Selection of optimal lag length in cointegrated VARs models with common cyclical features

Carlos Enrique Carrasco Gutiérrez

December 2007

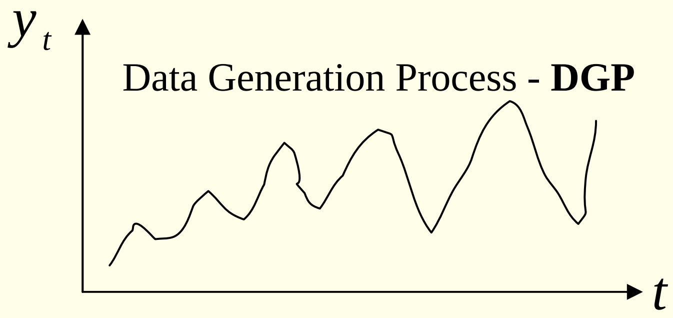
Motivation

AR(1) ➔ $y_t = \phi y_{t-1} + \varepsilon_t \qquad t = 1,....,\ T$

Let

$\phi = 0,6;\ y_0 = 0\ \ and \qquad \varepsilon_t \propto iid\ (0, \sigma^2)$

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$



Identification: What is the best model? That is, what is the value of the <u>parameter</u> ***p***?

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + \varepsilon_t$$

$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$? $\qquad y_t = \phi_1 y_{t-1} + \varepsilon_t$?

Usar: R^2 or $\bar{R}^2 = 1 - \dfrac{ESS/(n-k)}{TSS/(n-1)}$

Motivation

VAR(p) ➔ $y_t = A_1 y_{t-1} + A_2 y_{t-2} + \ldots + A_p y_{t-p} + \varepsilon_t$ $\qquad E(\varepsilon_t \varepsilon_s^T) = \begin{cases} \sum & if \quad t = s \\ 0 & if \quad t \neq s \end{cases}$

$$AIC(p) = \ln\left|\sum(p)\right| + \frac{2}{T} n^2 p$$

$$HQ(p) = \ln\left|\sum(p)\right| + \frac{2 \ln \ln T}{T} n^2 p$$

$$SC(p) = \ln\left|\sum(p)\right| + \frac{\ln T}{T} n^2 p$$

Standard Criteria IC(p)

Literature has recommended the use of these information criterion to select *p* in stationary VARs models and in cointegrated VARs models.

If the DGP is a VAR model with Cointegration and Common Cyclical Feature, how we should to select the parameter *p*?

Objective

The objective of this paper is investigated the performance of information criterion in selecting the lag order of a VAR model with cointegration and WF restrictions

Two procedure are compared:

a) The use of standard criteria as proposed by Vahid and Engle (1993):
IC(p) == > Standard criteria model selecting

b) the use of an alternative procedure of model selection criterion consisting in selecting the lag order and rank of short-run restrictions simultaneously:
IC(p,s) == > Alternative criteria

The question investigated here is; is the performance of IC(p,s) superior to that of IC(p)?

Importance of the correct model specification

In several contributions, the effect of lag length selection has been demonstrated.

1. 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.

2. Braun and Mittnik (1993) show that estimates of a VAR whose lag length differs from the true lag length are inconsistent as are the impulse response functions and variance decompositions derived from the estimated VAR.

3. Johansen(1991) and Gonzalo(1994), who point out that VAR-order selection may affect proper inference on cointegrating vectors and rank.

Literature

In line with our analysis

1. Vahid and Issler (2002), analyzed a covariance-stationary VAR model with SCCF. They showed that there exist gains for selecting p and s using IC(p,s) instead of IC(p).

2. In the same direction Guillén, Issler and Athanasopoulos (2005) analyzed an non-stationary VAR model with cointegration and SCCF restrictions. They showed that there exist gains for selecting p using IC(p,s) instead of IC(p).

3. **In this work, it is analyzed an non stationary VAR model with cointegration and WF restrictions. No one has dedicated to study this model.**

Contributions

We investigate the performance of IC(p,s) and IC(p) in selecting the lag length of the VAR model when it is restricted to cointegration and WF common cyclical.

Results indicate **gains for selecting p using IC(p,s) instead of IC(p)**

Econometric Model

In VAR (p) model: $$y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \varepsilon_t \qquad t = 1...T$$

The vast literature on cointegration has focused on long-run comovements, Engle and Granger (1987), Johansen (1998).

