ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Co-integration rank determination in partial systems using information criteria

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Giuseppe Cavaliere, Luca De Angelis, Luca Fanelli (2018). Co-integration rank determination in partial systems using information criteria. OXFORD BULLETIN OF ECONOMICS AND STATISTICS, 80(1), 65-89 [10.1111/obes.12195].

Availability:
This version is available at: https://hdl.handle.net/11585/601120 since: 2020-02-29

Published:
DOI: http://doi.org/10.1111/obes.12195

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

07 October 2026

This is the peer reviewed version published in OXFORD BULLETIN OF ECONOMICS AND STATISTICS: *Co-integration rank determination in partial systems using information criteria,* Giuseppe Cavaliere, Luca De Angelis and Luca Fanelli, 2017.

The final form was published in https://doi.org/10.1111/obes.12195

This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

*This item was downloaded from IRIS Università di Bologna (*https://cris.unibo.it/*)*

When citing, please refer to the published version.

Co-integration rank determination in partial systems using information criteria*

28th March 2017

Abstract

We investigate the asymptotic and finite sample properties of the most widely used information criteria for co-integration rank determination in 'partial' systems, i.e. in co-integrated Vector Autoregressive (VAR) models where a sub-set of variables of interest is modelled conditional on another sub-set of variables. The asymptotic properties of the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC) and the Hannan-Quinn Information Criterion (HQC) are established, and consistency of BIC and HQC is proved. Notably, the consistency of BIC and HQC is robust to violations of weak exogeneity of the conditioning variables with respect to the co-integration parameters. More precisely, BIC and HQC recover the true co-integration rank from the partial system analysis also when the conditional model does not convey all information about the co-integration parameters. This result opens up interesting possibilities for practitioners who can now determine the co-integration rank in partial systems without being concerned about the weak exogeneity of the conditioning variables. A Monte Carlo experiment based on a large dimensional data generating process shows that BIC and HQC applied in partial systems perform reasonably well in small samples and comparatively better

*After the first draft of this paper was written, Prof. Jean-Pierre Urbain passed away. Jean-Pierre was a leading expert of weak exogeneity in co-integrated models and this work has largely benefited from his contribution to the field. We dedicate this paper to his memory. We are grateful to the Editor, Anindya Banerjee, and two anonymous referees for their helpful comments on the earlier version of this paper. We thank the conference participants at the 'Workshop on Time Series Econometrics' at the Department of Economic Sciences, University of Salento, Lecce, Italy (June 2016), the 'Econometric Society Meeting', Geneve, Switzerland (August 2016), and the 'Latin American Econometric Society Annual Meeting', Medellín, Colombia (November 2016). We also thank Anders Rahbek, Peter Boswijk and Alain Hecq for their helpful comments and suggestions. We are solely responsible for all errors. This work was supported by the Italian Ministry of Education, University and Research (MIUR), PRIN project "Multivariate statistical models for risk assessment". We also acknowledge RFO grants from the University of Bologna.

than 'traditional' methods for co-integration rank determination. We further show the usefulness of our approach and the benefits of the conditional system analysis in two empirical illustrations, both based on the estimation of VAR systems on U.S. quarterly data. Overall, our analysis shows the gains of combining information criteria with partial systems analysis.

JEL Classification numbers: C30, C32, C52.

1 Introduction

Co-integration rank determination in vector autoregressive (VAR) systems where some variables of interest, Y_t, are modelled conditional on some other variables, Z_t, has been addressed in Harbo, Johansen, Nielsen and Rahbek (1998) [HJNR hereafter], see also Johansen (1992*a*, 1992*b*, 1996). The idea of the conditional analysis is that the dimensionality of the system is reduced, and, when conditioning is valid, co-integration tests may have better power properties. The 'partial system' approach proves to be useful in situations where (i) the dimension of the whole VAR system $p := \dim(X_t)$, $X_t' := (Y_t',\ Z_t')$, is large relative to the sample size T; (ii) the practitioner is essentially interested in modelling Y_t and it is known that Z_t contributes to the long-run equilibrium of the system but either the theory is silent about the short run properties of Z_t, or the scope of the analysis does not require modelling Z_t explicitly; (iii) the time series used for Z_t are considered poor proxies of the corresponding theoretical variables they should measure, especially as concerns their transitory dynamics.

In general, the parameters of the (conditional) model for Y_t given Z_t $(Y_t \mid Z_t)$ and of the marginal model for Z_t are interrelated, which means that a full system analysis is needed to draw efficient inference about the parameters of the two models. The special case where the partial (conditional) model for Y_t given Z_t contains as much information about the parameters of interest as the full system, is when Z_t is weakly exogenous, see Engle, Hendry and Richard (1983). Under the weak exogeneity of Z_t for the co-integrating parameters, HJNR derive the asymptotic distribution of the likelihood ratio test for co-integrating rank in the partial model, and provide tables of critical values. When Z_t is not weakly exogenous for the co-integrating parameters, the conditional model for Y_t given Z_t conveys only partial information about the long-run structure of the system.

Actually, the requirement of weak exogeneity of Z_t can nullify the benefits of the partial system analysis as suggested in HJNR and limit its usefulness in applied work. On the one hand, when weak exogeneity of Z_t fails, the practitioner cannot apply the asymptotic critical values tabulated in HJNR and must resort to the full system analysis. On the other hand, it is difficult to check the hypothesis of weak exogeneity of Z_t efficiently without a full system

analysis or disregarding the model for Z_t. Urbain (1992) shows that 'traditional' orthogonality tests (Pesaran and Smith, 1990) are not sufficient to address the issue of weak exogeneity in error-correction models, see also Boswijk and Urbain (1997). It is not surprising, therefore, that despite conditional (structural) error-correction models have been widely applied in applied work, only seldom the partial system approach has been used for inference on the co-integration rank, see Doornik, Hendry and Nielsen (1998), Bårdsen and Fisher (1999), Bruggeman, Donati and Warne (2003), and Johansen (1992*b*) for an early example. In empirical studies where a (structural) conditional system for Y_t given Z_t is used, the co-integration rank is typically either assumed to be known, or is assumed to be inferred from a full system analysis in a previous stage, see e.g. Boswijk (1995) and Ericsson (1995).

Information-based methods are a well-established alternative to approaches based on (sequential) Neyman-Pearson type tests for co-integration rank determination. In particular, Aznar and Salvador (2002) and Cavaliere, De Angelis, Rahbek, Taylor (2015) [CDRT hereafter], among others, show that standard information criteria such as the familiar Bayesian Information Criterion (BIC) and Hannan-Quinn Information Criterion (HQC) provide powerful alternatives to the traditional Johansen's co-integration rank tests, as they are weakly consistent estimators of the co-integration rank. Conversely, the Akaike Information Criterion (AIC) does not deliver a consistent estimate of the co-integration rank as its penalty does not satisfy the required rate condition (formally defined in Section 3 below). Therefore, it is an asymptotically upward biased estimator of the co-integration rank.

In this paper, we put forth an information criteria-based approach for co-integration rank determination in the partial (conditional) model for Y_t given Z_t, which does not require the hypothesis of weak exogeneity of Z_t with respect to the co-integration parameters. Our main result is that the use of either BIC or HQC in partial systems yields weakly consistent estimates of the true co-integration rank (the same result does not apply to the AIC). Compared to HJNR, the suggested approach does not require the use of tables of critical values: the selected co-integration rank is the one which minimizes BIC or HQC across rank $r = 0, 1, ..., p_Y$, where p_Y:=dim(Y_t). Notably, we prove the consistency of BIC and HQC irrespective of whether Z_t is weakly exogeneous or not. The main implication of this result is that the practitioner can in fact disregard the marginal model for Z_t without pre-testing or being concerned with the weak exogeneity of Z_t. BIC and HQC select asymptotically the 'true' co-integration rank from the conditional model of Y_t given Z_t, and perform reasonably well in finite samples. Thus, our approach to co-integration rank determination fully restores the main benefits of the partial system analysis.

We document the advantages of our approach in finite samples by a set of Monte Carlo experiments. We investigate co-integration rank determination by information criteria in the

partial system approach and compare results with those obtained with the full system approach. We consider two scenarios: one where the conditioning variables Z_t are weakly exogenous with respect to the co-integration parameters, and the other where the weak exogeneity constraint does not hold. We show that, under weak exogeneity, BIC tends to outperform 'traditional' co-integration rank tests in both full and conditional systems. The selection frequency of the null hypothesis of co-integration rank equal to the true co-integration rank delivered by HJRN's sequential test is remarkably inferior to the corresponding selection frequencies provided by BIC when Z_t is weakly exogenous for the co-integration parameters. Conversely, when the hypothesis of weak exogeneity is not valid, BIC tends to underestimate the true co-integration rank in small sample sizes. Instead, HQC in partial systems outperforms its analogue in full systems, irrespective of Z_t being weakly exogenous or not. That is, its selection frequencies are roughly the same in the two cases and the finite sample performance tends to be better than Johansen's (1996) sequential procedure in the full (unconditional) VAR system. Quite surprisingly, the results from the Monte Carlo experiments also show that the violation of the hypothesis of weak exogeneity has only a limited impact on the finite sample performance of HJNR's sequential test, although the associated critical values are tabulated under the maintained hypothesis of exogeneity. In particular, when the conditioning variables are not weakly exogenous for the co-integration parameters, HQC and HJNR's sequential procedure perform in partial systems and in samples of lengths typically available to practitioners no worse than their full system analogues. Thus, despite ignoring the marginal model for Z_t may imply a loss of information on the adjustment mechanisms at work in the whole system, the benefits of reduced dimensionality tends to compensate possible power losses when HQC is used. Interestingly, our Monte Carlo experiments also suggest that the finite sample performance of HQC and BIC (as well as HJNR's sequential procedure) is still satisfactory when VAR disturbances are driven by GARCH-type components. This robustness finding lines up with what Cavaliere, Rahbek and Taylor (2010) and CDRT have shown to be the case in full system analyses.

We also provide two empirical illustrations. In the first example, we emphasize the benefits of combining information criteria and the partial system approach by focusing on a relatively large system including seven ($p = 7$) U.S. macroeconomic variables observed at quarterly frequency on the period 1984-2008. We show that co-integration rank determination from the conditional model for Y_t given Z_t delivers, for suitable choices of Z_t, reliable results relative to the full system analysis. In the second example, we consider the co-integration/common-trend implications of a version of the small-scale dynamic stochastic general equilibrium (DSGE) monetary model of Benati and Surico (2009). We show that in this case, even if the dimension of X_t is not large relative to the sample size, it may be convenient to marginalize Z_t out of the analysis. Our empirical results confirm that co-integration rank determination by information criteria,

especially by the HQC, delivers reasonable and reliable results.

The paper is organized as follows. Section 2 describes the reference co-integrated VAR model. In Section 3, we outline the information criteria for co-integration rank determination in partial systems and we establish their asymptotic properties. The results from our Monte Carlo simulations are reported in Section 4, while Section 5 shows the usefulness of the suggested approach for the two empirical illustrations mentioned above. Section 6 contains some concluding remarks. The proof of our main result, Theorem 1, is given in Appendix A.

2 A conditional co-integrated VAR model

Consider the p-dimensional process $\{X_t\}$ (the full system) which satisfies the k-th order reduced rank VAR model:

$$\Delta X_t = \alpha\beta' X_{t-1} + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \alpha\rho' D_t + \phi d_t + \varepsilon_t, \quad t = 1, ..., T \tag{2.1}$$

where $X_t := (X_{1t}, ..., X_{pt})'$ and the initial values, $X_{1-k}, ..., X_0$, are fixed in the statistical analysis. In the context of (2.1), we assume that the standard 'I$(1, r_0)$ conditions', where $r_0 \in \{0, ..., p\}$ denotes the co-integration rank of the system, defined in Cavaliere, Rahbek and Taylor (2012) hold; that is, the characteristic polynomial associated with (2.1) has $p - r_0$ roots equal to 1 with all other roots lying outside the unit circle, and where α and β have full column rank r_0. The innovation process $\varepsilon_t := (\varepsilon_{1t}, ..., \varepsilon_{pt})'$ is assumed independent and distributed as $N_p(0, \Sigma)$, where Σ is a $(p \times p)$ positive definite and symmetric matrix. Conditional heteroskedaticity can be allowed as in Cavaliere, Rahbek and Taylor (2010). The deterministic variables in (2.1) are taken to satisfy one of the following cases (see, e.g., Johansen, 1996): (i) $D_t = 0$, $d_t = 0$ (no deterministic components); (ii) $D_t = 1$, $d_t = 0$ (restricted constant); or (iii) $D_t = t$, $d_t = 1$ (restricted linear trend).

