Econometrica, Vol. 74, No. 6 (November, 2006), 1699–1714

BOOTSTRAP ALGORITHMS FOR TESTING AND DETERMINING THE COINTEGRATION RANK IN VAR MODELS[1]

BY ANDERS RYGH SWENSEN

In this paper a bootstrap algorithm for a reduced rank vector autoregressive model with a restricted linear trend and independent, identically distributed errors is analyzed. For testing the cointegration rank, the asymptotic distribution under the hypothesis is the same as for the usual likelihood ratio test, so that the bootstrap is consistent. It is furthermore shown that a bootstrap procedure for determining the rank is asymptotically consistent in the sense that the probability of choosing the rank smaller than the true one converges to zero.

KEYWORDS: Vector autoregressive (VAR) models, reduced rank, bootstrap, likelihood ratio test, determination of rank.

1. INTRODUCTION

AN IMPORTANT AND MUCH USED MODEL for analyzing multivariate nonstationary economic data is the reduced rank vector autoregressive (VAR) model where the p-dimensional observations $X_{k+1}, \dots, X_T$ satisfy the kth order autoregressive model

$$\Delta X_t = \alpha\beta' X_{t-1} + \Gamma_1 \Delta X_{t-1} + \cdots + \Gamma_{k-1}\Delta X_{t-(k-1)} + \alpha\rho_1 t + \mu_0 + \varepsilon_t. \tag{1}$$

The errors ε_t are independent and identically distributed. The lag length k and the initial observations $X_1, \dots, X_k$ are considered as given. The $p \times r$ matrices α and β have full rank. Denote models of this form $H^\#(r)$.

A test for the hypothesis $H^\#(r)$ versus $H^\#(p)$ is to reject for small values of the likelihood ratio $L(H^\#(r))/L(H^\#(p))$. It is well known (see, e.g., Johansen (1994, 1995)), that if the errors are Gaussian, the so-called trace statistic $Q_r^\#$ has the form

$$Q_r^\# = -2\log\left(\frac{L(H^\#(r))}{L(H^\#(p))}\right) = -(T-k)\sum_{i=r+1}^{p} \log(1-\hat\lambda_i),$$

where $\hat\lambda_1 > \cdots > \hat\lambda_p > \hat\lambda_{p+1} = 0$ are the ordered roots of $\det[\lambda S_{11} - S_{10}S_{00}^{-1} \times S_{01}] = 0$. The matrices S_{ij}, $i, j = 0, 1$, are defined by $S_{ij} = (T-k)^{-1}\sum_{t=k+1}^{T} R_{it}R_{jt}'$,

[1]The co-editor and two anonymous referees are thanked for helpful comments. A very preliminary version of this paper was presented at the first conference of the European Science Foundation Network: Econometric Methods for the Modeling of Nonstationary Data, Policy Analysis and Forecasting (EMM) in Arona, Italy, 19–21 September, 2002. Another version of this paper was presented at the 59th European Meeting of the Econometric Society in Madrid, 20–24 August, 2004. Comments from the participants, in particular Carsten Trenkler, and partial financial support from the EMM network and the Bank of Norway are gratefully acknowledged.

where R_{0t} and R_{1t}, are the residuals of the differences ΔX_t and the variables $X^{\#}_{t-1} = (X'_{t-1}, t)'$ after regressing on $\Delta X_{t-1}, \ldots, \Delta X_{t-k+1}$ and 1.

The most common use of this test is as a part of a sequential procedure to determining the rank, which consists of testing repeatedly $H^{\#}(0), H^{\#}(1), \ldots$ and determining the rank as the number of the first test that is not rejected; see, e.g., Johansen (1992 or 1995, Chap. 12). As is well known, the asymptotic approximations to the distribution of both these procedures are not always accurate. Bootstrap methods may therefore be an alternative; see e.g., Park (2003) for promising results in a univariate context.

The bootstrap algorithm for testing $H^{\#}(r)$ versus $H^{\#}(p)$ that we consider is based on the recursive scheme defined in (1) using estimated coefficients, errors sampled from the residuals, and the initial observations as starting values. Variants of this procedure have been considered earlier in simulation studies by van Giersbergen (1996), Harris and Judge (1998), and Mantalos and Shukur (2001).

First, we study methods for testing the rank of the matrix $\Pi = \alpha\beta'$ and provide conditions under which the bootstrap based tests are consistent. More precisely, we establish weak consistency by showing that the bootstrap statistics converge weakly in probability to the usual asymptotic distributions. This result is no surprise. It is known that the usual asymptotic arguments apply in this situation (see, e.g., Johansen (2002, p. 1930)). Next, we consider the procedure for determining the rank and we show that the bootstrap version is, in fact, consistent so that the probability of choosing the rank smaller than the true one converges to zero. This problem is more intriguing and presents some new aspects that are nonstandard in a bootstrap context, because we have to do the resampling for different values of the rank of the estimated reduced rank matrix. The dimension of the cointegration space in the generated observations will therefore not correspond to the true cointegration rank, but to the imposed rank in the simulated version of (1).

The plan for the paper is as follows. In the next section we describe the bootstrap algorithms. In Section 3 we state the conditions under which the algorithms yield estimators that have the same asymptotic properties as the estimators constructed directly from the observations. Section 4 contains results from some Monte Carlo simulations. The proofs can be found in the Appendix.

If an $m \times n$ matrix a, where $n \leq m$, has full rank, $a_\perp$ denotes an $m \times (m-n)$ matrix of full rank such that $a'_\perp a = 0$. The matrix $a(a'a)^{-1}$ is defined as $\bar{a}$, so that $a'\bar{a} = I_n$ and $\bar{a}a'$ is a projection matrix.

2. BOOTSTRAP AALGORITHMS

We consider the following bootstrap algorithm for constructing samples of bootstrapped or pseudo-observations $X^*_1, \ldots, X^*_T$ to test $H^{\#}(r)$ versus $H^{\#}(p)$.

ALGORITHM 1:

(i) Let $\hat{\Gamma}_1, \dots, \hat{\Gamma}_{k-1}, \hat{\mu}_0$ be the unrestricted ordinary least squares (OLS) estimators and let $\hat{\varepsilon}_{k+1}, \dots, \hat{\varepsilon}_T$ be the OLS residuals, i.e., estimate (1) with $r = p$.

(ii) Compute the estimators $\hat{\beta}^{\#} = (\hat{\beta}', \hat{\rho}_1)'$ and $\hat{\alpha}$ under the assumption that the reduced rank is r and that $\alpha_\perp \mu_1' = 0$, as defined in Johansen (1995, p. 96).

