EUROPEAN UNIVERSITY INSTITUTE

DEPARTMENT OF ECONOMICS

EUI Working Paper ECO No. 99/10

A Bartlett Correction Factor for Tests on the Cointegrating Relations

Søren JOHANSEN

BADIA FIESOLANA, SAN DOMENICO (FI)

Printed in Italy in March 1999
European University Institute
Badia Fiesolana
I-50016 San Domenico (FI)
Italy

A Bartlett correction factor for tests on the cointegrating relations

Søren Johansen,
European University Institute, Florence

March 1999

Abstract

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 the test on the cointegrating relations in the vector autoregressive model for the $n-$dimensional process X_t given by

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

where ε_t are 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, like constant, linear term etc. The matrices α and β are $(n \times r)$ and the matrices $\Gamma_1, \ldots, \Gamma_{k-1}$, are $(n \times n$, Φ $(n \times n_d)$ and $\rho\,(n_D \times 1)$. For the analysis of this paper we assume that X_t is $I(1)$. A typical example of the deterministic terms is to let $d_t = 1$, and $D_t = t$, corresponding to having a linear trend in the process and trend stationary combinations $\beta' X_t$. The formulation below also allows, with minor modifications, the possibility that n_D and n_d are zero.

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. The finite sample distribution, however, is not always well approximated by the asymptotic distribution, see for instance Fachin (1997), Jacobsen and Gredenhoff (1998), and Jacobson, Vredin, and Warne (1998), Haug (1998), to mention a few of the many studies of the finite sample properties of the test of restrictions on β.

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 Bartlett (1937). For a recent survey of the theory of this type of correction see Cribaro-Neto and Cordeiro (1996). Briefly the method consists of calculating the expectation of the likelihood ratio (LR) statistic in the form $-2LogLR$ 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

$$E[-2 \log LR] \sim A(1 + \frac{B(\theta)}{T}),$$

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

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

has a distribution that is closer to the limit distribution, see Lawley (1956) for a proof of this statement under classical i.i.d. assumptions, that are not satisfied in the $I(1)$ model.

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 above characteristic numbers. We find for instance for the test that $\beta = H\psi$, ($H(n \times s)$) that $A = r(n-s)$, and for $m = n+s-r+1+2n_D$

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

The coefficients $v(\alpha), c(\alpha)$, and $c_d(\alpha)$ are given in Theorem 5 below. This result implies that one can see for which combinations of the parameters the usual χ^2 approximation breaks down, and more constructively when it is useful. In between there is an area where the Bartlett correction can serve as an improvement to the usual asymptotic results.

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 an expansion is given of the estimators of this reduced rank regression, and then an expansion is given of the log likelihood ratio test statistic. In Section 4 the main result on the Bartlett corrections are given and the results specialized to the models discussed in Section 2, and finally in Section 5 some simulation experiments are conducted which 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 analysed by reduced rank regression, see Johansen

(1996) for a detailed analysis of the models. The models allow deterministic terms of a suitably simple type, that covers many of the usual situations. We show how the correction term for the test of each of the models can be calculated simply if we have the correction term for a simple hypothesis, and 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.

- $\mathcal{M}_0$ **Unrestricted cointegrating space**

  The model is given by the 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 which can be expressed as

$$\beta = H\psi,$$

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

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

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

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

  with $\psi(s + n_D, r)$. 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

$$\begin{pmatrix} \beta \\ \rho \end{pmatrix} = \left(\begin{pmatrix} \beta_1^0 \\ \rho_1^0 \end{pmatrix} \psi_1, \begin{pmatrix} \psi_2 \\ \rho_2 \end{pmatrix} \right),$$

  where the matrices β_1^0 $(n \times r_1)$ of rank r_1 and ρ_1^0 $(n_D \times r_1)$ are known and the matrices ψ_1 $(r_1 \times r_1)$, ψ_2 $(n \times r_2)$, and ρ_2 $(n_D \times r_2)$ are unknown

$(r = r_1 + r_2)$, corresponding to prespecified coefficients β_1^0 and ρ_1^0 in some of the cointegrating relations. The likelihood ratio test of $\mathcal{M}_2$ in $\mathcal{M}_0$, satisfied

$$-2\log LR(\mathcal{M}_2|\mathcal{M}_0) \overset{w}{\to} \chi^2(r_1(n+n_D-r)).$$

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 below would have to be modified.

In the following sections we derive a correction factor for 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 a parameters $\beta^0 = H\psi_0$ 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

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

such that

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

Hence we see that the correction for the test we are really interested in, namely $\mathcal{M}_1$ in $\mathcal{M}_0$, can be found as the difference of the corrections to two tests of simple hypotheses on β and ρ in $\mathcal{M}_0$ and $\mathcal{M}_2$. Thus, if we can find a general result which allows us to derive a correction for a simple hypothesis on β and ρ in these various models, then we can derive the corrections by subtraction.

2.1 The deterministic terms

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

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

for some matrix M with the property that

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

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

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

and M has eigenvalues equal to 1. If s_1, s_2, and s_3 are quarterly dummies we can consider combinations like $d_t' = (1, t, s_1(t), s_2(t), s_3(t))$. 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 $\pm 1, \pm i$. Note that intervention dummies are not covered by this formulation and will give rise to more complicated formulae.

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), we find that the process can be represented by

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

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

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

It follows that

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

say. Taking expectations in (1) we find

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

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 10, shows that the sums are convergent, since C_i are exponentially decreasing. ■

We next show how the simple hypotheses on β and ρ in $\mathcal{M}_0$, $\mathcal{M}_1$, and $\mathcal{M}_2$, give rise to regression equations which 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

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

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

It is convenient for the calculations to reparametrize the model defining new parameters and regressors which involve the true value. In the following subsections we therefore need a notation for the true value of the parameters, as well as 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

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

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 non stationary components:

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

We find, using the true value of the parameters,

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

We choose new parameters

$$\begin{array}{lcll} \psi' & = & \beta'(I_n - C\Gamma^0)\bar{\beta}^0 & (r \times r) \\ \delta_1' & = & \beta'\beta_\perp^0 & (r \times (n-r)) \\ \delta_2' & = & \rho' - \psi'\rho^{0\prime} & (r \times n_D) \end{array}$$

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

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

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

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

Notice that the model is overparametrized since

$$\alpha\eta', \psi\eta^{-1}, \delta\eta^{-1}$$

give the same probability measure as (α, ψ, δ) for any η $(r \times r)$ of full rank. We can achieve just identification by choosing $\psi = I_r$, that is, by absorbing $\psi(r \times r)$ into $\alpha(n \times r)$ and adjusting 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

$$\begin{aligned} V_{t-1} &= \beta^{0\prime}X_{t-1}-E_0(\beta^{0\prime}X_{t-1}), && (7)\\ Z_{t-1} &= (\Delta X_{t-1}^{\prime}-E_0(\Delta X_{t-1}^{\prime}),\ldots,\Delta X_{t-k+1}^{\prime}-E_0(\Delta X_{t-k+1}^{\prime}))^{\prime}. && (8)\end{aligned}$$

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

$$\alpha_{\perp}^{\prime}(X_t-X_0)=\alpha_{\perp}^{\prime}\sum_{i=1}^{k-1}\Gamma_i(X_{t-i}-X_{-i})+\alpha_{\perp}^{\prime}\sum_{i=1}^{t}(\varepsilon_i+\Phi d_i).$$

By subtracting $\sum_{i=1}^{k-1}\alpha_{\perp}^{\prime}\Gamma_iX_t$ on both sides and replacing t by $t-1$ we get

$$\alpha_{\perp}^{\prime}\Gamma X_{t-1}=\alpha_{\perp}^{\prime}X_0+\alpha_{\perp}^{\prime}\sum_{i=1}^{k-1}\Gamma_i(X_{t-i-1}-X_{t-1}-X_{-i})+\alpha_{\perp}^{\prime}\sum_{i=1}^{t-1}(\varepsilon_i+\Phi d_i).$$

Because we are correcting for lagged differences in the 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 non stationary regressor given by the common trends

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

where A_0 depends on initial conditions.

The model equation (6) in the new variables and with suitably redefined parameters becomes

$$\underset{(n)}{\Delta X_t}=\alpha\underset{(r)}{V_{t-1}}+\alpha\delta^{\prime}\underset{(n-r+n_D)}{A_{t-1}}+\Psi\underset{((k-1)n)}{Z_{t-1}}+\Phi\underset{(n_d)}{d_t}+\underset{(n)}{\varepsilon_t}, \tag{10}$$

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 $(V_{t-1}^{\prime},A_{t-1}^{\prime})$ corrected for Z_{t-1} and d_t. Under the hypothesis $\delta=0$ the parameters can be found by regression of ΔX_t on Z_{t-1} and d_t.

Later we shall 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 $\beta = H\psi$ $(\psi(s \times r))$ and equation (1) is

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

We consider again the simple hypothesis $\beta = H\psi^0$, $\rho = \rho^0$, corresponding to a point in $\mathcal{M}_1$. We want to show that by introducing the true parameters β^0, $\rho^0, \ldots$ as before we can reformulate the equations to have the form (10), such that a test of a simple hypothesis is a test that $\delta = 0$.

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

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

In this case we have that $\beta^0 = H\psi^0$ implies that $\beta_\perp^0 = (H_\perp, H(H'H)^{-1}\psi_\perp^0)$ and hence

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

We introduce the new parameters

$$\begin{array}{rcll} \psi_1' & = & \psi' H'(I_n - C^0\Gamma^0)\bar{\beta}^0 & (r \times r) \\ \delta_1' & = & \psi' \psi_\perp^0 & (r \times (s-r)) \\ \delta_2' & = & \rho' - \psi_1' \rho^{0\prime} & (r \times n_D) \end{array}$$

since then

$$\begin{aligned} &\psi' H' X_{t-1} + \rho' D_t \\ &= \psi_1'(\beta^{0\prime} X_{t-1} + \rho^{0\prime} D_t) + \delta_1'(0_{(s-r)\times(n-s)}, I_{s-r})(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1} + \delta_2' D_t. \end{aligned}$$

The hypothesis is formulated as $\delta = 0$, $\psi_1 = I_r$. We let V_{t-1} and Z_{t-1} be defined by (7) and (8), and replace in this case the $(s-r)-$dimensional non stationary 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 ψ_1' is absorbed in α and the remaining parameters are adjusted accordingly such that also Z_{t-1}, and V_{t-1} have mean zero. The hypothesis of interest is $\delta = 0$, which corresponds to $\rho = \rho^0,\ \beta = \beta^0 = H\psi^0$.

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 $\mathcal{M}_2$

Again we investigate a simple hypothesis on β and ρ which can be formulated as $\psi_2 = \beta_2^0,\ \rho_2 = \rho_2^0$. The parameter α is decomposed corresponding to the cointegrating parameter into $\alpha = (\alpha_1, \alpha_2)$, such that

$$\alpha(\beta' X_{t-1} + \rho' D_t) = \alpha_1\psi_1'(\beta_1^{0\prime} X_{t-1} + \rho_1^{0\prime} D_t) + \alpha_2(\psi_2' X_{t-1} + \rho_2' D_t).$$

In this case we absorb ψ_1' into α_1 and include the regressor $\beta_1^{0\prime} X_{t-1}$ with the lagged differences Z_{t-1} instead of with V_{t-1}. We then decompose the second component $\psi_2' X_{t-1} + \rho_2' D_t$ of the process as

$$\begin{aligned} &\psi_2' X_{t-1} + \rho_2' D_t \\ &= (\psi_2'(I_n - C^0\Gamma^0)\bar{\beta}^0)\beta^{0\prime} X_{t-1} + \psi_2'\beta_\perp^0(\alpha_\perp^{0\prime}\Gamma^0\beta_\perp^0)^{-1}\alpha_\perp^{0\prime}\Gamma^0 X_{t-1} + \rho_2' D_t. \end{aligned}$$

Now

$$\psi_2'(I_n - C^0\Gamma^0)\bar{\beta}^0\beta^{0\prime} X_{t-1} = (\phi_1', \phi_2') \begin{pmatrix} \beta_1^{0\prime} X_{t-1} \\ \beta_2^{0\prime} X_{t-1} \end{pmatrix} = \phi_1'\beta_1^{0\prime} X_{t-1} + \phi_2'\beta_2^{0\prime} X_{t-1},$$

so that

$$\psi_2' X_{t-1} + \rho_2' D_t = \phi_1'(\beta_1^{0\prime} X_{t-1} + \rho_1^{0\prime} D_t) + \phi_2'(\beta_2^{0\prime} X_{t-1} + \rho_2^{0\prime} D_t) + \delta' A_{t-1}.$$

We have defined the parameters

$$\begin{array}{lcll} (\phi_1', \phi_2') & = & \psi_2'(I_n - C^0\Gamma^0)\bar{\beta}^0 & (r_2 \times r_1), (r_2 \times r_2) \\ \delta_1' & = & \psi_2'\beta_\perp^0 & (r_2 \times (n-r)) \\ \delta_2' & = & \rho' - \phi_1'\rho_1^{0\prime} - \phi_2'\rho_2^{0\prime} & (r_2 \times n_D). \end{array}$$

The hypothesis of interest in the new parameters is $\delta = 0$, $\phi_1 = 0$, $\phi_2 = I_{r_2}$. The process A_{t-1} is defined by (9) such that the equation becomes

$$\underset{(n)}{\Delta X_t} = \alpha_2 \underset{(r_2)}{V_{t-1}} + \alpha_2\delta' \underset{(n-r+n_D)}{A_{t-1}} + \Psi \underset{(r_1+(k-1)n)}{Z_{t-1}} + \Phi \underset{(n_d)}{d_t} + \underset{(n)}{\varepsilon_t}, \tag{14}$$

where

$$V_{t-1} = \beta_2^{0\prime} X_{t-1} \tag{15}$$

$$Z_{t-1} = (X_{t-1}'\beta_1^0, \Delta X_{t-1}', \ldots, \Delta X_{t-k+1}')' \tag{16}$$

both corrected for their mean, and where again α, Ψ, and Φ have been redefine to accommodate the change in regressors. It is seen that equation (14) is of the form (10), with a changed definition of V_{t-1} and Z_{t-1}, since the assumed stationary combinations $\beta_1^{0\prime} X_{t-1} - E_0(\beta_1^{0\prime} X_{t-1})$ are moved to the lagged differences. The hypothesis can be tested as $\delta = 0$, as the other restrictions are absorbed in α and Ψ.