We decompose X_t as $X_t' = (Y_t',\ Z_t')$, where Y_t is of dimension p_Y and Z_t is of dimension p_Z, with $p_Y + p_Z =: p$. Therefore, $\varepsilon_t' = (\varepsilon_t^{Y\prime},\ \varepsilon_t^{Z\prime})$ in (2.1) are i.i.d. Gaussian with mean zero and covariance matrix

$$\Sigma := \begin{pmatrix} \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix}.$$

All other matrices of parameters in system (2.1) are partitioned conformably. As in HJNR, we define the partial and marginal models assuming, for simplicity and without loss of generality, that $k = 2$ (extension to the case $k > 2$ is straightforward).

The conditional (or partial) model is

$$\Delta Y_t = \omega \Delta Z_t + \alpha_c \beta' X_{t-1} + \Gamma_c \Delta X_{t-1} + \rho_c' D_t + \phi_c d_t + \varepsilon_{ct} \tag{2.2}$$

where $\omega := \Sigma_{YZ}\Sigma_{ZZ}^{-1}$, $\alpha_c := \alpha_Y - \omega\alpha_Z$, $\Gamma_c := \Gamma_{Y1} - \omega\Gamma_{Z1}$ $\Gamma_1' := (\Gamma_{Y1}', \Gamma_{Z1}')$, $\rho_c' := \alpha_Y\rho_Y' - \omega\alpha_Z\rho_Z'$, $\phi_c := \phi_Y - \omega\phi_Z$ and $\varepsilon_{ct} := \varepsilon_t^Y - \omega\varepsilon_t^Z$, with covariance matrix $\Sigma_c = \Sigma_{YY.Z} := \Sigma_{YY} - \Sigma_{YZ}\Sigma_{ZZ}^{-1}\Sigma_{ZY}$.[1] The marginal model is

$$\Delta Z_t = \alpha_Z\beta' X_{t-1} + \Gamma_{Z1}\Delta X_{t-1} + \alpha_Z\rho_Z' D_t + \phi_Z d_t + \varepsilon_t^Z. \tag{2.3}$$

In general, both the conditional model in (2.2) and the marginal model in (2.3) incorporate information about the co-integration relations. Only when $\alpha_Z = 0$, the conditional model (2.2) solely incorporates information about the co-integration relations.

The statistical analysis of the likelihood function of the partial model in (2.2) can be performed by reduced rank regression of the process ΔY_t on X_t, corrected for lagged differences and the conditioning (stationary) variable ΔZ_t (see Chapter 8 in Johansen, 1996, and HJNR).[2] By concentrating the likelihood function of the partial model in (2.2) with respect to the parameters ω, α_Y, β, Γ_c and Σ_c, the residuals R_{0t}, R_{1t}, and $R_{\varepsilon t}$ corresponding to ΔY_t, X_{t-1}, and ε_{ct}, respectively, are obtained, as well as the corresponding residual product moments

$$S_{ij} := T^{-1}\sum_{t=1}^{T} R_{it}R_{jt}', \quad i,j = 0,1,\varepsilon.$$

Up to a constant term which does not depend on r, the maximized pseudo log-likelihood function associated with (2.1) is $\ell_T(r) := -\frac{T}{2}\log|\hat{\Sigma}_{c,r}|$, where $\hat{\Sigma}_{c,r}$ is the residual covariance matrix estimated from the conventional reduced rank regression based on (2.2) with rank r (see Johansen, 1996). Moreover, $|\hat{\Sigma}_{c,r}| := |S_{00}|\prod_{i=1}^{r}(1-\hat{\lambda}_i)$, and hence $\ell_T(r) := -\frac{T}{2}\log|S_{00}| - \frac{T}{2}\sum_{i=1}^{r}\log(1-\hat{\lambda}_i)$, where $\hat{\lambda}_1 > ... > \hat{\lambda}_{p_Y}$ are the p_Y largest solutions of the eigenvalue problem

$$\left|\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}\right| = 0. \tag{2.4}$$

Note that since S_{11} is $p \times p$ and S_{00} is $p_Y \times p_Y$, the rank of $S_{10}S_{00}^{-1}S_{01}$ is at most p_Y. Therefore, the eigenvalue problem in (2.4) has $p_Z = p - p_Y$ zero solutions, thus reducing the possible choice of the (maximum) number of co-integration relations from p to p_Y. Henceforth we call the sequential testing procedure based on the conditional model (2.2) and the implied eigenvalue problem in (2.4) 'HJNR's testing procedure'.

Johansen (1992*a*) shows that when $r = r_0 \leq p_Y$, if Z_t is weakly exogenous for β, the conditional model in (2.2) conveys full information on the co-integration relations and the maximum likelihood estimator for (α, β) in the partial (conditional) model (2.2) has the same asymptotic efficiency of the full system (unconditional) maximum likelihood estimator. In this

[1] We implicitly assume that $\alpha_Y \neq \omega\alpha_Z$, which implies $\alpha_c \neq 0$. As stated in HJNR, the special case $\alpha_c = 0$ is problematic since conducting inference on β from the conditional model (2.2) alone would be not only inefficient due to the violation of the hypothesis of weak exogeneity, but also inconsistent.

[2] For simplicity, here we consider the case of no deterministic components, i.e. case (i) below (2.1).

case, the asymptotic distribution of HJNR's testing procedure for co-integration rank is free of nuisance parameters. Asymptotic critical values have been tabulated under such conditions in HJNR. This is no longer the case if the assumption of weak exogeneity, $\alpha_Z = 0$, does not hold. More precisely, when $\alpha_Z \neq 0$, the conditional model does not convey all the information on the co-integration relations and maximum likelihood inference on β based on system (2.2) is neither fully efficient nor of standard type. However, we will show below that when $0 \leq r_0 \leq p_Y$, it is possible to consistently recover the co-integration rank also with $\alpha_Z \neq 0$ by using information criteria.

Next, we discuss an information criteria-based approach to consistent co-integration rank determination that does not require the investigator to take any stand on the weak exogeneity condition of Z_t with respect to β.

3 Information criteria for co-integration rank determination in partial systems

We now focus on the conditional partial model (2.2) and the implied eigenvalue problem discussed in the previous section. Let $\mathrm{IC}(r) := -2\ell_T(r) + p_T$ denote an information criterion, where $p_T := c_T\pi(r)$ is a penalty function that depends upon the number of parameters $\pi(r)$ and on a term c_T, to be specified below. We have that

$$\mathrm{IC}(r) := T\log|\hat{\Sigma}_{c,r}| + c_T\pi(r) = T\log|S_{00}| + T\sum_{i=1}^{r}\log(1-\hat{\lambda}_i) + c_T\pi(r) \tag{3.1}$$

where the eigenvalues $\hat{\lambda}_i$, $i = 1, ..., r$ are determined from (2.4) and $\pi(r) := r(p_Y + p - r)$ when no deterministic components are involved, $\pi(r) := r(p_Y + p - r + 1)$ in the case of a restricted constant and $\pi(r) := r(p_Y + p - r + 1) + p_Y$ in the restricted trend case. Different values of the coefficient c_T yield different information criteria through the resulting penalty function, p_T. In particular, as is well-known, $c_T = 2$, $\log T$, and $2\log\log T$ yields the AIC, BIC, and HQC, respectively.

The co-integration rank estimator is then given, in generic form, by

$$\hat{r} := \arg\min_{r=0,1,...,p_Y} \mathrm{IC}(r). \tag{3.2}$$

Importantly, the determination of the co-integration rank based on the estimator in (3.2) exempts practitioners from the use of tables of asymptotic critical values which are instead required when sequential testing procedures as in HJNR are applied.

In the following theorem, we state our main result about the consistency of the information

criteria. The key requirement is that the penalty term c_T diverges at a rate lower than T, as $T \to \infty$.

THEOREM 1: Let $\{X_t\}$ be generated as in (2.1) with the true parameters satisfying the I(1, r_0) conditions.

(a) If $r_0 \leq p_Y$, it holds that, as $T \to \infty$:

(a.i) for $r_0 < r \leq p_Y$, $\mathrm{P}(\mathrm{IC}(r) > \mathrm{IC}(r_0)) \to 1$, provided $c_T \to +\infty$;

(a.ii) for $0 \leq r < r_0$, $\mathrm{P}(\mathrm{IC}(r) > \mathrm{IC}(r_0)) \to 1$, provided $c_T/T \to 0$.

(b) If $r_0 > p_Y$, for all $r < p_Y$, as $T \to \infty$, it holds that $\mathrm{P}(\mathrm{IC}(r) > \mathrm{IC}(p_Y)) \to 1$, provided $c_T/T \to 0$.

REMARK 1: The results in case (a) of Theorem 1 imply that, when $r_0 \leq p_Y$, either BIC or HQC penalties in the partial system (2.2) will yield a weakly consistent estimate of r_0. The AIC will not deliver a consistent estimate of the co-integration rank as its penalty does not satisfy condition (a.i) in Theorem 1. Therefore, this criterion will not be considered further.

REMARK 2: When $r_0 > p_Y$ (case (b) of Theorem 1), i.e. the dimension of the conditional system is lower than the true number of co-integration relations in the full system (2.1), only certain linear combinations of the co-integrating vectors can be estimated on the basis of the partial model in (2.2). In such cases, the co-integration rank estimator in (3.2) will tend to select the maximum possible rank, $\hat{r} = p_Y$. Therefore, the results in Theorem 1 imply that, in general, $\hat{r}$ will converge to $\min(r_0, p_Y)$. Hence, in the limit $\hat{r} = p_Y$ whenever $r_0 \geq p_Y$.

REMARK 3: Importantly, in Theorem 1 the determination of the co-integration rank based on the estimator in (3.2) does not require α_Z to be zero in (2.3). That is, our approach can be also applied when the variables in ΔZ_t adjust to $\beta' X_{t-1}$, which is a concrete possibility in applied work. The practitioner can determine the co-integration rank without being concerned about the possible adjustment of ΔZ_t to $\beta' X_{t-1}$. The robustness of BIC and HQC to possible violations of weak exogeneity of the conditioning variables allows one to fully benefit from reductions of the dimensionality of the system. When $r_0 \leq p_Y$ (case (a) of Theorem 1), the intuition at the basis of this result is that the asymptotic behaviour of the eigenvalues from the determinantal problem (2.4) is such that, for $0 \leq r \leq p_Y$, the condition $\mathrm{IC}(r) > \mathrm{IC}(r_0)$ holds, under the conditions (a.i) and (a.ii) of Theorem 1, irrespective of Z_t being weakly exogenous for β or not. More precisely, as formally shown in the proof of Theorem 1 (Appendix A), when $r > r_0$, information criteria rely on a consistent estimator of the error variance matrix Σ_{c,r_0} because the eigenvalues $\hat{\lambda}_{r_0+1}, ..., \hat{\lambda}_r$ of the determinantal problem (2.4) are $O_p(T^{-1})$. This is

valid regardless of the weak exogeneity condition. In this case, the difference $\mathrm{IC}(r) - \mathrm{IC}(r_0)$ involves the term $T\sum_{i=r_0+1}^{r}\log(1-\hat{\lambda}_i)$ which is $O_p(1)$, plus a term which diverges under the condition (a.i) of Theorem 1, leading to $\mathrm{IC}(r_0) < \mathrm{IC}(r)$. Instead, when $r < r_0 \leq p_Y$, albeit when the weak exogeneity condition fails nuisance parameters are involved in the limit distribution of the eigenvalues of the determinantal problem (2.4), there are still $r_0 - r$ positive solutions. In this case, the difference $\mathrm{IC}(r) - \mathrm{IC}(r_0)$ involves the term $-T\sum_{i=r+1}^{r_0}\log(1-\hat{\lambda}_i)$ which is $O_p(T)$, plus a term which converges to zero under the condition (a.ii) of Theorem 1, leading again to $\mathrm{IC}(r_0) < \mathrm{IC}(r)$. When $r_0 > p_Y$ (case (b) of Theorem 1), the argument is similar. Finally, notice that since we do not have any particular restriction on the conditioning variables, Z_t could also contain co-integrated variables.