(iii) Check whether the roots of the equation $\det[\hat{A}(z)] = 0$, where

$$\hat{A}(z) = (1-z)I_p - \hat{\alpha}\hat{\beta}'z - \hat{\Gamma}_1(1-z)z - \cdots - \hat{\Gamma}_{k-1}(1-z)z^{k-1},$$

are equal to 1 or outside the unit circle and whether $\hat{\alpha}_\perp' \hat{\Gamma} \hat{\beta}_\perp$ is nonsingular, where $\hat{\Gamma} = I - \hat{\Gamma}_1 - \cdots - \hat{\Gamma}_{k-1}$.

(iv) If so, compute X_t^*, $t = k+1, \dots, T$, recursively from

$$\Delta X_t^* = \hat{\alpha}\hat{\beta}' X_{t-1}^* + \hat{\Gamma}_1 \Delta X_{t-1}^* + \cdots + \hat{\Gamma}_{k-1} \Delta X_{t-(k-1)}^* + \hat{\alpha}\hat{\rho}_1 t + \hat{\mu}_0 + \varepsilon_t^*, \tag{2}$$

with sampled residuals ε_t^* drawn with replacement from the estimated residuals $\hat{\varepsilon}_{k+1}, \dots, \hat{\varepsilon}_T$ and with the initial observations $X_1, \dots, X_k$ as starting values. Set $X_1^*, \dots, X_k^*$ equal to $X_1, \dots, X_k$.

The residuals and pseudo-observations depend on T, but we do not indicate this dependency when no confusion can arise. As usual in the literature, we use an asterisk ($*$) to denote the bootstrap distributions, so P^* is the conditional distribution of $X_1^*, \dots, X_T^*$ given the observations $X_1, \dots, X_T$.

REMARK 1: It may happen that some of the zeros of the estimated determinant of the characteristic equation are outside the unit circle (see Johansen (1995, p. 71)). The bootstrap sample will then become explosive. The purpose of checking the conditions in step (iii) is therefore to ensure that the pseudo-observations from the recursive scheme are in fact $I(1)$ variables. In the case when the conditions are not satisfied, it is an indication that the model is misspecified and that another more appropriate recursive scheme that reflects the properties of the observed data should be used. From a numerical point of view, the requirement that $\hat{\alpha}_\perp' \hat{\Gamma} \hat{\beta}_\perp$ is nonsingular is innocuous. Whereas $\hat{\Gamma}$ is the unrestricted OLS estimate, it will be true with probability 1. However, to stress the point that we want the bootstrap samples to be $I(1)$, we keep the requirement.

REMARK 2: The residuals that are used to construct the bootstrap samples are obtained by fitting an unrestricted VAR model of order k using OLS. Also the unrestricted OLS estimators $\hat{\Gamma}_1, \dots, \hat{\Gamma}_{k-1}$ and $\hat{\mu}_0$ are used in the recursion (2). An alternative is to use the residuals and estimators after having imposed the reduced rank r. The contrast between using the unrestricted estimators and imposing all restrictions corresponds to the univariate case to base the bootstrap replications either on the OLS residuals or on the differences of the observations. For unit root tests, Paparoditis and Politis (2003, 2005) considered this problem and found that there may be certain advantages to using the unrestricted estimators when the null hypothesis of a unit root is not true.

The bootstrap version of the statistic $Q_r^{\#}$, which we denote by $Q_r^{\#*}$, is now computed by regressing ΔX_t^* and $X_{t-1}^{\#*} = (X_{t-1}^{*\prime}, t)'$ on $\Delta X_{t-1}^*, \ldots, \Delta X_{t-k+1}^*$ and 1. Denote the moment matrices computed from the residuals S_{ij}^*, $i, j = 0, 1$, where the value 0 as before refers to differences and the value 1 refers to levels. Then $\hat{\lambda}_1^* > \cdots > \hat{\lambda}_p^* > \hat{\lambda}_{p+1}^* = 0$ are the ordered roots of $\det[\lambda S_{11}^* - S_{10}^* S_{00}^{*-1} S_{01}^*] = 0$ and $Q_r^{\#*} = -(T-k)\sum_{i=r+1}^{p} \log(1-\hat{\lambda}_i^*)$.

The bootstrap distribution of $Q_r^{\#*}$ is the conditional distribution of $Q_r^{\#*}$ given the observations $X_1, \ldots, X_T$. Denote the cumulative distribution function by $F_{r,T}^*$. The bootstrap test of $H^{\#}(r)$ versus $H^{\#}(p)$ at level γ consists of rejecting $H^{\#}(r)$ if $Q_r^{\#}$ is larger than the $(1-\gamma)$ quantile of the bootstrap distribution; that is, if $1 - F_{r,T}^*(Q_r^{\#}) \leq \gamma$. Because the bootstrap distribution is usually a very complicated function of the observations, it must be approximated. A feasible bootstrap test consists of rejecting if

$$\frac{1}{B}\,\mathrm{card}\{Q_r^{\#*} > Q_r^{\#}\} \leq \gamma, \tag{3}$$

where card$\{\cdot\}$ denotes the number of elements in the set $\{\cdot\}$ and B denotes the number of pseudo-samples that are generated. When B increases, expression (3) will converge toward $1 - F_{r,T}^*(Q_r^{\#}) \leq \gamma$.

As mentioned in the Introduction the most important use of the likelihood ratio test for the cointegration rank is as part the procedure to determine the rank. Since pseudo-samples are produced for each rank, a modification of Algorithm 1 is needed. The following enhancement seems natural.

ALGORITHM 2:

(i) Same as in Algorithm 1.

(ii) Let $\hat{\beta}_j^{\#} = (\hat{\beta}_j', \hat{\rho}_{1j})'$, $j = 0, \ldots, p-1$, be the estimators of the cointegration vectors that assume the reduced rank is j and that impose the condition $\alpha_\perp \mu_1' = 0$ as defined in Johansen (1995, p. 96). Let the norming be the usual. Define $\hat{\alpha}_j = S_{01}\hat{\beta}_j^{\#}(\hat{\beta}_j^{\#\prime} S_{11}\hat{\beta}_j^{\#})^{-1}$, $j = 0, \ldots, p-1$.

