BLOCK BOOTSTRAP THEORY FOR MULTIVARIATE INTEGRATED AND COINTEGRATED PROCESSES

CARSTEN JENTSCH, EFSTATHIOS PAPARODITIS, AND DIMITRIS N. POLITIS

Abstract. We develop some asymptotic theory for applications of block bootstrap resampling schemes to multivariate integrated and cointegrated time series. It is proved that a multivariate, continuous-path block bootstrap scheme applied to a full rank integrated process, succeeds in estimating consistently the distribution of the least squares estimators in both, the regression and the spurious regression case. Furthermore, it is shown that the same block resampling scheme does not succeed in estimating the distribution of the parameter estimators in the case of cointegrated time series. For this situation, a modified block resampling scheme, the so-called residual based block bootstrap, is investigated and its validity for approximating the distribution of the regression parameters is established.

JEL classification: C15; C32

1. Introduction

The block bootstrap methodology, Künsch (1989), Liu and Singh (1992), is a general resampling scheme applicable to time series data obeying a weak dependence structure; see Lahiri (2003) for an overview. It has been successfully applied to a wide range of inference problems in statistics and econometrics including the important area of estimation and testing when the underlying process is a multivariate integrated and/or cointegrated stochastic process studied extensively in econometrics literature. In the univariate case, applications of block resampling schemes to integrated processes have been considered by Paparoditis and Politis (2003), Parker et al. (2006) and Phillips (2010). For the multivariate case, Li and Maddala (1997) and Badillo et al. (2010) discussed applications of block bootstrap methods for estimation and testing purposes. However, despite these applied papers, in the multivariate context few is known about the theoretical properties of the block bootstrap used for statistical inference.

The aim of this paper is to fill this gap and to provide some asymptotic theory that justifies applications of block bootstrap methods to multivariate integrated and/or cointegrated processes. We first consider a so-called, full rank, multivariate integrated process and investigate properties of a block bootstrap resampling scheme that is based on block bootstrapping of the centered differences of the observed multivariate time series. This resampling scheme is a multivariate version of the so-called Continuous-Path Block bootstrap (CBB), proposed in the univariate context by Paparoditis and Politis (2001). Based on CBB-generated multivariate bootstrap observations, we first establish a functional limit theorem for the bootstrap partial sum process. Using this basic result, we then prove asymptotic validity of the CBB method applied to the matrix of the least squares estimators in the full rank integrated case. We also show that by means of the same bootstrap procedure, the distribution of the least squares estimators for the spurious regression as coined by Phillips (1986) can be also successfully approximated. This complements the results in Phillips (2001), who investigated the applicability of several bootstrap approaches in this framework. Further, we show that the same bootstrap method based on

Date: July 9, 2013.
Key words and phrases. Block bootstrap, bootstrap consistency, spurious regression, functional limit theorem, continuous-path bootstrap, residual block bootstrap.

resampling blocks of the (centered) differences does not work, however, in the case of multivariate cointegrated processes. This is due to the fact that in this case the block resampling scheme based on differences does not mimic correctly the cointegration structure of the underlying process. For this kind of processes a Residual-based Block Bootstrap (RBB) is more appropriate. This procedure is based on block bootstrapping of the (centered) residuals of a regression fit obtained using the cointegration relations. We prove that the RBB is asymptotically valid in approximating the distribution of the least squares estimators of the cointegration matrix.

The paper is organized as follows. Section 2 describes the CBB procedure and establishes a basic functional limit theorem for the block bootstrap generated partial sum process. Applications of the CBB bootstrap scheme to multivariate integrated processes are then considered and asymptotic validity for the full rank integrated case and in the spurious regression case are established. Section 3 deals with cointegrated multivariate processes. It is first demonstrated that the CBB applied to the least squares estimator fails. A residual based block bootstrap (RBB) resampling scheme is then discussed. It is shown, that for cointegrated processes the RBB resampling scheme leads to an asymptotically valid approximation of the distribution of the least squares cointegration matrix. Section 4 summarizes our findings while all proofs are deferred to Section 5.

