Research Report

Department of Statistics

The Order of the Asymptotic Error Term for Moments of the Log Likelihood Ratio Test for Cointegration

by

Rolf Larsson

1998:9

Stockholm University, S-106 91 Stockholm, Sweden

The Order of the Asymptotic Error Term for Moments of the Log Likelihood Ratio Test for Cointegration

by

Rolf Larsson

1998:9

Stockholm University
Department of Statistics

ISBN 91-7474-191-8
ISSN 0280-7564

October 1998

The order of the asymptotic error term for moments of the log likelihood ratio test for cointegration.

Rolf Larsson
Stockholm University*

October 1, 1998

Abstract

The purpose of this paper is to prove that all moments of the log likelihood ratio test for cointegration in a vector-autoregressive (VAR) model of arbitrary order with fixed starting values asymptotically equal the corresponding moments in a VAR model of order one, plus an error term of order T^{-1}, where T is the sample size. This generalizes the corresponding result of Larsson (1998a) for unit root tests. We also discuss the implications of our theorem to cointegration testing in panels.

Key words: Asymptotic error, cointegration.

1. Introduction

Consider a p-variate stochastic process $\{X_t\}$ satisfying

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{m-1} \Gamma_i \Delta X_{t-i} + \varepsilon_t, \tag{1.1}$$

where $\{\varepsilon_t\}$ is a sequence of independent p-variately normal random variables with expectation zero and covariance matrix Ω. Furthermore, assume that we have

*Rolf Larsson, Department of Statistics, Stockholm University, Sweden, voice +46 8 16 29 93, e-mail rolf.larsson@stat.su.se

observations $X_1, ..., X_T$ and that the initial values $X_0, X_{-1}, ..., X_{-m+1}$ are fixed. If the Π matrix has reduced rank $r < p$, the process is said to be cointegrated. In Johansen (1995), the log likelihood ratio test of the hypothesis $H(r) : \text{rank}(\Pi) \leq r$ against $H(p) : \text{rank}(\Pi) \leq p$ is studied. Under certain assumptions on the system (see assumption A of the next section), the asymptotic distribution of this test, $-2\log Q_T^{(m)}$ say, if $\text{rank}(\Pi) = r$ equals the distribution of

$$Z \equiv tr\left\{\int_0^1 dW(t)\, W(t)' \left(\int_0^1 W(t)\, W(t)'\, dt\right)^{-1} \int_0^1 W(t)\, dW(t)'\right\},$$

where $W(t)$ is a $p - r$ dimensional Wiener process with initial value zero. This is a non-standard distribution, but it may be simulated, approximating the Wiener process by a random walk (cf e.g. Johansen (1995), chap. 15).

If, instead, we have observations from a panel, i.e. if we observe N parallell series $\{X_{j,t}\}$, $j = 1, ..., N$, where $X_{j,t}$ fulfills

$$\Delta X_{j,t} = \Pi_j X_{j,t-1} + \sum_{i=1}^{m-1} \Gamma_{j,i} \Delta X_{j,t-i} + \varepsilon_{j,t},$$

the asymptotics may, assuming that all the $\varepsilon_{j,t}$ are independent, be greatly simplified. Following Larsson, Lyhagen and Löthgren (1998), the log likelihood ratio test of the null hypothesis $rank(\Pi_j) \leq r$ for all j against the alternative $rank(\Pi_j) = p$ for all j is simply $\sum_j -2\log Q_{jT}$, where $-2\log Q_{jT}$ is the log likelihood test for series j. Further, the average of these statistics, $N^{-1}\sum_j -2\log Q_{jT} \equiv \overline{-2\log Q_T}$, obeys the central limit theorem, i.e., as $N \to \infty$,

$$\sqrt{N}\frac{\overline{-2\log Q_T} - E\left(-2\log Q_{jT}\right)}{Var\left(-2\log Q_{jT}\right)} \xrightarrow{w} N(0,1), \tag{1.2}$$

where $\xrightarrow{w}$ denotes weak convergence and $N(0,1)$ denotes the standard normal distribution. Unfortunately, the finite sample moments, i.e. the moments of $-2\log Q_{jT} = -2\log Q_T^{(m)}$, (emphasizing the dependency of the number of lags m), are not so easy to obtain. However, replacing these by the moments of Z (the asymptotic moments), we have normal convergence as N *and* T tend to infinity in such a way that the error caused by the moment approximations tends to zero. Indeed, if for $n = 1, 2$

$$E\left\{\left(-2\log Q_T^{(m)}\right)^n\right\} = E\left(Z^n\right) + O\left(T^{-\kappa}\right) \tag{1.3}$$

for some $\kappa > 0$, the condition on N and T is that $\sqrt{N}T^{-\kappa} \to 0$ as N and T tend to infinity. In a sense, the main result of the present paper, relating the VAR(m) moments to the VAR(1) moments via

$$E\left\{\left(-2\log Q_T^{(m)}\right)^n\right\} = E\left\{\left(-2\log Q_T^{(1)}\right)^n\right\} + O\left(T^{-1}\right), \tag{1.4}$$

consists a first step to prove (1.3). At any rate, it provides us with an easy way of approximating the finite sample moments of $-2\log Q_T^{(m)}$ in (1.2), because to simulate the moments of the VAR(1) test statistic, we need not take the Γ_i parameters into account. We end this discussion by pointing out that, in the unit root testing case (in a sense the univariate special case of cointegration testing), (1.4) was proved in Larsson (1998a), while for $n = m = 1$, a proof of (1.3), together with a numerical evaluation of $E(Z)$ and the first order error term, which is $O_p(1)$ in this case, is found in Larsson (1998b).

In section 2, we present the precise mathematical formulation of our theorem, while the proof of it, in principle following the lines of Larsson (1998a), is contained in section 3. Finally, some concluding remarks are given in section 4.

2. The theorem

Equation (1.1) may also be written in the form

$$A(L)X_t = \varepsilon_t, \quad A(L) \equiv (1-L)I_p - \Pi L - \sum_{i=1}^{m-1}\Gamma_i(1-L)L^i,$$

where L is the lag operator and I_p is the p-dimensional identity matrix. Now, assume for a moment that Π has reduced rank $r < p$, or equivalently, that there are $p \times r$ matrices α and β of full rank satisfying $\Pi = \alpha\beta'$. Then, we may define $p \times (p-r)$ matrices $\alpha_\perp$, $\beta_\perp$ such that $(\alpha, \alpha_\perp)$ has full rank, $\alpha'\alpha_\perp = 0$ and similarly for $\beta_\perp$. Moreover, introduce $\Gamma \equiv I_p - \sum_{i=1}^{m-1}\Gamma_i$. This notation is needed to formulate the technical assumption(s)

Assumption A: $|A(z)| = 0$ implies $z = 1$ or $|z| > 1$. Moreover, the matrix $\alpha_\perp'\Gamma\beta_\perp$ has full rank.

