Exogeneity in Vector Error Correction Models with Purely Exogenous Long-Run Paths

Jacqueline PRADEL & Christophe RAULT

03 – 10
Exogeneity in vector error correction models with purely exogenous long-run paths

Jacqueline PRADEL∗(EUREQua, University of Paris I),
Christophe RAULT†(EPEE, University of Evry and EUREQua)

May 22, 2003

Abstract

Existing exogeneity conditions of literature are only sufficient and imply “overly strong” constraints on long-run parameters. This paper presents some new results on exogeneity in VECM models. A key concept of the analysis is the “purely exogenous long-run path”, i.e. a cointegrating vector only including “exogenous” variables. Extending earlier results of Johansen (1992b) and of Toda and Phillips (1991) among others, we propose a framework based on two canonical representations of the long-run matrix, which can constitute a suitable basis to formulate a necessary and sufficient condition for non-causality as well as a condition for strong exogeneity. An interesting property is that the statistics involved in the sequential procedures for testing these conditions are distributed as $\chi^2$ 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 the sequential test procedures are analyzed with Monte-Carlo experiments.

Keywords : Cointegration, canonical representation, strong exogeneity, non-causality, power, size distortions.

JEL Classification : C15; C22.

Paper number of words : 7200

∗Maison des Sciences de l’Économie, 106-112 bd. de L’Hôpital, 75647 Paris Cedex 13, France.
†Département d’économie, Boulevard François Mitterrand, 91025 Évry Cedex, France. E-mail : chrault@hotmail.com, web-site : http://www.multimania.com/chrault/index.html
1 Introduction

Vector error correction models (VECM) have now become an essential part of time series econometricians’ toolkit since they are based on the desire to develop a macro-econometric model which has transparent theoretical foundations, providing insights into the behavioral relationships which underlie the functioning of macro-economy, and which has flexible dynamics that fit the historical time series data well. As it is now well established two different families of approaches can be distinguished within this framework:

- the full system modelling approach is based on complete VECM in which a set on interesting statistically and economically meaningful hypotheses can be raised (cf. Johansen and Juselius, 1992). The optimality of this approach has been pointed out for instance by Phillips (1991),

- the partial system modelling approach which rests upon the estimation of conditional ECMs in which the generating process of some variables is not modelled (the marginal model). Conditional ECMs are very popular empirically (cf. Davidson, Hendry, Srba and Yeo, 1978; Hendry and Ericsson, 1991; Ericsson and Iron, 1994; Urbain, 1995; Rault et al., 2003), since they have same optimality properties than complete VECM when exogeneity conditions are satisfied (cf. Johansen, 1995). It means in other terms that if exogeneity is valid, neglecting the marginal model is without loss of information and conditional ECMs are invariably much simpler to model than the whole system.

This practical appealing aspect might explain why considerable theoretical interest has been devoted this last decade to the analysis of the exogeneity concept in linear VECM with variables at most integrated of order 1 (see e.g. Giannini and Mosconi, 1992; Hendry and Mizon, 1993; Johansen, 1992a, 1992b, 1995; Urbain, 1992, 2001; Ericsson et al., 1998). Our analysis in this paper is based on two observations on recent theoretical works in VECM.

- The first one is that the usual exogeneity conditions which can be expressed in term of coefficient nullities are only sufficient conditions. They are easily testable but sometimes imply “overly strong” restrictions. The conditions of
Johansen (1992a) and Urbain (1992) for instance, which are widely used in applied works, forbid the existence of long run relationships in the equations of the marginal model. These equations are thus a VAR model in first differences. This implies that under these conditions the cointegration properties of the full VECM are determined solely by the conditional model. Besides Johansen makes the assumption that macro-economists 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 the existence of \( r \) cointegrating vectors, whereas according to economic theory there should only exist \( m \), with \( m < r \). Already in his careful discussion of Boswijk’s paper (1995) on structural ECMs, Ericsson (1995) had noted that this was 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”.

- The second observation is that the necessary and sufficient Granger non-causality condition of Toda and Phillips (1991) which implies non-linear restrictions on long-run parameters appears difficult to test in applied studies. Indeed, as it is now well-known test statistics have only an asymptotic chi-squared distribution under some regularity conditions which turn out in general hard to establish empirically. That’s why another set of conditions have been proposed by Giannini and Mosconi (1992). However, these authors don’t clearly distinguish in their theorems the arbitrary part (namely the nullity of some parameter blocks) one can always achieve without any loss of generality, of the nullity of the parameter blocks resulting from the non causality property.

Therefore given these observations, we propose in this paper some extensions of the existing exogeneity conditions, which are based on two canonical decompositions of the long-run matrix \( \Pi \) according to the property one is interested in testing (non-causality or strong exogeneity\(^1\)). These canonical representations exploit the fact that the \( \beta \) cointegrating and \( \alpha \) loading factor matrices

\(^1\)In this paper, we shall confine ourselves to the concepts of weak and strong exogeneity proposed by Richard (1980) and Engle et al. (1983).
are not unique in the extent where \( \Pi = \alpha \beta' = (\alpha \Psi^{-1})(\Psi \beta') \) for any \( r \times r \) non singular matrix \( \Psi \). An interesting feature of our conditions is that they are necessary and sufficient. An appealing aspect of these conditions for the practitioner is that they can be tested using asymptotically chi-squared distributed test statistics which can easily be computed in a user friendly menu driven environment with most statistical computer packages as CATS in Rats or MALCOLM 2.4 (Mosconi, 1998) for instance, so that our approach is fully operational empirically. One key concept of our analysis is the **purely exogenous long-run path**, that is to say, a cointegrating relationship which only involves “exogenous” variables. Note that we do not address in this paper the identification problem of cointegrating vectors nor the issue of testing over-identifying restrictions on long-run parameters even if these issues are of course crucial in any dynamic structural econometric modelling, since our primary aim here is to give intrinsic exogeneity conditions.

The rest of the paper is organized as follows. Section II sets out the general VECM framework. Section III introduces two canonical representations of the long run matrix \( \Pi \) and proposes a new set of non-causality and strong exogeneity conditions. Section IV deals with inference and testing which are conducted within the setting proposed by Johansen. Section V reports some Monte Carlo results and analyses the asymptotic and finite sample properties of the sequential procedure developed in section IV. Finally, concluding remarks are presented in section VI and specific recommendations are provided for applied researchers. Proofs of important results are relegated to the appendix.

## 2 Cointegrated vector autoregressions

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

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

e.i.e. the system of reduced-rank dynamic equations which has been analyzed by Johansen (1988, 1991). \( \Gamma, \alpha, \beta \) 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 $\alpha$ is the matrix of adjustment or feedback coefficients, which measure how strongly the $r$ stationary variables $\beta' X_{t-1}$ feedback onto the system. $\varepsilon_t$ is an i.i.d normal distributed vector of errors, with a zero mean and a positive definite covariance matrix $\Sigma$; and $p$ is a constant integer. To keep the notation as simple as possible we omit deterministic components.

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

$$
\begin{align*}
&\text{conditional model} \\
&\Delta Y_t = \sum_{i=1}^{p-1} \Gamma_{YY,i}^+ \Delta Y_{t-i} + \sum_{i=0}^{p-1} \Gamma_{YZ,i}^+ \Delta Z_{t-i} + \alpha_Y^+ \left[ \begin{array}{c} \beta_Y' \\ \beta_Z' \end{array} \right] \left[ \begin{array}{c} Y_{t-1} \\ Z_{t-1} \end{array} \right] + \eta_{Y,t} \\
&\text{marginal model} \\
&\Delta Z_t = \sum_{i=1}^{p-1} \Gamma_{ZY,i}^+ \Delta Y_{t-i} + \sum_{i=1}^{p-1} \Gamma_{ZZ,i}^+ \Delta Z_{t-i} + \alpha_Z^+ \left[ \begin{array}{c} \beta_Y' \\ \beta_Z' \end{array} \right] \left[ \begin{array}{c} Y_{t-1} \\ Z_{t-1} \end{array} \right] + \varepsilon_{Z,t}
\end{align*}
$$

\(^2\)A review of the econometric analysis of I (2) variables is provided in Haldrup (1998).
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 Necessary and sufficient conditions for non-causality and strong exogeneity in VECM**

An important issue concerns Granger non-causality testing in VECM. This constitutes the purpose of the first sub-section which provides a necessary and sufficient condition for non-causality. This condition is easier to implement empirically than that of Toda and Phillips (1991) since it can always be tested in any cases by means of traditional chi-squared statistics. Moreover, unlike the condition proposed by Giannini and Mosconi (1992), our condition permits to clearly distinguish the nullity of the parameter blocks one can always achieve without any loss of generality of the nullity of the parameter blocks resulting from the non causality property.

