

SUPPLEMENT TO “BOOTSTRAP ALGORITHMS FOR TESTING AND DETERMINING THE COINTEGRATION RANK IN VAR MODELS”
(*Econometrica*, Vol. 74, No. 6, November 2006, 1699–1714)

BY ANDERS RYGH SWENSEN

PROOF OF EQUATION (A.1): The first part of the proof follows the proof of Theorem 4.2 in Johansen (1995). We shall provide some details, but for a full account we refer to Johansen (1995).

We use the result from Lemma 4.1 in Johansen (1995): if $\Psi(z) = \sum_{i=0}^{\infty} \Psi_i z^i$ is convergent for $|z| < 1 + \delta$ for some $\delta > 0$, and $\Psi_i^\#$ and $\Psi^\#(z)$ are defined by $\Psi_i^\# = \sum_{j=i+1}^{\infty} \Psi_j$ and $\Psi^\#(z) = \sum_{i=0}^{\infty} \Psi_i^\# z^i$, respectively, then $\Psi^\#(z)$ is also convergent for $|z| < 1 + \delta$ and $\Psi(z) = \Psi(1) + (1 - z)\Psi^\#(z)$. If $\Psi(z)$ is a polynomial, so is $\Psi^\#(z)$.

From the proof of Theorem 4.2 in Johansen (1995) it also follows that if $\tilde{X}_t^* = (X_t^{*'} \hat{\beta}, \Delta X_t^{*'} \hat{\beta}_\perp)'$, the recursion defined in Algorithm 1 may be expressed as

$$\tilde{A}(L)\tilde{X}_t^* = (\tilde{\alpha}, \tilde{\alpha}_\perp)'(\varepsilon_t^* + \hat{\mu}_0 + \hat{\alpha}\hat{\rho}_1 t)$$

for a suitable polynomial.

We then use the fact that the zeros of $\det[\hat{A}(z)] = 0$ are equal to 1 or are outside the unit circle when T is large enough. This follows from the fact that Assumption 1 implies that $\det[A(z)] = 0$ has exactly $p - r$ solutions at 1 and the rest of the solutions are outside the unit circle; see Corollary 4.3 in Johansen (1995). The definition of $\hat{A}(z)$ implies that $\det[\hat{A}(z)] = 0$ has at least $p - r$ solutions at 1; see Johansen (1995, p. 16). Because the estimators of the coefficients are consistent, the solutions of $\det[\hat{A}(z)] = 0$ must converge to those of $\det[A(z)] = 0$. Thus when T is large enough, $p - r$ of the solutions $\det[\hat{A}(z)] = 0$ are equal to 1 and the rest are outside the unit circle.

Lemma 4.1 of Johansen (1995) can now be applied to $\tilde{C}(z) = \tilde{A}(z)^{-1}$. The reason is that $\det[\hat{A}(z)] = 0$ and $\det[\tilde{A}(z)] = 0$ have the same roots except for $z = 1$ and that $\det[\tilde{A}(1)] \neq 0$. An argument is in Johansen (1995, p. 51). Because, as noted previously, $\det[\hat{A}(z)] = 0$ has no root inside the unit circle, all the roots of $\det[\tilde{A}(z)] = 0$ must be outside the unit circle when T is large enough. Now, because $\det[\tilde{A}(z)]$ is a polynomial, $\det[\tilde{A}(z)] = 0$ has a finite number of solutions. Since all roots are outside the unit circle, there must exist a $\delta > 0$ such that $\det[\tilde{A}(z)] \neq 0$ when $|z| < 1 + \delta$. Hence, we can argue as in Theorem 11.3.1 in Brockwell and Davis (1991) to verify that $\tilde{C}(z) = \tilde{A}(z)^{-1}$

satisfies the condition of Lemma 4.1 of Johansen (1995). Then

$$\begin{aligned}
 \text{(S1)} \quad \tilde{X}_t^* &= \begin{pmatrix} \hat{\beta}' X_t^* \\ \hat{\beta}'_{\perp} \Delta X_t^* \end{pmatrix} \\
 &= \tilde{A}(L)^{-1}(\tilde{\alpha}, \tilde{\alpha}_{\perp})'(\varepsilon_t^* + \hat{\mu}_0 + \hat{\alpha}\hat{\rho}_1 t) \\
 &= \tilde{C}(L)(\tilde{\alpha}, \tilde{\alpha}_{\perp})'(\varepsilon_t^* + \hat{\mu}_0 + \hat{\alpha}\hat{\rho}_1 t) \\
 &= \tilde{C}(1)(\tilde{\alpha}, \tilde{\alpha}_{\perp})'(\varepsilon_t^* + \hat{\mu}_0 + \hat{\alpha}\hat{\rho}_1 t) \\
 &\quad + \tilde{C}^{\#}(L)(\tilde{\alpha}, \tilde{\alpha}_{\perp})'\Delta(\varepsilon_t^* + \hat{\mu}_0 + \hat{\alpha}\hat{\rho}_1 t). \qquad \text{Q.E.D.}
 \end{aligned}$$

