

Giuseppe Cavaliere, Anders Rahbek,
A.M. Robert Taylor

Bootstrap Determination of the
Co-integration Rank in VAR models

Quaderni di Dipartimento

Serie Ricerche 2011, n. 9

ISSN 1973-9346



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Dipartimento di Scienze Statistiche “Paolo Fortunati”

Bootstrap Determination of the Co-integration Rank in VAR Models*

Giuseppe Cavaliere, Anders Rahbek and A.M.Robert Taylor

March 2011

Abstract

This paper discusses a consistent bootstrap implementation of the likelihood ratio [LR] co-integration rank test and associated sequential rank determination procedure of Johansen (1996). The bootstrap samples are constructed using the restricted parameter estimates of the underlying VAR model which obtain under the reduced rank null hypothesis. A full asymptotic theory is provided which shows that, unlike the bootstrap procedure in Swensen (2006) where a combination of unrestricted and restricted estimates from the VAR model is used, the resulting bootstrap data are $I(1)$ and satisfy the null co-integration rank, regardless of the true rank. This ensures that the bootstrap LR test is asymptotically correctly sized and that the probability that the bootstrap sequential procedure selects a rank smaller than the true rank converges to zero. Monte Carlo evidence suggests that our bootstrap procedures work very well in practice.

Keywords: Bootstrap; Co-integration; Trace statistic; Rank determination.

1 Introduction

Consider the case where the p -dimensional observations $\{X_t\}$ satisfy the k th order reduced rank vector autoregressive (VAR) model

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

with $\{\varepsilon_t\}$ independent and identically distributed (i.i.d.) with mean zero and full-rank variance matrix Ω , and where the initial values X_{1-k}, \dots, X_0 are fixed in the statistical analysis.

*We thank Iliyan Georgiev, Niels Haldrup, Bent Nielsen, Heino Bohn Nielsen, Søren Johansen and Anders Swensen for many useful discussions on this work. Parts of this paper were written while Cavaliere and Rahbek both visited CREATES whose hospitality is gratefully acknowledged. Rahbek is also affiliated with CREATES funded by the Danish National Research Foundation. An earlier version of this paper appeared under the title "Bootstrap Sequential Determination of the Co-integration Rank in VAR Models" as CREATES Research Paper 2010-7.

If: (a) the characteristic polynomial associated with (1) has $p - r$ roots equal to 1 and all other roots outside the unit circle, and (b) α and β have full column rank r , then the parameters in (1) will be said to satisfy the ‘ $I(1, r)$ conditions’. Under the above conditions X_t is $I(1)$ with co-integration rank r , such that the co-integrating relations $\beta'X_t - E(\beta'X_t)$ are stationary. The deterministic variables are assumed to satisfy one of the following cases (see, e.g., Johansen, 1996): (i) $d_t = 0$, $D_t = 0$ (no deterministic component); (ii) $D_t = 1$, $d_t = 0$ (restricted constant), or (iii) $D_t = t$, $d_t = 1$ (restricted linear trend).

The well-known (pseudo-) likelihood ratio [LR] test of Johansen (1996) for the hypothesis of co-integration rank (less than or equal to) r in (1), denoted $H(r)$, against $H(p)$, rejects for large values of the trace statistic, $Q_{r,T} := -T \sum_{i=r+1}^p \log(1 - \hat{\lambda}_i)$, where $\hat{\lambda}_1 > \dots > \hat{\lambda}_p$ are the largest p solutions to the eigenvalue problem,

$$(2) \quad |\lambda S_{11} - S_{10} S_{00}^{-1} S_{01}| = 0,$$

where $S_{ij} := T^{-1} \sum_{t=1}^T R_{it} R'_{jt}$, $i, j = 0, 1$, with R_{0t} and R_{1t} respectively denoting ΔX_t and $(X'_{t-1}, D_t)'$, corrected (by OLS) for $\Delta X_{t-1}, \dots, \Delta X_{t-k+1}$ and d_t . The sequential testing procedure based on $Q_{r,T}$ involves, starting with $r = 0$, testing in turn $H(r)$ against $H(p)$ for, $r = 0, \dots, p - 1$, until, for a given value of r , the asymptotic p -value associated with $Q_{r,T}$, exceeds a chosen (marginal) significance level. It is now well understood that the finite sample properties of the LR test, whether used sequentially or not, when based on asymptotic inference, can be quite poor; see Johansen (2002). In such circumstances, the bootstrap, when correctly implemented, can be an important device for improving upon the finite sample size properties of asymptotic tests; see, among others, van Giersbergen (1996), Swensen (2006) and Trenkler (2009), all of whom consider bootstrap procedures based around i.i.d. re-sampling, and Cavaliere *et al.* (2010a, 20010b) who use the wild bootstrap.

Of these bootstrap approaches the most complete treatment, particularly in terms of the underlying asymptotic theory, has been given by Swensen (2006). A key feature of the bootstrap algorithms proposed in Swensen (2006) is that they estimate (1) under both $H(r)$ and $H(p)$ and combine the parameter estimates from these to create the bootstrap sample data. Specifically, the bootstrap sample is generated using the estimates of α , β and ρ imposing reduced rank r , coupled with the unrestricted estimates of the short-run coefficients, Γ_j , $j = 1, \dots, k - 1$ and ϕ . Even though the latter are consistent irrespective of the true co-integration rank, it has been pointed out in Swensen (2009) that where the null hypothesis imposes a co-integration rank r smaller than the true rank, r_0 say, then so the potential arises for the resulting bootstrap samples to be explosive or to admit too many roots on the unit circle, thereby violating the $I(1, r)$ conditions and, hence, invalidating the use of the bootstrap. Moreover, where this happens it is not merely a finite sample problem and pertains even asymptotically. Swensen (2009) shows that to avoid this happening a number of auxiliary conditions must hold on the (unknown) parameters of the data generating process (DGP). These conditions rule out a non-empty set of perfectly plausible $I(1)$ DGPs.

In this paper we show that this problem can be solved by considering an alternative bootstrap scheme where the bootstrap data are generated using parameter estimates of α , β , ρ and $\Gamma_1, \dots, \Gamma_{k-1}, \phi$, all of which are obtained from estimating (1) under the null hypothesis, $H(r)$.¹ Although the resulting estimates of the dynamics are consistent only when $r = r_0$

¹This approach was mentioned in Remark 2 of Swensen (2006, pp.1701-02), although it was not pursued further there other than noting it as a possible alternative.

(i.e., where the null co-integration rank being tested coincides with the true rank), in Lemma 1 we establish the key result that, nonetheless, even when $r < r_0$ these estimates converge to pseudo-true values which ensure that the resulting bootstrap data are (at least in large samples) $I(1)$ with co-integrating rank r . As a consequence we are able, in Proposition 2, to prove that our bootstrap tests are asymptotically valid, attaining the same first-order limit null distribution as the original LR statistic both when $r = r_0$ and, crucially, when $r < r_0$, without the need for any auxiliary conditions to hold on the DGP. Given that the original LR statistic $Q_{r,T}$ diverges when $r < r_0$ this result ensures that the associated bootstrap analogue of Johansen's sequential procedure will be consistent in the usual sense that the probability of choosing a rank smaller than the true rank will converge to zero. We also report results from a small Monte Carlo study which suggest substantial improvements not only on the finite sample properties of the asymptotic LR test, but also relative to the corresponding bootstrap procedures of Swensen (2006) and the Bartlett-corrected LR tests of Johansen (2002).

The paper is organised as follows. In section 2 we outline our proposed bootstrap algorithms (Algorithms 1 and 2), demonstrating how these differ from Algorithms 1 and 2 in Swensen (2006). The large sample properties of the bootstrap procedure are established in section 3. The results of our Monte Carlo study are given in section 4. Section 5 concludes. Mathematical proofs are contained in the Appendix. In the following \xrightarrow{w} denotes weak convergence, \xrightarrow{p} convergence in probability, and $\xrightarrow{w_p}$ weak convergence in probability, in each case as $T \rightarrow \infty$; $\mathbb{I}(\cdot)$ denotes the indicator function; $x := y$ indicates that x is defined by y ; $\lfloor \cdot \rfloor$ denotes the integer part of its argument; I_k denotes the $k \times k$ identity matrix and $0_{j \times k}$ the $j \times k$ matrix of zeroes; the space spanned by the columns of any $m \times n$ matrix a is denoted as $\text{col}(a)$; if a is of full column rank $n < m$, then $\bar{a} := a(a'a)^{-1}$ and a_{\perp} is an $m \times (m - n)$ full column rank matrix satisfying $a'_{\perp} a = 0$; for any square matrix, a , $|a|$ is used to denote its determinant, $\|a\|$ the norm $\|a\|^2 := \text{tr}\{a'a\}$ and $\rho(a)$ its spectral radius (that is, the maximal modulus of the eigenvalues of a); for any vector, x , $\|x\|$ denotes the usual Euclidean norm, $\|x\| := (x'x)^{1/2}$. Finally, P^* denotes the bootstrap probability measure, i.e. conditional on the original sample; similarly, E^* denotes expectation under P^* .

2 Bootstrap Algorithms

The bootstrap implementations of the trace tests and associated sequential procedure from section 1 which we advocate are based around a bootstrap recursion which mimics the DGP in (1) under $H(r)$. To that end, let $\hat{v} := (\hat{v}_1, \hat{v}_2, \dots, \hat{v}_p)$ denote the eigenvectors from (2), viz,

$$(3) \quad \hat{v}' S_{11} \hat{v} = I_p, \quad \hat{v}' S_{10} S_{00}^{-1} S_{01} \hat{v} = \hat{\Lambda}_p := \text{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \dots, \hat{\lambda}_p).$$

The (uniquely defined) Gaussian QMLE of β , $\hat{\beta}^{(r)}$, may then be written as $\hat{\beta}^{(r)} := \hat{v} K_p^{(r)}$, where $K_p^{(r)} := (I_r, 0_{r \times (p-r)})'$, is a selection matrix indexed by r and p . When deterministic terms are included, $\hat{\beta}^{\#(r)} := (\hat{\beta}^{(r)'}, \hat{\rho}^{(r)'})' = \hat{v} K_{p+1}^{(r)}$. The remaining estimators $\hat{\alpha}^{(r)}, \hat{\Gamma}_1^{(r)}, \dots, \hat{\Gamma}_{k-1}^{(r)}$ and $\hat{\phi}^{(r)}$ are then obtained by OLS regression, as in Johansen (1996).