Besides, when analyzing strong exogeneity one deals simultaneously with weak exogeneity and non-causality. The second sub-section addresses this issue and develops a framework based on a canonical decomposition of the long-run matrix $\Pi$ in which a simple necessary and sufficient exogeneity condition for
strong exogeneity can be formulated.

It must be underlined that the necessary and sufficient conditions for non-causality and strong exogeneity proposed here can be expressed as minimum conditions on the parameters of canonical decompositions of the long-run matrix, in which the nullity of some parameter blocks implies no loss of generality. Actually, these representations exploit the indeterminacy existing on the $\alpha$ and $\beta$ matrices: it is indeed now well-known that the parameters of these matrices are not separately identified without $r^2$ additional restrictions (cf. Bauwens and Lubrano, 1994), since for any non-singular $r \times r$ matrix $\Psi$, we could define $\Pi = (\alpha \Psi^{-1}) \left( \Psi \beta \right)$, and $\alpha^* = \alpha \Psi^{-1}$, $\beta^* = \beta \Psi$ would be equivalent matrices of adjustment coefficients and cointegrating vectors.

3.1 Non-causality

As it is now generally admitted the non-causality hypothesis can turn out to be difficult to test empirically in VECM because test statistics usually require special conditions to be asymptotically chi-squared distributed (see e.g. Toda et Phillips, 1991, theorem 2): indeed, unless $r = 1$, long run non causality from $Y$ to $Z \left\{ \alpha_Z \beta_Y = 0 \right\}$ implies non linear restrictions on long run parameters. The aim of this section is to remedy this difficulty by developing a framework that permits to formulate a necessary and sufficient condition for non-causality which can be tested with statistics having in any cases an asymptotic chi-squared distribution. To this end, we consider the following theorem proved in Rault (2000):

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

(i) If we define $r_1 = \text{rank} (\beta_Y)$, then the $\alpha$ and $\beta$ matrices can always be repara-
metrised as follows:

\[
\beta = [\beta_1 \beta_2] = \begin{bmatrix} \beta_{YY} & 0 \\ \beta_{ZY} & \beta_{ZZ} \end{bmatrix}
\]

\[
\alpha = [\alpha_1 \alpha_2] = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ} \\ \alpha_{ZY} & \alpha_{ZZ} \end{bmatrix}
\]

where \(\beta_{YY}, \alpha_{YY}, \beta_{ZY}, \alpha_{YZ}, \beta_{ZZ}, \alpha_{ZZ}\) are respectively \(g \times r_1, g \times r_1, k \times r_1, k \times r_1, k \times r_1, k \times r_1\) sub-matrices, with rank \((\beta_{YY}) = r_1\) and rank \((\beta_{ZZ}) = r - r_1\).

(ii) \(r_1\) is uniquely defined and is invariant to the chosen reparametrisation. It is such as\(^3\)

\[
\max(0, r - k) \leq r_1 \leq \min(g, r).
\]

(See Rault, 2000, for Proof).

Under the reparametrisation of the \(\alpha\) and \(\beta\) matrices of theorem 1, the conditional and marginal models (cf. equation 2) become:

\[
\begin{align*}
\text{conditional model:} \\
\Delta Y_t &= \sum_{i=0}^{p-1} \Gamma_{YY,i}^+ \Delta Y_{t-i} + \sum_{i=1}^{p-1} \Gamma_{YZ,i}^+ \Delta Z_{t-i} + \alpha_{YY}^+ \beta_{YY}^* X_{t-1} + \alpha_{YZ}^+ \beta_{YZ}^* Z_{t-1} + \eta_{Y,t} \\

\text{marginal model:} \\
\Delta Z_t &= \sum_{i=1}^{p-1} \Gamma_{YZ,i} \Delta Y_{t-i} + \sum_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} + \alpha_{YZ} \beta_{YY}^* X_{t-1} + \alpha_{ZZ} \beta_{ZZ}^* Z_{t-1} + \varepsilon_{Z,t}
\end{align*}
\]

(2a)

Theorem 1 enables to determine the minimum number of long-run relationships containing necessarily the \(Y_t\) variables. It is proved using a basis change in the cointegrating space which permits to separate the cointegrating vectors into two sub-groups: a first one of dimension \(r_1\) containing both \(Z_t\) and \(Y_t\) variables, and another one of dimension \(r - r_1\) only composed of \(Z_t\) variables, denoted hereafter “purely exogenous long-run paths”. To this partition of the \(\beta\) matrix corresponds a new \(\alpha\) separated into \(\alpha = [\alpha_1 \alpha_2]\). Theorem 1 implies

\(^3\)This property is derived from rank \((\beta) = r\).
no loss of generality and only requires the determination of the $r_1$ rank of the upper block of the $\beta$ matrix, i.e. $\beta_{Y}$.

Given this canonical decomposition we can state the following proposition:

**Proposition 1**: Necessary and sufficient non-causality condition.

$Y$ doesn’t cause $Z$ in Granger sense (1969), if and only if

\[ \frac{1}{2} \Gamma_{Z,Y,i} = 0, \quad i = 1, \ldots, p - 1 \]

\[ a_{Z,Y} = 0 \]

in the canonical representation given by theorem 1.

The proof follows the same line of arguments as those presented in Toda and Phillips (1991) and is omitted here to save space. It is obvious that if $r_1 = 0$, then the $\beta_1$ vectors vanish completely, which means in other terms that there only exist purely exogenous long-run paths. In this case, the above condition reduces itself to $\Gamma_{Z,Y,i} = 0$.

The non-causality condition given above is very convenient to use empirically since long run non-causality hypothesis is equivalent to $\alpha_{Z,Y} = 0$, and hence can be tested using asymptotic $\chi^2$ tests (cf. section 4). It is indeed not a non-linear condition like Toda and Phillips’s one (1991) since it only implies nullity of some loading factors in the $\alpha$ matrix. Moreover, unlike the conventional sufficient long-run non-causality condition ($\alpha_{Z} = 0$), which has for consequence to eliminate the error-correction mechanism $(\alpha_{Z} \beta X_{t-1})$ in the marginal model, our condition doesn’t constrain this model to be a VAR in first differences, since it can also include purely exogenous long-run paths. Notice that this condition remains identical if one considers the parameters of the first $g$ equations of the conventional VECM partitioned into $Y_t$ and $Z_t$.

### 3.2 Strong exogeneity

This sub-section deals with the strong exogeneity hypothesis and introduces a necessary and sufficient condition for strong exogeneity, less restrictive and easier to handle than the existing one ($\Gamma_{Z,Y} (L) = 0$ and $\alpha_{Z} \beta_Y = 0$). Indeed, our condition allows both the marginal and conditional models to contain purely exogenous long-run paths. Before stating this condition we need a more general decomposition of the $\Pi$ matrix and we give it in the following theorem.
Theorem 2 Let \( \Pi = \alpha \beta' \) be a \( n \times n \) reduced rank matrix of rank \( r \) (\( 0 < r < n \)) and consider the reparametrisation
\[
\beta = \begin{bmatrix} \beta_{YY} & 0 \\ \beta_{ZY} & \beta_{ZZ} \end{bmatrix}, \quad \alpha = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ} \\ \alpha_{ZY} & \alpha_{ZZ} \end{bmatrix}
\]
given in theorem 1. Then:

(i) there exists an integer \( r_2 \) so that the \( \alpha \) et \( \beta \) matrices can always be reparametrised as follows:
\[
\alpha = [\alpha_1 \alpha_21 \alpha_{22}] = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ_1} & \alpha_{YZ_2} \\ \alpha_{ZY} & 0 & \alpha_{ZZ_1} \end{bmatrix},
\beta = [\beta_1 \beta_{21} \beta_{22}] = \begin{bmatrix} \beta_{YY} & 0 & 0 \\ \beta_{ZY} & \beta_{ZZ_1} & \beta_{ZZ_2} \end{bmatrix}
\]
where \( \alpha_{YY}, \beta_{YY}, \alpha_{ZY}, \beta_{ZY}, \alpha_{YZ_1}, \beta_{YZ_1}, \alpha_{YZ_2}, \alpha_{ZZ_2}, \beta_{ZZ_2} \) are respectively \( g \times r_1, g \times r_1, k \times r_1, k \times r_1, g \times r^*, k \times r^*, g \times r_2, k \times r_2, k \times r_2 \) submatrices, with \( r_1 + r_2 + r^* = r \) and \( \text{rank}(\alpha_{ZZ_2}) = r_2 \geq 0 \).

