A joint Initiative of Ludwig-Maximilians-Universität and Ifo Institute for Economic Research

CESifo Working Papers

SEPARATION, WEAK EXOGENEITY AND P-T DECOMPOSITION IN COINTEGRATED VAR SYSTEMS

A. Hecq
B. F. C. Palm
C. J.-P. Urbain*

CESifo Working Paper No. 660 (10)

February 2002

Category 10: Empirical and Theoretical Methods

CESifo
Center for Economic Studies & Ifo Institute for Economic Research
Poschingerstr. 5, 81679 Munich, Germany
Phone: +49 (89) 9224-1410 - Fax: +49 (89) 9224-1409
e-mail: office@CESifo.de
ISSN 1617-9595

* Support from METEOR through the research project "Dynamic and Nonstationary Panels: Theoretical and Empirical Issues" is gratefully acknowledged. The authors would like to thank Jörg Breitung, Geert Dhaene, Niels Haldrup, Peter Kugler, Peter Pedroni, Pravin Trivedi and an unknown referee for helpful comments and suggestions. Any shortcomings or errors are naturally ours.

CESifo *Working Paper No. 660*
February 2002

SEPARATION, WEAK EXOGENEITY AND P-T DECOMPOSITION IN COINTEGRATED VAR SYSTEMS

Abstract

The aim of this paper is to study the concept of separability in multiple nonstationary time series displaying both common stochastic trends and common stochastic cycles. When modeling the dynamics of multiple time series for a panel of several entities such as countries, sectors, firms, imposing some form of separability and commonalities is often required to restrict the dimension of the parameter space. For this purpose we introduce the concept of common feature separation and investigate the relationships between separation in cointegration and separation in serial correlation common features. Loosely speaking we investigate whether a set of time series can be partitioned into subsets such that there are serial correlation common features in the sub-groups and that there do not exist linear combinations of the common cycles which are white noise. The paper investigates three issues. First, it provides conditions for separating joint cointegrating vectors into marginal cointegrating vectors as well as separating joint short-term dynamics into marginal short-term dynamics. Second, conditions for making permanent-transitory decompositions based on marginal systems are given. Third, issues of weak exogeneity are considered. Likelihood ratio type tests for the different hypotheses under study are proposed. An empirical analysis of the link between economic fluctuations in the U.S. and Canada shows the practical relevance of the approach proposed in this paper.

JEL Classification: C32.

Keywords: separation, cointegration, common features, weak exogeneity, P-T decomposition, consumption function.

A. Hecq
Department of Quantitative Economics
University Maastricht
P.O.Box 616
6200 MD Maastricht
The Netherlands

F. C. Palm
Department of Quantitative Economics
University Maastricht
P.O.Box 616
6200 MD Maastricht
The Netherlands
f.palm@ke.unimaas.nl

J.-P. Urbain
Department of Quantitative Economics
University Maastricht
P.O.Box 616
6200 MD Maastricht
The Netherlands

1 Introduction

When modeling multiple time series, for instance for a panel of countries, sectors or firms, it is often appropriate to cluster the series into groups of variables which have strong intragroup relationships and weak or no intergroup interactions. For the analysis of high dimensional multiple time series, such a grouping can be sensible and very appealing from both a theoretical and practical point of view to deal with the "curse" of dimensionality. For example, Abadir, Hadri and Tzavalis (1999) and Gonzalo and Pitirakis (2000) show that an increase in the dimension of a cointegrated VAR model can lead to very undesirable properties for both the usual test statistics and estimators. Empirical studies in macroeconomics or international finance aimed to detect (dis)similarities across countries face the dimensionality problem once the study involves several variables and several countries. The solution that is usually adopted is to carry out a country by country analysis, e.g. extract the common stochastic trends components on a country by country basis. These "national" common stochastic trends are then compared in a subsequent stage of the analysis - see for example Banerjee, Marcellino and Orbat (2001), Haffner, Kutan and Zhou (1997), Hoffman (1999). While this seems the most practical solution given the problems mentioned above, this also raises issues as to whether the extracted components and time series properties are affected by these marginalizations. Ideally, one would like to cluster the series into groups of variables so that no loss of information occurs if the components extraction is based on single country analyses.

The concept of separation in cointegration provides a useful way of formally describing such a situation. Separation in cointegration, introduced by Konishi (1993), Konishi, Ramsey and Granger (1994), Konishi and Granger (1993) and later extended by Granger and Haldrup (1997) implies that common trends can be extracted from sub-systems of *I(1)* time series. These authors consider situations where subsets of cointegration relationships exist between economic time series which have no variables in common. Examples considered by these authors are the presence of long-run relationships between real sector variables only, and between monetary variables only. Other examples arise in panel data for several countries when sets of variables for countries taken separately are cointegrated. Under complete separation, the common trends extracted from a sub-group in a sub-system analysis do corre-

spond to those that would have been extracted from the complete system. Useful conditions for empirical work under which this correspondence holds are given in Granger and Haldrup (1997). These authors however limit their attention to Permanent-Transitory (P-T) decompositions à la Gonzalo-Granger (1995) and are essentially interested in the extraction of the common stochastic trend components. In the presence of common stochastic cycles, see Engle and Kozicki (1993) or Vahid and Engle (1993), and/or when other P-T decompositions are considered, their results need to be extended and generalized.

The major aim of this paper is to study the concept of separability in multiple time series displaying both common stochastic trends and common stochastic cycles.

The paper investigates three issues. First, it provides conditions for separating joint cointegrating vectors into marginal cointegrating vectors as well as separating joint short-term dynamics into marginal short-term dynamics. Second, conditions for making permanent-transitory decompositions based on marginal systems are given. Third, issues of weak exogeneity under separation are considered. For these purposes we introduce the concept of common feature separation and investigate the relationships between separation in cointegration and separation in serial correlation common features. We provide conditions under which a set of time series can be partitioned into subsets with subset-specific common features only (see e.g. Engle and Kozicki, 1993). While complete separation provides conditions under which common trends and common cycles can be extracted on a sub-system basis, these conditions do not yet ensure that extraction based on these (marginal) sub-systems will be fully efficient. In this paper we provide a set of sufficient conditions under which valid inference can be conducted in marginal separated systems.

The paper is organized as follows. Section 2 defines common features in the context of a cointegrated VAR(p). Section 3 recalls the notion of separation in cointegration, shows the implications that separation in cointegration has in terms of common features and reversely. A similar definition of separation for the common features matrix is introduced. The conditions under which maximum likelihood (ML) inference within a separated sub-system is as efficient as ML inference based on the joint system are analyzed in Section 4. Section 5 presents a set of likelihood ratio type statistics useful for testing particular separation hypotheses. In Section 6, an empirical analysis shows the usefulness of the concepts discussed in this study

when we analyze the presence of distinct common factors in consumption functions for the US and Canada. Section 7 concludes.

2 VAR Models with Cointegration and Common Features

Consider a Gaussian Vector Autoregression of finite order p (VAR(p)) model for an n-dimensional *I(1)* vector time series $\{x_t\}$:

$$x_t = \sum_{i=1}^{p} \Pi_i x_{t-i} + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{1}$$

with fixed initial values of $x_{-p+1}, ..., x_0$ and where ε_t is a n-dimensional homoskedastic Gaussian mean innovation process relative to $X_{t-1} = \{x_{t-1}, x_{t-2}, \ldots, x_0\}$ with nonsingular covariance matrix Ω. Let L denote the lag operator and define $\Pi(L) = I_n - \sum_{i=1}^{p} \Pi_i L^i$. We make the following assumption

Assumption 1. (Cointegration): *In the VAR model (1), we assume that*

1. $\text{rank}(\Pi(1)) = r, 0 < r < n$, *so that* $\Pi(1)$ *can be expressed as* $\Pi(1) = -\alpha\beta'$, *with* α *and* β *both* $(n \times r)$ *matrices of full column rank* r*;*

2. *the characteristic equation* $|\Pi(\xi)| = 0$ *has* $n - r$ *roots equal to* 1 *and all other roots outside the unit circle.*

Assumption 1 implies (see Johansen, 1995) that the process x_t is cointegrated of order (1,1). The columns of β span the space of cointegrating vectors, and the elements of α are the corresponding adjustment coefficients or factor loadings. Decomposing the matrix lag polynomial $\Pi(L) = \Pi(1)L + \Gamma^*(L)(1-L)$, with $\Gamma^*(L) = I_n - \sum_{j=1}^{p-1} \Gamma_j L^j = I_n - \Gamma(L)$, $\Gamma_j = -\sum_{k=j+1}^{p} \Pi_k$ $(j = 1, \ldots, p-1)$ and defining $\Delta = (1 - L)$, we obtain the vector error-correction model (VECM):

$$\Delta x_t = \alpha\beta' x_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta x_{t-i} + \varepsilon_t, \qquad t = 1, \ldots, T. \tag{2}$$

Throughout this paper we will also assume that p is known.

Serial correlation common features (SCCF, see Engle and Kozicki, 1993) hold for the

VECM (2), if there exists a $(n \times s)$ matrix $\widetilde{\beta}$, called common feature or cofeature matrix,[1] whose columns span the cofeature space, such that $\widetilde{\beta}'\Delta x_t = \widetilde{\beta}'\varepsilon_t$ is a s-dimensional vector mean innovation process with respect to the information available at time t.

Consequently, serial correlation common features hold if the cofeature matrix $\widetilde{\beta}'$ satisfies the following two conditions:

$$\textbf{Assumption 2:} \qquad \widetilde{\beta}'\Gamma_j = 0_{(s\times n)}, \qquad j = 1 \ldots p-1 \tag{3}$$

$$\textbf{Assumption 3:} \qquad \widetilde{\beta}'\Pi(1) = -\widetilde{\beta}'\alpha\beta' = 0_{(s\times n)} \tag{4}$$

Assumption 2 implies that $\widetilde{\beta}'$ must lie in the intersection of the left null spaces of the matrices describing the short-run dynamics Γ_j, $j = 1 \ldots p-1$. Given that $\Gamma_j = -\sum_{k=j+1}^{p} \Pi_k$, $j = 1, \ldots, p-1$ and $\Pi(1) = I_n - \sum_{j=1}^{p} \Gamma_j$, Assumption 3 implies that $\widetilde{\beta}'(I_n - \Pi_1) = 0_{(s\times n)}$, e.g. Π_1 must have eigenvalues equal to one with multiplicity equal to s and the corresponding eigenvectors must lie in the intersection of the left null spaces of the Γ_j matrices. Cointegrated VAR models satisfying both Assumptions 2 and 3 are considered in detail in Vahid and Engle (1993). To distinguish between models that satisfy either both Assumptions 2 and 3 or Assumption 2 only, Hecq, Palm and Urbain (2001) introduce the two following definitions:

Definition 1 (Strong Form Reduced Rank Structure) *If in addition to Assumption 1 (cointegration) both Assumptions 2 and 3 hold then implied reduced rank structure of the VECM (2), will be labelled a strong form reduced rank structure (SF). Under SF, there exists a $(n \times s)$ matrix $\widetilde{\beta}$, whose columns span the cofeature space, such that $\widetilde{\beta}'\Delta x_t = \widetilde{\beta}'\varepsilon_t$ is a $s-$dimensional vector mean innovation process with respect to X_{t-1}.*

Definition 2 (Weak Form Reduced Rank Structure) *If in addition to Assumption 1 (cointegration) only Assumption 2 holds then implied reduced rank structure of the VECM (2), will be labelled a weak form reduced rank structure (WF). Under WF, there exists a $(n \times s)$ matrix $\widetilde{\beta}$, whose columns span the cofeature space, such that $\widetilde{\beta}'(\Delta x_t - \alpha\beta' x_{t-1}) = \widetilde{\beta}'\varepsilon_t$ is a $s-$dimensional vector mean innovation process with respect to X_{t-1}.*

The implications of these two classes of models in terms of the nature of the dynamic common factors are discussed in more detail in Hecq, Palm and Urbain (2000a, 2001) where

[1] Notice that we will use indifferently the terms common features or cofeature in the sequel of this paper.

inferential issues are investigated and a mixed form is also proposed. At this stage it is already useful to note an important difference between SF and WF. In the latter case, the possible number of cofeature vectors s may be greater than $n-r$ but has to remain $\leq n-1$ and the corresponding $n-s$ common factors consist of linear combinations of the lagged first differences only. Notice that we could easily extend these definitions to the case where only part of the short-run components of the models are annihilated when premultiplied by $\widetilde{\beta}'$. This type of reduced rank structures has been previously mentioned by a.o. Ahn and Reinsel (1988) for stationary processes, Tiao and Tsay (1989) for vector-autoregressive moving average models and by Reinsel and Ahn (1992) and Ahn (1997) for cointegrated VAR processes. Remark that the WF restrictions are generally not invariant to alternative vector error-correction representations such as that where x_{t-p} appears in levels instead of x_{t-1}. The implications of the lack of invariance are discussed in more details in Hecq et al. (2001).

3 Separation in Cointegration and in Cofeature

3.1 Definitions

The notion of separation put forward by Konishi (1993), Konishi and Granger (1993) helps to identify sub-systems that could be investigated independently from each other, reducing thereby substantially the complexity and size of the modeling problem at hand.

Definition 3 *Consider the n dimensional cointegrated vector time series $x_t = (x'_{1t}, x'_{2t})'$ generated by the VECM (2), with x_{1t} and x_{2t} are distinct sub-vectors of dimension $n_1 \times 1$ and $n_2 \times 1$ respectively with $n_1 + n_2 = n$. If the matrix $-\Pi(1) = \alpha\beta'$ can be factorized such that the matrix of cointegrating vectors β' is block-diagonal:*

$$\beta' = \begin{pmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{pmatrix}',$$

where the sub-matrices β_{11} and β_{22} are respectively $n_1 \times r_1$ and $n_2 \times r_2$ full column rank matrices, with $0 < r_i < n_i$, $i = 1, 2$, $r = r_1 + r_2$, then the system is said to be subject to separate cointegration.

