Further results on weak-exogeneity in vector error correction models

Christophe RAULT*(EPEE, University of Evry and EUREQua, Sorbonne University)

December 16, 2003

Abstract

This paper provides a necessary and sufficient condition for weak exogeneity in vector error correction models. An interesting property is that the statistics involved in the sequential procedure for testing this condition are distributed as χ^2 variables and can therefore easily be calculated with usual statistical computer packages, which makes our approach fully operational empirically. Finally, the power and size distortions of this sequential test procedure are analysed with Monte-Carlo simulations.

I Introduction

This last decade considerable interest has been shown in the issue of weak exogeneity testing in a linear Vector Error Correction Model (VECM) with I (1) variables (see for instance Ericsson and alii,1998; Hecq *et al.*, 2000; Hendry and Mizon, 1993; Johansen, 1992, 1995; Urbain, 1992, 1995; Rault *et al*, 2003). Weak-exogeneity has also been extensively discussed in the two special issues of the *Journal of Policy Modeling (1992), vol 14, n° 3* and of the *Journal of Business and Economic Statistics (1998), vol 6, n° 4*, and is now widely recognized as a crucial concept for applied economic modeling[1].

The motivation of this paper rests upon two key observations on recent theoretical works in $VECM$.

- Firstly, the usual weak-exogeneity conditions which can be expressed in term of coefficient nullities are easily testable but sometimes imply "overly strong" restrictions. The conditions of Johansen (1992, cf. theorem 1) and Urbain (1992, cf. proposition 1) for instance, which are widely used in applied works, forbid the existence of long run relationships in the equations describing the evolution of the (weakly) exogenous variables. These equations are thus a VAR model in first differences. Besides Johansen makes the assumption that macro-economists have a potential economic interest in all cointegrating relations existing between the variables being

*Paris I, Maison des Sciences de l'Economie, 106-112 bd. de L'Hôpital, 75647 Paris Cedex 13, France. E-mail : chrault@hotmail.com, web-site : http://www.multimania.com/chrault/index.html

[1]In this paper, we shall confine ourselves to the concept of weak exogeneity proposed by Richard (1980) and Engle *et al* (1983).

investigated. But it is actually far from being always the case and a typical difficulty sometimes arises when cointegration tests suggest in empirical applications the existence of r cointegrating vectors, whereas according to economic theory there should only exist m, with $m < r$.

- Secondly, the sufficient weak exogeneity conditions of Hendry and Mizon (1993) and of Ericsson *et al.* (1998, cf. lemma 2) which only consider a r_1 subset of the cointegrating relationships as parameters on interest, give *a priori* the partition of the r cointegrating vectors into r_1 and r_2. The first r_1 vectors then belong to the equations of the endogenous variables and the r_2 last appear in the equations of the exogenous variables. Furthermore, only the long-run parameters of the conditional model are considered as possible economic parameters of interest for macro-economists. Yet in some applied studies, they can also be interested in short-run parameters. Indeed, modeling the short run adjustment structure, i.e. the feedbacks to deviations from the long-run equilibrium, is an important step, because it can reveal information on the underlying economic structure.

To address the above issues we propose in this note an extension of the existing weak exogeneity conditions, which is based on a canonical decomposition of the long-run matrix Π. This representation exploits the fact that the β cointegrating and α loading factor matrices are not unique in so far as $\Pi = \alpha\beta^{'} = \left(\alpha\Psi^{-1}\right)\left(\Psi\beta^{'}\right)$ for any $r \times r$ non singular matrix Ψ. An interesting feature of this representation is that it enables us to give a necessary and sufficient condition for weak-exogeneity. An appealing aspect of this condition for the practitioner is that it can be tested using asymptotically chi-squared distributed test statistics which can easily be computed with most statistical computer packages.

The plan of the paper is as follows. Section II sets out the general $VECM$ framework. Section III introduces the canonical representation of the long run matrix Π and proposes a necessary and sufficient condition for weak exogeneity. Section IV deals with inference and testing which are conducted within the setting proposed by Johansen and reports some Monte Carlo simulations to analyse the asymptotic and finite sample properties of the sequential procedure developed here. Finally, concluding remarks are presented in section V and specific recommendations are provided for applied researchers.

II Cointegrated vector autoregressions

We begin by setting out the basic framework and thus consider an n-dimensional $VECM(p)$ process $\{X_t\}$, generated by

$$\Delta X_t = \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i} + \alpha\beta^{'} X_{t-1} + \varepsilon_t, t = 1, ..., T, \quad (1)$$

where Γ_i, α, β are, respectively $n \times n$, $n \times r$, $n \times r$, $0 < r < n$ matrices such that $\Pi = \alpha\beta^{'}$; The r linear combinations of X_t, the cointegrating vectors, $\beta^{'} X_t$, are often interpreted as deviations from equilibrium and α is the matrix of adjustment or feedback coefficients, which measure how strongly the r stationary variables $\beta^{'} X_{t-1}$ feedback onto the system. ε_t is an i.i.d normal distributed vector of errors, with a zero mean and a positive definite covariance matrix Σ; and p is a constant integer. To keep the notation as simple as possible we omit (without any loss of generality) deterministic components.

It is assumed in addition that (i) $\left|(I_n - \sum_{i=1}^{p-1} \Gamma_i z^i)(1-z) + \alpha\beta^{'} z\right| = 0$ implies either $|z| > 1$ or $z = 1$, and that (ii) the matrix $\alpha_{\perp}^{\prime}(I_n - \sum_{i=1}^{p} \Gamma_i)\beta_{\perp}$ is invertible, where $\beta_{\perp}$ and $\alpha_{\perp}$ are both full rank $n \times n - r$ matrices satisfying $\alpha_{\perp}^{\prime}\alpha_{\perp} = \beta_{\perp}^{'}\beta_{\perp} = 0$, which rules out the possibility that one or more elements of X_t are $I(2)$. These two conditions ensure that $\{X_t\}$ and $\left\{\beta^{'} X_t\right\}$ are respectively $I(1)$ and $I(0)$ and that the conditions of the Granger theorem (1987) are satisfied.

