View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Research Papers in Economics

ISSN 1518-3548

BANCO CENTRAL DO BRASIL

Working Paper Series 205

Model selection, estimation and forecasting in VAR models with short-run and long-run restrictions

George Athanasopoulos, Osmani Teixeira de Carvalho Guillén,
João Victor Issler and Farshid Vahid

April, 2010

ISSN 1518-3548
CGC 00.038.166/0001-05

Working Paper Series	Brasília	n. 205	Apr.	2010	p. 1-49

Working Paper Series

Edited by Research Department (Depep) – E-mail: workingpaper@bcb.gov.br

Editor: Benjamin Miranda Tabak – E-mail: benjamin.tabak@bcb.gov.br
Editorial Assistant: Jane Sofia Moita – E-mail: jane.sofia@bcb.gov.br
Head of Research Department: Adriana Soares Sales – E-mail: adriana.sales@bcb.gov.br

The Banco Central do Brasil Working Papers are all evaluated in double blind referee process.

Reproduction is permitted only if source is stated as follows: Working Paper n. 205.

Authorized by Carlos Hamilton Vasconcelos Araújo, Deputy Governor for Economic Policy.

General Control of Publications

Banco Central do Brasil

Secre/Surel/Cogiv

SBS – Quadra 3 – Bloco B – Edifício-Sede – 1º andar

Caixa Postal 8.670

70074-900 Brasília – DF – Brazil

Phones: +55 (61) 3414-3710 and 3414-3565

Fax: +55 (61) 3414-3626

E-mail: editor@bcb.gov.br

The views expressed in this work are those of the authors and do not necessarily reflect those of the Banco Central or its members.

Although these Working Papers often represent preliminary work, citation of source is required when used or reproduced.

As opiniões expressas neste trabalho são exclusivamente do(s) autor(es) e não refletem, necessariamente, a visão do Banco Central do Brasil.

Ainda que este artigo represente trabalho preliminar, é requerida a citação da fonte, mesmo quando reproduzido parcialmente.

Consumer Complaints and Public Enquiries Center

Banco Central do Brasil

Secre/Surel/Diate

SBS – Quadra 3 – Bloco B – Edifício-Sede – 2º subsolo

70074-900 Brasília – DF – Brazil

Fax: +55 (61) 3414-2553

Internet: http//www.bcb.gov.br/?english

Model selection, estimation and forecasting in VAR models with short-run and long-run restrictions

George Athanasopoulos[*] Osmani Teixeira de Carvalho Guillén[†] João Victor Issler[‡]
Farshid Vahid[§]

The Working Papers should not be reported as representing the views of the Banco Central do Brasil. The views expressed in the papers are those of the author(s) and not necessarily reflect those of the Banco Central do Brasil.

Abstract

We study the joint determination of the lag length, the dimension of the cointegrating space and the rank of the matrix of short-run parameters of a vector autoregressive (VAR) model using model selection criteria. We consider model selection criteria which have data-dependent penalties as well as the traditional ones. We suggest a new two-step model selection procedure which is a hybrid of traditional criteria and criteria with data-dependant penalties and we prove its consistency. Our Monte Carlo simulations measure the improvements in forecasting accuracy that can arise from the joint determination of lag-length and rank using our proposed procedure, relative to an unrestricted VAR or a cointegrated VAR estimated by the commonly used procedure of selecting the lag-length only and then testing for cointegration. Two empirical applications forecasting Brazilian inflation and U.S. macroeconomic aggregates growth rates respectively show the usefulness of the model-selection strategy proposed here. The gains in different measures of forecasting accuracy are substantial, especially for short horizons.

Keywords: Reduced rank models, model selection criteria, forecasting accuracy.
JEL Classification: C32, C53.

---

[*] Department of Econometrics and Business Statistics, Monash University, Clayton, Victoria 3800, Australia

[†] Banco Central do Brasil and Ibmec-RJ, Brazil.

[‡] Graduate School of Economics – EPGE, Getulio Vargas Foundation, Brazil. Corresponding author. E-mail: joao.Issler@fgv.br

[§] Department of Econometrics and Business Statistics, Monash University, Clayton, Victoria 3800, Australia

1 Introduction

There is a large body of literature on the effect of cointegration on forecasting. Engle and Yoo (1987) compare the forecasts generated from an estimated vector error correction model (VECM) assuming that the lag order and the cointegrating rank are known, with those from an estimated VAR in levels with the correct lag. They find out that the VECM only produces forecasts with smaller mean squared forecast errors (MSFE) in the long-run. Clements and Hendry (1995) note that Engle and Yoo's conclusion is not robust if the object of interest is differences rather than levels, and use this observation to motivate their alternative measures for comparing multivariate forecasts. Hoffman and Rasche (1996) confirm Clements and Hendry's observation using a real data set. Christoffersen and Diebold (1998) also use Engle and Yoo's setup, but argue against using a VAR in levels as a benchmark on the grounds that the VAR in levels not only does not impose cointegration, it does not impose any unit roots either. Instead, they compare the forecasts of a correctly specified VECM with forecasts from correctly specified univariate models, and find no advantage in MSFE for the VECM. They use this result as a motivation to suggest an alternative way of evaluating forecasts of a cointegrated system. Silverstovs et al. (2004) extend Christoffersen and Diebold's results to multicointegrated systems. Since the afore-mentioned papers condition on the correct specification of the lag length and cointegrating rank, they cannot provide an answer as to whether we should examine the cointegrating rank of a system in multivariate forecasting if we do not have any a priori reason to assume a certain form of cointegration.

Lin and Tsay (1996) examine the effect on forecasting of the mis-specification of the cointegrating rank. They determine the lag order using the AIC, and compare the forecasting performance of estimated models under all possible numbers of cointegrating vectors (0 to 4) in a four-variable system. They observe that, keeping the lag order constant, the model with the correct number of cointegrating vectors achieves a lower MSFE for long-run forecasts, especially relative to a model that over-specifies the cointegrating rank. Although Lin and Tsay do not assume the correct specification of the lag length, their study also does not address the uncertainty surrounding the number of cointegrating vectors in a way that can lead to a modelling strategy for forecasting possibly cointegrated variables. Indeed, the results of their example with real data, in which they determine the cointegrating rank using a sequence of hypothesis tests, do not accord with their simulation results.

At the same time, there is an increasing amount of evidence of the advantage of considering rank restrictions for short-term forecasting in stationary VAR (and VARMA) models (see, for example, Ahn and Reinsel 1988, Vahid and Issler 2002, Athanasopoulos and Vahid 2008). One feature of these papers is that they do not treat lag-length and rank uncertainty, differently. Their quest is to identify the dimension of the most parsimonious state vector that can represent the dynamics of a system. Here, we add the cointegrating rank to the menu of unknowns and evaluate model selection criteria that determine all of these unknowns simultaneously. Our goal is to determine a modelling strategy that is useful for multivariate forecasting.

There are other papers in the literature that evaluate the performance of model selection criteria for determining lag-length and cointegrating rank, but they do not evaluate the forecast performance of the resulting models. Gonzalo and Pitarakis (1999) show that in large systems the usual model selection procedures may severely underestimate the cointegrating rank. Chao and Phillips (1999) show that the posterior information criterion (PIC) performs well in choosing the lag-length and the cointegrating rank simultaneously.

In this paper we evaluate the performance of model selection criteria in the simultaneous choice of the lag-length p, the rank of the cointegrating space q, and the rank of other parameter matrices r in a vector error correction model. We suggest a hybrid model selection strategy that selects p and r using a traditional model selection criterion, and then chooses q based on PIC. We then evaluate the forecasting performance of models selected using these criteria.

Our simulations cover the three issues of model building, estimation, and forecasting. We examine the performances of model selection criteria that choose p, r and q simultaneously (IC(p, r, q)), and compare their performances with a procedure that chooses p using a standard model selection criterion (IC(p)) and determines the cointegrating rank using a sequence of likelihood ratio tests proposed by Johansen (1988). We provide a comparison of the forecasting accuracy of fitted VARs when only cointegration restrictions are imposed, when cointegration and short-run restrictions are jointly imposed, and when neither are imposed. These comparisons take into account the possibility of model misspecification in choosing the lag length of the VAR, the number of cointegrating vectors, and the rank of other parameter matrices. In order to estimate the parameters of a model with both long-run and short-run restrictions, we propose a simple iterative procedure similar to the one proposed by Centoni et al. (2007).

It is very difficult to claim that any result found in a Monte Carlo study is general, especially in multivariate time series. There are examples in the VAR literature of Monte Carlo designs which led to all model selection criteria overestimating the true lag in small samples, therefore leading to the conclusion that the Schwarz criterion is the most accurate. The most important feature of these designs is that they have a strong propagation mechanism.[1] There are other designs with weak propagation mechanisms that result in all selection criteria underestimating the true lag and leading to the conclusion that AIC's asymptotic bias in overestimating the true lag may actually be useful in finite samples (see Vahid and Issler 2002, for references). We pay particular attention to the design of the Monte Carlo to make sure that we cover a wide range of data generating processes in terms of the strength of their propagation mechanisms.

---

[1] Our measure of the strength of the propagation mechanism is proportional to the trace of the product of the variance of first differences and the inverse of the variance of innovations.

The outline of the paper is as follows. In Section 2 we study finite VARs with long-run and short-run restrictions and motivate their empirical relevance. In Section 3, we outline an iterative procedure for computing the maximum likelihood estimates of parameters of a VECM with short-run restrictions. We provide an overview of model selection criteria in Section 4, and in particular we discuss model selection criteria with data dependent penalty functions. Section 5 describes our Monte Carlo design. Section 6 presents the simulation results and Section 8 concludes.

2 VAR models with long-run and short-run common factors

We start from the triangular representation of a cointegrated system used extensively in the cointegration literature (some early examples are Phillips and Hansen 1990, Phillips and Loretan 1991, Saikkonen 1992). We assume that the K-dimensional time series

$$y_t = \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix}, \quad t = 1, ..., T$$

where y_{1t} is $q \times 1$ (implying that y_{2t} is $(K-q) \times 1$) is generated from:

$$\begin{aligned} y_{1t} &= \beta y_{2t} + u_{1t} \\ \Delta y_{2t} &= u_{2t} \end{aligned} \tag{1}$$

where β is a $q \times (K-q)$ matrix of parameters, and

$$u_t = \begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix}$$

is a strictly stationary process with mean zero and positive definite covariance matrix. This is a data generating process (DGP) of a system of K cointegrated I(1) variables with q cointegrating vectors, also referred to as a system of K I(1) variables with $K-q$ common stochastic trends (some researchers also refer to this as a system of K variables with $K-q$ unit roots, which can be ambiguous if used out of context, and we therefore do not use it here).[2] The extra feature that we add to this fairly general DGP is that u_t is generated from a VAR of finite order p and rank r $(< K)$.

In empirical applications, the finite VAR(p) assumption is routine. This is in contrast to the theoretical literature on testing for cointegration, in which u_t is assumed to be an infinite VAR, and a finite VAR(p) is used as an approximation (e.g. Saikkonen 1992). Here, our emphasis is on building multivariate forecasting models rather than hypothesis testing. The finite VAR assumption is also routine when the objective is studying the maximum likelihood estimator of the cointegrating vectors, as in Johansen (1988).

---

[2] While in theory every linear system of K cointegrated I(1) variables with q cointegrating vectors can be represented in this way, in practice the decision on how to partition K-variables into y_{1t} and y_{2t} is not trivial, because y_{1t} are variables which must definitely have a non-zero coefficient in the cointegrating relationships.

The reduced rank assumption is considered for the following reasons. Firstly, this assumption means that all serial dependence in the K-dimensional vector time series u_t can be characterised by only $r < K$ serially dependent indices. This is a feature of most macroeconomic models, in which the short-run dynamics of the variables around their steady states are generated by a small number of serially correlated demand or supply shifters. Secondly, this assumption implies that there are $K - r$ linear combinations of u_t that are white noise. Gourieroux and Peaucelle (1992) call such time series "codependent," and interpret the white noise combinations as equilibrium combinations among stationary variables. This is justified on the grounds that, although each variable has some persistence, the white noise combinations have no persistence at all. For instance, if an optimal control problem implies that the policy instrument should react to the current values of the target variables, then it is likely that there will be such a linear relationship between the observed variables up to a measurement noise. Finally, many papers in multivariate time series literature provide evidence of the usefulness of reduced rank VARs for forecasting (see, for example, Velu et al. 1986, Ahn and Reinsel 1988). Recently, Vahid and Issler (2002) have shown that failing to allow for the possibility of reduced rank structure can lead to developing seriously misspecified vector autoregressive models that produce bad forecasts.

The dynamic equation for u_t is therefore given by (all intercepts are suppressed to simplify the notation)

$$u_t = B_1 u_{t-1} + B_2 u_{t-2} + \cdots + B_p u_{t-p} + \varepsilon_t \tag{2}$$

where $B_1, B_2, ..., B_p$ are $K \times K$ matrices with $rank \begin{bmatrix} B_1 & B_2 & ... & B_p \end{bmatrix} = r$, and ε_t is an i.i.d. sequence with mean zero and positive definite variance-covariance matrix and finite fourth moments. Note that the rank condition implies that each B_i has rank at most r, and the intersection of the null-spaces of all B_i is a subspace of dimension $K - r$. The following lemma derives the vector error correction representation of this data generating process.

Lemma 1 *The data generating process given by equations (1) and (2) has a reduced rank vector error correction representation of the type*

$$\Delta y_t = \gamma \begin{pmatrix} I_q & -\beta \end{pmatrix} y_{t-1} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \cdots + \Gamma_p \Delta y_{t-p} + \eta_t, \tag{3}$$

in which $rank \begin{bmatrix} \Gamma_1 & \Gamma_2 & ... & \Gamma_p \end{bmatrix} \leq r$.

Proof. Refer to the working paper version of the current paper. ■

This lemma shows that the triangular DGP (1) under the assumption that the dynamics of its stationary component (i.e. u_t) can be characterised by a small number of common factors, is equivalent to a VECM in which the coefficient matrices of lagged differences have reduced rank and their left null-spaces overlap. Hecq et al. (2006) call such a structure a VECM with weak serial correlation common features (WSCCF).

We should note that the triangular structure (1) implies $K-q$ common Beveridge-Nelson (BN) trends, but the reduced rank structure assumed for u_t does not imply that deviations from the BN trends (usually refereed to as BN cycles) can be characterised as linear combinations of r common factors. Vahid and Engle (1993) show that a DGP with common BN trends and cycles is a special case of the above under some additional restrictions and therefore a stricter form of comevement. Hecq et al. (2006) show that the uncertainty in determining the rank of the cointegrating space can adversely affect inference on common cycles, and they conclude that testing for weak common serial correlation features is a more accurate means of uncovering short-run restrictions in vector error correction models.

Our objective is to come up with a model development methodology that allows for cointegration and weak serial correlation common features. For stationary time series, Vahid and Issler (2002) show that allowing for reduced rank models is beneficial for forecasting. For partially non-stationary time series, there is an added dimension of cointegration. Here, we examine the joint benefits of cointegration and short-run rank restrictions for forecasting partially non-stationary time series.

3 Estimation of VARs with short-run and long-run restrictions

The maximum likelihood estimation of the parameters of a VAR written in error-correction form

$$\Delta y_t = \Pi\, y_{t-1} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \cdots + \Gamma_p \Delta y_{t-p} + \eta_t \tag{4}$$

under the long-run restriction that the rank of Π is q, the short-run restriction that rank of $[\ \Gamma_1 \quad \Gamma_2 \quad \ldots \quad \Gamma_p\ ]$ is r and the assumption of normality, is possible via a simple iterative procedure that uses the general principle of the estimation of reduced rank regression models Anderson (1951). Noting that the above model can be written as

$$\Delta y_t \quad = \quad \gamma\, \alpha' y_{t-1} + C\left[D_1 \Delta y_{t-1} + D_2 \Delta y_{t-2} + \cdots + D_p \Delta y_{t-p}\right] + \eta_t, \tag{5}$$

where α is a $K \times q$ matrix of rank q and C is a $K \times r$ matrix of rank r, one realises that if α was known, C and $D_i, i = 1, \ldots, p$, could be estimated using a reduced rank regression of Δy_t on $\Delta y_{t-1}, \cdots, \Delta y_{t-p}$ after partialling out $\alpha' y_{t-1}$. Also, if $D_i, i = 1, \ldots, p$, were known, then γ and α could be estimated using a reduced rank regression of Δy_t on y_{t-1} after controlling for $\sum_{i=1}^{p} D_i \Delta y_{t-i}$. This points to an easy iterative procedure for computing maximum likelihood estimates for all parameters.

Step 0. Estimate $[\hat{D}_1, \hat{D}_2, \ldots, \hat{D}_p]$ from a reduced rank regression of Δy_t on $(\Delta y_{t-1}, ..., \Delta y_{t-p})$ controlling for y_{t-1}. Recall that these estimates are simply coefficients of the canonical variates corresponding to the r largest squared partial canonical correlations (PCCs) between Δy_t and $(\Delta y_{t-1}, ..., \Delta y_{t-p})$, controlling for y_{t-1}.

Step 1. Compute the PCCs between Δy_t and y_{t-1} conditional on $[\hat{D}_1\Delta y_{t-1} + \hat{D}_2\Delta y_{t-2} + \cdots + \hat{D}_p\Delta y_{t-p}]$. Take the q canonical variates $\hat{\alpha}' y_{t-1}$ corresponding to the q largest squared PCCs as estimates of cointegrating relationships. Regress Δy_t on $\hat{\alpha}' y_{t-1}$ and $[\hat{D}_1\Delta y_{t-1} + \hat{D}_2\Delta y_{t-2} + \cdots + \hat{D}_p\Delta y_{t-p}]$, and compute $\ln|\hat{\Omega}|$, the logarithm of the determinant of the residual variance matrix.

