

Trenkler, Carsten

**Working Paper**

## Bootstrapping systems cointegration tests with a prior adjustment for deterministic terms

SFB 649 Discussion Paper, No. 2006,012

**Provided in Cooperation with:**

Collaborative Research Center 649: Economic Risk, Humboldt University Berlin

*Suggested Citation:* Trenkler, Carsten (2006) : Bootstrapping systems cointegration tests with a prior adjustment for deterministic terms, SFB 649 Discussion Paper, No. 2006,012, Humboldt University of Berlin, Collaborative Research Center 649 - Economic Risk, Berlin

This Version is available at:

<https://hdl.handle.net/10419/25095>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

SFB 649 Discussion Paper 2006-012

# Bootstrapping Systems Cointegration Tests with a Prior Adjustment for Deterministic Terms

Carsten Trenkler\*



\* Institute of Statistics and Econometrics,  
School of Business and Economics,  
Humboldt-Universität zu Berlin, Germany

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

<http://sfb649.wiwi.hu-berlin.de>  
ISSN 1860-5664

SFB 649, Humboldt-Universität zu Berlin  
Spandauer Straße 1, D-10178 Berlin



SFB 649 ECONOMIC RISK BERLIN

February 10, 2006

**Bootstrapping Systems Cointegration Tests  
With a Prior Adjustment for Deterministic Terms**

Carsten Trenkler\*  
Humboldt-Universität zu Berlin  
School of Business and Economics  
Institute of Statistics and Econometrics  
Spandauer Str. 1  
D-10178 berlin  
Germany  
Tel.: +49-30-2093-5711  
Fax.: +49-30-2093-5712  
email: trenkler@wiwi.hu-berlin.de

**Abstract**

In this paper we analyse bootstrap procedures for systems cointegration tests with a prior adjustment for deterministic terms suggested by Saikkonen & Lütkepohl (2000b) and Saikkonen, Lütkepohl & Trenkler (2006). The asymptotic properties of the bootstrap test procedures are derived and their small sample properties are studied. The simulation study also considers the standard asymptotic test versions and the Johansen cointegration test for comparison.

*Keywords:* Bootstrap, Systems cointegration tests, VEC models

*JEL classification:* C12, C13, C15, C32

---

\*This research was supported by the Deutsche Forschungsgemeinschaft (DFG) through the SFB 649 “Economic Risk”. Moreover, I am grateful to Anders Rygh Swensen, Helmut Lütkepohl, and participants of the “Unit Root and Cointegration Testing Conference” at the University of Algarve for many helpful comments and suggestions.

# 1 Introduction

It is known that systems cointegration tests suffer quite often from very poor small sample properties (compare e.g. Hubrich, Lütkepohl & Saikkonen 2001). In other words, the asymptotic test distributions may provide a bad approximation for small sample sizes prevailing in applied work. Hence, the inference on the number of cointegrating relations among a system of variables may be distorted. This applies both to cointegration tests based on Johansen (1988) but also to test procedures with a prior adjustment for deterministic terms as suggested by Saikkonen & Lütkepohl (2000b) and Saikkonen et al. (2006).

In the literature, the use of bootstrap methods has been suggested as a possible way to improve the small sample properties of a test procedure. This usually refers to a test's size, since it is not clear what effect bootstrap methods may have on the test's power (compare e.g. Davidson & MacKinnon 2006). Improved approximation to the small sample distribution can be expected if the test statistic is asymptotical pivotal (compare e.g. Horowitz 2001). This is the case for the cointegration test statistics considered in this paper. However, doubts have been raised about the usefulness of bootstrap methods in relation to time series data (compare e.g. Horowitz 2003, MacKinnon 2002). Obviously, the applicability and properties of bootstrap procedures are not as favourable as in the case of i.i.d. data due to the dependence structure in time series. As a response to that one may focus on time series models with i.i.d. error terms.

Nevertheless, different types of bootstrap procedures have been suggested and successfully applied in the time series framework in order to test for a unit root (see e.g. Basawa, Mallik, McCormic, Reeves & Taylor 1991, Paparoditis & Politis 2003, Park 2003). In addition, Park (2003) shows that bootstraps offer asymptotic refinements in finite samples for the ADF unit root test. The bootstrap algorithm he considers can be regarded as an univariate counterpart of the algorithms analysed later on. With respect to systems cointegration tests much less is known about bootstrapping. Based on simulation studies, van Giersbergen (1996), Harris & Judge (1998), and Mantalos & Shukur (2001) analyse different variants of so-called recursive and block bootstraps for Johansen cointegration tests. The results of these studies, although partly mixed, do not show that an application of bootstrap techniques generally improves the tests' small sample properties. However, a theoretical justification for the use of bootstrap methods to test for the cointegrating rank has been recently given by Swensen (2005a). He shows that recursive bootstrap procedures for the Johansen test have the same limiting distribution as the asymptotic test versions. Furthermore, his simulation findings demonstrate that bootstrap procedures can lead to an improvement in the tests' sizes in small samples.

We will follow Swensen (2005a) and analyse different variants of recursive bootstrap algorithms for systems cointegration tests with a prior adjustment for deterministic terms which have been suggested by Saikkonen & Lütkepohl (2000b) and Saikkonen et al. (2006). Both tests assume that the data generating process can be decomposed into an unobservable zero-mean stochastic part, which follows a vector autoregressive process, and a deterministic component, which consists of a level term and a linear trend. Their common idea is to estimate the deterministic terms in a first step and to adjust the original time series by these estimated terms. Then, a likelihood ratio type test in the spirit of Johansen (1988) is applied to the adjusted data. This model setup requires to introduce bootstrap equivalents of both the deterministic and the stochastic components. We define these equivalents such that the sum of them produces the same bootstrap data as are obtained by applying the recursive bootstrap scheme directly to the observable data. It turns out that the bootstrap tests have the same limiting distributions under the cointegrating rank null hypothesis as the corresponding asymptotic test procedures by Saikkonen & Lütkepohl (2000b) and Saikkonen et al. (2006). Furthermore, the small sample properties of the bootstrap procedures are analysed and compared with the ones of the asymptotic tests. The comparison also includes the Johansen cointegration test. Thereby we can compare our findings with the ones of Swensen (2005a, 2005b).

The rest of the paper is structured as follows. The next section describes the system cointegration tests and Section 3 introduces the bootstrap procedures. Their asymptotic properties are discussed in Section 4. Some simulation results are presented in Section 5 and Section 6 summarizes and concludes. The theoretical derivations and explanations on the response surfaces are deferred to two Appendices.

The following notation is used throughout. If  $A$  is an  $(n \times m)$  matrix of full column rank ( $n > m$ ), its orthogonal complement  $A_{\perp}$  is an  $(n \times (n - m))$  matrix of full column rank and such that  $A'A_{\perp} = 0$ . Furthermore, let  $\bar{A} = A(A'A)^{-1}$ . A  $(n \times n)$  identity matrix is denoted by  $I_n$ . LS and GLS are used to abbreviate least squares and generalized least squares respectively, LR relates to likelihood ratio, RR is short for reduced rank, and DGP refers to data generating process. The abbreviations r.h.s. and l.h.s. stand for right- and left-hand-side, respectively.

## 2 Test Procedures

Let us consider a  $n$ -dimensional times series  $y_t = (y_{1t}, \dots, y_{nt})'$  ( $t = 1, \dots, T$ ), which is generated by

$$y_t = \mu_0 + \mu_1 t + x_t, \quad t = 1, 2, \dots, \quad (2.1)$$

where  $\mu_0$  and  $\mu_1$  are unknown  $(n \times 1)$  parameter vectors. Hence, the deterministic part consists of a constant and a linear trend. The term  $x_t$  is an unobservable stochastic error process, for which we make the following assumption.

**Assumption 1.** The process  $x_t$  is integrated of order at most  $I(1)$  with cointegrating rank  $r$  and follows a vector autoregressive process of order  $p$ , VAR( $p$ ),

$$x_t = A_1 x_{t-1} + \cdots + A_p x_{t-p} + \varepsilon_t, \quad t = 1, 2, \dots, \quad (2.2)$$

where  $A_j$  are  $(n \times n)$  coefficient matrices and the initial values are such that  $x_t = 0, t \leq 0$ . For the error terms we assume  $\varepsilon_t \sim \text{i.i.d.}(0, \Omega)$  with positive definite covariance matrix  $\Omega$  and  $E(\varepsilon_t^4) < \infty$ .  $\square$

The initial value condition  $x_t = 0, t \leq 0$ , is imposed for convenience to construct the test by Saikkonen & Lütkepohl (2000b) and the bootstrap versions of this procedure. However, the asymptotic tests remain valid if the initial values have a fixed probability distribution, which does not depend on the sample size. Under Assumption 1, the process  $x_t$  has the usual vector error correction model (VECM) form

$$\Delta x_t = \Pi x_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \Delta x_{t-j} + \varepsilon_t, \quad t = 1, 2, \dots,$$

where  $\Pi$  and  $\Gamma_j$  ( $j = 1, \dots, p-1$ ) are  $(n \times n)$  unknown parameter matrices. Because the cointegrating rank is  $r$ , the matrix  $\Pi$  can be written as  $\Pi = \alpha\beta'$ , where  $\alpha$  and  $\beta$  are  $(n \times r)$  matrices of full column rank. Obviously, the cointegrating rank is equal to the rank of the matrix  $\Pi$ . As is well-known,  $\beta'x_t$  and  $\Delta x_t$  are zero mean  $I(0)$  processes. Defining  $\Gamma = I_n - \Gamma_1 - \cdots - \Gamma_{p-1} = I_n + \sum_{j=1}^{p-1} jA_{j+1}$  and  $C = \beta_{\perp}(\alpha'_{\perp}\Gamma\beta_{\perp})^{-1}\alpha'_{\perp}$ , we have

$$x_t = C \sum_{j=1}^t \varepsilon_j + \xi_t, \quad t = 1, 2, \dots,$$

where  $\xi_t$  is a zero mean  $I(0)$  process.

Finally, we derive the VECM representation for  $y_t$ . Multiplying (2.1) by  $A(L) = I_n - A_1L - \cdots - A_pL^p = I_n\Delta - \Pi L - \Gamma_1\Delta L - \cdots - \Gamma_{p-1}\Delta L^{p-1}$  and rearranging yields

$$\Delta y_t = \nu + \alpha(\beta'y_{t-1} - \phi(t-1)) + \sum_{j=1}^{p-1} \Gamma_j \Delta y_{t-j} + \varepsilon_t, \quad t = p+1, p+2, \dots, \quad (2.3)$$

where  $\nu = -\Pi\mu_0 + \Gamma\mu_1 = -\alpha\theta + \Gamma\mu_1$ ,  $\theta = \beta'\mu_0$ , and  $\phi = \beta'\mu_1$ .

In the following we consider the so-called trace test version, i.e. we aim to test the pair of hypotheses

$$H_0(r_0) : \text{rk}(\Pi) = r_0 \quad \text{vs.} \quad H_1(r_0) : \text{rk}(\Pi) > r_0. \quad (2.4)$$

The idea of the test proposal of Saikkonen & Lütkepohl (2000b) is to estimate the deterministic terms in a first step by a feasible GLS procedure. Then,  $y_t$  is adjusted by these estimated terms. In a second step, a Johansen LR-type test is performed on the adjusted time series. More precisely, defining

$$a_{0t} = \begin{cases} 1 & \text{for } t \geq 1 \\ 0 & \text{for } t \leq 0 \end{cases} \quad \text{and} \quad a_{1t} = \begin{cases} t & \text{for } t \geq 1 \\ 0 & \text{for } t \leq 0 \end{cases}$$

and multiplying (2.1) by  $A(L)$  we obtain

$$A(L)y_t = H_{0t}\mu_0 + H_{1t}\mu_1 + \varepsilon_t, \quad t = 1, \dots, T, \quad (2.5)$$

where  $H_{it} = A(L)a_{it}$  ( $i = 0, 1$ ), and  $\varepsilon_t = A(L)x_t$  (see (2.2)). Furthermore, define  $Q$  such that  $QQ' = \Omega^{-1}$  and multiply (2.5) by  $Q'$ , then we have

$$Q'A(L)y_t = Q'H_{0t}\mu_0 + Q'H_{1t}\mu_1 + \eta_t, \quad t = 1, \dots, T, \quad (2.6)$$

where  $\eta_t = Q'\varepsilon_t$ . The matrix  $Q$  can be chosen as  $Q = [\Omega^{-1}\alpha(\alpha'\Omega^{-1}\alpha)^{-1/2} : \alpha_{\perp}(\alpha'_{\perp}\Omega^{-1}\alpha_{\perp})^{-1/2}]$ . Hence, the error term  $\eta_t$  has a zero mean and a unit covariance matrix as it is required for a GLS transformation. To make the GLS estimation feasible, Saikkonen & Lütkepohl (2000b) propose to use the RR estimators  $\tilde{\alpha}$ ,  $\tilde{\beta}$ ,  $\tilde{\Gamma}_j$  ( $j = 1, \dots, p-1$ ) and  $\tilde{\Omega}$  which are obtained from (2.3) by applying  $r_0$ , the rank under the null hypothesis of the cointegration test. Based on the RR estimators, one can compute the estimators  $\tilde{Q}$  and  $\tilde{H}_{it}$  ( $i = 0, 1$ ). Thus, feasible GLS estimators of  $\mu_0$  and  $\mu_1$ , say  $\tilde{\mu}_0$  and  $\tilde{\mu}_1$ , are obtained by a multivariate LS estimation of the model

$$\tilde{Q}'\tilde{A}(L)y_t = \tilde{Q}'\tilde{H}_{0t}\mu_0 + \tilde{Q}'\tilde{H}_{1t}\mu_1 + \tilde{\eta}_t, \quad t = 1, \dots, T. \quad (2.7)$$

The estimated deterministic terms are used to adjust  $y_t$ , what gives the sample analogue  $\tilde{x}_t = y_t - \tilde{\mu}_0 - \tilde{\mu}_1 t$  of  $x_t$ . Then, an LR-type test is performed with respect to

$$\Delta\tilde{x}_t = \Pi\tilde{x}_{t-1} + \sum_{j=1}^{p-1} \Gamma_j\Delta\tilde{x}_{t-j} + e_t, \quad t = p+1, \dots, T, \quad (2.8)$$

where  $e_t$  is an error term. Since  $\tilde{x}_t$  is adjusted by the deterministic terms, a test version without deterministic terms like in Johansen (1988) is applied. Thus,  $\Delta\tilde{x}_t$  and  $\tilde{x}_{t-1}$  are regressed on  $(\Delta\tilde{x}'_{t-1}, \dots, \Delta\tilde{x}'_{t-p-1})'$  to obtain the residuals  $\tilde{R}_{0t}$  and  $\tilde{R}_{1t}$ , respectively. Next, consider

$\tilde{S}_{ij} = T^{-1} \sum_{t=1}^T \tilde{R}_{it} \tilde{R}'_{jt}$  ( $i, j = 0, 1$ ). Solving  $\det(\lambda \tilde{S}_{11} - \tilde{S}_{10} \tilde{S}_{00}^{-1} \tilde{S}_{01}) = 0$  we obtain the ordered generalized eigenvalues  $\tilde{\lambda}_1 \geq \dots \geq \tilde{\lambda}_n$ . Finally, the LR test statistic for the pair of hypotheses in (2.4) is<sup>1</sup>

$$GLS(r_0) = -T \sum_{j=r_0+1}^n \log(1 - \tilde{\lambda}_j).$$

As stated in Theorem 1 of Saikkonen & Lütkepohl (2000b), the level parameter  $\mu_0$  is not consistently estimated in the direction of  $\beta_{\perp}$  because  $\mu_0$  is not identified in that direction in (2.3). Therefore, Saikkonen et al. (2006) suggest to avoid the estimation of the level parameter in the first stage. Instead, only  $\mu_1$  is estimated and the effect of the level parameter is taken into account when the test is performed. Defining  $\phi_* = \beta'_{\perp} \mu_1$  and noting that  $\phi_* = \beta'_{\perp} C(\nu - \Gamma \beta (\beta' \beta)^{-1} \phi)$ , Saikkonen et al. (2006) propose to estimate  $\mu_1$  by

$$\hat{\mu}_1 = \tilde{\beta} (\tilde{\beta}' \tilde{\beta})^{-1} \tilde{\phi} + \tilde{\beta}_{\perp} (\tilde{\beta}'_{\perp} \tilde{\beta}_{\perp})^{-1} \tilde{\phi}_*, \quad (2.9)$$

where  $\tilde{\beta}$ ,  $\tilde{\beta}_{\perp}$ ,  $\tilde{\phi}$ , and  $\tilde{\phi}_* = \tilde{\beta}'_{\perp} \tilde{C}(\tilde{\nu} - \tilde{\Gamma} \tilde{\beta} (\tilde{\beta}' \tilde{\beta})^{-1} \tilde{\phi})$  are obtained from a RR estimation of (2.3). Using  $\hat{\mu}_1$  we adjust  $y_t$  for a linear trend and obtain

$$\hat{y}_t^c = y_t - \hat{\mu}_1 t.$$

Then, we set up an auxiliary VECM for  $\hat{y}_t^c$

$$\Delta \hat{y}_t^c = \alpha (\beta' y_{t-1}^c - \theta) + \sum_{j=1}^{p-1} \Gamma_j \Delta \hat{y}_{t-j}^c + e_t^c, \quad (2.10)$$

where  $e_t^c$  is an error term. Finally, an LR-type test, now including a constant term, can be performed with respect to (2.10). Denoting the corresponding ordered generalized eigenvalues as  $\hat{\lambda}_1 \geq \dots \geq \hat{\lambda}_n$ , we obtain the LR test statistic for the pair of hypotheses in (2.4)<sup>2</sup>

$$SLT(r_0) = -T \sum_{j=r_0+1}^n \log(1 - \hat{\lambda}_j).$$

Saikkonen & Lütkepohl (2000b) and Saikkonen et al. (2006) derive the limiting distributions for  $GLS(r_0)$  and  $SLT(r_0)$  given in Theorem 1.

---

<sup>1</sup>Note that the generalized eigenvalue problem described here is slightly different from the one in Saikkonen & Lütkepohl (2000b). However, the eigenvalue problems can be transformed into each other by an appropriate redefinition of the respective eigenvalues. Hence, the LR statistics based on the two different sets of eigenvalues are identical apart from minor numerical differences.

<sup>2</sup>Saikkonen et al. (2006) allow in addition for a level shift at unknown time. The test statistic considered here can be obtained by deleting all terms associated with the level shift. This has no effect on the limiting distribution.



**Theorem 1.** Under Assumption 1 and if  $H_0(r_0)$  is true, then

$$P_{r_0,T}(GLS(r_0) \leq x) \longrightarrow P(\text{tr}(D_{GLS}) \leq x) \text{ and } P_{r_0,T}(SLT(r_0) \leq x) \longrightarrow P(\text{tr}(D_{SLT}) \leq x)$$

for all  $x$  as  $T \rightarrow \infty$ , where

$$D_{GLS} = \left( \int_0^1 B_*(s) dB_*(s)' \right)' \left( \int_0^1 B_*(s) B_*(s)' ds \right)^{-1} \left( \int_0^1 B_*(s) dB_*(s)' \right),$$

$$D_{SLT} = \left( \int_0^1 B_+(s) dB_*(s)' \right)' \left( \int_0^1 B_+(s) B_+(s)' ds \right)^{-1} \left( \int_0^1 B_+(s) dB_*(s)' \right),$$

$B(s)$  is a  $(n - r_0)$ -dimensional standard Brownian motion,  $B_*(s) = B(s) - sB(1)$  is a  $(n - r_0)$ -dimensional Brownian bridge,  $B_+(s) = [B_*(s)', 1]'$  and  $dB_*(s) = dB(s) - dsB(1)$ .  $\square$

Several remarks are worth making regarding this theorem. The probability measures  $P_{r_0,T}$  refer to the joint distributions of the sample  $y_1, \dots, y_T$  for situations where  $\text{rk}(\Pi) = r_0$  (compare Swensen 2005a). With respect to the bootstrap framework we will introduce a further probability measure conditional on the observations. The limiting distributions for  $GLS(r_0)$  and  $SLT(r_0)$  differ in that the process  $B_+(s)$  appears in place of the Brownian bridge  $B_*(s)$ . Of course, the reason is that an intercept term is included in the auxiliary model on which  $LST(r_0)$  is based. On the other hand, the limiting distribution of  $LST(r_0)$  is formally similar to its counterpart in Theorem 6.3 of Johansen (1995) where a standard Brownian motion appears instead of the Brownian bridge in Theorem 1.

### 3 Bootstrap Algorithms

Swensen (2005a) proposes two variants of a recursive bootstrap related to the VECM (2.3). The first algorithm to generate the so-called pseudo or bootstrapped observations  $y_1^*, \dots, y_T^*$  works as follows.

#### Algorithm 1.

- (1) Estimate model (2.3) by a RR regression under the rank hypothesis  $H_0(r_0) : \text{rk}(\Pi) = r_0$  in order to obtain estimates  $\tilde{\alpha}$ ,  $\tilde{\beta}$ ,  $\tilde{\nu}$ ,  $\tilde{\phi}$ , and  $\tilde{\Gamma}_j$  ( $j = 1, \dots, p - 1$ ).
- (2) Check whether the roots of the equation  $\det[\tilde{A}(z)] = 0$ , where

$$\tilde{A}(z) = (1 - z)I_n - \tilde{\alpha}\tilde{\beta}'z - \tilde{\Gamma}_1(1 - z)z - \dots - \tilde{\Gamma}_{p-1}(1 - z)z^{p-1},$$

are equal to 1 or outside the unit circle and whether  $\tilde{\alpha}'_{\perp} \tilde{\Gamma}_{\perp} \tilde{\beta}_{\perp}$  is nonsingular.

(3) If so, compute the residuals  $\tilde{\varepsilon}_{p+1}, \dots, \tilde{\varepsilon}_T$  from the estimated VECM (2.3).

(4) Compute  $y_t^*$ ,  $t = p + 1, \dots, T$ , recursively by

$$\Delta y_t^* = \tilde{\nu} + \tilde{\alpha}(\tilde{\beta}' y_{t-1}^* - \tilde{\phi}(t-1)) + \sum_{j=1}^{p-1} \tilde{\Gamma}_j \Delta y_{t-j}^* + \varepsilon_t^*, \quad (3.1)$$

using the estimates obtained under (1) and sampled residuals  $\varepsilon_t^*$  drawn with replacement from the estimated residuals  $\tilde{\varepsilon}_{p+1}, \dots, \tilde{\varepsilon}_T$ . The starting values of the recursion,  $y_1^*, \dots, y_p^*$ , are set equal to  $y_1, \dots, y_p$ .

As usual in the literature, the asterisk refers to quantities related to the bootstrap procedure. Accordingly, the probability measure  $P_{r_0, T}^*$  refers to the conditional distribution of  $y_1^*, \dots, y_T^*$  given the observations  $y_1, \dots, y_T$ . Now, the subscript  $r_0$  indicates that the pseudo observations are generated by imposing the rank  $r_0$ . Moreover, note that the pseudo observations  $y_1^*, \dots, y_T^*$  depend on the sample size  $T$ . Requirement (ii) assures that the generated pseudo observations are indeed  $I(1)$ . If this condition is not satisfied, one may refer to a more appropriate resampling scheme as pointed out by Swensen (2005a).

The bootstrap test version is then computed by running the same test procedures as described above with respect to the pseudo series  $y_1^*, \dots, y_T^*$ . Let  $GLS_1^*(r_0)$  and  $SLT_1^*(r_0)$  be the respective bootstrap test statistics and denote the cumulative distribution functions of  $GLS_1^*(r_0)$  and  $SLT_1^*(r_0)$  by  $F_{r_0; GLS, 1}^*$  and  $F_{r_0; SLT, 1}^*$  respectively. The latter are conditional distributions given the observations  $y_1, \dots, y_T$ . Then, we reject the null hypothesis of  $r_0$  cointegrating relations at some chosen significance level  $\delta$  if

$$1 - F_{r_0; GLS, 1}^*(GLS(r_0)) \leq \delta \quad (3.2)$$

in case of the test by Saikkonen & Lütkepohl (2000b) and if

$$1 - F_{r_0; SLT, 1}^*(SLT(r_0)) \leq \delta \quad (3.3)$$

for the test by Saikkonen et al. (2006). The bootstrap distributions are usually very complicated functions of the observations. Therefore, they are approximated by repeating the bootstrap algorithm a large number of times, say  $M$ .<sup>3</sup> Then, we count the number of instances, for which  $GLS_1^*(r_0) > GLS(r_0)$ , say  $MR_{GLS, 1}$ , and for which  $SLT_1^*(r_0) > SLT(r_0)$ , say  $MR_{SLT, 1}$ . Finally, we reject  $H_0(r_0)$  for the bootstrap version of Saikkonen & Lütkepohl (2000b) if

$$MR_{GLS, 1}/M \leq \delta \quad (3.4)$$

---

<sup>3</sup>Davidson & MacKinnon (2000) suggest and analyse procedures to determine  $M$  in order to assure an appropriate approximation.

and for the bootstrap test of Saikkonen & Lütkepohl (2000b) if

$$MR_{SLT,1}/M \leq \delta, \quad (3.5)$$

respectively. For  $M \rightarrow \infty$ , the expressions (3.4) and (3.5) converge to (3.2) and (3.3) respectively.

The systems cointegration tests are sequentially applied in order to determine the cointegrating rank. Usually, one starts with a cointegration rank of zero under the null hypothesis. The rank chosen is the one for which  $H_0(r_0)$  cannot be rejected the first time given a certain significance level  $\delta$ . Although Algorithm 1 can be shown to be consistent with respect to a specific rank null hypothesis, Swensen (2005a) was not able to prove that Algorithm 1 consistently estimates the cointegrating rank if it is sequentially applied in connection with the Johansen test. Therefore, he considers a modification of the algorithm. The problem to be dealt with is that the bootstrap samples have to be generated for other ranks than the true one if a sequence of tests is performed. We do not explore the issue of sequential testing in the current framework. Nevertheless, we also consider the second algorithm proposed by Swensen (2005a), since it is also valid for a specific rank hypothesis.