Consider now the partition of the n dimensional cointegrated vector time series $X_t = (Y_t^{'}, Z_t^{'})^{'}$ generated by equation (1), where Y_t and Z_t are distinct sub-vectors of dimension g $\times$ 1 and k $\times$ 1 respectively with $g + k = n$. In this writing Y_t and Z_t denote respectively the dependent and explanatory variables. Equation (1) can then be rewritten without loss of generality as a conditional model for Y_t given Z_t and a marginal model for Z_t, that is :

$$\begin{cases} \text{conditional model} \\ \Delta Y_t = \sum_{i=1}^{p-1} \Gamma_{YY,i}^{+} \Delta Y_{t-i} + \sum_{i=0}^{p-1} \Gamma_{YZ,i}^{+} \Delta Z_{t-i} + \alpha_Y^{+} \ \beta^{'} X_{t-1} + \eta_{Y,t} \\ \text{marginal model} \\ \Delta Z_t = \sum_{i=1}^{p-1} \Gamma_{ZY,i} \Delta Y_{t-i} + \sum_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} + \alpha_Z \ \beta^{'} X_{t-1} + \varepsilon_{Z,t} \end{cases} \tag{2}$$

$$\text{with} \begin{cases} \Gamma_{YY}^{+}(L) = \Gamma_{YY}(L) - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Gamma_{ZY}(L) = I_g - \sum_{i=1}^{p-1} \Gamma_{YY,i}^{+} L^i \\ \Gamma_{YZ}^{+}(L) = \Gamma_{YZ}(L) - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Gamma_{ZZ}(L) = -\sum_{i=0}^{p-1} \Gamma_{YZ,i}^{+} L^i \\ \alpha_Y^{+} = \alpha_Y - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\alpha_Z \\ \eta_{Yt} = \varepsilon_{Yt} - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\varepsilon_{Zt} \\ \Sigma_{YY}^{+} = \Sigma_{YY} - \Sigma_{YZ}\Sigma_{ZZ}^{-1} \end{cases}$$

where L denotes the lag operator

and $\begin{pmatrix} \eta_{Yt} \\ \varepsilon_{Zt} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{YY}^{+} & 0 \\ 0 & \Sigma_{ZZ} \end{pmatrix}\right]$

with the partitioning of the matrices Γ_i, α and β being conformable to that of X_t.

Equation (2) is known as the $VECM$ block recursive form and its main interest is to provide the analytic expression of the conditional error correction model. Note that the disturbance orthogonalization doesn't affect the equations describing the evolution of the Z_t variables, i.e. the marginal model.

III A necessary and sufficient condition for weak exogeneity in VECM models

As it is now well-admitted, the presence or lack of weak exogeneity[2] depends crucially on what parameters the focus of attention is, but contrary to what it is often assumed in a cointegrated framework, there is no obvious reason for the investigator to be necessarily interested in all cointegrating vectors (as it is assumed in Johansen, 1992, theorem 1)[3], or even in a structural partition of the cointegrating vectors made *a priori* (as it is the case in Ericsson *et al.*, 1998, lemma 2). Indeed, when dealing with $VECM$ models, the parameters of interest might be for instance only the cointegrating vectors that enter the conditional model, or both short-run and long-run parameters of the conditional model. There are two arguments for this.

Firstly, applied economists are usually interested in the parameters of the conditional model and not necessarily in those of the marginal model because the former represents short and long-run behavioral parameters of interest such as supply and demand elasticities, propension to consume or save, etc.... Indeed, when economists undertake practical modeling they are typically interested in building a model of either a single variable or a small subset of the variables. Many of the variables are there because economists think they are relevant to the determination of the variables they want to model but they are not interested in explaining them.

Secondly, cointegration tests often suggest in empirical applications the existence of r cointegrating vectors, whereas according to economic theory there should only exist m, with $m < r$. In this case the typical difficulty arises of how to interpret in an economic way the $(r - m)$ remaining statistical cointegrating relationships, which in many situations turn out to appear only in the equations of the conditioning variables.

In this section we propose a necessary and sufficient condition for weak exogeneity in the setting of a canonical decomposition of the Π matrix which takes the issues discussed above into account. Before going into the CNS condition for weak-exogeneity we need to consider the following preliminary theorem :

Theorem 1 *Let* $\Pi = \alpha\beta^{'}$ *be a* $n \times n$ *reduced rank matrix of rank* r $(0 < r < n)$ *and partition* α *into* $\begin{bmatrix} \alpha_Y \\ \alpha_Z \end{bmatrix}$.

[2] Let's remember that Engle *et al* (1983) define a vector of Z_t variables to be weakly-exogenous for the parameters of interest, if (i) the parameters of interest only depend on those of the conditional model, (ii) the parameters of the conditional and marginal models are variation free, i.e. there exists a sequential cut of the two parameters spaces (cf. Florens and Mouchart, 1980).

[3] In his careful discussion of Boswijk's paper (1995) on structural ECMs, Ericsson (1995) has already noted that this is an "overly strong hypothesis", since according to him, "any individual empirical investigation might reasonably restrict its focus to only a subset of the cointegrating vectors in the economy".