Using these estimates, the bootstrap algorithm we consider in this paper is then based on the recursion

$$(4) \quad \Delta X_{r,t}^* = \hat{\alpha}^{(r)} \hat{\beta}^{(r)'} X_{r,t-1}^* + \sum_{i=1}^{k-1} \hat{\Gamma}_i^{(r)} \Delta X_{r,t-i}^* + \hat{\alpha}^{(r)} \hat{\rho}^{(r)'} D_t + \hat{\phi}^{(r)} d_t + \varepsilon_{r,t}^*,$$

where the bootstrap shocks $\varepsilon_{r,t}^*$ in (4) are obtained by re-sampling from the corresponding restricted residuals obtained from estimating (1) under rank r .

The recursive scheme in (4) differs from the corresponding bootstrap recursion in Swensen (2006) which takes the form

$$(5) \quad \Delta X_{r,t}^* = \hat{\alpha}^{(r)} \hat{\beta}^{(r)'} X_{r,t-1}^* + \sum_{i=1}^{k-1} \hat{\Gamma}_i^{(p)} \Delta X_{r,t-i}^* + \hat{\alpha}^{(r)} \hat{\rho}^{(r)'} D_t + \hat{\phi}^{(p)} d_t + \varepsilon_{p,t}^*$$

where $\hat{\Gamma}_1^{(p)}, \dots, \hat{\Gamma}_{k-1}^{(p)}$ and $\hat{\phi}^{(p)}$ are now the estimates of the short run matrices $\Gamma_1, \dots, \Gamma_{k-1}$ and ϕ , respectively, from estimating (1) unrestrictedly, i.e. under $H(p)$. This difference is crucial since showing that the bootstrap test of $H(r)$ is consistent when $r < r_0$, requires that the bootstrap data still satisfy the $I(1, r)$ conditions in large samples.² As acknowledged in Swensen (2009), this is not guaranteed when using the recursion in (5), unless a number of auxiliary restrictions, labelled Assumption 2 in Swensen (2009), hold on the parameters of (1); see also Remark 6, below. In contrast, as we will formally establish in section 3, these restrictions are rendered redundant if the bootstrap recursion in (4) is used. This is because (4) always delivers an $I(1)$ system with $r \leq r_0$ co-integrating vectors in the limit, regardless of the true co-integration rank, r_0 .

We now detail in Algorithm 1 our bootstrap implementation of the LR test for $H(r)$ against $H(p)$.

ALGORITHM 1:

- (i) Estimate model (1) under $H(r)$ using Gaussian QMLE yielding the estimates $\hat{\beta}^{(r)}$, $\hat{\alpha}^{(r)}$, $\hat{\rho}^{(r)}$, $\hat{\Gamma}_1^{(r)}$, \dots , $\hat{\Gamma}_{k-1}^{(r)}$ and $\phi^{(r)}$, together with the corresponding residuals, $\hat{\varepsilon}_{r,t}$.
- (ii) Check that the equation $|\hat{A}^{(r)}(z)| = 0$, with $\hat{A}^{(r)}(z) := (1-z)I_p - \hat{\alpha}^{(r)}\hat{\beta}^{(r)'}z - \sum_{i=1}^{k-1} \hat{\Gamma}_i^{(r)}(1-z)z^i$, has $p-r$ roots equal to 1 and all other roots outside the unit circle. If so, proceed to step (iii).
- (iii) Construct the bootstrap sample recursively from (4) initialised at $X_{r,j}^* = X_j$, $j = 1-k, \dots, 0$, and with the T bootstrap errors $\varepsilon_{r,t}^*$ generated using the re-centred residuals, $\hat{\varepsilon}_{r,t}^c := \hat{\varepsilon}_{r,t} - T^{-1} \sum_{i=1}^T \hat{\varepsilon}_{r,i}$, for either:
 - (a) the i.i.d. bootstrap, such that $\varepsilon_{r,t}^* := \hat{\varepsilon}_{r,t}^c \mathcal{U}_t$, where \mathcal{U}_t , $t = 1, \dots, T$ is an i.i.d. sequence of discrete uniform distributions on $\{1, 2, \dots, T\}$, or
 - (b) the wild bootstrap, where for each $t = 1, \dots, T$, $\varepsilon_{r,t}^* := \hat{\varepsilon}_{r,t}^c w_t$, where w_t , $t = 1, \dots, T$, is an i.i.d.N(0,1) sequence.
- (iv) Using the bootstrap sample, $\{X_{r,t}^*\}$, and denoting by $\hat{\lambda}_1^* > \dots > \hat{\lambda}_p^*$ the ordered solutions to the bootstrap analogue of the eigenvalue problem in (2), compute the bootstrap LR statistic $Q_{r,T}^* := -T \sum_{i=r+1}^p \log(1 - \hat{\lambda}_i^*)$. Define the corresponding p -value as $p_{r,T}^* := 1 - G_{r,T}^*(Q_{r,T})$, $G_{r,T}^*(\cdot)$ denoting the conditional (on the original data) cdf of $Q_{r,T}^*$.

²A further difference between (5) and (4) is that the the bootstrap shocks $\varepsilon_{p,t}^*$ are re-sampled from the unrestricted residuals which result from estimating (1) under $H(p)$.

(v) The bootstrap test of $H(r)$ against $H(p)$ at level η rejects $H(r)$ if $p_{r,T}^* \leq \eta$.

REMARK 1. Although, as we show in section 3, Algorithm 1 ensures that the bootstrap data satisfy the $I(1, r)$ conditions in the limit, this could fail in small samples. Consequently, step (ii) checks that the bootstrap samples are indeed $I(1)$ with co-integration rank r . Analogous conditions are checked in step (iii) of Algorithm 1 in Swensen (2006) for the recursion in (5).

REMARK 2. A simplified version of the recursion in (4) can be obtained by excluding the estimated deterministic term, $\hat{\alpha}^{(r)} \hat{\rho}^{(r)'} D_t + \hat{\phi}^{(r)} d_t$, and initializing the recursion at $X_{r,1-k}^* = \dots = X_{r,0}^* = 0$. Due to the similarity of $Q_{r,T}$ with respect to ρ and asymptotic similarity with respect to ϕ and the initial values, these two approaches are asymptotically equivalent.

REMARK 3. The bootstrap recursion in (4) uses estimates of the short-run matrices, $\{\Gamma_j\}_{j=1}^{k-1}$, obtained under $H(r)$. These will therefore be more efficient than the unrestricted estimates used in (5) when $H(r)$ holds. Consequently, one would expect the bootstrap tests from (4) to display superior finite sample size properties to the tests obtained from (5); see Park (2003) for the univariate case. When $r < r_0$, although the estimates of $\{\Gamma_j\}_{j=1}^{k-1}$ used in (4) will now, in contrast to those from (5), be inconsistent this does not imply that the bootstrap tests from (5) will necessarily be more powerful than those from (4), since both bootstrap recursions are misspecified when $r < r_0$. Both of these predictions are supported by the Monte Carlo results presented in section 4.

REMARK 4. In practice, the cdf $G_{r,T}^*(\cdot)$ required in Step (iv) of Algorithm 1 will be unknown, but can be approximated in the usual way through numerical simulation; cf. Hansen (1996) and Andrews and Buchinsky (2000). This is achieved by generating B (conditionally) independent bootstrap statistics, $Q_{r,T:b}^*$, $b = 1, \dots, B$, computed as in Algorithm 1 above. The simulated bootstrap p -value is then computed as $\tilde{p}_{r,T}^* := B^{-1} \sum_{b=1}^B \mathbb{I}(Q_{r,T:b}^* > Q_{r,T})$, and is such that $\tilde{p}_{r,T}^* \xrightarrow{a.s.} p_{r,T}^*$ as $B \rightarrow \infty$.

We conclude this section by outlining in Algorithm 2 the bootstrap sequential algorithm for determining the co-integrating rank.

ALGORITHM 2: Starting from $r = 0$ perform the following steps:

(i)–(iv) Same as in Algorithm 1.

(v) If $p_{r,T}^*$ exceeds the significance level, η , set $\hat{r} = r$, otherwise repeat steps (i)–(iv) testing the null of rank $(r + 1)$ against rank p if $r + 1 < p$, or set $\hat{r} = p$ if $r + 1 = p$.

3 Asymptotic Analysis

In this section we prove that the bootstrap LR tests from Algorithm 1 are asymptotically valid, i.e. that they are asymptotically correctly sized under the null and consistent under the alternative. The results established in this section hold for any $I(1)$ DGP satisfying the assumptions stated in Swensen (2006). These coincide with those made for the standard asymptotic test in Johansen (1996), Assumptions 1 and 2(i) below, coupled with a finite fourth order moment condition and the additional assumption from Swensen (2006, Lemma 3), stated as Assumption 3 below, that eigenvalues from (2) are distinct in the limit. Crucially, we do not require any further conditions, such as Assumption 2 of Swensen (2009), to hold.

ASSUMPTION 1: The parameters in (1) satisfy the $I(1, r_0)$ conditions.

ASSUMPTION 2: (i) The innovations $\{\varepsilon_t\}$ in (1) form an i.i.d. sequence with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_t') = \Omega$, with Ω positive definite, and (ii) $E\|\varepsilon_t\|^4 \leq K < \infty$.

ASSUMPTION 3: The limiting non-zero roots of (2) are distinct.

Our first result concerns the limiting behaviour of the QMLE, used to generate the bootstrap samples in Algorithm 1. In Lemma 1 we show the key result that in the limit this satisfies the $I(1, r)$ conditions, even when an incorrect rank $r < r_0$ is imposed.

Lemma 1 *Let $\{X_t\}$ be generated as in (1) under Assumptions 1, 2(i) and 3. Furthermore, let $\hat{\theta}^{(r)} := \{\hat{\alpha}^{(r)}, \hat{\beta}^{(r)}, \hat{\Gamma}_1^{(r)}, \dots, \hat{\Gamma}_{k-1}^{(r)}, \hat{\rho}^{(r)}, \hat{\phi}^{(r)}, \hat{\Omega}^{(r)}\}$ denote the QMLE for (1) under $H(r)$. Then: (i) for any $r \leq r_0$ and as $T \rightarrow \infty$, $\hat{\theta}^{(r)} \xrightarrow{p} \theta_0^{(r)}$, with the vector of pseudo-true parameters, $\theta_0^{(r)} := \{\alpha_0^{(r)}, \beta_0^{(r)}, \Gamma_{1,0}^{(r)}, \dots, \Gamma_{k-1,0}^{(r)}, \rho_0^{(r)}, \phi_0^{(r)}, \Omega_0^{(r)}\}$, defined in the Appendix; and (ii) the pseudo-true parameters $\theta_0^{(r)}$ satisfy the $I(1, r)$ conditions.*

REMARK 5. Lemma 1 implies that, in sufficiently large samples, the estimates $\hat{\alpha}^{(r)}, \hat{\beta}^{(r)}, \hat{\Gamma}_1^{(r)}, \dots, \hat{\Gamma}_{k-1}^{(r)}$ satisfy the $I(1, r)$ conditions, even if r is lower than the true rank r_0 . As a consequence, in the limit, the root check in step (ii) of Algorithms 1 and 2 becomes redundant, since the parameters used in the bootstrap recursions will always satisfy the $I(1, r)$ conditions.

