Common Cyclical Features Analysis in VAR Models with Cointegration

Alain Hecq, Franz C. Palm+ and Jean-Pierre Urbain *
University Maastricht, Department of Quantitative Economics

September 29, 2003

Abstract

The paper considers n-dimensional VAR models for variables exhibiting cointegration and common cyclical features. Two specific reduced rank vector error correction models are discussed. In one, named the "strong form" and denoted by SF, the collection of all coefficient matrices of a VECM has rank less than n, in the other, named the "weak form" and denoted by WF, the collection of all coefficient matrices except the matrix of coefficient of error correction terms has rank less than n. The paper explores the theoretical connections between these two forms, suggests asymptotic tests for each form and examines the small sample properties of these tests by Monte Carlo simulations.

The paper proposes a sequential test procedure that is aimed at uncovering strong forms by examining weak forms. For GDP series for five Latin American countries, 1950-1999, the WF appears to be supported by the data. Imposing the WF parameter restrictions leads to an improvement of forecast accuracy for these data series.

Keywords : Serial correlation common features, reduced rank structure, cointegration.
JEL Classification : C32

*The authors want to thank Peter Boswijk, Jörg Breitung, Bertrand Candelon, Niels Haldrup, Cheng Hsiao, Joao Issler, Paolo Paruolo, Horst Zank, an associate editor and three anonymous referees for very useful comments on a previous draft of this paper. The usual disclaimer applies.
+ Corresponding author: Franz Palm, University Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands. f.palm@ke.unimaas.nl

1 Introduction and Motivation

The interest in comovements between economic variables leading to common cyclical features has arisen for instance because economic theory predicts such comovements and many economic variables exhibit strong correlation at various frequencies. The vast literature on cointegration has focussed on long-run comovements. More recently, the existence of short-run comovements between stationary time series, e.g. between first differenced cointegrated I(1), series has been analyzed (see Engle and Kozicki 1993; Vahid and Engle, 1993; Gouriéroux and Peaucelle, 1989; Tiao and Tsay, 1989). Among these approaches, the concept of serial correlation common features (SCCF hereafter) introduced by Engle and Kozicki (1993) appears to be useful. It means that stationary time series move together in a way such that there exist linear combinations of these variables which yield white noise processes and that their impulse response functions are collinear. In general, imposing these common features restrictions when they are appropriate will increase estimation efficiency (Lütkepohl, 1991) and accuracy of forecasts (Vahid and Issler, 2002) as will be shown in the empirical analysis of gross domestic product in five Latin American countries.

Parametric restrictions implied by economic theory can also lead to testable hypotheses within a common feature context. This is for example the case for real business cycle models (Engle and Issler, 1995; Issler and Vahid, 2001; Hecq, Palm and Urbain 2000a), models for heterogeneous consumers who are either myopic and liquidity constrained or rational (Vahid and Engle, 1993; Hecq, Palm and Urbain, 2000b; 2002), the efficient market hypothesis (Hecq, 2000) and more generally for rational expectation and present value models.

The aim of this paper is to analyze common cyclical features[1] in relation with cointegration. The strong assumption that some linear combination of the first differences of the variables in the model is white noise will be called a strong form reduced rank structure (SF). It corresponds to the case of serial correlation common features of the variables in first differences and assumes that the left null spaces of the short-run dynamics matrices and cointegrating matrix overlap. When SCCF appears to be too strong, one can test for the existence of cofeatures in the form of linear combinations of the variables differenced once, that are not white noise but have lower order dynamics than the individual variables. Tiao and Tsay (1989) for example, study this type of structure in a multivariate ARMA model.

[1]To avoid confusion, it should be noticed from the outset that the term common cyclical features refers to a particular type of commonality leading to specific reduced rank structures. This concept should not be confused with the concept of cycle used in business cycle analyses (see the discussion in Cubadda, 1999). On the other hand, the concept of common cycles (in contrast to common cyclical features) refers to the common transitory component in particular permanent-transitory decompositions (see Vahid and Engle, 1993; Hecq, Palm and Urbain, 2000a).

They call it scalar component model (SCM) while Vahid and Engle (1997) study this concept under the name of "codependent cycles" in a VAR. Cubadda and Hecq (2001) analyze another reduced rank concept named polynomial SCCF (PSCCF) which assumes the existence of a polynomial matrix which when premultiplying a VECM, reduces it to a vector white noise process.

We consider a natural weaker alternative assumption under which the common cyclical part is reduced to a white noise by taking a linear combination of the variables in the first differences adjusted for long-run effects. This case will be termed weak form reduced rank structures (WF). In a Vector Error Correction model (VECM), the WF implies that the collection of all coefficient matrices except the matrix coefficient of error correction terms has reduced rank and intersecting null spaces. The WF is attractive as it allows for different common factors generating respectively the long-run and short-run dynamics of economic variables. It is a necessary condition for the existence of first order codependent cycles in a VAR(2) as studied by Vahid and Engle (1997). As it is also a necessary condition for the SCCF, it is a natural hypothesis to be tested in sequential model specification.

Our framework is an extension of that of Vahid and Engle (1993) as we explicitly consider the WF and use it in reduced rank testing aimed at recovering the SF. In the presence of the WF only, the lower bound to the number of common cycles is one whereas under SF, there have to be at least r common cycles in the system, with r being equal to the cointegration rank. We study both the WF and the SF, taking into account the implications of the WF for the SF in modeling. Thereby, we do not impose a nesting structure on the null spaces of the model dynamics. Notice that Reinsel and Ahn (1992) briefly discuss a form similar to the WF. In general, they impose a nesting structure on the null spaces of the model dynamics. They do not discuss all the implications for the admissible number of common features.

The paper is organized as follows. In Section 2 we present different forms of reduced rank structures that arise in empirical work. We focus on the partially non-stationary vector autoregression that will be reparametrized as a VECM. The relationships between the strong and weak form reduced rank structures will be analyzed. The mixed form (MF) combining SF and WF will also be considered. Section 3 presents simple statistical procedures based on a two-step canonical correlation and maximum likelihood analysis that allow to test various kinds of reduced rank structures, in particular to check whether short and long-run matrices have a common left null space. In Section 4, we study the small sample behavior of common feature tests using Monte Carlo simulations. We show why the number of common feature vectors can be artificially bounded by a wrong assumption about the nature of the reduced rank structure. We present a testing strategy that allows us to study cointegration and other common features of unknown order in an integrated framework and we provide simulation results for cases where the number and the parameters of cointegrating vectors are estimated. Finally, Section 5 illustrates the relevance of different forms of reduced rank structures for the analysis of the co-movements among the GDP of five Latin American countries for the period 1950-1999. It also shows the impact on forecast accuracy of imposing WF or SF restrictions that were not rejected by the data. A final section concludes.

2 Reduced rank structures

Let us consider a Gaussian Vector Autoregression of finite order p (VAR(p)) model for an n-vector time series $\{y_t, t = 1, \ldots, T\}$:

$$y_t = \sum_{i=1}^{p} \Phi_i y_{t-i} + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{1}$$

for fixed values of $y_{-p+1}, ..., y_0$ and where ε_t is a n-dimensional homoscedastic Gaussian mean innovation process relative to $\Im_t = \{y_{t-1}, y_{t-2}, \ldots, y_1\}$ with nonsingular covariance matrix Ω. Let L denote the lag operator and define $\Phi(L) = I_n - \sum_{i=1}^{p} \Phi_i L^i$. We make the following assumption

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

1. $\operatorname{rank}(\Phi(1)) = r, 0 < r < n$, *so that* $\Phi(1)$ *can be expressed as* $\Phi(1) = -\alpha\beta'$, *with* α *and* β *both* $(n \times r)$ *matrices of full column rank* r;
2. *the characteristic equation* $|\Phi(\xi)| = 0$ *has* $n - r$ *roots equal to* 1 *and all other roots outside the unit circle.*

Assumption 1 implies (see Johansen, 1995) that the process y_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 $\Phi(L) = \Phi(1)L + \Phi^*(L)(1 - L)$, and defining $\Delta = (1 - L)$, we obtain the vector error correction model:

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{j=1}^{p-1} \Phi_j^* \Delta y_{t-j} + \varepsilon_t, \qquad t = 1, \ldots, T, \tag{2}$$

where $\Phi_0^* = I_n$, $\Phi_j^* = -\sum_{k=j+1}^{p} \Phi_k$ $(j = 1, \ldots, p-1)$. Note that for notational convenience, deterministic terms (constants, trends, ...) are omitted at this level of presentation. With the exception of some simulation results in Section 4, throughout this paper we will also assume that p is known. Serial correlation common feature (see Engle and Kozicki, 1993) holds for the VECM (2), if there exists a $(n \times s)$ matrix $\tilde{\beta}$, whose columns span the cofeature space, such that $\tilde{\beta}'\Delta y_t = \tilde{\beta}'\varepsilon_t$ is a s-dimensional vector mean innovation process with respect to the information available at time t, $\Im_t$.