(i) If we define $m_1 = rank(\alpha_Y)$ with $m_1 > 0$ and $r - m_1 > 0$[4]*, then the α and β matrices can always be reparametrised as follows :*

$$\beta = [\beta_1 \ \beta_2] = \begin{bmatrix} \beta_{Y1} & \beta_{Y2} \\ \beta_{Z1} & \beta_{Z2} \end{bmatrix}$$

$$\alpha = [\alpha_1 \ \alpha_2] = \begin{bmatrix} \alpha_{Y1} & 0_{(g,r-m_1)} \\ \alpha_{Z1} & \alpha_{Z2} \end{bmatrix}$$

, where β_{Y1}, α_{Y1}, β_{Z1}, α_{Z1}, β_{Y2}, β_{Z2}, α_{Z2} are respectively $g \times m_1$, $g \times m_1$, $k \times m_1$, $k \times m_1$, $g \times r$-m_1, $k \times r$-m_1, $k \times r$-m_1 with rank $(\alpha_{Y1}) = m_1$ and rank $(\alpha_{Z2}) = r - m_1$.
(ii) m_1 is uniquely defined and is invariant to the chosen reparametrisation. It is such as[5]

$$max(0, r-k) \leq m_1 \leq min(g, r).$$

Proof. If α_Y has reduced rank m$_1$, then $\alpha_Y = \alpha_{Y1}\eta'$, for some g $\times$ m$_1$ matrix α_{Y1} and some n $\times$ m$_1$ matrix η. Now define new parameters $\beta_1' = \eta'\beta$ and $\beta_2' = \eta_\perp'\beta$ and $\alpha_{Z1} = \alpha_Z\eta(\eta'\eta)^{-1}, \alpha_{Z2} = \alpha_Z\eta_\perp(\eta_\perp'\eta_\perp)^{-1}$. Then $\alpha_Y\beta' = \alpha_{Y1}\eta'\beta = \alpha_{Y1}\beta_1 = (\alpha_{Y1}, 0)(\beta_1, \ \beta_2)'$ and $\alpha_Z\beta' = \alpha_Z(\eta(\eta'\eta)^{-1}\eta' + \eta_{\perp\perp}(\eta_\perp'\eta_\perp)^{-1}\eta_\perp')\beta' = \alpha_{Z1}\beta_1' + \alpha_{Z2}\beta_2' = (\alpha_{Z1}, \ \alpha_{Z2})(\beta_1, \ \beta_2)'$, which shows that theorem 1 is satisfied. ■

Under the reparametrisation of the α and β matrices, the conditional and marginal models (cf. equation 2) become :

$$\begin{cases} \text{conditional model} \\ \Delta Y_t = \sum_{i=1}^{p-1} \Gamma_{YY,i}^{+} \Delta Y_{t-i} + \sum_{i=0}^{p-1} \Gamma_{YZ,i}^{+} \Delta Z_{t-i} + \alpha_{Y1}^{+} \beta_1' X_{t-1} + \eta_{Y,t} \\ \text{marginal model} \\ \Delta Z_t = \sum_{i=1}^{p-1} \Gamma_{ZY,i} \Delta Y_{t-i} + \sum_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} + \alpha_{Z1}\beta_1' X_{t-1} + \alpha_{Z2}\beta_2' X_{t-1} + \varepsilon_{Z,t} \end{cases} \tag{2.a}$$

The canonical representation given in theorem 1 exploits the indeterminacy existing on the α and β matrices : it is indeed now well-known that the parameters of these two matrices are not separately identified without r^2 additional restrictions (cf. Bauwens and Lubrano, 1994), since for any non-singular matrix Ψ of dimensions (r, r), we could define $\Pi = \left(\alpha\Psi^{-1}\right)\left(\Psi\beta'\right)$, and $\alpha^* = \alpha\Psi^{-1}$, $\beta^* = \beta\Psi'$ would be equivalent matrices of adjustment coefficients and cointegrating vectors. Theorem 1 implies no loss of generality, and only requires the determination of the m_1 rank of the upper block of the α matrix, denoted α_Y[6] and reparametrised into $[\alpha_{Y1} \ 0_{(g,r-m_1)}]$.

We are in a position to state the following result :

[4]We assume that β_1 and β_2 each contain at least one cointegrating vector to exclude the case where β_1=β, which entails that β_2 is a null set.
[5]This condition is derived from $rank\,(\alpha) = r$.
[6]The way this rank can be determined in applied studies is discussed in section 4.

Proposition 2 *: **Necessary and sufficient weak exogeneity condition**. Suppose that the investigator's parameters of interest are those of the conditional model, i.e.* Ψ $=(\Gamma^{+}_{YY,i}$, $i = 1, ..., p-1$; $\Gamma^{+}_{YZ,i}, i = 0, ..., p-1$; α^{+}_{Y1};$\beta_1^{\prime})$*, then* Z_t *is weakly exogenous for* Ψ *if and only if* $\alpha_{Z1} = 0$ *in the canonical representation given by theorem 1.*

The proof follows the same line of arguments as those presented in Johansen (1992) and is omitted here to save space. Note that in our framework as in Johansen (1992) and Hendry and Mizon (1993), the parameters of interest are chosen prior to testing for weak exogeneity in the sense that they are the parameters of the conditional model which represents the subset of Y_t variables the investigator is interested in modeling conditionally on Z_t other variables (cf. the discussion above). Consequently, our approach also makes economic sense with economic theory typically providing the parameters of interest to the empirical researcher prior to the modeling exercise. Of course, a major difference with Hendry and Mizon's weak exogeneity condition (1993) which gives *a priori* the partition of β into $[\beta_1 \;\; \beta_2]$, so as β_1 and β_2 appear respectively in the conditional and marginal models, is that we determine explicitly this partition, exploiting the fact that the α and β are not unique (cf. infra). But the gain of doing this is that we are then able to provide a necessary and sufficient condition for weak-exogenity which is very convenient to use empirically since it only implies the nullity of some loading factors in the α matrix. One could object that in certain applied studies, the investigator might not consider all m_1 cointegrating vectors entering the conditional model as parameters of interest. This makes of course sense in some cases, and in such situations it is only the corresponding part of α_{Z1} which is required to vanish for weak exogeneity. Note that the argument is the same if only the parameters of specific equations in the conditional model are of structural interest for the purpose of the analysis.

