Econometric Theory, **23**, 2007, 615–637. Printed in the United States of America.
DOI: 10.1017/S0266466607070272

THE LIKELIHOOD RATIO TEST FOR COINTEGRATION RANKS IN THE I(2) MODEL

HEINO BOHN NIELSEN AND ANDERS RAHBEK
University of Copenhagen

This paper presents the likelihood ratio (LR) test for the number of cointegrating relations in the I(2) vector autoregressive model. It is shown that the asymptotic distribution of the LR test for the cointegration ranks is identical to the asymptotic distribution of the much applied test statistic based on the two-step estimation procedure in Johansen (1995, *Econometric Theory* 11, 25–59), Paruolo (1996, *Journal of Econometrics* 72, 313–356), and Rahbek, Kongsted, and Jørgensen (1999, *Journal of Econometrics* 90, 265–289). By construction the LR test statistic is smaller than the non-LR test statistic from the two-step procedure, 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 use of the LR test for rank determination in I(2) analysis.

1. INTRODUCTION

For many OECD countries for the postwar period, the first differences of nominal variables such as inflation rates or money growth seem to be well described as autoregressive processes integrated of order one, I(1). This implies that the levels of the nominal variables are integrated of order two, I(2), as reflected in several empirical applications: see, e.g., the analyses of money demand and open economy price determination in, inter alia, Juselius (1998, 1999), Diamandis, Georgoutsos, and Kouretas (2000), Banerjee, Cockerell, and Russell (2001), Banerjee and Russell (2001), Fliess and MacDonald (2001), Nielsen (2002), Kongsted (2003), and Bacchiocchi and Fanelli (2005).

For the multivariate p-dimensional vector autoregressive (VAR) model with no deterministic terms, Johansen (1992) showed that two reduced rank restric-

Discussions with Søren Johansen, Hans Christian Kongsted, and Bent Nielsen are gratefully acknowledged. We also thank the editor and three anonymous referees for very constructive comments that have led to a much improved version of the paper. This research was supported by Danish Social Sciences Research Council grant 2114-04-0001. Address correspondence to Heino Bohn Nielsen, Department of Economics, University of Copenhagen, Studiestræde 6, DK-1455 Copenhagen K; e-mail: heino.bohn.nielsen@econ.ku.dk.

© 2007 Cambridge University Press 0266-4666/07 $15.00

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

tions imposed allow for I(2), I(1), and also I(0) variables. Moreover, the variables in levels and first differences may be cointegrated corresponding to the cointegration ranks, or the integration indices, (r, s). The rank r determines the number of stationary relations, whereas the rank s determines the number of linear combinations of the variables in levels that are I(1). The remaining dimension, $p - r - s$, determines the number of common I(2) trends in the system. The analysis was extended to include deterministic terms such as the linear trend in Rahbek, Kongsted, and Jørgensen (1999) and constant terms in Paruolo (1996). See also Johansen (1995, 1997, 2006), Paruolo and Rahbek (1999), Kurita (2006), and Paruolo (2000) and the survey by Haldrup (1998) for further references on I(2).

Inference on the cointegration ranks (r, s) is crucial for empirical applications, and the main contribution of this paper is to present the likelihood ratio (LR) test for the number of cointegration ranks in I(2) VAR models and derive the asymptotic distributions. The LR test is new, as existing literature has applied a non-LR test, referred to as the two-step test here. The two-step test is based on the sum of two sets of canonical correlations from a non–maximum likelihood (non-ML) estimation procedure in two steps for I(2) VAR models; see Johansen (1995), Paruolo (1996), and Rahbek et al. (1999). The two-step-based test was developed as an approximation to the LR test that is straightforward to implement as a sequential application of reduced rank regression (RRR) well known from I(1) analysis. The LR test for finite samples is smaller than (or equal to) the corresponding two-step test and decreases rejection frequencies for small samples; see the included examples in Section 4.2 where the difference between the LR and the two-step-based statistic is quite large. Regarding asymptotics, we find that the asymptotic distributions of the LR tests are identical to the two-step-based tests' asymptotic distributions, and hence existing critical values published for the two-step tests apply to LR tests also. Based on a study of existing empirical examples in the literature and related Monte Carlo simulations we examine some further finite-sample properties of the tests. In particular, size properties of the proposed LR test are excellent.

By definition, the LR test is based on ML estimation of the parameters in the I(2) models. ML inference with known ranks, (r, s), was proposed by Johansen (1997) and Boswijk (2000) based on different parametrizations. In line with the results for the two-step analysis in Rahbek et al. (1999), ML inference has been extended to the linear trend case in Rahbek and Nielsen (2006). When deriving the limiting distributions of the LR test statistics, we apply the parametrizations of Johansen (1997) and Rahbek and Nielsen (2006). Note that the LR test and general ML estimation of the I(2) models already are implemented in the new version of CATS in RATS; see Dennis (2006). Ox code for calculating the LR test for the cointegration ranks can be obtained from the authors.

The rest of the paper is organized as follows. Section 2 presents the I(2) models and the LR test in addition to a brief review of the two-step test statistics. The asymptotic distributions of the LR test, and their use, are considered

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

in Section 3. Section 4 investigates small-sample properties of the rank tests based on empirical examples and Monte Carlo simulations.

Throughout, use is made of the following notation: for any $p \times r$ matrix α 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(α) and let $\bar{\alpha} = \alpha(\alpha'\alpha)^{-1}$. 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. MODELS, REPRESENTATION, AND ESTIMATION

This section reviews parametrization, representation, and estimation for I(2) models based on Johansen (1995), Paruolo (1996), and Rahbek et al. (1999), where more details can be found.

2.1. Model with No Deterministic Terms

Consider first the p-dimensional VAR model of order k with no deterministic terms as given by

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

Here Π, Γ, and Ψ_i are $p \times p$ matrices, ϵ_t is a p-dimensional independent and identically distributed (i.i.d.) $N_p(0, \Omega)$ sequence, $\Omega > 0$, and the initial values $X_{-k+1}, \ldots, X_0$ are fixed for the statistical analysis. The I(2) model, denoted $H(r, s)$, is then defined by two reduced rank restrictions given by

$$\Pi = \alpha\beta', \tag{2}$$

$$\alpha_\perp' \Gamma \beta_\perp = \xi\eta', \tag{3}$$

where α and β are $p \times r$ matrices and ξ and η are $(p - r) \times s$ matrices with $r \leq p$ and $s \leq p - r$. We use the notation $H(r) = H(r, p - r)$ to denote I(1) models, and $H(p)$ denotes the unrestricted vector autoregression.

2.1.1. Representation. To state a representation for X_t, denote by $H_0(r, s)$ the assumption that the characteristic polynomial, $A(z) = I_p(1 - z)^2 - \Pi z + \Gamma(1 - z)z - \Psi_1(1 - z)^2 z - \cdots - \Psi_{k-2}(1 - z)^2 z^{k-2}$, 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 imply reduced rank $r < p$ in (2) and reduced rank $s < p - r$ in (3).

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

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

$$X_t = C_2 \sum_{s=1}^{t} \sum_{i=1}^{s} \epsilon_i + C_1 \sum_{i=1}^{t} \epsilon_i + \gamma_c + \gamma_l t + X_{st}, \tag{4}$$

