Selection of Optimal Lag Length in Cointegrated VAR Models with Weak Form of Common Cyclical Features*

CARLOS ENRIQUE CARRASCO GUTIÉRREZ†

28th November 2007

Abstract

An important aspect of empirical research based on the vector autoregressive (VAR) model is the choice of the lag order, since all inference in the VAR model depends on the correct model specification. Literature has shown important studies of how to select the lag order of a nonstationary VAR model subject to cointegration restrictions. In this work, we consider an additional *weak form* (WF) restriction of common cyclical features in the model in order to analyze the appropriate way to select the correct lag order. Two methodologies have been used: the traditional information criteria (AIC, HQ and SC) and an alternative criterion ($IC(p, s)$) which select simultaneously the lag order p and the rank structure s due to the WF restriction. A Monte-Carlo simulation is used in the analysis. The results indicate that the cost of ignoring additional WF restrictions in vector autoregressive modelling can be high, especially when SC criterion is used.

KEY WORDS: Cointegration; Common Cyclical Features; Reduced Rank Model; Estimation; Information Criteria.

JEL Codes: C32, C53.

---

*Acknowledgments: I am grateful to comments and suggestions given by João Victor Issler, Wagner Gaglianone, Ricardo Cavalcanti, Luiz Renato Lima and participants of the Brazilian Econometric Meeting 2006 and European Meeting of the Econometric Society 2007. Special thanks are due to Alain Hecq for solving doubts and comments. The authors are responsible for any remaining errors in this paper. Carlos Enrique C. Gutiérrez acknowledges the support of CAPES-Brazil.

†Graduate School of Economics, EPGE-FGV. Praia de Botafogo, 190. CEP 22250-900, Rio de Janeiro, RJ, BRAZIL. E-mail: cgutierrez@fgvmail.br

INTRODUCTION

In the modelling of economic and financial time series, the vectorial autoregressive (VAR) model became a standard linear model used in empirical works. An important aspect of empirical research in the specification of the VAR models is the determination of the lag order of the autoregressive lag polynomial, since all inference in the VAR model depends on the correct model specification. In several contributions, the effect of lag length selection has been demonstrated: Lütkepohl (1993) indicates that selecting a higher order lag length than the true lag length causes an increase in the mean square forecast errors of the VAR and that underfitting the lag length often generates autocorrelated errors. Braun and Mittnik (1993) show that impulse response functions and variance decompositions are inconsistently derived from the estimated VAR when the lag length differs from the true lag length. When cointegration restrictions are considered in the model, the effect of lag length selection on the cointegration tests has been demonstrated. For example, Johansen (1991) and Gonzalo (1994) point out that VAR order selection may affect proper inference on cointegrating vectors and rank.

Recently empirical works have considered another kind of restrictions on the VAR model (e.g., Engle and Issler, 1995; Caporale, 1997; Mamingi and Sunday, 2003). Engle and Kozicki (1993) showed that VAR models can have another type of restrictions, called common cyclical features, which are restrictions on the short-run dynamics. These restrictions are defined in the same way as cointegration restrictions, while cointegration refers to relations among variables in the long-run, the common cyclical restrictions refer to relations in the short-run. Vahid and Engle (1993) proposed the *Serial Correlation Common Feature* (SCCF) as a measure of common cyclical feature. SCCF restrictions might be imposed in a covariance stationary VAR model or in a cointegrated VAR model. When short-run restrictions are imposed in cointegrated VAR models it is possible to define a weak version of SCCF restrictions. Hecq, Palm and Urbain (2006) defined a weak version of SCCF restrictions which they denominated it as *weak-form* (WF) common cyclical restrictions. A fundamental difference between SCCF and WF restrictions is in the form which each one imposes restrictions on the Vector Error Correction Model (VECM) representation[1]. When SCCF are imposed, all matrices of a VECM have rank less than the number of variables analyzed. On the other hand with WF restrictions all matrices, except the long-run matrix, have rank less than a number of variables in analysis. Hence, WF restrictions impose less restriction on VECM parameters. Some advantages emerge when WF restrictions are considered. First, due to the fact that WF restrictions does not impose restrictions on the cointegration space; the rank of common cyclical features is not limited by the choice of cointegrating rank. Another advantage is that WF restrictions is invariant over reparametrization in VECM representation.

[1] When a VAR model has cointegration restriction it can be represented as a VECM. This representation is also known as Granger Representation Theorem (Engle and Granger, 1987).

The literature has shown how to select an adequate lag order of a covariance stationary VAR model and an adequate lag order of a VAR model subject to cointegration restrictions. Among the classical procedures, there are the information criteria such as Akaike (AIC), Schwarz (SC) and Hannan-Quinn (HQ) (Lütkepohl, 1993). Kilian (2001) study the performance of traditional AIC, SC and HQ criterion of a covariance stationary VAR model. Vahid and Issler (2002) analyzed the standard information criterion in a covariance stationary VAR model subject to SCCF restriction and more recently Guillén, Issler and Athanasopoulos (2005) studied the standard information criterion in VAR models with cointegration and SCCF restrictions. However, when cointegrated VAR models contain additional weak form of common cyclical feature, there are no reported work on how to appropriately determine the VAR model order.

The objective of this paper is to investigate the performance of information criterion in selecting the lag order of a VAR model when the data are generated from a true VAR with cointegration and WF restrictions that is referred as the correct model. It will be carried out following two procedures: $a)$ the use of standard criteria as proposed by Vahid and Engle (1993), referred here as IC (p), and $b)$ the use of an alternative procedure of model selection criterion (see, Vahid and Issler, 2002; Hecq *et al.*, 2006) consisting in selecting simultaneously the lag order p and the rank s do to the weak form of common cyclical feature, which is referred to as IC(p, s)[2]. The most relevant results can be summarized as follows. The information criterion that selects simultaneously the pair (p, s) has better performance than the model chosen by conventional criteria. The cost of ignoring additional WF restrictions in vector autoregressive modelling can be high specially when SC criterion is used.

The remaining of this work is organized as follows. Section 2 shows the econometric model. In section 3 the information criteria are mentioned. Monte Carlo simulation is shown in section 4 and the results in section 5. Finally, the conclusions are shown in section 6.

THE ECONOMETRIC MODEL

We show the VAR model with short-run and long-run restrictions. First, we consider a Gaussian vector autoregression of finite order p, so-called VAR(p), such that:

$$y_t = \sum_{i=1}^{p} A_i y_{t-i} + \varepsilon_t \tag{1}$$

where, y_t is a vector of n first order integrated series, $I(1)$, A_i, $i = 1, \ldots, p$ are matrices of dimension $n \times n$, $\varepsilon_t \sim Normal\,(0, \Omega)$ and $\{\Omega,\;$ if $t = \tau$ and $0_{n \times n},\;$ if $t \neq \tau$, where Ω is non singular$\}$. The model (1) could be written equivalently as; $\Pi\,(L)\; y_t = \varepsilon_t$ where L represents the lag operator and

[2] This is quite recent in the literature (see, Hecq *et al.*, 2006).

