Further Results on Weak Exogeneity in Vector Error Correction Models*

Christophe Rault**

Abstract

This paper extends the result for non-causality and strong exogeneity of Pradel and Rault and Pradel (2003) Exogeneity in VAR-ECM models with purely exogenous long-run paths, *Oxford Bulletin of Economics and Statistics* to weak exogeneity. More precisely, it 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 be easily calculated with usual statistical computer packages, which makes our approach fully operational empirically.

Keywords: VAR-ECM, Canonical Representation, Weak Exogeneity.

JEL Codes: C15, C22.

*Submitted in August 2004. Revised in July 2005. I am very grateful to two anonymous referees for helpful remarks and suggestions on an earlier draft of this paper. The usual disclaimer applies.

**EPEE, University of Evry-Val d'Essonne, Département d'économie, Boulevard François Mitterrand, 91025 Evry Cedex, France, and IZA. E-mail: chrault@hotmail.com, site: <http://www.multimania.com/chrault/index.html>.

1. Introduction

In 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 et al. (1998); Hecq et al. (2002); Hendry and Mizon (1993); Johansen (1992, 1995); Urbain (1992, 1995); Rault and Pradel (2003)). Weak exogeneity has also been extensively discussed in two special issues of the *Journal of Policy Modeling* (1992), vol 14, issue 3 and of the *Journal of Business and Economic Statistics* (1998), vol 6, issue 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. In addition, Johansen makes the assumption that macroeconomists have a potential economic interest in all cointegrating relations existing between the variables being 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 an r_1 subset of the cointegrating relationships as parameters of 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 last r_2 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 macroeconomists. 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 in this paper, we propose an extension of the existing weak exogeneity conditions, which is based on a canonical decomposition of the

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

long-run matrix Π . This representation exploits the fact that the β cointegrating and α loading factor matrices are not unique insofar as $\Pi = \alpha\beta' = (\alpha\Psi^{-1})(\Psi\beta')$ 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 2 sets out the general *VECM* framework. Section 3 introduces the canonical representation of the long-run matrix Π and proposes a necessary and sufficient condition for weak exogeneity. Section 4 deals with inference and testing which are conducted within the setting proposed by Johansen. Finally, concluding remarks are presented in section 5 and specific recommendations are provided for applied researchers.

2. Cointegrated Vector Autoregressions

We begin by setting out the basic framework and thus consider, as in Rault and Pradel (2003), an n -dimensional *VECM*(p) process $\{X_t\}$, generated by

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

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

²Note that an $I(0)$ process might be defined as a process with the infinite sum of the autocovariances being finite. Alternatively, it may be defined in the frequency domain as a process with spectral density function that is positive and finite at the zero frequency. An $I(1)$ process is then defined as a process that requires first differences to achieve $I(0)$.

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 subvectors 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 :

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

$$\text{with } \left\{ \begin{array}{l} \Gamma_{YY}^+(L) = \Gamma_{YY}(L) - \Sigma_{YZ} \Sigma_{ZZ}^{-1} \Gamma_{ZY}(L) = I_g - \sum_{i=1}^{p-1} \Gamma_{YY,i}^+ L^i \\ \Gamma_{YZ}^+(L) = \Gamma_{YZ}(L) - \Sigma_{YZ} \Sigma_{ZZ}^{-1} \Gamma_{ZZ}(L) = - \sum_{i=0}^{p-1} \Gamma_{YZ,i}^+ L^i \\ \alpha_Y^+ = \alpha_Y - \Sigma_{YZ} \Sigma_{ZZ}^{-1} \alpha_Z \\ \eta_{Y,t} = \varepsilon_{Y,t} - \Sigma_{YZ} \Sigma_{ZZ}^{-1} \varepsilon_{Z,t} \\ \Sigma_{YY}^+ = \Sigma_{YY} - \Sigma_{YZ} \Sigma_{ZZ}^{-1} \Sigma_{ZY} \end{array} \right.$$

where L denotes the lag operator and $\begin{pmatrix} \eta_{Y,t} \\ \varepsilon_{Z,t} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Sigma_{YY}^+ & 0 \\ 0 & \Sigma_{ZZ} \end{pmatrix} \right]$, with the partitioning of the matrices Γ_i, α and β being conformable to that of $\bar{X}_t$.

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.