### Algorithm 2.

- (1) Estimate model (2.3) by a RR regression under the rank hypothesis  $H_0(r_0) : \text{rk}(\Pi) = r_0$  in order to obtain the estimates  $\tilde{\alpha}$ ,  $\tilde{\beta}$ , and  $\tilde{\phi}$ . The remaining parameters are estimated via OLS from an unrestricted version of (2.3) setting  $r = n$ . Let  $\hat{\nu}$  and  $\hat{\Gamma}_j$  ( $j = 1, \dots, p-1$ ) be the corresponding estimators.

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

$$\tilde{A}(z) = (1 - z)I_n - \tilde{\alpha}\tilde{\beta}'z - \hat{\Gamma}_1(1 - z)z - \dots - \hat{\Gamma}_{p-1}(1 - z)z^{p-1},$$

are equal to 1 or outside the unit circle and whether  $\tilde{\alpha}'_{\perp}\tilde{\Gamma}\tilde{\beta}_{\perp}$  is nonsingular.

- (3) If so, compute the unrestricted residuals  $\hat{\varepsilon}_{p+1}, \dots, \hat{\varepsilon}_T$  based on the OLS estimation in (1).

- (4) Compute  $y_t^*$ ,  $t = p+1, \dots, T$ , recursively from

$$\Delta y_t^* = \hat{\nu} + \tilde{\alpha}(\tilde{\beta}'y_{t-1}^* - \tilde{\phi}(t-1)) + \sum_{j=1}^{p-1} \hat{\Gamma}_j \Delta y_{t-j}^* + \varepsilon_t^*,$$

with sampled residuals  $\varepsilon_t^*$  drawn with replacement from the estimated residuals  $\hat{\varepsilon}_{p+1}, \dots, \hat{\varepsilon}_T$ . The starting values of the recursion,  $y_1^*, \dots, y_p^*$ , are set equal to  $y_1, \dots, y_p$ .

Note, that the pseudo observations and sampled residuals have the same notational form as in Algorithm 1 in order to avoid the introduction of additional symbols. Keep in mind that  $\tilde{\alpha}'_{\perp} \hat{\Gamma} \tilde{\beta}_{\perp}$  is nonsingular is true with probability 1, since the estimator  $\hat{\Gamma}$  is based on the unrestricted OLS estimators  $\hat{\Gamma}_j$  ( $j = 1, \dots, p - 1$ ) (see Swensen 2005a). The bootstrap tests based on Algorithm 2 are abbreviated as  $GLS_2^*(r_0)$  and  $SLT_2^*(r_0)$

Instead of considering Algorithms 1 and 2 one may also think of applying a bootstrap directly to (2.1). However, we have to remember that the unobserved error term  $x_t$  follows a VAR process. If this is ignored such that no finite-dimensional parametric model is used in order to reduce the DGP to independent random sampling, bootstrapping time-series data involves a number of partly unsolved problem (see e.g. Härdle, Horowitz & Kreiß 2003, van Giersbergen 1996). Most of the problems arise because bootstraps are not able to exactly replicate the dependence structure in the data. This is even relevant for stationary time series. Nevertheless, we may also employ a different bootstrap scheme specific to the test by Saikkonen & Lütkepohl (2000b). For example, one can base a bootstrap on the residuals from an estimation of model (2.7). The pseudo observation could then be generated via equations (2.1) and (2.2) using the relevant parameter estimates and exploiting the relationship  $\eta_t = Q' \varepsilon_t$  and the zero initial value assumption for  $x_t$ . However, the simulation results for this algorithm are rather similar to the results for Algorithm 1. Moreover, the theoretical properties will be much more difficult to derive. Therefore, we do not consider this algorithm in more detail.

## 4 Asymptotic Distributions of Bootstrap Tests

In this section we discuss the asymptotic properties of the bootstrap tests  $SLT_1^*(r_0)$ ,  $SLT_2^*(r_0)$ ,  $GLS_1^*(r_0)$  and  $GLS_2^*(r_0)$  by referring to weak convergence in  $P_{r_0, T}$  probability. As described by Swensen (2005a), this means that the conditional distributions converge in  $P_{r_0, T}$  probability given the observations. That is,  $E_{P_{r_0, T}^*}^*[f(X^*)] \rightarrow E_{P_{r_0, T}}[f(X)]$  in probability for all bounded continuous functions  $f$ , where  $E_{P_{r_0, T}^*}^*$  is the expectation conditional on the observations with respect to the measure  $P_{r_0, T}^*$ . In the following, we use  $E^*$  as an abbreviation for  $E_{P_{r_0, T}^*}^*$ .

In the following, we describe how to obtain the asymptotic distributions of the  $SLT_1^*(r_0)$  and  $GLS_1^*(r_0)$  test statistics making use of the results in Swensen (2005a), Saikkonen et al. (2006), and Saikkonen & Lütkepohl (2000b). To this end, it is helpful to briefly review the proof for the asymptotic null distributions of  $SLT(r_0)$  and  $GLS(r_0)$ . The proofs have mainly three steps.

First, it is shown that the estimators of the deterministic parameters  $\hat{\mu}_1$ ,  $\tilde{\mu}_0$ , and  $\tilde{\mu}_1$  have the

following properties.

$$\begin{aligned}
\beta'(\hat{\mu}_1 - \mu_1) &= O_p(T^{-3/2}) \\
(T-p)^{1/2}\beta'_\perp(\hat{\mu}_1 - \mu_1) &\xrightarrow{d} N(0, \beta'_\perp C \Omega C' \beta_\perp) \\
\beta'(\tilde{\mu}_0 - \mu_0) &= O_p(T^{-1/2}) \\
\beta'_\perp(\tilde{\mu}_0 - \mu_0) &= O_p(1) \\
\beta'(\tilde{\mu}_1 - \mu_1) &= O_p(T^{-3/2}) \\
(T-p)^{1/2}\beta'_\perp(\tilde{\mu}_1 - \mu_1) &\xrightarrow{d} N(0, \beta'_\perp C \Omega C' \beta_\perp),
\end{aligned}$$

where  $C = \beta_\perp(\alpha'_\perp \Gamma \beta_\perp)^{-1} \alpha'_\perp$  as before. The precondition for this to hold is that the RR estimators based on the VECM (2.3) are consistently estimated with certain rates. Second, the convergence properties of  $\hat{\mu}_1$ ,  $\tilde{\mu}_0$ , and  $\tilde{\mu}_1$  assure that the adjusted series  $\hat{y}_t^c = y_t - \hat{\mu}_1 t$  and  $\tilde{x}_t = y_t - \tilde{\mu}_0 - \tilde{\mu}_1 t$  can appropriately be used instead of  $y_t^c$  and  $x_t$  taking account of some adjustments. To be precise, an estimation error regarding the trend components remains in terms of the standardized sample moments involving  $\beta'_\perp \hat{y}_t^c$  and  $\beta'_\perp \tilde{x}_t$ . This is the reason why Brownian bridges instead of standard Brownian motions enter the limiting distributions in Theorem 1. Furthermore, note that it suffices that the estimator  $\tilde{\mu}_0$  is only bounded in probability in the direction of  $\beta_\perp$  for proving the asymptotic properties of  $GLS(r_0)$  (see Saikkonen & Lütkepohl 2000b). Thirdly, the asymptotic distributions are proven as in the case of the standard Johansen test (compare Johansen 1995, Chs. 10, 11), but now with respect to (2.8) and (2.10) respectively. Of course, one has to adjust the respective asymptotic expressions.

Saikkonen & Lütkepohl (2000b) do not follow exactly the proof of Johansen (1995) regarding the  $GLS$  test. Instead, they base their considerations on an auxiliary model related to an LM cointegration test. However, the framework of Swensen (2005a) is the direct bootstrap counterpart of Johansen (1995). Therefore, we follow Johansen (1995) in order to easily apply the results of Swensen (2005a).

Lemma 1 in Swensen (2005a) shows that a Granger representation theorem exists for the pseudo data generated by Algorithms 1 or 2 and that the random process  $S_T(s) = (T-p)^{-1/2} \sum_{i=k+1}^{[sT]} \varepsilon_i^*$ ,  $0 \leq s \leq 1$ , converges in  $P_{r_0, T}^*$  distributions toward a  $n$ -dimensional Brownian motion  $W$  with covariance matrix  $\Omega$ . Furthermore, according to Lemmas 2 and 3 in Swensen (2005a), the relevant sample moments based on the bootstrapped series converge weakly in  $P_{r_0, T}$  probability to the same limit expressions as the sample moments defined in terms of the observations. Thus, the limiting distribution results for the asymptotic test also hold for the bootstrapped Johansen test. Note that Lemmas 1-3 are actually proven in conjunction with Algorithm 2. However, they are also valid in case of Algorithm 1 as described in Remark 9 of

Swensen (2005a).

The first important step in combining the results of Swensen (2005a) on the one side and Saikkonen et al. (2006) and Saikkonen & Lütkepohl (2000b) on the other side is to express the bootstrap framework also in terms of the process  $x_t$  and the deterministic components. This is necessary because the test procedures of Saikkonen et al. (2006) and Saikkonen & Lütkepohl (2000b) involve  $x_t$  and  $y_t^c$  and their sample analogs  $\tilde{x}_t$  and  $\hat{y}_t^c$ . Accordingly, the asymptotic derivations refer to these quantities. The goal is to obtain pseudo series  $x_t^*$  and  $y_t^{c*}$  such that  $y_t^* = x_t^* + \mu_0^* + \mu_1^*t = y_t^{c*} + \mu_1^*t$ . This allows one to apply the *SLT* and *GLS* test procedures to the pseudo data  $y_t^*$  while having appropriate bootstrap versions of  $y_t^c$  and  $x_t$ . The quantities  $\mu_0^*$  and  $\mu_1^*$  are estimators of  $\mu_0$  and  $\mu_1$  based on the original observations. They have to follow certain restrictions as described below. Conditional on the observations we may regard  $\mu_0^*$  and  $\mu_1^*$  as parameters.

In a second step, we have to derive Granger representations for  $x_t^*$  and  $y_t^{c*}$ . Based on these representations and the properties of the bootstrap estimators of  $\mu_0^*$  and  $\mu_1^*$ , we can derive the limiting distributions of  $SLT_1^*(r_0)$  and  $GLS_1^*(r_0)$ . The proofs have the same structure as in the case of the asymptotic test procedures and we will, in fact, obtain the same limiting distributions as in Theorem 1.

Let us now consider the bootstrap algorithm related to  $x_t$ . Define the starting values  $x_t^* = y_t - \mu_0^* - \mu_1^*t$  for  $(t = 1, \dots, p)$ . Then, based on  $x_1^*, \dots, x_p^*$ , the pseudo series  $x_t^*$  is recursively generated according to

$$\Delta x_t^* = \tilde{\alpha}\tilde{\beta}'x_{t-1}^* + \sum_{j=1}^{p-1} \tilde{\Gamma}_j \Delta x_{t-j}^* + \varepsilon_t^*, \quad t = p+1, \dots, T, \quad (4.1)$$

where the same RR estimators  $\tilde{\alpha}$ ,  $\tilde{\beta}$ ,  $\tilde{\Gamma}_i$  ( $i = 1, \dots, p$ ), and sampled residuals  $\varepsilon_t^*$  are used as for producing  $y_t^*$ . It can be shown that  $x_t^* + \mu_0^* + \mu_1^*t = y_t^*$  ( $t = 1, \dots, T$ ) if  $\mu_0^*$  and  $\mu_1^*$  satisfy the following conditions

$$\begin{aligned} \tilde{\nu} &= -\tilde{\alpha}\tilde{\beta}'\mu_0^* + \tilde{\Gamma}\mu_1^* = -\tilde{\alpha}\theta^* + \tilde{\Gamma}\mu_1^* \\ \tilde{\phi} &= \tilde{\beta}'\mu_1^*. \end{aligned} \quad (4.2)$$

These conditions exactly correspond to the restrictions underlying the parameters  $\nu$  and  $\phi$  of the VECM (2.3). Similarly, we can uniquely recover  $\theta^* = \tilde{\beta}'\mu_0^*$  and  $\mu_1^*$  from the parameter estimates of (2.3). We obtain  $\theta^* = \tilde{\theta} = (\tilde{\alpha}'\tilde{\alpha})^{-1}\tilde{\alpha}'(\tilde{\Gamma}\hat{\mu}_1 - \tilde{\nu})$  and  $\mu_1^* = \hat{\mu}_1$  as defined in (2.9). However,  $\mu_0^*$  is not identified in the direction of  $\tilde{\beta}_\perp$  by the conditions (4.2). This corresponds to the fact that  $\beta'_\perp\mu_0$  is not identified in (2.3). The reason is that  $\nu$  does not change if  $\beta'_\perp\mu_0$  changes

(compare Saikkonen & Lütkepohl 2000b). Accordingly, the bootstrap level parameter  $\mu_0^*$  is not identified since the bootstrap algorithm 1 is based on the estimated VECM for  $y_t$ . To obtain a unique  $\mu_0^*$  and, thus, a unique version of  $x_t^*$  we have to suggest an estimator for  $\theta_* = \beta'_\perp \mu_0$ , say  $\tilde{\theta}_*$ . The asymptotic analysis presented later on shows that it is sufficient to consider an estimator that is bounded in  $P_{r_0, T}$  probability.

The bootstrap parameter  $\theta_*$  can be estimated from the first  $p$  observations based on (2.7), but only from the first  $p$  observations since  $\tilde{H}_{0t} \tilde{\beta}_\perp (\tilde{\beta}'_\perp \tilde{\beta}_\perp)^{-1} = 0$  for  $t \geq p + 1$ . Therefore, the estimator cannot be consistent but is bounded in  $P_{r_0, T}$  probability (see further below). To be precise, (2.7) is rewritten as

$$\tilde{Q}' \tilde{A}(L) y_t = \tilde{Q}' \tilde{H}_{0t} \tilde{\beta}_\perp (\tilde{\beta}'_\perp \tilde{\beta}_\perp)^{-1} \tilde{\beta}'_\perp \mu_0 + \tilde{Q}' \tilde{H}_{0t} \tilde{\beta} (\tilde{\beta}' \tilde{\beta})^{-1} \tilde{\beta}' \mu_0 + \tilde{Q}' \tilde{H}_{1t} \mu_1 + \tilde{\eta}_t.$$

Following Saikkonen & Lütkepohl (2000b, Proof of Theorem 1), the LS estimator  $\tilde{\theta}_* = \tilde{\beta}'_\perp \tilde{\mu}_0$  of  $\theta_* = \beta'_\perp \mu_0$  is derived from this regression model. We do not give an explicit expression for  $\tilde{\theta}_*$  because it will look rather complicated. Since  $\beta_\perp$  is consistently estimated by  $\tilde{\beta}_\perp$ ,  $\tilde{\theta}_*$  can also be regarded as an estimator of  $\theta_*$ . Using the decomposition (compare Lütkepohl & Saikkonen 2000, Equation (4.6))

$$\hat{\mu}_0 = \tilde{\beta}_\perp (\tilde{\beta}'_\perp \tilde{\beta}_\perp)^{-1} \tilde{\theta}_* + \tilde{\beta} (\tilde{\beta}' \tilde{\beta})^{-1} \tilde{\theta} \quad (4.3)$$

a version of  $\mu_0^*$  is obtained that satisfies (4.2). Furthermore, we define  $\theta_*^* = \tilde{\theta}_* = \tilde{\beta}'_\perp \mu_0^*$  in line with  $\theta^* = \tilde{\theta} = \tilde{\beta}' \mu_0^*$ . Hence, we have the following bootstrap model framework summarized in Assumption 2.

**Assumption 2.** Let the bootstrap series  $y_t^*$  be generated by  $y_t^* = x_t^* + \mu_0^* + \mu_1^* t$ ,  $t = 1, \dots, T$ , where  $x_t^*$  is determined by (4.1). Moreover,  $\mu_0^* = \hat{\mu}_0$  and  $\mu_1^* = \hat{\mu}_1$  are explicit parameter vectors conditional on the observations which satisfy the conditions (4.2).  $\square$

One may also generate  $y_t^*$  by  $y_t^* = y_t^{*c} + \mu_1^* t$  using the pseudo series  $y_t^{*c}$ . The latter one can be obtained via a corresponding VECM. This approach, however, is only useful with respect to the  $SLT^*$  tests since the  $GLS^*$ -framework requires a bootstrap equivalent of  $\mu_0^*$ . Moreover, note that the deterministic parameters chosen for the initialization of the bootstrap series  $x_t^*$  and used for  $y_t^* = x_t^* + \mu_0^* + \mu_1^* t$  ( $t = 1, \dots, T$ ) are exactly the same. Initializing the bootstrap with the true parameters  $\mu_0$  and  $\mu_1$ , i.e. using  $x_1, \dots, x_p$  as starting values, and adding the estimated deterministic  $\mu_0^* + \mu_1^* t$  to  $x_t^*$  does not produce  $y_t^*$  ( $t = 1, \dots, T$ ) of Algorithm 1. Instead, some other bootstrap version of  $y_t$  is obtained, which the test procedures would have been applied to.

However, this approach is not possible since we do not know  $x_1, \dots, x_p$ . Therefore, we refer to the algorithm described above, which is a feasible framework to generate bootstrap observations  $x_t^*$  and  $y_t^{c*}$  that have the following Granger representations stated in the next Lemma.

**Lemma 1.** Under Assumptions 1 and 2, the bootstrap elements  $x_t^*$  and  $y_t^{c*}$  have the following representations

$$x_t^* = \tilde{C} \sum_{i=p+1}^t \varepsilon_i^* + \sqrt{T} R_{t,T}^{0*}, \quad t = p+1, \dots, T, \quad (4.4)$$

and

$$y_t^{c*} = \tilde{C} \sum_{i=p+1}^t \varepsilon_i^* + \tau_0^* + \sqrt{T} R_{t,T}^{c*}, \quad t = p+1, \dots, T, \quad (4.5)$$

where for all  $\eta > 0$ ,  $P^*(\max_{p+1 \leq t \leq T} |R_{t,T}^{c*}| > \eta) \rightarrow 0$  and  $P^*(\max_{p+1 \leq t \leq T} |R_{t,T}^{0*}| > \eta) \rightarrow 0$  in  $P_{r_0,T}$  probability as  $T \rightarrow \infty$ ,  $|\cdot|$  is the Euclidean norm,  $\tilde{C} = \tilde{\beta}_\perp (\tilde{\alpha}'_\perp \tilde{\Gamma} \tilde{\beta}_\perp)^{-1} \tilde{\alpha}'_\perp$ , and  $\tau_0^* = \tilde{\beta} \theta^*$ . Moreover,  $E^*[\varepsilon_t^* \varepsilon_t^{*'}] = \tilde{\Omega}_T \rightarrow \Omega$  in  $P_{r_0,T}$  probability as  $T \rightarrow \infty$ .  $\square$

The proof of Lemma 1 in Appendix A shows that (4.4) is obtained from Swensen (2005a, Lemma1) by setting the parameter estimates related to the deterministic terms to zero. This is the case although  $\tilde{\beta}'_\perp x_p^*$  enters the term  $\sqrt{T} R_{t,T}^{0*}$ . Note that  $\tilde{\beta}'_\perp (x_p^* - x_p)$  is only bounded in  $P_{r_0,T}$  probability since  $x_p^* = x_p - (\mu_0^* - \mu_0) - (\mu_1^* - \mu_1)p$  and  $\tilde{\beta}'_\perp \mu_0^* = \theta_*^* = \theta_* + O_p(1)$  (see Saikkonen & Lütkepohl 2000b, Eq. (A2)). As explained in the proof, however,  $\tilde{\beta}'_\perp x_p^*$  needs only to be bounded in  $P_{r_0,T}$  probability to assure that the result for  $R_{t,T}^{0*}$  holds. Finally, the representation for  $y_t^{c*}$  is derived by adding  $\mu_0^*$  to  $x_t^*$ .

Given Lemma 1, we can now discuss the derivation of the limiting distribution of the bootstrap cointegration tests. Let us start with the  $SLT_1^*$  test. Applying this test procedure to the bootstrap data  $y_t^*$  delivers the estimator  $\hat{\mu}_1^* = \tilde{\beta}^* (\tilde{\beta}^{*'} \tilde{\beta}^*)^{-1} \tilde{\phi}^* + \tilde{\beta}'_\perp (\tilde{\beta}'_\perp \tilde{\beta}_\perp^*)^{-1} \tilde{\phi}_*^*$  for the 'bootstrap trend parameter'  $\mu_1^*$ , where  $\tilde{\beta}^*$ ,  $\tilde{\beta}_\perp^*$ ,  $\tilde{\phi}^*$ , and  $\tilde{\phi}_*^*$  are RR estimators based on the VECM (3.1). Accordingly, the series  $y_t^c = y_t - \mu_1 t$  and  $\hat{y}_t^c = y_t - \hat{\mu}_1 t$  have their respective bootstrap counterparts defined by  $y_t^{c*} = y_t^* - \mu_1^* t$  and  $\hat{y}_t^{c*} = y_t^* - \hat{\mu}_1^* t$ . Consequently,  $y_t^{c*}$  and  $\hat{y}_t^{c*}$  replace  $y_t^c$  and  $\hat{y}_t^c$  in the proof of the limiting distribution. From these relations it is obvious that  $\hat{\mu}_1^*$  is the relevant estimator of  $\mu_1^*$  and that the asymptotic properties of  $\hat{\mu}_1^*$  in relation to  $\mu_1^*$  have to be determined. To this end, we first derive the necessary asymptotic results for the RR estimators of the parameters in the VECM (3.1) conditional on the observations, i.e. with respect to the measure  $P_{r_0,T}^*$ . Given that, we obtain the conditional convergence and distributional properties of the estimator  $\hat{\mu}_1^*$  in relation to  $\mu_1^*$ . These properties corresponds exactly to the ones of  $\hat{\mu}_1$  re-



garding  $\mu_1$ . Hence, we can appropriately work with the adjusted bootstrap series  $\hat{y}_t^{c*}$  and apply an LR-type test to

$$\Delta \hat{y}_t^{c*} = \tilde{\alpha}(\tilde{\beta}' \hat{y}_{t-1}^{c*} - \theta^*) + \sum_{j=1}^{p-1} \tilde{\Gamma}_j \Delta \hat{y}_{t-j}^{c*} + e_t^{c*}, \quad (4.6)$$

where  $e_t^{c*}$  is an error term defined after (A.40) in Appendix A. Then, using Lemma 1 and the results of Swensen (2005a) we derive the asymptotic properties of the relevant quantities regarding the measure  $P_{r_0, T}$ . To this end, we follow again the setup of Johansen (1995, Chs. 10, 11) and obtain the limiting distribution given in Proposition 1. The details of the proof can be found in Appendix A.

**Proposition 1.** Under Assumptions 1 and 2 and if  $H_0(r_0)$  is true, then

$$P_{r_0, T}^*(SLT_1^*(r_0) \leq x) \longrightarrow P(\text{tr}(D_{SLT}) \leq x),$$

in  $P_{r_0, T}$  probability for all  $x$  as  $T \rightarrow \infty$ , where  $D_{GLS}$  and  $D_{SLT}$  are defined as in Theorem 1.  $\square$

Let us turn to the  $GLS_1^*$  test. First, we have to discuss whether the inconsistency of  $\tilde{\beta}'_{\perp} x_p^*$  with respect to  $\tilde{\beta}'_{\perp} x_p^*$  and the inconsistent estimation of the level of  $y_t^*$  have an impact on the limiting distribution of  $GLS_1^*(r_0)$ . Given that  $\tilde{\beta}'_{\perp} x_p^*$  is bounded in  $P_{r_0, T}$  probability it should be clear that it is bounded conditionally on the observations even if  $T \rightarrow \infty$ . In addition, we will not be able to estimate the level of  $y_t^*$  consistently in the direction of  $\tilde{\beta}_{\perp}$  conditional on the observations, i.e.  $\tilde{\theta}_*^* = \tilde{\beta}'_{\perp} \tilde{\mu}_0^*$  is only a bounded estimator of  $\theta_*^*$  in  $P_{r_0, T}^*$  probability. Thus, with respect to the estimated version of  $x_t^*$ , say  $\tilde{x}_t^*$ , an error that is bounded conditional on the observation and a bounded error in  $P_{r_0, T}^*$  probability are added up: the estimation errors regarding  $\tilde{\beta}'_{\perp} (x_p^* - x_p)$  and  $\tilde{\beta}'_{\perp} (\tilde{x}_t^* - x_t^*)$ . However, this does not cause additional problems since the joint error remains bounded in  $P_{r_0, T}^*$  probability. In this case, the applied normalizations of the sample moments involving  $\tilde{\beta}'_{\perp} \tilde{x}_t^*$  assure that the joint error has no asymptotic effect conditional on the observations. Accordingly, the limit expressions of the respective sample moments regarding the measure  $P_{r_0, T}$  are not either affected. Hence, it suffices if the estimation errors regarding the levels of  $\beta'_{\perp} y_t$  and  $\tilde{\beta}'_{\perp} y_t^*$  are bounded in  $P_{r_0, T}$  and  $P_{r_0, T}^*$  probability respectively.