PROOF OF EQUATION (A.2): We first consider the linear part and start with the decomposition

$$\text{(S2)} \quad X_t^* = \tilde{\beta}_{\perp} \hat{\beta}'_{\perp} \sum_{i=k+1}^t \Delta X_i^* + \tilde{\beta} \hat{\beta}' X_t^* + \tilde{\beta}_{\perp} \hat{\beta}'_{\perp} X_k^*.$$

Whereas $(0, \tilde{\beta}_{\perp}) \tilde{C}(1)(\tilde{\alpha}, \tilde{\alpha}_{\perp})' = \hat{C} = \hat{\beta}_{\perp}(\hat{\alpha}'_{\perp} \hat{\Gamma} \hat{\beta}_{\perp})^{-1} \hat{\alpha}'_{\perp}$, the first term in (S2) is equal to

$$\begin{aligned}
 \text{(S3)} \quad &\tilde{\beta}_{\perp} \hat{\beta}'_{\perp} \sum_{i=k+1}^t \Delta X_i^* \\
 &= \hat{C} \sum_{i=k+1}^t (\varepsilon_i^* + \hat{\mu}_0 + \hat{\alpha}\hat{\rho}_1 i) + (t-k)(0, \tilde{\beta}_{\perp}) \tilde{C}^{\#}(1)(\tilde{\alpha}, \tilde{\alpha}_{\perp})' \hat{\alpha}\hat{\rho}_1 \\
 &\quad + (0, \tilde{\beta}_{\perp}) \tilde{C}^{\#}(L)(\tilde{\alpha}, \tilde{\alpha}_{\perp})'(\varepsilon_t^* - \varepsilon_k^*).
 \end{aligned}$$

For $z \neq 1$ and $\tilde{A}(z)$ nonsingular,

$$\begin{aligned}
 \text{(S4)} \quad \tilde{C}^{\#}(z) \tilde{A}(z) &= \frac{\tilde{A}(z)^{-1} - \tilde{A}(1)^{-1}}{1-z} \tilde{A}(z) \\
 &= \frac{I - \tilde{A}(1)^{-1} \tilde{A}(z)}{1-z} \\
 &= \frac{-\tilde{A}(1)^{-1} [\tilde{A}(z) - \tilde{A}(1)]}{1-z} \\
 &= \tilde{A}(1)^{-1} \tilde{A}^{\#}(z).
 \end{aligned}$$

Therefore, $\tilde{C}^\#(1) = \tilde{A}(1)^{-1} \tilde{A}^\#(1) \tilde{A}(1)^{-1}$. However, $\tilde{A}^\#(1) = -\frac{d}{dz} \tilde{A}(z)|_{z=1} = -\dot{\tilde{A}}(1)$. The expression $(0, \tilde{\beta}_\perp) \tilde{C}^\#(1) (\tilde{\alpha}, \tilde{\alpha}_\perp)' \hat{\alpha} \hat{\rho}_1$ equals $\hat{C} \hat{\Gamma} \hat{\beta} \hat{\rho}_1$, which follows from Equation 37 in Rahbek and Mosconi (1999). The contribution to the linear term from (S3) is thus $(t-k)(\hat{C} \hat{\mu}_0 + \hat{C} \hat{\Gamma} \hat{\beta} \hat{\rho}_1)$, and the contribution from $\tilde{\beta} \hat{\beta}' X_t^*$ is $(\tilde{\beta}, 0) \tilde{C}(1) (\tilde{\alpha}, \tilde{\alpha}_\perp)' \hat{\alpha} \hat{\rho}_1 t = \tilde{\beta} \tilde{\alpha}' (\hat{\Gamma} \hat{C} - I) \hat{\alpha} \hat{\rho}_1 t = -\tilde{\beta} \hat{\rho}_1 t$. Summing, this is seen to equal $(t-k)\hat{\tau} - \tilde{\beta} \hat{\rho}_1 k$.