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 the those given in (10). But we still need to preserve the properties that under the null hypothesis the process $(V_{t-1}', Z_{t-1}')'$ is a mean zero stationary autoregressive process, and that V_{t-1} and $\delta' A_{t-1}$ have the factor α (or α_2) in front. All models (10), (13), and (14) have the property that they can be solved by reduced rank regression and that under the null hypothesis, $\delta = 0$, the model is solved by simple regression.

3 A general reduced rank equation and an expansion of the estimators and the test statistic

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

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_2	$n-r+n_D$	$r_1+(k-1)n$	α_2

Table 1: The choice of dimensions in the general regression model (17) which corresponds to the hypotheses discussed in Section 1.

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

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 the regressors V_{t-1}, A_{t-1}, and Z_{t-1}, and their dimensions, see Table 1.

In all cases the variables V_{t-1} and Z_{t-1} are, under the hypothesis $\delta = 0$, stationary with mean zero and A_{t-1} is a linear function of $\xi_\perp^{0\prime}\sum_{i=1}^t(\varepsilon_i + \Phi^0 d_i)$, with $\xi^0 = \alpha^0$ or α_2^0. 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.

The very detailed calculations in this paper continues the work in Johansen (1999) where the correction was found for the model where $\text{sp}(\xi)$ is known. The result derived there provide the main term of the Bartlett correction in the situation where ξ is unknown and we therefore briefly discuss this situation in the next subsection. We then give an expansion of the reduced rank estimator around the regression estimator valid under the null hypothesis and finally we derive an expansion of a simple hypothesis for $\delta = 0$.

3.1 The analysis for fixed $\xi = \xi^0$

Note that if $\xi = \xi^0\phi$, where ξ_0 is known the model equation is

$$\Delta X_t = \xi^0\phi' V_{t-1} + \xi^0\delta' A_{t-1} + \Psi Z_{t-1} + \Phi d_t + \varepsilon_t,$$

which implies, for $\bar{\xi}^0 = \xi^0(\xi^{0\prime}\xi^0)^{-1}$, that

$$\begin{aligned}\bar{\xi}^{0\prime}\Delta X_t &= \phi' V_{t-1} + \delta' A_{t-1} + \bar{\xi}^{0\prime}\Psi Z_{t-1} + \bar{\xi}^{0\prime}\Phi d_t + \bar{\xi}^{0\prime}\varepsilon_t \\ \xi_\perp^{0\prime}\Delta X_t &= \xi_\perp^{0\prime}\Psi Z_{t-1} + \xi_\perp^{0\prime}\Phi d_t + \xi_\perp^{0\prime}\varepsilon_t.\end{aligned}$$

Hence the model for $\bar{\xi}^{0\prime}\Delta X$ given $\xi_{\perp}^{0\prime}\Delta X_t$ and the past is

$$\bar{\xi}^{0\prime}\Delta X_t = \omega B_t + \phi' V_{t-1} + \delta' A_{t-1} + (\bar{\xi}^{0\prime} - \omega\xi_{\perp}^{0\prime})\Psi Z_{t-1} + (\bar{\xi}^{0\prime} - \omega\xi_{\perp}^{0\prime})\Phi d_t + \tilde{\varepsilon}_t, \tag{18}$$

for

$$B_t = \xi_{\perp}^{0\prime}\Delta X_t,\ \omega = \bar{\xi}^{0\prime}\Omega\xi_{\perp}^{0\prime}(\xi_{\perp}^{0\prime}\Omega\xi_{\perp}^{0})^{-1},$$

and

$$\tilde{\varepsilon}_t = \bar{\xi}^{0\prime}\varepsilon_t - \omega\xi_{\perp}^{0\prime}\varepsilon_t = (\xi^{0\prime}\Omega^{-1}\xi^0)^{-1}\xi^{0\prime}\Omega^{-1}\varepsilon_t.$$

We define the normalized error

$$U_t = (\xi^{0\prime}\Omega^{0-1}\xi^0)^{-\frac{1}{2}}\xi^{0\prime}\Omega^{0-1}\varepsilon_t,$$

such that for the true value of the parameters

$$\tilde{\varepsilon}_t = (\xi^{0\prime}\Omega^{0-1}\xi^0)^{-\frac{1}{2}}U_t.$$

We 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 X, U, and V

$$M_{uv.x} = M_{uv} - M_{ux}M_{xx}^{-1}M_{xv},$$

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

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

These moment matrices are natural when the likelihood function is concentrated with respect to Ψ and Φ. The maximum likelihood estimator of δ (for known $\xi = \xi^0$) is found by regression in (18)

$$\begin{aligned}\tilde{\delta}(\xi^0) &= M_{aa.v,z,b,d}^{-1}M_{a0.v,z,b,d}\bar{\xi}^0 = S_{aa.v,b}^{-1}S_{a0.v,b}\bar{\xi}^0\\ &= \delta + S_{aa.v,b}^{-1}S_{a\varepsilon.v,b}\Omega^{-1}\xi^0(\xi^{0\prime}\Omega^{-1}\xi^0)^{-1}.\end{aligned}$$

The test for the hypothesis $\delta = 0$, (still for known ξ^0) is under the hypothesis, $\delta = 0$, equal to

$$\begin{aligned}
& LR^{-2/T}(\delta = 0|\xi^0 \text{ known}) \\
= & |\bar{\xi}^{0\prime} S_{00.a,v,b}\bar{\xi}^0|/|\bar{\xi}^{0\prime} S_{00.v,b}\bar{\xi}^0| = |S_{uu.a,v,b}|/|S_{uu.v,b}| \\
= & |I_n - S_{uu.v,b}^{-1} S_{ua.v,b} S_{aa.v,b}^{-1} S_{au.v,b}| \\
= & |I_n - T^{-1}Q|,
\end{aligned} \tag{19}$$

say, such that

$$-2\log LR(\delta = 0|\xi_0 \text{ known}) \overset{1}{=} tr\{Q\} + \frac{1}{2T} tr\{Q^2\}. \tag{20}$$

We use here the notation $\overset{d}{=}$ to indicate that we have kept terms of order T^{-d}. An approximation to the expectation of $-2\log LR$ given by (20) was derived in Johansen (1999) and turns out to give the main contribution to the expectation derived in this paper.

3.2 The first order conditions for the estimation of ξ, δ, and Ω

In the rest of the paper we refer to the true value, the one for which we calculate the expectation, without the superscript since that simplifies the notation. We express the results below in the notation for the concentrated model, where the parameters Ψ and Φ have been eliminated, that is, we use the moment matrices S rather than M.

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

$$\tilde{\xi} = (S_{0v} + S_{0a}\tilde{\delta})(S_{vv} + \tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})^{-1} \tag{21}$$

$$\tilde{\Omega} = T^{-1}(S_{00} - \tilde{\xi}(S_{vv} + \tilde{\delta}' S_{av} + S_{va}\tilde{\delta} + \tilde{\delta}' S_{aa}\tilde{\delta})\tilde{\xi}') \tag{22}$$

$$\tilde{\delta} = S_{aa}^{-1}(S_{a0}\tilde{\Omega}^{-1}\tilde{\xi}(\tilde{\xi}'\tilde{\Omega}^{-1}\tilde{\xi})^{-1} - S_{av}). \tag{23}$$

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

Under the null hypothesis $\delta = 0$, the estimators are

$$\begin{aligned} \hat{\xi} &= S_{0v}S_{vv}^{-1} = \xi + S_{\varepsilon v}S_{vv}^{-1} \\ \hat{\Omega} &= T^{-1}(S_{00} - S_{0v}S_{vv}^{-1}S_{v0}) = T^{-1}S_{\varepsilon\varepsilon.v}. \end{aligned}$$

We next need a result about regression estimators for stationary processes and the type of deterministic terms we consider.

Lemma 2 *Let* $S_t = \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i}$ *with* θ_i *decreasing exponentially be a stationary process, and let* d_t *satisfy* $d_{t+1} = Md_t$*, with* $|eig(M)| = 1$*, and let*

$$\hat{\eta} = M_{dd}^{-1}M_{ds},$$

then

$$\hat{\eta}' M_{dd}\hat{\eta} = M_{sd}M_{dd}^{-1}M_{ds} \in O_P(1). \tag{24}$$

This result follows from Lemma 11. 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 3 *The estimators* $\tilde{\xi}$, $\tilde{\Omega}$, *and* $\tilde{\delta}$ *can be expanded around* $\hat{\xi}$, $\hat{\Omega}$ *and* 0 *respectively:*

$$\tilde{\xi} - \hat{\xi} = [S_{\varepsilon a.v}\tilde{\delta} - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta} - \hat{\xi}\tilde{\delta}' S_{av}]S_{vv}^{-1} + O_P(T^{-2}) \tag{25}$$

$$\begin{aligned} &(\hat{\Omega} - \tilde{\Omega}) \\ &= T^{-1}[S_{\varepsilon a.v}\tilde{\delta}\hat{\xi}' + \hat{\xi}\tilde{\delta}' S_{a\varepsilon.v} - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta}\hat{\xi}'] \\ &+T^{-1}[S_{\varepsilon a.v}\tilde{\delta} - \hat{\xi}\tilde{\delta}' S_{av} - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta}]S_{vv}^{-1}[\tilde{\delta}' S_{a\varepsilon.v} - S_{va}\tilde{\delta}\hat{\xi}' - \tilde{\delta}' S_{aa}\tilde{\delta}\hat{\xi}'] + O_P(T^{-\frac{5}{2}}) \\ &= T^{-1}\hat{\Omega}Q_0 + T^{-2}\hat{\Omega}Q_1 + O_P(T^{-\frac{5}{2}}) \end{aligned} \tag{26}$$

$$\tilde{\delta} = S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} + O_P(T^{-1}S_{aa}^{-\frac{1}{2}}). \tag{27}$$

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

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

since

$$\hat{\xi}\tilde{\delta}' = P(\hat{\xi}, \hat{\Omega})S_{\varepsilon a.v}S_{aa}^{-1} + O_P(T^{-1}S_{aa}^{-\frac{1}{2}}). \tag{28}$$

Proof. *Proof of (25):* From (21) we expand and find with $\hat{\xi} - \xi = S_{\varepsilon v}S_{vv}^{-1}$

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

We further have

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

since

$$S_{0a} - \hat{\xi}S_{va} = S_{\varepsilon a} - (\hat{\xi} - \xi)S_{va} = S_{\varepsilon a} - S_{\varepsilon v}S_{vv}^{-1}S_{va} = S_{\varepsilon a.v}, \tag{29}$$

which proves the result.

Proof of (26): From (22) we find

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

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

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

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

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

where we have used (29). The term of order T^{-1} is

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

Proof of (27): From (23) we have

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

Since $\tilde{\xi} - \hat{\xi}$ 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 that

$$\tilde{\delta} = S_{aa}^{-1}(S_{a0}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} - S_{av}) + O_P(T^{-1}S_{aa}^{-\frac{1}{2}}).$$

The main term can be reduced as follows:

$$\begin{aligned}
&S_{aa}^{-1}(S_{a0}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} - S_{av}) \\
&= S_{aa}^{-1}((S_{a\varepsilon} + S_{av}\xi)\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} - S_{av}) \\
&= S_{aa}^{-1}S_{a\varepsilon}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} + S_{aa}^{-1}S_{av}(\xi - \hat{\xi})\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} \\
&= S_{aa}^{-1}S_{a\varepsilon}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} - S_{aa}^{-1}S_{av}S_{vv}^{-1}S_{v\varepsilon}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} \\
&= S_{aa}^{-1}(S_{a\varepsilon} - S_{av}S_{vv}^{-1}S_{v\varepsilon})\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1} \\
&= S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1}.
\end{aligned}$$

This completes the proof of Theorem 3 on the expansion of the estimators.
■

We conclude this section by stating the theorem on the expansion of the likelihood ratio test for $\delta = 0$, in (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

$$\begin{aligned} -2\log LR &= -T\log|I_n - \hat{\Omega}^{-1}(\hat{\Omega}-\tilde{\Omega})| \\ &\overset{1}{=} Ttr\{\hat{\Omega}^{-1}(\hat{\Omega}-\tilde{\Omega})\} + \tfrac{T}{2}tr\{(\hat{\Omega}^{-1}(\hat{\Omega}-\tilde{\Omega}))^2\}. \end{aligned}$$

We apply Theorem 3 and find that, see (26),

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

such that

$$-2\log LR \overset{1}{=} tr\{Q_0\} + T^{-1}(tr\{Q_1\} + \frac{1}{2}tr\{Q_0^2\}). \tag{30}$$

Let further

$$\begin{aligned} \Sigma_{vv.z} &= \operatorname{Var}(V_t|Z_t) = \Sigma_{vv} - \Sigma_{vz}\Sigma_{zz}^{-1}\Sigma_{zv} \\ \kappa_\xi^2 &= \operatorname{Var}(\bar{\xi}'\varepsilon_t|\xi_\perp'\varepsilon_t) = (\xi'\Omega^{-1}\xi)^{-1}. \end{aligned}$$

We can then prove

Theorem 4 *An expansion of the log likelihood ratio test for $\delta = 0$ based upon (17) is given by*

$$\begin{aligned} -2\log LR = \; & Ttr\{S_{aa.b,v}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} + \tfrac{1}{2T}tr\{(S_{ua}S_{aa}^{-1}S_{au})^2\} \\ & +2tr\{S_{aa}^{-1}S_{ua.v,b}\kappa_\xi S_{vv}^{-1}S_{vb}S_{ba.v}\} \\ & +T^{-1}tr\{\kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}S_{au}\}, \\ & +tr\{S_{ba}S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_\xi^2 S_{vv}^{-1}S_{vb}\} \\ & -tr\{S_{ua}S_{aa}^{-1}S_{au}\kappa_\xi S_{vv}^{-1}S_{vb}S_{bv}S_{vv}^{-1}\kappa_\xi\} \\ & -2tr\{S_{ba}S_{aa}^{-1}S_{au}\kappa_\xi S_{vv}^{-1}S_{vu}\kappa_\xi S_{vv}^{-1}S_{vb}\} \end{aligned}$$

Note that the first two terms are the test statistic for $\delta = 0$ if ξ were known, see (19) and (20), the next term is of the order $O_P(T^{-\frac{1}{2}})$, and the last four terms are of the order $O_P(T^{-1})$. The proof of Theorem 4 based upon the expansions in Theorem 3 is given in the Appendix.

