A Bartlett correction factor for tests on the cointegrating relations

Johansen, Søren

Published in:
Econometric Theory

Publication date:
2000

Document version
Publisher's PDF, also known as Version of record

Citation for published version (APA):
Johansen, S. (2000). A Bartlett correction factor for tests on the cointegrating relations. *Econometric Theory*, *16*(5), 740-778.

Econometric Theory, **16**, 2000, 740–778. Printed in the United States of America.

A BARTLETT CORRECTION FACTOR FOR TESTS ON THE COINTEGRATING RELATIONS

SØREN JOHANSEN
European University Institute

Likelihood ratio tests for restrictions on cointegrating vectors are asymptotically χ^2 distributed. For some values of the parameters this asymptotic distribution does not give a good approximation to the finite sample distribution. In this paper we derive the Bartlett correction factor for the likelihood ratio test and show by some simulation experiments that it can be a useful tool for making inference.

1. INTRODUCTION

In this paper we derive a Bartlett correction for tests on cointegrating relations in the vector autoregressive model for the n-dimensional process X_t given by

$$\Delta X_t = \alpha(\beta' X_{t-1} + \rho' D_t) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi d_t + \varepsilon_t, \qquad t = 1, \dots, T, \tag{1}$$

where ε_t are independent and identically distributed (i.i.d.) $N_n(0, \Omega)$, the initial values are fixed, and d_t $(n_d \times 1)$ and D_t $(n_D \times 1)$ are deterministic terms, such as constant, linear term, etc. The matrices α and β are $(n \times r)$ and the matrices $\Gamma_1, \dots, \Gamma_{k-1}$, are $(n \times n)$, Φ $(n \times n_d)$, and ρ $(n_D \times r)$. For the calculation of the expectation in this paper we assume that X_t is $I(1)$. Conditions for this in terms of the parameters are given in Johansen (1996, Theorem 4.2).

The likelihood ratio test for hypotheses on β has been treated in Johansen and Juselius (1990) and Ahn and Reinsel (1990), and it is known that it is asymptotically χ^2 distributed, despite the fact that the asymptotic distribution of the estimator is mixed Gaussian. Finite sample distributions, however, are not always well approximated by asymptotic distributions (see, e.g., to mention a few of many studies of finite sample properties of tests of restrictions on β, Fachin, 1997; Gredenhoff and Jacobson, 1998; Jacobson, Vredin, and Warne, 1998; and Haug, 1998).

We derive here a correction term to the likelihood ratio test statistic for hypotheses on β with the purpose of improving the approximation to the asymptotic χ^2 distribution. The correction is the so-called Bartlett correction (see

Address correspondence to: Søren Johansen, Department of Economics, European University Institute, Badia Fiesolana, San Domenico di Fiesole, via dei Roccettini 9, I-50016, Italy.

0266-4666/00 $9.50

Bartlett, 1937). For a recent survey of the theory of this type of correction see Cribari-Neto and Cordeiro (1996). Briefly, the method consists of calculating the expectation of the likelihood ratio (LR) statistic in the form $-2\log LR$ for a given parameter point θ under the null hypothesis. Usually it is not possible to do this explicitly, and one can instead find an approximation of the form

$$E_\theta[-2\log LR] = f\left(1 + \frac{B(\theta)}{T}\right) + O(T^{-3/2}),$$

where f is equal to the degrees of freedom for the test and $B(\theta)$ shows how the remaining parameters under the null hypothesis distort the mean and hence the distribution of the test statistic. The idea is that the quantity

$$\frac{-2\log LR}{1 + \dfrac{B(\hat\theta)}{T}}$$

has expectation $f + O_P(T^{-3/2})$, thus presumably a distribution that is closer to the limit distribution. Note that estimation of θ gives an extra error that is absorbed into the term $O_P(T^{-3/2})$. Lawley (1956) proved that under assumptions of i.i.d. variables, the same correction improves not only the mean but all moments, thus giving a mathematical explanation of why the correction works so well in practice. No similar theorem has been proved in the case of $I(1)$ variables. Still it is of interest to calculate the Bartlett correction to see how it works in practice.

The model (1) is characterized by dimension (n), cointegrating rank (r), lag length (k), the number of deterministic terms restricted to the cointegrating space (n_D), the number of unrestricted terms (n_d) and finally of course the value of all the parameters and the sample size (T).

The main result presented in Section 4 is that the Bartlett correction is a function of the parameters through only two functions, and various combinations of the preceding numbers. We find, for instance, for the test for the same restrictions on all cointegrating relations β, that is, $\beta = H\tau(H(n \times s))$, that $f = r(n-s)$ and with $m = n + s - r + 1 + 2n_D$ we get

$$B(\theta) = \left[\frac{1}{2}m + n_d + kn\right] + \frac{1}{r}\left[(n - 2(r+1) + m)v(\alpha) + 2(c(\alpha) + c_d(\alpha))\right].$$

The coefficients $v(\alpha), c(\alpha)$, and $c_d(\alpha)$ are given in Theorem 4, which follows.

It will be apparent from the examples and the simulations in Section 5 that the influence of the parameters is crucial. For some parameter values the usual χ^2 approximation works well, whereas for others the correction factor is a useful improvement. Finally there are parameters points close to the boundary where the order of integration or the number of cointegrating relations change, and where the correction does not work well.

The plan of the paper is first to establish in Section 2 that a number of hypotheses can be given a general formulation as tests in a reduced rank regression model. In Section 3 an expansion is given of the estimators of this reduced rank regression. In Section 4 the main results on an expansion of the log likelihood ratio test and the Bartlett correction are given and the results specialized to the models discussed in Section 2, and finally in Section 5 some simulation experiments are conducted that show that the Bartlett correction is a useful addition to the usual asymptotic analysis. The very long and tedious proofs are given in an Appendix.

2. THE MODELS AND THE HYPOTHESES

We define in this section three models by restrictions on the cointegrating relations. All models can be analyzed by reduced rank regression (for a detailed analysis, see Johansen, 1996). The models allow deterministic terms of a suitably simple type, covering many usual situations. We show how correction terms for the tests of each of the models can be calculated if we have the correction term for a simple hypothesis. We show for each of the models how to formulate the test of a simple hypothesis as a test in a reduced rank regression, such that all the tests can be given the same uniform formulation. We first define the three models.

$\mathcal{M}_0$ Unrestricted Cointegrating Space. The model is given by equation (1) with unrestricted parameters.

$\mathcal{M}_1$ Same Restriction on All Cointegrating Relations. The model is defined as a submodel of $\mathcal{M}_0$ by the same restrictions on all cointegrating relations. This is expressed as

$$\beta = H\tau,$$

where H is $(n \times s)$ of rank s and known, $r \leq s < n$, and τ is $(s \times r)$ and unknown. The likelihood ratio test of $\mathcal{M}_1$ in $\mathcal{M}_0$ satisfies

$$-2 \log LR(\mathcal{M}_1 | \mathcal{M}_0) \xrightarrow{w} \chi^2(r(n-s)).$$

The restrictions on β can also be expressed as restrictions on (β, ρ) in the form

$$\begin{pmatrix} \beta \\ \rho \end{pmatrix} = \begin{pmatrix} H & 0_{n \times n_D} \\ 0_{n_D \times s} & I_{n_D} \end{pmatrix} \tau,$$

with τ $((s + n_D) \times r)$. One could also define a model by restricting simultaneously both β and ρ, but the present choice seems more relevant for the applications.

$\mathcal{M}_2$ Some Cointegrating Relations Known. The model is defined by the restrictions

$$\begin{pmatrix} \beta \\ \rho \end{pmatrix} = \left(\begin{pmatrix} \beta_1 \\ \rho_1 \end{pmatrix}, \begin{pmatrix} \beta_2^0 \\ \rho_2^0 \end{pmatrix} \right),$$

where the matrices β_2^0 $(n \times r_2)$ of rank r_2 and ρ_2^0 $(n_D \times r_2)$ are known and the matrices β_1 $(n \times r_1)$ and ρ_1 $(n_D \times r_1)$ are unknown $(r = r_1 + r_2)$, corresponding to prespecified coefficients β_2^0 and ρ_2^0 in some of the cointegrating relations. The likelihood ratio test of $\mathcal{M}_2$ in $\mathcal{M}_0$ satisfies

$$-2 \log LR(\mathcal{M}_2 | \mathcal{M}_0) \xrightarrow{w} \chi^2(r_2(n + n_D - r)).$$

The degrees of freedom is calculated in Johansen (1996, Theorem 7.3). It would also be relevant to formulate here the restriction that only β was partly known. This model, however, can not be estimated by reduced rank regression, and the analysis given subsequently would have to be modified.

In the following sections we derive a correction factor for the test of a simple hypothesis on β and ρ in each of the models $\mathcal{M}_0$, $\mathcal{M}_1$, and $\mathcal{M}_2$ and apply these to derive a correction factor for the test of $\mathcal{M}_1$ in $\mathcal{M}_0$ and $\mathcal{M}_2$ in $\mathcal{M}_0$ using the following trick.

To test $\mathcal{M}_1$ in $\mathcal{M}_0$, say, we take parameters $\beta^0 = H\tau^0$ and ρ^0 corresponding to a parameter point in $\mathcal{M}_1$. We define the concentrated likelihood function $L(\beta, \rho)$ and find the likelihood ratio test

$$\begin{aligned} LR(\mathcal{M}_1 | \mathcal{M}_0) &= \frac{\max_{\beta = H\tau, \rho} L(\beta, \rho)}{\max_{\beta, \rho} L(\beta, \rho)} \\ &= \frac{\max_{\beta=\beta^0, \rho=\rho^0} L(\beta, \rho)}{\max_{\beta, \rho} L(\beta, \rho)} \Bigg/ \frac{\max_{\beta=\beta^0, \rho=\rho^0} L(\beta, \rho)}{\max_{\beta = H\tau, \rho} L(\beta, \rho)} \\ &= LR(\beta = \beta^0, \rho = \rho^0 | \mathcal{M}_0)/LR(\beta = \beta^0, \rho = \rho^0 | \mathcal{M}_1), \end{aligned}$$

such that

$$\begin{aligned} &-2 \log LR(\mathcal{M}_1 | \mathcal{M}_0) \\ &\quad = -2 \log LR(\beta = \beta^0, \rho = \rho^0 | \mathcal{M}_0) + 2 \log LR(\beta = \beta^0, \rho = \rho^0 | \mathcal{M}_1). \end{aligned}$$

Hence we see that the correction for the test of $\mathcal{M}_1$ in $\mathcal{M}_0$ is the difference between the corrections to two tests of simple hypotheses on β and ρ in $\mathcal{M}_0$ and $\mathcal{M}_1$.

2.1. The Deterministic Terms

The correction will depend on the deterministic terms, and to get reasonably simple expressions we assume that they satisfy the relation

$$d_{t+h} = M^h d_t, \qquad h = \ldots, -1, 0, 1, \ldots \tag{2}$$

for some matrix M with the property that

$$|\mathrm{eig}(M)| = 1. \tag{3}$$

Further we assume that

$$\Delta D_t = K' d_t \tag{4}$$

for some $(n_d \times n_D)$ matrix K. Finally we assume that $(D_t, d_t)_{t=1}^T$ are linearly independent. Thus we allow, for instance, $d_t' = (1, t, t^2)$ and $D_t = t^3$, in which case

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}, \quad K = \begin{pmatrix} 1 \\ -3 \\ 3 \end{pmatrix}$$

and M has eigenvalues equal to 1. If s_1, s_2, and s_3 are quarterly dummies we can consider combinations such as $d_t' = (1, t, s_1(t), s_2(t), s_3(t))$. In this case we have $s_1(t+1) = s_4(t) = 1 - s_1(t) - s_2(t) - s_3(t)$, $s_2(t+1) = s_1(t)$, and $s_3(t+1) = s_2(t)$ such that

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & -1 & -1 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

which has eigenvalues ± 1, $\pm i$. Note that intervention dummies are not covered by this formulation and will give rise to more complicated formulae. The formulation also allows, with minor modifications, the possibility that n_D and n_d are zero.

LEMMA 1. *If X_t is $I(1)$ and given by equation* (1) *and if* (2), (3), *and* (4) *hold, then $E(\beta' X_{t-1} + \rho' D_t)$ and $E(\Delta X_t)$ are linear functions of d_t.*

Proof. From Granger's representation theorem (see Johansen, 1996, Theorem 4.2), we find that because the process X_t is $I(1)$ it has the representation

$$X_t = C \sum_{i=1}^{t} (\varepsilon_i + \Phi d_i) + C(L)(\varepsilon_t + \alpha\rho' D_t + \Phi d_t) + A,$$

where $C(z) = \sum_{i=0}^{\infty} C_i z^i$, A depends on initial conditions, $\beta' A = 0$, and

$$C = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1}\alpha'_\perp.$$

Note that the condition that the process is $I(1)$ implies that $|\alpha'_\perp \Gamma \beta_\perp| \neq 0$, such that C is well defined. It follows that

$$\begin{aligned} E(\Delta X_t) &= C\Phi d_t + C(L)(\alpha\rho' K' + \Phi\Delta)d_t \\ &= \left[C\Phi + \sum_{i=0}^{\infty} C_i(\alpha\rho' K' M^{-i} + \Phi(M^{-i} - M^{-i-1}))\right] d_t = \tilde{K}' d_t, \end{aligned}$$

say. Taking expectations in (1) we find

$$\tilde{K}' d_t = \alpha E(\beta' X_{t-1} + \rho' D_t) + \sum_{i=1}^{k-1} \Gamma_i \tilde{K}' M^{-i} d_t + \Phi d_t,$$

which shows the result for $E(\beta' X_{t-1} + \rho' D_t)$. Note that the result that M^h grows at most as a polynomial in h (see Lemma 9 in the Appendix), shows that the sums are convergent, because C_i are exponentially decreasing. ■

We next show how simple hypotheses on β and ρ in $\mathcal{M}_0$, $\mathcal{M}_1$, and $\mathcal{M}_2$ give rise to regression equations that can be given the same formulation. This allows us to derive all the results from one general reduced rank regression equation.

2.2. A Simple Hypothesis on β and ρ in $\mathcal{M}_0$

The model equation is given by (1), and we consider the hypothesis $\beta = \beta^0, \rho = \rho^0$, such that under the null hypothesis

$$\Delta X_t = \alpha(\beta^{0\prime} X_{t-1} + \rho^{0\prime} D_t) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi d_t + \varepsilon_t,$$

which is easily estimated by regression of ΔX_t on $\beta^{0\prime} X_{t-1} + \rho^{0\prime} D_t$, lagged differences, and d_t.

It is convenient for the calculations to reparametrize the model by defining new parameters and regressors that involve the true value. In the rest of this section we therefore need a notation for the true value of the parameters and also for the parameters of the model. We also need a notation for the estimator under the null hypothesis and one for the estimator under the alternative. Thus, for instance, we let α denote the parameter, α^0 the true value of the parameter, for which we calculate the expectations, $\tilde{\alpha}$ the reduced rank estimator in the model, and $\hat{\alpha}$ the regression estimator under the null hypothesis.

We use the notation

$$\Psi = (\Gamma_1, \dots, \Gamma_{k-1}), \qquad \Gamma = I_n - \sum_{i=1}^{k-1} \Gamma_i.$$

Note that

$$(I_n - C\Gamma)\beta_\perp = (I_n - \beta_\perp(\alpha_\perp' \Gamma \beta_\perp)^{-1}\alpha_\perp' \Gamma)\beta_\perp = 0,$$

such that for $\bar{\beta} = \beta(\beta'\beta)^{-1}$

$$(I_n - C\Gamma) = (I_n - C\Gamma)(\bar{\beta}\beta' + \beta_\perp \bar{\beta}_\perp') = (I_n - C\Gamma)\bar{\beta}\beta'.$$

We therefore decompose the process into stationary and nonstationary components using the true values of the parameters:

$$X_t = (I_n - C^0\Gamma^0)\bar{\beta}^0\beta^{0\prime}X_t + C^0\Gamma^0 X_t.$$

It follows that

$$\beta' X_{t-1} = \beta'(I_n - C^0\Gamma^0)\bar{\beta}^0\beta^{0\prime}X_{t-1} + \beta'\beta_\perp^0(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1}. \tag{5}$$

We define new parameters ϕ and $\delta = (\delta_1', \delta_2')'$ by

$$\phi' = \beta'(I_n - C^0\Gamma^0)\bar{\beta}^0 \qquad (r \times r),$$

$$\delta_1' = \beta'\beta_\perp^0 \qquad (r \times (n-r)),$$

$$\delta_2' = \rho' - \beta'(I_n - C^0\Gamma^0)\bar{\beta}^0\rho^{0\prime} \qquad (r \times n_D),$$

such that the old parameters in terms of the new are given by

$$\beta' = \delta_1'(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 + \phi'\beta^{0\prime}, \qquad \rho' = \delta_2' + \phi'\rho^{0\prime}.$$

The hypothesis $\beta = \beta^0$, $\rho = \rho^0$ is expressed in the new parameters as $\delta = 0$ and $\phi = I_r$. Equation (1) with the new parameters is

$$\Delta X_t = \alpha\phi'(\beta^{0\prime}X_{t-1} + \rho^{0\prime}D_t) + \alpha(\delta_1'(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1} + \delta_2' D_t)$$
$$+ \sum_{i=1}^{k-1}\Gamma_i \Delta X_{t-i} + \Phi d_t + \varepsilon_t. \tag{6}$$

We absorb $\phi(r \times r)$ into $\alpha(n \times r)$ and adjust δ accordingly. The hypothesis of interest is then $\delta = 0$.