(iii) Starting with $j = 0$, check whether the roots of the equation $\det[\hat{A}_j(z)] = 0$, where

$$\hat{A}_j(z) = (1-z)I_p - \hat{\alpha}_j\hat{\beta}_j' z - \hat{\Gamma}_1(1-z)z - \cdots - \hat{\Gamma}_{k-1}(1-z)z^{k-1},$$

are equal to 1 or located outside the unit circle, and whether $\hat{\alpha}'_{j\perp}\hat{\Gamma}\hat{\beta}_{j\perp}$ is nonsingular with $\hat{\Gamma} = I - \hat{\Gamma}_1 - \cdots - \hat{\Gamma}_{k-1}$.

(iv) If so, compute $X^*_{j,t}$, $t = k+1, \dots, T$, recursively from

$$\Delta X^*_{j,t} = \hat{\alpha}_j\hat{\beta}'_j X^*_{j,t-1} + \hat{\Gamma}_1\Delta X^*_{j,t-1} + \cdots + \hat{\Gamma}_{k-1}\Delta X^*_{j,t-(k-1)} + \hat{\mu}_0 + \hat{\alpha}_j\hat{\rho}_{1j}t + \varepsilon^*_t, \quad t = k+1, \dots, T, \tag{4}$$

with the sampled residuals ε^*_t drawn with replacement from the estimated residuals $\hat{\varepsilon}_{k+1}, \dots, \hat{\varepsilon}_T$. Use the initial observations $X_1, \dots, X_k$ as starting values. Set $X^*_{j,1}, \dots, X^*_{j,k}$ equal to $X_1, \dots, X_k$.

(v) Check if the fraction $\frac{1}{B}\operatorname{card}\{Q^{\#*}_j > Q^{\#}_j\}$ exceeds a fixed level γ. If it does, estimate the rank as j and stop, else perform steps (iii)–(v) with rank equal to $j+1$.

(vi) If none of the fractions $\frac{1}{B}\operatorname{card}\{Q^{\#*}_0 > Q^{\#}_0\}, \dots, \frac{1}{B}\operatorname{card}\{Q^{\#*}_{p-1} > Q^{\#}_{p-1}\}$ exceeds γ, let the estimated rank be p.

3. ASYMPTOTIC RESULTS

We now consider the asymptotic distribution of the bootstrap distribution introduced in the previous section for the case that $\{X_t\}$ is an $I(1)$ process defined by (1). More precisely, we make the following assumption:

ASSUMPTION 1:

(i) *Define the characteristic polynomial*

$$A(z) = (1-z)I_p - \Pi z - \Gamma_1(1-z)z - \cdots - \Gamma_{k-1}(1-z)z^{k-1},$$

where $\Pi = \alpha\beta'$. *Assume that the roots of* $\det[A(z)] = 0$ *are located outside the complex unit circle or at* 1. *Also assume that the matrices* α *and* β *have full rank* r *and that* $\alpha'_\perp\Gamma\beta_\perp$ *has full rank* $p-r$, *where* $\Gamma = I_p - \Gamma_1 - \cdots - \Gamma_{k-1}$.

(ii) *The random variables* ε_i, $i = 0, \pm1, \dots$, *are independent*, *identically distributed with expectation* 0 *and covariance matrix* Ω.

From Johansen (1994, 1995) it follows that under Assumption 1, X_t has the representation

$$X_t = C\sum_{i=k+1}^{t}\varepsilon_i + \tau(t-k) + Y_t + A, \tag{5}$$

where $C = \beta_\perp(\alpha'_\perp\Gamma\beta_\perp)^{-1}\alpha'_\perp$, $\tau = \bar{\beta}_\perp(\bar{\alpha}'_\perp\Gamma\bar{\beta}_\perp)^{-1}\bar{\alpha}'_\perp\mu_0 + (\bar{\beta}_\perp(\bar{\alpha}'_\perp\Gamma\bar{\beta}_\perp)^{-1}\bar{\alpha}'_\perp\Gamma - I)\bar{\beta}\rho_1 = C\mu_0 + (C\Gamma - I)\bar{\beta}\rho_1$, Y_t is a stationary process, and A is a term that depends only on the initial values so that $\beta' A = 0$. This implies, in particular, that if $\beta^{\#} = (\beta', \rho_1)'$, the process $\beta^{\#\prime} X_t^{\#}$ will be stationary for a suitable initial distribution.

The random process $\{\frac{1}{\sqrt{T}} \sum_{i=1}^{[uT]} \varepsilon_i, 0 \le u \le 1\}$, where $[uT]$ is the integer value of uT, converges weakly toward the Brownian motion $\{W(u), 0 \le u \le 1\}$ with covariance matrix Ω. From this result the asymptotic behavior of X_t can be deduced. In Johansen (1994, 1995) this is used to derive the asymptotic distribution of the likelihood ratio test in the case that the errors are Gaussian. As he points out, the proof is valid also for non-Gaussian variables.

Consider the likelihood ratio test for $H^{\#}(r)$ versus $H^{\#}(p)$ constructed on the pseudo-data generated by Algorithm 1. It is not surprising that a representation of the form (5) is valid for each bootstrap replication. The remainder terms will depend on the realization, however. In a bootstrap context it is necessary to pay closer attention to such terms. The following representation can be established by an elaboration of the proof of (5).

LEMMA 1: *Under Assumption 1 and if in addition $E(\varepsilon_1^4) < \infty$, the generated pseudo-observations have the representation*

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

where for all $\eta > 0$, $P^*(\max_{k+1 \le t \le T} |R_{t,T}^*| > \eta) \to 0$ in probability as $T \to \infty$, $|\cdot|$ is the usual Euclidean distance, $\hat{C} = \hat{\beta}_\perp (\hat{\alpha}_\perp' \hat{\Gamma} \hat{\beta}_\perp)^{-1} \hat{\alpha}_\perp'$, and $\hat{\tau} = \hat{C}\hat{\mu}_0 + (\hat{C}\hat{\Gamma} - I)\bar{\hat{\beta}}\hat{\rho}_1$. Furthermore, $E^*[\varepsilon_i^* \varepsilon_i^{*\prime}] = \hat{\Omega}_T \to \Omega$ in probability as $T \to \infty$.