REMARK 4: Theorem 1 assumes that the system lag length k in (2.1) is known. This assumption, which is often unreasonable in practice, can be relaxed and information criteria can be used to determine both the autoregressive lag length and the co-integration rank following either a jointly or a sequential procedure; see Cavaliere, De Angelis, Rahbek and Taylor (2016). Moreover, Aznar and Salvador (2002) provide an approach based on information criteria to also determine the form of the deterministic component. While we do not consider the determination of the system lag order and the choice of the form of the deterministic component in this paper, the strategies outlined by Cavaliere, De Angelis, Rahbek and Taylor (2016) and Aznar and Salvador (2002), respectively, can also be extended to the present framework.

REMARK 5: Given the correct choice of the co-integration rank, the maximum likelihood estimator of β in the conditional model (2.2) is still super-consistent (albeit inefficient), but the asymptotic distribution theory becomes very difficult, not to say impossible, without the assumption of weak exogeneity, see Johansen (1992*a*), Boswijk (1995), Phillips (1991, 1994) and Phillips and Hansen (1990). This complicates the problem of testing hypotheses concerning β from the conditional model (2.2). Johansen (1992*a,* point (iii), Sect. 4) suggests simulating the distribution involved in the asymptotic tests of hypotheses concerning β in these cases.

4 Simulation results

In this section, we investigate the finite sample performance of BIC and HQC in partial systems through a Monte Carlo experiment which uses large scale systems of $p = \dim(X_t) = 7$ variables as data generating processes.[3] In particular, we consider a DGP where the seven-dimensional

[3]All calculations in this and in the next section have been performed in OxMetrics and MATLAB. Codes are available from the authors upon request.

process $\{X_t\}$ satisfies the following VAR(2) model with 'true' co-integration rank $r_0 = 1$:

$$\Delta X_t = \alpha\beta' X_{t-1} + \Gamma_1 \Delta X_{t-1} + \alpha\rho' + \varepsilon_t, \quad \varepsilon_t \sim iidN(0, \Sigma), \quad t = 1, ..., T \tag{4.1}$$

where $\alpha' := (a, 0, 0, 0, c, c/2, c/3)$, $\beta' := (1, 0, 0, 0, 0, 0, 0)$ and $\Gamma_1 := \gamma I_7$.[4] For the sake of comparison and with the intention of gathering additional insights, the DGP is chosen by extending the Monte Carlo set-up already exploited by Cavaliere, Rahbek and Taylor (2012), CDRT and Cavaliere, De Angelis, Rahbek and Taylor (2016) to the case of larger dimensional VARs**.** Moreover, in order to fully exploit the potential of the conditional analysis we specify a 'full' (and rather arbitrary) structure for the covariance matrix Σ such that the resultant ω matrix in (2.2) is non-zero for all DGPs considered.[5] The initial conditions X_{-1} and X_0 are fixed to zero. Data are generated by setting $a = -0.4$ and $\gamma = 0.5$. The settings for the parameter c are discussed below because in our Monte Carlo design the condition $c \neq 0$ is related to the failure of the weak exogeneity of the conditional variables. Samples of length $T = 100$, 200 and 1000 are generated $M = 5,000$ times from VARs of the form (4.1) with $k = 2$ lags.[6]

For each generated sample, the partial (conditional) system (2.2) is estimated and the procedure for co-integration rank determination discussed in Section 3 is applied. In particular, to evaluate the impact of the dimension of the conditioning variable vector Z_t on the ability of the information criteria to determine the true co-integration rank, we consider the partition $X_t' = (Y_t', Z_t')$, where the dimension of Y_t ranges from two to four, i.e. $p_Y = 2$, 3 and 4, which implies a dimension for Z_t of $p_Z = 5$, 4 and 3, respectively. Thus, it is seen that the scalar parameter c in the α matrix in eq. (4.1) determines whether Z_t is weakly exogenous ($c = 0$) or not ($c \neq 0$) for β.

We consider two scenarios. In the first scenario, 'A' in the following, the parameter c is set to zero hence Z_t is weakly exogenous for β. In the second scenario, 'B' in the following, the parameter c is fixed to $\frac{3}{4}$.[7] Consequently, the asymptotic distribution of HJNR's test statistic will depend on nuisance parameters and tabulated critical values are no longer valid. Within

[4]Note that the assumption $\beta := (1, 0, ..., 0)'$ does not imply any loss of generality. Indeed, any process X_t satisfying (2.1) under the I(1, r_0) conditions with $r_0 = 1$ and $\beta \neq (1, 0, ..., 0)'$ can be suitably linearly transformed into the new process X_t^* such that $\beta^{*\prime} X_t^* \sim I(0)$ with $\beta^* := (1, 0, ..., 0)'$. The likelihood ratio test is invariant to such rotation.

[5]The matrix Σ has been obtained using the following spectral decomposition: $\Sigma = GDG'$, where the diagonal elements of D are generated randomly from a truncated standard normal and the elements of G are generated randomly using the standard normal distribution and such that $GG' = I_p$. The specific matrix Σ used in our simulations is available upon request to the authors. We also considered Monte Carlo experiments based on specifications of Σ leading to $\omega = 0$ in (2.2). Also in this case results are available upon request to the authors.

[6]For each replication, a sample of $T + 200$ observations is generated and the first 200 observations are then discarded.

[7]For given β, $a = -0.4$ and $\gamma = 0.5$, the parameter c cannot be fixed arbitrarily but such that the character-

Scenario B, we consider two additional situations which are of potential interest for applied work. One is the case where, in addition to the violation of the assumption of weak exogeneity, the conditioning set Z_t contains co-integrated variables. The other is the situation where, in addition to the violation of the assumption of weak exogeneity of Z_t, VAR disturbances are non-Gaussian and driven by GARCH-type components.

Scenario A: weak exogeneity

The results for the case of weak exogeneity ($c = 0$ in α of (4.1)) are reported in Table 1, which summarizes the selection frequencies, i.e. the percentage of times BIC and HQC select co-integration rank $r = 0, ..., p_Y$ in the partial systems out of the $M = 5,000$ Monte Carlo replications.[8] The finite sample performance of these information criteria are compared with that of HJNR's sequential test at the 5% nominal significance level. The asymptotic critical values for this test are taken from Table 3 of HJNR. Results are also compared with those obtained from the (unconditional) full VAR analysis. Specifically, for the full system we consider the BIC and HQC as outlined in CDRT and Johansen's (1996) sequential trace test for co-integration rank, again at the 5% nominal significance level.

TABLE 1 ABOUT HERE

Results in Table 1 show that, as expected, both information criteria and the HJNR tests applied to the partial system overall outperform their 'standard' counterparts from the (unconditional) full system when the hypothesis of weak exogeneity is valid. Indeed, comparing the results in the three left-panels of Table 1 (partial systems, $p_Y = 2$, 3 and 4) with the results in the right-panel of Table 1 (full system, $p_Y = p = 7$ and $p_Z = 0$) we observe that the three methods perform better in partial models compared to the (unconditional) full system, especially so for HQC and HJNR's sequential test. In particular, when $T = 100$ ($T = 200$), in the full system, HQC and Johansen's sequential test respectively select the 'true' co-integration rank, $r = r_0 = 1$, only 48.8% and 44.4% (83.5% and 75.3%) of the time, while these percentages increase to 71.9% and 79.1% (89.7% and 88.9%) for HQC and HJNR's sequential test, respectively, in the case of a partial system with $p_Y = 2$ and $p_Z = 5$ (left-panel of Table 1). These

istic equation associated with the VAR in (4.1) does not generate explosive roots. The value c =3/4 respects such a condition.

[8] We do not consider AIC in the Monte Carlo experiments since its poor performance in finite samples is documented in many contributions in the literature (see e.g., Kapetanios, 2004; Wang and Bessler, 2005; CDRT; Cavaliere et al., 2016). Additional simulations show that, also in the case of partial system analysis, this criterion tends to severely overestimate the true co-integration rank. These results are available upon request to the authors.

results highlight the benefits from analysing the partial system in place of the (unconditional) full VAR system in small samples under weak exogeneity.

In line with the results in CDRT, the BIC delivers the best overall performance and correctly estimates the 'true' co-integration rank more than 86% of the time in both partial and full systems, even when the sample size is relatively small ($T = 100$) and irrespectively of the dimension of the set of conditioning variables Z_t. Specifically, when $T = 100$, BIC correctly estimates $r = r_0$ 96.8%, 95.6% and 94.8% of the time for $p_Y = 2$, 3 and 4 ($p_Z = 5$, 4 and 3), respectively, while the percentages for $T = 200$ and $T = 1000$ are substantially constant over the different dimensions of the conditioning set. Conversely, HQC appears to be more affected than BIC by the dimension of the conditioning set when the sample size is relatively small, namely when $T = 100$ and $T = 200$: the larger the dimension of Z_t, p_Z, the better its finite sample performance. HJNR's sequential test performs similarly to HQC in partial systems. For instance, when $T = 200$, HJNR's sequential test correctly estimates the 'true' co-integration rank 88.9%, 84.2% and 82.4% of the time, against 89.7%, 82.7% and 81% of the time for HQC, when $p_Y = 2$, 3 and 4, respectively.

Scenario B: no weak exogeneity

In this scenario, we evaluate the impact of the lack of weak exogeneity of the conditioning variables on the performance of the methods for co-integration rank determination in partial systems. Since such an hypothesis is difficult to test in practice, it is interesting to quantify possible drawbacks associated with a partial system analysis, when the conditioning variables Z_t are not weakly exogenous. Table 2 reports, for the case of no weak exogeneity ($c = \frac{3}{4}$ in α of (4.1)), the estimated selection probabilities, i.e. the percentage of times BIC and HQC select $r = 0, .., p_Y$, respectively, for $p_Y = 2$, 3 and 4. Also in this case we sketch the estimated selection probabilities associated with HJNR's sequential test and, as in the previous scenario, we also consider co-integration rank determination using the (unconditional) full system (right-panel of Table 2).

TABLE 2 ABOUT HERE

Although it clearly outperforms both HQC and Johansen's sequential procedure in the full system analysis (see the results in the right-panel of Table 2), the results in the left-panels of Table 2 show that, when weak exogeneity does not hold, the performance of BIC deteriorates considerably when the sample size is relatively small, as it tends to underestimate the true co-integration rank. In particular, when $T = 100$ ($T = 200$) the selection frequency of the 'true' rank for BIC drops to 41.4%, 33.4% and 27.7% (80.1%, 71.1% and 63.2%) for $p_Y = 2$, 3 and 4, respectively. As predicted by the results in Theorem 1, this effect tends to disappear as

T increases: similarly to the case of weak exogeneity in Table 1, when $T = 1000$ BIC always selects $r = r_0 = 1$, even when the weak exogeneity hypothesis is not valid. Conversely, the HQC performs similarly to the case of weak exogeneity and still outperforms its counterpart in the (unconditional) full system (right-panel of Table 2). In particular, HQC selects $r = r_0 = 1$ only 49.7% (82.8%) of the time in the full system when $T = 100$ ($T = 200$), while these frequencies are 69.5%, 57.2% and 52.5% (89.6%, 84% and 83%) in partial systems for $p_Y = 2$, 3 and 4, respectively. Moreover, these results are very close to (and, in same cases, even outperform) the corresponding results under weak exogeneity (cf. Table 1). These results highlight that the HQC is not severely affected by the violation of the weak exogeneity hypothesis of the conditioning set. Quite surprisingly, also the performance of the HJNR's sequential test does not appear to be strongly affected by the lack of weak exogeneity of Z_t as we would expect, given that the critical values reported in Table 5 of HJNR are tabulated under the weak exogeneity assumption. Moreover, HJNR's sequential test displays a less marked tendency to overestimate the 'true' co-integration rank than Johansen's sequential procedure in the (unconditional) full system, thus leading to a better overall performance also in the case of no weak exogeneity. In particular, Johansen's sequential test selects $r = r_0 = 1$ only 45.5% (74%) of the time when $T = 100$ ($T = 200$), while the corresponding selection frequencies in partial systems for HJNR's sequential test are 69.3%, 62.1% and 57.9% (88.4%, 85.4% and 81.8%) in the case for $p_Y = 2$, 3 and 4, respectively. Furthermore, as in the case of weak exogeneity (Table 1), the empirical performance of HJNR's sequential test and HQC in partial systems are similar. As such, these two methods usefully complement each other in practice and are rather robust to violations of the hypothesis of weak exogeneity.[9]

TABLE 3 ABOUT HERE

We consider a further case which, in addition to the violation of the weak exogeneity hypothesis, allows Z_t to contain co-integrated variables. In particular, we add a further co-integration relation in model (4.1), *ceteris paribus*, by specifying

$$\alpha' := \begin{pmatrix} a & 0 & 0 & 0 & c & c/2 & c/3 \\ -a/2 & 0 & 0 & 0 & -0.5 & -0.3 & 0.25 \end{pmatrix} \text{ and } \beta' := \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix},$$

[9]It is also interesting to stress that the finite sample behaviour of BIC is the one which is more sensitive to the extent of the violation from weak exogeneity: The larger the value of c in α in (4.1), the worse its performance for small T. However, as expected from a theoretical viewpoint and in line with the results in Table 2, the tendency for BIC to underestimate the true co-integration rank in partial systems with no weak exogeneity disappears as T increases. Conversely, HQC and HJNR's sequential test are robust in finite samples to departures from the weak exogeneity assumption. Results for different values of c are available from the authors upon request.

where, as in the previous case of Scenario B, we set $a = -0.4$ and $c = \frac{3}{4}$. In this case, we have that the last two variables in X_t (and in Z_t) co-integrate among themselves, so that the true co-integration rank is $r_0 = 2$. The selection frequencies of $r = 0, ..., p_Y$ for BIC, HQC and HJNR's sequential procedure in partial systems of dimension $p_Y = 2$, 3 and 4 are reported in the left-panels of Table 3, whereas the results obtained from the (unconditional) full VAR analysis are reported in the right-panel of Table 3. Results in Table 3 show that, also in the case when Z_t contains co-integrated variables, HQC and HJNR's sequential procedure provide satisfactory results in partial systems, especially when p_Z is large, outperforming their analogues in the (unconditional) full system, even when the sample size is small. Conversely, when $T = 100$ and 200, BIC tends to underestimate the true co-integration rank and its performance in partial systems is worse than that obtained from the full system analysis.