In the reduced rank regression (6) we use the result that $\beta^{0\prime}X_{t-1} + \rho^{0\prime}D_t$ and ΔX_t have a mean that is linear in d_t (see Lemma 1), and that Φ enters unrestrictedly, to replace the regressors $\beta^{0\prime}X_{t-1} + \rho^{0\prime}D_t$ and the lagged differences with the stationary regressors

$$V_{t-1} = \beta^{0\prime}X_{t-1} - E_0(\beta^{0\prime}X_{t-1}), \tag{7}$$

$$Z_{t-1} = (\Delta X_{t-1}' - E_0(\Delta X_{t-1}'), \dots, \Delta X_{t-k+1}' - E_0(\Delta X_{t-k+1}'))'. \tag{8}$$

We also want to replace the regressor $(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1}$ by something simpler without changing the statistical model and hence the test that $\delta = 0$. We find by summing equation (1) that

$$\alpha_\perp'(X_t - X_0) = \alpha_\perp' \sum_{i=1}^{k-1} \Gamma_i (X_{t-i} - X_{-i}) + \alpha_\perp' \sum_{i=1}^{t} (\varepsilon_i + \Phi d_i).$$

By subtracting $\sum_{i=1}^{k-1} \alpha_\perp' \Gamma_i X_t$ on both sides and replacing t by $t-1$ we get

$$\alpha_\perp' \Gamma X_{t-1} = \alpha_\perp' X_0 + \alpha_\perp' \sum_{i=1}^{k-1} \Gamma_i (X_{t-i-1} - X_{t-1} - X_{-i}) + \alpha_\perp' \sum_{i=1}^{t-1} (\varepsilon_i + \Phi d_i).$$

Because we are correcting for lagged differences in regression (6) we can replace $(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1}$ and D_t by the nonstationary regressor given by the common trends extended by D_t:

$$A_{t-1} = \begin{pmatrix} A_0 + \alpha_\perp^{0\prime} \sum_{i=1}^{t-1} (\varepsilon_i + \Phi^0 d_i) \\ D_t \end{pmatrix}, \tag{9}$$

where A_0 depends on initial conditions.

Equation (6) in the new variables and with suitably redefined parameters becomes

$$\underset{(n)}{\Delta X_t} = \alpha \underset{(r)}{V_{t-1}} + \alpha\delta' \underset{(n-r+n_D)}{A_{t-1}} + \Psi \underset{((k-1)n)}{Z_{t-1}} + \Phi \underset{(n_d)}{d_t} + \underset{(n)}{\varepsilon_t}, \tag{10}$$

where the dimensions are indicated below each variable. The estimators for the parameters δ, α, Ψ, Φ, and Ω can be found by reduced rank regression of ΔX_t on (V_{t-1}', A_{t-1}') corrected for Z_{t-1} and d_t. Under the hypothesis $\delta = 0$ the parameters can be found by regression of ΔX_t on V_{t-1}, Z_{t-1}, and d_t.

Later we choose A_{t-1} such that it is orthogonal to the deterministic term d_t, which simplifies some notation. Note that if $D_t = 0$ then, of course, we do not extend the process, and A_{t-1} is defined entirely in terms of the random walks and initial values. Note also that if d_t contains a constant then, when correcting for d_t, the initial values disappear.

2.3. A Simple Hypothesis on β and ρ in Model $\mathcal{M}_1$

In model $\mathcal{M}_1$ the cointegrating vectors β are restricted as $\beta = H\tau(\tau(s \times r))$, and equation (1) is

$$\Delta X_t = \alpha(\tau' H' X_{t-1} + \rho' D_t) + \sum_{i=1}^{k-1} \Gamma_i \Delta X_{t-i} + \Phi d_t + \varepsilon_t. \tag{11}$$

We consider again the simple hypothesis $\beta = H\tau^0$, $\rho = \rho^0$, corresponding to a point in $\mathcal{M}_1$. We want to show that by introducing the true parameters β^0,

$\rho^0, \dots$ as before we can reformulate the equations to have the form (10), such that a test of a simple hypothesis is again a test that $\delta = 0$.

We decompose the process X_{t-1} using the true value of the parameters and find

$$\beta' X_{t-1} = \tau' H'(I_n - C^0\Gamma^0)\bar{\beta}^0\beta^{0\prime}X_{t-1} + \tau' H'\beta_\perp^0(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1}.$$

In this case $\beta^0 = H\tau^0$, which implies that $\beta_\perp^0 = (H_\perp, H(H'H)^{-1}\tau_\perp^0)$ and hence

$$\beta'\beta_\perp^0 = \tau' H'\beta_\perp^0 = \tau'\tau_\perp^0(0_{(s-r)\times(n-s)}, I_{s-r}).$$

We introduce the new parameters ϕ and $\delta = (\delta_1', \delta_2')'$ by

$$\begin{aligned}
\phi' &= \tau' H'(I_n - C^0\Gamma^0)\bar{\beta}^0 && (r \times r),\\
\delta_1' &= \tau'\tau_\perp^0 && (r \times (s-r)),\\
\delta_2' &= \rho' - \tau' H'(I_n - C^0\Gamma^0)\bar{\beta}^0\rho^{0\prime} && (r \times n_D),
\end{aligned}$$

because then

$$\begin{aligned}
&\tau' H' X_{t-1} + \rho' D_t\\
&\quad = \phi'(\beta^{0\prime}X_{t-1} + \rho^{0\prime}D_t) + \delta_1'(0_{(s-r)\times(n-s)}, I_{s-r})(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1}\\
&\qquad + \delta_2' D_t.
\end{aligned}$$

The hypothesis is formulated as $\delta = 0$ and $\phi = I_r$. We let $V_{t-1}$and Z_{t-1} be defined by (7) and (8) and replace in this case the $(s-r)$-dimensional nonstationary regressor $(0_{(s-r)\times(n-s)}, I_{s-r})(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1}$ with $s-r$ linear combinations K_1 of the common trends extended by D_t:

$$A_{t-1} = \begin{pmatrix} K_0 + K_1\alpha_\perp^{0\prime}\sum_{i=1}^{t-1}(\varepsilon_i + \Phi^0 d_i) \\ D_t \end{pmatrix} \tag{12}$$

for some matrices K_0 $((s-r)\times 1)$ depending on initial conditions and K_1 $((s-r)\times(n-r))$. Equation (11) then becomes

$$\underset{(n)}{\Delta X_t} = \alpha \underset{(r)}{V_{t-1}} + \alpha\delta' \underset{(s-r+n_D)}{A_{t-1}} + \Psi \underset{((k-1)n)}{Z_{t-1}} + \Phi \underset{(n_d)}{d_t} + \underset{(n)}{\varepsilon_t}, \tag{13}$$

where ϕ' is absorbed in α and the remaining parameters are adjusted accordingly such that Z_{t-1} and V_{t-1} also have mean zero. The hypothesis of interest is $\delta = 0$, which corresponds to $\rho = \rho^0$, $\beta = \beta^0 = H\tau^0$.

This equation has the same structure as (10) except that the dimension of A_{t-1} is changed to $s - r + n_D$.

2.4. A Simple Hypothesis on β and ρ in Model $\mathcal{M}_2$

We decompose α corresponding to β into $\alpha = (\alpha_1, \alpha_2)$ and find that the relevant part of (1) is

$$\alpha(\beta' X_{t-1} + \rho' D_t) = \alpha_1(\beta_1' X_{t-1} + \rho_1' D_t) + \alpha_2(\beta_2^{0\prime} X_{t-1} + \rho_2^{0\prime} D_t).$$

A simple hypothesis on β and ρ is formulated as $\beta_1 = \beta_1^0$, $\rho_1 = \rho_1^0$. Without loss of generality we can choose β_1 orthogonal to β_2^0 and adjust ρ_1 accordingly.

In this case we include the regressor $(\beta_2^{0\prime} X_{t-1} + \rho_2^{0\prime} D_t)$ with the lagged differences Z_{t-1} instead of with V_{t-1}. We then decompose the first component $\beta_1' X_{t-1} + \rho_1' D_t$ of the process as before:

$$\begin{aligned}\beta_1' X_{t-1} &+ \rho_1' D_t \\ &= \beta_1'(I_n - C^0\Gamma^0)\bar{\beta}^0\beta^{0\prime} X_{t-1} + \beta_1'\beta_\perp^0(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1} + \rho_1' D_t.\end{aligned}$$

Now define $\phi = (\phi_1, \phi_2)$ by

$$\beta_1'(I_n - C^0\Gamma^0)\bar{\beta}^0\beta^{0\prime} X_{t-1} = (\phi_1', \phi_2')\begin{pmatrix}\beta_1^{0\prime} X_{t-1} \\ \beta_2^{0\prime} X_{t-1}\end{pmatrix} = \phi_1'\beta_1^{0\prime} X_{t-1} + \phi_2'\beta_2^{0\prime} X_{t-1},$$

so that with A_{t-1} defined by (9)

$$\beta_1' X_{t-1} + \rho_1' D_t = \phi_1'(\beta_1^{0\prime} X_{t-1} + \rho_1^{0\prime} D_t) + \phi_2'(\beta_2^{0\prime} X_{t-1} + \rho_2^{0\prime} D_t) + \delta' A_{t-1},$$

where we have defined the new parameters $\phi = (\phi_1', \phi_2')'$ and $\delta = (\delta_1', \delta_2')'$ by

$$\begin{aligned}(\phi_1', \phi_2') &= \beta_1'(I_n - C^0\Gamma^0)\bar{\beta}^0 && (r_1 \times r_1), (r_1 \times r_2), \\ \delta_1' &= \beta_1'\beta_\perp^0 && (r_1 \times (n - r)), \\ \delta_2' &= \rho_1' - \phi_1'\rho_1^{0\prime} - \phi_2'\rho_2^{0\prime} && (r_1 \times n_D).\end{aligned}$$

Note that $\delta = 0$ implies that β_1 is proportional to β_1^0, because β_1 is assumed orthogonal to β_2^0. Hence $(\phi_1, \phi_2)' = (\beta_1'\bar{\beta}_1^0, 0)$ such that the hypothesis of interest in the new parameters is $\delta = 0$ and $\phi_1 = I_{r_1}$. The equation becomes

$$\underset{(n)}{\Delta X_t} = \alpha_1 \underset{(r_1)}{V_{t-1}} + \alpha_1\delta' \underset{(n - r + n_D)}{A_{t-1}} + \Psi \underset{(r_2 + (k-1)n)}{Z_{t-1}} + \Phi \underset{(n_d)}{d_t} + \underset{(n)}{\varepsilon_t}, \tag{14}$$

where

$$V_{t-1} = \beta_1^{0\prime} X_{t-1}, \tag{15}$$

$$Z_{t-1} = (X_{t-1}'\beta_2^0, \Delta X_{t-1}', \ldots, \Delta X_{t-k+1}')', \tag{16}$$

both corrected for their mean, and where again α_1, Ψ, and Φ have been redefined to absorb ϕ_1 and to accommodate the changed regressors. It is seen that equation (14) is of the form (10), with a changed definition of V_{t-1} and Z_{t-1}, because the assumed stationary combinations $\beta_2^{0\prime} X_{t-1} - E_0(\beta_2^{0\prime} X_{t-1})$ are moved to the lagged differences. The hypothesis can be tested as $\delta = 0$.

Thus in a general formulation that covers all the hypotheses we are interested in, we need to allow the dimensions of the variables entering the equation to be different from those given in (10). But we still need to preserve the properties that under the null hypothesis the process $Y_{t-1} = (V'_{t-1}, Z'_{t-1})'$ is a mean zero stationary autoregressive process and that V_{t-1} and $\delta' A_{t-1}$ have the factor α (or α_1) in front. All models (10), (13), and (14) have the property that they are estimated by reduced rank regression and that under the null hypothesis, $\delta = 0$, the models are estimated by simple regression.

3. A REDUCED RANK EQUATION AND SOME EXPANSIONS

To cover the different cases considered in Section 2, we discuss expansion and Bartlett correction of the likelihood ratio test for the hypothesis $\delta = 0$ in the equation

$$\underset{(n)}{\Delta X_t} = \xi \underset{(n_v)}{V_{t-1}} + \xi\delta' \underset{(n_a)}{A_{t-1}} + \Psi \underset{(n_z)}{Z_{t-1}} + \Phi \underset{(n_d)}{d_t} + \underset{(n)}{\varepsilon_t}, \tag{17}$$

where ε_t are i.i.d. $N_n(0, \Omega)$ and the parameters $(\xi, \delta, \Psi, \Phi, \Omega)$ vary freely.

This notation covers the different situations considered for suitable choices of ξ and the regressors V_{t-1}, A_{t-1}, and Z_{t-1} and their dimensions (see Table 1).

In all cases A_{t-1} is a function of D_t and $\xi'_\perp \sum_{i=1}^{t-1}(\varepsilon_i + \Phi d_i)$, with $\xi = \alpha$ or α_1. The variables V_{t-1} and Z_{t-1} are, under the hypothesis $\delta = 0$, stationary with mean zero.

Note that the stacked $(r + (k-1)n)$–dimensional process $Y_t = (V'_t, Z'_t)'$ is the same for all cases and contains $\beta' X_t$ and the lagged differences corrected for their mean.

In the rest of the paper we refer to the true value, the one for which we calculate the expectation, without the superscript because that simplifies the notation. We introduce some notation for the many product moments that are

Table 1. The choice of dimensions and ξ in the general regression model (17) that corresponds to the hypotheses discussed in Section 2

Model	n_v	n_a	n_z	ξ
$\mathcal{M}_0$	r	$n - r + n_D$	$(k-1)n$	α
$\mathcal{M}_1$	r	$s - r + n_D$	$(k-1)n$	α
$\mathcal{M}_2$	r_1	$n - r + n_D$	$r_2 + (k-1)n$	α_1

needed in the analysis of the expectation, but first we introduce the normalized errors

$$B_t = (\xi_\perp' \Omega \xi_\perp)^{-1/2} \xi_\perp' \varepsilon_t, \qquad U_t = (\xi' \Omega^{-1} \xi)^{-1/2} \xi' \Omega^{-1} \varepsilon_t \tag{18}$$

and define the product moment matrices $M_{..}$ for the variables ΔX_t, B_t, ε_t, U_t, and d_t at time t but V_{t-1}, A_{t-1}, and Z_{t-1} lagged one period. Thus, for instance,

$$\sum_{t=1}^{T} \begin{pmatrix} \Delta X_t \\ V_{t-1} \\ \varepsilon_t \end{pmatrix} \begin{pmatrix} \Delta X_t \\ V_{t-1} \\ \varepsilon_t \end{pmatrix}' = \begin{pmatrix} M_{00} & M_{0v} & M_{0\varepsilon} \\ M_{v0} & M_{vv} & M_{v\varepsilon} \\ M_{\varepsilon 0} & M_{\varepsilon v} & M_{\varepsilon\varepsilon} \end{pmatrix}.$$

We also use the notation for any three process A_{t-1}, U_t, and V_{t-1}, say,

$$M_{uv.a} = M_{uv} - M_{ua} M_{aa}^{-1} M_{av},$$

and in particular we use a notation for the moment matrices corrected for the lagged differences Z_{t-1} and d_t, because many results look a bit simpler:

$$S_{uv} = M_{uv.z,d} = M_{uv} - M_{ud} M_{dd}^{-1} M_{dv} - M_{uz.d} M_{zz.d}^{-1} M_{zv.d}.$$

These moment matrices are natural when the likelihood function is concentrated with respect to Ψ and Φ.