Step 2. Compute the PCCs between Δy_t and $(\Delta y_{t-1}, ..., \Delta y_{t-p})$ conditional on $\hat{\alpha}' y_{t-1}$. Take the r canonical variates $[\hat{D}_1\Delta y_{t-1} + \hat{D}_2\Delta y_{t-2} + \cdots + \hat{D}_p\Delta y_{t-p}]$ corresponding to the largest r PCCs as estimates of $[D_1\Delta y_{t-1} + D_2\Delta y_{t-2} + \cdots + D_p\Delta y_{t-p}]$. Regress Δy_t on $\hat{\alpha}' y_{t-1}$ and $[\hat{D}_1\Delta y_{t-1} + \hat{D}_2\Delta y_{t-2} + \cdots + \hat{D}_p\Delta y_{t-p}]$, and compute $\ln|\hat{\Omega}|$, the logarithm of the determinant of the residual variance matrix. If this is different from the corresponding value computed in Step 1, go back to Step 1. Otherwise, stop.

The value of $\ln|\hat{\Omega}|$ becomes smaller at each stage until it achieves its minimum, which we denote by $\ln|\hat{\Omega}_{p,r,q}|$. The values of $\hat{\alpha}$ and $[\hat{D}_1, \hat{D}_2, \ldots, \hat{D}_p]$ in the final stage will be the maximum likelihood estimators of α and $[D_1, D_2, \ldots, D_p]$. The maximum likelihood estimates of other parameters are simply the coefficient estimates of the final regression. Note that although γ and α, and also C and $[D_1, D_2, \ldots, D_p]$, are only identified up to appropriate normalisations, the maximum likelihood estimates of Π and $[\Gamma_1, \Gamma_2, \ldots, \Gamma_p]$ are invariant to the choice of normalisation. Therefore, the normalisation of the canonical correlation analysis is absolutely innocuous, and the "raw" estimates produced from this procedure can be linearly combined to produce any desired alternative normalisation. Also, the set of variables that are partialled out at each stage should include constants and other deterministic terms if needed.

4 Model selection

The modal strategy in applied work for modelling a vector of I(1) variables is to use a model selection criterion for choosing the lag length of the VAR, then test for cointegration conditional on the lag-order, and finally estimate the VECM. Hardly ever any further step is taken to simplify the model, and if any test of the adequacy of the model is undertaken, it is usually a system test. For example, to test the adequacy of the dynamic specification, additional lags of all variables are added to all equations, and a test of joint significance for K^2 parameters is used. For stationary time series, Vahid and Issler (2002) show that model selection criteria severely underestimate the lag order in weak systems, i.e. in systems where the propagation mechanism is weak. They also show that using model selection criteria (suggested in Lütkepohl 1993, p. 202) to choose lag order and rank simultaneously can remedy this shortcoming significantly.

In the context of VECMs, one can consider selecting (p, r) with these model selection criteria first, and then use a sequence of likelihood ratio tests to determine the rank of the cointegrating space q. Specifically, these are the analogues of Akaike information criterion (AIC), the Hannan and Quinn criterion (HQ) and the Schwarz criterion (SC), and are defined as

$$\begin{aligned} AIC(p,r) &= T\sum_{i=K-r+1}^{K}\ln\left(1-\hat{\lambda}_i^2(p)\right)+2(r(K-r)+rKp) && (6)\\ HQ(p,r) &= T\sum_{i=K-r+1}^{K}\ln\left(1-\hat{\lambda}_i^2(p)\right)+2(r(K-r)+rKp)\ln\ln T && (7)\\ SC(p,r) &= T\sum_{i=K-r+1}^{K}\ln\left(1-\hat{\lambda}_i^2(p)\right)+(r(K-r)+rKp)\ln T, && (8)\end{aligned}$$

where K is the dimension of (number of series in) the system, r is the rank of $[\ \Gamma_1\ \ \Gamma_2\ \ ...\ \ \Gamma_p\ ]$, p is the number of lagged differences in the VECM, T is the number of observations, and $\hat{\lambda}_i^2(p)$ are the sample squared PCCs between Δy_t and the set of regressors $(\Delta y_{t-1}, ..., \Delta y_{t-p})$ after the linear influence of y_{t-1} (and deterministic terms such as a constant term and seasonal dummies if needed) is taken away from them, sorted from the smallest to the largest. Traditional model selection criteria are special cases of the above when rank is assumed to be full, i.e. when r is equal to K. Here, the question of the rank of Π, the coefficient of y_{t-1} in the VECM, is set aside, and taking the linear influence of y_{t-1} away from the dependent variable and the lagged dependent variables concentrates the likelihood on $[\ \Gamma_1\ \ \Gamma_2\ \ ...\ \ \Gamma_p\ ]$. Then, conditional on the p and the r that minimise one of these criteria, one can use a sequence of likelihood ratio tests to determine q. While in the proof of Theorem 2 we show that the estimators of p and r based on HQ and SC are consistent, the estimator of q from the sequential testing method with a fixed level of significance is obviously not. Moreover, the asymptotic distribution of the likelihood ratio test statistic for q conditional on selected p and r may be far from that when the true p and r are known Leeb and Potscher (2005). Here, we study model selection criteria which choose p, r and q.

We consider two classes of model selection criteria. First, we consider direct extensions of the AIC, HQ and SC to the case where the rank of the cointegrating space, which is the same as the rank of Π, is also a parameter to be selected by the criteria. Specifically, we consider

$$\begin{aligned} AIC(p,r,q) &= T\ln|\hat{\Omega}_{p,r,q}|+2(q(K-q)+Kq+r(K-r)+rKp) && (9)\\ HQ(p,r,q) &= T\ln|\hat{\Omega}_{p,r,q}|+2(q(K-q)+Kq+r(K-r)+rKp)\ln\ln T && (10)\\ SC(p,r,q) &= T\ln|\hat{\Omega}_{p,r,q}|+(q(K-q)+Kq+r(K-r)+rKp)\ln T, && (11)\end{aligned}$$

where $\ln|\hat{\Omega}_{p,r,q}|$ (the minimised value of the logarithm of the determinant of the variance of the residuals of the VECM of order p, with Π having rank q and $[\ \Gamma_1\ \ \Gamma_2\ \ ...\ \ \Gamma_p\ ]$ having rank r) is computed by the iterative algorithm described above in Section 3. Obviously, when $q = 0$ or $q = K$, we are back in the straightforward reduced rank regression framework, where one set of eigenvalue calculations for each p provides the value of the log-likelihood function for $r = 1, ..., K$. Similarly, when $r = K$, we are back in the usual VECM estimation, and no iterations are needed.

We also consider a model selection criterion with a data dependent penalty function. Such model selection criteria date back at least to Poskitt (1987), Rissanen (1987) and Wallace and Freeman (1987). The model selection criterion that we consider in this paper is closer to those inspired by the "minimum description length (MDL)" criterion of Rissanen (1987) and the "minimum message length (MML)" criterion of Wallace and Freeman (1987). Both of these criteria measure the complexity of a model by the minimum length of the uniquely decipherable code that can describe the data using the model. Rissanen (1987) establishes that the closest the length of the code of any emprical model can possibly get to the length of the code of the true DGP P_θ is at least as large as $\frac{1}{2}\ln|E_\theta(\mathrm{FIM}_M(\hat{\theta}))|$, where $\mathrm{FIM}_M(\hat{\theta})$ is the Fisher information matrix of model M (i.e., $[-\partial^2 lnl_M/\partial\theta\partial\theta']$, the second derivative of the log-likelihood function of the model M) evaluated at $\hat{\theta}$, and E_θ is the mathematical expectation under P_θ. Rissanen uses this bound as a penalty term to formulate the MDL as a model selection criterion,

$$\mathrm{MDL} = -\ln l_M(\hat{\theta}) + \frac{1}{2}\ln|\mathrm{FIM}_M(\hat{\theta})|.$$

Wallace and Freeman's MML is also based on coding and information theory but is derived from a Bayesian perspective. The MML criterion is basically the same as the MDL plus an additional term that is the prior density of the parameters evaluated at $\hat{\theta}$ (see Wallace 2005, for more details and a summary of recent advances in this line of research). While the influence of this term is dominated by the other two terms as sample size increases, it plays the important role of making the criterion invariant to arbitrary linear transformations of the regressors in a regression context.

Based on their study of the asymptotic form of the Bayesian data density, Phillips (1996) and Phillips and Ploberger (1996) design the posterior information criterion (PIC), which is similar to MML and MDL criteria. Their important contribution has been to show that such criteria can be applied to partially nonstationary time series as well.[3] Chao and Phillips (1999) use the PIC for simultaneous selection of the lag length and cointegration rank in VARs.

There are practical difficulties in working with PIC that motivates simplifying this criterion. One difficulty is that $\mathrm{FIM}_M(\hat{\theta})$ must be derived and coded for all models considered (The details of the Fisher information matrix for a reduced rank VECM is given in the appendix). A more important one is the large dimension of $\mathrm{FIM}_M(\hat{\theta})$. For example, if we want to choose the best VECM allowing for up to 4 lags in a six variable system, we have to compute determinants of square matrices of dimensions as large as 180. These calculations are likely to push the boundaries of numerical accuracy of computers, in particular when these matrices are ill-conditioned[4]. This, and the favourable results of the HQ criterion in selecting lag p and rank of stationary dynamics r, led us to consider a two step procedure.

---

[3] Ploberger and Phillips (2003) generalised Rissanen's result to show that even for trending time series, the distance between any empirical model and the P_θ is larger or equal to $\frac{1}{2}\ln|E_\theta(\mathrm{FIM}_M)|$ almost everywhere on the parameter space. They use the outer-product formulation of the information matrix, which has the same expected value as the negative of the second derivative under P_θ.

4.1 A two-step procedure for model selection

In the first step, the linear influence of y_{t-1} is removed from Δy_t and $(\Delta y_{t-1}, ..., \Delta y_{t-p})$, then HQ$(p, r)$, as defined in (7), is used to determine p and r. Then PIC is calculated for the chosen values of p and r, for all q from 0 to K. This reduces the task to $K + 1$ determinant calculations only.

Theorem 2 *If the data generating process is*

$$\Delta y_t = c + \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \Gamma_2 \Delta y_{t-2} + \cdots + \Gamma_{p_0} \Delta y_{t-p_0} + \eta_t$$

in which
(i) all roots of the characteristic polynomial of the implied VAR for y_t are on or outside the unit circle and all those on the unit circle are $+1$;
(ii) the rank of Π is $q_0 \leq K$, which implies that Π can be written as $\gamma\alpha'$ where γ and α are full rank $K \times q_0$ matrices;
(iii) $\gamma'_\perp \left(I - \sum_{i=1}^{p_0} \Gamma_i\right) \alpha_\perp$ has full rank where $\gamma_\perp$ and $\alpha_\perp$ are full rank $K \times (K - q_0)$ matrices such that $\gamma'_\perp \gamma = \alpha'_\perp \alpha = 0$;
(iv) the rank of $\begin{bmatrix} \Gamma_1 & \Gamma_2 & \ldots & \Gamma_{p_0} \end{bmatrix}$ is $r_0 \leq K$;
(v) the rank of Γ_{p_0} is not zero;
(vi) $E(\eta_t \mid \mathcal{F}_{t-1}) = 0$ and $E(\eta_t \eta'_t \mid \mathcal{F}_{t-1}) = \Omega$ positive definite where $\mathcal{F}_{t-1}$ is the σ-field generated by $\{\eta_{t-1}, \eta_{t-2}, \ldots\}$, and $E(\eta_{it}^4) < \infty$ for $i = 1, 2, \ldots, K$,
and the maximum possible lag considered $p_{\max} \geq p_0$, then the estimators of p, r and q obtained from the two step procedure explained above are consistent.

Proof. See Appendix B. ■

5 Monte-Carlo design

To make the Monte-Carlo simulation manageable, we use a three-dimensional VAR. We consider VARs in levels with lag lengths of 2 and 3, which translates to 1 and 2 lagged differences in the VECM. This choice allows us to study the consequences of both under- and over-parameterisation of the estimated VAR.

---

[4] In our simulations, we came across one case where the determinant was returned to be a small negative number even though the matrix was symmetric positive definite. This happened both using GAUSS and also using MATLAB.

For each p_0, r_0 and q_0 we draw many sets of parameter values from the parameter space of cointegrated VARs with serial correlation common features that generate difference stationary data. In order to ensure that the DGPs considered do not lie in a subset of the parameter space that implies only very weak or only very strong propagation mechanisms we choose 50 DGPs with system R^2s (as defined in Vahid and Issler 2002) that range between 0.3 and 0.65, with a median between 0.4 and 0.5 and 50 DGPs with system R^2s that range between 0.65 and 0.9, with a median between 0.7 and 0.8.

From each DGP, we generate 1,000 samples of 100, 200 and 400 observations (the actual generated samples were longer, but the initial part of each generated sample is discarded to reduce the effect of initial conditions). In summary, our results are based on 1,000 samples of 100 different DGPs — a total of 100,000 different samples — for each of $T = 100$, 200 or 400 observations.

The Monte-Carlo procedure can be summarised as follows. Using each of the 100 DGPs, we generate 1,000 samples (with 100, 200 and 400 observations). We record the lag length chosen by traditional (full-rank) information criteria, labelled IC(p) for IC={AIC, HQ, SC}, and the corresponding lag length chosen by alternative information criteria, labelled IC(p, r, q) for IC={AIC, HQ, SC, PIC, HQ-PIC} where the last is the hybrid procedure we propose in Section 4.1.

We should note that although we present the results averaged over all 100 DGPs we have also analysed the results for the DGPs with low and high R^2s separately. We indeed found that any advantage of model selection criteria with a relatively smaller (larger) penalty factor was accentuated when only considering DGPs with relatively weaker (stronger) propagation mechanisms. In order to save space we do not present these results here but they are available upon request.

For choices made using the traditional IC(p) criteria, we use Johansen's (1988, 1991) trace test at the 5% level of significance to select q, and then estimate a VECM with no short-run restrictions. For choices made using IC(p, r, q), we use the two step procedure of Section 4.1 to obtain the triplet (p, r, q), and then estimate the resulting VECM with SCCF restrictions using the algorithm of Section 3. For each case we record the out-of-sample forecasting accuracy measures for up to 16 periods ahead. We then compare the out-of-sample forecasting accuracy measures for these two types of VAR models.

5.1 Measuring forecast accuracy

We measure the accuracy of forecasts using the traditional trace of the mean-squared forecast error matrix (TMSFE) and the determinant of the mean-squared forecast error matrix |MSFE| at different horizons. We also compute Clements and Hendry's (1993) *generalized forecast error second moment* (GFESM). GFESM is the determinant of the expected value of the outer product of the vector of stacked forecast errors of all future times up to the horizon of interest. For example, if forecasts up to h quarters ahead are of interest, this measure will be:

$$\text{GFESM} = \left| E \begin{pmatrix} \tilde{\varepsilon}_{t+1} \\ \tilde{\varepsilon}_{t+2} \\ \vdots \\ \tilde{\varepsilon}_{t+h} \end{pmatrix} \begin{pmatrix} \tilde{\varepsilon}_{t+1} \\ \tilde{\varepsilon}_{t+2} \\ \vdots \\ \tilde{\varepsilon}_{t+h} \end{pmatrix}' \right|,$$

where $\tilde{\varepsilon}_{t+h}$ is the K-dimensional forecast error of our K-variable model at horizon h. This measure is invariant to elementary operations that involve different variables (TMSFE is not invariant to such transformations), and also to elementary operations that involve the same variable at different horizons (neither TMSFE nor |MSFE| is invariant to such transformations). In our Monte-Carlo, the above expectation is evaluated for every model, by averaging over replications.

There is one complication associated with simulating 100 different DGPs. Simple averaging across different DGPs is not appropriate, because the forecast errors of different DGPs do not have identical variance-covariance matrices. Lütkepohl (1985) normalises the forecast errors by their true variance-covariance matrix in each case before aggregating. Unfortunately, this would be a very time consuming procedure for a measure like GFESM, which involves stacked errors over many horizons. Instead, for each information criterion, we calculate the percentage gain in forecasting measures, comparing the full-rank models selected by IC(p), with the reduced-rank models chosen by IC(p, r, q). This procedure is done at every iteration and for every DGP, and the final results are then averaged.

6 Monte-Carlo simulation results

6.1 Selection of lag, rank, and the number of cointegrating vectors

Simulation results are reported in "three-dimensional" frequency tables. The columns correspond to the percentage of times the selected models had cointegrating rank smaller than the true rank ($q < q_0$), equal to the true rank ($q = q_0$) and larger than the true rank ($q > q_0$). The rows correspond to similar information about the rank of short-run dynamics r. Information about the lag-length is provided within each cell, where the entry is disaggregated on the basis of p. The three numbers provided in each cell, from left to right, correspond to percentages with lag lengths smaller than the true lag, equal to the true lag and larger than the true lag. The 'Total' column on the right margin of each table provides information about marginal frequencies of p and r only. The row titled 'Total' on the bottom margin of each table provides information about the marginal frequencies of p and q only. Finally, the bottom right cell provides marginal information about the lag-length choice only.

We report results of two sets of 100 DGPs. Table 1 summarises the model selection results for 100 DGPs that have one lag in differences with a short-run rank of one and cointegrating rank of two, i.e., $(p_0, r_0, q_0) = (1, 1, 2)$. Table 2 summarises the model selection results for 100 DGPs that have two lags in differences with a short-run rank of one and cointegrating rank of one $(p_0, r_0, q_0) = (2, 1, 1)$. These two groups of DGPs are contrasting in the sense that the second group of DGPs have more severe restrictions in comparison to the first one.

