A_\perp\Omega\xi)(\eta'\Omega^{-1}\alpha_\Omega'\Sigma_{\beta\beta}^{-1}\alpha_\Omega'\Omega^{-1}\phi)+(\kappa'\Omega A_\perp\phi)(\eta'\Omega^{-1}\alpha_\Omega\Sigma_{\beta\beta}^{-1}\alpha_\Omega'\xi)\\
&+(\eta' A_\perp\Omega\xi)(\kappa'\alpha_\Omega\Sigma_{\beta\beta}^{-1}\alpha_\Omega'\Omega^{-1}\phi)+(\eta' A_\perp\phi)(\kappa'\alpha_\Omega\Sigma_{\beta\beta}^{-1}\alpha_\Omega'\xi).
\end{aligned}
$$

Introducing $\Sigma_{\alpha\alpha}=\alpha_\Omega\Sigma_{\beta\beta}^{-1}\alpha_\Omega'$ we have proved (35), because $\hat{\alpha}$ and $\hat{\Omega}$ are asymptotically independent.

Next consider (36), where we again use the asymptotic independence of $\hat{\alpha}$ and $\hat{\Omega}$ to find

$$
\begin{aligned}
&asCov(\kappa'\hat{B}\eta;\xi'\hat{\Omega}\phi)\\
&=asCov(\kappa'\Omega A_\perp\hat{\alpha}\alpha_\Omega'\Omega^{-1}\eta+\kappa'\alpha_\Omega\hat{\alpha}'A_\perp\eta-\kappa' A\Omega^{-1}\hat{\Omega}A_\perp\eta;\xi'\hat{\Omega}\phi)\\
&=-asCov(\kappa' A\Omega^{-1}\hat{\Omega}A_\perp\eta,\xi'\hat{\Omega}\phi)=-(\kappa' A\Omega^{-1}\Omega\xi)(\eta' A_\perp\Omega\phi).
\end{aligned}
$$

■

We start the proof of Theorem **12** by (30). We apply (35) of Lemma 16 with $\kappa=\xi=e_1$, and $\eta=\phi=\psi$ to find the asymptotic variance of $\hat{b}_1$:

$$
\begin{aligned}
&asCov(e_1'\hat{B}\psi;e_1'\hat{B}\psi)\\
&=(\Omega A_\perp\Omega)_{11}\psi'\Omega^{-1}\Sigma_{\alpha\alpha}\Omega^{-1}\psi+\psi'\Omega^{-1}\Sigma_{\alpha\alpha}e_1e_1'\Omega A_\perp\psi\\
&+\psi' A_\perp\Omega e_1e_1'\Sigma_{\alpha\alpha}\Omega^{-1}\psi+(\Sigma_{\alpha\alpha}+A)_{11}\psi' A_\perp\psi,
\end{aligned}
$$

which shows (30).

Next we prove (33). We apply Lemma 14 with C replaced by B to find the expansion

$$
(db_{2.1})'\psi=-b_{2.1}'(d\Omega)B_1\psi+(e_2'-b_2'\Omega b_1(b_1'\Omega b_1)^{-1}e_1')(dB)B_{1\perp}\Omega^{-1}\psi-\psi' b_1(b_1'\Omega b_1)^{-1}e_1'(dB)\Omega b_{2.1},
$$

so that

$$
\begin{aligned}
asVar(\hat{b}_{2.1}'\psi)&=asVar(-b_{2.1}'\hat{\Omega}B_1\psi\\
&+(e_2'-b_2'\Omega b_1(b_1'\Omega b_1)^{-1}e_1')\hat{B}B_{1\perp}\Omega^{-1}\psi-\psi' b_1(b_1'\Omega b_1)^{-1}e_1'\hat{B}\Omega b_{2.1})\\
&=asVar(L_\Omega+L_1+L_2)
\end{aligned}
$$

9.4.1 The variances

We first find the variances:

$$\begin{aligned} asVar(L_\Omega) &= asVar(-b_{2.1}'\hat{\Omega}B_1\psi) = (b_{2.1}'\Omega b_{21})(\psi'B_1\Omega B_1\psi) + (b_{2.1}'\Omega B_1\psi)^2 \\ &= (b_{2.1}'\Omega b_{21})(\psi'B_1\psi), \end{aligned}$$

because $B_1\Omega B_1 = B_1$ and $b_{2.1}'\Omega B_1 = 0$.