Under separate cointegration,[2] the VECM (2) specialises to become:

$$\begin{pmatrix} \Delta x_{1t} \\ \Delta x_{2t} \end{pmatrix} = \begin{pmatrix} \Gamma_{11}(L) & \Gamma_{12}(L) \\ \Gamma_{21}(L) & \Gamma_{22}(L) \end{pmatrix} \begin{pmatrix} \Delta x_{1t} \\ \Delta x_{2t} \end{pmatrix} + \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \begin{pmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{pmatrix}' \begin{pmatrix} x_{1t-1} \\ x_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \tag{5}$$

with the partitioning of the matrices being conformable with that of x_t.

Granger and Haldrup (1997) further refined the idea of separation by considering other parameter restrictions that may arise on the remaining matrices of the systems. The following definitions can be introduced for a VECM under separate cointegration.

Definition 4 (Complete Separation in the long-run) *If* $\alpha_{12} = 0, \alpha_{21} = 0, \Gamma_{12}(1) = 0, \Gamma_{21}(1) = 0$ *then the system (5) is said to be completely separated in the long-run.*

For the case where the system is not completely separated in the long-run, Granger and Haldrup (1997) introduce two different *mutually exclusive* forms of partial separation that depend on the structure of the parameter matrices of (5):

Definition 5 (Partial separation of Type A) *If* $\alpha_{12} \neq 0$ *and/or* $\alpha_{21} \neq 0$ *but* $\Gamma_{12}(1) = 0, \Gamma_{21}(1) = 0$*, then the system (5) is said to be (partially) separated of Type A.*

Definition 6 (Partial separation of Type B) *If* $\alpha_{12} = 0, \alpha_{21} = 0$ *but* $\Gamma_{12}(1) \neq 0$ *and/or* $\Gamma_{21}(1) \neq 0$ *then the system (5) is said to be (partially) separated of Type B.*

Note that Type B partial separation in (5) implies block-diagonality of $\sum_{i=1}^{p} \Pi_i$ due to the block-diagonality of $\alpha\beta' = -(I - \sum_{i=1}^{p} \Pi_i)$. This does however not imply block-diagonality of the Γ_j's since $\Gamma_j = -\sum_{k=j+1}^{p} \Pi_k$. In the sequel our common feature analysis will require additional conditions of short-run Granger non-causality $\Gamma_{12}(L) = 0, \Gamma_{21}(L) = 0$ so that we also define:

Definition 7 (Partial separation of Type C) *If* $\alpha_{12} \neq 0$ *and/or* $\alpha_{21} \neq 0$ *but* $\Gamma_{12}(L) \equiv 0, \Gamma_{21}(L) \equiv 0$ *then the system (5) is said to be (partially) separated of Type C.*

[2]Notice that by considering separate cointegration we exclude cases where $r = n - 1$ or $r = 1$.

Note that in the sequel we shall delete the word partial and use the wording separation of Type A, B or C respectively. We call a system completely separated if and only if $\beta_{12} = 0$, $\beta_{21} = 0, \alpha_{12} = 0$, $\alpha_{21} = 0$, $\Gamma_{12}(L) \equiv 0, \Gamma_{21}(L) \equiv 0$. If in addition to complete separation, $\Omega_{12} = 0$ the subsystems are independent.

Once we allow for the existence of reduced ranks of the short-run dynamics matrices, we may in analogy to separation in the long-run define the concept of serial correlation common features separation or cofeature separation:[3]

Definition 8 (Cofeature Separation) *Consider a $s \times n$ cofeature matrix $\widetilde{\beta}'$ that satisfies either $\widetilde{\beta}'\Delta x_t = \widetilde{\beta}'\varepsilon_t$ or $\widetilde{\beta}'(\Delta x_t - \alpha\beta' x_{t-1}) = \widetilde{\beta}'\varepsilon_t$ where $\widetilde{\beta}'\varepsilon_t$ is a $s-$dimensional vector mean innovation process with respect to X_{t-1}. For the partitioning of x_t in x_{1t} and x_{2t}, we say that the vector process x_t has separate SF cofeatures (or separate WF cofeatures) if the cofeature matrix $\widetilde{\beta}'$, partitioned conformably to that of x_t, can be written as*

$$\widetilde{\beta}' = \begin{pmatrix} \widetilde{\beta}_{11} & 0 \\ 0 & \widetilde{\beta}_{22} \end{pmatrix}',$$

with $\widetilde{\beta}'_{ii}$ being full row rank $(s_i \times n_i)$ matrices, $0 < s_i \leq n_i - r_i$ (SF) or $0 < s_i \leq n_i - 1$ (WF), $i = 1, 2$, $s = s_1 + s_2$.

Notice that separation in common features as defined above only requires α and $\Gamma(L)$ to have a particular reduced rank structure. Remark that the rank conditions we impose of the $\widetilde{\beta}_{ii}$ matrices again imply the presence of at least one common features vector for each sub-system, so that separation in common features can only occur when $s = s_1 + s_2$ satisfies $2 \leq s \leq n - r$, with $n - r = (n_1 - r_1) + (n_2 - r_2)$ (SF), or when $2 \leq s \leq n - 2$, (WF). Since one of the goals of this paper is to find conditions under which distinct sub-systems can be analyzed separately for the purpose of common trend-common cycle decompositions, we will consider separation in cointegration as a maintained assumption in the sequel, i.e. we shall assume that the matrix of cointegration β is block-diagonal.

[3] Alternative forms of cofeature separation could be considered but they are not investigated in this paper. For instance, when $\widetilde{\beta}' = [\widetilde{\beta}_1 \;\; 0]$, only the first subsystem of the VECM is subject to cofeatures. When $\widetilde{\beta}'$ is lower block-triangular, some cofeatures are present in the first subsystem, whereas other cofeatures are present in the complete system.

3.2 Relation between cointegration separation and cofeature separation

Under separation in common feature and under a WF reduced rank structure, the VECM (2) satisfies:

$$\begin{pmatrix} \widetilde{\beta}_{11} & 0 \\ 0 & \widetilde{\beta}_{22} \end{pmatrix}' \begin{pmatrix} \Gamma_{11}(L) & \Gamma_{12}(L) \\ \Gamma_{21}(L) & \Gamma_{22}(L) \end{pmatrix} = 0, \tag{6}$$

while under a SF assumption we have in addition to (6)

$$\begin{pmatrix} \widetilde{\beta}_{11} & 0 \\ 0 & \widetilde{\beta}_{22} \end{pmatrix}' \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} = 0. \tag{7}$$

These sets of restrictions (6)-(7) form the basis of the analysis in this section. Note that the conditions (6) and (7) have to hold for separation in common features whether or not there is separation in cointegration present in (2), i.e. whether or not β is block-diagonal. In order to clarify the implications of cofeature separation in terms of restrictions on the matrices of the VECM (2) and the relation between cofeature and cointegration separation, we will first consider a simple cointegrated VAR(1) since this model is actually useful to illustrate the important characteristics we are interested in:

$$\Delta x_t = \alpha^* \beta^{*\prime} x_{t-1} + \varepsilon_t. \tag{8}$$

We parametrize it in terms of the matrices α^* and β^* which are of the order $n \times r$ and $r \times n$ respectively and of rank r. We first impose the restrictions under separation in cointegration. Next, we impose the restrictions from separation in SF common features. Provided that α^* is unrestricted, separation in cointegration implies (see also Hansen and Johansen, 1998) that the cointegration relations can be expressed as linear combinations of $\beta_{11}' x_{1t-1}$and $\beta_{22}' x_{2t-1}$, that is there exist $(r \times r_i), i = 1, 2$ and $r_1 + r_2 = r$ full column rank matrices $F_i, i = 1, 2$ such that $\alpha^* \beta^{*\prime} x_{t-1} = \alpha^* [F_1 \beta_{11}' x_{1t-1} + F_2 \beta_{22}' x_{2t-1}] = \alpha \beta' x_{t-1}$ with

$$\beta^{*\prime} = (F_1, F_2) \begin{pmatrix} \beta_{11} & 0 \\ 0 & \beta_{22} \end{pmatrix}' = F\beta', \qquad \alpha = \alpha^* F,$$

where $F = (F_1 : F_2)$ is a $r \times r$ full rank matrix. The notation $\beta^{*\prime}$ is used to stress that under separation in cointegration β^* needs not to be block-diagonal as long as it can be factorized into the product of F times a normalized block-diagonal cointegrating matrix β'. In (8), there exist by definition $n - r$ serial correlation common feature vectors (or SF common features) given by the rows of $\alpha'_\perp$ where $\alpha'_\perp$ is the orthogonal complement[4] of α - see for example Vahid and Engle (1993). By definition, SF common feature separation for $s = n - r$ requires block-diagonality of $\alpha'_\perp$:

$$\alpha'_\perp = \begin{pmatrix} \alpha_{11\perp} & 0 \\ 0 & \alpha_{22\perp} \end{pmatrix}'.$$

with $\alpha'_{ii\perp}$ being a $(n_i - r_i) \times n_i$ matrix of rank $n_i - r_i$. Since $rank(\alpha) = r$ by Assumption 1, for SF common features separation with $s = n - r$ to occur, it is sufficient that one can find a full rank $r \times r$ square matrix D such that the loading matrix α can be written as the product of a $(n \times r)$ block-diagonal matrix α^0 of rank r times D :

$$\alpha = \alpha^0 D \text{ with } \alpha^0 = \begin{pmatrix} \alpha^0_{11} & 0 \\ 0 & \alpha^0_{22} \end{pmatrix}. \tag{9}$$

Consequently, under separation in common features for $s = n - r$ and in cointegration, the VECM (8) can be rewritten as

$$\Delta x_t = \alpha^* \beta^{*\prime} x_{t-1} + \varepsilon_t = \alpha\beta' x_{t-1} + \varepsilon_t = \alpha^0 D\beta' x_{t-1} + \varepsilon_t, \tag{10}$$

where the matrices β, α^0 and $\alpha_\perp$ are block-diagonal and D is an $(r \times r)$ matrix of full rank. Notice that α itself need not be block-diagonal. Also notice that for SF common feature separation for $s = n - r$ implies SF common features separation for $s < n - r$. The matrix $-\Pi(1) = \alpha\beta' = \alpha^0 D\beta'$ will not be block-diagonal unless D is block-diagonal. If complete separation in the long-run part is assumed $-\Pi(1)$ is block-diagonal and hence D has to be block-diagonal too.

For the VAR(1), when D is block-diagonal and therefore α is block-diagonal, Type B separation implies SF common feature separation with $s = n - r$ common features and

[4]We denote the orthogonal complement of any $n \times s$-dimensional matrix B, with $n > s$ and rank$(B) = s$, by the $n \times (n - s)$ matrix $B_\perp$ such that $B' B_\perp = 0$ with rank$(B_\perp)$= $n - s$ and rank$(B : B_\perp)$= n.

$s_i = n_i - r_i$ common features per block. The reverse, that is common features separation with $s = n - r$ common features and $s_i = n_i - r_i$ common features per block implies separation of Type B, if D is block-diagonal. Results for VAR(p) are presented in the Propositions 1 and 2.

Proposition 1 *In the VECM (2) under both separate cointegration and a strong form reduced rank structure,*

(i) For $p > 1$, $s_i < n_i - r_i$, Type B separation does not imply common feature separation. For $s_i = n_i - r_i$, $i = 1, 2$, or $p = 1$, Type B separation implies common feature separation.

(ii) For $p > 1$, Type C separation does not imply common feature separation for $s \leq \min(n - r, n - s_j^*; j = 1, 2, \ldots, p-1)$, where s_j^* is the rank of Γ_j except if $s = n - s_{j^*}^* = argmin(n - s_j^*; j = 1, 2, \ldots, p-1)$.

(iii) For $p > 1$, Type A separation does not imply common feature separation.

(iv) For $p > 1$, complete separation does not imply common feature separation for $s < \min(n - r, n - s_j^*; j = 1, 2, \ldots, p-1)$ where s_j^* is the rank of Γ_j. For $s = argmin(n - r, n - s_j^*; j = 1, 2, \ldots, p-1)$, complete separation implies common feature separation.