Consequently, serial correlation common features arise if there exists a cofeature matrix $\tilde{\beta}'$ such that the following two conditions are satisfied:

$$\textbf{Assumption 2:} \qquad \tilde{\beta}'\Phi_j^* = 0_{(s\times n)}, \qquad j = 1 \dots p-1 \tag{3}$$

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

Assumption 2 implies that $\tilde{\beta}'$ must lie in the intersection of the left null spaces of the matrices describing the short-run dynamics. Given that $\Phi_j^* = -\sum_{k=j+1}^{p}\Phi_k$, $j = 1, \dots, p-1$ and $\Phi_p^* = -\Phi(1) = -(I_n - \sum_{j=1}^{p}\Phi_j)$, Assumption 3 implies that $\tilde{\beta}'(I_n - \Phi_1) = 0_{(s\times n)}$, e.g. Φ_1 must have eigenvalues equal to one with multiplicity s and the corresponding eigenvectors must lie in the intersection of the left null spaces of the Φ_j^* matrices. Note that if the ranges of the Φ_j^*'s matrices are nested, i.e. if $range(\Phi_{j+1}^*) \subseteq range(\Phi_j^*)$, a nested reduced rank structure arises (see e.g. Ahn and Reinsel, 1988). We consider the restrictions implied by (3) or by (3) and (4) without imposing further nesting of the ranges of the Φ_j^*'s. This leads us to distinguish the following two concepts*:*

Definition 1 (Strong Form Reduced Rank Structure): *If in addition to Assumption 1 (cointegration) both Assumptions 2 and 3 hold, the 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 $\tilde{\beta}$, whose columns span the cofeature space, such that $\tilde{\beta}'\Delta y_t = \tilde{\beta}'\varepsilon_t$ is a s-dimensional vector mean innovation process with respect to $\Im_t$.*

Definition 2 (Weak Form Reduced Rank Structure): *If in addition to Assumption 1 (cointegration) only Assumption 2 holds, the 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 $\tilde{\beta}$, whose columns span the cofeature space, such that $\tilde{\beta}'(\Delta y_t - \alpha\beta' y_{t-1}) = \tilde{\beta}'\varepsilon_t$ is a s-dimensional vector mean innovation process with respect to $\Im_t$.*

Remark (a) The SF is usually considered in the literature (see Engle and Kozicki, 1993, Vahid and Engle, 1993 among others). It leads to serial correlation common features (SCCF). In this paper we prefer to use the concept of SF in order to enable a formal comparison with the WF and to highlight the fact that the concept of SCCF generally applies to stationary vector processes irrespective of the presence or absence of cointegration. Under the SF, we may define a $(n(p-1)+r)\times 1$ vector $X^*_{t-1} = [\Delta y'_{t-1}, \ldots, \Delta y'_{t-p+1}, y'_{t-1}\beta]'$ and a $n\times(n(p-1)+r)$ matrix $\Phi^* = [\Phi^*_1, \ldots, \Phi^*_{p-1}, \alpha]$, so that (2) is written as

$$\Delta y_t = \Phi^* X^*_{t-1} + \varepsilon_t, \qquad t = 1, ..., T. \tag{5}$$

Under the assumption of a SF, Φ^* is of reduced rank $n-s$ and can be written as $\Phi^* = A^*[C^*_1, \ldots, C^*_{p-1}, C^*_p] = A^*C^*$, where A^* is $n \times (n-s)$ full column rank matrix and C^* is $(n-s) \times (n(p-1)+r)$ and $\tilde{\beta}' A^* C^* X^*_{t-1} = 0$, e.g. $\tilde{\beta} \in sp(A^*_\perp)$ where $A^*_\perp$ is the orthogonal complement[2] of A^*. Consequently, as pointed out by Vahid and Engle (1993), in a $n-$dimensional $I(1)$ vector process y_t with $r < n$ cointegrating vectors, if the elements of y_t have common cyclical features (given by $f_t = C^* X^*_{t-1}$) there can be at most $n-r$ linearly independent cofeature vectors that eliminate the common cyclical features since the cofeature matrix must lie in $sp(\alpha_\perp)$. The SF implies that $s \leq n-r$ and that the common dynamic factors f_t consist of linear combinations of the elements of X^*_{t-1}. The implications of the SF can be stated more formally as:

Lemma 1: For the SF, $sp(\alpha) \subseteq sp(\widetilde{\beta}_\perp)$.

The proof follows directly from the linear independence between the vectors β and $\widetilde{\beta}$ (see Vahid and Engle, 1993) so that rank $[\beta : \widetilde{\beta}] = r + s \leq n$. Hence we have that $\dim[sp(\alpha)] \leq \dim[sp(\widetilde{\beta}_\perp)]$ or that rank $(\alpha) \leq$ rank $(\widetilde{\beta}_\perp)$ implying that $r \leq n-s$.

Remark (b) In the case of WF, we analogously define a $n(p-1)\times 1$ vector $X_{t-1} = [\Delta y'_{t-1}, \ldots, \Delta y'_{t-p+1}]'$ and the $n \times n(p-1)$ matrix $\Phi = [\Phi^*_1, \ldots, \Phi^*_{p-1}]$, so that (2) becomes

$$\Delta y_t = \alpha\beta' y_{t-1} + \Phi X_{t-1} + \varepsilon_t, \qquad t = 1, ..., T. \tag{6}$$

Under the assumption of a WF, Φ is of reduced rank $n-s$ and can be written as $\Phi = A[C_1, \ldots, C_{p-1}] = AC$, where A is $n\times(n-s)$ full column rank matrix and C is $(n-s)\times n(p-1)$ such that $\tilde{\beta}' A C X_{t-1} = 0$. The cofeature matrix $\tilde{\beta}$ must lie in $space(A_\perp)$ but not necessarily in $space(\alpha_\perp)$.

It is important to stress the difference between SF and WF. Firstly, the assumption of a SF reduced rank rules out predictability at any frequency and hence implies common cycles

[2] In the sequel, space will be denoted by sp. We shall always 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. We then say that $B_\perp$ spans the null space of B and B' spans the left null space of $B_\perp$.

at all frequencies. On the contrary, by allowing for linear combinations that are predictable in the long run, the WF reduced rank structure only restricts the short-run dynamics. Secondly, in the WF case, both the possible number and the nature of the common cyclical features change: s may be greater than $n-r$ but has to remain $\leq n-1$ and the corresponding $n-s$ common dynamic factors consist of linear combinations of the elements of X_{t-1}, $f_t = CX_{t-1}$, which only contain lagged first differences of the process. It is important to notice that the existence of s weak form common feature vectors with $s > r$, implies the existence of $s-r$ strong form common features as is shown in Lemma 2.

Lemma 2: In the VAR model (1) under Assumption 1 with $s > r$, Assumption 2 implies the existence of $s-r$ SF common feature vectors.

Proof: Denote by $\widetilde{\beta}$ the $n \times s$ matrix of linearly independent WF common feature vectors. Any nonsingular transformation of $\widetilde{\beta}$, $\widetilde{\beta}A$, with A being an $s \times s$ nonsingular matrix, also forms a basis of the space spanned by the columns of $\widetilde{\beta}$ and therefore is also a basis of the WF common feature space. The matrix $\widetilde{\beta}'\Phi(1) = -\widetilde{\beta}'\alpha\beta'$ has rank $\min(r, s)$. Therefore, if $s > r$, there are $s-r$ linearly independent column vectors such that there is an $n \times (s-r)$ matrix B with full column rank such that $B'\alpha = 0$. B can be constructed as $B = \widetilde{\beta}A^*$ by choosing the $s \times (s-r)$ matrix A^* with rank $s-r$ such that B forms a basis for the left null space of α. Note that we can always normalize B such that the upper part equals I_{s-r}. $\square$

Remark (c) The interpretation of the WF differs from that of the SF. The WF implies that the serial correlation pattern in Δy_t and $\alpha\beta' y_{t-1}$ are the same, their impulse response functions are collinear, and their dynamics are similar. Interpreting $\alpha\beta' y_{t-1}$ as deviations from fundamentals, the WF implies that the dynamics of these deviations are similar to those of the change in y_t. Alternatively, under the WF, the short-run and long-run dynamics of y_t are unrelated. The absence of common determinants of both types of dynamics could be the result of differences between short-run (cyclical) and long run (structural) adjustment costs, etc.