Under assumption A, the Granger representation theorem (Johansen (1995), theorem 4.2) states that there is a choice of initial distribution such that the linear combinations $\beta' X_t$ are stationary. These linear combinations are called the cointegrating relations, and r is called the cointegrating rank. Johansen (1995) shows that a consistent way of estimating this rank is to apply a sequential testing procedure based on tests of $H(r)$: rank$(\Pi) \leq r$ against $H(p)$: rank$(\Pi) \leq p$. Furthermore, he shows that the likelihood ratio test of $H(r)$ against $H(p)$, $Q_T^{(m)}$ say (to stress the dependency of the sample size T and of the order of the VAR process, m), satisfies

$$-2\log Q_T^{(m)} = -T \sum_{i=r+1}^{p} \log\left(1 - \widehat{\lambda}_i\right), \tag{2.1}$$

where $\widehat{\lambda}_1 > \widehat{\lambda}_2 > ... > \widehat{\lambda}_p$ are the ordered solutions of the eigenvalue problem

$$|S(\lambda)| = 0, \quad S(\lambda) \equiv \lambda S_{11} - S_{10} S_{00}^{-1} S_{01}, \tag{2.2}$$

with

$$S_{ij} \equiv M_{ij} - M_{i2} M_{22}^{-1} M_{2j}, \quad M_{ij} \equiv T^{-1} \sum_{t=1}^{T} Z_{it} Z_{jt}', \tag{2.3}$$

letting $Z_{0t} \equiv \Delta X_t$, $Z_{1t} \equiv X_{t-1}$ and $Z_{2t} \equiv (\Delta X_{t-1}, ..., \Delta X_{t-m+1})'$. The goal of the present article is to prove

Theorem 2.1. *Assume that $\Pi = \alpha\beta'$ and that assumption A holds. Then, for all positive integers n,*

$$E\left\{\left(-2\log Q_T^{(m)}\right)^n\right\} = E\left\{\left(-2\log Q_T^{(1)}\right)^n\right\} + O\left(T^{-1}\right).$$

3. The proof

In the following, we will generalize the corresponding proof for the univariate case, given in Larsson (1998a). The main idea is at first, assuming that[1] $\widehat{\lambda}_i = O_p\left(T^{-1}\right)$ for all $i \geq r+1$ (which is proved in the lemma below), to Taylor expand in (2.1) so that

$$-2\log Q_T^{(m)} = T \sum_{i=r+1}^{p} \left\{\widehat{\lambda}_i + O_p\left(\widehat{\lambda}_i^2\right)\right\} = T \sum_{i=r+1}^{p} \widehat{\lambda}_i + O_p\left(T^{-1}\right), \tag{3.1}$$

[1]Following Mann and Wald (1943), the notation $Y_T = O_p\left(T^{\kappa}\right)$ means that for each $\delta > 0$, there is a constant $A_\delta > 0$ such that $P\left(|Y_T| \leq A_\delta T^{\kappa}\right) \geq 1 - \delta$ for all T.

then to write $\sum_{i=r+1}^{p}\widehat{\lambda}_i$ as a trace of a certain product of matrices, and finally to investigate the order of the terms composing this matrix product in detail. To isolate all $\widehat{\lambda}_i$ with $i \geq r+1$, we at first need

Lemma 3.1. *If $\Pi = \alpha\beta'$ and assumption A holds, then for T sufficiently large the $p-r$ smallest solutions $\widehat{\lambda}_{r+1},...,\widehat{\lambda}_p$ to (2.2) equal the solutions to $|S^*(\lambda)| = 0$, where*

$$S^*(\lambda) \equiv \beta_\perp' S(\lambda)\beta_\perp - \beta_\perp' S(\lambda)\beta\left\{\beta' S(\lambda)\beta\right\}^{-1}\beta' S(\lambda)\beta_\perp.$$

Furthermore, $\widehat{\lambda}_i = O_p(T^{-1})$ for $r+1 \leq i \leq p$.

Proof. From (2.2), we have

$$0 = \left|\begin{pmatrix}\beta' \\ \beta_\perp'\end{pmatrix} S(\lambda)\begin{pmatrix}\beta & \beta_\perp\end{pmatrix}\right| = \left|\beta' S(\lambda)\beta\right| |S^*(\lambda)|,$$

with $S^*(\lambda)$ as above. But from lemma 10.3 of Johansen (1995), S_{00}, $\beta' S_{11}\beta$ and $\beta' S_{10}$ have well-defined limits (in probability), Σ_{00}, $\Sigma_{\beta\beta}$ and $\Sigma_{\beta 0} = \Sigma_{0\beta}'$, say. Hence,

$$\left|\beta' S(\lambda)\beta\right| \xrightarrow{w} \left|\lambda\Sigma_{\beta\beta} - \Sigma_{\beta 0}\Sigma_{00}^{-1}\Sigma_{0\beta}\right|,$$

showing that the r solutions of $\left|\beta' S(\lambda)\beta\right| = 0$ are $O_p(1)$ as $T \to \infty$.

Furthermore, as in the proof of lemma 10.3 in Johansen (1995), it follows that $\beta_\perp' S_{11}\beta_\perp$ is $O_p(T)$, while $\beta_\perp' S_{10}$ and $\beta_\perp' S_{11}\beta$ are $O_p(1)$. Consequently,

$$\beta_\perp' S(\lambda)\beta_\perp = \lambda\beta_\perp' S_{11}\beta_\perp - \beta_\perp' S_{10}S_{00}^{-1}S_{01}\beta_\perp = \lambda O_p(T) + O_p(1), \tag{3.2}$$

and similarly, $\beta_\perp' S(\lambda)\beta = O_p(1)$, showing that

$$S^*(\lambda) = \lambda O_p(T) + O_p(1),$$

and so, the solutions $\widehat{\lambda}_{r+1},...,\widehat{\lambda}_p$ to $|S^*(\lambda)| = 0$ have to be $O_p(T^{-1})$. ■

Thanks to lemma 3.1, we may now approximate the test statistic by the trace of a matrix product, as in the following lemma.

Lemma 3.2. *If $\Pi = \alpha\beta'$ and assumption A holds, we have*

$$-2\log Q_T^{(m)} = Ttr\left\{\left(\beta_\perp' S_{11}\beta_\perp\right)^{-1}\beta_\perp' S_{10}NS_{01}\beta_\perp\right\} + O_p\left(T^{-1}\right),$$

where

$$N = S_{00}^{-1} - S_{00}^{-1}S_{01}\beta\left(\beta' S_{10}S_{00}^{-1}S_{01}\beta\right)^{-1}\beta' S_{10}S_{00}^{-1}.$$

Proof. Since $\widehat{\lambda}_i = O_p\left(T^{-1}\right)$ for $i \geq r+1$, we may write $\lambda = T^{-1}\rho$, and as in the previous proof, we find

$$\beta_\perp' S(\lambda)\beta = T^{-1}\rho\beta_\perp' S_{11}\beta - \beta_\perp' S_{10}S_{00}^{-1}S_{01}\beta = -\beta_\perp' S_{10}S_{00}^{-1}S_{01}\beta + O_p\left(T^{-1}\right)$$

and

$$\beta' S(\lambda)\beta = T^{-1}\rho\beta' S_{11}\beta - \beta' S_{10}S_{00}^{-1}S_{01}\beta = -\beta' S_{10}S_{00}^{-1}S_{01}\beta + O_p\left(T^{-1}\right).$$

In conjunction with (3.2), these facts imply

$$S^*(\lambda) = T^{-1}\rho\beta_\perp' S_{11}\beta_\perp - \beta_\perp' S_{10}NS_{01}\beta_\perp + O_p\left(T^{-1}\right)$$

with N as defined above. Hence, application of (3.1) yields the result of lemma. ■

Going on in the fashion of Larsson (1998a), we now need some moving average representations.