Proof:

(i) Under SF, there exists an $(s \times n)$ matrix $\widetilde{\beta}'$ of rank s, $2 \leq s \leq n-2$ such that $\widetilde{\beta}'\alpha = 0, \widetilde{\beta}'\Gamma_j = 0; j = 1, 2, \ldots, p-1$. As the rank of α is r, there exists an $(n-r) \times n$ matrix D of rank$(n-r)$ such that $D'\alpha = 0$. Under separation of Type B, α is block-diagonal. We therefore have

$$D'\alpha = 0 \Leftrightarrow \begin{pmatrix} D'_{11}\alpha_{11} & D'_{12}\alpha_{22} \\ D'_{21}\alpha_{11} & D'_{22}\alpha_{22} \end{pmatrix} = 0$$

which the partitioning of D being commensurate with that of α. As α_{11} has rank r_1, there are $n_1 - r_1$ independent linear relationships

$$\underbrace{F'_{11}\alpha_{11}}_{(n_1-r_1)\times n_1} = 0 \Leftrightarrow \begin{pmatrix} D'_{11} \\ D'_{21} \end{pmatrix} \alpha_{11} = 0$$

(similarly for α_{22}). Therefore D can be expressed as $D = A \begin{pmatrix} F'_{11} & 0 \\ 0 & F'_{22} \end{pmatrix} = AF'$, with A being a $(n-r) \times (n-r)$ matrix of rank $(n-r)$, and $F'\alpha = 0$. The matrix $\widetilde{\beta}'$ which annihilates α can be expressed in terms of linear combinations of D, i.e. $\widetilde{\beta}' = GD = GAF'$, G being an $s \times (n-r)$ matrix of rank s. The same arguments can be applied to $\widetilde{\beta}'\Gamma_j = 0, j = 1, 2, \ldots, p-1$. The matrix $\widetilde{\beta}'$ is not block-diagonal, except in the case where $s_i = n_i - r_i$ and $s = n - r$, where GA is an $(n-r) \times (n-r)$ full rank matrix, so that we can take F' as the cofeature matrix, which is block diagonal. $\square$

(ii) Under Type C separation (note that it is only defined when $p > 1$) $\Gamma(L)$ is block-diagonal,

$$\Gamma_j = \begin{pmatrix} \Gamma_{j\,11} & 0 \\ 0 & \Gamma_{j\,22} \end{pmatrix}$$

with $\Gamma_{j\,ii}$ being an $n_i \times n_i$ matrix of rank $s^*_{j(i)}$, $i = 1, 2$. Under a SF reduced rank structure, there exists a $(s \times n)$ matrix $\widetilde{\beta}'$ of rank s, $s < \min(n-r, n-s^*_j; j = 1, 2, \ldots, p-1)$, s^*_j being the rank of Γ_j, such that $\widetilde{\beta}'\alpha = 0$ and $\widetilde{\beta}'\Gamma_j = 0; j = 1, 2, \ldots, p-1$. As Γ_j has rank $s^*_j = s^*_{j(1)} + s^*_{j(2)}$, there exists an $(n - s^*_j) \times n$ matrix D_j with rank $n - s^*_j$, such that $D'_j\Gamma_j = 0; j = 1, 2, \ldots, p-1$. As Γ_j is block-diagonal, for $\Gamma_{j\,ii}, i = 1, 2$, there exists a $(n_i - s^*_{j(i)}) \times n_i$ matrix $F'_{j\,ii}$ of rank $n_i - s*_{j(i)}$, such that $F'_{j\,ii}\Gamma_{j\,ii} = 0$.

Therefore, D'_j can be expressed as $D'_j = A_j F'_j$ with

$$\underbrace{F'_j}_{(n-s^*_j)\times n} = \begin{pmatrix} F'_{j\,11} & 0 \\ 0 & F'_{j\,22} \end{pmatrix},$$

where A_j is an $(n-s^*_j) \times (n-s^*_j)$ matrix of rank $n - s^*_j$. We therefore have $D'_j\Gamma_j = 0 \Leftrightarrow A_j F'_j \Gamma_j = 0$ with F'_j being block-diagonal. As α has rank r, there exists an $(n-r) \times n$ matrix D_0 of rank $n-r$ such that $D'_0\alpha = 0$. Any set of linearly independent row vectors annihilating α and Γ_j, $j = 1, 2, \ldots, p-1$, must be linear combinations of the matrices D_j, $j = 0, 1, \ldots, p-1$ respectively. Therefore, there exists $s \times (n - s^*_j)$ matrices G_j of rank s and a $s \times (n-r)$ matrix G_0 of rank s as well such that the $(s \times n)$ cofeature

matrix $\widetilde{\beta}'$ can be expressed as

$$\widetilde{\beta}' = G_0 D_0', \widetilde{\beta}' = G_j A_j F_j', \ j = 1, 2, \ldots, p-1.$$

The matrix $\widetilde{\beta}'$ will in general not be block-diagonal except when for at least one $j = 1, 2, \ldots, p-1$, denoted by j^*, $s = n - s_{j^*}^* = \text{argmin}(n-r, n-s_j^*; j = 1, 2, \ldots, p-1)$, so that $\widetilde{\beta}'$ premultiplied by the inverse of $G_{j^*} A_{j*}$ yields the block-diagonal matrix F_{j^*}'. If j^* is not unique, the resulting block-diagonal matrix $\widetilde{\beta}'$ may not be unique. $\square$

(iii) Type A separation is only relevant for a VAR(p); with $p > 1$. It arises if $\Gamma(1)$ is block-diagonal. Under SF (and under WF) reduced rank structure, we have $\widetilde{\beta}'\Gamma_j = 0, j = 1, 2, \ldots, p-1$, implying that $\widetilde{\beta}'\Gamma(1) = 0$. Under block-diagonality of $\Gamma(1)$, this is equivalent to

$$\begin{pmatrix} \widetilde{\beta}_{11}'\Gamma_{11}(1) & \widetilde{\beta}_{12}'\Gamma_{22}(1) \\ \widetilde{\beta}_{21}'\Gamma_{11}(1) & \widetilde{\beta}_{22}'\Gamma_{22}(1) \end{pmatrix} = 0.$$

As there are no restrictions on the ranks of the $\Gamma_{ii}(1)$, from

$$\underset{(s \times n_1)(n_1 \times n_1)}{\begin{pmatrix} \widetilde{\beta}_{11}' \\ \widetilde{\beta}_{21}' \end{pmatrix} \Gamma_{11}(1) = 0} \qquad \underset{(s \times n_2)(n_2 \times n_2)}{\begin{pmatrix} \widetilde{\beta}_{12}' \\ \widetilde{\beta}_{22}' \end{pmatrix} \Gamma_{22}(1) = 0,}$$

we can not infer that $\widetilde{\beta}_{21}'$ (and $\widetilde{\beta}_{12}'$) have to be zero. $\square$

(iv) The proof is similar to that of statement (ii), except that the block-diagonality of α under complete separation, can imply common feature separation when $s = \text{argmin}(n - r, n - s_j^*; j = 1, 2, \ldots, p-1)$. $\square$

Proposition 2 *In the VECM (2), with $p > 1$, under both separate cointegration and weak form reduced rank structure,*

(i) Type B separation does not imply separation in common features.

(ii) Type C separation (and complete separation) does not imply common feature separation for $s < \min(n - s_j^*; j = 1, 2; \ldots, p-1)$ where s_j^* is the rank of Γ_j. For $s = n - s_{j^*}^* = \text{argmin}(n - s_j^*; j = 1, 2; \ldots, p-1)$, Type C separation (complete separation) implies common feature separation.

(iii) Type A separation does not imply separation in common features.

Proof:

(i) The statement is obvious, as Type B separation implies restrictions on α but not on $\Gamma_j, j = 1, 2, \ldots, p-1$. $\square$

(ii) The proof follows the same lines as the proof of Proposition 1 (ii). $\square$

(iii) The proof is similar to that of Proposition 1 (iii). $\square$

Note, as pointed out when discussing the example of a VAR(1) above, that under the conditions of the Propositions 1 and 2, separation in common features is compatible with but it does not imply separation of Type A,B or C, but only restrictions on some parameter matrices.

It is important to observe that the above results hinge on a crucial assumption that the separation in cointegration and cofeature hold for the same partition of $x_t = (x'_{1t}, x'_{2t})'$. This assumption may be realistic in dynamic panel data. Of course, one can imagine cases where separation holds in cointegration for a given partitioning of x_t, while cofeature separation holds for another partitioning of the same x_t vector. An example would be an analysis of 4 different times series related to real consumption and real income for two countries 1 and 2 (see the empirical applications below) where (i) cointegration separation holds so that consumption and income cointegrate only within a country, but (ii) separate cofeature relationships exist that relate the two changes in consumption or the changes in real income only, reflecting for instance a common business cycle. It is easily seen that in this situation, the results presented in the Propositions just mentioned no longer hold. In particular, under a WF reduced rank structure separation in cofeature is then only compatible with Type B separation in cointegration while separate cofeature in a VECM with a SF reduced rank structure is incompatible with cointegration separation of any form.

3.3 Separation and P-T decomposition

While the preceding subsection clarified the relationships between separation in cointegration and separation in common features we now briefly reconsider the issue from a Permanent-

Transitory (P-T) decomposition point of view and summarize the conditions under which common trends and common cycles may be extracted from separated sub-systems. Although such P-T decompositions are not unique, most of them take the form:

$$x_t = Gx_t^P + Ax_t^T, \tag{11}$$

where G and A are loading matrices, x_t^P is the common stochastic trend or permanent components and x_t^T the (possibly common) transitory or cyclical component.

Granger and Haldrup (1997) focus on the Gonzalo and Granger (1995) [GG] P-T decomposition which, in the absence of common features, is given by

$$x_t = \underbrace{\beta_\perp(\alpha_\perp'\beta_\perp)^{-1}}_{G}\underbrace{\alpha_\perp' x_t}_{x_t^P} + \underbrace{\alpha(\beta'\alpha)^{-1}}_{A}\underbrace{\beta' x_t}_{x_t^T}. \tag{12}$$

This decomposition requires the matrix $(\beta \vdots \alpha_\perp)$ to be invertible. While this always holds in a VAR(1) cointegrated model, it generally does carry over to higher order cointegrated VAR systems (see for example Exercise 4.3 in Johansen, 1995). In this decomposition the common trends are given by $\alpha_\perp^{'} x_t$ while the transitory component is simply $\beta' x_t$.

Alternative decompositions exist. An appealing alternative is the Beveridge-Nelson-Stock-Watson [BNSW] decomposition [see King, Plosser, Stock and Watson, 1991; Proietti, 1997; Hecq et al, 2000a] that is valid under less restrictive conditions. As shown by Proietti (1997), Hecq et al. (2000a), this decomposition takes the form (11) with

$$Gx_t^P = (I-P)(\Gamma^*(1)-\alpha\beta^{'})^{-1}\alpha_\perp(\alpha_\perp^{'}\alpha_\perp)^{-1}\alpha_\perp^{'}\Gamma^*(L)x_t, \tag{13}$$

$$Ax_t^T = Px_t - (I-P)(\Gamma^*(1)-\alpha\beta^{'})^{-1}\Gamma^{**}(L)\Delta x_t, \tag{14}$$

with

$$P = (\Gamma^*(1)-\alpha\beta^{'})^{-1}\alpha[\beta^{'}(\Gamma^*(1)-\alpha\beta^{'})^{-1}\alpha]^{-1}\beta^{'}$$

being a $(n \times n)$ matrix satisfying a number of interesting properties given in Proietti (1997), Hecq et al. (2000a). The matrix polynomials $\Gamma^{**}(L)$ and the matrix $\Gamma^*(1)$ are obtained from the decomposition of $\Gamma^*(L) = \Gamma^*(1) + (1-L)\Gamma^{**}(L)$, where $\Gamma^*(L) = I_n - \sum_{j=1}^{p-1}\Gamma_j L^j =$

$I_n - \Gamma(L)$, $\Gamma_j = -\sum_{k=j+1}^{p} \Pi_k$ are given in (2) and $\Gamma^{**}(L) = \sum_{j=0}^{p-2} \Gamma_j^{**} L^j$, $\Gamma_j^{**} = \sum_{i=j+1}^{p-1} \Gamma_j^*$.

In this decomposition, the permanent part is taken as a multivariate random walk process, the common trends are thus given by $\alpha_\perp^{'} \Gamma^*(L) x_t$ (see *inter alia* Johansen, 1995) and the corresponding loadings are $(I - P)(\Gamma^*(1) - \alpha\beta^{'})^{-1} \alpha_\perp (\alpha_\perp^{'} \alpha_\perp)^{-1}$. The cyclical component Ax_t^T has two components. A first part of the transitory component which is linked to the process of adjustment to equilibrium has the cointegrating relationships as cycle generators $Px_t = (\Gamma^*(1) - \alpha\beta^{'})^{-1} \alpha [\beta^{'} (\Gamma^*(1) - \alpha\beta^{'})^{-1} \alpha]^{-1} \beta^{'} x_t$. A second component linked to the short-run fluctuations in a stricter sense is given by $-(I - P)(\Gamma^*(1) - \alpha\beta^{'})^{-1} \Gamma^{**}(L) \Delta x_t$. Hecq et al. (2000a) show how on can further refine these decompositions in the presence of cofeatures, but the details can be omitted here since the presentation above is sufficient to discuss the conditions under which common trends and common cycles may be extracted from separated sub-systems. Note also that, as shown by Proietti (1997) and Hecq et al. (2000a), a decomposition in the sense of Gonzalo-Granger (1995) is obtained by adding the second part of the cyclical component to the permanent component, which yields $x_t = (I - P)x_t + Px_t$. The common trends are then given by $\alpha_\perp^{'} \Gamma^*(1) x_t$ and not by $\alpha_\perp^{'} x_t$ as in Gonzalo-Granger (1995), for details see Hecq et al (2000a). Once we introduce SF common feature restrictions with $s + r = n$, both expressions coincide since the columns of $\widetilde{\beta}'$ and $\alpha_\perp$ span the same space and hence $\widetilde{\beta}' \Gamma^*(1) x_t = \widetilde{\beta}' x_t = \alpha_\perp' x_t$. Consequently, the BNSW decomposition (13)-(14) and Gonzalo-Granger (1995) decompositions coincide under SF when $s + r = n$. Both then also corresponds to the decomposition proposed in the presence of common features by Vahid and Engle (1993). For the cases where $s + r \neq n$, either under SF or WF, one has to rely on (13)-(14) which, although related to the GG decomposition, exists under weaker conditions.

From these decompositions and the results mentioned above a number of conclusions can be drawn:

- Complete separation in cointegration as defined by Granger and Haldrup (1997), implying that $I - P$ is block-diagonal, is a sufficient condition for sub-group common stochastic trend extraction based of GG. In the presence of common features, complete separation is also sufficient for sub-group common stochastic trend extraction based on BNSW if and only if the common features are of the strong form with $s + r = n$.

- Common features separation, while necessary to ensure that the common cycles for each

sub-group are linearly independent, is not sufficient to sustain sub-group extraction of common trends and common cycles (BNSW), [see the Section 6 for an illustration].

- In the presence of SF common features with $s + r < n$, necessary conditions for sub-group extraction of common trends and common cycles based on BNSW are given by the conditions of Granger and Haldrup (1997), (e.g. $\beta_{12} = 0, \beta_{21} = 0, \alpha_{21} = 0, \alpha_{12} = 0, \Gamma_{12}(1) = 0$ and $\Gamma_{21}(1) = 0$), to which we have to add the further condition of Granger Noncausality in both directions in order to achieve the block-diagonality of $\Gamma^*(L)$ and $\Gamma^{**}(L)$, -see (13)-(14) .