Therefore, X_t^* consists of a stochastic trend and a deterministic trend in addition to the stationary part. The random process $\{\frac{1}{\sqrt{T}} \sum_{i=k+1}^{[uT]} \varepsilon_i^*, 0 \le u \le 1\}$ converges weakly in probability as functions in $D[0,1]^p$ toward a p-dimensional Brownian motion W, with covariance matrix Ω, which implies the following proposition:

PROPOSITION 1: *Let the bootstrap samples be generated by Algorithm 1. Under Assumption 1 and if, in addition, $E(\varepsilon_1^4) < \infty$, as $T \to \infty$ and for all x,*

$$P^*(Q_r^{*\#} \le x) \to P\left(\mathrm{tr}\left(\int_0^1 (dB)F' \left[\int_0^1 FF'\right]^{-1} \int_0^1 F(dB)'\right) \le x\right) \tag{6}$$

in probability, where B is a $(p-r)$-dimensional standard Brownian motion on the unit interval and $F(u) = ((B(u) - \int_0^1 B(u)\,du)', u - 1/2)'$.

PROOF: The proof follows the arguments in the proof of Theorem 11.1 in Johansen (1995). The details can be found on the *Econometrica* supplementary material website (Swensen (2006)). *Q.E.D.*

REMARK 3: By inspecting the proof of Proposition 1 one can see that the result is still valid when the pseudo-observations are generated using the restricted residuals and estimators described in Remark 2. What is needed is that the mean of the squared residuals converges in probability toward Ω and that the estimators are consistent. These properties are also true for the restricted residuals and estimators.

The bootstrap procedure to determine the rank has the following consistency property in common with the usual procedure.

PROPOSITION 2: *Let the bootstrap samples be generated by Algorithm 2. Then, under Assumption 1 and if, in addition, $E(\varepsilon_1^4) < \infty$,*

$$F^*_{j,T}(Q^\#_j) \to 1, \quad j = 0, \ldots, r-1,$$

in probability as $T \to \infty$.

Combining Propositions 1 and 2 yields a corollary:

COROLLARY 1: *If R^* is the estimator of the rank described in Algorithm 2, under Assumption 1 and if $E(\varepsilon_1^4) < \infty$,*

$$\begin{aligned}
&\lim_{T\to\infty} P(R^* = j) = 0 \quad \textit{for} \quad j = 0, \ldots, r-1,\\
&\lim_{T\to\infty} P(R^* = r) = 1 - \gamma \quad \textit{for} \quad j = r,\\
&\limsup_{T\to\infty} P(R^* = j) \leq \gamma \quad \textit{for} \quad j = r+1, \ldots, p.
\end{aligned}$$

PROOF: The statement in the first line is a reformulation of Proposition 2. The statement in the second line is the probability of not rejecting $H^\#(r)$ against $H^\#(p)$, and follows from Proposition 1. Whereas $\sum_{j=0}^{p} P(R^* = j) = 1$, the statement in the third line must also hold. *Q.E.D.*

PROOF OF PROPOSITION 2: In Lemma 12.4 in Johansen (1995) it is proved that $Q^\#_0, \ldots, Q^\#_{r-1} \to \infty$ in probability, because all these statistics contain a contribution from an estimated eigenvalue that is positive in the limit. Hence the proposition will follow if we show that the random variables $Q^{\#*}_0, \ldots, Q^{\#*}_{r-1}$ converge weakly in probability.

To do that we proceed as in the proof of Proposition 1 using the following generalization of Lemma 1.

LEMMA 2: *Under the same assumptions as in Proposition 2, the generated pseudo-observations from Algorithm 2, for each $j = 0, \ldots, r-1$, have the representations*

$$X^*_{j,t} = \hat{C}_j \sum_{i=k+1}^{t} \varepsilon^*_i + \hat{\tau}_j(t-k) + \sqrt{T} R^*_{j,t,T}, \quad t = k+1, \ldots, T,$$

where for all $\eta > 0$, $P^*(\max_{k+1\le t\le T} |R^*_{j,t,T}| > \eta) \to 0$ *in probability as* $T \to \infty$, $|\cdot|$ *is the usual Euclidean distance*, $\hat{C}_j = \hat{\beta}_{j\perp}(\hat{\alpha}'_{j\perp}\hat{\Gamma}\hat{\beta}_{j\perp})^{-1}\hat{\alpha}'_{j\perp}$, *and* $\hat{\tau}_j = \hat{C}_j\hat{\mu}_0 + (\hat{C}_j\hat{\Gamma} - I)\bar{\hat{\beta}}_j\hat{\rho}_{1j}$.

By inspecting the proof of Lemma 1 it can be seen that it is essential to verify that the equation $\det[\hat{A}_j(z)] = 0$, which corresponds to the recursion (4), that is used to generate the pseudo-data, has roots in 1 or located outside the unit circle, that $\hat{\Pi}_j = \hat{\alpha}_j\hat{\beta}'_j$ has full rank j, and that $\hat{\alpha}'_{j\perp}\hat{\Gamma}\hat{\beta}_{j\perp}$ has full rank $p - j$ when T is large enough. These properties will also guarantee that the bootstrapped data are $I(1)$ and have cointegration rank j, $j = 0, \dots, r-1$. For a proof, use the consistency of the least squares estimators $\hat{\Gamma}_1, \dots, \hat{\Gamma}_{k-1}, \hat{\mu}_0$; see, e.g., Sims, Stock, and Watson (1990), and the following lemma.

LEMMA 3: *Let Assumption* 1 *be satisfied and define* Σ_{00} *as the limit in probability of* S_{00}. *Let* $\kappa_1 > \cdots > \kappa_r$ *be the nonzero roots of* $\det[\kappa\Sigma_{00} - (\Sigma_{00} - \Omega)] = 0$, *let* $w_1, \dots, w_r$ *be the corresponding eigenvectors that satisfy* $W'\Sigma_{00}W = I$, *where* W *is the matrix with columns* $w_1, \dots, w_r$, *let* W_j *be the* $p \times j$ *matrix that has columns* $w_1, \dots, w_j$, *and let* $\Pi^\# = (\Pi, \alpha\rho_1)$. *Then*, *as* $T \to \infty$,

$$\hat{\alpha}_j\hat{\beta}^{\#\prime}_j = \hat{\Pi}^\#_j \to (\Sigma_{00} - \Omega)W_j[W'_j(\Sigma_{00} - \Omega)W_j]^{-1}W'_j\Pi^\# = \Pi^\#_j$$

in probability, $j = 0, 1, \dots, r-1$.