It is important to observe that before testing for weak exogeneity, researchers usually impose identifying assumptions on the cointegrating vectors β. Then testing for weak exogeneity means testing zero restrictions on the α matrix. We cannot proceed likewise here since the canonical decomposition of the Π matrix given in theorem 1 entails a reparameterization of the cointegrating vectors resulting from the defined rotation of the cointegration space, but nothing guarantees that the resulting cointegrating vectors are of economic interest, even if some of the original just-identified cointegrating vectors were economically interpretable[7]. That is the reason why we have to test first for weak-exogeneity and then if weak-exogeneity of the conditioning variables Z_t is empirically satisfied estimation and identification of adjustments coefficients and cointegrating vectors can be carried out from the conditional model alone in the setting recently

[7]For instance, the two cointegrating vectors in Hendry and Mizon (1993) are interpretable as a money-demand relationship and an output relationship. A rotation, however, generates new cointegrating vectors that are linear combinations of those money-demand and output relationships, and those linear combinations are generally economically uninteresting because they confound the underlying economically interesting relationships.

proposed by Pesaran *et al* (2000), which is well-adapted for our purpose here. Indeed, Pesaran *et al* have generalized the existing cointegration analysis literature in two respects. Firstly, they examined the problem of efficient estimation of vector error correction models containing exogenous I(1) variables. Secondly, they considered the efficient estimation of vector error correction models subject to restrictions on the short-run dynamics, i.e. allowing the short-run dynamics to differ within and between equations.

For us, this implies that having determined empirically the correct split of the X_t vector under study into Y_t and (weakly exogenous) Z_t variables we can then follow Pesaran *et al* in imposing a set of just-identifying restrictions on the cointegrating vectors entering the conditional model. Then, the complete dynamic model may be estimated and the dynamics can be simpified at the same time as the over-identifying restrictions on the cointegrating vectors are tested using likelihood ratio tests.

Note that our CNS condition for weak exogeneity remains of course unchanged if only a subset of the parameters of the conditional model (for instance the long-run parameters) are of economic interest for the applied researcher. Furthermore, if the parameters of interest are given by the first g equations of the conventional $VECM$ partitioned into Y_t and Z_t, it only requires in addition that $\Sigma_{YZ} = 0$[8]. This latter case may turn out not to be very useful in practice since actually it seldom occurs that the empirical researcher has some structural interest in the unrestricted short-run dynamic parameters of the reduced form $VECM$, exception maybe in case of separation analyses (cf. Granger and Haldrup, 1997).

IV Inference and testing

The necessary and sufficient condition for weak exogeneity introduced in proposition 2 first requires to rewrite the Π matrix under the canonical decomposition given in theorem 1. Then, in this framework this condition has been expressed in term of coefficient nullities of the α matrix, which permits to use the conventional chi-squared statistics (see Johansen, 1995). As we have already noticed it, this representation requires the determination of the m_1 rank of the α_Y sub-matrix. The first subsection thus develops a sequential procedure to determine this specific rank. Then, the second subsection describes how to test the weak exogeneity condition introduced in proposition 2 and the third one reports some Monte Carlo simulations to analyse the size distortions and power of the sequential procedure of rank tests.

IV.1 Determination of the m_1 rank

[8] The proof is straightforward since as $\alpha_{Y1}^{+} = \alpha_{Y1} - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\,\alpha_{Z1}, \alpha_{Z1} = 0 \Longrightarrow \alpha_{Y1}^{+} = \alpha_{Y1}$.

The m_1 rank of the α_Y sub-matrix can be determined using a sequential test procedure very similar to that proposed by Rault (2000) (where it is the rank of a sub-matrix extracted from the β matrix which is investigated). This procedure is based on asymptotically chi-squared distributed LR statistics whose properties have been analyzed with Monte-Carlo experiments for known r.

More precisely, following Rault (2000) the m_1 rank of the α_Y sub-matrix can be determined as follow. First, define $m_a = min(g, r)$, $m_b = max(0, r-k)$ and then consider the following sequences of null hypotheses :

$$\begin{cases} H_{0,1} \begin{cases} \text{There exists a basis of the adjustment space such as} \\ \alpha = (H_1\theta_{r-m_a+1}, \kappa_{r-m_a+1}) \\ \text{with } H_1 = \begin{pmatrix} 0_{(g,k)} \\ I_k \end{pmatrix}, \text{ that is } m_b \leq rank\,(\alpha_Y) \leq m_a - 1. \end{cases} \\ \vdots \\ \text{for } j = 2, ..., m_a - m_b, \text{ as long as } H_{0,j-1} \text{ is not rejected,} \\ H_{0,j} \begin{cases} \text{There exists a basis of the adjustment space such as} \\ \alpha = (H_1\theta_{r-m_a+j}, \kappa_{r-m_a+j}) \\ \text{with } H_2 = \begin{pmatrix} 0_{(g,k)} \\ I_k \end{pmatrix}, \text{ that is } m_b \leq rank\,(\alpha_Y) \leq m_a - j, \end{cases} \end{cases}$$