The first three panels of the tables correspond to all model selection based on the traditional model selection criteria. The additional bottom row for each of these three panels provides information about the lag-length and the cointegrating rank, when the lag-length is chosen using the simple version of that model selection criterion and the cointegrating rank is chosen using the Johansen procedure, and in particular the sequential trace test with 5% critical values that are adjusted for sample size. Comparing the rows labelled 'AIC+J', 'HQ+J' and 'SC+J', we conclude that the inference about q is not sensitive to whether the selected lag is correct or not. In Table 1 all three criteria choose the correct q approximately 54%, 59% and 59% of the time for sample sizes 100, 200 and 400, respectively. In Table 2 all three criteria choose the correct q approximately 70%, 82% and 82% of the time for sample sizes 100, 200 and 400, respectively.

From the first three panels of Table 1 we can clearly see that traditional model selection criteria do not perform well in choosing p, r and q jointly in finite samples. The percentages of times the correct model is chosen are only 22%, 26% and 29% with the AIC, 39%, 52% and 62% with HQ, and 42%, 63% and 79% with SC, for sample sizes of 100, 200 and 400, respectively. Note that when we compare the marginal frequencies of (p, r), HQ is the most successful for choosing both p and r, a conclusion that is consistent with results in Vahid and Issler (2002).

The main reason for not being able to determine the triplet (p, r, q) correctly is the failure of these criteria to choose the correct q. Ploberger and Phillips (2003) show that the correct penalty for free parameters in the long-run parameter matrix is larger than the penalty considered by traditional model selection criteria. Accordingly, all three criteria are likely to over-estimate q in finite samples, and of them SC is likely to appear relatively most successful because it assigns a larger penalty to all free parameters, even though the penalty is still less than ideal. This is exactly what the simulations reveal.

The fourth panel of Table 1 includes results for the PIC. The percentages of times the correct model is chosen increase to 52%, 77% and 92% for sample sizes of 100, 200 and 400, respectively. Comparing the margins, it becomes clear that this increased success relative to HQ and SC is almost entirely due to improved precision in the selection of q. The PIC chooses q correctly 76%, 91% and 97% of the time for sample sizes 100, 200 and 400, respectively. Furthermore, for the selection of p and r only, PIC does not improve upon HQ.

Similar conclusions can be reached from the results for the $(2, 1, 1)$ DGPs presented in Table 2. We note that in this case, even though the PIC improves on HQ and SC in choosing the number of cointegrating vectors, it does not improve on HQ or SC in choosing the exact model, because it severely underestimates p. This echoes the findings of Vahid and Issler (2002) in the stationary case that the Schwarz criterion (recall that the PIC penalty is of the same order as the Schwarz penalty in the stationary case) severely underestimates the lag length in small samples in reduced rank VARs.

Our Monte-Carlo results show that the advantage of PIC over HQ and SC is in the determination of the cointegrating rank. Indeed, HQ seems to have an advantage over PIC in selecting the correct p and r in small samples. These results coupled with the practical difficulties in computing the PIC we outline in Section 4 motivated us to consider the two-step alternative procedure to improve the model selection task.

The final panels in Tables 1 and 2 summarise the performance of our two-step procedure. In both tables we can see that the hybrid HQ-PIC procedure improves on all other criteria in selecting the exact model. The improvement is a consequence of the advantage of HQ in selecting p and r better, and PIC in selecting q better.

Note that our hybrid procedure results in over-parameterised models more often than just using PIC as the model selection criterion. We examined whether this trade-off has any significant consequences for forecasting and found that it does not. In all simulation settings, models selected by the hybrid procedure with HQ-PIC as the model selection criteria forecast better than models selected by PIC. Again, we do not present these results here, but they are also available upon request.

6.2 Forecasts

Recall that the forecasting results are expressed as the percentage improvement in forecast accuracy measures of possibly rank reduced models over the unrestricted VAR model in levels selected by SC. Also, note that the object of interest in this forecasting exercise is assumed to be the first difference of variables, although GFESM gives a measure of accuracy that is the same for levels or differences.

We label the models chosen by the hybrid procedure proposed in the previous section and estimated by the iterative process of Section 3 as VECM(HQ-PIC). We label the models estimated by the usual Johansen method with AIC as the model section criterion for the lag order as VECM(AIC+J).

Table 3 presents the forecast accuracy improvements in a $(1, 1, 2)$ setting. In terms of the trace and determinant of the MSFE matrix, there is some improvement in forecasts over unrestricted VAR models at all horizons. With only 100 observations, GFESM worsens for horizons 8 and longer. This means that if the object of interest was some combination of differences across different horizons (for example, the levels of all variables or the levels of some variables and first differences of others), there may not have been any improvement in the MSFE matrix. With 200 or more observations, all forecast accuracy measures show some improvement, with the more substantial improvements being for the one-step-ahead forecasts. Also note that the forecasts of the models selected by the hybrid procedure are almost always better than those produced by the model chosen by the AIC plus Johansen method, which only pays attention to lag-order and long-run restrictions.

Table 4 presents the forecast accuracy improvements in a $(2, 1, 1)$ setting. This set of DGPs have more severe rank reductions than the $(1, 1, 2)$ DGPs, and, as a result, the models selected by the hybrid procedure show more substantial improvements in forecasting accuracy over the VAR in levels, in particular for smaller sample sizes. Forecasts produced by the hybrid procedure are also substantially better than forecasts produced by the AIC+Johansen method, which does not incorporate short-run rank restrictions. Note that although the AIC+Johansen forecasts are not as good as the HQ-PIC forecasts, they are substantially better than the forecasts from unrestricted VARs at short horizons.

Following a request from a referee in Tables 3 and 4 we have also presented Diebold and Mariano (1995) tests for equal predictive accuracy between the rank reduced specifications and the unrestricted VARs for the TMSFE. In general the results are as expected. Models that incorporate reduced rank restrictions rarely forecast significantly worse than the unrestricted models. They either perform the same or significantly better than the unrestricted VARs.

7 Empirical example

The techniques discussed in this paper are applied to two different data sets in forecasting exercises. The first data set contains Brazilian inflation, as measured by three different types of consumer-price indices, available on a monthly basis from 1994:9 through 2009:11, with a span of more than 15 years (183 observations). It was extracted from IPEADATA – a public database with downloadable Brazilian data (http://www.ipeadata.gov.br/). The second data set being analyzed consist of real U.S. per-capita private output[5], personal consumption per-capita, and fixed investment per-capita, available on a quarterly basis from 1947:1 through 2009:03, with a span of more than 62 years (251 observations). It was extracted from FRED's database of the Federal Reserve Bank of St. Louis (http://research.stlouisfed.org/fred2/). Considering that we will keep some observations for forecast evaluation (90 observations), the size of these data bases are close to the number of simulated observations in the Monte-Carlo exercise for $T = 100$ and $T = 200$ respectively.

7.1 Forecasting Brazilian Inflation

The Brazilian data set consists of three alternative measures of CPI price indices. The first is the official consumer price index used in the Brazilian Inflation-Targeting Program. It is computed by IBGE, the statistics bureau of the Brazilian government, labelled here as CPI-IBGE. The second is the consumer price index computed by Getulio Vargas Foundation, a traditional private institution which computes several Brazilian price indices since 1947, labelled here as CPI-FGV. The third is the consumer price index computed by FIPE, an institute of the Department of Economics of the University of São Paulo, labelled here as CPI-FIPE.

---

[5] Private Output is GNP minus Real Federal Consumption Expenditures and Gross Investment.

These three indices capture different aspects of Brazilian consumer-price inflation. First, they differ in terms of geographical coverage. CPI-FGV collects prices in 12 different metropolitan areas in Brazil, 11 of which are also covered by CPI-IBGE[6]. On the other hand, CPI-FIPE only collects prices in São Paulo – the largest city in Brazil – also covered by the other two indices. Tracked consumption bundles are also different across indices. CPI-FGV focus on bundles of the representative consumer with income between 1 and 33 minimum wages. CPI-IBGE focus on bundles of consumers with income between 1 and 40 minimum wages, while CPI-FIPE focus on consumers with income between 1 and 20 minimum wages.

Although all three indices measure consumer-price inflation in Brazil, Granger Causality tests confirm the usefulness of conditioning on alternative indices to forecast any given index in the models estimated here. Despite the existence of these forecasting gains, one should expect a similar pattern for impulse-response functions across models, reflecting a similar response of different price indices to shocks to the dynamic system.

We compare the forecasting performance of (i) the VAR in (log) levels, with lag length chosen by the standard Schwarz criterion; (ii) the VECM, using standard AIC for choosing the lag length and Johansen's test for choosing the cointegrating rank; and (iii) the reduced rank model, with rank and lag length chosen simultaneously using the Hannan-Quinn criterion and cointegrating rank chosen using PIC, estimated by the iterative process described in Section 3. All forecast comparisons are made using the first difference of the (log) levels of the price indices, i.e., price inflation.

For all three models, the estimation sample starts from 1994:09 through 2001:02, with 78 observations. With these initial estimates, we compute the applicable choices of p, r, and q for each model and forecast inflation up to 16 months ahead. Keeping the initial observation fixed (1994:9), we add one observation at the end of the estimation sample, choose potentially different values for p, r, and q for each model, and forecast inflation again up to 16 months ahead. This procedure is then repeated until the final estimation sample reaches 1994:9 through 2008:7, with 167 observations. Then, we have a total of 90 out-of-sample forecasts for each horizon (1 to 16 months ahead), which are used for forecast evaluation. Thus, the estimation sample varies from 78 to 167 observations and mimics closely the simulations labelled $T = 100$ in the Monte-Carlo exercise.

Results of the exercise described above are presented in Table 5. For any horizon, there are substantial forecasting gains of the VECM(HQ-PIC) over the VAR in levels: for example, for 4 months ahead, $GFESM$, $TMSFE$ and $|MSFE|$ show gains of 91.0%, 45.2% and 32.4% respectively. Results for 8 months ahead are even more impressive. Upon comparison, the forecasting gains of the VECM(AIC+J) over the latter are not as large for the shorter horizons, but it increases substantially for horizons 12 and 16 as expected. The comparison between VECM(HQ-PIC) and VECM(AIC+J) shows gains for the former almost everywhere, with substantial improvement in shorter horizons.

---

[6] There are no metropolitan areas covered by CPI-IBGE that are not covered by CPI-FGV.

For 4 months ahead, $GFESM$, $TMSFE$ and $|MSFE|$ show gains of 79.9%, 39.0% and 22.4% respectively.

Table 5 also shows Diebold-Mariano statistics for equal forecast variances among pair of forecasts for individual series and all horizons. As a rule, forecasting using the VAR in levels produces significant higher variances than either the VECM(HQ-PIC) or the VECM(AIC+J). Testing the equality of the variances of the forecast errors using the VECM(HQ-PIC) and the VECM(AIC+J) show significant differences at moderate horizons (4 and 8), although, most of the time, we cannot reject the null of equal variances. It should be noticed that there is no case where either the VAR in levels or the VECM(AIC+J) generate a smaller significant variance vis-a-vis the VECM(HQ-PIC) for any inflation series at all horizons.

It is also worth reporting the final choices of p, r, and q for the best models studied here as the estimation sample goes from 1994:09-2001:02 all the way to 1994:9-2008:7. While the VECM(HQ-PIC) chose $p = 1$[7], $r = 1$ or 2, and $q = 0$[8], most of the time, the VECM(AIC+J) chose $p = 1$[9] and $q = 1$[10], most of the time. Hence, the superior performance of the VECM(HQ-PIC) vis-a-vis the VECM(AIC+J) may be due to either imposing a reduced-rank structure or to ignoring potential cointegration relationships. This is especially true for the shorter horizons.

7.2 Forecasting U.S. Macroeconomic Aggregates

The data being analyzed here is well known. It consists of (log) real U.S. per-capita private output – y, personal consumption per-capita – c, and fixed investment per-capita – i, extracted from FRED's database on a quarterly frequency[11] from 1947:01 through 2009:03.

Again, we compare the forecasting performance of (i) the VAR in (log) levels, with lag length chosen by the standard Schwarz criterion; (ii) the VECM, using standard AIC for choosing the lag length and Johansen's test for choosing the cointegrating rank; and (iii) the reduced rank model, with rank and lag length chosen simultaneously using the Hannan-Quinn criterion and cointegrating rank chosen using PIC, estimated by the iterative process of Section 3. All forecast comparisons are made using the first difference of the (log) levels of the data, i.e., using $\Delta \log (y_t)$, $\Delta \log (c_t)$, and $\Delta \log (i_t)$. For all three models, the estimation sample starts from 1947:01 through 1983:02, with 146 observations. As before, we keep rolling the estimation sample until it reaches 1947:01 through 2005:03, with 235 observations, with a total of 90 out-of-sample forecasts for each horizon used for forecast evaluation. Since the estimation sample varies from 146 to 235 observations it mimics closely the simulations labelled $T = 200$ in the Monte-Carlo exercise.

---

[7] On ocasion it chose $p = 3$.

[8] On ocasion it chose $q = 1$.

[9] On ocasion it chose $p = 5$.

[10] On ocasion it chose $q = 3$.

[11] Using FRED's mnemonics (2010) for the series, the precise definitions are: PCECC96 - consumption, FPIC96 - investment, and (GNP96 - FGCEC96) - output. Population series mnemonics is POP, which is only available from 1952 on in FRED. To get a complete series starting in 1947:01 it was chained with the same series available in DRI database, whose mnemonics is GPOP.

Results of the exercise described above are presented in Table 6. For any horizon, there are considerable forecasting gains for the VECM(HQ-PIC) over the VAR in levels: at 8 quarters ahead, $GFESM$, $TMSFE$ and $|MSFE|$ show gains of 169.2%, 25.3% and 8.4% respectively. The forecasting gains of the VECM(AIC+J) over the VAR in levels are not as large for all horizons. The comparison between VECM(HQ-PIC) and VECM(AIC+J) shows gains for the former everywhere, with higher improvement in shorter horizons. For example, at 4 quarters ahead, $GFESM$, $TMSFE$ and $|MSFE|$ show gains of 44.8%, 33.3% and 21.4% respectively. Despite that, the Diebold-Mariano statistics for equal variances for the forecast errors predicting $\Delta \log (y_t)$, $\Delta \log (c_t)$, and $\Delta \log (i_t)$ are insignificant everywhere, when VECM(HQ-PIC) and VECM(AIC+J) are confronted.

Finally, we investigate the final choices of p, r, and q as the estimation sample progresses from 1947:01-1983:02 to 1947:01-2005:03. For the VECM(HQ-PIC) they are: $p = 1$, $r = 2$, and $q = 0$, everywhere, while the VECM(AIC+J) chose $p = 1$ and $q = 1$[12], most of the time. Hence, the superior performance of the VECM(HQ-PIC) vis-a-vis the VECM(AIC+J) may be due to either imposing a reduced-rank structure or to ignoring potential cointegration relationships.

8 Conclusion

Motivated by the results of Vahid and Issler (2002) on the success of the Hannan-Quinn criterion in selecting the lag length and rank in stationary VARs, and the results of Ploberger and Phillips (2003) and Chao and Phillips (1999) on the generalisation of Rissanen's theorem to trending time series and the success of PIC in selecting the cointegrating rank in VARs, we propose a combined HQ-PIC procedure for the simultaneous choice of the lag-length and the ranks of the short-run and long-run parameter matrices in a VECM and we prove its consistency. Our simulations show that this procedure is capable of selecting the correct model more often than other alternatives such as pure PIC or SC.

In this paper we also present forecasting results that show that models selected using this hybrid procedure produce better forecasts than unrestricted VARs selected by SC and cointegrated VAR models whose lag length is chosen by the AIC and whose cointegrating rank is determined by the Johansen procedure. We have chosen these two alternatives for forecast comparisons because we believe that these are the model selection strategies that are most often used in the empirical literature. However, we have considered several other alternative model selection strategies and the results are qualitatively the same: the hybrid HQ-PIC procedure leads to models that generally forecast better than VAR models selected using other procedures.

A conclusion we would like to highlight is the importance of short-run restrictions for forecasting.

---

[12] On ocasion it chose $q = 0$.

We believe that there has been much emphasis in the literature on the effect of long-run cointegrating restrictions on forecasting. Given that long-run restrictions involve the rank of only one of the parameter matrices of a VECM, and that inference on this matrix is difficult because it involves inference about stochastic trends in variables, it is puzzling that the forecasting literature has paid so much attention to cointegrating restrictions and relatively little attention to lag-order and short-run restrictions in a VECM. The present paper fills this gap and highlights the fact that the lag-order and the rank of short-run parameter matrices are also important for forecasting. Our hybrid model selection procedure and the accompanying simple iterative procedure for the estimation of a VECM with long-run and short-run restrictions provide a reliable methodology for developing multivariate autoregressive models that are useful for forecasting.

How often restrictions of the type considered in this paper are present in VAR approximations to real life data generating processes is an empirical question. Macroeconomic models in which trends and cycles in all variables are generated by a small number of dynamic factors fit in this category. Also, empirical papers that study either regions of the same country or similar countries in the same region often find these kinds of long-run and short-run restrictions. We illustrate the usefulness of the model-selection strategy discussed above in two empirical applications: forecasting Brazilian inflation and U.S. macroeconomic aggregates growth rates. We find gains of imposing short- and long-run restrictions in VAR models, since the VECM(HQ-PIC) and the VECM(AIC+J) outperform the VAR in levels everywhere. Tests of equal variance confirm that these gains are significant. Moreover, ignoring short-run restrictions usually produce inferior forecasts with these data, since the VECM(HQ-PIC) outperforms the VECM(AIC+J) almost everywhere, but these gains are not always significant in tests of equal variance.