Remark (d) As pointed out, the WF has an interest in its own as it is a necessary condition for the existence of first order codependent cycles in a VAR(2) (see e.g. Vahid and Engle, 1997; Hecq, 2000) and of the SF. The WF restrictions are generally not invariant to alternative vector error correction representations such as that where y_{t-p} appears in levels instead of y_{t-1}. The implications of the lack of invariance are that the results from a reduced rank analysis of short-run dynamics are parametrization-specific. Invariance may be obtained at the price of assuming a SCCF or that the ranges of Φ_j^*'s are nested (see e.g. Ahn and Reinsel, 1988). The methods put forward in this paper can be applied to any of these alternative parametrizations. We present the analysis for the VECM (2) with y_{t-1} appearing in levels, first, because this parametrization is frequently used in empirical work; second because if a reduced rank structure is found it will imply a lower order SCM than for other parametrizations; third, the WF is more likely to be appropriate as it applies to the coefficients of the higher order lags of Δy_t in the VECM, which are usually less significant than those of small order lags of Δy_t (for non-seasonal processes). Alternatively, when modeling series for which there are no strong reasons to a priori prefer any of the VECM parametrizations, one can test the WF restrictions for each parametrization. Note however that finding WF common features for each parametrization (even if the number of common features is the same for each parametrization) is a necessary but not sufficient condition for the existence of SF cofeatures. To illustrate this consider the simple case of a VAR(2), $y_t = \Phi_1 y_{t-1} + \Phi_2 y_{t-2} + \varepsilon_t$ that can be reparametrized as two observationally equivalent VECMs:

$$\Delta y_t = (\Phi_1 + \Phi_2 - I_n) y_{t-1} - \Phi_2 \Delta y_{t-1} + \varepsilon_t, \quad (7)$$

$$\Delta y_t = (\Phi_1 + \Phi_2 - I_n) y_{t-2} + (\Phi_1 - I) \Delta y_{t-1} + \varepsilon_t. \quad (8)$$

If the WF holds for (7), then there exists some $(n \times s)$ matrix Λ_1 such that $\Lambda_1' \Phi_2 = 0$. If the WF holds for (8), with the same value s, then there exists some $(n \times s)$ matrix Λ_2 such that $\Lambda_2'(\Phi_1 - I_n) = 0$. This does however not imply that there exists a $(n \times s)$ matrix Λ_3 such that $\Lambda_3' \Phi_2 + \Lambda_3'(\Phi_1 - I_n) = 0$. For this to be the case, the null spaces of $(\Phi_1 - I_n)$ and of (Φ_2) have to have a non empty intersection which differs from the zero vector. However, if SF holds for one parametrization of the WF, then the SF also holds for all parametrizations of the WF.

Next, one can test the SF restrictions for those parametrizations for which the WF restrictions are not rejected. This sequential testing is likely to lead to detecting useful structures in the data. Rejecting SF common features for each parametrization should be taken as strong evidence against the existence of SF cofeatures.

When $n > 2$, and $s - r > 0$, besides the $s - r$ SF common features implied by s WF common features, the mixed form (MF) reduced rank restrictions may arise. They combine the SF and the WF in the following way.

Definition 3 (Mixed Form Reduced Rank Structure): *If in addition to Assumption 1 (cointegration) Assumption 2 holds for s common feature vectors $\widetilde{\beta} = (\widetilde{\beta}_1, \widetilde{\beta}_2)$, with $\widetilde{\beta}_1$ and $\widetilde{\beta}_2$ being $n \times s_1$ and $n \times s_2$ full rank matrices respectively, with $s_1 + s_2 = s$, and in addition Assumption 3 holds for s_1 common feature vectors $\widetilde{\beta}_1$ with $s > s_1$ and $n - r > s_1 > \max(0, s - r)$, then the implied reduced rank structure of the VECM (2) will be labelled a mixed* *form reduced rank structure (MF). Under MF, the* $(n \times s)$ *matrix* $\widetilde{\beta}$ *spans the co-feature space, such that* $\widetilde{\beta}_1' \Delta y_t = \widetilde{\beta}_1' \varepsilon_t$ *is a* s_1*-dimensional vector mean innovation process with respect to* $\Im_t$ *and* $\widetilde{\beta}_2'(\triangle y_t - \alpha\beta' y_{t-1}) = \widetilde{\beta}_2' \varepsilon_t$ *is a* s_2*-dimensional vector mean innovation process to* $\Im_t$*.*

Remark (e) Under the MF, there are $s_1 - \max(0, s-r) > 0$ SF common feature vectors which are not implied by the WF and yield testable restrictions on the parameters of the VECM (2). The matrix $\widetilde{\beta}_1$ consists of $s - r$ columns which are linear combinations of $\widetilde{\beta}$ and $s_1 - \max(0, s-r)$ columns of $\widetilde{\beta}$ which satisfy Assumption 3.

Remark (f) Note that in the mixed case s_1 and s_2 have to satisfy the inequalities $s_1 + s_2 \leq n - 1$ and $s_1 \leq n - r$. Also, along the lines of lemma 1, we get $sp(\alpha) \supseteq sp(\widetilde{\beta}_\perp)$.

Notice that we could easily extend these representations in order to analyze models in which only a part of short-run components disappears. This type of reduced rank structures has been studied by Ahn and Reinsel (1988) for stationary processes, Tiao and Tsay (1989) for VARMA models and by Reinsel and Ahn (1992) and Ahn (1997) and Cubadda and Hecq (2001) for partially non-stationary processes.

3 Testing Different Forms of Reduced Rank Structures

3.1 Reduced rank hypotheses

The difference between the SF and the WF can be illustrated in terms of two competing models where we assume both cointegration and the existence of a $(n \times s)$ common feature matrix $\tilde{\beta}$. Under the assumption of SF the following model holds

$$\tilde{\beta}' \Delta y_t = \tilde{\beta}' \varepsilon_t, \qquad t = 1, ..., T, \tag{9}$$

while under WF we have

$$\tilde{\beta}' (\Delta y_t - \alpha\beta' y_{t-1}) = \tilde{\beta}' \varepsilon_t, \qquad t = 1, ..., T. \tag{10}$$

Let us first assume that the cointegrating rank r is known and fixed. For a given maintained reduced rank structure (WF or SF), we may consider the sequence of hypotheses (or models) separately in order to test $H_0 : rank(\widetilde{\beta}) \geq s$ against $H_a : rank(\widetilde{\beta}) < s$ for the different values of s starting with $s = 1$ (against the model without common features $s = 0$). In the SF case, the maximum number of common feature vectors is $n - r$. For the WF s has an upper bound[3]

[3] $s = n$ implies that $\Delta y_t - \alpha\beta' y_{t-1}$ is already a n-dimensional vector white noise process.

of $n-1$. Denote the largest values for which the null hypothesis is not rejected by s^*_{WF} and s^*_{SF} respectively ($s^*_{WF} \geq s^*_{SF}$).

For each value of s ($\leq n-r$) we can also compare the SF against the nesting alternative of a WF. But as the WF implies the strong form for values of $s > r$, it is sensible to compare the two for values of s, starting with $s = \max[1, s^*_{WF} - r + 1]$ up to and including $s = \min[s^*_{WF}, n-r]$. If r is unknown, it has to be determined first or the analysis has to be carried out for values of r taken as given.

3.2 Testing

Given that the hypotheses to be tested are nested, we rely on ML estimation of the underlying models following the approaches by Reinsel and Ahn (1992), Ahn (1997), Ahn and Reinsel (1988), Reinsel (1993) among others. Usually, when r and s are unknown, it appears impossible to find an explicit solution for the likelihood equations (see Johansen, 1995; Ahn, 1997). There are essentially two approaches to the determination of r, s and to the estimation of the parameters of interest. The first approach proposed and investigated by Ahn (1997), Ahn and Reinsel (1988) is to exploit the nested reduced rank structures and to compute numerically a Gaussian reduced-rank estimator based on iterative solution of approximate Newton-Raphson equations. Alternatively, one may follow a two-step approach in which r is first determined, while ignoring restrictions on the short-run dynamics of the model. Once r is determined and the cointegrating matrix β is estimated, s can be determined using the approach proposed by Vahid and Engle (1993) for example. The rationale behind this simple two-step analysis is that the determination of r and the efficiency of estimation of β are not affected asymptotically by the presence of the reduced rank structure on the short-run dynamics (see also Ahn, 1997; Phillips, 1991).

We use the two-step approach, although one may reasonably suspect small sample efficiency losses compared to using a one-step full information estimation method. As pointed out by various authors, a convenient way to test for reduced rank structures within the VECM is based on canonical correlation analysis. Let us first assume that r and β are known or that superconsistent estimates are available so that we may essentially consider them to be fixed and given.