These different hypotheses can be tested using the following sequential test procedure :

$$\begin{cases} \text{Step 1 : test } H_{0,1} \text{ with the } \xi_1 \text{ statistic at the } \alpha_1 \text{ level} \\ \text{and reject } H_{0,1} \ (\Longrightarrow rank\,(\alpha_Y) = m_a) \ \text{ if } \xi_1 \geq \chi^2_{1-\alpha_1}(v_1) \\ \vdots \\ \text{for } j = 2, ..., m_a - m_b, \text{ as long as } H_{0,j-1} \text{ is not rejected} \\ \text{Step j : test } H_{0,j} \text{ with the } \xi_j \text{ statistic at the } \alpha_j \text{ level} \\ \text{and reject } H_{0,j} \ (\Longrightarrow rank\,(\alpha_Y) = m_a - j + 1) \text{ if } \xi_j \geq \chi^2_{1-\alpha_j}(v_j), \\ \text{else accept } H_{0,j} \ (\Longrightarrow rank\,(\alpha_Y) = m_a - j) \text{ if } \xi_j < \chi^2_{1-\alpha_j}(v_j). \end{cases}$$

$$\text{where } \nu_j = (g - r + j)j$$

As in Rault (2000), each statistic is a likelihood ratio test :

$$\xi_j = -2lnQ(H_j/H_1) = T\left[\sum_{i=1}^{j} ln(1-\widehat{\rho}_i) + \sum_{i=1}^{r-j} ln(1-\widehat{\lambda}_i) - \sum_{i=1}^{r} ln(1-\widetilde{\lambda}_i)\right] \tag{3}$$

which is asymptotically distributed under $H_{0,j}$ as a χ^2_{vj} with $\nu_j = (g-r+j)j$ degrees of freedom. H_1 corresponds to the cointegrating hypothesis $\Pi = \alpha\beta^{'}$, $\widetilde{\lambda}_i$ denotes the eigenvalues of the unrestricted $VECM$ and $\widehat{\rho}_i$, $\widehat{\lambda}_i$ correspond to the eigenvalues associated respectively to the j restricted and the $r-j$ unrestricted vectors of the adjustment space.

IV.2 Weak exogeneity testing

Having determined the m_1 rank of the α_Y sub-matrix, the weak exogeneity hypothesis implies the following parametric restrictions :

$$H_{0,we} : \alpha_{Z1} = 0.$$

As these restrictions only correspond to coefficient nullities in the marginal model several conventional tests can be carried out (Likelihood Ratio test, Lagrange Multiplier (LM) test, Wald test). Such tests can easily be implemented in empirical applications using most statistical computer packages. Note that the LR test is generally preferable to the Wald and LM tests in this situation as the restrictions are nonlinear in Π, even if they are linear in α. The LR test is at least invariant to how those restrictions on Π are expressed.

IV.3 The Monte Carlo design and results

We now report some Monte Carlo replications and analyses the size distortions and power of the sequential procedure of rank tests introduced above. Artificial data were generated from five data generation processes (DGPs) depicted in Table 1 (cf. Appendix 1), each containing 11 variables ($g = 5$, $k = 6$), integrated of order one, cointegrated of order 4, expressed in $VECM$ forms. They only differ from the others by the rank m_1 of the α_Y matrix, which varies from 0 to 4, and have no short run dynamics. For each Monte Carlo simulation, we generated 10000 series of length $T + 100 + p$, where p denotes the lag length in the estimated $VECM$. We discarded the first 100 observations to eliminate startup effects. The vector of innovations ε_t was a gaussian eleven dimensional white noise, with zero mean and covariance matrix I_{11}. The initial values ($t = 0$) have been set to zero for all variables in the model, that is $X_0 = 0_{11}$, and $X_1 = e_1 \tilde{} N(0_{11}, I_{11})$. All simulations were carried out on a 266 Pentium II, using the matrix programming language GAUSS, the ε_t were generated by the function "RNDN" and the nominal level of all tests was 5%. Some routines are partly adapted from Sam Ouliaris's COINT GAUSS program. For each DGP, five sample sizes were included; $T \in \{50, 100, 200, 500, 1000\}$, and the adjusted LR tests statistics was used for $T \leq 100$. In each replication, the dimension of the cointegrating rank r is (in a first step) treated as known, so that we can assess our sequential test procedure itself and not other factors affecting its performance[9]. The tabulated results of the experiments are reported respectively in Tables 2 and 3 (cf. Appendix 1). These two tables contain the estimated empirical size and power of the $H_{o,j}$ null hypothesis tests, and the global sample empirical size of the sequential test procedure.

[9] What motivates the detailed discussion of the case where the cointegrating rank (r) has correctly been estimated using conventional likelihood ratio tests is that the performance of our sequential procedure remains almost unchanged when r is over-estimated (cf. supra).

We now discuss the performance of our sequential test procedure. Note that actually the results are very close to those reported in Rault (2000) for the sequential procedure of rank tests related to non-causality testing[10]. These results can be summarized as follows : all $H_{0,j}$ null hypothesis tests ($j = 1, .., 4$) suffer from size distortion in small samples ($T = 50, 100$). As the sample size increases they approximate quite well the correct size. It must be underlined that the asymptotic is reached all the later as the number of tested restrictions is important : for the least restricted null hypothesis ($H_{0,1}$), the empirical size is close to the nominal size of the tests for samples of size larger or equal to 500 (5.09 % for $T = 500$), whereas for the most restricted null hypothesis ($H_{0,4}$) the empirical size is still of 5.19 % for samples of size 1000. Furthermore, the percentage of null hypothesis rejection when they are not true goes to 100 % for all sample size considered in the experiment, indicating both finite distance and asymptotic power equal to one.

As far as the sequential test procedure is concerned, the multiplicity of tests lead to a global size problem in for small samples ($T = 50, 100$), since the sequential procedure estimated global size turns out to be highly dependent on the number of tests necessary to conclude (respectively 6.23 %, 7.14 %, 9.74 %, 11.2 % for $m_1 = 3, ..., 1$ and $T = 100$). On the contrary for large samples ($T = 200, 500$ or 1000), the estimated global size doesn't seem to vary a lot, indicating that the test procedure doesn't suffer from size distortion in large samples : for any possible m_1 true rank, the estimated global size is always very close to 5 % (respectively 5.05 %, 5.20 %, 5.53 %, 5.66 % for $m_1 = 1, ..., 4$ and $T = 500$). This result is due to the fact that the $H_{0,j}$ null hypothesis tests ($j = 1, .., 4$) are extremely powerful and never reject any null hypothesis $H_{0,j}$ when it's true. Note that a "success" is obtained by both rejecting and not rejecting certain hypotheses in combination, and therefore the probability of success is controlled by adjusting the size of the tests to match the available power. It means in other terms that if we need to perform up to j tests to determine the m_1 rank of the α_Y matrix, the global size of the sequential test procedure is simply given in large samples by $\alpha = 1 - (1 - \alpha_j)$. Note that this result is not true for finite samples.