Applying the GLS test procedure to the bootstrap data results in the following framework

$$y_t^* = \mu_0^* + \mu_1^* t + x_t^*, \quad (4.7)$$

$$\tilde{A}(L)y_t^* = \tilde{A}(L)\mu_0^* + \tilde{A}(L)\mu_1^* t + \varepsilon_t^*, \quad (4.8)$$

$$\tilde{Q}'\tilde{A}(L)y_t^* = \tilde{Q}'\tilde{H}_{0t}\mu_0^* + \tilde{Q}'\tilde{H}_{1t}\mu_1^* + \eta_t^*, \quad \eta_t^* = \tilde{Q}'\varepsilon_t^*, \quad (4.9)$$

$$\tilde{Q}^{*'}\tilde{A}^*(L)y_t^* = \tilde{Q}^{*'}\tilde{H}_{0t}^*\mu_0^* + \tilde{Q}^{*'}\tilde{H}_{1t}^*\mu_1^* + \tilde{\eta}_t^*, \quad (4.10)$$

which comprises the bootstrap analogues to (2.1), (2.5), (2.6) and (2.7) respectively. The error terms  $\tilde{\eta}_t^*$  are defined in Appendix A after equation (A.50). Note, that we apply the estimators  $\tilde{Q}^*$ ,  $\tilde{A}^*(L)$ ,  $\tilde{H}_{it}^*$  ( $i = 1, 2$ ) in (4.10), which are computed using the generated pseudo data. Thus, we estimate  $\mu_0^*$  and  $\mu_1^*$  from (4.10) and denote the respective estimators by  $\tilde{\mu}_0^*$  and  $\tilde{\mu}_1^*$ .

The time index for (4.7)-(4.10) covers  $t = 1, \dots, T$ . Thus, (4.8) is another representation of the pseudo observations with respect to the bootstrap error terms  $\varepsilon_t^*$  ( $t = p + 1, \dots, T$ ) of Algorithm 1 and the implicit bootstrap error terms  $\varepsilon_1^*, \dots, \varepsilon_p^*$ . The latter are obtained by assuming  $x_t^* = 0$  for  $t = 0, \dots, -p + 1$ . In fact, we require bootstrap error terms for the first  $p$  data points and need to know their properties in order to derive the properties of the estimator  $\tilde{\theta}_*^*$  conditional on the observations. Corresponding to the setup of the asymptotic test, the bootstrap level parameter can be estimated in the direction of  $\tilde{\beta}_\perp$  only from the first  $p$  (transformed) bootstrap observations. Given that

$$x_t^* = x_t - (\mu_0^* - \mu_0) - (\mu_1^* - \mu_1)t, \quad t = 1, \dots, p, \quad (4.11)$$

and the zero initial value assumption for both  $x_t$  and  $x_t^*$ , we have

$$\begin{aligned} \tilde{A}(L)x_t^* &= \varepsilon_t^* = \tilde{A}(L)x_t - \tilde{H}_{0t}(\mu_0^* - \mu_0) - \tilde{H}_{1t}(\mu_1^* - \mu_1), & t = 1, \dots, p \\ &= \varepsilon_t - (A(L) - \tilde{A}(L))x_t - \tilde{H}_{0t}(\mu_0^* - \mu_0) - \tilde{H}_{1t}(\mu_1^* - \mu_1). \end{aligned} \quad (4.12)$$

Obviously, the implicit bootstrap error terms  $\varepsilon_1^*, \dots, \varepsilon_p^*$  are not random conditional on the observations although their true values are not known. Hence,  $E^*(\varepsilon_t^*) = \varepsilon_t^* \neq 0$  in general for  $t = 1, \dots, p$ . However, the non-randomness should not cause any problems for two reasons. First, we only require  $\tilde{\theta}_*^*$  to be bounded in  $P_{r_0, T}^*$  probability. Note that  $\tilde{\theta}_*^*$  depends on  $\varepsilon_1^*, \dots, \varepsilon_p^*$  but is still random conditional on the observations because the transformation in (4.10) involves estimators that are random conditional on the observations. Second,  $\tilde{\theta}_*^*$  has no effect on the asymptotic properties of the other estimators  $\tilde{\beta}^{*'}\tilde{\mu}_0^*$  and  $\tilde{\mu}_1^*$  (compare Saikkonen & Lütkepohl 2000b, Proof of Theorem 1). Therefore, we can ignore the first  $p$  observations regarding these estimators.

Now, the proof regarding  $GLS_1^*(r_0)$  runs as follows. Based on the necessary asymptotic results for the RR estimators of the parameters in the VECM (3.1) we obtain the conditional convergence and distributional properties of the estimators  $\tilde{\mu}_0^*$  and  $\tilde{\mu}_1^*$ . These properties correspond exactly to the ones of  $\tilde{\mu}_0$  and  $\tilde{\mu}_1$  in relation to  $\mu_0$  and  $\mu_1$ , e.g.  $\tilde{\beta}'_{\perp}(\tilde{\mu}_0^* - \mu_0^*) = O_p^*(1)$  is obtained. Hence, we can appropriately work with the adjusted bootstrap series  $\tilde{x}_t^* = y_t^* - \tilde{\mu}_0^* - \tilde{\mu}_1^*t$  and apply an LR-type test to

$$\Delta\tilde{x}_t^* = \tilde{\alpha}\tilde{\beta}'\tilde{x}_{t-1}^* + \sum_{j=1}^{p-1} \tilde{\Gamma}_j\Delta\tilde{x}_{t-j}^* + e_t^*, \quad (4.13)$$

where  $e_t^*$  is an error term defined in the Appendix. Then, as for  $SLT_1^*(r_0)$ , we refer to Lemma 1 and the results of Swensen (2005a) and obtain the asymptotic properties of the relevant quantities regarding the measure  $P_{r_0,T}$  leading to

**Proposition 2.** Under Assumptions 1 and 2, assuming  $x_t^* = 0$  for  $t = 0, \dots, -p + 1$ , and if  $H_0(r_0)$  is true, then

$$P_{r_0,T}^*(GLS_1^*(r_0) \leq x) \longrightarrow P(\text{tr}(D_{GLS}) \leq x),$$

in  $P_{r_0,T}$  probability for all  $x$  as  $T \rightarrow \infty$ , where  $D_{GLS}$  and  $D_{SLT}$  are defined as in Theorem 1.  $\square$

The details of the proof of Proposition 2 are given in Appendix A.

When applying Algorithm 2 we have to use the unrestricted estimators  $\hat{\nu}$  and  $\hat{\Gamma}_i$  ( $i = 1, \dots, p - 1$ ) and the residuals  $\hat{\varepsilon}_t$  in order to generate  $x_t^*$ . Furthermore, the conditions for the estimators  $\theta^*$  and  $\mu_1^*$  have to be adjusted. The first equation of (4.2) now reads as  $\hat{\nu} = -\tilde{\alpha}\theta^* + \hat{\Gamma}\mu_1^*$  such that  $\hat{\theta} = (\tilde{\alpha}'\tilde{\alpha})^{-1}\tilde{\alpha}'(\hat{\Gamma}\hat{\mu}_1 - \hat{\nu})$  can be used for  $\theta^*$  and to obtain an appropriate version of  $\mu_0^*$ . Accordingly, we adjust Assumption 2 and Lemma 1. Otherwise, we go through to the same steps as above to show that  $SLT_2^*(r_0)$  and  $GLS_2^*(r_0)$  have the same limiting distribution as  $SLT_1^*(r_0)$  and  $GLS_1^*(r_0)$  respectively.

## 5 Monte Carlo Study

In this section we analyse and compare the small sample properties of the bootstrap and asymptotic test versions. To this end, we first use a VAR(1) process suggested by Toda (1994, 1995), which was also applied by a couple of other authors in order to study the properties of cointegration tests (e.g. Hubrich et al. 2001, Trenkler 2002). The process has the following structure

$$x_t = A_1 x_{t-1} + \varepsilon_t = \begin{bmatrix} A & 0 \\ 0 & I_{n-r} \end{bmatrix} x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \text{iid } N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} I_r & \Theta \\ \Theta' & I_{n-r} \end{bmatrix} \right), \quad (5.1)$$

where  $A = \text{diag}(a_1, \dots, a_r)$  is a  $(r \times r)$  diagonal matrix and  $\Theta$  is a  $(r \times (n-r))$  matrix describing the correlation between the stationary and nonstationary components. Other VAR(1) processes of interest can be obtained from (5.1) by linear transformations which leave the analysed LR tests invariant (see Toda 1994, 1995). Specifically, I will consider bivariate processes. In this case, if  $r = 0$ ,  $A$  and  $\Theta$  vanish and the process consists of two nonstationary components. If the cointegrating rank is 1,  $A = a_1$  with  $|a_1| < 1$  and  $\Theta = \theta$  is a scalar which represents the instantaneous correlation between the two components.

In the simulations the parameter values of the constant and the linear trend are zero, i.e.,  $\mu_i = 0$  ( $i = 0, 1$ ) throughout. This is done because the results of the tests, with the exception of the bootstrap tests based on Algorithm 2, are invariant to the specific parameter values of the deterministic components they allow for. The bootstrap tests based on Algorithm 2 are not invariant to the value of  $\mu_0$  since the constant is estimated using an unrestricted VAR model. In other words, the constant is not estimated under the appropriate rank specification so that changes in  $\mu_0$  are not perfectly captured by the corresponding estimator. Nevertheless, the differences in the results in response to changes in  $\mu_0$  are very small. Therefore, we keep  $y_t = x_t$  for all tests in the simulations.

The process (5.1) is a rather simple one. On the one hand, this allows to control the parameters of interest and, thereby, the properties of the DGP rather easily. On the other hand, (5.1) may be a quite unrealistic setup compared to empirical applications where richer autoregressive dynamics and error term cross-correlation structures are often present. Therefore, we use a data-based DGP in order to study the test properties for more complex processes. Our DGP refers to an empirical study by King, Plosser, Stock & Watson (1991), in which they analyse a small macroeconomic model for the U.S.. It consists of per-capita private real GNP (total GNP less real total government purchases of goods and services), per-capita real consumption, and per-capita gross private domestic fixed investment. Based on their recommendations we estimate a VECM with one lag and two cointegrating relationships using quarterly data in logarithms for the period 1949:1-1988:4. In other words, we consider a three-dimensional VAR(2) model. The two cointegrating relations describe the differences between GNP on the one side and consumption and investment on the other side. Hence, we specify restricted cointegrating relationships as implied by the real-business-cycle model to which King et al. (1991) refer. In order to re-

duce the impact of estimates of possibly insignificant parameters, a subset VECM with zero restrictions has been estimated. The subset restrictions have been searched for by a *Top-Down* strategy employing the Akaike Information Criterion (AIC) (see e.g. Brüggemann 2004, Chapter 3). Using the AIC avoids the occurrence of too many zero restrictions. All computations have been done with the econometric software program JMULTi (Lütkepohl & Krätzig 2004). We obtain the following estimated process, which is used for the simulations<sup>4</sup>

$$\begin{aligned} \Delta y_t = & \begin{pmatrix} -0.038 \\ -0.186 \\ 0.032 \end{pmatrix} + \begin{pmatrix} 0 & -0.026 \\ 0.217 & -0.150 \\ 0.126 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \end{pmatrix} y_{t-1} + \begin{pmatrix} 0 & 0 & 0.154 \\ 0 & 0.282 & 0.660 \\ 0.272 & 0.162 & 0 \end{pmatrix} \Delta y_{t-1} \\ & + \varepsilon_t, \quad \varepsilon_t \sim \text{iid } N \left( \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, 10^{-4} \begin{bmatrix} 0.588 & 0.821 & 0.465 \\ 0.821 & 4.870 & 1.688 \\ 0.465 & 1.688 & 1.376 \end{bmatrix} \right). \end{aligned} \quad (5.2)$$

The computations for the simulation study are performed by using programs written in GAUSS V6 for Windows. The RNDNS function with a fixed seed has been used to generate standard normally distributed random numbers. The number of replications is  $R = 1000$ , i.e. we generate 1000 different sets of multivariate random numbers. Note that we use the same random numbers for each set of parameter values. Accordingly, the test statistics are applied to the same time series for a given set of parameter values. Thus, the rejection frequencies reported later on are not independent, but they can be compared directly. We consider sample sizes of  $T = 50$ ,  $T = 100$ , and  $T = 200$ . In addition to these sample sizes we use a zero starting value for the two components of the Toda processes (5.1). With respect to the data-dependent KPSW-DGP (5.2) two starting values for each of the three series are taken from the original data.

The simulations are performed with respect to 1%, 5%, and 10% significance levels. However, the presentation focuses on the 5% level, since the relative results are very similar for the other significance levels. The rejection frequencies of the asymptotic *GLS* and *JOH* tests are based on asymptotic critical values computed from response surfaces given in Trenkler (2004) and Doornik (1998) respectively. These response surfaces make use of the fact that the asymptotic null distributions can be quite well approximated by Gamma distributions. The coefficients of a corresponding response surface for the *SLT* test are presented in Appendix B.

---

<sup>4</sup>For the simulation program the parameters are stored as 8-bit, double precision. This assures that the characteristic roots obey the restrictions imposed on the DGP.

We use the response surface critical values, since they deliver a more accurate approximation of the tests' asymptotic properties than the standard tabulated critical values (see Doornik 1998). With respect to the bootstrap tests, we use the upper  $\delta\%$ -quantile ( $\delta = 1, 5, 10$ ) of the empirical bootstrap distribution based on  $M = 1000$  bootstrap replications. We hope that this number of replications is large enough for a reasonable approximation of the relevant bootstrap distributions  $F_{r_0;GLS,1}^*$ ,  $F_{r_0;SLT,1}^*$ , etc..

Note, that the tests were not performed sequentially in the simulations. Thus, the results for testing  $H_0(1) : \text{rk}(\Pi) = 1$  are not conditioned on the outcome of the test of  $H_0(0) : \text{rk}(\Pi) = 0$  etc.. Finally, we do not report size adjusted rejection frequencies with respect to the power of the tests since such an adjustment is not possible in applied work.

The results for different versions of the bivariate Toda-DGP (5.1) are collected in Tables 1 and 2 and for the data-dependent KPSW-DGP (5.2) in Table 3. Tables 1 and 3 (Panel A) contain the rejection frequencies of the tests for a correct null hypothesis. We will use the term *size* with respect to these frequencies, although this use does not coincide with the exact definition of the size of a test. The definition would require to maximize the power function over the whole parameter space associated with the null hypothesis. This parameter space contains the nuisance parameters in the test statistics' small sample distributions. However, such a maximization is not done in our simulation framework.

In order to evaluate whether the tests keep the nominal size and whether the rejection frequencies of two tests are significantly different we may refer to an approach suggested by Paruolo (2001). He notes that for an increasing number of replications  $R$ , the Monte Carlo rejection frequencies of two tests  $(\hat{p}_1, \hat{p}_2)$  converge to a bivariate normal distribution

$$\sqrt{R} \left( \begin{bmatrix} \hat{p}_1 \\ \hat{p}_2 \end{bmatrix} - \begin{bmatrix} p_1 \\ p_2 \end{bmatrix} \right) \xrightarrow{d} N \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} p_1(1-p_1) & \tau \\ \tau & p_2(1-p_2) \end{bmatrix} \right), \quad (5.3)$$

where  $p_1$  and  $p_2$  are the true rejection probabilities of the tests with respect to the DGP considered. These are, in fact, the probabilities we aim to approximate by the Monte Carlo experiment. For the covariance we have  $\tau = \gamma - p_1 p_2$ , where  $\gamma$  is the probability of joint rejections of both tests.

Based on (5.3) we test whether the tests keep the nominal size of 5%. It follows that the standard error of the estimator  $\hat{p}_1$  for the true rejection probability  $p_1$  is  $s_{p_1} = \sqrt{p_1(1-p_1)/R}$ . Thus, if a Monte Carlo rejection frequency for a true null rank hypothesis is outside the band  $[0.0365, 0.0635]$ , it is not in line with the nominal size level of 5%.

We also use (5.3) for a comparison of the rejection frequencies of the corresponding asymptotic and bootstrap tests. First note that  $\sqrt{R}(\hat{p}_1 - \hat{p}_2)$  has an asymptotic variance equal to  $p_1(1 - p_1) + p_2(1 - p_2) - 2\tau$ . Using the Monte Carlo rejection frequencies  $\hat{p}_1$ ,  $\hat{p}_2$ , and  $\hat{\tau}$  as (consistent) estimators we test whether  $p_1$  and  $p_2$  are equal by considering the statistic

$$z = \left( \frac{R}{\hat{p}_1(1 - \hat{p}_1) + \hat{p}_2(1 - \hat{p}_2) - 2\hat{\tau}} \right)^{\frac{1}{2}} (\hat{p}_1 - \hat{p}_2) \xrightarrow{d} N(0, 1), \quad (5.4)$$

(see Paruolo 2001). We can apply (5.4) because all tests are based on the same generated time series for a given set of parameter values. However, (5.3) and (5.4) ignore the uncertainty regarding the computation of the critical values based on the bootstrap and response surface simulations. Therefore, we should regard (5.4) only as a test approximation, since we are not able to incorporate the mentioned simulation uncertainty into the statistic  $z$ .

To decide whether a bootstrap test has a preferable small sample power, we just have to test whether its power is significantly higher than the small sample power of the corresponding asymptotic test version. Of course, a significant deterioration is tested for in an analogous way. In order to study the relative size properties we have to take account of the nominal significance level as the benchmark. Referring to a 5% nominal level, bootstrapping significantly reduces the difference between the true and nominal probabilities of rejecting a correct null hypothesis in the following three cases.

- a) *Both size values are not larger than 0.05.* The size of the bootstrap test is significantly higher than the size of the asymptotic test.
- b) *Both size values are not smaller than 0.05.* The size of the bootstrap test is significantly lower than the size of the asymptotic test.
- c) *One size value is above and the other one is below 0.05.* The size values differ significantly and the size of the asymptotic test is significantly different from the nominal level while the size of the bootstrap test is not.

In c) one may skip the requirement that the size values differ significantly. Of course, one can setup a corresponding set of conditions for a significant deterioration in the tests' sizes. We apply a 5% level to test for significant differences between two rejection frequencies.

With respect to the Toda-DGPs we see in Table 1 that significant improvements in the tests' sizes are possible if the bootstrap test versions are applied. This is especially apparent for processes with a true cointegrating rank of  $r = 1$  for which the sizes of the asymptotic tests quite

**Table 1.** Rejection Frequencies of Tests for Bivariate Toda-DGP (5.1), VAR Order  $p = 1$ , Nominal Significance Level 0.05

	Panel A: $a_1 = 1 (r = 0), H_0 : r_0 = 0$			Panel B: $a_1 = 0.9 (r = 1), H_0 : r_0 = 1$			Panel C: $a_1 = 0.7 (r = 1), H_0 : r_0 = 1$		
	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$
	<i>GLS</i>	0.049	0.041	0.048	<b>0.019</b>	<b>0.023</b>	<b>0.025</b>	0.047	<b>0.034</b>
<i>GLS</i> * <sub>1</sub>	0.047	0.039	0.052	<b>0.028</b> <sup>+</sup>	0.040 <sup>+</sup>	0.040 <sup>+</sup>	0.056	0.042 <sup>+</sup>	0.057
<i>GLS</i> * <sub>2</sub>	0.050	0.039	0.050	0.042 <sup>+</sup>	0.060 <sup>+</sup>	<b>0.072</b>	<b>0.067</b> <sup>-</sup>	0.043 <sup>+</sup>	0.063
<i>SLT</i>	0.058	0.064	0.059	<b>0.031</b>	0.049	<b>0.066</b>	<b>0.067</b>	0.046	0.062
<i>SLT</i> * <sub>1</sub>	0.052 <sup>+</sup>	0.056 <sup>+</sup>	0.056	<b>0.028</b>	0.042 <sup>-</sup>	0.059 <sup>+</sup>	0.061 <sup>+</sup>	0.042	0.058 <sup>+</sup>
<i>SLT</i> * <sub>2</sub>	0.052 <sup>+</sup>	0.058 <sup>+</sup>	0.055	<b>0.026</b> <sup>-</sup>	0.042 <sup>-</sup>	0.058 <sup>+</sup>	0.057 <sup>+</sup>	0.044	0.060
<i>JOH</i>	0.060	0.052	<b>0.067</b>	<b>0.021</b>	<b>0.034</b>	<b>0.069</b>	<b>0.066</b>	0.063	<b>0.070</b>
<i>JOH</i> * <sub>1</sub>	0.052 <sup>+</sup>	0.048	0.062 <sup>+</sup>	<b>0.020</b>	<b>0.032</b>	<b>0.064</b> <sup>+</sup>	0.051 <sup>+</sup>	0.056 <sup>+</sup>	<b>0.067</b>
<i>JOH</i> * <sub>2</sub>	0.056	0.053	0.062 <sup>+</sup>	<b>0.018</b>	<b>0.030</b>	0.062 <sup>+</sup>	0.055 <sup>+</sup>	0.055 <sup>+</sup>	<b>0.065</b> <sup>+</sup>

*Note:* Bold entries indicate significant size differences with respect to a nominal level of 5%. Significant improvements of the bootstrap tests' rejection rates with respect to the corresponding asymptotic tests' rejections are indicated by <sup>+</sup> and significant deteriorations by <sup>-</sup>. Significance refers to a 5% level. The tests are based on an approximation ignoring the uncertainty involved in the bootstrap simulations and the computation of the critical values.



**Table 2.** Rejection Frequencies of Tests for Bivariate Toda-DGP (5.1) with True Cointegrating Rank  $r = 1$ , Rank under  $H_0$  is  $r_0 = 0$ , VAR Order  $p = 1$ ,  $\theta = 0.8$ , Nominal Significance Level 0.05

	Panel A: $a_1 = 0.9$			Panel B: $a_1 = 0.7$		
	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$
$GLS$	0.153	0.474	0.928	0.722	0.964	1.000
$GLS_1^*$	0.152	0.480	0.929	0.722	0.963	1.000
$GLS_2^*$	0.155	0.471	0.927	0.723	0.963	1.000
$SLT$	0.175	0.518	0.963	0.801	0.993	1.000
$SLT_1^*$	0.161 <sup>-</sup>	0.503 <sup>-</sup>	0.961	0.783 <sup>-</sup>	0.992	1.000
$SLT_2^*$	0.155 <sup>-</sup>	0.501 <sup>-</sup>	0.959	0.778 <sup>-</sup>	0.993	1.000
$JOH$	0.138	0.395	0.929	0.738	0.998	1.000
$JOH_1^*$	0.123 <sup>-</sup>	0.387	0.928	0.715 <sup>-</sup>	0.997	1.000
$JOH_2^*$	0.126 <sup>-</sup>	0.375 <sup>-</sup>	0.928	0.720 <sup>-</sup>	0.998	1.000

*Note:* Significant improvements at the 5% level of the bootstrap tests' rejection rates with respect to the corresponding asymptotic tests' rejections are indicated by <sup>+</sup> and significant deteriorations by <sup>-</sup>. The tests are based on an approximation ignoring uncertainty involved in the bootstrap simulations and the computation of the critical values.

often deviate significantly from the 5% level (see Panels B and C in Table 1). By contrast, for DGPs with  $r = 0$  the sizes of the asymptotic test are much more in line with the nominal level. Nevertheless, we find that the bootstrap tests reject sometimes too often such that significant deteriorations may occur as in case of  $GLS_2^*$  (see Panels B and C in Table 1).

We have also performed simulations with  $a_1$  equal to 0.8, 0.6 and 0.5 in order to evaluate whether the parameter  $a_1$  has a systematic impact on whether and how the application of bootstrap tests may improve the tests' sizes. Our overall conclusion is that the values of  $a_1$  seem to be less important in these respects than e.g. the sample size  $T$ . Interestingly, the most pronounced improvements occur for  $T = 50$  in case of  $SLT_1^*$ ,  $SLT_2^*$ ,  $JOH_1^*$ , and  $JOH_2^*$ , although we also observe a number of significant improvements in case of  $T = 200$ . Regarding  $GLS_1^*$  and  $GLS_2^*$  most of the improvements are observed for  $T = 100$ , while  $GLS_2^*$  tend to overreject for larger values of  $a_1$ .

Table 2 displays the small sample power for the case of a true cointegrating rank of  $r = 1$  ( $a_1 = 0.9$  and  $a_1 = 0.7$ ) if the null hypothesis  $H_0(r_0) : r_0 = 0$  is tested. It seems to be

obvious that no gain in small sample power should be expected when applying the bootstrap tests. In case of  $SLT_1^*$ ,  $SLT_2^*$ ,  $JOH_1^*$ , and  $JOH_1^*$  even significant reductions occur. Of course, a size adjustment would slightly improve the relative performance of the bootstrap test versions, since their sizes for the corresponding DGPs are in general a bit lower than the ones of the asymptotic tests (compare Table 1, Panel A). Interestingly, the instances of significant size reductions regarding  $SLT_1^*$ ,  $SLT_2^*$ ,  $JOH_1^*$ , and  $JOH_1^*$  correspond quite well to the cases where significant power losses occur.

The relative performance of the bootstrap tests becomes somewhat worse if we consider the results for the data-dependent KPSW-DGP (5.2), which are summarized in Table 3. In particular, bootstrap Algorithm 2 produces inferior results both for the size and the small sample power. This is especially apparent for  $GLS_2^*$ , which suffers from excessive size distortions and a very low power. A closer inspection of the respective bootstrap iterations points to problems with the estimation of the deterministic terms. The adjusted series have a mean clearly different from zero. Compared to the simpler Toda-DGPs and the asymptotic  $GLS$  test these deviations are much more pronounced. That may explain the poor performance of  $GLS_2^*$ . Nevertheless, the size for  $GLS_1^*$  are preferable compared to  $GLS$ . This also applies to the small sample power if the case  $r_0 = 1$  (Panel B) is considered.

The respective sizes of the  $SLT$  and  $SLT_1^*$  and the  $JOH$  and  $JOH_1^*$  tests are approximately the same. In general, we observe a lower power level for the bootstrap tests. In case of  $SLT_2^*$  and  $JOH_2^*$  the lower power might be attributable to their lower sizes. However, taken into account the results for  $SLT_1^*$  and  $JOH_1^*$  the size differences seem only to explain some portion of the drop in the powers of the tests based on Algorithm 2.

Comparing the different test proposals we find that no test outperforms the other ones. The only notable exception might be that  $SLT$  and  $JOH$  are generally preferable to  $GLS$  in terms of the small sample power. This also applies to the corresponding bootstrap test versions.