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

By Johansen (1992), X_{st} is a stationary mean-zero I(0) process, and the notation $\alpha_1 = \bar{\alpha}_\perp \xi$, $\beta_1 = \bar{\beta}_\perp \eta$, $\alpha_2 = (\alpha, \alpha_1)_\perp$, $\beta_2 = (\beta, \beta_1)_\perp$, and $\Theta = \Gamma\bar{\beta}\bar{\alpha}'\Gamma + I_p - \sum_{i=1}^{k-2}\Psi_i$ has been used. The coefficients γ_c and γ_l depend on the initial values of the process and satisfy $(\beta, \beta_1)'\gamma_l = 0$ and $\beta'\gamma_c - \delta\beta_2'\gamma_l = 0$, where $\delta = \bar{\alpha}'\Gamma\bar{\beta}_2$. It follows from (4) 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, there are r mean-zero stationary linear combinations,

$$\beta' X_t - \delta\beta_2' \Delta X_t. \tag{5}$$

2.1.2. Variation-Free Parametrization. To maximize the likelihood function for known ranks (r,s) under the hypothesis $H(r,s)$, Johansen (1997) suggested a reparametrization of the model based on the reduced ranks in (2) and (3) and also on the cointegration term in (5). The reparametrization is given by

$$\Delta^2 X_t = \alpha[\rho'\tau' X_{t-1} + \psi'\Delta X_{t-1}] + \tilde{\alpha}_{\perp\Omega}\kappa'\tau'\Delta X_{t-1} + \sum_{i=1}^{k-2}\Psi_i \Delta^2 X_{t-i} + \epsilon_t, \tag{6}$$

where $\tilde{\alpha}_{\perp\Omega} = \Omega\alpha_\perp(\alpha_\perp'\Omega^{-1}\alpha_\perp)^{-1}$. Furthermore, α, ρ $((r+s)\times r)$, τ $(p \times (r+s))$, ψ $(p \times r)$, κ $((r+s)\times(p-r))$, Ψ_i $(p \times p)$, $i = 1, \ldots, k-2$, and Ω are all freely varying parameters. The parameters in the previous formulation in (1)–(3) can be derived from the new parameters by the identities $\tau = (\beta, \beta_1)$, $\beta = \tau\rho$, $\tilde{\alpha}_{\Omega^{-1}} = \Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1}$, $\psi' = -\tilde{\alpha}_{\Omega^{-1}}'\Gamma$, and $\kappa' = -\alpha_\perp'\Gamma(\bar{\beta}, \bar{\beta}_1) = -(\alpha_\perp'\Gamma\bar{\beta}, \xi)$, using the skew-projection identity,

$$\alpha\tilde{\alpha}_{\Omega^{-1}} + \tilde{\alpha}_{\perp\Omega}\alpha_\perp' = I_p. \tag{7}$$

No closed form solution for the ML estimators of the parameters in (6) exists, but estimates are obtained by a switching algorithm: for fixed τ, the parameters $\alpha_\perp$ and α can be obtained by solving an eigenvalue problem, and the remaining parameters can be found from regression. For fixed values of these parameters, τ 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 implementations we use the two-step estimators, reviewed next, as starting values for the ML iterations.

In the two-step estimation, first the parameters α and β in (2) are estimated by RRR as applied in I(1) models ignoring the reduced rank restriction (3). In the second step, ξ and η in (3) are estimated by RRR with r and the estimated α and β from the first step fixed; see Johansen (1995).

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

2.2. Models with Linear Trends and Deterministics

It is often important to include deterministic linear trends and a constant level in empirical applications. Rahbek et al. (1999) propose an I(2) VAR model that allows deterministic linear trends and a constant level in all components and linear combinations of the process. This includes the cointegrating relations (5) and avoids at the same time the possibility of quadratic trends. In terms of the parameterization in (6), this model, denoted $H^*(r,s)$, can be represented as

$$\Delta^2 X_t = \alpha[\rho'\tau^{*\prime}X^*_{t-1} + \psi^{*\prime}\Delta X^*_{t-1}] + \widetilde{\alpha}_{\perp\Omega}\kappa'\tau^{*\prime}\Delta X^*_{t-1} + \sum_{i=1}^{k-2}\Psi_i\Delta^2 X_{t-i} + \epsilon_t, \quad \textbf{(8)}$$

where $\tau^* = (\tau', \tau_l)'$ ($(p+1)\times(r+s)$), $\psi^* = (\psi', \psi_c)'$ ($(p+1)\times r$), and, finally, $X^*_t = (X'_t, t)'$ is $p+1$-dimensional. As in Section 2.1.1, 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 (8) has the representation in (4) but with γ_l and γ_c being functions of τ_l and ψ_c in addition to initial values. In particular, $\Delta^2 X_t$,

$$\rho'\tau^{*\prime}X^*_t + \psi^{*\prime}\Delta X^*_t = \beta' X_t + \psi'\Delta X_t + \rho'\tau_l t + \psi_c, \quad \text{and}$$

$$\tau^{*\prime}\Delta X^*_t = (\beta, \beta_1)'\Delta X_t + \tau_l$$

are mean-zero stationary processes. As a result, X_t is as emphasized an I(2) process with linear trends in all linear combinations of the process. With X^*_t and ΔX^*_t replacing X_t and ΔX_t, respectively, ML estimation of the parameters under $H^*(r,s)$ is identical to the case of no deterministic terms. Asymptotic distributions of the ML estimators can be found in Rahbek and Nielsen (2006), and two-step-based inference is discussed in Rahbek et al. (1999).

Further classes of I(2) models with varying deterministic specifications have been considered in the literature, such as, e.g., the model with a constant level in all linear combinations of the process, including the cointegrating. In this case the model can be represented as in (8), with $X^*_t = (X'_t, 1)'$ and $\Delta X^*_t = \Delta X_t$. Models with this and also with different specifications of the parameter of the constant regressor are considered in Paruolo (1996, 2000).

3. RANK TESTING IN I(2) MODELS

This section analyzes and proposes the LR test for cointegration ranks in the I(2) VAR models using the parametrizations of Johansen (1997) and Rahbek and Nielsen (2006). Asymptotic distributions of the LR test statistics under the I(2) hypotheses $H_0(r,s)$ and $H^*_0(r,s)$ are addressed in Theorems 1 and 2, which follow. Also the conventionally applied two-step rank test statistics in Johansen (1995) and Rahbek et al. (1999) are briefly reviewed as they are used for comparison. Finally implementation of rank determination based on sequential testing is briefly discussed.

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

3.1. Test Statistics

For rank determination we propose to use the LR test for $H(r,s)$ against the unrestricted alternative, $H(p)$, as defined by

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

where $\check{\Omega}$ and $\hat{\Omega}$ denote the covariance matrices estimated under $H(r,s)$ and $H(p)$, respectively. The conventionally applied two-step rank test statistic is the sum of two test statistics:

$$S_{r,s}^{2S} = Q_r + Q_{r,s}, \tag{10}$$

where $Q_r = -T\sum_{i=r+1}^{p}\log(1-\lambda_i)$ and $Q_{r,s} = -T\sum_{i=s+1}^{p-r}\log(1-\zeta_i)$. Here λ_i are the $(p-r)$ smallest eigenvalues from the RRR in the first step, and ζ_i are the $(p-r-s)$ smallest eigenvalues from the RRR in the second step of the two-step estimation, corresponding to the reduced ranks of Π and $\alpha_\perp'\Gamma\beta_\perp$ in (2) and (3), respectively. Likewise for $H^*(r,s)$ and $H^*(p)$, in which case the LR test is denoted $S_{r,s}^{LR^*}$ and the two-step test is denoted $S_{r,s}^{2S^*}$.

3.2. Asymptotic Distributions of the LR Tests

In this section the asymptotic distributions are given for the LR tests of cointegration ranks in the I(2) models $H(r,s)$ and $H^*(r,s)$. To present the limiting distributions define for X_u, Y_u, and Z_u of dimension p_x, p_y, and p_z on the unit interval $u \in [0,1]$,

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

$$S(X,Y,Z) = \int_0^1 dXY'\left(\int_0^1 Y_u Y_u'\,du\right)^{-1}\int_0^1 YdZ'. \tag{11}$$

Consider first the hypothesis $H(r,s)$ against the unrestricted $H(p)$.

THEOREM 1. *Under $H_0(r,s)$ (cf. Section 2.1.1), 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

$$Q_r^\infty = \mathrm{tr}\{S(W,G,W)\} \quad \textit{and} \quad Q_{r,s}^\infty = \mathrm{tr}\{S(W_2,W_2,W_2)\}.$$

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,

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

$$G_u = \left(\left(W_{1u}', \int_0^u W_{2v}' \, dv \right)' | W_2 \right).$$

The proof is given in Appendix A. 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, Thm. 7). In particular, the two components, Q_r^∞ and $Q_{r,s}^\infty$, have the asymptotic distributions of the statistics from the first and second steps 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.

The result in Theorem 1 is extended to the model $H^*(r,s)$ in Theorem 2. The proof given in Appendix B is based on asymptotic theory in Rahbek and Nielsen (2006) for the ML estimators of the parameters in $H^*(r,s)$. An important implication is that the LR test and 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. Critical values are given in, inter alia, Rahbek et al. (1999) and Doornik (1998).