As in most practical applications it is inappropriate to assume that the cointegrating rank (r) is *a priori* known, we finally conducted additional simulations in the case r is unknown and determined using Johansen's trace test (which had not been considered in Rault, 2000). The results of the simulation experiments reported in Table 4 show that restricting the cointegrating rank has little impact on the performance of the sequential test procedure, as least as long as we do not restrict it to be less than the true rank. More precisely, if r is over-estimated the sequential procedure estimated size is very close to the case where the cointegrating rank

[10] In Rault (2000), it is the rank of a sub-matrix extracted from the β matrix which is investigated and not an α sub-matrix rank, as it is the case here.

is correctly specified. This finding should not surprise us since, if one supposes for instance that $r = 5$ instead of $r = 4$, it is then possible to produce by linear combination a column of zeros in the β and α matrices, which only adds a supplementary step in the sequential procedure of rank tests, but doesn't alter its performance since the $H_{0,j}$ null hypothesis tests are very powerful. However the performance of the sequential test procedure is severely distorted by underestimating the cointegrating rank. A similar result concerning the effectiveness of restriction testing on long-run parameters in the Johansen's framework has also been obtained by Greenslade *et al* (1999) when r is underestimated. This is a useful and significant result for the practitioner as it suggests that the sequential procedure may be conducted under the assumption of full rank of the Π matrix without affecting its performance.

V Concluding remarks

In this paper we have provided a necessary and sufficient condition for weak exogeneity in a VECM model. This condition has been given in the setting of a canonical decomposition of the Π matrix and requires the determination of a specific sub-matrix rank, which can easily be done for the practitioner using a simple sequential test procedure based on asymptotically χ^2 statistics, whose properties have been analyzed with Monte-Carlo experiments.

Our Monte-Carlo exercises have shown that the performance of the sequential test procedure is heavily dependent on the choice of the rank of the cointegrating matrix (Π). Indeed, provided this rank is correctly selected or under-estimated, sequential testing to determine the "true" α_Y rank has asymptotically a frequency of success comparable to linear restriction testing on cointegrating parameters by usual Johansen's tests (1991). By contrast, the performance of the sequential procedure is distorted by under-estimating the cointegrating rank and performs poorly with respect to size distortion, whatever the size of the sample is.

Our conclusions therefore are to recommend to investigate the α_Y matrix rank under the assumption of full rank of the cointegrating matrix since Monte Carlo simulations have shown that in small samples of the sort typically used by the applied researcher (about 100 quarterly observations say), there is in this case very much prospect of successfully detecting the true α_Y matrix rank. More precisely, our guideline for the practitioner is : (i) apply the standard Johansen tests for detecting the number of cointegrating vectors in the full system, (ii) investigate the rank of the α_Y matrix using our sequential test procedure in the way advocated above, (iii) decide on the endogeneity and weak exogeneity status of the variables keeping in mind that weak exogeneity is not invariant to the marginalisation of the model. Indeed, it is not an absolute property of a variable, rather a property of a particular model.

References

[1] Anderson T.W. [1951], "Estimating linear restrictions on regression coefficients for multivariate normal distributions", *Annals of Mathematical Statistics*, n° 85, pp 327-351.

[2] Barndorff and Nielsen [1978], "Information and exponential families in statistical theory", *John Wiley, Chichester.*

[3] Bauwens L., Lubrano M. [1994], "Identification restrictions and posterior densities in cointegrated gaussian VAR systems", *discussion paper*, Université Catholique de Louvain.

[4] Boswijk P. [1995], "Efficient inference on cointegration parameters in structural error correction models", *Journal of Econometrics*, vol 69, P 133-158.

[5] Engle R.F., Hendry D.F. and Richard J.F. [1983]," Exogeneity", *Econometrica*, vol 51, pp 277-304.

[6] Engle R.F. and Granger C.W.J. [1987], "Cointegration and Error-Correction : Representation, Estimation and Testing", *Econometrica*, vol 55, pp 251-76.

[7] Ericsson N.R. [1995], "Conditional and structural error correction models", *Journal of Econometrics*, vol 69, pp 159-171.

[8] Ericsson N.R., Hendry D.F., Mizon G.E. [1998], " Exogenity, cointegration, and economic policy analysis : an overview", *Journal of Business and Economic Statistics*, vol 6, n°4.

[9] Florens J.P and Mouchart M. [1980], "Initial and sequential reduction of bayesian experiments", *CORE Discussion Paper* n° 8015, Université Catholique de Louvain.

[10] Greenslade J.F, Hall S.G., Henry S.G.B. [1999], "On the identification of cointegrated systems in small samples : practical procedures with an application to UK wages and prices", *LBS Working paper.*

[11] Hecq A., Palm F.C, Urbain J.P [2000], "Separation, Weak Exogeneity and P-T Decomposition in Cointegrated VAR Systems with Common Features", *Econometric Reviews (forthcoming).*

[12] Hendry D.F and Mizon G.E [1993] ,"Evaluating dynamic models by encompassing the VAR", in P.C.B Phillips, ed., *Models, methods, and applications of econometrics* (Basil, Blackwell, Oxford), pp 272-300.

[13] Johansen S. [1988], "Statistical analysis of cointegration vectors", *Journal of Economic Dynamics and Control*, vol 12, pp 231-254.

[14] Johansen S. [1992] , "Cointegration in partial systems and the efficiency of single-equation analysis", *Journal of Econometrics,* vol 52, pp 389-402.