REMARK 6. It is important to note that analogous results to those given in Lemma 1 cannot be established for the corresponding estimates used in (5), as in Algorithm 1 of Swensen (2006); that is, where $\Gamma_1, \dots, \Gamma_{k-1}$ are estimated unrestrictedly under rank p . As recognized in Swensen (2009), the estimator there converges to $\theta_0^{(r,p)} := \{\alpha_0^{(r)}, \beta_0^{(r)}, \Gamma_{1,0}, \dots, \Gamma_{k-1,0}, \rho_0^{(r)}, \phi_0, \Omega_0\}$ which, in general, is not guaranteed to satisfy the $I(1, r)$ conditions. Indeed, Assumption 2 in Swensen (2009) requires $\theta_0^{(r,p)}$ to satisfy these conditions. To see this, note that to establish the result in Lemma 1(ii) one needs, in addition to $p - r$ unit roots, that $\rho(\Phi^{(r)}) < 1$, where $\Phi^{(r)}$ is defined in the proof of Lemma 1 by

$$(6) \quad \Phi^{(r)} := \begin{pmatrix} I_r + \beta_0^{(r)'} \alpha_0^{(r)} & \beta_0^{(r)'} \Gamma_{1,0}^{(r)} & \cdots & \cdots & \beta_0^{(r)'} \Gamma_{k-1,0}^{(r)} \\ \alpha_0^{(r)} & \Gamma_{1,0}^{(r)} & \cdots & \cdots & \Gamma_{k-1,0}^{(r)} \\ 0 & I_p & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_p & 0 \end{pmatrix}.$$

Assumption 2 of Swensen (2009) requires that $\rho(\Phi_{SW}^{(r)}) < 1$ for all $r < r_0$, with $\Phi_{SW}^{(r)}$ being defined as in (6) but with $\Gamma_{i,0}$ replacing $\Gamma_{i,0}^{(r)}$ ($i = 1, \dots, k - 1$). Thus, by definition, this is a re-statement that the parameters used in the bootstrap recursion (5) satisfy the $I(1, r)$ conditions in the limit for all $r < r_0$, which is clearly a strong assumption on the dynamics of the DGP in (1), since it is clear that there are non-empty parameter sets satisfying the $I(1, r_0)$ conditions, but for which $\rho(\Phi_{SW}^{(r)}) < 1$ would fail for $r < r_0$.

A direct implication of Lemma 1 for the bootstrap recursion in (4) is stated in the following proposition, which establishes that for any $r \leq r_0$ the bootstrap sample generated by (4) is $I(1)$ with co-integration rank r in large samples. This proposition holds irrespective of whether an i.i.d. or wild bootstrap re-sampling design is used.

Proposition 1 *Let $\{X_t\}$ be generated as in (1) under Assumptions 1,2 and 3, and let the bootstrap sample be generated as in Algorithm 1, for any $r \leq r_0$. Then it holds that*

$$(7) \quad X_{r,t}^* = \hat{C}^{(r)} \sum_{i=1}^t \varepsilon_{r,i}^* + \hat{\tau}_{r,t} + S_{r,t} T^{1/2},$$

where $P^*(\max_{t=1,\dots,T} \|S_{r,t}\| > \epsilon) \xrightarrow{P} 0$ for all $\epsilon > 0$. Furthermore, $T^{-1/2} \hat{C}^{(r)} \sum_{i=1}^{\lfloor Tu \rfloor} \varepsilon_{r,i}^* \xrightarrow{w} C_0^{(r)} W(u)$, $u \in [0, 1]$, where W is a p -dimensional Brownian motion with covariance matrix $\Omega_0^{(r)}$, $C_0^{(r)} := \beta_{0\perp}^{(r)} (\alpha_{0\perp}^{(r)'} \Gamma_0^{(r)} \beta_{0\perp}^{(r)})^{-1} \alpha_{0\perp}^{(r)'}$ is of rank $(p-r) \geq (p-r_0)$ and $\Gamma_0^{(r)} := \sum_{i=1}^{k-1} \Gamma_{0,i}^{(r)} - I_p$. If there are no deterministic, or a restricted constant in (1), then $\hat{\tau}_{r,t} = 0$; in the restricted linear trend case, $T^{-1} \hat{\tau}_{r,\lfloor Tu \rfloor} \xrightarrow{w} \tau_0^{(r)} u$, where $\tau_0^{(r)} := C_0^{(r)} \phi_0^{(r)} + (C_0^{(r)} \Gamma_0^{(r)} - I_p) \bar{\beta}_0^{(r)} \rho_0^{(r)'}$.

REMARK 7. The proof of Proposition 1 exploits the fact that, by Lemma 1(i), for any rank $r \leq r_0$ the bootstrap recursion in (4) coincides, in the limit, with the recursion $\Delta X_{r,t}^* = \alpha_0^{(r)} \beta_0^{(r)'} X_{r,t-1}^* + \sum_{i=1}^{k-1} \Gamma_{i,0}^{(r)} \Delta X_{t-i}^* + \alpha_0^{(r)} \rho_0^{(r)'} D_t + \phi_0^{(r)} d_t + \varepsilon_{r,t}^*$ which, by Lemma 1(ii), satisfies the $I(1, r)$ conditions. This property implies that the bootstrap sample is asymptotically $I(1)$ with r co-integrating relations.

Using Lemma 1 and Proposition 1, we now establish the asymptotic behaviour of the bootstrap trace statistic $Q_{r,T}^*$ of Algorithm 1. The stated results hold for any $r \leq r_0$.

Proposition 2 *Let the bootstrap statistic $Q_{r,T}^*$ be generated as in Algorithm 1. Then, under the conditions of Proposition 1, and for any $r \leq r_0$, $Q_{r,T}^* \xrightarrow{w} \text{tr}(\mathcal{Q}_{r,\infty})$, where $\mathcal{Q}_{r,\infty} := \int_0^1 dB_{p-r}(u) F_{p-r}(u)' (\int_0^1 F_{p-r}(u) F_{p-r}(u)' du)^{-1} \int_0^1 F_{p-r}(u) dB_{p-r}(u)'$, with B_{p-r} a $(p-r)$ dimensional standard Brownian motion, and where either: (i) in the no deterministic case, $F_{p-r} := B_{p-r}$; (ii) in the restricted constant case, $F_{p-r} := (B_{p-r}', 1)'$, or, (iii) in the restricted linear trend case, $F_{p-r} := (B_{p-r}', u|1)'$, where $a|b$ denotes the projection residuals of a onto b .*

REMARK 8. An immediate consequence of Proposition 2 is that the bootstrap test based on $Q_{r,T}^*$ will be asymptotically correctly sized under the null hypothesis ($r = r_0$), and will be consistent for all $r < r_0$. These two results follow using the results from Johansen (1996) that, under Assumptions 1 and 2(i), $Q_{r_0,T} \xrightarrow{w} \text{tr}(\mathcal{Q}_{r_0,\infty})$ while $Q_{r,T}$ diverges at rate T rate when $r < r_0$. In view of this, $p_{r_0,T}^* \xrightarrow{w} U[0, 1]$ and $p_{r,T}^* := 1 - G_{r,T}^*(Q_{r,T}) \xrightarrow{P} 0$, for all $r < r_0$.

We conclude this section by stating the following corollary of Proposition 2 which shows that the bootstrap sequential procedure in Algorithm 2 is consistent, replicating the first order asymptotic properties of the corresponding procedure based on the asymptotic LR tests given in Johansen (1996, Chapter 12).

Corollary 1 *Let \hat{r} denote the estimator of the co-integration rank as obtained in Algorithm 2. Then, under the conditions of Proposition 1: $\lim_{T \rightarrow \infty} P(\hat{r} = r) = 0$ for all $r = 0, 1, \dots, r_0 - 1$; $\lim_{T \rightarrow \infty} P(\hat{r} = r_0) = 1 - \eta \cdot \mathbb{I}(r_0 < p)$, and $\lim_{T \rightarrow \infty} \sup_{r \in \{r_0+1, \dots, p\}} P(\hat{r} = r) \leq \eta$.*

4 Numerical Results

Using Monte Carlo simulation we compare the finite sample performance of the bootstrap procedures from Algorithms 1 and 2 respectively, using the i.i.d. version of the re-sampling scheme (almost identical results were obtained using the wild bootstrap scheme and, hence, are not reported) in step (iii), with the corresponding asymptotic procedures of Johansen (1996), Bartlett-corrected procedures of Johansen (2002) and bootstrap procedures from Swensen (2006). As our simulation DGP we consider the VAR(2) process of dimension $p = 4$,

$$(8) \quad \Delta X_t = \alpha \beta' X_{t-1} + \Gamma_1 \Delta X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. } N(0, I_4) \quad (t = 1, \dots, T)$$

with $X_0 = \Delta X_0 = 0$, and $T \in \{50, 100, 200\}$. The long-run parameter vectors are set to $\beta := (1, 0, 0, 0)'$, $\alpha := (a, 0, 0, 0)'$. All experiments were run over 10,000 Monte Carlo replications. For the bootstrap tests, any replications violating the root check conditions (step (ii) in Algorithms 1 and 2 and step (iii) in Algorithms 1 and 2 of Swensen, 2006) were discarded and the experiment continued until 10,000 valid replications were obtained. For each bootstrap procedure we report the frequency with which such violations occurred.³

We first report results for the case of a single co-integration vector, setting $a = -0.4$ and

$$\Gamma_1 := \begin{bmatrix} \gamma & \delta & 0 & 0 \\ \delta & \gamma & 0 & 0 \\ 0 & 0 & \gamma & 0 \\ 0 & 0 & 0 & \gamma \end{bmatrix}$$

with $\gamma = 0.8$ and $\delta \in \{0, 0.1, 0.2, 0.3\}$. For all of these parameter combinations, X_t is $I(1)$ with co-integrating rank $r_0 = 1$. The role of the parameter δ is to isolate the violation or otherwise of the auxiliary conditions given in Assumption 2 of Swensen (2009); in particular, these conditions are satisfied only for $\delta = 0$ or $\delta = 0.1$; cf. Remark 6.⁴

Empirical rejection frequencies [ERF] of the tests for $r = 1$ are reported in Panel (a) of Table I for tests run at the nominal 5% level. It is seen from these results that the standard asymptotic test for $r = 1$, $Q_{1,T}$, displays very poor finite sample size control with ERFs of around 45% for $T = 50$, improving somewhat to around 13% for $T = 200$. In contrast, the ERFs of our proposed bootstrap test, $Q_{1,T}^*$, all lie very close to the nominal 5% level, even for $T = 50$. In line with the prediction in Remark 3, $Q_{1,T}^*$ also appears to be rather better sized than the corresponding bootstrap test of Swensen (2006), $Q_{1,T}^{*SW}$, the latter being too liberal. The $Q_{1,T}^*$ test also displays better size control than the Bartlett-corrected LR test, $Q_{1,T}^{BC}$, which is somewhat undersized for the smaller sample sizes.

