DISCUSSION PAPERS
Institute of Economics
University of Copenhagen

03-42

# Likelihood Ratio Testing for Cointegration Ranks in I(2) Models

Heino Bohn Nielsen and Anders Rahbek

Studiestræde 6, DK-1455 Copenhagen K., Denmark
Tel. +45 35 32 30 82 - Fax +45 35 32 30 00
http://www.econ.ku.dk

# LIKELIHOOD RATIO TESTING FOR COINTEGRATION RANKS IN I(2) MODELS

Heino Bohn Nielsen* and Anders Rahbek†

*Institute of Economics, University of Copenhagen. heino.bohn.nielsen@econ.ku.dk
†Dept. of Applied Mathematics and Statistics, University of Copenhagen. rahbek@math.ku.dk

Abstract: This paper presents the likelihood ratio (LR) test for the number of cointegrating and multi-cointegrating relations in the I(2) vector autoregressive model. It is shown that the asymptotic distribution of the LR test for the (multi-) cointegration ranks is identical to the asymptotic distribution of the much applied test statistic based on the Two-Step procedure in Johansen (1995), Paruolo (1996), and Rahbek, Kongsted, and Jørgensen (1999). By construction the LR test statistic is smaller than the non-LR test statistic from the Two-Step procedure as the latter ignores some of the restrictions concerning the hypothesis of I(2), and application of the LR test may change rank selection in empirical work. Based on a study of existing empirical applications and related Monte Carlo simulations we conclude that the LR test has much better size properties when compared to the Two-Step based test. Overall, we propose to use of the LR test for rank determination in I(2) analysis as the Two-Step based statistic was developed as a feasible approximation to the then unobtainable LR test.

Keywords: Vector Autoregression, Error Correction Model, Cointegration, I(2), Likelihood Ratio Test, Monte Carlo, Reduced Rank, Rank Testing.
JEL Classification: C32.

## 1 Introduction and Summary

For many OECD countries for the post-war period, the first difference of nominal variables, e.g. inflation rates or money growth, seem to behave as unit root processes, implying that the levels of the nominal variables are integrated of second order, I(2). Johansen (1992) shows that a model for I(2) variables can be parameterized as a vector autoregressive (VAR) model with two reduced rank restrictions imposed, see also Johansen (1995),

Discussions with Søren Johansen and Hans Christian Kongsted are gratefully acknowledged.

Paruolo (1996), Johansen (1997), Rahbek, Kongsted, and Jørgensen (1999), Paruolo and Rahbek (1999), Paruolo (2000) and the survey by Haldrup (1998). Several applications have appeared in the empirical literature on e.g. money demand and open economy price determination, see *inter alia* Juselius (1998), Juselius (1999), Diamandis, Georgoutsos, and Kouretas (2000), Banerjee, Cockerell, and Russell (2001), Banerjee and Russell (2001), Fliess and MacDonald (2001), Nielsen (2002) and Kongsted (2003).

In this paper we present the likelihood ratio (LR) test for the (multi-) cointegration ranks in the I(2) model and derive the asymptotic distribution. Inference on the cointegration ranks $(r, s)$ is central in empirical applications of the multivariate $p-$dimensional I(2) model as it determines the number of multi-cointegrating stationary relations, $r$, and the number of I(1) trends, $s$. The remaining dimension, $p - r - s$, equals the number of common I(2) trends in the system.

Existing literature up to now has applied alone the non-LR test based on the sum of two sets of canonical correlations from the Two-Step estimation procedure in I(2) VAR models, see Johansen (1995), Paruolo (1996), and Rahbek, Kongsted, and Jørgensen (1999). While the Two-Step procedure is possible to implement as a sequential application of reduced rank regression (RRR) well-known from I(1) analysis, it does not make use of the rich structure of the I(2) parameterization. In particular, it ignores a set of restrictions in the first step of the procedure and as a result the value of the LR test for finite samples is in general smaller than the corresponding Two-Step based test. Regarding asymptotics, we find that the asymptotic distribution of the LR test is identical to the one of the Two-Step based test, which implies that existing critical values published for the Two-Step test apply to the LR test as well. Based on a study of existing empirical examples in the literature and related Monte Carlo simulations we examine some finite sample properties of the two tests. First of all, we find that the size properties of the LR test are excellent. Secondly, for ranks $(r, s)$, where $r > 0$ or $p - r - s > 0$, the Two-Step statistic is always larger than the LR statistic, which increases rejection frequencies. This shows in some of the examples where the difference between the LR and the Two-Step statistics is quite large. Overall we propose that the LR test should be used rather than the Two-Step rank test, not only because of the better size properties, but also because the Two-Step based statistic was developed as an approximation to the then unobtainable LR test.

Maximum likelihood (ML) estimation of the parameters in the I(2) model with known ranks, $(r, s)$, have been proposed by Johansen (1997) and Boswijk (2000) based on different parameterizations. We base our derivations on the parametrization in Johansen (1997). Ox code for calculating the LR test for the (multi-) cointegration ranks can be obtained from the authors. Note also that ML estimation of the I(2) model is implemented in the new version of CATS in RATS and in the Matlab program *me2* by Omtzigt.

The rest of the paper is organized as follows: Section 2 presents the I(2) model and the used notation. Section 3 presents the LR and the Two-Step test statistics, and the asymptotic distribution of the LR test is then derived in Section 4. Section 5 illustrates some small sample properties of the rank tests based on an empirical example and a Monte Carlo simulation.

Throughout the paper use is made of the following notation: for any $p \times r$ matrix $\alpha$ of rank $r$, $r < p$, let $\alpha_\perp$ indicate a $p \times (p-r)$ matrix whose columns form a basis of the orthogonal complement of span$(\alpha)$. Hence $\alpha_\perp = 0$ if $r = p$ and $\alpha_\perp = I_p$ if $\alpha = 0$. Define also $\overline{\alpha} = \alpha(\alpha'\alpha)^{-1}$ and let $P_\alpha = \overline{\alpha}\alpha' = \alpha\overline{\alpha}'$ denote the orthogonal projection matrix onto span$(\alpha)$. Finally, the symbols $\xrightarrow{D}$, $\xrightarrow{P}$ and $\overset{D}{=}$ are used to indicate weak convergence, convergence in probability and equality in distribution respectively.

# 2 Model and Representation

This section introduces the notation used throughout and briefly reviews the I(2) VAR model.

## 2.1 The I(2) Model

Consider the $p-$dimensional vector autoregressive model of order $k$ as given by,

$$X_t = \Pi_1 X_{t-1} + \ldots + \Pi_k X_{t-k} + \epsilon_t, \quad t = 1, 2, ..., T,$$

or, in a parameterization convenient for the presentation of I(2) analysis,

$$\Delta^2 X_t = \Pi X_{t-1} - \Gamma \Delta X_{t-1} + \sum_{i=1}^{k-2} \Psi_i \Delta^2 X_{t-i} + \epsilon_t. \tag{1}$$

Here the $p \times p$ dimensional matrices are related by the identities $\Pi = \sum_{i=1}^{k} \Pi_i - I$, $\Gamma = I + \sum_{i=2}^{k} (i-1)\Pi_i$, and $\Psi_j = \sum_{i=j+2}^{k} (i-j-1)\Pi_i$. Finally, $\epsilon_t$ is a $p$-dimensional iid $N(0, \Omega)$ sequence with $\Omega$ positive definite, and the initial values $X_{-k+1}, ..., X_0$ are fixed.

The I(2) model, denoted $H(r, s)$, is then defined by two reduced rank restrictions given by,

$$\begin{aligned} \Pi &= \alpha\beta', & (2)\\ \alpha_\perp' \Gamma \beta_\perp &= \xi\eta', & (3) \end{aligned}$$

where $\alpha$ and $\beta$ are $p \times r$ matrices and $\xi$ and $\eta$ are $(p-r) \times s$ matrices with $r \leq p$ and $s \leq p - r$. Note, that (3) alternatively may be stated as $P_{\alpha_\perp} \Gamma P_{\beta_\perp} = \alpha_1 \beta_1'$, where $\alpha_1 = \overline{\alpha}_\perp \xi$ and $\beta_1 = \overline{\beta}_\perp \eta$ are $p \times s$ matrices, where by definition span$(\alpha_1) \subset$ span$(\alpha_\perp)$ and span$(\beta_1) \subset$ span$(\beta_\perp)$.

We use the notation $H(r) = H(r, p-r)$ to denote I(1) models, in which case $\alpha_1$ and $\beta_1$ are $p \times (p-r)$ matrices, and $H(p)$ denotes the unrestricted VAR.

## 2.2 Representation

Corresponding to the reduced ranks of $\Pi$ and $\alpha_\perp'\Gamma\beta_\perp$ in (2) and (3), consider the assumption that the characteristic polynomial, $A(z) = I_p - \Pi_1 z - \ldots - \Pi_k z^k$, has exactly $2(p-r)-s$ roots at $z=1$ and the remaining roots outside the unit circle. It relates to $H(r,s)$ as the $2(p-r)-s$ unit roots are equivalent to the reduced rank $r<p$ in (2) and the reduced rank $s<p-r$ in (3). When this assumption holds, $X_t$ under $H(r,s)$ is referred to as satisfying $H_0(r,s)$.

