A note on weak exogeneity in VAR cointegrated models

Juan J. Dolado *

Bank of Spain, Madrid, Spain

Received 11 October 1991
Accepted 6 December 1991

In this note an extension of the traditional definition of weak exogeneity when the variables are I(1) and cointegrated is presented. In particular, the concept of long-run weak exogeneity is introduced when the parameters of interest are the elements of the cointegrating vectors. This definition extends some previously available definitions in this framework.

1. Introduction

This note addresses the issue of which definition of exogeneity is needed to carry out efficient estimation in the context of single equation conditional dynamic models estimated by least squares. In particular, we examine if the concept of weak exogeneity, introduced by Engle et al. (1983) as a sufficient condition for valid inference on the parameters of a conditional distribution in a framework of I(0) variables, still holds when variables are I(1) and there is cointegration. It is shown that the traditional definition is too strong and some examples are provided where there is no weak exogeneity but conditional inference is still efficient. This is an interesting issue since economic systems often have so many potentially useful variables that the system gets very large and then it may be worth to consider a single equation framework, where a variable is modelled conditional on the remaining variables.

The paper is organized as follows. Section 1 discusses the data generation process (DGP) and introduces the relevant assumptions and notation. Section 2 analyses the concept of long-run weak exogeneity associated to the parameters of the cointegrating vector. Finally, Section 3 extends the traditional definition of weak exogeneity to a nonstationary framework.

2. The vector autoregressive model

Let us consider an n dimensional vector time series $\{x_t,\ t=1, 2\dots\}$, which follows a vector autoregressive (VAR) process of order k,

$$x_t = \sum_{j=1}^{k} \Pi_j x_{t-j} + \epsilon_t, \tag{1}$$

Correspondence to: Juan J. Dolado, Research Department, Bank of Spain, Alcalá 50, 28014-Madrid, Spain.
* I am grateful to Manuel Arellano, Peter Boswijk and Soren Johansen for helpful comments.

where $\{\epsilon\}$ is a Gaussian innovation process relative to the past history of $\{x_t\}$ with variance matrix Ω. We define the matrix polynomial

$$\Pi(z) = I - \Pi_1 z - \ldots - \Pi_k z^k. \tag{2}$$

Following Johansen (1990), we state:

Assumption (Cointegration). System (1) verifies

(i) $\Pi(z)| = 0$ has all its roots outside the unit circle or equal to 1.
(ii) $\Pi(1) = -\alpha\beta'$, where α and β are $n \times r$ matrices of rank r, $0 \leq r \leq n$, and rank $\Pi(1) = r$.
(iii) rank $\alpha'_{\perp}(\Pi_1(1))\beta_{\perp} = n - r$, where $\alpha_{\perp}$ and $\beta_{\perp}$ are the $n \times (n-r)$ orthogonal complement matrices of α and β, of rank $n - r$, such that $\alpha'_{\perp}\alpha = 0$, $\beta'_{\perp}\beta = 0$, and $\Pi_1(z) = \partial\Pi(z)/\partial z$.

Together these assumptions imply that the components of x_t are I(1), whereas the linear combinations $\beta' x_t$ are stationary. Hence the components of x_t are cointegrated with cointegrating matrix β and cointegrating rank r. Moreover, system (1) can be expressed as an error correction model (ECM),

$$\Delta x_t = \sum_1^{k-1} \Gamma_j \Delta x_{t-j} + \alpha\beta' x_{t-k} + \epsilon_t,$$

with

$$\Gamma_j = \sum_1^{j} \Pi_i. \tag{3}$$

In order to discuss the problem of exogeneity in the context of a single equation, let us use the $(1 \times n)$ vector $e_1' = (1, 0, \ldots, 0)$ to select, without loss of generality, the first equation and E_2' be a $(n+1 \times n)$ marix of vectors orthogonal to e_1, which selects the remaining $(n-1)$ variables. Partitioning ϵ_t and Ω conformably, given that $\{\epsilon_t\}$ is a Gaussian process, we have

$$\mathrm{E}(\epsilon_t / F_{t-1}) = \sigma_{12}' \Omega_{22}^{-1} \epsilon_{21},$$

$$V(\epsilon_t / F_{t-1}) = \sigma_{11} - \sigma_{12}' \Omega_{22}' \sigma_{12} = \sigma_{11.2},$$

where F_{t-1} is the σ-field generated by $\{\epsilon_{2t},\ X_{t-j},\ j = 1, 2 \ldots\}$ and σ_{12} is an $(n - 1 \times 1)$ column vector.