Our proposed bootstrap test also performs well in terms of empirical power. The ERFs of the bootstrap test for $r = 0$, $Q_{0,T}^*$, are nowhere smaller than those of the $Q_{0,T}^{BC}$ test, with the most significant gains seen for the smaller sample sizes considered. Moreover, in most cases,

³All computations were performed in Gauss 9.0, except the Bartlett corrected statistics which were computed in Ox using programs kindly supplied to us by Heino Bohn Nielsen. The Gauss procedure for computing the bootstrap algorithms is available from the authors upon request.

⁴Specifically, when rank $r = 0$ is tested, it must hold that $\rho(\Phi_{SW}^{(0)}) < 1$ where $\Phi_{SW}^{(0)} = \Gamma_1$. For $\delta = 0$ and $\delta = 0.1$ this is indeed the case. However, for $\delta = 0.2$, $\rho(\Phi_{SW}^{(0)}) = \rho(\Gamma_1) = 1.0$, while for $\delta = 0.3$, $\rho(\Gamma_1) = 1.1$. In the former case the bootstrap DGP will be integrated of order two in the limit, while for the latter it will be explosive.

the power of the bootstrap $Q_{0,T}^*$ test is not lower than that of Swensen's bootstrap $Q_{0,T}^{*SW}$ test. The only exceptions occur for $\delta = 0, 0.1$ when $T = 50$, where the empirical power of $Q_{0,T}^*$ is slightly inferior to that of $Q_{0,T}^{*SW}$. However, the latter is strongly oversized under the null here, having size about 19% when $\delta = 0$ and $r = 0$; cf. Panel (b) of Table I for $\gamma = 0.8$.

The associated results for the sequential procedures are reported in Table II. Among the two bootstrap algorithms considered, our Algorithm 2 appears to have the best performance in terms of its ability to select the true co-integration rank, $r_0 = 1$. As expected, our Algorithm 2 is not affected by the value of δ , whereas contrastingly the behaviour of Swensen's sequential Algorithm 2 is clearly heavily affected by the value of δ . For $T = 50$, Algorithm 2 is also superior to the Bartlett-corrected procedure, although the reverse appears to be the case for $T = 100$; this latter result is, however, largely an artefact of the undersizing observed in $Q_{1,T}^{BC}$.

We now compare the frequency with which the bootstrap recursions fail to generate valid $I(1)$ bootstrap samples. Taking the sequential procedures to illustrate, the fraction of times our Algorithm 2 generates explosive bootstrap samples is remarkably small; in particular, it never exceeds 0.3%, 0.1% and 0.05% for $T = 50, 100$ and 200 , respectively. In contrast, Algorithm 2 of Swensen (2006) displays a higher number of failures of the $I(1, r)$ conditions, even when it is asymptotically valid ($\delta = 0.0$ or $\delta = 0.1$). For instance, when $T = 50$ ($T = 100$) and $\delta = 0.1$, explosive bootstrap samples are generated 8.4% (2.7%) of the time. For $I(1)$ DGPs with $\delta \geq 0.2$, this failure rate increases dramatically; e.g. when $\delta = 0.3$ explosive samples are generated 49.5% (91.6%) of the time for $T = 50$ ($T = 200$).

To conclude this section we consider the case where $a = 0$, so that the DGP reduces to the non-co-integrated VAR(2), $\Delta X_t = \Gamma \Delta X_{t-1} + \varepsilon_t$, $t = 1, \dots, T$. As in Johansen (2002, section 3.1), we set $\Gamma_1 := \gamma I_4$, so that the $I(1, r)$ conditions are met with $r = 0$, provided $|\gamma| < 1$. The empirical size of the various trace tests for $r = 0$ are reported for $\gamma \in \{0.0, 0.5, 0.8, 0.9\}$ in Table I, Panel (b). As γ increases our $Q_{0,T}^*$ test is clearly preferable to the corresponding $Q_{0,T}^{*SW}$ test, again in line with the prediction from Remark 3. Interestingly, for large γ and small T the bootstrap test tends to be slightly oversized whereas (in line with Johansen, 2002, Table V), the Bartlett correction over-corrects, so the corresponding $Q_{0,T}^{BC}$ test is undersized.

5 Conclusions

In this paper we have discussed bootstrap implementations of the likelihood ratio co-integration rank test and associated sequential procedure of Johansen (1996) based on restricted estimates of the underlying VAR model. We have shown that, in contrast to what has been established for the bootstrap procedures advocated in Swensen (2006), this approach is asymptotically correctly sized and consistent. A small Monte Carlo experiment suggested that this procedure works very well in finite samples, outperforming not only asymptotic-based procedures, but also the corresponding procedures from Swensen (2006). Finally, although our analysis was based on the assumption of i.i.d. errors, the extension to the case of martingale difference sequences with heteroskedasticity of unknown form is straightforward using the results obtained in Cavaliere, Rahbek and Taylor (2010a,b).

Department of Statistical Sciences, University of Bologna; giuseppe.cavaliere@unibo.it
Department of Economics, University of Copenhagen; anders.rahbek@econ.ku.dk
School of Economics, University of Nottingham; robert.taylor@nottingham.ac.uk

Appendix

PROOF OF LEMMA 1. Let $\alpha_0, \beta_0, \Psi_0 := (\Gamma_{1,0}, \dots, \Gamma_{k-1,0}), \rho_0, \phi_0$ and Ω_0 denote the true parameters in (1). We first present a convenient normalisation of the co-integration parameters, which allows us to prove part (i) of the Lemma, convergence to the pseudo-true parameter vector $\theta_0^{(r)}$. We then prove part (ii) by establishing that $\theta_0^{(r)}$ does indeed satisfy the $I(1, r)$ conditions for $r \leq r_0$. The proofs are first obtained for the case (i) of no deterministic and later generalized to the cases of (ii) a restricted constant and (iii) a restricted linear trend.

Normalisation. By Johansen (1996), the r_0 largest eigenvalues $(\hat{\lambda}_i)_{i=1, \dots, r_0}$ from (2) satisfy, as $T \rightarrow \infty$, the population eigenvalue problem, $|\lambda \Sigma_{\beta\beta} - \Sigma_{\beta 0} \Sigma_{00}^{-1} \Sigma_{0\beta}| = 0$, where $\Sigma_{\beta\beta} := \text{Var}(\beta' X_{t-1} | \Delta \mathbb{X}_{2t}), \Sigma_{00} := \text{Var}(\Delta X_t | \Delta \mathbb{X}_{2t}), \Sigma_{\beta 0} := \text{Cov}(\beta' X_{t-1}, \Delta X_t | \Delta \mathbb{X}_{2t})$ and $\Delta \mathbb{X}_{2t} := (\Delta X'_{t-1}, \dots, \Delta X'_{t-k-1})'$. Let $\kappa := (\kappa_1, \dots, \kappa_{r_0})$ denote eigenvectors corresponding to the eigenvalues $\lambda_1 > \lambda_2 > \dots > \lambda_{r_0} > 0$, such that $\kappa' \Sigma_{\beta\beta} \kappa = I_{r_0}$. We can then define $\beta_0 := \beta \kappa$ and $\alpha_0 := \alpha (\kappa')^{-1}$. Observe that, $\alpha \beta' = \alpha_0 \beta_0'$, while also

$$(A.1) \quad \Sigma_{\beta_0 \beta_0} = I_{r_0} \quad \text{and} \quad \Sigma_{\beta_0 0} \Sigma_{00}^{-1} \Sigma_{0 \beta_0} = \text{diag}(\lambda_1, \dots, \lambda_{r_0}),$$

with $\Sigma_{\beta_0 \beta_0} := \text{Var}(\beta_0' X_{t-1} | \Delta \mathbb{X}_{2t})$ and $\Sigma_{\beta_0 0} := \text{Cov}(\beta_0' X_{t-1}, \Delta X_t | \Delta \mathbb{X}_{2t})$. Indeed, the relations in (A.1) are the population equivalents of the sample normalisations in (3).

Convergence to pseudo-true values. First, with $\hat{\beta} = \hat{\beta}^{(r)}$ the QMLE under the true rank r_0 , then as in Johansen (1996, proof of Lemma 13.1), $\bar{\beta}_0'(\hat{\beta} - \beta_0) \xrightarrow{p} 0$ and $T^{1/2} \beta_{0\perp}'(\hat{\beta} - \beta_0) \xrightarrow{p} 0$. Therefore, since $\hat{\beta}^{(r)} = \hat{v} K_p^{(r)} = \hat{\beta} K_{r_0}^{(r)}$, it follows that, using continuity of the eigenvectors and eigenvalues (as the latter are distinct by Assumption 3),

$$(A.2) \quad \bar{\beta}_0'(\hat{\beta}^{(r)} - \beta_0^{(r)}) \xrightarrow{p} 0 \quad \text{and} \quad T^{1/2} \beta_{0\perp}'(\hat{\beta}^{(r)} - \beta_0^{(r)}) \xrightarrow{p} 0$$

where $\beta_0^{(r)} := \beta_0 K_{r_0}^{(r)}$. Next, consider the QMLE $\hat{\alpha}^{(r)}$:

$$(A.3) \quad \hat{\alpha}^{(r)} = S_{01} \hat{\beta}^{(r)} = S_{01} (\beta_0 \bar{\beta}_0' + \beta_{0\perp} \bar{\beta}_{0\perp}') \hat{\beta}^{(r)} \xrightarrow{p} \Sigma_{0\beta_0} K_{r_0}^{(r)} = \alpha_0 K_{r_0}^{(r)} =: \alpha_0^{(r)}.$$