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 and then give the calculations of some coefficients in the cointegrated VAR that are needed to formulate the main result. Finally we state the main result and specialize it to the various situations covered by the general formulation as indicated in Section 2.

We choose to calculate the conditional expectation of the likelihood ratio test statistic conditioning on the process $\xi_\perp' \varepsilon_t$. The argument for that is, that it is easier to do so since 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 conditional mean does not depend on the conditioning variable such that what we find is also the unconditional mean.

When $\xi_\perp' \varepsilon_t$ is fixed so is the regressor A_t which we denote with a_{t-1}. We further define

$$b_t = (\xi_\perp' \Omega \xi_\perp)^{-\frac{1}{2}} \xi_\perp' \varepsilon_t,$$

of dimension $n_b = n - n_v$.

4.1 The conditioning variables

The fixed regressors a_{t-1} and b_t are defined in terms of $\xi_\perp' \sum_{i=1}^{t-1} \varepsilon_i$, 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.

When we do not condition on $\xi_\perp' \varepsilon_t$ we have the following relations

$$T^{-\frac{1}{2}} \sum_{t=1}^{T} b_t \in O_P(1), \tag{31}$$

$$\left(\sum_{t=1}^{T} a_{t-1} a_{t-1}' \right)^{-1} \sum_{t=1}^{T} a_{t-1} a_{t-1-k}' \xrightarrow{P} I_{n_a}, \text{ for all } k, \tag{32}$$

$$T^{-1}\sum_{t=1}^{T} b_t b_t' \xrightarrow{P} I_{n_b}, \tag{33}$$

$$T^{-1}\sum_{t=1}^{T} b_t b_{t-i-1}' \xrightarrow{P} 0,\ i = 0, 1, \ldots \tag{34}$$

$$T^{-1}\sum_{t=1}^{T} b_t (b_{t-i-1}' K b_{t-i-1}) b_t' \xrightarrow{P} tr\{K\} I_{n_b} \tag{35}$$

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

$$M_{ba} M_{aa}^{-1} M_{ab} \xrightarrow{w} \int_0^1 (dW) F' \left(\int_0^1 F F' du \right)^{-1} \int_0^1 F (dW)', \tag{36}$$

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

$$(\xi_\perp' \Omega \xi_\perp)^{-\frac{1}{2}} \xi_\perp' T^{-\frac{1}{2}} \sum_{i=1}^{[Tu]} \varepsilon_i \xrightarrow{w} W(u),$$

of dimension $n_b = n - n_v$. The process F is defined as the limit of A_{t-1}. If for instance $d_t = 1$ and $D_t = 0$ then

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

such that in the direction $\alpha_\perp' \Phi$ ($\neq 0$), the process grows linearly and orthogonal to that it behaves like a random walk. In this case F is of dimension $n_b = n_a = n - r$ and

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

If instead $D_t = t - \bar{t}$ and $d_t = 1$, and

$$A_{t-1} = \begin{pmatrix} \alpha_\perp' (\sum_{i=1}^{t-1} \varepsilon_i - T^{-1} \sum_{t=1}^{T} \sum_{i=1}^{t-1} \varepsilon_i) + \Phi (t - \bar{t}) \\ t - \bar{t} \end{pmatrix},$$

then a non singular linear transformation of A_{t-1}, which leaves the statistic invariant, removes the coefficient Φ, such that in this case the process F is of dimension $n_a = n - r + 1$, and given by

$$F(u) = \begin{pmatrix} W(u) - \int_0^1 W(s) ds \\ u - \frac{1}{2} \end{pmatrix}.$$

Finally of A_{t-1} is given by (12) with $D_t = t$, both corrected for their mean, and $d_t = 1$, it is possible to prove that

$$F_i(u) = W_i(u) - \int_0^1 W_i(s)ds, i = 1, \ldots, n-s,$$
$$F_{n-s+1} = u - \tfrac{1}{2}.$$

When conditioning on the sequence $\xi_\perp' \varepsilon_t$ we assume that relations (31)-(36) hold for the sequence we are fixing. That is, we replace $\sum_{t=1}^T a_{t-1}a_{t-1-k}'$ by $\sum_{t=1}^T a_{t-1}a_{t-1}' = M_{aa}$, $T^{-1}\sum_{t=1}^T b_t b_{t-k}'$ by I_{n_b} or 0, etc. in order to simplify the expressions.

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. It is in all cases of dimension $n_y = n_v + n_z = r + (k-1)n$ and is a stationary autoregressive process given by the equation

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

where

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

We find the representation

$$Y_t = \sum_{\nu=0}^{\infty} P^\nu Q \varepsilon_{t-\nu} = \sum_{\nu=0}^{\infty} \theta_\nu U_{t-\nu-1} + \psi_\nu b_{t-\nu-1},$$

where we have decomposed ε_t into the components U_t and b_t, and

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

Note that by the definition of Σ

$$\sum_{\nu=0}^{\infty}(\theta_\nu\theta_\nu' + \psi_\nu\psi_\nu') = \mathrm{Var}(Y_t) = \Sigma. \tag{37}$$

We find

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

Since

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

we find, when ξ is either α or $\alpha_2 = \alpha(0_{r_2\times r_1}, I_{r_2})'$, that

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

such that

$$\theta = -(I_{n_v}, 0_{n_v\times n_z})'\kappa_\xi = -\tilde{\kappa}_\xi, \tag{38}$$

where

$$\kappa_\xi^2 = \mathrm{Var}(\bar{\xi}'\varepsilon_t|\xi_\perp'\varepsilon_t) = \mathrm{Var}((\xi'\Omega^{-1}\xi)^{-1}\xi'\Omega^{-1}\varepsilon_t) = (\xi'\Omega^{-1}\xi)^{-1}.$$

4.3 The main results

We can finally state the main result about the Bartlett correction factor. The proof is left for the Appendix and we give here some corollaries, which show explicitly how the correction can be used for the tests mentioned in Section 2.

Theorem 5 *The conditional expectation of the log likelihood ratio test for the hypothesis $\delta = 0$ in (17) is given by*

$$\begin{aligned}&E[-2\log LR(\delta=0)|\xi_\perp'\varepsilon]\\ &\overset{1}{=} n_v n_a + \tfrac{n_v n_a}{T}[\tfrac{1}{2}(n_v+n_a+1)+n_d+n+n_z]\\ &+\tfrac{n_a}{T}[(n-n_v+n_a-1)v(\xi)+2(c(\xi)+c_d(\xi))]\end{aligned}$$

where

$$\begin{aligned}v(\xi) &= tr\{V_\xi\}, V_\xi = \tilde{\kappa}_\xi\tilde{\kappa}_\xi'\Sigma^{-1}\\ c(\xi) &= tr\{P(I_{n_y}+P)^{-1}V_\xi\} + tr\{[P\otimes(I_{n_y}-P)V_\xi][I_{n_y}\otimes I_{n_y} - P\otimes P]^{-1}\}\\ c_d(\xi) &= tr\{[M\otimes(I_{n_y}-P)V_\xi][I_{n_y}\otimes I_{n_y} - M\otimes P]^{-1}\}\end{aligned}$$

It will be seen from the proof that the correction term is the one derived in the situation where ξ were known, see Johansen (1999), apart from a term equal to $\frac{n_a(n-n_v)}{T}v(\xi)$. The proof of Theorem 5 is given in the Appendix.

Note that the coefficients v, c, and c_d depend on the choice of ξ. If $\xi = \alpha$, then

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

with $\Sigma_{\beta\beta} = \mathrm{Var}(\beta' X_t|\Delta X_t, \dots, \Delta X_{t-k+2})$. If, however, $\xi = \alpha_2$ then

$$v(\alpha_2) = tr\{(\alpha_2'\Omega^{-1}\alpha_2)^{-1}\Sigma_{\beta_2\beta_2.\beta_1}^{-1}\},$$

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

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

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

If d_t contains seasonal dummies then $tr\{M^h\}$ is a periodic function and a more complicated expression can be found. In order to understand the parameter function $v(\xi)$ that enter the expressions, note that the long-run variance of Y_t conditional on the common trends is given by $\theta\theta'$. Thus the matrix V_ξ measures the "ratio" between the unconditional variance of Y_t and the conditional long-run variance.

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

Corollary 6 *The Bartlett correction for the test of a simple hypothesis* $\beta = \beta^0,\ \rho = \rho^0$ *in model* $\mathcal{M}_0$*, is given by*

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

where $v(\alpha)$, $c(\alpha)$, *and* $c_d(\alpha)$ *are given in Theorem 5.*

Proof. This follows from Theorem 5 by substituting $n_v = r$, $n_a = n - r + n_D$, $n_z = (k-1)n$, $\xi = \alpha$, see Table 1. ■

Corollary 7 *The Bartlett correction factor for the test of* $\mathcal{M}_1 : \beta = H\phi$, *with* $H\,(n \times s)$ *is given by*

$$\begin{aligned} & E[-2\log LR(\mathcal{M}_1|\mathcal{M}_0)]/r(n-s) \\ & \overset{1}{=} 1 + \tfrac{1}{T}[\tfrac{1}{2}(n+s-r+1+2n_D) + n_d + kn] \\ & + \tfrac{1}{Tr}[(2n+s-3r-1+2n_D)v(\alpha) + 2(c(\alpha)+c_d(\alpha))]. \end{aligned}$$

Proof. From Corollary 6 we use the result for a simple hypothesis on β and ρ in the unrestricted cointegrating model $\mathcal{M}_0$. We apply Theorem 5 to a simple hypothesis on β and ρ in $\mathcal{M}_1$. The dimensions are given by $n_v = r$, $n_a = s - r + n_D$, $n_z = (k-1)n$, $\xi = \alpha$.

$$\begin{aligned} & E[-2\log LR(\beta = \beta^0,\ \rho = \rho^0|\mathcal{M}_1)] \\ & \overset{1}{=} r(s-r+n_D) + \tfrac{r(s-r+n_D)}{T}[\tfrac{1}{2}(s+n_D+1) + n_d + kn] \\ & + \tfrac{(s-r+n_D)}{T}[(n+s-2r-1+n_D)v(\alpha) + 2(c(\alpha)+c_d(\alpha))]. \end{aligned}$$

Note that since V_{t-1} and Z_{t-1} have the same definitions in both cases, the matrix Σ and P have not changed, and that $\xi = \alpha$ has the same meaning in both models. Thus the coefficients $v(\alpha)$, $c_d(\alpha)$, and $c_d(\alpha)$ are the same as in Corollary 6. Subtracting the expressions we find the required result. ■

Corollary 8 *The Bartlett correction factor for the test of* $\mathcal{M}_2$:

$$\begin{pmatrix} \beta \\ \rho \end{pmatrix} = \left(\begin{pmatrix} \beta_1^0 \\ \rho_2^0 \end{pmatrix} \psi_1, \begin{pmatrix} \psi_2 \\ \rho_2 \end{pmatrix} \right),$$

where the matrix $\beta_1^0\ (n \times r_1)$ *is of rank* r_1, *and* ρ_2^0 *is* $(1 \times r_1)$, *is given by*

$$\begin{aligned} & E[-2\log LR(\mathcal{M}_2|\mathcal{M}_0)]/r_1(n-r+n_D) \\ & \overset{1}{=} 1 + \tfrac{1}{T}[\tfrac{1}{2}(n+n_D+1-r_2) + n_d + kn] \\ & + \tfrac{1}{Tr_1}[(2n-2r+n_D-1)(v(\alpha)-v(\alpha_2)) - r_1 v(\alpha_2) + 2(c(\alpha)-c(\alpha_2))], \end{aligned}$$

where the coefficients $c(.)$, $c_d(.)$, *and* $v(.)$ *are defined in Theorem 5.*

Proof. From Corollary 6 we use the result for a simple hypothesis on β and ρ in the unrestricted cointegration model $\mathcal{M}_0$. We apply Theorem 5 to find the result for a simple hypothesis on β and ρ in model $\mathcal{M}_2$. Note that the dimensions have changed as have the definitions of V_{t-1} and Z_{t-1} and that $\xi = \alpha_2$. In both cases the stacked vector (V_{t-1}, Z_{t-1}) is the same and hence the matrix P and the variance matrix Σ has the same meaning in both expressions. We apply Theorem 5 and find for $n_v = r_2$, $n_a = n - r + n_D$, $n_z = r_1 + (k-1)n$, $\xi = \alpha_2$.

$$
\begin{aligned}
&E[-2\log LR(\beta = \beta^0, \rho = \rho^0 | \mathcal{M}_2)] \\
&\overset{1}{=} r_2(n - r + n_D) + \tfrac{r_2(n-r+n_D)}{T}[\tfrac{1}{2}(n - r_1 + 1 + n_D) + r_1 + n_d + kn] \\
&+ \tfrac{(n-r+n_D)}{T}[(2n - r_2 - r - 1 + n_D)v(\alpha_2) + 2(c(\alpha_2) + c_d(\alpha_2))]
\end{aligned}
$$

Subtracting we find the result. ■

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, since we can get complete information on how the correction works. Then we present a few results where the DGP has been chosen so as to match the results obtained for real data, analysed elsewhere.