Therefore, the conditional distribution for the first equation can be written as

$$\Delta x_{1t} = e_1' \Delta x_t = \left(e_1 - E_2 \Omega_{22}^{-1} \sigma_{12}\right)' \left(\sum_1^{k-1} \Gamma_i \Delta x_{t-1} + \alpha\beta' x_{t-k} \right) + \sigma_{12}' \Omega_{22}^{-1} \Delta x_t + \eta_t, \tag{4}$$

with $\eta_t \sim \mathrm{nid}(0, \sigma_{11.2})$.

The marginal distribution for the remaining $(n-1)$ elements in Δx_t is given by

$$\Delta x_{2t} = E_2'\Delta x_t = \sum_{1}^{k-1} E_2'\Gamma_i \Delta x_{t-1} + E_2'\alpha\beta' x_{t-k} + E_2' e_t, \tag{5}$$

i.e., the distribution of x_{2t} given x_{1t-1}, x_{2t-1}, etc.

3. Long-run weak exogeneity

Engle, Hendry and Richard (1983) introduced the concept of weak exogeneity with respect to the parameters of the conditional distribution (4) as a condition which allows to carry out optimal inference with respect to a set of parameters disregarding the information on the process generating the marginal distribution. If $D(x_t/X_{t-1}, \lambda)$ denotes the joint distribution of the x_t's conditional on the past history $X_{t-1} = \{x_{t-j};\ j = 1, 2\ldots\}$, then D can be factorized as follows:

$$D(x_t/X_{t-1};\ \lambda) = D(x_{1t}/x_{2t},\ X_{t-1};\ \lambda_1) D(x_{2t}/X_{t-1};\ \lambda_2). \tag{6}$$

Then, Engle et al. (1983) define x_{2t} as weakly exogenous for a set of parameters of interest Ψ if and only if there exists a partition (λ_1, λ_2) of λ such that: (i) ψ is a function of λ_1 alone, and (ii) λ_1 and λ_2 are variation free, so that there are no restrictions linking λ_1 with λ_2.

Following the previous definition, Boswijk (1991) and Johansen (1990) find x_{2t} as weakly exogenous for $\Psi = (\alpha, \beta)$ if $E_2'\alpha = 0$, that is,

$$\alpha = \begin{pmatrix} \alpha_1 \\ 0 \end{pmatrix}_{n-1}^{1}.$$

Because $E_2'\alpha = 0$, the marginal model of x_{2t} conditional on past history does not contain information on the cointegrating vectors in β and the error correction coefficient α_1.

However, as we will prove in what follows, the previous condition is too strong. To prove a weaker condition, we rely upon a recent result by Phillips (1991) on optimal inference in cointegrated systems. Phillips' Theorem 2 can be used to show that the ML estimator of β in (4) is asymptotically equivalent to the FIML estimator in (3) iff $\Omega_{12}^1 = 0$, where Ω^1 is the long-run variance between η_t and Δx_{2t}. Hence, there is no loss of efficiency in estimating β from the conditional model. To obtain the off-diagonal terms in the long-run variance–covariance matrix, we use the MA representation of the cointegrated system, as derived by Johansen (1989, Theorem 3.1),

$$\Delta x_t = C(L)\epsilon_t, \tag{7}$$

where $c(1) = \beta_\perp(\alpha'_\perp \Pi_1(1)\beta_\perp)^{-1}\alpha'_\perp$ which exists by assumption (ii).

Since $\eta_t = (e_1 - E_2\Omega_{22}^{-1}\sigma_{12})'\epsilon_t$ then $\Omega_{12}^1 = 0$ implies that

$$\begin{aligned}&\left(e_1 - E_2 Q_{22}^{-1}\sigma_{12}\right)' \Omega\alpha_\perp \left(\beta'_\perp \Pi_1'(1)\alpha_\perp\right)^{-1}\beta'_\perp E_2 \\ &\quad = (\sigma_{11.2}, 0)'\alpha_\perp\left(\beta'_\perp \Pi_1'(1)\alpha_\perp\right)^{-1}\beta'_\perp E_2 = (\sigma_{11.2}, 0)' C(1)' E_2 = 0\end{aligned} \tag{8}$$

and this is a necessary and sufficient condition for x_{2t} being weakly exogenous with respect to β in this framework. We will denote this property as *long-run weak exogeneity*.