Define the $T \times n$ matrices $W_1 = \Delta Y = (\Delta y_1, \ldots, \Delta y_T)'$, $Y_{-1} = (y_0, \ldots, y_{T-1})'$, $Z_1 = \Delta Y^*$ with ΔY^* being the LS residuals from the multivariate regression of ΔY on $Y_{-1}\beta$ and the $T \times (n(p-1)+r)$ matrix $W_2 = [Z_2, Y_{-1}\beta]$ with Z_2 being the $T \times n(p-1)$ matrix $(\Delta Y^*_{-1}, \ldots \Delta Y^*_{-p+1})$. Under the maintained hypothesis of a SF reduced rank structure, the sequence of common feature Gaussian likelihood ratio test statistics for $H_0 : rank(\Phi^*) \leq n-s$ against $H_a : rank(\Phi^*) > n-s$, where Φ^* is defined in (5), or equivalently for $H_0 : rank(\widetilde{\beta}) \geq s$ against $H_a : rank(\widetilde{\beta}) < s$ can be shown (see Lütkepohl, 1991; Velu et al, 1986)) to be

$$\xi_S = -T \sum_{i=1}^{s} \log(1-\lambda_i), \qquad s = 1, ..., n-r, \tag{11}$$

where $0 \leq \lambda_1 \leq \lambda_2 \leq ... \leq \lambda_{n-r} < 1$ are the ordered eigenvalues of the symmetric matrix $(W_1'W_1)^{-1/2}W_1'W_2(W_2'W_2)^{-1}W_2'W_1(W_1'W_1)^{-1/2}$. The test statistic (11) can also be interpreted as the minimum of the objective function of the GMM estimator of $\widetilde{\beta}$ subject to the normalization $(1/T)\ \widetilde{\beta}'W_1'W_1\widetilde{\beta} = I_s$ (see Anderson and Vahid, 1998). For known r and β, under the null the test statistic ξ_S is asymptotically χ^2-distributed with $s(n(p-1)+r) - s(n-s)$ degrees of freedom (Vahid and Engle, 1993).

In the case of WF reduced rank structure, this likelihood ratio test for $H_0 : rank(\widetilde{\beta}) \geq s$ against $H_a : rank(\widetilde{\beta}) < s$ reads as

$$\xi_W = -T \sum_{i=1}^{s} \log(1-\tilde{\lambda}_i), \qquad s = 1, ..., n-1, \tag{12}$$

where $0 \leq \widetilde{\lambda}_1 \leq \widetilde{\lambda}_2 \leq ... \leq \widetilde{\lambda}_{n-1} < 1$ are the ordered eigenvalues of the symmetric matrix $(Z_1'Z_1)^{-1/2}Z_1'Z_2(Z_2'Z_2)^{-1}Z_2'Z_1(Z_1'Z_1)^{-1/2}$. This statistic has an asymptotic χ^2-distribution with $s(n(p-1)) - s(n-s)$ degrees of freedom under the null. If β and r are unknown and have to be estimated, concentrating out the cointegrating vectors has no effect on the asymptotic distribution provided r is not underestimated (see Paruolo, 2002 for a formal proof). Intuitively, the $T-$consistency of the cointegrating vectors allows to consider them fixed before performing the common feature analysis in a second step. Similarly, overestimating the cointegrating rank leads to the presence of additional I(1) series whose weights in the multivariate LS regression converge to zero in probability in such an unbalanced system. Moreover, the Monte Carlo simulations in Section 4 also show good small sample behavior of the test statistics (11) and (12) when r is overestimated.

A MF reduced rank structure hypothesis $H_0 : rank(\tilde{\beta}_1) \geq s_1$, for $min(n-r, s) \geq s_1 > \max(0, s-r)$ and $rank(\tilde{\beta}_2) \geq s_2$ against $H_a : rank(\tilde{\beta}_1) < s_1$ or $rank(\tilde{\beta}_2) < s_2$, with $s_1+s_2 = s$, can be tested in several ways. One way is to test SF restrictions for $s_1 = 1, \ldots, n-r$, using the statistic ξ_S in (11). As this test ignores the restrictions implied by the existence of s_2 weak form common features, some power might be lost as will be illustrated in the next section.

Alternatively, the parameters $\tilde{\beta}$ and α from the WF can be estimated jointly by FIML for given s and β, e.g. by maximizing the likelihood function based on the $(s \times 1)$ subsystem

(10), normalized on the first s variables of Δy_t by setting $\tilde{\beta}' = (I_s \ \tilde{\beta}^{*\prime}_{s\times(n-s)})$, and completed by adding $(n-s)$ "reduced form" equations for the remaining $(n-s)$ variables in Δy_t

$$B'\Delta y_t \quad = \quad \begin{pmatrix} 0_{s\times n} & 0_{s\times n} & \dots & 0_{s\times n} & \alpha_1 \\ \Phi^*_{21} & \Phi^*_{22} & \dots & \Phi^*_{2p-1} & \alpha_2 \end{pmatrix} \hat{X}^*_{t-1} + B'\varepsilon_t, \tag{13}$$

with

$$B' = \begin{pmatrix} I_s & \tilde{\beta}^{*\prime}_{s\times(n-s)} \\ 0_{(n-s)\times s} & I_{n-s} \end{pmatrix},$$

$\hat{X}^*_{t-1} = X^*_{t-1}$ with β replaced by the first stage superconsistent estimate, the Φ^*_{2i} matrices, $i = 1, \dots, p-1$, indicate the $n-s$ bottom rows of the Φ^*_i matrices in (2) and $(\alpha'_1 \ \alpha'_2)$ is the partition of $\alpha' B$. Under a MF structure, for given β, s and s_1, we can specify a similar pseudo-structural system:

$$B'\Delta y_t = \begin{pmatrix} 0_{s_1\times n} & 0_{s_1\times n} & \dots & 0_{s_1\times n} & 0_{s_1\times r} \\ 0_{s_2\times n} & 0_{s_2\times n} & \dots & 0_{s_2\times n} & \alpha_2 \\ \Phi^*_{31} & \Phi^*_{32} & \dots & \Phi^*_{3p-1} & \alpha_3 \end{pmatrix} \hat{X}^*_{t-1} + B'\varepsilon_t, \tag{14}$$

where

$$B' = \begin{pmatrix} I_{s_1} & \tilde{\beta}^{*\prime}_{1,s_1\times(n-s_1)} \\ 0_{s_2\times s_1} & \tilde{\beta}'_{2,s_2\times(n-s_1)} \\ 0_{(n-s)\times s_1} & A_{(n-s)\times(n-s_1)} \end{pmatrix},$$

$\tilde{\beta}'_{2,s_2\times(n-s_1)} = (I_{s_2} \ \tilde{\beta}^{*\prime}_{2,s_2\times(n-s)})'$, $A_{(n-s)\times(n-s_1)} = (0_{(n-s)\times s_2} \ I_{n-s})$, the Φ^*_{3i} matrices, $i = 1, \dots, p-1$, indicate the $n-s$ bottom rows of the Φ^*_i matrices in (2) and $\alpha' B = (0'_{s_1\times r} \ \alpha'_2 \ \alpha'_3)$ with α_2 and α_3 of dimension $(s_2 \times r)$ and $(n-s)\times r$ respectively.

For given β and s, the MF with s_1 SF vectors and s_2 WF vectors can tested against the WF by testing for the validity of the additional parameter restrictions implied by (14) using a standard LR test statistics denoted by ξ_M. No efficiency loss arises if a superconsistent estimate is substituted for the cointegrating vectors β. Under the null of the MF, ξ_M is asymptotically χ^2-distributed with degrees of freedom given by the number of additional parametric restrictions imposed under (14), i.e. $s_1 r - s_2 s_1$. This estimation procedure has been used in the empirical analysis reported in Section 5.

For given r, a likelihood ratio test statistic for the null hypothesis of a SF against the alternative of a WF, for each possible common feature rank $s = \max(1, s^*_{WF} - r + 1), \ldots, \min(n - r, s^*_{WF})$, is given by

$$\xi_{SW} = -T \sum_{i=1}^{s} \log\{\frac{(1-\lambda_i)}{(1-\tilde{\lambda}_i)}\}, \tag{15}$$

where the $\widetilde{\lambda}_i$'s and the λ_i's are defined as above. Conditional on known r and β, all variables involved are weakly stationary both under the null and the alternative, so that standard asymptotic theory applies. ξ_{SW} has an asymptotic χ^2-distribution with degrees of freedom equal to the number of restrictions rs imposed under the H_0. If the null hypothesis is rejected, one can proceed further in determining s by testing the number of zero squared canonical correlations between Z_1 and Z_2. Note that the test statistics (11), (12) and (15) only enable to formally compare nested models. For model comparisons involving non-nested hypotheses, we propose to select the model which, for given p, r and β, minimizes one of the well-known model selection criteria (AIC, SBC, HQC) where, given that we have omitted deterministic terms, the number of parameters is $n(n(p-1)+r) - s(n(p-1)+r) + s(n-s)$ under the SF and $n(n(p-1)+r) - s(n(p-1)) + s(n-s)$ under the WF. These model selection criteria can be also used to select the optimal values for r and s given p (as we assumed in the preceding section) and have also recently been considered for common feature analysis by Vahid and Issler (2002) with unknown s and p.