5.1 The model with 1 cointegrating vector and lag 1

We first consider the model with only one lag, one cointegration relation and no deterministic terms, that is, the model

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

In this situation we have $Z_t = 0$, and

$$\beta' X_t = \sum_{i=0}^{\infty} (1 + \beta'\alpha)^i \beta' \varepsilon_{t-i},$$

such that

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

If we want to test a simple hypothesis on β we find the coefficients

$$
\begin{aligned}
v(\alpha) &= V_\alpha = \Sigma^{-1}\kappa_\alpha^2 = -\frac{\beta'\alpha(2+\beta'\alpha)}{\alpha'\Omega^{-1}\alpha\beta'\Omega\beta} \\
P &= 1+\beta'\alpha,\ c(\alpha) = -2\frac{\beta'\alpha(1+\beta'\alpha)}{\alpha'\Omega^{-1}\alpha\beta'\Omega\beta}.
\end{aligned}
$$

With this notation we find from Corollary 7

Corollary 9 *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*

$$
\begin{aligned}
&E[-2\log LR(\beta = \beta^0|\mathcal{M}_0)]/(n-1) \\
&\stackrel{1}{=} 1 + \tfrac{1}{2T}(3n+1) - \tfrac{1}{T}\tfrac{\beta'\alpha}{\alpha'\Omega^{-1}\alpha\beta'\Omega\beta}[(2n-3)(2+\beta'\alpha)+4(1+\beta'\alpha)].
\end{aligned}
$$

In order to simplify the simulations we transform the problem linearly, by defining $v_1 = \beta(\beta'\Omega\beta)^{-\frac{1}{2}}$,

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

and finally vectors $v_3, \ldots, v_n$ such that

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

The new variables $\tilde{X}_t = v' X_t$ satisfy the equations

$$
\begin{aligned}
&\Delta\tilde{X}_{1t} = (\beta'\Omega\beta)^{-\frac{1}{2}}\beta'\alpha(\beta'\Omega\beta)^{\frac{1}{2}}_{1t-1}\tilde{X} + \delta_{1t} = \eta\tilde{X}_{1t-1} + \delta_{1t} \\
&\Delta\tilde{X}_{2t} = -\left(\alpha'\Omega^{-1}\alpha - \alpha'\beta(\beta'\Omega\beta)^{-1}\beta'\alpha\right)^{\frac{1}{2}}(\beta'\Omega\beta)^{\frac{1}{2}}_{1t-1}\tilde{X} + \delta_{2t} = \xi\tilde{X}_{1t-1} + \delta_{2t} \\
&\Delta\tilde{X}_{it} = \delta_{it}, i = 3, \ldots, n.
\end{aligned}
$$

where $\delta_t = v'\varepsilon_t$ are i.i.d. $N_n(0, I_n)$, and

$$
\begin{aligned}
\eta &= (\beta'\Omega\beta)^{-\frac{1}{2}}\beta'\alpha(\beta'\Omega\beta)^{\frac{1}{2}} \\
\xi &= -\left(\alpha'\Omega^{-1}\alpha - \alpha'\beta(\beta'\Omega\beta)^{-1}\beta'\alpha\right)^{\frac{1}{2}}(\beta'\Omega\beta)^{\frac{1}{2}}
\end{aligned}
$$

Thus only two parameters enter the DGP and it is possible for a given n to tabulate the effect of the Bartlett correction as a function of just two variables ξ and η, see Tables 2-4. In this formulation β is a unit vector and $\alpha' = (\eta, \xi, 0, 0, 0)$. We find the coefficients

$$
\begin{aligned}
v(\alpha) &= -\tfrac{\eta(2+\eta)}{\eta^2+\xi^2}, \\
c(\alpha) &= -2\tfrac{\eta(1+\eta)}{\eta^2+\xi^2}.
\end{aligned}
$$

$\xi \backslash \eta$	-0.1	-0.2	-0.4	-0.6	-0.8	-1.0
0.0	$\frac{1.3}{16.5}$	$\frac{3.7}{12.2}$	$\frac{4.7}{8.6}$	$\frac{5.0}{7.3}$	$\frac{5.2}{6.8}$	$\frac{5.0}{6.2}$
-0.1	$\frac{3.8}{13.5}$	$\frac{4.2}{11.0}$	$\frac{4.8}{8.7}$	$\frac{4.9}{7.6}$	$\frac{4.9}{7.0}$	$\frac{5.0}{6.5}$
-0.2	$\frac{5.4}{10.8}$	$\frac{4.9}{9.9}$	$\frac{4.9}{8.6}$	$\frac{5.2}{7.7}$	$\frac{5.1}{7.0}$	$\frac{5.1}{6.4}$
-0.4	$\frac{6.1}{8.2}$	$\frac{5.6}{8.3}$	$\frac{5.4}{8.0}$	$\frac{5.3}{7.6}$	$\frac{5.2}{7.1}$	$\frac{5.2}{6.5}$
-0.6	$\frac{5.8}{7.3}$	$\frac{5.5}{7.6}$	$\frac{5.8}{7.6}$	$\frac{5.4}{7.3}$	$\frac{5.5}{7.0}$	$\frac{5.4}{6.7}$
-0.8	$\frac{5.6}{6.9}$	$\frac{5.5}{7.2}$	$\frac{5.8}{7.4}$	$\frac{5.5}{7.3}$	$\frac{5.4}{6.9}$	$\frac{5.5}{6.7}$
-1.0	$\frac{5.6}{6.7}$	$\frac{5.5}{6.9}$	$\frac{5.9}{7.0}$	$\frac{5.7}{7.0}$	$\frac{5.6}{6.9}$	$\frac{5.4}{6.6}$

Table 2: Simulation of $T = 50$ observations from an AR(1) process in 2 dimensions with $r = 1$ cointegrating relations. The number of simulations is 10.000. The table gives the corrected p-value over the uncorrected p-value for a nominal 5% test. The simulation standard error is 0.2%

We then find some results in Table 2 ($T = 50$, $n = 2$), Table 3 ($T = 50$, $n = 5$) and Table 4 ($T = 100$, $n = 5$). 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 $\eta + \xi < -0.2$) the Bartlett correction factor gives a useful correction.

Note that for $\xi = 0$, both coefficients have a factor η^{-1}, such that for small η, the correction factor tends to infinity. The DGP where both ξ 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 $\xi \neq 0$, corresponds to a DGP generating an $I(2)$ process, and the derivation of the correction factor is not valid in this case.

For $n = 5$, it appears from Table 3, that the situation is worse and the actual size can be vary large indeed. The region where the Bartlett correction is useful is approximately given by $\eta + \xi < -0.4$. Obviously the situation improves if T is 100, see Table 4.

Usually the test for β is proceeded by a test for the rank, and if η and ξ are sufficiently small the hypothesis of 1 cointegrating relation will not be accepted, thus for small values of ξ and η the Bartlett correction is not needed.

$\xi \backslash \eta$	-0.1	-0.2	-0.4	-0.6	-0.8
0.0	$\frac{0.02}{78.56}$	$\frac{1.79}{64.64}$	$\frac{6.84}{40.34}$	$\frac{6.67}{25.08}$	$\frac{6.20}{17.60}$
-0.1	$\frac{1.68}{67.93}$	$\frac{3.69}{59.18}$	$\frac{7.02}{38.51}$	$\frac{6.87}{24.58}$	$\frac{6.15}{17.52}$
-0.2	$\frac{9.47}{44.93}$	$\frac{7.43}{46.79}$	$\frac{7.48}{34.10}$	$\frac{6.71}{23.03}$	$\frac{6.12}{17.15}$
-0.4	$\frac{8.30}{19.98}$	$\frac{7.80}{24.09}$	$\frac{7.00}{23.55}$	$\frac{6.21}{19.21}$	$\frac{6.02}{15.48}$
-0.6	$\frac{6.68}{13.40}$	$\frac{6.55}{15.87}$	$\frac{6.06}{17.13}$	$\frac{5.97}{15.91}$	$\frac{5.72}{13.80}$
-0.8	$\frac{6.11}{11.50}$	$\frac{6.01}{13.16}$	$\frac{5.82}{13.94}$	$\frac{5.78}{13.76}$	$\frac{5.59}{12.80}$

Table 3: Simulation of $T = 50$ observations from an AR(1) process in 5 dimensions with $r = 1$ cointegrating relations. The number of simulations is 10.000. The table gives the corrected p-value over the uncorrected p-value for a nominal 5% test. The simulation standard error is 0.2%

$\xi \backslash \eta$	-0.1	-0.2	-0.4	-0.6	-0.8
0	$\frac{1.54}{62.91}$	$\frac{5.96}{39.42}$	$\frac{5.84}{18.45}$	$\frac{5.56}{12.18}$	$\frac{5.50}{9.67}$
-0.1	$\frac{6.79}{44.60}$	$\frac{6.37}{33.02}$	$\frac{5.84}{17.41}$	$\frac{5.30}{12.00}$	$\frac{5.17}{9.66}$
-0.2	$\frac{7.72}{21.53}$	$\frac{6.65}{22.10}$	$\frac{5.65}{15.59}$	$\frac{5.38}{11.57}$	$\frac{5.10}{9.67}$
-0.4	$\frac{5.82}{10.25}$	$\frac{5.74}{11.39}$	$\frac{5.24}{11.21}$	$\frac{5.23}{9.94}$	$\frac{5.18}{9.19}$
-0.6	$\frac{5.41}{8.33}$	$\frac{5.43}{9.01}$	$\frac{5.32}{9.26}$	$\frac{5.11}{8.91}$	$\frac{5.10}{8.36}$
-0.8	$\frac{5.28}{7.54}$	$\frac{8.24}{5.35}$	$\frac{5.09}{8.45}$	$\frac{5.21}{7.97}$	$\frac{5.04}{7.67}$

Table 4: Simulation of $T = 100$ observations from an AR(1) process in 5 dimensions with $r = 1$ cointegrating relations. The number of simulations is 10.000. The table gives the corrected p-value over the uncorrected p-value for a nominal 5% test. The simulation standard error is 0.2%

5.2 Some real life examples

As a perhaps more interesting case consider the data set discussed in Johansen (1996) of a four variable system consisting of m_t (log real M2), y_t (log real income), i_t^b(bond rate), and finally i_t^d (deposit rate) observed quarterly from 1974:1 to 1987:3. We take as DGP the parameters determined by the estimation, and simulate 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.

The Bartlett factor in this case is given by Corollary 7, since we have a hypothesis only on β, which we formulate as

$$\beta = H\phi = \beta^0\phi,$$

with $\phi\,(1,1)$. We find with $n = 4$, $r = 1$, $s = 1$, $n_D = 1$, $n_d = 0$, $k = 2$, such the degrees of freedom is are $r(n - s) = 3$, and

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

We find for a test of nominal size 5% a simulated p-value of 10.3% (10000 observations) and a corrected p-value of 3.1%.

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

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

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 7 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 has an actual size of 9.9% whereas the size for the corrected test is 3.1%.

As another example consider the Australian data consisting of consumer price indices (in logarithms) for Australia p_t^{au} and US p_t^{us} and the exchange rate $exch_t$ together with the five year treasure bond rate for both countries i_t^{au} and i_t^{us}. The data is observed quarterly from 1972:1 to 1991:1, which gives an effective number of observation 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 are $r(n - s) = 6$. Since $d_t = 1$, we find that

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

The Bartlett factor can be found from Corollary 7 and is given by

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

We found that a nominal 5% test in reality corresponds to a test size of 21%. The correction of the test gives a size of 6.3%. The result is based upon 10.000 simulations.

Next consider the test for price homogeneity given by the restriction

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

and $H = R_\perp$. In this example $s = 4$, such that the degrees of freedom are $r(n - s) = 2$. We find the Bartlett correction from Corollary 7 as given by

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

By performing 10.000 simulations we see that a nominal 5% test correspond to a test of size 10.5%, and that the Bartlett correction gives the size as 3.37%.

6 Conclusion

In this paper we have derived an approximation of the log likelihood ratio statistic for various hypotheses on the cointegrating coefficients in a VAR model. Despite the rather tedious calculations it turns out that the final result depends on the obvious quantities like, dimension, lag length, cointegrating rank, number of restricted deterministic terms, and number of unrestricted terms, as well as on the hypothesis, that we want to test. The effect of the parameters is focussed in two or sometimes three functions, which can be easily calculated once the parameters of the model has been estimated.

The usefulness of the results is demonstrated by some simulation experiments. Table 2 and 3 give the results for all models with one lag, one cointegrating vector and no deterministic terms in case $n = 2$, and $T = 50$, and Table 4 for $n = 5$ and $T = 100$.

7 References

Ahn, S. K. and 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. *Proceeding of the Royal Society of London* Series A, Vol. 160, 268-282.

Cribari-Neto, F. and Cordeiro, G. (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. and Jacobson, T. (1998) Bootstrap testing and approximate finite sample distributions for test of linear restrictions on cointegrating vectors. Discussion paper, Stockholm School of Economics.

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. Vredin, A., and Warne, A. (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. Discussion paper, EUI.