THEOREM 2. *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 $Q_r^{*\infty} = \mathrm{tr}\{S(W, G_1, W)\}$ and $Q_{r,s}^{*\infty} = \mathrm{tr}\{S(W_2, G_2, W_2)\}$. The Brownian motion W_u is defined in Theorem 1, and

$$G_{1u} = \left(\left(W_{1u}', \int_0^u W_{2v}' \, dv, u' \right) | W_2 \right) \quad \text{and} \quad G_{2u} = (W_{2u}', 1)'.$$

The results in Theorems 1 and 2 can be extended to further classes of I(2) models. Specifically, Paruolo (2000) showed that the two-step estimators are asymptotically efficient for a wide class of models including the case of a constant level in all linear combinations of the process. Hence mimicking the proofs of Theorems 1 and 2 and using Paruolo (1996), the limiting distribution of the LR test for cointegration ranks is conjectured to be the same as reported in Theorem 2, but with $G_1 = (W_{1u}', \int_0^u W_{2v}' \, dv, u, 1)' | W_2)$ and $G_2 = W_{2u}$.

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, and we discuss rank selection for the model with no deterministic terms for simplicity.

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 $(\hat{r}, \hat{s})$ of the ranks

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

TABLE 1. Partial nesting structure for $p = 4$

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

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

$$(\hat{r}, \hat{s}) = \{(r,s) | S_{r,s} \leq c_{r,s}; S_{r_0,s_0} > c_{r_0,s_0} \text{ for } (r_0 < r, s_0 \leq p - r_0) \quad \text{and} \quad (r_0 = r, s_0 < s)\},$$

where $S_{r,s}$ denote either of the test statistics and $c_{r,s}$ their asymptotic critical values. The estimator $(\hat{r}, \hat{s})$ will select the correct ranks with a limiting probability $(1 - \pi)$, where π is the size of each test in the sequence as discussed in Johansen (1995).

4. 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 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.

4.1. The LR and Two-Step Statistics 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 for the case of $H(r,s)$. Note that the same result holds for $H^*(r,s)$.

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

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. First note that the LR test, $S_{r,s}^{LR}$, is given in (9) as $-T\log|\breve{\Omega}^{-1}\hat{\Omega}|$, where $\breve{\Omega}$ and $\hat{\Omega}$ denote the covariance matrices found by ML estimation under $H(r,s)$ and $H(p)$, respectively. The two-step-based test, $S_{r,s}^{2S}$, in (10) can be rewritten as

$$S_{r,s}^{2S} = Q_r + Q_{r,s} = -T\log|\breve{\Omega}^{-1}\hat{\Omega}| + (-T\log|\ddot{\Omega}^{-1}\breve{\Omega}|) = -T\log|\ddot{\Omega}^{-1}\hat{\Omega}|,$$

where $\ddot{\Omega}$ is the two-step-based estimator of Ω under $H(r,s)$ and $\breve{\Omega}$ is the (ML) estimator of Ω under $H(r) = H(r,p-r)$. Thus, $S_{r,s}^{LR} \leq S_{r,s}^{2S}$ is equivalent to $|\breve{\Omega}| \leq |\ddot{\Omega}|$. This follows by showing that (i) the two-step-based estimators of the parameters are not ML estimators and, at the same time, (ii) the parameters of the two-step parametrization lie in the parameter set over which the ML estimation is performed. To see this, note that the estimator $\breve{\Omega}$ is found by maximizing the likelihood function as defined by (6), with freely varying parameters in the parametrization as given by $\theta_{LR} := (\alpha, \rho, \tau, \psi, \kappa, \Psi_i, \Omega)$. The estimator $\ddot{\Omega}$ is best discussed in the variation-dependent parametrization given by $\theta_{2S} := (\alpha, \beta, \xi, \eta, \alpha_\perp'\Gamma\bar{\beta}, \bar{\alpha}'\Gamma, \Psi_i, \Omega)$. Fixing β and α at their first step RRR values, $\ddot{\Omega}$ is found by maximizing the likelihood function over the remaining parameters in θ_{2S}. Setting $\alpha = \alpha$, $\tau = (\beta, \bar{\beta}_\perp\eta)$, $\rho = (I,0)'$, $\Omega = \Omega$, $\psi' = -\tilde{\alpha}_{\Omega^{-1}}'\Gamma$, $\kappa' = -(\alpha_\perp'\Gamma\bar{\beta}, \xi)$, and $\Psi_i = \Psi_i$, this corresponds to a specific parameter point in θ_{LR}, and, by definition, $|\breve{\Omega}| \leq |\ddot{\Omega}|$. Note that in two special cases, the two-step procedure actually maximizes the likelihood function under $H(r,s)$: first, in I(1) models, where $s = p-r$ and $\alpha_\perp'\Gamma\beta_\perp$ is nonsingular; and second, if $r = 0$ where the second step of the two-step procedure is conditional on $\alpha = \beta = 0$. ■

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, whereas the type II error probabilities, for tests of more restricted models, are potentially larger.

Note that the magnitude of the difference depends on the number of restrictions in (3), which is ignored in the first step of the two-step estimator. The number of ignored restrictions is largest for models on the diagonal of Table 1, where $s = 0$, and the LR and the two-step statistics coincide for models located in the first row and last column of Table 1. Also note that in the model $H^*(r,s)$ a restriction on the parameter of the constant regressor is ignored in the first

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

step of the two-step procedure. Finally, note that in Proposition 1 the special case of $r(p-r-s)=0$ is implied by either no I(2) trends ($p-r-s=0$) or by no I(0) relations ($r=0$). Under the assumption of I(2), the case of $r=0$ implies that Γ has reduced rank $s<p$ and can be factorized as $\Gamma=\xi\eta'$, and hence this case mimics the statistical analysis of an I(1) VAR model in first differences.

4.2. Empirical Study and Monte Carlo Simulations

To analyze the small-sample properties and compare the LR and two-step tests, three approaches were taken. First, we revisited the rank determination in a number of published I(2) applications in the literature using both types of statistics. We reconsidered the money demand analyzed by I(2) VAR models in, inter alia, Johansen (1992) and Rahbek et al. (1999), the domestic pricing behavior in Banerjee et al. (2001), the Danish export pricing behavior in Nielsen (2002), and the Danish import price determination in Kongsted (2003). Next, Monte Carlo simulations were performed with the estimated models as data generating processes (DGPs); i.e., for each empirical example a DGP was defined as the model $H^*(r,s)$ with parameter values set to the ML estimates. With this choice of DGPs for the Monte Carlo study they were empirically relevant, and we checked that the parameters satisfied the assumptions underlying the theoretical results. Finally, we constructed a larger study based on different 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 above the nominal size. This is particularly the case if $s=0$ in the DGP. In some cases the LR test still overrejects true hypotheses, as is also known from the LR test in I(1) models. If the size distortion of the LR test is deemed important, the challenging task of deriving a Bartlett correction of the LR test can be considered; see, e.g., Johansen (2002) for the I(1) case. To keep the presentation concise and focus on the main points we report results only for the Danish import price determination. The full set of results can be obtained from the authors.[1]

4.2.1. Empirical Illustration. In this section we revisit the rank determination in Kongsted (2003) based on a $p=4$-dimensional set of data including Danish import prices, domestic prices, a measure of competing prices, and an interest rate. Kongsted (2003) estimated a VAR(2) model for the effective sample 1975:3–1995:4 and included a linear trend in all directions, corresponding to the model $H^*(r,s)$, in addition to an unrestricted impulse dummy for the observation 1992:2. Here we adopt the same specification, noting that the

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

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 LR statistics for the cointegration ranks are reported on 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, whereas $H^*(1,2)$ has a p-value of 11%. If this model is nevertheless rejected, the next potential model is $H^*(2,0)$, comprising two stationary relations and two I(2) trends, again with a p-value of around 10%. The test statistic for the theoretically preferred model, $H^*(2,1)$, is 26 corresponding to a p-value of 35%. The LR rank tests thus point toward the models $H^*(1,2)$, $H^*(2,0)$, and $H^*(2,1)$ as potentially relevant for the current data. The eigenvalues of the companion matrices corresponding to these models, i.e., the inverse roots of the characteristic polynomials, are given in Table 3. The processes have $2(p-r)-s$ unit roots, and for all three choices of parameter values the remaining roots seem to be far from the unit circle in accordance with the assumptions for Theorem 2.