3. A Necessary and Sufficient Condition for Weak Exogeneity in *VECM* Models

It is now well-admitted that the presence or lack of weak exogeneity³ depends crucially on what parameters the focus of attention is, but contrary to what 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

³Let'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 parameter spaces (cf. Florens and Mouchat (1980)).

in Johansen (1992), theorem 1),⁴ or even in a structural partition of the cointegrating vectors made *a priori* (as 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, propensity to consume or save, etc. Indeed, when economists undertake practical modeling they are not equally interested in describing the behavior of all the variables of the system. 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. For instance, when modeling wage and price determination, economists often include unemployment because they believe that unemployment affects wages, but they don't include the many variables which they think might explain unemployment. It would therefore be quite surprising if a cointegrated vector existed which would explain unemployment. Even though in the real world we have no doubt that wages affect unemployment, in such models unemployment can often be treated as weakly exogenous for the parameters of the subsystem composed by wage and price equations.⁵

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 difficulty lies in how to economically interpret 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 NSC condition for weak exogeneity we need to consider the following preliminary theorem:

Theorem 3.1 *Let $\Pi = \alpha\beta'$ be an $n \times n$ reduced rank matrix of rank r ($0 < r < n$) and partition α into $\begin{bmatrix} \alpha_Y \\ \alpha_Z \end{bmatrix}$.*

⁴In 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".

⁵Note also that the marginal distribution of policy variables may be hard to model empirically, because of changes in the policy regime. In such cases, valid conditioning on those variables can considerably assist empirical modeling for the estimation of parameters of interest.

(i) If we define $m_1 = \text{rank}(\alpha_Y)$ with $m_1 > 0$ and $r - m_1 > 0$,⁶ then the α and β matrices can always be reparameterized as follows:

$$\begin{aligned}\beta &= [\beta_1 \ \beta_2] = \begin{bmatrix} \beta_{Y1} & \beta_{Y2} \\ \beta_{Z1} & \beta_{Z2} \end{bmatrix} \\ \alpha &= [\alpha_1 \ \alpha_2] = \begin{bmatrix} \alpha_{Y1} & 0_{(g, r-m_1)} \\ \alpha_{Z1} & \alpha_{Z2} \end{bmatrix}\end{aligned}$$

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 $\text{rank}(\alpha_{Y1}) = m_1$ and $\text{rank}(\alpha_{Z2}) = r - m_1$.

(ii) m_1 is uniquely defined and is invariant to the chosen reparameterization. It is such as⁷

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

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 $\alpha_Z \beta' = \alpha_Z (\eta (\eta' \eta)^{-1} \eta' + \eta_\perp \eta'_\perp (\eta'_\perp \eta_\perp)^{-1} \eta'_\perp) \beta' = \alpha_{Z1} \beta'_1 + \alpha_{Z2} \beta'_2 = (\alpha_{Z1}, \alpha_{Z2})(\beta_1, \beta_2)'$, which shows that theorem 1 is satisfied. ■

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

$$\begin{cases} \text{conditional model} \\ \Delta Y_t = \sum_{i=1}^{p-1} \Gamma_{Y,i}^+ \Delta Y_{t-i} + \sum_{i=0}^{p-1} \Gamma_{Y,i}^+ \Delta Z_{t-i} + \alpha_{Y1}^+ \beta'_1 X_{t-1} + \eta_{Y,t} \\ \text{marginal model} \\ \Delta Z_t = \sum_{i=1}^{p-1} \Gamma_{Z,i} \Delta Y_{t-i} + \sum_{i=1}^{p-1} \Gamma_{Z,i} \Delta Z_{t-i} + \alpha_{Z1} \beta'_1 X_{t-1} + \alpha_{Z2} \beta'_2 X_{t-1} + \varepsilon_{Z,t} \end{cases} \quad (2.a)$$

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 (Bauwens and Lubrano, 1996), since for any non-singular matrix Ψ of dimensions (r, r) , we could define $\Pi = (\alpha \Psi^{-1}) (\Psi \beta')$, and $\alpha^* = \alpha \Psi^{-1}$, $\beta^* = \beta \Psi'$ would be

⁶We assume that β_1 and β_2 each contain at least one cointegrating vector to exclude the case where $\beta_1 = \beta$, which entails that β_2 is a null set.

⁷This condition is derived from $\text{rank}(\alpha) = r$.