Under $H_0(r,s)$, $X_t$ in (1) is an I(2) process with the representation,

$$X_t = C_2 \sum_{s=1}^{t}\sum_{i=1}^{s} \epsilon_i + C_1 \sum_{i=1}^{t} \epsilon_i + \gamma_1 + \gamma_2 t + \tilde{X}_t, \tag{4}$$

where, see Johansen (1992),

$$C_2 = \beta_2(\alpha_2'\Theta\beta_2)^{-1}\alpha_2', \quad \beta' C_1 = \bar{\alpha}'\Gamma C_2, \quad \beta_1' C_1 = \bar{\alpha}_1'(I - \Theta C_2), \tag{5}$$

and $\tilde{X}_t$ is a stationary I(0) process. Here $\Theta = \Gamma\overline{\beta}\overline{\alpha}'\Gamma + I_p - \sum_{i=1}^{k-2}\Psi_i$, $\alpha_2 = (\alpha,\alpha_1)_\perp$, and $\beta_2 = (\beta,\beta_1)_\perp$ such that $\alpha_2'\Theta\beta_2$ has full rank $p-r-s$ under $H_0(r,s)$. The coefficients $\gamma_1$ and $\gamma_2$ depend on the initial values of the process and satisfy $(\beta,\beta_1)'\gamma_2 = 0$, and $\beta'\gamma_1 - \delta\beta_2'\gamma_2 = 0$, where $\delta = \bar{\alpha}'\Gamma\bar{\beta}_2$. Also note that by definition $(\alpha,\alpha_1,\alpha_2)$ and $(\beta,\beta_1,\beta_2)$ are square non-singular matrices with orthogonal blocks, and that $\beta_2$ is a function of $(\beta,\beta_1)$.

It follows from (4) and (5) that $(\beta,\beta_1)'X_t$ is I(1) as $(\beta,\beta_1)'C_2 = 0$, whereas the $(p-r-s)$ linear combinations $\beta_2'X_t$ are I(2). Moreover, the $r$ linear combinations

$$\beta' X_t - \delta\beta_2'\Delta X_t \tag{6}$$

are stationary, i.e. multi-cointegrating. Note also that since $\beta'X_t - \delta\beta_2'\Delta X_t$ and $(\beta,\beta_1)'\Delta X_t$ are stationary, also

$$\beta' X_t - \bar{\alpha}'\Gamma\Delta X_t = \beta' X_t - \bar{\alpha}'\Gamma(P_{\beta,\beta_1} + P_{\beta_2})\Delta X_t \tag{7}$$

is a stationary process.

# 3 Estimation and Rank Test Statistics

This section presents the LR test for cointegration ranks in the I(2) VAR model. The presentation is based on the ML estimation for known ranks $(r,s)$ in the parametrization of Johansen (1997). Also the conventionally applied Two-Step rank test based on the estimator of Johansen (1995) is reviewed. Finally implementation of rank determination based on sequential testing is briefly discussed.

## 3.1 Likelihood Ratio Test

In order to maximize the likelihood function for known ranks $(r, s)$ under the hypothesis $H(r, s)$, Johansen (1997) proposes a reparameterization of the model based on the reduced ranks in (2) and (3) as well as the multi-cointegration term in (7). The parametrization is given by,

$$\begin{aligned} \Delta^2 X_t &= \alpha[\rho'\tau' X_{t-1} + \psi'\Delta X_{t-1}] + \Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\kappa'\tau'\Delta X_{t-1} \\ &\quad + \sum_{i=1}^{k-2} \Psi_i \Delta^2 X_{t-i} + \epsilon_t, \end{aligned} \tag{8}$$

where $\alpha$ $(p \times r)$, $\rho$ $((r+s) \times r)$, $\tau$ $(p \times (r+s))$, $\psi$ $(p \times r)$, $\kappa$ $((r+s) \times (p-r))$, $\Psi_i$ $(p \times p)$, $i = 1, ..., k-2$, and $\Omega$ $(p \times p)$ all are freely varying parameters. The parameters in the previous formulation in (1)-(3) can be derived from the new parameters by the identities $\tau = (\beta, \beta_1)$, $\beta = \tau\rho$, $\psi' = -(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1}\Gamma$, and $\kappa' = -\alpha_\perp'\Gamma(\overline{\beta}, \overline{\beta}_1) = -(\alpha_\perp'\Gamma\overline{\beta}, \xi)$, using the projection identity,

$$\alpha(\alpha'\Omega^{-1}\alpha)^{-1}\alpha'\Omega^{-1} + \Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\alpha_\perp' = I_p.$$

No closed form solution for the ML estimators of the parameters in (8) exists, but estimates can be obtained through an iterative algorithm that switches between two steps: For fixed $\tau$, the parameters $\alpha_\perp$ and $\alpha$ can be obtained by solving an eigenvalue problem and the remaining parameters can be found from regression. For fixed values of these parameters, $\tau$ can be estimated by generalized least squares. Convergence to the global maximum of the likelihood function is not guaranteed, but the value of the likelihood function increases in each iteration. In our implementation we use the Two-Step estimates, presented below, as starting values for the ML iterations.

The LR test for $H(r, s)$ against the unrestricted alternative, $H(p)$, is given by

$$S_{r,s}^{LR} = -2\log Q\left(H(r, s) \mid H(p)\right) = -T\log\left|\breve{\Omega}^{-1}\hat{\Omega}\right|, \tag{9}$$

where $\breve{\Omega}$ and $\hat{\Omega}$ denote the covariance matrices estimated under $H(r, s)$ and $H(p)$ respectively.

## 3.2 Two-Step Rank Test

The Two-Step estimator is based on the parameterization in (1). In the first step, alone the reduced rank restriction of $\Pi$ in (2) is imposed which can stated as the equation,

$$\Delta^2 X_t = \alpha\beta' X_{t-1} - \Gamma\Delta X_{t-1} + \sum_{i=1}^{k-2} \Psi_i \Delta^2 X_{t-i} + \epsilon_t. \tag{10}$$

Ignoring the second restriction (3), the parameters $\alpha$ and $\beta$ are estimated by RRR as applied in I(1) models. The test for reduced rank of $\Pi$ is the familiar *trace* test of $H(r)$ in $H(p)$ :

$$Q_r = -T \sum_{i=r+1}^{p} \log(1-\lambda_i),$$

where $1 \geq \lambda_1 \geq ... \geq \lambda_p \geq 0$ are the eigenvalues from the corresponding RRR, see Johansen (1996, Chapter 6).

The second step is conditional on $r$ and the estimated $\alpha$ and $\beta$ from the first step. Given these, the model is decomposed into a marginal model for $\alpha'_\perp \Delta^2 X_t$ and a conditional model for $\overline{\alpha}'\Delta^2 X_t$ given $\alpha'_\perp \Delta^2 X_t$. The restriction in (3) concerns alone the marginal equation as given by,

$$\alpha'_\perp \Delta^2 X_t = -\xi\eta'\overline{\beta}'_\perp \Delta X_{t-1} - \alpha'_\perp \Gamma \overline{\beta}\beta' \Delta X_{t-1} + \sum_{i=1}^{k-2} \alpha'_\perp \Psi_i \Delta^2 X_{t-i} + \alpha'_\perp \epsilon_t,$$

for which the parameters $\xi$ and $\eta$ are estimated by RRR. Conditional on $r$, the test statistic for reduced rank, $s$, of $\alpha'_\perp \Gamma \beta_\perp$, can be written as

$$Q_{r,s} = -T \sum_{i=s+1}^{p-r} \log(1-\zeta_i),$$

where $1 \geq \zeta_1 \geq ... \geq \zeta_{p-r} \geq 0$ are the second step eigenvalues. Finally, the Two-Step test statistic proposed for the joint hypothesis $H(r,s)$ against the unrestricted model $H(p)$ is given by,

$$S_{r,s}^{2S} = Q_r + Q_{r,s}.$$

## 3.3 Rank Determination

Different values of the cointegration ranks define the sequence of partially nested models illustrated in Table 1 for the $p = 4$ dimensional case.

In the determination of the cointegration ranks in empirical applications, economic theory may provide some guidance. With a particular model suggested by theory, the corresponding hypothesis, $H(r,s)$, can be tested directly against $H(p)$. Alternatively, or if little is known *a priori*, an estimate $(\widehat{r}, \widehat{s})$ of the ranks can be obtained by a sequential application of the rank tests. The idea is to start testing the most restricted model, $H(0,0)$, and in case of rejection to proceed left-to-right and top-to-bottom in Table 1. The estimator can be written as

$$\begin{aligned}(\widehat{r}, \widehat{s}) = \ & \{(r,s) \mid S_{r,s} \leq c_{r,s}; \\ & S_{r_0,s_0} > c_{r_0,s_0} \text{ for the indices } (r_0 < r, s_0 \leq p - r_0) \text{ and } (r_0 = r, s_0 < s)\},\end{aligned}$$

where $S_{r,s}$ denote either of the test statistics, and $c_{r,s}$ their asymptotic critical values, see the next section. The estimator $(\widehat{r}, \widehat{s})$ will select the correct ranks with a limiting

| $r$ | Models | | | | |
|---|---|---|---|---|---|
| 0 | $H(0,0)\ \subset$ | $H(0,1)\ \subset$ | $H(0,2)\ \subset$ | $H(0,3)$ | $\subset\ H(0,4)=H(0)$ |
| | | | | | $\cap$ |
| 1 | | $H(1,0)\ \subset$ | $H(1,1)\ \subset$ | $H(1,2)$ | $\subset\ H(1,3)=H(1)$ |
| | | | | | $\cap$ |
| 2 | | | $H(2,0)\ \subset$ | $H(2,1)$ | $\subset\ H(2,2)=H(2)$ |
| | | | | | $\cap$ |
| 3 | | | | $H(3,0)$ | $\subset\ H(3,1)=H(3)$ |
| | | | | | $\cap$ |
| 4 | | | | | $H(4,0)=H(4)$ |
| $p-r-s$ | 4 | 3 | 2 | 1 | 0 |

*Table 1: Partial nesting structure for p=4.*

probability $(1-\pi)$, where $\pi$ is the size of each test in the sequence as discussed in Johansen (1995).