Finally, we study the correlation between the asymptotic and bootstrap tests' decisions. This should help to answer the question whether applied researchers have to expect mixed signals when using both the bootstrap and the asymptotic test versions. To assess the relationship between the rejections of two tests we first compute the proportion of joint rejections ( $\hat{\gamma}$ ) with respect to the lower rejection frequency,  $\hat{p}_l = \min(\hat{p}_1, \hat{p}_2)$ . Hence, we obtain the ratio  $\hat{r} = \hat{\gamma}/\hat{p}_l$ , which measures the degree to which the higher rejection frequency of one test is due to additional rejections. Thus, a low value of  $\hat{r}$  means that the test with the lower number of rejections rejects the null hypothesis in many simulation iterations in which the other test does not.

**Table 3.** Rejection Frequencies of Tests for KPSW-DGP (5.2) with True Cointegrating Rank  $r = 2$ , VAR Order  $p = 2$ , Nominal Significance Level 0.05

	Panel A: $H_0 : r_0 = 2$			Panel B: $H_0 : r_0 = 1$			Panel C: $H_0 : r_0 = 0$		
	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$
<i>GLS</i>	0.001	0.006	0.017	0.022	0.040	0.074	0.172	0.487	0.939
<i>GLS</i> <sub>1</sub> <sup>*</sup>	0.023 <sup>+</sup>	0.027 <sup>+</sup>	0.043 <sup>+</sup>	0.033 <sup>+</sup>	0.062 <sup>+</sup>	0.118 <sup>+</sup>	0.100 <sup>-</sup>	0.423 <sup>-</sup>	0.930 <sup>-</sup>
<i>GLS</i> <sub>2</sub> <sup>*</sup>	0.067	0.113	0.083	0.000	0.000 <sup>-</sup>	0.010 <sup>-</sup>	0.000 <sup>-</sup>	0.008 <sup>-</sup>	0.362 <sup>-</sup>
<i>SLT</i>	0.011	0.025	0.051	0.058	0.122	0.357	0.276	0.630	0.990
<i>SLT</i> <sub>1</sub> <sup>*</sup>	0.009	0.027	0.046	0.026 <sup>-</sup>	0.096 <sup>-</sup>	0.338 <sup>-</sup>	0.126 <sup>-</sup>	0.498 <sup>-</sup>	0.982 <sup>-</sup>
<i>SLT</i> <sub>2</sub> <sup>*</sup>	0.001 <sup>-</sup>	0.009 <sup>-</sup>	0.027 <sup>-</sup>	0.013 <sup>-</sup>	0.059 <sup>-</sup>	0.314 <sup>-</sup>	0.109 <sup>-</sup>	0.497 <sup>-</sup>	0.981 <sup>-</sup>
<i>JOH</i>	0.004	0.010	0.029	0.039	0.097	0.366	0.285	0.596	0.994
<i>JOH</i> <sub>1</sub> <sup>*</sup>	0.003	0.011	0.028	0.014 <sup>-</sup>	0.077 <sup>-</sup>	0.332 <sup>-</sup>	0.109 <sup>-</sup>	0.485 <sup>-</sup>	0.991
<i>JOH</i> <sub>2</sub> <sup>*</sup>	0.000 <sup>-</sup>	0.006 <sup>-</sup>	0.013 <sup>-</sup>	0.007 <sup>-</sup>	0.044 <sup>-</sup>	0.299 <sup>-</sup>	0.099 <sup>-</sup>	0.461 <sup>-</sup>	0.992

*Note:* Bold entries indicate significant size differences with respect to nominal level of 5% (Panel A). Significant improvements of the bootstrap tests' rejection rates with respect to the corresponding asymptotic test rejections are indicated by <sup>+</sup> and significant deteriorations by <sup>-</sup>. Significance refers to a 5% level. The tests are based on an approximation ignoring uncertainty involved in the bootstrap simulations and the computation of the critical values.

**Table 4.** Relationship Between Test Rejections for Bivariate Toda-DGP with True Cointegrating Rank  $r = 1$  ( $a_1 = 0.9$ ), VAR Order  $p = 1$ ,  $\theta = 0.8$ , Nominal Significance Level 0.05

$\hat{r}/\hat{\rho}$	$H_0 : r_0 = 1$			$H_0 : r_0 = 0$		
	$T = 50$	$T = 100$	$T = 200$	$T = 50$	$T = 100$	$T = 200$
$GLS/GLS_1^*$	1.00/0.82	1.00/0.75	1.00/0.78	0.95/0.93	0.99/0.96	1.00/0.96
$GLS/GLS_2^*$	1.00/0.67	1.00/0.61	1.00/0.56	0.95/0.94	0.98/0.96	1.00/0.96
$GLS_1^*/GLS_2^*$	1.00/0.81	0.95/0.77	0.93/0.67	1.00/0.97	0.99/0.96	0.95/0.93
$SLT/SLT_1^*$	0.96/0.91	1.00/0.92	0.97/0.91	1.00/0.94	0.98/0.94	1.00/0.95
$SLT/SLT_2^*$	0.96/0.88	1.00/0.92	0.97/0.90	1.00/0.93	0.99/0.94	1.00/0.90
$SLT_1^*/SLT_2^*$	0.89/0.85	0.98/0.98	0.97/0.96	0.96/0.93	0.98/0.96	1.00/0.92
$JOH/JOH_1^*$	0.95/0.93	0.97/0.94	1.00/0.96	0.96/0.93	0.98/0.96	1.00/0.93
$JOH/JOH_2^*$	0.94/0.87	0.90/0.84	1.00/0.94	0.98/0.92	0.99/0.94	1.00/0.95
$JOH_1^*/JOH_2^*$	0.89/0.84	0.87/0.83	0.97/0.95	1.00/0.94	0.99/0.95	0.98/0.96

*Note:* The first entry measures the relative portion of joint rejections with respect to the test with the lower rejection rate ( $\hat{r}$ ). The second entry describes the correlation between the rejections ( $\hat{\rho}$ ).

The second measure we consider is the correlation of the rejections of two tests given by

$$\hat{\rho} = \frac{\hat{\tau}}{\sqrt{\hat{p}_1(1 - \hat{p}_1)} \cdot \sqrt{\hat{p}_2(1 - \hat{p}_2)}}.$$

The ratio  $\hat{\rho}$  can also be regarded as an estimator of the limit correlation of the rejection frequencies  $\hat{p}_1$  and  $\hat{p}_2$  (compare (5.3)). The correlation measure  $\hat{\rho}$  automatically takes account of differing rejection frequencies. To be precise,  $\hat{\rho} < 1$  if  $\hat{p}_1 \neq \hat{p}_2$ . Note that  $\hat{\rho}$  might get very small if  $\hat{p}_1$  and  $\hat{p}_2$  are both close to 1 and  $\hat{\tau}$  is just slightly smaller than  $\min(\hat{p}_1, \hat{p}_2)$ .

In Tables 4 and 5 we report figures for  $\hat{r}$  and  $\hat{\rho}$  for some representative cases. We obtain the following main results. The correlations of the rejections of  $GLS_1^*$ ,  $SLT_1^*$ ,  $SLT_2^*$ ,  $JOH_1^*$ , and  $JOH_2^*$  with the rejections of the corresponding asymptotic tests are generally quite high, i.e. above 0.8. In case of the KPSW-DGP (5.2) values below 0.60 can occur for the correlation measure  $\hat{\rho}$  owing to the differing rejection frequencies of asymptotic and bootstrap tests. The rejections of  $GLS_2^*$  are strongly correlated with the rejections of  $GLS$  for the Toda-DGPs. However, the correlations dramatically fall when using the KPSW-DGP.

The correlations between the rejections of the bootstrap tests based on Algorithms 1 and 2 are usually quite large as long as the rejection frequencies are close to each other. If they deviate, it is  $\hat{\rho}$  that falls more than  $\hat{r}$ . Hence, higher rejections of one test are due to additional

**Table 5.** Relationship Between Test Rejections for KPSW-DGP with True Cointegrating Rank  $r = 2$ , Rank under  $H_0$  is  $r_0 = 1$ , VAR Order  $p = 2$ , Nominal Significance Level 0.05

$\hat{r}/\hat{\rho}$	$T = 50$	$T = 100$	$T = 200$
$GLS/GLS_1^*$	0.96/0.77	1.00/0.79	0.99/0.76
$GLS/GLS_2^*$	0.00/ —	0.00/ —	0.60/0.20
$GLS_1^*/GLS_2^*$	0.00/ —	0.00/ —	0.70/0.18
$SLT/SLT_1^*$	1.00/0.66	0.99/0.86	1.00/0.96
$SLT/SLT_2^*$	1.00/0.46	1.00/0.67	1.00/0.91
$SLT_1^*/SLT_2^*$	0.92/0.65	1.00/0.77	1.00/0.94
$JOH/JOH_1^*$	1.00/0.59	0.99/0.87	0.99/0.92
$JOH/JOH_2^*$	1.00/0.42	1.00/0.66	1.00/0.86
$JOH_1^*/JOH_2^*$	1.00/0.71	1.00/0.74	0.99/0.91

*Note:* The first entry measures the relative portion of joint rejections with respect to the test with the lower rejection rate ( $\hat{r}$ ). The second entry describes the correlation between the rejections ( $\hat{\rho}$ ).

rejections. However, the correlations of  $GLS_1^*$  and  $GLS_2^*$  are very low for the KPSW-DGP in any case. To sum up, corresponding asymptotic and bootstrap tests do not differ systematically in their test decisions. However, the applied researcher has to be aware that there are situations where disagreements may occur.

Finally, we relate our findings to the results of the simulation studies performed by Swensen (2005a, 2005b) in terms of the Johansen test. Swensen (2005a) considers a five-dimensional VAR(1) process with one cointegrating relation as DGP, which is similar to a five-dimensional version of the Toda process (5.1). This process has also been employed by Johansen (2002). In addition, Swensen (2005b) also employs the Toda process and another DGP that has been used in the literature before. There are a couple of issues worth mentioning.

First, the results of Swensen (2005a, 2005b) on the tests's sizes are slightly more encouraging for the bootstrap tests and, second, he found the differences between the Algorithms 1 and 2 to be rather small. These findings may be DGP-specific. Nevertheless, our results are in line with these findings if we focus our attention on the simple Toda-DGP and the *JOH*- and *SLT*-tests. In fact, most of the deteriorations with respect to the bootstrap tests and most of the pronounced differences between the algorithms are confined to the more complex KPSW-DGP (5.2). Probably, it is rather difficult to replicate the autoregressive structure of this DGP by appropriate parameter estimates. As a consequence of this, we may obtain autocorrelated

residuals. It is known from other studies that this negatively affects the bootstrap tests (compare e.g. Härdle et al. 2003, MacKinnon 2002).

However, our results regarding the loss in small sample power and the rather high correlations between the asymptotic and bootstrap test decisions are confirmed by the findings of Swensen (2005a, 2005b). Performing a sequential test application, Swensen (2005a, 2005b) figured out that the asymptotic tests may find more often the true cointegrating rank than the bootstrap tests, although the latter have superior size properties. In other words, the bootstrap test's sequence terminates too early due to a lower small sample power.

## 6 Summary

In this paper we have analysed two recursive bootstrap procedures for systems cointegration tests with a prior adjustment for deterministic terms suggested by Saikkonen & Lütkepohl (2000b) and Saikkonen et al. (2006). Their small sample properties are studied and compared to the properties of the corresponding asymptotic tests within a simulation study. This simulation study also includes the Johansen cointegration test.

Our overall conclusions are that bootstrap methods can help to improve the size properties of systems cointegration tests in small samples if simple processes are considered. However, in case of more complex DGPs the situation is less favourable. With respect to the tests' small sample power no increase should be expected. In fact, we often observe significant reductions in the small sample power if the bootstrap test versions are applied. Our mixed findings correspond quite well to the outcomes of different simulation studies on the Johansen cointegration test. While Swensen (2005a, 2005b) also detects size improvements for the bootstrap test versions in case of DGPs with a simple autoregressive structure, van Giersbergen (1996) and Harris & Judge (1998) report deteriorations in terms of small sample size and power for more complex DGPs. Moreover, some of our findings seem to support a remark of Johansen (2002, p. 1930). He does not necessarily expect an improved approximation of the tests' distributions by the use of bootstrap methods. The reason is the nonuniformity in the convergence with respect to the nuisance parameters in the null distributions of the test statistics. By contrast, Johansen (2002) proposes a small sample correction following Bartlett (1937). However, Swensen (2005a) has shown that bootstrap approximations work better than the correction suggested by Johansen (2002) in a number of situations.

When using bootstrap tests, we have a clear preference for bootstrap Algorithm 1, in which all parameters are estimated imposing the cointegrating rank under the null hypothesis. In case

of the more complex data-dependent DGP, the tests based on Algorithm 1 clearly outperform the corresponding versions related to Algorithm 2 with respect to the small sample properties. The latter algorithm uses residuals and estimators for the short-run parameters from an unrestricted VAR. Furthermore, we have found that the rejections of corresponding bootstrap and asymptotic tests are usually strongly correlated.

It should be rather straightforward to adjust our asymptotic results such that different assumptions about the presence of deterministic terms can be made. Hence, one may apply valid bootstrap versions of cointegration tests considered by Saikkonen & Lütkepohl (2000a), which allow for level shifts at known time, or of a test procedure excluding a linear trend as e.g. suggested by Saikkonen et al. (2006). Finally, an extension of the bootstrap algorithms in order to consistently determine the cointegration rank may be conducted in future research. The asymptotic results of Swensen (2005a) regarding the Johansen test indicate that such a consistent sequential procedure can also be found for tests with a prior adjustment for deterministic terms. However, a sequential bootstrap procedure would probably not work better than the asymptotic version due to the inferior small sample power of the bootstrap tests.

van Giersbergen (1996) pointed out the difficulties of both bootstrap and asymptotic tests to deal with autocorrelated error terms in case of an inappropriate VAR order choice. Given that, a more promising line of research may be to design and evaluate an overall bootstrap procedure for choosing the lag order and the cointegrating rank. The hope is to improve the model specification resulting in a more reliable inference on the cointegrating rank.

## Appendix A: Proofs

### Proof of Lemma 1

According to Lemma 1 and Remark 9 of Swensen (2005a) there exists a Granger representation for  $y_t^*$ , which is given by

$$y_t^* = \tilde{C} \sum_{i=p+1}^t \varepsilon_i^* + \tilde{\tau}(t-p) + \sqrt{T}R_{t,T}^*, \quad t = p+1, \dots, T, \quad (\text{A.1})$$

where  $\tilde{C} = \tilde{\beta}_\perp (\tilde{\alpha}'_\perp \tilde{\Gamma} \tilde{\beta}_\perp)^{-1} \tilde{\alpha}'_\perp$ ,  $\tilde{\tau} = \tilde{C}(\tilde{\nu} + \tilde{\alpha}\tilde{\phi}) - (\tilde{C}\tilde{\Gamma} - I)\tilde{\beta}\tilde{\phi}$ , and  $\sqrt{T}R_{t,T}^*$  is defined in (A.4).

Note that  $x_t^*$  and  $y_t^*$  are generated using the same bootstrap residuals and parameter estimators. Therefore it is sufficient to show that the deterministic part  $\mu_0^* + \mu_1^*t$  can be extracted from

(A.1) in order to prove (4.4) of Lemma 1. Hence, we have to show

$$x_t^* + \mu_0^* + \mu_1^* t = \tilde{C} \sum_{i=p+1}^t \varepsilon_i^* + \sqrt{T} R_{t,T}^{0*} + \mu_0^* + \mu_1^* t, \quad t = p+1, \dots, T, \quad (\text{A.2})$$

where  $y_t^* = x_t^* + \mu_0^* + \mu_1^* t$  and  $\sqrt{T} R_{t,T}^{c*}$  as defined in (A.9). Then, we can subtract  $\mu_0^* + \mu_1^* t$  from both sides of (A.2) and obtain the Granger representation for  $x_t^*$  as proposed in Lemma 1.

In the following, we first repeat some steps of the proof in Swensen (2005a) in order to introduce a number of expressions in line with the notation used in this paper. Let  $\tilde{Y}_t^* = (y_t^* \tilde{\beta}', \Delta y_t^* \tilde{\beta}'_{\perp})'$ . Then, according to Johansen (1995) and Swensen (2005a), the recursion defined in Algorithm 1 can be stated as

$$\tilde{A}(L) \tilde{Y}_t^* = (\tilde{\alpha}, \tilde{\alpha}_{\perp})' (\varepsilon_t^* + \tilde{\nu} + \tilde{\alpha} \tilde{\phi} - \tilde{\alpha} \tilde{\phi} t)$$

for some suitable lag polynomial  $\tilde{A}(L)$ . Further, following the arguments in Swensen (2005a), we can define  $\tilde{C}(z) = \tilde{A}(z)^{-1}$  and obtain

$$\begin{aligned} \tilde{Y}_t^* &= \begin{pmatrix} \tilde{\beta}' y_t^* \\ \tilde{\beta}'_{\perp} \Delta y_t^* \end{pmatrix} = \tilde{A}(L)^{-1} (\tilde{\alpha}, \tilde{\alpha}_{\perp})' (\varepsilon_t^* + \tilde{\nu} + \tilde{\alpha} \tilde{\phi} - \tilde{\alpha} \tilde{\phi} t) \\ &= \tilde{C}(L) (\tilde{\alpha}, \tilde{\alpha}_{\perp})' (\varepsilon_t^* + \tilde{\nu} + \tilde{\alpha} \tilde{\phi} - \tilde{\alpha} \tilde{\phi} t) \\ &= \tilde{C}(1) (\tilde{\alpha}, \tilde{\alpha}_{\perp})' (\varepsilon_t^* + \tilde{\nu} + \tilde{\alpha} \tilde{\phi} - \tilde{\alpha} \tilde{\phi} t) + \tilde{C}^{\#}(L) (\tilde{\alpha}, \tilde{\alpha}_{\perp})' \Delta (\varepsilon_t^* + \tilde{\nu} + \tilde{\alpha} \tilde{\phi} - \tilde{\alpha} \tilde{\phi} t). \end{aligned} \quad (\text{A.3})$$

The derivation of the Granger representation for  $y_t^*$  is based on the decomposition

$$y_t^* = \tilde{\beta}_{\perp} \tilde{\beta}'_{\perp} \sum_{i=p+1}^t \Delta y_i^* + \tilde{\beta} \tilde{\beta}' y_t^* + \tilde{\beta}_{\perp} \tilde{\beta}'_{\perp} y_p^*.$$

Using this equation in connection with (A.3), Swensen (2005a) derives the expressions for  $\tilde{\tau}$  and the remainder term. In our notation we have for the latter one

$$\begin{aligned} \sqrt{T} R_{t,T}^* &= (0, \tilde{\beta}_{\perp}) \tilde{C}^{\#}(L) (\tilde{\alpha}, \tilde{\alpha}_{\perp})' (\varepsilon_t^* - \varepsilon_p) + (\tilde{\beta}, 0) \tilde{C}(1) (\tilde{\alpha}, \tilde{\alpha}_{\perp})' \varepsilon_t^* \\ &\quad + (\tilde{\beta}, 0) \tilde{C}^{\#}(L) (\tilde{\alpha}, \tilde{\alpha}_{\perp})' \Delta \varepsilon_t^* \\ &\quad + (\tilde{\beta}, 0) \tilde{C}(1) (\tilde{\alpha}, \tilde{\alpha}_{\perp})' (\tilde{\nu} + \tilde{\alpha} \tilde{\phi}) - (\tilde{\beta}, 0) \tilde{C}^{\#}(1) (\tilde{\alpha}, \tilde{\alpha}_{\perp})' \tilde{\alpha} \tilde{\phi} \\ &\quad + p \tilde{\beta} \tilde{\phi} + \tilde{\beta}_{\perp} \tilde{\beta}'_{\perp} y_p. \end{aligned} \quad (\text{A.4})$$

Thus, in order to derive (4.5) from (A.1) we have to analyse the last four terms of (A.4) and  $\tilde{\tau}(t-p)$ , which all involve deterministic components. Since we assume that  $\mu_0^*$  and  $\mu_1^*$  satisfy the conditions (4.2) the estimators  $\tilde{\nu}$  and  $\tilde{\phi}$  are replaced accordingly.



Let us start with the term  $\tilde{\tau}(t - p)$ . Given the definition of  $\tilde{\tau}$  and the conditions (4.2) we obtain

$$\begin{aligned}
\tilde{\tau} &= \tilde{C}'(\tilde{\alpha}\theta^* + \tilde{\Gamma}\mu_1^* + \tilde{\alpha}\tilde{\beta}'\mu_1^*) - (\tilde{C}\tilde{\Gamma} - I)\tilde{\beta}\tilde{\beta}'\mu_1^* \\
&= \tilde{C}\tilde{\Gamma}\mu_1^* - (\tilde{C}\tilde{\Gamma} - I)\tilde{\beta}\tilde{\beta}'\mu_1^* \\
&= \tilde{C}\tilde{\Gamma}(I - \tilde{\beta}\tilde{\beta}')\mu_1^* + \tilde{\beta}\tilde{\beta}'\mu_1^* \\
&= \tilde{C}\tilde{\Gamma}\tilde{\beta}_\perp\tilde{\beta}'_\perp\mu_1^* + \tilde{\beta}\tilde{\beta}'\mu_1^* \\
&= [\tilde{\beta}_\perp(\tilde{\beta}'_\perp\tilde{\beta}_\perp)^{-1}\tilde{\beta}'_\perp + \tilde{\beta}(\tilde{\beta}'\tilde{\beta})^{-1}\tilde{\beta}']\mu_1^* \\
&= \mu_1^*.
\end{aligned} \tag{A.5}$$

Furthermore,

$$\begin{aligned}
p\tilde{\beta}\tilde{\phi} + \tilde{\beta}_\perp\tilde{\beta}'_\perp y_p &= p\tilde{\beta}\tilde{\beta}'\mu_1^* + \tilde{\beta}_\perp\tilde{\beta}'_\perp(x_p^* + \mu_0^* + p\mu_1^*) \\
&= \tilde{\beta}_\perp\tilde{\beta}'_\perp x_p^* + \tilde{\beta}_\perp\tilde{\beta}'_\perp\mu_0^* + p(\tilde{\beta}\tilde{\beta}' + \tilde{\beta}_\perp\tilde{\beta}'_\perp)\mu_1^* \\
&= \tilde{\beta}_\perp\tilde{\beta}'_\perp x_p^* + \tilde{\beta}_\perp\tilde{\beta}'_\perp\mu_0^* + p\mu_1^*.
\end{aligned} \tag{A.6}$$

Then,

$$\begin{aligned}
(\tilde{\beta}, 0)\tilde{C}(1)(\tilde{\alpha}, \tilde{\alpha}_\perp)'(\tilde{\nu} + \tilde{\alpha}\tilde{\phi}) &= (\tilde{\beta}, 0)\tilde{C}(1)(\tilde{\alpha}, \tilde{\alpha}_\perp)'(-\tilde{\alpha}\theta^* + \tilde{\Gamma}\mu_1^* + \tilde{\alpha}\tilde{\beta}'\mu_1^*) \\
&= \tilde{\beta}\tilde{\alpha}'(\tilde{\Gamma}\tilde{C} - I)(-\tilde{\alpha}\tilde{\beta}'\mu_0^* + \tilde{\Gamma}\mu_1^* + \tilde{\alpha}\tilde{\beta}'\mu_1^*) \\
&= \tilde{\beta}\tilde{\alpha}'(\tilde{\Gamma}\tilde{C} - I)\tilde{\Gamma}\mu_1^* + \tilde{\beta}\tilde{\beta}'\mu_0^* - \tilde{\beta}\tilde{\beta}'\mu_1^* \\
&= \tilde{\beta}\tilde{\alpha}'\tilde{\Gamma}(\tilde{C}\tilde{\Gamma} - I)\mu_1^* + \tilde{\beta}\tilde{\beta}'\mu_0^* - \tilde{\beta}\tilde{\beta}'\mu_1^*,
\end{aligned} \tag{A.7}$$

where the second equality follows from the proof of Lemma 1 in Swensen (2005a). Finally, we have

$$\begin{aligned}
-(\tilde{\beta}, 0)\tilde{C}^\#(1)(\tilde{\alpha}, \tilde{\alpha}_\perp)'\tilde{\alpha}\tilde{\phi} &= -(\tilde{\beta}, 0)\tilde{C}^\#(1)(\tilde{\beta}'\mu_1^*, 0)' \\
&= -\tilde{\beta} \left[ -I - \tilde{\alpha}'\tilde{\Gamma}\tilde{\beta} + \tilde{\alpha}'\tilde{\Gamma}\tilde{\beta}_\perp(\tilde{\alpha}'_\perp\tilde{\Gamma}\tilde{\beta}_\perp)^{-1}\tilde{\alpha}'_\perp\tilde{\Gamma}\tilde{\beta} \right] \tilde{\beta}'\mu_1^* \\
&= \tilde{\beta}\tilde{\beta}'\mu_1^* + \tilde{\beta}\tilde{\alpha}'\tilde{\Gamma}\tilde{\beta}\tilde{\beta}'\mu_1^* - \tilde{\beta}\tilde{\alpha}'\tilde{\Gamma}\tilde{C}\tilde{\Gamma}\tilde{\beta}\tilde{\beta}'\mu_1^* \\
&= \tilde{\beta}\tilde{\beta}'\mu_1^* - \tilde{\beta}\tilde{\alpha}'\tilde{\Gamma}(\tilde{C}\tilde{\Gamma} - I)\tilde{\beta}\tilde{\beta}'\mu_1^*,
\end{aligned} \tag{A.8}$$

where the second equality follows from the results in Rahbek & Mosconi (1999, p. 89). Note in this respect that the expression  $-\dot{D}(1)$  in Rahbek & Mosconi (1999) corresponds to  $\tilde{C}^\#(1)$ .