- In the presence of WF common features, necessary conditions for sub-group extraction of common trends and common cycles based on BNSW are again cointegration separation and Granger Noncausality in both directions (Type C).

To briefly summarize, when common features are present, common features separation ensures that the common stochastic cycles are group-specific while cointegration separation ensures that the common trends are group-specific but to ensure that these only involve variables from the individual sub-groups, the condition of Granger Noncausality in both directions is the additional requirement which constitutes a rather strong requirement from a practical point of view, at least much stronger than the one derived in Granger and Haldrup (1997). While common feature separation implies the existence of independent common cycles between the two separated sub-groups it does however generally not imply that these can be derived independently from each other unless two-direction Granger Noncausality holds.

4 Separated cointegrated versus partial systems

In the previous section we have pointed out that common trends and common cycles can only be extracted from a subsystem under rather stringent conditions on the dynamics of the complete system. To compute these components, which, when based on a Gonzalo-Granger decomposition or on a Beveridge-Nelson-Stock-Watson decomposition, are expressed explicitly in terms of observable variables, consistent and preferably efficient estimates of the different parameter matrices entering both components are required. Hence, while at the model representation level the aforementioned conditions appear to be useful for the

extraction on a group by group basis, they do not yet enable us to assess whether this extraction can be done efficiently. The efficiency issue is obviously closely related to (weak) exogeneity conditions in this extended class of cointegrated models. Since the concept of weak exogeneity is specific to well-defined parameters of interest and is basically a statement about the parameter space of a given statistical model, see Engle, Hendry and Richard (1983), the standard conditions for weak exogeneity in cointegrated systems need to be adapted to cover VAR models with separate cointegration and common feature restrictions to make sure in these cases that the analysis of conditional subsystems can be carried out without loss of information.

In their discussion of separation in cointegrated systems, Granger and Haldrup (1997, p.453) claim that the block-diagonality of the loading matrix α implied by complete or Type B separation is equivalent to weak exogeneity of x_{1t} (resp. x_{2t}) w.r.t. β'_{22} (resp. β'_{11}), [see *inter alia* Johansen, 1992, 1995, Urbain, 1992], and hence that efficient maximum likelihood (ML) inference based on the partial systems can be achieved under complete (or Type B) separation in cointegration. This claim overlooks the model specific nature of weak exogeneity conditions. As pointed out in Hansen and Johansen (1998), their statement is not true unless we have block-diagonality of the covariance matrix of the errors of the VECM (2). This naturally also carries over to cointegrated systems with both separate common features and complete separation in cointegration. Indeed, under block-diagonality of the disturbance covariance matrix, complete separation implies that the likelihood function factorizes into the product of the likelihood functions of the marginal processes. This is naturally a very stringent condition.

Let us illustrate this issue with a simple VAR(2) model, results for higher model being easily derived from this simple case. We will first derive sufficient conditions for weak exogeneity in the case of a VECM with a WF reduced rank structure but without assuming any form of separation. This will enable us to derive sufficient conditions for weak exogeneity for the cases of separation of various forms. Finally, conditions under which the conditional model reduces to a marginal model will be mentioned.

4.1 Weak exogeneity in cointegrated VAR with common features

We first consider the case of cointegrated systems with r cointegration relations and s common features restrictions and partition $x_t = (x'_{1t}, x'_{2t})'$ where x_{it} are $(n_i \times 1), i = 1, 2$ with $n_1 + n_2 = n$. We assume that the partition has been chosen such that $n_1 \geq \max(r, n-s)$. This dimensionality assumption is required to ensure that the coefficients of the cointegrating and common feature relationships can be identified from the partial systems only. For the sake of simplicity of the notation and without loss of generality,[5] we assume that order p of the VAR (1) is 2 so that in VECM form it reads as :

$$\begin{pmatrix} \Delta x_{1t} \\ \Delta x_{2t} \end{pmatrix} = \begin{pmatrix} \Gamma_1 \\ \Gamma_2 \end{pmatrix} \Delta x_{t-1} + \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \beta' x_{t-1} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \tag{15}$$

with

$$\begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \Omega_{11} & \Omega_{12} \\ \Omega_{21} & \Omega_{22} \end{pmatrix} \right].$$

We first study the presence of WF common feature vectors and reparametrize the VECM in terms of $\widetilde{\beta}_\perp$, with $\widetilde{\beta}'_\perp \widetilde{\beta} = 0$, where $\widetilde{\beta}$ denotes the cofeature matrix and $\widetilde{\beta}_\perp$ denotes the orthogonal complement of $\widetilde{\beta}$, and in terms of the common factors $f_t = C\Delta x_{t-1}$ where C is a $(n-s) \times n$ full row rank matrix (see Ahn, 1997; Hecq et al, 2001, Vahid and Engle, 1993). For the identification of $\widetilde{\beta}_\perp$ and C from $\Gamma = \widetilde{\beta}_\perp C$ a normalization has to be imposed. Usually a $(n-s) \times (n-s)$ submatrix of C is set equal to I_{n-s}. Partitioning $\widetilde{\beta}_\perp$ and α in accordance with the partitioning of x_t, we get:

$$\begin{pmatrix} \Delta x_{1t} \\ \Delta x_{2t} \end{pmatrix} = \begin{pmatrix} \widetilde{\beta}_{1\perp} \\ \widetilde{\beta}_{2\perp} \end{pmatrix} C\Delta x_{t-1} + \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \beta' x_{t-1} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}. \tag{16}$$

As above, let X_{t-1} denote the information set containing all information available up to period $t-1$. Notice that if the dimensionality assumption $n_1 \geq \max(r, n-s)$ holds, after imposing the normalization, the number of parameters in $\widetilde{\beta}_{1\perp}$ and C to be estimated [i.e. $(n_1 + s)(n-s)$] is smaller than that in Γ_1 [i.e. $n_1 \times n$]. Similarly, in this case, the number of parameters to be estimated in α_1 and β [i.e. $(n + n_1 - r)r$] is smaller than the number

[5] The generalization to the VAR(p) is straightforward by replacing Γ by $\Gamma(L) = \sum_{i=1}^{p-1} \Gamma_i L^i$.

of parameters in the unrestricted matrix $\alpha_1\beta'$. A necessary condition for identification of the parameters in the first subsystem from that subsystem is therefore satisfied.

Conditions for weak exogeneity of a given set of variables require that the parameters of interest are specified. We first assume that the parameters of interest ψ_0 are given by the parameter matrices $(\alpha, \beta', \widetilde{\beta}_\perp, C)$. From the joint model, the partial (conditional) model for $\Delta x_{1t}|\Delta x_{2t}, X_{t-1}, \theta_c$ is easily obtained from these equations:

$$\Delta x_{1t} = \pi_0 \Delta x_{2t} + [\widetilde{\beta}_{1\perp} - \pi_0 \widetilde{\beta}_{2\perp}] C \Delta x_{t-1} + [\alpha_1 - \pi_0 \alpha_2] \beta' x_{t-1} + \varepsilon_{1.2t},$$

where $\pi_0 = \Omega_{12}\Omega_{22}^{-1}$, $\varepsilon_{1.2t} = \varepsilon_{1t} - \pi_0 \varepsilon_{2t}$.

The marginal subsystem for Δx_{2t} reads as

$$\Delta x_{2t} = \widetilde{\beta}_{2\perp} C \Delta x_{t-1} + \alpha_2 \beta' x_{t-1} + \varepsilon_{2t}.$$

Let us denote the parameters of the conditional by $\theta_c' = (\pi_0, [\widetilde{\beta}_{1\perp} - \pi_0 \widetilde{\beta}_{2\perp}], C, [\alpha_1 - \pi_0 \alpha_2], \beta', \Omega_{11.2})$ and those of the marginal model by $\theta_m' = (\widetilde{\beta}_{2\perp}, C, \alpha_2, \beta', \Omega_{22})$, with $\Omega_{11.2} = \Omega_{11} - \Omega_{12}\Omega_{22}^{-1}\Omega_{21}$. It is seen that β and C enter both the conditional and the marginal models so that, whether or not β is block-diagonal, valid maximum likelihood inference on the parameters of interest from the conditional model can only be achieved under some rather restrictive conditions stated in proposition 3.

Proposition 3 *In the VECM (2) with WF reduced rank structure, then x_{2t} is weakly exogenous for the parameters of interest $\psi_0 = (\alpha, \beta', \widetilde{\beta}_\perp, C)$ if and only if $\alpha_2 = 0$ and $\widetilde{\beta}_{2\perp} = 0$.*

The proof follows the line of arguments presented in the case of a cointegrated system without common features and is omitted here to save space (see inter alia, Johansen, 1995, Urbain, 1992). It is obvious that the two conditions are sufficient for weak exogeneity. They are also necessary as the condition $\widetilde{\beta}_{2\perp} C = 0$ implies that $\widetilde{\beta}_{2\perp} = 0$ as C is by assumption of full rank $n - s$. Notice that, as the $rank[\widetilde{\beta}] = n - s$, $\widetilde{\beta}_{2\perp} = 0$ can only hold for n_2 such that $n_1 \geq n - s$. Similarly, as $rank[\alpha] = r$, $\alpha_2 = 0$ can only hold for n_2 such that $n_1 \geq r$. Note that weak exogeneity of Δx_{2t} implies that x_{2t} is a random walk.

Consider now the same specification but with a SF reduced rank structure. The only difference in the analysis is that the loading (error-correction) matrix α has the form $\alpha = \widetilde{\beta}_\perp B$,

with B being a $(n-s)\times r$ full column rank matrix. Under a SF reduced rank structure, the model (15) can be expressed as

$$\begin{pmatrix} \Delta x_{1t} \\ \Delta x_{2t} \end{pmatrix} = \begin{pmatrix} \widetilde{\beta}_{1\perp} \\ \widetilde{\beta}_{2\perp} \end{pmatrix} C\Delta x_{t-1} + \begin{pmatrix} \widetilde{\beta}_{1\perp} \\ \widetilde{\beta}_{2\perp} \end{pmatrix} B\beta' x_{t-1} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}. \tag{17}$$

Hence the standard condition for weak exogeneity of variables in a cointegrated systems for the long-run parameters $\alpha_2 = 0$ may be equivalently be written as $\widetilde{\beta}_{2\perp}B = 0$ which can only occur when $\widetilde{\beta}_{2\perp} = 0$. Note that the parameters of interest are $\psi_0 = (\alpha, \beta', \widetilde{\beta}_\perp, C)$ or equivalently $(\beta', \widetilde{\beta}_\perp, C, B)$ since the parameter $\alpha = \widetilde{\beta}_\perp B$ under a SF. Hence:

Proposition 4 *In the VECM (2) with SF reduced rank structure, if $\widetilde{\beta}_{2\perp} = 0$, then x_{2t} is weakly exogenous for the parameter of interest $\psi_0 = (\beta', \widetilde{\beta}_\perp, C, B)$.*

The proof is similar to that of Proposition 3 and hence is also omitted here to save space.

4.2 Weak exogeneity and separability

Under WF, the model (16) can be expressed as:

$$\begin{aligned} \begin{pmatrix} \Delta x_{1t} \\ \Delta x_{2t} \end{pmatrix} &= \underbrace{\begin{pmatrix} \widetilde{\beta}_{11\perp} & \widetilde{\beta}_{12\perp} \\ \widetilde{\beta}_{21\perp} & \widetilde{\beta}_{22\perp} \end{pmatrix}}_{\widetilde{\beta}_\perp,(n\times n-s)} \underbrace{\begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix}}_{C,(n-s\times n)} \begin{pmatrix} \Delta x_{1t-1} \\ \Delta x_{2t-1} \end{pmatrix} \\ &\quad + \underbrace{\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix}}_{\alpha,\ (n\times r)} \underbrace{\begin{pmatrix} \beta'_{11} & \beta'_{12} \\ \beta'_{21} & \beta'_{22} \end{pmatrix}}_{\beta',(r\times n)} \begin{pmatrix} x_{1t-1} \\ x_{2t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}, \end{aligned} \tag{18}$$

where $\widetilde{\beta}_{ij\perp}$ is $n_i \times (n_j - s_j), C_{ij}$ is $(n_i - s_i)\times n_j$, α_{ij} is $n_i \times r_j, \beta'_{ij}$ is $r_i \times n_j$. We again denote $\Omega_{12}\Omega_{22}^{-1}$ by π_0 so that the partial (conditional) model for $\Delta x_{1t}|\Delta x_{2t}, X_{t-1}, \theta_c$ is easily obtained from these equations:

$$\begin{aligned} \Delta x_{1t} &= \pi_0 \Delta x_{2t} + \left[(\widetilde{\beta}_{11\perp} - \pi_0\widetilde{\beta}_{21\perp})C_{11} + (\widetilde{\beta}_{12\perp} - \pi_0\widetilde{\beta}_{22\perp})C_{21}\right]\Delta x_{1t-1} \\ &\quad + \left[(\widetilde{\beta}_{11\perp} - \pi_0\widetilde{\beta}_{21\perp})C_{12} + (\widetilde{\beta}_{12\perp} - \pi_0\widetilde{\beta}_{22\perp})C_{22}\right]\Delta x_{2t-1} \\ &\quad + \left[(\alpha_{11} - \pi_0\alpha_{21})\beta'_{11} + (\alpha_{12} - \pi_0\alpha_{22})\beta'_{21}\right] x_{1t-1} \\ &\quad + \left[(\alpha_{11} - \pi_0\alpha_{21})\beta'_{12} + (\alpha_{12} - \pi_0\alpha_{22})\beta'_{22}\right] x_{2t-1} + \varepsilon_{1.2t}, \end{aligned} \tag{19}$$

where $\varepsilon_{1.2t} = \varepsilon_{1t} - \Omega_{12}\Omega_{22}^{-1}\varepsilon_{2t}$. Similarly, the marginal subsystem for Δx_{2t} reads as

$$\begin{aligned}\Delta x_{2t} &= (\widetilde{\beta}_{21\perp}C_{11} + \widetilde{\beta}_{22\perp}C_{21})\Delta x_{1t-1} + (\widetilde{\beta}_{21\perp}C_{12} + \widetilde{\beta}_{22\perp}C_{22})\Delta x_{2t-1} \\ &\quad +(\alpha_{21}\beta_{11}^{'} + \alpha_{22}\beta_{21}^{'})x_{1t-1} + (\alpha_{21}\beta_{12}^{'} + \alpha_{22}\beta_{22}^{'})x_{2t-1} + \varepsilon_{2t}.\end{aligned}$$

From this specifications we are able to derive the different cases of interest under the assumption of WF and SF reduced rank structures respectively. These different cases will depend on (i) the parameters of interest and (ii) the type of separation that underlies the model. We will now consider various cases of separability in cointegration and for each of these sufficient conditions will be derived. The results are summarized in Table 1.