# 4 Asymptotic Distribution of the LR Test

In this section the asymptotic distribution is given for the LR test of cointegration ranks in the I(2) model $H(r,s)$. We also discuss the inclusion of deterministic terms such as linear trends in the analysis. To present the limiting distributions, introduce the following notation: for two stochastic processes $X_u$ and $Y_u$ on the unit interval $u \in [0,1]$, define the process $X_u$ corrected for $Y_u$ by

$$X_u|\,Y = X_u - \left(\int_0^1 X_s Y_s' ds\right)\left(\int_0^1 Y_s Y_s' ds\right)^{-1} Y_u.$$

## 4.1 No Deterministic Components

Consider the hypothesis $H(r,s)$ against the unrestricted $H(p)$. We state the asymptotic distribution of the LR statistic in (9).

**Theorem 1** *Under $H_0(r,s)$, then as $T \to \infty$,*

$$S_{r,s}^{LR} = -2\log Q(\,H(r,s)|\,H(p)) \xrightarrow{D} Q_r^{\infty} + Q_{r,s}^{\infty},$$

*where*

$$Q_r^{\infty} = tr\left\{\int_0^1 dW_u G_u' \left(\int_0^1 G_u G_u' du\right)^{-1} \int_0^1 G_u dW_u'\right\}$$
$$Q_{r,s}^{\infty} = tr\left\{\int_0^1 dW_{2u} W_{2u}' \left(\int_0^1 W_{2u} W_{2u}' du\right)^{-1} \int_0^1 W_{2u} dW_{2u}'\right\}.$$

*Here $W_u = (W_{1u}', W_{2u}')'$ is a $(p-r)$ dimensional standard Brownian motion on the unit interval, $u \in [0,1]$, with $W_{1u}$ of dimension $s$ and $W_{2u}$ of dimension $p-r-s$. Furthermore,*

$$G_u = \left( \begin{array}{c} W_{1u} \\ \int_0^u W_{2v} dv \end{array} \middle| W_2 \right).$$

The proof is given in the appendix. Note, that the asymptotic distribution of the LR test is identical to the asymptotic distribution of the Two-Step rank test derived in Johansen (1995, Theorem 7). In particular, the two components, $Q_r^\infty$ and $Q_{r,s}^\infty$, have the asymptotic distributions of the statistics from the first and second step of the Two-Step rank test, $Q_r$ and $Q_{r,s}$. This implies that the critical values for the Two-Step rank test can be applied to the LR test.

## 4.2 Linear Trends and Deterministics

In empirical applications it is often important to include deterministic linear trends and a constant level. Rahbek, Kongsted, and Jørgensen (1999) propose a specification which allows deterministic linear trends and a constant level in all components and linear combinations of the process. This includes the multi-cointegrating relations (6), and avoids at the same time the possibility of quadratic trends. An important feature of the model, denoted $H^*(r,s)$ in the following, is that the LR test as well as the Two-Step test, are asymptotically similar with respect to the parameters of the deterministic terms, see also Nielsen and Rahbek (2000) for a discussion of the implication of similarity.

In terms of the parameterization in (8), $H^*(r,s)$ can be represented as:

$$\begin{aligned} \Delta^2 X_t &= \alpha[\rho'\tau^{*\prime} X_{t-1}^* + \psi^{*\prime}\Delta X_{t-1}^*] + \Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\kappa'\tau^{*\prime}\Delta X_{t-1}^* \\ &\quad + \sum_{i=1}^{k-2} \Psi_i \Delta^2 X_{t-i} + \epsilon_t, \end{aligned} \tag{11}$$

where $\tau^* = (\tau', \tau_1')'$ $((p+1)\times(r+s))$, $\psi^* = (\psi', \psi_1')'$ $((p+1)\times r)$, and, finally, $X_{t-1}^* = (X_{t-1}', t)'$ is $p+1$ dimensional. Like in Section 2.2, the model with exactly $2(p-r)-s$ unit roots and the remaining roots with modulus larger than one, is referred to as $H_0^*(r,s)$. Under $H_0^*(r,s)$ the process $X_t$ in (11) has the representation in (4) but with $\gamma_1$ and $\gamma_2$ being functions of $\tau_1$ and $\psi_1$ in addition to initial values, see Rahbek, Kongsted, and Jørgensen (1999). As a result, $X_t$ is as emphasized an I(2) process with linear trends in all linear combinations of the process, including the multi-cointegrating ones.

The result in Theorem 1 is extended to the model $H^*(r,s)$ in Theorem 2. As the proof mimics the proof of Theorem 1, we state the result without proof, simply noting that the asymptotic distributions of the ML estimators are identical to the asymptotic distributions of the Two-Step estimators in Rahbek, Kongsted, and Jørgensen (1999).

With respect to the Two-Step based analysis, $H^*(r,s)$ in terms of the parametrization in (1), corresponds to restricting a constant and linear regressor, $\mu_1 + \mu_2 t$, as follows:

$$\mu_2 = \alpha b_1' \quad \text{and} \quad \alpha_\perp' \mu_1 = -\alpha_\perp' \Gamma \overline{\beta} b_1' - \xi \eta_1', \tag{12}$$

where $b_1$ is $1 \times r$ and $\eta_1$ is $1 \times s$. Each restriction is simple to impose in the two separate steps, see Rahbek, Kongsted, and Jørgensen (1999) for further details.

**Theorem 2** *Under $H_0^*(r,s)$, then as $T \to \infty$,*

$$-2 \log Q(H^*(r,s) |\, H^*(p)) \xrightarrow{D} Q_r^{*\infty} + Q_{r,s}^{*\infty},$$

*where*

$$\begin{aligned} Q_r^{*\infty} &= tr\left\{ \int_0^1 dW_u G_{1u}' \left( \int_0^1 G_{1u} G_{1u}' du \right)^{-1} \int_0^1 G_{1u} dW_u' \right\} \\ Q_{r,s}^{*\infty} &= tr\left\{ \int_0^1 dW_{2u} G_{2u}' \left( \int_0^1 G_{2u} G_{2u}' du \right)^{-1} \int_0^1 G_{2u} dW_{2u}' \right\}. \end{aligned}$$

*Here $W_u = (W_{1u}', W_{2u}')'$ is a $(p-r)$ dimensional standard Brownian motion on the unit interval, $u \in [0,1]$, with $W_{1u}$ of dimension $s$ and $W_{2u}$ of dimension $p-r-s$. Furthermore,*

$$G_{1u} = \left( \begin{array}{l|l} W_{1u} & \\ \int_0^u W_{2v} dv & G_2 \\ u & \end{array} \right) \quad \text{and} \quad G_{2u} = \begin{pmatrix} W_{2u} \\ 1 \end{pmatrix}.$$

Critical values are given in *inter alia* Rahbek, Kongsted, and Jørgensen (1999) and in Doornik (1998).

The results in Theorems 1 and 2 can be extended to further classes of I(2) models considered in the literature. A leading example is the model which allows for a constant level in all linear combinations of the process, including the multi-cointegrating, where the limiting distribution of the LR test for cointegration ranks is the same as reported in Theorem 2, but with $G_1$ and $G_2$ replaced by,

$$\left( \begin{array}{l|l} W_{1u} & \\ \int_0^u W_{2v} dv & W_2 \\ 1 & \end{array} \right) \quad \text{and} \quad W_{2u}, \tag{13}$$

respectively. This model can be represented as in (11), with $X_{t-1}^* = (X_{t-1}', 1)'$ and $\Delta X_t^* = \Delta X_t$ corresponding to (13). Models with this and different sophisticated specifications of trend and level parameters are considered for the Two-Step analysis in Paruolo (1996).

# 5 Finite Sample Properties

From the previous section it is clear that asymptotically the LR and the Two-Step tests are identical and this section explores their finite sample difference. Initially we characterize the relation between the LR statistic and the Two-Step based statistic. Next, we then discuss some main points based on a detailed study of existing empirical applications and corresponding Monte Carlo simulations. For illustration, details of one of the empirical applications are given.

## 5.1 The LR and Two-Step Statistic in Finite Samples

The asymptotic equivalence of the LR test and the Two-Step rank test does not hold for finite samples in general. The relation between the test statistics is stated in the following proposition:

**Proposition 1** *Consider the hypothesis $H(r,s)$ in $H(p)$. In finite samples,*

$$S_{r,s}^{LR} \leq S_{r,s}^{2S}.$$

*Equality holds if $(p-r-s)\,r = 0$.*

Proof: Note first that under the alternative, $H(p)$, estimation is identical for the ML and the Two-Step analysis. However, under the null, $H(r,s)$, the Two-Step procedure does not maximize the likelihood function. One way to see this is by comparing (10) and (8) to obtain

$$\begin{aligned} \Gamma &= -\alpha\psi' - \Omega\alpha_{\perp}\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1}\kappa'\tau' \\ &= \alpha\left(\alpha'\Omega^{-1}\alpha\right)^{-1}\alpha'\Omega^{-1}\Gamma + \Omega\alpha_{\perp}\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1}\alpha_{\perp}'\Gamma\overline{\beta}\beta' + \Omega\alpha_{\perp}\left(\alpha_{\perp}'\Omega\alpha_{\perp}\right)^{-1}\xi\eta'\overline{\beta}_{\perp}'. \end{aligned}$$

The first step of the Two-Step procedure ignores the restriction imposed on the last term, and instead of the $2s\,(p-r) - s^2$ free parameters in $\xi\eta'$ the Two-Step procedure allows for $(p-r)^2$ parameters. In other words, the estimators of $\alpha$ and $\beta$ from the first step are not the MLEs in general. Hence, maximizing the likelihood function in the second step with $\alpha$ and $\beta$ fixed at their first step non-MLE values, will result in a smaller value when compared to the ML estimation. Equivalently, $S_{r,s}^{LR} \leq S_{r,s}^{2S}$ holds. In two special cases, the Two-Step procedure maximizes the likelihood function under $H(r,s)$. First in I(1) models, where $s = p - r$ and $\alpha_{\perp}'\Gamma\beta_{\perp}$ is non-singular. And secondly if $r = 0$ where the second step of the Two-Step procedure is conditional on $\alpha = \beta = 0$. ■

Note, that the magnitude of the difference depends on the number of redundant parameters as well as the sample correlation between the terms in (10) and the redundant terms.