Collecting the results of (A.5) - (A.8) yields

$$\begin{aligned}
& \mu_1^*(t-p) + [p\mu_1^* + \bar{\beta}_\perp \tilde{\beta}'_\perp \mu_0^*] + [\bar{\beta} \tilde{\alpha} \tilde{\Gamma}' (\tilde{C} \tilde{\Gamma} - I) \mu_1^* + \bar{\beta} \tilde{\beta}' \mu_0^* - \bar{\beta} \tilde{\beta}' \mu_1^*] \\
& + [\bar{\beta} \tilde{\beta}' \mu_1^* - \bar{\beta} \tilde{\alpha}' \tilde{\Gamma}' (\tilde{C} \tilde{\Gamma} - I) \bar{\beta} \tilde{\beta}' \mu_1^*] \\
& = (\bar{\beta}_\perp \tilde{\beta}'_\perp + \bar{\beta} \tilde{\beta}') \mu_0^* + \mu_1^* t + [\bar{\beta} \tilde{\alpha} \tilde{\Gamma}' (\tilde{C} \tilde{\Gamma} - I)] (I - \bar{\beta} \tilde{\beta}') \mu_1^* \\
& = \mu_0^* + \mu_1^* t + \bar{\beta} \tilde{\alpha} \tilde{\Gamma}' (\tilde{C} \tilde{\Gamma} - I) \bar{\beta}_\perp \tilde{\beta}'_\perp \mu_1^* \\
& = \mu_0^* + \mu_1^* t + \bar{\beta} \tilde{\alpha} \tilde{\Gamma}' (\bar{\beta}_\perp \tilde{\beta}'_\perp - \bar{\beta}_\perp \tilde{\beta}'_\perp) \mu_1^* \\
& = \mu_0^* + \mu_1^* t
\end{aligned}$$

for the deterministic components on the r.h.s. of (A.1). Note that the third equality is obtained by plugging in the definition of  $\tilde{C}$ . Hence, as mentioned above, we can subtract  $\mu_0^* + \mu_1^* t$  from both sides of (4.5) and obtain the Granger representation for  $x_t^*$  as given in Lemma 1 with

$$\begin{aligned}
\sqrt{T} R_{t,T}^{0*} &= (0, \bar{\beta}_\perp) \tilde{C}^\#(L) (\bar{\alpha}, \bar{\alpha}_\perp)' (\varepsilon_t^* - \varepsilon_p) + (\bar{\beta}, 0) \tilde{C}(1) (\bar{\alpha}, \bar{\alpha}_\perp)' \varepsilon_t^* \\
&+ (\bar{\beta}, 0) \tilde{C}^\#(L) (\bar{\alpha}, \bar{\alpha}_\perp)' \Delta \varepsilon_t^* + \bar{\beta}_\perp \tilde{\beta}'_\perp x_p^*.
\end{aligned} \tag{A.9}$$

As mentioned in Section 4,  $\tilde{\beta}'_\perp(x_p^* -)$  does not converge to zero. However, it is bounded in  $P_{r_0,T}$  probability and bounded with respect to the measure  $P_{r_0,T}^*$ . To see this, we write  $\tilde{\beta}'_\perp x_p^* = \tilde{\beta}'_\perp [x_p - (\mu_0^* - \mu_0) - (\mu_1^* - \mu_1)p]$  in line with (4.11). It is obvious, that this expression is bounded conditional on the observations since all terms are nonrandom in this case. Regarding the measure  $P_{r_0,T}$  we have  $\tilde{\beta}'_\perp(\mu_1^* - \mu_1) = O_p(T^{-1/2})$  if  $\mu_1^* = \hat{\mu}_1$  and, from Saikkonen & Lütkepohl (2000b, Proof of Theorem 1),  $\tilde{\beta}'_\perp(\mu_0^* - \mu_0) = O_p(1)$  if  $\mu_0^* = \hat{\mu}_0$  from (4.3) is applied. Thus,  $\bar{\beta}_\perp \tilde{\beta}'_\perp x_p^* = \bar{\beta}_\perp \tilde{\beta}'_\perp x_p + O_p(1)$  is bounded in  $P_{r_0,T}$  probability due to the consistency of  $\tilde{\beta}_\perp$ . Thus, the result  $P_{r_0,T}^*(\max_{p+1 \leq t \leq T} |R_{t,T}^{0*}| > \eta) \rightarrow 0$  for all  $\eta > 0$  in  $P_{r_0,T}$  probability as  $T \rightarrow \infty$  follows from the proof of Lemma 1 in Swensen (2005a) because for all  $\eta > 0$ ,  $P_{r_0,T}^*(|\bar{\beta}_\perp \tilde{\beta}'_\perp x_p^*|/\sqrt{T} > \eta) \rightarrow 0$  in  $P_{r_0,T}$  probability as  $T \rightarrow \infty$ . In fact, for proving the corresponding result for  $R_{t,T}^*$ , Swensen (2005a) only considers our remainder term  $\sqrt{T} R_{t,T}^{c*}$  with  $y_p^*$  instead of  $x_p^*$  since the estimators of the deterministic parameters of the VECM (2.3) are bounded in probability due to their consistency.

$E^*[\varepsilon_t^* \varepsilon_t^{*'}] = \tilde{\Omega}_T \rightarrow \Omega$  in  $P_{r_0,T}$  probability as  $T \rightarrow \infty$  follows directly from Lemma 1 and Remark 9 of Swensen (2005a) since we use the same bootstrap residuals for generating  $x_t^*$  as for  $y_t^*$ .

The representation (4.5) given for  $y_t^{c*}$  in Lemma 1 follows from the foregoing derivations. First, add  $\mu_0^*$  on both sides of (4.4). Second, decompose  $\mu_0^*$  into  $\mu_0^* = (\bar{\beta} \tilde{\beta}' + \bar{\beta}_\perp \tilde{\beta}'_\perp) \mu_0^* = \tau_0^* + \bar{\beta}_\perp \tilde{\beta}'_\perp \mu_0^*$ . Then, we have  $\bar{\beta}_\perp \tilde{\beta}'_\perp x_p^* + \bar{\beta}_\perp \tilde{\beta}'_\perp \mu_0^* = \bar{\beta}_\perp \tilde{\beta}'_\perp y_p^{c*}$ . Hence, (4.5), i.e.  $y_t^{c*} = x_t^* + \mu_0^* =$

$\tilde{C} \sum_{i=p+1}^t \varepsilon_t^* + \tau_0^* + \sqrt{T} R_{t,T}^{c*}$ , is obtained with

$$\begin{aligned} \sqrt{T} R_{t,T}^{c*} &= (0, \bar{\beta}_\perp) \tilde{C}^\#(L) (\bar{\alpha}, \bar{\alpha}_\perp)' (\varepsilon_t^* - \varepsilon_p) + (\bar{\beta}, 0) \tilde{C}(1) (\bar{\alpha}, \bar{\alpha}_\perp)' \varepsilon_t^* \\ &\quad + (\bar{\beta}, 0) \tilde{C}^\#(L) (\bar{\alpha}, \bar{\alpha}_\perp)' \Delta \varepsilon_t^* + \bar{\beta}_\perp \bar{\beta}'_\perp y_p^{c*}. \end{aligned} \quad (\text{A.10})$$

Note that  $\bar{\beta}_\perp \bar{\beta}'_\perp y_p^{c*} = \bar{\beta}_\perp \bar{\beta}'_\perp (y_p^c - (\mu_1^* - \mu_1)p) = \bar{\beta}_\perp \bar{\beta}'_\perp y_p^c + O_p(T^{-1/2})$  such that  $\bar{\beta}_\perp \bar{\beta}'_\perp y_p^{c*}$  is bounded in  $P_{r_0,T}$  probability. The result  $P_{r_0,T}^*(\max_{p+1 \leq t \leq T} |R_{t,T}^{c*}| > \eta) \rightarrow 0$  for all  $\eta > 0$  in  $P_{r_0,T}$  probability as  $T \rightarrow \infty$  follows from the argumentation with respect to the remainder term  $R_{t,T}^{0*}$  above. This completes the proof of Lemma 1.

## Proof of Proposition 1

We now present the results needed to prove Proposition 1. They are collected in different Lemmas. The idea is to mimic the proof of Theorem 4.1 for  $SLT(r_0)$  given in Saikkonen et al. (2006).<sup>5</sup> To this end, we first show a number of results conditional on the observations, i.e. with respect to the measure  $P_{r_0,T}^*$ . Then, we obtain convergence and distribution results regarding the measure  $P_{r_0,T}$ . In order to accomplish that we have to make use of Lemma 1 and of Lemmas 1-3 in Swensen (2005a) several times. Note again that Lemmas 1-3 are actually proven in conjunction with Algorithm 2 but are also valid in case of Algorithm 1 (see Remark 9 of Swensen 2005a).

In the following we will use the bootstrap stochastic order symbols which have been introduced by Park (2003) and Chang & Park (2003). Similar to Chang & Park (2003), they are defined here as follows. Let  $(X_n^*)$  be a sequence of bootstrap statistics. Then,  $X_n^* = O_p^*(T^\lambda)$  if for every  $\epsilon > 0$  there exists a finite  $K$  and  $N(K, \epsilon) \in \mathbb{N}$ , such that  $P_{r_0,T}^*\{|T^{-\lambda} X_n^*| > K\} < \epsilon$  for all  $T \geq N(K, \epsilon)$ . Furthermore,  $X_n^* = o_p^*(T^\lambda)$  if for every  $\epsilon \in P_{r_0,T}^*\{|T^{-\lambda} X_n^*| > \epsilon\}$  converges to zero in  $P_{r_0,T}$  probability as  $T \rightarrow \infty$ . These definitions can be adopted for sequences of random vectors. Furthermore,  $\xrightarrow{d^*}$  signifies convergence in  $P_{r_0,T}^*$  distributions (compare Swensen 2005a) and,  $\xrightarrow{p^*}$  and  $\xrightarrow{p}$  refer to convergence in  $P_{r_0,T}^*$  and  $P_{r_0,T}$  probability, respectively.

The framework, on which the derivation of the asymptotic distribution is based, is described in Section 4. As mentioned there, the bootstrap estimator of  $\mu_1^*$  has to satisfy certain convergence and distributional properties. In deriving them we need convergence results for the RR estimators of the parameters of the VECM (3.1), which is based on the pseudo data  $y_t^*$ . They will be stated conditional on the observations  $y_1, \dots, y_T$  in Lemma A1. The asterisk indicates

<sup>5</sup>Note again that we do not consider a level shift at unknown time in our framework. Therefore, the model setup and the corresponding proofs in Saikkonen et al. (2006) simplify accordingly. Nevertheless, the asymptotic test distribution is not affected by excluding the level shift.

the estimators that are based on the bootstrap data. For the derivations, we have to assume that  $\tilde{\beta}^*$  has been made unique by the normalization  $\tilde{\beta}^*[(\tilde{\beta}'\tilde{\beta})^{-1}\tilde{\beta}'\tilde{\beta}^*]^{-1} = \tilde{\beta}^*(\tilde{\beta}'\tilde{\beta}^*)^{-1}$ . A corresponding normalization follows for  $\tilde{\alpha}^*$ . Note that all other relevant quantities are invariant to these normalizations (compare Saikkonen & Lütkepohl 2000b). Therefore, we do not explicitly indicate the normalized version of the estimators in the following.

**Lemma A1.** Under Assumption 1, we have

$$\tilde{\alpha}^* - \tilde{\alpha} = O_p^*(T^{-1/2}) \quad (\text{A.11})$$

$$\tilde{\beta}^* - \tilde{\beta} = O_p^*(T^{-1}) \quad (\text{A.12})$$

$$\tilde{\phi}^* - \tilde{\phi} = O_p^*(T^{-3/2}) \quad (\text{A.13})$$

$$\tilde{\nu}^* - \tilde{\nu} = O_p^*(T^{-1/2}) \quad (\text{A.14})$$

$$\tilde{\Gamma}_i^* - \tilde{\Gamma}_i = O_p^*(T^{-1/2}) \quad (i = 1, \dots, p-1) \quad (\text{A.15})$$

$$\tilde{\Omega}^* - \tilde{\Omega} = O_p^*(T^{-1/2}). \quad (\text{A.16})$$

□

**Proof** To proof Lemma A1 we follow Johansen (1995, Ch. 13). Note, however, that he considers a linear trend orthogonal to the space spanned by the cointegration vectors. Thus, we have to adjust the treatment of the deterministic terms. This basically results in the use of different norming matrices. Furthermore, keep in mind that estimators based on the observations are not random if we condition on the observations.

Define  $Z_{0t}^* = \Delta y_t^*$ ,  $Z_{1t}^* = (y_{t-1}^{*'}, t-1)'$ , and  $Z_{2t}^* = (\Delta y_{t-1}^{*'}, \dots, \Delta y_{t-p+1}^{*'}, 1)'$ . Then, regress  $Z_{0t}^*$  and  $Z_{1t}^*$  on  $Z_{2t}^*$  to obtain the residuals  $R_{0t}^*$  and  $R_{1t}^*$ , respectively. Using these residuals we define

$$S_{ij}^* = (T-p)^{-1} \sum_{t=p+1}^T R_{it}^* R_{jt}^{*'} \quad i, j = 0, 1.$$

In the same way we can obtain the moment matrices  $S_{ij}$  ( $i, j = 0, 1$ ) from  $Z_{0t}$ ,  $Z_{1t}$ ,  $Z_{2t}$ ,  $R_{0t}$ , and  $R_{1t}$ , which are defined in terms of the original observations.

We first show the consistency of the estimators  $\tilde{\alpha}^*$ ,  $\tilde{\beta}^*$ ,  $\tilde{\phi}^*$ , and  $\tilde{\Omega}^*$ . Define  $\tilde{\beta}^{*\#} = (\tilde{\beta}^{*'}, -\tilde{\phi}^{*'})'$ . The estimator  $\tilde{\beta}^{*\#}$  is based on the eigenvectors corresponding to the  $r$  largest eigenvalues of

$$|\lambda S_{11}^* - S_{10}^* S_{00}^{*-1} S_{01}^*| = 0, \quad (\text{A.17})$$

(see Swensen 2005a). Given  $\tilde{\beta}^{*\#}$ , we can obtain the estimators  $\tilde{\alpha}^* = S_{01}^* \tilde{\beta}^{*\#} (\tilde{\beta}^{*\#'} S_{11}^* \tilde{\beta}^{*\#})^{-1}$  and  $\tilde{\Omega}^* = S_{00}^{*-1} - S_{01}^* \tilde{\beta}^{*\#} (\tilde{\beta}^{*\#'} S_{11}^* \tilde{\beta}^{*\#})^{-1} \tilde{\beta}^{*\#'} S_{10}^*$  corresponding to Johansen (1995). Similar to

Johansen (1995, Ch. 13) we transform equation (A.17) to

$$|\lambda \tilde{A}'_T S_{11}^* \tilde{A}_T - \tilde{A}'_T S_{10}^* S_{00}^{*-1} S_{01}^* \tilde{A}_T| = 0, \quad (\text{A.18})$$

where

$$\tilde{A}_T = \begin{bmatrix} \tilde{\beta} & T^{-1/2} \tilde{\beta}_\perp & 0 \\ -\phi & T^{-1/2} \tilde{\tau}' \tilde{\beta}_\perp & T^{-1} \end{bmatrix} = \begin{bmatrix} \tilde{\beta}^{\# \#} & T^{-1/2} \tilde{C}'_T \end{bmatrix}$$

with

$$\tilde{C}_T = \begin{bmatrix} \tilde{\beta}_\perp & 0 \\ \tilde{\tau}' \tilde{\beta}_\perp & T^{-1/2} \end{bmatrix}$$

as in Swensen (2005a). The eigenvalues of (A.17) and (A.18) are identical and the eigenvectors of (A.18) can be obtained from those of (A.17) by premultiplying with

$$\tilde{A}_T^{-1} = \begin{bmatrix} \tilde{\beta}' & 0 \\ T^{1/2} \tilde{\beta}'_\perp & 0 \\ T \tilde{\tau}' & T \end{bmatrix}.$$

Next, we study the weak limit of (A.18) conditional on the observations. Given the definition of  $\tilde{A}_T$  we can write (A.18) as

$$\left| \lambda \begin{bmatrix} \tilde{\beta}^{\# \#} S_{11}^* \tilde{\beta}^{\# \#} & T^{-1/2} \tilde{\beta}^{\# \#} S_{11}^* \tilde{C}'_T \\ T^{-1/2} \tilde{C}'_T S_{11}^* \tilde{\beta}^{\# \#} & \tilde{C}'_T S_{11}^* \tilde{C}_T \end{bmatrix} - \begin{bmatrix} \tilde{\beta}^{\# \#} S_{10}^* S_{00}^{*-1} S_{01}^* \tilde{\beta}^{\# \#} & T^{-1/2} \tilde{\beta}^{\# \#} S_{10}^* S_{00}^{*-1} S_{01}^* \tilde{C}'_T \\ \tilde{C}'_T S_{10}^* S_{00}^{*-1} S_{01}^* \tilde{\beta}^{\# \#} & T^{-1} S_{10}^* S_{00}^{*-1} S_{01}^* \tilde{C}_T \end{bmatrix} \right| = 0.$$

Hence, we need a number of conditional convergence results to determine the weak limit of (A.18). They are summarized in the next lemma, which is the counterpart of Lemma 10.3 in Johansen (1995).

**Lemma A2.** Under Assumption 1, we have

$$S_{00}^* \xrightarrow{p^*} S_{00} \quad (\text{A.19})$$

$$S_{01}^* \tilde{\beta}^{\# \#} \xrightarrow{p^*} S_{01} \tilde{\beta}^{\# \#} \quad (\text{A.20})$$

$$\tilde{\beta}^{\# \#} S_{11}^* \tilde{\beta}^{\# \#} \xrightarrow{p^*} \tilde{\beta}^{\# \#} S_{11} \tilde{\beta}^{\# \#} \quad (\text{A.21})$$

$$(T - p)^{-1} \tilde{C}'_T S_{11}^* \tilde{C}_T \xrightarrow{d^*} \int_0^1 G^*(s) G^*(s)' ds \quad (\text{A.22})$$

$$\tilde{C}'_T (S_{10}^* - S_{11}^* \tilde{\beta}^{\# \#} \tilde{\alpha}) \xrightarrow{d^*} \int_0^1 G^*(s) (dW^*(s))' \quad (\text{A.23})$$

$$\tilde{C}'_T S_{11}^* \tilde{\beta}^{\# \#} = O_p^*(1), \quad (\text{A.24})$$

where  $G(s)^* = [(W^*(s) - \bar{W}^*(s))' \tilde{C}' \tilde{\beta}_\perp, s - 1/2]'$ ,  $\bar{W}^*(s) = \int_0^1 W^*(s) ds$ , and  $W^*(s)$  is a  $n$ -dimensional Brownian motion with covariance matrix  $\tilde{\Omega}$ .  $\square$

**Proof** (A.19) - (A.21) follow from the proof of Lemma 2 in Swensen (2005a). Consider for example (A.19). Note that  $S_{00}^* = M_{00}^* - M_{02}^* M_{22}^{*-1} M_{20}^*$  where  $M_{ij}^* = (T-p)^{-1} \sum_{t=p+1}^T Z_{it}^* Z_{jt}^{*'} (i, j = 0, 2)$ . Similarly, we can write  $S_{00}$  by the help of  $M_{ij}$  ( $i, j = 0, 2$ ) defined with respect to the original observations. Swensen (2005a) shows that the first differences of the observations and the bootstrap data have invertible moving average representations given by  $\Delta y_{t-j} = \tilde{\nu}_j + \sum_{i=0}^{\infty} \tilde{\xi}_{i,j}' \tilde{\varepsilon}_{t-j}$  and  $\Delta y_{t-j}^* = \tilde{\nu}_j + \sum_{i=0}^{\infty} \tilde{\xi}_{i,j}' \varepsilon_{t-j}^*$  respectively, where  $\tilde{\nu}_j$  and  $\tilde{\xi}_{i,j}$  are a  $(n \times 1)$  vector and a  $(n \times n)$  matrix of parameter estimates described in Swensen (2005a). Note that  $\Delta y_{t-j}$  is expressed using the true error terms  $\varepsilon_t$  in Swensen (2005a). However, we can equally use the residuals  $\tilde{\varepsilon}_t$  by replacing the true parameter vector  $\nu_j$  by the estimated vector  $\tilde{\nu}_j$ . This is possible because we condition on the observations.

Given the moving average representations it is straightforward to see that  $E^*[(\Delta y_{t-j} - \tilde{\nu}_j)(\Delta y_{t-k} - \tilde{\nu}_k)'] = \sum_{i=0}^{\infty} \tilde{\xi}_{i,j}' \tilde{\Omega} \tilde{\xi}_{i,k} = E^*[(\Delta y_{t-j}^* - \tilde{\nu}_j)(\Delta y_{t-k}^* - \tilde{\nu}_k)']$  ( $j, k = 0, \dots, p-1$ ), because  $E^*[\varepsilon_t^* \varepsilon_t^{*'}] = (T-p)^{-1} \sum_{i=p+1}^T \tilde{\varepsilon}_t \tilde{\varepsilon}_t' = \tilde{\Omega}$ . Based on this result we obtain  $E^*(M_{ij}^*) = E^*(M_{ij})$ . Furthermore, we conclude from the argumentation of Swensen (2005a) related to equation (22) in the proof of Lemma 2 that the conditional variance of  $M_{ij}^*$  ( $i, j = 0, 2$ ) converges to zero in  $P_{r_0, T}^*$  probability. Thus, (A.19) will follow. (A.20) and (A.21) can be obtained in a similar way by noting that  $\tilde{\beta}^{\#'}[y'_{t-1}, (t-1)']$  is a stationary series with respect to the measure  $P_{r_0, T}^*$ .

Note that  $S_T^*(s) \xrightarrow{d} W^*(s)$  (see Swensen 2005a), where  $S_T^*(s) = (T-p)^{-1} \sum_{i=k+1}^{[sT]} \varepsilon_i^*$ ,  $0 \leq s \leq 1$ . Here, we consider the Brownian motion  $W^*$  with the estimated covariance matrix  $\tilde{\Omega}$ . We may also consider a Brownian motion  $W$  with covariance matrix  $\Omega$ , since  $\tilde{\Omega}$  converges to  $\Omega$  for  $T \rightarrow \infty$ . Nevertheless, we use  $W^*$  to stress that we refer to convergence conditional on the observations. Now, we replace the unconditional distribution result  $S_T^*(s) \xrightarrow{d} W(s)$  used in Lemma 3 of Swensen (2005a) by the conditional result mentioned before. Then, (A.22) and (A.23) follow from Lemma 3 and its proof in Swensen (2005a).

Finally, we turn to (A.24). Define  $\bar{Z}_1^* = (T-p)^{-1} \sum_{t=p+1}^T Z_{1t}^*$ ,  $Z_{2t}^{l*} = (\Delta y_{t-1}^{*'}, \dots, \Delta y_{t-p+1}^{*'})'$ ,  $\bar{Z}_2^* = (T-p)^{-1} \sum_{t=p+1}^T Z_{2t}^{l*}$ ,  $S_{yy}^* = (T-p)^{-1} \sum_{t=p+1}^T (Z_{1t}^* - \bar{Z}_1^*)(Z_{1t}^* - \bar{Z}_1^*)'$ ,  $S_{yz}^* = (T-p)^{-1} \sum_{t=p+1}^T (Z_{1t}^* - \bar{Z}_1^*)(Z_{2t}^{l*} - \bar{Z}_2^*)'$ , and  $S_{zz}^* = (T-p)^{-1} \sum_{t=p+1}^T (Z_{2t}^{l*} - \bar{Z}_2^*)(Z_{2t}^{l*} - \bar{Z}_2^*)'$ . Then, we have  $S_{11}^* = S_{yy}^* - S_{yz}^* S_{zz}^{*-1} S_{zy}^*$  with  $S_{yz}^* = S_{yz}^{*'}$  corresponding to Johansen (1995, pp. 146-147). Hence,

$$\tilde{C}'_T S_{11}^* \tilde{\beta}^{\#} = \tilde{C}'_T S_{yy}^* \tilde{\beta}^{\#} - \tilde{C}'_T S_{yz}^* S_{zz}^{*-1} S_{zy}^* \tilde{\beta}^{\#}. \quad (\text{A.25})$$

From Lemma 1 of Swensen (2005a) we conclude that  $\tilde{\beta}^{\#'}(Z_{1t}^* - \bar{Z}_1^*)$  and  $(Z_{2t}^{l*} - \bar{Z}_2^*)$  are zero-mean  $I(0)$  series and that  $\tilde{C}'_T(Z_{1t}^* - \bar{Z}_1^*)$  is a zero-mean  $I(1)$  series with respect to the measure

$P_{r_0, T}^*$ . Now, we can apply Theorem B.13 of Johansen (1995) to determine the conditional convergence orders of the product moment matrices involved in (A.25). We obtain that  $\tilde{C}'_T S_{yy}^* \tilde{\beta}^\#$ ,  $\tilde{C}'_T S_{yz}^*$ ,  $S_{zz}^{*-1}$ , and  $S_{zy}^* \tilde{\beta}^\#$  are all  $O_p^*(1)$ . Thus,  $\tilde{C}'_T S_{11}^* \tilde{\beta}^\#$  is also  $O_p^*(1)$  confirming (A.24). This completes the proof of Lemma A2.  $\square$

Using Lemma A.2, the weak limit of (A.18) can be derived. Following Johansen (1995) we see that the ordered eigenvalues of (A.18) converge, conditional on the observations, to those of

$$\left| \lambda \begin{bmatrix} \tilde{\beta}^{\#'} S_{11} \tilde{\beta}^\# & 0 \\ 0 & \int_0^1 G^*(u) G^*(u)' du \end{bmatrix} - \begin{bmatrix} \tilde{\beta}^{\#'} S_{10} S_{00}^{-1} S_{01} \tilde{\beta}^\# & 0 \\ 0 & 0 \end{bmatrix} \right| = 0$$

$$|\lambda \tilde{\beta}^{\#'} S_{11} \tilde{\beta}^\# - \tilde{\beta}^{\#'} S_{10} S_{00}^{-1} S_{01} \tilde{\beta}^\#| |\lambda \int_0^1 G^*(u) G^*(u)' du| = 0,$$

and the space spanned by the  $r$  first eigenvectors of (A.18), given by  $sp(\tilde{A}_T^{-1} \tilde{\beta}^{\#\#})$ , converges to the space spanned by vectors with zeros in the last  $(n - r + 1)$  coordinates. Thus, we have

$$\tilde{A}_T^{-1} \tilde{\beta}^{\#\#} = \begin{bmatrix} I_r \\ T^{1/2} \tilde{\beta}'_{\perp} \tilde{\beta}^* \\ T(\tilde{\tau}' \tilde{\beta}^* - \tilde{\phi}^*) \end{bmatrix} = \begin{bmatrix} I_r \\ T^{1/2} U_T \end{bmatrix} = \begin{bmatrix} I_r \\ o_p^*(1) \\ o_p^*(1) \end{bmatrix},$$

with  $U_T = (\tilde{\beta}^* \tilde{\beta}'_{\perp}, T^{1/2}(\tilde{\beta}^* \tilde{\tau} - \tilde{\phi}^*))'$ . Hence,  $\tilde{\beta}^*$  is a consistent estimator of  $\tilde{\beta}$  conditional on the observations. Using this result, it is straightforward to show that  $\tilde{\tau}' \tilde{\beta}^* \xrightarrow{p^*} \tilde{\tau}' \tilde{\beta} = \tilde{\phi}$ . Thus,  $\tilde{\phi}^*$  is consistent as well and we have  $(\tilde{\phi}^* - \tilde{\phi}) = O_p^*(T^{-1})$  and  $(\tilde{\beta}^* - \tilde{\beta}) = O_p^*(T^{-1/2})$ .