Lemma 3.3. *(a): If $\Pi = \alpha\beta'$ and assumption A holds, then, defining $\widetilde{X}_t' \equiv \left(Z_t', U_t'\right)$ with $Z_t \equiv \beta' X_t$ and $U_t \equiv \beta_\perp' \Delta X_t$, we have*

$$\widetilde{X}_t = C_t(L)\varepsilon_t + a_t, \tag{3.3}$$

where

$$a_t \equiv \sum_{s=1}^{m-1} c_{t-s}e_s, \quad e_s \equiv \widetilde{\Gamma}_s X_0 + \ldots + \widetilde{\Gamma}_{m-1}X_{s-m+1}, \quad C_t(L) \equiv \sum_{i=0}^{t-1} c_i L^i,$$

for some matrices $\widetilde{\Gamma}_s, \ldots, \widetilde{\Gamma}_{m-1}$, where L is the lag operator, $c_0 = I_p$ and c_n is defined recursively through

$$c_n = \sum_{i=0}^{(m-1)\wedge n} c_{n-i}\widetilde{\Gamma}_i, \quad n = 1, 2, \ldots.$$

Moreover, we have the representation

$$C_t(L) = C_t(1) + (1-L)C_t^{(1)}(L), \quad C_t^{(1)}(L) = \sum_{i=0}^{t-1} c_i^{(1)} L^i,$$

where

$$c_i^{(1)} \equiv -\sum_{j=i+1}^{t-1} c_j, \quad i = 0, 1, ..., t-1.$$

Furthermore, for some $\delta > 0$, *the sums*

$$\begin{aligned} C_\infty (L) &\equiv \lim_{t\to\infty} C_t (L) = \sum_{i=0}^{\infty} c_i L^i, \\ C_\infty^{(1)} (L) &\equiv \lim_{t\to\infty} C_t^{(1)} (L) = \sum_{i=0}^{\infty} c_i^{(1)} L^i, \end{aligned}$$

are absolutely convergent for $|L| < 1 + \delta$, *and*[2] $\|c_i\|$ *and* $\left\|c_i^{(1)}\right\|$ *tend to zero exponentially fast as* $i \to \infty$.

(b): If $\Pi = \alpha\beta'$ *and assumption A holds, we also have*

$$\Delta X_t = D_t (L) \varepsilon_t + a_t^*,$$

where a_t^* *has a similar representation as* a_t, *and where*

$$D_t (L) \equiv \sum_{i=0}^{2t-2} d_i L^i,$$

with $d_0 = I_p$, *and where the* $D_t (L)$ *polynomial fulfills the corresponding expansion and convergence properties as* $C_t (L)$ *of part (a).*

Proof. Johansen (1995) in his proof of the Granger representation theorem derives the representation

$$\widetilde{A} (L) \widetilde{X}_t = \begin{pmatrix} \overline{\alpha}' \\ \overline{\alpha}_\perp' \end{pmatrix} \varepsilon_t,$$

and shows it to be invertible. Hence, as in the proof of Theorem 2.1 of Johansen (1995), we may write

$$\widetilde{X}_t = C_t (L) \varepsilon_t + a_t, \quad C_t (L) \equiv \widetilde{C}_t (L) \begin{pmatrix} \overline{\alpha}' \\ \overline{\alpha}_\perp' \end{pmatrix},$$

[2] We define the norm $\|A\|$ of a matrix A through $\|A\| \equiv tr\left(AA'\right)$.

and the properties of $C_t(L)$ and a_t as stated in part (a) of our lemma follow as in Johansen (1995), theorems 2.1 and 2.2.

As for (b), rewrite (1.1) as

$$B(L)\Delta X_t = \alpha\beta' X_{t-1} + \varepsilon_t \equiv \varepsilon_t^*, \quad B(L) \equiv I_p - \sum_{i=1}^{m-1} \Gamma_i L^i.$$

But because of the Granger representation theorem, the process ε_t^* is stationary (under a suitable choice of initial distribution), hence the equation $|B(z)| = 0$ can only have roots z with $|z| > 1$. Hence, we may as in part (a) write

$$\Delta X_t = D_t^*(L)\,\varepsilon_t^* + a_t^{(1)}, \quad D_t^*(L) \equiv \sum_{i=0}^{t-1} d_i^* L^i,$$

where $a_t^{(1)}$ may be expressed in terms of the d_i^* sequence in the same way as a_t was expressed in terms of the c_i's of part (a). Moreover, because of part (a), it follows that

$$D_t^*(L)\,\varepsilon_t^* = D_t^*(L)\,\varepsilon_t + D_t^*(L)\begin{pmatrix} \alpha & 0 \end{pmatrix}\left(\sum_{i=0}^{t-2} c_i\varepsilon_{t-1-i} + a_{t-1}^{(2)}\right) = \sum_{j=0}^{t-1} d_j\varepsilon_{t-j} + a_t^{(3)},$$

where

$$d_i \equiv d_i^* + \sum_{j=0}^{i-1} d_j^*\begin{pmatrix} \alpha & 0 \end{pmatrix} c_{i-j-1}, \quad i = 1, ..., 2t-2, \quad a_t^{(3)} \equiv \sum_{i=0}^{t-1} d_i^*\begin{pmatrix} \alpha & 0 \end{pmatrix} a_{t-1-i}^{(2)}.$$

As a consequence of these representations, the d_i and $a_t^* = a_t^{(1)} + a_t^{(3)}$ sequences fulfill the corresponding expansion and convergence properties as the c_i and a_t sequences of part (a), completing the proof of (b). ■

For later use, we need to specify some first order asymptotic results, similar to those of Johansen (1995). To this end, we define the random walk $S_t \equiv \sum_{i=1}^{t} \varepsilon_t$. Also, from now on, $\sum_t$ will denote summation over $\{1 \le t \le T\}$, and all ε_t with a t not belonging to this interval will be interpreted as zero.