We now have considered the linear part of the lemma and have shown that

$$X_t^* = \hat{C} \sum_{i=k+1}^t \varepsilon_i^* + \hat{\tau}(t-k) + \sqrt{T} R_{t,T}^*,$$

where

$$\begin{aligned} \sqrt{T} R_{t,T}^* &= (0, \tilde{\beta}_\perp) \tilde{C}^\#(L) (\tilde{\alpha}, \tilde{\alpha}_\perp)' (\varepsilon_t^* - \varepsilon_k^*) \\ &\quad + (\tilde{\beta}, 0) \tilde{C}(1) (\tilde{\alpha}, \tilde{\alpha}_\perp)' (\varepsilon_t^* + \hat{\mu}_0) \\ &\quad + (\tilde{\beta}, 0) \tilde{C}^\#(L) (\tilde{\alpha}, \tilde{\alpha}_\perp)' \Delta (\varepsilon_t^* + \hat{\mu}_0 + \hat{\alpha} \hat{\rho}_1 t) \\ &\quad - \tilde{\beta} \hat{\rho}_1 k + \tilde{\beta}_\perp \hat{\beta}' X_k^* \\ &= (0, \tilde{\beta}_\perp) \tilde{C}^\#(L) (\tilde{\alpha}, \tilde{\alpha}_\perp)' (\varepsilon_t^* - \varepsilon_k^*) + (\tilde{\beta}, 0) \tilde{C}(1) (\tilde{\alpha}, \tilde{\alpha}_\perp)' \varepsilon_t^* \\ &\quad + (\tilde{\beta}, 0) \tilde{C}^\#(L) (\tilde{\alpha}, \tilde{\alpha}_\perp)' \Delta \varepsilon_t^* - \tilde{\beta} \hat{\rho}_1 k \\ &\quad + (\tilde{\beta}, 0) \tilde{C}(1) (\tilde{\alpha}, \tilde{\alpha}_\perp)' \hat{\mu}_0 \\ &\quad + (\tilde{\beta}, 0) \tilde{C}^\#(1) (\tilde{\alpha}, \tilde{\alpha}_\perp)' \hat{\alpha} \hat{\rho}_1 + \tilde{\beta}_\perp \hat{\beta}' X_k^*. \end{aligned} \quad Q.E.D.$$

PROOF OF EQUATION (A.3): When $\hat{\mu}_0 = \hat{\mu}_1 = 0$,

$$\begin{aligned} \sqrt{T} R_{t,T}^* &= (0, \tilde{\beta}_\perp) \tilde{C}^\#(L) \tilde{A}(L) (\tilde{X}_t^* - \tilde{X}_k^*) + \tilde{\beta} \hat{\beta}' X_t^* + \tilde{\beta}_\perp \hat{\beta}' X_k^* \\ &= \tilde{\beta} \hat{\beta}' X_t^* + (0, \tilde{\beta}_\perp) \tilde{A}(1)^{-1} \tilde{A}^\#(L) (\tilde{X}_t^* - \tilde{X}_k^*) + \tilde{\beta}_\perp \hat{\beta}' X_k^*, \end{aligned}$$

where the first equality follows from (S1) and the second follows from (S4).

Note that $\tilde{A}^\#(z)$ is a polynomial, where the coefficients are functions of the parameters, so that $R_{t,T}^*$ only involves a finite number of the generated X_t^* as $T \rightarrow \infty$. Q.E.D.

PROOF OF EQUATION (A.4): By a well-known argument (see, e.g., Hall and Heyde (1980, p. 53)),

$$\begin{aligned}
& P^* \left(\max_{k+1 \leq t \leq (T-k)} |\tilde{X}_t^*| / \sqrt{(T-k)} > \eta \right) \\
&= P^* \left(\frac{1}{T-k} \sum_{t=k+1}^T \tilde{X}_t^* \tilde{X}_t^* I[|\tilde{X}_t^*| > \eta \sqrt{(T-k)}] > \eta^2 \right) \\
&\leq \frac{1}{\eta^2} E^* \left[\frac{1}{T-k} \sum_{t=k+1}^T \tilde{X}_t^* \tilde{X}_t^* I(|\tilde{X}_t^*| > \eta \sqrt{(T-k)}) \right] \\
&\leq \frac{1}{\eta^4 (T-k)^2} \sum_{t=k+1}^T E^* (\tilde{X}_t^* \tilde{X}_t^*)^2,
\end{aligned}$$

so it suffices to show that

$$(S5) \quad \frac{1}{(T-k)^2} \sum_{t=k+1}^T E^* (\tilde{X}_t^* \tilde{X}_t^*)^2 \rightarrow 0$$

in probability.