Further, note that  $\tilde{\beta}^{\#\#'} S_{11}^* \tilde{\beta}^{\#\#} = (\tilde{\beta}^\# + \tilde{C}'_T U_T)' S_{11}^* (\tilde{\beta}^\# + \tilde{C}'_T U_T)$  and  $\tilde{\beta}^{\#\#'} S_{10}^* = (\tilde{\beta}^\# + \tilde{C}'_T U_T)' S_{10}^*$  corresponding to Johansen (1995, p. 181). Using (A.19) - (A.21) and the fact that  $U_T = o_p^*(T^{-1/2})$  we obtain  $\tilde{\beta}^{\#\#'} S_{11}^* \tilde{\beta}^{\#\#} \xrightarrow{p^*} \tilde{\beta}^{\#'} S_{11} \tilde{\beta}^\#$  and  $\tilde{\beta}^{\#\#'} S_{10}^* \xrightarrow{p^*} \tilde{\beta}^{\#'} S_{10}$ . Given the definition of  $\tilde{\alpha}^*$  and  $\tilde{\Omega}^*$ , the latter two results in conjunction with (A.19) imply that  $\tilde{\alpha}^*$  and  $\tilde{\Omega}^*$  converge towards  $\tilde{\alpha}$  and  $\tilde{\Omega}$  in  $P_{r_0, T}^*$  probability respectively.

Having established the consistency of  $\tilde{\beta}^*$ ,  $\tilde{\phi}^*$ ,  $\tilde{\alpha}^*$  and  $\tilde{\Omega}^*$ , we now proceed to prove (A.11)-(A.13) following the proof of Johansen (1995, Lemma 13.2). In our case the first order conditions for  $\tilde{\beta}^{\#\#}$  and  $\tilde{\alpha}^*$  read as

$$\tilde{\alpha}^* \tilde{\Omega}^{*-1} (S_{01}^* - \tilde{\alpha}^* \tilde{\beta}^{\#\#'} S_{11}^*) = 0 \quad (\text{A.26})$$

$$(S_{01}^* - \tilde{\alpha}^* \tilde{\beta}^{\#\#'} S_{11}^*) \tilde{\beta}^{\#\#} = 0. \quad (\text{A.27})$$

Further, define

$$S_{\epsilon 1}^* = (T - p)^{-1} \sum_{t=p+1}^T \tilde{\epsilon}_t \tilde{R}'_{1t} = S_{01}^* - \tilde{\alpha} \tilde{\beta}^{\#\#'} S_{11}^*, \quad (\text{A.28})$$

where  $\tilde{\varepsilon}_t^* = \tilde{R}_{0t} - \tilde{\alpha}\tilde{\beta}^{\#\prime}\tilde{R}_{1t}$  (compare Johansen 1995, p. 91).

Insert (A.28) and  $\tilde{C}_T U_T = \tilde{\beta}^{*\#} - \tilde{\beta}^\#$  into (A.26) and multiply by  $\tilde{C}_T$  from the right. Then, using Lemma A.2 and the consistency of  $\tilde{\alpha}^*$  and  $\tilde{\Omega}^*$  we find that  $TU_T = O_p^*(1)$  in line with Johansen (1995, p. 182-183). Recalling the definition of  $U_T$ , we conclude that  $(\tilde{\beta}^* - \tilde{\beta}) = O_p^*(T^{-1})$  and  $(\tilde{\phi}^* - \tilde{\phi}) = O_p^*(T^{-3/2})$ . Now, insert (A.28) into (A.27) and multiply by  $(T-p)^{1/2}$ . Then, using (A.12) and (A.21) gives  $\tilde{\alpha}^* - \tilde{\alpha} = O_p^*(T^{-1/2})$  corresponding to Johansen (1995, p. 183), which shows (A.11).

(A.14) and (A.15) can be obtained from an analog of Theorem 13.5 in Johansen (1995). We do not show this in detail but consider instead the convergence result for  $\tilde{\Omega}^*$  in (A.16), which is not treated by Johansen (1995). The proof of (A.16) closely follows the idea of the proof of Theorem 13.5 in Johansen (1995). Thus, (A.14) and (A.15) are derived accordingly. Write (3.1) as

$$\Delta y_t^* = \tilde{\vartheta} Z_t(\tilde{\beta}^\#) + \varepsilon_t^*, \quad (\text{A.29})$$

where  $Z_t(\tilde{\beta}^\#) = (1, y_{t-1}^{\#\prime}\tilde{\beta}^\#, \Delta y_{t-1}^{\#\prime}, \dots, \Delta y_{t-p+1}^{\#\prime})'$  and  $\tilde{\vartheta} = (\tilde{\nu}, \tilde{\alpha}, \tilde{\Gamma}_1, \dots, \tilde{\Gamma}_{p-1})$ . Note that  $Z_t(\tilde{\beta}^\#)$  is a  $I(0)$  series conditional on the observations. Let  $\Omega^* = (T-p)^{-1} \sum_{t=p+1}^T \varepsilon_t^* \varepsilon_t^{*\prime}$  and  $\tilde{\Omega}^* = (T-p)^{-1} \sum_{t=p+1}^T \tilde{\varepsilon}_t^* \tilde{\varepsilon}_t^{*\prime}$ , where  $\tilde{\varepsilon}_t^*$  are the residuals obtained from estimating (A.29) given  $\tilde{\beta}^\#$ . Tedious but straightforward derivations show that

$$\begin{aligned} (T-p)^{1/2}(\tilde{\Omega}^* - \Omega^*) &= (T-p)^{-1/2} \sum_{t=p+1}^T \varepsilon_t^* (Z_t(\tilde{\beta}^\#) - \bar{Z}_T(\tilde{\beta}^\#))' \\ &\quad \times \left( (T-p)^{-1} \sum_{t=p+1}^T (Z_T(\tilde{\beta}^\#) - \bar{Z}_T(\tilde{\beta}^\#))(Z_t(\tilde{\beta}^\#) - \bar{Z}_T(\tilde{\beta}^\#))' \right)^{-1} \\ &\quad \times (T-p)^{-1} \sum_{t=p+1}^T (Z_T(\tilde{\beta}^\#) - \bar{Z}_T(\tilde{\beta}^\#)) \varepsilon_t^{*\prime}, \end{aligned} \quad (\text{A.30})$$

where  $\bar{Z}_T(\tilde{\beta}^\#) = (T-p)^{-1} \sum_{t=p+1}^T Z_t(\tilde{\beta}^\#)$ . From Theorem B.13 in Johansen (1995) we conclude that the first and second term on the right hand side of (A.30) are  $O_p^*(1)$  and the third term is  $o_p^*(1)$ . Thus,  $\tilde{\Omega}^* - \Omega^* = o_p^*(T^{-1/2})$ , which means that  $\tilde{\Omega}^*$  and  $\Omega^*$  have the same probability limit conditional on the observations. Given (A.12) and (A.13), it is clear that this result also holds if  $\tilde{\beta}^\#$  is replaced by  $\tilde{\beta}^{*\#}$  (compare Johansen 1995, p. 188).

Given that the bootstrap residuals  $\varepsilon_t^*$  are randomly drawn with replacement, they can be regarded as i.i.d. samples from the empirical distribution given by the estimated residuals  $\tilde{\varepsilon}_t$ . In other words, the bootstrap samples  $\varepsilon_t^*$  become i.i.d. with covariance matrix  $\tilde{\Omega} = (T -$



$p)^{-1} \sum_{t=p+1}^T \tilde{\varepsilon}_t \tilde{\varepsilon}_t'$  under the probability measure  $P_{r_0, T}^*$  (compare Park 2002, 2003 for the univariate case). Because we assume that the fourth moments of the error terms exists, a central limit theorem shows that  $\Omega^* - \tilde{\Omega} = O_p^*(T^{-1/2})$  (compare Hamilton (1994, p. 340-343) for the standard asymptotic situation). Hence, we finally get  $\tilde{\Omega}^* - \tilde{\Omega} = O_p^*(T^{-1/2})$ , which proves (A.16). Hence, we have shown Lemma A.2.  $\square$

Saikkonen et al. (2006) consider a transformed version of the relevant VECM in order to introduce certain parameters needed for the derivations later on. In our case we have to use the decomposition  $y_t^* \rightarrow \mu_0^* + \mu_1^* t + x_t^*$  with respect to (3.1) assuming that the conditions (4.2) hold. This yields the transformed VECM

$$\Delta x_t^* = \tilde{\nu}^0 + \tilde{\alpha}(\tilde{\beta}' x_{t-1}^* - \tilde{\phi}^{(0)}(t-1)) + \sum_{j=1}^{p-1} \tilde{\Gamma}_j \Delta x_{t-j}^* + \varepsilon_t^*, \quad t = p+1, p+2, \dots, \quad (\text{A.31})$$

where  $\tilde{\nu}^{(0)} = \tilde{\nu} - \tilde{\Gamma} \mu_1^* + \tilde{\alpha} \tilde{\beta}' \mu_0^*$  and  $\tilde{\phi}^{(0)} = \tilde{\phi} - \tilde{\beta}' \mu_1^*$ . Note that the true values of  $\tilde{\nu}^{(0)}$  and  $\tilde{\phi}^{(0)}$  are zero due to the conditions (4.2).

Define the estimators  $\tilde{\phi}^{*(0)} = \tilde{\phi}^* - \tilde{\beta}^{*'} \mu_1^*$  and  $\tilde{\nu}^{*(0)} = \tilde{\nu}^* - \tilde{\Gamma}^* \mu_1^* + \tilde{\alpha}^* \tilde{\beta}^{*'} \mu_0^*$ , which are obtained by transforming the RR estimators of the VECM (3.1) corresponding to Saikkonen et al. (2006). Their convergence properties are summarized in the next lemma.

**Lemma A3.** Under Assumptions 1 and 2, we have

$$\tilde{\phi}^{*(0)} = O_p^*(T^{-3/2}) \quad (\text{A.32})$$

$$\tilde{\nu}^{*(0)} = O_p^*(T^{-1/2}). \quad (\text{A.33})$$

$\square$

**Proof** The estimators  $\tilde{\phi}^{*(0)}$  and  $\tilde{\nu}^{*(0)}$  can also be obtained from the transformed VECM (A.31). This is a model with a restricted linear trend and unrestricted mean term as model (3.1). Thus, we can apply (A.13) of Lemma A.1 for the estimators of the trend parameter  $\tilde{\phi}^{(0)}$  and the mean term  $\tilde{\nu}^{(0)}$ . Since they have true values of zero, (A.32) follows.  $\square$

We are now in the position to state the properties of the estimator of  $\mu_1^*$  based on the bootstrap data.

**Lemma A4.** Under Assumptions 1 and 2, we have

$$\tilde{\beta}'(\hat{\mu}_1^* - \mu_1^*) = O_p^*(T^{-3/2}) \quad (\text{A.34})$$

$$(T-p)^{-1/2} \tilde{\beta}'_{\perp}(\hat{\mu}_1^* - \mu_1^*) \xrightarrow{d^*} N(0, \tilde{\beta}'_{\perp} \tilde{C} \tilde{\Omega} \tilde{C}' \tilde{\beta}_{\perp}). \quad (\text{A.35})$$

□

**Proof** The proof of Lemma A4 is analogous to the corresponding part of the proof of Lemma A.14 in Saikkonen et al. (2006). There, the properties of the estimator  $\hat{\mu}_1$  based on the original observations are shown. Given the definition of  $\tilde{\nu}^{(0)}$  we have  $\tilde{\beta}'_{\perp} \tilde{C} \tilde{\nu}^{(0)} = \tilde{\beta}'_{\perp} \tilde{C} \tilde{\nu} - \tilde{\beta}'_{\perp} \tilde{C} \tilde{\Gamma} \mu_1^*$ . From Saikkonen et al. (2006, Section 4)  $\tilde{\beta}'_{\perp} \tilde{C} \tilde{\nu} = \tilde{\beta}'_{\perp} \hat{\mu}_1 + \tilde{\beta}'_{\perp} \tilde{C} \tilde{\Gamma} \tilde{\beta} \tilde{\phi}$  is obtained. Using the latter equality and the definition of  $\tilde{\phi}^{(0)}$  results in  $\tilde{\beta}'_{\perp} \tilde{C} \tilde{\nu}^{(0)} = \tilde{\beta}'_{\perp} \hat{\mu}_1 - \tilde{\beta}'_{\perp} \tilde{C} \tilde{\Gamma} \mu_1^* + \tilde{\beta}'_{\perp} \tilde{C} \tilde{\Gamma} \tilde{\beta} \tilde{\phi}' \mu_1^* + \tilde{\beta}'_{\perp} \tilde{C} \tilde{\Gamma} \tilde{\beta} \tilde{\phi}^{(0)}$  such that  $\tilde{\beta}'_{\perp}(\hat{\mu}_1 - \mu_1^*) = \tilde{\beta}'_{\perp} \tilde{C}(\tilde{\nu}^{(0)} - \tilde{\Gamma} \tilde{\beta} \tilde{\phi}^{(0)})$ . Since the same relation applies to the estimators based on the pseudo observations one has

$$\tilde{\beta}^{*'}_{\perp}(\hat{\mu}_1^* - \mu_1^*) = \tilde{\beta}^{*'}_{\perp} \tilde{C}^* (\tilde{\nu}^{*(0)} - \tilde{\Gamma} \tilde{\beta}^* \tilde{\phi}^{*(0)}).$$

Then, by Lemma A.2,

$$\tilde{\beta}^{*'}_{\perp}(\hat{\mu}_1^* - \mu_1^*) = \tilde{\beta}^{*'}_{\perp} \tilde{C}^* \tilde{\nu}^{*(0)} + o_p^*(T^{-1/2})$$

follows. Corresponding to Saikkonen et al. (2006) the estimator  $\tilde{\nu}^{*(0)}$  can be viewed as the LS estimator of  $\tilde{\nu}^{(0)}$  in the auxiliary model

$$\Delta x_t^* = \tilde{\nu}^{(0)} + \tilde{\alpha}(\tilde{\beta}^{*'} x_{t-1}^* - \tilde{\phi}^{*(0)}(t-1)) + \sum_{j=1}^{p-1} \tilde{\Gamma}_j \Delta x_{t-j}^* + u_t^*, \quad (\text{A.36})$$

which is obtained from (A.31) by replacing  $\tilde{\beta}$  and  $\tilde{\phi}^{(0)}$  by their bootstrap analogs  $\tilde{\beta}^*$  and  $\tilde{\phi}^{*(0)}$  and by noting that  $u_t^* = \varepsilon_t^* - \tilde{\alpha}[(\tilde{\beta}^* - \tilde{\beta})' x_{t-1}^* + \tilde{\phi}^{*(0)}(t-1)]$ ,  $\tilde{\beta}' y_{t-1}^{c*} = \tilde{\beta}' x_{t-1}^* + \theta^*$ ,  $\tilde{\beta}^{*'} x_{t-1}^* = \tilde{\beta}^{*'} y_{t-1}^{c*} - \tilde{\theta}^*$ , and that  $\Delta y_{t-i}^{c*} = \Delta x_{t-i}^*$  ( $i = 0, \dots, p-1$ ). Keep in mind, that the estimator  $\tilde{\theta}^*$  can be recovered from the RR estimators of (3.1).

Equation (A.36) implies that the estimator  $\tilde{\beta}^{*'}_{\perp} \tilde{C}^* \tilde{\nu}^{*(0)}$  can be obtained by LS from

$$\tilde{\beta}^{*'}_{\perp} \tilde{C}^* \Delta x_t^* = \tilde{\Lambda} p_t^* + \tilde{u}_t^*, \quad (\text{A.37})$$

where  $p_t^* = [1, \Delta x_{t-1}^*, \dots, \Delta x_{t-p+1}^*]$ ,  $\tilde{\Lambda}$  is a conformable coefficient matrix and  $\tilde{u}_t^* = \tilde{\beta}^{*'}_{\perp} \tilde{C}^* \varepsilon_t^* - \tilde{\beta}^{*'}_{\perp} \tilde{C}^* \tilde{\alpha}[(\tilde{\beta}^* - \tilde{\beta})' x_{t-1}^* + \tilde{\phi}^{*(0)}(t-1)]$ . By the definition of  $\tilde{C}$  and Lemma A.2,  $\tilde{\beta}^{*'}_{\perp} \tilde{C}^* \tilde{\alpha} = O_p^*(T^{-1/2})$ . Using this fact, Lemma A.2 and the assumptions, it is straightforward to show that conditional on the observations the asymptotic properties of the LS estimator of  $\tilde{\Lambda}$  in the auxiliary regression model (A.37) can be obtained by assuming that the error term equals  $\tilde{\beta}'_{\perp} \tilde{C} \varepsilon_t^*$ .

Following Saikkonen et al. (2006), it is also easy to show that the estimation of the intercept term in (A.37) is asymptotically independent of the estimation of the other regression coefficients. Thus, we asymptotically obtain  $\tilde{\beta}'_{\perp} \tilde{C}' \tilde{\nu}^{*(0)} = (T-p)^{-1} \tilde{\beta}'_{\perp} \tilde{C}' \sum_{t=p+1}^T \varepsilon_t^*$  since the true value of  $\tilde{\nu}^{*(0)}$  is zero. Therefore, we can conclude that

$$(T-p)^{1/2} \tilde{\beta}'_{\perp} (\hat{\mu}_1^* - \mu_1^*) = \tilde{\beta}'_{\perp} \tilde{C}' (T-p)^{-1/2} \sum_{t=p+1}^T \varepsilon_t^* + o_p^*(1).$$

This, Lemma 1 in Swensen (2005a), and a standard central limit theorem yield

$$(T-p)^{1/2} \tilde{\beta}'_{\perp} (\hat{\mu}_1^* - \mu_1^*) \xrightarrow{d^*} N(0, \tilde{\beta}'_{\perp} \tilde{C}' \tilde{\Omega} \tilde{C}' \tilde{\beta}_{\perp}). \quad (\text{A.38})$$

In order to obtain (A.35) we need to show that  $\tilde{\beta}_{\perp}^*$  on the l.h.s. of (A.38) can be replaced by  $\tilde{\beta}_{\perp}$ .

First note that

$$(\tilde{\beta}_{\perp}^* - \tilde{\beta}_{\perp})' (\hat{\mu}_1^* - \mu_1^*) = (\tilde{\beta}_{\perp}^* - \tilde{\beta}_{\perp})' \tilde{\beta}^* (\tilde{\beta}'^* \tilde{\beta}^*)^{-1} \tilde{\beta}'^* (\hat{\mu}_1^* - \mu_1^*) + (\tilde{\beta}_{\perp}^* - \tilde{\beta}_{\perp})' \tilde{\beta}_{\perp}^* (\tilde{\beta}'_{\perp} \tilde{\beta}_{\perp}^*)^{-1} \tilde{\beta}'_{\perp} (\hat{\mu}_1^* - \mu_1^*) \quad (\text{A.39})$$

Due to the consistency of the estimator  $\tilde{\beta}^*$  with respect to  $\tilde{\beta}$  and the result (A.38) the latter term on the r.h.s. is of order  $o_p^*(T^{-1/2})$ . The same is true for the former term because  $\tilde{\beta}'^* (\hat{\mu}_1^* - \mu_1^*) = \tilde{\phi}^{*(0)} = O_p^*(T^{-3/2})$  given the definition of  $\tilde{\phi}^{*(0)}$ , the conditions (4.2), and Lemma A.3. Thus, (A.35) follows and (A.34) is obtained if  $\tilde{\beta}^*$  in the latter equality can be replaced by  $\tilde{\beta}$ . To this end, write

$$(\tilde{\beta}^* - \tilde{\beta})' (\hat{\mu}_1^* - \mu_1^*) = (\tilde{\beta}^* - \tilde{\beta})' \tilde{\beta}^* (\tilde{\beta}'^* \tilde{\beta}^*)^{-1} \tilde{\beta}'^* (\hat{\mu}_1^* - \mu_1^*) + (\tilde{\beta}^* - \tilde{\beta})' \tilde{\beta}_{\perp}^* (\tilde{\beta}'_{\perp} \tilde{\beta}_{\perp}^*)^{-1} \tilde{\beta}'_{\perp} (\hat{\mu}_1^* - \mu_1^*).$$

From the arguments made in relation to (A.39) we can deduce that both terms on the r.h.s. are  $o_p^*(T^{-1/2})$  such that (A.34) is shown. This completes the proof of Lemma A.4.  $\square$

As described in Section 4, the  $SLT_1^*$  test is based on the VECM for the adjusted bootstrap series  $\hat{y}_t^{c*} = y_t^* - \hat{\mu}_1^* t$  given by

$$\Delta \hat{y}_t^{c*} = \tilde{\alpha} (\tilde{\beta}' \hat{y}_{t-1}^{c*} - \theta^*) + \sum_{j=1}^{p-1} \tilde{\Gamma}_j \Delta \hat{y}_{t-j}^{c*} + e_t^{c*}, \quad (\text{A.40})$$

where  $e_t^{c*} = \varepsilon_t^* - \tilde{\alpha} \tilde{\beta}' (\hat{y}_{t-1}^{c*} - y_{t-1}^{c*}) (t-1) + \Delta \hat{y}_t^{c*} - \Delta y_t^{c*} - \sum_{j=1}^{p-1} \tilde{\Gamma}_j (\Delta \hat{y}_{t-j}^{c*} - \Delta y_{t-j}^{c*}) = \varepsilon_t^* + \tilde{\alpha} \tilde{\beta}' (\hat{\mu}_1^* - \mu_1^*) (t-1) - \tilde{\Gamma} (\hat{\mu}_1^* - \mu_1^*)$ . Define  $\tilde{Z}_{0t}^{c*} = \Delta \hat{y}_t^{c*}$ ,  $\tilde{Z}_{1t}^{c*} = (\hat{y}_{t-1}^{c*}, 1)'$ , and  $\tilde{Z}_{2t}^{c*} = (\Delta \hat{y}_{t-1}^{c*}, \dots, \Delta \hat{y}_{t-p+1}^{c*}, 1)'$ . Then, regress  $\tilde{Z}_{0t}^{c*}$  and  $\tilde{Z}_{1t}^{c*}$  on  $\tilde{Z}_{2t}^{c*}$  to obtain the residuals  $\tilde{R}_{0t}^{c*}$  and  $\tilde{R}_{1t}^{c*}$ , respectively. Using these residuals we write

$$\tilde{S}_{ij}^{c*} = (T-p)^{-1} \sum_{t=p+1}^T \tilde{R}_{it}^{c*} \tilde{R}_{jt}^{c*} \quad i, j = 0, 1.$$

In the same way we can obtain the moment matrices  $S_{ij}^{c*}$  ( $i, j = 0, 1$ ) from  $Z_{0t}^{c*}$ ,  $Z_{1t}^{c*}$ ,  $Z_{2t}^{c*}$ ,  $R_{0t}^{c*}$ , and  $R_{1t}^{c*}$ , which are defined in terms of  $y_t^{c*} = y_t^* - \mu_1^* t$ . Furthermore, let  $S_{ij}^c$  ( $i, j = 0, 1$ ) be the moment matrices based on the original observations adjusted by the trend component, i.e.  $y_t^c = y_t - \mu_1 t$ . Finally, define  $\tilde{\beta}^+ = (\tilde{\beta}', -\theta^*)'$ .

In order to proof Proposition 1 we follow Johansen (1995, Ch. 11). Thus, we need corresponding results to Lemma 2 and 3 of Swensen (2005a) with respect to model (A.40). To this end, we consider first the following lemma.