Proposition 2 will then follow from the representation in Lemma 2 and an argument similar to that used to prove Proposition 1. It thus suffices to show that the equation $\det[A_j(z)] = 0$, where

$$A_j(z) = (1-z)I_p - \Pi_j z - \Gamma_1(1-z)z - \cdots - \Gamma_{k-1}(1-z)z^{k-1},$$

has roots that are either equal to 1 or located outside the unit circle for $j = 0, \dots, r-1$. The matrix $\Pi^\#_j$ is defined in Lemma 3 and can be expressed as $\Pi^\#_j = \alpha_j\beta^{\#\prime}_j$, where $\beta^\#_j = (\beta'_j, \rho_{1j})'$. Then $\Pi_j = \alpha_j\beta'_j$. Furthermore, $\alpha'_{j\perp}\Gamma\beta_{j\perp}$ must have full rank. These facts follow from the next lemma.

LEMMA 4: *Let Assumption* 1 *be satisfied*. *Then the roots of the equations* $\det[A_j(z)] = 0$ *are either equal to* 1 *or have moduli larger than* 1 *and* $\alpha'_{j\perp}\Gamma\beta_{j\perp}$ *has full rank* $p - j$, $j = 0, 1, \dots, r-1$.

Hence, Proposition 2 is proved. *Q.E.D.*

4. NUMERICAL SIMULATIONS

The data generating processes (DGP) we simulate is the reduced rank VAR model of dimension 5 considered by Johansen (2002),

$$\Delta X_t = \alpha\beta' X_{t-1} + \varepsilon_t, \quad t = 1, \dots, T,$$

where $\alpha' = (a_1, a_2, 0, 0, 0)$ and $\beta' = (1, 0, 0, 0, 0)$. The errors ε_t are independent Gaussian with diagonal covariance matrix $\Omega = I_5$.

For each value of α, the DGP was simulated $S = 10{,}000$ times using the random number generator ran2 of Press, Teukolsky, Vetterling, and Flannery (1992), and models of lag length $k = 1$ that contain a constant and a restricted linear term were estimated. For each of the S realizations, the estimated parameters and the OLS residuals were then used to generate $B = 5{,}000$ sets of pseudo-observations, and the empirical distribution of the statistics based on these pseudo-observations was computed.

The initial observation, in this case X_0, was generated by drawing a random variable with a Gaussian distribution that has the variance of the stationary process $\bar{\beta}\beta' X_t$. The nonstationary part of X_0, $\bar{\beta}_\perp \beta'_\perp X_0$, was set equal to 0.

In Table I some results for the trace statistic for $T = 100$ are displayed for four parameter combinations. What is tabulated is the fraction of the generated samples where the test statistics exceed the asymptotic critical values tabulated in Osterwald-Lenum (1992). The corresponding columns are denoted by a prefix a. The prefix b denotes the empirical levels of the bootstrap tests, so that corresponding columns show the fractions for which (3) is satisfied for various levels γ.

The bootstrapped statistic has distributions shifted to the left compared to the asymptotic distribution. This is most pronounced for the combination $(a_1, a_2) = (-0.1, -0.1)$, where the bootstrap test has a level much below the nominal. As explained in Johansen (2002), the time series is in this case close to a DGP with lower rank, i.e., rank 0. For the other combinations considered,

TABLE I

EMPIRICAL LEVELS FOR TESTS BASED ON TRACE STATISTIC FOR VARIOUS VALUES OF (a_1, a_2) AND $T = 100$

	$(-0.1, -0.1)$		$(-0.8, -0.1)$		$(-0.1, -0.8)$		$(-0.8, -0.8)$	
Level	btr	atr	btr	atr	btr	atr	btr	atr
0.50	0.33	0.45	0.50	0.59	0.53	0.79	0.51	0.62
0.20	0.08	0.14	0.20	0.26	0.23	0.48	0.20	0.28
0.10	0.03	0.05	0.10	0.14	0.12	0.31	0.11	0.16
0.05	0.01	0.02	0.05	0.07	0.06	0.19	0.06	0.13
0.025	0.00	0.01	0.02	0.04	0.03	0.12	0.03	0.05
0.01	0.00	0.00	0.01	0.02	0.01	0.06	0.01	0.02

TABLE II

PROBABILITIES FOR CHOOSING VARIOUS RANKS WHEN $a_1 = a_2 = -0.4$

		$r=0$	$r=1$	$r=2$	$r=3$	$r=4–5$	Asymp.
$T = 100$	$r=0$	0.073	0	0	0	0	0.073
	$r=1$	0.051	0.800	0	0	0	0.851
	$r=2$	0	0.029	0.041	0	0	0.070
	$r=3$	0	0	0.003	0.003	0	0.006
	$r=4–5$	0	0	0	0	0	0
	Boot	0.124	0.829	0.044	0.003	0	
$T = 250$	$r=0$	0	0	0	0	0	0
	$r=1$	0	0.928	0	0	0	0.928
	$r=2$	0	0.023	0.043	0	0	0.066
	$r=3$	0	0	0.001	0.004	0	0.005
	$r=4–5$	0	0	0	0	0	0
	Boot	0	0.951	0.044	0.004	0	

the nominal and empirical levels of the bootstrapped trace statistic are much closer. Indeed, the empirical levels of the bootstrapped test show much less variation over the grid of parameters considered than the test based on the asymptotic critical values.

Now we turn to the procedures for determination of the rank and we let $a_1 = a_2 = -0.4$. The simultaneous distribution for choosing a particular rank by the bootstrap method and by the asymptotic procedure is displayed in Table II for $T = 100$ and $T = 250$ using a level of 0.05. Because the true rank in this case is $r = 1$, the ideal method has a distribution concentrated in $r = 1$. As can be expected, because both procedures are consistent, the concentration in $r = 1$ is stronger for $T = 250$. Notice also the strong dependence between choosing a particular rank by the bootstrap method and by the asymptotic procedure. For $T = 100$, 0.917 of the mass is concentrated on the diagonal. The association is even stronger for $T = 250$: then 0.975 of the mass is on the diagonal.

Also the probability of choosing a too large rank, i.e., $r = 2, 3, 4,$ or 5, is for the bootstrap method bounded by the nominal level 0.05, but the asymptotic probability is not much larger. However, the fact that the bootstrap approximation may be closer than the asymptotic one does not entail that the method for determining the rank based on the bootstrap distributions is superior.

5. CONCLUSION

In this paper we have considered a recursive bootstrap algorithm for testing and determining the rank in a reduced rank VAR model. Under reasonable conditions these methods have the same asymptotic properties as the traditional ones. A small numerical study shows some improvements for the distribution of the test statistic.

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

Manuscript received March, 2006; final revision received April, 2006.