In Equation 4.13 in Johansen (1995) it is shown that the relationship between $\hat{A}(z)$ and $\tilde{A}(z)$ is given by $(\tilde{\alpha}, \tilde{\alpha}_\perp)' \hat{A}(z) = \tilde{A}(z) (\hat{\beta}, \hat{\beta}_\perp (1-z))'$ when $z \neq 1$. Therefore, $\tilde{A}(0) = (\tilde{\alpha}, \tilde{\alpha}_\perp)' (\hat{\beta}, \hat{\beta}_\perp)^{-1} = (\tilde{\alpha}, \tilde{\alpha}_\perp)' (\hat{\beta}, \hat{\beta}_\perp)$, so that $\tilde{A}(L) \tilde{X}_t^* = (\tilde{\alpha}, \tilde{\alpha}_\perp)' \varepsilon_t^*$ may alternatively be represented as $B(L) \tilde{X}_t^* = (\hat{\beta}, \hat{\beta}_\perp)' \varepsilon_t^*$, where $\hat{B}(z)$ is a polynomial of order k with $\hat{B}(0) = I$.

Introduce $\underline{\tilde{X}}_t^* = (\tilde{X}_t^*, \dots, \tilde{X}_{t-k+1}^*)'$, $\kappa_t^* = (\hat{\beta}, \hat{\beta}_\perp)' \varepsilon_t^*$, $\underline{\kappa}_t^* = (\kappa_t^*, 0, \dots, 0)'$, and the matrix

$$\hat{B} = \begin{pmatrix} \hat{B}_1 & \cdots & \cdots & \cdots & \hat{B}_k \\ I & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & \cdots & I & 0 \end{pmatrix}.$$

Then

$$\underline{\tilde{X}}_t^* = \hat{B} \underline{\tilde{X}}_{t-1}^* + \underline{\kappa}_t^*, \quad t = k+1, \dots, T.$$

Solving the equation backward, we therefore have

$$\underline{\tilde{X}}_t^* = \sum_{j=0}^{t-(k+1)} \hat{B}^j \underline{\kappa}_{t-j}^* + \hat{B}^{(t-k)} \underline{\tilde{X}}_k^* = \underline{V}_{1,t}^* + \underline{V}_{2,t}^*.$$

Whereas $E^*(\tilde{X}_t^* \tilde{X}_t^*)^2 \leq E^*(\tilde{X}_t^* \tilde{X}_t^*)^2 \leq 8(E^*(\underline{V}_{1,t}^* \underline{V}_{1,t}^*)^2 + (\underline{V}_{2,t}^* \underline{V}_{2,t}^*)^2)$, (S5) and hence (A.4) will follow from

$$(S6) \quad \frac{1}{(T-k)^2} \sum_{t=k+1}^T E^*(\underline{V}_{l,t}^* \underline{V}_{l,t}^*)^2 \rightarrow 0, \quad l = 1, 2,$$

in probability. Consider first the case $l = 1$.

The equations $\det[\hat{B}(z)] = 0$ and $\det[\tilde{A}(z)] = 0$ have the same solutions, and the eigenvalues of \hat{B} equal the inverse of these solutions. Therefore all the eigenvalues of \hat{B} must be smaller in modulus than 1 when T is large enough, because the coefficients of the polynomial $\tilde{A}(z)$ tend to those of $\tilde{A}(z)$, and the solutions of $\det[\tilde{A}(z)] = 0$ have moduli larger than 1 by assumption (see, e.g., Johansen (1995, p. 51)). Then