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

8 Appendix

8.1 Some technical results

Lemma 10 *Under the assumptions (2) and (3) the powers M^h grow at most as a polynomial in h.*

Proof. The Jordan form of the matrix M contains blocks of, for instance, the form

$$J_3(\lambda) = \begin{pmatrix} \lambda & 0 & 0 \\ 1 & \lambda & 0 \\ 0 & 1 & \lambda \end{pmatrix},$$

where $|\lambda| = 1$. This has the property that

$$J_3^h(\lambda) = \begin{pmatrix} \lambda^h & 0 & 0 \\ h\lambda^{h-1} & \lambda^h & 0 \\ \frac{1}{2}h(h-1)\lambda^{h-2} & h\lambda^{h-1} & \lambda^h \end{pmatrix}$$

This is bounded by a polynomial of degree 2 in h. In a similar way one can prove 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. ■

Lemma 11 *Let $S_t = \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i}$ with θ_i decreasing exponentially. Let*

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

Let d_t satisfy $d_{t+1} = Md_t$, with $|eig(M)| = 1$, then

$$tr\{E(M_{sd}M_{dd}^{-1}M_{ds})\} \to \sum_{h=-\infty}^{\infty} tr\{M^h\}tr\{\gamma(h)\}. \tag{39}$$

Proof.

$$\begin{aligned} & tr\{E(M_{sd}M_{dd}^{-1}M_{ds})\} \\ = \; & tr\{E\textstyle\sum_{s,t=1}^{T}\sum_{i,j=0}^{\infty}\theta_i\varepsilon_{t-i}d'_tM_{dd}^{-1}d_s\varepsilon'_{s-j}\theta'_j\} \\ = \; & \textstyle\sum_{i,j,t} tr\{\theta_i\Omega\theta'_j\}tr\{d'_tM_{dd}^{-1}d_{t-i+j}\} \\ = \; & \textstyle\sum_{i,j}\sum_{t=1}^{T} tr\{d'_tM_{dd}^{-1}d_tM^{-i+j}\}tr\{\Omega\theta'_j\theta_i\} \\ \to \; & \textstyle\sum_{h=-\infty}^{\infty} tr\{M^h\}tr\{\gamma(h)\}. \end{aligned}$$

■

We next give an expansion of a projection matrix which will be used in the detailed calculations below. Recall from Theorem 3 that $\hat{P}(\hat{\xi},\hat{\Omega}) = \hat{\xi}(\hat{\xi}'\hat{\Omega}^{-1}\hat{\xi})^{-1}\hat{\xi}'\hat{\Omega}^{-1}$, and 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 12

$$\begin{aligned} P(\hat{\xi},\hat{\Omega}) = \; & P(\xi,\hat{\Omega}) + \bar{P}(\xi,\hat{\Omega})(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1} \\ & +\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\bar{P}(\xi,\hat{\Omega}) \\ & +\bar{P}(\xi,\hat{\Omega})(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\bar{P}(\xi,\hat{\Omega} \\ & -\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\bar{P}(\xi,\hat{\Omega})(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1} \\ & -\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\bar{P}(\xi,\hat{\Omega}) \\ & -\bar{P}(\xi,\hat{\Omega})(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1} \\ & +O_P(T^{-\frac{3}{2}}). \end{aligned}$$

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

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

say, and

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

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

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

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

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

When substituting u and v we find the result. ■

8.2 Proof of Theorem 4

We start with (30)

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

and evaluate each term starting with the easy ones.

8.2.1 Calculation of $tr\{Q_1\}$

We find from Theorem 3, that

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

Because $tr\{Q_1\}$ is multiplied by T^{-1} we need only retain the main term in each of the matrices. Thus we can replace $P(\xi, \hat{\Omega})$ by $P = P(\xi, \Omega)$, $\tilde{\delta}$ by $S_{aa}^{-1}S_{au}\kappa_\xi$, $\hat{\xi}\tilde{\delta}'$ by $PS_{\varepsilon a}S_{aa}^{-1}$, $S_{\varepsilon a.v}$ by $S_{\varepsilon a}$, and finally $T^{-1}S_{vv} = T^{-1}M_{vv.z,d}$ by $\Sigma_{vv.z} = \mathrm{Var}(V_t|Z_t)$. We find

$$\begin{aligned}
tr\{Q_1\} &\overset{0}{=} tr\{\Sigma_{vv.z}^{-1}[\kappa_\xi S_{ua}S_{aa}^{-1}S_{a\varepsilon}(I_n - P)' - S_{va}S_{aa}^{-1}S_{a\varepsilon}P']\Omega^{-1} \\
&\qquad \times[(I_n - P)S_{\varepsilon a}S_{aa}^{-1}S_{au}\kappa_\xi - PS_{\varepsilon a}S_{aa}^{-1}S_{av}]\} \\
&\overset{0}{=} tr\{\Sigma_{vv.z}^{-1}\kappa_\xi S_{ua}S_{aa}^{-1}S_{a\varepsilon}\Omega^{-1}(I_n - P)S_{\varepsilon a}S_{aa}^{-1}S_{au}\kappa_\xi\} \\
&\qquad + tr\{\Sigma_{vv.z}^{-1}S_{va}S_{aa}^{-1}S_{a\varepsilon}\Omega^{-1}PS_{\varepsilon a}S_{aa}^{-1}S_{av}\} \\
&= tr\{\kappa_\xi \Sigma_{vv.z}^{-1}\kappa_\xi S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}S_{au}\} \\
&\qquad + tr\{\Sigma_{vv.z}^{-1}S_{va}S_{aa}^{-1}S_{au}S_{ua}S_{aa}^{-1}S_{av}\},
\end{aligned} \tag{40}$$

where we have used the properties of projections

$$\begin{aligned}
P'\Omega^{-1}(I_n - P) &= 0,\ P'\Omega^{-1}P = P'\Omega^{-1} \\
(I_n - P)'\Omega^{-1}(I_n - P) &= \Omega^{-1}(I_n - P) = \xi_\perp(\xi_\perp'\Omega\xi_\perp)^{-1}\xi_\perp' \\
S_{a\varepsilon}\Omega^{-1}PS_{\varepsilon a} &= S_{au}S_{ua},\ S_{a\varepsilon}\Omega^{-1}(I_n - P)S_{\varepsilon a} = S_{ab}S_{ba}.
\end{aligned}$$

8.2.2 Calculation of $tr\{Q_0^2\} = T^{-1}tr\{(\hat{\Omega}^{-1}[S_{\varepsilon a.v}\tilde{\delta}\hat{\xi}' + \hat{\xi}\tilde{\delta}'S_{a\varepsilon.v} - \hat{\xi}\tilde{\delta}'S_{aa}\tilde{\delta}\hat{\xi}'])^2\}$

The factor T^{-1} allows us to replace each matrix with its limit. We replace $\hat{\xi}\tilde{\delta}'$ by $PS_{\varepsilon a}S_{aa}^{-1}$ and $S_{\varepsilon a.v}$ by $S_{\varepsilon a}$, and find

$$Q_0 \overset{0}{=} \Omega^{-1}[S_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}P' + PS_{\varepsilon a}S_{aa}^{-1}S_{a\varepsilon}(I_n - P')].$$

Hence

$$\begin{array}{rcl}\frac{1}{2}tr\{Q_0^2\} & \overset{0}{=} & \frac{1}{2}tr\{[S_{a\varepsilon}\Omega^{-1}S_{\varepsilon a}]S_{aa}^{-1}[S_{a\varepsilon}\Omega^{-1}PS_{\varepsilon a}]S_{aa}^{-1}\} \\ & + & \frac{1}{2}tr\{[S_{a\varepsilon}\Omega^{-1}PS_{\varepsilon a}]S_{aa}^{-1}[S_{a\varepsilon}\Omega^{-1}(I_n-P)S_{\varepsilon a}]S_{aa}^{-1}\} \\ & = & \frac{1}{2}tr\{(S_{ua}S_{aa}^{-1}S_{au})^2\}+tr\{S_{au}S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}\},\end{array} \tag{41}$$

using

$$S_{a\varepsilon}\Omega^{-1}S_{\varepsilon a}=S_{a\varepsilon}\Omega^{-1}PS_{\varepsilon a}+S_{a\varepsilon}\Omega^{-1}(I_n-P)S_{\varepsilon a}=S_{au}S_{ua}+S_{ab}S_{ba}, \tag{42}$$

where

$$U_t=(\xi'\Omega^{-1}\xi)^{-\frac{1}{2}}\xi'\Omega^{-1}\varepsilon_t, b_t=(\xi_\perp'\Omega\xi_\perp)^{-\frac{1}{2}}\xi_\perp'\varepsilon_t$$

8.2.3 The main term $tr\{Q_0\}$

This term is of the order of $T^{-\frac{1}{2}}$, and hence we have to keep more terms in the expansions.

From (28) we find that $\hat{\xi}\tilde{\delta}'$ is of the order of $S_{aa}^{-\frac{1}{2}}$ and that

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

for some $\tilde{\delta}_1\in O_P(S_{aa}^{-\frac{1}{2}})$, 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 we can replace $\hat{\xi}\tilde{\delta}'$ by $\hat{P}S_{\varepsilon a.v}S_{aa}^{-1}$ introducing errors of at most the order $o_P(T^{-1})$ 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

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

This term, however, does not give a contribution since

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

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

$$\begin{aligned} \hat{\Omega}Q_0 &= S_{\varepsilon a.v}\tilde{\delta}\hat{\xi}' + \hat{\xi}\tilde{\delta}' S_{a\varepsilon.v} - \hat{\xi}\tilde{\delta}' S_{aa}\tilde{\delta}\hat{\xi}' \\ &\overset{1}{=} S_{\varepsilon a.v}S_{aa}^{-1}S_{a\varepsilon.v}\hat{P}' + \hat{P}S_{\varepsilon a.v}S_{aa}^{-1}S_{a\varepsilon.v} - \hat{P}S_{\varepsilon a.v}S_{aa}^{-1}S_{a\varepsilon.v}\hat{P}'. \end{aligned} \tag{43}$$

Hence

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

where we use the property

$$\hat{P}'\hat{\Omega}^{-1}(I_n - \hat{P}) = 0.$$

We next expand around ξ, but keep $\hat{\Omega}$. We find from Lemma **??**, using $\bar{P}(\xi, \hat{\Omega}) = \xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp'$,

$$\begin{aligned} & tr\{Q_0\} = tr\{S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}S_{\varepsilon a.v}\} \\ & +2tr\{S_{aa}^{-1}S_{a\varepsilon.v}\xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp'(\hat{\xi} - \xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}S_{\varepsilon a.v}\} \\ & +tr\{S_{aa}^{-1}S_{a\varepsilon.v}\xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp'(\hat{\xi} - \xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi} - \xi)'\xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp' S_{\varepsilon a.v}\} \\ & -tr\{S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi} - \xi)'\xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp'(\hat{\xi} - \xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}S_{\varepsilon a.v}\} \\ & -2tr\{S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi} - \xi)'\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi} - \xi)'\xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp' S_{\varepsilon a.v}\} \\ & = A_1 + A_2 + A_3 + A_4 + A_5. \end{aligned}$$

In order to simplify these expressions we need the identities

$$\begin{aligned} & S_{a\varepsilon.v}\hat{\Omega}^{-1}\xi = T(S_{au.v}, S_{ab.v}) \begin{pmatrix} S_{uu.v} & S_{ub.v} \\ S_{ub.v} & S_{bb.v} \end{pmatrix}^{-1} \begin{pmatrix} (\xi'\Omega^{-1}\xi)^{\frac{1}{2}} \\ 0 \end{pmatrix} \\ & = TS_{au.v,b}S_{uu.v,b}^{-1}(\xi'\Omega^{-1}\xi)^{\frac{1}{2}}, \end{aligned} \tag{44}$$

$$S_{a\varepsilon.v}\xi_\perp = S_{ab.v}(\xi_\perp'\Omega\xi_\perp)^{\frac{1}{2}}, \tag{45}$$

$$\begin{aligned} & (\xi'\Omega^{-1}\xi)^{\frac{1}{2}}(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\xi'\Omega^{-1}\xi)^{\frac{1}{2}} \\ & = T^{-1}\left(\begin{pmatrix} I \\ 0 \end{pmatrix}' \begin{pmatrix} S_{uu.v} & S_{ub.v} \\ S_{ub.v} & S_{bb.v} \end{pmatrix}^{-1} \begin{pmatrix} I \\ 0 \end{pmatrix}\right)^{-1} = T^{-1}S_{uu.v,b}, \end{aligned} \tag{46}$$

$$(\xi_\perp'\Omega\xi_\perp)^{\frac{1}{2}}(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}(\xi_\perp'\Omega\xi_\perp)^{\frac{1}{2}}$$
$$= T\left(\begin{pmatrix} 0 \\ I \end{pmatrix}'\begin{pmatrix} S_{uu.v} & S_{ub.v} \\ S_{ub.v} & S_{bb.v} \end{pmatrix}\begin{pmatrix} 0 \\ I \end{pmatrix}\right)^{-1} = TS_{bb.v}^{-1}. \tag{47}$$

We then take the terms one by one

The term $\mathbf{A}_1$ We apply first (44) and then (46) and find

$$\begin{aligned} A_1 &= T^2 tr\{S_{aa}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}(\xi'\Omega^{-1}\xi)^{\frac{1}{2}}(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\xi'\Omega^{-1}\xi)^{\frac{1}{2}}S_{uu.b,v}^{-1}S_{ua.b,v}\} \\ &= Ttr\{S_{aa}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\}. \end{aligned}$$

Now we want to modify this as follows

$$Ttr\{S_{aa}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\}$$
$$= Ttr\{S_{aa.b,v}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} + Ttr\{(S_{aa}^{-1} - S_{aa.b,v}^{-1})S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\}.$$