Therefore, the story we learn from Table 2 and Table 3 is that when the conditioning variables are not weakly exogenous for the co-integration parameters and in samples of lengths typically available to practitioners, HQC and HJNR's sequential procedures (but not BIC) perform in partial systems no worse than their analogues perform in full systems. Thus, despite ignoring the marginal model for Z_t may imply a loss of information on the adjustment mechanisms at work in the whole system, with HQC the reduction of dimensionality tends to fully compensate the possible loss of power in co-integration rank determination from the conditional system.

As we noticed in Section 2, co-integration rank determination in partial systems can potentially be extended to the case of conditional heteroskedasticity along the lines of Cavaliere, Rahbek and Taylor (2010) and CDRT.[10] When the actual unconditional probability distribution of ε_t in system (2.1) is non-Gaussian, it is still in principle possible to factorize the density of X_t as the product of the conditional density for Y_t given Z_t and the marginal density for Z_t. However, the dynamic structure of the conditional system might not necessarily conform to that in eq. (2.2) and the parametric restrictions under which Z_t is weakly exogenous for β might become more involved. It turns out that checking the robustness of the HQC and BIC to conditional heteroskedasticity in partial systems is crucial for applied work. We perform such robustness check by considering the same DGP as in Table 2 but specifying the innovation

[10] Cavaliere, Rahbek and Taylor (2010) show that the extension to conditional heteroskedasticity is allowed as long as that the innovations $\{\varepsilon_t\}$ in (2.1) (i) form a martingale difference sequence with respect to the filtration $\mathcal{F}_t := \sigma(X_t, ..., X_{t-1})$, where $\mathcal{F}_{t-1} \subseteq \mathcal{F}_t$ for $t = ..., -1, 0, 1, 2, ...$, (ii) satisfy the condition $\mathrm{E}\,\|\varepsilon_t\|^4 \leq K < \infty$, $\|\cdot\|$ being the Euclidean norm, and (iii) fulfill the global homoskedasticity condition:

$$\frac{1}{T}\sum_{t=1}^{T} \mathrm{E}\left(\varepsilon_t \varepsilon_t' | \mathcal{F}_{t-1}\right) \xrightarrow{p} \Sigma > 0.$$

process ε_t in (4.1) as a CCC-GARCH (Bollerslev, 1990) process, i.e.,

$$\begin{aligned}
\varepsilon_t &= H_t^{1/2}\nu_t, \quad \nu_t \sim iidN(0, I_p) \\
H_t &:= D_t R D_t = (\rho_{ij}\sqrt{h_{it}h_{jt}}) \\
D_t &:= \text{diag}(h_{1t}^{1/2}, ..., h_{pt}^{1/2}) \\
h_{it} &= d_{0,i} + d_1\varepsilon_{it-1}^2 + d_2 h_{it-1}, \quad i = 1, ..., p \\
R &:= (\rho_{ij}), \quad \rho_{ii} = 1, \qquad i, j = 1, ..., p
\end{aligned}$$

where $H_t = E(\varepsilon_t\varepsilon_t'|\mathcal{F}_{t-1})$ is the conditional variance of ε_t, $\mathcal{F}_t := \sigma(X_t, ..., X_{t-1})$ is the filtration at time t and R is a symmetric positive definite matrix containing the constant conditional correlations ρ_{ij}.[11] Results are reported in Table 4. By comparing selection frequencies in Table 4 with those in Table 2, we observe that the CCC-GARCH(1,1) disturbances have only a limited impact on the finite sample ability of information criteria (as well as HJNR's sequential procedure) to correctly select co-integration rank. This evidence confirms, as expected, the results already documented by Cavaliere, Rahbek and Taylor (2010) and CDRT in the full system framework.

TABLE 4 ABOUT HERE

In summary, the results of our Monte Carlo simulations in scenarios A and B show that the benefits of combining information criteria with partial system analyses can be substantial and usefully complement HJNR's sequential test approach in practice. Moreover, the results in scenarios A and B also show that the HJNR's sequential test does not appear to be severely affected by violations of the weak exogeneity hypothesis, even though the asymptotic distribution of the involved test statistics depends on nuisance parameters and tabulated critical values are no longer valid. It is worth remarking that the HQC and BIC in partial systems perform reasonably well in small samples also under deviations from normality induced by GARCH-type innovations.

5 Empirical illustrations

Despite conditional (possibly uni-equational) error-correction models have been used extensively in the literature (e.g. Urbain, 1992), only seldom the partial system approach to co-integration rank determination has been applied in empirical works. Doornik, Hendry and

[11] Data are generated by using the same specification of α, β and Γ_1 as in Table 2. As concerns the CCC-GARCH part, we draw the coefficients ρ_{ij} $(i \neq j)$ from independent uniform distributions with support (-1,1) and set the GARCH parameters in (4.2) to $d_1 = 0.05$, $d_2 = 0.94$ and $d_{0,i} = 1 - 0.1(i-1)$, $i = 1, ..., p$.

Nielsen (1998), Bårdsen, Jansen and Nymoen (2003) and Bruggeman, Donati and Warne (2003) are some of the few examples where HJNR's testing procedure is implemented. As already observed, two main difficulties characterize the use of HJNR's sequential test in applied work. First, it is generally difficult to select a priori the vector Z_t of weakly exogenous conditioning variables. Second, testing weak exogeneity of Z_t efficiently either requires a full-system approach or the analysis of the marginal model for Z_t, see e.g. Urbain (1992), Boswjik (1995), Ericsson (1995) and Boswjik and Urbain (1997). In both cases the benefits of the partial system analysis are nullified.

In the previous sections we have shown that we can use BIC and/or HQC to determine co-integration rank in the conditional model Y_t given Z_t in (2.2), without being concerned about the weak exogeneity of Z_t. Moreover, the Monte Carlo results show that the finite sample performance of BIC and HQC in the conditional model Y_t given Z_t is generally not worse than the empirical performance of the BIC and HQC in the full system for $X_t' := (Y_t',\ Z_t')$ regardless of whether VAR disturbances are Gaussian or driven by GARCH-type components. In this section, we illustrate the advantages of our approach by two empirical examples, both based on the estimation of VAR models on U.S. quarterly data.

In the first empirical example (Section 5.1), we consider a relatively large-scale VAR system estimated on quarterly data relative to the 'Great Moderation' period, 1984Q1-2008Q2. The idea in this case is to show that, other than reducing the dimensionality of the system, the partial system approach combined with the use of information criteria can be useful for co-integration rank determination. In particular, it is argued that in the absence of clear theoretical guidances about the number of stochastic trends in $X_t' := (Y_t',\ Z_t')$, the typical situation researchers face when dealing with large systems, one useful empirical specification strategy is to start the analysis by including in Z_t time series that are not supposed to co-integrate. Indeed, althought Z_t can be co-integrated in our approach (see the Remark 3 and the Monte Carlo results in Table 3), in large systems it can be relatively easier to make conjectures on which time series should not co-integrate rather than on the variables which should be weakly exogenous. Given different choices of the candidate Z_t vector, the empirical illustration shows that our empirical specification strategy, combined with BIC and HQC, simplifies co-integration rank determination.

In the second empirical example (Section 5.2), we investigate the co-integration/common-trend implications of a small-scale DSGE model for the U.S. business cycle. We show that our approach can be advantageous compared to the full-system analysis when the researcher does not want to take any explicit modelling stand on one of the time series that contributes to the system's long-run equilibrium. More specifically, the DSGE model we consider features the natural rate of output, which is a variable that captures a highly theoretic concept, difficult to

characterize in applied work. We show that it is convenient to treat (proxy of the) natural rate of output as a conditioning variable, disregarding whether it is weakly exogeneous or not.

5.1 Large-dimensional VAR

We consider a vector X_t containing seven U.S. quarterly macroeconomic variables, which includes (the log of) non-durable personal real consumption, denoted cnd_t, (the log of) durable personal real consumption, denoted cd_t, (the log of) fixed-private real investments, denoted i_t, (the log of) the real gross domestic product (GDP), denoted y_t, the inflation rate, denoted inf_t, the federal funds rate, denoted f_t, and the 10 year-Treasury Bill rate, denoted R_t. The macroeconomic literature has recently documented a dramatic fall in the variances of the main macroeconomic time series, which has been termed 'Great Moderation'. Kim and Nelson (1999) and Stock and Watson (2002) offer support for a break in the macroeconomic volatilities around 1984. McConnell and Perez-Quiros (2000) identify 1984Q1 as the breakdate of the variance of the U.S. real GDP. Boivin and Giannoni (2006) also detect a break in the coefficients of a reduced-form VAR for the U.S. economy in the early 1980s. In line with this literature, we identify the beginning and the end of the 'Great Moderation' sample with the dates 1984Q1 and 2008Q2, respectively, for a total of $T + k = 98$ quarterly observations, where k is the VAR lag order for $X_t := (cnd_t, cd_t, i_t, y_t, inf_t, f_t, R_t)'$. The end of the sample, 2008Q2, is justified by our decision to avoid dealing with the acceleration of the financial crisis and, above all, with the binding zero-lower bound (ZLB) that has kicked in since the end of 2008.

FIGURE 1 ABOUT HERE

Figure 1 plots the 'great ratios' $cnd_t - y_t$, $cd_t - y_t$ and $i_t - y_t$, the (ex-post) real interest rate, $R_t - inf_t$, the interest rate differential $R_t - f_t$ and the (ex-post) real version of the policy rate, $f_t - inf_t$, on the period 1984Q1-2008Q2. According to neoclassical theory, under balanced growth there should be a common stochastic trend driving cnd_t, cd_t, i_t and y_t, and the 'great ratios' $cnd_t - y_t$, $cd_t - y_t$ and $i_t - y_t$ should be stationary (King, Plosser, Stock and Watson 1991). Also the interest rates R_t, f_t and the inflation rate inf_t should be driven by a single (common) stochastic trend and the (ex-post) real interest rates and the interest rate differential should be stationary. The simple graphical inspection of Figure 1 suggests, however, that aside perhaps from $cd_t - y_t$, the 'great ratios' do not seem fully consistent with stationary (or trend-stationary) patterns on the period 1984Q1-2008Q2. In particular, the fluctuations of $cnd_t - y_t$ around a possible linear trend and of $i_t - y_t$ around a constant level do not line up with the typical persistence of stationary, mean-reverting processes. Moreover, only the interest rate differential $R_t - f_t$ appears to be stationary (although highly persistent) around a constant level,

while the (ex-post) real interest rates $R_t - inf_t$ and $f_t - inf_t$ display some trending behavior (which can partially be explained in terms of the 'aggressive' monetary policy undertaken by the Fed to fight inflation since the end of the seventies). All this evidence is consistent with the view that it was the change in conduct of monetary policy undertaken by the Fed at the end of the seventies/beginning of the eighties that 'stabilized' the U.S. economy through an aggressive behaviour against inflation. In line with the 'Great Moderation' explanation of the U.S. business cycle (Clarida, Galì and Gertler 2000), one can hardly claim that there are only two common stochastic trends (i.e. five co-integrating relations) driving the variables in X_t. Rather, it would seem reasonable to expect the presence of not more than three co-integrating relationships in the 1984Q1-2008Q2 period, corresponding to the existence of at least four common stochastic trends in X_t. Given our purposes, we prefer not to take any stand about the possible additional co-integration relation other than $cd_t - y_t$ and $R_t - f_t$.