APPENDIX: PROOFS

PROOF OF LEMMA 1: First we use an argument as in the proof of Theorem 4.2 in Johansen (1995) to establish a representation of the pseudo-observations. Let $\tilde{X}_t^* = (X_t^{*\prime}\hat{\beta}, \Delta X_t^{*\prime}\hat{\beta}_\perp)'$. The recursion defined in Algorithm 1 may be expressed as $\tilde{\hat{A}}(L)\tilde{X}_t^* = (\bar{\hat{\alpha}}, \bar{\hat{\alpha}}_\perp)'(\varepsilon_t^* + \hat{\mu}_0 + \hat{\alpha}\hat{\rho}_1 t)$ for a suitable polynomial. From the consistency of the estimators it follows that $\det[\hat{A}(z)] = 0$ implies that $z = 1$ or $|z| > 1$, when T is large enough. Then $\tilde{X}_t^*$ may be represented as

$$\tilde{X}_t^* = \tilde{\hat{C}}(1)(\bar{\hat{\alpha}}, \bar{\hat{\alpha}}_\perp)'(\varepsilon_t^* + \hat{\mu}_0 + \hat{\alpha}\hat{\rho}_1 t) + \tilde{\hat{C}}^{\#}(L)(\bar{\hat{\alpha}}, \bar{\hat{\alpha}}_\perp)'\Delta(\varepsilon_t^* + \hat{\mu}_0 + \hat{\alpha}\hat{\rho}_1 t), \tag{A.1}$$

where $\tilde{\hat{A}}(z)^{-1} = \tilde{\hat{C}}(z) = \tilde{\hat{C}}(1) + (1 - z)\tilde{\hat{C}}^{\#}(z)$.

This may be expressed as

$$X_t^* = \hat{C} \sum_{i=k+1}^{t} \varepsilon_i^* + \hat{\tau}(t - k) + \sqrt{T}R_{t,T}^*, \tag{A.2}$$

where

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

The terms that involve $\hat{\mu}_0$ and $\hat{\rho}_1$ in the previous expression are bounded in probability because the estimators are consistent. Whereas the terms that involve ε_t^* have the same form as if the recursion in Algorithm 1 contains no deterministic terms, we may therefore in the rest of the proof assume that $\hat{\mu}_0 = \hat{\mu}_1 = 0$.

The remainder term may then be expressed by the pseudo-observations $X_k^*, \dots, X_T^*$ as

$$\sqrt{T}R_{t,T}^* = \bar{\hat{\beta}}\hat{\beta}' X_t^* + (0, \bar{\hat{\beta}}_\perp)\tilde{\hat{A}}(1)^{-1}\tilde{\hat{A}}^{\#}(L)(\tilde{X}_t^* - \tilde{X}_k^*) + \bar{\hat{\beta}}_\perp \hat{\beta}_\perp' X_k^*. \tag{A.3}$$

Hence, the first part of the lemma follows if, for all $\eta > 0$,

$$P^*\Big(\max_{k+1 \le t \le T} |\tilde{X}_t^*|/\sqrt{T} > \eta\Big) \to 0 \tag{A.4}$$

in probability. Equation (A.4) follows by expressing $\{\tilde{X}_t^*\}$ as a stationary autoregressive process.

Additional details for the proof of equations (A.1)–(A.4) can be found in the supplement (Swensen (2006)). *Q.E.D.*

PROOF OF LEMMA 3: Let $\hat{\lambda}_1 > \cdots > \hat{\lambda}_p > \hat{\lambda}_{p+1} = 0$ be the ordered roots of $\det[\lambda S_{11} - S_{10}S_{00}^{-1}S_{01}] = 0$ and let $\hat{v}_1, \dots, \hat{v}_{p+1}$ be the corresponding eigenvectors such that $\hat{V}'S_{11}\hat{V} = I$, where $\hat{V}$ is the matrix with columns $\hat{v}_1, \dots, \hat{v}_{p+1}$. Then

$$S_{11}\hat{V}\hat{\Lambda}_{p+1} = S_{10}S_{00}^{-1}S_{01}\hat{V}, \tag{A.5}$$

where $\hat{\Lambda}_{p+1}$ is a diagonal matrix with diagonal elements $\hat{\lambda}_1, \dots, \hat{\lambda}_{p+1}$. Define the $(p+1) \times p$ matrices $\hat{D}_1 = (\hat{\Lambda}_p^{-1}, 0)'$ and $\hat{D}_2 = (\hat{\Lambda}_p^{-1/2}, 0)'$. Using (A.5), the p first columns of V may be written

$$S_{11}^{-1}S_{10}S_{00}^{-1}S_{01}\hat{V}\hat{D}_1 = S_{11}^{-1}S_{10}S_{00}^{-1}S_{01}\hat{V}\hat{D}_2\hat{\Lambda}_p^{-1/2} = S_{11}^{-1}S_{10}\hat{W}\hat{\Lambda}_p^{-1/2}, \tag{A.6}$$

where $\hat{W} = S_{00}^{-1}S_{01}\hat{V}\hat{D}_2$. However, from $S_{00}\hat{W}\hat{\Lambda}_p = S_{01}\hat{V}\hat{D}_2\hat{\Lambda}_p$ and

$$\begin{aligned} S_{01}S_{11}^{-1}S_{10}\hat{W} &= S_{01}S_{11}^{-1}S_{10}S_{00}^{-1}S_{01}\hat{V}\hat{D}_2 = S_{01}S_{11}^{-1}S_{11}\hat{V}\hat{\Lambda}_{p+1}\hat{D}_2 \\ &= S_{01}\hat{V}\hat{\Lambda}_{p+1}\hat{D}_2 = S_{01}\hat{V}\hat{D}_2\hat{\Lambda}_p, \end{aligned}$$

it follows that

$$S_{00}\hat{W}\hat{\Lambda}_p = S_{01}S_{11}^{-1}S_{10}\hat{W}. \tag{A.7}$$

Also using the definition of $\hat{W}$ and (A.5), $\hat{W}'S_{00}\hat{W} = \hat{D}_2'\hat{V}'S_{10}S_{00}^{-1}S_{00}S_{00}^{-1}S_{01} \times \hat{V}\hat{D}_2 = \hat{D}_2'\hat{V}'S_{11}\hat{V}\hat{\Lambda}_{p+1}\hat{D}_2 = I$.