The number of additional parameters is largest for models on the diagonal of Table 1, where $s = 0$, while the LR and the Two-Step statistics coincide for model located in the first row and last column of Table 1. Also note that in the model $H^*(r, s)$, the second restriction in (12) is also ignored in the first step of the Two-Step procedure, which introduces an additional distortion.

Proposition 1 implies that rejection frequencies for hypothesis $H(r, s)$ in $H(p)$, where $r(p - r - s) > 0$, are lower for the LR test than for the Two-Step test. As a consequence the Type I error probability is lower for the LR test, while the Type II error probabilities, for test of more restricted models, are potentially larger.

## 5.2 Empirical Study and Monte Carlo

To analyze small sample properties and compare the LR and Two-Step tests, rank determination in a number of published I(2) applications in the literature was reconsidered. In each case we revisited the rank determination based on the original data using both types of statistics. Furthermore, Monte Carlo simulations were made with the estimated models as data generating processes (DGPs), that is estimated parameter values were used in the DGP definitions. This way the results from the Monte Carlo studies have empirical relevance. Specifically we considered the Danish import price determination in Kongsted (2003), the domestic pricing behavior in Banerjee, Cockerell, and Russell (2001), the money demand analyzed by I(2) VAR models in *inter alia* Johansen (1992) and Rahbek, Kongsted, and Jørgensen (1999), and the Danish export pricing behavior in Nielsen (2002). In addition, we also constructed simulations based on DGPs to cover a broader range of the parameter space of the I(2) model.

The main result emerging from the studies is that the size properties of the LR test are clearly preferable to the Two-Step test, and the rejection frequency of the LR test for a true null hypothesis is often close to the nominal size of the test. The Two-Step test, on the other hand, is often severely size distorted, with rejection frequencies far from the nominal size. This is particular the case if $s = 0$ in the true model. In some cases the LR test still over-rejects true hypotheses, as it is also known from the LR test in I(1) models. If the size distortion of the LR test is deemed important, a Bartlett correction of the likelihood ratio may be developed along the lines of Johansen (2002).

For presentation of the main points we report results only for the Danish import price determination, while the remaining results can be obtained from the authors.

### 5.2.1 Empirical Illustration

We revisit the rank determination based on the original import price data in Kongsted (2003) and then use the estimated parameters to define DGPs for a small Monte

Carlo study. Data consists of import prices, domestic prices, a measure of competeting prices, and an interest rate. Kongsted (2003) estimates a VAR(2) for the effective sample $1975:3-1995:4$. The deterministic specification includes a linear trend in all directions, corresponding to the model $H^*(r,s)$, and an unrestricted impulse dummy for $1992:2$, and we adopt the same specification noting that the impulse dummy has no effect asymptotically. According to Kongsted (2003), economic theory suggests the model $H^*(2,1)$ with two stationary relations, one I(1) trend and one I(2) trend.

The Two-Step rank test statistics, also reported in Kongsted (2003), are given in the left hand side of Table 2. Starting from the most restricted model $H^*(0,0)$ it is possible to reject all models with $r=0$ at a 5% level. In the second row, $H^*(1,0)$ and $H^*(1,1)$ are easily rejected, while $H^*(1,2)$ has a $p-$value of 9%. To achieve that the first model not rejected is the preferred $H^*(2,1)$ one has to use what appears to be a 10% level.

The LR statistics are reported in the right hand side of Table 2. As noted, all test statistics in first row and last column are identical to the Two-Step results, and the remaining test statistics are all lower. Using the LR test there are less evidence for rejecting $H^*(1,2)$. If this model is nevertheless rejected, the next potential model is $H^*(2,0)$, comprising two I(2) trends, again with a $p-$value of 10%. The LR statistic of 45 for this model is markedly lower than the corresponding Two-Step statistic of 75.

Thus based alone on asymptotic inference with an asymptotic $p-$value of 5%, both statistics lead to different results than expected from economic theory. In the next we use Monte Carlo simulations to investigate further the differences of the two tests in particular with respect to size properties.

**Simulations based on the estimated model.** To analyze the difference, we set up a small Monte Carlo simulation. As DGP we use $H^*(2,1)$ with parameters set to the ML estimates, and generate time series, $X_{-101},...,X_1,...X_T$, for $T$ ranging between 50 and 1000, by replacing $\epsilon_t$ with random draws from an iid Gaussian distribution with the estimated covariance matrix. The initial values $X_{-101}$ and $X_{-100}$ are taken from the actual data and the first 100 observations are discarded to eliminate the importance of this choice.

The rejection frequencies at a nominal 5% level are reported in Table 3 for the two tests. The rejection frequencies of $H^*(r,s)$ in $H^*(p)$ are identical for the Two-Step rank test and the LR test for models with $r=0$. For the models $H^*(1,0), H^*(1,1), H^*(1,2), H^*(2,0)$ and $H^*(2,1)$, the finite sample distributions, and the rejection frequencies, differ.

For the LR test the rejection frequencies for the true model $H^*(2,1)$ are very close to the nominal 5% for all sample lengths. The rejection frequencies of some of the more restricted hypotheses, e.g. $H^*(1,2)$ and $H^*(2,0)$, are relatively low in small samples, indicating a low power.

The Two-Step test is clearly size distorted, with rejection frequencies for the true model $H^*(2,1)$ around 30% in small samples. The actual size converges very slowly to the nominal 5%, and for $T = 500$ the rejection frequency is still twice the nominal size. At the same time the rejection frequencies for more restricted models are higher than the LR test. For a small sample, $T = 50$, the rejection frequency for $H^*(2,0)$ is 94% compared to a rejection frequency of 22% for the LR test.

The distributions of the test statistics are reported in Figure 1 for the case $T = 75$. The graphs are organized according to the nesting structure in Table 1. The lower right graph reports the results for tests of the true model, $H^*(2,1)$. The distribution of the LR statistics almost coincides with the asymptotic distribution, while the distribution of the Two-Step statistics is more dispersed and displaced to the right. For some of the more restricted hypotheses, there are big differences between the distributions. The difference is largest if $s = 0$.

**Further simulations based on the estimated model.** To illustrate the difference for the case of $s = 0$ in the DGP, such that a large distortion appears in the Two-Step estimation of the model with correct ranks, we consider the model $H^*(2,0)$ as the DGP. Based on the original data, this model was not rejected by the LR test.

The rejection frequencies are reported in Table 4. Note, that the true model $H^*(2,0)$ is almost always rejected using the Two-Step rank test – and this is the case for all relevant sample lengths. For $T = 200$, corresponding to 50 years of quarterly observations, the actual size is 66%. The LR tests, on the other hand, have good size properties, with rejection frequencies close to the nominal size.

| $r$ | Two-Step rank test | | | | | LR test | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 343.53 | 238.76 | 161.88 | 122.79 | 101.00 | 343.53 | 238.76 | 161.88 | 122.79 | 101.00 |
| | [.00] | [.00] | [.00] | [.00] | [.00] | [.00] | [.00] | [.00] | [.00] | [.00] |
| 1 | | 187.53 | 96.95 | 51.17 | 48.48 | | 142.37 | 88.22 | 49.94 | 48.48 |
| | | [.00] | [.00] | [.09] | [.01] | | [.00] | [.00] | [.11] | [.01] |
| 2 | | | 75.15 | 27.61 | 24.52 | | | 45.36 | 26.10 | 24.52 |
| | | | [.00] | [.27] | [.07] | | | [.10] | [.35] | [.07] |
| 3 | | | | 13.02 | 10.42 | | | | 11.72 | 10.42 |
| | | | | [.37] | [.11] | | | | [.49] | [.11] |
| $p-r-s$ | 4 | 3 | 2 | 1 | 0 | 4 | 3 | 2 | 1 | 0 |