We start our empirical analysis by estimating a VAR for $X_t := (cnd_t, cd_t, i_t, y_t, inf_t, f_t, R_t)'$ ($p = 7$) with four lags ($k = 4$) and a restricted linear trend, which corresponds to the case (iii) $D_t = t$, $d_t = 1$ in system (2.1). The model with restricted linear trend allows for trend-stationary co-integration relationships, which can be an interesting case given the time series patterns observed in Figure 1. Notice that the full system analysis is carried out for comparative purposes only, namely to check whether we obtain significant differences in the co-integration rank determination with respect to the partial system approach. The right-panel of Table 5reports results from Johansen's (1996) sequential co-integration rank procedure at the 5% nominal (asymptotic) level and using BIC and HQC. We select co-integration ranks of 4 (Johansen's test), 2 (BIC) and 4 (HQC), respectively.

TABLE 5 ABOUT HERE

Given the relatively large dimension of X_t and the relatively small number of observations, we investigate whether a partial system analysis delivers a different co-integration rank. Consider, for instance, the following partition of X_t: $X_t' = (Y_t', Z_t')'$, with $Y_t := (cnd_t, cd_t, R_t, y_t, f_t)$ ($p_Y = 5$) and $Z_t := (i_t, inf_t)'$ ($p_Z = 2$). In this case, we have included in Z_t as initial step of the analysis a set of conditioning variables (the fixed-private investments and the inflation rate) which are expected to be I(1) but not co-integrated. Albeit Z_t can potentially be co-integrated in our setup (see Remark 3 and Table 3), our suggestion is that it may be convenient to start the analysis with this choice because it is often easier to make conjectures on which variables are likely to be no co-integrated, rather than on which variables should not adjust to the long run equilibria. Modelling Y_t conditional on Z_t requires considering a five-dimensional system as in (2.2) and, notably, we can apply BIC and HQC ignoring the weak exogeneity condition of Z_t. Results are reported in the middle panel of Table 4, which also reports, for

completeness, the outcome of HJNR's sequential procedure. Although the asymptotic critical values associated with HJNR's test are computed under the assumption of weak exogeneity of Z_t, the Monte Carlo evidence summarized in Tables 1-3 has shown that HJNR's test performs reasonably well in samples of length $T = 100$ irrespective of whether Z_t is weakly exogenous or not (or co-integrated or not). The results in the middle panel of Table 5show that we select co-integration ranks of 3 (HJNR), 1 (BIC) and 3 (HQC), respectively. Therefore, we observe a systematic reduction of the co-integration rank compared to the (unconditional) full system analysis.

We repeat the co-integration rank determination analysis by changing the choice of Y_t and Z_t, in particular using $Y_t := (cd_t, R_t, y_t, f_t)'$ $(p_Y = 4)$ and $Z_t := (cnd_t, i_t, inf_t)'$ $(p_Z = 3)$. In this case, the partial system for Y_t given Z_t in (2.2) is four-dimensional. Results are summarized in the left-panel of Table 4. We still select co-integration rank 2 (HJNR), 1 (BIC) and 3 (HQC), respectively.

Overall, the partial system approach combined with the use of BIC and HQC suggests that the selected co-integration rank does not exceed the rank a practitioner would select by using the full system. The inference we have obtained with the conditional model seems in line with the idea that there should be no more than three co-integration relationships in X_t over the period 1984Q1-2008Q2. More precisely, the application of the HQC on the conditional models Y_t given Z_t provides clear and robust indications that, as expected, X_t is driven by at least four common stochastic trends. It is worth stressing that this conclusion is obtained without any assumption about the weak exogeneity of the conditioning variables included in Z_t.

To complete our analysis, we test whether the hypothesis of weak exogeneity of $Z_t := (i_t, inf_t)'$ $(p_Z = 2)$ and then of $Z_t := (cnd_t, i_t, inf_t)'$ $(p_Z = 3)$ is supported by the data when the co-integration rank is fixed at $r = 3$ as suggested by our partial system approach. To do so, we estimate the full VAR for $X_t := (cnd_t, cd_t, i_t, y_t, inf_t, f_t, R_t)'$, fix the co-integration rank to $r = 3$ and then test the corresponding zero restrictions on the α matrix, given β. The likelihood-ratio test for the null hypothesis of weak exogeneity of $Z_t := (i_t, inf_t)'$ is 37.4 and implies the rejection of the null at the 5% nominal level. Likewise, the likelihood-ratio test for the null hypothesis of weak exogeneity of $Z_t := (cnd_t, i_t, inf_t)'$ equals 55.9 and also leads to reject the null considering the 5% nominal level.

Finally, we observe that the empirical evidence summarized in Table 5 is consistent with the Monte Carlo evidence discussed in Section 4. For instance, the results for the full system analysis reported in Table 2 (Scenario B) suggest that for small values of r_0 relative to p and in samples of length $T = 100$, HQC and Johansen's test tend to select co-integration rank larger than the true one, while BIC tends to select the true co-integration rank more accurately. On the other hand, HQC and HJNR's procedure improve considerably their selection performance

in the partial system approach and become more powerful than BIC, which in turn tends to underestimate the true co-integration rank. Based on these considerations, the conclusion from the partial system approach that the co-integration rank is 3 appears a sensible one.

5.2 Small-scale DSGE model

We consider the small-scale DSGE monetary model of Benati and Surico (2009), also estimated in Bårdsen and Fanelli (2015). The structural model features the following equations:

$$\tilde{x}_t = \gamma E_t \tilde{x}_{t+1} + (1-\gamma)\tilde{x}_{t-1} - \delta(i_t - E_t\pi_{t+1}) + \omega_{\tilde{x},t} \tag{5.1}$$

$$\pi_t = \frac{\varrho}{1+\varrho\chi} E_t\pi_{t+1} + \frac{\chi}{1+\varrho\chi}\pi_{t-1} + \kappa\tilde{x}_t + \omega_{\pi,t} \tag{5.2}$$

$$i_t = \rho i_{t-1} + (1-\rho)(\varphi_\pi \pi_t + \varphi_{\tilde{x}}\tilde{x}_t) + \omega_{i,t} \tag{5.3}$$

$$\omega_{a,t} = \rho_a \omega_{a,t-1} + u_{a,t}, \quad -1 < \rho_a < 1, \quad u_{a,t} \sim \text{WN}(0, \sigma_a^2), \quad a = \tilde{x}, \pi, i \tag{5.4}$$

and expectations are conditional on the agents' information set $\mathcal{F}_t$, i.e. $E_t\cdot := E(\cdot \mid \mathcal{F}_t)$. The first equation is a forward-looking IS curve, the second is the New Keynesian Phillips curve and the third is the policy rule through which the Central Bank fixes the short term interest rate. $\tilde{x}_t = (x_t - x_t^n)$ is the output gap, where x_t is the log of output and x_t^n the natural rate of output; π_t is the inflation rate and i_t is the nominal policy interest rate; $\omega_{a,t}$, $a = \tilde{x}, \pi, i$ are stochastic disturbances autocorrelated of order one, while $u_{a,t}$, $a = \tilde{x}, \pi, i$, can be interpreted as demand, supply and monetary shocks, respectively. The structural parameters are collected in the vector $\theta = (\gamma,\, \delta,\, \varrho,\, \varkappa,\, \kappa, \rho,\, \varphi_\pi,\, \varphi_{\tilde{x}},\, \rho_{\tilde{x}},\, \rho_\pi,\, \rho_i,\, \sigma^2_{\tilde{x}}, \sigma^2_\pi, \sigma^2_i)'$ and their economic interpretation may be found in Benati and Surico (2009). Notice that system (5.1)-(5.4) does not specify how the natural rate of output is generated. To complete the model, we assume that x_t^n obeys the equation

$$\Delta x_t^n = \psi(i_{t-1} - \pi_{t-1}) + \omega_{x^n,t} \quad , \quad \omega_{x^n,t} \sim \text{WN}(0, \sigma^2_{x^n}) \tag{5.5}$$

where $\psi \leq 0$ is an additional auxiliary parameter. When $\psi = 0$, x_t^n in (5.5) follows a driftless random walk which captures the effects of technology shocks and represents the I(1) stochastic trend driving system. With $\psi < 0$, instead, x_t^n is still I(1) but the natural growth rate of output is negatively tied to the level of the (ex-post) real interest rate.

System (5.1)-(5.5) with $\psi = 0$ is analysed in Bårdsen and Fanelli (2015) to investigate the testable implications of small DSGE models, including their common trend/co-integration properties. The appealing feature of system (5.1)-(5.5) is that, under particular conditions, the rational expectations solution of the DSGE model can be approximated by a four-dimensional ($p = 4$) co-integrated VAR system for $X_t := (x_t, \pi_t, i_t, x_t^n)'$. More specifically, the testable

co-integration implications of this DSGE model are captured by the term

$$\tilde{\beta}' X_{t-1} := \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_{t-1} \\ \pi_{t-1} \\ i_{t-1} \\ x^n_{t-1} \end{pmatrix} \tag{5.6}$$

which implies (true) co-integration rank $r_0 = 3$, and requires the 'output gap' $x_t - x^n_t$, the inflation rate π_t and the policy rate i_t to be jointly stationary. Moreover, it can be shown that when $\psi = 0$ in (5.5), the variable x^n_t does not adjust to the three co-integrating relations, i.e. it is weakly exogenous with respect to $\beta = \tilde{\beta}$ in (5.6). In contrast, when $\psi < 0$, x^n_t adjusts to the inflation rate and the policy rate. Equation (5.6) suggests that if the solution of the DSGE model (5.1)-(5.5) is captured by a VAR model for $X_t = (x_t, \pi_t, i_t, x^n_t)'$, the system should be driven by a *single* stochastic trend representing the cumulated effects of technology shocks.

In this model, the natural rate of output x^n_t enters the co-integration relations in (5.6). However, a practitioner might not be interested in modelling x^n_t directly within the VAR framework. Indeed (see also the previous example), the natural rate of output is a variable which captures a highly theoretic concept: it refers to the highest level of real output that can be sustained over the long term, and there is some disagreement among economists about how to characterize such a concept in applied work. In many empirical business cycle models x^n_t is treated as an unobserved (latent) state variable, while in empirical applications x^n_t is often proxied with measures of economic activity provided by official institutions or national statistical institutes. Thus, albeit in this example the dimension of X_t is not particularly large, x^n_t represents a 'natural' conditioning variable for the purpose of testing the co-integration implications of the DSGE model. In addition, the practitioner might not be interested in modelling x^n_t directly within the VAR framework due to poor confidence on the short run properties of the time series used to proxy x^n_t. Hence, it seems reasonable to investigate whether the data support the co-integration/common trend implications predicted by the DSGE model by considering the partition $X_t' = (Y_t',\ Z_t')$, with $Y_t = (x_t, \pi_t, i_t)'$ ($p_Y = 3$) and $Z_t = x^n_t$ ($p_Z = 1$), and then apply the BIC and HQC information criteria on the conditional model for Y_t given Z_t. For comparative purposes, we also compute the sequential co-integration rank testing procedure of HJNR which requires Z_t to be weakly exogenous with respect to $\tilde{\beta}$ matrix in (5.6), namely that $\psi = 0$ in (5.5). As in the previous example, in order to show the usefulness of the partial system approach relative to the (unconditional) full system analysis, we also determine the co-integration rank of X_t through the full VAR system approach.

