EUI WORKING PAPERS

ECO No. 2005/15

Problems Related to Over-Identifying Restrictions for Structural Vector Error Correction Models

HELMUT LÜTKEPOHL

EUROPEAN UNIVERSITY INSTITUTE
Department of Economics

EUROPEAN UNIVERSITY INSTITUTE
DEPARTMENT OF ECONOMICS

Problems Related to Over-Identifying Restrictions for Structural Vector Error Correction Models

HELMUT LÜTKEPOHL

EUI Working Paper **ECO** No. 2005/15

ISSN 1725-6704

Printed in Italy
European University Institute
Badia Fiesolana
I – 50016 San Domenico di Fiesole (FI)
Italy

http://www.iue.it/
http://cadmus.iue.it/dspace/index.jsp

November 4, 2005

Problems Related to Over-identifying Restrictions for Structural Vector Error Correction Models

by

Helmut Lütkepohl [1]
Department of Economics
European University Institute
Via della Piazzuola 43
I-50133 Firenze
ITALY

Abstract. Structural vector autoregressive (VAR) models are in frequent use for impulse response analysis. If cointegrated variables are involved, the corresponding vector error correction models offer a convenient framework for imposing structural long-run and short-run restrictions. Occasionally it is desirable to impose over-identifying restrictions in this context. Some related problems are pointed out. They result from the fact that the over-identifying restrictions have to be in the admissible parameter space which is not always obvious. Conditions are given that can help in avoiding the problems.

Key Words: Cointegration, vector autoregressive process, vector error correction model, impulse responses

JEL classification: C32

[1] I thank Ralf Brüggemann for helpful comments on an earlier draft of this paper.

1 Introduction

In structural vector autoregressive (SVAR) modelling long-run restrictions are often used in addition to short-run restrictions to identify the shocks and impulse responses of interest. In particular, if cointegrated variables are present, the cointegration properties may be useful in specifying the number of shocks with permanent and transitory effects. Vector error correction models (VECMs) and the framework laid out by King, Plosser, Stock & Watson (1991) offer a possible setup for imposing identifying restrictions. In this note I will argue that these restrictions require some care in doing inference for impulse responses. In particular, certain over-identifying restrictions are not possible because they imply a singular residual covariance matrix which is usually ruled out by assumption and is also not plausible from a theoretical point of view. In other words, some over-identifying restrictions may not be possible because they are outside the admissible parameter space. This of course also means that associated t-ratios cannot be interpreted in the usual way. Unfortunately, it is not always obvious which over-identifying restrictions are possible and which ones are not admissible. Therefore I will discuss conditions that will help to see more easily which restrictions are not feasible. It may be worth pointing out that the problem also affects the impulse responses. In particular, the interpretation of confidence intervals around impulse response functions needs some care. This issue will also be discussed in the following.

The study is structured as follows. In the next section the model setup for structural modelling with cointegrated VAR processes will be presented. Estimation of the models is discussed in Section 3. An example based on U.S. macroeconomic data from King et al. (1991) is presented in Section 4 and conclusions follow in Section 5. The structural VECM framework of the present article was proposed by King et al. (1991) and is also discussed in detail in Lütkepohl (2005, Chapter 9).

A variable will be called *integrated of order d* ($I(d)$) if stochastic trends or unit roots can be removed by differencing the variable d times and a stochastic trend still remains after differencing only $d-1$ times. A variable without a stochastic trend or unit root is sometimes called $I(0)$. To simplify matters, in the following all variables are assumed to be either $I(0)$ or $I(1)$. A set of $I(1)$ variables is called *cointegrated* if a linear combination exists which is $I(0)$. A K-dimensional system of variables y_t is called $I(1)$ if at least one component is $I(1)$. In that case, any linear combination $c'y_t$ which is $I(0)$ is called a cointegration relation. Using this terminology it can happen that a linear combination of $I(0)$ variables is called a cointegration relation. In the present context, this terminology is convenient, however.