The maximum likelihood estimators based upon (17) will be denoted by $\tilde{\delta}$, $\tilde{\xi}$, and $\tilde{\Omega}$. The first order conditions for the estimators in model (17) can be solved for each of the variables as

$$\tilde{\xi} = (S_{0v} + S_{0a} \tilde{\delta})(S_{vv} + \tilde{\delta}' S_{av} + S_{va} \tilde{\delta} + \tilde{\delta}' S_{aa} \tilde{\delta})^{-1}, \tag{19}$$

$$\tilde{\Omega} = T^{-1}(S_{00} - \tilde{\xi}(S_{vv} + \tilde{\delta}' S_{av} + S_{va} \tilde{\delta} + \tilde{\delta}' S_{aa} \tilde{\delta})\tilde{\xi}'), \tag{20}$$

$$\tilde{\delta} = S_{aa}^{-1}(S_{a0} - S_{av} \tilde{\xi}') \tilde{\Omega}^{-1} \tilde{\xi} (\tilde{\xi}' \tilde{\Omega}^{-1} \tilde{\xi})^{-1}. \tag{21}$$

Note that the equations cannot be solved simultaneously, because the estimators are expressed in terms of each other.

Under the null hypothesis $\delta = 0$, the estimators are denoted $\hat{\xi}$ and $\hat{\Omega}$:

$$\hat{\xi} = S_{0v} S_{vv}^{-1} = \xi + S_{\varepsilon v} S_{vv}^{-1},$$

$$\hat{\Omega} = T^{-1}(S_{00} - S_{0v} S_{vv}^{-1} S_{v0}) = T^{-1} S_{\varepsilon\varepsilon.v}.$$

We next expand the estimators $\tilde{\xi}$, $\tilde{\Omega}$, and $\tilde{\delta}$, not around the parameter point $(\xi, \Omega, 0)$ but around the estimator under the null $(\hat{\xi}, \hat{\Omega}, 0)$.

THEOREM 2. *The estimators* $\tilde{\xi}$, $\tilde{\Omega}$, *and* $\tilde{\delta}$ *can be expanded around* $\hat{\xi}$, $\hat{\Omega}$, *and* 0, *respectively*:

$$\tilde{\xi} - \hat{\xi} = [S_{\varepsilon a.v}\tilde{\delta} - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta} - \hat{\xi}\tilde{\delta}' S_{av}]S_{vv}^{-1} + O_P(T^{-2}), \tag{22}$$

$$(\hat{\Omega} - \tilde{\Omega}) = T^{-1}\hat{\Omega}Q_0 + T^{-2}\hat{\Omega}Q_1 + O_P(T^{-5/2}), \tag{23}$$

$$\hat{\Omega}Q_0 = S_{\varepsilon a.v}\tilde{\delta}\hat{\xi}' + \hat{\xi}\tilde{\delta}' S_{a\varepsilon.v} - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta}\hat{\xi}',$$

$$\hat{\Omega}Q_1 = [S_{\varepsilon a.v}\tilde{\delta} - \hat{\xi}\tilde{\delta}' S_{av} - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta}]\,(T^{-1}S_{vv})^{-1} \times [\tilde{\delta}' S_{a\varepsilon.v} - S_{va}\tilde{\delta}\hat{\xi}' - \tilde{\delta}' S_{aa}\tilde{\delta}\hat{\xi}'],$$

$$\tilde{\delta} = S_{aa}^{-1} S_{a\varepsilon.v}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} + O_P(T^{-1}S_{aa}^{-1/2}). \tag{24}$$

The proof of Theorem 2 is given in the Appendix. Notice that if $n_D = 0$, then $S_{aa} \in O_P(T^2)$, but if D_t is in the model the order of magnitude depends on D_t. Thus, we use S_{aa} to indicate an order of magnitude, such that, for example, $\tilde{\delta} \in O_P(S_{aa}^{-1/2})$.

We conclude this section by giving the likelihood ratio test for $\delta = 0$ in model (17). We find

$$LR^{2/T}(\delta = 0) = \frac{|\tilde{\Omega}|}{|\hat{\Omega}|} = \frac{|\hat{\Omega} - (\hat{\Omega} - \tilde{\Omega})|}{|\hat{\Omega}|} = |I_n - \hat{\Omega}^{-1}(\hat{\Omega} - \tilde{\Omega})|$$

such that by expanding $-2\log LR$ and applying (23) we get

$$\begin{aligned} -2\log LR &= -T\log|I_n - \hat{\Omega}^{-1}(\hat{\Omega} - \tilde{\Omega})| \\ &\overset{1}{=} T\,\mathrm{tr}\{\hat{\Omega}^{-1}(\hat{\Omega} - \tilde{\Omega})\} + \frac{T}{2}\,\mathrm{tr}\{(\hat{\Omega}^{-1}(\hat{\Omega} - \tilde{\Omega}))^2\}, \\ &\overset{1}{=} \mathrm{tr}\{Q_0\} + T^{-1}(\mathrm{tr}\{Q_1\} + \frac{1}{2}\,\mathrm{tr}\{Q_0^2\}). \end{aligned} \tag{25}$$

We have used the notation $\overset{d}{=}$ to indicate that we have kept terms of the order T^{-d}.

4. THE BARTLETT CORRECTION FACTOR

In this section we give the main result on the Bartlett correction. We first discuss briefly the idea of conditioning on the common trends. Then we calculate some coefficients that are needed in the formulation of the main result. Next we give a stochastic expansion of $-2\log LR$, and finally we state the main result on the Bartlett correction. The result is specialized to the various situations covered by the general formulation as discussed in Section 2.

We choose first to calculate the expectation of the likelihood ratio test statistic conditioning on the process $\xi_\perp' \varepsilon_t$. The argument for this is that it is easier to do so because many of the expressions derived involve ratios of quadratic forms and turn out to be possible to calculate, if we first condition on $\xi_\perp'\varepsilon_t$. Another argument is that the asymptotic distribution of $\hat{\beta}$ is mixed Gaussian, where the mixing variable is just the limit of $\sum_{i=1}^t \alpha_\perp' \varepsilon_i$, which are fixed when we condition on $\xi_\perp'\varepsilon_t$. The end result is that the first order correction term in the conditional mean does not depend on the conditioning variable, such that what we find is also a correction term for the unconditional mean.

4.1. The Conditioning Variables

The processes A_{t-1} and B_t (see (9), (12), and (18)) are defined in terms of D_t, $\xi_\perp' \sum_{i=1}^{t-1} \varepsilon_i$, and $\xi_\perp'\varepsilon_t$. It is convenient to orthogonalize A_{t-1} on the deterministic terms d_t such that in the following $M_{ad} = 0$. Note that if d_t contains a constant, then A_{t-1} no longer depends on the initial values, but in any case the limit of $T^{-1/2}A_{[Tu]}$ does not involve the initial values.

When we do not condition on $\xi_\perp'\varepsilon_t$ we have the following relations (see Chan and Wei, 1988):

$$\left(\sum_{t=1}^T A_{t-1}A_{t-1}'\right)^{-1} \sum_{t=1}^T A_{t-1}A_{t-1-k}' \xrightarrow{\text{a.s.}} I_{n_a}, \quad \text{for all } k, \tag{26}$$

$$T^{-1}\sum_{t=1}^T B_t B_{t-k}' \xrightarrow{\text{a.s.}} I_{n_b}, \quad \text{if } k = 0, \quad \text{and} \quad 0 \text{ if } k \neq 0, \tag{27}$$

$$T^{-1}\sum_{t=1}^T B_t(B_{t-k}'KB_{t-k})B_t' \xrightarrow{\text{a.s.}} \operatorname{tr}\{K\}I_{n_b}, \qquad k \neq 0 \tag{28}$$

for any $n_b \times n_b$ matrix K. Finally we have the weak limit

$$M_{ba}M_{aa}^{-1}M_{ab} \xrightarrow{w} \int_0^1 (dW)F'\left(\int_0^1 FF'du\right)^{-1}\int_0^1 F(dW)', \tag{29}$$

where the Brownian motion $W(u)$ is defined by

$$T^{-1/2}\sum_{t=1}^{[Tu]} B_t = (\xi_\perp'\Omega\xi_\perp)^{-1/2}\xi_\perp' T^{-1/2}\sum_{t=1}^{[Tu]}\varepsilon_t \xrightarrow{w} W(u),$$

of dimension $n_b = n - n_v$. The process F is defined as a linear transformation of the limit of A_{t-1} (see Johansen, 1996). If, for instance, $\xi = \alpha$, $d_t = 1$, and $D_t = 0$ then

$$A_{t-1} = \alpha_\perp'\left(\sum_{i=1}^{t-1}\varepsilon_i - T^{-1}\sum_{t=1}^T\sum_{i=1}^{t-1}\varepsilon_i\right) + \alpha_\perp'\Phi(t - \bar{t}).$$

For $\tau = \alpha_\perp'\Phi$ $(\neq 0)$ we see that the univariate process $\tau' A_{t-1}$ is dominated by the linear term, whereas the $n - r - 1$–dimensional process $\tau_\perp' A_{t-1}$ has only the random walk components. In this case F is of dimension $n_b = n_a = n - r$ and is given by

$$F_i(u) = W_i(u) - \int_0^1 W_i(u)du, \qquad i = 1, \ldots, n_a - 1, \qquad F_{n_a}(u) = u - \frac{1}{2}.$$

When conditioning on the sequence $\xi_\perp' \varepsilon_t$ we assume that relations (26)–(28) hold for the sequence we are fixing. We keep the notation $M_{ba} M_{aa}^{-1} M_{ab}$ because the terms involving this quantity cancel in the final result.

4.2. The Autoregressive Model

Before we formulate the main result we need some notation for the vector autoregressive process given in model (1), which is the basis for all the calculations. Under the null hypothesis the model is estimated by ordinary least squares of ΔX_t on V_{t-1}, Z_{t-1}, and d_t, and we therefore introduce the stacked process $Y_t = (X_t'\beta, \Delta X_t', \ldots, \Delta X_{t-k+2}')'$ corrected for its mean, that is, $Y_t = (V_t', Z_t')'$. It is in all cases of dimension $n_y = n_v + n_z = r + (k-1)n$ (see Table 1) and is a stationary autoregressive process given by the equation

$$Y_t = PY_{t-1} + Q\varepsilon_t,$$

where

$$P = \begin{pmatrix} I_r + \beta'\alpha & \beta'\Gamma_1 & \ldots & \beta'\Gamma_{k-2} & \beta'\Gamma_{k-1} \\ \alpha & \Gamma_1 & \ldots & \Gamma_{k-2} & \Gamma_{k-1} \\ 0 & I_n & \ldots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \ldots & I_n & 0 \end{pmatrix}, \qquad Q = \begin{pmatrix} \beta' \\ I_n \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{30}$$

We find the representation

$$Y_t = \sum_{\nu=0}^{\infty} P^\nu Q \varepsilon_{t-\nu} = \sum_{\nu=0}^{\infty} (\theta_\nu U_{t-\nu} + \psi_\nu B_{t-\nu}),$$

where we have decomposed ε_t into the components U_t and B_t (see (18)) and defined

$$\theta_\nu = P^\nu Q\xi(\xi'\Omega^{-1}\xi)^{-1/2}, \qquad \psi_\nu = P^\nu Q\Omega\xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1/2}.$$

We define the variance

$$\Sigma = \operatorname{Var}(Y_t) = \sum_{\nu=0}^{\infty} (\theta_\nu\theta_\nu' + \psi_\nu\psi_\nu') = \Sigma_{\theta\theta} + \Sigma_{\psi\psi}, \tag{31}$$

say, and find

$$\theta = \sum_{\nu=0}^{\infty} \theta_\nu = (I_{n_y} - P)^{-1} Q\xi(\xi'\Omega^{-1}\xi)^{-1/2}, \tag{32}$$

$$\psi = \sum_{\nu=0}^{\infty} \psi_\nu = (I_{n_y} - P)^{-1} Q\Omega\xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1/2}. \tag{33}$$

From the definition of P and Q it follows that

$$(I_{n_y} - P)(I_r, 0_{r\times(k-1)n})' = -Q\alpha.$$

By multiplying by the matrix $(I_{r_1}, 0_{r_1\times r_2})'$ we find because $\alpha_1 = \alpha(I_{r_1}, 0_{r_1\times r_2})'$

$$(I_{n_y} - P)(I_{r_1}, 0_{r_1\times(r_2+(k-1)n)})' = -Q\alpha_1,$$

such that with ξ either α or α_1 we have

$$(I_{n_y} - P)(I_{n_v}, 0_{n_v\times n_z})' = -Q\xi.$$

From (32) we then get with $\kappa_\xi = (\xi'\Omega^{-1}\xi)^{-1/2}$ that

$$\theta = -(I_{n_v}, 0_{n_v\times n_z})'\kappa_\xi = -\tilde{\kappa}_\xi, \tag{34}$$

say. Note that because $Y_t = (V_t', Z_t')'$ we get

$$\tilde{\kappa}_\xi'\Sigma^{-1}\tilde{\kappa}_\xi = \kappa_\xi(I_{n_v}, 0_{n_v\times n_z})\Sigma^{-1}(I_{n_v}, 0_{n_v\times n_z})'\kappa_\xi = \kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi. \tag{35}$$

4.3. The Main Results

We can finally state the main results. We start in Theorem 3 with a stochastic expansion of the likelihood ratio test for $\delta = 0$, which forms the basis for the calculation of the Bartlett correction factor in Theorem 4. The proofs are left for the Appendix, but we give here some corollaries to show explicitly how the correction can be used for the tests mentioned in Section 2. We first give the assumptions that are needed for the results to hold.

Assumption 1. We assume throughout that the process X_t given by (1) is $I(1)$, that the deterministic terms satisfy (2)–(4), and that the conditioning variables $\xi_\perp'\varepsilon_t$ and $\xi_\perp'\sum_{i=1}^{t}\varepsilon_i$ satisfy (26)–(28).

THEOREM 3. *Under Assumption* 1, *an expansion of the log likelihood ratio test for* $\delta = 0$ *based upon* (17) *is given by*

$$-2\log LR(\delta=0)$$

$$\stackrel{1}{=} T\,\mathrm{tr}\{S_{aa.b,v}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} + \frac{1}{2T}\,\mathrm{tr}\{(S_{ua}S_{aa}^{-1}S_{au})^2\}$$

$$+\ 2T\,\mathrm{tr}\{S_{aa}^{-1}S_{au.v,b}\kappa_\xi S_{vv}^{-1}S_{vb}S_{bb}^{-1}S_{ba.v}\}$$

$$+\ \mathrm{tr}\{\kappa_\xi S_{vv}^{-1}\kappa_\xi S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}S_{au}\}$$

$$+\ \mathrm{tr}\{S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_\xi^2 S_{vv}^{-1}S_{vb}S_{ba}\}$$

$$-\ \mathrm{tr}\{S_{aa}^{-1}S_{au}\kappa_\xi S_{vv}^{-1}S_{vb}S_{bv}S_{vv}^{-1}\kappa_\xi S_{ua}\}$$

$$-\ 2\,\mathrm{tr}\{S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_\xi S_{uv}S_{vv}^{-1}\kappa_\xi S_{ua}\}.$$

The first two terms are the test statistic for $\delta = 0$ if ξ were known (see Johansen, 1999), the next term is of the order $O_P(T^{-1/2})$ but has expectation $O(T^{-1})$, and the last four terms are of the order $O_P(T^{-1})$. The proof of Theorem 3, based upon the expansions in Theorem 2, is given in the Appendix.

THEOREM 4. *Under Assumption* 1, *the conditional expectation given* $\xi_\perp'\varepsilon_t$ *of the log likelihood ratio test for the hypothesis* $\delta = 0$ *in* (17) *is approximated as*

$$E[-2\log LR(\delta=0)|\xi_\perp'\varepsilon] \stackrel{1}{=} n_v n_a + \frac{n_v n_a}{T}\left[\frac{1}{2}(n_v+n_a+1)+n_d+n+n_z\right]$$

$$+\ \frac{n_a}{T}\left[(n-n_v+n_a-1)v(\xi)+2(c(\xi)+c_d(\xi))\right]$$

where

$$v(\xi) = \mathrm{tr}\{V_\xi\},\qquad V_\xi = \tilde\kappa_\xi\tilde\kappa_\xi'\Sigma^{-1},$$

$$c(\xi) = \mathrm{tr}\{P(I_{n_y}+P)^{-1}V_\xi\} + \mathrm{tr}\{[P\otimes(I_{n_y}-P)V_\xi][I_{n_y}\otimes I_{n_y}-P\otimes P]^{-1}\},$$

$$c_d(\xi) = \mathrm{tr}\{[M\otimes(I_{n_y}-P)V_\xi][I_{n_y}\otimes I_{n_y}-M\otimes P]^{-1}\}.$$

Here $\tilde\kappa_\xi$ *is given by* (34), P *is given by* (30), *and* M *is defined in Section* 2.1.