2. Block Bootstrap for Integrated Processes

2.1. Preliminaries.

Suppose we have m-variate time series data $X_1, \ldots, X_n$ plus one additional pre-sample value X_0 at hand stemming from a stochastic process $\{X_t, t \in \mathbb{N}_0\}$ where the X_t's are $\mathbb{R}^m$-valued random variables. We assume that $\{X_t, t \in \mathbb{N}_0\}$ is a so-called, m-dimensional, full-rank random walk, i.e., X_t follows the model

$$X_t = X_{t-1} + U_t, \quad t \in \mathbb{N}, \tag{2.1}$$

where $X_0 = O_{as}(1)$ is random following a certain fixed distribution and $\{U_t, t \in \mathbb{N}\}$ is a stationary process satisfying either Assumption 2.1 or Assumption 2.2 below. By these assumptions, $\{U_t, t \in \mathbb{N}\}$ is allowed to be either a linear process or a process fulfilling some general strong mixing conditions; see also Phillips (1988) and Paparoditis and Politis (2003).

Assumption 2.1.
The process $\{U_t, t \in \mathbb{N}\}$ is a sequence of m-variate random variables satisfying

$$U_t = \sum_{\nu=0}^{\infty} \psi_\nu \epsilon_{t-\nu},$$

where $\psi_\nu = (\psi_{\nu,ij})_{i,j=1,\ldots,m}$, $\sum_{\nu=1}^{\infty} \nu|\psi_{\nu,ij}| < \infty$ for each $i, j = 1, \ldots, m$, $(\epsilon_t, t \in \mathbb{Z})$ is an m-variate sequence of i.i.d. random variables with $E(\epsilon_t) = 0$, $E(\epsilon_{it}^4) < \infty$ for $i = 1, 2, \ldots, m$ and $E(\epsilon_t \epsilon_t') = \Sigma > 0$. Furthermore,

$$\Omega = 2\pi f(0) = \left(\sum_{\nu=0}^{\infty} \psi_\nu\right) \Sigma \left(\sum_{\nu=0}^{\infty} \psi_\nu\right)' > 0,$$

where $f(\omega)$ denotes the spectral density matrix of $\{U_t, t \in \mathbb{N}\}$.

Assumption 2.2.
The process $\{U_t, t \in \mathbb{N}\}$ is a sequence of strictly stationary, strong mixing m-variate random variables satisfying $E(U_t) = 0$, $E|U_{i1}|^{\beta+\epsilon} < \infty$, $i = 1, \ldots, m$ for some $\beta > 2$, $\epsilon > 0$ and $\sum_{k=1}^{\infty} \alpha(k)^{1-2/\beta} < \infty$. The α-mixing coefficient $\alpha(k)$ is defined by

$$\alpha(k) = \sup_{A \in \mathcal{F}_{t+k}^{\infty}, B \in \mathcal{F}_{-\infty}^{t}} |P(A \cap B) - P(A)P(B)|,$$

where $\mathcal{F}_{t+k}^{\infty} = \sigma(X_{t+k}, X_{t+k+1}, \ldots)$, $\mathcal{F}_{-\infty}^{t} = \sigma(X_t, X_{t-1}, \ldots)$ *and* $\sigma(Y)$ *denotes the* σ*-algebra generated by the random variable* Y*. Furthermore, it holds true that* $E(n^{-1}S_nS_n') \to \Omega > 0$ *as* $n \to \infty$*, where* $S_t = \sum_{j=1}^{t} U_j$*.*

Some remarks on the two sets of assumptions are in order. Assumption 2.1 imposes linearity of the process $\{U_t\}$ with sufficiently fast decaying coefficient matrix entries. In contrast, $\{U_t\}$ is assumed to fulfill a suitable strong mixing condition in Assumption 2.2 that is just strong enough to prove limit theorems. In both conditions, the long-run variance Ω of $\{U_t\}$ is assumed to be non-singular to ensure the existence of its inverse.

The following notation borrowed e.g. from Phillips (1988) is used in the sequel. We set

$$\Omega = \Gamma(0) + \sum_{k=1}^{\infty} \Gamma(h) + \sum_{k=1}^{\infty} \Gamma(-h) = \Omega_0 + \Omega_1 + \Omega_1', \tag{2.2}$$

where $\Gamma(h) = Cov(U_{t+h}, U_t) = E(U_{t+h}U_t')$ denotes the covariance matrix at lag $h \in \mathbb{Z}$ of the innovation process $\{U_t\}$.