The following general notation will be used. The natural logarithm is abbreviated as log. For a suitable matrix A, $\text{rk}(A)$, $\det(A)$ and $A_\perp$ denote the rank, the determinant and an orthogonal complement of A, respectively. Moreover, vec is the column stacking operator which stacks the columns of a matrix in a column vector and vech is the column stacking operator for symmetric square matrices which stacks the columns from the main diagonal downwards only. The $(n \times n)$ identity matrix is signified as I_n and $0_{n\times m}$ denotes an $(n \times m)$ zero matrix.

2 The Model Setup

As mentioned earlier, it is assumed that all variables are at most $I(1)$ and that the data generation process can be represented as a VECM of the form

$$\Delta y_t = \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1}\Delta y_{t-p+1} + u_t, \quad t = 1, 2, \ldots, \tag{2.1}$$

where y_t is a K-dimensional vector of observable variables and α and β are $(K \times r)$ matrices of rank r. More precisely, β is the cointegration matrix and r is the cointegrating rank of the process. The term $\alpha\beta' y_{t-1}$ is sometimes referred to as error correction term. The Γ_j's, $j = 1, \ldots, p-1$, are $(K \times K)$ short-run coefficient matrices and u_t is a white noise error vector with mean zero and nonsingular covariance matrix Σ_u, $u_t \sim (0_{K\times 1}, \Sigma_u)$. Moreover, $y_{-p+1}, \ldots, y_0$ are assumed to be fixed initial conditions.

Although in practice there will usually also be deterministic terms such as nonzero means or polynomial trends, it will be assumed in the following that such terms are absent. They do not play a role in impulse response analysis which is the focus of this study. The main results are unaffected by such terms. Therefore they are omitted.

2.1 The Identification Problem

Impulse responses are often used to study the relationships between the variables of a dynamic model such as (2.1). In this context, identifying structural innovations which induce responses of the variables reflecting the actual ongoings in a system is an important task. In the present VECM framework, the so-called B-model setup is typically used (Lütkepohl (2005, Chapter 9)). It is assumed that the structural innovations, say ε_t, have zero mean and identity covariance matrix, $\varepsilon_t \sim (0_{K\times 1}, I_K)$, and they are linearly related to the u_t such that

$$u_t = B\varepsilon_t.$$

Hence, $\Sigma_u = BB'$. This relation represents $\frac{1}{2}K(K+1)$ independent equations because the covariance matrix is symmetric. For a unique specification of the K^2 elements of B we need at least $\frac{1}{2}K(K-1)$ further restrictions. Some of them may be obtained via a more detailed examination of the cointegration structure of the model, as will be seen in the following.

According to Granger's representation theorem (see Johansen (1995)), the process y_t has the representation

$$y_t = \Xi \sum_{i=1}^{t} u_i + \sum_{j=0}^{\infty} \Xi_j^* u_{t-j} + y_0^*, \quad t = 1, 2, \ldots, \tag{2.2}$$

where the term y_0^* contains the initial values and the Ξ_j^*'s are absolutely summable so that $\sum_{j=0}^{\infty} \Xi_j^* u_{t-j}$ represents a stationary process where shocks have transitory effects only, that is, $\Xi_j^* \to 0$ for $j \to \infty$. The term $\sum_{i=1}^{t} u_i$, $t = 1, 2, \ldots$, is a K-dimensional random walk. Thus, the long-run effects of shocks are represented by the term $\Xi \sum_{i=1}^{t} u_i$ which captures the common stochastic trends. The matrix Ξ is of the form

$$\Xi = \beta_\perp \left[\alpha'_\perp \left(I_K - \sum_{i=1}^{p-1} \Gamma_i \right) \beta_\perp \right]^{-1} \alpha'_\perp$$

and has rank $K-r$. Thus, there are $K-r$ independent common trends. Substituting $B\varepsilon_i$ for u_i in the common trends term in (2.2) gives $\Xi \sum_{i=1}^{t} u_i = \Xi B \sum_{i=1}^{t} \varepsilon_i$ so that the long-run effects of the structural innovations are given by ΞB.