Whereas $\hat{\beta}_j^{\#}$ is defined as the first j columns of $\hat{V}$, it follows from (A.6) that it may alternatively be expressed as $\hat{\beta}_j^{\#} = S_{11}^{-1}S_{10}\hat{W}_j\hat{\Lambda}_j^{-1/2}$, where $\hat{W}_j$ is the $p\times j$ matrix that consists of the j first columns of $\hat{W}$ and $\hat{\Lambda}_j$ is the $j\times j$ diagonal matrix with diagonal elements $\hat{\lambda}_1,\dots,\hat{\lambda}_j$. Then $\hat{\alpha}_j = S_{01}\hat{\beta}_j^{\#}(\hat{\beta}_j^{\#\prime}S_{11}\hat{\beta}_j^{\#})^{-1}$ and

$$\hat{\Pi}_j^{\#} = \hat{\alpha}_j\hat{\beta}_j^{\#\prime} = S_{01}S_{11}^{-1}S_{10}\hat{W}_j(\hat{W}_j'S_{01}S_{11}^{-1}S_{10}\hat{W}_j)^{-1}\hat{W}_j'S_{01}S_{11}^{-1}. \tag{A.8}$$

Equation (A.7) shows that $\hat{\lambda}_1,\dots,\hat{\lambda}_p$ and $\hat{W}$ are defined from the equation $\det[\lambda S_{00} - S_{01}S_{11}^{-1}S_{10}] = 0$ and the normalization $\hat{W}'S_{00}\hat{W} = I$. Note that when $T\to\infty$, $S_{00}\to\Sigma_{00}$ in probability and that the OLS estimator $S_{01}S_{11}^{-1}$ of $\Pi^{\#}$ is consistent. The conclusion of the lemma will follow from (A.8) if $S_{01}S_{11}^{-1}S_{10}\to\Sigma_{00}-\Omega$ in probability as $T\to\infty$. However, $S_{00} - S_{01}S_{11}^{-1}S_{10} = \frac{1}{T-k}\sum_{t=k+1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t'$, where $\hat{\varepsilon}_t$, $t=k+1,\dots,T$, are the OLS residuals and, by the consistency of the OLS estimators, $\frac{1}{T-k}\sum_{t=k+1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t'\to\Omega$ in probability as $T\to\infty$. *Q.E.D.*

PROOF OF LEMMA 4: Because we consider the zeroes of the determinants of characteristic polynomials, we may assume that $\mu_1=\mu_0=0$.

Write $A_j(z) = (1-z)I_p - \Pi_j z - \Gamma_1(1-z)z - \cdots - \Gamma_{k-1}(1-z)z^{k-1}$ as $A_j(z) = I - A_{1,j}z - A_2z^2 - \cdots - A_kz^k$, where $A_{1,j} = I + \Pi_j + \Gamma_1$, $A_i = \Gamma_i - \Gamma_{i-1}$, $i = 2,\dots,k-1$, and $A_k = -\Gamma_{k-1}$.

For each j the solutions to the equation $\det[A_j(z)]=0$ are equal to the inverses of the eigenvalues of the companion matrix

$$\begin{pmatrix} A_{1,j} & A_2 & \cdots & \cdots & A_k \\ I & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & \cdots & I & 0 \end{pmatrix}.$$

Now suppose that $a = (a_1',\dots,a_k')'$ is an eigenvector that corresponds to an eigenvalue equal to 1. Then $A_{1,j}a_1 + A_2a_2 + \cdots + A_ka_k = a_1$, $a_i = a_{i-1}$, $i = 2,\dots,k$, so that $A_j(1)a_1 = a_1 - A_{1,j}a_1 - A_2a_1 - \cdots - A_ka_1 = 0$, but $A_j(1) = -\Pi_j$, which has rank j. The space of solutions of $A_j(1)a_1 = 0$ has therefore dimension $p-j$ and the equation $\det[A_j(z)]=0$ has $pk-(p-j) = p(k-1)+j$ solutions that are different from 1. We will show that they all have modulus larger than 1.

Let $\{X_{j,t}\}$ be a process generated by the recursion defined by the characteristic polynomial $A_j(z)$. As explained in Hansen and Johansen (1999), the stochastic process $(X_{j,t}'\beta_j, \Delta X_{j,t-1}',\dots,\Delta X_{j,t-k+1}')'$ can be represented as an

AR(1) process using the matrix

$$\Phi_j = \begin{pmatrix} \beta_j'\alpha_j + I & \beta_j'\Gamma_1 & \cdots & \cdots & \beta_j'\Gamma_{k-1} \\ \alpha_j & \Gamma_1 & \cdots & \cdots & \Gamma_{k-1} \\ 0 & I & \cdots & \cdots & \cdots \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & \cdots & \cdots & I & 0 \end{pmatrix}. \tag{A.9}$$

Let $\Phi = \Phi_r$. Then matrices Φ_j, $j = 0, 1, \ldots, r-1$, are all submatrices of the matrix Φ. More precisely, Φ_j is Φ with rows and columns $j+1, \ldots, r$ deleted. Let x_j be an eigenvector of Φ_j and let λ be the corresponding eigenvalue. If $\tilde{x}_j$ is the $r + p(k-1)$ vector that has the same j first and $p(k-1)$ last elements as x_j and elements $j+1, \ldots, r$ equal to zero, then $\Phi\tilde{x}_j = \lambda\tilde{x}_j$.

All eigenvalues of Φ_j are therefore eigenvalues of Φ. Whereas these have moduli less than 1, the same is the case for the eigenvalues of Φ_j. Let λ be such an eigenvalue and let $b = (b_1', \ldots, b_k')'$ be the corresponding eigenvector. Then

$$\begin{aligned} &(\lambda - 1)b_1 - \beta_j'\alpha_j b_1 - \beta_j'\Gamma_1 b_2 - \cdots - \beta_j'\Gamma_{k-1}b_k = 0, \\ &-\alpha_j b_1 + (\lambda I - \Gamma_1)b_2 - \Gamma_2 b_3 - \cdots - \Gamma_{k-1}b_k = 0, \\ &\lambda b_i - b_{i-1} = 0, \quad i = 3, \ldots, k. \end{aligned}$$

This means that $b_i = \lambda^{-i+2}b_2$, $i = 3, \ldots, k$, and the preceding first two equations may be written

$$(\lambda - 1)b_1 - \beta_j'\alpha_j b_1 - \beta_j'\Gamma_1 b_2 - \cdots - \beta_j'\Gamma_{k-1}b_2\lambda^{-k+2} = 0 \tag{A.10}$$

and

$$-\alpha_j b_1 + (\lambda I - \Gamma_1)b_2 - \Gamma_2 b_2\lambda^{-1} - \cdots - \Gamma_{k-1}b_2\lambda^{-k+2} = 0. \tag{A.11}$$