Next we consider $asVar(L_1) = asVar((e_2' - b_2'\Omega b_1(b_1'\Omega b_1)^{-1}e_1')\hat{B}B_{1\perp}\Omega^{-1}\psi)$ which follows from (35) for $\kappa = \xi = e_2' - b_2'\Omega b_1(b_1'\Omega b_1)^{-1}e_1' = e_{2.1}'$, say, and $\eta = \phi = B_{1\perp}\Omega^{-1}\psi$

$$\begin{aligned} &asVar(L_1) \\ &= (\psi'\Omega^{-1}B_{1\perp}\Omega^{-1}\Sigma_{\alpha\alpha}\Omega^{-1}B_{1\perp}\Omega^{-1}\psi)(e_{2.1}'\Omega A_\perp\Omega e_{2.1}) + (e_{2.1}'\Sigma_{\alpha\alpha}e_{2.1})(\psi'A_\perp\psi) \\ &+ 2(\psi'\Omega^{-1}B_{1\perp}\Omega^{-1}\Sigma_{\alpha\alpha}e_{2.1})(e_{2.1}'\Omega A_\perp\psi) + (e_{2.1}'Ae_{2.1})(\psi'A_\perp\psi), \end{aligned}$$

using $A_\perp B_{1\perp} = A_\perp\Omega$.

Similarly we find $asVar(L_2) = asVar(\psi'b_1(b_1'\Omega b_1)^{-1}e_1'\hat{B}\Omega b_{2.1})$ from (35) for $\kappa = \xi = e_1(b_1'\Omega b_1)^{-1}b_1'\psi$ and $\eta = \phi = \Omega b_{2.1}$

$$\begin{aligned} &asVar(L_2) \\ &= (b_{2.1}'\Omega\Omega^{-1}\Sigma_{\alpha\alpha}\Omega^{-1}\Omega b_{2.1})(\psi'b_1(b_1'\Omega b_1)^{-1}e_1'\Omega A_\perp\Omega e_1(b_1'\Omega b_1)^{-1}b_1'\psi) \\ &+ (\psi'b_1(b_1'\Omega b_1)^{-1}e_1'\Sigma_{\alpha\alpha}e_1(b_1'\Omega b_1)^{-1}b_1'\psi)(b_{2.1}'\Omega A_\perp\Omega b_{2.1}) \\ &+ 2(b_{2.1}'\Omega\Omega^{-1}\Sigma_{\alpha\alpha}e_1(b_1'\Omega b_1)^{-1}b_1'\psi)(\psi'b_1(b_1'\Omega b_1)^{-1}e_1'\Omega A_\perp\Omega b_{2.1}) \\ &+ (\psi'b_1(b_1'\Omega b_1)^{-1}e_1'Ae_1(b_1'\Omega b_1)^{-1}b_1'\psi)(b_{2.1}'\Omega A_\perp\Omega b_{2.1}) \\ &= (b_{2.1}'\Sigma_{\alpha\alpha}b_{2.1})(\psi'b_1(b_1'\Omega b_1)^{-1}e_1'\Omega A_\perp\Omega e_1(b_1'\Omega b_1)^{-1}b_1'\psi), \end{aligned}$$

because $A_\perp\Omega b_{2.1} = 0$, so that only the first term is non-zero.

9.4.2 The covariances

We first take the covariance $asCov(L_1; L_\Omega) = -asCov(e_{2.1}'\hat{B}B_{1\perp}\Omega^{-1}\psi; b_{2.1}'\hat{\Omega}B_1\psi)$ which we calculate from (36) for $\kappa = e_{2.1},\ \xi = b_{2.1},\ \eta = B_{1\perp}\Omega^{-1}\psi$, and $\phi = B_1\psi$

$$asCov(L_\Omega; L_1) = -(e_{2.1}'Ab_{2.1})(\psi'\Omega^{-1}B_{1\perp}A_\perp\Omega B_1\psi) = 0.$$

because $A_\perp\Omega B_1 = 0$.