Note that the conditional expectation does not depend on the conditioning variable and hence equals the expectation. It will be seen from the proof that the correction term is the one derived in the situation where ξ is known (see Johansen, 1999), apart from a term equal to $T^{-1}n_a(n-n_v)v(\xi)$. Note that the coefficients v, c, and c_d depend on the choice of ξ. If $\xi = \alpha$, then from (35)

$$v(\alpha) = \mathrm{tr}\{(\alpha'\Omega^{-1}\alpha)^{-1}\Sigma_{\beta\beta}^{-1}\},$$

with $\Sigma_{\beta\beta} = \mathrm{Var}(\beta' X_t | \Delta X_t, \dots, \Delta X_{t-k+2})$. If, however, $\xi = \alpha_1$ then

$$v(\alpha_1) = \mathrm{tr}\{(\alpha_1' \Omega^{-1} \alpha_1)^{-1} \Sigma_{\beta_1\beta_1.\beta_2}^{-1}\},$$

with $\Sigma_{\beta_1\beta_1.\beta_2} = \mathrm{Var}(\beta_1' X_t | \beta_2' X_t, \Delta X_t, \dots, \Delta X_{t-k+2})$, corresponding to having moved $\beta_2' X_{t-1}$ from V_{t-1} to Z_{t-1}.

The coefficient $c_d(\xi)$ can be calculated simply in some cases, such as $d_t = (1, t)'$, because then $\mathrm{tr}\{M^h\} = n_d = 2$ for all h. This means that

$$\begin{aligned} c_d(\xi) &= \mathrm{tr}\{[M \otimes (I_{n_y} - P)V_\xi][I_{n_y} \otimes I_{n_y} - M \otimes P]^{-1}\} \\ &= \sum_{v=0}^{\infty} \mathrm{tr}\{[M \otimes (I_{n_y} - P)V_\xi][M^v \otimes P^v]\} \\ &= \sum_{v=0}^{\infty} \mathrm{tr}\{M^{v+1}\}\mathrm{tr}\{(I_{n_y} - P)V_\xi P^v\} = n_d\, \mathrm{tr}\{V_\xi\} = n_d v(\xi). \end{aligned}$$

Note that M can be replaced by M' because only $\mathrm{tr}\{M^{v+1}\}$ enters the result. If d_t contains seasonal dummies then $\mathrm{tr}\{M^h\}$ is a periodic function and a more complicated expression can be found. To understand the parameter function V_ξ that enters the expressions, note that the long-run variance of Y_t conditional on the common trends is given by $\theta\theta' = \tilde{\kappa}_\xi \tilde{\kappa}_\xi'$ (see (34)). Thus the matrix $V_\xi = \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1}$ measures the "ratio" between the conditional long-run variance and the unconditional variance of Y_t.

We specialize the result to the hypotheses discussed in Section 2.

COROLLARY 5. *The Bartlett correction for the test of a simple hypothesis $\beta = \beta^0$, $\rho = \rho^0$ in model $\mathcal{M}_0$, is given by*

$$\begin{aligned} &E[-2 \log LR(\beta = \beta^0, \rho = \rho^0 | \mathcal{M}_0)] \\ &\quad \stackrel{1}{=} r(n - r + n_D) + \frac{r(n - r + n_D)}{T}\left[\frac{1}{2}(n + n_D + 1) + n_d + kn\right] \\ &\qquad + \frac{(n - r + n_D)}{T}[(2(n - r) + n_D - 1)v(\alpha) + 2(c(\alpha) + c_d(\alpha))], \end{aligned}$$

where $v(\alpha)$, $c(\alpha)$, and $c_d(\alpha)$ are given in Theorem 4.

Proof. This follows from Theorem 4 by substituting $n_v = r$, $n_a = n - r + n_D$, $n_z = (k - 1)n$, $\xi = \alpha$ (see Table 1). ■

COROLLARY 6. *The Bartlett correction factor for the test of $\mathcal{M}_1 : \beta = H\tau$, with H $(n \times s)$, is given by*

$$E[-2\log LR(\mathcal{M}_1|\mathcal{M}_0)]/r(n-s)$$

$$\overset{1}{=} 1 + \frac{1}{T}\left[\frac{1}{2}(n+s-r+1+2n_D) + n_d + kn\right]$$

$$+ \frac{1}{\mathrm{Tr}}[(2n+s-3r-1+2n_D)v(\alpha) + 2(c(\alpha)+c_d(\alpha))],$$

where $v(\alpha)$, $c(\alpha)$, *and* $c_d(\alpha)$ *are given in Theorem* 4.

COROLLARY 7. *The Bartlett correction factor for the test of* $\mathcal{M}_2$,

$$\begin{pmatrix}\beta\\ \rho\end{pmatrix} = \left(\begin{pmatrix}\beta_1\\ \rho_1\end{pmatrix}, \begin{pmatrix}\beta_2^0\\ \rho_2^0\end{pmatrix}\right),$$

where the matrix β_2^0 $(n \times r_2)$ *is of rank* r_2 *and* ρ_2^0 *is* $(1 \times r_2)$, *is given by*

$$E[-2\log LR(\mathcal{M}_2|\mathcal{M}_0)]/r_2(n-r+n_D)$$

$$\overset{1}{=} 1 + \frac{1}{T}\left[\frac{1}{2}(n+n_D+1-r_1) + n_d + kn\right]$$

$$+ \frac{1}{\mathrm{Tr}_2}[(2(n-r)+n_D-1)(v(\alpha)-v(\alpha_1)) - r_2 v(\alpha_1)]$$

$$+ \frac{2}{\mathrm{Tr}_2}[(c(\alpha)-c(\alpha_1)+c_d(\alpha)-c_d(\alpha_1))],$$

where the coefficients $c(\cdot)$, $c_d(\cdot)$, *and* $v(\cdot)$ *are defined in Theorem* 4.

The proofs of these corollaries follow from Theorem 4 by using the trick explained in Section 2.

5. SIMULATION EXPERIMENTS

We report here some simple simulation experiments to illustrate the usefulness of the correction. We first give the result for the model with only one lag and one cointegrating relation, because we get complete information on how the correction works. We then present a few results where the data generating process (DGP) has been chosen so as to match the results obtained for real data analyzed elsewhere. Throughout the number of simulations is 10,000. When calculating the Bartlett correction we assume that we know the lag length and the cointegrating rank as given by the DGP, but we estimate the remaining parameters and hence the Bartlett factor.

5.1. The Model with One Cointegrating Vector and One Lag

The model is

$$\Delta X_t = \alpha\beta' X_{t-1} + \varepsilon_t.$$

We find the coefficients

$$\operatorname{Var}(\beta' X_t) = \Sigma = \frac{\beta'\Omega\beta}{1-(1+\beta'\alpha)^2}, \qquad \kappa_\alpha^2 = (\alpha'\Omega^{-1}\alpha)^{-1},$$

$$v(\alpha) = V_\alpha = \Sigma^{-1}\kappa_\alpha^2 = -\frac{\beta'\alpha(2+\beta'\alpha)}{\alpha'\Omega^{-1}\alpha\beta'\Omega\beta},$$

$$P = 1+\beta'\alpha,\ c(\alpha) = -2\,\frac{\beta'\alpha(1+\beta'\alpha)}{\alpha'\Omega^{-1}\alpha\beta'\Omega\beta}.$$

With this notation we find the following corollary from Corollary 5.

COROLLARY 8. *In the model $\Delta X_{t-1} = \alpha\beta' X_{t-1} + \varepsilon_t$ with one cointegrating relation, the Bartlett correction factor for the hypothesis $\beta = \beta^0$ is*

$$E[-2\log LR(\beta=\beta^0 \mid \mathcal{M}_0)]/(n-1)$$
$$\stackrel{1}{=} 1 + \frac{1}{2T}(3n+1) - \frac{1}{T}\,\frac{\beta'\alpha}{\alpha'\Omega^{-1}\alpha\beta'\Omega\beta}$$
$$\times [(2n-3)(2+\beta'\alpha) + 4(1+\beta'\alpha)].$$

It is seen that the parameters enter through two functions $\alpha'\beta$ and $\alpha'\Omega^{-1}\alpha\beta'\Omega\beta$. This result holds not only for the correction term but for the distribution of the likelihood ratio statistic. To see this we transform the problem linearly by defining $\widetilde{X}_t = v' X_t$, where $v_1 = \beta(\beta'\Omega\beta)^{-1/2}$,

$$v_2 = -(\Omega^{-1} - \beta(\beta'\Omega\beta)^{-1}\beta')\alpha(\alpha'\Omega^{-1}\alpha - \alpha'\beta(\beta'\Omega\beta)^{-1}\beta'\alpha)^{-1/2},$$

and finally $v_3,\dots,v_n$ are such that

$$v_i'\Omega v_j = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases}, \qquad i,j = 1,\dots,n.$$

The new variables satisfy the equations

$$\begin{aligned} \Delta\widetilde{X}_{1t} &= \eta\widetilde{X}_{1t-1} + \tilde{\varepsilon}_{1t}, \\ \Delta\widetilde{X}_{2t} &= \gamma\widetilde{X}_{1t-1} + \tilde{\varepsilon}_{2t}, \\ \Delta\widetilde{X}_{it} &= \tilde{\varepsilon}_{it}, \qquad i = 3,\dots,n, \end{aligned} \tag{36}$$

where $\tilde{\varepsilon}_t = v'\varepsilon_t$ are i.i.d. $N_n(0, I_n)$ and

$$\eta = \beta'\alpha,\ \gamma = -(\alpha'\Omega^{-1}\alpha\beta'\Omega\beta - (\alpha'\beta)^2)^{1/2},$$

such that

$$v(\alpha) = -\frac{\eta(2+\eta)}{\gamma^2+\eta^2}, \qquad c(\alpha) = -2\,\frac{\eta(1+\eta)}{\gamma^2+\eta^2}.$$

Thus any DGP given by (α, β, Ω) is completely described by just two numbers and can be transformed into the "canonical" form (36) with $\beta' = (1,0,\ldots,0)$ and $\alpha' = (\eta, \gamma, 0, \ldots, 0)$. It is therefore possible for a given n to tabulate the effect of the Bartlett correction as a function of two variables γ and η (see Table 2, where we give some results for $T = 50$, $n = 2$, and $n = 5$). It is seen that for $n = 2$ a nominal 5% test can have an actual size up to 16% and that in many cases (roughly $\eta + \gamma < -0.2$) the Bartlett correction factor gives a useful correction.

Note that for $\gamma = 0$, both coefficients $v(\alpha)$ and $c(\alpha)$ have a factor η^{-1}, which tends to infinity for small η. The DGP where both γ and η are zero corresponds to no cointegration, and the test on β does not have a meaning in such a situation. The model with $\eta = 0$, and $\gamma \neq 0$, corresponds to a DGP generating an $I(2)$ process, and the derivation of the correction factor is not valid in this case.

For $n = 5$ the situation is worse, and the actual size can be very large indeed. The region where the Bartlett correction is useful is approximately given by $\eta + \gamma < -0.4$. Obviously the situation improves if T increases.

Usually the test for β is preceded by a test for the rank, and if η and γ are sufficiently small the hypothesis of one cointegrating relation will not be accepted; thus for small values of γ and η the Bartlett correction is not needed.

5.2. Some Real Life Examples

In the examples that follow, taken from Johansen (1996), we use as the DGP the AR(2) model fitted to the real data. We have left out the seasonal dummies and started the process at the initial values for the data.

We first consider the Danish data consisting of the four variables m_t (log real M2), y_t (log real income), i_t^b (bond rate), and i_t^d (deposit rate) observed quarterly from 1974:1 to 1987:3. We fitted a model with a restricted constant term and simulated a time series with 53 observations, which was the number of observations in the example.

We first give the result for a simple test on β, that is, that the coefficients are those of the DGP. The Bartlett factor in this case is given by Corollary 6, because we have a hypothesis that involves β and not ρ. We find with $n = 4$, $r = 1$, $s = 1$, $n_D = 1$, $n_d = 0$, $k = 2$ that the degrees of freedom is $r(n - s) = 3$ and that

$$E[-2\log LR(\beta = \beta^0 \,|\, \mathcal{M}_0)]/3 \stackrel{1}{=} 1 + \frac{23}{2T} + \frac{1}{T}[7v(\alpha) + 2c(\alpha)].$$

TABLE 2. Corrected size over the uncorrected size for a nominal 5% test (T = 50, r = 1, and n = 2 and 5)

	$n = 2$				$n = 5$			
$\gamma\backslash\eta$	−0.1	−0.2	−0.4	−0.8	−0.1	−0.2	−0.4	−0.8
0.0	1.3/16.5	3.7/12.2	4.7/8.6	5.2/6.8	0.02/78.6	1.8/64.6	6.8/40.3	6.2/17.6
−0.1	3.8/13.5	4.2/11.0	4.8/8.7	4.9/7.0	1.7/67.9	3.7/59.2	7.0/38.5	6.2/17.5
−0.2	5.4/10.8	4.9/9.9	4.9/8.6	5.1/7.0	9.5/44.9	7.4/46.8	7.5/34.1	6.1/17.2
−0.4	6.1/8.2	5.6/8.3	5.4/8.0	5.2/7.1	8.3/20.0	7.8/24.1	7.0/23.6	6.0/15.5
−0.8	5.6/6.9	5.5/7.2	5.8/7.4	5.4/6.9	6.1/11.5	6.0/13.2	5.8/13.9	5.6/12.8

We find for a test of nominal size 5% an actual size of 20.7% and a corrected size of 8.0%.

Another test of the form $\beta = H\tau$ is given by the matrix H:

$$H = \begin{pmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -1 \end{pmatrix}$$

with τ (2×1) corresponding to the test that m_t and y_t enter with the same coefficient with opposite sign and that the same holds for i_t^b and i_t^d.

We find again from Corollary 6 with $n = 4$, $r = 1$, $s = 2$, $k = 2$, $n_D = 1$, $n_d = 0$ that the Bartlett factor is

$$E[-2 \log LR(\mathcal{M}_1|\mathcal{M}_0)]/2 \overset{1}{=} 1 + \frac{12}{T} + \frac{1}{T}[8v(\alpha) + 2c(\alpha)].$$

We find that a nominal 5% test gives an actual size of 19.5%, whereas the size for the corrected test is 8.9%.

As another example consider the data consisting of consumer price indices for Australia, p_t^{au}, and the United States, p_t^{us}, and the exchange rate $exch_t$ (all in logarithms) together with the five year treasury bond rate for both countries, i_t^{au} and i_t^{us}. The data are observed quarterly from 1972:1 to 1991:1, which gives an effective number of observations of 75. We fitted a model with two lags and unrestricted constant and found two cointegrating relations. We first test a simple hypothesis on the two cointegrating relations. In this case we have $n = 5$, $r = 2$, $s = 2$, $k = 2$, $n_D = 0$, $n_d = 1$, such that the degrees of freedom is $r(n - s) = 6$. Because $d_t = 1$, we find that

$$c_d(\alpha) = n_d v(\alpha) = v(\alpha).$$

The Bartlett factor can be found from Corollary 6 and is given by

$$E[-2 \log LR(\mathcal{M}_1|\mathcal{M}_0)]/6 \overset{1}{=} 1 + \frac{14}{T} + \frac{1}{2T}[7v(\alpha) + 2c(\alpha)].$$

We found that a nominal 5% test in reality corresponds to a size of 21%. The correction of the test gives a size of 6.3%.

Next consider the test for price homogeneity given by the restriction $R'\beta = 0$, where

$$R = (1,1,0,0,0)$$

and $H = R_\perp$. In this example $s = 4$ such that the degrees of freedom is $r(n - s) = 2$. We find the Bartlett factor from Corollary 6 given by

$$E[-2 \log LR(\mathcal{M}_1|\mathcal{M}_0)]/2 \overset{1}{=} 1 + \frac{15}{T} + \frac{1}{2T}[9v(\alpha) + 2c(\alpha)].$$

A nominal 5% test corresponds to a test of size 10.5%, and the Bartlett correction gives the size 3.4%.

6. CONCLUSION

In this paper we have derived a correction to the log likelihood ratio statistic for various hypotheses on the cointegrating coefficients in a cointegrated vector autoregressive model. The correction depends on the parameters under the null hypothesis and on dimension, lag length, cointegrating rank, number of restricted deterministic terms, and number of unrestricted terms. The effect of the parameters is summarized in two functions, which are easily calculated once the parameters of the model have been estimated.

The usefulness of the results is demonstrated by some simulation experiments.

REFERENCES

Ahn, S.K. & C.G. Reinsel (1990) Estimation for partially non-stationary multivariate autoregressive models. *Journal of the American Statistical Association* 85, 813–823.