A few words of explanation are in order. In the first column of Table 1 the restrictions imposed on the VECM (19) are given. For instance, in the case III.b, in addition to the restrictions imposed on the VECM (19) in the case III.a. (i.e. separation in cointegration and WF reduced rank structure), we also impose separation of Type C. For the WF and SF, we first consider the model without separation in cointegration. Next, in addition to separation in cointegration, separation of Type C and complete separation are considered respectively. Sufficient conditions for weak exogeneity of x_{2t} for the parameters of interest are given in the third column. The cases I and II correspond to the models considered in Propositions 3 and 4. These conditions follow directly when the corresponding restrictions are imposed on the model in the form (19). The necessary and sufficient conditions for weak exogeneity given in Proposition 3 appear to be overly strong when for instance only part of the cointegrating vectors are parameters of interest (for a similar analysis, albeit in a different set-up, see Hendry and Mizon, 1993; Ericsson, 1995; Ericsson, Hendry and Mizon, 1998). Under the conditions presented in (III.a) the marginal model for Δx_{2t} becomes a VECM of the form

$$\begin{aligned}\Delta x_{2t} &= \widetilde{\beta}_{22\perp}C_{21}\Delta x_{1t-1} + \widetilde{\beta}_{22\perp}C_{22}\Delta x_{2t-1} + \alpha_{22}\beta_{22}^{'}x_{2t-1} + \varepsilon_{2t} \\ &= \widetilde{\beta}_{22\perp}C_{2}\Delta x_{t-1} + \alpha_{22}\beta_{22}^{'}x_{2t-1} + \varepsilon_{2t},\end{aligned}$$

Table 1: Conditions for Weak Exogeneity

Restrictions on VECM	Parameters of interest ψ_0	Conditions implying weak exogeneity of x_{2t}
Absence of separation in cointegration		
I. WF	α, β, $\widetilde{\beta}_{\perp}$, C	1. $\alpha_2 = 0$ 2. $\widetilde{\beta}_{2\perp} = 0$
II. SF and $r + s \leq n$ $\alpha = \widetilde{\beta}_{\perp} B$	$\beta, \widetilde{\beta}_{\perp}, B, C$	$\widetilde{\beta}_{2\perp} = 0$
Separation in cointegration $\beta_{12} = 0,\ \beta_{21} = 0$		
III.a. WF	α_{11}, β_{11}, $\widetilde{\beta}_{1\perp}$, C_1	1. $\alpha_{21} = 0$, $\alpha_{12} - \Omega_{12}\Omega_{22}^{-1}\alpha_{22} = 0$ 2. $\widetilde{\beta}_{21\perp} = 0$, $\widetilde{\beta}_{12\perp} - \Omega_{12}\Omega_{22}^{-1}\widetilde{\beta}_{22\perp} = 0$
III.b. WF and separation of Type C: $\widetilde{\beta}_{11\perp}C_{12} + \widetilde{\beta}_{12\perp}C_{22} = 0$ $\widetilde{\beta}_{21\perp}C_{11} + \widetilde{\beta}_{22\perp}C_{21} = 0$	α_{11}, β_{11}, $\widetilde{\beta}_{11\perp}$, C_{11}	1. $\alpha_{21} = 0$, $\alpha_{12} = 0$ 2. $\Omega_{12} = 0$
III.c. WF and complete separation: $\alpha_{12} = 0, \alpha_{21} = 0$ $\widetilde{\beta}_{11\perp}C_{12} + \widetilde{\beta}_{12\perp}C_{22} = 0$ $\widetilde{\beta}_{21\perp}C_{11} + \widetilde{\beta}_{22\perp}C_{21} = 0$	α_{11}, β_{11}, $\widetilde{\beta}_{11\perp}$, C_{11}	$\Omega_{12} = 0$
IV.a. SF	$\beta_{11}, \widetilde{\beta}_{1\perp}, C_1, B_{11}$	1. $\widetilde{\beta}_{22\perp}B_{21} = 0$, $(\widetilde{\beta}_{11\perp}B_{12} + \widetilde{\beta}_{12\perp}B_{22})$ $-\Omega_{12}\Omega_{22}^{-1}\widetilde{\beta}_{22\perp}B_{22} = 0$ 2. $\widetilde{\beta}_{21\perp} = 0$, $\widetilde{\beta}_{12\perp} - \Omega_{12}\Omega_{22}^{-1}\widetilde{\beta}_{22\perp} = 0$
IV.b. SF and separation of Type C: $\widetilde{\beta}_{11\perp}C_{12} + \widetilde{\beta}_{12\perp}C_{22} = 0$ $\widetilde{\beta}_{21\perp}C_{11} + \widetilde{\beta}_{22\perp}C_{21} = 0$	$\beta_{11}, \widetilde{\beta}_{11\perp}, C_{11}, B_{11}$	1. $\widetilde{\beta}_{21\perp}B_{11} + \widetilde{\beta}_{22\perp}B_{21} = 0$ $\widetilde{\beta}_{11\perp}B_{12} + \widetilde{\beta}_{12\perp}B_{22} = 0$ 2. $\Omega_{12} = 0$
IV.c. SF and complete separation: $\widetilde{\beta}_{11\perp}B_{12} + \widetilde{\beta}_{12\perp}B_{22} = 0$ $\widetilde{\beta}_{21\perp}B_{11} + \widetilde{\beta}_{22\perp}B_{21} = 0$ $\widetilde{\beta}_{11\perp}C_{12} + \widetilde{\beta}_{12\perp}C_{22} = 0$ $\widetilde{\beta}_{21\perp}C_{11} + \widetilde{\beta}_{22\perp}C_{21} = 0$	$\beta_{11}, \widetilde{\beta}_{11\perp}, C_{11}, B_{11}$	$\Omega_{12} = 0$

while the conditional model reads as

$$\begin{aligned}\Delta x_{1t} &= \pi_0 \Delta x_{2t} + \widetilde{\beta}_{11\perp} C_{11} \Delta x_{1t-1} + \widetilde{\beta}_{11\perp} C_{12} \Delta x_{2t-1} + \alpha_{11} \beta_{11}^{'} x_{1t-1} + \varepsilon_{1.2t} \\ &= \pi_0 \Delta x_{2t} + \widetilde{\beta}_{11\perp} C_1 \Delta x_{t-1} + \alpha_{11} \beta_{11}^{'} x_{1t-1} + \varepsilon_{1.2t},\end{aligned}$$

where $C_i = (C_{i1} : C_{i2}), i \in \{1, 2\}$, so that no cross-equation restrictions remain. Remark that these are conditions concerning the efficiency of an analysis based on a conditional sub-system but do not yet enable one to consider separated marginal systems as x_{1t-1} and x_{2t-1} both appear in this sub-system. Also note that these are sufficient conditions and show that the conditions in Proposition 3 are not necessary once the parameters of interest are only sub-matrices of $\alpha, \beta^{'}$ and $\widetilde{\beta}_{\perp}$. Alternative sets of conditions could be derived. For example, if $\Omega_{12} = 0$ (conditional independence), then both α and $\widetilde{\beta}_{\perp}$ have to be block-diagonal. The cases (III.b) and (III.c) summarize these implications. If in addition to complete separation in the Granger-Haldrup sense (III.c), $\Omega_{12} = 0, \Gamma_{12}(L) \equiv 0, \Gamma_{21}(L) \equiv 0$, the block-diagonality of $\widetilde{\beta}_{\perp}$ results and conditions for weak exogeneity (III.c) are satisfied.

Under the SF reduced rank structure assumption, as mentioned above, the matrix of error-correction terms takes the form $\alpha = \widetilde{\beta}_{\perp} B$, with B being a $(n - s) \times r$ full column rank matrix, with here $n \geq s + r$. With the partitioning introduced earlier this yields

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} = \begin{pmatrix} \widetilde{\beta}_{11\perp} & \widetilde{\beta}_{12\perp} \\ \widetilde{\beta}_{21\perp} & \widetilde{\beta}_{22\perp} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix}.$$

After substituting the resulting expressions for α_{ij}, $i, j \in \{1, 2\}$ into the conditional model for Δx_{1t} in (19) and the marginal model for Δx_{2t}, it is fairly straightforward to verify that the conditions given in Table 1 for the cases II, IVa-IVc (SF reduced rank) indeed imply weak exogeneity for the parameters of interest. Notice that the condition for weak exogeneity under SF reduced rank structure $\widetilde{\beta}_{2\perp} = 0$ implies that $\alpha_2 = \widetilde{\beta}_{2\perp} B = 0$, so that under Type B separation and under complete separation when $\alpha_{12} = 0$ too, $r = r_1$ [of course we then have to require that $n_1 \geq \max(r, n - s)$]. Therefore, under SF reduced rank structure and under Type B or complete separation, weak exogeneity requires that the cointegrating relationships in the model include the variables x_{1t} and appear only in the first subsystem.

Finally, notice that under separation of Type B part of the requirements for weak exogene-

ity (e.g. $\alpha_{21} = 0$) are satisfied by assumption, so that one has to check whether the remaining requirements of case I (for the WF) or case II (for the SF) are also satisfied. Under separation of Type A one has to proceed in a similar way. Note that if $\alpha_{12} = 0$, weak exogeneity as considered in cases I or II cannot hold because it requires $\alpha_{21} = 0$, which would contradict Type A separation.

The results presented in the preceding sub-sections summarize conditions for efficient maximum likelihood (ML) based estimation and inference on the parameters of interest within a partial model. It does not yet imply that the system may be split into two separated subsystems, i.e. *independent marginal* systems that can be treated independently. Indeed, let us again consider the VECM (16) under separation in cointegration. In general even under weak exogeneity of x_{2t}, separation (here *marginal* systems in contrast to *partial* or *conditional* systems) will only enable us to recover the parameters of interest if additional restrictions are imposed such as the block-diagonality of C, i.e. when $C_{12} = 0, C_{21} = 0$ which corresponds to Type C separation in cointegration. In that case, if in addition to $\alpha_{21} = 0$, $\widetilde{\beta}_{2\perp} = 0$ and $\alpha_{12} - \Omega_{12}\Omega_{22}^{-1}\alpha_{22} = 0$ for the weak form for example, then the common cyclical factors entering the partial, here marginal, system for Δx_{1t} are only made of linear combinations of lagged Δx_{1t} and $\beta_1' x_{1t-1}$.

These results have to be related to the general discussion of conditioning vs. marginalization in dynamic econometric models. Valid marginalization is much more demanding than valid conditioning and requires that the joint likelihood of the complete system can be written as the product of the marginal likelihoods of the subsystems which entails not only weak exogeneity but also Granger-Noncausality (given common feature restrictions) and conditional independence. This last requirement of conditional independence is valid irrespectively of the presence or absence of common feature restrictions in a VECM as also pointed out by Hansen and Johansen (1998, page 68).

5 Inference on cofeature separation and modeling strategy

The results derived in Section 3 and 4 have obvious implications for the way in which separation should be tested in a VAR model in practice. The following sequence of increasingly restrictive hypotheses comes naturally to one's mind.

- Within a complete system (2) first test for the number of cointegration relationships r and fix the cointegration rank to this value in subsequent analyses.
- Second, test for separation in cointegration, that is for the block-diagonality of β, e.g. using tests proposed by Konishi and Granger (1993).
- Third, investigate whether α is block-diagonal (Type B separation).
- Fourth, test the WF reduced rank structure.
- Fifth, test the strong form.
- Sixth, test for separation in common features, i.e. block-diagonality of $\widetilde{\beta}$.
- Seventh, if block-diagonality of $\widetilde{\beta}$ is not rejected, check whether $\Gamma(L)$ is block-diagonal (Type C separation).
- Eight, check whether the findings are consistent with the implications of the Propositions 1 and 2:
  - Under block-diagonality of β', SF with $s_i = n_i - r_i$, Type B separation implies common features separation. When $p > 1$, and $s = \text{argmin}(n - s_j^*; j = 1, 2, \ldots, p-1)$, Type C separation implies common feature separation. When $p > 1$ and $s = \text{argmin}(n - r, n - s_j^*; j = 1, 2, \ldots, p-1)$, complete separation implies common feature separation.
  - Under block-diagonality of β', for $p > 1$, WF with $s = \text{argmin}(n - s_j^*; j = 1, 2, \ldots, p-1)$, Type C separation (or complete separation) implies common feature separation.
- Ninth, test for weak exogeneity by checking the appropriate conditions as given in Table 1.
- Tenth, test for complete separation including the block-diagonality of Ω.

The first three steps of our sequential testing strategy are standard and discussed in detail in a.o. Johansen (1995). Johansen and Juselius (1990) give likelihood ratio tests for linear restrictions on both α and β. Harbo, Johansen, Nielsen and Rahbek (1998) consider

asymptotic inference on cointegration rank in partial systems. Johansen and Swensen (1999) provide asymptotically χ^2 distributed likelihood ratio tests based on canonical correlations for nonlinear restrictions on α and linear restrictions on β. The remaining steps need to be described in some detail.