It is true that discovering the "true" model is a different objective from model selection for forecasting. However, in the context of partially non-stationary variables, there are no theoretical results that lead us to a definite model selection strategy for forecasting. Using a two variable example, Elliott (2006) shows that, ignoring estimation uncertainty, whether or not considering cointegration will improve short-run or long-run forecasting depends on all parameters of the DGP, even the parameters of the covariance matrix of the errors. In addition there is no theory that tells us whether finite sample biases of parameter estimates will help or hinder forecasting in partially non-stationary VARs. Given this state of knowledge, when one is given the task of selecting a single model for forecasting it is reasonable to use a model selection criterion that is more likely to pick the "true" model and in this paper we verify that VARs selected by our hybrid model selection strategy are likely to produce better forecasts than unrestricted VARs and VARs that only incorporate cointegration restrictions.

References

Ahn, S. K. & Reinsel, G. C. (1988), 'Nested reduced-rank autoregressive models for multiple time series', *Journal of the American Statistical Association* **83**, 849–856.

Anderson, H. & Vahid, F. (1998), 'Testing multiple equation systems for common nonlinear components', *Journal of Econometrics* **84**, 1–36.

Anderson, T. (1951), 'Estimating linear restrictions on regression coefficients for multivariate normal distributions', *Annals of Mathematical Statistics* **22**, 327–351.

Athanasopoulos, G. & Vahid, F. (2008), 'VARMA versus VAR for macroeconomic forecasting', *Journal of Business and Economic Statistics* **26**, 237–252.

Aznar, A. & Salvador, M. (2002), 'Selecting the rank of the cointegation space and the form of the intercept using an information criterion', *Econometric Theory* **18**, 926–947.

Centoni, M., Cubbada, G. & Hecq, A. (2007), 'Common shocks, common dynamics and the international business cycle', *Economic Modelling* **24**, 149–166.

Chao, J. & Phillips, P. (1999), 'Model selection in partially nonstationary vector autoregressive processes with reduced rank structure', *Journal of Econometrics* **91**, 227–271.

Christoffersen, P. & Diebold, F. (1998), 'Cointegration and long-horizon forecasting', *Journal of Business and Economic Statistics* **16**, 450–458.

Clements, M. P. & Hendry, D. F. (1993), 'On the limitations of comparing mean squared forecast errors (with discussion)', *Journal of Forecasting* **12**, 617–637.

Clements, M. P. & Hendry, D. F. (1995), 'Forecasting in cointegrated systems', *Journal of Applied Econometrics* **10**, 127–146.

Diebold, F. X. & Mariano, R. S. (1995), 'Comparing predictive accuracy', *Journal of Business and Economic Statistics* **13**, 253–263.

Elliott, G. (2006), Forecasting with trending data, *in* G. Elliott, C. Granger & A. Timmermann, eds, 'Handbook of Economic Forecasting', Vol. 1, Elsevier, chapter 11, pp. 555–604.
URL: *http://ideas.repec.org/h/eee/ecofch/1-11.html*

Engle, R. F. & Yoo, S. (1987), 'Forecasting and testing in cointegrated systems', *Journal of Econometrics* **35**, 143–159.

Gonzalo, J. & Pitarakis, J. (1995), 'Specification via model selection in vector error correction models', *Economic Letters* **60**, 321–328.

Gonzalo, J. & Pitarakis, J.-Y. (1999), Dimensionality effect in cointegration tests, *in* R. Engle & H. White, eds, 'Cointegration, Causality and Forecasting: A Festschrift in Honour of Clive W. J. Granger', New York: Oxford University Press, chapter 9, pp. 212–229.

Gourieroux, C. & Peaucelle, I. (1992), 'Series codependantes application a l'hypothese de parite du pouvoir d'achat', *Revue d'Analyse Economique* **68**, 283–304.

Hecq, A., Palm, F. & Urbain, J.-P. (2006), 'Common cyclical features analysis in VAR models with cointegration', *Journal of Econometrics* **132**, 117–141.

Hoffman, D. & Rasche, R. (1996), 'Assessing forecast performance in a cointegrated system', *Journal of Applied Econometrics* **11**, 495–517.

Johansen, S. (1988), 'Statistical analysis of cointegrating vectors', *Journal of Economic Dynamics and Control* **12**, 231–254.

Johansen, S. (1991), 'Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models', *Econometrica* **59**, 1551–1580.

Leeb, H. & Potscher, B. (2005), 'Model selection and inference: Facts and fiction', *Econometric Theory* **21**, 21–59.

Lin, J. L. & Tsay, R. S. (1996), 'Cointegration constraints and forecasting: An empirical examination', *Journal of Applied Econometrics* **11**, 519–538.

Lütkepohl, H. (1985), 'Comparison of criteria for estimating the order of a vector autoregressive process', *Journal of Time Series Analysis* **9**, 35–52.

Lütkepohl, H. (1993), *Introduction to Multiple Time Series Analysis*, 2nd edn, Springer-Verlag, Berlin-Heidelberg.

Magnus, J. R. & Neudecker, H. (1988), *Matrix Differential Calculus with Applications in Statistics and Econometrics*, New York: John Wiley and Sons.

Paulsen, J. (1984), 'Order determination of multivariate autoregressive time series with unit roots', *Journal of Time Series Analysis* **5**, 115–127.

Phillips, P. C. B. (1996), 'Econometric model determination', *Econometrica* **64**, 763–812.

Phillips, P. C. B. & Hansen, B. (1990), 'Statistical inference in instrumental variables regression with I(1) processes', *Review of Economic Studies* **57**, 99–125.

Phillips, P. C. B. & Loretan, M. (1991), 'Estimating long-run economic equilibria', *Review of Economic Studies* **58**, 407–436.

Phillips, P. & Ploberger, W. (1996), 'An asymptotic theory of Bayesian inference for time series', *Econometrica* **64**, 381–413.

Ploberger, W. & Phillips, P. C. B. (2003), 'Empirical limits for time series econometric models', *Econometrica* **71**, 627–673.

Poskitt, D. S. (1987), 'Precision, complexity and bayesian model determination', *Journal of the Royal Statistical Society B* **49**, 199–208.

Quinn, B. (1980), 'Order determination for a multivariate autoregression', *Journal of the Royal Statistical Society B* **42**, 182–185.

Reinsel, G. C. (1997), *Elements of Multivariate Time Series*, 2nd edn, New York: Springer-Verlag.

Rissanen, J. (1987), 'Stochastic complexity', *Journal of the Royal Statistical Society B* **49**, 223–239.

Saikkonen, P. (1992), 'Estimation and testing of cointegrated systems by an autoregressive approximation', *Econometric Theory* **8**, 1–27.

Silverstovs, B., Engsted, T. & Haldrup, N. (2004), 'Long-run forecasting in multicointegrated systems', *Journal of Forecasting* **23**, 315–335.

Sims, C., Stock, J. & Watson, M. (1990), 'Inference in linear time series models with some unit roots', *Econometrica* **58**, 113–144.

Tsay, R. (1984), 'Order selection in nonstationary autoregressive models', *Annals of Statistics* **12**, 1425–1433.

Vahid, F. & Engle, R. F. (1993), 'Common trends and common cycles', *Journal of Applied Econometrics* **8**, 341–360.

Vahid, F. & Issler, J. V. (2002), 'The importance of common cyclical features in VAR analysis: A Monte-Carlo study', *Journal of Econometrics* **109**, 341–363.

Velu, R., Reinsel, G. & Wickern, D. (1986), 'Reduced rank models for multiple time series', *Biometrika* **73**, 105–118.

Wallace, C. (2005), *Statistical and Inductive Inference by Minimum Message Length*, Berlin: Springer.

Wallace, C. & Freeman, P. (1987), 'Estimation and inference by compact coding', *Journal of the Royal Statistical Society B* **49**, 240–265.

A The Fisher information matrix of the reduced rank VECM

Assuming that the first observation in the sample is labelled observation $-p+1$ and that the sample contains $T+p$ observations, we write the K-variable reduced rank VECM

$$\Delta y_t = \gamma' \begin{pmatrix} I_q & \beta' \end{pmatrix} y_{t-1} + \begin{pmatrix} I_r \\ C' \end{pmatrix} \left[D_1 \Delta y_{t-1} + D_2 \Delta y_{t-2} + \cdots + D_p \Delta y_{t-p} \right] + \mu + e_t,$$

or in stacked form

$$\Delta Y = Y_{-1} \begin{pmatrix} I_q \\ \beta \end{pmatrix} \gamma + WD \begin{pmatrix} I_r & C \end{pmatrix} + \iota_T \mu' + E,$$

where

$$\begin{aligned} \underset{T\times K}{\Delta Y} &= \begin{bmatrix} \Delta y_1' \\ \vdots \\ \Delta y_T' \end{bmatrix}, \quad \underset{T\times K}{Y_{-1}} = \begin{bmatrix} y_0' \\ \vdots \\ y_{T-1}' \end{bmatrix}, \quad \underset{T\times K}{E} = \begin{bmatrix} e_1' \\ \vdots \\ e_T' \end{bmatrix} \\ \underset{T\times Kp}{W} &= \begin{pmatrix} \Delta Y_{-1} & \cdots & \Delta Y_{-p} \end{pmatrix} = \begin{bmatrix} \Delta y_0' & \cdots & \Delta y_{-p+1}' \\ \vdots & \vdots & \vdots \\ \Delta y_{T-1}' & \cdots & \Delta y_{T-p}' \end{bmatrix} \\ \underset{Kp\times r}{D} &= \begin{pmatrix} D_1' \\ \vdots \\ D_p' \end{pmatrix}, \end{aligned}$$

and ι_T is a $T \times 1$ vector of ones. When e_t are $N(0, \Omega)$ and serially uncorrelated, the log-likelihood function, conditional on the first p observations being known, is:

$$\begin{aligned} \ln l(\theta, \omega) &= -\frac{KT}{2} \ln(2\pi) - \frac{T}{2} \ln|\Omega| - \frac{1}{2} \sum_{t=1}^{T} e_t' \Omega^{-1} e_t \\ &= -\frac{KT}{2} \ln(2\pi) - \frac{T}{2} \ln|\Omega| - \frac{1}{2} tr\left(E\Omega^{-1}E'\right), \end{aligned}$$

where

$$\theta = \begin{pmatrix} vec(\beta) \\ vec(\gamma) \\ vec(D) \\ vec(C) \\ \mu \end{pmatrix}$$

is a $(K-q)q + Kq + Kpr + r(K-r) + K$ matrix of mean parameters, and $\omega = vech(\Omega)$ is a $K(K+1)/2$ vector of unique elements of the variance matrix. The differential of the log-likelihood is (see Magnus and Neudecker 1988)

$$\begin{aligned} d \ln l(\theta, \omega) &= -\frac{T}{2} tr \Omega^{-1} d\Omega + \frac{1}{2} tr\left(\Omega^{-1} d\Omega \Omega^{-1} E'E\right) - \frac{1}{2} tr\left(\Omega^{-1} E' dE\right) - \frac{1}{2} tr\left(\Omega^{-1} dE' E\right) \\ &= \frac{1}{2} tr\left(\Omega^{-1}\left(E'E - T\Omega\right)\Omega^{-1} d\Omega\right) - tr\left(\Omega^{-1} E' dE\right), \end{aligned}$$

and the second differential is:

$$\begin{aligned} d^2 \ln l(\theta, \omega) &= tr\left(d\Omega^{-1}\left(E'E - T\Omega\right)\Omega^{-1} d\Omega\right) + \frac{1}{2} tr\left(\Omega^{-1}\left(2E'dE - Td\Omega\right)\Omega^{-1} d\Omega\right) \\ &\quad - tr\left(d\Omega^{-1} E' dE\right) - tr\left(\Omega^{-1} dE' dE\right). \end{aligned}$$

Since we eventually want to evaluate the Fisher information matrix at the maximum likelihood estimator, and at the maximum likelihood estimator $\hat{E}'\hat{E} - T\hat{\Omega} = 0$, and also $\hat{\Omega}^{-1}\hat{E}'dE/d\theta = 0$ (these are apparent from the first differentials), we can delete these terms from the second differential, and use $tr\,(AB) = vec\,(A')'\,vec\,(B)$ to obtain

$$\begin{aligned} d^2 \ln l\,(\theta, \omega) &= -\frac{T}{2} tr\left(\Omega^{-1} d\Omega \Omega^{-1} d\Omega\right) - tr\left(\Omega^{-1} dE' dE\right) \\ &= -\frac{T}{2}\,(d\omega)'\,\mathbf{D}_K'\left(\Omega^{-1} \otimes \Omega^{-1}\right)\mathbf{D}_K d\omega - (vec\,(dE))'\left(\Omega^{-1} \otimes I_T\right) vec\,(dE)\,, \end{aligned}$$

where $\mathbf{D}_K$ is the "duplication matrix". From the model, we can see that

$$dE = -Y_{-1}\begin{pmatrix} 0 \\ d\beta \end{pmatrix}\gamma - Y_{-1}\begin{pmatrix} I_q \\ \beta \end{pmatrix} d\gamma - WdD\begin{pmatrix} I_r & C \end{pmatrix} - WD\begin{pmatrix} 0 & dC \end{pmatrix} - \iota_T d\mu',$$

and therefore

$$vec\,(dE) = -\left[\begin{array}{ccccc} \gamma' \otimes Y_{-1}^{(2)} & I_K \otimes Y_{-1}\begin{pmatrix} I_q \\ \beta \end{pmatrix} & \begin{pmatrix} I_r \\ C' \end{pmatrix} \otimes W & \begin{pmatrix} 0 \\ I_{K-r} \end{pmatrix} \otimes WD & I_K \otimes \iota_T \end{array}\right] d\theta.$$

Hence, the elements of the Fisher information matrix are:

$$\begin{aligned}
FIM_{11} &= \gamma\Omega^{-1}\gamma' \otimes Y_{-1}^{(2)\prime}Y_{-1}^{(2)}, \quad FIM_{12} = \gamma\Omega^{-1} \otimes Y_{-1}^{(2)\prime}Y_{-1}\begin{pmatrix} I_q \\ \beta \end{pmatrix}, \\
FIM_{13} &= \gamma\Omega^{-1}\begin{pmatrix} I_r \\ C' \end{pmatrix} \otimes Y_{-1}^{(2)\prime}W, \quad FIM_{14} = \gamma\Omega^{-1}\begin{pmatrix} 0 \\ I_{K-r} \end{pmatrix} \otimes Y_{-1}^{(2)\prime}WD \\
FIM_{15} &= \gamma\Omega^{-1} \otimes Y_{-1}^{(2)\prime}\iota_T \\
FIM_{22} &= \Omega^{-1} \otimes \begin{pmatrix} I_q & \beta' \end{pmatrix} Y_{-1}'Y_{-1}\begin{pmatrix} I_q \\ \beta \end{pmatrix}, \quad FIM_{23} = \Omega^{-1}\begin{pmatrix} I_r \\ C' \end{pmatrix} \otimes \begin{pmatrix} I_q & \beta' \end{pmatrix} Y_{-1}'W \\
FIM_{24} &= \Omega^{-1}\begin{pmatrix} 0 \\ I_{K-r} \end{pmatrix} \otimes \begin{pmatrix} I_q & \beta' \end{pmatrix} Y_{-1}'WD, \quad FIM_{25} = \Omega^{-1} \otimes \begin{pmatrix} I_q & \beta' \end{pmatrix} Y_{-1}'\iota_T \\
FIM_{33} &= \begin{pmatrix} I_r & C \end{pmatrix}\Omega^{-1}\begin{pmatrix} I_r \\ C' \end{pmatrix} \otimes W'W, \quad FIM_{34} = \begin{pmatrix} I_r & C \end{pmatrix}\Omega^{-1}\begin{pmatrix} 0 \\ I_{K-r} \end{pmatrix} \otimes W'WD \\
FIM_{35} &= \begin{pmatrix} I_r & C \end{pmatrix}\Omega^{-1} \otimes W'\iota_T \\
FIM_{44} &= \begin{pmatrix} 0 & I_{K-r} \end{pmatrix}\Omega^{-1}\begin{pmatrix} 0 \\ I_{K-r} \end{pmatrix} \otimes D'W'WD, \quad FIM_{45} = \begin{pmatrix} 0 & I_{K-r} \end{pmatrix}\Omega^{-1} \otimes D'W'\iota_T \\
FIM_{55} &= \Omega^{-1} \otimes \iota_T'\iota_T = \Omega^{-1} \times T
\end{aligned}$$