[15] Johansen S. [1995], "*Likelihood-based inference in co-integrated vector autoregressive models*", Oxford University Press, 267 p.

[16] Rault C. [2000], "Non-causality in VAR-ECM models with purely exogenous long run paths", *Economics Letters*, n°67-2.

[17] Rault C. and Pradel J. [2003], "Exogeneity in VAR-ECM models with purely exogenous long-run paths", forthcoming in *Oxford Bulletin of Economics and Statistics* in December.

[18] Richard J.F. [1980], "Models with several regimes and changes in exogeneity", *Review of Economic Studies*, vol 7, pp 1-20.

[19] Sims, C. A. [1980], "Macroeconomics and reality" *Econometrica,* vo 48, pp 1-48.

[20] Urbain J.P. [1992], "On weak exogeneity in error correction models", *Oxford Bulletin of Economics and Statistics*, vol 54, n° 2, pp 187-207.

[21] Urbain J.P. [1995], "Partial versus full system modelling of cointegrated systems : an empirical illustration", *Journal of Econometrics*, vol 69, P 177-210.

Appendix 1 : Simulation results

Table 1: Data Generation Processes (DGP) (n = 11, g = 5, k = 6)[1]

DGP (1) : $m_1 = 4$	DGP (2) : $m_1 = 3$	DGP (3) : $m_1 = 2$	DGP (4) : $m_1 = 1$	DGP (5) : $m_1 = 0$
Beta	Beta	Beta	Beta	b Beta
0.10 1.70 0.60 -0.10	0.10 1.70 0.60 -0.10	0.10 1.70 0.60 -0.10	0.10 1.70 0.60 -0.10	0.10 1.70 0.60 -0.10
-0.10 -2.00 -0.40 -0.40	-0.10 -2.00 -0.40 -0.40	-0.10 -2.00 -0.40 -0.40	-0.10 -2.00 -0.40 -0.40	-0.10 -2.00 -0.40 -0.40
-0.30 0.50 -0.20 0.50	-0.30 0.50 -0.20 0.50	-0.30 0.50 -0.20 0.50	-0.30 0.50 -0.20 0.50	-0.30 0.50 -0.20 0.50
-1.00 0.10 0.20 0.20	-1.00 0.10 0.20 0.20	-1.00 0.10 0.20 0.20	-1.00 0.10 0.20 0.20	-1.00 0.10 0.20 0.20
0.10 -1.00 -0.80 0.30	0.10 -1.00 -0.80 0.30	0.10 -1.00 -0.80 0.30	0.10 -1.00 -0.80 0.30	0.10 -1.00 -0.80 0.30
0.20 -0.10 -0.20 0.10	0.20 -0.10 -0.20 0.10	0.20 -0.10 -0.20 0.10	0.20 -0.10 -0.20 0.10	0.20 -0.10 -0.20 0.10
0.10 0.20 0.10 0.20	0.10 0.20 0.10 0.20	0.10 0.20 0.10 0.20	0.10 0.20 0.10 0.20	0.10 0.20 0.10 0.20
0.10 -0.20 -0.30 -0.20	0.10 -0.20 -0.30 -0.20	0.10 -0.20 -0.30 -0.20	0.10 -0.20 -0.30 -0.20	0.10 -0.20 -0.30 -0.20
0.20 -0.10 0.20 -0.10	0.20 -0.10 0.20 -0.10	0.20 -0.10 0.20 -0.10	0.20 -0.10 0.20 -0.10	0.20 -0.10 0.20 -0.10
0.60 0.50 0.30 0.00	0.60 0.50 0.30 0.00	0.60 0.50 0.30 0.00	0.60 0.50 0.30 0.00	0.60 0.50 0.30 0.00
0.10 -0.30 -0.30 0.00	0.10 -0.30 -0.30 0.00	0.10 -0.30 -0.30 0.00	0.10 -0.30 -0.30 0.00	0.10 -0.30 -0.30 0.00
alpha	alpha	alpha	alpha	alpha
-0.50 -0.30 -0.40 0.30	-0.50 -0.30 -0.40 0.00	-0.50 -0.30 -0.60 0.00	-0.90 -0.30 -0.60 0.00	0.00 0.00 0.00 0.00
0.30 0.20 0.60 0.50	0.30 0.20 0.60 0.00	0.30 0.20 0.40 0.00	0.60 0.20 0.40 0.00	0.00 0.00 0.00 0.00
-0.20 -0.20 -0.20 0.00	-0.20 -0.20 -0.20 0.00	-0.20 -0.20 -0.40 0.00	-0.60 -0.20 -0.40 0.00	0.00 0.00 0.00 0.00
0.70 0.10 0.50 0.10	0.70 0.10 0.50 0.00	0.70 0.10 0.20 0.00	0.30 0.10 0.20 0.00	0.00 0.00 0.00 0.00
-0.90 -0.50 -1.10 0.00	-0.90 -0.50 -1.10 0.00	-0.90 -0.50 -1.00 0.00	-1.50 -0.50 -1.00 0.00	0.00 0.00 0.00 0.00
-1.00 0.00 -0.20 0.10	-1.00 0.00 -0.20 0.10	-1.00 0.00 -0.20 0.10	-1.00 0.00 -0.20 0.10	-1.00 0.00 -0.20 0.10
-0.20 0.40 0.00 0.10	-0.20 0.40 0.00 0.10	-0.20 0.40 0.00 0.10	-0.20 0.40 0.00 0.10	-0.20 0.40 0.00 0.10
-0.50 -0.50 0.60 0.00	-0.50 -0.50 0.60 0.00	-0.50 -0.50 0.60 0.00	-0.50 -0.50 0.60 0.00	-0.50 -0.50 0.60 0.00
0.10 0.50 0.10 -0.20	0.10 0.50 0.10 -0.20	0.10 0.50 0.10 -0.20	0.10 0.50 0.10 -0.20	0.10 0.50 0.10 -0.20
0.10 0.30 -0.30 0.20	0.10 0.30 -0.30 0.20	0.10 0.30 -0.30 0.20	0.10 0.30 -0.30 0.20	0.10 0.30 -0.30 0.20
-0.90 0.40 0.30 -0.50	-0.90 0.40 0.30 -0.50	-0.90 0.40 0.30 0.50	-0.90 0.40 0.30 -0.50.	-0.20 0.50 0.30 0.30