Lemma 3.4. *Under assumption A and if* $\Pi = \alpha\beta'$, *we have for fixed and arbitrary* k *that*

$$\sum_t \varepsilon_{t-k}\varepsilon_t' = \begin{cases} O_p(T) & \text{if } k = 0, \\ O_p\left(T^{1/2}\right) & \text{otherwise,} \end{cases} \tag{3.4}$$

$$\sum_t S_{t-k-1}\varepsilon_t' = O_p(T), \quad (3.5)$$

$$\sum_t S_{t-1}S_{t-1}' = O_p\left(T^2\right). \quad (3.6)$$

Proof. Because of the law of large numbers, $T^{-1}\sum_t \varepsilon_t\varepsilon_t'$ converges in probability to Ω, showing that $\sum_t \varepsilon_t\varepsilon_t'$ is $O_p(T)$. Further, for k nonzero, the fact that $\sum_t \varepsilon_{t-k}\varepsilon_t'$ is $O_p\left(T^{1/2}\right)$ follows from the univariate central limit theorem and the central limit theorem for m-dependent sequences (cf Chung (1974)).

Regarding (3.5), we have from the limit theory of e.g. Johansen (1995) that as $T \to \infty$,

$$T^{-1}\sum_t S_{t-1}\varepsilon_t' \xrightarrow{w} \int_0^1 W_t dW_t', \quad (3.7)$$

where $\xrightarrow{w}$ denotes weak convergence and W_t is a standard Wiener process, proving (3.5) in the special case $k = 0$. In general, we may write

$$\sum_t S_{t-k-1}\varepsilon_t' = \sum_t S_{t-1}\varepsilon_t' - \sum_{i=1}^{k}\sum_t \varepsilon_{t-i-1}\varepsilon_t',$$

which, from (3.4) and (3.7), is $O_p(T)$, proving (3.5).

Finally, by Johansen (1995) again,

$$T^{-2}\sum_t S_{t-1}S_{t-1}' \xrightarrow{w} \int_0^1 W_t W_t' dt,$$

as $T \to \infty$, showing (3.6). ■

Now, we go on in the style of Larsson (1998a) to prove

Lemma 3.5. *Under assumption A and if* $\Pi = \alpha\beta'$,

$$\sum_t C_{t-1}(1) S_{t-1}\varepsilon_t' = C_T(1)\sum_t S_{t-1}\varepsilon_t' + O_p(1), \quad (3.8)$$

$$\sum_t C_{t-1}(1) S_{t-1}S_{t-1}'C_{t-1}(1)' = C_T(1)\sum_t S_{t-1}S_{t-1}'C_T(1)' + O_p(1), \quad (3.9)$$

$$\beta'\sum_t X_{t-1}\varepsilon_t' = \begin{pmatrix} I_r & 0 \end{pmatrix}\sum_{i=0}^{T-2} c_i \sum_t \varepsilon_{t-1-i}\varepsilon_t' + O_p(1), \quad (3.10)$$

$$\sum_t \Delta X_{t-k}\varepsilon_t' = \sum_{i=0}^{2(T-k-1)} d_i \sum_t \varepsilon_{t-k-i}\varepsilon_t' + O_p(1), \quad (3.11)$$

for k arbitrary. Furthermore, $\beta' \sum_t X_{t-1}\varepsilon_t'$ and $\sum_t \Delta X_{t-k}\varepsilon_t'$, $k > 0$, are both $O_p\left(T^{1/2}\right)$, while the quantities $\beta_\perp' M_{12}$, $\beta' M_{12}$, M_{22} and M_{02} are all $O_p(1)$.

Proof. Looking at first at (3.8), we find

$$\sum_t C_{t-1}(1)\, S_{t-1}\varepsilon_t' = \sum_{i=0}^{T-2} c_i \sum_{t=i+2}^{T} S_{t-1}\varepsilon_t' = C_{T-1}(1) \sum_t S_{t-1}\varepsilon_t' - r_1, \tag{3.12}$$

where

$$r_1 \equiv \sum_{i=0}^{T-2} c_i \sum_{t=1}^{i+1} S_{t-1}\varepsilon_t' = \sum_{t=1}^{T-1} \sum_{i=t-1}^{T-2} c_i S_{t-1}\varepsilon_t'.$$

Thus, because[3] $\|c_i\| \sim \gamma^i$ for some γ with $|\gamma| < 1$,

$$\|r_1\| \le \sum_{t=1}^{T-1} \sum_{i=t-1}^{T-2} \|c_i\| \left\|S_{t-1}\varepsilon_t'\right\| = \sum_{t=1}^{T-1} O\left(\gamma^t\right) \left\|S_{t-1}\varepsilon_t'\right\| = O_p(1). \tag{3.13}$$

Further, because of the exponential decay of the c_i coefficients, replacing $C_{T-1}(1)$ by $C_T(1)$ in (3.12) will only cause an exponentially small error, and so, (3.8) follows.

The proof of (3.9) follows similar lines, and is omitted.

Regarding $\beta' \sum_t X_{t-1}\varepsilon_t'$, lemma 3.3 yields

$$\beta' \sum_t X_{t-1}\varepsilon_t' = \begin{pmatrix} I_r & 0 \end{pmatrix} \left\{ \sum_t C_{t-1}(L)\, \varepsilon_{t-1}\varepsilon_t' + r_2 \right\}, \quad r_2 \equiv \sum_t a_{t-1}\varepsilon_t'.$$

Here,

$$\sum_t C_{t-1}(L)\, \varepsilon_{t-1}\varepsilon_t' = \sum_{i=0}^{T-2} c_i \sum_t \varepsilon_{t-i-1}\varepsilon_t', \tag{3.14}$$

and because by lemma 3.4, $\sum_t \varepsilon_{t-i-1}\varepsilon_t'$ is $O_p\left(T^{1/2}\right)$, and we may argue as in (3.13) above to conclude that the r.h.s. of (3.14) has to be (at most) $O_p\left(T^{1/2}\right)$. Further, because r_2 is a linear combination of normal r.v's with expectation zero and covariance matrix $a_-\left(\Omega \otimes I_T\right) a_-'$, where $a_- \equiv (a_0, ..., a_{T-1})$, r_2 is $O_p(1)$, and (3.10) and the fact that $\beta' \sum_t X_{t-1}\varepsilon_t'$ is $O_p\left(T^{1/2}\right)$ follow.

[3] The notation $a_n \sim b_n$ means that as $n \to \infty$, $a_n / b_n \to 1$.

By arguing in a similar fashion, we may show that for k arbitrary, (3.11) holds and that for $k > 0$, $\sum_t \Delta X_{t-k}\varepsilon_t'$ is $O_p\left(T^{1/2}\right)$.