$\Pi(L) = I_n - \sum_{i=1}^{p} A_i L^i$ that when $L = 1$, $\Pi(1) = I_n - \sum_{i=1}^{p} A_i$. If cointegration is considered in (1) the $(n \times n)$ matrix $\Pi(\cdot)$ satisfies two conditions: a) Rank $(\Pi(1)) = r$, $0 < r < n$, such that $\Pi(1)$ can be expressed as $\Pi(1) = -\alpha\beta'$, where α and β are $(n \times r)$ matrices with full column rank, r. b) The characteristic equation $|\Pi(L)| = 0$ has $n - r$ roots equal to 1 and all other are outside the unit circle. These assumptions imply that y_t is cointegrated of order $(1, 1)$. The elements of α are the adjustment coefficients and the columns of β span the space of cointegration vectors. We can represent a VAR model as VECM. Decomposing the polynomial matrix $\Pi(L) = \Pi(1) L + \Pi^*(L) \Delta$, where $\Delta \equiv (1 - L)$ is the difference operator, a Vector Error Correction Model (VECM) is obtained:

$$\Delta y_t = \alpha\beta' y_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta y_{t-i} + \varepsilon_t \tag{2}$$

where: $\alpha\beta' = -\Pi(1)$, $\Gamma_j = -\sum_{k=j+1}^{p} A_k$ for $j = 1, ..., p-1$ and $\Gamma_0 = I_n$. The VAR(p) model can include additional short-horizon restrictions as shown by Vahid and Engle (1993). We consider an interesting WF restriction (as defined by Hecq, Palm and Urbain, 2006) that does not impose restrictions over long-run relations.

Definition 1 *Weak Form-WF holds in (2) if, in addition to assumption 1 (cointegration), there exists a* $(n \times s)$ *matrix* $\tilde{\beta}$ *of rank s, whose columns span the cofeature space, such that* $\tilde{\beta}'(\Delta y_t - \alpha\beta' y_{t-1}) = \tilde{\beta}'\varepsilon_t$ *, where* $\tilde{\beta}'\varepsilon_t$ *is a s-dimensional vector that constitutes an innovation process with respect to information prior to period t.*

Consequently we considerate WF restrictions in the VECM if there exists a cofeature matrix $\tilde{\beta}$ that satisfies the following assumption:

Assumption 1 : $\tilde{\beta}' \, \Gamma_j = 0_{s \times n}$ *for* $j = 1, ..., p-1$.

Imposing WF restrictions is convenient because it allows the study of both cointegration and common cyclical feature without the constraint $r + s \leq n$. We can rewrite the VECM with WF restrictions as a model of reduced-rank structure. In (2) let $X_{t-1} = [\Delta y'_{t-1}, ...\Delta y'_{t-p+1}]'$ and $\Phi = [\Gamma_1, ..., \Gamma_{p-1}]$, therefore we get:

$$\Delta y_t = \alpha\beta' y_{t-1} + \Phi X_{t-1} + \varepsilon_t \tag{3}$$

If assumption (1) holds matrices $\Gamma_i, i = 1, ..., p$ are all of rank $(n - s)$ then we can write $\Phi = \tilde{\beta}_\perp \Psi = \tilde{\beta}_\perp [\Psi_1, ..., \Psi_{p-1}]$, where, $\tilde{\beta}_\perp$ is $n \times (n - s)$ full column rank matrix, Ψ is of dimension $(n - s) \times n(p - 1)$, the matrices $\Psi_i, i = 1, ..., p - 1$ all of rank $(n - s) \times n$. Hence, given assumption (1), there exists $\tilde{\beta}$ of $n \times s$ such that $\tilde{\beta}'\tilde{\beta}_\perp = 0$. That is, $\tilde{\beta}_\perp$ $n \times (n - s)$ is a full column rank orthogonal to the complement of $\tilde{\beta}$ with $rank(\tilde{\beta}, \tilde{\beta}_{\perp}) = n$. Rewriting model (3) we have:

$$\Delta y_t = \alpha\beta' y_{t-1} + \tilde{\beta}_{\perp} (\Psi_1, \Psi_2, ..., \Psi_{p-1}) X_{t-1} + \varepsilon_t \tag{4}$$

$$= \alpha\beta' y_{t-1} + \tilde{\beta}_{\perp} \Psi X_{t-1} + \varepsilon_t \tag{5}$$

Estimation of (5) is carried out via the switching algorithms (see, Hecq, 2006) that use the procedure in estimating reduced-rank regression models suggested by Anderson (1951). There is a formal connection between a reduced-rank regression and the canonical analysis as noted by Izenman (1975), Box and Tiao (1977), Tso (1980) and Veleu *et al.* (1986). When the multivariate regression has all of its matrix coefficients of full rank, it may be estimated by usual Least Square or Maximum-Likelihood procedures. But when the matrix coefficients are of reduced-rank they have to be estimated using the reduced-rank regression models of Anderson (1951). The use of canonical analysis may be regarded as a special case of reduced-rank regression. More specifically, the maximum-likelihood estimation of the parameters of the reduced-rank regression model may result in solving a problem of canonical analysis[3]. Therefore, we can use the expression $CanCorr\{X_t, Z_t|W_t\}$ that denotes the partial canonical correlations between X_t and Z_t: both sets concentrate out the effect of W_t that allows us to obtain canonical correlation, represented by the eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \hat{\lambda}_3 .... > \hat{\lambda}_n$. The Johansen test statistic is based on canonical correlation. In model (2) we can use the expression $CanCorr\{\Delta y_t, y_{t-1}|W_t\}$ where $W_t = [\Delta y_{t-1}, \Delta y_{t-2}, ..., \Delta y_{t+p-1}]$ that summarizes the reduced-rank regression procedure used in the Johansen approach. It means that one extracts the canonical correlations between Δy_t and y_{t-1}: both sets concentrated out the effect of lags of W_t. In order to test for the significance of the r largest eigenvalues, one can rely on Johansen's trace statistic (6):

$$\xi_r = -T \sum_{i=r+1}^{n} Ln\,(1 - \hat{\lambda}_i^2) \qquad i = 1, ..., n \tag{6}$$