Now

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

such that

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

and hence we find for A_1

$$\begin{aligned} A_1 = \ & Ttr\{S_{aa}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} \\ \stackrel{1}{=} \ & Ttr\{S_{aa.b,v}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} \\ & -Ttr\{S_{aa}^{-1}(S_{ab}S_{bb}^{-1}S_{ba} + S_{av.b}S_{vv.b}^{-1}S_{va.b})S_{aa}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} \\ \stackrel{1}{=} \ & Ttr\{S_{aa.b,v}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} \\ & -tr\{S_{aa}^{-1}(S_{ab}S_{ba} + S_{av.b}\Sigma_{vv.z}^{-1}S_{va.b})S_{aa}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} \\ \stackrel{1}{=} \ & Ttr\{S_{aa.b,v}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} - T^{-1}tr\{S_{au}S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}\} \\ & -T^{-1}tr\{\Sigma_{vv.z}^{-1}S_{va}S_{aa}^{-1}S_{au}S_{ua}S_{aa}^{-1}S_{av}\}. \end{aligned} \tag{48}$$

The term $\mathbf{A}_2$ To simplify the term A_2 we use (45) and then (46) and (47) to find

$$\begin{aligned} A_2 &= 2tr\{S_{aa}^{-1}S_{a\varepsilon.v}\xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp'(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}S_{\varepsilon a.v}\} \\ &= 2Ttr\{S_{aa}^{-1}S_{ab.v}(\xi_\perp'\Omega\xi_\perp)^{\frac{1}{2}}(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}(\xi_\perp'\Omega\xi_\perp)^{\frac{1}{2}} \\ &\quad \times S_{bv}S_{vv}^{-1}(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\xi'\Omega^{-1}\xi)^{\frac{1}{2}}S_{uu.v,b}^{-1}S_{ua.v,b}\} \\ &= 2Ttr\{S_{aa}^{-1}S_{ab.v}S_{bb.v}^{-1}S_{bv}S_{vv}^{-1}\kappa_\xi S_{ua.v,b}\}. \end{aligned}$$

This term is of the order of $T^{-\frac{1}{2}}$, and since

$$T^{-1}S_{bb.v} = T^{-1}S_{bb} - T^{-1}S_{bv}S_{vv}^{-1}S_{vb} = T^{-1}S_{bb} + O_P(T^{-1}),$$

we find that we can replace $T^{-1}S_{bb.v}$ by $T^{-1}S_{bb}$ and get

$$A_2 \overset{1}{=} 2Ttr\{S_{aa}^{-1}S_{ab.v}S_{bb}^{-1}S_{bv}S_{vv}^{-1}\kappa_\xi S_{ua.v,b}\} \tag{49}$$

8.2.4 The term $\mathbf{A}_3$

We find since the term is of the order of T^{-1} that we can replace each matrix with its limit to simplify the expression

$$\begin{aligned} A_3 = &\; tr\{S_{aa}^{-1}S_{a\varepsilon.v}\xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp'(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1} \\ &\quad \times(\hat{\xi}-\xi)'\xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1}\xi_\perp'S_{\varepsilon a.v}\} \\ \overset{1}{=} &\; tr\{S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_\xi^2 S_{vv}^{-1}S_{vb}S_{ba.v}\} \\ \overset{1}{=} &\; tr\{S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_\xi^2 S_{vv}^{-1}S_{vb}S_{ba}\} \end{aligned} \tag{50}$$

where we have used

$$T^{-1}S_{ba.v} = T^{-1}S_{ba} + O_P(T^{-\frac{1}{2}}).$$

8.2.5 The term $\mathbf{A}_4$

Again we can replace each matrix with its limit to simplify the expression

$$\begin{aligned} A_4 = &\; -tr\{S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\xi_\perp(\xi_\perp'\hat{\Omega}\xi_\perp)^{-1} \\ &\quad \times\xi_\perp'(\hat{\xi}-\xi)(\xi'\hat{\Omega}^{-1}\xi)^{-1}\xi'\hat{\Omega}^{-1}S_{\varepsilon a.v}\} \\ \overset{1}{=} &\; -tr\{S_{aa}^{-1}S_{au}\kappa_\xi S_{vv}^{-1}S_{vb}S_{bv}S_{vv}^{-1}\kappa_\xi S_{ua}\} \end{aligned} \tag{51}$$

8.2.6 The term $\mathbf{A}_5$

Using the same relations as before we find

$$\begin{aligned} A_5 = & -2tr\{S_{aa}^{-1}S_{a\varepsilon.v}\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1}(\hat{\xi}-\xi)'\hat{\Omega}^{-1}\xi(\xi'\hat{\Omega}^{-1}\xi)^{-1} \\ & \times(\hat{\xi}-\xi)'\xi_{\perp}(\xi_{\perp}'\hat{\Omega}\xi_{\perp})^{-1}\xi_{\perp}'S_{\varepsilon a.v}\} \\ \overset{1}{=} & -2tr\{S_{aa}^{-1}S_{au}\kappa_{\xi}S_{vv}^{-1}S_{vu}\kappa_{\xi}S_{vv}^{-1}S_{vb}S_{ba}\} \end{aligned} \quad (52)$$

Inserting (40), (41), (48) - (52) into the expression for the likelihood ratio test (30) we find

$$\begin{aligned} -2\log LR = & Ttr\{S_{aa.b,v}^{-1}S_{au.b,v}S_{uu.b,v}^{-1}S_{ua.b,v}\} + \tfrac{1}{2T}tr\{(S_{ua}S_{aa}^{-1}S_{au})^2\} \\ & +2Ttr\{S_{aa}^{-1}S_{ua.v,b}\kappa_{\xi}S_{vv}^{-1}S_{vb}S_{bb}^{-1}S_{ba.v}\} \\ & +T^{-1}tr\{\kappa_{\xi}\Sigma_{vv.z}^{-1}\kappa_{\xi}S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}S_{au}\}, \\ & +tr\{S_{ba}S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_{\xi}^2S_{vv}^{-1}S_{vb}\} \\ & -tr\{S_{ua}S_{aa}^{-1}S_{au}\kappa_{\xi}S_{vv}^{-1}S_{vb}S_{bv}S_{vv}^{-1}\kappa_{\xi}\} \\ & -2tr\{S_{ba}S_{aa}^{-1}S_{au}\kappa_{\xi}S_{vv}^{-1}S_{vu}\kappa_{\xi}S_{vv}^{-1}S_{vb}\} \end{aligned} \quad (53)$$

Note that two terms from A_1, as given in (48), cancel a term in the expression for Q_1 (40) and Q_0^2 (41).

This completes the proof of the representation of the likelihood ratio test statistic given in Theorem 4.

8.3 Proof of Theorem 5

We use the result of (53) 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. Below we shall indicate by $E_{\xi_{\perp}}[\ldots]$ the expectation formation, 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. This also holds in the case of (14) where we condition on $\alpha_{2\perp}'\varepsilon_t$ rather than $\alpha_{\perp}'\varepsilon_t$.

8.3.1 The main terms $\mathbf{K}_1+\mathbf{T}^{-1}\mathbf{K}_2$

We have

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

This is the correction term given in (Johansen 1999, Theorem 3.3) based upon the regression equation

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

for the test that $\delta = 0$, when ξ is known.

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

where the coefficients $v(\xi), c(\xi), c_d(\xi)$ are given in Theorem 5.

The rest of the proof of Theorem 5 deals with the problem of evaluating

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

We first consider the terms K_4, K_5, K_6, and K_7. Since they are of the order of T^{-1}, such that we can replace each matrix by its limit.

8.3.2 The term $\mathbf{K}_4 = E_{\xi_\perp}[tr\{\kappa_\xi \Sigma_{vv.z}^{-1}\kappa_\xi S_{ua}S_{aa}^{-1}S_{ab}S_{ba}S_{aa}^{-1}S_{au}\}]$

We replace $S_{ua}S_{aa}^{-1}S_{ab} = M_{ua.z,d}M_{aa.z,d}^{-1}M_{ab.z,d}$ by $M_{ua}M_{aa}^{-1}M_{ab}$ and find

$$K_4 \overset{0}{=} E_{\xi_\perp}[tr\{\kappa_\xi \Sigma_{vv.z}^{-1}\kappa_\xi\, M_{ua}M_{aa}^{-1}M_{ab}M_{ba}M_{aa}^{-1}M_{au}\}].$$

Since M_{ua} is Gaussian $N_{n_v \times n_a}(0, I_{n_v} \otimes M_{aa})$ given $\alpha'_\perp\varepsilon_t$, we have

$$K_4 \overset{0}{=} tr\{\kappa_\xi \Sigma_{vv.z}^{-1}\kappa_\xi\}tr\{M_{ba}M_{aa}^{-1}M_{ab}\}. \tag{54}$$

8.3.3 The term $\mathbf{K}_5 = TE_{\xi_\perp}[tr\{S_{ba}S_{aa}^{-1}S_{ab}S_{bv}S_{vv}^{-1}\kappa_\xi^2 S_{vv}^{-1}S_{vb}\}]$

We replace $S_{ba}S_{aa}^{-1}S_{ab}$ by $M_{ba}M_{aa}^{-1}M_{ab}$. We next consider $\kappa_\xi S_{vv}^{-1}S_{vb}$. We use the identity,

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

Thus

$$\kappa_\xi S_{vv}^{-1} S_{vb} = \tilde{\kappa}_\xi' M_{yy.d}^{-1} M_{yb.d}.$$

Hence we replace $\kappa_\xi S_{vv}^{-1} S_{vb}$ by $T^{-1}\tilde{\kappa}_\xi \Sigma^{-1} M_{yb}$ since

$$M_{yb.d} = M_{yb} - M_{yd} M_{dd}^{-1} M_{db} = M_{yb} + O_P(1).$$

We have to find

$$\begin{aligned} K_5 = \; & T^{-1} E_{\xi_\perp}[tr\{M_{ba} M_{aa}^{-1} M_{ab} M_{by} \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} M_{yb}\}] \\ = \; & T^{-1} E_{\xi_\perp}[tr\{M_{ba} M_{aa}^{-1} M_{ab} \textstyle\sum_{s,t,i,j} b_t (U_{t-i-1}'\theta_i' + b_{t-i-1}'\psi_i') \\ & \quad \times \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} (\theta_j U_{s-j-1} + \psi_j b_{s-j-1}) b_s'\}]. \end{aligned}$$

We get a non zero mean if the number of stochastic factors is even. We find for two stochastic factors that for $t = s$ and $i = j$, we get

$$\begin{aligned} & T^{-1} E_{\xi_\perp}[tr\{M_{ba} M_{aa}^{-1} M_{ab} \textstyle\sum_{s,t,i,j} b_t U_{t-i-1}' \theta_i' \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} \theta_j U_{s-j-1} b_s'\}] \\ & \overset{0}{=} T^{-1} tr\{M_{ba} M_{aa}^{-1} M_{ab} \textstyle\sum_{t,i,j} b_t b_{t+j-i}'\} tr\{\theta_i' \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} \theta_j\} \\ & \overset{0}{=} tr\{M_{ba} M_{aa}^{-1} M_{ab}\} tr\{\textstyle\sum_i \theta_i' \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} \theta_i\} \\ & = tr\{M_{ba} M_{aa}^{-1} M_{ab}\} tr\{(\textstyle\sum_{i=0}^\infty \theta_i \theta_i') \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1}\}, \end{aligned} \tag{55}$$

since

$$T^{-1} \sum_{t=1}^T b_t b_{t+j-i} \overset{P}{\to} I_{n_b},$$

if $i = j$ and 0 otherwise, see (33) and (34). With no stochastic terms we find

$$T^{-1} tr\{M_{ba} M_{aa}^{-1} M_{ab} \sum_{s,t,i,j} b_t b_{t-i-1}' \psi_i' \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} \psi_j b_{s-j-1} b_s'\}.$$

For $t = s$ and $i = j$ we find

$$\begin{aligned} & T^{-1} tr\{M_{ba} M_{aa}^{-1} M_{ab} \textstyle\sum_{t,i} b_t b_{t-i-1}' \psi_i' \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} \psi_i b_{t-i-1} b_t'\} \\ & \overset{0}{=} T^{-1} tr\{M_{ba} M_{aa}^{-1} M_{ab} \textstyle\sum_t b_t b_t'\} tr\{\sum_i \psi_i' \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1} \psi_i\} \\ & \overset{0}{=} tr\{M_{ba} M_{aa}^{-1} M_{ab}\} tr\{(\textstyle\sum_{i=0}^\infty \psi_i \psi_i') \Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}_\xi' \Sigma^{-1}\}. \end{aligned} \tag{56}$$

Thus we find from (55) and (56)

$$K_5 \overset{0}{=} tr\{M_{ba} M_{aa}^{-1} M_{ab}\} tr\{\kappa_\xi \Sigma_{vv.z}^{-1} \kappa_\xi\}, \tag{57}$$

using

$$\sum_{i=0}^\infty (\psi_i \psi_i' + \theta_i \theta_i') = \Sigma.$$

8.3.4 The term $\mathbf{K}_6 = -TE_{\xi_\perp}[tr\{S_{ua}S_{aa}^{-1}S_{au}\kappa_\xi S_{vv}^{-1}S_{vb}S_{bv}S_{vv}^{-1}\kappa_\xi\}]$

We replace $S_{au} = M_{au.z,d}$ by M_{au}, and $S_{vv}^{-1}S_{vb} = (I_{n_y}, 0_{n_v \times n_z})M_{yy.d}^{-1}M_{yb.d}$ by $T^{-1}(I_{n_y}, 0_{n_v \times n_z})\Sigma^{-1}M_{yb}$ and find

$$
\begin{aligned}
K_6 &\overset{0}{=} -T^{-1}E_{\xi_\perp}[tr\{M_{ua}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb}M_{by}\Sigma^{-1}\tilde{\kappa}_\xi\} \\
&= -T^{-1}E_{\xi_\perp}[tr\{M_{aa}^{-1}\textstyle\sum_{s,t,m,r,i,j} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}(\theta_j U_{s-j-1} + \psi_j b_{s-j-1})b_s' \\
&\qquad \times b_m(U_{m-i-1}'\theta_i' + b_{m-i-1}'\psi_i')\Sigma^{-1}\tilde{\kappa}_\xi U_r a_{r-1}'\}].
\end{aligned}
$$

For four stochastic terms we get

$$
-T^{-1}E_{\xi_\perp}[tr\{M_{aa}^{-1}\sum_{s,t,r,m,i,j} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_j U_{s-j-1}b_s'b_m U_{m-i-1}'\theta_i'\Sigma^{-1}\tilde{\kappa}_\xi U_r a_{r-1}'\}].
$$

We find for $s-j-1 = t, m-i-1 = r$, the sums M_{ab} and M_{ba} which normalized by T^{-1} and M_{aa}^{-1} tend to zero.