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 ⁸ and reparameterized into $[\alpha_{Y1} \ 0_{(g,r-m_1)}]$.

We are in a position to state the following result:

Proposition 3.1 *Necessary and sufficient weak exogeneity condition. Suppose that the investigator's parameters of interest are those of the conditional model, i.e. $\Psi = (\Gamma_{YY,i}^+, i = 1, \dots, p-1; \Gamma_{YZ,i}^+, i = 0, \dots, p-1; \alpha_{Y1}^+; \beta_1')$, then Z_t is weakly exogenous for Ψ if and only if $\alpha_{Z1} = 0$ in the canonical representation given by theorem 3.1.*

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 the 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 explicitly determine this partition, exploiting the fact that the α and β are not unique (cf. *infra*). But the gain in doing this is that we are then able to provide a necessary and sufficient condition for weak exogeneity 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 obviously makes 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.

4. Inference and Testing

The necessary and sufficient condition for weak exogeneity introduced in proposition 3.1 first requires that matrix Π be rewritten under the canonical decomposition given in theorem 3.1. Then, in this framework, this condition has been expressed in terms of coefficient nullities of the α matrix, which permits the use of conventional chi-squared statistics (Johansen, 1995).

⁸The way this rank can be determined in applied studies is discussed in section 4.

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

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

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

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

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

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

$$\xi_j = -2 \ln Q(H_j/H_1) = T \left[\sum_{i=1}^j \ln(1 - \hat{\rho}_i) + \sum_{i=1}^{r-j} \ln(1 - \hat{\lambda}_i) - \sum_{i=1}^r \ln(1 - \tilde{\lambda}_i) \right] \quad (3)$$

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

Having determined the m_1 rank of the α_Y 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.

As in most practical applications it is inappropriate to assume that the cointegrating rank (r) is *a priori* known, we finally conducted simulations in 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 the Appendix show that restricting the cointegrating rank has little impact on the performance of the sequential test procedure presented above, at least as long as we do not restrict it to be less than the true rank. More precisely, if r is overestimated the estimated size of the sequential procedure is very close to the case where the cointegrating rank is correctly specified (i.e. the sequential procedure does not suffer from size distortion in large samples, cf. Rault (2000)). 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 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.

5. 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 submatrix 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 underestimated, 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 underestimating the cointegrating rank and performs poorly with respect to size distortion, whatever the size of the sample.

In conclusion, we therefore recommend investigating 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), there is, in this case, a high probability of successfully detecting the true α_Y matrix rank. More precisely, our guidelines for the practitioner are: (i) to apply the standard Johansen's 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 marginalization of the model. Indeed, it is not an absolute property of a variable, rather a property of a particular model.

References

- Bauwens, L. & Lubrano, M. (1996). Identification restrictions and posterior densities in cointegrated Gaussian VAR systems. *Advances in Econometrics*, II:3–28. Part B.
- Boswijk, P. (1995). Efficient inference on cointegration parameters in structural error correction models. *Journal of Econometrics*, 69:133–158.
- Engle, R. F. & Granger, C. W. J. (1987). Cointegration and error-correction: Representation, estimation and testing. *Econometrica*, 55:251–276.
- Engle, R. F., Hendry, D. F., & Richard, J. F. (1983). Exogeneity. *Econometrica*, 51:277–304.
- Ericsson, N. R. (1995). Conditional and structural error correction models. *Journal of Econometrics*, 69:159–171.
- Ericsson, N. R., Hendry, D. F., & Mizon, G. E. (1998). Exogeneity, cointegration and economic policy analysis: An overview. *Journal of Business and Economic Statistics*, 6(4).
- Florens, J. P. & Mouchat, M. (1980). Initial and sequential reduction of Bayesian experiments. CORE Discussion Paper no. 8015, Université Catholique de Louvain.
- 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. *Computing in Economics and Finance*, 643. Society for Computational Economics.