Note that $E_2'\alpha = 0$ is a sufficient condition for (8) to hold, since $E_2'\alpha = 0$ implies $\alpha_2 = 0$. Then $\alpha_\perp$ is any matrix of the form (0, Q), where Q is full rank, and the first three terms in (8) can be written as

$$(1, -\sigma_{12}\Omega_{22}^{-1})\begin{pmatrix} e_1' \\ E_2' \end{pmatrix}\Omega\begin{pmatrix} 0 \\ Q \end{pmatrix} = (\sigma_{11.2}, 0)\begin{pmatrix} 0 \\ Q \end{pmatrix} = 0.$$

However, it is not necessary since it could happen that $\beta'_\perp E_2$ has a deficient column rank and

$$(\sigma_{11.2}, 0)\alpha_\perp(\beta'_\perp \Pi_1(1)'\alpha_\perp)^{-1}$$

happens to lie in its null space. The latter restriction holds in the following example. Let $n = 3$ and

$$\beta = (0, 1, 1)'; \quad \alpha = (-0.1, 0.1, 0.1);$$

$$\beta'_\perp = \begin{pmatrix} 1 & 1 & -1 \\ 2 & -1 & 1 \end{pmatrix}; \quad \alpha'_\perp = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 1 & 0 \end{pmatrix}; \quad \alpha'_\perp \Pi_1(1)\beta_\perp = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

It is easy to check that this example verifies assumptions (i)–(iii) and that (8) holds in spite of $\alpha_2 = 0$.

4. A new definition of weak exogeneity when variables are I(1)

In the light of the previous discussion, the traditional definition of weak exogeneity in an I(0) framework can be amended in the following way. Let us define two sets of parameters of interest: the long-run parameters $\Psi_1 = \{\beta\}$ and the short-run parameters

$$\Psi_s = \{\alpha_1, \sigma_{12}'\Omega_{22}^{-1}, (e_1' - \sigma_{12}'\Omega_{22}^{-1}E_2')\Gamma_i, \sigma_{11.2}\}.$$

Retaining the assumptions (i)–(iii) the conditional and marginal densities in (6) can be reformulated as

$$D(\Delta x_t/X_{t-1}; \lambda) = D(\Delta x_{1t}/\Delta x_{2t}, X_{t-1}; \lambda_1)D(\Delta x_{2t}/X_{t-1}; \lambda_2).$$

Then, x_{2t} is weakly exogenous with respect to Ψ if: (i) Ψ_1 is a function of λ_1 and satisfies long-run weak exogeneity and (ii) Ψ_s is a function of λ_1 and λ_1 and λ_2 are variation free.

The following example serves as an illustration of the previous concepts. Let the DGP be

$$x_{1t} = \beta x_{2t} + \epsilon_{1t}, \tag{9}$$

$$\Delta x_{2t} = \delta_1 \Delta\epsilon_{1t-1} + \epsilon_{2t}, \tag{10}$$

with $\epsilon_{it} \sim \text{nid}(0, \sigma_i^2)$, $\text{E}(\epsilon_{1t}\epsilon_{2s}) = 0$.

Then, the OLS estimator of β in (9) is given by

$$T(\hat\beta - \beta) = T^{-1}\sum x_{2t}\epsilon_{1t}/T^{-2}\sum x_{2t}^2 \Rightarrow \sigma_1\int_0^1 B_2(r)\,\mathrm{d}B_1(r)/\sigma_2\int_0^1 B_2^2(r)\,\mathrm{d}r, \tag{11}$$

convergence. Phillips (1991) has shown that (11) is a 'mixture of normals' and hence the OLS estimator of β in the conditional model is asymptotically equivalent to the ML estimator of β in the full system. Thus, no asymptotic efficiency is lost. However, it is important to note that since $\delta_1 \Delta\epsilon_{1t-1}$ appears in the marginal density, x_{2t} is not weakly exogenous in the traditional sense, though it is long-run weakly exogenous. This can be easily seen since $\Delta\epsilon_{1t-1}$ vanishes in the long-run variance–covariance matrix and hence $\Omega^{\text{l}}_{12} = 0$, as required [see Phillips (1991, p. 300, remark (m))]. Thus, full 'variation freeness' is too strong.

References

Boswijk, P.H., 1991, Testing for cointegration in structural models, Report AE 7/91 (University of Amsterdam).

Engle, R.F., D.F. Hendry and J.F. Richard, 1983, Exogeneity, Econometrica 51, 277–304.

Johansen, S., 1989, Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, Preprint 1989, no. 3 (Institute of Mathematical Statistics, University of Copenhagen); forthcoming in Econometrica.

Johansen, S., 1990, Cointegration in partial systems and the efficiency of single equation analysis, Preprint 1990 (Institute of Mathematical Statistics, University of Copenhagen).

Phillips, P.C.B., 1991, Optimal inference in cointegrated systems, Econometrica 59, 283–306.