To treat $\beta_\perp' M_{12}$, we observe that

$$T\beta_\perp' M_{12} = \beta_\perp' \sum_t X_{t-1} \left(\begin{array}{ccc} \Delta X_{t-1}' & \dots & \Delta X_{t-m+1}' \end{array} \right).$$

In the following, we will focus on the first component of the r.h.s. of this expression, and it will become clear that the others may be treated in a similar way. To this end, we have by lemma 3.3 that, with $b_t \equiv \sum_{i=1}^t a_t$,

$$\begin{aligned} \beta_\perp' \sum_t X_{t-1}\Delta X_{t-1}' &= \left(\begin{array}{cc} 0 & I_{p-r} \end{array} \right) \sum_t \left\{ C_{t-1}(L) S_{t-1} + b_t \right\} \left\{ D_{t-1}(L)\varepsilon_{t-1} + a_{t-1}^* \right\}' \\ &= \left(\begin{array}{cc} 0 & I_{p-r} \end{array} \right) \sum_t C_{t-1}(L) S_{t-1}\varepsilon_{t-1}' D_{t-1}(L)' + r_3, \end{aligned}$$

where r_3, not written out explicitly here, is $O_p(1)$ in the usual fashion. Moreover, it follows that

$$\sum_t C_{t-1}(L) S_{t-1}\varepsilon_{t-1}' D_{t-1}(L)' = \sum_{i=0}^{T-2} \sum_{j=0}^{2(T-2)} c_i \sum_t S_{t-i-1}\varepsilon_{t-j-1}' d_j', \tag{3.15}$$

and because of lemma 3.4, $\sum_t S_{t-i-1}\varepsilon_{t-j-1}'$ is $O_p(T)$, and it is a simple matter to generalize the argument of (3.13) to prove that (3.15) is $O_p(T)$. This completes the proof of the fact that $\beta_\perp' \sum_t X_{t-1}\Delta X_{t-1}'$ is $O_p(T)$, which was what we wanted to show to realize that $\beta_\perp' M_{12}$ has to be $O_p(1)$.

To see that $\beta' M_{12}$, M_{22} and M_{02} are all $O_p(1)$ is very similar, and the proofs of these facts are omitted here. ■

The following lemma writes the main term of lemma 3.2 in a more tractable form for further analysis.

Lemma 3.6. *Under assumption A and if* $\Pi = \alpha\beta'$,

$$-2\log Q_T^{(m)} = Ttr\left\{ \left(\beta_\perp' S_{11}\beta_\perp\right)^{-1} \beta_\perp' S_{10}\alpha_\perp \left(\alpha_\perp' S_{00}\alpha_\perp\right)^{-1} \alpha_\perp' S_{01}\beta_\perp \right\} + R + O_p\left(T^{-1}\right), \tag{3.16}$$

where

$$R \equiv -2Ttr\left\{ \left(\beta_\perp' S_{11}\beta_\perp\right)^{-1} \beta_\perp' S_{10} R_0 S_{01}\beta_\perp \right\}, \tag{3.17}$$

$$R_0 \equiv \left\{\alpha - \alpha_\perp \left(\alpha_\perp' S_{00}\alpha_\perp\right)^{-1} \alpha_\perp' S_{00}\alpha\right\} \left(\alpha'\alpha\right)^{-1} S_{\beta\beta}^{-1} R_1' \times\alpha_\perp \left(\alpha_\perp' S_{00}\alpha_\perp\right)^{-1} \alpha_\perp', \tag{3.18}$$

with $S_{\beta\beta} \equiv \beta' S_{11}\beta$ *and*

$$R_1 \equiv S_{01}\beta - \alpha S_{\beta\beta} = T^{-1}\sum_{i=0}^{T-2}\sum_t \varepsilon_t\varepsilon_{t-i-1}' c_i' \begin{pmatrix} I_r \\ 0 \end{pmatrix} - T^{-1}Y' M_{22}^{-1} M_{21}\beta = O_p\left(T^{-1/2}\right), \tag{3.19}$$

where

$$Y' \equiv \sum_t \varepsilon_t Z_{2t}' = \sum_t \varepsilon_t \begin{pmatrix} \Delta X_{t-1}' & \dots & \Delta X_{t-m+1}' \end{pmatrix}.$$

Proof. Using the representation (cf (1.1))

$$Z_{0t} = \alpha\beta' Z_{1t} + \Psi Z_{2t} + \varepsilon_t, \tag{3.20}$$

where as before, $Z_{0t} = \Delta X_t$, $Z_{1t} = X_{t-1}$, $Z_{2t} = (\Delta X_{t-1}, \dots, \Delta X_{t-m+1})'$ and where $\Psi \equiv (\Gamma_1, \dots, \Gamma_{m-1})$, we find from (2.3) that

$$\begin{aligned} S_{01} - \alpha\beta' S_{11} &= M_{01} - \alpha\beta' M_{11} - \left(M_{02}M_{22}^{-1}M_{21} - \alpha\beta' M_{12}M_{22}^{-1}M_{21}\right) \\ &= T^{-1}\sum_t \left(Z_{0t} - \alpha\beta' Z_{1t}\right) Z_{1t}' - T^{-1}\sum_t \left(Z_{0t} - \alpha\beta' Z_{1t}\right) Z_{2t}' M_{22}^{-1} M_{21} \\ &= T^{-1}\sum_t (\varepsilon_t + \Psi Z_{2t}) \left(Z_{1t}' - Z_{2t}' M_{22}^{-1} M_{21}\right) \\ &= T^{-1}\sum_t \varepsilon_t Z_{1t}' - T^{-1}\sum_t \varepsilon_t Z_{2t}' M_{22}^{-1} M_{21}, \end{aligned} \tag{3.21}$$

in conjunction with (3.10) of lemma 3.5 proving the second equality of (3.19). Also, the third equality of (3.19) follows from lemma 3.5.

Now, inserting the identity $S_{01}\beta = \alpha S_{\beta\beta} + R_1$ into the expression for N of lemma 3.2, we find

$$N = S_{00}^{-1} - S_{00}^{-1} (\alpha S_{\beta\beta} + R_1) A^{-1} \left(S_{\beta\beta}\alpha' + R_1'\right) S_{00}^{-1}, \tag{3.22}$$

where $A = A_0 + R_2$ with

$$\begin{aligned} A_0 &\equiv S_{\beta\beta}\alpha' S_{00}^{-1}\alpha S_{\beta\beta}, \\ R_2 &\equiv S_{\beta\beta}\alpha' S_{00}^{-1} R_1 + R_1' S_{00}^{-1}\alpha S_{\beta\beta} + R_1' S_{00}^{-1} R_1. \end{aligned} \tag{3.23}$$

But, knowing that R_1 is $O_p\left(T^{-1/2}\right)$ and that, by Johansen (1995), lemma 10.3, $S_{\beta\beta}$ and S_{00} are $O_p(1)$, we see that A_0 is $O_p(1)$ and that R_2 is $O_p\left(T^{-1/2}\right)$. Hence, we have

$$A^{-1} = A_0^{-1} - A_0^{-1} R_2 A_0^{-1} + O_p\left(T^{-1}\right),$$

and simplification of (3.22) leads us to

$$N = N_0 + R_3 + O_p\left(T^{-1}\right),$$

where

$$\begin{aligned} N_0 &\equiv S_{00}^{-1} - S_{00}^{-1}\alpha\left(\alpha' S_{00}^{-1}\alpha\right)^{-1}\alpha S_{00}^{-1}, & (3.24)\\ R_3 &\equiv -S_{00}^{-1}\alpha\left(\alpha' S_{00}^{-1}\alpha\right)^{-1} S_{\beta\beta}^{-1} R_1' S_{00}^{-1} - S_{00}^{-1} R_1 S_{\beta\beta}^{-1}\left(\alpha' S_{00}^{-1}\alpha\right)^{-1}\alpha' S_{00}^{-1} \\ &\quad + S_{00}^{-1}\alpha\left(\alpha' S_{00}^{-1}\alpha\right)^{-1} S_{\beta\beta}^{-1} R_2 S_{\beta\beta}^{-1}\left(\alpha' S_{00}^{-1}\alpha\right)^{-1}\alpha' S_{00}^{-1}. & (3.25) \end{aligned}$$