- Hecq, A., Palm, F. C., & Urbain, J. P. (2002). Separation weak exogeneity and P-T decomposition in cointegrated VAR systems with common features. *Econometric Reviews*, 21:273–307.
- Hendry, D. F. & Mizon, G. E. (1993). Evaluating dynamic models by encompassing the VAR. In Phillips, P. C. B., editor, *Models, Methods and Applications of Econometrics*, pages 272–300. Basil, Blackwell, Oxford.
- Johansen, S. (1991). Estimation and hypothesis testing of co-integration vectors in Gaussian vectors autoregressive models. *Econometrica*, 6:1551–1580.
- Johansen, S. (1992). Cointegration in partial systems and the efficiency of single-equation analysis. *Journal of Econometrics*, 52:389–402.
- Johansen, S. (1995). *Likelihood-Based Inference in Co-Integrated Vector Autoregressive Models*. Oxford University Press. 267 p.
- Rault, C. (2000). Non-causality in VAR-ECM models with purely exogenous long run paths. *Economics Letters*, 67(2).
- Rault, C. & Pradel, J. (2003). Exogeneity in VAR-ECM models with purely exogenous long-run paths. *Oxford Bulletin of Economics and Statistics*, 66(1):629–653.
- Richard, J. F. (1980). Models with several regimes and changes in exogeneity. *Review of Economic Studies*, 7:1–20.
- Urbain, J. P. (1992). On weak exogeneity in error correction models. *Oxford Bulletin of Economics and Statistics*, 54(2):187–207.
- Urbain, J. P. (1995). Partial versus full system modelling of cointegrated systems: An empirical illustration. *Journal of Econometrics*, 69:177–210.

Appendix

Simulation Results

Table A.1
Data generation processes (*DGP*) ($n = 11, g = 5, k = 6$)

<i>DGP</i> (1) : $m_1 = 4$				<i>DGP</i> (2) : $m_1 = 3$				<i>DGP</i> (3) : $m_1 = 2$			
Beta				Beta				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	-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
-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
0.20	-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	-0.30	-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
0.60	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
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
0.30	0.20	0.60	0.50	0.30	0.20	0.60	0.00	0.30	0.20	0.40	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
0.70	0.10	0.50	0.10	0.70	0.10	0.50	0.00	0.70	0.10	0.20	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
-1.00	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
-0.50	-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	0.30	-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
<i>DGP</i> (4) : $m_1 = 1$				<i>DGP</i> (5) : $m_1 = 0$							
Beta				Beta							
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				
-0.30	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				
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				
0.10	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				
0.20	-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				
0.10	-0.30	-0.30	0.00	0.10	-0.30	-0.30	0.00				
Alpha				Alpha							
-0.90	-0.30	-0.60	0.00	0.00	0.00	0.00	0.00				
0.60	0.20	0.40	0.00	0.00	0.00	0.00	0.00				
-0.60	-0.20	-0.40	0.00	0.00	0.00	0.00	0.00				
0.30	0.10	0.20	0.00	0.00	0.00	0.00	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				
-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				
0.10	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				
-0.90	0.40	0.30	-0.50	-0.20	0.50	0.30	0.30				

Notes: *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 Y columns of these two *DGP*s are not linearly independent since they are respectively linked by $C_3 = 2C_2$, for *DGP* (3) and by $C_3 = 2C_2$, $C_1 = C_2 + C_3$ for *DGP*(4).

Table A.2

Empirical size and power of the H_{0j} null hypothesis tests ($j = 1, \dots, 4$) (rejection per 100), with 10,000 replications at the 5% nominal level of significance¹

Hypothesis tested	$H_{0,1} : \{\text{rank}(\alpha_Y) \leq 3\}$ versus $\{\text{rank}(\alpha_Y) = 4\}$	$H_{0,1} : \{\text{rank}(\alpha_Y) \leq 3\}$ versus $\{\text{rank}(\alpha_Y) = 4\}$	$H_{0,1} : \{\text{rank}(\alpha_Y) \leq 3\}$ versus $\{\text{rank}(\alpha_Y) = 4\}$	$H_{0,1} : \{\text{rank}(\alpha_Y) \leq 3\}$ versus $\{\text{rank}(\alpha_Y) = 4\}$
Sample size	$\alpha_1 W_1 = \Psi_1 > A_1$ ²	$\alpha_2 W_2 = \Psi_2 > A_2$	$\alpha_3 W_3 = \Psi_3 > A_3$	$\alpha_4 W_4 = \Psi_4 > A_4$
$DGP(1) : m_1 = 4$	50 100 200 500 1000	100 100 100 100 100	100 100 100 100 100	100 100 100 100 100
$DGP(2) : m_1 = 3$	50 100 200 500 1000	7.87 6.35 5.23 5.09 5.03	100 100 100 100 100	100 100 100 100 100
$DGP(3) : m_1 = 2$	50 100 200 500 1000	0.00 0.00 0.00 0.00 0.00	12.4 7.22 6.12 5.27 5.11	100 100 100 100 100
$DGP(4) : m_1 = 1$	50 100 200 500 1000	0.10 0.00 0.00 0.00 0.00	0.21 0.00 0.00 0.00 0.00	17.5 9.57 7.02 5.58 5.11
$DGP(5) : m_1 = 0$	50 100 200 500 1000	0.20 0.00 0.00 0.00 0.00	0.42 0.00 0.00 0.00 0.00	1.30 1.05 0.40 0.00 0.00
				20.4 9.86 7.12 5.72 5.19