Let us first consider test statistics for weak form common features in the VECM (2). Tests carried out in this paper are based on the sample canonical correlations of two sets of random vectors $\Delta X = (\Delta x_1, \ldots, \Delta x_T)$, $\Delta Z = (\Delta X'_{-1}, \ldots \Delta X'_{-p+1})'$, both corrected for deterministic components and cointegrating relationships $\beta' X_{-1}$, where $X_{-1} = (x_0, \ldots, x_{T-1})$, with β known or superconsistently estimated. Assuming that the deterministic component is simply specified as a set of functions of time t, including also an unrestricted constant term (and possibly seasonal dummies) and denoted by $g(t)$, we can summarize the procedure as follows:

$$CanCor(\Delta X, \Delta Z \mid g(t), \beta' X_{-1})$$

Squared canonical correlations are found as solutions of the usual eigenvalue problem

$$\mid \lambda I - S_{XX}^{*-1/2} S_{XZ}^{*} S_{ZZ}^{*-1} S_{ZX}^{*} S_{XX}^{*-1/2} \mid = 0,$$

where S_{XZ}^{*} denotes the covariance matrix between the elements of ΔX and ΔZ both corrected by $g(t)$ and long-run relationships. We denote these by ΔX^* and ΔZ^*. Note that in the SF model, the analysis consists in a canonical correlation analysis between ΔX and $\Delta W = (\Delta Z', X'_{-1}\beta)'$ both sets corrected for the deterministic components $g(t)$, i.e.: $CanCor(\Delta X, \Delta W \mid g(t))$. The empirical squared canonical correlations $\widehat{\lambda}_i^w$ (for the WF), $\widehat{\lambda}_i^s$ (for the SF), ordered as follows $0 \leq \widehat{\lambda}_1^w \leq \widehat{\lambda}_2^w \leq \ldots \leq \widehat{\lambda}_s^w \leq 1$ (similarly for SF) allow us to test the null that the first s linear combinations are white noises using the statistics $\xi_W = -T \sum_{i=1}^{s} \log(1 - \widehat{\lambda}_i^w)$, $s = 1, ..., n$ and $\xi_S = -T \sum_{i=1}^{s} \log(1 - \widehat{\lambda}_i^s)$, $s = 1, ..., n - r$ which have asymptotic χ^2 distributions under the null (see Hecq et al., 2001 for details). The number of degrees of freedom are $s[n(p-1) + r] - s(n-s)$ for the SF and $sn(p-1) - s(n-s)$ for the WF. If we define $S_{0.XX}^{*} = S_{XX}^{*} - S_{XZ}^{*} S_{ZZ}^{*-1} S_{ZX}^{*}$, the maximized (log-)likelihood for the WF case is given by:

$$\mathcal{L}^{WF} = -\frac{T}{2} \left\{ \ln\{\det(S_{0.XX}^{*})\} - \sum_{i=1}^{s} \ln(1 - \widehat{\lambda}_i^w) \right\}.$$

The corresponding expression for the log-likelihood function of the strong form model ($\mathcal{L}^{SF}$) is immediate.

To get intuitively into separation in common features let us reformulate the problem in terms of moment restrictions. The existence of a strong form common feature space means $E[\widetilde{\beta}'\Delta x_t \otimes \Delta W_t] = 0$, where ΔW_t is a set of instruments composed of the $p-1$ lags of the n variables of Δx_t and the cointegrating vectors. In the WF framework the condition becomes $E[\widetilde{\beta}'\Delta x_t^* \otimes \Delta Z_t^*] = 0$, where ΔZ_t^* is a set of instruments composed of the lags of all variables in Δx_t. In the SF case the variables Δx_t and ΔW_t have to be detrended by $g(t)$. In the model under WF, Δx_t^* and ΔZ_t^* have been concentrated on $g(t)$ and on the error-correction terms using super-consistent estimates of the cointegrating vectors obtained from a reduced rank regression in a first step. Now, under separation in common features we know that under the null the condition $E[(\widetilde{\beta}_1'\Delta x_{1t} + \widetilde{\beta}_2'\Delta x_{2t}) \otimes \Delta W_t] = 0$ or $E[(\widetilde{\beta}_1'\Delta x_{1t}^* + \widetilde{\beta}_2'\Delta x_{2t}^*) \otimes \Delta Z_t^*] = 0$, where $\widetilde{\beta}_1$ and $\widetilde{\beta}_2$ are common feature vectors for the mutually exclusive n_1 and n_2 sets, should also hold.

In order to determine whether $\widetilde{\beta}'$ is block-diagonal we need a statistic to test different restrictions on common feature vectors. To do so, along the line of Johansen (1995), Konishi and Granger (1993) for cointegration a likelihood ratio test is computed using a switching algorithm for the restrictions:

$$\widetilde{\beta} = \left\{\widetilde{\beta}_1, \widetilde{\beta}_2\right\} = \{H_1\Psi_1, H_2\Psi_2\},$$

where H_1 and H_2 are matrices of known constants respectively of dimension $(n \times n_1)$ and $(n \times n_2)$, while Ψ_1 and Ψ_2 are matrices containing unknown parameters of dimension $(n_1 \times s_1)$ and $(n_2 \times s_2)$.

Example 1 *With $n_1 = n_2 = 2$ and $s_1 = s_2 = 1$, $r_1 = r_2 = 1$, the situation we face in the empirical section of this paper, these matrices can be parametrized such that:*

$$\widetilde{\beta}_1 = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ 0 \\ 0 \end{pmatrix}, \quad \widetilde{\beta}_2 = \begin{pmatrix} 0 \\ 0 \\ \varphi_3 \\ \varphi_4 \end{pmatrix}, \quad H_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\Psi_1 = \begin{pmatrix} \varphi_1 \\ \varphi_2 \end{pmatrix}, \quad \Psi_2 = \begin{pmatrix} \varphi_3 \\ \varphi_4 \end{pmatrix}$$

The log-likelihood under common feature restrictions has been previously labelled $\mathcal{L}^{WF}$ or $\mathcal{L}^{SF}$. If convergence is achieved using a switching algorithm,[6] the log-likelihood functions under separate restrictions on different common feature vectors are labelled $\mathcal{L}^{WF}_{sep}$ or $\mathcal{L}^{SF}_{sep}$. In practice, as for Johansen's analyssis of cointegration, we propose the following steps:[7]

Step 1. Choose an unrestricted vector $\widetilde{\beta}_1$,

Step 2. Given $\widetilde{\beta}_1$, solve $CanCor(H_2'\Delta X^*, \Delta Z^* \mid \widetilde{\beta}_1'\Delta X^*)$ to obtain $\widetilde{\beta}_2$,

Step 3. Given $\widetilde{\beta}_2$, compute $CanCor(H_1'\Delta X^*, \Delta Z^* \mid \widetilde{\beta}_2'\Delta X^*)$ to obtain a new $\widetilde{\beta}_1$,

Step 4. Continue this process until convergence to the maximum of the likelihood function.

Under cofeature separation, the maximized log-likelihood is given by

$$\mathcal{L}^{WF}_{sep} = -\frac{T}{2}\left\{\ln\{\det(S^*_{0.XX})\} - \sum_{i=1}^{s_1}\ln(1-\widehat{\lambda}^w_{i,\widetilde{\beta}_1}) - \sum_{j=1}^{s_2}\ln(1-\widehat{\lambda}^w_{j,\widetilde{\beta}_2})\right\},$$

where $S^*_{0.XX}$ is defined above, $\widehat{\lambda}^w_{i,\widetilde{\beta}_1}$ and $\widehat{\lambda}^w_{j,\widetilde{\beta}_2}$ are the eigenvalues obtained after convergence under respectively the restrictions $\widetilde{\beta}_1 = H_1\Psi_1$ and $\widetilde{\beta}_2 = H_2\Psi_2$. A likelihood ratio test for separation in common features is then obtained by computing $LR^{SF}_{sep} = -2(\mathcal{L}^{SF}_{sep} - \mathcal{L}^{SF})$ for the strong form model or $LR^{WF}_{sep} = -2(\mathcal{L}^{WF}_{sep} - \mathcal{L}^{WF})$ for the weak form. These statistics have an asymptotic χ^2 distribution under the null. The number of degrees of freedom is equal to the number of restrictions (beyond those from normalization) imposed by separation. [see e.g. Johansen, 1995, p.115, for more details on the computation of the number of degrees of freedom]. In the example 1, after normalization, with $n = 4$, $s = 2$, the unrestricted matrix $\widetilde{\beta}$ contains $(n \times s) - s^2 = 4$ free parameters. After normalization, the restricted $\widetilde{\beta}$ contains two free parameters. The number of degrees of freedom equals 2.

[6] Note that due to the absence of cross-restrictions between the different common feature vectors, the convergence is very fast.

[7] We describe the algorithm for WF, the results for SF differ only by the fact that we do not concentrate with respect to the cointegrating vectors.

The last two steps are more straightforward. Weak exogeneity analyses depend on the chosen parameters of interest but are easily conducted once the system (with cointegration and common features restrictions) is written in pseudo-structural form in terms of the matrix $\widetilde{\beta}_{\perp}$, see Vahid and Engle (1993), and estimated by FIML. The last step, e.g. testing the block-diagonality of the covariance matrix could be investigated by standard variable addition tests. Note finally that a global LR test statistic of the independence of the subsystems can also be conducted by comparing the log-likelihood of the complete model with the sum of the log-likelihoods of the two marginal models. Such a LR type test has a χ^2 distribution under the null.[8]

It is worth making two remarks at this stage:

Remark 1 The major drawback of the analysis we propose here relates to the need of specifying and estimating complete systems to assess the appropriateness (in terms of P-T decomposition) and efficiency of the sub-group based analysis. While the purpose of the aforementioned testing strategy is to sustain sub-systems analysis, it requires full system estimation to compute the different test statistics. This is naturally the most important drawback here but it is not specific to the analysis we propose and occurs in (weak) exogeneity and Granger Noncausality analyses in general. Hence, while feasible, our testing strategy implicitly supposes a low number of variables and sub-groups.

Remark 2 An important final remark should be made concerning the restrictiveness of the specifications we investigate here. While the class of models we consider may at the first sight appear overly restrictive, it should be noted that the restrictions we propose to investigate are actually systematically imposed without further investigation in the recent literature on panel data cointegration (see the surveys by Phillips and Moon, 1999; Baltagi and Kao, 2000) and panel common cycle analysis (Hecq et al., 2000b). While few papers allow for a non-diagonal covariance matrix between the subgroups of the panels, cointegration separation is

[8]It is important to recognize that the sequential approach proposed above, while being appealing as we consider the different restrictions sequentially, naturally raises the crucial issue of size distortion that is likely to occur in any empirical application since the choice of the hypotheses tested is dependent on the outcome of the tests of the previous hypotheses in the sequence.

indeed systematically imposed by excluding cointegration across members of a panel. With the exception of a few studies like Banerjee et al. (2001), Groen and Kleibergen(1999), Hall, Lazarova and Urga (1999) and Larsson and Lyhagen (1999) this restriction is usually not explicitly discussed and studied.

6 Empirical Analysis

To show the feasibility of the analysis we illustrate the testing strategy within a small dimensional problem that is also used in Hecq et al. (2000b) where the notion of serial correlation common features is introduced in a nonstationary panel data context.

6.1 The Economic Model

We focus on the permanent income hypothesis versus the heterogenous consumer model proposed by Campbell and Mankiw (1990, 1991), also known as the "λ" *model*. These authors consider two groups of agents who receive respectively a disposable income Y_{1t} and Y_{2t} in fixed proportions of the total income Y_t, such that $Y_{1t} = \lambda Y_t$, $Y_{2t} = (1-\lambda)Y_t$ and $Y_t = Y_{1t} + Y_{2t}$. Agents in the first group are subject to liquidity constraints, consume their current income while agents in the second group consume their permanent income. We get the following system:

$$\begin{cases} C_{1t} = Y_{1t} = \lambda Y_t \\ C_{2t} = Y_{2t}^P = (1-\lambda)Y_t^P \\ Y_{1t} = Y_{1t}^P + Y_{1t}^T \\ Y_{2t} = Y_{2t}^P + Y_{2t}^T, \end{cases} \tag{20}$$

where C_{it} is the consumption of agent i and Y_{it}^P and Y_{it}^T are the permanent and transitory components of income of the agent i and are assumed to be *I(1)* and *I(0)*, respectively. Aggregating over agents we get $C_t = Y_{1t}^P + Y_{1t}^T + Y_{2t}^P = Y_t^P + Y_{1t}^T$, and thus:

$$\begin{cases} C_t = Y_t^P + \lambda Y_t^T \\ Y_t = Y_t^P + Y_t^T, \end{cases} \tag{21}$$

which shows that aggregate consumption and income share a common trend Y_t^P. It is also easily seen that if $\lambda = 0$ we get the permanent income model. In order to stress the common cycle component let us take the first difference of aggregate consumption $C_t = C_{1t} + C_{2t}$. By substituting the shares of income in the total income and taking first differences we obtain $\Delta C_t = \lambda \Delta Y_t + (1 - \lambda)\Delta Y_t^P$. Consequently, assuming that the permanent income is a martingale, the consumption function can be tested by using the regression $\Delta C_t = \lambda \Delta Y_t + (1 - \lambda)\varepsilon_t$, where ε_t is a martingale difference sequence which is not orthogonal by construction to ΔY_t except under the hypothesis of common features between ΔC_t and ΔY_t with $[1 \; -\lambda]$ being the associated common features vector. Empirical studies have shown that λ is usually significantly different from zero with a value in the range 0.3 to 0.5 for most countries.