*Table 2: Rank determination for the data in Kongsted (2003). Figures in square brackets are asymptotic $p-$values according to the $\Gamma-$approximation of Doornik (1998).*

| $T$ | Models $H^*(r,s)$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $H^*(0,0)$ | $H^*(0,1)$ | $H^*(0,2)$ | $H^*(0,3)$ | $H^*(0,4)$ | $H^*(1,0)$ | $H^*(1,1)$ | $H^*(1,2)$ | $H^*(1,3)$ | $H^*(2,0)$ | $H^*(2,1)$ |
| | Two-Step rank test | | | | | | | | | | |
| 50 | 100.0 | 100.0 | 98.8 | 92.0 | 83.6 | 100.0 | 92.9 | 53.6 | 43.2 | 94.2 | **33.5** |
| 75 | 100.0 | 100.0 | 100.0 | 100.0 | 98.1 | 100.0 | 99.8 | 76.9 | 72.4 | 99.5 | **32.9** |
| 100 | 100.0 | 100.0 | 100.0 | 100.0 | 99.8 | 100.0 | 100.0 | 89.5 | 91.3 | 100.0 | **28.7** |
| 200 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **16.8** |
| 500 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **10.1** |
| 1000 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **7.5** |
| | LR test | | | | | | | | | | |
| 50 | 100.0 | 100.0 | 98.8 | 92.0 | 83.6 | 97.5 | 66.7 | 27.7 | 43.2 | 21.5 | **4.0** |
| 75 | 100.0 | 100.0 | 100.0 | 100.0 | 98.1 | 100.0 | 98.6 | 53.3 | 72.4 | 51.8 | **5.5** |
| 100 | 100.0 | 100.0 | 100.0 | 100.0 | 99.8 | 100.0 | 100.0 | 77.7 | 91.3 | 81.2 | **6.3** |
| 200 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **5.7** |
| 500 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **5.6** |
| 1000 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **5.5** |

*Table 3: Rejection frequencies in a simulation based on Kongsted (2003). DGP is $H^*(2,1)$. The tests are not calculated sequentially. Bold indicates rejection frequencies for tests of the correct model (empirical size). Based on 5000 replications and a nominal 5% level.*

| $T$ | Models $H^*(r,s)$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $H^*(0,0)$ | $H^*(0,1)$ | $H^*(0,2)$ | $H^*(0,3)$ | $H^*(0,4)$ | $H^*(1,0)$ | $H^*(1,1)$ | $H^*(1,2)$ | $H^*(1,3)$ | $H^*(2,0)$ |
| | Two-Step rank test | | | | | | | | | |
| 50 | 100.0 | 100.0 | 96.2 | 84.3 | 82.8 | 100.0 | 86.6 | 33.3 | 39.3 | **91.0** |
| 75 | 100.0 | 100.0 | 100.0 | 99.1 | 97.0 | 100.0 | 99.1 | 46.6 | 59.5 | **91.1** |
| 100 | 100.0 | 100.0 | 100.0 | 100.0 | 99.7 | 100.0 | 99.9 | 64.6 | 79.5 | **86.3** |
| 200 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 98.1 | 99.6 | **66.3** |
| 500 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **38.3** |
| 1000 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **23.0** |
| 2500 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **11.2** |
| 5000 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **8.0** |
| | LR test | | | | | | | | | |
| 50 | 100.0 | 100.0 | 96.2 | 84.3 | 82.8 | 94.3 | 47.7 | 17.7 | 39.3 | **6.6** |
| 75 | 100.0 | 100.0 | 100.0 | 99.1 | 97.0 | 100.0 | 90.2 | 31.2 | 59.5 | **7.5** |
| 100 | 100.0 | 100.0 | 100.0 | 100.0 | 99.7 | 100.0 | 99.2 | 55.7 | 79.5 | **7.5** |
| 200 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 97.6 | 99.6 | **5.9** |
| 500 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **5.9** |
| 1000 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **6.1** |
| 2500 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **5.6** |
| 5000 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | **5.2** |

*Table 4: Rejection frequencies in a simulation based on Kongsted (2003). DGP is $H^*(2,0)$. See also notes to Table 3.*

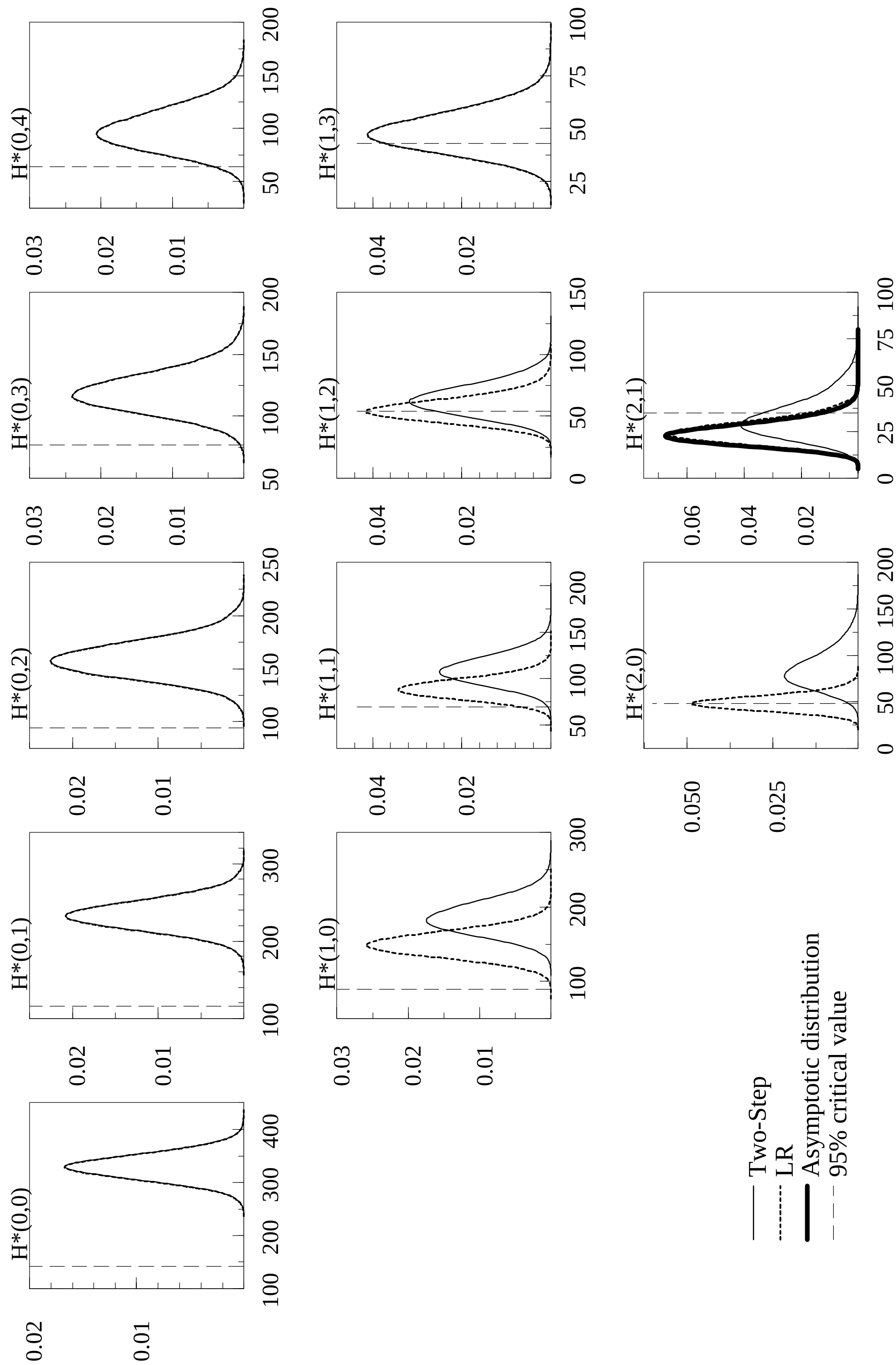


*Figure 1: Distributions of the two test statistics for the case T=75. Graphs are organized according the partial nesting structure. Based on 5000 replications.*

# A Proof of Theorem 1

In this appendix we derive the asymptotic distribution of the LR test for cointegration ranks in the I(2) model. In order to motivate the notation and ease the presentation of the I(2) test we start by considering the well-known likelihood ratio test for cointegration rank in the I(1) model.

Throughout we use the notation $\breve{\theta}$ and $\hat{\theta}$ to denote the ML estimates of a parameter $\theta$ under the null hypothesis and under the alternative respectively.

## A.1 Asymptotics for the I(1) LR Test

Consider the well-known I(1) model and the rank test in this case. The present rederivation of the LR test is not based on the conventional representation in terms of eigenvalues and canonical correlations, see e.g. Johansen (1996, Chapter 11), but is instead based on a linear regression type formulation. It is assumed here that the reader is familiar with the well-established literature on I(1) VAR models.

For simplicity and without loss of generality consider the simplest case of the $p-$dimensional I(1) VAR(1) model as given by

$$\Delta X_t = \Pi X_{t-1} + \epsilon_t, \tag{14}$$

with the hypothesis $H(r)$ parameterized as $\Pi = \alpha\beta'$. The aim is to derive the asymptotic distribution of the likelihood ratio test,

$$-2\log Q(\,H(r)|\,H(p)) = -T\log\left|\breve{\Omega}^{-1}\hat{\Omega}\right|,$$

where $\breve{\Omega}$, $\hat{\Omega}$ are the ML estimates of the covariance matrix $\Omega$ under the hypothesis $H(r)$ and the alternative respectively. Denote by $H_0(r)$ the model with exactly $p-r$ unit roots in the characteristic polynomial, $A(z)$, with the remaining roots outside the unit circle.

**Lemma 1** *Under $H_0(r)$, then as $T \to \infty$, the LR statistic $-2\log Q(\,H(r)|\,H(p))$ converges in distribution to*

$$tr\left\{\left(\int_0^1 W_u dW_u'\right)'\left(\int_0^1 W_u W_u' du\right)^{-1}\int_0^1 W_u dW_u'\right\}, \tag{15}$$

*where $W_u$ is a $(p-r)$-dimensional standard Brownian motion on the unit interval, $u \in [0,1]$.*

Proof: Recall that under $H(r)$ the parameters $\alpha$ and $\beta$ are non-identified. Identification is obtained by normalization on the known $p \times r$ matrix $c$ such that $\beta_c = \beta(c'\beta)^{-1}$ and

$\alpha_c = \alpha\beta' c$, and hence $\Pi = \alpha\beta' = \alpha_c\beta_c'$. Denote by $\alpha_0$, $\beta_0$ and $\Omega_0$ the true parameters corresponding to the null, $H_0(r)$.