Next, observing that

$$N_0 \begin{pmatrix} \alpha & S_{00}\alpha_\perp \end{pmatrix} = \begin{pmatrix} 0 & \alpha_\perp \end{pmatrix},$$

we may, because $\begin{pmatrix} \alpha & S_{00}\alpha_\perp \end{pmatrix}$ and $\begin{pmatrix} \alpha & \alpha_\perp \end{pmatrix}$ both have full rank, rewrite (3.24) as

$$N_0 = \begin{pmatrix} 0 & \alpha_\perp \end{pmatrix} \left\{ \begin{pmatrix} \alpha' \\ \alpha_\perp' \end{pmatrix} \begin{pmatrix} \alpha & S_{00}\alpha_\perp \end{pmatrix} \right\}^{-1} \begin{pmatrix} \alpha' \\ \alpha_\perp' \end{pmatrix} = \alpha_\perp \left(\alpha_\perp' S_{00}\alpha_\perp\right)^{-1} \alpha_\perp',$$

the last equality following from the well-known partitioned inverse formula. This yields the main term of (3.16), and for (3.17), (3.23) and (3.25) imply

$$R_3 = \widetilde{R} + \widetilde{R}' + O_p\left(T^{-1}\right),$$

where

$$-\widetilde{R} \equiv S_{00}^{-1}\alpha\left(\alpha' S_{00}^{-1}\alpha\right)^{-1} S_{\beta\beta}^{-1} R_1' \left(S_{00}^{-1} - S_{00}^{-1}\alpha\left(\alpha' S_{00}^{-1}\alpha\right)^{-1}\alpha' S_{00}^{-1}\right).$$

But as above, it follows that

$$S_{00}^{-1} - S_{00}^{-1}\alpha\left(\alpha' S_{00}^{-1}\alpha\right)^{-1}\alpha' S_{00}^{-1} = \alpha_\perp\left(\alpha_\perp' S_{00}\alpha_\perp\right)^{-1}\alpha_\perp'$$

and

$$S_{00}^{-1}\alpha\left(\alpha' S_{00}^{-1}\alpha\right)^{-1}=\left\{\alpha-\alpha_{\perp}\left(\alpha_{\perp}' S_{00}\alpha_{\perp}\right)^{-1}\alpha_{\perp}' S_{00}\alpha\right\}\left(\alpha'\alpha\right)^{-1}.$$

As a simple consequence of this and of the cyclicity of the tr operator, (3.17) and (3.18) are deduced. ■

Our next lemma relates the quantities building up the expression of lemma 3.6 to the corresponding quantities in the $m=1$ case.

Lemma 3.7. *Under assumption A and if* $\Pi=\alpha\beta'$,

$$T S_{00}\alpha_{\perp} = \sum_t \varepsilon_t\varepsilon_t'\alpha_{\perp}+R_1^*\alpha_{\perp}+O_p(1), \quad (3.26)$$

$$T\beta_{\perp}' S_{10}\alpha_{\perp} = \begin{pmatrix} 0 & I_{p-r}\end{pmatrix} C_T(1)\sum_t S_{t-1}\varepsilon_t'\alpha_{\perp}+R_2^*\alpha_{\perp}+O_p(1), \quad (3.27)$$

$$T\beta_{\perp}' S_{11}\beta_{\perp} = \begin{pmatrix} 0 & I_{p-r}\end{pmatrix} C_T(1)\sum_t S_{t-1}S_{t-1}' C_T(1)'\begin{pmatrix} 0 \\ I_{p-r}\end{pmatrix} + O_p(T), \quad (3.28)$$

where

$$R_1^* \equiv \sum_{i=1}^{T-1} d_i\sum_t \varepsilon_{t-i}\varepsilon_t' - M_{02}M_{22}^{-1}Y = O_p\left(T^{1/2}\right), \quad (3.29)$$

$$R_2^* \equiv \begin{pmatrix} 0 & I_{p-r}\end{pmatrix}\sum_{i=1}^{T-2} c_i^{(1)}\sum_t \varepsilon_{t-i}\varepsilon_t' - \beta_{\perp}' M_{12}M_{22}^{-1}Y = O_p\left(T^{1/2}\right). \quad (3.30)$$

Proof. We will start by proving (3.28). To this end, recall that from (2.3),

$$\beta_{\perp}' S_{11}\beta_{\perp}=\beta_{\perp}' M_{11}\beta_{\perp}-\beta_{\perp}' M_{12}M_{22}^{-1}M_{21}\beta_{\perp}. \quad (3.31)$$

Now, from lemma 3.5 it is seen that the second term of the r.h.s. of (3.31) is $O_p(1)$, and we are left to consider $\beta_{\perp}' M_{11}\beta_{\perp}$. For this, we have by lemma 3.3

$$T\beta_{\perp}' M_{11}\beta_{\perp}=\beta_{\perp}'\sum_t X_{t-1}X_{t-1}'\beta_{\perp}$$

$$=\begin{pmatrix} 0 & I_{p-r}\end{pmatrix}\left\{\sum_t C_{t-1}(L)\,S_{t-1}S_{t-1}' C_{t-1}(L)'+r_4\right\}\begin{pmatrix} 0 \\ I_{p-r}\end{pmatrix}, \quad (3.32)$$

with

$$r_4 \equiv \sum_t \sum_{i=0}^{t-2} c_i S_{t-i-1} b_{t-1}' + \sum_t \sum_{i=0}^{t-2} b_{t-1} S_{t-i-1}' c_i' + \sum_t b_{t-1} b_{t-1}',$$

which by arguing in the usual manner may be seen to be $O_p(1)$. Now, using lemma 3.3, we have

$$C_t(L) = C_t(1) + (1-L) C_t^{(1)}(L), \quad C_t^{(1)}(L) = \sum_{i=0}^{t-1} c_i^{(1)} L^i,$$

so the main term of (3.32) satisfies

$$\sum_t C_{t-1}(L) S_{t-1} S_{t-1}' C_{t-1}(L)' = \sum_t C_{t-1}(1) S_{t-1} S_{t-1}' C_{t-1}(1)' + r_5, \qquad (3.33)$$

where

$$\begin{aligned} r_5 \equiv & \sum_t \sum_{i=0}^{t-2} \sum_{j=0}^{t-2} c_i^{(1)} \varepsilon_{t-i} S_{t-j-1}' c_j' + \sum_t \sum_{i=0}^{t-2} \sum_{j=0}^{t-2} c_i S_{t-i-1} \varepsilon_{t-j}' c_j^{(1)'} \\ & + \sum_t \sum_{i=0}^{t-2} \sum_{j=0}^{t-2} c_i^{(1)} \varepsilon_{t-i} \varepsilon_{t-j}' c_j^{(1)'}. \end{aligned} \qquad (3.34)$$

But the main term of the r.h.s. of (3.33) is handled in lemma 3.5, and we are left to show that r_5 is $O_p(T)$. Now, looking at the first term on the r.h.s. of (3.34),

$$\sum_t \sum_{i=0}^{t-2} \sum_{j=0}^{t-2} c_i^{(1)} \varepsilon_{t-i} S_{t-j-1}' c_j' = \sum_{i=0}^{T-2} \sum_{j=0}^{T-2} c_i^{(1)} \left(\sum_t \varepsilon_{t-i} S_{t-j-1}' \right) c_j',$$

where by lemma 3.4, the stochastic sum of the r.h.s. is (at most) $O_p(T)$, making the whole expression $O_p(T)$ by the usual arguments. It is clear that the remaining terms of (3.34) may be handled in a similar way. Hence, via (3.31)-(3.33), (3.28) follows.