4 Monte Carlo Results

In this section we present evidence on the finite sample behavior of the sequential test procedures put forward in Section 3.2. One should indeed be careful when interpreting the outcome of the three sequences of LR tests ξ_S, ξ_W and ξ_{SW}. Given that s is unknown, and given the sequential nature of the testing procedure, the significance levels of the individual tests in the sequence must be distinguished from the overall Type I error of the sequential testing procedure. Also, the above sequential procedures are essentially based on asymptotic properties such as the irrelevance of the reduced rank structure for the optimal estimation of β and the determination of r. A Monte Carlo experiment should shed some light on the finite sample behavior of the sequences of common features LR tests presented in the preceding section. We concentrate on three issues which we believe are particularly relevant for applications:

1. the size and power in finite samples of the common feature LR tests,

2. the possible effect of incorrectly specifying the number of cointegrating vectors and/or the lag length,

3. the performance of the two-step approach used in the empirical application in which the cointegrating vectors and their number are estimated in a first step by ML methods.

In order to address these issues we consider a simple trivariate data generating process (DGP) with $p = 2$ where we assume the existence of two common feature vectors, i.e. $s = 2$. Throughout the simulations,[4] p is fixed either to its true value $p = 2$ or to 4. Strong and weak form reduced rank structures are considered here. The DGP is a Gaussian VAR of order two written in VECM form. In order to provide some motivation for the choice of the DGP, we label the three variables as c_t, i_t, y_t for consumption, investment and real output. In line with a simple form of a neo-classical model (see King et al, 1988; Hecq et al., 2000a; Issler and Vahid, 2001), we assume the existence of two long-run relationships: $c_t - y_t$ and $i_t - y_t$.

$$\begin{bmatrix} \Delta c_t \\ \Delta i_t \\ \Delta y_t \end{bmatrix} = \begin{bmatrix} 0.2 & 0.1 & 0.1 \\ 0.8 & 0.4 & 0.4 \\ 0.4 & 0.2 & 0.2 \end{bmatrix} \begin{bmatrix} \Delta c_{t-1} \\ \Delta i_{t-1} \\ \Delta y_{t-1} \end{bmatrix} + \alpha\beta' \begin{bmatrix} c_{t-1} \\ i_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \tag{16}$$

$$\begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1.0 & 0.6 & 0.6 \\ 0.6 & 1.0 & 0.6 \\ 0.6 & 0.6 & 1.0 \end{bmatrix} \right).$$

The cofeature matrix associated with the DGP in (16) is given by

$$\tilde{\beta}' = \begin{bmatrix} 1 & -0.25 & 0 \\ 1 & 0 & -0.5 \end{bmatrix}.$$

It yields two linear combinations of the variables in the model that annihilate the short-run dynamics. In our experiments, the nature of the reduced rank structure depends on the choice of the values for α and β. Tables 1 and 2 illustrate the simulated rejection frequencies when the DGP has a SF reduced rank structure with $r = 1$ and $s = 2$. Tables 3 and 4 present the rejection frequencies when the DGP has a WF reduced rank structure with $s = 2$ and $r = 1$ or $r = 2$. The four tables report the rejection frequencies of the statistics (11), (12) and (15) for models assuming $r = 1$, $r = 2$ and r being determined using Johansen's trace test, using the correct lag length $p = 2$ (Tables 1 and 3) or setting the lag length equal to 4 (Tables 2 and 4). Notice that for the SF and with $n = 3$, the number of cointegrating vectors is by definition

[4] The size and power of codependence tests in the presence of either incorrectly specified lag length of the model, omission of a cointegrating vector, non-normal errors, or temporal aggregation have been extensively analyzed by Beine and Hecq (1999).

bounded to be equal to one in the DGP. We therefore present simulation results for models with the correct specification of the cointegrating rank as well as with over-specification of r.

In each case, we use 10,000 replications and a sample size of T=1000 and 100. The cointegration coefficients β are set equal to their estimated values obtained by ML estimation in a first stage (see Johansen, 1995). Conditionally on these estimates for β, α and $\tilde{\beta}$ are estimated by ML as described in Section 3.2. All simulations have been performed with GAUSS and the first 50 observations initialize the processes. The empirical (size unadjusted) power and size are given as percentage rejection frequencies. The nominal size used to obtain these rejection frequencies is fixed at 5% for each individual test.

Insert Tables 1-2 about here

Tables 1 and 2 report simulation results for a DGP with SF. Several remarks are worth to be made:

- In general, the differences between the results in Tables 1 and 2 are small. The inefficiency resulting from choosing too long lags is small.
- When the DGP has a SF and the number of cointegrating vectors is correctly specified, both ξ_S and ξ_W behave fairly well in detecting the two cofeature vectors. Note also that the sequence of the LR tests ξ_{SW} does not show any significant size distortion.
- If we estimate the number of cointegrating vectors or fix it at a value higher than the true r, the rejection frequency of ξ_S is distorted. ξ_W still behaves very well in detecting the correct number of common feature vectors. However the LR tests for SF versus WF display significant size distortions reaching 50% instead of the 5% chosen nominal level.
- Overall, the tests appear to reject too frequently the null hypothesis when the model is misspecified in some way (with the exception of lag length). The tests therefore tend to favor accepting models with fewer restrictions than the true model, implying thereby a loss of efficiency, but not a misspecification.

In the Tables 3 and 4, rejection frequencies for a DGP under WF restrictions are given.

Insert Tables 3-4 about here

We draw some conclusions from Tables 3 and 4 for the DGP with WF common features:

- Again, the effect of overfitting the lag length is small.

- When the DGP with $r = 1$ has a WF reduced rank structure, ξ_W determines without size distortions the correct number of common feature vectors, whether r is fixed at the true value, estimated or fixed at 2. When the true value of r equals 2, ξ_W performs very well except when r is fixed at 1.

- The statistic ξ_S detects a SF reduced rank structure implied by a WF reduced rank structure ($s - r > 0$) with a rejection frequency of approximately 5% when r is correctly specified (panel one). When r is fixed at a value larger than the true one or when it is estimated, the size of ξ_S is much larger than the nominal size of 5%. For ξ_{SW}, the rejection frequencies are similar.

- It is interesting to note that the sequence of ξ_W still selects the correct number of common feature vectors without size distortions when we overspecify the number of cointegrating vectors. This is not surprising since the coefficient of a non significant $I(1)$ variable in a $I(0)$ model converges in probability to zero. ξ_S still rejects the presence of any cofeature vector since this case excludes the existence of an implied SF ($s - r = 0$).

- Overall, the likelihood ratio statistics ξ_{SW} for the null of SF against the WF has high power close to one in most cases. When $s - r > 0$ in the DGP, there are $(s - r)$ implied SF common feature vectors and the rejection frequencies for $s = 1$ in Tables 3 and 4 have to be interpreted as an empirical size of the test. In these cases, the statistic ξ_{SW} rejects too frequently the (implied) null hypothesis.

- When r is determined from the data using Johansen's trace test, the rejection frequencies are close to those for the cases where r equals its true value.

Results for the statistics presented above with a small sample correction as suggested by Reinsel and Ahn (1992) for cointegration tests, where ξ_W and ξ_S are respectively premultiplied by the factors $(T - n(p-1))/T$ and $(T - n(p-1) - r)/T$, (for further details see Hecq, 2000), have been obtained as well. They are available from the authors upon request. Overall, the results are similar to the corresponding results given in Tables 1-4. For $T = 100$, in some instances, the corrected version of the tests performs better than the uncorrected ones.