where the eigenvalues $0 < \hat{\lambda}_n < ... < \hat{\lambda}_1$ are the solution of : $|\lambda m_{11} - m_{10}^{-1} m_{00} m_{01}| = 0$, where $m_{ij},\ i,\ j = 0.1$, are the second moment matrices: $m_{00} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{0t}\tilde{u}_{0t}'$, $m_{10} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{1t}\tilde{u}_{0t}'$, $m_{01} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{0t}\tilde{u}_{1t}'$, $m_{11} = \frac{1}{T}\sum_{t=1}^{T} \tilde{u}_{1t}\tilde{u}_{1t}'$ of the residuals $\tilde{u}_{0t}$ and $\tilde{u}_{1t}$ obtained in the multivariate least squares regressions $\Delta y_t = (\Delta y_{t-1}, ...\Delta y_{t-p+1}) + u_{0t}$ and $y_{t-1} = (\Delta y_{t-1}, ...\Delta y_{t-p+1}) + u_{1t}$ respectively (see, Hecq *et al.,* 2006; Johansen, 1995). The result of Johansen test is a superconsistent estimated β. Moreover, we could also use a canonical correlation approach to determine the rank of the common features space due to WF restrictions. It is a test for the existence of cofeatures in the form of linear combinations of the variables in the first differences, corrected for long-run effects which are white noise (i.e., $\tilde{\beta}'(\Delta y_t - \alpha\beta' y_{t-1}) = \tilde{\beta}'\varepsilon_t$ where $\tilde{\beta}'\varepsilon_t$ is a white noise). Canonical analysis is

[3]This estimation is referred as *Full Information Maximum Likelihood* - FIML

adopted in the present work in estimating, testing and selecting lag-rank of VAR models as shown in next sections.

MODEL SELECTION CRITERIA

In model selection we use two procedures to identify the VAR model order. The standard selection criteria, IC(p) and the modified informational criteria, IC(p, s), novelty in the literature, which consists on identifying p and s simultaneously.

The model estimation following the standard selection criteria, IC(p), used by Vahid and Engle (1993) entails the following steps:

1. Estimate p using standard informational criteria: Akaike (AIC), Schwarz (SC) and Hanna-Quinn (HQ). We choose the lag length of the VAR in levels that minimize the information criteria.
2. Using the lag length chosen in the previous step, find the number of cointegration vector, r using Johansen cointegration test[4].
3. Conditional on the results of cointegration analysis, a final VECM is estimated and then the multi-step ahead forecast is calculated.

The above procedure is followed when there is evidence of cointegration restrictions. We check the performance of IC(p) when WF restrictions contain the true model. Additionally we check the performance of alternative selection criteria IC(p, s). Vahid and Issler (2002) analyzed a covariance-stationary VAR model with SCCF restrictions. They showed that the use of IC(p, s) has better performance than IC(p) in VAR model lag order selection. In the present work we analyze cointegrated VAR model with WF restrictions in order to analyze the performance of IC(p) and IC(p, s) for model selection. The question investigated is: is the performance of IC(p, s) superior to that of IC(p)? This is an important question we aim to answer in this work.

The procedure of selecting the lag order and the rank of the structure of short-run is carried out by minimizing the following modified information criteria (see Hecq, 2006).

$$AIC\,(p, s) = \sum_{i=n-s+1}^{T} \ln(1 - \lambda_i^2(p)) + \frac{2}{T} \times N \tag{7}$$

$$HQ(p, s) = \sum_{i=n-s+1}^{T} \ln(1 - \lambda_i^2(p)) + \frac{2\ln(\ln T)}{T} \times N \tag{8}$$

[4]Cointegration rank and vectors are estimated using the FIML as shown in Johansen (1991).

$$SC(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{\ln T}{T} \times N \tag{9}$$

$$N = [n \times (n \times (p-1)) + n \times r] - [s \times (n \times (p-1) + (n-s))]$$

The number of parameters N is obtained by subtracting the total number of mean parameters in the VECM (i.e., $n^2 \times (p-1) + nr$), for given r and p, from the number of restrictions the common dynamics imposes from $s \times (n \times (p-1)) - s \times (n-s)$. The eigenvalues λ_i are calculated for each p. To calculate the pair (p, s) we assume that no restriction of cointegration exists, that is, $r = n$ (see Hecq, 2006). We fix p in model (3) and then find λ_i $i = 1, 2...n$ using the program $cancorr(\Delta y_t, X_{t-1} \mid y_{t-1})$. This procedure is followed for every p and in the end we choose the p and s that minimizes the IC(p, s). After selecting the pair (p, s) we can test the cointegration relation using the procedure of Johansen. Finally we estimate the model using the switching algorithms as shown in the next chapter. Notice that in this simultaneous selection, testing the cointegration relation is the last procedure to follow, so we are inverting the hierarquical procedure followed by Vahid and Engle (1993) where the first step is the selection of the number of cointegration relations. It may be an advantage specially when r is over-estimated. Few works have been dedicated to analyze the order of the VAR models considering modified IC(p, s). As mentioned, Vahid and Issler (2002) suggested the use of IC(p, s) to simultaneously choose the order p and a number of reduced rank structure s on covariance stationary VAR model subject to SCCF restrictions. However, no work has analyzed the order of the VAR model with cointegration and WF restrictions using a modified criterion, which is exactly the contribution of this paper.

To estimate the VAR model considering cointegration and WF restrictions we use the switching algorithms model as considered by Hecq (2006). Consider the VECM given by:

$$\Delta y_t = \alpha \beta' y_{t-1} + \tilde{\beta}_\perp \Psi X_{t-1} + \varepsilon_t \tag{10}$$

A full description of switching algorithms is presented below in four steps:

$Step1$: Estimation of the cointegration vectors β.

Using the optimal pair $(\bar{p}, \bar{s})$ chosen by information criteria (7), (8) or (9), we estimate β (and so its rank, $r = \bar{r}$) using Johansen cointegration test.

$Step2$: Estimation of $\tilde{\beta}_\perp$ and Ψ.

Taking $\hat{\beta}$ estimated in step one, we proceed to estimate $\tilde{\beta}_\perp$ and Ψ. Hence, we run a regression of Δy_t and of X_{t-1}on $\hat{\beta}' y_{t-1}$. We labeled the residuals as u_0 and u_1, respectively.

Therefore, we obtain a reduced rank regression:

$$u_0 = \tilde{\beta}_{\perp}\Psi u_1 + \varepsilon_t \tag{11}$$

where Ψ can be written as $\Psi = \left(C_1, ..., C_{(\bar{p}-1)}\right)$ of $(n-\bar{s})\times n(\bar{p}-1)$ and $\tilde{\beta}_{\perp}$ of $n\times(n-\bar{s})$. We estimate (11) by FIML. Thus, we can obtain $\tilde{\beta}_{\perp}$ and $\hat{\Psi}$.

$Step 3$: Estimate of the Maximum Likelihood (ML) function.