- Cointegration (low frequency) == > Common Trend (long-run)

Engle and Kozicki (1993) and Vahid and Engle (1993), showed that the VAR model can have additional restrictions. Common cyclical features refer to short-run comovements

- Common cyclical features (high frequency) == > Common Cycles (short-run)

SCCF === > Strong restrictions

WF === > Weak restrictions

Econometric Model

How these restrictions are imposing in the model?

Consider a Gaussian Vector Autoregression of order p (VAR(p))

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \varepsilon_t \qquad t = 1....T \qquad (1)$$

- y_t is a vector of n first order integrated series, $I(1)$;
- A_i, $i = 1, \ldots, p$ are matrices of dimension $n \times n$;
- $\varepsilon_t \sim Normal\,(0, \Omega)$;
- $E\,(\varepsilon_t) = 0$;
- $E\,(\varepsilon_t \varepsilon_\tau) = \{\Omega,$ if $t = \tau$ and $0_{n\times n}$, if $t \neq \tau$, where Ω is no singular$\}$.

$$\Pi\,(L) = I_n - \left(A_1 L + A_2 L^2 + ... + A_p L^p\right)$$

Econometric Model

The model (1) could be written equivalently as:

$$\Pi(L)y_t = \varepsilon_t \tag{2}$$

For L =1; $\Pi(1) = I_n - (A_1 + A_2 + ... + A_p)$

Assumption 1 (Cointegration): *The $(n \times n)$ matrix $\Pi(\cdot)$ satisfies:*

1. 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.*

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

Assumption 1 implies (see Johansen, 1995) that the process y_t is cointegrated of order (1,1)

Econometric Model

Representing VAR model as VECM

(Granger Representation theorem. Granger and Weiss (1983))

Decomposing the matrix lag polynomial $\Pi(L) = \Pi(1)L + \Pi^*(L)(1-L)$

We obtain the vector error correction model (VECM) :

$$\Delta y_t = \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1}\Delta y_{t-p+1} + \varepsilon_t \qquad (3)$$

where

$$\alpha\beta' = -\Pi(1) = -\left[I_n - \left(A_1 + A_2 + \ldots.. + A_p\right)\right]$$

$$\Gamma_j = -\sum_{k=j+1}^{p} A_k \quad (j = 1, \ldots., p-1) \quad and \quad \Gamma_0 = I_n$$

Econometric Model

Common Cyclical Features = = = > Short-run restrictions

There are two type of common cyclical features:

- **SCCF** : Serial Correlation Common Feature (or SF)
- **WF** : Weak Form common cyclical feature

Econometric Model

SCCF restriction

The VECM

$$\Delta y_t = \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1}\Delta y_{t-p+1} + \varepsilon_t \qquad (3)$$

has Serial Correlation Common Feature (SCCF) restriction, if there exist a matrix of $(n \; x \; s)$ $\tilde{\beta}$ such that:

$$\tilde{\beta}'(\Delta y_t) = \tilde{\beta}'\varepsilon_t \qquad (4)$$

Equivalently we assume:

Assumption 3.5 : $\tilde{\beta}'\Gamma_j = 0_{s\times n} \quad for \;\; j = 1,\ldots\ldots,p-1$

Assumption 3.6 : $\tilde{\beta}'\alpha\beta' = 0_{s\times n}$

Econometric Model

WF restriction

The VECM

$$\Delta y_t = \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1} + \varepsilon_t \qquad (3)$$

has Weak Form (WF) restriction, if there exist a matrix of (*n x s*) $\tilde{\beta}$ such that:

$$\tilde{\beta}'\left(\Delta y_t - \alpha\beta' y_{t-1}\right) = \tilde{\beta}'\varepsilon_t \qquad (5)$$

Equivalently we assume:

Assumption 3.8 : $\tilde{\beta}'\Gamma_j = 0_{s\times n} \quad for \;\; j = 1,\ldots\ldots,p-1$

Therefore, SCCF can be renamed as Strong Form (SF) because impose more restrictions over VECM parameters !

Econometric Model

Differences between SF and WF in the VECM model

- Imposing SF, the collection of all coefficient matrices of a VECM has rank less than *n*.
- With WF restrictions the collection of all coefficients matrices, except the long-run matrix has rank less than *n*.

Advantages of WF

- WF allows the study of both cointegration and common cyclical feature without the constraint $r + s \leq n$
- WF is invariant over reparametrization in VECM representation

Therefore, WF impose less restrictions over VECM representation. It may result in a hypothesis less strong on the model especially in economical applications.