The structural innovations ε_t have nonsingular covariance matrix and, hence, the matrix B must also be nonsingular. Thus, $\text{rk}(\Xi B) = K - r$ and there can be at most r zero columns in the matrix ΞB. In other words, at most r of the structural innovations can have transitory effects and at least $K-r$ of them must have permanent effects. From this fact it follows that a just-identified system can be obtained by imposing $r(r-1)/2$ additional restrictions on the transitory shocks and $(K-r)((K-r)-1)/2$ restrictions on the permanent shocks (see, e.g., King et al. (1991), Gonzalo & Ng (2001)). The transitory shocks may be identified, for example, by placing zero restrictions on B directly and thereby specifying that certain shocks have no instantaneous impact on some of the variables. Generally, identifying restrictions are often of the form

$$C_{\Xi B} \text{vec}(\Xi B) = c_l \quad \text{and} \quad C_s \text{vec}(B) = c_s, \tag{2.3}$$

where $C_{\Xi B}$ and C_s are appropriate selection matrices to specify the long-run and contemporaneous restrictions, respectively, and c_l and c_s are vectors of suitable dimensions. In practice, the latter vectors are typically zero. In other words, zero restrictions are specified in (2.3) for ΞB and B. The first set of restrictions can be written alternatively as

$$C_l \text{vec}(B) = c_l, \tag{2.4}$$

where $C_l \equiv C_{\Xi B}(I_K \otimes \Xi)$ is a matrix of long-run restrictions on B. Precise conditions for local just-identification may be found in Lütkepohl (2005, Proposition 9.4).

3 Estimation

Assuming that the lag order, $p-1$, the cointegrating rank, r, and structural identifying restrictions are given, a VECM can be estimated by concentrating out the reduced form parameters and then estimating B as described in the following.

Estimators of the reduced form parameters of the VECM (2.1) are available via the Johansen (1995) Gaussian maximum likelihood (ML) procedure. Replacing the reduced form parameters by their ML estimators gives the concentrated log-likelihood function

$$\log l_c(B) = \text{constant} - \frac{T}{2}\log|B|^2 - \frac{T}{2}\text{tr}(B'^{-1}B^{-1}\widetilde{\Sigma}_u), \tag{3.1}$$

where $\widetilde{\Sigma}_u = T^{-1}\sum_{t=1}^{T}\widehat{u}_t\widehat{u}_t'$ and the $\widehat{u}_t$'s are the estimated reduced form residuals. Maximization of this function with respect to B subject to the structural restrictions has to be done by numerical methods because a closed form solution is usually not available (see Lütkepohl (2005, Chapter 9) for details).

Under usual assumptions, the ML estimator of B, $\widehat{B}$ say, is consistent and asymptotically normal,

$$\sqrt{T}\text{vec}(\widehat{B} - B) \xrightarrow{d} \mathcal{N}(0, \Sigma_{\widehat{B}}). \tag{3.2}$$

Expressions for the covariance matrix of the asymptotic distribution in terms of the model parameters can be obtained by working out the corresponding information matrix (see Vlaar (2004)). For practical purposes, bootstrap methods are in common use for inference in this context.

The result in (3.2) implies that the t-ratios of elements with regular asymptotic distributions can be used for assessing the significance of individual parameters, provided the corresponding over-identifying zero restriction is a valid one. In other words, the zero value of the corresponding parameter must be within the admissible parameter space. This is not always obvious as we will argue in the following. Clearly, the asymptotic distribution of $\widehat{B}$ is singular, because of the identifying restrictions that have been imposed on B. Therefore F-tests will in general not be valid and have to be interpreted cautiously.