B Proof of Theorem 2

The first three assumptions ensure that $\Delta\, y_t$ is covariance stationary and y_t are cointegrated with cointegrating rank q_0. These together with assumption (vi) ensure that all sample means and covariances of Δy_t consistently estimate their population counterparts and the least squares estimator of parameters is consistent. Assumptions (iv) and (v) state that the true rank is r_0 and the true lag-length is p_0 (or the lag order of the implied VAR in levels is $p_0 + 1$). For any (p, r) pair, the second step of the analysis produces the least squares estimates of $\Gamma_1, \ldots, \Gamma_p$ with rank r when no restrictions are imposed on Π Anderson (1951). Reinsel (1997) contains many of the results that we use in this proof). Under the assumption of normality, these are the ML estimates of $\Gamma_1, \ldots, \Gamma_p$ with rank r with Π unrestricted and the resulting $\hat{\Omega}_{p,r}$ used in the HQ procedure is the corresponding ML estimate of Ω. Note that normality of the true errors is not needed for the proof. We use the results of Sims et al. (1990) who show that in the above model the least squares estimates of $\Gamma_1, \ldots, \Gamma_p$ have the standard asymptotic properties as in stationary VARs, in particular that they consistently estimate their population counterparts and that their rate of convergence is the same as $T^{-\frac{1}{2}}$. Let $z_t, z_{t-1}, \ldots, z_{t-p}$ denote $\Delta y_t, \Delta y_{t-1}, ..., \Delta y_{t-p}$ after the influence of the constant and y_{t-1} is removed from them and let $Z, Z_{-1}, \ldots, Z_{-p}$ denote $T \times K$ matrices with $z'_t, z'_{t-1}, \ldots, z'_{t-p}$ in their row $t = 1, \ldots, T$ (we assume that the sample starts from $t = -p_{\max} + 1$), and let $W_p = [Z_{-1} \vdots \cdots \vdots Z_{-p}]$ and $B_p = [\Gamma_1 \vdots \cdots \vdots \Gamma_p]'$. The estimated model in the second step can be written as:

$$Z = W_p \hat{B}_p + \hat{U}_p$$

where $\hat{U}_p$ is the $T \times K$ matrix of residuals when the lag length is p. In an unrestricted regression $\ln|\frac{1}{T}\hat{U}'_p\hat{U}_p| = \ln|\frac{1}{T}(Z'Z - Z'W_p(W'_pW_p)^{-1}W'_pZ)| = \ln|\frac{1}{T}Z'Z| + \ln|I_K - (Z'Z)^{-1}Z'W_p(W'_pW_p)^{-1}W'_pZ|$ $= \ln|\frac{1}{T}Z'Z| + \sum_{i=1}^{K}\ln(1 - \hat{\lambda}^2_i(p))$, where $\hat{\lambda}^2_1(p) \leq \hat{\lambda}^2_2(p) \leq \ldots \leq \hat{\lambda}^2_K(p)$, the eigenvalues of $(Z'Z)^{-1}Z'W_p(W'_pW_p)^{-1}W'_pZ$ are the ordered sample partial canonical correlations between Δy_t and $\Delta y_{t-1}, ..., \Delta y_{t-p}$ after the influence of a constant and y_{t-1} has been removed. Under the restriction that the rank of B is r, the log-determinant of the squared residuals matrix becomes $\ln|\frac{1}{T}\hat{U}'_{p,r}\hat{U}_{p,r}| = \ln|\frac{1}{T}Z'Z| + \sum_{i=K-r+1}^{K}\ln(1 - \hat{\lambda}^2_i(p))$. Further, note that $W_p = [W_{p-1} \vdots Z_{-p}]$ and from the geometry of least squares we know
$Z'W_p(W'_pW_p)^{-1}W'_pZ = Z'W_{p-1}(W'_{p-1}W_{p-1})^{-1}W'_{p-1}Z + Z'Q_{p-1}Z_{-p}(Z'_{-p}Q_{p-1}Z_{-p})^{-1}Z'_{-p}Q_{p-1}Z$ where $Q_{p-1} = I_T - W_{p-1}(W'_{p-1}W_{p-1})^{-1}W'_{p-1}$.
(i) Consider $p = p_0$ and $r = r_0 - 1$: $\ln|\frac{1}{T}\hat{U}'_{p_0,r_0-1}\hat{U}_{p_0,r_0-1}| - \ln|\frac{1}{T}\hat{U}'_{p_0,r_0}\hat{U}_{p_0,r_0}| = -\ln(1 - \hat{\lambda}^2_{K-r_0+1}(p_0))$. $\hat{\lambda}^2_{K-r_0+1}(p_0)$ converges in probability to its population counterpart, the r_0-th largest eigenvalue of $\Sigma_z^{-1}B'_{p_0}\Sigma_w B_{p_0}$, where Σ_x denotes the population second moment of the vector x. This population canonical correlation is strictly greater than zero because B_{p_0} has rank r_0. Therefore
$p\lim\ (\ln|\frac{1}{T}\hat{U}'_{p_0,r_0-1}\hat{U}_{p_0,r_0-1}| - \ln|\frac{1}{T}\hat{U}'_{p_0,r_0}\hat{U}_{p_0,r_0}|) = -\ln(1 - \lambda^2_{K-r_0+1}(p_0)) > 0$.
(ii) Consider $p = p_0 - 1$ and $r = r_0$:

$$\begin{aligned} (Z'Z)^{-1}Z'W_{p_0}(W'_{p_0}W_{p_0})^{-1}W'_{p_0}Z &= (Z'Z)^{-1}Z'W_{p_0-1}(W'_{p_0-1}W_{p_0-1})^{-1}W'_{p_0-1}Z \\ &\quad + (Z'Z)^{-1}Z'Q_{p_0-1}Z_{-p_0}(Z'_{-p_0}Q_{p_0-1}Z_{-p_0})^{-1}Z'_{-p_0}Q_{p_0-1}Z. \end{aligned}$$

Since the second matrix on the right side is positive semi-definite, it follows that $\hat{\lambda}^2_i(p_0 - 1) \leq \hat{\lambda}^2_i(p_0)$ for all $i = 1, \ldots, K$.[13] We know that the probability limits of the smallest $K - r_0$ eigenvalues $\hat{\lambda}^2_i(p_0)$

---

[13] Some textbooks define positive definiteness and associated inequalities concerning ordered eigenvalues for symmetric matrices only. Note that since the eigenvalues of any square matrix A is the same as the eigenvalues of GAG^{-1}

are zero. Therefore, the probability limits of the smallest $K - r_0$ eigenvalues $\hat{\lambda}_i^2(p_0 - 1)$ must also be zero. Moreover, the trace of the matrix on the left is equal to the sum of the traces of the two matrices on the right of the equal sign. The probability limit of the last matrix on the right side is $\Sigma_z^{-1}\Gamma'_{p_0}\Sigma_{z.w}\Gamma_{p_0}$ where $\Sigma_{z.w} = p\lim(\frac{1}{T}Z'_{-p_0}Q_{p_0-1}Z_{-p_0})$, and since $rank(\Gamma_{p_0}) > 0$ by assumption, the probability limit of the trace of the second matrix on the right hand side will be strictly positive (note that even when $\Gamma_{p_0}$is nilpotent (i.e. has all zero eigenvalues even though its rank is not zero), $\Sigma_z^{-1}\Gamma'_{p_0}\Sigma_{z.w}\Gamma_{p_0}$ will not be nil-potent). Therefore it must be that $p\lim \hat{\lambda}_i^2(p_0 - 1) < p\lim \hat{\lambda}_i^2(p_0)$ for at least one $i = r_0 + 1, \ldots, K$. This implies that $p\lim\ (\ln|\frac{1}{T}\hat{U}'_{p_0-1,r_0}\hat{U}_{p_0-1,r_0}| - \ln|\frac{1}{T}\hat{U}'_{p_0,r_0}\hat{U}_{p_0,r_0}|) = \sum_{i=K-r_0+1}^{K}(\ln(1-\lambda_i^2(p_0-1)) - \ln(1-\lambda_i^2(p_0))) > 0$.

(i) and (ii), together with the fact that $|\hat{U}'_{p_1,r_1}\hat{U}_{p_1,r_1}| \geq |\hat{U}'_{p_2,r_2}\hat{U}_{p_2,r_2}|$ whenever $p_1 \leq p_2$ and $r_1 \leq r_2$ (i.e., for all nested models the less restrictive cannot fit worse) imply that the probability limit of $\ln|\frac{1}{T}\hat{U}'_{p_0,r_0}\hat{U}_{p_0,r_0}|$ is *strictly smaller* than the probability limit of $\ln|\frac{1}{T}\hat{U}'_{p,r}\hat{U}_{p,r}|$ for all ($p \leq p_0$ and $r < r_0$) or ($p < p_0$ and $r \leq r_0$). Although the penalty favours the smaller models, the reward for parsimony increases at rate $\ln\ln T$ while the reward for better fit increases at rate T and therefore dominates. Hence, the probability of choosing a model with ($p \leq p_0$ and $r < r_0$) or ($p < p_0$ and $r \leq r_0$) goes to zero asymptotically.

(i') In (i), replace $p = p_0$ with $p = \tilde{p} \geq p_0$. The model now includes redundant lags whose true coefficients are zero and these coefficients are consistently estimated. Moreover, adding these zero parameters does not change the rank. Therefore all arguments in (i) apply to this case also, and we can therefore deduce that the probability of under-estimating r with this procedure goes to zero asymptotically.

(ii') In (ii), replace $r = r_0$ with $r = \tilde{r} \geq r_0$. The model now does not impose all rank restrictions that the true data generating process includes, but the extra eigenvalues will converge to their true value of zero asymptotically and all arguments in (ii) apply to this case also. Therefore, we can conclude that the probability of under-estimating p with this procedure goes to zero asymptotically.

(iii) Consider $p = \tilde{p} \geq p_0$ and $r = \tilde{r} \geq r_0$ with at least one of the inequalities strict. These are all models that are larger than the true model and nest the true model. The probability limit of $\ln|\frac{1}{T}\hat{U}'_{\tilde{p},\tilde{r}}\hat{U}_{\tilde{p},\tilde{r}}|$ for these models is the same as the probability limit of $\ln|\frac{1}{T}\hat{U}'_{p_0,r_0}\hat{U}_{p_0,r_0}|$. However, we know that $T(\ln|\frac{1}{T}\hat{U}'_{p_0,r_0}\hat{U}_{p_0,r_0}| - \ln|\frac{1}{T}\hat{U}'_{\tilde{p},\tilde{r}}\hat{U}_{\tilde{p},\tilde{r}}|)$ is the likelihood ratio statistic of testing general linear restrictions that reduce the $\tilde{p},\tilde{r}$ model to the p_0, r_0 model. Since these restrictions are true, $T(\ln|\frac{1}{T}\hat{U}'_{p_0,r_0}\hat{U}_{p_0,r_0}| - \ln|\frac{1}{T}\hat{U}'_{\tilde{p},\tilde{r}}\hat{U}_{\tilde{p},\tilde{r}}|) = O_p(1)$. While the reward for better fit from larger models is bounded in probability, the penalty terms for extra parameters increases without bound. Hence, the probability of choosing a larger model that nests the true model goes to zero asymptotically. This

---

for any invertible matrix G with the same dimensions as A (see Magnus and Neudecker 1988, Chapter 1) one can choose $G = (Z'Z)^{\frac{1}{2}}$ and make all matrices on both sides of the inequality symmetric without changing any of their eigenvalues. Indeed this is a useful transformation for calculating canonical correlations because computer procedures for computation of eigenvalues of symmetric matrices are more accurate than those for general matrices.

completes the proof that the first step of the procedure consistently estimates p_0 and r_0.

For the consistency of the second step estimator of q_0, we note that Chao and Phillips (1999) show that the PIC can be written as the sum (Chao and Phillips 1999, express PIC as product of the likelihood and penalty term, here we refer to the logarithmic transformation of the PIC expressed in their paper) of two parts, one that comprises the log-likelihood of q given p and its associated penalty, and the other that comprises the log-likelihood of p without any restrictions on q and a penalty term involving the lag-length. With similar steps one can write the PIC in our case as the sum of one part related to q given p and r and another that involves p and r. Hence, plugging in p and r that are estimated via another consistent procedure does not alter the consistency of the estimator of q. The main reason that the choice of p and r does not affect the consistency of q is that the smallest $K - q_0$ sample squared canonical correlations between Δy_t and y_{t-1} converge to zero in probability and the remaining q_0 converge to positive limits, *regardless of any finite stationary elements that are partialed out.* Therefore, for a given (p, r) when $q < q_0$, T times the difference in log-likelihood values dominates the penalty term, and hence the probability of underpredicting q goes to zero and $T \to \infty$. Also, when $q > q_0$, T times the difference in log-likelihood values remains bounded in probability, but the magnitude of the penalty for lack of parsimony grows without bound as $T \to \infty$, therefore the probability of overestimating q goes to zero asymptotically also. Note that the fact that the asymptotic distribution of the likelihood ratio statistic is not χ^2 or that it may depend on nuisance parameters does not matter. What is important is that it is $O_p(1)$. Hence the second step produces a consistent estimator of q_0, and this completes the proof.

Remark 3 *The above proof is not exclusive to HQ and applies to any model selection criterion in which* $c_T \to \infty$ *and* $\frac{c_T}{T} \to 0$ *as* $T \to \infty$, *where* c_T *is the penalty for each additional parameter in the first stage of the procedure. The consistency of model selection criteria with this property for determining* p *in vector autoregressions has been established in Quinn (1980), and in autoregressions with unit roots in Paulsen (1984) and Tsay (1984). Consistency of such criteria for selection of cointegrating rank* q *and the lag order* p *has been established in Gonzalo and Pitarakis (1995) and Aznar and Salvador (2002). Consistency of PIC for selection of cointegrating rank* q *and the lag order* p *has been established in Chao and Phillips (1999). The contribution here is proving the consistency when* r *is added to the set of parameters to be estimated, and showing that this can be achieved with a two-step procedure.*

Remark 4 *As with all models selected with any consistent model selection criterion, the warning of Leeb and Potscher (2005) applies to models selected with our procedure as well in the sense that there is no guarantee that any inference made based on asymptotic distributions conditional on* p, q, r *selected by this procedure will necessarily be more accurate than that based on an unrestricted autoregression of order* $p_{\max}$.

Remark 5 *Let* $\tilde{\alpha}_1$ *be a full rank* $K \times (K - r_0)$ *matrix such that* $\tilde{\alpha}_1' [\begin{array}{cccc} \Gamma_1 & \Gamma_2 & \ldots & \Gamma_{p_0} \end{array}] = 0$. *Such a matrix exists because rank* $[\begin{array}{cccc} \Gamma_1 & \Gamma_2 & \ldots & \Gamma_{p_0} \end{array}] = r_0$ *but it is not unique. We can augment* $\tilde{\alpha}_1$ *with* r_0 *additional linearly independent vectors arranged as columns of matrix* $\tilde{\alpha}_2$ *to form a basis for* $\mathbb{R}^n$, *and to achieve uniqueness we can choose these matrices such that* $(\tilde{\alpha}_1 \vdots \tilde{\alpha}_2)' \Omega (\tilde{\alpha}_1 \vdots \tilde{\alpha}_2) = I_K$. *The DGP can be alternatively written as*

$$\begin{aligned} \tilde{\alpha}_1' \Delta y_t &= c_1 + \Pi_{(1)} y_{t-1} + \eta_{(1),t} \\ \tilde{\alpha}_2' \Delta y_t &= c_2 + \Pi_{(2)} y_{t-1} + \Gamma_{(2),1} \Delta y_{t-1} + \Gamma_{(2),2} \Delta y_{t-2} + \cdots + \Gamma_{(2),p_0} \Delta y_{t-p_0} + \eta_{(2),t} \end{aligned}$$

where for any vector or matrix X, $X_{(i)} = \tilde{\alpha}_i' X$, $i = 1, 2$. *While we have presented the model selection criteria as penalised log-likelihoods and have referred to maximum likelihood estimators and likelihood ratio tests in our proof to conform with the previous literature, all arguments could be phrased in the context of GMM estimation of the above structural model and test statistics for testing overidentifying restrictions in the first block of this structure Anderson and Vahid (1998). Therefore, there is no need for any assumption of normality at any stage.*

C Tables

Table 1: Performance of IC(p,r,q) in a $(1,1,2)$ design and its comparison with the usual application of the Johansen method

	T=100				T=200				T=400			
	$q<q_0$	$q=q_0$	$q>q_0$	Total	$q<q_0$	$q=q_0$	$q>q_0$	Total	$q<q_0$	$q=q_0$	$q>q_0$	Total
AIC												
$r<r_0$	0,0,0	2,0,0	4,0,0	6,0,0	0,0,0	1,0,0	1,0,0	2,0,0	0,0,0	0,0,0	0,0,0	1,0,0
$r=r_0$	0,0,0	0,22,9	0,31,13	0,54,23	0,0,0	0,26,7	0,37,10	0,63,17	0,0,0	0,29,6	0,38,10	0,67,15
$r>r_0$	0,0,0	0,5,3	0,7,4	0,11,6	0,0,0	0,6,2	0,7,3	0,13,5	0,0,0	0,5,2	0,8,2	0,13,4
Total	0,0,0	2,27,12	4,38,17	6,65,29	0,0,0	1,32,9	1,44,13	2,76,22	0,0,0	0,34,8	0,46,12	1,80,19
AIC+J	1,9,1	10,41,3	6,26,3	17,76,7	0,2,0	4,53,3	2,34,2	6,89,5	0,0,0	1,55,3	1,38,2	2,94,4
HQ												
$r<r_0$	0,0,0	10,0,0	8,0,0	19,0,0	0,0,0	5,0,0	3,0,0	8,0,0	0,0,0	2,0,0	1,0,0	3,0,0
$r=r_0$	0,3,0	0,39,3	0,30,3	0,72,6	0,1,0	0,52,2	0,33,1	0,86,3	0,0,0	0,62,1	0,31,1	0,94,2
$r>r_0$	0,0,0	0,2,0	0,1,0	0,3,0	0,0,0	0,2,0	0,1,0	0,2,0	0,0,0	0,1,0	0,1,0	0,2,0
Total	0,3,0	10,41,4	8,31,3	19,75,6	0,1,0	5,54,2	3,34,1	8,89,3	0,0,0	2,63,1	1,32,1	3,95,2
HQ+J	2,8,0	20,34,0	14,22,0	37,63,0	0,2,0	11,48,0	8,31,0	19,81,0	0,0,0	4,55,0	3,38,0	7,93,0
SC												
$r<r_0$	3,0,0	24,0,0	9,0,0	36,0,0	0,0,0	15,0,0	4,0,0	19,0,0	0,0,0	7,0,0	1,0,0	8,0,0
$r=r_0$	0,8,0	0,42,1	0,13,0	0,63,1	0,4,0	0,63,0	0,14,0	0,80,0	0,1,0	0,79,0	0,12,0	0,92,0
$r>r_0$	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0
Total	3,8,0	24,42,1	9,13,0	36,63,1	0,4,0	15,63,0	4,14,0	19,81,0	0,1,0	7,79,0	1,12,0	8,92,0
SC+J	4,5,0	31,23,0	24,14,0	58,42,0	0,2,0	22,36,0	16,23,0	38,62,0	0,0,0	11,47,0	9,33,0	20,80,0
PIC												
$r<r_0$	7,0,0	24,0,0	1,0,0	32,0,0	1,0,0	14,0,0	0,0,0	15,0,0	0,0,0	5,0,0	0,0,0	5,0,0
$r=r_0$	0,14,0	0,52,0	0,2,0	0,68,0	0,6,0	0,77,0	0,2,0	0,85,0	0,2,0	0,92,0	0,1,0	0,95,0
$r>r_0$	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0
Total	7,14,0	24,52,0	1,2,0	32,68,0	1,6,0	14,77,0	0,2,0	15,85,0	0,2,0	5,92,0	0,1,0	5,95,0
HQ-PIC												
$r<r_0$	4,0,0	14,0,0	1,0,0	19,0,0	0,0,0	8,0,0	0,0,0	8,0,0	0,0,0	3,0,0	0,0,0	3,0,0
$r=r_0$	0,15,1	0,54,5	0,2,0	0,72,6	0,6,0	0,79,3	0,2,0	0,86,3	0,2,0	0,90,2	0,1,0	0,93,2
$r>r_0$	0,1,0	0,3,0	0,0,0	0,3,0	0,0,0	0,2,0	0,0,0	0,2,0	0,0,0	0,2,0	0,0,0	0,2,0
Total	4,15,1	14,57,5	1,2,0	19,75,6	0,6,0	8,81,3	0,2,0	8,89,3	0,2,0	3,92,2	0,1,0	3,95,2

Note: The total of the three entries a,b,c in each cell show the percentage of times the selected model falls in the category identified by the column and row labels. Entry a shows the percentage where $p<p_0$, b shows the percentage where $p=p_0$ and c the percentage where $p>p_0$. The row labeled X+J shows this information for the method commonly used in practice, where the lag-length p is chosen by model selection criterion X, and then the Johansen procedure is used for determining q.