Given the parameters estimated in steps 1 and 2 we use a recursive algorithm to estimate the Maximum Likelihood (ML) function. We calculate the eigenvalues associated with $\hat{\Psi}$, $\hat{\lambda}_i^2$ $i = 1, ..., \bar{s}$ and the matrix of residuals $\sum_{\bar{r},\, s=\bar{s}}^{\max}$. Hence, we compute the ML function:

$$L^0_{\max,\ \bar{r}<n,\ s=\bar{s}} = -\frac{T}{2}\left[\ln\left|\sum_{\bar{r}<n,\ s=\bar{s}}^{\max}\right| - \sum_{i=1}^{\bar{s}}\ln\left(1-\hat{\lambda}_i^2\right)\right] \tag{12}$$

If $\bar{r} = n$, we use instead of (12) the derived log-likelihood: $L_{\max,\ r=n,\ s=\bar{s}} = -\frac{T}{2}\ln\left|\sum_{\bar{r}=n,\ s=\bar{s}}^{\max}\right|$. The determinant of the covariance matrix for $\bar{r} = n$ cointegration vector is calculated by

$$\ln\left|\sum_{\bar{r}=n,\ s=\bar{s}}^{\max}\right| = \ln\left|m_{00} - m_{01}m_{11}^{-1}m_{10}\right| - \sum_{i=1}^{\bar{s}}\ln\left(1-\hat{\lambda}_i^2\right) \tag{13}$$

where m_{ij} refers to cross moment matrices obtained in multivariate least square regressions from Δy_t and X_{t-1} on y_{t-1}. In this case, estimation does not imply an iterative algorithm yet because the cointegrating space spans R^n.

$Step 4$: Reestimation of β.

We reestimate β to obtain a more appropriated value for the parameters. In order to reestimate β we use the program $CanCorr\left[\Delta y_t, y_{t-1} \mid \hat{\Psi}X_{t-1}\right]$ and thus using the new $\hat{\beta}$ we can repeat step 2 to reestimate $\tilde{\beta}_{\perp}$ and Ψ. Then, we can calculate the new value of the ML function in the step 3. Henceforth, we obtain $L^1_{\max,\ r=\bar{r},\ s=\bar{s}}$ for calculating $\Delta L = \left(L^1_{\max,\ r=\bar{r},\ s=\bar{s}}\ \text{-}\ L^0_{\max,\ r=\bar{r},\ s=\bar{s}}\right)$.

We repeat steps 1 to 4 to choose $\tilde{\beta}_{\perp}$and Ψ until convergence is reached (i.e., $\Delta L < 10^{-7}$). In the end, optimal parameters $\bar{p}$, $\bar{r}$ and $\bar{s}$ are obtained and it can be used for estimation and forecasting of a VECM with WF restrictions.

MONTE-CARLO DESIGN

One of the critical issues regarding Monte-Carlo experiments is the data generating processes. To build the data generating processes we consider a VAR model with three variables, one cointegration vector, and two cofeatures vectors (i.e., $n = 3$, $r = 1$ and $s = 2$, respectively). β and $\tilde{\beta}$ satisfy:

$$\beta = \begin{bmatrix} 1.0 \\ 0.2 \\ -1.0 \end{bmatrix}, \tilde{\beta} = \begin{bmatrix} 1.0 & 0.1 \\ 0.0 & 1.0 \\ 0.5 & -0.5 \end{bmatrix}$$

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

Consider the VAR(3) model: $y_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + \varepsilon_t$. The VECM respresentation as a function of the VAR level parameters can be written as:

$$\Delta y_t = (A_1 + A_2 + A_3 - I_3) y_{t-1} - (A_2 + A_3)\Delta y_{t-1} - A_3 \Delta y_{t-2} + \varepsilon_t \tag{14}$$

The VAR coefficients must simultaneously obey the restrictions: a) The cointegration restrictions: $\alpha\beta' = (A_1 + A_2 + A_3 - I_3)$; b) WF restrictions: $\tilde{\beta}' A_3 = 0$ (iii) $\tilde{\beta}'(A_2 + A_3) = 0$ and c) covariance-stationary condition. Considering the cointegration restrictions we can rewrite (14) as the following VAR(1):

$$\xi_t = F\, \xi_{t-1} + v_t \tag{15}$$

$$\xi_t = \begin{bmatrix} \triangle y_t \\ \triangle y_{t-1} \\ \beta' y_t \end{bmatrix}, F = \begin{bmatrix} -(A_2 + A_3) & -A_3 & \alpha \\ I_3 & 0 & 0 \\ -\beta(A_2 + A_3) & -\beta' A_3 & \beta'\alpha + 1 \end{bmatrix} \text{ and } v_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \beta' \varepsilon_t \end{bmatrix}$$

Thus, the equation (15) will be covariance-stationary if all eigenvalues of matrix F lie inside the unit circle. An initial idea to design the Monte-Carlo experiment may consist of constructing the companion matrix (F) and verify whether the eigenvalues of the companion matrix all lie inside the unit circle. This may be carried out by selecting their values from a uniform distribution, and then verifying whether or not the eigenvalues of the companion matrix all lie inside the unit circle. However, this strategy could lead to a wide spectrum of search for adequate values for the companion matrix. Hence, we follow an alternative procedure. We propose an analytical solution to generate a covariance-stationary VAR, based on the choice of the eigenvalues, and then on the generation of the respective companion matrix. In the appendix we present a detailed discussion of the final choice of these free parameters, including analytical solutions. In our simulation, we constructed 100 data generating processes and for each of these we generate 1000 samples containing 1000 observations. In order to reduce the impact of initial values, we consider only the last 100 and 200 observations. All the experiments were conducted in the MatLab environment.

RESULTS

Values in Table I represent the percentage of time that the model selection criterion, IC(p), chooses that cell corresponding to the lag and number of cointegration vectors in 100000 realizations. The true lag-cointegrating vectors are identified by bold numbers and the selected lag-cointegration vectors chosen more times by the criterion are underlined. The results show that, in general, the AIC criterion choose more frequently the correct lag length for 100 and 200 observations. For example, for 100 observations, the AIC, HQ and SC criteria chose the true lag, p, 54.08%, 35.62% and 17.49% of the times respectively. Note that all three criteria chose more frequently the correct rank of cointegration $(r = 1)$. When 200 observations are considered, the correct lag length was chosen 74.72%, 57.75% and 35.28% of the time for AIC, HQ and SC respectively. Again all three criteria selected the true cointegrated rank $r = 1$. Tables II contains the percentage of time that the simultaneous model selection criterion, IC(p, s), chooses that cell, corresponding to the lag-rank and number of cointegrating vectors in 100,000 realizations. The true lag-rank-cointegration vectors are identified by bold numbers and the best lag-rank combination chosen more times by each criterion are underlined. The results show that, in general, the AIC criterion chooses more frequently lag-rank for 100 and 200 observations. For instance, for 100 observations, the AIC, HQ and SC criteria choose more frequently the true pair $(p, s) = (3, 1)$, 56.34%, 40.85% and 25.20% of the times respectively. For 200 observations, AIC, HQ and SC criteria choose more frequently the true pair $(p, s) = (3, 1)$, 77.07%, 62.58% and 45.03% of the times respectively. Note that all three criteria choose more frequently the correct rank of cointegration $(r = 1)$ in both samples.