Bartlett, M.S. (1937) Properties of sufficiency and statistical tests. *Proceedings of the Royal Statistical Society of London, Series A*, 160, 268–282.

Chan, N.H. & C.Z. Wei (1988) Limiting distributions of least squares estimators of unstable autoregressive processes. *Annals of Statistics* 16, 367–401.

Cribari-Neto, F. & G. Cordeiro (1996) On Bartlett and Bartlett-type corrections. *Econometric Reviews* 15, 339–367.

Fachin, S. (1997) Bootstrap and Asymptotic Tests of Long-Run Relationships in Cointegrated Systems. Discussion paper, University of Rome.

Gredenhoff, M. & T. Jacobson (1998) Bootstrap testing and approximate finite sample distributions for test of linear restrictions on cointegrating vectors. *Journal of Business and Economics Statistics*, forthcoming.

Haug, A. (1998) Testing Linear Restrictions on Cointegrating Vectors: Size and Power of Wald Tests in Finite Samples. Discussion paper, University of Canterbury.

Jacobson, T. (1995) Simulating small sample properties of the maximum likelihood cointegration model: Estimation and testing. *Finnish Economic Papers* 8, 96–107.

Jacobson, T., A. Vredin, & A. Warne (1998) Are real wages and unemployment related? *Economica* 65, 69–96.

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

Johansen, S. (1999) A small sample correction for test of hypotheses on the cointegrating vectors. *Journal of Econometrics*, forthcoming.

Johansen, S. & K. Juselius (1990) Maximum likelihood estimation and inference on cointegration—with applications to the demand for money. *Oxford Bulletin of Economics and Statistics* 52, 169–210.

Lawley, D.N. (1956) A general method for approximating to the distribution of likelihood ratio criteria. *Biometrika* 43, 296–303.

APPENDIX

Some technical results. The deterministic terms are modelled by the matrix M, which enters in the proof of Lemma 1. We there need the following result.

LEMMA 9. *Under the assumptions* (2) *and* (3) *the powers* M^h *grow at most as a polynomial in* h.

Proof. The matrix M has a Jordan form that is block diagonal with blocks of the form

$$J_k(\lambda) = \lambda I_k + J, \qquad J_{ij} = 1, \qquad i = j+1, \qquad J_{ij} = 0, \qquad i \neq j+1,$$

where λ is an eigenvalue of M, assumed to satisfy $|\lambda| = 1$. In the powers of J the subdiagonal moves down, and $J^k = 0$. We find that the Jordan form for M^h has the blocks

$$J_k(\lambda)^h = (\lambda I_k + J)^h = \sum_{i=0}^{k-1} \binom{h}{i} \lambda^{h-i} J^i.$$

This is bounded by a polynomial of degree $k-1$ in h. This proves that M^h is bounded by a polynomial of degree at most equal to the order (minus one) of the largest Jordan block in M. Because M^{-1} has eigenvalues $\bar{\lambda}$, the same result holds for negative powers of M. ■

When regressing a stationary variable on deterministic terms d_t, the order of magnitude of the product moments depends on d_t. We formulate a result that is used in the proof of Theorem 4.

LEMMA 10. *Let* $S_t = \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i}$ *with* θ_i *decreasing exponentially. Let*

$$\gamma(h) = \mathrm{Cov}(S_t, S_{t+h}) = \sum_{i=0}^{\infty} \theta_i \Omega \theta_{i+h}'.$$

Let d_t *satisfy* $d_{t+1} = M d_t$, *with* $|\mathrm{eig}(M)| = 1$; *then*

$$\mathrm{tr}\{E(M_{sd} M_{dd}^{-1} M_{ds})\} \to \sum_{h=-\infty}^{\infty} \mathrm{tr}\{M^h\} \mathrm{tr}\{\gamma(h)\}. \tag{A.1}$$

Proof.

$$\mathrm{tr}\{E(M_{sd} M_{dd}^{-1} M_{ds})\} = \mathrm{tr}\left\{ E \sum_{s,t=1}^{T} \sum_{i,j=0}^{\infty} \theta_i \varepsilon_{t-i} d_t' M_{dd}^{-1} d_s \varepsilon_{s-j}' \theta_j' \right\}$$

$$= \sum_{i,j,t} \mathrm{tr}\{\theta_i \Omega \theta_j'\} \mathrm{tr}\{d_t' M_{dd}^{-1} d_{t-i+j}\}$$

$$= \sum_{i,j} \sum_{t=1}^{T} \mathrm{tr}\{d_t' M_{dd}^{-1} d_t M^{-i+j}\} \mathrm{tr}\{\Omega \theta_j' \theta_i\}$$

$$\to \sum_{h=-\infty}^{\infty} \mathrm{tr}\{M^h\} \mathrm{tr}\{\gamma(h)\}.$$ ■

Proof of Theorem 2 and some expansions of a projection matrix

Proof of Theorem 2.

Proof of (22). We expand (19) and find with $S_{0a} = \xi S_{va} + S_{\varepsilon a}$, $S_{0v} = \xi S_{vv} + S_{\varepsilon v}$, and $\hat{\xi} - \xi = S_{\varepsilon v} S_{vv}^{-1}$ that

$$\tilde{\xi} = \hat{\xi} + [S_{0a}\tilde{\delta} - \hat{\xi}(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})] S_{vv}^{-1} + O_P(T^{-2})$$

$$= \hat{\xi} + [S_{\varepsilon a.v}\tilde{\delta} - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta} - \hat{\xi}\tilde{\delta}' S_{av}] S_{vv}^{-1} + O_P(T^{-2}),$$

where we have used the identity

$$S_{0a} - \hat{\xi} S_{va} = S_{\varepsilon a} - (\hat{\xi} - \xi) S_{va} = S_{\varepsilon a} - S_{\varepsilon v} S_{vv}^{-1} S_{va} = S_{\varepsilon a.v}. \tag{A.2}$$

■

Proof of (23). From (19) and (20) we find

$$T\tilde{\Omega} = S_{00} - (S_{0v} + S_{0a}\tilde{\delta})(S_{vv} + \tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})^{-1}(S_{v0} + \tilde{\delta}' S_{a0}).$$

We now expand the last term and keep terms of order T^{-1}. Throughout we use $\hat{\xi} = S_{0v} S_{vv}^{-1}$. Then

$$(S_{0v} + S_{0a}\tilde{\delta})(S_{vv} + \tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})^{-1}(S_{v0} + \tilde{\delta}' S_{a0})$$

$$\stackrel{1}{=} (S_{0v} + S_{0a}\tilde{\delta})[S_{vv}^{-1} - S_{vv}^{-1}(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta}) S_{vv}^{-1}$$
$$+ S_{vv}^{-1}\{(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta}) S_{vv}^{-1}\}^2](S_{v0} + \tilde{\delta}' S_{a0})$$
$$= S_{0v} S_{vv}^{-1} S_{v0} + \hat{\xi}\tilde{\delta}' S_{a0} + S_{0a}\tilde{\delta}\hat{\xi}' - \hat{\xi}(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})\hat{\xi}'$$
$$+ S_{0a}\tilde{\delta} S_{vv}^{-1}\tilde{\delta}' S_{a0} - \hat{\xi}(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta}) S_{vv}^{-1}\tilde{\delta}' S_{0a}$$
$$- S_{a0}\tilde{\delta} S_{vv}^{-1}(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})\hat{\xi}'$$
$$+ \hat{\xi}(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta}) S_{vv}^{-1}(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})\hat{\xi}'.$$

The main term combines with S_{00} to give $T\hat{\Omega} = S_{00} - S_{0v} S_{vv}^{-1} S_{v0}$. The term of order T^0 is

$$\hat{\xi}\tilde{\delta}' S_{a0} + S_{0a}\tilde{\delta}\hat{\xi}' - \hat{\xi}(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})\hat{\xi}' = \hat{\xi}\tilde{\delta}' S_{a\varepsilon.v} + S_{\varepsilon a.v}\tilde{\delta}\hat{\xi}' - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta}\hat{\xi}',$$

where we have used (A.2). The term of order T^{-1} is

$$[S_{0a}\tilde{\delta} - \hat{\xi}(\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})] S_{vv}^{-1}[\tilde{\delta}' S_{a0} - (\tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})\hat{\xi}']$$
$$= [S_{\varepsilon a.v}\tilde{\delta} - \hat{\xi}\tilde{\delta}' S_{av} - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta}] S_{vv}^{-1}[\tilde{\delta}' S_{a\varepsilon.v} - S_{va}\tilde{\delta}\hat{\xi}' - \tilde{\delta}' S_{aa}\tilde{\delta}\hat{\xi}'],$$

where we again use (A.2). ■

Proof of (24). Because $\tilde{\xi} - \hat{\xi}$ and $\tilde{\Omega} - \hat{\Omega}$ are both in $O_P(T^{-1})$ and $S_{0a}S_{aa}^{-1}S_{a0} \in O_P(1)$, we find from (21) and (A.2) that

$$\tilde{\delta} = S_{aa}^{-1}(S_{a0} - S_{av}\hat{\xi}')\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} + O_P(T^{-1}S_{aa}^{-1/2})$$
$$= S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} + O_P(T^{-1}S_{aa}^{-1/2}). \qquad \blacksquare$$

This completes the proof of Theorem 2 on the expansion of the estimators. ■

The expansions can conveniently be expressed in terms of a projection matrix

$$\hat{P} = P(\hat{\xi},\hat{\Omega}) = \hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1}\hat{\xi}'\hat{\Omega}^{-1} \overset{0}{=} P = P(\xi,\Omega),$$

because from (24)

$$\hat{\xi}\tilde{\delta}' = P(\hat{\xi},\hat{\Omega})S_{\varepsilon a.v}S_{aa}^{-1} + O_P(T^{-1}S_{aa}^{-1/2}). \tag{A.3}$$

Further we have

$$S_{\varepsilon a.v} = S_{\varepsilon a} - S_{\varepsilon v}S_{vv}^{-1}S_{va} = S_{\varepsilon a} + O_P(T^{1/2}T^{-1}S_{aa}^{1/2}) = S_{\varepsilon a} + O_P(T^{-1/2}S_{aa}^{1/2}),$$

and we find the intermediate results, where $\kappa_\xi = (\xi'\Omega^{-1}\xi)^{-1/2}$

$$\begin{aligned}
\hat{\xi}\tilde{\delta}'S_{av} &\overset{0}{=} \hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1}\hat{\xi}'\hat{\Omega}^{-1}S_{\varepsilon a.v}S_{aa}^{-1}S_{av} \overset{0}{=} PS_{\varepsilon a}S_{aa}^{-1}S_{av}\\
\hat{\xi}\tilde{\delta}'S_{aa}\tilde{\delta} &\overset{0}{=} PS_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} \overset{0}{=} PS_{\varepsilon a}S_{aa}^{-1}S_{au}\kappa_\xi\\
S_{\varepsilon a.v}\tilde{\delta} &\overset{0}{=} S_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} \overset{0}{=} S_{\varepsilon a}S_{aa}^{-1}S_{au}\kappa_\xi,\\
\hat{\xi}\tilde{\delta}'S_{a\varepsilon.v} &\overset{0}{=} PS_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}\\
\hat{\xi}\tilde{\delta}'S_{aa}\tilde{\delta}\hat{\xi}' &\overset{0}{=} PS_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}P'.
\end{aligned} \tag{A.4}$$

We next give an expansion of the projection matrix $\hat{P}$, which will be used in the detailed calculations that follow. We define

$$\bar{P}(\hat{\xi},\hat{\Omega}) = \hat{\xi}_\perp(\hat{\xi}_\perp'\hat{\Omega}\hat{\xi}_\perp)^{-1}\hat{\xi}_\perp' = \hat{\Omega}^{-1} - \hat{\Omega}^{-1}\hat{P}(\hat{\xi},\hat{\Omega}).$$

Note that we only expand as a function of $\hat{\xi}$ but keep $\hat{\Omega}$.

LEMMA 11. *The expansion of $P(\hat{\xi},\hat{\Omega})$ around the point $(\xi,\hat{\Omega})$ is given by*

$$\begin{aligned}
P(\hat{\xi},\hat{\Omega}) \overset{1}{=} {} & P(\xi,\hat{\Omega}) + \bar{P}(\xi,\hat{\Omega})(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}\\
&+ \hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\bar{P}(\xi,\hat{\Omega})\\
&+ \bar{P}(\xi,\hat{\Omega})(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\bar{P}(\xi,\hat{\Omega})\\
&- \hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\bar{P}(\xi,\hat{\Omega})(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}\\
&- \hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\bar{P}(\xi,\hat{\Omega}).\\
&- \bar{P}(\xi,\hat{\Omega})(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}.
\end{aligned}$$

Proof. Let $u = \hat{\Omega}^{-1/2}\xi$, such that $\xi = \hat{\Omega}^{1/2}u$, and define $v = \hat{\Omega}^{-1/2}(\hat{\xi} - \xi)$, such that $(\hat{\xi} - \xi) = \hat{\Omega}^{1/2}v$. Then

$$P(\xi,\hat{\Omega}) = \xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1} = \hat{\Omega}^{1/2}u(u'u)^{-1}u'\hat{\Omega}^{-1/2} = \hat{\Omega}^{1/2}P_0\hat{\Omega}^{-1/2},$$

say, and

$$\bar{P}(\xi,\hat{\Omega}) = \hat{\Omega}^{-1} - \hat{\Omega}^{-1}P(\xi,\hat{\Omega}) = \hat{\Omega}^{-1} - \hat{\Omega}^{-1/2}u(u'u)^{-1}u'\hat{\Omega}^{-1/2} = \hat{\Omega}^{-1/2}\bar{P}_0\hat{\Omega}^{-1/2}.$$

Then we find using $\bar{u} = u(u'u)^{-1}$, such that $\bar{u}u' = P_0$

$$\begin{aligned}
\hat{\Omega}^{-1/2}P(\hat{\xi},\hat{\Omega})\hat{\Omega}^{1/2} &= (u+v)[(u+v)'(u+v)]^{-1}(u+v)' \\
&= (u+v)[u'u + u'v + v'u + v'v]^{-1}(u+v)' \\
&= (u+v)[(u'u)^{-1} - (u'u)^{-1}(u'v + v'u + v'v)(u'u)^{-1} \\
&\qquad\qquad + (u'u)^{-1}(u'v + v'u)(u'u)^{-1}(u'v + v'u)(u'u)^{-1}] \\
&\quad \times (u+v)' + O(|v|^3) \\
&= P_0 + L_1 + L_2 + O(|v|^3).
\end{aligned}$$

The first order term is given by

$$\begin{aligned}
L_1 &= \bar{u}v' + v\bar{u}' - \bar{u}(u'v + v'u)\bar{u}' = \bar{u}v'(I_n - \bar{u}u') + (I_n - \bar{u}u')v\bar{u}' \\
&= \bar{u}v'\bar{P}_0 + \bar{P}_0 v\bar{u}'.
\end{aligned}$$

The quadratic term is

$$\begin{aligned}
L_2 &= -\bar{u}(u'v + v'u)(u'u)^{-1}v' - v(u'u)^{-1}(u'v + v'u)\bar{u}' + v(u'u)^{-1}v' \\
&\quad - \bar{u}v'v\bar{u}' + \bar{u}(u'v + v'u)(u'u)^{-1}(u'v + v'u)\bar{u}' \\
&= \bar{P}_0 v(u'u)^{-1}v'\bar{P}_0 - \bar{u}v'\bar{P}_0 v\bar{u}' - \bar{P}_0 v\bar{u}'v\bar{u}' - \bar{u}v'\bar{u}v'\bar{P}_0.
\end{aligned}$$

When substituting u and v we find the result. ■

The last four terms of the expansion of $P(\hat{\xi},\hat{\Omega})$ are of the order T^{-1}, and we can simplify them by replacing $\hat{\Omega}$ by Ω. We use $U_t = (\xi'\Omega^{-1}\xi)^{-1/2}\xi'\Omega^{-1}\varepsilon_t$, $B_t = (\xi_\perp'\Omega\xi_\perp)^{-1/2}\xi_\perp'\varepsilon_t$, $\hat{\xi} - \xi = S_{\varepsilon v}S_{vv}^{-1}$, $\kappa_\xi = (\xi'\Omega^{-1}\xi)^{-1/2}$, and $S_{v\varepsilon}\bar{P}(\xi,\Omega)S_{\varepsilon v} = S_{vb}S_{bv}$, and we find the following corollary.