Regarding $\hat{\Psi}^{(r)}$, with $\Upsilon_{02} := \text{Cov}(\Delta X_t, \Delta \mathbb{X}_{2t}), \Upsilon_{\beta_0 2} := \text{Cov}(\beta_0' X_{t-1}, \Delta \mathbb{X}_{2t}), \Upsilon_{22} := \text{Var}(\Delta \mathbb{X}_{2t})$ and, similarly, $M_{02} := T^{-1} \sum_{t=1}^T \Delta X_t \Delta \mathbb{X}_{2t}', M_{12} := T^{-1} \sum_{t=1}^T X_{t-1} \Delta \mathbb{X}_{2t}'$, we have

$$\hat{\Psi}^{(r)} = \left(M_{02} - \hat{\alpha}^{(r)} \hat{\beta}^{(r)'} M_{12} \right) M_{22}^{-1} \xrightarrow{p} \Psi_0^{(r)} := \left(\Upsilon_{02} - \alpha_0 K_{r_0}^{(r)} K_{r_0}^{(r)'} \Upsilon_{\beta_0 2} \right) \Upsilon_{22}^{-1}$$

with $K_{r_0, \perp}^{(r)} = (0, I_{r_0-r})'$. Using the projection identity, $I_r = K_{r_0, \perp}^{(r)} K_{r_0, \perp}^{(r)'} + K_{r_0}^{(r)} K_{r_0}^{(r)'}$, which will be applied repeatedly in this Appendix, it follows that

$$(A.4) \quad \Psi_0^{(r)} = (\Upsilon_{02} - \alpha_0 \Upsilon_{\beta_0 2}) \Upsilon_{22}^{-1} + \alpha_0 K_{r_0, \perp}^{(r)} K_{r_0, \perp}^{(r)'} \Upsilon_{\beta_0 2} \Upsilon_{22}^{-1} = \Psi_0 + \alpha_0 K_{r_0, \perp}^{(r)} K_{r_0, \perp}^{(r)'} \Upsilon_{\beta_0 2} \Upsilon_{22}^{-1}$$

as $(\Upsilon_{02} - \alpha_0 \Upsilon_{\beta_0 2}) \Upsilon_{22}^{-1} = \Psi_0$. Finally, for $\hat{\Omega}^{(r)}$ it holds that

$$(A.5) \quad \hat{\Omega}^{(r)} = S_{00} - \hat{\alpha}^{(r)} \hat{\alpha}^{(r)'} \xrightarrow{p} \Omega_0^{(r)} := \Sigma_{00} - \alpha_0 K_{r_0}^{(r)} K_{r_0}^{(r)'} \alpha_0' = \Omega_0 + \alpha_0 K_{r_0, \perp}^{(r)} K_{r_0, \perp}^{(r)'} \alpha_0' > 0.$$

Pseudo true values satisfy the $I(1, r)$ conditions, with $r \leq r_0$. Rewrite the DGP as

$$(A.6) \quad \Delta X_t = \alpha_0^{(r)} \beta_0^{(r)'} X_{t-1} + \Psi_0^{(r)} \Delta \mathbb{X}_{2t} + \varepsilon_{r,t},$$

see (A.3)–(A.4), with $\varepsilon_{r,t}$ given by

$$(A.7) \quad \varepsilon_{r,t} = \varepsilon_t + \alpha_0 K_{r_0, \perp}^{(r)} K_{r_0, \perp}^{(r)'} (\beta_0' X_{1t-1} - \Upsilon_{\beta_0 2} \Upsilon_{22}^{-1} \Delta \mathbb{X}_{2t}).$$

Observe that $\beta_0^{(r)'} X_{t-1}$ and $\Delta \mathbb{X}_{2t}$ in (A.6) are uncorrelated with $\varepsilon_{r,t}$. This property, which is central to the proof of Lemma 1(ii), is a vital difference between our Algorithm 1 and Algorithm 2 of Swensen (2006); in the latter, $\Psi_0^{(r)}$ is replaced by Ψ_0 (this resulting from the use of the unrestricted estimator of Ψ_0 in the bootstrap recursion) and the uncorrelatedness result stated above no longer holds.

With $\mathbb{X}_t := (X_t', \dots, X_{t-k+1}')'$ the system can be written in companion form as

$$(A.8) \quad \Delta \mathbb{X}_t = \mathbb{A}^{(r)} \mathbb{B}^{(r)'} \mathbb{X}_{t-1} + \mathbb{E}_t^{(r)}$$

with $\mathbb{E}_t^{(r)} := (\varepsilon_{r,t}', 0, \dots, 0)'$, \mathbb{X}_0 fixed, and $\Psi_0^{(r)} = (\Gamma_{1,0}^{(r)}, \dots, \Gamma_{k-1,0}^{(r)})$,

$$(A.9) \quad \mathbb{A}^{(r)} := \begin{pmatrix} \alpha_0^{(r)} & \Psi_0^{(r)} \\ 0 & I_{p(k-1)} \end{pmatrix}, \quad \mathbb{B}^{(r)} := \begin{pmatrix} \beta_0^{(r)} & I_p & 0 & \cdots & 0 \\ 0 & -I_p & I_p & \cdots & 0 \\ 0 & 0 & -I_p & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -I_p \end{pmatrix}.$$

Observe first, that by Assumption 1, $\mathbb{Y}_t := \mathbb{B}^{(r)'} \mathbb{X}_t$ is covariance stationary with covariance $\Sigma_{YY} > 0$, which from (A.8), using the aforementioned uncorrelatedness, is the solution to,

$$(A.10) \quad \Sigma_{YY} = \Phi^{(r)} \Sigma_{YY} \Phi^{(r)'} + \Sigma_{EE},$$

where $\Phi^{(r)} = (I_{r+p(k-1)} + \mathbb{B}^{(r)'} \mathbb{A}^{(r)})$ and $\Sigma_{EE} = \text{Var}(\mathbb{B}^{(r)'} \mathbb{E}_t^{(r)})$. From the definition of $\varepsilon_{r,t}$ in (A.7), $\text{Var}(\varepsilon_{r,t}) > 0$ and, as $\mathbb{B}^{(r)'} \mathbb{E}_t^{(r)} = (\varepsilon_{r,t}' \beta_0^{(r)}, \varepsilon_{r,t}', 0, \dots, 0)'$ it follows that $\Sigma_{EE} \geq 0$, with $V' \Sigma_{EE} V = 0$, and $V_{\perp}' \Sigma_{EE} V_{\perp} > 0$, where

$$(A.11) \quad V = \begin{pmatrix} I_r & 0 & \cdots & 0 \\ -\beta_0^{(r)} & 0 & \cdots & 0 \\ 0 & I_p & & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 0 & \cdots & 0 & I_p \end{pmatrix}, \quad V_{\perp} = \begin{pmatrix} \beta_0^{(r)'} \\ I_p \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

As $\Phi^{(r)}$ solves (A.10), and since $\Sigma_{YY} > 0$ and $\Sigma_{EE} \geq 0$, the spectral radius of $\Phi^{(r)}$ satisfies $\rho(\Phi^{(r)}) \leq 1$. Suppose, for λ an eigenvalue of $\Phi^{(r)}$, $|\lambda| = 1$. Then, using (A.10), the corresponding eigenvector v would be in the space spanned by V , $v \in \text{col}(V)$. However, as

$$\Phi^{(r)'} V = (I_{r+p(k-1)} + \mathbb{A}^{(r)'} \mathbb{B}^{(r)}) V = \begin{pmatrix} I_r & 0 & \cdots & 0 \\ 0 & I_p & & \vdots \\ 0 & 0 & \ddots & \\ \vdots & & \ddots & I_p \\ 0 & \cdots & & 0 \end{pmatrix},$$

any $v \in \text{col}(V)$ is not an eigenvector of $\Phi^{(r)'}$. Thus, we conclude that $\rho(\Phi^{(r)}) < 1$.

Finally, consider the eigenvalue problem,

$$(A.12) \quad \left| (1-z)I_{pk} - \mathbb{A}^{(r)}\mathbb{B}^{(r)'}z \right| = 0.$$

Observe that $|\mathbb{B}^{(r)'}\mathbb{A}^{(r)}| \neq 0$, since $\rho(\Phi^{(r)}) < 1$ and $\mathbb{B}^{(r)'}\mathbb{A}^{(r)} = \Phi^{(r)} - I_{r+p(k-1)}$. Next, the full rank of $\mathbb{B}^{(r)'}\mathbb{A}^{(r)}$ implies, with $\mathbb{A}_\perp^{(r)} = (I_p, -\Psi_0^{(r)})'\alpha_{0,\perp}^{(r)}$, that $N := (\mathbb{B}^{(r)}, \mathbb{A}_\perp^{(r)})$ has full rank, since $|\mathbb{B}^{(r)'}\mathbb{A}^{(r)}| \neq 0$ implies $|(\mathbb{A}^{(r)}, \mathbb{A}_\perp^{(r)})'N| \neq 0$. Multiplying (A.12) from the left and right by N and its transpose, respectively, shows that there are $(p-r)$ roots at $z=1$, while the remaining satisfy $|z| > 1$ since $\rho(\Phi^{(r)}) < 1$. This completes the proof of Lemma 1 for the case of no deterministic terms.

Case (ii) of a restricted constant: In this case, replace β by $\beta^\#$, and X_{t-1} by $X_{t-1}^\#$, with $X_{t-1}^\# = (X_{t-1}', 1)'$ and $\beta^\# = (\beta', \rho)'$. Define $\gamma_0^\# := (\beta_{0\perp}', 0)'$ and $\xi_0^\# := (-\rho_0\bar{\beta}_0', 1)'$ such that $(\beta_0^\#, \gamma_0^\#, \xi_0^\#)$ spans $\mathcal{R}^{p^\#}$, $p^\# = p+1$. As for the case of no deterministic terms, it holds that $\bar{\beta}_0^{\#'}(\hat{\beta}^\# - \beta_0^\#)' \xrightarrow{p} 0$, $(T^{1/2}\bar{\gamma}_0^\#, \bar{\xi}_0^\#)'(\hat{\beta}^\# - \beta_0^\#)' \xrightarrow{p} 0$. Therefore, as $\hat{\beta}^{\#(r)} = \hat{v}K_{p^\#}^{(r)} = \hat{\beta}^\#K_{r_0}^{(r)}$, we find,

$$\bar{\beta}_0^{\#'}(\hat{\beta}^{\#(r)} - \beta_0^{\#(r)})' \xrightarrow{p} 0, \quad \text{and} \quad (T^{1/2}\bar{\gamma}_0^{\#}, \bar{\xi}_0^{\#})'(\hat{\beta}^{\#(r)} - \beta_0^{\#(r)})' \xrightarrow{p} 0.$$