The most relevant results can be summarized as follows:

− All criteria (AIC, HQ and SC) choose the correct parameters more often when using IC(p, s).

− The AIC criterion has better performance in selecting the true model more frequently for both the IC(p, s) and the IC(p) criteria.

− When the size of the sample decreases the true value p is less frequently selected by all the traditional criteria.

− Table I shows that ignoring WF restrictions the standard SC has the worst performance in choosing the true value of p.

It is known that literature suggests the use of the traditional SC and HQ criteria in VAR model selection. The results of this work indicate that if additional WF restrictions are ignored, the standard SC and HQ criteria select few times the true value of p. That is, there is a cost of ignoring additional WF restrictions in the model specially when SC criterion is used. In general, the standard Schwarz or Hannan-Quinn selection criteria should not be used for this purpose in small samples due to the tendency of identifying an underparameterized model. In general, the use of these alternative criteria of selection, IC(p, s) has better performance than the usual criteria, IC(p), when the cointegrated VAR model has additional WF restriction.

CONCLUSIONS

In this work, we considered an additional *weak form* restriction of common cyclical features in a cointegrated VAR model in order to analyze the appropriate way for selecting the correct lag order. These additional WF restrictions are defined in the same way as cointegration restrictions, while cointegration refers to relations among variables in the long-run, the common cyclical restrictions refer to relations in the short-run. Two methodologies have been used for selecting lag length; the traditional information criterion, IC(p), and an alternative criterion (IC(p, s)) that selects simultaneously the lag order p and the rank structure s due to the WF restriction.

The results indicate that information criterion that selects the lag length and the rank order simultaneously has better performance than the model chosen by conventional criteria. When the WF restrictions are ignored there is a non trivial cost in selecting the true model with standard information criteria. In general, the standard Schwarz or Hannan-Quinn criteria selection criteria should not be used for this purpose in small samples due to the tendency of identifying an under-parameterized model.

In applied work, when the VAR model contains WF and cointegration restrictions, we suggest the use of AIC(p, s) criteria for simultaneously choosing the lag-rank, since it provides considerable gains in selecting the correct VAR model. Since no work in the literature has been dedicated to analyze a VAR model with WF common cyclical restrictions, the results of this work provide new insights and incentives to proceed with this kind of empirical work.

APPENDIX A : TABLES

Table I. Performance of information criterion, IC (p) in selecting the lag order p

Frequency of lag(p) and cointegrating vectors (r) choice by different criteria for trivariate VAR model in levels when the true model have parameters: p = 3 and r = 1.

		Number of observations = 100 Selected cointegrated vectors				Number of observations = 200 Selected cointegrated vectors			
		0	**1**	2	3	0	1	2	3
	Selected Lag								
	1	0,000	0,996	0,359	0,031	0,000	0,095	0,016	0,003
	2	0,002	32,146	1,136	0,048	0,000	17,073	0,686	0,033
	3	2,792	**54,082**	0,902	0,041	0,012	**74,721**	1,488	0,108
AIC(p)	4	0,737	4,068	0,091	0,003	0,005	4,177	0,081	0,006
	5	0,392	0,987	0,031	0,000	0,013	0,828	0,020	0,000
	6	0,219	0,333	0,014	0,000	0,023	0,257	0,005	0,000
	7	0,166	0,173	0,006	0,000	0,039	0,133	0,002	0,000
	8	0,133	0,107	0,005	0,000	0,060	0,115	0,001	0,000
	1	0,000	3,884	1,915	0,165	0,000	1,098	0,243	0,021
	2	0,002	52,593	1,907	0,080	0,000	37,390	1,614	0,098
	3	2,600	**35,617**	0,612	0,027	0,012	**57,749**	1,146	0,082
HQ(p)	4	0,065	0,189	0,007	0,000	0,001	0,158	0,004	0,000
	5	0,059	0,037	0,000	0,000	0,009	0,082	0,001	0,000
	6	0,073	0,025	0,000	0,000	0,016	0,076	0,000	0,000
	7	0,059	0,019	0,001	0,000	0,030	0,070	0,000	0,000
	8	0,053	0,011	0,000	0,000	0,044	0,055	0,001	0,000
	1	0,000	8,344	6,609	0,511	0,000	3,964	1,385	0,093
	2	0,003	61,966	2,279	0,105	0,000	55,156	2,776	0,169
	3	2,042	**17,485**	0,313	0,015	0,012	**35,283**	0,728	0,044
SC(p)	4	0,049	0,045	0,000	0,000	0,001	0,083	0,002	0,000
	5	0,071	0,025	0,000	0,000	0,007	0,076	0,001	0,000
	6	0,057	0,016	0,000	0,000	0,013	0,063	0,000	0,000
	7	0,036	0,009	0,000	0,000	0,025	0,056	0,000	0,000
	8	0,017	0,003	0,000	0,000	0,027	0,035	0,001	0,000

Numbers represent the percentage times that the model selection criterion choice that cell corresponding to the lag and number of cointegration vectors in 100,000 realizations. The true lag-cointegrating vectors are indentified by bold numbers

Table II. Performance of information criterion, IC(p, s) in selecting p and s simultaneously