The two-step rank test statistics, also reported in Kongsted (2003), are given on the right-hand side of Table 2. As noted, all test statistics in the first row and last column are identical to the LR results, and the remaining test statistics are all higher. Using the two-step test there is marginally more evidence for rejecting $H^*(1,2)$. Furthermore, the model $H^*(2,0)$ is clearly rejected, and the two-step statistic of 75 is markedly higher than the corresponding LR statistic of 45. Thus based on asymptotic inference with a 5% significance level, both statistics lead to different results than expected from economic theory, and the two tests give different indications on the potentially relevant models. In the next section we use Monte Carlo simulations to investigate further the differences of the two tests in particular with respect to size properties.

4.2.2. Simulations Based on the Estimated Model. To analyze the differences, we set up a small Monte Carlo simulation with the DGPs defined by the parameters from ML estimation under $H^*(2,1)$; see Table 4 for specific parameter values. For each simulation i, we generate time series, $X_1^i, \ldots X_T^i$, for T ranging between 50 and 1,000 and by replacing ϵ_t with random draws from an i.i.d. Gaussian distribution with the estimated covariance matrix. Each series is initiated with initial values X_0^i and ΔX_0^i; to ensure that each nonstationary series is initiated the same way across all simulations, the nonstationary I(1) and I(2) directions of the process are initiated from the corresponding observed initial values. That is, with $\beta_2 = \tau_\perp$, $\beta_2' X_0^i = \beta_2' X_0^{\text{obs}}$, $\beta_2' \Delta X_0^i = \beta_2' \Delta X_0^{\text{obs}}$, and, correspondingly, $\beta_1' X_0^i = \beta_1' X_0^{\text{obs}}$. The stationary directions, $\rho' \tau^{*\prime} X_t^{i*} + \psi^{*\prime} \Delta X_t^{i*}$, $\tau^{*\prime} \Delta X_t^{i*}$, and $\beta_2' \Delta^2 X_t^i$, are on the other hand initiated from the invariant distribution by discarding 100 initial simulations of the stationary series, thus eliminating the importance of initial values in the stationary series.

The rejection frequencies at a nominal 5% level are reported in Table 4 for the two tests. The rejection frequencies of $H^*(r,s)$ in $H^*(p)$ are identical for

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

TABLE 2. Rank determination for the data in Kongsted (2003)

r	LR test					Two-step test				
0	343.53 [0.00]	238.76 [0.00]	161.88 [0.00]	122.79 [0.00]	101.00 [0.00]	343.53 [0.00]	238.76 [0.00]	161.88 [0.00]	122.79 [0.00]	101.00 [0.00]
1		142.37 [0.00]	88.22 [0.00]	49.94 [0.11]	48.48 [0.01]		187.53 [0.00]	96.95 [0.00]	51.17 [0.09]	48.48 [0.01]
2			45.36 [0.10]	26.10 [0.35]	24.52 [0.07]			75.15 [0.00]	27.61 [0.27]	24.52 [0.07]
3				11.72 [0.49]	10.42 [0.11]				13.02 [0.37]	10.42 [0.11]
$p - r - s$	4	3	2	1	0	4	3	2	1	0

Note: Figures in square brackets are asymptotic p-values according to the Γ-approximation of Doornik (1998).

TABLE 3. Eigenvalues of the companion matrices corresponding to the models $H^*(1,2)$, $H^*(2,0)$, and $H^*(2,1)$ estimated for the data in Kongsted (2003)

Model	Eigenvalues						
$H^*(1,2)$	1	1	1	1	$0.280 \pm 0.225i$	-0.394	-0.260
$H^*(2,0)$	1	1	1	1	$0.248 \pm 0.289i$	-0.407	-0.190
$H^*(2,1)$	1	1	1	0.513	$0.261 \pm 0.315i$	-0.402	-0.009

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, whereas 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 $s = 0$.

To illustrate the difference when $s = 0$ in the true model, we performed the simulations again taking ML estimation under $H^*(2,0)$ as defining the DGP; see Table 5 for the parameter values. In this case a large distortion appears in the two-step estimation of the model with correct ranks. The rejection frequencies are reported in Table 5. 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 of the two-step test is 66%. The LR tests, on the other hand, have good size properties, with rejection frequencies close to the nominal size.

NOTE

1. See www.econ.ku.dk/okohn.

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

TABLE 4. Rejection frequencies in a simulation based on Kongsted (2003)

	Models $H^*(r,s)$										
T	$H^*(0,0)$	$H^*(0,1)$	$H^*(0,2)$	$H^*(0,3)$	$H^*(0,4)$	$H^*(1,0)$	$H^*(1,1)$	$H^*(1,2)$	$H^*(1,3)$	$H^*(2,0)$	$H^*(2,1)$
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**
1,000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**5.5**
Two-step test											
50	⋯	⋯	⋯	⋯	⋯	100.0	92.9	53.6	⋯	94.2	**33.5**
75	⋯	⋯	⋯	⋯	⋯	100.0	99.8	76.9	⋯	99.5	**32.9**
100	⋯	⋯	⋯	⋯	⋯	100.0	100.0	89.5	⋯	100.0	**28.7**
200	⋯	⋯	⋯	⋯	⋯	100.0	100.0	100.0	⋯	100.0	**16.8**
500	⋯	⋯	⋯	⋯	⋯	100.0	100.0	100.0	⋯	100.0	**10.1**
1,000	⋯	⋯	⋯	⋯	⋯	100.0	100.0	100.0	⋯	100.0	**7.5**

Note: Results are based on 5,000 replications and a nominal 5% level. The DGP is $H^*(2,1)$, i.e., (8) with $\epsilon_t \sim N(0,\Omega)$ and parameters and initial values given by

$$\alpha = \begin{pmatrix} -0.3058 & 0.0852 \\ 0.3172 & 0.0615 \\ 0.1495 & 0.0726 \\ -0.0394 & -0.0046 \end{pmatrix}, \quad \rho = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \tau^* = \begin{pmatrix} 1 & -0.4297 & 0.3329 \\ -0.5770 & 1 & -0.8600 \\ -0.5674 & -1.0475 & 1 \\ -0.0633 & 2.4535 & 1.0062 \\ 0.3222 & 0.6808 & -0.6143 \end{pmatrix}, \quad \psi^* = \begin{pmatrix} 1.6311 & -5.8530 \\ -1.0495 & 0.4409 \\ -0.7871 & -9.6780 \\ -0.4651 & 5.7352 \\ -33.6453 & 21.5573 \end{pmatrix},$$

$$\tilde{\alpha}_{\perp\Omega}\kappa' = \begin{pmatrix} -0.2117 & -0.2254 & 0.0222 \\ -0.4834 & -0.4892 & 0.0831 \\ 0.2582 & -0.0280 & -0.4106 \\ 0.0107 & -0.1815 & -0.2453 \end{pmatrix}, \quad 10^3 \cdot \Omega = \begin{pmatrix} 0.1635 & 0.0999 & -0.0078 & 0.0178 \\ 0.0999 & 0.3368 & -0.0147 & 0.0205 \\ -0.0078 & -0.0147 & 0.1154 & 0.0064 \\ 0.0178 & 0.0205 & 0.0064 & 0.0193 \end{pmatrix}, \quad \text{and} \quad (X_{-1}, X_0) = \begin{pmatrix} -0.3803 & -0.3863 \\ -0.7854 & -0.8151 \\ -0.3797 & -0.3320 \\ 0.0846 & 0.0841 \end{pmatrix}.$$

The tests are not calculated sequentially. Bold indicates rejection frequencies for tests of the correct model (empirical size), whereas ⋯ indicates that the results are identical to the LR test by construction. The standard error of the rejection frequency corresponding to the true probability π is $s_\pi = \sqrt{\pi(1-\pi)/5{,}000}$, e.g., $s_{0.05} = 0.003$.