2.2. **The Continuous-path Block Bootstrap (CBB).**

We first describe a multivariate version of the so-called Continuous-path Block Bootstrap (CBB) as proposed by Paparoditis and Politis (2003). This general block bootstrap resampling scheme will be applied in the sequel in order to approximate the distribution of statistics calculated using time series data $X_0, X_1, \ldots, X_n$ stemming from the process (2.1), where $\{U_t\}$ satisfies Assumption 2.1 or Assumption 2.2. More precisely, the multivariate CBB consists of the following four steps.

Step 1. Compute the differences $U_t = \Delta X_t = X_t - X_{t-1}$, $t = 1, \ldots, n$.

Step 2. Choose a block length $b < n$ and let k be the smallest number of blocks needed to get a bootstrap sample of length $l = kb$, such that $l \geq n$. Let $i_0, \ldots, i_{k-1}$ be i.i.d. random variables uniformly distributed on the set $\{0, 1, 2, \ldots, n-b\}$.

Step 3. Let $U_1^*, \ldots, U_l^*$ be a moving block bootstrap sample, where for $j = 1, 2, \ldots, b$ and $m = 0, 1, 2, \ldots, k-1$,

$$U_{mb+j}^* = \widehat{U}_{i_m+j} := U_{i_m+j} - E^*\left(U_{i_m+j}\right) = U_{i_m+j} - \frac{1}{n-b+1}\sum_{\tau=0}^{n-b} U_{\tau+j}. \tag{2.3}$$

Step 4. Generate then the bootstrap pseudo-time series $X_1^*, \ldots, X_l^*$, as

$$X_t^* = \begin{cases} X_0, & t = 0 \\ X_{t-1}^* + U_t^*, & t \in \{1, \ldots, l\}. \end{cases} \tag{2.4}$$

Remark 2.1. *Note that the centering of the bootstrap sample in* (2.3) *is tailor-made for the moving block bootstrap and adjusted centering has to be applied for other approaches as e.g. non-overlapping block bootstrap, cyclical block bootstrap or stationary bootstrap.*

2.3. **A Functional Limit Theorem (FLT) for the bootstrap partial sum process.**

We first establish a basic result which is useful to prove consistency of the CBB resampling scheme. In fact, the asymptotic theory for establishing bootstrap consistency for statistics based on time series $X_1, X_2, \ldots, X_n$ relies to a large extent on the asymptotic behavior of the standardized partial sum process $\{S_l^*(\nu), 0 \leq \nu \leq 1\}$, which is defined by

$$\begin{aligned} S_l^*(\nu) &= \frac{1}{\sqrt{l}} (\Omega_l^*)^{-1/2} \sum_{t=1}^{j-1} U_t^*, \quad \text{for} \quad \frac{j-1}{l} \leq \nu < \frac{j}{l}, \quad j = 1, \ldots, l, \\ S_l^*(\nu) &= \frac{1}{\sqrt{l}} (\Omega_l^*)^{-1/2} \sum_{t=1}^{l} U_t^*, \quad \text{for} \quad \nu = 1. \end{aligned} \tag{2.5}$$

By convention, summations over empty sets are zero and $(\Omega_l^*)^{1/2}$ is the symmetric (positive semi-definite) square root of

$$\Omega_l^* = Var^* \left(\frac{1}{\sqrt{l}} \sum_{t=1}^{l} U_t^* \right) = \frac{1}{b} \sum_{s_1, s_2 = 1}^{b} \frac{1}{n-b+1} \sum_{t=0}^{n-b} \widehat{U}_{t+s_1} \widehat{U}'_{t+s_2}, . \tag{2.6}$$

Notice that by Assumption 2.1 or Assumption 2.2, Ω_l^* is positive definite with probability tending to one.