Table 2: Performance of IC(p, r, q) in a $(2, 1, 1)$ design and its comparison with the usual application of the Johansen method

	T=100				T=200				T=400			
	$q < q_0$	$q = q_0$	$q > q_0$	Total	$q < q_0$	$q = q_0$	$q > q_0$	Total	$q < q_0$	$q = q_0$	$q > q_0$	Total
AIC												
$r < r_0$	1,0,0	1,0,0	1,0,0	3,0,0	0,0,0	1,0,0	1,0,0	2,0,0	0,0,0	0,0,0	1,0,0	1,0,0
$r = r_0$	0,0,1	1,11,4	4,34,14	6,45,19	0,0,1	1,14,4	2,41,11	3,57,15	0,0,1	0,16,3	1,44,10	1,60,14
$r > r_0$	1,1,3	1,3,2	2,9,6	3,13,11	0,0,3	1,3,3	1,8,5	2,11,10	0,1,2	1,3,3	1,8,5	2,11,11
Total	2,1,4	3,14,6	7,43,20	12,58,30	0,0,4	3,17,7	4,49,16	7,68,25	0,1,3	1,19,6	3,52,15	4,71,25
AIC+J	2,8,1	23,43,4	6,11,2	32,62,6	0,1,0	11,68,4	4,2,13	13,82,5	0,0,0	4,74,4	1,15,1	4,91,5
HQ												
$r < r_0$	0,0,0	3,0,0	3,0,0	6,0,0	0,0,0	2,0,0	1,0,0	3,0,0	0,0,0	1,0,0	0,0,0	1,0,0
$r = r_0$	0,1,1	8,37,4	6,25,3	14,64,8	0,0,0	3,56,3	2,25,2	6,79,5	0,0,0	2,65,2	1,21,2	2,85,4
$r > r_0$	0,0,1	1,1,1	1,2,1	3,4,3	0,0,1	0,1,1	1,1,1	1,3,3	0,0,1	0,1,1	0,1,2	1,3,4
Total	0,1,2	12,39,5	10,27,4	23,67,10	0,1,1	6,57,3	4,27,3	10,82,8	0,0,1	3,66,3	1,22,4	4,88,8
HQ+J	3,5,0	47,25,0	14,6,0	64,36,0	0,1,0	32,51,0	7,9,0	39,61,0	0,0,0	12,71,0	3,15,0	15,85,0
SC												
$r < r_0$	2,0,0	10,0,0	3,0,0	15,0,0	0,0,0	7,0,0	1,0,0	8,0,0	0,0,0	3,0,0	0,0,0	4,0,0
$r = r_0$	2,8,0	21,42,1	3,6,0	26,55,2	0,3,0	12,71,1	1,4,0	13,77,1	0,0,0	8,86,0	0,2,1	4,90,1
$r > r_0$	0,0,0	0,0,0	0,0,0	1,0,0	0,0,0	0,0,0	0,0,0	1,0,0	0,0,0	0,0,0	0,0,0	0,1,0
Total	4,8,0	32,42,1	7,6,0	42,56,2	0,3,0	19,71,1	2,4,0	22,77,1	0,0,0	11,86,1	0,2,1	8,90,1
SC+J	5,2,0	62,7,0	22,2,0	89,11,0	0,0,0	55,26,0	14,5,0	69,31,0	0,0,0	34,48,0	8,10,0	41,59,0
PIC												
$r < r_0$	4,0,0	11,0,0	1,0,0	16,0,0	0,0,0	7,0,0	0,0,0	7,0,0	0,0,0	3,0,0	0,0,0	3,0,0
$r = r_0$	4,3,0	37,28,0	1,1,0	42,41,0	1,4,0	25,62,0	0,0,0	26,66,0	0,0,0	10,87,0	0,0,0	9,88,0
$r > r_0$	0,0,0	0,0,0	0,0,0	1,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0	0,0,0
Total	8,3,0	48,28,0	2,1,0	59,41,0	1,4,0	32,62,0	1,0,0	34,66,0	0,0,0	13,87,0	0,0,0	12,88,0
HQ-PIC												
$r < r_0$	1,0,0	5,0,0	0,0,0	6,0,0	0,0,0	3,0,0	0,0,0	3,0,0	0,0,0	1,0,0	0,0,0	1,0,0
$r = r_0$	2,13,3	12,49,4	0,1,0	14,63,7	0,4,1	5,77,4	0,0,0	6,79,5	0,0,0	2,86,4	0,0,0	2,86,4
$r > r_0$	1,2,2	2,2,1	0,0,0	2,4,3	0,0,1	1,2,2	0,0,0	1,3,3	0,0,0	1,2,4	0,0,0	1,2,4
Total	4,15,5	19,51,5	1,1,0	23,67,10	1,4,2	9,79,6	0,0,0	10,82,8	0,0,0	4,88,8	0,0,0	4,88,8

Note: The total of the three entries a, b, c in each cell show the percentage of times the selected model falls in the category identified by the column and row labels. Entry a shows the percentage where $p < p_0$, b shows the percentage where $p = p_0$ and c the percentage where $p > p_0$. The row labeled X+J shows this information for the method commonly used in practice, where the lag-length p is chosen by model selection criterion X, and then the Johansen procedure is used for determining q.

Table 3: Percentage improvement in forecast accuracy measures for possibly reduced rank models over unrestricted VARs in a (1,1,2) setting.

Horizon	T=100			T=200			T=400		
(h)	TMSFE	|MSFE|	GFESM	TMSFE	|MSFE|	GFESM	TMSFE	|MSFE|	GFESM
VECM(HQ-PIC) for all DGPs									
1	1.4 (44,46,10)[a]	3.8	3.8	1.4 (49,49,2)	4.0	4.0	0.9 (53,47,0)	2.7	2.7
4	0.7 (23,77,0)	1.6	3.7	0.7 (46,54,0)	2.4	10.2	0.3 (27,73,0)	1.1	6.3
8	0.7 (19,80,1)	1.8	-7.2	0.1 (5,91,4)	0.1	8.0	0.1 (4,96,0)	0.5	6.8
12	0.2 (3,93,4)	0.5	-19.4	0.4 (14,86,0)	0.9	7.8	0.1 (4,96,0)	0.2	6.6
16	0.2 (5,94,1)	0.6	-31.3	0.4 (18,82,0)	1.0	3.7	0.1 (4,95,1)	0.2	7.2
VECM(AIC+J) for all DGPs									
1	0.9 (28,63,9)	2.3	2.3	0.8 (30,67,3)	2.3	2.3	0.4 (27,71,2)	1.0	1.0
4	0.4 (14,86,0)	0.6	2.0	0.2 (13,86,1)	0.8	5.5	0.1 (8,92,0)	0.4	2.2
8	0.5 (21,78,1)	1.4	-5.5	0.0 (2,91,7)	-0.2	4.2	0.1 (2,98,0)	0.2	1.9
12	0.1 (5,92,3)	0.4	-12.5	0.2 (12,88,0)	0.5	4.1	0.0 (0,98,2)	-0.1	1.4
16	0.1 (5,92,3)	0.4	-20.4	0.3 (18,82,0)	0.7	1.5	0.0 (3,97,0)	0.0	1.8

VECM(HQ-PIC) are models selected by the model selection process proposed in Section 4.1 and estimated by the algorithm proposed in Section 3. VECM(AIC+J) are estimated by the usual Johansen procedure with AIC as the model selection criterion for the lag length.

[a] We perform Diebold and Mariano (1995) tests at the 5% level of significance for equal predictive accuracy between the reduced rank models and unrestricted VARs. For cell (x,y,z), y denotes the percentage of DGPs for which the Null of equal forecast accuracy is not rejected and entries x and z denote the percentage of DGPs for which the Null is rejected with a positive statistic (i.e., the reduced rank model is significantly more accurate than the unrestricted VAR) and a negative statistic (i.e., the reduced rank model is significantly less accurate than the unrestricted VAR) respectively.

Table 4: Percentage improvement in forecast accuracy measures for possibly reduced rank models over unrestricted VARs in a (2,1,1) setting.

Horizon	T=100			T=200			T=400		
(h)	TMSFE	|MSFE|	GFESM	TMSFE	|MSFE|	GFESM	TMSFE	|MSFE|	GFESM
VECM(HQ-PIC) for all DGPs									
1	7.8 (87,13,0)[a]	21.8	21.8	4.5 (90,10,0)	12.9	12.9	2.5 (95,5,0)	7.5	7.5
4	2.2 (69,31,0)	8.1	37.8	2.0 (78,22,0)	5.2	30.6	0.9 (47,53,0)	2.3	17.5
8	1.0 (24,76,0)	2.7	38.5	0.6 (22,78,0)	2.3	34.1	0.6 (32,68,0)	2.2	25.7
12	0.4 (12,87,1)	0.8	29.8	0.8 (27,73,0)	2.4	36.8	0.9 (82,18,0)	2.9	29.5
16	0.8 (16,84,0)	1.8	25.5	0.3 (16,59,25)	0.3	32.8	0.7 (39,61,0)	2.4	32.7
VECM(AIC+J) for all DGPs									
1	5.4 (81,19,0)	14.1	14.1	3.2 (81,19,0)	8.7	8.7	1.4 (72,28,0)	4.1	4.1
4	1.3 (29,71,0)	4.8	21.6	1.2 (61,39,0)	3.0	21.3	0.6 (35,65,0)	1.8	10.7
8	0.7 (15,85,0)	1.9	21.5	0.6 (23,77,0)	2.3	26.1	0.4 (14,86,0)	1.7	16.8
12	0.5 (11,89,0)	0.9	14.5	0.6 (19,81,0)	1.9	29.6	0.7 (65,35,0)	2.4	19.2
16	0.6 (13,87,0)	1.4	11.0	0.2 (16,84,0)	0.3	27.4	0.6 (38,62,0)	2.2	22.0

VECM(HQ-PIC) are models selected by the model selection process proposed in Section 4.1 and estimated by the algorithm proposed in Section 3. VECM(AIC+J) are estimated by the usual Johansen procedure with AIC as the model selection criterion for the lag length.

[a] Refer to note in Table 3.

Table 5: Percentage improvement in forecast accuracy measures for reduced ranked models and unrestricted VARs for Brazilian inflation.

Horizon (h)	VECM(HQ-PIC) versus VAR in levels			VECM(AIC+J) versus VAR in levels			VECM(HQ-PIC) versus VECM(AIC+J)		
	GFESM	|MFSE|	TMFSE	GFESM	|MFSE|	TMFSE	GFESM	|MFSE|	TMFSE
1	69.6 (*, *,-)	69.6	36.9	3.4 (-,-,-)	3.4	2.0	66.2 (-,**,-)	66.2	34.9
4	91.0 (**,**,**)	45.2	32.4	11.1 (**,-,-)	11.1	10.0	79.9 (**,**,**)	34.1	22.4
8	107.9 (**,**,-)	32.9	24.6	45.7 (**,**,-)	26.5	15.9	62.1 (-,*,-)	6.4	8.8
12	120.3 (**,*,-)	38.4	33.6	52.3 (**,**,**)	41.9	32.8	68.0 (-,-,-)	-3.6	0.8
16	142.7 (*,*,-)	40.2	36.4	81.9 (**,**,**)	42.9	39.8	60.8 (-,-,-)	-2.7	-3.4

VECM(HQ-PIC) is the model selected by the model selection process proposed in Section 4.1 and estimated by the algorithm proposed in Section 3. VECM(AIC+J) is the model estimated by the usual Johansen procedure with AIC as the model selection criterion for the lag length. See Section 7 for further details. The triplet $(\cdot, \cdot, \cdot)$ presents tests of equal variance for forecast errors predicting $\Delta \ln (CPI - IBGE_t)$, $\Delta \ln (CPI - FGV_t)$, and $\Delta \ln (CPI - FIPE_t)$ respectively. The symbols **, * and - denote, respectively, significance at the 5% level, at the 10% level, and not significant at the 10% level.

Table 6: Percentage improvement in forecast accuracy measures for reduced ranked models and unrestricted VARs for U.S. macroeconomic aggregates.

Horizon (h)	VECM(HQ-PIC) versus VAR in levels			VECM(AIC+J) versus VAR in levels			VECM(HQ-PIC) versus VECM(AIC+J)		
	GFESM	|MFSE|	TMFSE	GFESM	|MFSE|	TMFSE	GFESM	|MFSE|	TMFSE
1	60.4 (**,**,**)	60.4	35.1	39.5 (**,**,**)	39.5	16.7	20.9 (-,-,-)	20.9	18.4
4	134.7 (**,**,**)	83.5	56.3	90.0 (**,**,**)	50.2	35.1	44.7 (-,-,-)	33.2	21.1
8	169.2 (**,**,-)	25.3	8.4	121.0 (**,**,-)	17.7	7.3	48.2 (-,**,-)	7.6	1.1
12	176.3 (*,**,-)	20.0	1.5	119.0 (**,**,-)	9.1	0.2	57.3 (-,-,-)	11.0	1.3
16	147.3 (**,**,-)	26.0	3.6	79.7 (**,**,-)	16.0	3.2	67.7 (-,-,-)	10.0	0.3

VECM(HQ-PIC) is the model selected by the model selection process proposed in Section 4.1 and estimated by the algorithm proposed in Section 3. VECM(AIC+J) is the model estimated by the usual Johansen procedure with AIC as the model selection criterion for the lag length. See Section 7 for further details. The triplet $(\cdot, \cdot, \cdot)$ presents tests of equal variance for forecast errors predicting $\Delta \ln (y_t)$, $\Delta \ln (c_t)$, and $\Delta \ln (i_t)$ respectively. The symbols **, * and - denote, respectively, significance at the 5% level, at the 10% level, and not significant at the 10% level.