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

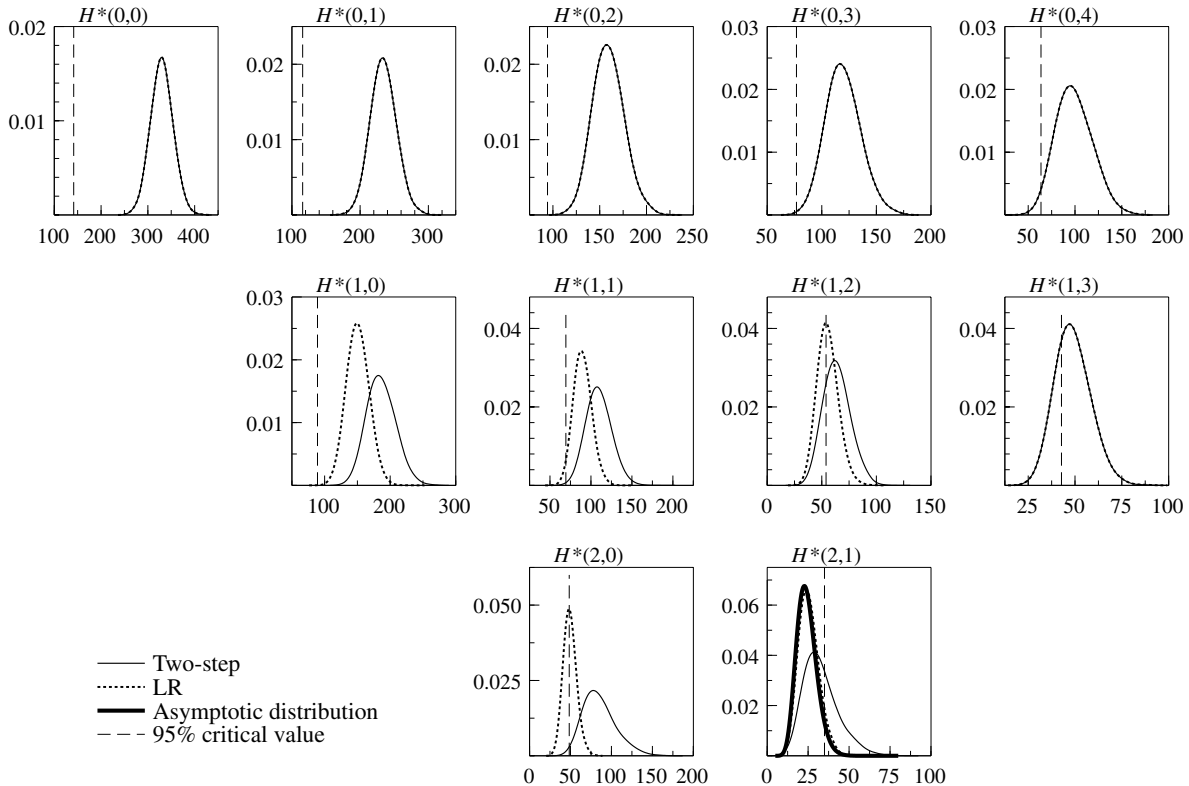


FIGURE 1. Distributions of the two test statistics for the case $T = 75$. Graphs are organized according to the partial nesting structure. Results are based on 5,000 replications.

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

TABLE 5. Rejection frequencies in a simulation based on Kongsted (2003)

	Models $H^*(r,s)$									
T	$H^*(0,0)$	$H^*(0,1)$	$H^*(0,2)$	$H^*(0,3)$	$H^*(0,4)$	$H^*(1,0)$	$H^*(1,1)$	$H^*(1,2)$	$H^*(1,3)$	$H^*(2,0)$
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**
1,000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**6.1**
2,500	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**5.6**
5,000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	**5.2**
Two-step test										
50	⋯	⋯	⋯	⋯	⋯	100.0	86.6	33.3	⋯	**91.0**
75	⋯	⋯	⋯	⋯	⋯	100.0	99.1	46.6	⋯	**91.1**
100	⋯	⋯	⋯	⋯	⋯	100.0	99.9	64.6	⋯	**86.3**
200	⋯	⋯	⋯	⋯	⋯	100.0	100.0	98.1	⋯	**66.3**
500	⋯	⋯	⋯	⋯	⋯	100.0	100.0	100.0	⋯	**38.3**
1,000	⋯	⋯	⋯	⋯	⋯	100.0	100.0	100.0	⋯	**23.0**
2,500	⋯	⋯	⋯	⋯	⋯	100.0	100.0	100.0	⋯	**11.2**
5,000	⋯	⋯	⋯	⋯	⋯	100.0	100.0	100.0	⋯	**8.0**

Note: The DGP is $H^*(2,0)$ with parameters

$$\alpha = \begin{pmatrix} -0.0394 & 0.0618 \\ 0.0244 & 0.0453 \\ 0.0411 & 0.0614 \\ 0.0584 & 0.0072 \end{pmatrix}, \quad \rho = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tau^* = \begin{pmatrix} 1 & 0.0850 \\ -0.8765 & 1 \\ -0.1135 & -1.9275 \\ -1.0899 & 2.2898 \\ -0.0266 & 1.2046 \end{pmatrix}, \quad \psi^* = \begin{pmatrix} 2.4169 & -7.4260 \\ 0.1976 & 1.2738 \\ -0.1551 & -14.1581 \\ -11.4035 & 0.7272 \\ -29.0551 & 14.9474 \end{pmatrix}, \quad \widetilde{\alpha}_{\perp\Omega}\kappa' = \begin{pmatrix} -0.6906 & -0.3606 \\ 0.4199 & -0.3213 \\ 0.3333 & 0.2402 \\ -0.1602 & -0.0680 \end{pmatrix},$$

$$10^3 \cdot \Omega = \begin{pmatrix} 0.1854 & 0.0856 & -0.0160 & 0.0220 \\ 0.0856 & 0.3511 & -0.0144 & 0.0112 \\ -0.0160 & -0.0144 & 0.1149 & 0.0033 \\ 0.0220 & 0.0112 & 0.0033 & 0.0194 \end{pmatrix}, \quad \text{and} \quad (X_{-1}, X_0) = \begin{pmatrix} -0.3803 & -0.3863 \\ -0.7854 & -0.8151 \\ -0.3797 & -0.3320 \\ 0.0846 & 0.0841 \end{pmatrix}.$$

The tests are not calculated sequentially. Bold indicates rejection frequencies for tests of the correct model (empirical size), whereas ⋯ indicates that the results are identical to the LR test by construction. The standard error of the rejection frequency corresponding to the true probability π is $s_\pi = \sqrt{\pi(1-\pi)/5{,}000}$, e.g., $s_{0.05} = 0.003$.

REFERENCES

Bacchiocchi, E. & L. Fanelli (2005) Testing the purchasing power parity through I(2) cointegration techniques. *Journal of Applied Econometrics* 20, 749–770.

Banerjee, A., L. Cockerell, & B. Russell (2001) An I(2) analysis of inflation and the markup. *Journal of Applied Econometrics* 16, 221–240.

Banerjee, A. & B. Russell (2001) The relationship between the markup and inflation in the G7 economies and Australia. *Review of Economics and Statistics* 83, 377–387.

Boswijk, H.P. (2000) Mixed normality and ancillarity in I(2) systems. *Econometric Theory* 16, 878–904.

Dennis, J. (2006) *CATS in RATS—Cointegration Analysis of Time Series, Version 2*. Estima.

Diamandis, P., D. Georgoutsos, & 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, 573–593.

Fliess, N. & R. MacDonald (2001) The instability of the money demand function: An I(2) interpretation. *Oxford Bulletin of Economics and Statistics* 63, 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.

Johansen, S. (1995) A statistical analysis of cointegration for I(2) variables. *Econometric Theory* 11, 25–59.

Johansen, S. (1997) Likelihood analysis of the I(2) model. *Scandinavian Journal of Statistics* 24, 433–462.

Johansen, S. (2002) A small sample correction for the test of cointegration rank in the vector autoregressive model. *Econometrica* 70, 1929–1961.

Johansen, S. (2006) Statistical analysis of hypotheses on the cointegrating relations in the I(2) model. *Journal of Econometrics* 132, 81–115.

Juselius, K. (1998) A structured VAR for Denmark under changing monetary regimes. *Journal of Business & Economic Statistics* 16, 400–412.

Juselius, K. (1999) Price convergence in the medium and long run: An I(2) analysis of six price indices. In R.F. Engle & H. White (eds.), *Cointegration, Causality, and Forecasting—A Festschrift in Honour of Clive W.J. Granger*. Oxford University Press.

Kongsted, H.C. (2003) An I(2) cointegration analysis of small-country import price determination. *Econometrics Journal* 6, 53–71.

Kurita, T. (2006) Likelihood Analysis of Weak Exogeneity in I(2) Systems and Reduced Econometric Representations. Working paper, Nuffield College, Oxford.

Nielsen, B. & A. Rahbek (2000) Similarity issues in cointegration analysis. *Oxford Bulletin of Economics and Statistics* 62, 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, 449–472.