$$\begin{aligned} & E^*(\underline{V}_{1,t}^* \underline{V}_{1,t}^*)^2 \\ & \leq E^* \left(\sum_{i,j=0}^{t-(k+1)} \underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^i \underline{\kappa}_{t-i}^* \right)^2 \\ & \leq E^* \left(\sum_{j=0}^{t-(k+1)} \underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^j \underline{\kappa}_{t-j}^* \underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^j \underline{\kappa}_{t-j}^* \right. \\ & \quad + 2 \sum_{i<j}^{t-(k+1)} \underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^i \underline{\kappa}_{t-i}^* \underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^i \underline{\kappa}_{t-i}^* \\ & \quad + 2 \sum_{i<j}^{t-(k+1)} \underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^i \underline{\kappa}_{t-i}^* \underline{\kappa}_{t-i}^* \hat{B}^i \hat{B}^j \underline{\kappa}_{t-j}^* \\ & \quad \left. + 2 \sum_{i<j}^{t-(k+1)} \underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^j \underline{\kappa}_{t-j}^* \underline{\kappa}_{t-i}^* \hat{B}^i \hat{B}^i \underline{\kappa}_{t-i}^* \right). \end{aligned}$$

For a vector $x = (x_1, \dots, x_n)'$, consider the norm $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ and for an $n \times n$ matrix $C = \{c_{ij}\}$, let $\|C\|_\infty$ be the induced norm, which equals $\max_{1 \leq i \leq n} \sum_{j=1}^n |c_{ij}|$. Then $\|x\|_\infty^2 \leq x'x \leq n\|x\|_\infty^2$ and if all the eigenvalues of C have modulus less than 1, $\|C^m\|_\infty \leq \text{const}|\lambda|^{m/2}$, where λ is an eigenvalue that has maximal modulus; see Corollary A.2 in Johansen (1995).

Because in our case $n = pk$, then

$$\begin{aligned} E^*(\underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^j \underline{\kappa}_{t-j}^* \underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^j \underline{\kappa}_{t-j}^*) & = E^*[(\underline{\kappa}_{t-j}^* \hat{B}^j \hat{B}^j \underline{\kappa}_{t-j}^*)^2] \\ & \leq (pk)^2 E^*(\|\hat{B}^j \underline{\kappa}_{t-j}^*\|_\infty^4) \end{aligned}$$

$$\begin{aligned}
&\leq (pk)^2 \|\hat{B}\|_\infty^{4j} E^*(\|\underline{\kappa}_{t-j}^*\|_\infty^4) \\
&\leq (pk)^2 \|\hat{B}\|_\infty^{4j} E^*[(\underline{\kappa}_{t-j}^{*'} \underline{\kappa}_{t-j}^*)^2] \\
&= (pk)^2 \|\hat{B}\|_\infty^{4j} E^*(|\underline{\kappa}_{t-j}^*|^4),
\end{aligned}$$

but $\|\hat{B}\|_\infty^{4j}$ is less than the product of the eigenvalue of \hat{B} that has the largest modulus to the power $2j$ and a constant, and this eigenvalue value converges in probability to a number with modulus less than 1. The existence of the fourth moments of the elements of ε_t will assure that $E^*(|\underline{\kappa}_t^*|^4)$ converges and therefore that (S6) holds. A similar argument works for the case $l = 2$, because the eigenvalues of \hat{B} have moduli that are less than 1.

The last statement in the lemma follows by the consistency of the estimators because $E^*[\varepsilon_t^* \varepsilon_t^{*'}] = \frac{1}{T-k} \sum_{i=k+1}^T \hat{\varepsilon}_i \hat{\varepsilon}_i'$. *Q.E.D.*

PROOF OF PROPOSITION 1: The proof in Johansen (1995) is based on a slightly different model, i.e., with unrestricted constant and no linear drift so that $\rho_1 = 0$ in model (1). The main ideas apply also in the situation we consider. The first step is to show that the $p - r + 1$ smallest solutions of

$$(S7) \quad \det[\lambda S_{11}^* - S_{10}^* S_{00}^{*-1} S_{01}^*] = 0$$

converge to zero. This follows from Lemma S1, which follows.

To find the asymptotic distribution, we use the representation in Lemma 1 and an argument similar to Johansen (1995) to show that the $p - r + 1$ smallest solutions multiplied by T converge weakly in probability toward the distribution of the solutions of

$$(S8) \quad \left| \rho \int_0^1 GG' - \int_0^1 G(dW)' \alpha_\perp (\alpha_\perp' \Omega \alpha_\perp)^{-1} \alpha_\perp' \int_0^1 (dW)G' \right| = 0.$$

This is a consequence of the continuous mapping theorem and Lemma S2 (following text). Proposition 1 will now follow by defining $B = (\alpha_\perp' \Omega \alpha_\perp)^{-1/2} \alpha_\perp' W$, which is a standard Brownian motion.