Rearranging (A.11), premultiplying by β_j', and inserting (A.10) yield $\lambda\beta_j' b_2 = (\lambda - 1)b_1$. Therefore, if $\lambda \neq 1$, (A.11) may be written

$$\begin{aligned} &-(\lambda/(\lambda - 1))\alpha_j\beta_j' b_2 \\ &\quad + (\lambda I - \Gamma_1)b_2 - \Gamma_2 b_2\lambda^{-1} - \cdots - \Gamma_{k-1}b_2\lambda^{-k+2} = 0, \end{aligned}$$

which after multiplying with $\lambda - 1$ and rearranging becomes

$$\begin{aligned} &-\lambda\alpha_j\beta_j' b_2 + (\lambda - 1)\lambda b_2 \\ &\quad - (\lambda - 1)[\Gamma_1 + \Gamma_2\lambda^{-1} + \cdots + \Gamma_{k-1}\lambda^{-k+2}]b_2 = 0. \end{aligned}$$

This equals

$$\text{(A.12)}\quad \lambda^2 b_2 - \lambda(\alpha_j\beta_j' + I + \Gamma_1)b_2 + (\Gamma_1 - \Gamma_2)b_2 + \cdots + \lambda^{-k+3}(\Gamma_{k-2} - \Gamma_{k-1})b_2 + \lambda^{-k+2}\Gamma_{k-1}b_2 = 0.$$

Note, however that $A_j(z)$ can also be expressed as

$$\begin{aligned} A_j(z) &= (1-z)I - \Pi_j z - \sum_{i=1}^{k-1}\Gamma_i(1-z)z^i \\ &= I - (I + \Pi_j + \Gamma_1)z \\ &\quad - (\Gamma_2 - \Gamma_1)z^2 - \cdots - (\Gamma_{k-1} - \Gamma_{k-2})z^{k-1} + \Gamma_{k-1}z^k, \end{aligned}$$

which has the same form as (A.12). Hence, for eigenvalues of the matrix (A.9) not equal to 0 or 1, $\lambda^2 A_j(1/\lambda)b_2 = 0$, so that $1/\lambda$ is a solution of $\det[A_j(z)] = 0$, but as already pointed out, these eigenvalues all have modulus less than 1. Also there are $p(k-1)+j$ such eigenvalues. Hence, we have demonstrated that all solutions of $\det[A_j(z)] = 0$ that are not equal to 1, have modulus larger than 1.

That $\alpha_{j\perp}\Gamma\beta_{j\perp}$ is nonsingular for each $j = 0, \ldots, r-1$ follows from the necessary condition for stationarity in Theorem 4.2 in Johansen (1995). The argument is as follows. We start out with an $I(1)$ process that can be represented as described in Theorem 4.2 in Johansen (1995) using the reduced rank matrix Π. Then we consider new processes $X_{j,t}$ with reduced rank matrices $\Pi_j = \alpha_j\beta_j'$. They are not necessarily $I(1)$, but, as already argued, $(\beta_j' X_{j,t}, \Delta X_{j,t-1}, \ldots, \Delta X_{j,t-k+1})$ may be represented by an AR(1) process using (A.9), which has eigenvalues with moduli less than 1. Thus, $\beta_j' X_{j,t}$, $j = 0, \ldots, r-1$, are stationary processes and hence $\alpha_{j\perp}\Gamma\beta_{j\perp}$ is nonsingular for each $j = 0, \ldots, r-1$. *Q.E.D.*

REFERENCES

HANSEN, H., AND S. JOHANSEN (1999): "Some Tests for Parameter Constancy in Cointegrated VAR-Models," *Econometrics Journal*, 2, 306–333. [1711]

HARRIS, R. I. D., AND G. JUDGE (1998): "Small Sample Testing for Cointegration Using the Bootstrap Approach," *Economics Letters*, 58, 31–37. [1700]

JOHANSEN, S. (1992): "Determination of Cointegration Rank in the Presence of a Linear Trend," *Oxford Bulletin of Economics and Statistics*, 54, 383–397. [1700]

——— (1994): "The Role of the Constant and Linear Terms in Cointegration Analysis of Nonstationary Variables," *Econometric Reviews*, 13, 205–229. [1699,1703,1704]

——— (1995): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford, U.K.: Oxford University Press. [1699-1705,1709,1713]

——— (2002): "A Small Sample Correction for the Test of Cointegration Rank in the Vector Autoregressive Model," *Econometrica*, 70, 1929–1961. [1700,1707]

MANTALOS, P., AND G. SHUKUR (2001): "Bootstrapped Johansen Tests for Cointegration Relationships: A Graphical Analysis," *Journal of Statistical Computation and Simulation*, 68, 351–371. [1700]

OSTERWALD-LENUM, M. (1992): "A Note with Quantiles of the Asymptotic Distribution of the Maximum Likelihood Cointegration Rank Test Statistics," *Oxford Bulletin of Economics and Statistics*, 54, 461–471. [1707]

PAPARODITIS, E., AND D. N. POLITIS (2003): "Residual-Based Block Bootstrap for Unit Root Testing," *Econometrica*, 71, 813–855. [1702]

——— (2005): "Bootstraping Unit Root Tests for Autoregressive Time Series," *Journal of the American Statistical Association*, 100, 545–553. [1702]

PARK, J. Y. (2003): "Bootstrap Unit Root Tests," *Econometrica*, 71, 1845–1895. [1700]

PRESS, W. H., S. A. TEUKOLSKY, W. T. VETTERLING, AND B. P. FLANNERY (1992): *Numerical Recipes in FORTRAN* . The Art of Scientific Computing. (Second Ed.). Cambridge, U.K.: Cambridge University Press. [1707]

SIMS, C. A., J. H. STOCK, AND M. W. WATSON (1990): "Inference in Linear Time Series Models with Some Unit Roots," *Econometrica*, 58, 113–144. [1706]

SWENSEN, A. R. (2006): "Supplement to 'Bootstrap Algorithms for Testing and Determining the Cointegration Rank in VAR Models'," *Econometrica Supplementary Material*, 74, http://www.econometricsociety.org/ecta/supmat/5756proofs.pdf. [1704,1710]

VAN GIERSBERGEN, N. P. A. (1996): "Bootstrapping the Trace Statistics in VAR Models: Monte Carlo Results and Applications," *Oxford Bulletin of Economics and Statistics*, 58, 391–408. [1700]