Econometric Model

WF restriction

Representing VECM as reduced rank structure

$X_{t-1} = [\Delta y'_{t-1}, \ldots . \Delta y'_{t-p+1}]'$ $\Phi = [\Gamma_1, \ldots, \Gamma_{p-1}]$

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

$$\begin{aligned}\Delta y_t &= \alpha\beta' y_{t-1} + \tilde{\beta}_\perp (\Psi_1, \Psi_2, \ldots, \Psi_{p-1}) X_{t-1} + \varepsilon_t \\ &= \alpha\beta' y_{t-1} + \tilde{\beta}_\perp \Psi X_{t-1} + \varepsilon_t\end{aligned}$$

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

Standard Criteria IC(p) and IC(p,s)

- Standard criteria IC(p)
- Modify informational criteria IC(p,s)

Standard criteria IC(p). See Lütkepohl (1993)

$$y_t = A_1 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p} + \varepsilon_t$$

$$AIC = \ln\left|\sum(p)\right| + \frac{2}{T} n^2 p$$

$$HQ = \ln\left|\sum(p)\right| + \frac{2 \ln \ln T}{T} n^2 p$$

$$SC = \ln\left|\sum(p)\right| + \frac{\ln T}{T} n^2 p$$

The model estimation - Standard Criteria

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

1. Estimate p using IC(p)
2. Find the number of cointegration vector, r using Johansen cointegration test.
3. Conditional the results of cointegration analysis, a final VECM is estimated and then the multi-step ahead forecast is calculated.

Alternative Criteria IC(p,s)

Tso (1981) and Vahid & Issler (2002)

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

$$AIC(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{2}{T} \times [number\ of\ parameters] \qquad (4\text{-}1)$$

$$HQ(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{2\ln(\ln T)}{T}[number\ of\ parameters] \qquad (4\text{-}2)$$

$$SC(p,s) = \sum_{i=n-s+1}^{T} \ln(1-\lambda_i^2(p)) + \frac{\ln T}{T}[number\ of\ parameters] \qquad (4\text{-}3)$$

$$number\ of\ parameter = [n \times (n \times (p-1) + r)] - [s \times (n \times (p-1) + (n-s))]$$

Estimation using Standard Criteria Switching Algorithms. See Hecq (2006).

$$\Delta y_t = \alpha\beta' y_{t-1} + \tilde{\beta}_{\perp}\Psi' X_{t-1} + \varepsilon_t \qquad (4\text{-}5)$$

*Step*1 : Estimation of the cointegration vectors β.

Using the optimal pair $(\bar{p}, \bar{s})$ chosen by information criteria (4-1), (4-2) or (4-3), we estimate β (and so its rank, $r = \bar{r}$) using Johansen cointegration test.

*Step*2 : 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:

$$u_0 = \tilde{\beta}_{\perp}\Psi u_1 + \varepsilon_t \qquad (4\text{-}6)$$

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 (4-6) by FIML. Thus, we can obtain $\tilde{\beta}_{\perp}$ and $\hat{\Psi}$.

Estimation using Standard Criteria Switching Algorithms.

$Step3$: 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:

$$L^0_{\max,\ \bar{r}<n,\ s=\bar{s}} = -\frac{T}{2}\left[\ln\left|\sum_{\bar{r}<n,\ s=\bar{s}}^{\max}\right| - \sum_{i=1}^{\bar{s}} \ln\left(1-\hat{\lambda}_i^2\right)\right] \qquad (4\text{-}7)$$

$Step4$: Restimation of β.

We restimate β to obtain a more appropriated value for the parameters. In order to restimate 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 restimate $\tilde{\beta}_\perp$ 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 $\Delta L = \left(L^1_{\max,\ r=\bar{r},\ s=\bar{s}} - L^0_{\max,\ r=\bar{r},\ s=\bar{s}}\right)$

We repeat the algorithm from steps 1 to 4 to choose $\tilde{\beta}_\perp$and Ψ until convergence is reached (i.e., $\Delta L < 10^{-7}$). At the end, the algorithm optimal

Monte-Carlo design

Data Generator Process – DGP

VAR(p), p=3, n = 3, r = 1 , s = 2

$$\beta = \begin{bmatrix} 1.0 \\ 0.2 \\ -1.0 \end{bmatrix}_{n\times r}, \ \tilde{\beta} = \begin{bmatrix} 1.0 & 0.1 \\ 0.0 & 1.0 \\ 0.5 & -0.5 \end{bmatrix}_{n\times s}$$