Table 5 contains some illustrative simulation results for a DGP with a MF reduced rank. For this purpose, the DGP is slightly modified and extended in order to account for a MF. The selected DGP is a VAR(2) with $n = 4$, $r = 2$ and $s = 3$. From Lemma 2 there is one implied cofeature vector ($s - r = 1$). The loading matrix α is chosen such that the DGP displays one additional cofeature vectors, i.e. $s_1 = 2$. The following matrices are retained:

$$\tilde{\beta} = \begin{bmatrix} 1 & 1 & 1 \\ -.25 & 0 & 0 \\ 0 & -.5 & 0 \\ 0 & 0 & -.4 \end{bmatrix}, \tilde{\beta}_\perp = \begin{bmatrix} -.1 \\ -.4 \\ -.2 \\ -.25 \end{bmatrix}, \alpha = \begin{bmatrix} -.2 & .2 \\ -.8 & 0 \\ -.4 & .8 \\ -.5 & 0 \end{bmatrix}, \beta = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1.2 & -.8 \\ -1 & -1 \end{bmatrix}.$$

This particular choice of α implies the existence of $\tilde{\beta}_1'$ satisfying[5] $\tilde{\beta}_1'\alpha = 0$. As discussed in the preceding section, we report results for ξ_S and a likelihood ratio tests of the mixed form denoted by ξ_M.

In Table 5 we report rejection frequencies, based on 10,000 replications, under the correct assumption of a mixed form with $s_1 = 2$ (size of the tests) as well as those obtained when we let the parameter $\alpha_{3,1}$ successively take the values -0.45, -0.5 which implies the existence of a weak form[6]. In all the cases, the empirical power is not size adjusted.

Insert Table 5 about here

From Table 5, we observe that ξ_S and ξ_M do not suffer from serious size distortion. With respect to the empirical powers, it appears that ξ_M performs substantially better than ξ_S. Remark that r is assumed known while β is estimated and thus the cointegrating rank is correctly specified.

The limited Monte Carlo evidence presented in this section leads us to propose the following model selection strategy.

1. Start by determining the lag length p and the number of cointegrating vectors, trying to avoid underestimation of r. In practice, Johansen's ML statistics complemented by a visual inspection may prove useful to determine an upper bound for r,

2. compute the sequences of common feature LR tests ξ_S and ξ_W and select s for the SF and WF respectively (denoted by s^*_{SF} and s^*_{WF}), check whether the number of WF common features exceeds r, in which case the WF implies $s^*_{WF} - r$ SF common features,

[5]The columns of the 4×2 matrix $\tilde{\beta}_1$ may simply be constructed by adding the first and second column of $\tilde{\beta}$ on the one hand and by adding the second and the third one on the other hand.

[6]Remark that the values of $\alpha_{3,1}$ chosen for the computation of the empirical powers only imply small deviations from the mixed form. For other values the empirical power rapidly reaches 1.

3. for the cases where $s = \max(1, s^*_{WF} - r + 1), ..., \min(n - r, s^*_{WF}$ compute ξ_{SW} to select the appropriate reduced rank structure,

4. for the cases where $s = \max(1, s^*_{WF} - r + 1), ..., \min(n - r, s^*_{WF})$, compute a likelihood ratio MF test,

Alternatively, one can use information criteria to compare the various forms.

5 Common Cycles and Common Trends in Latin America

In this section we test for the presence of comovements in annual GDP series for five Latin American countries: Brazil, Argentina, Mexico, Peru and Chile. The series are derived from the Total Economy Database[7] and span the period 1950-1999. We provide evidence for the importance of imposing short-run reduced rank restrictions when forecasting. We examine the presence of WF and SF reduced rank structures in a VECM model for the real gross domestic product of these five major Latin American economies. Figures 1 and 2 present both the log-levels and the growth rates of these variables. From Figure 1, it appears that these series display some similar trending behavior. However, by visual inspection of Figure 2 it is difficult to detect the form of short-run comovements. Cointegration and common feature tests have been performed to determine the number of stochastic trends and cycles which these series share.

INSERT FIGURES 1 AND 2 ABOUT HERE

An unrestricted VAR with three lags seems to capture appropriately the dynamics of this multivariate process. Johansen's ML tests detect two cointegrating vectors for the model with a constant term only and three cointegrating vectors for the model with a deterministic trend constrained in the long-run. The time trend is used to capture the possible presence of a deterministic convergence process or of trend stationary variables. The modules of the largest roots of the companion matrix are respectively (0.9904, 0.9843, 0.9843, 0.9280, 0.9280, 0.8528) and (0.9826, 0.9826, 0.8970, 0.8970, 0.7733, 0.7733) for the model with a constant and the model with a deterministic trend, suggesting the presence of respectively three and two common trends in these two specifications. Exclusion restriction tests of the hypothesis that

[7]University of Groningen and The Conference Board, GGDC Total Economy Database, 2002, www.eco.rug.nl/ggdc. The variables are expressed in US dollars for the base-year 1990 and converted at "Geary-Khamis" purchasing power parities.

one of the countries can be excluded from respectively the two and the three cointegrating vectors lead to $p-$values smaller than 0.001.

For the analysis, the determination of the number of common feature vectors is important. Tables 6 and 7 report the eigenvalues, the value of the log-likelihood, the $p-value$ associated with the asymptotic test statistic for the null hypothesis that there exist at least s cofeature vectors, as well as the value of the HQ criterion. From the Tables, it turns out that with $r = 2$, we cannot reject, both using LR tests and information criteria, the presence of three co-feature vectors of each kind (WF and SF). The likelihood ratio statistic for the null of SF against WF for $r = 2$ and $s = 3$ is given by $\xi_{SW} = \xi_S - \xi_W = -2(\text{loglik}^S\text{-loglik}^W) = 3.65$. ξ_{SW} follows a $\chi^2_{(6)}$ asymptotic distribution, so we do not reject the SF model. Because the SF imposes more restrictions than the WF, it is selected as our favorite parsimonious model. Notice that one of the SF vectors has been implied by the WF because $s - r = 1$. Once we include a linear trend in the long-run relations however, we do not reject the hypothesis of three cointegrating vectors. Consequently, due to the constraint $r + s \leq n$ under SF, we are only able to detect two co-feature vectors. However, as shown in the simulations, the WF test is robust to an overestimation of the number of long-run relationships and still provides evidence in favor of three co-feature vectors. We would reach the same conclusion when using the HQ criterion.

Insert Tables 6 and 7 about here

Our interest is also in the impact of reduced rank structures on forecast accuracy. Both Monte Carlo simulations and empirical analyzes in Vahid and Issler (2002) show substantial gains in forecast accuracy from imposing the restrictions implied by the reduced-rank structures. We compare the 1-step ahead RMSE of different specifications.[8] For the model with the unrestricted intercept we consider (i) the VECM with $r = 2$, (ii) the VAR in first differences with two lags (DVAR),[9] (iii) the SF and (iv) the WF with in both cases $r = 2$ and $s = 3$. For the model with a time trend in the long run we compare (i) the VECM with $r = 3$, (ii) the DVAR, (iii) the SF with $r = 3$ and $s = 2$, (iv) the WF with $r = 3$ and $s = 3$ and (v) a MF in which we cannot reject the null hypothesis ($p-$value is 0.36) that two of the WF relationships are also of the SCCF type.

[8] Notice that we keep $p = 3$ fixed in our analysis. See Vahid and Issler (2002) for an alternative approach.

[9] Notice that although the DVAR is misspecified due to the omission of long-run relationships, Clements and Hendry (1999) show that DVAR forecasts appear to provide immunity against parameter non-constancy that could seriously bias VECM forecasts.

The model is estimated for the period 1950-1991. The last 8 years are used to evaluate the model forecasts. During the period to be forecast two major financial crises hit Latin American countries in 1994 and 1998. Tables 8 and 9 report the RMSE for the 1-step ahead forecasts. In order to separate the issue of the non-constancy of the long-run relationships from the impact of imposing additional reduced rank restrictions, the cointegrating vectors were obtained using the full sample.

Insert Tables 8 -9

From Tables 8 and 9 we can conclude that imposing additional short-run restrictions improves the forecast accuracy. The only exception is Chile for which the VAR in first differences outperforms the other three (or four) specifications. This may be an indication of parameter non-constancy in the long run. For the specification with $r = 2$, the SF gives smaller RMSE's errors for Argentina and Peru while the WF is preferred for Brazil and Mexico. For the model with $r = 3$ and the restricted trend, the WF gives more accurate forecasts than the SF in almost all cases. The MF provides even better results in terms of forecasting performance.

6 Conclusion

In this paper, we studied a linear Gaussian VAR model with nonstationary but cointegrated variables that have common cyclical features.

We introduced the concepts of strong, weak and mixed form reduced rank structures and discussed their implications for VAR modeling. SF reduced rank structures arise when the common features are such that there exists one or several linear combinations of the set of variables under investigation expressed in first differences which are white noise. The existence of a WF reduced rank structure implies that linear combinations of the first differences of the variables in the model in deviation from the long-run relationships are white noise. We showed that the constraint that the number of common features plus the number of cointegrating relationships should be less than or equal to the number of variables no longer applies under the WF. This allows to consider a larger number of common feature relationships between the variables in first differences. Imposing the restrictions implied by WF leads to an efficiency increase for the estimates, resulting from the reduction in the number of free parameters to be estimated.