The problem related to the asymptotic properties of $\widehat{B}$ comes about because it is not obvious which over-identifying restrictions are admissible, that is, which over-identifying restrictions result in a nonsingular matrix B, due to the way the long-run restrictions are set up. It may be instructive to look at an example to see this problem more clearly.

Consider a three-dimensional system ($K = 3$) where all variables are $I(1)$ and which has cointegrating rank $r = 1$. Then there can be at most one transitory shock which is identified without further restrictions (except that its position and sign must be specified). If there is indeed a transitory shock, there are just two permanent shocks which are identified by one further restriction. Suppose the corresponding identifying restrictions are specified as follows:

$$\Xi B = \begin{bmatrix} * & * & 0 \\ * & * & 0 \\ * & * & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} * & 0 & * \\ * & * & * \\ * & * & * \end{bmatrix}. \tag{3.3}$$

In these matrices the asterisks denote unrestricted elements. Thus, the last element in ε_t is the transitory shock and the first two elements are permanent shocks. One restriction is placed on B to identify the two permanent shocks. The way it is specified, the second shock does not have an instantaneous effect on the first variable.

In this example it can be shown that any zero restriction placed on the last column of B will make the matrix singular and is therefore inadmissible (see Proposition 1 below). One implication of this result is that asymptotic (or bootstrap) t-ratios attached to the elements in the last column of B cannot be used to test whether the corresponding parameters are significantly different from zero. As a further implication, all instantaneous responses to the transitory shock must be nonzero as a consequence of the identifying restrictions imposed in (3.3). Thus, zero instantaneous responses are ruled out and, hence, the asymptotic or bootstrap confidence intervals cannot be used to assess whether there is no instantaneous response of some variable even if zero is included in the confidence interval set up in the usual way. Clearly, the latter situation is possible and can even occur if bootstrap methods are used. An example will be provided in Section 4. I will now present a criterion for deciding on inadmissible restrictions. A proof is given in the Appendix.

Proposition 1

In the model (2.1), suppose the $(K \times (K - r))$ matrix $\alpha_\perp$ is such that all sets of $K - r$ rows are nonsingular $((K - r) \times (K - r))$ matrices and there are $r^* \leq r$ transitory shocks. Then the number of admissible zero restrictions placed on a column of B associated with a transitory shock cannot be greater than $r - 1$. $\square$

Clearly this proposition confirms what was discussed in the context of the previous example. If $r = 1$ and there is one transitory shock (i.e., $r^* = r$), then there cannot be any zero restriction on the column of B corresponding to the transitory shock. If $r = 2$ and there are two transitory shocks, there can be at most one zero restriction on each of the two columns of B corresponding to the transitory shocks. For example, in a three-dimensional system with just-identifying restrictions

$$\Xi B = \begin{bmatrix} * & 0 & 0 \\ * & 0 & 0 \\ * & 0 & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} * & * & * \\ * & * & 0 \\ * & * & * \end{bmatrix} \tag{3.4}$$

the last two shocks are transitory. Hence, there can only be at most one zero restriction on each of the last two columns of B. Thus, no further zero restriction can be imposed on the last column because there is already one identifying zero restriction on this column. Moreover, only one zero restriction can be imposed on the second column of B.

If $r = 2$ and there is only one transitory shock ($r^* = 1$), say the last one in ε_t, then we may have just identifying restrictions of the form

$$\Xi B = \begin{bmatrix} * & * & 0 \\ * & * & 0 \\ * & * & 0 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} * & 0 & * \\ 0 & * & * \\ * & * & * \end{bmatrix}. \tag{3.5}$$

In this case again only one over-identifying zero restriction can be imposed on the last column of B.