(ii) if in addition \( \alpha_{ZY} = 0 \) (or \( r_1 = 0 \)), then \( r_2 \) is uniquely defined and is invariant to the chosen reparametrisation. It is such as:
\[
\max(0, r - r_1 - g) \leq r_2 \leq \min(g, k, r).
\]

(See Appendix for Proof).

Given the new expression of the \( \alpha \) and \( \beta \) matrices given in theorem 2 (i), the conditional and marginal models (cf. equation 2) can be rewritten as:

conditional model:
\[
\Delta Y_t = \sum_{i=1}^{p-1} \Gamma_{YY,i}^+ \Delta Y_{t-i} + \sum_{i=0}^{p-1} \Gamma_{YY,i}^+ \Delta Z_{t-i} + \alpha_{YY} \beta_1 X_{t-1} + \left( \alpha_{YZ_1} \beta_{ZZ_1} + \alpha_{YZ_2} \beta_{ZZ_2} \right) Z_{t-1} + \eta_{Y,t}
\]

marginal model:
\[
\Delta Z_t = \sum_{i=1}^{p-1} \Gamma_{ZY,i} \Delta Y_{t-i} + \sum_{i=1}^{p-1} \Gamma_{ZZ,i} \Delta Z_{t-i} + \alpha_{ZY} \beta_1 X_{t-1} + \alpha_{ZZ_1} \beta_{ZZ_1} Z_{t-1} + \varepsilon_{Z,t}
\]

\(^4\text{Remind that rank}(\alpha) = r\).
Theorem 2 is proved using a basis change in the dual space of the cointegrating space, namely in the adjustment space (generated by the rows of the $\alpha$ matrix), which leads to partitioning the $(r - r_1)$ long run relations involving the $Z_t$ variables alone into two sub-groups of respectively dimension $r^*$ and $r_2$: a first sub-group which only belongs to the conditional model and a second one which both appears in the marginal and conditional models. To this second partition on the $\alpha$ adjustment space corresponds a new $\beta$ whose expression is given by theorem 2. The fact that the $\alpha$ and $\beta$ matrices are of rank $r$ entails in particular that the $\text{rank}(\alpha^+_{YZ_1}) = \text{rank}(\beta_{ZZ_1}) = r^* \geq 0$ and that $\text{rank}(\beta_{ZZ_2}) = r_2 \geq 0$. The table below summarizes in which equations these three distinctive groups of long-run relations appear:

<table>
<thead>
<tr>
<th>Cointegrating relations</th>
<th>Number of cointegrating vectors</th>
<th>Coefficients in the conditional model</th>
<th>Coefficients in the marginal model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(r_1 + r_2 + r^* = r)$</td>
<td>$[\beta_{Y}, \beta_{Z}] = [\eta_{it} \sim I(0)]$</td>
<td>$r_1$</td>
<td>$\alpha_{Y}$</td>
</tr>
<tr>
<td>$\beta_{ZZ_1}Z_t = \eta_{2t} \sim I(0)$</td>
<td>$r^*$</td>
<td>$\alpha^+_{YZ_1}$</td>
<td>$\alpha_{YZ}$</td>
</tr>
<tr>
<td>$\beta_{ZZ_2}Z_t = \eta_{2t} \sim I(0)$</td>
<td>$r_2$</td>
<td>$\alpha^+_{YZ_2}$</td>
<td>$0$</td>
</tr>
</tbody>
</table>

**Remark 1** - The canonical decomposition given in theorem 2 can be applied to any singular $\Pi$ matrix of rank $r$.

**Remark 2** - Some rank conditions are explicitly given in theorem 2. Nevertheless one must keep in mind that the ranks of the different blocks of the $\alpha$ and $\beta$ matrices are yet always linked by the two following conditions$^5$:

- the $\Pi = \alpha\beta$ matrix is of rank $r$,
- the $\alpha'\left(I_n - \sum_{i=1}^{p} \Gamma_i\right)\beta$ matrix is of full rank $(n - r)$.

Consequently the different parameter blocks cannot be equal to zero independently of the short-run coefficients $\Gamma$, and of the rank conditions given above.

$^5$Note that the second condition rules out the possibility of variables integrated of order 2 and assures that the variables are at most integrated of order 1. All our results are valid as long as this condition holds. However this condition is not a relation between the parameters and it corresponds to concrete economic situations. Our method can be applied under the assumption of stability of this situation, i.e. no variable is supposed to go from the state I (1) to the state I (2). Actually, this is a relatively usual condition in the econometric literature.
The strong exogeneity hypothesis for the parameters of interest combines non-causality with weak exogeneity. Let’s remember that Engle and al ‘s (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 Mouchart, 1980). We can now established the following proposition:

**Proposition 2 : Necessary and sufficient condition for strong exogeneity.** If $r_2 < k$ then $Z_t$ is strongly exogenous for all the parameters of the conditional model if and only if

$$\begin{cases} 
\alpha_{ZY} = 0 \\
\alpha_{YZ} = 0 \\
\Gamma_{ZY,i} = 0, \ i = 1, ..., p - 1 
\end{cases}$$

in the canonical representation given by theorem 2.

The proof follows directly from the previous results and hence is omitted here to save space. If $r_2 = k$, this means that there exists as many purely exogenous long-run paths as $Z_t$ variables, which entails that the $Z_t$ are stationary\(^6\). In this case there is no $\beta_{ZZ}$ parameter to estimate since this sub-matrix is equal to $I_k$; therefore, these variables can appear as well in the conditional model as in the marginal one.

Like the non-causality condition given in section 3.1, the condition in proposition 1 doesn’t prohibit the marginal model from including cointegration relations that involve purely exogenous long-run paths. Thus, it doesn’t entail this model to be a VAR in first differences, since in addition to the nullity of the parameters $\Gamma_{ZY,i} = 0, \ i = 1, ..., p - 1$ associated to variables in first differences, it is only the corresponding part of $\alpha$ (and not full lines) which is required to vanish for strong exogeneity. What is more, the conditional and marginal models now contain two completely separate sets of purely exogenous long run paths. This means that the parameters of interest of the investigator can now also include a subset of purely exogenous long run paths in addition to the long

---

\(^6\)Notice that the condition $\text{rank} (\beta_{ZZ}) = r - r_1$ implies that $k \geq r - r_1 = r_2 + r^*$. In the case where $r_2 = k$ then $r^* = 0$. 

12
run relationships involving both $Y_t$ and $Z_t$. The strong exogeneity of the $Z_t$ variables for all parameters of the conditional model implies that:

(i) none of the $r_2$ purely exogenous long-run paths appearing in the marginal model belong to the conditional model,

(ii) $Y$ doesn’t cause $Z$ in Granger sense (1969), which implies that:

a) none of the $r_1$ long-run paths containing both $Y_t$ and $Z_t$ appear in the marginal model,

b) $\Gamma_{ZY,i} = 0$, $i = 1, ..., p - 1$.

Strong exogeneity entails a complete separation of the process generating $\{Y_t/Z_t\}$ and $\{Z_t\}$, that is between the conditional and marginal models. Under this hypothesis one can make valid forecasts of $Y$ from the conditional model, given forecasts of $Z$ from the marginal model. Theorem 3 provides an important result concerning the non-causality hypothesis.

**Theorem 3 (deduced from Giannini and Mosconi, 1992, theorem 2).**

If $X_t$ is an $n$-dimensional process integrated of order 1, cointegrated of order $r$, and generated by equation (1), then the necessary and sufficient non-causality condition ($\alpha_{ZY} = 0$ and $\Gamma_{ZY,i} = 0$, $i = 1, ..., p - 1$) entails $r^* = 0$ in the canonical representation given by Theorem 2.