The results then follow proceeding as in the case of no deterministic terms. Specifically, in (A.3)-(A.7) replace β_0 by $\beta_0^\#$, and $\beta_0^{(r)'}X_{t-1}$ by $\beta_0^{\#(r)'}X_{t-1}^\#$. For (A.8) and onwards, while $\mathbb{A}^{(r)}$ and $\mathbb{B}^{(r)}$ remain unaltered, $\mathbb{E}_t^{(r)}$ should be replaced by $\mathbb{E}_t^{(r)} + \mathbb{A}^{(r)}\mathbb{R}^{(r)'}$,

$$(A.13) \quad \mathbb{R}^{(r)} := (\rho_0^{(r)}, 0, \dots, 0).$$

Case (iii) of a restricted linear trend: Here, $\beta^\# = (\beta', \rho)'$, while $X_{t-1}^\# = (X_{t-1}', t)'$ and $d_t = 1$ enters in (1). For the asymptotic analysis in this case, Rahbek *et al.* (1999, proof of Theorem 4.2) apply the non-orthogonal basis $(\beta_0^\#, \gamma_0^\#, \xi_0^\#)$ for $\mathcal{R}^{p^\#}$, where $\gamma_0^\# := (\beta_{0\perp}', -\gamma_0')'$ and $\xi_0^\# := (0, 1)'$, with $\gamma_0 := \beta_{0\perp}'C_0(\phi_0 + \Gamma_0\bar{\beta}_0\rho_0')$ and $C_0 := \beta_{0\perp}(\alpha_{0\perp}'\Gamma_0\beta_{0\perp})^{-1}\alpha_{0\perp}'$. With, $\hat{\beta}^\# - \beta_0^\# := \beta_0^{\#}\hat{b} + \gamma_0^{\#}\hat{b}_\gamma + \xi_0^{\#}\hat{b}_\xi$, then by Rahbek *et al.* (1999), $\hat{b} = o_p(1)$, $\hat{b}_\gamma = o_p(T^{1/2})$, $\hat{b}_\xi = o_p(T)$, such that $\hat{\beta}^{\#(r)} = \hat{\beta}^\#K_{r_0}^{(r)}$ converges as desired. Proceed as in the case of a restricted constant, replacing β_0 by $\beta_0^\#$ in (A.3)-(A.7), using $\Delta\mathbb{X}_{2t}$ corrected by a constant by OLS in the definitions of M_{i2} , $i=0,1,2$, and correspondingly, in (A.7), subtracting $\alpha_0K_{r_0,\perp}^{(r)'}K_{r_0,\perp}^{(r)}E(\beta^{\#'}X_{t-1} - \Upsilon_{\beta_0 2}\Upsilon_{22}^{-1}\Delta\mathbb{X}_{2t})$. Finally, add the term $\mathbb{A}^{(r)}\mathbb{R}^{(r)'}t + \mathbb{F}^{(r)}$ in (A.8), with $\mathbb{R}^{(r)}$ defined in (A.13) and $\mathbb{F}^{(r)} := (\phi_0^{(r)'}, 0, \dots, 0)'$, such that the covariance stationary process \mathbb{Y}_t is defined by $\mathbb{Y}_t := \mathbb{B}^{(r)'}\mathbb{X}_t + \mathbb{R}^{(r)'}t$. \square

PROOF OF PROPOSITION 1: Consider first the case (i) of *no deterministic terms*: For $r = r_0$ the result is established in Lemma A.4 of Cavaliere *et al.* (2010a). Next, for $r < r_0$, set $\mathbb{X}_{r,t}^* := (X_{r,t}^*, \dots, X_{r,t-k+1}^*)'$ and $\mathbb{X}_{r,0}^* := \mathbb{X}_0$, and use the companion form in (A.8) with $\varepsilon_{r,t}$ replaced by $\varepsilon_{r,t}^*$, to see directly that $X_{r,t}^* = (I_p, 0, \dots, 0)\mathbb{X}_{r,t}^*$ has the representation,

$$(A.14) \quad X_{r,t}^* = \hat{C}^{(r)}\sum_{i=1}^t \varepsilon_{r,t}^* + S_{r,t}T^{1/2},$$

where $\hat{C}^{(r)} = \hat{\beta}_\perp^{(r)} \left(\hat{\alpha}_\perp^{(r)'} \hat{\Gamma}^{(r)} \hat{\beta}_\perp^{(r)} \right)^{-1} \hat{\alpha}_\perp^{(r)'}$, $\hat{\Gamma}^{(r)} = \sum_{i=1}^{k-1} \hat{\Gamma}_i^{(r)} - I$, and $S_{r,t} = S_{r,t}^1 + S_{r,t}^2$, with

$$(A.15) \quad S_{r,t}^1 = (\hat{\alpha}^{(r)}, \hat{\Psi}^{(r)}) (\hat{\mathbb{B}}' \hat{\mathbb{A}})^{-1} \sum_{i=0}^{t-1} \hat{\Phi}^i (T^{-1/2} \hat{\mathbb{B}}' e_{r,t-i}^*),$$

$$(A.16) \quad S_{r,t}^2 = T^{-1/2} [\hat{C}^{(r)} (I_p, -\hat{\Psi}^{(r)}) + (\hat{\alpha}^{(r)}, \hat{\Psi}^{(r)}) (\hat{\mathbb{B}}' \hat{\mathbb{A}}) \hat{\Phi}^t \hat{\mathbb{B}}'] \mathbb{X}_{r,t}^*.$$

Here $\hat{\Phi} := (I_{r+p(k-1)} + \hat{\mathbb{B}}' \hat{\mathbb{A}})$ with the matrices $\hat{\mathbb{A}}$ and $\hat{\mathbb{B}}$ defined as $\mathbb{A}^{(r)}$ and $\mathbb{B}^{(r)}$ of (A.9) with $\alpha_0^{(r)}$ and $\beta_0^{(r)}$ replaced by the corresponding estimators $\hat{\alpha}^{(r)}$, $\hat{\beta}^{(r)}$, and $e_{r,t}^* := (\varepsilon_{r,t}^{*'}, 0, \dots, 0)'$. Consider first $S_{r,t}^1$. Note that

$$\max_{t=1, \dots, T} \|S_{r,t}^1\| \leq \max_{t=1, \dots, T} \|(\hat{\alpha}^{(r)}, \hat{\Psi}^{(r)}) (\hat{\mathbb{B}}' \hat{\mathbb{A}})^{-1} \sum_{i=0}^{t-1} \hat{\Phi}^i (T^{-1/2} \hat{\mathbb{B}}' e_{r,t-i}^*)\| \leq \psi_T \max_{t=1, \dots, T} \|T^{-1/2} \eta_t^*\|,$$

where $\eta_t^* = \hat{\mathbb{B}}' e_{r,t}^* = \left(\hat{\beta}^{(r)}, I_p, 0, \dots, 0 \right)' \varepsilon_{r,t}^*$ and $\psi_T = \|(\hat{\alpha}^{(r)}, \hat{\Psi}^{(r)}) (\hat{\mathbb{B}}' \hat{\mathbb{A}})^{-1}\| \|\sum_{i=0}^{T-1} \hat{\Phi}^{(r)i}\|$. It follows that $\psi_T \xrightarrow{p} \psi = \left\| (\alpha_0^{(r)}, \Psi_0^{(r)}) (\mathbb{B}^{(r)'} \mathbb{A}^{(r)})^{-1} \right\| \left\| (\mathbb{B}^{(r)'} \mathbb{A}^{(r)})^{-1} \right\|$ by using the consistency of the estimators established in Lemma 1 and the definition of $\hat{\Phi}$. In particular, note that for sufficiently large T we have, by continuity, that $\rho(\hat{\Phi}) < 1$, which implies that $\|\hat{\Phi}^i\| \leq \text{const} \cdot \rho^i$ for some $0 < \rho < 1$, uniformly over i . Finally, showing that $P^* (\max_{t=1, \dots, T} \|T^{-1/2} \eta_t^*\| > \eta)$ is of order $o_p(1)$ implies $P^* (\max_{t=1, \dots, T} \|S_{r,t}^1\| > \eta) \xrightarrow{p} 0$. This holds by showing that $P^* (T^{-1/2} \max_{t=1, \dots, T} \|\varepsilon_{r,t}^*\| > \eta) = o_p(1)$. For the wild bootstrap, observe that with $\bar{\varepsilon}_r := T^{-1} \sum_{t=1}^T \hat{\varepsilon}_{r,t}$, $\mathbb{E}^* (\varepsilon_{r,t}^{*'} \varepsilon_{r,t}^*)^2 = [(\hat{\varepsilon}_{r,t} - \bar{\varepsilon}_r)' (\hat{\varepsilon}_{r,t} - \bar{\varepsilon}_r)]^2 = (\hat{\varepsilon}_{r,t}' \hat{\varepsilon}_{r,t} - \bar{\varepsilon}_r' \bar{\varepsilon}_r)^2$, while for the i.i.d. bootstrap, $\mathbb{E}^* (\varepsilon_{r,t}^{*'} \varepsilon_{r,t}^*)^2 = T^{-1} \sum_{t=1}^T (\hat{\varepsilon}_{r,t}' \hat{\varepsilon}_{r,t} - \bar{\varepsilon}_r' \bar{\varepsilon}_r)^2$. Thus in both cases one has, by Chebychev's inequality,

$$P^* (T^{-1/2} \max_{t=1, \dots, T} \|\varepsilon_{r,t}^*\| > \eta) \leq \frac{1}{\eta^4 T^2} \sum_{t=1}^T \mathbb{E}^* (\varepsilon_{r,t}^{*'} \varepsilon_{r,t}^*)^2 = \frac{3}{\eta^4 T^2} \sum_{t=1}^T (\hat{\varepsilon}_{r,t}' (\hat{\varepsilon}_{r,t} - \bar{\varepsilon}_r))^2 \xrightarrow{p} 0$$

since, under the assumption that ε_t has bounded fourth moment, $T^{-1} \sum_{t=1}^T (\hat{\varepsilon}_{r,t}' \hat{\varepsilon}_{r,t})^2 = O_p(1)$. To see this note that by definition, cf. (A.7),

$$\hat{\varepsilon}_{r,t} = \Delta X_t - \hat{\alpha}^{(r)} \hat{\beta}^{(r)'} X_{1t-1} - \hat{\Psi}^{(r)} \Delta \mathbb{X}_{2t} = \varepsilon_t - \left(\hat{\alpha}^{(r)} \hat{\beta}^{(r)'} - \alpha_0 \beta_0' \right) X_{1t} - \left(\hat{\Psi}^{(r)} - \Psi_0 \right) \Delta \mathbb{X}_{2t}.$$