To prove (3.26), observe that

$$S_{00} \alpha_\perp = \left(S_{00} - S_{01} \beta \alpha' \right) \alpha_\perp, \qquad (3.35)$$

and moreover, it follows as in (3.21) that via (3.20),

$$T \left(S_{00} - S_{01} \beta \alpha' \right) = \sum_t Z_{0t} \varepsilon_t' - M_{02} M_{22}^{-1} \sum_t Z_{2t} \varepsilon_t'. \qquad (3.36)$$

Now, by lemma 3.3,

$$\sum_t Z_{0t}\varepsilon_t' = \sum_t \Delta X_t\varepsilon_t' = \sum_t D_t(L)\,\varepsilon_t\varepsilon_t' + \sum_t a_t^*\varepsilon_t', \tag{3.37}$$

where in the usual fashion, $\sum_t a_t^*\varepsilon_t' = O_p(1)$ while

$$\sum_t D_t(L)\,\varepsilon_t\varepsilon_t' = \sum_t \varepsilon_t\varepsilon_t' + \sum_{i=1}^{T-1} d_i \sum_t \varepsilon_{t-i}\varepsilon_t', \tag{3.38}$$

where by lemma 3.4, the first term of the r.h.s. is $O_p(T)$ and the second term is $O_p\left(T^{1/2}\right)$. Moreover, observe that via lemma 3.5, the second term of the r.h.s. of (3.36) is $O_p\left(T^{1/2}\right)$. Hence, (3.35)-(3.38) yield (3.26).

Finally, as for (3.27) it follows in the same fashion as above that

$$\beta_\perp' S_{10}\alpha_\perp = \beta_\perp'\left(S_{10} - S_{11}\beta\alpha'\right)\alpha_\perp, \tag{3.39}$$

where

$$T\beta_\perp'\left(S_{10} - S_{11}\beta\alpha'\right) = \sum_t \beta_\perp' X_{t-1}\varepsilon_t' - \beta_\perp' M_{12}M_{22}^{-1}\sum_t Z_{2t}\varepsilon_t', \tag{3.40}$$

the second term of the r.h.s. being $O_p\left(T^{1/2}\right)$ by lemma 3.5. As for the first term, lemma 3.3 gives

$$\sum_t \beta_\perp' X_{t-1}\varepsilon_t' = \begin{pmatrix} 0 & I_{p-r} \end{pmatrix}\left\{\sum_t C_{t-1}(L)\,S_{t-1}\varepsilon_t' + \sum_t b_{t-1}\varepsilon_t'\right\}, \tag{3.41}$$

where as usual, $\sum_t b_{t-1}\varepsilon_t'$ is $O_p(1)$ and where

$$\sum_t C_{t-1}(L)\,S_{t-1}\varepsilon_t' = \sum_t C_{t-1}(1)\,S_{t-1}\varepsilon_t' + \sum_{i=1}^{T-2} c_i^{(1)}\sum_t \varepsilon_{t-i}\varepsilon_t'. \tag{3.42}$$

Here, the first term is handled in lemma 3.5, and the second term is $O_p\left(T^{1/2}\right)$ by lemma 3.4. Thus, (3.39)-(3.42) yield (3.27). ■

Observe that, when inserting the main terms of this lemma into the result of lemma 3.6, we are left with the main term of the corresponding expression for $-2\log Q_T^{(1)}$. (In the $m = 1$ case, there are no $O_p\left(T^{-1/2}\right)$ rest terms.) Hence, our theorem is proved if we can show that the $O_p\left(T^{1/2}\right)$ terms of lemma 3.6 give rise only to a $O\left(T^{-1}\right)$ term when calculating moments. For this, we need

Lemma 3.8. *The correlations between the elements of* $\sum_{i=0}^{T-2} c_i \sum_t \varepsilon_{t-k-i}\varepsilon_t'$ *for* $k > 0$ *and the elements of* $\sum_t S_{t-1}\varepsilon_t'$ *and* $\sum_t S_{t-1}S_{t-1}'$ *respectively, are all of order* $T^{-1/2}$ *as* $T \to \infty$*. Replacing* c_i *by* $c_i^{(1)}$ *or* d_i*, the same statement holds true.*

Proof. Denote the elements of Ω, the covariance matrix of the ε_t's, by ω_{ij}, and the elements of the vectors ε_t and S_t by $\varepsilon_t^{(i)}$and $S_t^{(i)}$, $i = 1, ..., p$, respectively. Then, the covariance between the (m, n)'th element of $\sum_t S_{t-1}\varepsilon_t'$ and the (r, s)'th element of $\sum_t \varepsilon_{t-k-i}\varepsilon_t'$ is, the latter having expectation zero,

$$\begin{aligned} E\left(\sum_t S_{t-1}^{(m)}\varepsilon_t^{(n)} \sum_u \varepsilon_{u-k-i}^{(r)}\varepsilon_u^{(s)}\right) &= \sum_t\sum_u E\left(S_{t-1}^{(m)}\varepsilon_t^{(n)}\varepsilon_{u-k-i}^{(r)}\varepsilon_u^{(s)}\right) \\ &= \sum_t E\left(S_{t-1}^{(m)}\varepsilon_{t-k-i}^{(r)}\right) E\left(\varepsilon_t^{(n)}\varepsilon_t^{(s)}\right) \\ &= T\omega_{mr}\omega_{ns}. \end{aligned} \tag{3.43}$$

Now, denoting the elements of c_i by $c_i^{(r,q)}$, the (r, s)'th element of $\sum_{i=0}^{T-2} c_i \sum_t \varepsilon_{t-k-i}\varepsilon_t'$ is

$$a^{(r,s)} \equiv \sum_{q=1}^{p}\sum_{i=0}^{T-2} c_i^{(r,q)} \sum_t \varepsilon_{t-k-i}^{(q)}\varepsilon_t^{(s)},$$

and by (3.43),

$$E\left(\sum_t S_{t-1}^{(m)}\varepsilon_t^{(n)} a^{(r,s)}\right) = T\sum_{q=1}^{p}\sum_{i=0}^{T-2} c_i^{(r,q)}\omega_{mq}\omega_{ns},$$

which, because of the exponential decay of the c_i's, is $O(T)$. Hence, because the variances of $\sum_t S_{t-1}^{(m)}\varepsilon_t^{(n)}$ and $a^{(r,s)}$ are $O(T^2)$ and $O(T)$, respectively, the correlation is $O\left(T^{-1/2}\right)$.