Johansen Tested coint. Vectors (r)		0			**1**			2			3		
	Selected rank (s)	1	2	3	**1**	2	3	1	2	3	1	2	3
	Selected lag (p)												
Sample size = 100													
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,002	0,000	0,000	39,049	0,001	0,000	1,218	0,000	0,000	0,056	0,000	0,000
	3	0,301	0,000	0,000	**56,341**	0,003	0,000	1,559	0,000	0,000	0,053	0,000	0,000
AIC(p,s)	4	0,004	0,000	0,000	1,186	0,001	0,000	0,070	0,000	0,000	0,001	0,000	0,000
	5	0,000	0,000	0,000	0,114	0,001	0,000	0,012	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,020	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,006	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,002	0,000	0,000	55,563	0,000	0,000	1,888	0,000	0,000	0,081	0,000	0,000
	3	0,267	0,000	0,000	**40,855**	0,000	0,000	1,207	0,000	0,000	0,043	0,000	0,000
HQ(p,s)	4	0,000	0,000	0,000	0,088	0,000	0,000	0,005	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,004	0,000	0,000	70,971	0,000	0,000	2,574	0,000	0,000	0,113	0,000	0,000
	3	0,221	0,000	0,000	**25,204**	0,000	0,000	0,887	0,000	0,000	0,025	0,000	0,000
SC(p,s)	4	0,000	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
Sample size = 200													
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,000	0,000	0,000	18,797	0,000	0,000	0,681	0,000	0,000	0,038	0,000	0,000
	3	0,000	0,000	0,000	**77,065**	0,002	0,000	2,260	0,000	0,000	0,145	0,000	0,000
AIC(p,s)	4	0,000	0,000	0,000	0,908	0,000	0,000	0,035	0,000	0,000	0,001	0,000	0,000
	5	0,000	0,000	0,000	0,063	0,000	0,000	0,002	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,003	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,000	0,000	0,000	33,952	0,000	0,000	1,370	0,000	0,000	0,086	0,000	0,000
	3	0,000	0,000	0,000	**62,576**	0,000	0,000	1,877	0,000	0,000	0,111	0,000	0,000
HQ(p,s)	4	0,000	0,000	0,000	0,027	0,000	0,000	0,001	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	-	-	-	-	-	-	-	-	-	-	-	-
	2	0,000	0,000	0,000	50,983	0,000	0,000	2,351	0,000	0,000	0,146	0,000	0,000
	3	0,000	0,000	0,000	**45,028**	0,000	0,000	1,416	0,000	0,000	0,076	0,000	0,000
SC(p,s)	4	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	5	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	6	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	7	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	8	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

Numbers represent the percentage times that the simultaneous model selection criterion IC(p,s) choice that cell, corresponding to the lag-rank and number of cointegrating vectors in 100,000 realizations. The true lag-rank-cointegration vectors are identified by bold numbers and the best lag-rank-cointegration vectors chosen by criterias are identified by underline lines.

APPENDIX B : VAR RESTRICTIONS FOR THE DGPs

Let's consider the VAR(3) model :

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + A_3 y_{t-3} + \varepsilon_t \quad (16)$$

with parameters: $A_1 = \begin{bmatrix} a^1_{11} & a^1_{12} & a^1_{12} \\ a^1_{21} & a^1_{22} & a^1_{22} \\ a^1_{31} & a^1_{32} & a^1_{32} \end{bmatrix}$, $A_2 = \begin{bmatrix} a^2_{11} & a^2_{12} & a^2_{12} \\ a^2_{21} & a^2_{22} & a^2_{22} \\ a^2_{31} & a^2_{32} & a^2_{32} \end{bmatrix}$ and $A_3 = \begin{bmatrix} a^3_{11} & a^3_{12} & a^3_{12} \\ a^3_{21} & a^3_{22} & a^3_{22} \\ a^3_{31} & a^3_{32} & a^3_{32} \end{bmatrix}$

We consider the cointegration vectors $\beta = \begin{bmatrix} \beta_{11} \\ \beta_{21} \\ \beta_{31} \end{bmatrix}$, the cofeatures vectors $\tilde{\beta} = \begin{bmatrix} \tilde{\beta}_{11} & \tilde{\beta}_{12} \\ \tilde{\beta}_{21} & \tilde{\beta}_{22} \\ \tilde{\beta}_{31} & \tilde{\beta}_{32} \end{bmatrix}$ and the adjustament matrix $\alpha = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{bmatrix}$. The long-run relation is defined by $\alpha\beta' = (A_1 + A_2 + A_3 - I_3)$.
The VECM respresentation is:

$$\Delta y_t = \alpha\beta' y_{t-1} - (A_2 + A_3)\Delta y_{t-1} - A_3 \Delta y_{t-2} + \varepsilon_t \quad (17)$$

Considering the cointegration restrictions we can rewrite (17) as the following VAR(1)

$$\xi_t = F\,\xi_{t-1} + v_t \quad (18)$$

where $\xi_t = \begin{bmatrix} \triangle y_t \\ \triangle y_{t-1} \\ \beta' y_t \end{bmatrix}$, $F = \begin{bmatrix} -(A_2 + A_3) & -A_3 & \alpha \\ I_3 & 0 & 0 \\ -\beta(A_2 + A_3) & -\beta' A_3 & \beta'\alpha + 1 \end{bmatrix}$ and $v_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \beta'\varepsilon_t \end{bmatrix}$

1) Short-run restrictions (WF)

Let us, $G = -[R_{21}K + R_{31}]$, $K = [(R_{32} - R_{31})/(R_{21} - R_{22})]$, $R_{j1} = \tilde{\beta}_{j1}/\tilde{\beta}_{11}$, $R_{j2} = \tilde{\beta}_{j2}/\tilde{\beta}_{12}$ $(j = 2, 3)$ and $S = \beta_{11}G + \beta_{21}K + \beta_{31}$

(i) $\tilde{\beta}' A_3 = 0 ==> A_3 = \begin{bmatrix} -Ga^3_{31} & -Ga^3_{32} & -Ga^3_{33} \\ -Ka^3_{31} & -Ka^3_{32} & -Ka^3_{33} \\ -a^3_{31} & -a^3_{32} & -a^3_{33} \end{bmatrix}$

(ii) $\tilde{\beta}'(A_2 + A_3) = 0 ==> \tilde{\beta}' A_2 = 0 ==> A_2 = \begin{bmatrix} -Ga^2_{31} & -Ga^2_{32} & -Ga^2_{33} \\ -Ka^2_{31} & -Ka^2_{32} & -Ka^2_{33} \\ -a^2_{31} & -a^2_{32} & -a^2_{33} \end{bmatrix}$

2) Long-run restrictions (cointegration)

(iv) $\beta'(A_2 + A_3) = [-(a_{31}^2 + a_{31}^3)S \;\; -(a_{32}^2 + a_{32}^3)S \;\; -(a_{33}^2 + a_{33}^3)S]$ and $\beta' A_3 = [-a_{31}^3 S \;\; - a_{32}^3 S \;\; - a_{33}^3 S]$

(v) $\beta'\alpha + 1 = \beta = \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} \end{bmatrix} \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{bmatrix} + 1 = \beta_{11}\alpha_{11} + \beta_{21}\alpha_{21} + \beta_{31}\alpha_{31} + 1$