LEMMA S1: *Under the same assumptions as in Proposition 1 and if $\eta > 0$,*

$$P^*(\|S_{00}^* - \Sigma_{00}\| > \eta) \rightarrow 0,$$

$$P^*(\|S_{01}^* \hat{\beta}^\# - \Sigma_{0\beta}\| > \eta) \rightarrow 0,$$

and

$$P^*(\|\hat{\beta}^\# S_{11}^* \hat{\beta}^\# - \Sigma_{\beta\beta}\| > \eta) \rightarrow 0$$

in probability, where $\|\cdot\|$ denotes the matrix norm $\|C\| = [\text{tr}(C'C)]^{1/2}$. Here Σ_{00} , $\Sigma_{0\beta}$, and $\Sigma_{\beta\beta}$ are the limits in probability of S_{00} , $S_{01} \beta^\#$, and $\beta^\# S_{11} \beta^\#$, respectively.

LEMMA S2: *Under the same assumptions as in Proposition 1 with norming matrix $\hat{C}_T = \begin{pmatrix} \hat{\beta}_\perp & 0 \\ -\hat{\gamma}'\hat{\beta}_\perp & T^{-1/2} \end{pmatrix}$,*

$$\frac{1}{T}\hat{C}'_T S_{11}^* \hat{C}_T \rightarrow \int_0^1 GG'$$

and

$$\frac{1}{T}\hat{C}'_T (S_{10}^* - S_{11}^* \hat{\beta}^\# \hat{\alpha}') \rightarrow \int_0^1 G(dW)',$$

where $\bar{W} = \int_0^1 W(u) du$, $G(u) = ((W(u) - \bar{W})' C \bar{\beta}_\perp, u - 1/2)'$, and the convergence is weakly in probability. Q.E.D.

PROOF OF LEMMA S1: We consider only the first part; the others are proved in a similar manner. Note that $S_{00}^* = M_{00}^* - M_{02}^* M_{22}^{*-1} M_{20}^*$, where

$$M^* = \begin{pmatrix} M_{00}^* & M_{02}^* \\ M_{20}^* & M_{22}^* \end{pmatrix}$$

is the $(pk + 1) \times (pk + 1)$ matrix $\sum_{t=k+1}^T Z_t^* Z_t^{*'} / (T - k)$ and $Z_t^* = (\Delta X_t^{*'}, \dots, \Delta X_{t-k+1}^{*'}, 1)'$, $t = k + 1, \dots, T$. Let M be defined similarly in terms of the original observations, where $Z_t = (\Delta X_t', \dots, \Delta X_{t-k+1}', 1)'$, $t = k + 1, \dots, T$. By the ergodic theorem, $M \rightarrow \Sigma_M$, say, in probability, as $T \rightarrow \infty$. The lemma will follow if, for all $\eta > 0$,

$$(S9) \quad P^*(\|M^* - \Sigma_M\| > \eta) \rightarrow 0$$

in probability.

Whereas $\|M^* - \Sigma_M\|^2$ is equal to the sum of squares of the elements of $M^* - \Sigma_M$, it is sufficient to prove that (S9) is valid for each element.

We now verify that the variables Z_t and Z_t^* have a moving average representation. The model (1) may be written

$$\Delta X_t = \alpha \beta^{\#'} X_{t-1}^\# + \Gamma_1 \Delta X_{t-1} + \dots + \Gamma_{k-1} \Delta X_{t-(k-1)} + \mu_0 + \varepsilon_t.$$

After some manipulation, this is seen to imply

$$\begin{aligned} \beta' \Delta X_t + \rho_1(t + 1) &= (I + \beta' \alpha)(\beta' X_{t-1} + \rho_1 t) + \beta' \Gamma_1 \Delta X_{t-1} \\ &\quad + \dots + \beta' \Gamma_{k-1} \Delta X_{t-(k-1)} + \beta' \mu_0 + \beta' \varepsilon_t. \end{aligned}$$