We designed a modeling strategy and proposed likelihood ratio tests for the three types of reduced rank structures. We studied the small sample properties of the test using Monte Carlo simulations. It appeared that in particular under SF it is of great importance to correctly determine the cointegrating rank before testing SF against WF. The application shows the presence of both long-run and short-run relationships among the real gross domestic product of five Latin American countries. Moreover, it appears that imposing MF restrictions gives the smallest RMSE for all countries but Chile.

Finally, it is worth noticing three extensions to the present paper. Hecq, Palm and Urbain (2000b) apply a common feature analysis in a dynamic panel context to tackle the problem of the high dimensionality of the parameter space. Hecq, Palm and Urbain (2002) test for SF and WF common features together with separation both in the long run and the short run. Paruolo (2002) develops WF and SF reduced rank structures for processes integrated of order 2.

References

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

[2] 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.

[3] Anderson, H.M. and F. Vahid (1998), Testing Multiple Equation Systems for Common Nonlinear Components, *Journal of Econometrics*, 84, 1-36.

[4] Beine, M. and A. Hecq (1999), Inference in Codependence : Some Monte Carlo Results and Applications, *Annales d'Economie et de Statistique*, 54, 69-90.

[5] Clements, M.P. and D.F. Hendry (1999), *Forecasting Non-stationary Economic Time Series*, The MIT Press.

[6] Cubadda, G. (1999), Common Serial Correlation and Common Business Cycles: A Cautious Note, *Empirical Economics*, 24, 529-535.

[7] Cubadda, G. and A. Hecq (2001), On Non-contemporaneous Short-Run Comovements, *Economics Letters*, 73, 389-397.

[8] Engle, R. F. and J.V. Issler (1995), Estimatign Common Sectoral Cycles, *Journal of Monetary Economics*, 35, 83-113.

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

[10] Gouriéroux, Ch. and I. Peaucelle (1989), Detecting a Long-Run Relationship, CEPREMAP Discussion Paper 8902.

[11] Hecq, A. (2000), *Common Cyclical Features in Multiple Time Series and Panel Data: Methodological Aspects and Applications*, Ph. D. Thesis, University Maastricht.

[12] 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.

[13] Hecq, A., F.C. Palm and J.-P. Urbain (2000b), Testing for Common Cyclical Features in Nonstationary Panel data, in B.H. Baltagi (editor) *Advances in Econometrics: Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Vol. 15, JAI Press, 131-160.

[14] Hecq, A., F.C. Palm and J.-P. Urbain (2002), Separation, Weak Exogeneity and P-T Decomposition in Cointegrated VAR Systems with Common Features, *Econometric Reviews*, 21, 273-307.

[15] Issler, J. V. and F. Vahid (2001), Common Cycles and the Importance of Transitory Shocks to Macroeconomic Aggregates, *Journal of Monetary Economics*, 47, 449-475.

[16] King, R.G., Plosser, C.I. and S. Rebelo (1988), Production, Growth and Business Cycles II, *Journal of Monetary Economics*, 21, 309-341.

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

[18] Lütkepohl, H. (1991), *Introduction to Multiple Time Series Analysis*, (Springer Verlag: Berlin).

[19] Paruolo, P. (2002), Common Features and Common I(2) trends in VAR Systems, mimeo Dept. of Economics, University of Insubria.

[20] Phillips, P.C.B. (1991), Optimal Inference in Cointegrated Systems, *Econometrica*, 59, 283-306.

[21] Reinsel, G. C. (1993), *Elements of Multivariate Time Series Analysis*, (Springer Verlag: Berlin).

[22] 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.

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

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

[25] Vahid, F. and R. F. Engle (1997), Codependent Cycles, *Journal of Econometrics*, 80, 199-221.

[26] Vahid, F. and J.V. Issler (2002), The Importance of Common-Cyclical Features in VAR Analysis: A Monte-Carlo Study, *Journal of Econometrics*, 109, 341-363.

[27] Velu, R. P. , Reinsel, G. C. and D. W. Wichern (1986), Reduced Rank Models for Multivariate Time Series, *Biometrika*, 73, 105-118.

Table 1: Empirical Rejection Frequencies of the LR tests for SF

DGP: SF	**Estimated Model**						
$r = 1$	$r = 1, p = 2$						
$p = 2$		$T = 1000$			$T = 100$		
$s = 2$		ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
$\alpha = \begin{bmatrix} -0.10 \\ -0.40 \\ -0.20 \end{bmatrix}$	$s \geq 1$	0.24	0.41	0.78	0.33	0.53	0.92
	$s \geq 2$	4.90	5.00	5.37	6.72	6.53	6.13
$\beta' = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$	$s = 3$	100.00	100.00	100.00	100.00	100.00	93.96
	$r = 2, p = 2$						
		$T = 1000$			$T = 100$		
		ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
	$s \geq 1$	1.76	0.41	4.09	2.37	0.66	4.81
	$s \geq 2$	34.59	4.91	50.24	40.88	7.29	53.58
	$s = 3$	100.00	100.00	100.00	100.00	100.00	98.12
	$r = \hat{r}, p = 2$						
		$T = 1000$			$T = 100$		
		ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
	$s \geq 1$	0.52	0.42	1.34	0.78	0.61	1.72
	$s \geq 2$	14.19	5.02	15.85	18.26	7.12	18.87
	$s = 3$	100.00	100.00	100.00	100.00	100.00	94.84

- The rejection frequencies are based on 10,000 replications and calculated using asymptotic critical values. The nominal level is fixed at 5%.
- $\hat{r}$ has been determined in the first step using Johansen's trace test.

Table 2: Empirical Rejection Frequencies of the LR tests for SF

DGP: SF	**Estimated Model**						
$r = 1$, $p = 2$, $s = 2$, $\alpha = \begin{bmatrix} -0.10 \\ -0.40 \\ -0.20 \end{bmatrix}$, $\beta' = \begin{bmatrix} 0 & 1 & -1 \end{bmatrix}$	$r = 1, p = 4$						
		$T = 1000$			$T = 100$		
		ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
	$s \geq 1$	0.26	0.21	2.54	0.48	0.53	3.09
	$s \geq 2$	5.52	5.49	5.36	10.38	10.17	6.57
	$s = 3$	100.00	100.00	100.00	100.00	100.00	93.45
	$r = 2, p = 4$						
		$T = 1000$			$T = 100$		
		ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
	$s \geq 1$	1.10	0.23	12.94	2.20	0.77	13.99
	$s \geq 2$	21.97	5.61	49.97	32.78	11.06	51.62
	$s = 3$	100.00	100.00	100.00	100.00	100.00	97.77
	$r = \hat{r}, p = 4$						
		$T = 1000$			$T = 100$		
		ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
	$s \geq 1$	0.63	0.22	4.51	1.14	0.67	5.43
	$s \geq 2$	11.11	5.61	15.99	18.66	11.14	19.17
	$s = 3$	100.00	100.00	100.00	100.00	100.00	94.50

- The rejection frequencies are based on 10,000 replications and calculated using asymptotic critical values. The nominal level is fixed at 5%.
- $\hat{r}$ has been determined in the first step using Johansen's trace test.

Table 3: Empirical Rejection Frequencies of the LR tests for WF

DGP: WF	Estimated Model		$T = 1000$			$T = 100$		
			ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
$r = 1$, $p = 2$, $s = 2$, $\alpha = \begin{bmatrix} -0.50 \\ 0.10 \\ 0.20 \end{bmatrix}$, $\beta' = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$	$r = 1, p = 2$	$s \geq 1$	5.10	0.34	10.78	6.22	0.50	12.34
		$s \geq 2$	100.00	5.03	100.00	100.00	7.59	100.00
		$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00
	$r = 2, p = 2$	$s \geq 1$	30.11	0.35	40.57	36.08	0.60	46.32
		$s \geq 2$	100.00	4.88	100.00	100.00	7.56	100.00
		$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00
	$r = \hat{r}, p = 2$	$s \geq 1$	11.69	0.36	17.35	16.45	0.54	22.82
		$s \geq 2$	100.00	4.98	100.00	100.00	7.80	100.00
		$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00
$r = 2$, $p = 2$, $s = 2$, $\alpha = \begin{bmatrix} -0.50 & -0.20 \\ 0.10 & -0.30 \\ 0.20 & 0.20 \end{bmatrix}$, $\beta' = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$	$r = 2, p = 2$	$s \geq 1$	100.00	0.40	100.00	99.84	0.49	99.91
		$s \geq 2$	100.00	5.17	100.00	100.00	6.98	100.00
		$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00
	$r = 1, p = 2$	$s \geq 1$	100.00	5.18	100.00	98.00	5.24	98.61
		$s \geq 2$	100.00	100.00	100.00	100.00	97.11	99.96
		$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00
	$r = \hat{r}, p = 2$	$s \geq 1$	100.00	0.40	100.00	99.84	0.49	99.91
		$s \geq 2$	100.00	5.17	100.00	100.00	6.98	100.00
		$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00

- The rejection frequencies are based on 10,000 replications and calculated using asymptotic critical values. The nominal level is fixed at 5%
- $\hat{r}$ has been determined in the first step using Johansen's trace test.