Indeed, if $r^*$ were strictly positive, the non-causality condition would be sufficient for the existence of variables integrated of order 2, since the matrix $\alpha'_\perp (I_n - \sum_{i=1}^{p} \Gamma_i)\beta_\perp$ would have in this case reduced rank.

We mention now for completeness the following corollary which considers the parameters of the first $g$ equations of the partitioned $VECM$ as of economic interest.

**Corollary 1:** $Z_t$ is strongly exogenous for the parameters of the first $g$ equations of the partitioned $VECM$ into $Y_t$ and $Z_t$ if this reduced form is already orthogonalized (i.e. $\Sigma_{YZ} = 0$) and if $Z_t$ is strongly exogenous for the parameters of the conditional model. (See Appendix for Proof).
Note that this last 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).

4 Inference and testing

The necessary and sufficient conditions for non-causality and strong exogeneity introduced in section 3 require first and foremost to rewrite the $\Pi$ matrix under one of the canonical decompositions given by theorems 1 and 2. Then, in this framework these conditions have all been expressed in terms of coefficient nullities of the $\alpha$, $\beta$ and $\Gamma$ matrices, which permits to use the conventional chi-squared statistics (see Johansen, 1995). As we have already noticed, these two representations require the determination of one or two specific ranks of submatrices, according to the hypothesis under consideration. In this section we focus exclusively on the decomposition given in theorem 2 since non-causality testing in the framework of theorem 1 has already been carefully investigated in Rault (2000). We thus develop a sequential procedure of rank tests suitable to the remaining case, that is to the strong-exogeneity hypothesis. More precisely, we first provide a sequential procedure of rank tests in order to determine the $r_1$ and $r_2$ ranks of the $\beta_Y$ and $\alpha_{ZZ}$ matrices. These ranks are needed to rewrite the $\Pi$ matrix under the canonical form given in theorem 2. We then indicate how to test in this framework the necessary and sufficient condition for strong exogeneity given in proposition 2. Note that it is assumed here that the cointegrating rank is known, that is, has been estimated in the lines of Johansen (1988), at a preliminary step (cf. section 2).

The strong exogeneity hypothesis may be investigated using the following sequential test procedure composed of four stages:

1) first determine the $r_1$ rank of the $\beta_Y$ matrix which corresponds to the reparametrisation of the $\alpha$ and $\beta$ matrices given in theorem 1 using the simple
sequential test procedure proposed by Rault (2000). This procedure is based on asymptotically chi-squared distributed LR statistics whose properties have been analyzed with Monte-Carlo experiments for known \( r \).

2) then, for the value of the \( r_1 \) rank found in stage 1, test \( H_0 : \alpha_{ZY} = 0 \) using conventional asymptotic \( \chi^2 \) tests (cf. for instance Toda and Phillips (1991, 1993). If you accept the null hypothesis pass to step (3), or else stop. The latter case implies you also refuse the strong exogeneity hypothesis since as stated in proposition 2 it requires \( \alpha_{ZY} = 0 \).

3) finally, given the \( r_1 \) rank found in step 1 and \( \alpha_{ZY} = 0 \) consider the model given by the canonical decomposition in theorem 1:

\[
\beta = \begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix} = \begin{bmatrix} \beta_{YY} & 0 \\ \beta_{ZY} & \beta_{ZZ} \end{bmatrix}, \quad \alpha = \begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix} = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ} \\ \alpha_{ZY} & \alpha_{ZZ} \end{bmatrix}
\]

and for \( j = 1, \ldots, \min(g, r - r_1) \) test the restrictions:

\[
\alpha_{ZZ} = \begin{bmatrix} 0_{(k,j)} & \alpha_{ZZ_2} \end{bmatrix}
\]

(where \( \alpha_{ZZ} \) and \( \alpha_{ZZ_2} \) are respectively \( k \times r - r_1 \) and \( k \times r - r_1 - j \) sub-matrices)

under which the long-run parameters can be written as follows:

\[
\beta = \begin{bmatrix} \beta_1 & \beta_21 & \beta_{22} \end{bmatrix} = \begin{bmatrix} \beta_{YY} & 0 & 0 \\ \beta_{ZY} & \beta_{ZZ_1} & \beta_{ZZ_2} \end{bmatrix}, \\
\alpha = \begin{bmatrix} \alpha_1 & \alpha_{21} & \alpha_{22} \end{bmatrix} = \begin{bmatrix} \alpha_{YY} & \alpha_{YZ_1} & \alpha_{YZ_2} \\ \alpha_{YZ} & 0 & 0 \\ \alpha_{ZZ_2} & 0 & \alpha_{ZZ_2} \end{bmatrix}
\]

where \( \beta_{YY}, \alpha_{YY}, \beta_{ZY}, \beta_{ZZ_1}, \alpha_{YZ_1}, \beta_{ZZ_2}, \alpha_{YZ_2}, \alpha_{ZZ_2} \) are respectively \( g \times r_1, g \times r_1, k \times j, g \times j, k \times r - r_1 - j, g \times r - r_1 - j, k \times r - r_1 - j \) sub-matrices.

The above writing corresponds to a model defined as:

\[
H_{0,j} \begin{cases} 
\text{There exists a basis of the cointegrating and adjustment spaces such as } \beta = (\beta_1, H_1 \beta_{ZZZ}, H_1 \beta_{ZZZ}), \alpha = (H_2 \alpha_{YY}, H_2 \alpha_{YZ}, \alpha_{22}) \\
\text{with, } \text{rank} (\beta_{YY}) = r_1, \text{rank} (\beta) = \text{rank} (\alpha) = r 
\end{cases}
\]

where \( H_1 = \left( \begin{smallmatrix} 0_{(g,k)} \\ I_k \end{smallmatrix} \right) \) and \( H_2 = \left( \begin{smallmatrix} I_g \\ 0_{(g,k)} \end{smallmatrix} \right) \)

and \( \beta_1, \beta_{ZZ_1}, \beta_{ZZ_2}, \alpha_{YY}, \alpha_{YZ_1}, \alpha_{22}, \) are respectively \( n \times r_1, k \times j, k \times r - r_1 - j, g \times r_1, g \times j, n \times r - r_1 - j \) sub-matrices.
The parameters to be estimated are $\beta_1, \beta_{ZZ_1}, \beta_{ZZ_2}, \alpha_{YY}, \alpha_{YZ_1}, \alpha_{22}$, given that some of the constraints apparently imposed on long-run parameters, in particular those on the $\beta$ matrix are only identifying constraints in our model. The boundaries for $j$ are calculated from the rank conditions:

$$1 \leq j \leq k$$

$$j \leq g$$

$$0 \leq r - r_1 - j$$

To test the $H_{0,j}$ hypothesis, we proceed as in Johansen and Juselius (1992), Johansen (1995), Giannini and Mosconi (1992), Konishi and Granger (1993) and we use the appropriate version of the switching algorithms developed by these authors (cf. Hecq et al., 2001) to test our linear restrictions on the long-run parameters, the test statistic being in this case asymptotically chi-squared distributed. As pointed out by a referee this assumes that the switching algorithm reaches the global maximum of the likelihood function. Our way of proceeding is thus very close to that of Hecq et al., (2001) who test restrictions on the $\alpha$ matrix in the context of separability analyses.