Notes: ¹The adjusted version of the test statistic was used for $T = 50, 100$.

² $A_i, i = 1, \dots, 4$ denotes the critical value from the χ^2 distribution at the 5% level of significance.

Table A.3

Empirical size of the sequential test procedure (rejection per 100), with 10,000 replications at the 5% nominal level of significance in the case where the cointegrating rank (that is $r = 4$) is known

$DGP(2) : m_1 = 3, P(W_1)$					
Sample size T	50	100	200	500	1000
m_1 estimated= m_1	7.44	6.23	5.19	5.05	5.01
$DGP(3) : m_1 = 2, P(W_1 W_2)$ ¹					
Sample size T	50	100	200	500	1000
m_1 estimated= m_1	12.8	7.14	5.99	5.2	5.05
$DGP(4) : m_1 = 1, P(W_1 W_2 W_3)$					
Sample size T	50	100	200	500	1000
m_1 estimated= m_1	18.6	9.74	6.91	5.53	5.13
$DGP(5) : m_1 = 0, P(W_1 W_2 W_3 W_4)$					
Sample size T	50	100	200	500	1000
m_1 estimated= m_1	20.9	11.2	7.09	5.66	5.18

Notes: ¹ $P(W_1, W_2)$ represents the probability of being simultaneously in the acceptance region $\bar{W}_1$ of test 1 and in the critical region W_2 of test 2.

Table A.4

Empirical size of the sequential test procedure (rejection per 100), with 10,000 replications at the 5% nominal level of significance in the case where the cointegrating rank (that is $r = 4$) is not correctly selected

		$DGP(2) : m_1 = 3, P(W_1)$				
	Sample size T	50	100	200	500	1000
$r = 2$	m_1 estimated = m_1	24.7	21.8	19.6	16.3	13.8
$r = 3$	m_1 estimated = m_1	15.7	13.94	12.4	11.4	10.1
$r = 5$	m_1 estimated = m_1	7.95	6.63	5.50	5.26	5.12
$r = 6$	m_1 estimated = m_1	8.26	6.77	5.64	5.42	5.26
		$DGP(3) : m_1 = 2, P(W_1W_2)$				
$r = 2$	m_1 estimated = m_1	32.2	24.3	20.8	18.8	15.5
$r = 3$	m_1 estimated = m_1	22.1	15.9	13.9	12.4	11.7
$r = 5$	m_1 estimated = m_1	13.5	7.75	6.40	5.51	5.24
$r = 6$	m_1 estimated = m_1	14.0	8.12	6.65	5.71	5.36
		$DGP(4) : m_1 = 1, P(W_1W_2W_3)$				
$r = 2$	m_1 estimated = m_1	39.2	27.6	22.4	20.0	16.3
$r = 3$	m_1 estimated = m_1	28.9	18.5	14.8	13.9	12.1
$r = 5$	m_1 estimated = m_1	19.2	10.1	7.31	5.78	5.28
$r = 6$	m_1 estimated = m_1	19.7	10.5	7.66	5.99	5.42
		$DGP(5) : m_1 = 0, P(W_1W_2W_3W_4)$				
$r = 2$	m_1 estimated = m_1	43.2	30.4	25.3	22.1	18.1
$r = 3$	m_1 estimated = m_1	31.1	20.4	16.1	14.1	12.4
$r = 5$	m_1 estimated = m_1	22.1	12.0	7.78	6.11	5.32
$r = 6$	m_1 estimated = m_1	22.9	12.7	8.37	6.52	5.60