It may be worth commenting on the condition that all sets of $K - r$ rows of $\alpha_\perp$ have to be nonsingular $((K-r)\times(K-r))$ matrices. Because α is assumed to be estimated unrestrictedly, the condition will be satisfied for the corresponding estimator $\widehat{\alpha}_\perp$ with probability one. Thus, in practice the condition will be satisfied if no restrictions are imposed on α in the reduced form estimation procedure. The situation may be different, however, if weak exogeneity restrictions are imposed, for example.

So far short-run restrictions for instantaneous effects have been dealt with. A similar problem also arises for the long-run restrictions imposed on ΞB, however. The next proposition deals with this case. It is also proven in the Appendix.

Proposition 2

In the model (2.1), suppose that each $((K - r) \times (K - r))$ submatrix of $\beta_\perp$ is nonsingular and there are r transitory shocks. Then the number of admissible zero restrictions placed on a column of ΞB associated with a permanent shock cannot be greater than $K - r - 1$. □

Note that in this proposition it is assumed that the number of transitory shocks is identical to the cointegrating rank r. Otherwise placing zero restrictions on individual elements of ΞB may not be possible. Consider, for instance, the example in (3.5) where the cointegrating rank is two whereas there is only one transitory shock. Then the first two columns of ΞB form a matrix of rank one. Thus, no single element can be restricted to zero individually.

For the examples (3.3) and (3.4), where the number of transitory shocks is identical to the cointegrating rank, Proposition 2 applies, however. It implies that in (3.4) no further valid zero restriction can be imposed on ΞB because $K - r - 1 = 0$. Moreover, in (3.3) at most one zero restriction can be imposed on each of the two first columns of the ΞB matrix because in that case $K - r - 1 = 1$. Note that the nonsingularity condition for the submatrices of $\beta_\perp$ is not a critical one if an estimated β matrix is considered and no specific restrictions are placed on β for the reasons discussed in the context of Proposition 1 with respect to $\alpha_\perp$.

Clearly, the propositions also have implications for more general tests for over-identifying restrictions in structural VECMs. An LR test is a standard tool in this context. Suppose there are over-identifying restrictions for B. In that case, $\widehat{B}\widehat{B}'$ will not be equal to the reduced form white noise covariance estimator $\widetilde{\Sigma}_u$ and the LR statistic is

$$\lambda_{LR} = T(\log|\widehat{B}\widehat{B}'| - \log|\widetilde{\Sigma}_u|). \tag{3.6}$$

It has an asymptotic χ^2-distribution with degrees of freedom equal to the number of over-identifying restrictions, if the null hypothesis holds and the restrictions are admissible. If inadmissible restrictions are imposed, this will result in a zero determinant term $|\widehat{B}\widehat{B}'|$ and a program used to compute the LR statistic should return an error message because the log cannot be evaluated. Thus, making an error here is perhaps not likely. Still Propositions 1 and 2 can be helpful to indicate how to avoid problematic restrictions in the first place.

Another implication of the propositions is that inference for certain impulse responses may be problematic. In particular, the usual confidence intervals for the instantaneous responses may be misleading and cannot be interpreted in the standard way if the confidence intervals contain zero. In the next section these issues will be illustrated by means of an example based on real data. It will also be shown that the usual bootstrap confidence intervals may in fact contain zero even if zero is not an admissible value of the response to a particular impulse.

4 An Example

I use the King et al. (1991) data for the three quarterly U.S. variables log private output (q_t), consumption (c_t), and investment (i_t) (all multiplied by 100) to illustrate the theoretical points of the previous section.[2] Data are available for the period 1947Q1–1988Q4. These data are also used in Chapter 9 of Lütkepohl (2005) where a reduced form VECM with one lagged difference, cointegrating rank $r = 2$ and an unrestricted intercept term is fitted. I use the same model in the following.