More precisely, as our aim is to determine the $r_2$ rank of the $\alpha_{ZZ}$ sub-matrix, let us consider the following sequences of null hypotheses:

$$
\begin{cases}
H_{0,1} \text{ i.e. } \text{rank} (\alpha_{ZZ}) \leq \min (k, r - r_1 - 1), \\
\vdots \\
H_{0,j} \text{ i.e. } \text{rank} (\alpha_{ZZ}) \leq \min (k, r - r_1 - j), \text{ as long as } H_{0,j-1} \text{ is not rejected}, \\
\end{cases}
$$

To test these different hypotheses, we adopt the following sequential test procedure:\footnote{It must be underlined that the hypotheses about the asymptotic distribution of the statistic under the null are not rejected by the results of our simulations reported in section 5.}

\footnote{This procedure is a natural extension of the one presented in great details in Rault (2000).}

\footnote{Let’s point out (cf. the Monte Carlo experiment results) that this procedure performs quite well in term of $\alpha_{ZZ}$ rank selection, even if some $\alpha_{ZZ}$ columns are linear combination of others.}

16
Step 1: test $H_{0,1}$ with the $\xi_1$ statistic at the $\alpha_1$ level
and reject $H_{0,1}$ if $\xi_1 > \chi^2_{1-\alpha_1}(v_1) \implies \text{rank}(\alpha ZZ) = r - r_1$,

for $j = 2, ..., \min (g, r - r_1)$, as long as $H_{0,j-1}$ is not rejected,
Step j : test $H_{0,j}$ with the $\xi_j$ statistic at the $\alpha_j$ level
and reject $H_{0,j}$ if $\xi_j > \chi^2_{1-\alpha_j}(v_j) \implies \text{rank}(\alpha ZZ) = r - r_1 - j + 1$
else accept $H_{0,j}$ if $\xi_j < \chi^2_{1-\alpha_j}(v_j) \implies \text{rank}(\alpha ZZ) = r - r_1 - j$

Each statistic is a likelihood ratio test:

$$\xi_j = -2 \ln Q(L_{0,j}/L_1) = T \left[ \sum_{i=1}^{r_1} \ln(1 - \tilde{\rho}_{1i}) + \sum_{i=1}^{r_2} \ln(1 - \tilde{\rho}_{2i}) + \sum_{i=1}^{r_3} \ln(1 - \tilde{\rho}_{3i}) - \sum_{i=1}^{r} \ln(1 - \tilde{\lambda}_i) \right]$$

which is asymptotically distributed under $H_{0,j}$ as a $\chi^2_{1-\alpha_j}(v_j)$, where $\nu_j$ is the number of degrees of freedom calculated as the number of zero restrictions implied by $H_{0,j}$, that are beyond those from normalization. $L_1$ corresponds to the cointegrating hypothesis under the assumption of $r$ cointegrating vectors, $\tilde{\lambda}_i$ denotes the eigenvalues of the unrestricted VECM, and $\tilde{\rho}_{1i}, \tilde{\rho}_{2i}, \tilde{\rho}_{3i}$ correspond to the eigenvalues associated respectively to the $r_1, j, (r-r_1-j)$ restricted cointegrating relationships and their associated weights.

It should be emphasized that as the $\xi_j$ statistic is derived under asymptotic arguments it seems to provide a bad approximation of the finite sample distribution since it tends to over-reject true nulls in small samples (see Psaradakis, 1994). Therefore, as one often encounters small size samples in empirical applications, we also consider the adjusted LR tests statistics, which is given by replacing $T$ by $T - (s/n) + 0.5 [n - r(n - k)/(n + 1)]$ in equation (3), where $s$ denotes the number of parameters to be estimated in equation (1). This small sample correction performs better in terms of size distortion when testing for linear restrictions on a multivariate Gaussian model (see Anderson, 1984, ch 8 for the stationary case, and the simulations of Psaradakis, 1994 for the non-stationary case).

4) Finally, we can test the strong exogeneity hypothesis for given values of the $r_1$ and $r_2$ ranks. Indeed, in this case the $\Pi$ matrix can be rewritten under the
canonical expression given in theorem 2 and the strong exogeneity hypothesis leads in this framework to the following parametric restrictions (cf. proposition 2):

$$H_{0,se} : \begin{cases} \Gamma_{Z_Y; i} = 0, & i = 1, ..., p - 1 \\
\alpha_{Y Z_2}^+ = 0 \end{cases}$$

As these restrictions only correspond to coefficient nullities several conventional tests can be carried out (Likelihood ratio test, Lagrange Multiplier (LM) test, Wald test). As the underlying motivation to imposing strong exogeneity restrictions is to reduce the dimension of the estimated model (in a forecasting purpose), a LM test could probably be the more robust one to possible mispecification of the full VECM (cf. Zhu, 2001). Such a test can easily be implemented in empirical applications using most statistical computer packages.

5 The Monte Carlo design and results

This section reports some Monte Carlo replications and analyzes the size distortions and power of the sequential procedure of rank tests introduced in section 4. Artificial data were generated from four data generation processes (DGPs) depicted in Table 1 of Appendix 2, each containing 11 variables ($g = 5$, $k = 6$), integrated of order one, cointegrated of order 4, expressed in VECM forms. They are all of rank $r_1 = 1$ and have short-run dynamic, since $p$ is chosen to be equal to 2. Alternative $r_2$ ranks of the $\alpha_{ZZ}$ matrix were specified varying from 0 to 3.

For each Monte Carlo simulation, we generated 6000 series of length $T + 100 + p$, where $p$ denotes the lag length in the estimated VECM. We discarded the first 100 observations to eliminate start-up effects. The vector of innovations $\varepsilon_t$ 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 $X_1 = e_1 \sim N(0_{11}, I_{11})$. All simulations were carried out on a 266 Pentium II, using the matrix programming language GAUSS, the $\varepsilon_t$ 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; \( T = 50, 100, 200, 500, 1000 \), and the adjusted LR tests statistics was used for \( T \leq 100 \). In each replication, the lag length \( p \) and the dimension of the cointegrating rank \( r \) are treated as known in a first step, so that we can exclusively focus on the performance of our sequential test procedure.

We now comment the results of our sequential test procedure related to strong exogeneity testing. We first consider the situation where the \( r_1 \) rank of the \( \beta_Y \) matrix has correctly been selected as well as \( \alpha_{ZY} = 0 \), that is we suppose that the starting point is given by stage 3 of the procedure developed in section 4. Then, only the \( r_2 \) rank of \( \alpha_{ZZ} \) has to be determined. The tabulated results of the experiments are reported respectively in Tables 2 and 3 (cf. Appendix 2). Tables 2 contains the estimated empirical size and power of the \( H_{0,j} \) null hypothesis tests, while Tables 3 presents the sequential test procedure empirical size. The numbers in the body of Tables 2 are respectively the percentage of rejections of the null hypotheses and the percentage of acceptance of the true rank \( r_2 \) of \( \alpha_{ZZ} \) (that is the proportion of “success”of \( \alpha_{ZZ} \) actual rank determination) at the 5 \% level. 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\(^{10} \) and can be summarized as follows.

All \( H_{0,j} \) null hypothesis tests \((j = 1, \ldots, 4)\) suffer from size distortion in small samples \((T = 50, 100)\), but in large samples \((T = 500, 1000)\) they approximate quite well the correct size. Moreover, our simulation results indicate both finite distance and asymptotic power of the tests equal to 1.

As far as the sequential test procedure is concerned, our simulations show as in Rault (2000) that the multiplicity of tests leads to a global size problem in small samples \((T = 50, 100)\), since the sequential procedure estimated size turns out to be highly dependent on the number of tests necessary to conclude (respectively 7.17 \%, 9.63 \%, 10.7 \% for \( r_2 = 2, \ldots, 0 \) and \( T = 100 \)). However, it is no more the case in large samples \((T = 200, 500 \text{ or } 1000)\) since for any possible \( r_2 \) true rank,\(^{10} \)

\( ^{10} \)In Rault (2000), it is the rank of a sub-matrix extracted from the \( \beta \) matrix which is investigated and not an \( \alpha \) sub-matrix rank, as it is the case here.
the estimated size is now always very close to 5 % (respectively 5.23 %, 5.49 %, 5.64 % for \( r_2 = 2, ..., 0 \) 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.

Table 4 reports results in the situation where \( r_1 \) is treated as unknown and where the nullity of the parameters of the \( \alpha_{ZY} \) matrix has to be tested in addition to the \( r_2 \) rank. The performance of the tests is changed quite substantially compared to the previous situation, since now, to return a “success” in the experiment, our sequential procedure must make correct decisions concerning \( r_1 \), \( \alpha_{ZY} \) and \( r_2 \). It appears that even for large samples, the sequential procedure estimated size deteriorates. It is respectively of 11.7%, 11.9 %, 12.02 % for \( r_1 = 1, r_2 = 2, ..., 0 \) and \( T = 1000 \)). However the size of the procedure remains quite acceptable given the number of tests to be carried out to reach a decision.