Table 4: Empirical Rejection Frequencies of the LR tests for WF

DGP: WF

$r = 1$

$p = 2$

$s = 2$

$$\alpha = \begin{bmatrix} -0.50 \\ 0.10 \\ 0.20 \end{bmatrix}$$

$$\beta' = \begin{bmatrix} 1 & 0 & -1 \end{bmatrix}$$

Estimated Model

$r = 1, p = 4$

	$T = 1000$			$T = 100$		
	ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
$s \geq 1$	5.43	0.18	36.10	8.23	0.82	35.07
$s \geq 2$	100.00	5.68	100.00	100.00	12.85	99.98
$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00

$r = 2, p = 4$

	$T = 1000$			$T = 100$		
	ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
$s \geq 1$	18.56	0.19	57.14	25.35	0.62	58.96
$s \geq 2$	100.00	5.60	100.00	100.00	11.40	100.00
$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00

$r = \hat{r}, p = 4$

	$T = 1000$			$T = 100$		
	ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
$s \geq 1$	9.87	0.19	40.73	15.66	0.72	42.59
$s \geq 2$	100.00	5.69	100.00	100.00	12.39	100.00
$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00

DGP: WF

$r = 2$

$p = 2$

$s = 2$

$$\alpha = \begin{bmatrix} -0.50 & -0.20 \\ 0.10 & -0.30 \\ 0.20 & 0.20 \end{bmatrix}$$

$$\beta' = \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{bmatrix}$$

Estimated Model

$r = 2, p = 4$

	$T = 1000$			$T = 100$		
	ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
$s \geq 1$	100.00	0.18	100.00	99.42	0.59	99.95
$s \geq 2$	100.00	5.67	100.00	100.00	11.57	100.00
$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00

$r = 1, p = 4$

	$T = 1000$			$T = 100$		
	ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
$s \geq 1$	100.00	5.78	100.00	99.12	10.15	98.70
$s \geq 2$	100.00	100.00	100.00	100.00	99.68	99.96
$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00

$r = \hat{r}, p = 4$

	$T = 1000$			$T = 100$		
	ξ_S	ξ_W	ξ_{SW}	ξ_S	ξ_W	ξ_{SW}
$s \geq 1$	100.00	0.18	100.00	99.43	0.65	99.95
$s \geq 2$	100.00	5.67	100.00	100.00	12.40	100.00
$s = 3$	100.00	100.00	100.00	100.00	100.00	100.00

- The rejection frequencies are based on 10,000 replications and calculated using asymptotic critical values. The nominal level is fixed at 5%.
- $\hat{r}$ has been determined in the first step using Johansen's trace test.

Table 5: Empirical Rejection Frequencies of the LR tests for MF, $p = 2$

	T	**Size**	**Power**	**Power**
		$\alpha_{3,1} = -0.4$	$\alpha_{3,1} = -0.45$	$\alpha_{3,1} = -0.5$
ξ_S	100	8.22	15.04	28.29
	1000	5.15	78.20	99.35
ξ_M	100	6.87	19.28	38.93
	1000	5.03	91.42	100

- The nominal level is fixed at 5%.
- The statistics ξ_S and ξ_M use the estimated $\hat{\beta}$ under the assumption of known cointegrating rank $r = 2$.

Table 6: Common Features Test Statistics, $r = 2$ (constant only)

	SF			WF		
	λ_i	$p - val$	$loglik^S$	$\tilde{\lambda}_i$	$p - val$	$loglik^W$
$s \geq 1$	0.08	0.83	828.727	0.08	0.65	828.748
$s \geq 2$	0.21	0.64	823.219	0.17	0.49	824.113
$s \geq 3$	0.24	0.54	816.579	0.21	0.41	818.407
$s \geq 4$	0.72	(< 0.001)	786.177	0.59	<0.001	797.239
$s = 5$	0.82	(< 0.001)	746.136	0.69	< 0.001	769.381

Table 7: Common Features Test Statistics, $r = 3$ (deterministic trend)

	SF			WF		
	λ_i	$p - val$	$loglik^S$	$\tilde{\lambda}_i$	$p - val$	$loglik^W$
$s \geq 1$	0.09	0.86	843.73	0.08	0.68	844.069
$s \geq 2$	0.21	0.71	837.971	0.17	0.53	839.632
$s \geq 3$	0.55	(0.01)	819.051	0.27	0.27	832.141
$s \geq 4$	0.75	(<0.001)	786.267	0.62	<0.001	809.077
$s = 5$	0.82	(<0.001)	746.136	0.70	<0.001	780.489

Table 8: RMSE of 1-step ahead Forecasts (unrestricted constant)

	VECM $r = 2, s = 0$	DVAR $r = 0, s = 0$	SF $r = 2, s = 3$	WF $r = 3, s = 3$
Argentina	0.0693	0.0810	**0.0489**	0.0499
Brazil	0.0215	0.0247	0.0148	**0.0130**
Chile	0.0482	**0.0372**	0.0605	0.0594
Mexico	0.0414	0.0474	0.0321	**0.0318**
Peru	0.0472	0.0636	**0.0344**	0.0359

- Bold figures indicate the model with the lowest RMSE.

Table 9: RMSE of 1-step ahead Forecasts (linear trend in the long run)

	VECM $r = 3, s = 0$	DVAR $r = 0, s = 0$	SF $r = 3, s = 2$	WF $r = 3, s = 3$	MF
Argentina	0.0620	0.0810	0.0513	0.0485	**0.0473**
Brazil	0.0236	0.0247	0.0161	0.0167	**0.0155**
Chile	0.0482	**0.0372**	0.0467	0.0566	0.0596
Mexico	0.0339	0.0474	0.0340	0.0279	**0.0277**
Peru	0.0434	0.0636	0.0456	0.0308	**0.0303**

- Bold figures indicate the model with the lowest RMSE.

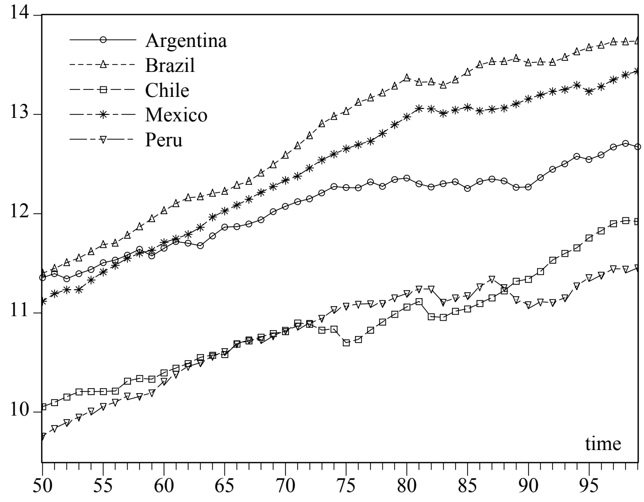

Figure 1: Log-levels of Real Gross Domestic Product (1950-1999)

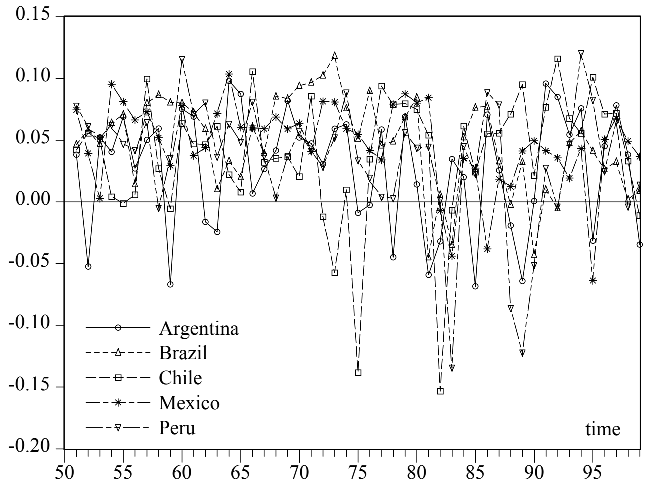

Figure 2: Growth Rates of Real Gross Domestic Product (1950-1999)