In order to derive the asymptotic distribution of $-2\log Q(H(r)|\, H(p))$ introduce the simple auxiliary null hypothesis $H_{\text{aux}}$ that $\beta$ is known, $\beta = \beta_0$, as given by the equation,

$$\Delta X_t = \alpha\beta_0' X_{t-1} + \epsilon_t. \tag{16}$$

Using that by definition $H_{\text{aux}} \subseteq H(r) \subseteq H(p)$, and therefore in particular $Q(H(r)|\, H(p)) = Q(H(r)|\, H_{\text{aux}}) \times Q(H_{\text{aux}}|\, H(p))$, it holds that

$$-2\log Q(H(r)|\, H(p)) = -2\log Q(H_{\text{aux}}|\, H(p)) - [-2\log Q(H_{\text{aux}}|\, H(r))]. \tag{17}$$

Consider first $-2\log Q(H_{\text{aux}}|\, H(r))$: Introduce the $(p-r)\times r$ dimensional parameter

$$B = \bar{\beta}_{0\perp}' (\beta - \beta_0), \tag{18}$$

where $\beta$ is normalized by $c = \bar{\beta}_0$. With this definition, $H(r)$ can be rewritten as,

$$\Delta X_t = \alpha\beta' X_{t-1} + \epsilon_t = \alpha\beta' \left(\bar{\beta}_{0\perp}\beta_{0\perp}' + \bar{\beta}_0\beta_0'\right) X_{t-1} + \epsilon_t = \alpha\left(B' Z_{1t} + Z_{0t}\right) + \epsilon_t,$$

where $\alpha$ and $\beta$ normalized on $c = \bar{\beta}_0$. Here $Z_{1t} = \beta_{0\perp}' X_{t-1}$ and $Z_{0t} = \beta_0' X_{t-1}$ are I(1) and I(0) processes respectively. Note that the hypothesis $H_{\text{aux}}$ is simply given by $B = 0$. Define the estimated residuals

$$\hat{\epsilon}_t = \Delta X_t - \hat{\alpha}\left(\hat{B}' Z_{1t} + Z_{0t}\right), \quad \breve{\epsilon}_t = \Delta X_t - \breve{\alpha} Z_{0t} \quad \text{and} \quad \hat{\epsilon}_{0t} = \Delta X_t - \hat{\alpha} Z_{0t}.$$

Then by definition, $-2\log Q(H_{\text{aux}}|\, H(r)) = -T\log\left|\breve{\Omega}^{-1}\hat{\Omega}\right|$ with $\breve{\Omega} = \frac{1}{T}\sum_{t=1}^{T}\breve{\epsilon}_t\breve{\epsilon}_t'$, and,

$$\hat{\Omega} = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_t\hat{\epsilon}_t' = \frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_{0t}\hat{\epsilon}_{0t}' - (Y_T + Y_T' - X_T),$$

where

$$X_T = \hat{\alpha}\hat{B}'\left(\frac{1}{T}\sum_{t=1}^{T} Z_{1t}Z_{1t}'\right)\hat{B}\hat{\alpha}' \quad \text{and} \quad Y_T = \left(\frac{1}{T}\sum_{t=1}^{T}(\Delta X_t - \hat{\alpha} Z_{0t})\, Z_{1t}'\right)\hat{B}\hat{\alpha}'.$$

Next, the asymptotic distribution of the ML estimates of $\alpha$ and $\beta$ under $H(r)$ and normalized by $c = \bar{\beta}_0$, $\hat{\alpha}$ and $\hat{\beta}$ respectively, is given in Johansen (1996) and it follows that,

$$T\hat{B} = T\bar{\beta}_{0\perp}'\left(\hat{\beta} - \beta_0\right) \xrightarrow{D} B^\infty = \left(\int_0^1 F_u F_u' du\right)^{-1}\int_0^1 F_u dV_u' \Omega_0^{-1}\alpha_0(\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}, \tag{19}$$

with $F_u = \beta_{0\perp}' M V_u$, where $V_u$ is a Brownian motion on $u \in [0,1]$ with covariance $\Omega_0$, and $M = \beta_{0\perp}(\alpha_{0\perp}'\beta_{0\perp})^{-1}\alpha_{0\perp}'$.

Using (19), the consistency of $\hat{\alpha}$ and the continuous mapping theorem, it follows that,

$$
\begin{aligned}
TX_T &\xrightarrow{D} \alpha_0 B^{\infty\prime}\left(\int_0^1 F_u F_u' du\right) B^\infty \alpha_0' \\
&= \alpha_0(\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}\alpha_0'\Omega_0^{-1}\int_0^1 dV_u F_u'\left(\int_0^1 F_u F_u' du\right)^{-1}\int_0^1 F_u dV_u'\Omega_0^{-1}\alpha_0(\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}\alpha_0'.
\end{aligned}
$$

Similarly, by convergence to stochastic integrals,

$$
\begin{aligned}
TY_T &= \left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t Z_{1t}' - (\hat{\alpha}-\alpha_0)\frac{1}{T}\sum_{t=1}^{T}\epsilon_t Z_{0t}'\right) T\hat{B}\hat{\alpha}' \\
&\xrightarrow{D} \int_0^1 dV_u F_u'\left(\int_0^1 F_u F_u' du\right)^{-1}\int_0^1 F_u dV_u'\Omega_0^{-1}\alpha_0(\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}\alpha_0'
\end{aligned}
$$

Therefore by joint convergence, and as $\frac{1}{T}\sum_{t=1}^{T}\hat{\epsilon}_{0t}\hat{\epsilon}_{0t}' \xrightarrow{P} \Omega_0$, $\breve{\Omega} \xrightarrow{P} \Omega_0$,

$$
\begin{aligned}
&-2\log Q(H_{\text{aux}}|\,H(r)) = -T\log\left|\breve{\Omega}^{-1}\hat{\Omega}\right| = Ttr\left\{\Omega_0^{-1}\left(Y_T + Y_T' - X_T\right)\right\} + o_P(1) \\
&\xrightarrow{D} tr\left\{\Omega_0^{-1}\alpha_0(\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}\alpha_0'\Omega_0^{-1}\int_0^1 dV_u F_u'\left(\int_0^1 F_u F_u' du\right)^{-1}\int_0^1 F_u dV_u'\right\}.
\end{aligned}
$$

For $-2\log Q(H_{\text{aux}}|\,H(p))$, the remaining term in (17), this also implies that,

$$
-2\log Q(H_{\text{aux}}|\,H(p)) \xrightarrow{D} tr\left\{\Omega_0^{-1}\int_0^1 dV_u F_u'\left(\int_0^1 F_u F_u' du\right)^{-1}\int_0^1 F_u dV_u'\right\}.
$$

Hence by the joint convergence of $-2\log Q(H_{\text{aux}}|\,H(p))$ and $-2\log Q(H_{\text{aux}}|\,H(r))$ under $H_0(r)$, and the fact that both distributions are defined in terms of the same underlying Brownian motion,

$$
\begin{aligned}
&-2\log Q(H(r)|\,H(p)) \\
&= -2\log Q(H_{\text{aux}}|\,H(p)) - [-2\log Q(H_{\text{aux}}|\,H(r))] \\
&\xrightarrow{D} tr\left\{\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'\int_0^1 dV_u F_u'\left(\int_0^1 F_u F_u' du\right)^{-1}\int_0^1 F_u dV_u'\right\} \\
&\overset{D}{=} tr\left\{\left(\int_0^1 W_u dW_u'\right)'\left(\int_0^1 W_u W_u' du\right)^{-1}\int_0^1 W_u dW_u'\right\},
\end{aligned}
$$

with $W_u = (\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1/2}V_u$ and the skew projection, $I_p = \alpha_0(\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}\alpha_0'\Omega_0^{-1} + \Omega_0\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'$ has been used. This completes the proof of Lemma 1. ■

## A.2 Asymptotics for the I(2) LR Test

The proof in the I(2) case is analogous to the proof in the I(1) case with the main difference being the more sophisticated parameterization. Again we consider, without loss of generality, the simplest case of the I(2) model, the VAR(2) model as given by

$$\Delta^2 X_t = \Pi X_{t-1} - \Gamma \Delta X_{t-1} + \epsilon_t.$$

The hypothesis of interest is $H(r,s)$ against the alternative $H(p)$. Under $H(r,s)$ we use the parameterization in (8),

$$\Delta^2 X_t = \alpha[\rho'\tau' X_{t-1} + \psi'\Delta X_{t-1}] + \Omega\alpha_\perp(\alpha_\perp'\Omega\alpha_\perp)^{-1}\kappa'\tau'\Delta X_{t-1} + \epsilon_t. \tag{20}$$

We want to derive the asymptotic distribution of the likelihood ratio test,

$$-2\log Q(\,H(r,s)|\,H(p)) = -T\log\left|\breve{\Omega}^{-1}\hat{\Omega}\right|.$$

Proof of Theorem 1: As before, with $\theta$ a parameter, the corresponding true parameter is denoted $\theta_0$. Henceforth, the parameters $\beta$ and $\tau$ under $H(r,s)$ are normalized on $c = \bar{\beta}_0$ and $c = \bar{\tau}_0$ respectively such that $\bar{\beta}_0'\beta = I_r$ and $\bar{\tau}_0'\tau = I_{r+s}$. Furthermore, set $\alpha_\perp = (I_p - \bar{\beta}_0(\alpha'\bar{\beta}_0)^{-1}\alpha')\bar{\beta}_{0\perp}$ such that all other parameters are identified, see Johansen (1997). Note in particular that $\rho = \bar{\tau}_0'\beta$ which is $(r+s)\times r$.