Therefore, considering short- and long-run restrictions, the companion matrix F is represented as:

$$F = \begin{bmatrix} -(A_2 + A_3) & -A_3 & \alpha \\ I_3 & 0 & 0 \\ -\beta(A_2 + A_3) & -\beta' A_3 & \beta'\alpha + 1 \end{bmatrix}$$

$$= \begin{bmatrix} -G(a_{31}^2 + a_{31}^3) & -G(a_{32}^2 + a_{32}^3) & -G(a_{33}^2 + a_{33}^3) & -Ga_{31}^3 & -Ga_{32}^3 & -Ga_{33}^3 & \alpha_{11} \\ -K(a_{31}^2 + a_{31}^3) & -G(a_{32}^2 + a_{32}^3) & -G(a_{33}^2 + a_{33}^3) & -Ka_{31}^3 & -Ka_{32}^3 & -Ka_{33}^3 & \alpha_{21} \\ -(a_{31}^2 + a_{31}^3) & -G(a_{32}^2 + a_{32}^3) & -(a_{33}^2 + a_{33}^3) & -a_{31}^3 & -a_{32}^3 & -a_{33}^3 & \alpha_{31} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -(a_{31}^2 + a_{31}^3)S & -(a_{32}^2 + a_{32}^3)S & -(a_{33}^2 + a_{33}^3)S & -a_{31}^3 S & -a_{32}^3 S & -a_{33}^3 S & b \end{bmatrix}$$

with $b = \beta'\alpha + 1 = \beta_{11}\alpha_{11} + \beta_{21}\alpha_{21} + \beta_{31}\alpha_{31} + 1$

3) Restrictions of covariance-stationary in equation (18)

The equation (18) will be covariance-stationary, all eigenvalues of matrix F lie inside the unit circle. Therefore, the eigenvalues of the matrix F is a number λ such that:

$$|F - \lambda I_7| = 0 \tag{19}$$

The solution of (19) is:

$$\lambda^7 + \Omega\lambda^6 + \Theta\lambda^5 + \Psi\lambda^4 = 0 \tag{20}$$

where the parameters Ω, Θ, and Ψ are: $\Omega = G(a_{31}^2 + a_{31}^3) + K(a_{32}^2 + a_{32}^3) + a_{33}^2 + a_{33}^3 - b$, $\Theta = Ga_{31}^3 + Ka_{32}^3 - (a_{33}^2 + a_{33}^3)b - Gb(a_{31}^2 + a_{31}^3) - Kb(a_{32}^2 + a_{32}^3) + \alpha_{31}S(a_{33}^2 + a_{33}^3) + S\alpha_{21}(a_{32}^2 + a_{32}^3) + S\alpha_{11}(a_{31}^2 + a_{31}^3) + a_{33}^3$ and $\Psi = -a_{33}^3 b - Ga_{31}^3 b - Ka_{32}^3 b + \alpha_{31}a_{33}^3 S + a_{32}^3 S\alpha_{21} + a_{31}^3 S\alpha_{11}$, and the first four roots are $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$.We calculated the parameters of matrices A_1, A_2 and A_3 as function of roots (λ_5, λ_6 and λ_7) and free parameters. Hence we have three roots satisfying equation (20)

$$\lambda^3 + \Omega\lambda^2 + \Theta\lambda + \Psi = 0 \tag{21}$$

for λ_5, we have: $\lambda_5^3 + \Omega\lambda_5^2 + \Theta\lambda_5 + \Psi = 0$ ....$Eq1$
for λ_6, we have: $\lambda_6^3 + \Omega\lambda_6^2 + \Theta\lambda_6 + \Psi = 0$ ....$Eq2$
for λ_7, we have: $\lambda_7^3 + \Omega\lambda_7^2 + \Theta\lambda_7 + \Psi = 0$ ....$Eq3$

Solving $Eq1, Eq2$ and $Eq3$ we have: $\Omega = -\lambda_7 - \lambda_6 - \lambda_5$, $\Theta = \lambda_6\lambda_7 + \lambda_6\lambda_5 + \lambda_5\lambda_7$ and $\Psi = -\lambda_5\lambda_6\lambda_7$. Equaling these parameters with relations above we have:

$a_{31}^2 = -(-Ka_{32}^2 - Ka_{32}^2 b + \alpha_{31}Sa_{33}^2 - \lambda^6\lambda^7 - \lambda^6 - \lambda^7 - a_{33}^2 b - \lambda^5\lambda^6\lambda^7 + b - \lambda^5\lambda^7 - \lambda^5\lambda^6 - a_{33}^2 + Sa_{32}^2\alpha_{21} - \lambda^5)/(S\alpha_{11} - G - Gb)$

$a_{32}^3 = (-S^2\lambda^7\alpha_{11}\alpha_{31} - b^2\lambda^7 G - \lambda^6 Gb^2 + b\lambda^7 S\alpha_{11} + \lambda^6 S\alpha_{11}b - a_{31}^3 S\alpha_{11}G + a_{31}^3 S^2\alpha_{11}^2 - Ga_{31}^3 bS\alpha_{11} - \lambda^5 Gb^2 + \lambda^5 S\alpha_{11}b - \lambda^7\lambda^6\alpha_{31}SG - \lambda^7\lambda^5\alpha_{31}SG - S^2\alpha_{11}\lambda^5\alpha_{31} - S^2\alpha_{11}\lambda^6\alpha_{31} + S\lambda^5 Gb\alpha_{31} + S\alpha_{31}\lambda^6 Gb - \lambda^5\lambda^7\lambda^6 G + \lambda^6\lambda^7 Gb + \lambda^5\lambda^7 Gb + \lambda^5\lambda^6 Gb - SGb^2\alpha_{31} + S^2\alpha_{11}b\alpha_{31} - S^2\alpha_{11}\alpha_{31}a_{33}^2 + S^2\alpha_{31}^2 a_{33}^2 G + SG^2 a_{31}^3\alpha_{31} + S\alpha_{11}a_{33}^2 b + Gb^3 - S\alpha_{11}b^2 - S^2\alpha_{11}Ka_{32}^2\alpha_{31} - S^2\alpha_{11}\alpha_{31}Ga_{31}^3 + S^2 a_{32}^2\alpha_{21}G\alpha_{31} - Sa_{32}^2\alpha_{21}Gb + S\alpha_{31}G^2 a_{31}^3 b - S\alpha_{31}a_{33}^2 Gb + S\alpha_{11}Ka_{32}^2 b + S\lambda^7 Gb\alpha_{31} - \lambda^5\lambda^6\alpha_{31}SG - \lambda^5\lambda^7\lambda^6\alpha_{31}SG + \lambda^5\lambda^7\lambda^6 S\alpha_{11})/(S\alpha_{11}K\alpha_{31} - KG\alpha_{31} + bG\alpha_{21} - K\alpha_{31}Gb - S\alpha_{11}\alpha_{21} + G\alpha_{21})/S$