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. The results of the simulation experiments reported in Table 5 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 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 \( \beta \) matrix, which only adds a supplementary step in the sequential procedure 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. This is a useful and significant result for the practitioner as it suggests that the sequential test procedure may be conducted under the assumption of full rank of the \( \Pi \) matrix without affecting its performance very markedly.
6 Concluding remarks

In this paper we have studied the non-causality and strong exogeneity hypotheses in the context of a vector error correction model and we have generalized the results obtained by Johansen (1992) and Toda and Phillips (1991). In particular, we have developed a general setting that constitutes a suitable basis in which a less restrictive strong exogeneity condition, as well as a simple testable non-causality condition can be formulated. This setting is obtained by two canonical representations of the long-run matrix $\Pi$ and requires the determination of the rank of one or two specific sub-matrices, according to the property one is dealing with (non-causality or strong exogeneity). The novelty of our approach is to exploit the indeterminacy of the cointegrating and loading factor matrices to give a necessary and sufficient condition for non-causality as well as a condition for strong exogeneity. These different sets of conditions permit to clearly distinguish the nullity of some parameter blocks one can always achieve without any loss of generality using basis changes of the nullity of the parameter blocks resulting from the hypothesis to be tested.

We showed in Monte Carlo experiments that providing the cointegrating rank is correctly selected or over-estimated, sequential testing to determine specific sub-matrix ranks can have asymptotically a frequency of success comparable to linear restriction testing on cointegrating parameters by usual Johansen’s tests (1992). We think furthermore that a fruitful approach may be to use the Bartlett correction factor recently suggested by Johansen (2000) to test hypotheses on cointegrating vectors, to improve the performance of our sequential procedure in small samples, but additional Monte Carlo investigations have to be carried out.

Finally, we’d like to provide the applied researcher with a brief guideline and specific comments.

First of all, we want to highlight that the exogeneity conditions introduced in this paper appear as a bit more general than the existing ones and may provide a solution to the two classical related issues often encountered in empirical
applications: (i) when there exists empirically more equilibrium relationships than stipulated by economic theory (ii) when one is willing to make inference on only the subset of cointegrating vectors in the conditional model. Indeed, when such situations occur, applied researchers are faced with the problem of what to do.

Secondly, special care should be taken in applied work to determine the true number of cointegrating vectors before testing for exogeneity. This point is of importance since our Monte Carlo experiment suggests that in data sets whose cointegrating rank has correctly been determined or over-estimated, there is a high probability that one can also discover the true ranks of sub-matrices connected to strong exogeneity testing. However, the penalty attached to wrongly determined $r$ is substantial when $r$ is under-estimated and in this case our sequential test procedure performs poorly with respect to size distortion, whatever the size of the sample is.

Thirdly, our simulation results highlight the importance of using large samples in applied research if possible, since our sequential test procedure related to strong exogeneity performs better in term of true sub-matrix rank selection in large samples.

Acknowledgements

We wish to thank the participants of the Fifth World Congress of the Bernoulli Society for Mathematical Statistics and Probability and 63rd Annual Meeting of the Institute of Mathematical Statistics, 15-20 May 2000, Guanajuato, Mexico, of the NBER/NSF Time Series Workshop, 22-23 September, 2000, Fort Collins, University of Colorado, USA, of the Royal Statistical Society International Conference, 11-15 September 2000, University of Reading, England, of the 2001 ESEM and EEA congresses in Lausanne, as well as M. Lubrano, J.P Urbain and two anonymous referees for helpful comments and suggestions on earlier drafts of this paper. The usual disclaimer applies.
References


[26] Konishi T., Granger C.W.J [1993], “Separation in cointegrated system”, manuscript UCSD.


Appendix 1: Mathematical details

Proof of theorem 2

(i) Given the expressions of the $\alpha$ and $\beta$ matrices reported in theorem 1, we make another reparametrisation, but this time in the adjustment space $(EC)$, spanned by the columns of the $\alpha$ matrix. To this end, we now consider $Im(\alpha_2)^{11}$ and we make a new basis change in order to have a basis of $Im(\alpha_2) \cap Im\left(\begin{pmatrix} 0 \\ I_k \end{pmatrix}\right)$. It is then easily shown that the transformation matrix from the $\alpha^*$ basis to the $\alpha^{**}$ basis can be written as:

$$P = \begin{bmatrix} I_{r_1} & 0 & 0 \\ 0 & I_{r^*} & B^* \\ 0 & 0 & I_{r_2} \end{bmatrix},$$

which leads to a new $\alpha$ of the form:

$$\alpha = [\alpha_1 \alpha_{21} \alpha_{22}] = \begin{bmatrix} \alpha_{Y1} & \alpha_{YZ} & \alpha_{YZ} \\ \alpha_{Z1} & \alpha_{Y} & 0 \\ \alpha_{Z2} & 0 & \alpha_{Z2} \end{bmatrix}$$

Finally the corresponding $\beta$ given by the expression $\beta^{**} = \beta^*(P^{-1})'$ can easily be checked to be the one reported in theorem 2:

$$\beta = [\beta_1 \beta_{21} \beta_{22}] = \begin{bmatrix} \beta_{1Y} & 0 & 0 \\ \beta_{2Y} & \beta_{2Y} & \beta_{2Z} \end{bmatrix}$$

(ii) Consider now a basis change in the cointegrating space of the form $\beta^* = \beta M$, which entails $\alpha^* = \alpha(M^r)^{-1}$ where $M$ is a non-singular $(r, r)$ matrix and let us determine the conditions that must hold on this transformation matrix $M$ to have:

$$\begin{cases} \text{rank}(\beta_{1Y}^r) = \text{rank}(\beta_{2Y}^r) = r_1 \\ \beta_{1Y}^r = \beta_{2Y}^r = 0 \end{cases}.$$ 

Partitioning $M$ into $\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$ enables us to write:

$$\beta^* = \begin{bmatrix} \beta_{1Y}^* & \beta_{1Z}^* \\ \beta_{2Y}^* & 0 \end{bmatrix} = \begin{bmatrix} \beta_{1Y} & \beta_{1Z} \\ \beta_{2Y} & \beta_{2Z} \end{bmatrix} \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} \beta_{1Y} M_{11} + \beta_{2Z} M_{21} & \beta_{1Y} M_{12} + \beta_{2Z} M_{22} \end{bmatrix}.$$ 

$^{11}Im(\alpha_2)$ represents the $\alpha_2$ image space.
We have then the following implications:

\[
\begin{cases}
\beta_{YY} M_{12} \Leftrightarrow M_{12} = 0, \\
\text{rank} (\beta_{YY} M_{11}) = r_1 \Leftrightarrow M_{11} \text{invertible} \\
M_{11} \text{invertible and } M \text{ invertible } \Leftrightarrow M_{12} \text{invertible}
\end{cases}
\]

Next, the corresponding \( \alpha \) given by \( \alpha^* = \alpha (M')^{-1} \) can easily be checked to have the following expression:

\[
\alpha^* = \begin{bmatrix}
\alpha_{YY} & \alpha_{YZ} \\
\alpha_{ZY} & \alpha_{ZZ}
\end{bmatrix}
\begin{bmatrix}
(M_{11}')^{-1} & -(M_{11}')^{-1} M_{21} (M_{22}')^{-1} \\
0 & (M_{22}')^{-1}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
\alpha_{YY} (M_{11}')^{-1} & -\alpha_{YY} (M_{11}')^{-1} M_{21} (M_{22}')^{-1} + \alpha_{YZ} M_{21} (-M_{22}')^{-1} \\
\alpha_{ZY} (M_{11}')^{-1} & -\alpha_{ZY} (M_{11}')^{-1} M_{21} (M_{22}')^{-1} + \alpha_{ZZ} (M_{22}')^{-1}
\end{bmatrix},
\]

that is

\[
\alpha^* = \begin{bmatrix}
\alpha_{YY}^* & \alpha_{YZ}^* \\
\alpha_{ZY}^* & \alpha_{ZZ}^*
\end{bmatrix}
\]

If \( \alpha_{ZY} = 0 \), then \( \alpha_{ZZ}^* = \alpha_{ZZ} (M_{22}')^{-1} \) and we have \( \text{rank}(\alpha_{ZZ}^*) = \text{rank}(\alpha_{ZZ}) = r_2 \), which means that \( r_2 \) is uniquely defined and is invariant to the chosen reparametrisation.

\[\Box\]

**Proof of Corollary 1**

The proof is straightforward since \( \alpha_{YZ}^+ = 0 \Leftrightarrow \alpha_{YZ} - \Sigma_{YZ} \Sigma_{ZYZ}^{-1} \alpha_{ZYZ} = 0. \]