Because $r = 2$, there can be two transitory shocks and it is assumed that in fact the last two components in the ε_t vector are transitory shocks. An identifying zero restriction is imposed on B as in (3.4). In other words, the just-identifying zero restrictions on ΞB and B are the same as in (3.4). The following ML estimates are obtained with standard errors based on 2000 bootstrap replications in parentheses underneath the estimates:

$$\widehat{\Xi}\widehat{B} = \begin{bmatrix} \underset{(0.91)}{-0.71} & 0 & 0 \\ \underset{(0.98)}{-0.76} & 0 & 0 \\ \underset{(0.89)}{-0.69} & 0 & 0 \end{bmatrix}, \qquad \widehat{B} = \begin{bmatrix} \underset{(0.18)}{0.08} & \underset{(0.25)}{1.03} & \underset{(0.57)}{-0.45} \\ \underset{(0.82)}{-0.60} & \underset{(0.10)}{0.43} & 0 \\ \underset{(0.42)}{0.26} & \underset{(0.37)}{1.96} & \underset{(0.50)}{1.00} \end{bmatrix}. \tag{4.1}$$

Although all standard errors of the elements in the first column of $\widehat{\Xi}\widehat{B}$ are large relative to the estimated long-run effects of the permanent shock, according to Proposition 2 we cannot conclude that any one of the effects is not significantly different from zero because no zero restriction can be imposed on the first column of ΞB. Moreover, because $r = 2$, there can be at most one zero restriction in each of the last two columns of B which correspond to the transitory shocks (see Proposition 1). As there is already one identifying zero restriction placed in the last column, we cannot test whether the other two elements are significantly different from zero. On the other hand, we can test the three elements in the second column individually.

To illustrate the implications for an impulse response analysis, the structural impulse responses are depicted in Figure 1 together with 95% confidence intervals generated by 2000 bootstrap replications. These are Efron percentile intervals in the terminology of Benkwitz, Lütkepohl & Wolters (2001) and may not be the best possible confidence intervals in the present context.[3] They are used here because they nicely illustrate the issues discussed in

[2]The data are available at the website http://www.wws.princeton.edu/mwatson/.

[3]Computations were done with JMulTi (Lütkepohl & Krätzig (2004)).