The corresponding statement with $\sum_t S_{t-1}S_{t-1}'$ in place of $\sum_t S_{t-1}\varepsilon_t'$ follows in a similar manner. Further, since the $\left\{c_i^{(1)}\right\}$ and $\{d_i\}$ sequences decay exponentially in the same way as the $\{c_i\}$ sequence does, we may replace $\{c_i\}$ by either $\left\{c_i^{(1)}\right\}$ or $\{d_i\}$ without altering the result. ■

At last, we are able to prove theorem 1.1.

Proof of theorem 1.1: Inserting the results of lemma 3.7 into the main term of (3.16) and Taylor expanding, we obtain

$$-2\log Q_T^{(m)} = U + V + O_p\left(T^{-1}\right), \tag{3.44}$$

where $U = -2\log Q_T^{(1)} = O_p(1)$ and where V is $O_p\left(T^{-1/2}\right)$, composed by deterministic terms, $\sum_t S_{t-1}\varepsilon_t^{'}$, $\sum_t S_{t-1}S_{t-1}^{'}$ and terms of the type $\sum_{i=0}^{T-2} c_i \sum_t \varepsilon_{t-k-i}\varepsilon_t^{'}$, $k > 0$. No other stochastic terms enter, because we may replace the quantities $S_{\beta\beta}$, M_{22}, $\beta^{'} M_{12}$ and M_{02} appearing in the rest terms R, R_1^* and R_2^* (cf (3.17), (3.29) and (3.30), respectively) by their (deterministic) limits in probability, making errors of order $T^{-1/2}$ which contribute to the overall error term to the order T^{-1}. Indeed, we e.g. have

$$S_{\beta\beta} = \beta^{'}\left(M_{11} - M_{12}M_{22}^{-1}M_{21}\right)\beta,$$

where by lemma 3.3,

$$\begin{aligned}
\beta^{'} M_{11}\beta &= T^{-1}\beta^{'}\sum_t X_{t-1}X_{t-1}^{'}\beta \\
&= T^{-1}\begin{pmatrix} I_r & 0 \end{pmatrix}\sum_t C_{t-1}(L)\,\varepsilon_t\varepsilon_t^{'}C_{t-1}(L)^{'}\begin{pmatrix} I_r \\ 0 \end{pmatrix} + O_p\left(T^{-1}\right),
\end{aligned}$$

with, in the usual fashion,

$$\begin{aligned}
T^{-1}\sum_t C_{t-1}(L)\,\varepsilon_t\varepsilon_t^{'}C_{t-1}(L)^{'} &= T^{-1}\sum_t\sum_{i=0}^{t-2}\sum_{j=0}^{t-2} c_i\varepsilon_{t-i}\varepsilon_{t-j}^{'}c_j^{'} \\
&= T^{-1}\sum_{i=0}^{T-2} c_i\sum_t \varepsilon_{t-i}\varepsilon_{t-i}^{'}c_i^{'} + O_p\left(T^{-1/2}\right) \\
&= \sum_{i=0}^{T-2} c_i\Omega c_i^{'} + O_p\left(T^{-1/2}\right),
\end{aligned}$$

where $\sum_{i=0}^{T-2} c_i\Omega c_i^{'}$ is $O_p(1)$. It is clear that $\beta^{'} M_{12}$ and M_{22} may be treated similarly, proving that $S_{\beta\beta}$ is $O_p(1)$ and has a deterministic limit. Further, the factor $\beta_\perp^{'} S_{10}$ (which is not treated in lemma 3.7) entering equation (3.17) causes no trouble, since by arguing as in (3.15), this factor is seen to be $O_p(1)$, with a dominating term composed by terms of the type $T^{-1}\sum_t S_{t-1}\varepsilon_t^{'}$. A similar statement is true regarding the factor $\beta_\perp^{'} M_{12}$ appearing in (3.30).

Now, a simple binomial expansion in (3.44) yields

$$E\left\{\left(-2\log Q_T^{(m)}\right)^n\right\} = E\left(U^n\right) + nE\left(U^{n-1}V\right) + O_p\left(T^{-1}\right),$$

where

$$\begin{aligned} E\left(U^{n-1}V\right) &= Corr\left(U^{n-1},V\right)\left\{Var\left(U^{n-1}\right)Var\left(V\right)\right\}^{1/2} \\ &= Corr\left(U,V\right)O\left(T^{-1/2}\right). \end{aligned}$$

But because of lemma 3.8, we must have $Corr\left(U,V\right) = O\left(T^{-1/2}\right)$ since applying any almost surely smooth function to two random variables which are correlated to a certain asymptotic order does not change this order. This completes the proof of our theorem. ■

4. Concluding remarks

The main result of this paper may be generalized in many interesting directions. For example, for a practitioner it is often necessary to include constants, linear terms etc. in the model, a practice which actually affects the limit distribution of the likelihood ratio test for cointegration (cf Johansen (1995), chap. 6), but maybe not the order of the type of asymptotic error term studied in this paper. Also, it would of course be of interest to explicitly study how the nuisance parameter matrices Γ_1,...,Γ_{m-1} affect the $O_p\left(T^{-1}\right)$ error term of our theorem. As for unit root testing in an AR(2) process, this task was brought up in section 3 of Larsson (1998a). More interesting however (for example in the context of panel data), would be to relate the moments of the VAR(1) likelihood ratio cointegration test to the asymptotic distribution, i.e. to prove (1.3), and to see to what extent α and β affect the so produced error term. In the univariate case, related problems are studied in e.g. Larsson (1997), Larsson (1998b) and Nielsen (1997).

5. References

Chung, K. L. (1974). *A Course in Probability Theory,* Academic Press, Orlando.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models,* Oxford University Press, Oxford.

Larsson, R. (1997). On the asymptotic expectations of some unit root tests in a first order autoregressive process in the presence of trend, *Annals of the Institute of Statistical Mathematics,* **49**, 585-599.

Larsson, R. (1998a). The order of the error term for moments of the log likelihood ratio unit root test in an autoregressive process, *Annals of the Institute of Statistical Mathematics,* **50**, 29-48.

Larsson, R. (1998b). Bartlett corrections for unit root test statistics, *Journal of Time Series Analysis,* **19**, 425-438.

Larsson, R., Lyhagen, J. and Löthgren, M. (1998). Likelihood-Based Cointegration Tests in Heterogeneous Panels, Working Paper in Economics and Finance No. 250, Stockholm School of Economics.

Mann, H.B. and Wald, A. (1943). On stochastic limit and order relationships, *Annals of Mathematical Statistics,* **14**, 217-226.

Nielsen, B. (1997). Bartlett correction of the unit root test in autoregressive models, *Biometrika,* **84**, 500-504.