Introduce the parameters defined in Johansen (1997):

$$B_0 = \bar{\beta}_{20}'(\psi - \psi_0), \quad B_1 = \bar{\beta}_{10}'(\beta - \beta_0), \quad B_2 = \bar{\beta}_{20}'(\beta - \beta_0), \quad C = \bar{\beta}_{20}'(\tau - \tau_0)\rho_\perp,$$

where $\rho_\perp = \left(I - \bar{\rho}_0(\rho'\bar{\rho}_0)^{-1}\rho\right)\bar{\rho}_{0\perp}$. Note that $\rho = \bar{\tau}_0'\beta = \rho_0 + \bar{\tau}_0'\beta_{10}B_1 = \rho(B_1)$ and define similarly $\rho_\perp(B_1)$.

Note initially, that from Johansen (1997, Lemma 1) it follows that for the ML estimators $\hat{B}_0, \hat{B}_1, \hat{B}_2$ and $\hat{C}$ derived under $H(r,s)$,

$$T\hat{B}_0 \xrightarrow{D} B_0^\infty, \quad T\hat{B}_1 \xrightarrow{D} B_1^\infty, \quad T^2\hat{B}_2 \xrightarrow{D} B_2^\infty \quad \text{and} \quad T\hat{C} \xrightarrow{D} C^\infty, \tag{21}$$

under $H_0(r,s)$. Here

$$B^\infty = (B_0^{\infty\prime}, B_1^{\infty\prime}, B_2^{\infty\prime})' = \left(\int_0^1 H_u H_u' du\right)^{-1}\int_0^1 H_u dV_{1u}' \tag{22}$$

$$C^\infty = \left(\int_0^1 H_{0u}H_{0u}' du\right)^{-1}\int_0^1 H_{0u}dV_{2u}', \tag{23}$$

where

$$H_u = \begin{pmatrix} H_{0u} \\ H_{1u} \\ H_{2u} \end{pmatrix} = \begin{pmatrix} \beta_2' C_2 V_u \\ \beta_1' C_1 V_u \\ \beta_2' C_2 \int_0^u V_s ds \end{pmatrix},$$

with $V_u$ a Brownian motion on $u \in [0,1]$ with covariance $\Omega_0$. Furthermore,

$$V_{1u} = \left(\alpha_0'\Omega_0^{-1}\alpha_0\right)^{-1}\alpha_0'\Omega_0^{-1}V_u \tag{24}$$

$$\begin{aligned} V_{2u} &= \left(\bar{\rho}_{0\perp}'\kappa_0\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\kappa_0'\bar{\rho}_{0\perp}\right)^{-1}\bar{\rho}_{0\perp}'\kappa_0\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'V_u \\ &= -\left(\xi_0'\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0\right)^{-1}\xi_0'\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'V_u, \end{aligned} \tag{25}$$

where the definition that $\xi = -\kappa'\bar{\rho}_\perp$ has been used, see Section 3.1. Now under $H(r,s)$ the model in (20) can be rewritten as

$$\Delta^2 X_t = A_0 Z_{0t} + A_1 Z_{1t} + A_2 Z_{2t} + \epsilon_t \tag{26}$$

where $Z_{0t}, Z_{1t}$ and $Z_{2t}$ are I(0), I(1) and I(2) regressors respectively, defined by

$$Z_{0t} = \begin{pmatrix} \beta_0' X_{t-1} + \psi_0' \Delta X_{t-1} \\ \tau_0' \Delta X_{t-1} \end{pmatrix}, \quad Z_{1t} = \begin{pmatrix} \beta_{20}' \Delta X_{t-1} \\ \beta_{10}' X_{t-1} \end{pmatrix} \quad \text{and} \quad Z_{2t} = \beta_{20}' X_{t-1}.$$

Finally,

$$A_0 = \left(\alpha, \alpha\left(\psi - \psi_0\right)' \bar{\tau}_0 + \Omega\alpha_\perp \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1} \kappa'\right) \tag{27}$$

$$A_1 = \left(\alpha B_0' + \Omega\alpha_\perp \left(\alpha_\perp' \Omega \alpha_\perp\right)^{-1} \kappa' \left[\bar{\rho}_\perp \left(B_1\right) C' + \bar{\rho}\left(B_1\right) B_2'\right], \alpha B_1'\right) \tag{28}$$

$$A_2 = \alpha B_2'. \tag{29}$$

Introduce next the auxiliary hypothesis, $H_{\text{aux}}$ where $\psi$ $(p\times r)$, $\beta$ $(p\times r)$ and $\tau$ $(p\times(r+s))$ are fixed at their true values, $\psi_0$, $\beta_0$ and $\tau_0$, corresponding to $B_0, B_1, B_2$ and $C$ all identically zero. Note that under $H_{\text{aux}}$, the model equation in (26) reduces to

$$\Delta^2 X_t = A_0 Z_{0t} + \epsilon_t,$$

and furthermore that $H_{\text{aux}} \subseteq H(r,s) \subseteq H(p)$. Hence,

$$-2\log Q(H(r,s)|\, H(p)) = -2\log Q(H_{\text{aux}}|\, H(p)) - \left[-2\log Q(H_{\text{aux}}|\, H(r,s))\right].$$

Turn first to $-2\log Q(H_{\text{aux}}|\, H(r,s))$ and define the corresponding estimated residuals,

$$\hat{\epsilon}_t = \Delta^2 X_t - \hat{A}_0 Z_{0t} - \hat{A}_1 Z_{1t} - \hat{A}_2 Z_{2t}, \quad \breve{\epsilon}_t = \Delta^2 X_t - \breve{A}_0 Z_{0t} \quad \text{and} \quad \hat{\epsilon}_{0t} = \Delta^2 X_t - \hat{A}_0 Z_{0t}.$$

Then $\breve{\Omega} = \frac{1}{T}\sum_{t=1}^T \breve{\epsilon}_t \breve{\epsilon}_t'$, and

$$\hat{\Omega} = \frac{1}{T}\sum_{t=1}^T \hat{\epsilon}_t \hat{\epsilon}_t' = \frac{1}{T}\sum_{t=1}^T \hat{\epsilon}_{0t} \hat{\epsilon}_{0t}' + X_T - Y_T - Y_T',$$

where

$$\begin{aligned} X_T &= \left(\hat{A}_1, \hat{A}_2\right) \left(\frac{1}{T}\sum_{t=1}^T \left(Z_{1t}', Z_{2t}'\right)' \left(Z_{1t}', Z_{2t}'\right)\right) \left(\hat{A}_1, \hat{A}_2\right)' \\ Y_T &= \left(\frac{1}{T}\sum_{t=1}^T \left(\Delta X_t - \hat{A}_0 Z_{0t}\right) \left(Z_{1t}', Z_{2t}'\right)\right) \left(\hat{A}_1, \hat{A}_2\right)'. \end{aligned}$$

With $D_T =$blockdiag$\left(\frac{1}{\sqrt{T}} I_{p-r}, \frac{1}{T^{3/2}} I_{p-r-s}\right)$ it follows that

$$D_T' \left(\frac{1}{T}\sum_{t=1}^T \left(Z_{1t}', Z_{2t}'\right)' \left(Z_{1t}', Z_{2t}'\right)\right) D_T \xrightarrow{D} \int_0^1 H_u H_u' du. \tag{30}$$

Next, using (21) together with the definitions of $A_i$ in (27)-(29) and consistency of the remaining parameters, it follows that

$$\sqrt{T}D_T^{-1}\left(\hat{A}_1,\hat{A}_2\right)\overset{D}{\to}\alpha_0\left(B_0^{\infty\prime},B_1^{\infty\prime},B_2^{\infty\prime}\right)-\left(\Omega_0\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0C^{\infty\prime},0,0\right). \tag{31}$$

Combining (30) and (31) and using the definitions (22)-(23) it follows that

$$TX_T\overset{D}{\to}X^\infty=\left[\alpha_0\left(B_0^{\infty\prime},B_1^{\infty\prime},B_2^{\infty\prime}\right)'-\left(\Omega_0\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0C^{\infty\prime},0,0\right)\right]\times$$
$$\left[\int_0^1 H_uH_u'du\right]\left[\alpha_0\left(B_0^{\infty\prime},B_1^{\infty\prime},B_2^{\infty\prime}\right)'-\left(\Omega_0\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0C^{\infty\prime},0,0\right)\right]'$$