Paruolo, P. (1996) On the determination of integration indices in I(2) systems. *Journal of Econometrics* 72, 313–356.

Paruolo, P. (2000) Asymptotic efficiency of the two stage estimator in I(2) systems. *Econometric Theory* 16, 524–550.

Paruolo, P. & A. Rahbek (1999) Weak exogeneity in I(2) VAR systems. *Journal of Econometrics* 93, 281–308.

Rahbek, A., H.C. Kongsted, & C. Jørgensen (1999) Trend-stationarity in the I(2) cointegration model. *Journal of Econometrics* 90, 265–289.

Rahbek, A. & H.B. Nielsen (2006) Likelihood Analysis of the I(2) Cointegration Model with Trend Stationarity. Working paper, University of Copenhagen.

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

APPENDIX A: Asymptotic Distribution of the LR Test

In this Appendix we give the proof of Theorem 1. We first introduce some notation. Similar to the definition of $S(X,Y,Z)$ in (11), define

$$R(Y,Z) \equiv \left(\int_0^1 Y_u Y_u' \, du \right)^{-1} \int_0^1 Y \, dZ'.$$

Define further product moment matrices M_{xy} for time series processes X_t and Y_t of dimension p_x and p_y, respectively, by $M_{xy} \equiv (1/T)\sum_{t=1}^T X_t Y_t'$. With θ a parameter, the corresponding true parameter is denoted θ_0. Furthermore, $\check{\theta}$ and $\hat{\theta}$ denote the ML estimator under the null hypothesis and the alternative, respectively.

Proof. Under $H(p)$ the VAR(k) model in (1) is given by

$$\Delta^2 X_t = \Pi X_{t-1} - \Gamma \Delta X_{t-1} + \Phi \Delta^2 \mathbb{X}_{t-1} + \epsilon_t, \tag{A.1}$$

where $\Phi = (\Psi_1, \ldots, \Psi_{k-2})$ and $\Delta^2 \mathbb{X}_{t-1} = (\Delta^2 X_{t-1}', \ldots, \Delta^2 X_{t-k+2}')'$. The hypothesis of interest is $H(r,s)$ given by (6),

$$\Delta^2 X_t = \alpha[\rho' \tau' X_{t-1} + \psi' \Delta X_{t-1}] + \tilde{\alpha}_{\perp\Omega} \kappa' \tau' \Delta X_{t-1} + \Phi \Delta^2 \mathbb{X}_{t-1} + \epsilon_t. \tag{A.2}$$

Henceforth, the parameters β, τ, and $\alpha_\perp$ under $H(r,s)$ are normalized on $c = \bar{\beta}_0$, $c = \bar{\tau}_0$, and $c = \bar{\alpha}_{0\perp}$, respectively, such that $\bar{\beta}_0'\beta = I_r$, $\bar{\tau}_0'\tau = I_{r+s}$, and $\tilde{\alpha}_{0\perp\Omega}'\alpha_\perp = I_{p-r}$. These are convenient normalizations and ensure identification of all parameters in the model. Note in particular that $\rho = \bar{\tau}_0'\beta$, which is $(r+s) \times r$.

We want to derive the asymptotic distribution of the LR test,

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

As in the likelihood analysis of the I(2) model in Johansen (1997) we introduce key parameters:

$$\begin{aligned} &B_0 = \bar{\beta}_{20}'(\psi - \psi_0), \qquad B_1 = \bar{\beta}_{10}'(\beta - \beta_0), \qquad B_2 = \bar{\beta}_{20}'(\beta - \beta_0), \\ &C = \bar{\beta}_{20}'(\tau - \tau_0)\rho_\perp, \end{aligned} \tag{A.3}$$

where $\rho_\perp = (I - \bar{\rho}_0(\rho'\bar{\rho}_0)^{-1}\rho)\bar{\rho}_{0\perp}$. Note that $\rho = \bar{\tau}_0'\beta = \rho_0 + \bar{\tau}_0'\beta_{10}B_1 = \rho(B_1)$ and define similarly $\rho_\perp(B_1)$. By Johansen (1997, Lem. 1) it follows that for the ML estimators $\hat{B}_0$, $\hat{B}_1$, $\hat{B}_2$, and $\hat{C}$ derived under $H(r,s)$,

$$(T\hat{B}_0', T\hat{B}_1', T^2\hat{B}_2',)' \xrightarrow{D} B^\infty = (B_0^{\infty\prime}, B_1^{\infty\prime}, B_2^{\infty\prime})' = R(H, V_1), \tag{A.4}$$

$$T\hat{C} \xrightarrow{D} C^\infty = R(H_0, V_2) \tag{A.5}$$

under $H_0(r,s)$. Here

$$H_u = (H_{0u}', H_{1u}', H_{2u}')' = \left(V_u' C_{20}' \beta_{20}, V_u' C_{10}' \beta_{10}, \int_0^u V_s' \, ds C_{20}' \beta_{20} \right), \tag{A.6}$$

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

with V_u a Brownian motion on $u \in [0,1]$ with covariance Ω_0. Furthermore,

$$V_{1u} = (\alpha_0'\Omega_0^{-1}\alpha_0)^{-1}\alpha_0'\Omega_0^{-1}V_u = \tilde{\alpha}_{0\Omega^{-1}}'V_u, \tag{A.7}$$

$$\begin{aligned} V_{2u} &= (\bar{\rho}_{0\perp}'\kappa_0(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\kappa_0'\bar{\rho}_{0\perp})^{-1}\bar{\rho}_{0\perp}'\kappa_0(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\alpha_{0\perp}'V_u \\ &= -(\xi_0'(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\xi_0)^{-1}\xi_0'(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\alpha_{0\perp}'V_u \equiv -\tilde{\xi}_{0\Omega_{\alpha_\perp\alpha_\perp}^{-1}}'\alpha_{0\perp}'V_u, \end{aligned} \tag{A.8}$$

where the identity $\xi_0 = -\kappa_0'\bar{\rho}_{0\perp}$ has been used (cf. Section 2.1.2; see also Johansen, 1997, Table 2) and $\Omega_{\alpha_\perp\alpha_\perp} \equiv \alpha_\perp'\Omega\alpha_\perp$. Now under $H(r,s)$ the model in (A.2) can be rewritten as

$$\Delta^2 X_t = \theta^{z_0}Z_{0t} + \theta^{z_1}Z_{1t} + \theta^{z_2}Z_{2t} + \epsilon_t, \tag{A.9}$$

in terms of the regressors Z_{0t}, Z_{1t}, and Z_{2t}, which are zero mean I(0), I(1), and I(2), respectively, and defined by

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

The parameters are given by

$$\begin{aligned} \theta^{z_0} &= (\alpha, \alpha(\psi - \psi_0)'\bar{\tau}_0 + \tilde{\alpha}_{\perp\Omega}\kappa', \Phi), \\ \theta^{z_1} &= (\alpha B_0' + \tilde{\alpha}_{\perp\Omega}\kappa'[\bar{\rho}_\perp(B_1)C' + \bar{\rho}(B_1)B_2'], \alpha B_1'), \\ \theta^{z_2} &= \alpha B_2'. \end{aligned} \tag{A.11}$$

Introduce the auxiliary hypothesis, H_{aux}, where ψ $(p \times r)$, β $(p \times r)$, and τ $(p \times (r+s))$ are fixed at their true values, ψ_0, β_0, and τ_0, corresponding to B_0, B_1, B_2, and C all identically zero. Note that $H_{\text{aux}} \subseteq H(r,s) \subseteq H(p)$ and that under H_{aux}, the model equation in (A.9) reduces to

$$\Delta^2 X_t = \theta^{z_0}Z_{0t} + \epsilon_t. \tag{A.12}$$

It follows by definition that

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

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

$$\begin{aligned} &\hat{\epsilon}_t = \Delta^2 X_t - \hat{\theta}^{z_0}Z_{0t} - \hat{\theta}^{z_1}Z_{1t} - \hat{\theta}^{z_2}Z_{2t}, \qquad \breve{\epsilon}_t = \Delta^2 X_t - \breve{\theta}^{z_0}Z_{0t}, \qquad \text{and} \\ &\hat{\epsilon}_{0t} = \Delta^2 X_t - \hat{\theta}^{z_0}Z_{0t}. \end{aligned} \tag{A.13}$$