Only for $t = r$, $s = m$, and $i = j$, we get something non zero

$$
\begin{aligned}
&-T^{-1}E_{\xi_\perp}[tr\{M_{aa}^{-1}\textstyle\sum_{s,t,i,j} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_j U_{s-j-1}b_s'b_{s-j+i}U_{s-j-1}'\theta_i'\Sigma^{-1}\tilde{\kappa}_\xi U_t a_{t-1}'\}] \\
&\overset{0}{=} -tr\{M_{aa}^{-1}\textstyle\sum_{t,i} a_{t-1}a_{t-1}'\}tr\{\tilde{\kappa}_\xi'\Sigma^{-1}\theta_i T^{-1}\sum_s b_s'b_s\theta_i'\Sigma^{-1}\tilde{\kappa}_\xi\} \\
&\overset{0}{=} -n_a n_b tr\{\tilde{\kappa}_\xi'\Sigma^{-1}(\textstyle\sum_{i=0}^\infty \theta_i\theta_i')\Sigma^{-1}\tilde{\kappa}_\xi\}.
\end{aligned}
$$

For two stochastic factors we find

$$
-T^{-1}E_{\xi_\perp}[tr\{M_{aa}^{-1}\sum_{s,t,m,r,i,j} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\psi_j b_{s-j-1}b_s'b_m b_{m-i-1}'\psi_i'\Sigma^{-1}\tilde{\kappa}_\xi U_r a_{r-1}'\}],
$$

and we take again $t = r,\ s = m,\ i = j$ and find

$$
\begin{aligned}
&\overset{0}{=} \quad -T^{-1}tr\{M_{aa}^{-1}\textstyle\sum_t a_{t-1}a_{t-1}'\}\sum_j tr\{\tilde{\kappa}_\xi'\Sigma^{-1}\psi_j\sum_s b_{s-j-1}b_s'b_s b_{s-j-1}'\psi_j'\Sigma^{-1}\tilde{\kappa}_\xi\} \\
&\overset{0}{=} \quad -n_a n_b tr\{\tilde{\kappa}_\xi'\Sigma^{-1}(\textstyle\sum_{j=0}^\infty \psi_j\psi_j')\Sigma^{-1}\tilde{\kappa}_\xi\},
\end{aligned}
$$

which together with the previous result gives

$$
K_6 \overset{0}{=} -n_a n_b tr\{\kappa_\xi \Sigma_{vv.z}^{-1}\kappa_\xi\}. \tag{58}
$$

8.3.5 The term $\mathbf{K}_7 = -2TE_{\xi_\perp}[tr\{S_{ba}S_{aa}^{-1}S_{au}\kappa_\xi S_{vv}^{-1}S_{vu}\kappa_\xi S_{vv}^{-1}S_{vb}\}]$

We find as before

$$\begin{aligned} K_7 = \quad & -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\textstyle\sum_{t,s,m,i,j} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}(\theta_j U_{s-j-1}+\psi_j b_{s-j-1})U_s' \\ & \times\tilde{\kappa}_\xi'\Sigma^{-1}(\theta_i U_{m-i-1}+\psi_i b_{m-i-1})b_m'\}]. \end{aligned}$$

For four stochastic terms we get

$$-2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{t,s,m,i,j} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_j U_{s-j-1}U_s'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_i U_{m-i-1}b_m'\}.$$

We notice that to get a contribution we need s, t, and m to be tied together which always gives something of the form $\sum_t a_{t-1}b_{t-k}'$ which normalized by $T^{-1}M_{ba}M_{aa}^{-1}$ goes to zero. For two stochastic terms we find with $t=s$,

$$\begin{aligned} & -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\textstyle\sum_{t,j,i,m} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\psi_j b_{t-j-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\psi_i b_{m-i-1}b_m'\}] \\ & \overset{0}{=} -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\textstyle\sum_{t,j,i,m} a_{t-1}b_{t-j-1}'\psi_j'\Sigma^{-1}\tilde{\kappa}_\xi U_t U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\psi_i b_{m-i-1}b_m'\}] \\ & = E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\textstyle\sum_{i,j}(\sum_t a_{t-1}b_{t-j-1}')\psi_j'\Sigma^{-1}\tilde{\kappa}_\xi\tilde{\kappa}_\xi'\Sigma^{-1}\psi_i(T^{-1}\sum_m b_{m-i-1}b_m')\}] \end{aligned}$$

which tends to zero since $T^{-1}\sum_m b_{m-i-1}b_m' \to 0$. Thus K_7 does not give a contribution.

Finally we consider the term K_3 which apparently is of higher order of magnitude, such that we have to take into account more terms when expanding.

8.3.6 The term $\mathbf{K}_3 = 2TE_{\xi_\perp}[tr\{S_{ba.v}S_{aa}^{-1}S_{au.b,v}\kappa_\xi S_{vv}^{-1}S_{vb}\}]$

We first consider

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

since $M_{ad}=0$. We replace it by M_{aa} and next expand using $M_{ad}=0$

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

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

We insert

$$\kappa_\xi S_{vv}^{-1}S_{vb} = \tilde{\kappa}_\xi M_{yy.d}^{-1}M_{yb.d}.$$

and find

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

such that K_3 is split into 5 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 $\mathbf{K}_{31}$ Let $b_{t.d} = b_t - M_{bd}M_{dd}^{-1}d_t$, then setting $t = s - r - 1$, we get

$$\begin{aligned}
K_{31} &= 2E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb.d}\}] \\
&= 2E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{t,s,i} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}(\theta_i U_{s-i-1} + \psi_i b_{s-i-1})b_{s.d}'\}] \\
&= 2E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{t,i} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_i U_t b_{t+i+1.d}'\}] \\
&= 2tr\{M_{ba}M_{aa}^{-1}\sum_{t,i} a_{t-1}b_{t+i+1.d}'\}tr\{\tilde{\kappa}_\xi'\Sigma^{-1}\theta_i\}.
\end{aligned}$$

In order to evaluate this we note that $d_{t+i+1} = M^{i+1}d_t$ such that

$$b_{t+i+1.d} = b_{t+i+1} - M_{bd}M_{dd}^{-1}d_{t+i+1} = b_{t+i+1} - M_{bd}M_{dd}^{-1}M^{i+1}d_t.$$

Hence

$$\begin{aligned} & \sum_t a_{t-1} b'_{t+i+1.d} \\ = & \sum_t a_{t-1} b'_{t+i+1} - \sum_t a_{t-1} d'_t M'^{i+1} M_{dd}^{-1} M_{db} \\ = & M_{ab} + O_P(T^{\frac{1}{2}}), \end{aligned}$$

since we have $M_{ad} = 0$. Thus we find the result

$$K_{31} \overset{0}{=} -2tr\{\tilde{\kappa}'_\xi \Sigma^{-1} \tilde{\kappa}_\xi\} tr\{M_{ba} M_{aa}^{-1} M_{ab}\}, \tag{59}$$

since $\theta = \sum_{i=0}^{\infty} \theta_i = -\tilde{\kappa}_\xi$.

The term $\mathbf{K}_{32}$

$$\begin{aligned} & K_{32} = -2T^{-1} E_{\xi_\perp}[tr\{M_{by} \Sigma^{-1} M_{ya} M_{aa}^{-1} M_{au} \tilde{\kappa}'_\xi \Sigma^{-1} M_{yb}\}] \\ & = -2T^{-1} E_{\xi_\perp}[tr\{\textstyle\sum_{t,s,r,m,j,i,k} b_t (U'_{t-j-1} \theta'_j + b'_{t-j-1} \psi'_j) \Sigma^{-1} \\ & \times (\theta_i U_{s-i-1} + \psi_i b_{s-i-1}) a'_{s-1} M_{aa}^{-1} a_{r-1} U'_r \tilde{\kappa}'_\xi \Sigma^{-1} (\theta_k U_{m-k-1} + \psi_k b_{m-k-1}) b'_m\}]. \end{aligned}$$

We try first four stochastic factors

$$\begin{aligned} K_{321} = \; & -2T^{-1} E_{\xi_\perp}[tr\{\textstyle\sum_{t,s,r,m,j,i,k} b_t U'_{t-j-1} \theta'_j \Sigma^{-1} \theta_i U_{s-i-1} a'_{s-1} \\ & \times M_{aa}^{-1} a_{r-1} U'_r \tilde{\kappa}'_\xi \Sigma^{-1} \theta_k U_{m-k-1} b'_m\}]. \end{aligned}$$

Since the factor in front is T^{-1} we need only terms of order 1 from this expectation. The only term where we shall get a summation of b with itself, which is of the order of T, is when $t = m$ and $k = j$. Then we must have $r = s - i - 1$, and we get

$$\begin{aligned} & K_{321} \\ = \; & -2T^{-1} E_{\xi_\perp}[tr\{\sum_{t,r,j,i} b_t U'_{t-j-1} \theta'_j \Sigma^{-1} \theta_i U_r a'_{r+i} M_{aa}^{-1} a_{r-1} U'_r \tilde{\kappa}'_\xi \Sigma^{-1} \theta_j U_{t-j-1} b'_t\}] \\ = \; & -2T^{-1} tr\{I_{n_a}\} tr\{\Sigma^{-1} \theta \tilde{\kappa}'_\xi \Sigma^{-1} \sum_{j=0}^{\infty} \theta_j \theta'_j\} tr\{S_{bb}\} \\ = \; & -2n_a n_b tr\{\Sigma^{-1} \theta \tilde{\kappa}'_\xi \Sigma^{-1} \sum_{j=0}^{\infty} \theta_j \theta'_j\}. \\ = \; & 2n_a n_b tr\{\Sigma^{-1} \tilde{\kappa}_\xi \tilde{\kappa}'_\xi \Sigma^{-1} \sum_{j=0}^{\infty} \theta_j \theta'_j\}. \end{aligned} \tag{60}$$

Next try two stochastic factors. There are three potential contributions (K_{322}, K_{323}, K_{324}) since the factor U_r is always present

$$\begin{aligned}
K_{322} &= -2T^{-1}E_{\xi_\perp}[tr\{ \\
&\quad \sum_{t,s,r,m,j,i,k} b_t U'_{t-j-1}\theta'_j \Sigma^{-1}\psi_i b_{s-i-1}a'_{s-1}M_{aa}^{-1}a_{r-1}U'_r\tilde{\kappa}'_\xi\Sigma^{-1}\psi_k b_{m-k-1}b'_m\}] \\
&= -2T^{-1}E_{\xi_\perp}[tr\{ \\
&\quad \sum_{s,m,j,i,k} b_{r+j+1}U'_r\theta'_j\Sigma^{-1}\psi_i b_{s-i-1}a'_{s-1}M_{aa}^{-1}a_{r-1}U'_r\tilde{\kappa}'_\xi\Sigma^{-1}\psi_k b_{m-k-1}b'_m\}] \\
&= -2T^{-1}E_{\xi_\perp}[tr\{ \\
&\quad \sum_{s,m,j,i,k} b_{r+j+1}b'_{s-i-1}\psi'_i\Sigma^{-1}\theta_j U_r a'_{s-1}M_{aa}^{-1}a_{r-1}U'_r\tilde{\kappa}'_\xi\Sigma^{-1}\psi_k b_{m-k-1}b'_m\}] \\
&= -2T^{-1}tr\{\sum_{s,m,j,i,k} b_{r+j+1}a'_{r-1}M_{aa}^{-1}a_{s-1}b'_{s-i-1}\psi'_i\Sigma^{-1}\theta_j\tilde{\kappa}'_\xi\Sigma^{-1}\psi_k b_{m-k-1}b'_m\} \\
&= -2T^{-1}tr\{M_{ba}M_{aa}^{-1}M_{ab}\psi'\Sigma^{-1}\theta\tilde{\kappa}'_\xi\Sigma^{-1}(\sum_{m,k}\psi_k b_{m-k-1}b'_m)\} \in o(1),
\end{aligned}$$

see (4.3).

$$\begin{aligned}
K_{323} &= -T^{-1}E_{\xi_\perp}[tr\{ \\
&\quad \sum_{t,s,r,m,j,i,k} b_t b'_{t-j-1}\psi'_j\Sigma^{-1}\psi_i b_{s-i-1}a'_{s-1}M_{aa}^{-1}a_{r-1}U'_r\tilde{\kappa}'_\xi\Sigma^{-1}\theta_k U_{m-k-1}b'_m\}] \\
&= -T^{-1}E_{\xi_\perp}[tr\{ \\
&\quad \sum_{t,s,r,j,i,k} b_t b'_{t-j-1}\psi'_j\Sigma^{-1}\psi_i b_{s-i-1}a'_{s-1}M_{aa}^{-1}a_{r-1}U'_r\tilde{\kappa}'_\xi\Sigma^{-1}\theta_k U_r b'_{r+k+1}\}] \\
&= -T^{-1}tr\{\sum_{t,s,r,j,i,k} b_t b'_{t-j-1}\psi'_j\Sigma^{-1}\psi_i b_{s-i-1}a'_{s-1}M_{aa}^{-1}a_{r-1}b'_{r+k+1}\}tr\{\tilde{\kappa}'_\xi\Sigma^{-1}\theta_k\} \\
&= -T^{-1}tr\{(\sum_{t,j} b_t b'_{t-j-1}\psi'_j)\Sigma^{-1}\psi M_{ba}M_{aa}^{-1}M_{ab}\}tr\{\tilde{\kappa}'_\xi\Sigma^{-1}\theta\} \in o(1),
\end{aligned}$$

see (4.3).