**Lemma A5.** Under Assumptions 1 and 2, we have

$$\begin{aligned}\tilde{S}_{00}^{c*} &\xrightarrow{p} \Sigma_{00}^c \\ \tilde{S}_{01}^{c*} \tilde{\beta}^+ &\xrightarrow{p} \Sigma_{0\beta}^c \\ \tilde{\beta}^{+'} \tilde{S}_{11}^{c*} \tilde{\beta}^+ &\xrightarrow{p} \Sigma_{\beta\beta}^c,\end{aligned}$$

where  $\Sigma_{00}^c$ ,  $\Sigma_{0\beta}^c$ , and  $\Sigma_{\beta\beta}^c$  are the respective probability limits of  $S_{00}^c$ ,  $S_{01}^c \beta^+$ , and  $\beta^{+'} S_{11}^c \beta^+$  with respect to the original observations as  $T \rightarrow \infty$ .  $\square$

**Proof** We first prove that  $\tilde{S}_{00}^{c*} \xrightarrow{p^*} S_{00}^{c*}$ ,  $\tilde{S}_{01}^{c*} \tilde{\beta}^+ \xrightarrow{p^*} S_{01}^{c*} \tilde{\beta}^+$ , and  $\tilde{\beta}^{+'} \tilde{S}_{11}^{c*} \tilde{\beta}^+ \xrightarrow{p^*} \tilde{\beta}^{+'} S_{11}^{c*} \tilde{\beta}^+$ . Hence, conditional on the observations, the use of the estimator  $\hat{\mu}_1^*$  instead of  $\mu_1^*$  for the bootstrap data adjustment has no asymptotic impact on the properties of the product moment matrices of interest. Then, Lemma A5 follows from Lemma 2 of Swensen (2005a) in connection with Lemma 1 above.

Using the definitions stated before Lemma A5, we can write in accordance with Johansen (1995, Ch. 6)

$$\begin{aligned}\tilde{S}_{00}^{c*} &= (T-p)^{-1} \left( \sum_{t=p+1}^T \tilde{Z}_{0t}^{c*} \tilde{Z}_{0t}^{c*'} - \sum_{t=p+1}^T \tilde{Z}_{0t}^{c*} \tilde{Z}_{2t}^{c*'} \left( \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} \right)^{-1} \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{0t}^{c*'} \right) \\ \tilde{S}_{01}^{c*} \tilde{\beta}^+ &= (T-p)^{-1} \left( \sum_{t=p+1}^T \tilde{Z}_{0t}^{c*} \tilde{Z}_{1t}^{c*'} - \sum_{t=p+1}^T \tilde{Z}_{0t}^{c*} \tilde{Z}_{2t}^{c*'} \left( \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} \right)^{-1} \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{1t}^{c*'} \right) \tilde{\beta}^+ \\ \tilde{\beta}^{+'} \tilde{S}_{11}^{c*} \tilde{\beta}^+ &= (T-p)^{-1} \tilde{\beta}^{+'} \left( \sum_{t=p+1}^T \tilde{Z}_{1t}^{c*} \tilde{Z}_{1t}^{c*'} - \sum_{t=p+1}^T \tilde{Z}_{1t}^{c*} \tilde{Z}_{2t}^{c*'} \left( \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} \right)^{-1} \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{1t}^{c*'} \right) \tilde{\beta}^+.\end{aligned}$$

Note that  $\sum_{t=p+1}^T \tilde{Z}_{0t}^{c*} \tilde{Z}_{0t}^{c*'} = \sum_{t=p+1}^T [Z_{0t}^{c*} - (\hat{\mu}_1^* - \mu_1^*)][Z_{0t}^{c*} - (\hat{\mu}_1^* - \mu_1^*)']$ , where  $Z_{0t}^{c*} = \Delta y_t^{c*}$  is a zero-mean stationary process with respect to the measure  $P_{r_0, T}^*$  according to Lemma 1. Using Lemma A4 and Johansen (1995, Theorem B.13), we have  $\sum_{t=p+1}^T \tilde{Z}_{0t}^{c*} \tilde{Z}_{0t}^{c*'} = \sum_{t=p+1}^T Z_{0t}^{c*} Z_{0t}^{c*'} +$

$O_p^*(1)$ . Similarly,  $\sum_{t=p+1}^T \tilde{Z}_{0t}^{c*} \tilde{Z}_{2t}^{c*'} = \sum_{t=p+1}^T Z_{0t}^{c*} Z_{2t}^{c*'} + O_p^*(1)$  and  $\sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} = \sum_{t=p+1}^T Z_{2t}^{c*} Z_{2t}^{c*'} + O_p^*(1)$  are obtained. Combining these results, yields

$$\begin{aligned} \tilde{S}_{00}^{c*} &= (T-p)^{-1} \left( \sum_{t=p+1}^T Z_{0t}^{c*} Z_{0t}^{c*'} - \sum_{t=p+1}^T Z_{0t}^{c*} Z_{2t}^{c*'} \left( \sum_{t=p+1}^T Z_{2t}^{c*} Z_{2t}^{c*'} \right)^{-1} \sum_{t=p+1}^T Z_{2t}^{c*} Z_{0t}^{c*'} \right) + o_p^*(1) \\ &= S_{00}^{c*} + o_p^*(1). \end{aligned}$$

Moreover,  $\sum_{t=p+1}^T \tilde{Z}_{0t}^{c*} \tilde{Z}_{1t}^{c*'} \tilde{\beta}^+ = \sum_{t=p+1}^T [Z_{0t}^{c*} - (\hat{\mu}_1^* - \mu_1^*)][Z_{1t}^{c*} - \{(\hat{\mu}_1^* - \mu_1^*)'(t-1) : 0\}'] \tilde{\beta}^+$ . Note that  $\tilde{\beta}^{+'} Z_{1t}^{c*} = \tilde{\beta}' y_{t-1}^{c*} - \theta^*$  is also a zero-mean stationary process with respect to the measure  $P_{r_0, T}^*$ . Again, we refer to Lemma A4 and Johansen (1995, Theorem B.13) to conclude that  $\sum_{t=p+1}^T \tilde{Z}_{0t}^{c*} \tilde{Z}_{1t}^{c*'} \tilde{\beta}^+ = \sum_{t=p+1}^T Z_{0t}^{c*} Z_{1t}^{c*'} \tilde{\beta}^+ + O_p^*(1)$ . In a similar fashion, we obtain  $\sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{1t}^{c*'} \tilde{\beta}^+ = \sum_{t=p+1}^T Z_{2t}^{c*} Z_{1t}^{c*'} \tilde{\beta}^+ + O_p^*(1)$  such that, together with the results found above,  $\tilde{S}_{01}^{c*} \tilde{\beta}^+ = S_{01}^{c*} \tilde{\beta}^+ + o_p^*(1)$  can be deduced. By an analogous argumentation,  $\sum_{t=p+1}^T \tilde{\beta}^{+'} \tilde{Z}_{1t}^{c*} \tilde{Z}_{1t}^{c*'} \tilde{\beta}^+ = \sum_{t=p+1}^T \tilde{\beta}^{+'} Z_{1t}^{c*} Z_{1t}^{c*'} \tilde{\beta}^+ + O_p^*(1)$  can be shown. Using this, we finally find  $\tilde{\beta}^{+'} \tilde{S}_{11}^{c*} \tilde{\beta}^+ = \tilde{\beta}^{+'} S_{11}^{c*} \tilde{\beta}^+ + o_p^*(1)$ , what completes the proof of Lemma A5.  $\square$

As an intermediate step in deriving the distributional properties of the relevant quantities we first show

**Lemma A6.** Under Assumptions 1 and 2, we have

$$(T-p)^{-1} \tilde{B}'_T \tilde{S}_{11}^{c*} \tilde{B}_T \xrightarrow{p^*} (T-p)^{-2} \tilde{B}'_T \sum_{t=p+1}^T \tilde{Z}_{1t}^{c*} \tilde{Z}_{1t}^{c*'} \tilde{B}_T \quad (\text{A.41})$$

$$\tilde{B}'_T \tilde{S}_{10}^{c*} \tilde{\alpha}_\perp \xrightarrow{p^*} (T-p)^{-1} \tilde{B}'_T \sum_{t=p+1}^T \tilde{Z}_{1t}^{c*} \tilde{\varepsilon}_t^{c*'} \tilde{\alpha}_\perp \quad (\text{A.42})$$

where  $\tilde{B}_T = \begin{pmatrix} \tilde{\beta}_\perp & 0 \\ 0 & T^{1/2} \end{pmatrix}$  and  $\tilde{\varepsilon}_t^{c*} = \varepsilon_t^* - \tilde{\Gamma} \tilde{\beta}_\perp \tilde{\beta}'_\perp (\hat{\mu}_1^* - \mu_1^*)$ .  $\square$

**Proof** The estimation error with respect to  $\hat{\mu}_1^*$  does not vanish on the r.h.s. of (A.41) because  $\tilde{\beta}'_\perp (\hat{\mu}_1^* - \mu_1^*) = O_p^*(T^{-1/2})$  in contrast to  $\tilde{\beta}' (\hat{\mu}_1^* - \mu_1^*) = O_p^*(T^{-3/2})$  according to Lemma A4. Thus, we only have to show that

$$(T-p)^{-2} \sum_{t=p+1}^T \tilde{B}'_T \tilde{Z}_{1t}^{c*} \tilde{Z}_{2t}^{c*'} \left( \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} \right)^{-1} \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{1t}^{c*'} \tilde{B}_T \xrightarrow{p^*} 0. \quad (\text{A.43})$$

Note that  $\tilde{\beta}'_\perp y_{t-1}^{c*}$  equals a zero-mean process plus a bounded term with respect to the measure  $P_{r_0, T}^*$ . This follows from the definition of  $\tau_0^*$  in Lemma 1 and  $\sqrt{T} R_{t, T}^{c*}$  in (A.10). Then, using

the results of Lemmas 1 and A4 in connection with Theorem B.13 of Johansen (1995) we find that  $\sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} = O_p^*(T)$ ,  $\sum_{t=p+1}^T \tilde{\beta}'_{\perp} \hat{y}_{t-1}^{c*} \tilde{Z}_{2t}^{c*'} = O_p^*(T)$ , and  $\sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} = O_p^*(T^{1/2})$ . Combining the latter two results yields  $\sum_{t=p+1}^T \tilde{B}'_T \tilde{Z}_{1t}^{c*} \tilde{Z}_{2t}^{c*'} = O_p^*(T)$  and (A.43) follows, which implies (A.41).

In line with Johansen (1995, p. 91) we can write  $\tilde{S}_{10}^{c*} \tilde{\alpha}_{\perp} = (T-p)^{-1} \sum_{t=p+1}^T \tilde{R}_{1t}^{c*} (\tilde{\alpha} \tilde{\beta}^{+'} \tilde{R}_{1t}^{c*} + \tilde{\epsilon}_t^{c*})' \tilde{\alpha}_{\perp} = (T-p)^{-1} \sum_{t=p+1}^T \tilde{R}_{1t}^{c*} \tilde{\epsilon}_t^{c*'} \tilde{\alpha}_{\perp}$ , where  $\tilde{\epsilon}_t^{c*} = \tilde{R}_{0t}^{c*} - \tilde{\alpha} \tilde{\beta}^{+'} \tilde{R}_{1t}^{c*}$ . Given the derivations used in Johansen (1995, pp. 90-91) and our notation with respect to the VEC model (A.40) it is straightforward to show that  $\tilde{\epsilon}_t^{c*} = e_t^{c*} \left( I - \tilde{Z}_{2t}^{c*'} \left( \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} \right)^{-1} \tilde{Z}_{2t}^{c*} \right)$  and  $\tilde{R}_{1t}^{c*} = \tilde{Z}_{1t}^{c*} - \sum_{t=p+1}^T \tilde{Z}_{1t}^{c*} \tilde{Z}_{2t}^{c*'} \left( \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} \right)^{-1} \tilde{Z}_{2t}^{c*}$ . Using these expressions, we have

$$\tilde{S}_{10}^{c*} \tilde{\alpha}_{\perp} = (T-p)^{-1} \left( \sum_{t=p+1}^T \tilde{Z}_{1t}^{c*} e_t^{c*'} - \sum_{t=p+1}^T \tilde{Z}_{1t}^{c*} \tilde{Z}_{2t}^{c*'} \left( \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} \right)^{-1} \sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} e_t^{c*'} \right) \tilde{\alpha}_{\perp}$$

Note again that  $e_t^{c*} = \varepsilon_t^* + \tilde{\alpha} \tilde{\beta}' (\hat{\mu}_1^* - \mu_1^*) (t-1) - \tilde{\Gamma} (\hat{\mu}_1^* - \mu_1^*)$ . Then, by Lemmas 1 and 4 together with Johansen (1995, Theorem B.13),  $\sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} e_t^{c*'} \tilde{\alpha}_{\perp} = O_p^*(T^{1/2})$ . Combining this with  $\sum_{t=p+1}^T \tilde{B}'_T \tilde{Z}_{1t}^{c*} \tilde{Z}_{2t}^{c*'} = O_p^*(T)$  and  $\sum_{t=p+1}^T \tilde{Z}_{2t}^{c*} \tilde{Z}_{2t}^{c*'} = O_p^*(T)$  shows that  $\tilde{B}'_T \tilde{S}_{10}^{c*} \tilde{\alpha}_{\perp} = (T-p)^{-1} \sum_{t=p+1}^T \tilde{B}'_T \tilde{Z}_{1t}^{c*} e_t^{c*'} \tilde{\alpha}_{\perp} + o_p^*(1)$ .

As in case of (A.41), the estimation error regarding  $\hat{\mu}_1^*$  does not vanish in  $\tilde{B}'_T \tilde{S}_{10}^{c*} \tilde{\alpha}_{\perp} = (T-p)^{-1} \sum_{t=p+1}^T \tilde{B}'_T \tilde{Z}_{1t}^{c*} e_t^{c*'} \tilde{\alpha}_{\perp}$ . However, we can write this expression more explicitly by noting that  $e_t^{c*'} \tilde{\alpha}_{\perp} = \{ \varepsilon_t^* - \tilde{\Gamma} [\tilde{\beta} (\tilde{\beta}' \tilde{\beta})^{-1} \tilde{\beta}' + \tilde{\beta}_{\perp} (\tilde{\beta}'_{\perp} \tilde{\beta}_{\perp})^{-1} \tilde{\beta}'_{\perp}] (\hat{\mu}_1^* - \mu_1^*) \}' \tilde{\alpha}_{\perp}$ . Since  $\tilde{\beta}' (\hat{\mu}_1^* - \mu_1^*) = O_p^*(T^{-3/2})$ , we obtain  $\tilde{B}'_T \tilde{S}_{10}^{c*} \tilde{\alpha}_{\perp} = (T-p)^{-1} \sum_{t=p+1}^T \tilde{B}'_T \tilde{Z}_{1t}^{c*} [\varepsilon_t^* - \tilde{\Gamma} \tilde{\beta}_{\perp} (\tilde{\beta}'_{\perp} \tilde{\beta}_{\perp})^{-1} \tilde{\beta}'_{\perp} (\hat{\mu}_1^* - \mu_1^*)]' \tilde{\alpha}_{\perp} + o_p^*(1)$ , what shows (A.42) and completes the proof of Lemma A6.  $\square$

We can now state

**Lemma A7.** Under Assumptions 1 and 2,

$$(T-p)^{-1} \tilde{B}'_T \tilde{S}_{11}^{c*} \tilde{B}_T \rightarrow \int_0^1 G_+(s) G_+(s)' ds \quad (\text{A.44})$$

$$\tilde{B}'_T \tilde{S}_{10}^{c*} \tilde{\alpha}_{\perp} \rightarrow \int_0^1 G_+(s) dG_*(s)' \alpha_{\perp}, \quad (\text{A.45})$$

where  $G_+(s) = [B_*(s)' \Omega^{1/2} C' \beta_{\perp}, 1]'$ ,  $dG_*(s) = \Omega^{1/2} dB_*(s)$ , and convergence is weakly in  $P_{r_0, T}$  probability.  $\square$

**Proof** We can exactly argue as in Swensen (2005a, Proof of Lemma 3). Our proof also relies on the continuous mapping theorem involving the same functionals as in Saikkonen et al. (2006)

and related work by Lütkepohl & Saikkonen (2000) and Hansen & Johansen (1998, pp. 118-121). Thus, analogous to Swensen (2005a), we only have to show that the process  $\{(T - p)^{-1/2} y_{[sT]}^{c*}, 0 \leq s \leq 1\}$  converges weakly in  $P_{r_0, T}$  probability as element in  $D[0, 1]^n$ .

From Lemma 1 we see, that the remainder term  $R_{t, T}^{c*}$  vanishes asymptotically. The constant term is treated as in Saikkonen et al. (2006). A more detailed description of this treatment in relation to the Johansen test can be found in Hansen & Johansen (1998, pp. 118-121). Therefore, we only have to analyse  $S_T^*(s) = (T - p)^{-1/2} \sum_{i=p+1}^{[sT]} \varepsilon_i^*$ . This is the same partial sum analysed by Swensen (2005a) except that the our bootstrap residuals are based on Algorithm 1 and not on Algorithm 2. However, according to Swensen (2005a, Remark 9) this does not change the asymptotic results.

Hence, following the arguments of Swensen (2005a), convergence in  $P_{r_0, T}$  probability is obtained by using the approach of Pollard (1984, Theorem V.19). It can be shown that  $E^*[f(S_T^*)] \rightarrow E[f(W)]$  in  $P_{r_0, T}$  probability for all bounded continuous functions  $f$  defined on  $D[0, 1]^n$ , the space of functions that are right continuous and have left limits. Thus,  $S_T^*$  converges weakly in  $P_{r_0, T}$  probability to a Brownian motion  $W$  with covariance matrix  $\Omega$ . Furthermore, note that the estimators  $\tilde{\beta}_\perp$ ,  $\tilde{\alpha}_\perp$ , and  $\tilde{\Gamma}$  involved in the l.h.s. of (A.44) and (A.45) are consistent in terms of the measure  $P_{r_0, T}$  (compare e.g. Johansen 1995). Moreover, we have  $\tilde{B}'_T \tilde{Z}_{1t}^{c*} = \{[\tilde{\beta}'_\perp y_{t-1}^{c*} - \tilde{\beta}'_\perp (\hat{\mu}_1^* - \mu_1^*)(t - 1)]', T^{1/2}\}'$ . Then, given the consideration of the (restricted) constant term as in Saikkonen et al. (2006) or Hansen & Johansen (1998, pp. 118-121), Lemma A7 follows from the relationship of the functionals on the r.h.s. of (A.41) and (A.42) with those in equations (A.10)/(A.11) and (A.14)/(A.15) of Lütkepohl & Saikkonen (2000). Note in this respect that the estimator  $\tilde{\mu}_1$  used in Lütkepohl & Saikkonen (2000) and  $\hat{\mu}_1$  have the same asymptotic properties. Thus, the asymptotic properties of  $\hat{\mu}_1^*$  also correspond to those of  $\tilde{\mu}_1$  in Lütkepohl & Saikkonen (2000).  $\square$

Finally, Proposition 1 follows from Lemmas A5 and A7 in the same way as in Johansen (1995, Ch. 11) and Saikkonen et al. (2006) with obvious modifications (compare Swensen 2005a).  $\square$

## Proof of Proposition 2

We now prove Proposition 2. First, we need

**Lemma A8.** Under Assumptions 1 and 3, we have

$$\tilde{\beta}'(\tilde{\mu}_0^* - \mu_0^*) = O_p^*(T^{-1/2}) \quad (\text{A.46})$$

$$\tilde{\beta}'_{\perp}(\tilde{\mu}_0^* - \mu_0^*) = O_p^*(1) \quad (\text{A.47})$$

$$\tilde{\beta}'(\tilde{\mu}_1^* - \mu_1^*) = O_p^*(T^{-3/2}) \quad (\text{A.48})$$

$$(T - p)^{-1/2} \tilde{\beta}'_{\perp}(\tilde{\mu}_1^* - \mu_1^*) \xrightarrow{d^*} N(0, \tilde{\beta}'_{\perp} \tilde{C} \tilde{\Omega} \tilde{C}' \tilde{\beta}_{\perp}) \quad (\text{A.49})$$

□

**Proof** First, it is shown that Lemma A8 holds if  $\tilde{\beta}$  and  $\tilde{\beta}_{\perp}$  are replaced by  $\tilde{\beta}^*$  and  $\tilde{\beta}_{\perp}^*$ . Then, the lemma follows from (A.12). Thus, we consider the following two sets of quantities:  $\underline{\theta}^* = \tilde{\beta}'^* \mu_0^*$ ,  $\underline{\theta}_*^* = \tilde{\beta}'_{\perp}^* \mu_0^*$ ,  $\underline{\tau}^* = \tilde{\beta}'^* \mu_1^*$ ,  $\underline{\tau}_*^* = \tilde{\beta}'_{\perp}^* \mu_1^*$  and  $\theta^* = \tilde{\beta}' \mu_0^*$ ,  $\theta_*^* = \tilde{\beta}'_{\perp} \mu_0^*$ ,  $\tau^* = \tilde{\beta}' \mu_1^*$ ,  $\tau_*^* = \tilde{\beta}'_{\perp} \mu_1^*$ . The second set can be regarded as the real parameters conditional on the observations and the first set comprises the analogs of the real parameters that are considered first.

The estimators  $\tilde{\theta}^*$ ,  $\tilde{\theta}_*^*$ ,  $\tilde{\tau}^*$ , and  $\tilde{\tau}_*^*$  are obtained by LS estimation of

$$\begin{aligned} \tilde{Q}^*{}' \tilde{A}^*(L) y_t^* &= \tilde{Q}^*{}' \tilde{H}_{0t}^* \tilde{\beta}^* (\tilde{\beta}'^* \tilde{\beta}^*)^{-1} \underline{\theta}^* + \tilde{Q}^*{}' \tilde{H}_{0t}^* \tilde{\beta}_{\perp}^* (\tilde{\beta}'_{\perp}^* \tilde{\beta}_{\perp}^*)^{-1} \underline{\theta}_*^* \\ &+ \tilde{Q}^*{}' \tilde{H}_{1t}^* \tilde{\beta}^* (\tilde{\beta}'^* \tilde{\beta}^*)^{-1} \underline{\tau}^* + \tilde{Q}^*{}' \tilde{H}_{1t}^* \tilde{\beta}_{\perp}^* (\tilde{\beta}'_{\perp}^* \tilde{\beta}_{\perp}^*)^{-1} \underline{\tau}_*^* + \tilde{\eta}_t^*, \quad t = 1, \dots, T, \end{aligned} \quad (\text{A.50})$$

which is a rewritten version of (4.10). For the error terms we have  $\tilde{\eta}_t^* = \tilde{Q}^*{}' n_t^* = \tilde{Q}^*{}' (\varepsilon_t^* - (\tilde{\alpha}^* \tilde{\beta}'^* - \tilde{\alpha} \tilde{\beta}') x_{t-1}^* - \sum_{j=1}^{p-1} (\tilde{\Gamma}_j^* - \tilde{\Gamma}_j) \Delta x_{t-j}^*)$  corresponding to (A.4) in Saikkonen & Lütkepohl (1997), which is the discussion paper version of Saikkonen & Lütkepohl (2000b).

The definition of  $\tilde{H}_{0t}^*$  leads to  $\tilde{H}_{0t}^* \tilde{\beta}_{\perp}^* (\tilde{\beta}'_{\perp}^* \tilde{\beta}_{\perp}^*)^{-1} = 0$  for  $t \geq p + 1$ . Corresponding to Lütkepohl & Saikkonen (1997, Proof of Lemma 2), which is the discussion paper version of Lütkepohl & Saikkonen (2000), this implies that the appropriately standardized moment matrix related to the LS estimator  $\tilde{\theta}_*^*$  is asymptotically block diagonal. For the upper left block we have

$$\sum_{t=1}^p (\tilde{\beta}'_{\perp}^* \tilde{\beta}_{\perp}^*)^{-1} \tilde{\beta}'_{\perp}^* \tilde{H}_{0t}^* \tilde{\Omega}^{*-1} \tilde{H}_{0t}^* \tilde{\beta}_{\perp}^* (\tilde{\beta}'_{\perp}^* \tilde{\beta}_{\perp}^*)^{-1} = \sum_{t=1}^p (\tilde{\beta}'_{\perp} \tilde{\beta}_{\perp})^{-1} \tilde{\beta}'_{\perp} \tilde{H}'_{0t} \tilde{\Omega}^{-1} \tilde{H}_{0t} \tilde{\beta}_{\perp} (\tilde{\beta}'_{\perp} \tilde{\beta}_{\perp})^{-1} + o_p^*(1)$$

owing to Lemma A1. Following the argumentation in Lütkepohl & Saikkonen (1997, Proof of Lemma 2), we see that this block is positive definite given the definition of  $\tilde{H}_{0t}$  and the requirement within bootstrap Algorithm 1 that  $\tilde{\alpha}'_{\perp} \tilde{\Gamma} \tilde{\beta}_{\perp}$  is nonsingular. The lower right block in the standardized moment matrix related to the LS estimator  $\tilde{\theta}_*^*$  is asymptotically nonsingular as will become evident from the explanations below in connection with Saikkonen & Lütkepohl (1997, Proof of Theorem 1). Furthermore,  $\sum_{t=1}^p (\tilde{\beta}'_{\perp}^* \tilde{\beta}_{\perp}^*)^{-1} \tilde{\beta}'_{\perp}^* \tilde{H}'_{0t} \tilde{\Omega}^{*-1} n_t^*$  is bounded in  $P_{r_0, T}^*$  probability. This follows from the definition of  $n_t^*$  above and the definition of  $\varepsilon_1^*, \dots, \varepsilon_p^*$  in



(4.12). Finally, the standardized sample moments of the other regressors and the error terms are also bounded in  $P_{r_0, T}^*$  probability. Again, this will be seen from the descriptions below and Saikkonen & Lütkepohl (1997, Proof of Theorem 1). Hence, it follows that

$$\tilde{\theta}_*^* = \tilde{\beta}_\perp^{*'} \tilde{\mu}_0^* = \tilde{\beta}_\perp^{*'} \mu_0^* + O_p^*(1) = \theta_*^* + O_p^*(1),$$

which implies (A.47).