Then $\breve{\Omega} = M_{\breve{\epsilon}\breve{\epsilon}}$, and

$$\hat{\Omega} = M_{\hat{\epsilon}\hat{\epsilon}} = M_{\hat{\epsilon}_0\hat{\epsilon}_0} + X^T - Y^T - (Y^T)',$$

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

where, with $Z_t = (Z'_{1t}, Z'_{2t})'$ and $\theta^z = (\theta^{z_1}, \theta^{z_2})$,

$$X^T = \hat{\theta}^z M_{zz} \hat{\theta}^{z\prime} \quad \text{and} \quad Y^T = (M_{\epsilon z} - (\hat{\theta}^{z_0} - \theta_0^{z_0}) M_{0z}) \hat{\theta}^{z\prime}.$$

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

$$D'_T M_{zz} D_T \xrightarrow{D} \int_0^1 H_u H'_u \, du. \tag{A.14}$$

Next, using (A.4), the definitions in (A.11), and consistency of the remaining parameters, it follows that

$$\sqrt{T} \hat{\theta}^z D_T^{-1} \xrightarrow{D} \theta^\infty = \alpha_0 (B_0^{\infty\prime}, B_1^{\infty\prime}, B_2^{\infty\prime}) - (\tilde{\alpha}_{0\perp\Omega} \xi_0 C^{\infty\prime}, 0, 0). \tag{A.15}$$

Combining (A.14) and (A.15) and using the definitions in (A.5) it follows that

$$TX^T \xrightarrow{D} X^\infty = \theta^\infty \int_0^1 H_u H'_u \, du \theta^{\infty\prime}.$$

Note that as $\alpha'_0 \alpha_{0\perp} = 0$ and $\tilde{\alpha}_{0\perp\Omega} = \Omega_0 \alpha_{0\perp} (\alpha'_{0\perp} \Omega_0 \alpha_{0\perp})^{-1}$ then by definition,

$$\text{tr}\{\Omega_0^{-1} X^\infty\} = \text{tr}\{\alpha'_0 \Omega_0^{-1} \alpha_0 S(V_1, H, V_1)\} + \text{tr}\{\xi'_0 (\alpha'_{0\perp} \Omega_0 \alpha_{0\perp})^{-1} \xi_0 S(V_2, H_0, V_2)\},$$

i.e., the cross-product terms vanish. Next, by convergence to stochastic integrals,

$$TY^T = \sqrt{T} M_{\epsilon z} D_T \sqrt{T} D_T^{-1} \hat{\theta}^{z\prime} + o_P(1) \xrightarrow{D} S(V, H, V_1) \alpha'_0 - S(V, H_0, V_2) \xi'_0 \tilde{\alpha}'_{0\perp\Omega}.$$

Therefore by joint convergence, and as $M_{\hat{\epsilon}_0 \hat{\epsilon}_0} \xrightarrow{P} \Omega_0$, $\check{\Omega} \xrightarrow{P} \Omega_0$, and furthermore using the definitions of V_1 and V_2 in (A.7) and (A.8),

$$\begin{aligned} -2 \log Q(H_{\text{aux}} | H(r,s)) = -T \log |\check{\Omega}^{-1} \hat{\Omega}| &\xrightarrow{D} \text{tr}\{\Omega_0^{-1} \alpha_0 \tilde{\alpha}'_{0\Omega^{-1}} S(V, H, V)\} \\ &+ \text{tr}\{(\alpha'_{0\perp} \Omega_0 \alpha_{0\perp})^{-1} \xi_0 \tilde{\xi}'_{0\Omega^{-1}_{\alpha_\perp \alpha_\perp}} \alpha'_{0\perp} S(V, H_0, V) \alpha_{0\perp}\}. \end{aligned} \tag{A.16}$$

Next, consider $-2 \log Q(H_{\text{aux}} | H(p))$. Under the alternative $H(p)$ the model is given by (A.1) and the ML estimators $\hat{\Pi}$, $\hat{\Gamma}$, $\hat{\Phi}$, and $\hat{\Omega}$ are given by ordinary least squares (OLS) regression of $\Delta^2 X_t$ on $(X'_{t-1}, \Delta X'_{t-1}, \Delta^2 \mathbb{X}'_{t-1})'$. Define Z_{it} as in (A.10), but redefine θ^{z_i} in (A.11) as

$$\begin{aligned} &\theta^{z_0} = (\Pi \bar{\beta}_0, -\Gamma \bar{\tau}_0 - \Pi \bar{\beta}_0 \psi'_0 \bar{\tau}_0, \Phi), \qquad \theta^{z_2} = \Pi \bar{\beta}_{20}, \\ &\theta^{z_1} = (-\Pi \bar{\beta}_0 \psi'_0 \bar{\beta}_{20} - \Gamma \bar{\beta}_{20}, \Pi \bar{\beta}_{10}). \end{aligned} \tag{A.17}$$

With these definitions, together with the identities $\alpha'_{0\perp} \Gamma_0 = -\kappa'_0 \tau'_0$ and (7), standard limiting arguments give that $\hat{\theta}^{z_0}$ is consistent, $\hat{\theta}^{z_0} \xrightarrow{P} \theta_0^{z_0} = (\alpha_0, \tilde{\alpha}_{0\perp\Omega} \kappa'_0, \Phi_0)$. Furthermore, $\theta_0^{z_1} = \theta_0^{z_2} = 0$, and similar to (A.15),

$$\sqrt{T} \hat{\theta}^z D_T^{-1} = (T \hat{\theta}^{z_1}, T^2 \hat{\theta}^{z_2}) \xrightarrow{D} R(V, H). \tag{A.18}$$

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

With these results, identical to the $H(r,s)$ case,

$$-2\log Q(H_{\text{aux}}|H(p)) \xrightarrow{D} \text{tr}\{\Omega_0^{-1} S(V,H,V)\}. \tag{A.19}$$

Next, collecting terms and using (7), it follows that

$$-2\log Q(H(r,s)|H(p)) = -2\log Q(H_{\text{aux}}|H(p)) - [-2\log Q(H_{\text{aux}}|H(r,s))] \tag{A.20}$$

$$\xrightarrow{D} \text{tr}\{\Omega_0^{-1}\tilde{\alpha}_{0\perp\Omega}\alpha_{0\perp}' S(V,H,V)\} - \text{tr}\{\alpha_{0\perp}(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\xi_0\tilde{\xi}'_{0\Omega^{-1}_{\alpha_\perp\alpha_\perp}}\alpha_{0\perp}' S(V,H_0,V)\}. \tag{A.21}$$

The first term in (A.21) can be rewritten as

$$\text{tr}\{\Omega_0^{-1}\tilde{\alpha}_{0\perp\Omega}\alpha_{0\perp}' S(V,H_0,V)\} + \text{tr}\{\Omega_0^{-1}\tilde{\alpha}_{0\perp\Omega}\alpha_{0\perp}' S(V,G,V)\},$$

where $G_u = ((H_{1u}', H_{2u}')' | H_0)$. Next, $I_{p-r} = \xi_0\tilde{\xi}'_{0\Omega^{-1}_{\alpha_\perp\alpha_\perp}} + \tilde{\xi}_{0\perp\Omega_{\alpha_\perp\alpha_\perp}}\xi_{0\perp}'$ and

$$\alpha_{0\perp}(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\xi_0\tilde{\xi}'_{0\Omega^{-1}_{\alpha_\perp\alpha_\perp}}\alpha_{0\perp}' = \alpha_{0\perp}(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\alpha_{0\perp}' - \alpha_{20}(\alpha_{20}'\Omega_0\alpha_{20})^{-1}\alpha_{20}', \tag{A.22}$$

where $\tilde{\xi}_{0\perp\Omega_{\alpha_\perp\alpha_\perp}} = (\alpha_{0\perp}'\Omega_0\alpha_{0\perp})\xi_{0\perp}(\xi_{0\perp}'\alpha_{0\perp}'\Omega_0\alpha_{0\perp}\xi_{0\perp})^{-1}$ and $\alpha_{20} = \alpha_{0\perp}\xi_{0\perp}$. Inserting (A.22) in the second term of (A.21) and collecting terms,

$$-2\log Q(H(r,s)|H(p)) \xrightarrow{D} \text{tr}\{\alpha_{0\perp}(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\alpha_{0\perp}' S(V,G,V)\}$$
$$+ \text{tr}\{\alpha_{20}(\alpha_{20}'\Omega_0\alpha_{20})^{-1}\alpha_{20}' S(V,H_0,V)\}, \tag{A.23}$$

and the result in Theorem 1 follows by defining the standard Brownian motion, W, by setting $W_u = (\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1/2}\alpha_{0\perp}' V_u$. ■

APPENDIX B. Asymptotics for the Linear Trend Case

Proof of Theorem 2. The proof mimics the proof of Theorem 1 with new parameters and notation as presented in the following discussion. The needed asymptotic distributions of the ML estimators in $H^*(r,s)$ are given in Rahbek and Nielsen (2006).