$a_{33}^3 = -(Kb^3 G - \lambda^5 Gb^2 K + S\alpha_{11}\lambda^6 K\lambda^7\lambda^5 + Kb\lambda^7 S\alpha_{11} - Kb^2\lambda^7 G - S^2\alpha_{21}\lambda^7\alpha_{11} + \lambda^6 GbS\alpha_{21} + S\alpha_{21}\lambda^7 Gb - \lambda^6 Gb^2 K + \lambda^6 S\alpha_{11}Kb - \lambda^6 S^2\alpha_{11}\alpha_{21} + \lambda^5 GbS\alpha_{21} + \lambda^5 S\alpha_{11}Kb - \lambda^5 S^2\alpha_{11}\alpha_{21} - \lambda^7\lambda^6 S\alpha_{21}G + Kb\lambda^7\lambda^6 G + Kb\lambda^7\lambda^5 G + Kb\lambda^5\lambda^6 G - \lambda^7\lambda^6 KG\lambda^5 - S^2\alpha_{11}\alpha_{21}Ka_{32}^2 + S^2\alpha_{11}\alpha_{21}b - S^2\alpha_{11}\alpha_{21}a_{33}^2 + S^2\alpha_{21}^2 a_{32}^2 G - S\alpha_{11}Kb^2 + S\alpha_{21}G^2 a_{31}^3 - S\alpha_{21}Gb^2 + S^2 a_{31}^3 K\alpha_{11}^2 - S^2\alpha_{11}\alpha_{21}Ga_{31}^3 + S^2\alpha_{21}a_{33}^2 G\alpha_{31} + S\alpha_{11}K\hat{\ }2ba_{32}^2 + S\alpha_{11}Kba_{33}^2 - S\alpha_{11}a_{31}^3 KG - S\alpha_{11}KbGa_{31}^3 - SKba_{33}^2 G\alpha_{31} + S\alpha_{21}G^2 a_{31}^3 b - S\alpha_{21}\lambda^5\lambda^6 G - S\alpha_{21}\lambda^5\lambda^7\lambda^6 G - S\alpha_{21}Ka_{32}^2 Gb - S\alpha_{21}\lambda^7\lambda^5 G)/(S\alpha_{11}K\alpha_{31} - KG\alpha_{31} + bG\alpha_{21} - K\alpha_{31}Gb - S\alpha_{11}\alpha_{21} + G\alpha_{21})/S$

We can calculate a_{31}^2, $a_{32}^3$and a_{33}^3 fixing the set $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$ and sort independently from uniform distributions $(-0.9; 0.9)$ the values of a_{31}^3, a_{32}^2, a_{33}^2, λ_5, λ_6 and λ_7. Hencefore, each parameter of the matrices A_1, A_2 and A_3 are defined and so we can generate the DGPs of VAR(3) model with cointegration and WF restrictions.

REFERENCES

Anderson TW. 1951. Estimating linear restrictions on regression coefficients for multivariate normal distributions. *Ann. Math. Statist*. **22:** 327-351. [Correction (1980) *Ann. Statist* **8** 1400.]

Bewley R, Yang M. 1998. On the size and power of system tests for cointegration. *The Review of Economics and Statistics* **80**: 675-679.

Box GE, Tiao GC. 1977. A canonical analysis of multiple time series. *Biometrika* **64**:355-365.

Brandner P, Kunst RM. 1990. Forecasting vector autoregressions - The influence of cointegration: A Monte Carlo Study. Research Memorandum N0 265, Institute for Advanced Studies, Vienna.

Braun PA, Mittnik S. 1993. Misspecifications in Vector Autoregressions and Their Effects on Impulse Responses and Variance Decompositions. *Journal of Econometrics* **59**, 319-41.

Caporale GM. 1997. Common Features and Output Fluctuations in the United Kingdom. *Economic-Modelling* **14**: 1-9.

Cubadda G. 1999. Common serial correlation and common business cycles: A cautious note. *Empirical Economics* **24**: 529-535.

Engle, Granger. 1987. Cointegration and error correction: representation, estimation and testing. *Econometrica* **55**:251-76.

Engle, Kozicki. 1993. Testing for common features. *Journal of Business and Economic Statistics* **11:** 369-395.

Engle, Issler JV. 1995. Estimating Common Sectoral Cycles. *Journal of Monetary Economics* **35:**83-113.

Guillén, Issler, Athanasopoulos. 2005. Forecasting Accuracy and Estimation Uncertainty using VAR Models with Short- and Long-Term Economic Restrictions: A Monte-Carlo Study. *Ensaios Econômicos EPGE* **589**.

Hamilton JD. 1994. Time Series Analysis. Princeton University Press, New Jersey.

Hecq A, Palm FC, Urbain JP. 2006. Testing for Common Cyclical Features in VAR Models with Cointegration, *Journal of Econometrics* **132**: 117-141.

Hecq A. 2006. Cointegration and Common Cyclical Features in VAR Models: Comparing Small Sample Performances of the 2-Step and Iterative Approaches. Mimeo.

Hecq A. 2000. Common Cyclical Features in Multiple Time Series and Panel Data: Methodological Aspects and Applications.

Izenman AJ. 1975. Reduced rank regression for the multivariate linear model, *Journal of. multivariate Analysis* **5:** 248-264.

Johansen S. 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control* **12**: 231-54.

Johansen S. 1995. Likelihood-based Inference in Cointegrated Vector Autoregressive Models. Oxford Univ. Press. Z .

Kilian L. 2001. Impulse response analysis in vector autoregressions with unknown lag order. *Journal of Forecasting* **20**: 161-179.

Lutkepohl H. 1993. Introduction to Multiple Time Series Analysis. Springer, New York. Z .

Mamingi N, Iyare SO. 2003. Convergence and Common Features in International Output: A Case Study of the Economic Community of West African States, 1975-1997. *Asian-African-Journal-of-Economics-and-Econometrics*. **3**: 1-16.

Reinsel GC, Velu RP. 1998. Multivariate Reduced-rank Regression. Springer, New York. Z.

Stock JH, Watson MW. 1988. Testing for common trends. *Journal of the American Statistical Association* **83**:97-107.

Tiao GC, Tsay RS. 1985. A canonical correlation approach to modeling multivariate time series., *in Proceedings of the Business and Economic Statistics Section, American statistical Association*, 112-120.

Tiao GC, Tsay RS. 1989. Model specification in multivariate time series (with discussion). J. Roy. Statist. Soc. **51:** 157-213.

Tsay RS, Tiao GC. 1985. Use of canonical analysis in time series model identification. Biometrika 72 299 315. Z .

Tso MK 1981. Reduced rank regression and Canonical analysis. *Journal of the Royal Statistical Society* **43:** 183-189.

Vahid F, Issler JV. 2002. The Importance of Common Cyclical Features in VAR Analysis: A Monte Carlo Study. *Journal of Econometrics* **109:** 341-363.

Vahid F, Engle.1993. Common trends and common cycles. *Journal of Applied Econometrics* **8**: 341-360.

Velu RP, Reinsel GC, Wichern DW. 1986. Reduced rank models for multiple times series. *Biometrika* **73:** 105-118.