Observe that $S_l^*(\cdot)$ is a random element taking values in the space $D[0,1]^m = D[0,1] \times \cdots \times D[0,1]$, the product metric space of all real valued vector functions on $[0,1]$ that are right continuous at each element of $[0,1]$ and possess finite left limits. Each component space $D[0,1]$ is endowed with the Skorohod metric, denoted by d, which ensures separability of the metric space. For the product space $D[0,1]^m$, the metric $\tilde{d}$ is defined by

$$\tilde{d}(x,y) = \max_{i \in \{1,2,\ldots,m\}} \Big\{ d(x_i, y_i) : \ x_i, y_i \in D[0,1] \Big\}.$$

In the following theorem, we prove that conditional on the sample $X_1, \ldots, X_n$, the standardized CBB bootstrap partial sum process (2.5), converges weakly to an m-dimensional standard Wiener process, where each element of $W(t)$ is a univariate Wiener process and the elements of $W(t)$ are independent from each other. The notation $S_l^* \Rightarrow W$ in probability, means that the distance between the law of $S_l^* = S_l^*(\cdot)$ and the law of W tends to zero in probability for any distance metrizing weak convergence.

Theorem 2.1. *Let $\{X_t, t \in \mathbb{N}_0\}$ be an m-dimensional stochastic process following* (2.1) *and assume that the process $\{U_t, t \in \mathbb{N}\}$ satisfies Assumption 2.1 or Assumption 2.2. If $b \to \infty$ as $n \to \infty$ such that $b/\sqrt{n} \to 0$, then*

$$S_l^* \Rightarrow W$$

in probability, respectively, where $\{W(t) = (W_1(t), \ldots, W_m(t))', t \in [0,1]\}$ is here and throughout this paper an m-dimensional standard Wiener process on $[0,1]$, i.e. each element $W_i(\cdot)$ is a univariate Wiener process and the elements of $W(\cdot)$ are independent.

We next discuss some cases where Theorem 2.1 is employed to establish asymptotic properties of block-bootstrap based, statistical inference procedures.

2.4. **Applications to Regression Estimators.**

2.4.1. *Regressing X_t on X_{t-1}.*
Consider the least squares (LS) estimator $\widehat{A}$ obtained by regressing X_t on X_{t-1}, that is,

$$\widehat{A} = \left(\sum_{t=2}^{n} X_t X'_{t-1} \right) \left(\sum_{t=2}^{n} X_{t-1} X'_{t-1} \right)^{-1}. \tag{2.7}$$

We want to approximate the unknown distribution of $n(\widehat{A} - I_m)$ by the bootstrap distribution of $l(\widehat{A}^* - I_m)$, where

$$\widehat{A}^* = \left(\sum_{t=2}^{l} X_t^* X_{t-1}^{*'} \right) \left(\sum_{t=2}^{l} X_{t-1}^* X_{t-1}^{*'} \right)^{-1}, \tag{2.8}$$

I_m is the $(m \times m)$ unity matrix and $X_1^*, \ldots, X_l^*$ is generated using the CBB scheme described in Section 2.2. As

$$n(\widehat{A} - I_m) \Rightarrow \left\{ \Omega^{1/2} \int_0^1 W(t) dW(t)' \Omega^{1/2} + \Omega_1 \right\} \left\{ \Omega^{1/2} \int_0^1 W(t) W(t)' dt \Omega^{1/2} \right\}^{-1}$$

holds under Assumptions 2.1 [cf. Lütkepohl (2006), Proposition C.18] or 2.2 [cf. Phillips and Durlauf (1986), Theorem 2.1], the following theorem establishes validity of the multivariate CBB procedure for approximating the distribution of the least squares estimator $\widehat{A}$.

Theorem 2.2. *Under the assumptions of Theorem 2.1, conditionally on $X_1, X_2, \ldots, X_n$, it holds true that*

$$l(\widehat{A}^* - I_m) \Rightarrow \left\{ \Omega^{1/2} \int_0^1 W(t) dW(t)' \Omega^{1/2} + \Omega_1 \right\} \left\{ \Omega^{1/2} \int_0^1 W(t) W(t)' dt \Omega^{1/2} \right\}^{-1}$$

in probability, where Ω and Ω_1 are defined in (2.2).