Next, observe first that,

$$(A.17) \quad \begin{aligned} \hat{\alpha}^{(r)} \hat{\beta}^{(r)'} - \alpha_0 \beta_0' &= \left(\hat{\alpha} K_{r_0}^{(r)} K_{r_0}^{(r)'} \hat{\beta}' - \alpha_0 K_{r_0}^{(r)} K_{r_0}^{(r)'} \beta_0' \right) - \alpha_0 K_{r_0, \perp}^{(r)} K_{r_0, \perp}^{(r)'} \beta_0' \\ &= (\hat{\alpha} - \alpha_0) K_{r_0}^{(r)} K_{r_0}^{(r)'} \beta_0' - \hat{\alpha} K_{r_0}^{(r)} K_{r_0}^{(r)'} \left(\hat{\beta} - \beta_0 \right)' - \alpha_0 K_{r_0, \perp}^{(r)} K_{r_0, \perp}^{(r)'} \beta_0'. \end{aligned}$$

$$(A.18) \quad \begin{aligned} \hat{\Psi}^{(r)} - \Psi_0 &= \left(M_{02} - \hat{\alpha}^{(r)} \hat{\beta}^{(r)'} M_{12} \right) M_{22}^{-1} - \left(\Upsilon_{02} - \alpha_0 \Upsilon_{\beta_0 2} \right) \Upsilon_{22}^{-1} \\ &= \left(M_{02} M_{22}^{-1} - \Upsilon_{02} \Upsilon_{22}^{-1} \right) - \left(\hat{\alpha}^{(r)} \hat{\beta}^{(r)'} - \alpha_0 \beta_0' \right) M_{12} M_{22}^{-1} \\ &\quad - \alpha_0 \left(\beta_0' M_{12} M_{22}^{-1} - \Upsilon_{\beta_0 2} \Upsilon_{22}^{-1} \right). \end{aligned}$$

Using (A.17) and (A.18), and collecting terms, it follows that,

$$(A.19) \quad \hat{\varepsilon}_{r,t} = \varepsilon_{r,t} + v_{r,t}$$

where $\varepsilon_{r,t}$ is given in (A.7), while

$$v_{r,t} = \left[\hat{\alpha} K_{r_0}^{(r)} K_{r_0}^{(r)'} \left(\hat{\beta} - \beta_0 \right)' - \left(\hat{\alpha} - \alpha_0 \right) K_{r_0}^{(r)} K_{r_0}^{(r)'} \beta_0' \right] R_{1t} \\ + \left[\alpha_0 \left(\beta_0' M_{12} M_{22}^{-1} - \Upsilon_{\beta_0 2} \Upsilon_{22}^{-1} \right) - \left(M_{02} M_{22}^{-1} - \Upsilon_{02} \Upsilon_{22}^{-1} \right) \right] \Delta \mathbb{X}_{2t}$$

with $R_{1t} = X_{t-1} - M_{12} M_{22}^{-1} \Delta \mathbb{X}_{2t}$, that is, X_{t-1} corrected for $\Delta \mathbb{X}_{2t}$ by OLS regression. Thus, as claimed above,

$$(A.20) \quad \frac{1}{T} \sum_{t=1}^T \left(\hat{\varepsilon}_{r,t}' \hat{\varepsilon}_{r,t} \right)^2 = \frac{1}{T} \sum_{t=1}^T \left(\varepsilon_{r,t}' \varepsilon_{r,t} \right)^2 + o_p(1),$$

by using standard arguments for the stationary processes ε_t , $\beta_0' X_{t-1}$ and $\Delta \mathbb{X}_{2t}$ in combination with consistency of the parameters in the definition of $v_{r,t}$, while for (cross) products in terms of the non-stationary $\beta_{0\perp}' X_{t-1}$, standard arguments for non-stationary processes in combination with super-consistency give the desired result.

For $S_{r,t}^2$, \mathbb{X}_0^* is fixed at initial values, and again applying the established consistency from Lemma 1, and that for sufficiently large T we have, $\rho(\hat{\Phi}) < 1$, it therefore holds that $P^* \left(\max_{t=1, \dots, T} \|T^{-1/2} S_{r,t}^2\| > \eta \right) \xrightarrow{P} 0$.

That $T^{-1/2} X_{r, \lfloor Tu \rfloor}^* \xrightarrow{w_p} C_0^{(r)} W(u)$, follows if, $T^{-1/2} \sum_{t=1}^{\lfloor Tu \rfloor} \varepsilon_{r,t}^* \xrightarrow{w_p} W(\cdot)$, which for the wild bootstrap, as in Cavaliere *et al.* (2010a), is implied by the pointwise convergence,

$$(A.21) \quad \frac{1}{T} \sum_{t=1}^{\lfloor Tu \rfloor} \hat{\varepsilon}_{r,t} \hat{\varepsilon}_{r,t}' = \frac{1}{T} \sum_{t=1}^{\lfloor Tu \rfloor} \varepsilon_{r,t} \varepsilon_{r,t}' + o_p(1) \xrightarrow{P} u \Omega_0^{(r)},$$

see also (A.20). For the i.i.d. bootstrap, the result follows as in Swensen (2006, Lemma S2) using (A.21) and finite fourth order moments of $\varepsilon_{r,t}$.

Consider next the case of *deterministic terms*: First in case (ii) of a *restricted constant*, the representation is as in (A.14), but with $\varepsilon_{r,t}^*$ replaced by $(\varepsilon_{r,t}^* + \hat{\alpha}^{(r)} \hat{\rho}^{(r)'})$, and $e_{r,t}^*$ in (A.15) replaced by $(e_{r,t}^* + (\hat{\rho}^{(r)} \hat{\alpha}^{(r)'}, 0, \dots, 0)')$. The results hold as before since $\hat{C}^{(r)}(\varepsilon_{r,t}^* + \hat{\alpha}^{(r)} \hat{\rho}^{(r)'}) = \hat{C}^{(r)} \varepsilon_{r,t}^*$, and the extra term in $S_{r,t}^1$ induced by $(\hat{\rho}^{(r)} \hat{\alpha}^{(r)'}, 0, \dots, 0)'$ is bounded in probability from the previous arguments. For the case (iii) of a *restricted linear trend*, $\varepsilon_{r,t}^*$ is replaced by $(\varepsilon_{r,t}^* + \hat{\alpha}^{(r)} \hat{\rho}^{(r)'} t + \hat{\phi}^{(r)})$, and $e_{r,t}^*$ in (A.15) by $[e_{r,t}^* + (\hat{\rho}^{(r)} \hat{\alpha}^{(r)'} t + \hat{\phi}^{(r)}, 0, \dots, 0)']$. Observe initially that, $\hat{C}^{(r)} \sum_{i=1}^t (\varepsilon_{r,i}^* + \hat{\alpha}^{(r)} \hat{\rho}^{(r)'} t + \hat{\phi}^{(r)}) = \hat{C}^{(r)} (\sum_{i=1}^t \varepsilon_{r,i}^* + \hat{\phi}^{(r)} t)$. Next, in terms of $S_{r,t}^1$, there is a linear trend part induced by, $(\hat{\alpha}^{(r)}, \hat{\Psi}^{(r)}) (\hat{\mathbb{B}}' \hat{\mathbb{A}})^{-1} (\hat{\Phi}_t(L) t) \hat{\mathbb{B}}' (\hat{\rho}^{(r)} \hat{\alpha}^{(r)'}, 0, \dots, 0)'$, where the polynomial, $\hat{\Phi}_t(z) := \sum_{i=0}^{t-1} \hat{\Phi}^i z^i$, $z \in \mathbb{C}$. An expansion of this around $z = 1$ (cf. Lemma 4.1, Johansen, 1996) gives, $\hat{\Phi}_t(z) = \hat{\Phi}_t(1) + \hat{\Phi}_t^*(z) (1 - z)$, with $\hat{\Phi}_t^*(z) = \sum_{i=0}^{t-2} \hat{\Phi}_i^* z^i$, $\hat{\Phi}_i^* = -\sum_{j=i+1}^{t-1} \hat{\Phi}^j$ and $\hat{\Phi}_i^*$ exponentially decreasing as above. Hence the linear trend induced in $S_{r,t}^1$ is given by, $(\hat{\alpha}^{(r)}, \hat{\Psi}^{(r)}) (\hat{\mathbb{B}}' \hat{\mathbb{A}})^{-1} \hat{\Phi}_t(1) \hat{\mathbb{B}}' (\hat{\rho}^{(r)} \hat{\alpha}^{(r)'}, 0, \dots, 0)'$. Collecting terms,

$$\hat{\tau}_{r,t} = \left[\hat{C}^{(r)} \hat{\phi}^{(r)} + (\hat{\alpha}^{(r)}, \hat{\Psi}^{(r)}) (\hat{\mathbb{B}}' \hat{\mathbb{A}})^{-1} \hat{\Phi}_t(1) \hat{\mathbb{B}}' \left(\hat{\rho}^{(r)} \hat{\alpha}^{(r)'}, 0, \dots, 0 \right)' \right] t,$$

and from the established consistency in Lemma 1, $T^{-1} \hat{\tau}_{r,[Tu]} \xrightarrow{w} \tau_0^{(r)} u$, with $\tau_0^{(r)}$ as given in the proposition, by standard algebraic manipulations. \square

PROOF OF PROPOSITION 2: As in the proof of Theorem 3 of Cavaliere *et al.* (2010a), this follows immediately by the results in Proposition 1 using standard arguments and defining $B_{p-r} := \left(\alpha_{\perp}^{(r)'} \Omega^{(r)} \alpha_{\perp}^{(r)} \right)^{-1/2} \alpha_{\perp}^{(r)'} W$. \square

References

- ANDREWS, D.W.K. AND M. BUCHINSKY (2000): A three-step method for choosing the number of bootstrap repetitions, *Econometrica* 68, pages 23–51.
- CAVALIERE, G., A. RAHBK AND A.M.R. TAYLOR (2010a): Co-integration rank testing under conditional heteroskedasticity, *Econometric Theory* 26, 1719–1760.
- (2010b): Testing for co-integration in vector autoregressions with non-stationary volatility, *Journal of Econometrics* 158, 7–24.
- HANSEN, B.E. (1996): Inference when a nuisance parameter is not identified under the null hypothesis, *Econometrica* 64, 413–430.
- JOHANSEN, S. (1996): *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford: Oxford University Press.
- (2002): A small sample correction of the test for cointegrating rank in the vector autoregressive model, *Econometrica* 70, 1929–1961.
- PARK, J.Y. (2003): Bootstrap unit root tests, *Econometrica* 71, 1845–1895.
- RAHBK, A., H.C. KONGSTED AND C. JØRGENSEN (1999): Trend stationarity in the I(2) cointegration model, *Journal of Econometrics* 90, 265–289.
- SWENSEN, A.R. (2006): Bootstrap algorithms for testing and determining the cointegration rank in VAR models, *Econometrica* 74, 1699–1714.
- (2009): Corrigendum to “Bootstrap algorithms for testing and determining the cointegration rank in VAR models”, *Econometrica* 77, 1703–1704.
- TRENKLER, C. (2009): Bootstrapping systems cointegration tests with a prior adjustment for deterministic terms, *Econometric Theory* 25, 243–269.
- VAN GIEBERSBERGEN, 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.