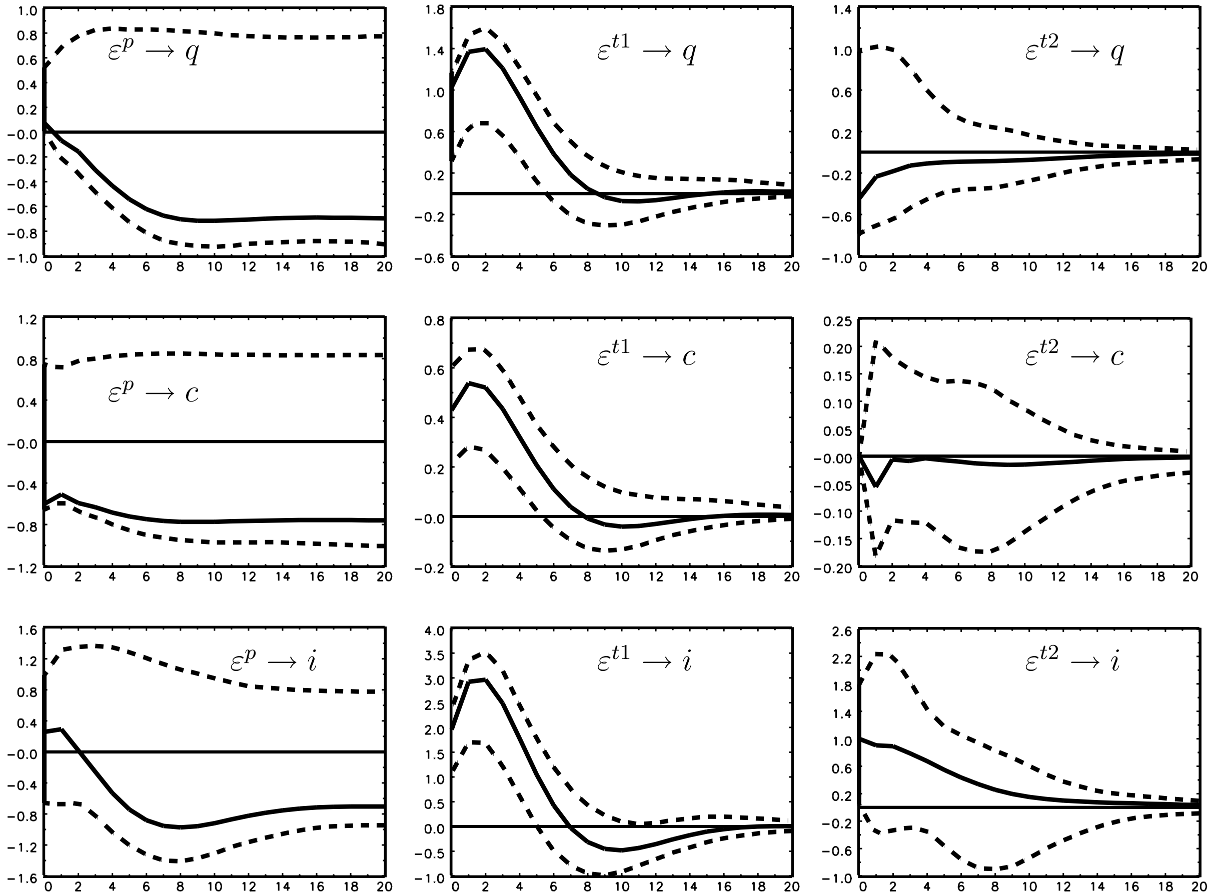

Figure 1: Responses of output (q), consumption (c), and investment (i) to a permanent shock (ε^p) and two transitory shocks (ε^{t1} and ε^{t2}) with 95% Efron percentile bootstrap confidence intervals based on 2000 bootstrap replications.

the previous section and they are perhaps the most commonly used confidence intervals for impulse responses in practice. Although most of the instantaneous effects have confidence intervals which include zero, some of them cannot be zero because a zero effect would imply a singular B matrix. For example, the second transitory impulse must have a nonzero instantaneous effect on the first variable ($\varepsilon^{t2} \to q$) although the corresponding confidence interval includes zero (see the last panel in the first row of Fig. 1). Clearly, the confidence interval is misleading or at least does not permit a standard interpretation. It may also be worth noting that one implication of the identifying zero restriction on the last column of B is that the instantaneous effects of the second transitory shock on the first and last variables ($\varepsilon^{t2} \to q$ and $\varepsilon^{t2} \to i$) are nonzero. This implication may not always be desired or apparent when one thinks about the identifying restrictions.

5 Conclusions

In this note I have pointed out a problem with imposing over-identifying restrictions in structural VECMs with cointegrated variables and long-run restrictions. It is shown that they may result in a singular reduced form residual covariance matrix without this being obvious. Therefore some care is necessary in imposing over-identifying restrictions in these models. Moreover, inference regarding the instantaneous and long-run effects of structural shocks can be problematic. In particular, interpreting t-ratios in the usual way as indicators of the significance of the instantaneous or long-run effects may not be meaningful. These results also have obvious implications for impulse response analysis. For example, confidence intervals of some impulse responses have to be interpreted with great care. Conditions were derived that can help in overcoming these problems and an empirical example is presented which illustrates the theoretical issues.

Appendix. Proofs

Proof of Proposition 1

Without loss of generality suppose that the r^* transitory shocks are the last r^* elements of ε_t. Thus, $\Xi B = [\Theta : 0_{K\times r^*}]$ and, taking into account the structure of Ξ, $\alpha'_\perp B = \alpha'_\perp[B_1 : B_2] = [\alpha'_\perp B_1 : 0_{K\times r^*}]$, where B_1 and B_2 are $(K \times (K - r^*))$ and $(K \times r^*)$ submatrices of B, respectively. Thus, $\alpha'_\perp B_2 = 0_{(K-r)\times r^*}$.