2.4.2. *Spurious Regression of X_{1t} on $X_{2t}, \ldots, X_{mt}$.*
Consider next the so-called spurious regression case, where the first component X_{1t} of X_t is regressed on the remaining set of random variables $X_{2t}, \ldots, X_{mt}$, while the underlying process is the full rank random walk (2.1). Let $\widehat{\alpha}_n$ and $\widehat{\beta}_n = (\widehat{\beta}_2, \ldots, \widehat{\beta}_m)$ be the least squares estimators of α and $\beta = (\beta_2, \ldots, \beta_m)$ when one fits the (spurious regression) model

$$X_{1t} = \alpha + \sum_{i=2}^{m} \beta_i X_{it} + \epsilon_t, \tag{2.9}$$

to the time series at hand. More specificaly,

$$\widehat{\beta}_n = \left\{ \frac{1}{n^2} \sum_{t=1}^{n} (Z_t - \overline{Z})(Z_t - \overline{Z})' \right\}^{-1} \left\{ \frac{1}{n^2} \sum_{t=1}^{n} (Z_t - \overline{Z})(X_{1t} - \overline{X}_1) \right\} \tag{2.10}$$

and $\widehat{\alpha}_n = \overline{X}_1 - \overline{Z}' \widehat{\beta}_n$, where $Z_t = (X_{2t}, X_{3,t}, \ldots, X_{mt})$, $\overline{X}_1 = \frac{1}{n} \sum_{t=1}^{n} X_{1t}$ and $\overline{Z} = \frac{1}{n} \sum_{t=1}^{n} Z_t$. From Phillips (1986), Theorem 2 and under Hamilton (1994), Proposition 18.2, we get under Assumptions 2.2 or 2.1, respectively, that

$$\frac{1}{\sqrt{n}} \widehat{\alpha}_n \Rightarrow b_1 - b_2' A_{22}^{-1} a_{21} \tag{2.11}$$

$$\widehat{\beta}_n \Rightarrow A_{22}^{-1} a_{21}, \tag{2.12}$$

where $a_{11}, b_1 \in \mathbb{R}$, $a_{12}, a_{21}, b_2 \in \mathbb{R}^{m-1}$ and $A_{22} \in \mathbb{R}^{(m-1)\times(m-1)}$ such that

$$\begin{aligned} A &= \begin{pmatrix} a_{11} & a_{12}' \\ a_{21} & A_{22} \end{pmatrix} = \Omega^{1/2} \left\{ \int_0^1 W(t) W'(t) dt - \int_0^1 W(t) dt \int_0^1 W'(t) dt \right\} \Omega^{1/2}, \\ b &= \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \Omega^{1/2} \int_0^1 W(t) dt. \end{aligned}$$

Further, let $\widehat{\alpha}_l^*$ and $\widehat{\beta}_l^* = (\widehat{\beta}_2^*, \ldots, \widehat{\beta}_m^*)$ be the bootstrap analogues of $\widehat{\alpha}_n$ and of $\widehat{\beta}_n$, i.e., the least squares estimators of α and $\beta = (\beta_2, \ldots, \beta_m)$ in running the regression

$$X_{1t}^* = \alpha + \sum_{i=2}^{m} \beta_i X_{it}^* + \epsilon_t$$

using the bootstrap pseudo-time series $X_1^*, \ldots, X_l^*$ generated by the CBB scheme of Section 2.2. Here, $Z_t^* = (X_{2t}^*, \ldots, X_{mt}^*)'$, $\overline{X}_1^* = \frac{1}{l}\sum_{t=1}^{l} X_{1t}^*$ and $\overline{Z}^* = \frac{1}{l}\sum_{t=1}^{l} Z_t^*$. The corresponding bootstrap estimator is then given by

$$\widehat{\beta}_l^* = \left\{ \frac{1}{l^2} \sum_{t=1}^{l} (Z_t^* - \overline{Z}^*)(Z_t^* - \overline{Z}^*)' \right\}^{-1} \left\{ \frac{1}{l^2} \sum_{t=1}^{l} (Z_t^* - \overline{Z}^*)(X_{1t}^* - \overline{X}_1^*) \right\} \tag{2.13}$$

and $\widehat{\alpha}_l^* = \overline{X}_1^* - \overline{Z}^{*\prime} \widehat{\beta}_l^*$. We then have the following theorem.

Theorem 2.3. *Under the assumptions of Theorem 2.1, it holds*

$$\begin{aligned} \frac{1}{\sqrt{l}} \widehat{\alpha}_l^* &\Rightarrow b_1 - b_2' A_{22}^{-1} a_{21} \\ \widehat{\beta}_l^* &\Rightarrow A_{22}^{-1} a_{21} \end{aligned}$$

in probability, respectively.

As the above theorem in comparison to (2.11) and (2.12) shows, the CBB procedure succeeds in approximating correctly the distribution of the parameter estimators $\widehat{\alpha}_n$ and $\widehat{\beta}_n$ in the spurious regression case.