6.2 Empirical Results

For illustrative purpose, we confine our analysis to the "λ *model*" for the USA and Canada using annual data for the period from 1950 to 1992. The data are taken from the Penn World Tables Mark 5.6 (see Summers and Heston, 1991) which due to their definition homogeneity are extremely useful and have been extensively used in empirical cross-country studies. It is interesting to note that Hoffman (1999)'s study points out that total permanent shocks seem to be relatively unimportant for output volatility for both Canada and US which motivates further analysis of short-run comovements between these two countries.

The data used are $Y =$"*RGDPL: Real GDP per capita (Laspeyres index) in 1985 international prices*" and $C =$ "*C: Real Consumption share of GDP in 1985 international prices*"$\times Y/100$. This last operation is necessary to get the consumption in level and not in percentage of income. Although we may expect some lack of power with this small sample of annual data, the use of an annual frequency avoids the pitfalls that one would probably encounter with seasonally adjusted monthly or quarterly data (see Hecq, 1998; Cubadda, 1999).

A first descriptive analysis using standard unit root tests and both Engle and Granger (1987) and Johansen (1995) cointegration analyses reveals the existence of a single cointegrating vector for both countries taken separately. In both cases, the cointegrating vectors

are close to the expected theoretical vectors, although the long-run unit income elasticity is more plausible for the Canada (1.00) than for the US (1.15).

6.2.1 Separation in Cointegration

We first analyze cointegration separation at the level of the complete system for the two countries. Using information criteria and the results from testing for the presence of residual auto-correlations, heteroskedasticity and non-normality, a VAR of order four in the levels of the time series is found to appropriately characterize the covariance structure of this four-dimensional vector of variables. We then apply Johansen's tests statistics with a deterministic linear trend restricted to lie in the cointegration space. The results are reported in Table 2.

Table 2: Johansen's ML Tests Statistics

	λ_i	Max.Eig.Test	95% cv	Trace Test	95% cv
$r = 0$	0.71	48.64*	31.5	83.34*	63.0
$r \leq 1$	0.43	21.98	25.5	34.70	42.4
$r \leq 2$	0.22	9.87	19.0	12.72	25.3
$r \leq 3$	0.07	2.85	12.2	2.85	12.2

Both the maximum eigenvalue and the trace test favor the existence of a single cointegrating vector. The smallness of our sample may however seriously affect the power of these tests so that the analysis is complemented by both a visual inspection of cointegrating vectors presented in Figure 1, and the analysis of the roots[9] of $(\alpha\beta' - I)$ that lead us to retain two cointegrating vectors.

The next step in our sequential analysis is to test for restrictions on the cointegrating space as well as on the loading matrix in order to analyze separation in cointegration. The results are presented in Table 3. The null hypothesis $H_0 : \beta_{12} = \beta_{21} = 0$ corresponds to the cointegration separation hypothesis that is tested using a LR test that is $\chi^2(3)$ distributed under the null of separation.[10]

[9] We obtain two pairs of complex roots whose moduli are respectively 1 and 0.207. For a formal discussion of the usefulness of analyzing the roots of the companion form in cointegrated systems, see Johansen (2001).

[10] One should notice that the number of degrees of freedom of the different statistics is equal to the number of restrictions imposed under the different forms of separation plus one since the linear trend coefficient is not significant in the long-run relationship for Canada.

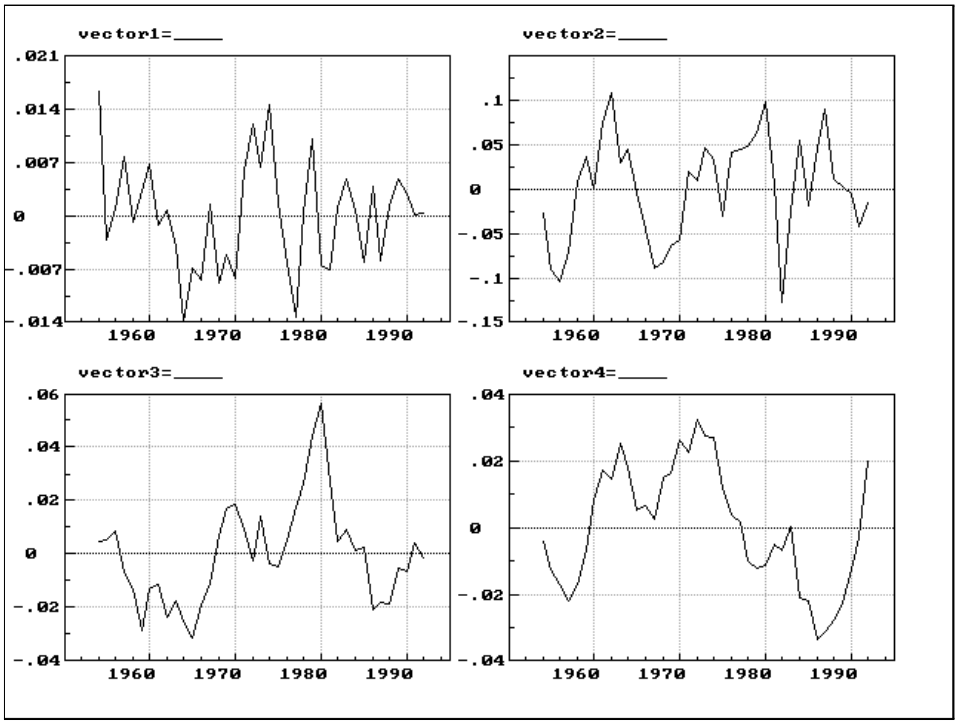

Figure 1: Cointegrating Vectors

The null $H_0 : \beta_{12} = \beta_{21} = 0$ & $\alpha_{12} = \alpha_{21} = 0$ corresponds to separation of Type B while the two remaining hypotheses are joint hypotheses of cointegration separation and block-triangularity of the loading matrix. It appears that separation in cointegration cannot be rejected.

Table 3: Hypotheses Testing

H_0	Test stat.	Distr.	P-values
$\beta_{12} = \beta_{21} = 0$	0.762	$\chi^2(3)$	0.858
$\beta_{12} = \beta_{21} = 0$ and $\alpha_{12} = \alpha_{21} = 0$	42.58	$\chi^2(5)$	<0.001
$\beta_{12} = \beta_{21} = 0$ and $\alpha_{21} = 0$	9.55	$\chi^2(4)$	0.048
$\beta_{12} = \beta_{21} = 0$ and $\alpha_{12} = 0$	22.19	$\chi^2(4)$	<0.001

On the contrary, we reject separation of Type B. Due to the small sample size we can however not clearly decide whether the loading matrix is block-triangular, i.e. whether α_{21} is zero or not but we guess that it is not. Indeed, the null hypothesis is clearly rejected when we do not restrict the trend to zero in the equation for Canada or when we do not impose separation in cointegration.

Assuming cointegration separation alone, e.g. only block-diagonality of β, we obtain:

$$\alpha\beta' x_t = \begin{pmatrix} 1.262 & -0.428 \\ -0.075 & 0.805 \\ 1.280 & -0.554 \\ 0.819 & -0.847 \end{pmatrix} \begin{pmatrix} -0.981 & 1.000 & 0.000 & 0.000 & 0.000 \\ 0.000 & 0.000 & -0.923 & 1.000 & -0.004 \end{pmatrix} \begin{pmatrix} Ca\_y_t \\ Ca\_c_t \\ USA\_y_t \\ USA\_c_t \\ trend \end{pmatrix},$$

which yields the following long-run consumption function, $c_t = 0.98 y_t$ for Canada and $c_t = 0.92\, y_t + .004\, trend$ for the USA. The estimated long-run income elasticity corresponds to that obtained from individual cointegration analyses in the case of Canada. The result is now more plausible for the USA. The possible explanation can be the omission of the trend and as well as Canadian variables in the static regression. This trend partially accounts for government spendings.

6.2.2 Separation in Common Features

Fixing the two cointegrating vectors to their estimated values and using test statistics based on canonical correlations proposed in Hecq et al. (2001), we determine the dimension of the common feature space both in the WF and the SF model using the statistics ξ_W and ξ_S presented in Section 5. Note that in the SF case the number of common feature vectors is bounded by $n - r$, i.e. by 2 in this analysis. Table 4 presents the results.

Table 4: Common Feature Tests

$r = 2$	λ_i^*		df		$Pb > \chi^2_{df}$		$Loglik$	
	ξ_S	ξ_W	ξ_S	ξ_W	ξ_S	ξ_W	ξ_S	ξ_W
$s \geq 1$	–		–		–		729.25	729.25
$s \geq 2$	0.401	0.217	11	9	0.045	0.386	719.245	724.467
$s \geq 3$	0.491	0.366	24	20	0.004	0.124	706.056	715.552
$s \geq 4$	0.523	0.498	39	33	(< 0.001)	0.011	691.588	702.084
$s = 5$	0.817	0.734	56	48	(< 0.001)	< 0.001	658.378	676.245

Under the SF model we do not retain any common feature vectors. In contrast, we obtain two weak form common features vectors. The normalized cofeature matrix is given by

$$\widetilde{\beta}'\Delta x_t^* = \begin{pmatrix} 1 & -0.671 & 0 & 0.257 \\ 0 & 0.047 & 1 & -0.719 \end{pmatrix} \begin{pmatrix} \Delta(Ca\_c_t)^* \\ \Delta(Ca\_y_t)^* \\ \Delta(USA\_c_t)^* \\ \Delta(USA\_y_t)^* \end{pmatrix},$$

where the superscript * indicates that variables have been taken in deviation from their (long-run) error-correction terms. The short-run elasticities are plausible, namely 0.67 for Canada and 0.72 for the US. Under common feature separation, the switching algorithm proposed in the previous section yields a log-likelihood value of $\mathcal{L}_{sep}^{WF} = 714.29$ that should be compared with the value of the log-likelihood under two unrestricted common feature vectors, see Table 4, namely 715.55. The LR test for separation in common feature, $\chi^2(2)$ under[11] the null, follows directly and yields a value of $LR_{sep}^{WF} = 2.52$. Consequently we cannot reject

[11] Indeed the 4×2 normalized cofeature matrix contains two zeros. Separation in common feature adds two zero restrictions.

the hypothesis of separation in weak form common features so that there apparently exist independent cycles in the short-run dynamics for the USA and Canada.

Finally, we estimate the complete system under the restrictions of separation in cointegration and separation in weak form common features. FIML estimates and associated asymptotic standard errors are given in Table 5. In the columns 2 and 4, we find the estimates of the coefficients of the equations for consumption premultiplied by the common feature vectors which are (1 -0.475 0 0) and (0 0 1 -0.682) for Canada and the USA respectively. Under separation in common feature, the short-run income elasticities for Canada and the US are respectively 0.48 and 0.68, namely slightly less than for the unrestricted model. In the columns 3 and 5, the estimates of the error-correction form of the equations for income in resp. Canada and the USA are presented. Due to the significance of adjustment to error-correction terms, e.g. a non-diagonal α matrix, a P-T decomposition (whether GG or BNSW) cannot be extracted for each country separately (for more details see Granger and Haldrup, 1997).

Statistics for various hypotheses of block-diagonality or triangularity of $\Gamma(L)$ were computed under the maintained hypotheses of separation in cointegration and separation in weak form common features. The LR test for the null of block-diagonality of $\Gamma(L)$, distributed as a $\chi^2(12)$ under the null, has a value of 44.83 which immediately shows that separation of Type C is rejected at any reasonable significance level. The finding is consistent with Propositions 1 and 2 that common feature separation does not necessarily imply Type C separation. The null hypotheses of upper or lower triangularity of $\Gamma(L)$ are similarly rejected using LR test statistics, with asymptotic $\chi^2(6)$ distributions under the null. The computed value of the statistic for upper triangularity (resp. lower) is 27.05 (resp. 33.64).

Table 5: FIML Estimates under Separation in Cointegration and in Weak Form Common Features

Dep.Var	$\Delta(Ca\_c_t)$ $Coef.(Std.Err.)$	$\Delta(Ca\_y_t)$ $Coef.(Std.Err.)$	$\Delta(USA\_c_t)$ $Coef.(Std.Err.)$	$\Delta(USA\_y_t)$ $Coef.(Std.Err.)$
$\mathbf{\Delta(Ca\_y}_t)$	.475(.107)	-	-	-
$\mathbf{\Delta(USA\_y}_t)$	-	-	.682(.103)	-
Const	$-0.316(-)$	$0.613(-)$	$-0.010(-)$	$0.606(-)$
$(Ca\_CI)_{t-1}$	$-0.629(0.171)$	$1.378(0.497)$	$-0.148(0.076)$	$1.355(0.394)$
$(USA\_CI)_{t-1}$	$0.699(0.228)$	$-1.075(0.551)$	$-0.133(0.141)$	$-1.012(0.502)$
$\mathbf{\Delta(Ca\_y}_{t-1})$		$1.132(0.374)$		$1.044(0.282)$
$\mathbf{\Delta(Ca\_y}_{t-2})$		$0.477(0.328)$		$0.767(0.271)$
$\mathbf{\Delta(Ca\_y}_{t-3})$		$0.405(0.308)$		$0.725(0.149)$
$\mathbf{\Delta(Ca\_c}_{t-1})$		$-0.745(0.285)$		$-0.939(0.231)$
$\mathbf{\Delta(Ca\_c}_{t-2})$		$-0.654(0.304)$		$-1.017(0.234)$
$\mathbf{\Delta(Ca\_c}_{t-3})$		$-0.627(0.319)$		$-0.642(0.245)$
$\mathbf{\Delta(USA\_y}_{t-1})$		$-2.514(0.583)$		$-2.165(0.459)$
$\mathbf{\Delta(USA\_y}_{t-2})$		$-1.616(0.477)$		$-1.321(0.403)$
$\mathbf{\Delta(USA\_y}_{t-3})$		$-.800(0.232)$		$-0.909(0.216)$
$\mathbf{\Delta(USA\_c}_{t-1})$		$2.833(0.610)$		$2.512(0.496)$
$\mathbf{\Delta(USA\_c}_{t-2})$		$2.104(0.547)$		$1.728(0.456)$
$\mathbf{\Delta(USA\_c}_{t-3})$		$1.253(0.268)$		$0.686(0.287)$