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

$$\Delta y_t = (A_1 + A_2 + A_3 - I_3)y_{t-1} - (A_2 + A_3)\Delta y_{t-1} - A_3 \Delta y_{t-2} + \varepsilon_t \qquad (5\text{-}3)$$

a) The cointegration restrictions :

(i) $\alpha\beta' = (A_1 + A_2 + A_3 - I_3)$:

b) WF restrictions :

(ii) $\tilde{\beta}' A_3 = 0$

(iii) $\tilde{\beta}'(A_2 + A_3) = 0$

Monte-Carlo design

c) Covariance-stationary condition.

$$\xi_t = F\ \xi_{t-1} + v_t \tag{5-4}$$

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

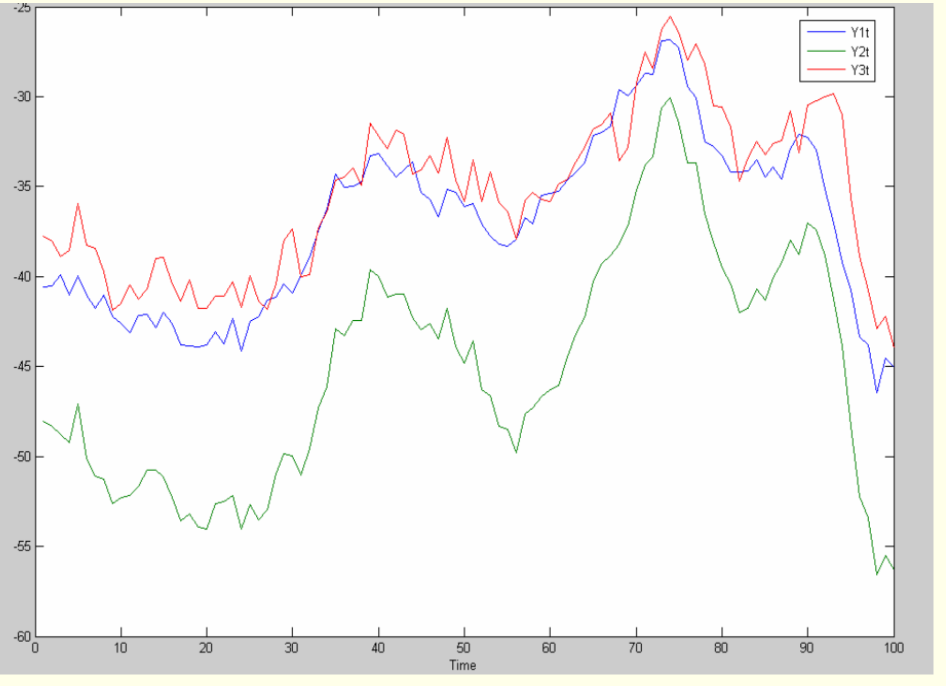
$$v_t = \begin{bmatrix} \varepsilon_t \\ 0 \\ \beta'\varepsilon_t \end{bmatrix}$$

Thus, the equation (5-4) will be covariance-stationary if all eigenvalues of matrix F lie inside the unit circle.

Monte-Carlo Simulation Results

100 DGPs, N = 1000 samples, T = 1000 observations ➔100,000 realizations

Figure 7.1: Simulated series VAR (3), n =3



Monte-Carlo Simulation Results

Table 7.1 Selecting lag order by using IC(*p*)

		Number of observations = 100 Selected cointegrated vectors				Number of observations = 200 Selected cointegrated vectors			
	Selected Lag	0	**1**	2	3	0	1	2	3
	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

Selection of optimal lag length in cointegrated VARs models with common cyclical features

Monte-Carlo Simulation Results

Table 7.2 Selecting lag order by using IC(*p,s*)

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)	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)	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)	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

Selection of optimal lag length in cointegrated VARs models with common cyclical features

Monte-Carlo Simulation Results

Table 7.2 Selecting lag order by using IC(*p,s*)

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 = 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)	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)	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)	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

Conclusions

✓ All criteria (AIC, HQ and SC) chose 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.

✓ Standard model-selection criteria that place a strong penalty on over-parameterization, such as the Schwarz or Hannan-Quinn criteria, may choose too small a lag-length when the true model has common cycles of the WF. However, they improve if the rank order is simultaneously selected with the lag length.

✓ In general, the SC and HQ selection criteria should not be used for this purpose in small samples due to the tendency of identifying an under parameterized model.

This confirms that 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.

Selection of optimal lag length in cointegrated VARs models with common cyclical features

Carlos Enrique Carrasco Gutiérrez

December 2007