TABLE I: EMPIRICAL REJECTION FREQUENCIES OF ASYMPTOTIC AND BOOTSTRAP CO-INTEGRATION RANK TESTS

| δ | | Panel (a): Co-integrated VAR(2) Model with Rank $\tau_0 = 1$ | | | | | | | | | | | | Panel (b): Non Co-integrated VAR(2) Model | | | | | | | |
|----------|-----|--|-------|-------|-----------------|-------|------|----------------|-----|-----|-----------------|------|----------|---|------|-----------------|-----|-----------------|------|------|-----|
| | | $Q_{0,T}^{BC}$ | | | $Q_{0,T}^{*SW}$ | | | $Q_{1,T}^{BC}$ | | | $Q_{1,T}^{*SW}$ | | | $Q_{0,T}^{BC}$ | | $Q_{0,T}^{*SW}$ | | $Q_{0,T}^{*SW}$ | | | |
| T | ERF | RC | ERF | RC | ERF | RC | ERF | RC | ERF | RC | ERF | RC | γ | T | ERF | RC | ERF | RC | ERF | RC | |
| 0.0 | 50 | 97.6 | 32.2 | 55.9 | 0.0 | 58.2 | 2.2 | 45.5 | 2.0 | 5.2 | 0.1 | 10.2 | 0.4 | 0 | 50 | 17.5 | 5.5 | 4.8 | 0.0 | 4.8 | 0.0 |
| | 100 | 100.0 | 98.8 | 99.4 | 0.0 | 99.4 | 0.1 | 22.6 | 3.8 | 5.4 | 0.0 | 7.6 | 0.0 | 100 | 9.8 | 5.4 | 4.7 | 0.0 | 4.8 | 0.0 | 0.0 |
| | 200 | 100.0 | 100.0 | 100.0 | 0.0 | 100.0 | 0.0 | 13.3 | 4.8 | 5.0 | 0.0 | 5.8 | 0.0 | 200 | 7.1 | 5.1 | 4.8 | 0.0 | 4.9 | 0.0 | 0.0 |
| 0.1 | 50 | 97.5 | 32.1 | 54.3 | 0.0 | 55.0 | 7.2 | 45.4 | 2.1 | 5.3 | 0.1 | 10.2 | 0.5 | 0.5 | 50 | 37.5 | 4.3 | 5.5 | 0.0 | 7.7 | 0.0 |
| | 100 | 100.0 | 98.8 | 99.3 | 0.0 | 99.1 | 2.5 | 22.6 | 3.9 | 5.4 | 0.0 | 7.7 | 0.0 | 100 | 17.1 | 5.0 | 5.2 | 0.0 | 6.0 | 0.0 | 0.0 |
| | 200 | 100.0 | 100.0 | 100.0 | 0.0 | 100.0 | 0.5 | 13.3 | 4.8 | 4.9 | 0.0 | 5.8 | 0.0 | 200 | 10.1 | 4.9 | 4.8 | 0.0 | 5.0 | 0.0 | 0.0 |
| 0.2 | 50 | 97.3 | 31.4 | 51.3 | 0.0 | 47.8 | 24.1 | 45.0 | 2.4 | 5.1 | 0.1 | 9.9 | 0.8 | 0.8 | 50 | 78.8 | 1.9 | 8.1 | 0.0 | 19.3 | 0.4 |
| | 100 | 100.0 | 98.6 | 99.1 | 0.0 | 98.0 | 28.9 | 22.3 | 4.2 | 5.3 | 0.0 | 8.3 | 0.0 | 100 | 44.3 | 3.1 | 6.6 | 0.0 | 11.4 | 0.0 | 0.0 |
| | 200 | 100.0 | 100.0 | 100.0 | 0.0 | 100.0 | 34.4 | 13.1 | 4.9 | 5.0 | 0.0 | 6.5 | 0.0 | 200 | 21.6 | 4.5 | 5.3 | 0.0 | 6.9 | 0.0 | 0.0 |
| 0.3 | 50 | 97.4 | 31.6 | 49.0 | 0.0 | 41.8 | 48.6 | 45.1 | 2.8 | 4.9 | 0.1 | 10.2 | 1.6 | 0.9 | 50 | 93.4 | 0.9 | 11.0 | 1.6 | 30.0 | 2.0 |
| | 100 | 100.0 | 98.3 | 98.7 | 0.0 | 94.8 | 73.3 | 21.7 | 4.2 | 5.3 | 0.0 | 9.3 | 0.0 | 100 | 74.2 | 1.1 | 8.1 | 0.0 | 21.7 | 0.2 | 0.2 |
| | 200 | 100.0 | 100.0 | 100.0 | 0.0 | 100.0 | 91.6 | 12.4 | 4.9 | 5.1 | 0.0 | 8.8 | 0.0 | 200 | 43.2 | 3.0 | 6.4 | 0.0 | 11.7 | 0.0 | 0.0 |

Notes: 'ERF' denotes the empirical rejection rates; 'RC' denotes the fraction of times the bootstrap algorithm generates explosive samples. All experiments conducted using 10,000 replications. The number B of bootstrap samples used in the bootstrap algorithms is 399. All tests are conducted at the nominal 5% significance level. The VAR model was fitted with a restricted constant (i.e. deterministic case (ii) in Section 2). For both the standard and the Bartlett-corrected likelihood ratio tests, asymptotic critical values as reported in Table 15.2 of Johansen (1996) are employed.

TABLE II: SEQUENTIAL PROCEDURES FOR DETERMINING THE CO-INTEGRATION RANK. VAR(2) MODEL WITH RANK $\tau_0 = 1$

| δ | T | Asymptotic | | | | | | | | | | | | Bartlett | | | | | | Bootstrap (SW) | | | | | |
|----------|-----|------------------|------|------|------|------------------|------|-----|------|------|------------------|-----|-----|----------|------|------------------|-----|-----|------|----------------|--|--|--|--|--|
| | | $\hat{\tau} = 0$ | 1 | 2 | 3, 4 | $\hat{\tau} = 0$ | 1 | 2 | 3, 4 | RC | $\hat{\tau} = 0$ | 1 | 2 | 3, 4 | RC | $\hat{\tau} = 0$ | 1 | 2 | 3, 4 | RC | | | | | |
| 0.0 | 50 | 2.4 | 52.1 | 32.8 | 12.7 | 67.8 | 30.3 | 1.7 | 0.2 | 44.1 | 50.6 | 4.4 | 0.9 | 0.3 | 41.8 | 48.1 | 9.0 | 1.1 | 3.5 | | | | | | |
| | 100 | 0.0 | 77.4 | 18.8 | 3.8 | 1.2 | 94.9 | 3.5 | 0.3 | 0.6 | 94.0 | 4.7 | 0.7 | 0.1 | 0.6 | 91.8 | 6.9 | 0.8 | 0.3 | | | | | | |
| | 200 | 0.0 | 86.7 | 11.5 | 1.8 | 0.0 | 95.2 | 4.6 | 0.2 | 0.0 | 95.0 | 4.3 | 0.7 | 0.0 | 0.0 | 94.2 | 5.2 | 0.5 | 0.0 | | | | | | |
| 0.1 | 50 | 2.5 | 52.1 | 32.8 | 12.6 | 67.9 | 30.0 | 1.8 | 0.3 | 45.7 | 49.0 | 4.3 | 1.0 | 0.3 | 45.0 | 45.0 | 8.7 | 1.3 | 8.4 | | | | | | |
| | 100 | 0.0 | 77.4 | 18.9 | 3.7 | 1.2 | 94.9 | 3.7 | 0.3 | 0.7 | 94.0 | 4.6 | 0.7 | 0.1 | 0.9 | 91.4 | 6.9 | 0.8 | 2.7 | | | | | | |
| | 200 | 0.0 | 86.7 | 11.5 | 1.8 | 0.0 | 95.2 | 4.6 | 0.2 | 0.0 | 95.1 | 4.1 | 0.8 | 0.0 | 0.0 | 94.2 | 5.3 | 0.5 | 0.6 | | | | | | |
| 0.2 | 50 | 2.7 | 52.3 | 32.8 | 12.2 | 68.6 | 29.0 | 2.1 | 0.2 | 48.7 | 46.3 | 4.1 | 0.9 | 0.3 | 52.2 | 38.1 | 8.3 | 1.4 | 25.2 | | | | | | |
| | 100 | 0.0 | 77.7 | 18.5 | 3.7 | 1.4 | 94.4 | 3.9 | 0.2 | 0.9 | 93.8 | 4.6 | 0.7 | 0.1 | 2.0 | 89.7 | 7.4 | 0.9 | 29.1 | | | | | | |
| | 200 | 0.0 | 86.9 | 11.4 | 1.7 | 0.0 | 95.1 | 4.6 | 0.2 | 0.0 | 95.0 | 4.4 | 0.7 | 0.0 | 0.0 | 93.5 | 5.8 | 0.7 | 34.4 | | | | | | |
| 0.3 | 50 | 2.6 | 52.3 | 33.2 | 11.9 | 68.4 | 28.9 | 2.5 | 0.2 | 51.0 | 44.1 | 4.0 | 0.9 | 0.3 | 58.2 | 31.9 | 8.1 | 1.8 | 49.5 | | | | | | |
| | 100 | 0.0 | 78.3 | 18.4 | 3.4 | 1.7 | 94.2 | 4.0 | 0.2 | 1.3 | 93.4 | 4.6 | 0.7 | 0.1 | 5.2 | 85.5 | 8.4 | 0.9 | 73.4 | | | | | | |
| | 200 | 0.0 | 87.6 | 10.9 | 1.5 | 0.0 | 95.1 | 4.7 | 0.2 | 0.0 | 94.9 | 4.5 | 0.7 | 0.0 | 0.0 | 91.2 | 7.7 | 1.1 | 91.6 | | | | | | |

Notes: 'CRT' denotes Algorithm 2 of Section 2, while 'SW' denotes Algorithm 2 of Swensen (2006). See also notes to Table 1.