Following the above arguments it is straightforward to see that  $\tilde{\theta}_*^*$  has no effect on the asymptotic properties of  $\tilde{\theta}^*$ ,  $\tilde{\tau}^*$ , and  $\tilde{\tau}_*^*$  conditional on the observations. Therefore, we can focus on the estimation of  $\underline{\theta}^*$ ,  $\underline{\tau}^*$ ,  $\underline{\tau}_*^*$  and ignore the first  $p$  (bootstrap) observations in (A.50). Then, the proof of (A.46), (A.48), and (A.49) follows exactly the structure of the proof of Theorem 1 in Saikkonen & Lütkepohl (2000b) and Saikkonen & Lütkepohl (1997), which deal with the corresponding properties of the estimators of the deterministic parameters in the asymptotic test set-up. To be precise, by conditioning on the observation we are able to replicate the proof of Saikkonen & Lütkepohl (2000b) for the bootstrap quantities since they take the role of the observations. Thus, we have to replace a number of quantities with their bootstrap counterparts in the proof by Saikkonen & Lütkepohl (2000b) and consider now convergence in  $P_{r_0, T}^*$  probability. This is analogous to the proof of Lemma A4. The following rules apply. The bootstrap series  $y_t^*$ ,  $x_t^*$  and the adjusted series  $\tilde{x}_t^*$  are used instead of  $y_t$ ,  $x_t$ , and  $\tilde{x}_t$ . Moreover, the true parameters are replaced by their estimators regarding the observations and the latter ones are interchanged with the estimators based on the pseudo data. Finally, the considered error terms have to be redefined accordingly in terms of the quantities which refer to the bootstrap framework, i.e.  $\tilde{\eta}_t^*$  defined above is used instead of  $\tilde{\eta}_t$ , which occur in (2.7). In addition, we conclude from Lemma 2 that  $\tilde{\beta}' x_t^*$  and  $\Delta x_t^*$  are zero mean stationary processes with respect to the measure  $P_{r_0, T}^*$  and that the random process  $S_T^*(s) = 1/\sqrt{T} \sum_{i=k+1}^{[sT]} \varepsilon_i^*$ ,  $0 \leq s \leq 1$  converges in  $P_{r_0, T}^*$  distributions toward a  $n$ -dimensional Brownian motion with covariance matrix  $\tilde{\Omega}$ .

Adjusting the proof of Theorem 1 in Saikkonen & Lütkepohl (2000b) and Saikkonen & Lütkepohl (1997) in the aforementioned way gives first

$$\begin{aligned} \tilde{\theta}^* &= \underline{\theta}^* + O_p^*(T^{-1/2}) \\ \tilde{\tau}^* &= \underline{\tau}^* + O_p^*(T^{-3/2}) \\ (T-p)^{-1/2}(\tilde{\tau}_*^* - \underline{\tau}_*^*) &= \tilde{\beta}_\perp^{*'} \tilde{C}^*(T-p)^{-1/2} \sum_{t=p+1}^T n_t^* + o_p^*(1) \\ &= \tilde{\beta}_\perp^{*'} \tilde{C}(T-p)^{-1/2} \sum_{t=p+1}^T \varepsilon_t^* + o_p^*(1) \xrightarrow{d^*} N(0, \tilde{\beta}_\perp^{*'} \tilde{C} \tilde{\Omega} \tilde{C}' \tilde{\beta}_\perp). \end{aligned} \tag{A.51}$$

The latter result is obtained by Lemma 2, Lemma A.1, and the definition of  $n_t^*$ . Note again that  $\tilde{\beta}'_{\perp} x_t^*$  does not have a zero mean asymptotically because  $x_p^*$  is not a consistent estimator of  $x_p$  in the direction of  $\tilde{\beta}_{\perp}$ . However, the estimation error is bounded. Inspecting the derivations in Saikkonen & Lütkepohl (2000b) and Saikkonen & Lütkepohl (1997), it is seen that this is sufficient to assure that the required standardized sample moments are all bounded in  $P_{r_0, T}^*$  probability. Accordingly, we obtain (A.51). Then, (A.46), (A.48), and (A.49) follow from (A.51) by applying Lemma A.1 as mentioned above. This completes the proof of Lemma A8.  $\square$

Given Lemma A8 we can base the  $GLS_1^*$  test on the VECM (4.13) for the adjusted bootstrap series  $\tilde{x}_t^* = y_t^* - \tilde{\mu}_0^* - \tilde{\mu}_1^* t$  with  $e_t^* = \varepsilon_t^* - \tilde{\alpha} \tilde{\beta}' (\tilde{x}_{t-1}^* - x_{t-1}^*) (t-1) + \Delta \tilde{x}_t^* - \Delta x_t^* - \sum_{j=1}^{p-1} \tilde{\Gamma}_j (\Delta \tilde{x}_{t-j}^* - \Delta x_{t-j}^*) = \varepsilon_t^* + \tilde{\alpha} \tilde{\beta}' (\tilde{\mu}_0^* - \mu_0^*) + \tilde{\alpha} \tilde{\beta}' (\tilde{\mu}_1^* - \mu_1^*) (t-1) - \tilde{\Gamma} (\tilde{\mu}_1^* - \mu_1^*)$ . Thus, we need analogous versions of Lemmas A5-A7 in relation to (4.13). To this end, we define the following quantities. Let  $\tilde{Z}_{0t}^{x*} = \Delta \tilde{x}_t^*$ ,  $\tilde{Z}_{1t}^{x*} = \tilde{x}_{t-1}^*$ , and  $\tilde{Z}_{2t}^{x*} = (\Delta \tilde{x}_{t-1}^{*'}, \dots, \Delta \tilde{x}_{t-p+1}^{*'})'$ . Then, regress  $\tilde{Z}_{0t}^{x*}$  and  $\tilde{Z}_{1t}^{x*}$  on  $\tilde{Z}_{2t}^{x*}$  to obtain the residuals  $\tilde{R}_{0t}^{x*}$  and  $\tilde{R}_{1t}^{x*}$ , respectively. Using these residuals we have

$$\tilde{S}_{ij}^{x*} = (T-p)^{-1} \sum_{t=p+1}^T \tilde{R}_{it}^{x*} \tilde{R}_{jt}^{x*} \quad i, j = 0, 1.$$

Similarly, we obtain the moment matrices  $S_{ij}^{x*}$  ( $i, j = 0, 1$ ) from  $Z_{0t}^{x*}$ ,  $Z_{1t}^{x*}$ ,  $Z_{2t}^{x*}$ ,  $R_{0t}^{x*}$ , and  $R_{1t}^{x*}$ , which are defined in terms of  $x_t^* = y_t^* - \mu_0^* - \mu_1^* t$ . Furthermore, let  $S_{ij}^x$  ( $i, j = 0, 1$ ) be the moment matrices based on the original observations adjusted by the level term and trend component, i.e based on  $x_t = y_t - \mu_0 - \mu_1 t$ .

**Lemma A9.** Under Assumptions 1 and 3, we have

$$\begin{aligned} \tilde{S}_{00}^{x*} &\xrightarrow{p} \Sigma_{00}^x \\ \tilde{S}_{01}^{x*} \tilde{\beta}^+ &\xrightarrow{p} \Sigma_{0\beta}^x \\ \tilde{\beta}' \tilde{S}_{11}^{x*} \tilde{\beta} &\xrightarrow{p} \Sigma_{\beta\beta}^x, \end{aligned}$$

where  $\Sigma_{00}^x$ ,  $\Sigma_{0\beta}^x$ , and  $\Sigma_{\beta\beta}^x$  are the respective probability limits of  $S_{00}^{x*}$ ,  $S_{01}^{x*} \beta$ , and  $\beta' S_{11}^{x*} \beta$  with respect to the original observations as  $T \rightarrow \infty$ .  $\square$

**Proof** The proof is analog to the one of Lemma A5. First, one has to show that  $\tilde{S}_{00}^{x*} \xrightarrow{p^*} S_{00}^{x*}$ ,  $\tilde{S}_{01}^{x*} \tilde{\beta} \xrightarrow{p^*} S_{01}^{x*} \tilde{\beta}$ , and  $\tilde{\beta}' \tilde{S}_{11}^{x*} \tilde{\beta} \xrightarrow{p^*} \tilde{\beta}' S_{11}^{x*} \tilde{\beta}$ . These results are again proven by writing  $\tilde{S}_{00}^{x*}$ ,  $\tilde{S}_{01}^{x*} \tilde{\beta}$ , and  $\tilde{\beta}' \tilde{S}_{11}^{x*} \tilde{\beta}$  in terms of  $\tilde{Z}_{0t}^{x*}$ ,  $\tilde{Z}_{1t}^{x*}$ , and  $\tilde{Z}_{2t}^{x*}$ . Then, use Lemma 2, Lemma A8 and Johansen (1995, Theorem B.13). Note in this respect that  $Z_{0t}^{x*}$ ,  $Z_{2t}^{x*}$ , and  $\tilde{\beta}' Z_{1t}^{x*}$  are zero-mean stationary

processes with respect to the measure  $P_{r_0, T}^*$  given Lemma 2 and the definition of  $\sqrt{T}R_{t, T}^{0*}$  in (A.9). Finally, it follows from Swensen (2005a, Lemma 2) in connection with Lemma 2 that  $S_{00}^{x*}$ ,  $S_{01}^{x*}\tilde{\beta}$ , and  $\tilde{\beta}'S_{11}^{x*}\tilde{\beta}$  converge to  $\Sigma_{00}^x$ ,  $\Sigma_{0\beta}^x$ , and  $\Sigma_{\beta\beta}^x$  in  $P_{r_0, T}$  probability.  $\square$

**Lemma A10.** Under Assumptions 1 and 3, we have

$$(T-p)^{-1}\tilde{\beta}'_{\perp}\tilde{S}_{11}^{x*}\tilde{\beta}_{\perp} \xrightarrow{p^*} (T-p)^{-2}\tilde{\beta}'_{\perp} \sum_{t=p+1}^T w_{t-1}^* w_{t-1}' \tilde{\beta}_{\perp} \quad (\text{A.52})$$

$$\tilde{\beta}'_{\perp}\tilde{S}_{10}^{x*}\tilde{\alpha}_{\perp} \xrightarrow{p^*} (T-p)^{-1}\tilde{\beta}'_{\perp} \sum_{t=p+1}^T w_{t-1}^* \bar{\varepsilon}_t' \tilde{\alpha}_{\perp} \quad (\text{A.53})$$

where  $w_{t-1}^* = x_{t-1}^* - (\tilde{\mu}_1^* - \mu_1^*)(t-1)$  and  $\bar{\varepsilon}_t' = \varepsilon_t^* - \tilde{\Gamma}'\tilde{\beta}_{\perp}\tilde{\beta}'_{\perp}(\tilde{\mu}_1^* - \mu_1^*)$ .  $\square$

**Proof** Note that  $\tilde{\beta}'_{\perp}x_{t-1}^*$  is a zero-mean process plus a bounded term with respect to the measure  $P_{r_0, T}^*$  in line with Lemma 2 because  $\tilde{\beta}'_{\perp}x_p^* = O^*(1)$ . Then, using Lemma 2, Lemma A8, and Theorem B.13 of Johansen (1995) we get  $(T-p)^{-1}\tilde{\beta}'_{\perp}\tilde{S}_{11}^{x*}\tilde{\beta}_{\perp} \xrightarrow{p^*} (T-p)^{-2} \sum_{t=p+1}^T \tilde{\beta}'_{\perp}\tilde{Z}_{1t}^{x*}\tilde{Z}_{1t}^{x*'}\tilde{\beta}_{\perp}$  corresponding to (A.41). Finally, (A.52) follows by applying (A.47) and (A.49) to  $(T-p)^{-2} \sum_{t=p+1}^T \tilde{\beta}'_{\perp}\tilde{Z}_{1t}^{x*}\tilde{Z}_{1t}^{x*'}\tilde{\beta}_{\perp}$  since  $\tilde{\beta}'_{\perp}\tilde{Z}_{1t}^{x*} = \tilde{\beta}'_{\perp}\tilde{x}_{t-1}^* = \tilde{\beta}'_{\perp}x_{t-1}^* - \tilde{\beta}'_{\perp}(\tilde{\mu}_0^* - \mu_0^*) - \tilde{\beta}'_{\perp}(\tilde{\mu}_1^* - \mu_1^*)(t-1)$ .

Analogously to (A.42), we first have  $\tilde{\beta}'_{\perp}\tilde{S}_{10}^{x*}\tilde{\alpha}_{\perp} \xrightarrow{p^*} (T-p)^{-1}\tilde{\beta}'_{\perp} \sum_{t=p+1}^T \tilde{Z}_{1t}^{x*}\bar{\varepsilon}_t'\tilde{\alpha}_{\perp}$  by taking into account the obvious modifications due to the fact that now the VECM (4.13) is considered instead of (A.40). Furthermore, we make use of  $(T-p)^{-1} \sum_{t=p+1}^T \tilde{\beta}'_{\perp}(\tilde{\mu}_0^* - \mu_0^*)\bar{\varepsilon}_t'\tilde{\alpha}_{\perp} = (T-p)^{-1}\tilde{\beta}'_{\perp}(\tilde{\mu}_0^* - \mu_0^*) \sum_{t=p+1}^T \left( \varepsilon_t^* \tilde{\alpha}_{\perp} + \tilde{\beta}'_{\perp}(\tilde{\mu}_0^* - \mu_0^*)(\tilde{\mu}_1^* - \mu_1^*)\tilde{\beta}'_{\perp}\tilde{\beta}_{\perp}\tilde{\Gamma}'\tilde{\alpha}_{\perp} \right) = o_p^*(1)$  obtained from Lemmas 2, A1, and A8. Connecting this result with the conditional limit expression for  $\tilde{\beta}'_{\perp}\tilde{S}_{10}^{x*}\tilde{\alpha}_{\perp}$  and the definition of  $\tilde{Z}_{1t}^{x*}$  gives (A.53). This completes the proof of Lemma A10.  $\square$

Finally, we turn to the analog of Lemma A7 for the  $GLS_1^{x*}$  test.

**Lemma A11.** Under Assumptions 1 and 3,

$$(T-p)^{-1}\tilde{\beta}'_{\perp}\tilde{S}_{11}^{x*}\tilde{\beta}_{\perp} \rightarrow \int_0^1 G_*(s)G_*(s)'ds$$

$$\tilde{\beta}'_{\perp}\tilde{S}_{10}^{x*}\tilde{\alpha}_{\perp} \rightarrow \int_0^1 G_*(s)dG_*(s)'\alpha_{\perp},$$

where  $G_*(s) = \beta'_{\perp}C\Omega^{1/2}B_*(s)$ ,  $dG_*(s) = \Omega^{1/2}dB_*(s)$ , and convergence is weakly in  $P_{r_0, T}$  probability.  $\square$

**Proof** Lemma A12 can be proven following the argumentation related to the proof of Lemma A7. Now, we consider the process  $\{(T - p)^{-1/2}x_{[sT]}^*, 0 \leq s \leq 1\}$  and refer to Lemma 2. From the proof of the latter we see that also the remainder term  $R_{t,T}^{0*}$  vanishes asymptotically since the estimation error related to  $\tilde{\beta}'_{\perp}x_p^*$  is bounded in  $P_{r_0,T}$  probability. Moreover, the functionals on the r.h.s. of (A.52) and (A.53) correspond exactly to the ones in equations (A.10)/(A.11) and (A.14)/(A.15) of Lütkepohl & Saikkonen (2000) since we do not have to consider a constant term in the VECM (4.13). Thus, Lemma A11 will hold.  $\square$

Finally, Proposition 2 follows given Lemmas A9 and A11. As for showing Proposition 1, we refer to the proof approach in Johansen (1995, Ch. 11) performing the necessary modifications due to the bootstrap framework (compare Swensen 2005a) and due to specific characteristics of the GLS testing framework (see Saikkonen & Lütkepohl 2000b and the derivations above).  $\square$

In line with Swensen (2005a) the limiting results of Propositions 1 and 2 also hold if the bootstrap Algorithm 2 is used. As mentioned in Section 4, the residuals and estimators for the short-run parameters are taken from an estimated unrestricted VAR model in first differences. Accordingly, the conditions (4.2), Assumptions 2 and 3, and Lemmas 1 and 2 are modified. A careful inspection of Lemmas A1-A11 and their proofs shows that the derivations are also valid in terms of  $SLT_2^*(r_0)$  and  $GLS_2^*(r_0)$  with obvious adjustments. Hence, it can be shown that  $SLT_2^*(r_0)$  and  $GLS_2^*(r_0)$  converge to the same limiting distributions in  $P_{r_0,T}$  probability as  $SLT_1^*(r_0)$  and  $GLS_1^*(r_0)$  respectively.

## Appendix B: Response Surface for the *SLT*-Test

The response surface for the *SLT* test by Saikkonen et al. (2006) is computed exactly in the same way as the one for the *GLS* test, which is presented in Trenkler (2004). The idea of the response surface is to approximate the limiting distribution given in Theorem 1 by a Gamma distribution with two parameters. These parameters can be related to the mean and variance of the distribution of interest. Therefore, the aim of the response surface simulations is to compute the mean and variance for discrete realizations of the limit distribution of *SLT*. To this end, we consider different sample sizes  $T$  and dimensions  $d = n - r_0$ . These simulated realizations are based on random walks, which are the discrete counterparts of Brownian motions. The dependency of the mean and variance on the sample size  $T$  is used to estimate the asymptotic moments for  $T \rightarrow \infty$ . Thus, we only need to present response surface coefficients with re-

**Table 6.** Response Surface for Mean and Variance of Asymptotic Distribution of the Trace Test Statistics  $SLT(r_0)$

	Mean	Variance
$d^2$	2.0046	3.0125
$d$	1.7392	1.9664
$\sqrt{d}$	1.0027	—
1 (constant)	-0.5442	1.4214

spect to the dimension  $d$  in order to estimate the asymptotic mean and variance of the relevant distributions. These coefficients are given in Table 6 and refer to polynomials of  $d$ . The calculated moments are then used to fit an approximating Gamma distribution, which can be used to compute  $p$ -values or arbitrary quantiles. Further details can be found in Trenkler (2004).

## References

- Bartlett, M. S. (1937). Properties of sufficiency and statistical tests, *Proceedings of the Royal Society of London, Series A* **160**: 268–282.
- Basawa, I. V., Mallik, A. K., McCormic, W. P., Reeves, J. H. & Taylor, R. L. (1991). Bootstrap test of significance and sequential bootstrap estimation for unstable first order autoregressive processes, *Communications in Statistics A* **20**: 1015–1026.
- Brüggemann, R. (2004). *Model Reduction Methods for Vector Autoregressive Processes*, Springer-Verlag, Berlin.
- Chang, Y. & Park, J. Y. (2003). A sieve bootstrap for the test of a unit root, *Journal of Time Series Analysis* **24**: 379–400.
- Davidson, R. & MacKinnon, J. G. (2000). Bootstrap tests: How many bootstraps?, *Econometric Reviews* **19**: 55–69.
- Davidson, R. & MacKinnon, J. G. (2006). The power of bootstrap and asymptotic tests, *Journal of Econometrics*, **forthcoming**.
- Doornik, J. A. (1998). Approximations to the asymptotic distributions of cointegration tests, *Journal of Economic Surveys* **12**: 573–593.

- Hamilton, J. D. (1994). *Time Series Analysis*, Princeton University Press, Princeton.
- Hansen, P. R. & Johansen, S. (1998). *Workbook on Cointegration*, Oxford University Press, Oxford.
- Härdle, W., Horowitz, J. & Kreiß, J.-P. (2003). Bootstrap methods for time series, *International Statistical Review* **71**: 435–459.
- Harris, R. I. D. & Judge, G. G. (1998). Small sample testing for cointegration using the bootstrap approach, *Economics Letters* **58**: 31–37.
- Horowitz, J. L. (2001). The bootstrap, in J. J. Heckmann & E. E. Leamer (eds), *Handbook of Econometrics*, Vol. 5, North-Holland, Amsterdam.
- Horowitz, J. L. (2003). The bootstrap in econometrics, *Statistical Science* **18**: 211–218.
- Hubrich, K., Lütkepohl, H. & Saikkonen, P. (2001). A review of systems cointegration tests, *Econometric Reviews* **20**: 247–318.
- Johansen, S. (1988). Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control* **12**: 231–254.
- Johansen, S. (1995). *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*, Oxford University Press, Oxford.
- Johansen, S. (2002). A small sample correction for the test of cointegrating rank in the vector autoregressive model, *Econometrica* **70**: 1929–1961.
- King, R. G., Plosser, C. I., Stock, J. H. & Watson, M. W. (1991). Stochastic trends and economic fluctuations, *American Economic Review* **81**: 819–840.
- Lütkepohl, H. & Krätzig, M. (2004). *Applied Time Series Econometrics*, Cambridge University Press, Cambridge.
- Lütkepohl, H. & Saikkonen, P. (1997). Testing for the cointegration rank of a VAR process with a time trend, *Discussion Paper 79*, SFB 373, Humboldt-Universität zu Berlin.
- Lütkepohl, H. & Saikkonen, P. (2000). Testing for the cointegration rank of a VAR process with a time trend, *Journal of Econometrics* **95**: 177–198.

- MacKinnon, J. G. (2002). Bootstrap inference in econometrics, *Canadian Journal of Economics* **35**: 615–645.
- Mantalos, P. & Shukur, G. (2001). Bootstrapped johansen tests for cointegration relationships: A graphical analysis, *Journal of Statistical Computations and Simulation* **68**: 351–371.
- Paparoditis, E. & Politis, D. N. (2003). Residual-based block bootstrap for unit root testing, *Econometrica* **71**: 813–855.
- Park, J. Y. (2002). An invariance principle for sieve bootstrap in time series, *Econometric Theory* **18**: 469–490.
- Park, J. Y. (2003). Bootstrap unit root tests, *Econometrica* **71**: 1845–1895.
- Paruolo, P. (2001). The power of lambda max, *Oxford Bulletin of Economics and Statistics* **63**: 395–403.
- Pollard, D. (1984). *Convergence of Stochastic Processes*, Springer, New York.
- Rahbek, A. & Mosconi, R. (1999). Cointegration rank inference with stationary regressors in VAR models, *Econometrics Journal* **2**: 76–91.
- Saikkonen, P. & Lütkepohl, H. (1997). Trend adjustment prior to testing for the cointegration rank of a VAR process, *Discussion paper 84*, SFB 373, Humboldt-Universität zu Berlin.
- Saikkonen, P. & Lütkepohl, H. (2000a). Testing for the cointegration rank of a VAR process with structural shifts, *Journal of Business & Economic Statistics* **18**: 451–464.
- Saikkonen, P. & Lütkepohl, H. (2000b). Trend adjustment prior to testing for the cointegration rank of a VAR process, *Journal of Time Series Analysis* **21**: 435–456.
- Saikkonen, P., Lütkepohl, H. & Trenkler, C. (2006). Break date estimation for VAR processes with level shift with an application to cointegration testing, *Econometric Theory*, **22**: 15–68.
- Swensen, A. R. (2005a). Bootstrap algorithms for testing and determining the cointegration rank in VAR models, *mimeo, revised version of the Statistical Research Report, No. 5, 2004, with the same titel*, Department of Mathematics, University of Oslo.

- Swensen, A. R. (2005b). Finite sample properties of a bootstrap procedure for estimation of rank in reduced rank VAR-models, *mimeo*, Department of Mathematics, University of Oslo.
- Toda, H. Y. (1994). Finite sample properties of likelihood ratio tests for cointegrating ranks when linear trends are present, *The Review of Econometrics and Statistics* **76**: 66–79.
- Toda, H. Y. (1995). Finite sample performance of likelihood ratio tests for cointegrating ranks in vector autoregressions, *Econometric Theory* **11**: 1015–1032.
- Trenkler, C. (2002). *Testing for the Cointegrating Rank in the Presence of Level Shifts*, Shaker Verlag, Aachen.
- Trenkler, C. (2004). Determining  $p$ -values for systems cointegration tests with a prior adjustment for deterministic terms, *CASE Discussion Paper 37*, Humboldt-Universität zu Berlin.
- van Giersbergen, N. P. A. (1996). Bootstrapping the trace statistic in VAR models: Monte Carlo results and applications, *Oxford Bulletin of Economics and Statistics* **58**: 391–408.



## SFB 649 Discussion Paper Series 2006

For a complete list of Discussion Papers published by the SFB 649, please visit <http://sfb649.wiwi.hu-berlin.de>.

- 001 "Calibration Risk for Exotic Options" by Kai Detlefsen and Wolfgang K. Härdle, January 2006.
- 002 "Calibration Design of Implied Volatility Surfaces" by Kai Detlefsen and Wolfgang K. Härdle, January 2006.
- 003 "On the Appropriateness of Inappropriate VaR Models" by Wolfgang Härdle, Zdeněk Hlávka and Gerhard Stahl, January 2006.
- 004 "Regional Labor Markets, Network Externalities and Migration: The Case of German Reunification" by Harald Uhlig, January/February 2006.
- 005 "British Interest Rate Convergence between the US and Europe: A Recursive Cointegration Analysis" by Enzo Weber, January 2006.
- 006 "A Combined Approach for Segment-Specific Analysis of Market Basket Data" by Yasemin Boztuğ and Thomas Reutterer, January 2006.
- 007 "Robust utility maximization in a stochastic factor model" by Daniel Hernández-Hernández and Alexander Schied, January 2006.
- 008 "Economic Growth of Agglomerations and Geographic Concentration of Industries - Evidence for Germany" by Kurt Geppert, Martin Gornig and Axel Werwatz, January 2006.
- 009 "Institutions, Bargaining Power and Labor Shares" by Benjamin Bental and Dominique Demougin, January 2006.
- 010 "Common Functional Principal Components" by Michal Benko, Wolfgang Härdle and Alois Kneip, January 2006.
- 011 "VAR Modeling for Dynamic Semiparametric Factors of Volatility Strings" by Ralf Brüggemann, Wolfgang Härdle, Julius Mungo and Carsten Trenkler, February 2006.
- 012 "Bootstrapping Systems Cointegration Tests with a Prior Adjustment for Deterministic Terms" by Carsten Trenkler, February 2006.

**SFB 649, Spandauer Straße 1, D-10178 Berlin**  
**<http://sfb649.wiwi.hu-berlin.de>**

This research was supported by the Deutsche  
Forschungsgemeinschaft through the SFB 649 "Economic Risk".