3. Block Bootstrap for Cointegrated Processes

3.1. **CBB applied to cointegrated processes.**

Consider now the case where the underlying process fulfills a cointegrated relation, that is, the multivariate time series $X_1, \ldots, X_n$ is generated by the $m = m_1 + m_2$-dimensional model

$$\begin{aligned} X_{1t} &= BX_{2t} + U_{1t}, \\ X_{2t} &= X_{2,t-1} + U_{2t}, \end{aligned} \tag{3.1}$$

where $X_0 = O_{as}(1)$ is random following a certain fixed distribution, X_{1t}, U_{1t} are m_1-dimensional, X_{2t}, U_{2t} are m_2-dimensional, B is an $(m_1 \times m_2)$ matrix and $\{U_t = (U_{1t}', U_{2t}')', t \in \mathbb{N}_0\}$ is a stationary process satisfying Assumption 2.1 or Assumption 2.2. Notice that $\{X_t, t \in \mathbb{N}\}$ is cointegrated with cointegration rank m_1. Our goal is to investigate the properties of the CBB bootstrap in estimating the distribution of the LS-estimator $\widehat{B}$ of B, where

$$\widehat{B} = \left(\sum_{t=1}^{n} X_{1t} X_{2t}' \right) \left(\sum_{t=1}^{n} X_{2t} X_{2t}' \right)^{-1}.$$

Under the imposed conditions [cf. Lütkepohl (2006), Lemma 7.1 or Phillips and Durlauf (1986), Theorem 4.1], we have

$$n(\widehat{B} - B) \Rightarrow \left[\Omega^{1/2} \int_0^1 W(t) dW'(t) \Omega^{1/2} + \Omega_0 + \Omega_1 \right]_{12} \left[\left[\Omega^{1/2} \int_0^1 W(t) W'(t) dt \Omega^{1/2} \right]_{22} \right]^{-1}, \tag{3.2}$$

where Ω_1, Ω and Ω_0 are defined in (2.2) and $[C]_{12}, [C]_{22}$ denote the $(m_1 \times m_2)$ upper-right and the $(m_2 \times m_2)$ lower-right part of an $(m \times m)$ matrix C with $m = m_1 + m_2$. The bootstrap analogue of $\widehat{B}$ is given by

$$\widehat{B}^* = \left(\sum_{t=1}^{l} X_{1t}^* X_{2t}^{*\prime} \right) \left(\sum_{t=1}^{l} X_{2t}^* X_{2t}^{*\prime} \right)^{-1}, \tag{3.3}$$

where $X_1^*, X_2^*, \ldots, X_l^*$ is generated according to the CBB algorithm described in Section 2.2. Now, recall that the CBB is based on resampling blocks of the differenced time series $X_t - X_{t-1}$, $t = 1, 2, \ldots, n$, which leads to

$$\begin{aligned} X_{1,t} - X_{1,t-1} &= BU_{2,t} + (U_{1,t} - U_{1,t-1}) \\ &\equiv V_{1t}, \end{aligned}$$

and $V_{1t} \neq U_{1,t} = X_{1,t} - BX_{2,t}$. Notice that $X_{2,t} - X_{2,t-1} = U_{2,t}$. Thus, the CBB uses the (centered) differences of the innovations $(V_{1,t}^{'}, U_{2,t}^{'}), t = 1, 2, \ldots, n$, to resample the blocks and not the innovations $(U_{1,t}^{'}, U_{2,t}^{'})$ as in the case of a full rank integrated process. Furthermore, the resampled blocks obtained from a centered version of $(V_{1,t}^{'}, U_{2,t}^{'}), t = 1, 2, \ldots, n$, that is the pseudo replicates $(V_{1,t}^{*\prime}, U_{2,t}^{*\prime})$, $t = 1, 2, \ldots, l$, are integrated to obtain the bootstrap observations $X_1^*, X_2^*, \ldots, X_l^*$. From this we get by using $M_t = \lceil \frac{t}{b} \rceil - 1$, $B_{t,m} = min(b, t - mb)$, $X_{1,t}^* = \sum_{j=1}^t V_{1,j}^*$, $X_{2,t}^* = \sum_{j=1}^t U_{2,j}^*$ and the definition of $V_{1,t}^*$ that