The variables in $X_t = (x_t,\ \pi_t,\ i_t,\ x^n_t)'$ are chosen as follows: x_t is the (logged) real GDP; π_t is measured by the quarterly growth rate of the GDP deflator; i_t is measured by the effective federal funds rate expressed in quarterly terms (averages of monthly values); x^n_t is approximated

with the official measure provided by the Congressional Budget Office (CBO). The data source is the web site of the Federal Reserve Bank of St. Louis. Also in this case the data cover the 'Great Moderation' period. To compare results with Bårdsen and Fanelli (2015), we focus on the sample 1985Q1-2008Q3 which comprises $T = 95$ quarterly observations. The partial (conditional) system (2.2) is estimated with $k = 2$ lags and a restricted constant (corresponding to case (ii) $D_t = 1$ and $d_t = 0$ in the in the full system (2.1)). Results are reported in the left-panel of Table 5, which summarizes the co-integration rank selected by the BIC and HQC, as well as HJNR's sequential procedure. While BIC selects co-integration rank $r = 1$, HQC selects co-integration rank $r = 3$ as predicted by the DSGE model. Instead, HJNR's sequential procedure selects rank $r = 2$ at the 5% nominal significance level. Recall, however, that the asymptotic critical values reported in Table 6have been tabulated under the maintained hypothesis of weak exogeneity of $Z_t = (x_t^n)$ which might not hold (see below). The right panel of Table 6summarizes the results obtained with the full VAR system analysis and reports the co-integration rank selected by BIC and HQC and Johansen's sequential procedure. We notice that in this case BIC selects $r = 1$, HQC selects $r = 2$ and Johansen's sequential procedure select rank $r = 1$ at the 5% nominal significance level. Aside from BIC and compared to the full system analysis, the co-integration ranks selected by HQC and HJNR's sequential procedure in the conditional model are in line with the prediction of the DSGE model.

TABLE 6 ABOUT HERE

It is interesting in this case to test whether for fixed co-integration rank $r = 3$, which is the value selected by the HQC in the partial system analysis, the variable $Z_t = (x_t^n)$ is weakly exogenous with respect to β in (2.1). This amounts to testing whether the last row of the $p \times 3$ matrix α is zero in system (2.1) for a given (exactly) identified specification of the $p \times 3$ matrix β. The likelihood-ratio test for the three zero restrictions in the α matrix is equal to 35.2 and when compared with asymptotic critical values taken from its null χ^2-distribution with 3 degree of freedom leads to a firm rejection of the null hypothesis of weak exogeneity. This evidence suggests that, although our Monte Carlo results show that HJNR's sequential procedure tends to perform reasonably well in small samples also when Z_t is not weakly exogenous, it must be interpreted with caution in applied work. Conversely, our approach based on information criteria, allows us to disregard any consideration about the weak exogeneity of Z_t and therefore is robust to whether the parameter ψ is $\psi = 0$ or $\psi < 0$ in the structural equations of our DSGE model.

We complete our investigation of the small scale DSGE model by comparing results with Bårdsen and Fanelli (2015). The analysis in Bårdsen and Fanelli (2015), based on a full system approach, suggests that if the co-integration rank is fixed at $r = 3$, a likelihood-ratio test for the

restrictions that $\beta = \tilde{\beta}$ as in (5.6) leads to a p-value of 0.009 when asymptotic critical values are used, and to a p-value of 0.04 when a non-parametric bootstrap version of the likelihood-ratio test is computed along the lines of Boswijk, Cavaliere, Rahbek and Taylor (2016). Overall, we can interpret these evidences as supportive of the small-DSGE model (5.1)-(5.5) on the period 1985Q1-2008Q3.

6 Concluding remarks

Information criteria have been recently used as valid alternative to co-integration rank tests in co-integrated VAR systems. There are cases where it is convenient to devote attention to the model that characterizes the dynamics of the variables of interest, conditional on some other variables whose dynamic characterization is not of primary interest. This approach is useful either in large systems or when the practitioner benefits from marginalizing out variables which contribute to the long-run equilibrium but do not play a key role in the rest of the analysis. In this paper we have extended the use of information criteria for co-integration rank determination to the case of partial systems. We have proved consistency of BIC and HQC in partial systems and shown the robustness of these criteria to violations of the weak exogeneity of the conditioning variables for the co-integration parameters. A set of Monte Carlo experiments and two empirical illustrations have shown that these criteria represent valid alternatives to the 'traditional' approaches to co-integration rank determination.

Since the focus of the paper is on co-integration rank determination in partial (conditional) systems, it can be correctly argued that, once the co-integration rank has been determined from the conditional model by information criteria, the inference on the co-integration parameters (e.g. testing hypotheses on the co-integration parameters) can be more involved relative to the full system analysis if Z_t is not weakly exogenous; see Johansen (1992a) and Boswijk (1995). In particular, if the conditioning variables are not weakly exogenous for the long-run parameters, the maximum likelihood estimator of the co-integration parameters is still super-consistent (albeit not fully efficient) but its asymptotic distribution is not mixed-Gaussian, making the construction of tests cumbersome; see also Phillips (1991, 1994). In this paper we have shown that the benefits of marginalizing Z_t out may compensate, in some specific problems and cases of interest, possible losses of efficiency of estimates and the problem of dealing with non-standard inference. As the examples in the empirical section of this paper show, our approach is particularly convenient when the primary concern of the investigator is co-integration rank determination. Obviously, the decision of whether it is preferable to rely on a full or partial system approach for the problem under investigation depends on the specific purposes of the analysis.

References

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

Bårdsen, G. and L. Fanelli (2015): Frequentist evaluation of small DSGE models, *Journal of Business and Economic Statistics* 33, 307–322.

Bårdsen, G. and P.G. Fisher (1999): Economic theory and econometric dynamics in modelling wages and prices in the United Kingdom, *Empirical Economics* 24, 483–507.

Bårdsen, G., E.S. Jansen and R. Nymoen (2003): Econometric inflation targeting, *Econometrics Journal* 6, 429–460.

Benati, L. and P. Surico (2009): VAR analysis and the Great Moderation, *American Economic Review*, 99, 1636–1652.

Boivin, J., and M. Giannoni (2006), Has monetary policy become more Effective?, *Review of Economics and Statistics*, 88, 445–462.

Bollerslev, T. (1990), Modeling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model. *Review of Economics and Statistics*, 72, 498–505.

Boswijk, P. (1995), Efficient inference on cointegration parameters in structural error correction models, *Journal of Econometrics* 69, 133–158.

Boswijk, P. and J-P Urbain (1997), Lagrange-multiplier tests for weak exogeneity: a synthesis, *Econometric Reviews* 16, 21-38.

Boswijk, P., G. Cavaliere, A. Rahbek and A.M.R. Taylor (2016), Inference on co-integration parameters in heteroskedastic vector autoregression, *Journal of Econometrics*, 192, 64–85.

Bruggeman, A., P. Donati and A. Warne (2003): Is the demand for Euro area M3 stable?, *ECB Working Paper Series* No. 255.

Cavaliere, G., A. Rahbek and A.M.R. Taylor (2010): Cointegration rank testing under conditional heteroskedasticity, *Econometric Theory* 26, 1719–1760.

Cavaliere, G., A. Rahbek and A.M.R. Taylor (2012): Bootstrap determination of the co-integration rank in VAR models, *Econometrica* 80, 1721–1740.

Cavaliere, G., L. De Angelis, A. Rahbek and A.M.R. Taylor (2015): A comparison of sequential and information-based methods for determining the co-integration rank in heteroskedastic VAR models, *Oxford Bulletin of Economics and Statistics* 77, 106–128.

Cavaliere, G., L. De Angelis, A. Rahbek and A.M.R. Taylor (2016): Determining the cointegration rank in heteroskedastic VAR models of unknown order, forthcoming in *Econometric Theory*.

Clarida R. J., Galì J, and M. Gertler (2000), Monetary policy rules and macroeconomic stability: evidence and some theory. *Quarterly Journal of Economics* 115, 147–180.

Doornik, J., D.F. Hendry and B. Nielsen (1998): Inference in cointegrating models: UK M1 revisited, *Journal of Economic Surveys* 12, 533–572.

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

Ericsson, N.R. (1995), Conditional and structural error correction models, *Journal of Econometrics* 69, 159–171.

Harbo, I., S. Johansen, B. Nielsen and A. Rahbek (1998): Asymptotic Inference on Cointegrating Rank in Partial Systems, *Journal of Business & Economic Statistics* 16, 388–399.

Johansen, S. (1992*a*), Cointegration in partial systems and the efficiency of single-equation analysis, *Journal of Econometrics* 52, 389–402.

Johansen, S. (1992*b*), Testing weak exogeneity and the order of cointegration in UK money demand data, *Journal of Policy Modeling* 14(3), 313–334.

Johansen, S. (1996): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.

Kim, C., and C. Nelson (1999), Has the U.S. economy become more stable? A Bayesian approach based on a markov-switching model of the business cycle, *Review of Economics and Statistics*, 81, 608–616.

King, R. G., Plosser, C. I., Stock J. H. and Watson, M. W. (1991), Stochastic trends and economic fluctuations, *American Economic Review* 81, 819–840.

McConnell, M., and G. Perez-Quiros (2000), Output fluctuations in the United States: What has changed since the early 1980s?, *American Economic Review*, 90, 1464–1476.

Pesaran, H.M. and R. Smith (1990), A unified approach to estimation and orthogonality tests in linear single equation econometric models, *Journal of Econometrics* 44, 41–66.

Phillips, P.C.B. (1991), Optimal inference in cointegrated systems, *Econometrica* 59, 283–306.

Phillips, P.C.B. (1994) Some exact distribution theory for maximum likelihood estimators of cointegrating coe cients in error correction models, *Econometrica* 62, 73–93.

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

Stock, J. H., and M. W. Watson (2002), Has the business cycle changed and why?, in M. Gertler and K. Rogoff (eds.): *NBER Macroeconomic Annual*, MIT Press, Cambridge, MA.

Urbain, J-P. (1992), On weak exogeneity in error correction models, *Oxford Bulletin of Economics and Statistics* 54, 187–207.

A Appendix

Without loss of generality, in what follows we assume that there are no deterministic components in the model and in estimation (case (i) in (2.1)), i.e. $D_t = 0$ and $d_t = 0$.

As in HJNR, define the conditional variance

$$\text{Var}\begin{pmatrix} \beta' X_{t-1} & \\ \Delta Y_t & |\Delta X_{t-1}, ..., \Delta X_{t-k+1} \\ \Delta Z_t & \end{pmatrix} = \begin{pmatrix} \Sigma_{\beta\beta} & \Sigma_{\beta Y} & \Sigma_{\beta Z} \\ \Sigma_{Y\beta} & \Sigma_{YY} & \Sigma_{YZ} \\ \Sigma_{Z\beta} & \Sigma_{ZY} & \Sigma_{ZZ} \end{pmatrix}$$

from which we obtain the variance of βX_{t-1} and ΔY_t, conditional on the past and ΔZ_t,

$$\begin{aligned} \text{Var}\begin{pmatrix} \beta' X_{t-1} & \\ \Delta Y_t & |\Delta Z_t, \Delta X_{t-1}, ..., \Delta X_{t-k+1} \end{pmatrix} &= \begin{pmatrix} \Sigma_{\beta\beta.Z} & \Sigma_{\beta Y.Z} \\ \Sigma_{Y\beta.Z} & \Sigma_{YY.Z} \end{pmatrix} \\ &= \begin{pmatrix} \Sigma_{\beta\beta} & \Sigma_{\beta Y} \\ \Sigma_{Y\beta} & \Sigma_{YY} \end{pmatrix} - \begin{pmatrix} \Sigma_{\beta Z} \\ \Sigma_{YZ} \end{pmatrix} \Sigma_{ZZ}^{-1} \begin{pmatrix} \Sigma_{Z\beta} \\ \Sigma_{ZY} \end{pmatrix}'. \end{aligned}$$

The following results are similar to those given by Johansen (1996, Ch. 10 and 11) and HJNR. The results do not assume weak exogeneity, i.e. $\alpha_Z = 0$ in (2.2) and (2.3), as this assumption is not required for the consistency of information criteria in partial systems. As in Section 3, we denote by $\hat{\lambda}_i$, $i = 1, ..., p$, the (ordered) solutions to the eigenvalue problem in (2.4).

Lemma A.1 *Let $\{X_t\}$ be generated as in (2.1) with the parameters satisfying the I(1, r_0) conditions. Then, when $r_0 \leq p_Y$, as $T \to \infty$, $\hat{\lambda}_i$, for $i = 1, ..., r_0$, converge to the nonzero roots of*

$$\left|\lambda \Sigma_{\beta\beta.Z} - \Sigma_{\beta Y.Z}\Sigma_{YY.Z}^{-1}\Sigma_{Y\beta.Z}\right| = 0,$$

with $\Sigma_{\beta\beta.Z} > 0$. Moreover, the remaining $p - r_0$ roots, $\hat{\lambda}_i$, $i = r_0 + 1, ..., p$, are of order $O_p(T^{-1})$.