$$
\begin{aligned}
K_{324} &= -2T^{-1}E_{\xi_\perp}[tr\{ \\
&\sum_{t,s,r,m,j,i,k} b_t b'_{t-j-1}\psi'_j\Sigma^{-1}\theta_i U_{s-i-1}a'_{s-1}M_{aa}^{-1}a_{r-1}U'_r\tilde{\kappa}'_\xi\Sigma^{-1}\psi_k b_{m-k-1}b'_m\}] \\
&= -2T^{-1}tr\{\sum_{t,r,m,j,i,k} b_t b'_{t-j-1}\psi'_j\Sigma^{-1}\theta_i\tilde{\kappa}'_\xi\Sigma^{-1}\psi_k b_{m-k-1}b'_m\}tr\{a'_{r+i}M_{aa}^{-1}a_{r-1}\} \\
&= -2T^{-1}tr\{\sum_{t,j,i} b_t b'_{t-j-1}\psi'_j\Sigma^{-1}\theta_i\tilde{\kappa}'_\xi\Sigma^{-1}\psi_j b_{t-j-1}b'_t\}tr\{I_{n_a}\} \\
&= -2T^{-1}n_a tr\{\sum_{t,j,i} b_t b'_{t-j-1}\psi'_j\Sigma^{-1}\theta\tilde{\kappa}'_\xi\Sigma^{-1}\psi_j b_{t-j-1}b'_t\}.
\end{aligned}
$$

We make the approximation (35)

$$
T^{-1}\sum_t b_t(b'_{t-j-1}\psi'_j\Sigma^{-1}\theta\tilde{\kappa}'_\xi\Sigma^{-1}\psi_j b_{t-j-1} - tr\{\psi'_j\Sigma^{-1}\theta\tilde{\kappa}'_\xi\Sigma^{-1}\psi_j\})b'_t = 0
$$

and find

$$
\begin{aligned}
K_{324} &\overset{0}{=} -2n_a n_b tr\{\Sigma^{-1}\theta\tilde{\kappa}'_\xi\Sigma^{-1}\textstyle\sum_{j=0}^\infty \psi_j\psi'_j\}, \\
&= 2n_a n_b tr\{\Sigma^{-1}\tilde{\kappa}_\xi\tilde{\kappa}'_\xi\Sigma^{-1}\textstyle\sum_{j=0}^\infty \psi_j\psi'_j\},
\end{aligned}
$$

which together with (60) gives the contribution

$$
K_{32} \overset{0}{=} 2n_a n_b tr\{\kappa'_\xi\Sigma_{vv.z}^{-1}\kappa'_\xi\}. \tag{61}
$$

The terms $\mathbf{K}_{33}$

$$
\begin{aligned}
K_{33} &= -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}M_{ab}M_{bu.d}\tilde{\kappa}'_\xi\Sigma^{-1}M_{yb}\}] \\
&= -2tr\{M_{ba}M_{aa}^{-1}M_{ab}T^{-1}\textstyle\sum_{t,r,k} b_{t.d}b'_{t+k+1}\}tr\{\tilde{\kappa}'_\xi\Sigma^{-1}\theta_k\}.
\end{aligned}
$$

Next we evaluate

$$
T^{-1}\sum_t b_{t.d}b'_{t+k+1} = T^{-1}\sum_t b_t b'_{t+k+1} - M_{bd}M_{dd}^{-1}T^{-1}\sum_t d_t b'_{t+k+1} \to 0.
$$

Thus we find

$$
K_{33} \overset{0}{=} 0.
$$

The term K34

$$\begin{aligned}
&K_{34} \\
=\; & -2T^{-1}E_{\xi_\perp}[\{tr\{M_{ba}M_{aa}^{-1}M_{ay}\Sigma^{-1}M_{yu}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb}\}] \\
=\; & -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{t,s,r,i,j,k} a_t(U_{t-i-1}'\theta_i' + b_{t-i-1}'\psi_i')\Sigma^{-1} \\
& \times(\theta_j U_{s-j-1} + \psi_j b_{s-j-1})U_s'\tilde{\kappa}_\xi'\Sigma^{-1}(\theta_k U_{r-k-1} + \psi_k b_{r-k-1})b_r'\}]
\end{aligned}$$

For four stochastic factors we find

$$\begin{aligned}
&K_{341} \\
=\; & -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{t,s,r,i,j,k} a_t U_{t-i-1}'\theta_i'\Sigma^{-1}\theta_j U_{s-j-1}U_s'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_k U_{r-k-1}b_r'\}]
\end{aligned}$$

since $s-j-1 \neq s$ we must tie all indices to s and that will involve the summation of a and b, which when normalized by $M_{ba}M_{aa}^{-1}$ is bounded, and hence the contribution is $o(1)$, because of the factor T^{-1}.

For two stochastic terms we find two potential contributions which are small due to (4.3)

$$\begin{aligned}
&K_{342} \\
=\; & -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{t,s,r,i,j,k} a_t U_{t-i-1}'\theta_i'\Sigma^{-1}\psi_j b_{s-j-1}U_s'\tilde{\kappa}_\xi'\Sigma^{-1}\psi_k b_{r-k-1}b_r'\}] \\
=\; & -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{s,r,i,j,k} a_{s+i+1}U_s'\theta_i'\Sigma^{-1}\psi_j b_{s-j-1}U_s'\tilde{\kappa}_\xi'\Sigma^{-1}\psi_k b_{r-k-1}b_r'\}] \\
& -2E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{i,j,k}(\sum_s a_{s+i+1}b_{s-j-1}')\psi_j'\Sigma^{-1}\theta_i\tilde{\kappa}_\xi'\Sigma^{-1}\psi_k(T^{-1}\sum_r b_{r-k-1}b_r'\}]
\end{aligned}$$

which goes to zero since $T^{-1}\sum_r b_{r-k-1}b_r' \to 0$.

$$\begin{aligned}
&K_{343} \\
=\; & -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{t,s,r,i,j,k} a_t b_{t-i-1}'\psi_i'\Sigma^{-1}\psi_j b_{s-j-1}U_s'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_k U_{r-k-1}b_r'\}] \\
=\; & -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}\sum_{t,s,i,j,k} a_t b_{t-i-1}'\psi_i'\Sigma^{-1}\psi_j b_{s-j-1}U_s'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_k U_s b_{s+k+1}'\}] \\
=\; & -2tr\{M_{ba}M_{aa}^{-1}\sum_{i,j,k}\sum_t a_t b_{t-i-1}'\psi_i'\Sigma^{-1}\psi_j(T^{-1}\sum_s b_{s-j-1}b_{s+k+1}')\}tr\{\tilde{\kappa}_\xi'\Sigma^{-1}\theta_k\}
\end{aligned}$$

which again tends to zero. Thus the term becomes

$$K_{34} \overset{0}{=} -2T^{-1}E_{\xi_\perp}[\{tr\{M_{ba}M_{aa}^{-1}M_{ay}\Sigma^{-1}M_{yu}\tilde{\kappa}_\xi'\Sigma^{-1}M_{yb}\}] \overset{0}{=} 0,$$

and hence does not give a contribution.

Finally we need

The term $\mathbf{K}_{35}$

$$
\begin{aligned}
& K_{35} = 2E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1}M_{au}\tilde{\kappa}_\xi'\Sigma^{-1}(\Sigma - T^{-1}M_{yy})\Sigma^{-1}M_{yb}\}] \\
= \quad & -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1} \sum_{t,s,r,l,j,m,n,k} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1} \\
& \times[\theta_m(U_{r-m-1}U_{r-n-1}' - \delta_{nm}I_{n_v})\theta_n' + \psi_m(b_{r-m-1}b_{r-n-1}' - \delta_{nm}I_{n_b})\psi_n' \\
& +\theta_m U_{r-m-1}b_{r-n-1}'\psi_n' + \psi_m b_{r-m-1}U_{r-n-1}'\theta_n']\Sigma^{-1}(\theta_k U_{l-k-1} + \psi_k b_{l-k-1})b_l'\}]
\end{aligned}
$$

There is one term with four stochastic factors

$$
\begin{aligned}
& K_{351} = -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1} \\
& \quad \times \sum_{t,s,r,l,j,m,n,k} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_m(U_{r-m-1}U_{r-n-1}' - \delta_{nm}I_r)\theta_n'\Sigma^{-1}\theta_k U_{l-k-1}b_l'\}]
\end{aligned}
$$

If $t = l-k-1 \neq r-m-1 = r-n-1$ then the expectation is zero. If $t = r-m-1 \neq r-n-1 = l-k-1 (= t+m-n)$ we get

$$
\begin{aligned}
& K_{3511} = -T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1} \\
& \quad \times \sum_{t,l,j,m,n,k} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_m U_t U_{t+m-n}'\theta_n'\Sigma^{-1}\theta_k U_{t+m-n}b_{t+m-n+k+1}'\}] \\
= \quad & -T^{-1}tr\{M_{ba}M_{aa}^{-1}M_{ba}\}tr\{\theta'\Sigma^{-1}\theta\}tr\{\tilde{\kappa}_\xi'\Sigma^{-1}\theta\} \in o(1)
\end{aligned}
$$

If $t = r-n-1 \neq r-m-1 = l-k-1 (= t-m+n)$ we find

$$
\begin{aligned}
K_{3512} \quad = \quad & -T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1} \\
& \times \sum_{t,j,m,n,k} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_m U_{t-m+n}U_t'\theta_n'\Sigma^{-1}\theta_k U_{t-m+n}b_{t-m+n+k+1}'\}]
\end{aligned}
$$

which is again $\in o(1)$ by the same arguments.

With two stochastic factors we find the terms K_{352}, K_{353}, and K_{354}

$$
\begin{aligned}
& K_{352} = -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1} \\
& \times \textstyle\sum_{t,r,l,m,n,k} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\psi_m(b_{r-m-1}b_{r-n-1}' - \delta_{nm}I_{n-r})\psi_n'\Sigma^{-1}\theta_k U_{l-k-1}\}b_l'].
\end{aligned}
$$

$$
\begin{aligned}
& K_{353} = -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1} \\
& \times \sum_{t,r,l,m,n,k} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\theta_m U_{r-m-1}b_{r-n-1}'\psi_n'\Sigma^{-1}\psi_k b_{l-k-1}b_l'\}].
\end{aligned}
$$

$$K_{354} = -2T^{-1}E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1} \times \sum_{t,r,l,m,n,k} a_{t-1}U_t'\tilde{\kappa}_\xi'\Sigma^{-1}\psi_m b_{r-m-1}U_{r-n-1}'\theta_n'\Sigma^{-1}\psi_k b_{l-k-1}b_l'\}].$$

For K_{352} we let $t = l - k - 1$ and get

$$K_{352} = -2T^{-1}tr\{M_{ba}M_{aa}^{-1}M_{ab}\} tr\{\tilde{\kappa}_\xi'\Sigma^{-1}\sum_{r,m,n}\psi_m(b_{r-m-1}b_{r-n-1}' - \delta_{nm}I_{n-r})\psi_n'\Sigma^{-1}\theta\}.$$

This is $o(1)$ since the term with $n = m$ is zero, and when $n \neq m$ the sum is $o_P(T)$. For K_{353} we let $t = r - m - 1$ and get

$$K_{353} = -2tr\{M_{ba}M_{aa}^{-1} \times \textstyle\sum_{m,n,k}\sum_t a_{t-1}b_{t+m-n}'\psi_n'\Sigma^{-1}\psi_k(T^{-1}\sum_l b_{l-k-1}b_l')\}tr\{\tilde{\kappa}_\xi'\Sigma^{-1}\theta_m\} \in o(1)$$

For K_{354} we let $t = r - n - 1$ and find

$$\begin{aligned} K_{354} &= -2E_{\xi_\perp}[tr\{M_{ba}M_{aa}^{-1} \\ &\times \sum_{m,n,k}\sum_t a_{t-1}b_{t+m-n}'\psi_m'\Sigma^{-1}\tilde{\kappa}_\xi\theta_n'\Sigma^{-1}\psi_k(T^{-1}\sum_l b_{l-k-1}b_l')\} \\ &= -2T^{-1}tr\{M_{ba}M_{aa}^{-1} \\ &\times \sum_{m,n,k}\sum_t a_{t-1}b_{t+m-n}'\psi_m'\Sigma^{-1}\tilde{\kappa}_\xi\theta_n'\Sigma^{-1}\psi_k(T^{-1}\sum_l b_{l-k-1}b_l')\}, \end{aligned}$$

which tends to zero.

Thus we find that K_{35} does not give a contribution, and hence the contribution from K_3 is found from (61) and (59)

$$K_3 \overset{0}{=} 2(n_an_b - tr\{M_{ba}M_{aa}^{-1}M_{ab}\})tr(\kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi\}. \quad (62)$$

This completes the calculations and it remains to compare (62) with (54), (57) and (58)

$$\begin{aligned} K_4 &\overset{0}{=} tr\{\kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi\}tr\{M_{ba}M_{aa}^{-1}M_{ab}\} \\ K_5 &\overset{0}{=} tr\{M_{ba}M_{aa}^{-1}M_{ab}\}tr\{\kappa_\xi^2\Sigma_{vv.z}^{-1}\} \\ K_6 &\overset{0}{=} -n_an_btr\{\kappa_\xi\Sigma_{vv.z}^{-1}\kappa_\xi\} \end{aligned}$$

which is seen to give

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