Note that, as $\alpha_0'\alpha_{0\perp}=0$ by definition,

$$tr\left\{\Omega_0^{-1}X^\infty\right\}=tr\left\{\alpha_0\int_0^1 dV_{1u}H_u'\left(\int_0^1 H_uH_u'du\right)^{-1}\int_0^1 H_udV_{1u}'\alpha_0'\right\}$$
$$+\,tr\left\{\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0\int_0^1 dV_{2u}H_u'\left(\int_0^1 H_uH_u'du\right)^{-1}\int_0^1 H_udV_{2u}'\xi_0'\right\},$$

that is, the cross product terms vanish. Next, by convergence to stochastic integrals,

$$TY_T=T\left(\frac{1}{T}\sum_{t=1}^T\left(\Delta X_t-\hat{A}_0Z_{0t}\right)\left(Z_{1t}',Z_{2t}'\right)D_T\right)TD_T^{-1}\left(\hat{A}_1,\hat{A}_2\right)'$$
$$=\left(\sum_{t=1}^T\frac{1}{\sqrt{T}}\epsilon_t\left(Z_{1t}',Z_{2t}'\right)D_T\right)\sqrt{T}D_T^{-1}\left(\hat{A}_1,\hat{A}_2\right)'+o_P(1)\overset{D}{\to}$$
$$\int_0^1 dV_uH_u'\left[\alpha_0\left(B_0^{\infty\prime},B_1^{\infty\prime},B_2^{\infty\prime}\right)-\left(\Omega_0\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0C^{\infty\prime},0,0\right)\right]'$$
$$\overset{D}{=}\int_0^1 dV_uH_u'\left(\int_0^1 H_uH_u'du\right)^{-1}\int_0^1 H_udV_{1u}'\alpha_0'-$$
$$\int_0^1 dV_uH_{0u}'\left(\int_0^1 H_{0u}H_{0u}'du\right)^{-1}\int_0^1 H_{0u}dV_{2u}'\xi_0'\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'\Omega_0.$$

Therefore by joint convergence, and as $\frac{1}{T}\sum_{t=1}^T\hat{\epsilon}_{0t}\hat{\epsilon}_{0t}'\overset{P}{\to}\Omega_0$, $\breve{\Omega}\overset{P}{\to}\Omega_0$, and furthermore using the definition of $V_1$ and $V_2$ in (24)-(25),

$$-2\log Q(H_{\text{aux}}|\,H(r,s))=-T\log\left|\breve{\Omega}^{-1}\hat{\Omega}\right|$$
$$\overset{D}{\to}tr\left\{\Omega_0^{-1}\alpha_0(\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}\alpha_0'\Omega_0^{-1}\int_0^1 dV_uH_u'\left(\int_0^1 H_uH_u'du\right)^{-1}\int_0^1 H_udV_u'\right\}$$
$$+\,tr\left\{\left[\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0\left(\xi_0'\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0\right)^{-1}\xi_0'\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'\right]\times\right.$$
$$\left.\int_0^1 dV_uH_{0u}'\left(\int_0^1 H_{0u}H_{0u}'du\right)^{-1}\int_0^1 H_{0u}dV_u'\right\}.$$

This implies directly that also

$$-2\log Q(H_{\text{aux}}|\,H(p)) \xrightarrow{D} tr\left\{\Omega_0^{-1}\int_0^1 dV_u H_u'\left(\int_0^1 H_u H_u' du\right)^{-1}\int_0^1 H_u dV_u'\right\}.$$

Using the projection $I_p = \alpha_0(\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}\alpha_0'\Omega_0^{-1} + \Omega_0\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'$ and collecting terms, it follows that

$$\begin{aligned}
&-2\log Q(H(r,s)|\,H(p)) \\
&= -2\log Q(H_{\text{aux}}|\,H(p)) - [-2\log Q(H_{\text{aux}}|\,H(r,s))] \\
&\xrightarrow{D} tr\left\{\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'\int_0^1 dV_u H_u'\left(\int_0^1 H_u H_u' du\right)^{-1}\int_0^1 H_u dV_u'\right\} - \qquad (32)\\
&tr\left\{\left[\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0\left(\xi_0'\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0\right)^{-1}\xi_0'\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'\right]\times\right. \qquad (33)\\
&\left.\int_0^1 dV_u H_{0u}'\left(\int_0^1 H_{0u}H_{0u}' du\right)^{-1}\int_0^1 H_{0u} dV_u'\right\}. \qquad (34)
\end{aligned}$$

The term in (32) can be rewritten as

$$\begin{aligned}
&tr\left\{\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'\int_0^1 dV_u H_{0u}'\left(\int_0^1 H_{0u}H_{0u}' du\right)^{-1}\int_0^1 H_{0u} dV_u'\right\} + \\
&tr\left\{\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'\int_0^1 dV_u G_u'\left(\int_0^1 G_u G_u' du\right)^{-1}\int_0^1 G_u dV_u'\right\}
\end{aligned}$$

where

$$G_u = \left(\begin{array}{c|c} H_{1u} & \\ H_{2u} & H_0\end{array}\right) = \left(\begin{array}{c} H_{1u} - \int_0^1 H_{1u}H_{0u}'du\left[\int_0^1 H_{0u}H_{0u}'du\right]^{-1}H_{0u}' \\ H_{2u} - \int_0^1 H_{2u}H_{0u}'du\left[\int_0^1 H_{0u}H_{0u}'du\right]^{-1}H_{0u}'\end{array}\right).$$

Note next that the term appearing in (33) can be written as

$$\begin{aligned}
&\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0\left(\xi_0'\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\xi_0\right)^{-1}\xi_0'\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}' \\
&= \alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}' - \alpha_{0\perp}\xi_{0\perp}\left(\xi_{0\perp}'\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\xi_{0\perp}\right)^{-1}\xi_{0\perp}'\alpha_{0\perp}'.
\end{aligned}$$

Hence,

$$\begin{aligned}
&-2\log Q(H(r,s)|\,H(p)) \xrightarrow{D} \\
&tr\left\{\alpha_{0\perp}\left(\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\right)^{-1}\alpha_{0\perp}'\int_0^1 dV_u G_u'\left(\int_0^1 G_u G_u' du\right)^{-1}\int_0^1 G_u dV_u'\right\} + \\
&tr\left\{\alpha_{20}\left(\alpha_{20}'\Omega_0\alpha_{20}\right)^{-1}\alpha_{20}'\int_0^1 dV_u H_{0u}'\left(\int_0^1 H_{0u}H_{0u}' du\right)^{-1}\int_0^1 H_{0u} dV_u'\right\},
\end{aligned}$$

and the result in Theorem 1 follows by defining the standard Brownian motion,

$$W_u = (\alpha'_{0\perp}\Omega_0\alpha_{0\perp})^{-1/2}\,\alpha'_{0\perp}V_u.$$

■

# References


Banerjee, A., L. Cockerell, and B. Russell (2001): "An I(2) Analysis of Inflation and the Markup," *Journal of Applied Econometrics*, 16, 221–240.

Banerjee, A., and B. Russell (2001): "The Relationship Between the Markup and Inflation in the G7 Economies and Australia," *The Review of Economics and Statistics*, 83(2), 377–387.

Boswijk, H. P. (2000): "Mixed Normality and Ancillarity in I(2) Systems," *Econometric Theory*, 16, 878–904.

Diamandis, P., D. Georgoutsos, and G. Kouretas (2000): "The Monetary Model in the Presence of I(2) Components: Long-Run Relationships, Short-Run Dynamics and Forecasting of the Greek Drachma," *Journal of International Money and Finance*, 19, 917–941.

Doornik, J. A. (1998): "Approximations to the Asymptotic Distribution of Cointegration Tests," *Journal of Economic Surveys*, 12(5), 573–593.

Fliess, N., and R. MacDonald (2001): "The Instability of the Money Demand Function: and I(2) Interpretation," *Oxford Bulletin of Economics and Statistics*, 63(4), 475–495.

Haldrup, N. (1998): "A Review of the Econometric Analysis of I(2) Variables," *Journal of Economic Surveys*, 12, 595–650.

Johansen, S. (1992): "A Representation of Vector Autoregressive Processes Integrated of Order 2," *Econometric Theory*, 8, 188–202.

——— (1995): "A Statistical Analysis of Cointegration for I(2) Variables," *Econometric Theory*, 11, 25–59.

——— (1996): *Likelihood-Based Inference in Cointegrated Autoregressive Models*. Oxford University Press, Oxford, 2nd edn.

——— (1997): "Likelihood Analysis of the I(2) Model," *Scandinavian Journal of Statistics*, 24, 433–462.

——— (2002): "A Small Sample Correction for the Test of Cointegration Rank in the Vector Autoregressive Model," *Econometrica*, 70(5), 1929–1961.

Juselius, K. (1998): "A Structured VAR for Denmark under Changing Monetary Regimes," *Journal of Business and Economic Statistics*, 16(4), 400–412.

——— (1999): "Price Convergence in the Medium and Long Run: An I(2) Analysis of Six Price Indices," in *Cointegration, Causality, and Forecasting - A Festschrift in Honour of Clive W.J. Granger*, ed. by R. F. Engle, and H. White. Oxford University Press, Oxford.

Kongsted, H. C. (2003): "An I(2) Cointegration Analysis of Small-Country Import Price Determination," *Econometrics Journal*, 6, 53–71.

Nielsen, B., and A. Rahbek (2000): "Similarity Issues in Cointegration Analysis," *Oxford Bulletin of Economics and Statistics*, 62(1), 5–22.

Nielsen, H. B. (2002): "An I(2) Cointegration Analysis of Price and Quantity Formation in Danish Manufactured Exports," *Oxford Bulletin of Economics and Statistics*, 64(5), 449–472.

Paruolo, P. (1996): "On the Determination of Integration Indices in I(2) Systems," *Journal of Econometrics*, 72, 313–356.

——— (2000): "Asymptotic Efficiency of the Two Stage Estimator in I(2) Systems," *Econometric Theory*, 16, 524–550.

Paruolo, P., and A. Rahbek (1999): "Weak Exogeneity in I(2) VAR Systems," *Journal of Econometrics*, 93, 281–308.

Rahbek, A., H. C. Kongsted, and C. Jørgensen (1999): "Trend-Stationarity in the I(2) Cointegration Model," *Journal of Econometrics*, 90, 265–289.