Proof of Lemma A.1:

Let $B_T' := \alpha_\perp'\Gamma$, where $\Gamma := I_p - \sum_{i=1}^{k-1}\Gamma_i$. By Granger representation theorem, we have that

$$\frac{1}{\sqrt{T}} B_T' X_{\lfloor T\cdot \rfloor} \xrightarrow{w} \alpha_\perp' W(\cdot) =: H(\cdot),$$

where

$$\frac{1}{\sqrt{T}} \sum_{i=1}^{\lfloor T\cdot \rfloor} \varepsilon_i \xrightarrow{w} W(\cdot)$$

and $W(\cdot)$ is the p-dimensional Brownian motion with variance matrix Σ. We also define $W = (W_Y', W_Z')'$, where W_Y is p_Y-dimensional and W_Z is p_Z-dimensional, and $W_c := W_Y - \omega W_Z$.

Similar to the results of Lemma 10.3 in Johansen (1996) and Lemma 6 in HJNR, we have that

$$T^{-1}B_T' S_{11} B_T \xrightarrow{w} \int_0^1 HH'du \tag{A.1}$$

$$B_T' S_{1\varepsilon_c} \xrightarrow{w} \int_0^1 H(dW_c)' \tag{A.2}$$

$$B_T' S_{11}\beta = O_p(1) \tag{A.3}$$

$$\beta' S_{11}\beta \xrightarrow{p} \Sigma_{\beta\beta.Z} \tag{A.4}$$

$$\beta' S_{10} \xrightarrow{p} \Sigma_{\beta Y.Z} \tag{A.5}$$

$$S_{00} \xrightarrow{p} \Sigma_{YY.Z} \tag{A.6}$$

By the continuous mapping theorem, (**??**) and (A.1) implies that

$$\beta' S_{10} S_{00}^{-1} S_{01}\beta = \Sigma_{\beta Y.Z}\Sigma_{YY.Z}^{-1}\Sigma_{Y\beta.Z}. \tag{A.7}$$

Also, from the conditional model (2.2) we have $R_{0t} = \alpha_c\beta' R_{1t} + R_{\varepsilon_t}$, which shows that $S_{01} = \alpha_c\beta' S_{11} + S_{\varepsilon 1}$ and $S_{01}B_T = \alpha_c\beta' S_{11}B_T + S_{\varepsilon 1}B_T$. Therefore, the limiting distribution of $S_{01}B_T$ involves a combination of two stochastic integrals, one given in (**??**) and the other being the limit of $\beta' S_{11}B_T$.

Let $S(\lambda) := \lambda S_{11} - S_{10}S_{00}^{-1}S_{01}$ so that the determinantal problem in (2.4) is $|S(\lambda)| = 0$. We first show that the r_0 largest solution to $|S(\lambda)| = 0$ converge to the roots of $\left|\lambda\Sigma_{\beta\beta.Z} - \Sigma_{\beta Y.Z}\Sigma_{YY.Z}^{-1}\Sigma_{Y\beta.Z}\right| = 0$. Let $A_T := (\beta,\ T^{-1/2}B_T)$ so that, by (**??**), (**??**), (**??**), (**??**) and (A.7), we have

$$|A_T'S(\lambda)A_T| = \left|\lambda\Sigma_{\beta\beta.Z} - \Sigma_{\beta Y.Z}\Sigma_{YY.Z}^{-1}\Sigma_{Y\beta.Z}\right| \left|\lambda \int_0^1 HH'du\right| = 0 \tag{A.8}$$

which has $(p - r_0)$ roots equals to zero and r_0 positive roots. Remind that there are p_Z zero roots in (2.4) and that the $(p_Y - r_0)$ smallest positive eigenvalues from (2.4) are of order T^{-1}. We now show that the remaining $p - r_0$ roots are of order $O_p(T^{-1})$. Similarly to Johansen (1996, p. 159) and Equation (A.7) in HJNR, we have

$$|A_T'S(\lambda)A_T| = |\beta'S(\lambda)\beta| \left|T^{-1}B_T' \left\{S(\lambda) - S(\lambda)\beta\left[\beta'S(\lambda)\beta\right]^{-1}\beta'S(\lambda)\right\} B_T\right|. \tag{A.9}$$

Let $T \to \infty$ with $\rho := T\lambda$ fixed. The first term in (A.9) is

$$\beta'S(\lambda)\beta = \frac{\rho}{T}\beta'S_{11}\beta - \beta'S_{10}S_{00}^{-1}S_{01}\beta = -\Sigma_{\beta Y.Z}\Sigma_{YY.Z}^{-1}\Sigma_{Y\beta.Z} + o_p(1)$$

by (A.7) and since $T^{-1}\beta'S_{11}\beta \xrightarrow{p} 0$ by (**??**). Moreover, since $T^{-1}B_T'S_{11}\beta \xrightarrow{p} 0$ by (**??**), the second term in (A.9) is

$$T^{-1}B_T' \left\{S(\lambda) - S(\lambda)\beta\left[\beta'S(\lambda)\beta\right]^{-1}\beta'S(\lambda)\right\} B_T = T^{-1}\rho B_T'S_{11}B_T - T^{-1/2}B_T'S_{10}N_TS_{01}B_TT^{-1/2} + o_p(1)$$

where $N_T := S_{00}^{-1} - S_{00}^{-1}S_{01}\beta(\beta'S_{10}S_{00}^{-1}S_{01}\beta)^{-1}\beta'S_{10}S_{00}^{-1}$. The distribution of the $p - r_0$ smallest solutions of (2.4) can be derived using the result in (**??**) and noting that, using (**??**)-(A.7), $N_T = N + o_p(1)$, where

$$N := \Sigma_{YY.Z}^{-1} - \Sigma_{YY.Z}^{-1}\Sigma_{Y\beta.Z}(\Sigma_{\beta Y.Z}\Sigma_{YY.Z}^{-1}\Sigma_{Y\beta.Z})^{-1}\Sigma_{\beta Y.Z}\Sigma_{YY.Z}^{-1}. \tag{A.10}$$

In the special case where the hypothesis of weak exogeneity is valid, i.e. $\alpha_Z = 0$ and $\alpha_c = \alpha_Y$, we have $\alpha_\perp = \begin{pmatrix} \alpha_{Y\perp} & 0 \\ 0 & I_{p_Z} \end{pmatrix}$, which implies that (A.10) reduces to $N = \alpha_{Y\perp}(\alpha_{Y\perp}'\Sigma_c\alpha_{Y\perp})^{-1}\alpha_{Y\perp}'$ (see Lemma 4 in HJNR) and $\alpha_{Y\perp}'S_{01}B_T = \alpha_{Y\perp}'S_{\varepsilon 1}B_T \xrightarrow{w} \alpha_{Y\perp}'\int_0^1(dW_c)H'$ (see HJNR, p. 398). Hence, under the assumption of weak exogeneity, the $p - r_0$ smallest solutions of (2.4), normalized by T, converge to those of the equation (A.8) in HJNR. In the general case of no weak exogeneity, the $p - r_0$ smallest eigenvalues (normalized by T) converge to those of a more complicated equation which involves the combination of stochastic integrals that defines the limiting distribution of $S_{01}B_T$ (see discussion above) and (A.10). Nevertheless, we have that $\hat{\lambda}_i$, $i = r_0 + 1, ..., p$, are all of order $O_p(T^{-1})$. □

Proof of Theorem 1

(a) When $r_0 \leq p_Y$, by (3.1), we have that for any $r \in \{r_0 + 1, ..., p_Y\}$,

$$\mathrm{IC}(r) - \mathrm{IC}(r_0) = T \sum_{i=r_0+1}^{r} \log(1 - \hat{\lambda}_i) + c_T(p_Y + p - r - r_0)(r - r_0). \tag{A.11}$$

As in Johansen (1996, p. 160), see also proof of Theorem 1 in HJNR, by Lemma A.1 the first term on the right hand side of (A.11) satisfies $T \sum_{i=r_0+1}^{r} \log(1 - \hat{\lambda}_i) = O_p(1)$. As $(p_Y + p - r - r_0)(r - r_0) > 0$ and $c_T \to \infty$, we have that, $\mathrm{P}(\mathrm{IC}(r) > \mathrm{IC}(r_0)) \to 1$, as $T \to \infty$, which proves part (a.i).

For any $r \in \{0, ..., r_0 - 1\}$, (3.1) implies that

$$\mathrm{IC}(r) - \mathrm{IC}(r_0) = -T \sum_{i=r+1}^{r_0} \log(1 - \hat{\lambda}_i) + c_T(p_Y + p - r - r_0)(r - r_0). \tag{A.12}$$

As in Lemma 12.1 in Johansen (1996) and in the proof of Theorem 1 in HJNR, by Lemma A.1 we have that the r_0 largest sample eigenvalues $\hat{\lambda}_i$, $i = 1, ..., r_0$, converge to the roots λ_i of the equation

$$|\lambda \Sigma_{\beta\beta.Z} - \Sigma_{\beta Y.Z} \Sigma_{YY.Z}^{-1} \Sigma_{Y\beta.Z}| = 0. \tag{A.13}$$

As $\Sigma_{\beta\beta.Z}$, $\Sigma_{YY.Z} > 0$, all the λ_i are positive and, by definition, smaller than one, we have that $-T \sum_{i=r+1}^{r_0} \log(1 - \hat{\lambda}_i)$ is of order $O_p(T)$, where the rate is sharp. As by assumption $c_T/T \to 0$ as $T \to \infty$, the first term on the right side of (A.12) dominates and, hence, $\mathrm{P}(\mathrm{IC}(r) > \mathrm{IC}(r_0)) \to 1$.

(b) When $r_0 > p_Y$, for any $r \in \{0, ..., p_Y - 1\}$, (3.1) implies that

$$\mathrm{IC}(r) - \mathrm{IC}(p_Y) = -T \sum_{i=r+1}^{p_Y} \log(1 - \hat{\lambda}_i) + c_T(p - r)(r - p_Y). \tag{A.14}$$

In this case, there are p_Y sample eigenvalues $\hat{\lambda}_i$, $i = 1, ..., p_Y$, which converge to the positive and, by definition, smaller than one roots λ_i of equation (A.13), whereas the remaining p_Z roots are zero by construction. Hence, we have that $-T \sum_{i=r+1}^{p_Y} \log(1 - \hat{\lambda}_i)$ is of order $O_p(T)$, where the rate is sharp. As by assumption $c_T/T \to 0$ as $T \to \infty$, the first term on the right side of (A.14) dominates and, hence, $\mathrm{P}(\mathrm{IC}(r) > \mathrm{IC}(p_Y)) \to 1$. This completes the proof. □

Table 1: Monte Carlo results for 'Scenario A': VAR(2) model with $r_0 = 1$ with weak exogeneity of Z_t.

	Partial: $p_Y = 2$ ($p_Z = 5$)			Partial: $p_Y = 3$ ($p_Z = 4$)			Partial: $p_Y = 4$ ($p_Z = 3$)			Full: $p_Y = p = 7$ ($p_Z = 0$)		
	BIC	HQC	HJNR	BIC	HQC	HJNR	BIC	HQC	HJNR	BIC	HQC	Johan
						$T = 100$						
$r = 0$	0.2	0.0	0.0	0.8	0.0	0.0	1.6	0.0	0.1	12.0	0.0	0.1
$\mathbf{r = 1}$	96.8	71.9	79.1	95.6	57.9	70.7	94.8	50.2	62.1	86.7	48.8	44.4
$r = 2$	3.0	28.1	20.9	3.4	35.3	26.2	3.5	37.4	31.1	1.4	36.2	36.6
$r = 3$	-	-	-	0.1	6.9	3.1	0.1	10.3	5.8	0.0	11.5	13.8
$r = 4$	-	-	-	-	-	-	0.0	2.1	0.9	0.0	2.7	3.4
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.4	1.1
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.2	0.4
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.1	0.1
						$T = 200$						
$r = 0$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$\mathbf{r = 1}$	99.7	89.7	88.9	99.8	82.7	84.2	99.9	81.0	82.4	100.0	83.5	75.3
$r = 2$	0.3	10.3	11.1	0.2	16.1	14.6	0.1	17.3	15.9	0.0	15.5	20.3
$r = 3$	-	-	-	0.0	1.3	1.2	0.0	1.6	1.5	0.0	1.0	3.6
$r = 4$	-	-	-	-	-	-	0.0	0.2	0.2	0.0	0.1	0.6
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.1
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
						$T = 1000$						
$r = 0$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$\mathbf{r = 1}$	100.0	98.2	93.9	100.0	96.8	93.0	100.0	97.1	91.9	100.0	98.6	91.1
$r = 2$	0.0	1.8	6.1	0.0	3.1	6.6	0.0	2.8	7.5	0.0	1.4	8.0
$r = 3$	-	-	-	0.0	0.1	0.3	0.0	0.1	0.5	0.0	0.0	0.7
$r = 4$	-	-	-	-	-	-	0.0	0.0	0.1	0.0	0.0	0.1
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0

Table 2: Monte Carlo results for 'Scenario B': VAR(2) model with $r_0 = 1$ without weak exogeneity of Z_t.