Banco Central do Brasil

Trabalhos para Discussão

Os Trabalhos para Discussão podem ser acessados na internet, no formato PDF,
no endereço: http://www.bc.gov.br

Working Paper Series

Working Papers in PDF format can be downloaded from: http://www.bc.gov.br

1	**Implementing Inflation Targeting in Brazil** *Joel Bogdanski, Alexandre Antonio Tombini and Sérgio Ribeiro da Costa Werlang*	Jul/2000
2	**Política Monetária e Supervisão do Sistema Financeiro Nacional no Banco Central do Brasil** *Eduardo Lundberg*	Jul/2000
	Monetary Policy and Banking Supervision Functions on the Central Bank *Eduardo Lundberg*	Jul/2000
3	**Private Sector Participation: a Theoretical Justification of the Brazilian Position** *Sérgio Ribeiro da Costa Werlang*	Jul/2000
4	**An Information Theory Approach to the Aggregation of Log-Linear Models** *Pedro H. Albuquerque*	Jul/2000
5	**The Pass-Through from Depreciation to Inflation: a Panel Study** *Ilan Goldfajn and Sérgio Ribeiro da Costa Werlang*	Jul/2000
6	**Optimal Interest Rate Rules in Inflation Targeting Frameworks** *José Alvaro Rodrigues Neto, Fabio Araújo and Marta Baltar J. Moreira*	Jul/2000
7	**Leading Indicators of Inflation for Brazil** *Marcelle Chauvet*	Sep/2000
8	**The Correlation Matrix of the Brazilian Central Bank's Standard Model for Interest Rate Market Risk** *José Alvaro Rodrigues Neto*	Sep/2000
9	**Estimating Exchange Market Pressure and Intervention Activity** *Emanuel-Werner Kohlscheen*	Nov/2000
10	**Análise do Financiamento Externo a uma Pequena Economia** **Aplicação da Teoria do Prêmio Monetário ao Caso Brasileiro: 1991–1998** *Carlos Hamilton Vasconcelos Araújo e Renato Galvão Flôres Júnior*	Mar/2001
11	**A Note on the Efficient Estimation of Inflation in Brazil** *Michael F. Bryan and Stephen G. Cecchetti*	Mar/2001
12	**A Test of Competition in Brazilian Banking** *Márcio I. Nakane*	Mar/2001

13	**Modelos de Previsão de Insolvência Bancária no Brasil** *Marcio Magalhães Janot*	Mar/2001
14	**Evaluating Core Inflation Measures for Brazil** *Francisco Marcos Rodrigues Figueiredo*	Mar/2001
15	**Is It Worth Tracking Dollar/Real Implied Volatility?** *Sandro Canesso de Andrade and Benjamin Miranda Tabak*	Mar/2001
16	**Avaliação das Projeções do Modelo Estrutural do Banco Central do Brasil para a Taxa de Variação do IPCA** *Sergio Afonso Lago Alves*	Mar/2001
	Evaluation of the Central Bank of Brazil Structural Model's Inflation Forecasts in an Inflation Targeting Framework *Sergio Afonso Lago Alves*	Jul/2001
17	**Estimando o Produto Potencial Brasileiro: uma Abordagem de Função de Produção** *Tito Nícias Teixeira da Silva Filho*	Abr/2001
	Estimating Brazilian Potential Output: a Production Function Approach *Tito Nícias Teixeira da Silva Filho*	Aug/2002
18	**A Simple Model for Inflation Targeting in Brazil** *Paulo Springer de Freitas and Marcelo Kfoury Muinhos*	Apr/2001
19	**Uncovered Interest Parity with Fundamentals: a Brazilian Exchange Rate Forecast Model** *Marcelo Kfoury Muinhos, Paulo Springer de Freitas and Fabio Araújo*	May/2001
20	**Credit Channel without the LM Curve** *Victorio Y. T. Chu and Márcio I. Nakane*	May/2001
21	**Os Impactos Econômicos da CPMF: Teoria e Evidência** *Pedro H. Albuquerque*	Jun/2001
22	**Decentralized Portfolio Management** *Paulo Coutinho and Benjamin Miranda Tabak*	Jun/2001
23	**Os Efeitos da CPMF sobre a Intermediação Financeira** *Sérgio Mikio Koyama e Márcio I. Nakane*	Jul/2001
24	**Inflation Targeting in Brazil: Shocks, Backward-Looking Prices, and IMF Conditionality** *Joel Bogdanski, Paulo Springer de Freitas, Ilan Goldfajn and Alexandre Antonio Tombini*	Aug/2001
25	**Inflation Targeting in Brazil: Reviewing Two Years of Monetary Policy 1999/00** *Pedro Fachada*	Aug/2001
26	**Inflation Targeting in an Open Financially Integrated Emerging Economy: the Case of Brazil** *Marcelo Kfoury Muinhos*	Aug/2001
27	**Complementaridade e Fungibilidade dos Fluxos de Capitais Internacionais** *Carlos Hamilton Vasconcelos Araújo e Renato Galvão Flôres Júnior*	Set/2001

28	**Regras Monetárias e Dinâmica Macroeconômica no Brasil: uma Abordagem de Expectativas Racionais** *Marco Antonio Bonomo e Ricardo D. Brito*	Nov/2001
29	**Using a Money Demand Model to Evaluate Monetary Policies in Brazil** *Pedro H. Albuquerque and Solange Gouvêa*	Nov/2001
30	**Testing the Expectations Hypothesis in the Brazilian Term Structure of Interest Rates** *Benjamin Miranda Tabak and Sandro Canesso de Andrade*	Nov/2001
31	**Algumas Considerações sobre a Sazonalidade no IPCA** *Francisco Marcos R. Figueiredo e Roberta Blass Staub*	Nov/2001
32	**Crises Cambiais e Ataques Especulativos no Brasil** *Mauro Costa Miranda*	Nov/2001
33	**Monetary Policy and Inflation in Brazil (1975-2000): a VAR Estimation** *André Minella*	Nov/2001
34	**Constrained Discretion and Collective Action Problems: Reflections on the Resolution of International Financial Crises** *Arminio Fraga and Daniel Luiz Gleizer*	Nov/2001
35	**Uma Definição Operacional de Estabilidade de Preços** *Tito Nícias Teixeira da Silva Filho*	Dez/2001
36	**Can Emerging Markets Float? Should They Inflation Target?** *Barry Eichengreen*	Feb/2002
37	**Monetary Policy in Brazil: Remarks on the Inflation Targeting Regime, Public Debt Management and Open Market Operations** *Luiz Fernando Figueiredo, Pedro Fachada and Sérgio Goldenstein*	Mar/2002
38	**Volatilidade Implícita e Antecipação de Eventos de *Stress*: um Teste para o Mercado Brasileiro** *Frederico Pechir Gomes*	Mar/2002
39	**Opções sobre Dólar Comercial e Expectativas a Respeito do Comportamento da Taxa de Câmbio** *Paulo Castor de Castro*	Mar/2002
40	**Speculative Attacks on Debts, Dollarization and Optimum Currency Areas** *Aloisio Araujo and Márcia Leon*	Apr/2002
41	**Mudanças de Regime no Câmbio Brasileiro** *Carlos Hamilton V. Araújo e Getúlio B. da Silveira Filho*	Jun/2002
42	**Modelo Estrutural com Setor Externo: Endogenização do Prêmio de Risco e do Câmbio** *Marcelo Kfoury Muinhos, Sérgio Afonso Lago Alves e Gil Riella*	Jun/2002
43	**The Effects of the Brazilian ADRs Program on Domestic Market Efficiency** *Benjamin Miranda Tabak and Eduardo José Araújo Lima*	Jun/2002

44	**Estrutura Competitiva, Produtividade Industrial e Liberação Comercial no Brasil** *Pedro Cavalcanti Ferreira e Osmani Teixeira de Carvalho Guillén*	Jun/2002
45	**Optimal Monetary Policy, Gains from Commitment, and Inflation Persistence** *André Minella*	Aug/2002
46	**The Determinants of Bank Interest Spread in Brazil** *Tarsila Segalla Afanasieff, Priscilla Maria Villa Lhacer and Márcio I. Nakane*	Aug/2002
47	**Indicadores Derivados de Agregados Monetários** *Fernando de Aquino Fonseca Neto e José Albuquerque Júnior*	Set/2002
48	**Should Government Smooth Exchange Rate Risk?** *Ilan Goldfajn and Marcos Antonio Silveira*	Sep/2002
49	**Desenvolvimento do Sistema Financeiro e Crescimento Econômico no Brasil: Evidências de Causalidade** *Orlando Carneiro de Matos*	Set/2002
50	**Macroeconomic Coordination and Inflation Targeting in a Two-Country Model** *Eui Jung Chang, Marcelo Kfoury Muinhos and Joanílio Rodolpho Teixeira*	Sep/2002
51	**Credit Channel with Sovereign Credit Risk: an Empirical Test** *Victorio Yi Tson Chu*	Sep/2002
52	**Generalized Hyperbolic Distributions and Brazilian Data** *José Fajardo and Aquiles Farias*	Sep/2002
53	**Inflation Targeting in Brazil: Lessons and Challenges** *André Minella, Paulo Springer de Freitas, Ilan Goldfajn and Marcelo Kfoury Muinhos*	Nov/2002
54	**Stock Returns and Volatility** *Benjamin Miranda Tabak and Solange Maria Guerra*	Nov/2002
55	**Componentes de Curto e Longo Prazo das Taxas de Juros no Brasil** *Carlos Hamilton Vasconcelos Araújo e Osmani Teixeira de Carvalho de Guillén*	Nov/2002
56	**Causality and Cointegration in Stock Markets: the Case of Latin America** *Benjamin Miranda Tabak and Eduardo José Araújo Lima*	Dec/2002
57	**As Leis de Falência: uma Abordagem Econômica** *Aloisio Araujo*	Dez/2002
58	**The Random Walk Hypothesis and the Behavior of Foreign Capital Portfolio Flows: the Brazilian Stock Market Case** *Benjamin Miranda Tabak*	Dec/2002
59	**Os Preços Administrados e a Inflação no Brasil** *Francisco Marcos R. Figueiredo e Thaís Porto Ferreira*	Dez/2002
60	**Delegated Portfolio Management** *Paulo Coutinho and Benjamin Miranda Tabak*	Dec/2002

61	**O Uso de Dados de Alta Freqüência na Estimação da Volatilidade e do Valor em Risco para o Ibovespa** *João Maurício de Souza Moreira e Eduardo Facó Lemgruber*	Dez/2002
62	**Taxa de Juros e Concentração Bancária no Brasil** *Eduardo Kiyoshi Tonooka e Sérgio Mikio Koyama*	Fev/2003
63	**Optimal Monetary Rules: the Case of Brazil** *Charles Lima de Almeida, Marco Aurélio Peres, Geraldo da Silva e Souza and Benjamin Miranda Tabak*	Feb/2003
64	**Medium-Size Macroeconomic Model for the Brazilian Economy** *Marcelo Kfoury Muinhos and Sergio Afonso Lago Alves*	Feb/2003
65	**On the Information Content of Oil Future Prices** *Benjamin Miranda Tabak*	Feb/2003
66	**A Taxa de Juros de Equilíbrio: uma Abordagem Múltipla** *Pedro Calhman de Miranda e Marcelo Kfoury Muinhos*	Fev/2003
67	**Avaliação de Métodos de Cálculo de Exigência de Capital para Risco de Mercado de Carteiras de Ações no Brasil** *Gustavo S. Araújo, João Maurício S. Moreira e Ricardo S. Maia Clemente*	Fev/2003
68	**Real Balances in the Utility Function: Evidence for Brazil** *Leonardo Soriano de Alencar and Márcio I. Nakane*	Feb/2003
69	**r-filters: a Hodrick-Prescott Filter Generalization** *Fabio Araújo, Marta Baltar Moreira Areosa and José Alvaro Rodrigues Neto*	Feb/2003
70	**Monetary Policy Surprises and the Brazilian Term Structure of Interest Rates** *Benjamin Miranda Tabak*	Feb/2003
71	**On Shadow-Prices of Banks in Real-Time Gross Settlement Systems** *Rodrigo Penaloza*	Apr/2003
72	**O Prêmio pela Maturidade na Estrutura a Termo das Taxas de Juros Brasileiras** *Ricardo Dias de Oliveira Brito, Angelo J. Mont'Alverne Duarte e Osmani Teixeira de C. Guillen*	Maio/2003
73	**Análise de Componentes Principais de Dados Funcionais – uma Aplicação às Estruturas a Termo de Taxas de Juros** *Getúlio Borges da Silveira e Octavio Bessada*	Maio/2003
74	**Aplicação do Modelo de Black, Derman & Toy à Precificação de Opções Sobre Títulos de Renda Fixa** *Octavio Manuel Bessada Lion, Carlos Alberto Nunes Cosenza e César das Neves*	Maio/2003
75	**Brazil's Financial System: Resilience to Shocks, no Currency Substitution, but Struggling to Promote Growth** *Ilan Goldfajn, Katherine Hennings and Helio Mori*	Jun/2003

76	**Inflation Targeting in Emerging Market Economies** *Arminio Fraga, Ilan Goldfajn and André Minella*	Jun/2003
77	**Inflation Targeting in Brazil: Constructing Credibility under Exchange Rate Volatility** *André Minella, Paulo Springer de Freitas, Ilan Goldfajn and Marcelo Kfoury Muinhos*	Jul/2003
78	**Contornando os Pressupostos de Black & Scholes: Aplicação do Modelo de Precificação de Opções de Duan no Mercado Brasileiro** *Gustavo Silva Araújo, Claudio Henrique da Silveira Barbedo, Antonio Carlos Figueiredo, Eduardo Facó Lemgruber*	Out/2003
79	**Inclusão do Decaimento Temporal na Metodologia Delta-Gama para o Cálculo do VaR de Carteiras Compradas em Opções no Brasil** *Claudio Henrique da Silveira Barbedo, Gustavo Silva Araújo, Eduardo Facó Lemgruber*	Out/2003
80	**Diferenças e Semelhanças entre Países da América Latina: uma Análise de *Markov Switching* para os Ciclos Econômicos de Brasil e Argentina** *Arnildo da Silva Correa*	Out/2003
81	**Bank Competition, Agency Costs and the Performance of the Monetary Policy** *Leonardo Soriano de Alencar and Márcio I. Nakane*	Jan/2004
82	**Carteiras de Opções: Avaliação de Metodologias de Exigência de Capital no Mercado Brasileiro** *Cláudio Henrique da Silveira Barbedo e Gustavo Silva Araújo*	Mar/2004
83	**Does Inflation Targeting Reduce Inflation? An Analysis for the OECD Industrial Countries** *Thomas Y. Wu*	May/2004
84	**Speculative Attacks on Debts and Optimum Currency Area: a Welfare Analysis** *Aloisio Araujo and Marcia Leon*	May/2004
85	**Risk Premia for Emerging Markets Bonds: Evidence from Brazilian Government Debt, 1996-2002** *André Soares Loureiro and Fernando de Holanda Barbosa*	May/2004
86	**Identificação do Fator Estocástico de Descontos e Algumas Implicações sobre Testes de Modelos de Consumo** *Fabio Araujo e João Victor Issler*	Maio/2004
87	**Mercado de Crédito: uma Análise Econométrica dos Volumes de Crédito Total e Habitacional no Brasil** *Ana Carla Abrão Costa*	Dez/2004
88	**Ciclos Internacionais de Negócios: uma Análise de Mudança de Regime Markoviano para Brasil, Argentina e Estados Unidos** *Arnildo da Silva Correa e Ronald Otto Hillbrecht*	Dez/2004
89	**O Mercado de *Hedge* Cambial no Brasil: Reação das Instituições Financeiras a Intervenções do Banco Central** *Fernando N. de Oliveira*	Dez/2004

90	**Bank Privatization and Productivity: Evidence for Brazil** *Márcio I. Nakane and Daniela B. Weintraub*	Dec/2004
91	**Credit Risk Measurement and the Regulation of Bank Capital and Provision Requirements in Brazil – a Corporate Analysis** *Ricardo Schechtman, Valéria Salomão Garcia, Sergio Mikio Koyama and Guilherme Cronemberger Parente*	Dec/2004
92	**Steady-State Analysis of an Open Economy General Equilibrium Model for Brazil** *Mirta Noemi Sataka Bugarin, Roberto de Goes Ellery Jr., Victor Gomes Silva, Marcelo Kfoury Muinhos*	Apr/2005
93	**Avaliação de Modelos de Cálculo de Exigência de Capital para Risco Cambial** *Claudio H. da S. Barbedo, Gustavo S. Araújo, João Maurício S. Moreira e Ricardo S. Maia Clemente*	Abr/2005
94	**Simulação Histórica Filtrada: Incorporação da Volatilidade ao Modelo Histórico de Cálculo de Risco para Ativos Não-Lineares** *Claudio Henrique da Silveira Barbedo, Gustavo Silva Araújo e Eduardo Facó Lemgruber*	Abr/2005
95	**Comment on Market Discipline and Monetary Policy by Carl Walsh** *Maurício S. Bugarin and Fábia A. de Carvalho*	Apr/2005
96	**O que É Estratégia: uma Abordagem Multiparadigmática para a Disciplina** *Anthero de Moraes Meirelles*	Ago/2005
97	**Finance and the Business Cycle: a Kalman Filter Approach with Markov Switching** *Ryan A. Compton and Jose Ricardo da Costa e Silva*	Aug/2005
98	**Capital Flows Cycle: Stylized Facts and Empirical Evidences for Emerging Market Economies** *Helio Mori e Marcelo Kfoury Muinhos*	Aug/2005
99	**Adequação das Medidas de Valor em Risco na Formulação da Exigência de Capital para Estratégias de Opções no Mercado Brasileiro** *Gustavo Silva Araújo, Claudio Henrique da Silveira Barbedo,e Eduardo Facó Lemgruber*	Set/2005
100	**Targets and Inflation Dynamics** *Sergio A. L. Alves and Waldyr D. Areosa*	Oct/2005
101	**Comparing Equilibrium Real Interest Rates: Different Approaches to Measure Brazilian Rates** *Marcelo Kfoury Muinhos and Márcio I. Nakane*	Mar/2006
102	**Judicial Risk and Credit Market Performance: Micro Evidence from Brazilian Payroll Loans** *Ana Carla A. Costa and João M. P. de Mello*	Apr/2006
103	**The Effect of Adverse Supply Shocks on Monetary Policy and Output** *Maria da Glória D. S. Araújo, Mirta Bugarin, Marcelo Kfoury Muinhos and Jose Ricardo C. Silva*	Apr/2006