COROLLARY 12. *The expansion of* $P(\hat{\xi},\hat{\Omega})$ *around the point* $(\xi,\hat{\Omega})$, *using* $\hat{\xi} - \xi = S_{\varepsilon v}S_{vv}^{-1}$, $\kappa_\xi = (\xi'\Omega^{-1}\xi)^{-1/2}$, *and* $S_{v\varepsilon}\bar{P}(\xi,\Omega)S_{\varepsilon v} = S_{vb}S_{bv}$, *is given by*

$$\begin{aligned}
P(\hat{\xi},\hat{\Omega}) \overset{1}{=} {} & P(\xi,\hat{\Omega}) + \bar{P}(\xi,\hat{\Omega})S_{\varepsilon v}S_{vv}^{-1}(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1} \\
& + \hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}S_{vv}^{-1}S_{v\varepsilon}\bar{P}(\xi,\hat{\Omega}) \\
& + \xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1/2}S_{bv}S_{vv}^{-1}\kappa_\xi^2 S_{vv}^{-1}S_{vb}(\xi_\perp'\Omega\xi_\perp)^{-1/2}\xi_\perp' \\
& - \Omega^{-1}\xi(\xi'\Omega^{-1}\xi)^{-1}S_{vv}^{-1}S_{vb}S_{bv}S_{vv}^{-1}(\xi'\Omega^{-1}\xi)^{-1}\xi'\Omega^{-1} \\
& - \Omega^{-1}\xi(\xi'\Omega^{-1}\xi)^{-1}S_{vv}^{-1}S_{vu}\kappa_\xi S_{vv}^{-1}S_{vb}(\xi_\perp'\Omega\xi_\perp)^{-1/2}\xi_\perp' \\
& - \xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1/2}S_{bv}S_{vv}^{-1}\kappa_\xi S_{uv}S_{vv}^{-1}(\xi'\Omega^{-1}\xi)^{-1}\xi'\Omega^{-1}.
\end{aligned}$$

Proof of Theorem 3

We start with (25):

$$-2\log LR \overset{1}{=} \mathrm{tr}\{Q_0\} + T^{-1}(\mathrm{tr}\{Q_1\} + \frac{1}{2}\,\mathrm{tr}\{Q_0^2\})$$

and evaluate each term starting with the ones that have a factor T^{-1}.

Calculation of tr$\{Q_1\}$. We find from Theorem 2 that

$$\mathrm{tr}\{Q_1\} = T\,\mathrm{tr}\{\hat\Omega^{-1}[S_{\varepsilon a.v}\tilde\delta - \hat\xi\tilde\delta' S_{av} - \hat\xi\tilde\delta' S_{aa}\tilde\delta]S_{vv}^{-1}[\tilde\delta' S_{a\varepsilon.v} - S_{va}\tilde\delta\hat\xi' - \tilde\delta' S_{aa}\tilde\delta\hat\xi']\}.$$

Because tr$\{Q_1\}$ is multiplied by T^{-1} we need only retain the main term in each of the matrices. The key quantities are calculated in (A.4), and we find

$$S_{\varepsilon a.v}\tilde\delta - \hat\xi\tilde\delta' S_{av} - \hat\xi\tilde\delta' S_{aa}\tilde\delta \overset{0}{=} (I_n - P)S_{\varepsilon a}S_{aa}^{-1}S_{au}\kappa_\xi - PS_{\varepsilon a}S_{aa}^{-1}S_{av},$$

such that with $\hat\Omega$ replaced by Ω

$$\begin{aligned}\mathrm{tr}\{Q_1\} &\overset{0}{=} T\,\mathrm{tr}\{\kappa_\xi S_{ua}S_{aa}^{-1}S_{a\varepsilon}\Omega^{-1}(I_n - P)S_{\varepsilon a}S_{aa}^{-1}S_{au}\kappa_\xi S_{vv}^{-1}\\ &\qquad + S_{va}S_{aa}^{-1}S_{a\varepsilon}\Omega^{-1}PS_{\varepsilon a}S_{aa}^{-1}S_{av}S_{vv}^{-1}\}\\ &= T\,\mathrm{tr}\{\kappa_\xi S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}S_{au}\kappa_\xi S_{vv}^{-1}\}\\ &\quad + T\,\mathrm{tr}\{S_{va}S_{aa}^{-1}S_{au}S_{ua}S_{aa}^{-1}S_{av}S_{vv}^{-1}\},\end{aligned} \tag{A.5}$$

where the expression has been reduced using the properties of the projection matrix

$$P'\Omega^{-1}(I_n - P) = 0, \qquad P'\Omega^{-1}P = P'\Omega^{-1}$$

$$(I_n - P)'\Omega^{-1}(I_n - P) = \Omega^{-1}(I_n - P) = \xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1}\xi_\perp'$$

and the identities (see (18) for the definition of U_t and B_t)

$$\begin{aligned}S_{a\varepsilon}\Omega^{-1}PS_{\varepsilon a} &= S_{a\varepsilon}\Omega^{-1}\xi(\xi'\Omega^{-1}\xi)^{-1}\xi'\Omega^{-1}S_{\varepsilon a} = S_{au}S_{ua},\\ S_{a\varepsilon}\Omega^{-1}(I_n - P)S_{\varepsilon a} &= S_{a\varepsilon}\xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1}\xi_\perp' S_{\varepsilon a} = S_{ab}S_{ba}.\end{aligned} \tag{A.6}$$

Calculation of tr$\{Q_0^2\}$. We find from Theorem 2 that from (A.4)

$$\begin{aligned}Q_0 &= \hat\Omega^{-1}[S_{\varepsilon a.v}\tilde\delta\hat\xi' + \hat\xi\tilde\delta' S_{a\varepsilon.v} - \hat\xi\tilde\delta' S_{aa}\tilde\delta\hat\xi']\\ &\overset{0}{=} \Omega^{-1}[(I_n - P)S_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}P' + PS_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}]\end{aligned}$$

and hence

$$\begin{aligned}Q_0^2 &\overset{0}{=} \Omega^{-1}(I_n - P)S_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}\Omega^{-1}PS_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}\\ &\quad + \Omega^{-1}PS_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}\Omega^{-1}(I_n - P)S_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}P'\\ &\quad + \Omega^{-1}PS_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}\Omega^{-1}PS_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon},\end{aligned}$$

using $P'\Omega^{-1}(I_n - P) = 0$. Hence again using (A.6),

$$\tfrac{1}{2}\mathrm{tr}\{Q_0^2\} \overset{0}{=} \tfrac{1}{2}\mathrm{tr}\{(S_{ua}S_{aa}^{-1}S_{au})^2\} + \mathrm{tr}\{S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}S_{au}\}. \tag{A.7}$$

The main term tr$\{\boldsymbol{Q_0}\}$. This term is of the order of one, and hence we have to keep more terms in the expansions. From (A.3) we find that $\hat{\xi}\tilde{\delta}'$ is of the order of $S_{aa}^{-1/2}$ and that

$$\hat{\xi}\tilde{\delta}' = \hat{P}S_{\varepsilon a.v}S_{aa}^{-1} + T^{-1}\hat{\xi}\tilde{\delta}_1' + o_P(T^{-1}S_{aa}^{-1/2}),$$

for some $\tilde{\delta}_1 \in O_P(S_{aa}^{-1/2})$, and hence

$$\hat{\xi}\tilde{\delta}'S_{a\varepsilon.v} \overset{1}{=} \hat{P}S_{\varepsilon a.v}S_{aa}^{-1}S_{a\varepsilon.v} + T^{-1}\hat{\xi}\tilde{\delta}_1'S_{a\varepsilon.v}.$$

We first want to show that the coefficient $\tilde{\delta}_1$ does not play a role because of the way the projection matrices enter, such that we can replace $\hat{\xi}\tilde{\delta}'$ by $\hat{P}S_{\varepsilon a.v}S_{aa}^{-1}$, introducing errors of at most the order $o_P(T^{-1}S_{aa}^{-1/2})$ in the expression for Q_0. We find from

$$\hat{\Omega}Q_0(\tilde{\delta}\hat{\xi}') = S_{\varepsilon a.v}\tilde{\delta}\hat{\xi}' + \hat{\xi}\tilde{\delta}'S_{a\varepsilon.v} - \hat{\xi}\tilde{\delta}'S_{aa}\tilde{\delta}\hat{\xi}'$$

that

$$\begin{aligned}
&\hat{\Omega}[Q_0(\hat{\xi}\tilde{\delta}') - Q_0(\hat{P}S_{\varepsilon a.v}S_{aa}^{-1})] \\
&\quad \overset{1}{=} T^{-1}[S_{\varepsilon a.v}\tilde{\delta}_1\hat{\xi}' + \hat{\xi}\tilde{\delta}_1'S_{a\varepsilon.v} - \hat{\xi}\tilde{\delta}_1'S_{aa}S_{aa}^{-1}S_{a\varepsilon.v}\hat{P}' - \hat{P}S_{\varepsilon a.v}S_{aa}^{-1}S_{aa}\tilde{\delta}_1\hat{\xi}'] \\
&\quad = T^{-1}[(I_n - \hat{P})S_{\varepsilon a.v}\tilde{\delta}_1\hat{\xi}' + \hat{\xi}\tilde{\delta}_1'S_{a\varepsilon.v}(I_n - \hat{P})'].
\end{aligned}$$

This term, however, does not give a contribution because

$$\mathrm{tr}\{\hat{\Omega}^{-1}(I_n - \hat{P})S_{\varepsilon a.v}\tilde{\delta}_1\hat{\xi}'\} = \mathrm{tr}\{\hat{\xi}'\hat{\Omega}^{-1}(I_n - \hat{P})S_{\varepsilon a.v}\tilde{\delta}_1\}$$

and $\hat{\xi}'\hat{\Omega}^{-1}(I_n - \hat{P}) = 0$. In the following we therefore replace $\tilde{\delta}\hat{\xi}'$ by $S_{aa}^{-1}S_{a\varepsilon.v}\hat{P}'$, and we find

$$\begin{aligned}
\mathrm{tr}\{Q_0\} &\overset{1}{=} \mathrm{tr}\{\hat{\Omega}^{-1}S_{\varepsilon a.v}S_{aa}^{-1}S_{a\varepsilon.v}\hat{P}'\} + \mathrm{tr}\{\hat{\Omega}^{-1}\hat{P}S_{\varepsilon a.v}S_{aa}^{-1}S_{a\varepsilon.v}(I - \hat{P})'\} \\
&= \mathrm{tr}\{\hat{\Omega}^{-1}S_{\varepsilon a.v}S_{aa}^{-1}S_{a\varepsilon.v}\hat{P}'\},
\end{aligned}$$

where we have used that $\hat{P}'\hat{\Omega}^{-1}(I_n - \hat{P}) = 0$.

We next expand around ξ but keep $\hat{\Omega}$. We find from Corollary 12, using the notation $\bar{P}(\xi, \hat{\Omega}) = \xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp'$, that

$$\begin{aligned}
\mathrm{tr}\{Q_0\} \overset{1}{=} {} & \mathrm{tr}\{S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}P(\xi, \hat{\Omega})S_{\varepsilon a.v}\} \\
& + 2\,\mathrm{tr}\{S_{aa}^{-1}S_{a\varepsilon.v}\bar{P}(\xi, \hat{\Omega})S_{\varepsilon v}S_{vv}^{-1}(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}S_{\varepsilon a.v}\} \\
& + \mathrm{tr}\{S_{aa}^{-1}S_{a\varepsilon.v}\xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1/2}S_{bv}S_{vv}^{-1}\kappa_\xi^2S_{vv}^{-1}S_{vb}(\xi_\perp'\Omega\xi_\perp)^{-1/2}\xi_\perp'S_{\varepsilon a.v}\} \\
& - \mathrm{tr}\{S_{aa}^{-1}S_{a\varepsilon.v}\Omega^{-1}\xi(\xi'\Omega^{-1}\xi)^{-1}S_{vv}^{-1}S_{vb}S_{bv}S_{vv}^{-1}(\xi'\Omega^{-1}\xi)^{-1}\xi'\Omega^{-1}S_{\varepsilon a.v}\} \\
& - 2\,\mathrm{tr}\{S_{aa}^{-1}S_{a\varepsilon.v}\xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1/2}S_{bv}S_{vv}^{-1}\kappa_\xi S_{uv}S_{vv}^{-1}(\xi'\Omega^{-1}\xi)^{-1}\xi'\Omega^{-1}S_{\varepsilon a.v}\} \\
= {} & A_1 + A_2 + A_3 + A_4 + A_5.
\end{aligned}$$

The last three terms are of the order T^{-1} and simplify using the relations

$$S_{a\varepsilon.v}\Omega^{-1}\xi(\xi'\Omega^{-1}\xi)^{-1} = S_{au.v}\kappa_\xi = S_{au} + O_P(S_{aa}^{1/2}T^{-1/2}),$$

$$S_{a\varepsilon.v}\xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1/2} = S_{ab.v} = S_{ab} + O_P(S_{aa}^{1/2}T^{-1/2}).$$

We find

$$\begin{aligned}
A_3 &\overset{1}{=} \operatorname{tr}\{S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_\xi^2 S_{vv}^{-1}S_{vb}S_{ba}\},\\
A_4 &\overset{1}{=} -\operatorname{tr}\{S_{aa}^{-1}S_{au}\kappa_\xi S_{vv}^{-1}S_{vb}S_{bv}S_{vv}^{-1}\kappa_\xi S_{ua}\},\\
A_5 &\overset{1}{=} -2\operatorname{tr}\{S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_\xi S_{uv}S_{vv}^{-1}\kappa_\xi S_{ua}\}.
\end{aligned} \tag{A.8}$$

We still need to simplify the first two terms, A_1 and A_2. We introduce the definitions

$$L' = \begin{pmatrix} (\xi'\Omega^{-1}\xi)^{-1/2}\xi'\Omega^{-1} \\ (\xi_\perp'\Omega\xi_\perp)^{-1/2}\xi_\perp' \end{pmatrix}, \qquad L'\varepsilon_t = \begin{pmatrix} U_t \\ B_t \end{pmatrix},$$

and from $\hat\Omega = T^{-1}S_{\varepsilon\varepsilon.v}$, we find

$$L'\hat\Omega L = T^{-1}\begin{pmatrix} S_{uu.v} & S_{ub.v} \\ S_{ub.v} & S_{bb.v} \end{pmatrix}, \qquad S_{a\varepsilon.v}L = \begin{pmatrix} S_{au.v} \\ S_{ab.v} \end{pmatrix}$$

such that

$$\begin{aligned}
S_{a\varepsilon.v}\hat\Omega^{-1}\xi &= S_{a\varepsilon.v}L(L'\hat\Omega L)^{-1}L'\xi\\
&= T(S_{au.v}, S_{ab.v})\begin{pmatrix} S_{uu.v} & S_{ub.v} \\ S_{ub.v} & S_{bb.v} \end{pmatrix}^{-1}\begin{pmatrix} (\xi'\Omega^{-1}\xi)^{1/2} \\ 0 \end{pmatrix}\\
&= TS_{au.v,b}S_{uu.v,b}^{-1}(\xi'\Omega^{-1}\xi)^{1/2},
\end{aligned}$$

and

$$\begin{aligned}
&(\xi'\Omega^{-1}\xi)^{1/2}(\xi'\hat\Omega^{-1}\xi)^{-1}(\xi'\Omega^{-1}\xi)^{1/2}\\
&\quad = [(\xi'\Omega^{-1}\xi)^{-1/2}\xi'L(L'\hat\Omega L)^{-1}L'\xi(\xi'\Omega^{-1}\xi)^{-1/2}]^{-1}\\
&\quad = T^{-1}\left(\begin{pmatrix} I_{n_u} \\ 0 \end{pmatrix}'\begin{pmatrix} S_{uu.v} & S_{ub.v} \\ S_{ub.v} & S_{bb.v} \end{pmatrix}^{-1}\begin{pmatrix} I_{n_u} \\ 0 \end{pmatrix}\right)^{-1} = T^{-1}S_{uu.v,b},
\end{aligned}$$

which combine to give

$$S_{a\varepsilon.v}\hat\Omega^{-1}\xi(\xi'\hat\Omega^{-1}\xi)^{-1} = S_{au.v,b}\kappa_\xi \tag{A.9}$$

and hence

$$
\begin{aligned}
A_1 &= \operatorname{tr}\{S_{aa}^{-1} S_{a\varepsilon.v} \hat{\Omega}^{-1} P(\xi, \hat{\Omega}) S_{\varepsilon a.v}\} \\
&= \operatorname{tr}\{S_{aa}^{-1} [S_{a\varepsilon.v} \hat{\Omega}^{-1} \xi][(\xi' \hat{\Omega}^{-1} \xi)^{-1}][\xi' \hat{\Omega}^{-1} S_{\varepsilon a.v}]\} \\
&= T \operatorname{tr}\{S_{aa}^{-1} S_{au.v,b} S_{uu.v,b}^{-1} S_{ua.v,b}\}.
\end{aligned}
$$

Now we want to replace S_{aa}^{-1} by $S_{aa.b,v}^{-1}$ and use the identity

$$
S_{aa.b,v} = S_{aa.b} - S_{av.b} S_{vv.b}^{-1} S_{va.b} = S_{aa} - S_{ab} S_{bb}^{-1} S_{ba} - S_{av.b} S_{vv.b}^{-1} S_{va.b}
$$

to find

$$
\begin{aligned}
S_{aa}^{-1} - S_{aa.b,v}^{-1} &= S_{aa}^{-1} - (S_{aa} - S_{ab} S_{bb}^{-1} S_{ba} - S_{av.b} S_{vv.b}^{-1} S_{va.b})^{-1} \\
&= -S_{aa}^{-1}(S_{ab} S_{bb}^{-1} S_{ba} + S_{av.b} S_{vv.b}^{-1} S_{va.b}) S_{aa}^{-1} + O_P(S_{aa}^{-1} T^{-2}),
\end{aligned}
$$

such that for A_1 we find the main term $T \operatorname{tr}\{S_{aa.b,v}^{-1} S_{au.b,v} S_{uu.b,v}^{-1} S_{ua.b,v}\}$.