Table 2: Empirical size and power of the Ho,j null hypothesis tests (j = 1,..,4) (rejection per 100), with 10000 replications at the 5 % nominal level of significance [2]

DGPS	DGP (1) : $m_1 = 4$					DGP (2) : $m_1 = 3$					DGP (3) : $m_1 = 2$					DGP (4) : $m_1 = 1$					DGP (5) : $m_1 = 0$					Hypothesis tested
Sample size T	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	
α_1 $W_1 = \Psi_1 > A_1$ [3]	100	100	100	100	100	**7.87**	**6.35**	**5.23**	**5.09**	**5.03**	0.00	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.20	0.00	0.00	0.00	0.00	$H_{0,1:}$ {rang $(\alpha_Y) \leq 3$} contre{rang $(\alpha_Y) = 4$}
α_2 $W_2 = \Psi_2 > A_2$	100	100	100	100	100	100	100	100	100	100	**12.4**	**7.22**	**6.12**	**5.27**	**5.11**	0.21	0.00	0.00	0.00	0.00	0.42	0.00	0.00	0.00	0.00	$H_{0,2:}$ {rang $(\alpha_Y) \leq 2$} contre{rang $(\alpha_Y) \geq 3$}
α_3 $W_3 = \Psi_3 > A_3$	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	**17.5**	**9.57**	**7.02**	**5.58**	**5.11**	1.30	1.05	0.40	0.00	0.00	$H_{0,3:}$ {rang $(\alpha_Y) \leq 1$} contre{rang $(\alpha_Y) \geq 2$}
α_4 $W_4 = \Psi_4 > A_4$	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	100	**20.4**	**9.86**	**7.12**	**5.72**	**5.19**	$H_{0,4:}$ {rang $(\alpha_Y) = 0$} contre{rang $(\alpha_Y) \geq 1$}

Table3: Empirical size of the sequential test procedure (rejection per 100), with 10000 replications at the 5 % nominal level of significance in the case where the cointegating rank (ie. r = 4) is known

DGPS	DGP (2) : $m_1 = 3$					DGP (3) : $m_1 = 2$					DGP (4) : $m_1 = 1$					DGP (5) : $m_1 = 0$				
	$P(W_1)$					$P(\overline{W_1}\ W_2)$ [4]					$P(\overline{W_1}\ \overline{W_2}\ W_3)$					$P(\overline{W_1}\ \overline{W_2}\ \overline{W_3}\ W_4)$				
Sample size T	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000
m_1 estimated = m_1	7.44	6.23	5.19	5.05	5.01	12.8	7.14	5.99	5.20	5.05	18.6	9.74	6.91	5.53	5.13	20.9	11.2	7.09	5.66	5.18

---

[1] DGP (1), (2) and (5) can easily be seen to be respectively of rank $m_1 = 4, 3, 0$. However the fact that DGP (3) and (4) are of rank $m_1 = 2$ and $m_1=1$ is less straightforward : it requires noticing that the α_Y columns of these two DGPs are not linearly independent since they are respectively linked by $C_3 = 2 C_2$, for DGP (3) and by $C_3 = 2 C_2$, $C_1 = C_2 + C_3$ for DGP (4).

[2] The adjusted version of the test statistic was used for T = 50, 100.

[3] Ai, i = 1,..,4 denotes the critical value from the χ^2 distribution at the 5 % level of significance.

[4] $P(\overline{W_1}\ W_2)$ represents the probability to be at the same time in the acceptance region $\overline{W_1}$ of test 1 and in the critical region W_2 of test 2.

Appendix 1 : Simulation results

Table 4. Empirical size of the sequential test procedure (rejection per 100), with 10000 replications at the 5 % nominal level of significance in the case where the cointegating rank (ie. r = 4) is not correctly selected

DGPS		DGP (2) : $m_l = 3$ $P(W_1)$					DGP (3) : $m_l = 2$ $P(\overline{W}_1 \; W_2)$[5]					DGP (4) : $m_l = 1$ $P(\overline{W}_1 \; \overline{W}_2 \; W_3)$					DGP (5) : $m_l = 0$ $P(\overline{W}_1 \; \overline{W}_2 \; \overline{W}_3 \; W_4)$				
Sample size T		50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000	50	100	200	500	1000
r=2	m_l estimated = m_l	24.7	21.8	19.6	16.3	13.8	32.2	24.3	20.8	18.8	15.5	39.2	27.6	22.4	20.0	16.3	43.2	30.4	25.3	22.1	18.1
r=3	m_l estimated = m_l	15.7	13.94	12.4	11.4	10.1	22.1	15.9	13.9	12.4	11.7	28.9	18.5	14.8	13.9	12.1	31.1	20.4	16.1	14.1	12.4
r=5	m_l estimated = m_l	7.95	6.63	5.50	5.26	5.12	13.5	7.75	6.40	5.51	5.24	19.2	10.1	7.31	5.78	5.28	22.1	12.0	7.78	6.11	5.32
r= 6	m_l estimated = m_l	8.26	6. 77	5.64	5.42	5.26	14.0	8.12	6.65	5.71	5.36	19.7	10.5	7.66	5.99	5.42	22.9	12.7	8.37	6.52	5.60

[5] $P(\overline{W}_1 \; W_2)$ represents the probability to be at the same time in the acceptance region $\overline{W}_1$ of test 1 and in the critical region W_2 of test 2.