Suppose b is an arbitrary column of B_2 and there are r zero elements in b. Moreover, suppose that Λ is a permutation matrix such that

$$\Lambda b = \begin{bmatrix} b_1^* \\ 0_{r\times 1} \end{bmatrix}$$

and note that Λ is an orthogonal matrix. Then

$$0_{(K-r)\times 1} = \alpha'_\perp \Lambda' \Lambda b = \alpha'_\perp \Lambda' \begin{bmatrix} b_1^* \\ 0_{r\times 1} \end{bmatrix} = \alpha'_\perp \Lambda' \begin{bmatrix} I_{K-r} & 0_{(K-r)\times r} \\ 0_{r\times(K-r)} & 0_{r\times r} \end{bmatrix} \begin{bmatrix} b_1^* \\ 0_{r\times 1} \end{bmatrix}$$

implies $b_1^* = 0_{(K-r)\times 1}$ because

$$\alpha'_\perp \Lambda' \begin{bmatrix} I_{K-r} \\ 0_{r\times(K-r)} \end{bmatrix}$$

is nonsingular by assumption. Hence, if there are r zero elements in b, it follows that $b = 0_{K\times 1}$ which contradicts the fact that B is nonsingular and, thus, cannot have a zero column. Thereby we have shown that none of the columns associated to transitory shocks can have r zero elements and, hence, Proposition 1 is proven.

Proof of Proposition 2

Define the $(K \times (K-r))$ matrix $\eta = \beta_\perp \left[\alpha'_\perp \left(I_K - \sum_{i=1}^{p-1} \Gamma_i\right) \beta_\perp\right]^{-1}$ and note that each $((K-r) \times (K-r))$ dimensional submatrix of η is nonsingular by the assumptions of Proposition 2. Suppose that the last r shocks are transitory. Hence,

$$\Xi B = \eta\alpha'_\perp B = \eta\alpha'_\perp[B_1 : B_2] = [\Theta : 0_{K\times r}],$$

where B_1 and B_2 are $(K\times(K-r))$ and $(K\times r)$ matrices, respectively, and Θ is $(K\times(K-r))$. Let θ be a column of Θ with $K-r$ zeros and denote by η^* the $((K-r)\times(K-r))$ submatrix of η consisting of the same rows where the zeros appear in θ. Then $\eta^*\alpha'_\perp b = 0$, where b is the column of B_1 corresponding to θ (i.e., $\Xi b = \theta$). Due to the nonsingularity of η^*, $\alpha'_\perp b = 0$ so that $b \in \text{span}(\alpha)$. However, B_2 is a basis of $\text{span}(\alpha)$. Hence, B must be singular because one of the columns of B_1 is in $\text{span}(B_2)$. This contradicts the assumptions of the model and thereby proves Proposition 2.

References

Benkwitz, A., Lütkepohl, H. & Wolters, J. (2001). Comparison of bootstrap confidence intervals for impulse responses of German monetary systems, *Macroeconomic Dynamics* **5**: 81–100.

Gonzalo, J. & Ng, S. (2001). A systematic framework for analyzing the dynamic effects of permanent and transitory shocks, *Journal of Economic Dynamics & Control* **25**: 1527–1546.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.

King, R. G., Plosser, C. I., Stock, J. H. & Watson, M. W. (1991). Stochastic trends and economic fluctuations, *American Economic Review* **81**: 819–840.

Lütkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*, Springer-Verlag, Berlin.

Lütkepohl, H. & Krätzig, M. (eds) (2004). *Applied Time Series Econometrics*, Cambridge University Press, Cambridge.

Vlaar, P. J. G. (2004). On the asymptotic distribution of impulse response functions with long-run restrictions, *Econometric Theory* **20**: 891–903.