28
### Appendix 2: Monte Carlo Results

Table 1. Data Generation Processes (DGP) \((n = 11, g = 5, k = 6, r_1 = 1)\)

<table>
<thead>
<tr>
<th>DGP ((1): r_1 = 3)</th>
<th>DGP ((2): r_2 = 2)</th>
<th>DGP ((3): r_1 = 1)</th>
<th>DGP ((4): r_2 = 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>beta</strong></td>
<td><strong>beta</strong></td>
<td><strong>beta</strong></td>
<td><strong>beta</strong></td>
</tr>
<tr>
<td>-1.50 -0.50 -1.00 0.00</td>
<td>-1.50 -0.50 -1.00 0.00</td>
<td>-1.50 -0.50 -1.00 0.00</td>
<td>-1.50 -0.50 -1.00 0.00</td>
</tr>
<tr>
<td>6.00 2.00 4.00 0.00</td>
<td>6.00 2.00 4.00 0.00</td>
<td>6.00 2.00 4.00 0.00</td>
<td>6.00 2.00 4.00 0.00</td>
</tr>
<tr>
<td>-1.35 -0.45 -0.90 0.00</td>
<td>-1.35 -0.45 -0.90 0.00</td>
<td>-1.35 -0.45 -0.90 0.00</td>
<td>-1.35 -0.45 -0.90 0.00</td>
</tr>
<tr>
<td>1.50 0.50 1.00 0.00</td>
<td>1.50 0.50 1.00 0.00</td>
<td>1.50 0.50 1.00 0.00</td>
<td>1.50 0.50 1.00 0.00</td>
</tr>
<tr>
<td>-1.50 -0.50 -1.00 0.00</td>
<td>-1.50 -0.50 -1.00 0.00</td>
<td>-1.50 -0.50 -1.00 0.00</td>
<td>-1.50 -0.50 -1.00 0.00</td>
</tr>
<tr>
<td>0.20 0.10 -0.20 -0.40</td>
<td>0.20 0.10 -0.20 -0.40</td>
<td>0.20 0.10 -0.20 -0.40</td>
<td>0.20 0.10 -0.20 -0.40</td>
</tr>
<tr>
<td>0.80 0.20 0.10 0.30</td>
<td>0.80 0.20 0.10 0.30</td>
<td>0.80 0.20 0.10 0.30</td>
<td>0.80 0.20 0.10 0.30</td>
</tr>
<tr>
<td>0.10 0.30 -0.30 0.60</td>
<td>0.10 0.30 -0.30 0.60</td>
<td>0.10 0.30 -0.30 0.60</td>
<td>0.10 0.30 -0.30 0.60</td>
</tr>
<tr>
<td>0.20 -0.50 0.70 -0.10</td>
<td>0.20 -0.50 0.70 -0.10</td>
<td>0.20 -0.50 0.70 -0.10</td>
<td>0.20 -0.50 0.70 -0.10</td>
</tr>
<tr>
<td>0.40 -0.20 0.50 0.00</td>
<td>0.40 -0.20 0.50 0.00</td>
<td>0.40 -0.20 0.50 0.00</td>
<td>0.40 -0.20 0.50 0.00</td>
</tr>
<tr>
<td>0.10 -0.30 0.40 0.20</td>
<td>0.10 -0.30 0.40 0.20</td>
<td>0.10 -0.30 0.40 0.20</td>
<td>0.10 -0.30 0.40 0.20</td>
</tr>
</tbody>
</table>

Table 2. Empirical size and power of the \(H_{0,j}\) null hypothesis tests \((j = 1,\ldots,4)\) \(^2\) (rejection per 100), with 6000 replications at the 5\% nominal level of significance \(^3\)

<table>
<thead>
<tr>
<th>Sample size (T)</th>
<th>DGP ((1): r_1 = 3)</th>
<th>DGP ((2): r_2 = 2)</th>
<th>DGP ((3): r_1 = 1)</th>
<th>DGP ((4): r_2 = 0)</th>
<th>Hypothesis tested</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha_1 W_1 = \Psi &gt; A_1)</td>
<td>(\alpha_1 W_1 = \Psi &gt; A_1)</td>
<td>(\alpha_1 W_1 = \Psi &gt; A_1)</td>
<td>(\alpha_1 W_1 = \Psi &gt; A_1)</td>
<td>(\alpha_1 W_1 = \Psi &gt; A_1)</td>
<td>(\alpha_1 W_1 = \Psi &gt; A_1)</td>
</tr>
<tr>
<td>100 100 100 100 100</td>
<td>7.97 6.65 5.43 5.12 5.09</td>
<td>0.10 0.00 0.10 0.00 0.00</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>(H_{0,1}) ((\text{range } \alpha_1 \leq 5)) against ((\text{range } \alpha_1 = 4))</td>
</tr>
<tr>
<td>(\alpha_2 W_2 = \Psi &gt; A_2)</td>
<td>(\alpha_2 W_2 = \Psi &gt; A_2)</td>
<td>(\alpha_2 W_2 = \Psi &gt; A_2)</td>
<td>(\alpha_2 W_2 = \Psi &gt; A_2)</td>
<td>(\alpha_2 W_2 = \Psi &gt; A_2)</td>
<td>(\alpha_2 W_2 = \Psi &gt; A_2)</td>
</tr>
<tr>
<td>100 100 100 100 100</td>
<td>12.6 7.32 6.19 5.38 5.14</td>
<td>0.21 0.00 0.00 0.00 0.00</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>(H_{0,2}) ((\text{range } \alpha_2 \leq 5)) against ((\text{range } \alpha_2 \leq 3))</td>
</tr>
<tr>
<td>(\alpha_3 W_3 = \Psi &gt; A_3)</td>
<td>(\alpha_3 W_3 = \Psi &gt; A_3)</td>
<td>(\alpha_3 W_3 = \Psi &gt; A_3)</td>
<td>(\alpha_3 W_3 = \Psi &gt; A_3)</td>
<td>(\alpha_3 W_3 = \Psi &gt; A_3)</td>
<td>(\alpha_3 W_3 = \Psi &gt; A_3)</td>
</tr>
<tr>
<td>100 100 100 100 100</td>
<td>20.6 9.63 7.15 5.68 5.23</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>(H_{0,3}) ((\text{range } \alpha_3 \leq 5)) against ((\text{range } \alpha_3 \leq 2))</td>
</tr>
<tr>
<td>(\alpha_4 W_4 = \Psi &gt; A_4)</td>
<td>(\alpha_4 W_4 = \Psi &gt; A_4)</td>
<td>(\alpha_4 W_4 = \Psi &gt; A_4)</td>
<td>(\alpha_4 W_4 = \Psi &gt; A_4)</td>
<td>(\alpha_4 W_4 = \Psi &gt; A_4)</td>
<td>(\alpha_4 W_4 = \Psi &gt; A_4)</td>
</tr>
<tr>
<td>100 100 100 100 100</td>
<td>20.6 9.63 7.15 5.68 5.23</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>0.00 0.00 0.00 0.00 0.00</td>
<td>(H_{0,4}) ((\text{range } \alpha_4 \leq 5)) against ((\text{range } \alpha_4 \leq 2))</td>
</tr>
</tbody>
</table>

---

1. DGP \((1)\) and \((2)\) can easily be seen to be respectively of rank \(r_1 = 3\) and \(r_2 = 2\). However the fact that DGP \((3)\) and \((4)\) are respectively of rank \(r_1 = 1\) and \(r_2 = 0\) is less straightforward: it requires noticing that the \(\beta_i\) columns of these two DGP are not linearly independent since they are respectively linked by \(C_1 = 2C_2\), for DGP \((3)\) and by \(C_1 = 2C_2, C_1 - C_2\) for DGP \((4)\). Note also that to save place the short run structure of the four DGP is not reproduced here but is available on request.

2. Note that \(r_1\) is supposed here to have correctly been selected, and \(\alpha_2 = 0\) is also supposed to have been accepted at a preliminary step.

3. The adjusted version of the test statistic was used for \(T = 50, 100\).

4. \(A_i, i = 1, \ldots, 4\) denotes the critical value from the Chi-square distribution at the 5\% level of significance.