104	**Extração de Informação de Opções Cambiais no Brasil** *Eui Jung Chang e Benjamin Miranda Tabak*	Abr/2006
105	**Representing Roommate's Preferences with Symmetric Utilities** *José Alvaro Rodrigues Neto*	Apr/2006
106	**Testing Nonlinearities Between Brazilian Exchange Rates and Inflation Volatilities** *Cristiane R. Albuquerque and Marcelo Portugal*	May/2006
107	**Demand for Bank Services and Market Power in Brazilian Banking** *Márcio I. Nakane, Leonardo S. Alencar and Fabio Kanczuk*	Jun/2006
108	**O Efeito da Consignação em Folha nas Taxas de Juros dos Empréstimos Pessoais** *Eduardo A. S. Rodrigues, Victorio Chu, Leonardo S. Alencar e Tony Takeda*	Jun/2006
109	**The Recent Brazilian Disinflation Process and Costs** *Alexandre A. Tombini and Sergio A. Lago Alves*	Jun/2006
110	**Fatores de Risco e o *Spread* Bancário no Brasil** *Fernando G. Bignotto e Eduardo Augusto de Souza Rodrigues*	Jul/2006
111	**Avaliação de Modelos de Exigência de Capital para Risco de Mercado do Cupom Cambial** *Alan Cosme Rodrigues da Silva, João Maurício de Souza Moreira e Myrian Beatriz Eiras das Neves*	Jul/2006
112	**Interdependence and Contagion: an Analysis of Information Transmission in Latin America's Stock Markets** *Angelo Marsiglia Fasolo*	Jul/2006
113	**Investigação da Memória de Longo Prazo da Taxa de Câmbio no Brasil** *Sergio Rubens Stancato de Souza, Benjamin Miranda Tabak e Daniel O. Cajueiro*	Ago/2006
114	**The Inequality Channel of Monetary Transmission** *Marta Areosa and Waldyr Areosa*	Aug/2006
115	**Myopic Loss Aversion and House-Money Effect Overseas: an Experimental Approach** *José L. B. Fernandes, Juan Ignacio Peña and Benjamin M. Tabak*	Sep/2006
116	**Out-Of-The-Money Monte Carlo Simulation Option Pricing: the Join Use of Importance Sampling and Descriptive Sampling** *Jaqueline Terra Moura Marins, Eduardo Saliby and Joséte Florencio dos Santos*	Sep/2006
117	**An Analysis of Off-Site Supervision of Banks' Profitability, Risk and Capital Adequacy: a Portfolio Simulation Approach Applied to Brazilian Banks** *Theodore M. Barnhill, Marcos R. Souto and Benjamin M. Tabak*	Sep/2006
118	**Contagion, Bankruptcy and Social Welfare Analysis in a Financial Economy with Risk Regulation Constraint** *Aloísio P. Araújo and José Valentim M. Vicente*	Oct/2006

119	**A Central de Risco de Crédito no Brasil: uma Análise de Utilidade de Informação** *Ricardo Schechtman*	Out/2006
120	**Forecasting Interest Rates: an Application for Brazil** *Eduardo J. A. Lima, Felipe Luduvice and Benjamin M. Tabak*	Oct/2006
121	**The Role of Consumer’s Risk Aversion on Price Rigidity** *Sergio A. Lago Alves and Mirta N. S. Bugarin*	Nov/2006
122	**Nonlinear Mechanisms of the Exchange Rate Pass-Through: a Phillips Curve Model With Threshold for Brazil** *Arnildo da Silva Correa and André Minella*	Nov/2006
123	**A Neoclassical Analysis of the Brazilian “Lost-Decades”** *Flávia Mourão Graminho*	Nov/2006
124	**The Dynamic Relations between Stock Prices and Exchange Rates: Evidence for Brazil** *Benjamin M. Tabak*	Nov/2006
125	**Herding Behavior by Equity Foreign Investors on Emerging Markets** *Barbara Alemanni and José Renato Haas Ornelas*	Dec/2006
126	**Risk Premium: Insights over the Threshold** *José L. B. Fernandes, Augusto Hasman and Juan Ignacio Peña*	Dec/2006
127	**Uma Investigação Baseada em Reamostragem sobre Requerimentos de Capital para Risco de Crédito no Brasil** *Ricardo Schechtman*	Dec/2006
128	**Term Structure Movements Implicit in Option Prices** *Caio Ibsen R. Almeida and José Valentim M. Vicente*	Dec/2006
129	**Brazil: Taming Inflation Expectations** *Afonso S. Bevilaqua, Mário Mesquita and André Minella*	Jan/2007
130	**The Role of Banks in the Brazilian Interbank Market: Does Bank Type Matter?** *Daniel O. Cajueiro and Benjamin M. Tabak*	Jan/2007
131	**Long-Range Dependence in Exchange Rates: the Case of the European Monetary System** *Sergio Rubens Stancato de Souza, Benjamin M. Tabak and Daniel O. Cajueiro*	Mar/2007
132	**Credit Risk Monte Carlo Simulation Using Simplified Creditmetrics’ Model: the Joint Use of Importance Sampling and Descriptive Sampling** *Jaqueline Terra Moura Marins and Eduardo Saliby*	Mar/2007
133	**A New Proposal for Collection and Generation of Information on Financial Institutions’ Risk: the Case of Derivatives** *Gilneu F. A. Vivan and Benjamin M. Tabak*	Mar/2007
134	**Amostragem Descritiva no Apreçamento de Opções Européias através de Simulação Monte Carlo: o Efeito da Dimensionalidade e da Probabilidade de Exercício no Ganho de Precisão** *Eduardo Saliby, Sergio Luiz Medeiros Proença de Gouvêa e Jaqueline Terra Moura Marins*	Abr/2007

135	**Evaluation of Default Risk for the Brazilian Banking Sector** *Marcelo Y. Takami and Benjamin M. Tabak*	May/2007
136	**Identifying Volatility Risk Premium from Fixed Income Asian Options** *Caio Ibsen R. Almeida and José Valentim M. Vicente*	May/2007
137	**Monetary Policy Design under Competing Models of Inflation Persistence** *Solange Gouvea e Abhijit Sen Gupta*	May/2007
138	**Forecasting Exchange Rate Density Using Parametric Models: the Case of Brazil** *Marcos M. Abe, Eui J. Chang and Benjamin M. Tabak*	May/2007
139	**Selection of Optimal Lag Length inCointegrated VAR Models with Weak Form of Common Cyclical Features** *Carlos Enrique Carrasco Gutiérrez, Reinaldo Castro Souza and Osmani Teixeira de Carvalho Guillén*	Jun/2007
140	**Inflation Targeting, Credibility and Confidence Crises** *Rafael Santos and Aloísio Araújo*	Aug/2007
141	**Forecasting Bonds Yields in the Brazilian Fixed income Market** *Jose Vicente and Benjamin M. Tabak*	Aug/2007
142	**Crises Análise da Coerência de Medidas de Risco no Mercado Brasileiro de Ações e Desenvolvimento de uma Metodologia Híbrida para o *Expected Shortfall*** *Alan Cosme Rodrigues da Silva, Eduardo Facó Lemgruber, José Alberto Rebello Baranowski e Renato da Silva Carvalho*	Ago/2007
143	**Price Rigidity in Brazil: Evidence from CPI Micro Data** *Solange Gouvea*	Sep/2007
144	**The Effect of Bid-Ask Prices on Brazilian Options Implied Volatility: a Case Study of Telemar Call Options** *Claudio Henrique da Silveira Barbedo and Eduardo Facó Lemgruber*	Oct/2007
145	**The Stability-Concentration Relationship in the Brazilian Banking System** *Benjamin Miranda Tabak, Solange Maria Guerra, Eduardo José Araújo Lima and Eui Jung Chang*	Oct/2007
146	**Movimentos da Estrutura a Termo e Critérios de Minimização do Erro de Previsão em um Modelo Paramétrico Exponencial** *Caio Almeida, Romeu Gomes, André Leite e José Vicente*	Out/2007
147	**Explaining Bank Failures in Brazil: Micro, Macro and Contagion Effects (1994-1998)** *Adriana Soares Sales and Maria Eduarda Tannuri-Pianto*	Oct/2007
148	**Um Modelo de Fatores Latentes com Variáveis Macroeconômicas para a Curva de Cupom Cambial** *Felipe Pinheiro, Caio Almeida e José Vicente*	Out/2007
149	**Joint Validation of Credit Rating PDs under Default Correlation** *Ricardo Schechtman*	Oct/2007

150	**A Probabilistic Approach for Assessing the Significance of Contextual Variables in Nonparametric Frontier Models: an Application for Brazilian Banks** *Roberta Blass Staub and Geraldo da Silva e Souza*	Oct/2007
151	**Building Confidence Intervals with Block Bootstraps for the Variance Ratio Test of Predictability** *Eduardo José Araújo Lima and Benjamin Miranda Tabak*	Nov/2007
152	**Demand for Foreign Exchange Derivatives in Brazil: Hedge or Speculation?** *Fernando N. de Oliveira and Walter Novaes*	Dec/2007
153	**Aplicação da Amostragem por Importância à Simulação de Opções Asiáticas Fora do Dinheiro** *Jaqueline Terra Moura Marins*	Dez/2007
154	**Identification of Monetary Policy Shocks in the Brazilian Market for Bank Reserves** *Adriana Soares Sales and Maria Tannuri-Pianto*	Dec/2007
155	**Does Curvature Enhance Forecasting?** *Caio Almeida, Romeu Gomes, André Leite and José Vicente*	Dec/2007
156	**Escolha do Banco e Demanda por Empréstimos: um Modelo de Decisão em Duas Etapas Aplicado para o Brasil** *Sérgio Mikio Koyama e Márcio I. Nakane*	Dez/2007
157	**Is the Investment-Uncertainty Link Really Elusive? The Harmful Effects of Inflation Uncertainty in Brazil** *Tito Nícias Teixeira da Silva Filho*	Jan/2008
158	**Characterizing the Brazilian Term Structure of Interest Rates** *Osmani T. Guillen and Benjamin M. Tabak*	Feb/2008
159	**Behavior and Effects of Equity Foreign Investors on Emerging Markets** *Barbara Alemanni and José Renato Haas Ornelas*	Feb/2008
160	**The Incidence of Reserve Requirements in Brazil: Do Bank Stockholders Share the Burden?** *Fábia A. de Carvalho and Cyntia F. Azevedo*	Feb/2008
161	**Evaluating Value-at-Risk Models via Quantile Regressions** *Wagner P. Gaglianone, Luiz Renato Lima and Oliver Linton*	Feb/2008
162	**Balance Sheet Effects in Currency Crises: Evidence from Brazil** *Marcio M. Janot, Márcio G. P. Garcia and Walter Novaes*	Apr/2008
163	**Searching for the Natural Rate of Unemployment in a Large Relative Price Shocks' Economy: the Brazilian Case** *Tito Nícias Teixeira da Silva Filho*	May/2008
164	**Foreign Banks' Entry and Departure: the recent Brazilian experience (1996-2006**) *Pedro Fachada*	Jun/2008
165	**Avaliação de Opções de Troca e Opções de *Spread* Européias e Americanas** *Giuliano Carrozza Uzêda Iorio de Souza, Carlos Patrício Samanez e Gustavo Santos Raposo*	Jul/2008

166	**Testing Hyperinflation Theories Using the Inflation Tax Curve: a case study** *Fernando de Holanda Barbosa and Tito Nícias Teixeira da Silva Filho*	Jul/2008
167	**O Poder Discriminante das Operações de Crédito das Instituições Financeiras Brasileiras** *Clodoaldo Aparecido Annibal*	Jul/2008
168	**An Integrated Model for Liquidity Management and Short-Term Asset Allocation in Commercial Banks** *Wenersamy Ramos de Alcântara*	Jul/2008
169	**Mensuração do Risco Sistêmico no Setor Bancário com Variáveis Contábeis e Econômicas** *Lucio Rodrigues Capelletto, Eliseu Martins e Luiz João Corrar*	Jul/2008
170	**Política de Fechamento de Bancos com Regulador Não-Benevolente: Resumo e Aplicação** *Adriana Soares Sales*	Jul/2008
171	**Modelos para a Utilização das Operações de Redesconto pelos Bancos com Carteira Comercial no Brasil** *Sérgio Mikio Koyama e Márcio Issao Nakane*	Ago/2008
172	**Combining Hodrick-Prescott Filtering with a Production Function Approach to Estimate Output Gap** *Marta Areosa*	Aug/2008
173	**Exchange Rate Dynamics and the Relationship between the Random Walk Hypothesis and Official Interventions** *Eduardo José Araújo Lima and Benjamin Miranda Tabak*	Aug/2008
174	**Foreign Exchange Market Volatility Information: an investigation of real-dollar exchange rate** *Frederico Pechir Gomes, Marcelo Yoshio Takami and Vinicius Ratton Brandi*	Aug/2008
175	**Evaluating Asset Pricing Models in a Fama-French Framework** *Carlos Enrique Carrasco Gutierrez and Wagner Piazza Gaglianone*	Dec/2008
176	**Fiat Money and the Value of Binding Portfolio Constraints** *Mário R. Páscoa, Myrian Petrassi and Juan Pablo Torres-Martínez*	Dec/2008
177	**Preference for Flexibility and Bayesian Updating** *Gil Riella*	Dec/2008
178	**An Econometric Contribution to the Intertemporal Approach of the Current Account** *Wagner Piazza Gaglianone and João Victor Issler*	Dec/2008
179	**Are Interest Rate Options Important for the Assessment of Interest Rate Risk?** *Caio Almeida and José Vicente*	Dec/2008
180	**A Class of Incomplete and Ambiguity Averse Preferences** *Leandro Nascimento and Gil Riella*	Dec/2008
181	**Monetary Channels in Brazil through the Lens of a Semi-Structural Model** *André Minella and Nelson F. Souza-Sobrinho*	Apr/2009

182	**Avaliação de Opções Americanas com Barreiras Monitoradas de Forma Discreta** *Giuliano Carrozza Uzêda Iorio de Souza e Carlos Patrício Samanez*	Abr/2009
183	**Ganhos da Globalização do Capital Acionário em Crises Cambiais** *Marcio Janot e Walter Novaes*	Abr/2009
184	**Behavior Finance and Estimation Risk in Stochastic Portfolio Optimization** *José Luiz Barros Fernandes, Juan Ignacio Peña and Benjamin Miranda Tabak*	Apr/2009
185	**Market Forecasts in Brazil: performance and determinants** *Fabia A. de Carvalho and André Minella*	Apr/2009
186	**Previsão da Curva de Juros: um modelo estatístico com variáveis macroeconômicas** *André Luís Leite, Romeu Braz Pereira Gomes Filho e José Valentim Machado Vicente*	Maio/2009
187	**The Influence of Collateral on Capital Requirements in the Brazilian Financial System: an approach through historical average and logistic regression on probability of default** *Alan Cosme Rodrigues da Silva, Antônio Carlos Magalhães da Silva, Jaqueline Terra Moura Marins, Myrian Beatriz Eiras da Neves and Giovani Antonio Silva Brito*	Jun/2009
188	**Pricing Asian Interest Rate Options with a Three-Factor HJM Model** *Claudio Henrique da Silveira Barbedo, José Valentim Machado Vicente and Octávio Manuel Bessada Lion*	Jun/2009
189	**Linking Financial and Macroeconomic Factors to Credit Risk Indicators of Brazilian Banks** *Marcos Souto, Benjamin M. Tabak and Francisco Vazquez*	Jul/2009
190	**Concentração Bancária, Lucratividade e Risco Sistêmico: uma abordagem de contágio indireto** *Bruno Silva Martins e Leonardo S. Alencar*	Set/2009
191	**Concentração e Inadimplência nas Carteiras de Empréstimos dos Bancos Brasileiros** *Patricia L. Tecles, Benjamin M. Tabak e Roberta B. Staub*	Set/2009
192	**Inadimplência do Setor Bancário Brasileiro: uma avaliação de suas medidas** *Clodoaldo Aparecido Annibal*	Set/2009
193	***Loss Given Default*: um estudo sobre perdas em operações prefixadas no mercado brasileiro** *Antonio Carlos Magalhães da Silva, Jaqueline Terra Moura Marins e Myrian Beatriz Eiras das Neves*	Set/2009
194	**Testes de Contágio entre Sistemas Bancários – A crise do *subprime*** *Benjamin M. Tabak e Manuela M. de Souza*	Set/2009
195	**From Default Rates to Default Matrices: a complete measurement of Brazilian banks' consumer credit delinquency** *Ricardo Schechtman*	Oct/2009

196	**The role of macroeconomic variables in sovereign risk** *Marco S. Matsumura and José Valentim Vicente*	Oct/2009
197	**Forecasting the Yield Curve for Brazil** *Daniel O. Cajueiro, Jose A. Divino and Benjamin M. Tabak*	Nov/2009
198	**Impacto dos *Swaps* Cambiais na Curva de Cupom Cambial: uma análise segundo a regressão de componentes principais** *Alessandra Pasqualina Viola, Margarida Sarmiento Gutierrez, Octávio Bessada Lion e Cláudio Henrique Barbedo*	Nov/2009
199	**Delegated Portfolio Management and Risk Taking Behavior** *José Luiz Barros Fernandes, Juan Ignacio Peña and Benjamin Miranda Tabak*	Dec/2009
200	**Evolution of Bank Efficiency in Brazil: A DEA Approach** *Roberta B. Staub, Geraldo Souza and Benjamin M. Tabak*	Dec/2009
201	**Efeitos da Globalização na Inflação Brasileira** *Rafael Santos e Márcia S. Leon*	Jan/2010
202	**Considerações sobre a Atuação do Banco Central na Crise de 2008** *Mário Mesquita e Mario Torós*	Mar/2010
203	**Hiato do Produto e PIB no Brasil: uma Análise de Dados em Tempo Real** *Rafael Tiecher Cusinato, André Minella e Sabino da Silva Pôrto Júnior*	Abr/2010
204	**Fiscal and monetary policy interaction: a simulation based analysis of a two-country New Keynesian DSGE model with heterogeneous households** *Marcos Valli and Fabia A. de Carvalho*	Apr/2010