The last two terms are of the order of T^{-1} and can be simplified using

$$
T^{-1} S_{bb} \sim I_{n_b}, \qquad S_{av.b} \sim S_{av}, \qquad S_{au.b,v} \sim S_{au}, \qquad T^{-1} S_{uu.b,v} \sim I_{n_u},
$$
$$
T^{-1} S_{vv.b} \sim T^{-1} S_{vv},
$$

and we find

$$
\begin{aligned}
A_1 &\stackrel{1}{=} T \operatorname{tr}\{S_{aa.b,v}^{-1} S_{au.b,v} S_{uu.b,v}^{-1} S_{ua.b,v}\} \\
&\quad - T^{-1} \operatorname{tr}\{S_{aa}^{-1} S_{ab} S_{ba} S_{aa}^{-1} S_{au} S_{ua}\} - \operatorname{tr}\{S_{aa}^{-1} S_{av} S_{vv}^{-1} S_{va} S_{aa}^{-1} S_{au} S_{ua}\}. \qquad \textbf{(A.10)}
\end{aligned}
$$

To find A_2 we note that $S_{a\varepsilon.v} \xi_\perp = S_{ab.v}(\xi_\perp' \Omega \xi_\perp)^{1/2}$ and that

$$
\begin{aligned}
(\xi_\perp' \Omega \xi_\perp)^{1/2} (\xi_\perp' \hat{\Omega} \xi_\perp)^{-1} (\xi_\perp' \Omega \xi_\perp)^{1/2} &= [(\xi_\perp' \Omega \xi_\perp)^{-1/2} \xi_\perp' S_{\varepsilon\varepsilon.v} \xi_\perp (\xi_\perp' \Omega \xi_\perp)^{-1/2}]^{-1} \\
&= T S_{bb.v}^{-1},
\end{aligned}
$$

which combine to give

$$
S_{a\varepsilon.v} \xi_\perp (\xi_\perp' \hat{\Omega} \xi_\perp)^{-1} \xi_\perp' S_{\varepsilon v} = T S_{ab.v} S_{bb.v}^{-1} S_{bv}. \qquad \textbf{(A.11)}
$$

We then find

$$
\begin{aligned}
A_2 &\stackrel{1}{=} 2 \operatorname{tr}\{S_{aa}^{-1} S_{a\varepsilon.v} \xi_\perp (\xi_\perp' \hat{\Omega} \xi_\perp)^{-1} \xi_\perp' S_{\varepsilon v} S_{vv}^{-1} (\xi' \hat{\Omega}^{-1} \xi)^{-1} \xi' \hat{\Omega}^{-1} S_{\varepsilon a.v}\} \\
&= 2T \operatorname{tr}\{S_{aa}^{-1} S_{ab.v} S_{bb.v}^{-1} S_{bv} S_{vv}^{-1} \kappa_\xi S_{ua.v,b}\},
\end{aligned}
$$

using (A.9) and (A.11). Again we can use the relation $T^{-1} S_{bb.v} = T^{-1} S_{bb} + O_P(T^{-1})$ to simplify so that we get

$$
A_2 \stackrel{1}{=} 2T \operatorname{tr}\{S_{aa}^{-1} S_{ab.v} S_{bb}^{-1} S_{bv} S_{vv}^{-1} \kappa_\xi S_{ua.v,b}\}. \qquad \textbf{(A.12)}
$$

Inserting the expressions (A.5), (A.7), (A.8), (A.10), and (A.11) into the expression for the likelihood ratio test (25) we find the expression in Theorem 3. Note that two terms from A_1, as given in (A.10), cancel a term in the expression for Q_1 (A.5) and one in the expression for Q_0^2 (A.7).

This completes the proof of the representation of the likelihood ratio test statistic given in Theorem 3. ∎

Proof of Theorem 4

Throughout we need the following result from the multivariate normal distribution.

LEMMA 13. *Let X and Y be multivariate normally distributed with mean zero and covariance matrix* $\Sigma \otimes \Omega$, *that is*,

$$E[\lambda_1' X \lambda_2 \lambda_3' Y \lambda_4] = \lambda_1' \Sigma \lambda_3 \lambda_2' \Omega \lambda_4.$$

Then, for A and B of suitable dimensions, we have

$$E[XAY'] = \Sigma \operatorname{tr}\{\Omega A'\}, \qquad E[XBY] = \Sigma B' \Omega.$$

Proof. We let $A = \sum_{i=1}^n \mu_i \nu_i'$, such that

$$E[\mu' XAY' \nu] = \sum_{i=1}^n E[\mu' X \mu_i \nu' Y \nu_i] = \mu' \Sigma \nu \sum_{i=1}^n \mu_i' \Omega \nu_i = \mu' \Sigma \nu \operatorname{tr}\{\Omega A'\}.$$

Next let $B = \sum_{i=1}^n \eta_i \xi_i'$; then

$$E[\mu' XBY \nu] = \sum_{i=1}^n E[\mu' X \eta_i \xi_i' Y \nu] = \sum_{i=1}^n \mu' \Sigma \xi_i \eta_i' \Omega \nu = \mu' \Sigma B' \Omega \nu.$$ ∎

We also need the calculation given in the following lemma.

LEMMA 14. *Let* $S_t = \sum_{i=0}^\infty \sigma_i U_{t-i}$ *and* $H_t = \sum_{i=0}^\infty \eta_i U_{t-i}$, *where* σ_i *and* η_i *are coefficient matrices that decrease exponentially fast. Let* f_t *and* g_t *be sequences of vectors*; *then*

$$E[\lambda_1' M_{fs} \lambda_2 \lambda_3' M_{gh} \lambda_4] = \lambda_1' E\left[\sum_{t,s,i,j} f_t U_{t-i}' \sigma_i' \lambda_2 \lambda_3' g_s U_{s-j}' \eta_j' \lambda_4\right]$$

$$= \sum_{i,j,t} \lambda_1' f_t g_{t-i+j} \lambda_3 \lambda_2' \sigma_i \eta_j' \lambda_4. \tag{A.13}$$

For suitable choices of coefficients we can obtain results about product moments such as M_{au}, M_{by}, etc. Subsequently we indicate by $E_{\xi_\perp}[\dots]$ the conditional expectation and leave out the conditioning variables $\xi_\perp' \varepsilon_t$. Notice that when we condition on $\xi_\perp' \varepsilon_t$ the processes A_{t-1} and B_t are fixed and denoted by a_{t-1} and b_t. This also holds in the case of (14), where we condition on $\alpha_{2\perp}' \varepsilon_t$ rather than $\alpha_\perp' \varepsilon_t$.

To systematize the calculations in the proof of Theorem 4 we define some linear functions

$$N_{ij} = M_{ii}^{-1/2}(M_{ij} - E_{\xi_\perp}(M_{ij})), \qquad i = a, b, d, \qquad j = u, y, \qquad \mu_{bT} = E_{\xi_\perp}(M_{by})$$

and some quadratic functions

$$Q_{yy} = T^{-1/2}(M_{yy} - T\Sigma), \qquad Q_{yu} = T^{-1/2} M_{yu}.$$

LEMMA 15. *Under Assumption* 1, *and conditional on* $\xi_\perp' \varepsilon$, *it holds that all the linear and quadratic functions*, $N_{..}$ *and* $Q_{..}$, *are asymptotically normally distributed with mean zero. The linear functions are asymptotically independent* (*a.i.*) *of the quadratic functions. Moreover we find that* N_{by} *is a.i. of* $(N_{au}, N_{bu}, N_{ay}, N_{du})$ *and* N_{au} *is a.i. of* N_{yd}. *Subsequently we give some results that are needed in the proof of Theorem* 4.

$$\mathrm{Var}_{\xi_\perp}(N_{au}) \overset{0}{=} I_{n_a} \otimes I_{n_u}, \tag{A.14}$$

$$\mathrm{Cov}_{\xi_\perp}(N_{au}, N_{ay}) \overset{0}{=} I_{n_a} \otimes \theta', \tag{A.15}$$

$$E_{\xi_\perp}[N_{ya} N_{au}] \overset{0}{=} n_a \theta, \tag{A.16}$$

$$\mathrm{Var}_{\xi_\perp}(N_{by}) \overset{0}{=} I_{n_b} \otimes \Sigma_{\theta\theta}, \tag{A.17}$$

$$T^{-1} \mu_{bT}' A \mu_{bT} \overset{0}{=} \mathrm{tr}\{A\} \Sigma_{\psi\psi}, \tag{A.18}$$

$$E_{\xi_\perp}[M_{yy}] \overset{0}{=} \Sigma, \tag{A.19}$$

$$T^{-1} E_{\xi_\perp}[M_{yb} M_{by}] \overset{0}{=} n_b \Sigma, \tag{A.20}$$

$$T^{-1} E_{\xi_\perp}[M_{by} \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} M_{yb}] \overset{0}{=} I_{n_b} \mathrm{tr}\{\tilde{\kappa}_\xi' \Sigma^{-1} \tilde{\kappa}_\xi\} \tag{A.21}$$

Proof. It follows from the central limit theorem that all variables are asymptotically normal. We shall prove the remaining statements by calculating variances and covariances, using Lemmas 13 and 14.

We find immediately the result (A.14), because M_{au} is Gaussian with mean zero and variance $M_{aa} \otimes I_{n_u}$. To prove (A.15) we find from Lemma 14 with $f_t = g_t = a_{t-1}, S_t = U_t, H_t = Y_{t-1}$

$$\begin{aligned} \mathrm{Cov}(\lambda_1' N_{au} \lambda_2, \lambda_3' N_{ay} \lambda_4) &= \sum_i \lambda_1' M_{aa}^{-1/2} \sum_t a_{t-1} a_{t+i}' M_{aa}^{-1/2} \lambda_3 \lambda_2' \theta_i' \lambda_4 \\ &\overset{0}{=} \lambda_1' \lambda_3 \lambda_2' \theta' \lambda_4, \end{aligned} \tag{A.22}$$

where we have used (26). Relation (A.16) follows from (A.15) and Lemma 13.

To prove (A.17) we use Lemma 14 with $f_t = g_t = b_t$ and $S_t = H_t = Y_{t-1}$:

$$\begin{aligned} \mathrm{Var}_{\xi_\perp}[\lambda_1' N_{by} \lambda_2] &= \sum_{i,j} \lambda_1' M_{bb}^{-1/2} \sum_t b_t b_{t-i+j}' M_{bb}^{-1/2} \lambda_1 \mathrm{tr}\{\lambda_2' \theta_j \theta_i' \lambda_2\} \\ &\overset{0}{=} \lambda_1' I_{n_b} \lambda_1 \lambda_2' \Sigma_{\theta\theta} \lambda_2. \end{aligned}$$

Here we used (27) and $\Sigma_{\theta\theta} = \sum_i \theta_i \theta_i'$ (see (31)). The statement (A.18) follows from (28) because

$$T^{-1} \mu_{bT}' A \mu_{bT} = T^{-1} \sum_{t,s,i,j} \psi_j b_{s-j-1} b_s' A b_t b_{t-i-1}' \psi_i' \overset{0}{=} \sum_i \mathrm{tr}\{A\} \psi_i \psi_i' = \mathrm{tr}\{A\} \Sigma_{\psi\psi}.$$

To prove (A.19) we use Lemma 13, (27), and (31) and find

$$T^{-1}E_{\xi_\perp}[M_{yy}] = T^{-1}\sum_{t,i,j}\theta_i E_{\xi_\perp}[U_{t-i-1}U'_{t-j-1}]\theta'_j + \sum_{t,i,j}\psi_i b_{t-i-1}b'_{t-j-1}\psi'_j$$

$$\overset{0}{=} \sum_i \theta_i\theta'_i + \sum_i \psi_i\psi'_i = \Sigma.$$

Relation (A.20) follows from Lemma 13, (A.17) and (A.18) and

$$T^{-1}E_{\xi_\perp}[M_{yb}M_{by}] = T^{-1}E_{\xi_\perp}[N_{yb}M_{bb}N_{by}] + T^{-1}\mu'_{bT}\mu_{bT}$$

$$\overset{0}{=} T^{-1}\Sigma_{\theta\theta}\,\mathrm{tr}\{M_{bb}\} + \mathrm{tr}\{I_{n_b}\}\Sigma_{\psi\psi} \overset{0}{=} n_b\Sigma.$$

The result (A.21) follows from (A.17), (A.18), and Lemma 13.

To complete the proof we need to show that some covariances are zero. The basic calculation is given in Lemma 14. We find that

$$\mathrm{Cov}_{\xi_\perp}[\lambda'_1 M_{dd}^{-1/2}M_{du}\lambda_2, \lambda'_3 M_{bb}^{-1/2}M_{by}\lambda_4]$$

$$= \sum_{t,i}\lambda'_1 M_{dd}^{-1/2}d_t b'_{t+i+1}M_{bb}^{-1/2}\lambda_3\lambda'_2\theta'_i\lambda_4,$$

$$\mathrm{Cov}_{\xi_\perp}[\lambda'_1 M_{aa}^{-1/2}M_{au}\lambda_2, \lambda'_3 M_{bb}^{-1/2}M_{by}\lambda_4]$$

$$= \sum_{t,i}\lambda'_1 M_{aa}^{-1/2}a_{t-1}b'_{t+i+1}M_{bb}^{-1/2}\lambda_3\lambda'_2\theta'_i\lambda_4,$$

$$\mathrm{Cov}_{\xi_\perp}[\lambda'_1 M_{bb}^{-1/2}M_{bu}\lambda_2, \lambda'_3 M_{bb}^{-1/2}M_{by}\lambda_4]$$

$$= \sum_{t,i}\lambda'_1 M_{bb}^{-1/2}b_t b'_{t+i+1}M_{bb}^{-1/2}\lambda_3\lambda'_2\theta'_i\lambda_4,$$

$$\mathrm{Cov}_{\xi_\perp}[\lambda'_1 M_{aa}^{-1/2}M_{ay}\lambda_2, \lambda'_3 M_{bb}^{-1/2}M_{by}\lambda_4]$$

$$= \sum_{t,i}\lambda'_1 M_{aa}^{-1/2}a_{t-1}b'_{t-i+j}M_{bb}^{-1/2}\lambda_3\lambda'_4\theta_j\theta'_i\lambda_2$$

are all $O(T^{-1/2})$ because of the factor $M_{bb}^{-1/2}$, and

$$\mathrm{Cov}_{\xi_\perp}[\lambda'_1 M_{aa}^{-1/2}M_{au}\lambda_2, \lambda'_3 M_{dd}^{-1/2}M_{dy}\lambda_4]$$

$$= \sum_{t,i}\lambda'_1 M_{aa}^{-1/2}a_{t-1}d'_{t-i+j}M_{dd}^{-1/2}\lambda_3\lambda'_4\theta_j\theta'_i\lambda_2 = 0,$$

because

$$\sum_t a_{t-1}d'_{t-i+j} = \sum_t a_{t-1}d'_t M^{-i+j} = 0,$$

as a_{t-1} has been orthogonalized on d_t. Finally we need the result that the linear variables are asymptotically uncorrelated with the quadratic variables. This follows from the fact that the Gaussian distribution has third moment zero. ∎

Proof of Theorem 4. We use the result of Theorem 3 in the form

$$E[-2\log LR|\xi_\perp'\varepsilon] \overset{1}{=} K_1 + T^{-1}(K_2 + K_3 + K_4 + K_5 + K_6 + K_7)$$

and evaluate each in turn.