29
Table 3. Empirical size of the sequential test procedure (rejection per 100), with 6000 replications at the 5 % nominal level of significance

<table>
<thead>
<tr>
<th>DGPS</th>
<th>DGP (2) : $t_2 = 2$</th>
<th>DGP (3) : $t_2 = 1$</th>
<th>DGP (4) : $t_2 = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P(\overline{W}_1)$</td>
<td>$P(\overline{W}_1 \overline{W}_2)$</td>
<td>$P(\overline{W}_1 \overline{W}_2 \overline{W}_3)$</td>
</tr>
<tr>
<td>Sample size T</td>
<td>50 100 200 500 1000</td>
<td>50 100 200 500 1000</td>
<td>50 100 200 500 1000</td>
</tr>
<tr>
<td>$r_2$ estimated = true $r_2$</td>
<td>12.9 7.17 6.12 5.05</td>
<td>18.2 9.63 6.87 5.49 5.09</td>
<td>20.5 10.7 7.01 5.64 5.16</td>
</tr>
</tbody>
</table>

Table 4. Empirical size of the sequential test procedure (rejection per 100), with 6000 replications at the 5 % nominal level of significance in the case where the cointegrating rank (ie. $r = 4$) is known

<table>
<thead>
<tr>
<th>DGPS</th>
<th>DGP (2) : $t_2 = 2$</th>
<th>DGP (3) : $t_2 = 1$</th>
<th>DGP (4) : $t_2 = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P(\overline{W}_1)$</td>
<td>$P(\overline{W}_1 \overline{W}_2)$</td>
<td>$P(\overline{W}_1 \overline{W}_2 \overline{W}_3)$</td>
</tr>
<tr>
<td>Sample size T</td>
<td>50 100 200 500 1000</td>
<td>50 100 200 500 1000</td>
<td>50 100 200 500 1000</td>
</tr>
<tr>
<td>$r_2$ estimated = true $r_2$</td>
<td>14.5 14.21 14.1 12.4 11.7</td>
<td>35.5 15.3 14.1 13.2 11.9</td>
<td>37.2 26.5 16.5 14.4 12.2</td>
</tr>
</tbody>
</table>

Table 5. Empirical size of the sequential test procedure (rejection per 100), with 6000 replications at the 5 % nominal level of significance in the case where the cointegrating rank (ie. $r = 4$) is not correctly selected

<table>
<thead>
<tr>
<th>DGPS</th>
<th>DGP (2) : $t_2 = 2$</th>
<th>DGP (3) : $t_2 = 1$</th>
<th>DGP (4) : $t_2 = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$P(\overline{W}_1)$</td>
<td>$P(\overline{W}_1 \overline{W}_2)$</td>
<td>$P(\overline{W}_1 \overline{W}_2 \overline{W}_3)$</td>
</tr>
<tr>
<td>Sample size T</td>
<td>50 100 200 500 1000</td>
<td>50 100 200 500 1000</td>
<td>50 100 200 500 1000</td>
</tr>
<tr>
<td>$r=2$ $r_2$ estimated = true $r_2$</td>
<td>100 100 100 100 100</td>
<td>56.3 38.9 35.8 35.1 30.2</td>
<td>60.1 51.3 39.7 35.9 29.6</td>
</tr>
<tr>
<td>$r=3$ $r_2$ estimated = true $r_2$</td>
<td>25.6 25.3 23.6 22.4 18.5</td>
<td>45.4 26.8 23.9 22.7 18.9</td>
<td>47.6 38.5 27.8 24.8 20.3</td>
</tr>
<tr>
<td>$r=5$ $r_2$ estimated = true $r_2$</td>
<td>14.9 14.8 14.7 12.6 11.8</td>
<td>36.8 16.2 14.9 13.8 12.3</td>
<td>38.1 27.0 17.2 14.9 12.6</td>
</tr>
<tr>
<td>$r=6$ $r_2$ estimated = true $r_2$</td>
<td>15.4 15.4 15.2 12.9 12.0</td>
<td>37.4 16.8 15.4 14.2 12.6</td>
<td>38.9 27.9 17.8 15.6 13.1</td>
</tr>
</tbody>
</table>

5 Note that $r_1$ is supposed here to have correctly been selected, and $\alpha_{zy}=0$ is also supposed to have been accepted at a preliminary step.
6 $P(\overline{W}_1 \overline{W}_2)$ represents the probability to be at the same time in the acceptance region $\overline{W}_1$ of test 1 and in the critical region $\overline{W}_2$ of test 2.
7 Now, to return a “success” in the experiment, the sequential procedure must make correct decisions concerning $r_1$, $\alpha_{zy}$ and $r_2$.
8 $P(\overline{W}_1 \overline{W}_2)$ represents the probability to be at the same time in the acceptance region $\overline{W}_1$ of test 1 and in the critical region $\overline{W}_2$ of test 2.
2003

0301 Basic Income/ Minimum Wage Schedule and the Occurrence of Inactivity Trap: Some Evidence on the French Labor Market
   Thierry LAURENT & Yannick L’HORTY

0302 Exonérations ciblées de cotisations sociales et évolution de l’emploi : un bilan à partir des études disponibles
   Philippe DE VREYER

0303 Incorporating Labour Market Frictions into an Optimizing-Based Monetary Policy Model
   Stéphane MOYEN & Jean-Guillaume SAHUC

0304 Indeterminacy in a Cash-in-Advance Two-Sector Economy
   Stefano BOSI, Francesco MAGRIS & Alain VENDITTI

0305 Competitive Equilibrium Cycles with Endogenous Labor
   Stefano BOSI, Francesco MAGRIS & Alain VENDITTI

0306 Robust European Monetary Policy
   Jean-Guillaume SAHUC

0307 Reducing Working Time and Inequality: What Has the French 35-Hour Work Week Experience Taught Us?
   Fabrice GILLES & Yannick L’HORTY

   Yannick L’HORTY & Christophe RAULT

0309 Inflation, Minimum Wage and Other Wages: an Econometric Study on French Macroeconomic Data
   Yannick L’HORTY & Christophe RAULT

0310 Exogeneity in Vector Error Correction Models with Purely Exogenous Long-Run Paths
   Jacqueline PRADEL & Christophe RAULT

2002

0201 Inflation, salaires et SMIC: quelles relations?
   Yannick L’HORTY & Christophe RAULT

0202 Le paradoxe de la productivité
   Nathalie GREENAN & Yannick L’HORTY

0203 35 heures et inégalités
   Fabrice GILLES & Yannick L’HORTY

0204 Droits connexes, transferts sociaux locaux et retour à l’emploi
   Denis ANNE & Yannick L’HORTY
0205 Animal Spirits with Arbitrarily Small Market Imperfection
Stefano BOSI, Frédéric DUFORT & Francesco MAGRIS

0206 Actualité du protectionnisme : l'exemple des importations américaines d'acier
Anne HANAUT

0207 The Fragility of the Fiscal Theory of Price Determination
Gaetano BLOISE

0208 Pervasiveness of Sunspot Equilibria
Stefano BOSI & Francesco MAGRIS

0209 Du côté de l'offre, du côté de la demande : quelques interrogations sur la politique française en direction des moins qualifiés
Denis FOUGERE, Yannick L'HORTY & Pierre MORIN

0210 A "Hybrid" Monetary Policy Model: Evidence from the Euro Area
Jean-Guillaume SAHUC

0211 An Overlapping Generations Model with Endogenous Labor Supply: a Dynamic Analysis
Carine NOURRY & Alain YENDITTI

0212 Rhythm versus Nature of Technological Change
Martine CARRE & David DROUOT

0213 Revisiting the "Making Work Pay" Issue: Static vs Dynamic Inactivity Trap on the Labor Market
Thierry LAURENT & Yannick L'HORTY

0214 Déqualification, employabilité et transitions sur le marché du travail : une analyse dynamique des incitations à la reprise d'emploi
Thierry LAURENT, Yannick L'HORTY, Patrick MAILLE & Jean-François OUVRARD

0215 Privatization and Investment: Crowding-Out Effect vs Financial Diversification
Guillaume GIRMENS & Michel GUILLARD

0216 Taxation of Savings Products: An International Comparison
Thierry LAURENT & Yannick L'HORTY

0217 Liquidity Constraints, Heterogeneous Households and Sunspots Fluctuations
Jean-Paul BARINCI, Arnaud CHERON & François LANGOT

0218 Influence of Parameter Estimation Uncertainty on the European Central Banker Behavior: an Extension
Jean-Guillaume SAHUC