Under $H^*(p)$ the VAR(k) model in (8) is given by

$$\Delta^2 X_t = \Pi X_{t-1} - \Gamma\Delta X_{t-1} + \Phi\Delta^2\mathbb{X}_{t-1} + \Pi_l(t-1) - \Gamma_c + \epsilon_t, \tag{B.1}$$

where Π_l and Γ_c are p-dimensional parameters corresponding to the linear trend and constant regressors, respectively. Under $H^*(r,s)$ given by (6), the right-hand side in (B.1) is replaced by

$$\alpha[\rho'(\tau' X_{t-1} + \tau_l(t-1)) + \psi'\Delta X_{t-1} + \psi_c] + \tilde{\alpha}_{\perp\Omega}\kappa'(\tau'\Delta X_{t-1} + \tau_l) + \Phi\Delta^2\mathbb{X}_{t-1} + \epsilon_t. \tag{B.2}$$

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

From the likelihood analysis in Rahbek and Nielsen (2006), introduce parameters for the deterministic terms

$$B_l' = \rho'(\tau_l - \tau_{l0}), \qquad B_c' = (\psi - \psi_0)'\bar{\tau}_0\tau_{l0} + (\psi_c - \psi_{c0}), \qquad C_l' = \rho_\perp'(\tau_l - \tau_{l0}),$$

in addition to B_0, B_1, and B_2 in (A.4) and C in (A.5). By Rahbek and Nielsen (2006) it follows that for the ML estimators derived under $H^*(r,s)$ and evaluated under $H_0^*(r,s)$,

$$(T\hat{B}_0', T\hat{B}_1', T^2\hat{B}_2', T^{3/2}\hat{B}_l', \sqrt{T}\hat{B}_c')' \xrightarrow{D} B^\infty = (B_0^{\infty\prime}, B_1^{\infty\prime}, B_2^{\infty\prime}, B_l^{\infty\prime}, B_c^{\infty\prime})' = R(H^*, V_1)$$

$$(T\hat{C}', \sqrt{T}\hat{C}_l')' \xrightarrow{D} (C^{\infty\prime}, C_l^{\infty\prime})' = R(H_0^*, V_2).$$

With H_u defined in (A.6), $H_u^* = (H_u', u, 1)'$ and $H_{0u}^* = (H_{0u}', 1)'$, whereas V_1 and V_2 are defined in (A.7) and (A.8). Now under $H^*(r,s)$ the model in (B.2) can be represented as

$$\Delta^2 X_t = \theta^{z_0} Z_{0t}^* + \theta^{z_1} Z_{1t}^* + \theta^{z_2} Z_{2t}^* + \theta^{z_l} Z_{lt}^* + \theta^{z_c} Z_{ct}^* + \epsilon_t, \tag{B.3}$$

where $Z_{lt}^* = t - 1$, $Z_{ct}^* = 1$, $Z_{2t}^* = \beta_{20}' X_{t-1}$, and

$$Z_{0t}^* = \begin{pmatrix} \beta_0' X_{t-1} + \psi_0' \Delta X_{t-1} + \rho_0' \tau_{l0} Z_{lt} + \psi_{c0} \\ \tau_0' \Delta X_{t-1} + \tau_{l0} \\ \Delta^2 \mathbb{X}_{t-1} \end{pmatrix}, \qquad Z_{1t}^* = \begin{pmatrix} \beta_{20}' \Delta X_{t-1} \\ \beta_{10}' X_{t-1} + \beta_{10}' \bar{\tau}_0 \tau_{l0} \end{pmatrix}. \tag{B.4}$$

The parameters θ^{z_0}, θ^{z_1}, and θ^{z_2} are given by (A.11), whereas

$$\theta^{z_l} = \alpha B_l', \qquad \theta^{z_c} = (\alpha B_c' + \tilde{\alpha}_{\perp\Omega}\kappa'[\bar{\rho}_\perp(B_1)C_l' + \bar{\rho}(B_1)B_l']). \tag{B.5}$$

The auxiliary hypothesis, H_{aux}^*, is given by ψ, β, τ, ψ_c, and τ_l fixed at their true values ψ_0, β_0, τ_0, ψ_{c0}, and τ_{l0}, such that under H_{aux}^*, the model equation in (B.3) reduces as before to (A.12). For $-2\log Q(H_{\text{aux}}^* | H^*(r,s))$ define $\check{\epsilon}_t$ and $\hat{\epsilon}_{0t}$ as in (A.13), whereas

$$\hat{\epsilon}_t = \Delta^2 X_t - \hat{\theta}^{z_0} Z_{0t}^* - \hat{\theta}^{z_1} Z_{1t}^* - \hat{\theta}^{z_2} Z_{2t}^* - \hat{\theta}^{z_l} Z_{lt}^* - \hat{\theta}^{z_c} Z_{ct}^*.$$

Furthermore, set $Z_t = (Z_{1t}^{*\prime}, Z_{2t}^{*\prime}, Z_{lt}^*, Z_{ct}^*)'$, $\theta^z = (\theta^{z_1}, \theta^{z_2}, \theta^{z_l}, \theta^{z_c})$ and define the normalization by $D_T = \text{blockdiag}((1/\sqrt{T})I_{p-r}, (1/T^{3/2})I_{p-r-s}, 1/T, 1)$. Then as in (A.16),

$$-2\log Q(H_{\text{aux}}^* | H^*(r,s)) \xrightarrow{D} \text{tr}\{\Omega_0^{-1}\alpha_0 \tilde{\alpha}_{0\Omega^{-1}}' S(V, H^*, V)\}$$
$$+ \text{tr}\{(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\xi_0\tilde{\xi}_{0\Omega_{\alpha_\perp\alpha_\perp}^{-1}}'\alpha_{0\perp}' S(V, H_0^*, V)\alpha_{0\perp}\}.$$

Next, turn to $-2\log Q(H_{\text{aux}}^* | H^*(p))$. Under the alternative $H^*(p)$ the ML estimators $\hat{\Pi}$, $\hat{\Gamma}$, $\hat{\Phi}$, $\hat{\Pi}_l$, $\hat{\Gamma}_c$, and $\hat{\Omega}$ are found by OLS regression of $\Delta^2 X_t$ on $(X_{t-1}', \Delta X_{t-1}', \Delta^2 \mathbb{X}_{t-1}')'$ and the constant and linear trend. Define Z_{it}^* as in (B.4) and redefine θ^{z_i} in (B.5) for $i = 0,1,2$ as in (A.17), whereas

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press

$$\theta^{z_l} = \Pi_l - \Pi\bar{\beta}_0\rho_0'\tau_{l0}, \qquad \theta^{z_c} = \Gamma\bar{\tau}_0\tau_{l0} - \Gamma_c - \Pi\bar{\beta}_0\psi_{c0} - \Pi(\bar{\beta}_{10}\beta_{10}' - \bar{\beta}_0\psi_0')\bar{\tau}_0\tau_{l0}. \tag{B.6}$$

Then as in (A.19), $-2\log Q(H^*_{\text{aux}}|H^*(p)) \xrightarrow{D} \text{tr}\{\Omega_0^{-1}S(V,H^*,V)\}$. Finally, as in (A.23),

$$-2\log Q(H^*(r,s)|H^*(p)) \xrightarrow{D} \text{tr}\{\alpha_{0\perp}(\alpha_{0\perp}'\Omega_0\alpha_{0\perp})^{-1}\alpha_{0\perp}'S(V,G^*,V)\} + \text{tr}\{\alpha_{20}(\alpha_{20}'\Omega_0\alpha_{20})^{-1}\alpha_{20}'S(V,H_0^*,V)\},$$

where $G_u^* = ((H_{1u}', H_{2u}', u)'|H_0^*)$. ■

https://doi.org/10.1017/S0266466607070272 Published online by Cambridge University Press