The Main Terms $K_1 + T^{-1}K_2$. We have

$$K_1 + T^{-1}K_2 = TE_{\xi_\perp}[\mathrm{tr}\{S_{au.v,b}S_{uu.v,b}^{-1}S_{ua.v,b}S_{aa.v,b}^{-1}\} + \tfrac{1}{2}T^{-1}\,\mathrm{tr}\{(S_{au}S_{ua}S_{aa}^{-1})^2\}].$$

This is the correction term given in Johansen (1999) based upon the regression equation

$$\underset{(n_v)}{\bar{\xi}'\Delta X_t} = \omega\,\underset{(n-n_v)}{\xi_\perp'\Delta X_t} + (\phi',(\bar{\xi}'-\omega\xi_\perp')\Psi)\,\underset{(n_v+n_z)}{(V_{t-1}',Z_{t-1}')'} + \delta'\underset{(n_a)}{A_{t-1}} + \bar{\xi}'\Phi\,\underset{(n_d)}{d_t} + \underset{(n_v)}{\tilde{\varepsilon}_t}$$

for the test that $\delta = 0$, when ξ is known ($\bar{\xi} = \xi(\xi'\xi)^{-1}$). Note that just as $\xi_\perp'\sum_{i=1}^{t-1}\varepsilon_i$ enters as a regressor in A_{t-1}, one can check that $\xi_\perp'\Delta X_t$ can be replaced by $B_t = (\xi_\perp'\Omega\xi_\perp)^{-1/2}\xi_\perp'\varepsilon_t$, because (17) shows that $\xi_\perp'\Delta X_t$ is a linear function of B_t, and (Z_{t-1}, d_t), which are corrected for in the regression. The result given there states that

$$K_1 + T^{-1}K_2 \overset{1}{=} n_v n_a + \frac{n_v n_a}{T}\left[\frac{1}{2}(n_v + n_a + 1) + (n_d + n_z + n)\right]$$
$$+ \frac{n_a}{T}[(n_a - 1)v(\xi) + 2(c(\xi) + c_d(\xi))],$$

with the coefficients $v(\xi), c(\xi)$, and $c_d(\xi)$ given in Theorem 4.

The rest of the proof of Theorem 4 deals with the problem of showing that

$$K_3 + K_4 + K_5 + K_6 + K_7 \overset{0}{=} v(\xi)n_a(n - n_v).$$

We first consider the terms K_4, K_5, K_6, and K_7. Because they have a factor of T^{-1}, we only need to consider the main terms.

The Term $K_4 = TE_{\xi_\perp}[\mathrm{tr}\{\kappa_\xi S_{vv}^{-1}\kappa_\xi S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}S_{au}\}]$. We reintroduce the product moments $M_{..}$ and find $T^{-1}S_{vv} = T^{-1}M_{vv.z,d}$, which converges toward $\mathrm{Var}(V_{t-1}|Z_{t-1}) = \Sigma_{vv.z}$, say. We find $S_{ua}S_{aa}^{-1}S_{ab} = M_{ua.z,d}M_{aa.z,d}^{-1}M_{ab.z,d}$ where the highest order term is $M_{ua}M_{aa}^{-1}M_{ab}$, and we find from (A.14) and Lemma 13 that

$$K_4 \overset{0}{=} E_{\xi_\perp}[\mathrm{tr}\{\kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi M_{ua}M_{aa}^{-1}M_{ab}M_{ba}M_{aa}^{-1}M_{au}\}]$$
$$= E_{\xi_\perp}[\mathrm{tr}\{\kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi N_{ua}M_{aa}^{-1/2}M_{ab}M_{ba}M_{aa}^{-1/2}N_{au}]$$
$$= \mathrm{tr}\{\kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi\}\mathrm{tr}\{M_{ab}M_{ba}M_{aa}^{-1}\}. \quad \textbf{(A.23)}$$

The Term $K_5 = TE_{\xi_\perp}[\mathrm{tr}\{S_{ba}S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_\xi^2 S_{vv}^{-1}S_{vb}\}]$. We replace $S_{ba}S_{aa}^{-1}S_{ab} = M_{ba.z,d}M_{aa.z,d}^{-1}M_{ab.z,d}$ *by* $M_{ba}M_{aa}^{-1}M_{ab}$. We next consider $\kappa_\xi S_{vv}^{-1}S_{vb}$. Because $Y_t = (V_t', Z_t')'$, we have the identity

$$(I_{n_v}, 0_{n_v \times n_z}) M_{yy.d}^{-1} M_{yb.d}$$

$$= (I_{n_v}, 0_{n_v \times n_z}) \begin{pmatrix} M_{vv.d} & M_{vz.d} \\ M_{zv.d} & M_{zz.d} \end{pmatrix}^{-1} \begin{pmatrix} M_{vb.d} \\ M_{zb.d} \end{pmatrix}$$

$$= M_{vv.z,d}^{-1} M_{vb.z,d} = S_{vv}^{-1} S_{vb}. \tag{A.24}$$

Thus

$$\kappa_\xi S_{vv}^{-1} S_{vb} = \kappa_\xi (I_{n_v}, 0_{n_v \times n_z}) M_{yy.d}^{-1} M_{yb.d} = \tilde{\kappa}_\xi' M_{yy.d}^{-1} M_{yb.d}$$

(see (34)). We find that $T^{-1} M_{yy.d}$ converges toward $\mathrm{Var}(Y_t) = \Sigma$ (see (A.19)), such that we can replace $\kappa_\xi S_{vv}^{-1} S_{vb}$ by $T^{-1} \tilde{\kappa}_\xi' \Sigma^{-1} M_{yb}$ because $M_{yb.d} = M_{yb} + O_P(1)$. We find from (A.21) that

$$K_5 \overset{0}{=} T^{-1} E_{\xi_\perp} [\mathrm{tr}\{M_{ba} M_{aa}^{-1} M_{ab} M_{by} \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} M_{yb}\}]$$

$$\overset{0}{=} \mathrm{tr}\{M_{ba} M_{aa}^{-1} M_{ab}\} \mathrm{tr}\{\tilde{\kappa}_\xi' \Sigma^{-1} \tilde{\kappa}_\xi\}. \tag{A.25}$$

The Term $K_6 = -TE_{\xi_\perp}[\mathrm{tr}\{S_{ua} S_{aa}^{-1} S_{au} \kappa_\xi S_{vv}^{-1} S_{vb} S_{bv} S_{vv}^{-1} \kappa_\xi\}]$. We replace $S_{au} = M_{au.z,d}$ by M_{au}, and from (A.24) we have that $S_{vv}^{-1} S_{vb}$ can be replaced by $T^{-1}(I_{n_y}, 0_{n_v \times n_z}) \Sigma^{-1} M_{yb}$. We find because $E_{\xi_\perp}[N_{ua} N_{au}] = n_a I_{n_u}$ (see (A.14) and Lemma 13) and $T^{-1} E_{\xi_\perp}[M_{yb} M_{by}] \overset{0}{=} n_b \Sigma$ (see (A.20)), that because N_{au} and N_{by} are asymptotically independent we get

$$K_6 \overset{0}{=} -T^{-1} E_{\xi_\perp} [\mathrm{tr}\{M_{ua} M_{aa}^{-1} M_{au} \tilde{\kappa}_\xi' \Sigma^{-1} M_{yb} M_{by} \Sigma^{-1} \tilde{\kappa}_\xi\}$$

$$\overset{0}{=} -T^{-1} \, \mathrm{tr}\{E_{\xi_\perp}[N_{ua} N_{au}] \tilde{\kappa}_\xi' \Sigma^{-1} E_{\xi_\perp}[M_{yb} M_{by}] \Sigma^{-1} \tilde{\kappa}_\xi\}$$

$$\overset{0}{=} -n_a n_b \, \mathrm{tr}\{\tilde{\kappa}_\xi' \Sigma^{-1} \tilde{\kappa}_\xi\}. \tag{A.26}$$

The Term $K_7 = -2TE_{\xi_\perp}[\mathrm{tr}\{S_{ba} S_{aa}^{-1} S_{au} \kappa_\xi S_{vv}^{-1} S_{vu} \kappa_\xi S_{vv}^{-1} S_{vb}\}]$. We find as before that we can use (A.24) to replace $\kappa_\xi S_{vv}^{-1} S_{vb}$ by $T^{-1} \tilde{\kappa}_\xi' \Sigma^{-1} M_{yb}$ and by a similar argument $\kappa_\xi S_{vv}^{-1} S_{vu}$ by $T^{-1} \tilde{\kappa}_\xi' \Sigma^{-1} M_{yu}$ and find

$$K_7 \overset{0}{=} -2T^{-1} E_{\xi_\perp} [\mathrm{tr}\{M_{ba} M_{aa}^{-1} M_{au} \tilde{\kappa}_\xi' \Sigma^{-1} M_{yu} \tilde{\kappa}_\xi' \Sigma^{-1} M_{yb}\}] \overset{0}{=} 0, \tag{A.27}$$

as (N_{au}, N_{yu}) and N_{by} are independent asymptotically.

The Term $K_3 = 2T^2 E_{\xi_\perp}[\mathrm{tr}\{S_{ba.v} S_{aa}^{-1} S_{au.b,v} \kappa_\xi S_{vv}^{-1} S_{vb} S_{bb}^{-1}\}]$. The stochastic variable is of the order $T^{1/2}$, and we need some more care, but it turns out to have an expectation that is of the order of a constant. We can replace a matrix by another, only if the relative error is $O_P(T^{-1})$. We want to replace S_{aa} by M_{aa} and expand as follows:

$$S_{aa} = M_{aa.z,d} = M_{aa.d} - M_{az.d} M_{zz.d}^{-1} M_{za.d} = M_{aa} - M_{az} M_{zz.d}^{-1} M_{za},$$

because $M_{ad} = 0$. Thus $M_{aa}^{-1}(S_{aa} - M_{aa}) \in O_P(T^{-1})$. Similarly we find

$$S_{bb} = M_{bb.z,d} = M_{bb.d} - M_{bz.d} M_{zz.d}^{-1} M_{zb.d} = M_{bb} - O_P(1)$$

and $M_{bb}^{-1}(S_{bb} - M_{bb}) \in O_P(T^{-1})$, such that we can replace S_{bb} and S_{aa} by M_{bb} and M_{aa}. Next we expand using $M_{ad} = 0$:

$$S_{au.v,b} = M_{au} - M_{ab}M_{bb.d}^{-1}M_{bu.d} - M_{ay.b,d}M_{yy.b,d}^{-1}M_{yu.b,d}, \tag{A.28}$$

$$S_{ba.v} = M_{ba.y,d} = M_{ba} - M_{by.d}M_{yy.d}^{-1}M_{ya}, \tag{A.29}$$

$$T^{-1}M_{yy.d} = (\Sigma - (\Sigma - T^{-1}M_{yy.d}))^{-1} \overset{\frac{1}{2}}{=} \Sigma^{-1} + \Sigma^{-1}(\Sigma - T^{-1}M_{yy})\Sigma^{-1}. \tag{A.30}$$

We insert from (A.24) $\kappa_\xi S_{vv}^{-1}S_{vb} = \tilde{\kappa}_\xi' M_{yy.d}^{-1}M_{yb.d}$ and use (A.28–A.30) to find

$$\begin{aligned}
&2T^2 S_{ba.v}S_{aa}^{-1}S_{au.v,b}\kappa_\xi S_{vv}^{-1}S_{vb}S_{bb}^{-1} \\
&\quad \overset{0}{=} 2T^2 M_{ba.y,d}M_{aa}^{-1}M_{au.y,b,d}\tilde{\kappa}_\xi' M_{yy.d}^{-1}M_{yb.d}M_{bb}^{-1} \\
&\quad \overset{0}{=} 2TM_{ba}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb.d}M_{bb}^{-1} \\
&\qquad - 2T^{-1}M_{by}\Sigma^{-1}M_{ya}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb} - 2T^{-1}M_{ba}M_{aa}^{-1}M_{ab}M_{bu.d}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb} \\
&\qquad - 2T^{-1}M_{ba}M_{aa}^{-1}M_{ay}\Sigma^{-1}M_{yu}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb} \\
&\qquad + 2M_{ba}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}(\Sigma - T^{-1}M_{yy})\Sigma^{-1}M_{yb},
\end{aligned}$$

where $T^{-1}M_{bb}$ has been replaced by I_{n_b} in the last four terms, because they are of the order of a constant. In this way K_3 is split into five terms:

$$K_3 = K_{31} + K_{32} + K_{33} + K_{34} + K_{35}.$$

This gives a number of contributions, which we investigate one by one. The term K_{31} is found from

$$\begin{aligned}
K_{31} &= 2TE_{\xi_\perp}[\mathrm{tr}\{M_{ba}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb.d}M_{bb}^{-1}\}] \\
&\overset{0}{=} 2TE_{\xi_\perp}[\mathrm{tr}\{M_{ba}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb}M_{bb}^{-1}\}] \\
&\quad - 2TE_{\xi_\perp}[\mathrm{tr}\{M_{ba}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yd}M_{dd}^{-1}M_{db}M_{bb}^{-1}\}].
\end{aligned}$$

The first term is of the order $T^{1/2}$, so asymptotic considerations are not enough and we calculate the mean directly. We find

$$\begin{aligned}
E_{\xi_\perp}[M_{aa}^{-1/2}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb}] &= E_{\xi_\perp}[M_{aa}^{-1/2}\sum_{s,t,i} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_i U_{s-i-1}b_s'] \\
&= \sum_{t,i} M_{aa}^{-1/2}a_{t-1}b_{t+i+1}'\,\mathrm{tr}\{\tilde{\kappa}_\xi'\Sigma^{-1}\theta_i\} \\
&\overset{0}{=} -M_{aa}^{-1/2}M_{ab}\,\mathrm{tr}\{\tilde{\kappa}_\xi'\Sigma^{-1}\tilde{\kappa}_\xi\},
\end{aligned}$$

because $\theta = -\tilde{\kappa}_\xi$ (see (32) and (34)). The second term is zero, because it is of the order of a constant, and N_{au} and N_{yd} are asymptotically independent. Thus we find

$$K_{31} \overset{0}{=} -2\,\mathrm{tr}\{\tilde{\kappa}_\xi'\Sigma^{-1}\tilde{\kappa}_\xi\}\mathrm{tr}\{M_{ba}M_{aa}^{-1}M_{ab}\}. \tag{A.31}$$

The term K_{32} is of the order of a constant, and we can apply Lemma 15 and find the two contributions using the fact that N_{au} and N_{ay} are independent of N_{by}. We get that $E_{\xi_\perp}[N_{ya}N_{au}] \overset{0}{=} n_a\theta = -n_a\tilde{\kappa}_\xi$ (see (A.16)) and $T^{-1}E_{\xi_\perp}[M_{yb}M_{by}] \overset{0}{=} n_b\Sigma$ (see (A.20)), such that

$$K_{32} = -2T^{-1}E_{\xi_\perp}[\mathrm{tr}\{\Sigma^{-1}M_{ya}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb}M_{by}\}]$$

$$\overset{0}{=} -2T^{-1}\,\mathrm{tr}\{\Sigma^{-1}E_{\xi_\perp}[N_{ya}N_{au}]\tilde{\kappa}_\xi'\Sigma^{-1}E_{\xi_\perp}[M_{yb}M_{by}]\}$$

$$\overset{0}{=} -2n_bn_a\,\mathrm{tr}\{\Sigma^{-1}\theta\tilde{\kappa}_\xi'\Sigma^{-1}\Sigma\} = 2n_bn_a\,\mathrm{tr}\{\tilde{\kappa}_\xi'\Sigma^{-1}\tilde{\kappa}_\xi\}. \quad \textbf{(A.32)}$$

The term K_{33} is zero because N_{bu} and N_{du} are asymptotically independent of N_{yb}. The term K_{34} is also zero because (N_{ay}, N_{by}) are asymptotically independent of M_{yu}; finally the term K_{35} is zero because (N_{au}, N_{by}) is asymptotically independent of N_{yy} (see Lemma 15).

Hence the contribution from K_3 is found from (A.32) and (A.31):

$$K_3 \overset{0}{=} 2(n_an_b - \mathrm{tr}\{M_{ba}M_{aa}^{-1}M_{ab}\})\mathrm{tr}(\kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi\}. \quad \textbf{(A.33)}$$

This completes the calculations, and it remains to compare (A.33) with (A.23) and (A.25)–(A.27), which gives

$$K_3 + K_4 + K_5 + K_6 + K_7 \overset{0}{=} n_an_b\,\mathrm{tr}(\kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi\}.$$

This completes the proof of Theorem 4. ■