	Partial: $p_Y = 2$ ($p_Z = 5$)			Partial: $p_Y = 3$ ($p_Z = 4$)			Partial: $p_Y = 4$ ($p_Z = 3$)			Full: $p_Y = p = 7$ ($p_Z = 0$)		
	BIC	HQC	HJNR	BIC	HQC	HJNR	BIC	HQC	HJNR	BIC	HQC	Johan
						$T = 100$						
$r = 0$	57.7	8.7	15.2	65.4	9.7	16.4	71.4	8.8	14.9	13.5	0.1	0.2
$\mathbf{r = 1}$	41.4	69.5	69.3	33.4	57.2	62.1	27.7	52.5	57.9	85.2	49.7	45.5
$r = 2$	1.0	21.7	15.4	1.2	28.0	19.4	1.0	30.4	22.8	1.2	36.3	36.1
$r = 3$	-	-	-	0.0	5.1	2.2	0.0	6.9	3.8	0.0	11.2	13.5
$r = 4$	-	-	-	-	-	-	0.0	1.4	0.6	0.0	2.0	3.4
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.4	0.9
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.1	0.3
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.2	0.1
						$T = 200$						
$r = 0$	19.7	0.4	0.8	28.7	1.0	1.9	36.7	1.1	3.0	0.0	0.0	0.0
$\mathbf{r = 1}$	80.1	89.6	88.4	71.1	84.0	85.4	63.2	83.0	81.8	100.0	82.8	74.0
$r = 2$	0.3	10.0	10.7	0.2	14.1	11.7	0.1	14.6	13.7	0.0	15.4	21.5
$r = 3$	-	-	-	0.0	0.9	0.9	0.0	1.2	1.2	0.0	1.6	3.6
$r = 4$	-	-	-	-	-	-	0.0	0.1	0.2	0.0	0.1	0.7
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.1
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
						$T = 1000$						
$r = 0$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$\mathbf{r = 1}$	100.0	97.6	93.6	100.0	96.7	92.9	100.0	97.1	91.9	100.0	98.6	91.3
$r = 2$	0.0	2.4	6.4	0.0	3.3	6.6	0.0	2.8	7.2	0.0	1.4	7.6
$r = 3$	-	-	-	0.0	0.1	0.5	0.0	0.1	0.8	0.0	0.0	0.9
$r = 4$	-	-	-	-	-	-	0.0	0.0	0.0	0.0	0.0	0.2
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0

Table 3: Monte Carlo results for 'Scenario B': VAR(2) model with $r_0 = 2$ without weak exogeneity of Z_t (and Z_t contains co-integrated variables).

	Partial: $p_Y = 2$ ($p_Z = 5$)			Partial: $p_Y = 3$ ($p_Z = 4$)			Partial: $p_Y = 4$ ($p_Z = 3$)			Full: $p_Y = p = 7$ ($p_Z = 0$)		
	BIC	HQC	HJNR	BIC	HQC	HJNR	BIC	HQC	HJNR	BIC	HQC	Johan
						$T = 100$						
$r = 0$	22.1	0.5	0.7	21.5	0.2	0.7	27.8	0.4	0.8	0.0	0.0	0.0
$r = 1$	57.1	25.8	33.5	53.6	14.5	28.0	54.8	16.3	28.9	52.2	1.2	3.7
$\mathbf{r = 2}$	20.8	73.7	65.8	24.1	67.9	61.1	16.8	57.4	56.9	46.9	56.2	58.7
$r = 3$	-	-	-	0.8	17.4	10.2	0.6	21.6	11.9	0.9	32.1	28.6
$r = 4$	-	-	-	-	-	-	0.0	4.2	1.5	0.0	8.2	7.0
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	1.7	1.5
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.4	0.3
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.2	0.2
						$T = 200$						
$r = 0$	0.8	0.0	0.0	0.3	0.0	0.0	0.7	0.0	0.0	0.0	0.0	0.0
$r = 1$	36.5	2.9	2.7	25.4	0.9	2.2	35.4	1.3	3.8	1.1	0.0	0.0
$\mathbf{r = 2}$	62.8	97.1	97.3	74.1	89.2	89.1	63.8	84.6	86.1	98.9	84.4	82.4
$r = 3$	-	-	-	0.2	10.0	8.7	0.1	13.1	9.3	0.0	14.2	15.1
$r = 4$	-	-	-	-	-	-	0.0	1.0	0.8	0.0	1.3	2.1
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.1	0.2
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.1
						$T = 1000$						
$r = 0$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$r = 1$	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
$\mathbf{r = 2}$	100.0	100.0	100.0	100.0	97.8	94.5	100.0	96.6	94.1	100.0	98.0	92.8
$r = 3$	-	-	-	0.0	2.2	5.5	0.0	3.3	5.5	0.0	2.0	6.5
$r = 4$	-	-	-	-	-	-	0.0	0.1	0.4	0.0	0.0	0.6
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.1
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0

Table 4: Monte Carlo results for 'Scenario B': VAR(2) model with $r_0 = 1$, CCC-GARCH(1,1) innovations and without weak exogeneity of Z_t.

	Partial: $p_Y = 2$ ($p_Z = 5$)			Partial: $p_Y = 3$ ($p_Z = 4$)			Partial: $p_Y = 4$ ($p_Z = 3$)			Full: $p_Y = p = 7$ ($p_Z = 0$)		
	BIC	HQC	HJNR	BIC	HQC	HJNR	BIC	HQC	HJNR	BIC	HQC	Johan
						$T = 100$						
$r = 0$	59.2	22.9	29.1	51.4	14.1	19.7	45.6	8.9	12.6	0.4	0.0	0.0
$\mathbf{r = 1}$	39.7	58.5	57.5	46.4	53.1	58.1	51.9	49.6	55.4	97.1	42.9	38.7
$r = 2$	1.1	18.6	13.3	2.1	27.2	19.5	2.4	30.8	26.7	2.4	38.4	38.2
$r = 3$	-	-	-	0.0	5.6	2.8	0.0	9.1	4.7	0.1	14.1	16.6
$r = 4$	-	-	-	-	-	-	0.0	1.6	0.6	0.0	3.2	4.3
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.9	1.5
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.4	0.6
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.2	0.2
						$T = 200$						
$r = 0$	50.6	23.5	24.8	39.9	13.6	15.0	32.2	8.5	9.3	0.0	0.0	0.0
$\mathbf{r = 1}$	49.1	68.4	66.5	59.7	71.7	71.6	67.6	73.5	73.2	99.8	79.9	70.5
$r = 2$	0.2	8.0	8.7	0.4	13.7	12.4	0.2	15.9	15.7	0.2	17.8	23.2
$r = 3$	-	-	-	0.0	1.0	1.0	0.0	2.0	1.6	0.0	2.1	5.0
$r = 4$	-	-	-	-	-	-	0.0	0.1	0.1	0.0	0.2	0.9
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.2
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.1
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
						$T = 1000$						
$r = 0$	23.1	10.2	8.4	11.8	3.2	2.7	6.1	1.3	1.1	0.0	0.0	0.0
$\mathbf{r = 1}$	76.9	87.1	84.2	88.1	92.2	87.7	93.9	93.3	87.3	100.0	95.7	84.1
$r = 2$	0.0	2.6	7.4	0.1	4.6	9.2	0.0	5.1	10.6	0.0	4.2	13.8
$r = 3$	-	-	-	0.0	0.0	0.5	0.0	0.3	1.0	0.0	0.1	1.7
$r = 4$	-	-	-	-	-	-	0.0	0.0	0.0	0.0	0.0	0.3
$r = 5$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 6$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0
$r = 7$	-	-	-	-	-	-	-	-	-	0.0	0.0	0.0

Table 5: Empirical results for the large-scale VAR on U.S. macroeconomic variables.

	Partial: $Z_t := (cnd_t, i_t, inf_t)'$ ($p_Z = 3$)				Partial: $Z_t := (i_t, inf_t)'$ ($p_Z = 2$)				Full: $p_Y = p = 7$ ($p_Z = 0$)			
	HJNR	(5% c.v.)	BIC	HQC	HJNR	(5% c.v.)	BIC	HQC	Johan	(5% c.v.)	BIC	HQC
$r = 0$	138.3	(89.2)	-1150.4	-1150.4	173.8	(108.0)	-1365.5	-1365.5	247.1	(146.7)	-1678.4	-1678.4
$r = 1$	70.5	(62.9)	**-1168.2**	-1184.9	105.0	(80.9)	**-1379.7**	-1397.9	168.7	(114.9)	-1693.2	-1714.5
$r = 2$	**37.9**	(39.9)	-1159.9	-1190.2	61.5	(56.3)	-1377.9	-1411.2	104.5	(86.9)	**-1702.9**	-1742.3
$r = 3$	12.2	(20.7)	-1153.9	**-1194.8**	**29.1**	(35.5)	-1373.9	**-1419.4**	64.9	(62.6)	-1697.0	-1751.5
$r = 4$	-	-	-1143.3	-1191.8	12.0	(17.9)	-1363.7	-1418.2	**33.4**	(42.2)	-1692.2	**-1758.9**
$r = 5$	-	-	-	-	-	-	-1357.5	-1418.2	19.9	(25.5)	-1678.4	-1754.2
$r = 6$	-	-	-	-	-	-	-	-	7.8	(12.4)	-1672.3	-1754.2
$r = 7 = p$	-	-	-	-	-	-	-	-	-	-	-1671.1	-1756.0

Notes. Results from the partial and full system analyses of $X_t := (cnd_t, cd_t, i_t, y_t, inf_t, f_t, \mathrm{R}_t)'$ ($p = 7$). VAR model with $k = 4$ lags and restricted trend. The 5% c.v. for HJNR's test are taken from Table 3 of Harbo et al. (1998) and the 5% c.v. for Johansen's test ('Johan') are taken from Tables 15.3 and 15.4 of Johansen (1996). The selected co-integration rank according to the test/criterion used are denoted in bold.

Table 6: Empirical results for the DSGE model.

	Partial system				Full system			
	BIC	HQC	HJNR	(5% c.v.)	BIC	HQC	Johan	(5% c.v.)
$r = 0$	-1806.1	-1806.1	71.1	(41.7)	-3699.3	-3699.3	108.7	(53.4)
$r = 1$	**-1816.6**	-1827.3	28.8	(25.5)	**-3701.7**	-3713.7	**33.0**	(34.8)
$r = 2$	-1811.5	-1829.7	**11.2**	(12.3)	-3692.9	**-3714.1**	15.1	(19.9)
$r = 3 = p_Y$	-1809.0	**-1831.8**	-	-	-3684.9	-3712.1	2.4	(9.13)
$r = 4 = p$	-	-	-	-	-3680.9	-3711.1	-	-

Notes. Results from the partial and full system analysis of $X_t := (x_t, \pi_t, i_t, x_t^n)'$ ($p = 4$) and from the partial system analysis of $Y_t := (x_t, \pi_t, i_t)'$ and $Z_t := (x_t^n)$ ($p_Y = 3$ and $p_Z = 1$). The 5% c.v. for HJNR's test are taken from Table 3 of Harbo et al. (1998) and the 5% c.v. for Johansen's test ('Johan') are taken from Table 15.2 of Johansen (1996) (restricted constant). The selected co-integration rank according to the test/criterion used are denoted in bold. The co-integration rank predicted by the theory under the null of the DSGE model is $r = r_0 = 3$.