- **Notes**
- The table reports FIML Estimates under Separation in Cointegration and in Weak Form Common Features,
- Asymptotic standard errors are reported in parentheses,
- $(Ca\_CI)_{t-1}$ denotes the cointegrating relationship found for the Canadian series,
- $(USA\_CI)_{t-1}$ denotes the cointegrating relationship found for the US series,

FIML-estimates of the model under separation in cointegration and in the WF common features are reported in Table 5. From the empirical results reported in Tables 3-5 and the conditions derived earlier it is immediately seen that none of the weak exogeneity conditions pointed out in Section 4 holds as the hypotheses of diagonality of the error-correction matrix α is already rejected at usual significance levels. As a by product, efficient inference requires estimation of the joint model and permanent-transitory decompositions may not be obtained separately for each country. Although the block-diagonality of the common feature matrix $\widetilde{\beta}$

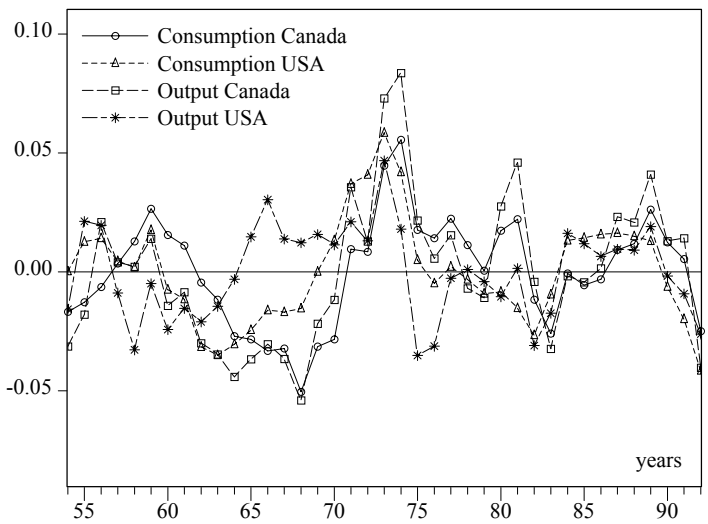

Figure 2: Full System, Decomposition of x_t

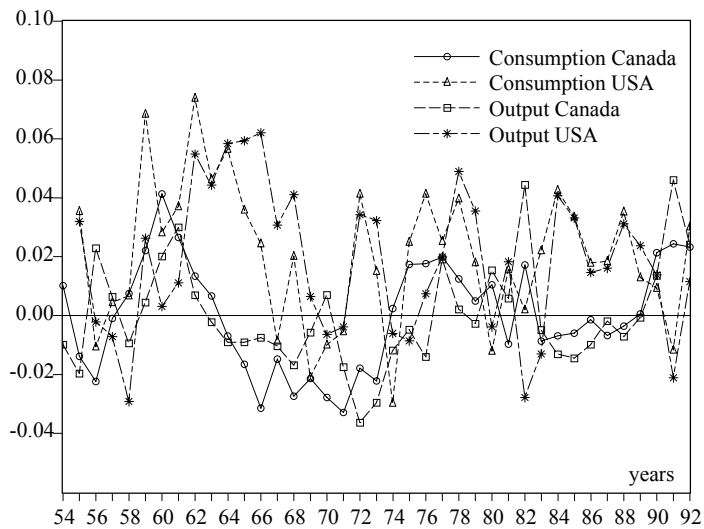

Figure 3: Marginal Systems, Decomposition of x_t

is not rejected by the data, so that the two common cycles appear to be country-specific, the latter cycles may not be extracted country by country as they involve linear combinations of the short-run dynamics for both Canada and the US.

Finally, the relevance of our analysis and the importance of satisfying separability conditions can be illustrated by the computation of P-T decompositions from both the full system and from the marginal sub-systems. If all the different separability conditions mentioned in this paper hold, both analyses should yield very similar decompositions up to slight differences that could arise from not imposing separability restrictions when estimating the joint system and from different normalizations when computing the P-T decompositions for the joint system and for the two marginal systems respectively. We therefore did compute the transitory components for each of the four series from both a complete system analysis and from a marginal country by country analysis.

Figure 2 presents the decomposition based on a full system analysis while Figure 3 was obtained from the decomposition of the marginal sub-systems. The surprising differences in the graphical representation illustrate in a convincing way the importance of testing and empirically validating separability restrictions before imposing them when computing P-T decompositions.

7 Concluding remarks

In this paper, we have considered the issue of separation in cointegrated VAR models with common features by extending the analysis proposed by Granger and Haldrup (1997). In particular, we have pointed out that their analysis needed to be extended to cover the case with common features. In this set-up, we have studied the relationships between various forms of separation in VAR models and the presence and form of common features. From this analysis, explicit testable parameter restrictions have been derived under which sub-group P-T analysis is equivalent to full system P-T analysis. These conditions entail both restrictions on the long-run matrices as well as more restrictive assumptions such as Granger Noncausality and conditional independence.

We have also clarified the relationships between weak exogeneity and separability in this class of cointegrated VAR models and pointed out that the standard conditions of weak

exogeneity (w.r.t. to the long-run parameters) usually considered in cointegrated VAR models need to be generalized in this class of models.

The separation restrictions considered arise in applications where panel data in the form of sets of time series for several entities such as households, firms, sectors, regions or countries are analyzed. Not only did we point out the implications between the existence of common features and forms of separation, we have also proposed and adopted a modelling strategy that allows an investigator to test these implications in a systematic way and to finally obtain a model that is statistically validated and has restricted dynamics which can be interpreted in terms of economic theory. The analysis has been illustrated using consumption and income series for both Canada and the USA which are found to satisfy separation in cointegration and which have two "common" cycles, but which are specific to each of the two countries so that a country by country BN decomposition of the series could lead to very misleading results.

References

[1] Abadir, K.M., K. Hadri and E. Tzavalis (1999). The Influence of VAR Dimensions on Estimator Biases, *Econometrica*, 67, 163-181.

[2] Ahn, S.K. (1997), Inference of Vector Autoregressive Models With Cointegration and Scalar Components, *Journal of the American Statistical Association*, 92, 350-356.

[3] Ahn, S.K. and G.C. Reinsel (1988), Nested Reduced Rank Autoregressive Models for Multiple Time Series, *Journal of the American Statistical Association*, 83, 849-856.

[4] Baltagi, B.H. and C. Kao (2000), Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey, in: B.H. Baltagi (editor), *Advances in Econometrics: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Vol. 15, (JAI Press: New York).

[5] Banerjee, A., Marcellino, M. and C. Osbat (2001), Testing for PPP: Should We Use Panel Methods?, IGIER working paper n.186, Università Bocconi.

[6] Campbell, J.Y. and N.G. Mankiw (1990), Permanent Income, Current Income, and Consumption, *Journal of Business and Economic Statistics*, 3, 265-279.

[7] Campbell, J.Y. and N.G. Mankiw (1991), The Response of Consumption to Income: A Cross-Country Investigation, *European Economic Review,* 35, 723-767.

[8] Cubadda, G. (1999), Common Cycles in Seasonal Non-Stationary Time Series, *Journal of Applied Econometrics*, 14, 273-291.

[9] Engle, R.F., D.F. Hendry and J.-F. Richard (1983), Exogeneity, *Econometrica*, 51, 277-304.

[10] Engle, R.F. and C.W.J. Granger (1987), Co-integration and Error Correction, Representation, Estimation and Testing, *Econometrica*, 55, 251-276.

[11] Engle, R.F. and S. Kozicki (1993), Testing for Common Features (with comments), *Journal of Business and Economic Statistics*, 11, 369-395.

[12] Ericsson, N.E. (1995), Conditional and Structural Error Correction Models, *Journal of Econometrics*, 69, 159-171.

[13] Ericsson, N.E., D.F. Hendry and G.E. Mizon (1998), "Exogeneity, Cointegration and Policy Analysis", *Journal of Business Economic and Statistics*, 16, 370-387.

[14] Gonzalo, J. and C.W.J. Granger (1995), Estimation of Common Long-Memory Components in Cointegrated Systems, *Journal of Business and Economic Statistics*, 33, 27-35.

[15] Gonzalo, J. and J.Y. Pitirakis (2000), Dimensionality Effect in Cointegration Analysis, in R. F. Engle and H. White (editors): *Cointegration, Causality and Forecasting, Festschrift in Honour of Clive W.J. Granger* , Oxford University Press.

[16] Granger, C.W.J. and N. Haldrup (1997), Separation in Cointegrated Systems and P-T Decompositions, *Oxford Bulletin of Economics and Statistics*, 59, 449-463.

[17] Groen, J.J. and F. Kleibergen (1999), Likelihood-Based Cointegration Analysis in Panels of Vector Error Correction Models, Discussion Paper TI 99-055/4, Tinbergen Institute, Erasmus University Rotterdam.

[18] Haffner, R.W., A.M. Kutan and S. Zhou (1997), Linkages in EMS Term Structures; Evidence from Common Trend and Transitory Components, *Journal on International Money and Finance*, 16, 595-607.

[19] Hall, S., S. Lazarova and G. Urga (1999) A Principal Components Analysis of Common Stochastic Trends in Heterogeneous Panel Data: Some Monte Carlo Evidence, *Oxford Bulletin of Economics and Statistics*, 61, 749 - 767

[20] Hansen, P.R. and S. Johansen (1998), *Workbook on Cointegration*, (Oxford University Press: Oxford).

[21] Harbo, I., S. Johansen, B. Nielsen and A. Rahbek (1998), Asymptotic Inference on Cointegration Rank in Partial Systems, *Journal of Business and Economic Statistics*, 16, 388-399.

[22] Hecq, A.(1998), Does Seasonal Adjustment Induce Common Cycles ?, *Economics Letters, 59, 289-297.*

[23] Hecq, A, F.C. Palm and J.P. Urbain (2000a), Permanent-Transitory Decomposition in VAR Models with Cointegration and Common Cycles, *Oxford Bulletin of Economics and Statistics*, 62, 511-532.

[24] Hecq, A, F.C. Palm and J.P. Urbain (2000b), Testing for Common Cyclical Features in Nonstationary Panel Data Models, in: B.H. Baltagi (editor), *Advances in Econometrics: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Vol. 15, (JAI Press: New York).

[25] Hecq, A, F.C. Palm and J.P. Urbain (2001), Testing for Common Cyclical Featurs in VAR Models with Cointegration, University Maastricht RM/01/002.

[26] Hendry, D.F. and G.E. Mizon (1993), Evaluating Dynamic Econometric Models by Encompassing the VAR, in: P.C.B.Phillips (editor), *Models, Methods and Applications of Econometrics*, (Basil Blackwell: Oxford).

[27] Hoffman, M. (1999), National Stochastic Trends and International Macroeconomic Fluctuations, European University Institute Florence Working Paper 99/26..

[28] Johansen, S. (1992), Cointegration in Partial Systems and the Efficiency of Single-Equation Analysis, *Journal of Econometrics,* 52, 389-402.

[29] Johansen, S. (1995), *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models* (Oxford University Press: Oxford)

[30] Johansen, S. (2001), The Asymptotic Variance of the Estimated Roots in a Cointegrated Vector Autoregressive Model, EUI Working Paper ECO No. 2001/1.

[31] Johansen, S. and K. Juselius (1990), Maximum Likelihood Estimation and Inference on Cointegration - With Applications to the Demand for Money, *Oxford Bulletin of Economics and Statistics*, 52, 169-210.

[32] Johansen, S. and A.R. Swensen (1999), Testing Exact Rational Expectations in Cointegrated Vector Autoregressive Models, *Journal of Econometrics*, 93, 73-91.

[33] King, R.G., Plosser, C.I., Stock, J.H. and M.W. Watson (1991), Stochastic Trends and Economic Fluctuations, *American Economic Review*, 81, 4, 819-840.

[34] Konishi, T. (1993), *Separation and Long-Run Non-Causality in a Cointegrated System*, PhD Dissertation, UCSD.

[35] Konishi, T. and C.W.J. Granger (1993), Separation in Cointegrated Systems, manuscript UCSD.

[36] Konishi, T., Ramsey, V.A. and C.W.J. Granger (1994), Stochastic Trends and Short-Run Relationships between Financial Variables and Real Activity, manuscript UCSD

[37] Larsson, R. and J.Lyhagen (1999) Likelihood-Based Inference in Multivariate Panel Cointegration Models, SSE Working Paper 331, Stockholm School of Economics.

[38] Phillips, P.C.B. and H. Moon (1999), Nonstationary Panel Data Analysis: An Overview of Some Recent Developments, mimeo, Yale University.

[39] Proietti, T. (1997), Short Run Dynamics in Cointegrated Systems, *Oxford Bulletin of Economics and Statistics*, 59, 405-422.

[40] Reinsel, G. C. and S. K Ahn (1992), Vector Autoregressive Models with Unit Roots and Reduced Rank Structure : Estimation, Likelihood Ratio Tests, and Forecasting, *Journal of Time Series Analysis*, 13, 353-375.

[41] Summers, R. and A. Heston (1991), The Penn World Table (Mark 5): An Expanded Set of International Comparisons, 1950-1988, *Quarterly Journal of Economics*, 327-368.

[42] Tiao, G.C. and R.S. Tsay (1989), Model Specification in Multivariate Time Series, *Journal of Royal Statistical Society*, series B, 51, 157-213.

[43] Urbain, J.-P. (1992), "On Weak Exogeneity in Error Correction Models", *Oxford Bulletin of Economics and Statistics*, 54, 187-207.

[44] Vahid, F. and R.F. Engle (1993), Common Trends and Common Cycles, *Journal of Applied Econometrics*, 8, 341-360.