Arguing as in Hansen and Johansen (1999, p. 311), the stochastic vector $(X'_t\beta^\#, \Delta X'_{t-1}, \dots, \Delta X'_{t-k+1})'$ can be represented as an AR(1) process

$$\begin{pmatrix} \beta^\# X_t^\# \\ \Delta X_t' \\ \vdots \\ \Delta X'_{t-k+1} \end{pmatrix} = \Phi \begin{pmatrix} \beta^\# X_{t-1}^\# \\ \Delta X_{t-1}' \\ \vdots \\ \Delta X'_{t-k} \end{pmatrix} + \begin{pmatrix} \beta' \mu_0 \\ \mu_0 \\ \vdots \\ 0 \end{pmatrix} + \begin{pmatrix} \beta' \varepsilon_t \\ \varepsilon_t \\ \vdots \\ 0 \end{pmatrix},$$

where the matrix Φ has the form

$$(S10) \quad \Phi = \begin{pmatrix} \beta' \alpha + I & \beta' \Gamma_1 & \cdots & \cdots & \beta' \Gamma_{k-1} \\ \alpha & \Gamma_1 & \cdots & \cdots & \Gamma_{k-1} \\ 0 & I & \cdots & \cdots & \cdots \\ \vdots & \vdots & & & \vdots \\ 0 & \cdots & \cdots & I & 0 \end{pmatrix}.$$

By also expanding the state space, $(X'_t\beta^\#, \Delta X'_{t-1}, \dots, \Delta X'_{t-k})'$ has such a representation. The process is stationary. Here we need the assumption that $\alpha'_\perp \Gamma \beta_\perp$ has full rank $p - r$. Furthermore, $(X'_t\beta^\#, \Delta X'_{t-1}, \dots, \Delta X'_{t-k})'$ can be represented in terms of the errors $\varepsilon_t, \varepsilon_{t-1}, \dots$. To see this, we note that the term Y_t in (5) corresponds to $\bar{\beta} Z_t$ in Equation (4.7) in the proof of Theorem 4.2 in Johansen (1995). Furthermore, from Equation (4.16) in the same proof, the term denoted by Z_t , and therefore Y_t here, is part of an autoregressive process that can be expressed by the errors $\varepsilon_t, \varepsilon_{t-1}, \dots$. This implies that all terms, apart from A , in (5) can be represented in this way. However, A cancels in $\Delta X'_{t-j}$, $j = 1, \dots, k$, and because $\beta' A = 0$, all terms in $(X'_t\beta^\#, \Delta X'_{t-1}, \dots, \Delta X'_{t-k})'$ can be represented in terms of the errors $\varepsilon_t, \varepsilon_{t-1}, \dots$. The process is therefore causal and has a characteristic polynomial with determinant that has zeros outside the unit circle; see, e.g., Theorems 3.1.1 and 11.3.1 in Brockwell and Davis (1991). Whereas the eigenvalues of the matrix in the AR(1) representation are the inverses of these roots, $(X'_t\beta^\#, \Delta X'_{t-1}, \dots, \Delta X'_{t-k})'$ has a moving average representation of the form $(X'_t\beta^\#, \Delta X'_{t-1}, \dots, \Delta X'_{t-k})' = \nu_1 + \sum_{i=0}^{\infty} \zeta_i \eta_{t-i}$, where $\eta_t = (\varepsilon'_t \beta, \varepsilon'_t, 0, \dots, 0)'$, $\zeta_0 = I$, and ζ_1, \dots are matrices such that the maximal eigenvalue of ζ_i is bounded by λ^i for some $0 < \lambda < 1$, uniformly in $i = 1, \dots$. The elements of the $[r + (k - 1)p]$ -dimensional vector can be expressed as functions of the parameters of (1). Thus, $Z_t = \nu_2 + \sum_{i=0}^{\infty} G \zeta_i \eta_{t-i}$ for a suitable matrix G and vector ν_2 . Each element of Z_t , which we also denote by Z_t (dropping the index), has a moving average representation

$$Z_t = \nu_3 + \sum_{i=0}^{\infty} \xi'_i \varepsilon_{t-i},$$

where $|\xi_i| < \text{const}\lambda^i$ uniformly in $i = 1, \dots$. Similarly,

$$Z_t^* = \hat{\nu}_3 + \sum_{i=0}^{\infty} \hat{\xi}_i' \varepsilon_{t-i}^*,$$

where $|\hat{\xi}_i| < \text{const}\lambda^i$ uniformly in $i = 1, \dots$ when T is large enough and $\hat{\nu}_3$ denotes ν_3 with the estimates for the unknown parameters plugged in.