Next $asCov(L_2; L_\Omega) = asCov(\psi'b_1e_1'\hat{B}\Omega b_{2.1}; b_{2.1}'\hat{\Omega}B_1\psi)$ can be found for

$$\kappa = e_1(b_1'\Omega b_1)^{-1}b_1'\psi,\ \xi = b_{2.1},\ \eta = \Omega b_{2.1},\ \phi = B_1\psi,$$

which gives

$$asCov(L_\Omega; L_2) = (\psi'(b_1'\Omega b_1)^{-1}b_1'e_1'Ab_{2.1})(b_{2.1}'\Omega A_\perp B_1\psi) = 0,$$

because $b_{2.1}'\Omega A_\perp = 0$.

$$
\begin{aligned}
&asCov(\kappa'\hat{B}\eta; \xi'\hat{B}\phi)\\
&= (\kappa'\Omega A_\perp\Omega\xi)(\eta'\Omega^{-1}\Sigma_{\alpha\alpha}\Omega^{-1}\phi) + (\eta'\Omega^{-1}\Sigma_{\alpha\alpha}\xi)(\kappa'\Omega A_\perp\phi)\\
&+ (\eta' A_\perp\Omega\xi)(\kappa'\Sigma_{\alpha\alpha}\Omega^{-1}\phi) + (\eta' A_\perp\phi)(\kappa'[\Sigma_{\alpha\alpha} + A]\xi)
\end{aligned}
$$

Finally $asCov(L_1; L_2) = Cov(e_{2.1}'\hat{B}B_{1\perp}\Omega^{-1}\psi; \psi' b_1(b_1'\Omega b_1)^{-1}e_1'\hat{B}\Omega b_{2.1})$ can be found from (35) for $\kappa' = e_{2.1}'$, $\xi' = \psi' b_1(b_1'\Omega b_1)^{-1}e_1'$, $\eta = B_{1\perp}\Omega^{-1}\psi$, $\phi = \Omega b_{2.1}$. We note that $A_\perp\phi = A_\perp\Omega b_{2.1} = 0$, so we only need to consider the terms

$$
\begin{aligned}
&(\kappa'\Omega A_\perp\Omega\xi)(\eta'\Omega^{-1}\Sigma_{\alpha\alpha}\Omega^{-1}\phi) + (\eta' A_\perp\Omega\xi)(\kappa'\Sigma_{\alpha\alpha}\Omega^{-1}\phi)\\
&= (\psi'\Omega^{-1}B_{1\perp}\Omega^{-1}\Sigma_{\alpha\alpha}b_{2.1})(e_{2.1}'\Omega A_\perp\Omega e_1)(b_1'\Omega b_1)^{-1}b_1'\psi\\
&\quad + (e_{2.1}'\Sigma_{\alpha\alpha}b_{2.1})(\psi'\Omega^{-1}B_{1\perp}A_\perp\Omega e_1(b_1'\Omega b_1)^{-1}b_1'\psi)
\end{aligned}
$$

Collecting the terms we have proved the result.■

10 References

Blanchard, O. F. and D. Quah 1989. The dynamic effects of aggregate demand and supply disturbances. *American Economic Review* 79, 655–673.

Fisher, F. M. 1966. *The Identification Problem in Econometrics.* McGraw-Hill. New York NY.

Johansen S., 1995. Identifying Restrictions of Linear Equations — with applications to simultaneous equations and cointegration. *Journal of Econometrics* 69, 111–132.

Johansen, S. 1996. *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models.* Oxford University Press, Oxford (2.ed.).

Lin. M.-T. 2007. Long-run propositions, long-run identification, and cointegrating relations. Discussion paper. National University of Singapore.

Paruolo P. 1997. Asymptotic inference on the moving average impact matrix in cointegrated $I(1)$ VAR systems, *Econometric Theory*, 13, 79–118.

Rado, R. 1942. A theorem on independence relations. *Quarterly Journal of Mathematics, Oxford* 13, 83–89.

van der Vaart, A. W. 1998. *Asymptotic Statistics.* Cambridge University Press, Cambridge.