Now $E(M) = E[(Z_t - \nu_3)^2] = \sum_{i=0}^{\infty} \xi_i' \Omega \xi_i$ and $E^*(M^*) = E^*[(Z_t^* - \hat{\nu}_3)^2] = \sum_{i=0}^{\infty} \hat{\xi}_i' \Omega_T \hat{\xi}_i$ so that $E^*(M^*) \rightarrow E(M)$ in probability. Whereas $\|M^* - \Sigma_M\| \leq \|M^* - E^*(M^*)\| + \|E^*(M^*) - \Sigma_M\|$, (S9) will follow from

$$P^*(\|M^* - E^*(M^*)\| > \eta) \rightarrow 0$$

in probability. Hence, by Chebychev's inequality it suffices to show that

$$(S11) \quad \frac{\text{Var}^*(\sum_{t=k+1}^T Z_t^{*2})}{(T-k)^2} = \frac{\sum_{t=-T+k-1}^{T-k+1} (1-t/T) \text{Cov}^*(Z_0^{*2}, Z_t^{*2})}{T-k} \rightarrow 0$$

in probability as $T \rightarrow \infty$. However,

$$\begin{aligned} \text{Cov}^*(Z_0^{*2}, Z_t^{*2}) &\leq \text{Var}^*(Z_0^{*2}) \leq E^*(Z_0^{*4}) \\ &= E^* \left[\sum_{i=0}^{\infty} (\hat{\xi}_i' \varepsilon_{-i}^*)^4 + 2 \sum_{0 \leq i < j} (\hat{\xi}_i' \varepsilon_{-i}^*)^2 (\hat{\xi}_j' \varepsilon_{-j}^*)^2 \right]. \end{aligned}$$

Now, we use that $(\hat{\xi}_i' \varepsilon_{-i}^*)^2 \leq |\hat{\xi}_i|^2 |\varepsilon_{-i}^*|^2$, that $|\hat{\xi}_i| < \text{const}\lambda^i$, $i = 1, \dots$, when T is large enough, and that $E^*(\|\varepsilon_t^*\|^4) \rightarrow E^*(\|\varepsilon_t\|^4) < \infty$ in probability by the weak law of large numbers and the existence of the fourth moments of the elements of ε_t . Then (S11) will follow. *Q.E.D.*

PROOF OF LEMMA S2: The proof relies on the continuous mapping theorem in the same way as the proof in Lemma 2 in Johansen (1994). The same functionals are involved, so it suffices to prove that the process $\{X_{[uT]}^*/\sqrt{T} : 0 \leq u \leq 1\}$ converges weakly in probability as element in $D[0, 1]^p$.

By the results from Lemma 1, it follows that the remainder term R_T^* vanishes, and it is sufficient to consider the linear part and $S_T^*(u) = \sum_{i=k+1}^{[uT]} \varepsilon_i^*/\sqrt{T}$. The linear term is treated as in Lemma 2 in Johansen (1994).

To prove that $S_T^* \rightarrow W$ weakly in probability, it is convenient to follow the approach of Pollard (1984) and exploit the fact that the limit is continuous. Hence, one can work with the uniform norm in $D[0, 1]^p$ and show that

$$E^*[f(S_T^*)] \rightarrow E[f(W)]$$

in probability for all bounded continuous functions f . This is explained in more detail for the one-dimensional case in Swensen (2003). *Q.E.D.*

Dept. of Mathematics, University of Oslo, P.O. Box 1053, N-0316 Blindern, Oslo, Norway; swensen@math.uio.no.

REFERENCES

- BROCKWELL, P. J., AND R. A. DAVIS (1991): *Time Series: Theory and Methods*. New York: Springer-Verlag. [1,8]
- HALL, P., AND C. C. HEYDE (1980): *Martingale Limit Theory and Its Applications*. New York: Academic Press. [4]
- HANSEN, H., AND S. JOHANSEN (1999): "Some Tests for Parameter Constancy in Cointegrated VAR-Models," *Econometrics Journal*, 2, 306–333. [8]
- JOHANSEN, S. (1994): "The Roll of the Constant and Linear Terms in Cointegration Analysis of Nonstationary Variables," *Econometric Reviews*, 13, 205–229. [9]
- (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford, U.K.: Oxford University Press. [1,2,4-6,8]
- POLLARD, D. (1984): *Convergence of Stochastic Processes*. New York: Springer-Verlag. [9]
- RAHBECK, A., AND R. MOSCONI (1999): "Cointegration Rank Inference with Stationary Regressors in VAR-Models," *Econometrics Journal*, 2, 76–91. [3]
- SWENSEN, A. R. (2003): "Bootstrapping Unit Root Tests for Integrated Processes," *Journal of Time Series Analysis*, 24, 99–126. [9]