$$\begin{aligned} X_{1,t}^* &= B \sum_{m=0}^{M_t} \sum_{s=1}^{B_{t,m}} \widehat{U}_{2,i_m+s} + \sum_{m=0}^{M_t} \sum_{s=1}^{B_{t,m}} (\widehat{U}_{1,i_m+s} - \widehat{U}_{1,i_m+s-1}) \\ &= BX_{2,t}^* + \sum_{m=0}^{M_t-1} (\widehat{U}_{1,i_m+b} - \widehat{U}_{1,i_m}) + (\widehat{U}_{1,i_{M_t}+(t-M_t b)} - \widehat{U}_{1,i_{M_t}}), \end{aligned} \tag{3.4}$$

where $\widehat{U}_{1,t}$ and $\widehat{U}_{2,t}$ are the centered differences of the component processes $U_{1,t}$ and $U_{2,t}$, see also (2.3). Note also that the two last terms in the last right-hand side above are the increments between the last and the first random variable in each randomly selected block. Further, (3.4) leads to

$$\begin{aligned} \sum_{t=1}^l X_{1,t}^* X_{2,t}^{*\prime} &= B \sum_{t=1}^l X_{2,t}^* X_{2,t}^{*\prime} + \sum_{t=1}^l U_{1,t}^* X_{2,t}^{*\prime} + \sum_{t=1}^l \sum_{m=0}^{M_t-1} (\widehat{U}_{1,i_m+b} - \widehat{U}_{1,i_m}) X_{2,t}^{*\prime} - \sum_{t=1}^l \widehat{U}_{1,i_{M_t}} X_{2,t}^{*\prime} \\ &= B \sum_{t=1}^l X_{2,t}^* X_{2,t}^{*\prime} + \sum_{t=1}^l U_{1,t}^* X_{2,t}^{*\prime} + R_{1,l} + R_{2,l}, \end{aligned}$$

with an obvious notation for $R_{1,l}$ and $R_{2,l}$. Notice that in obtaining the last equality above, $\sum_{t=1}^l U_{1,i_{M_t}+(t-M_t b)} X_{2,t}^{*\prime} = \sum_{t=1}^l U_{1,t}^* X_{2,t}^{*\prime}$ has been used. From the above and (3.3) we get that

$$l(\widehat{B}^* - B) = \Big(l^{-1} \sum_{t=1}^l U_{1,t}^* X_{2,t}^{*\prime}\Big) \Big(l^{-2} \sum_{t=1}^l X_{2t}^* X_{2t}^{*\prime}\Big)^{-1} + \widetilde{R}_{1,l} + \widetilde{R}_{2,l}, \tag{3.5}$$

where

$$\widetilde{R}_{1,l} = l^{-1} R_{1,l} \Big(l^{-2} \sum_{t=1}^l X_{2t}^* X_{2t}^{*\prime}\Big)^{-1} \quad \text{and} \quad \widetilde{R}_{2,l} = l^{-1} R_{2,l} \Big(l^{-2} \sum_{t=1}^l X_{2t}^* X_{2t}^{*\prime}\Big)^{-1}.$$

Notice that the terms $R_{1,l}$ and $R_{2,l}$ are due to the fact that integrating the block resampled $V_{1,t}$'s, the differences $\widehat{U}_{1,i_m+b} - \widehat{U}_{1,i_m}$ within each block do not cancel out. Now, comparing (3.5) with

$$n(\widehat{B} - B) = \Big(n^{-1} \sum_{t=1}^n U_{1,t} X_{2,t}^{\prime}\Big) \Big(n^{-2} \sum_{t=1}^n X_{2t} X_{2t}^{\prime}\Big)^{-1},$$

indicates that $\left(l^{-1}\sum_{t=1}^{l} U_{1,t}^{*}X_{2,t}^{*\prime}\right)\left(l^{-2}\sum_{t=1}^{l} X_{2t}^{*}X_{2t}^{*\prime}\right)^{-1}$ mimics in the bootstrap world the stochastic behavior of $\left(n^{-1}\sum_{t=1}^{n} U_{1,t}X_{2,t}^{\prime}\right)\left(n^{-2}\sum_{t=1}^{n} X_{2t}X_{2t}^{\prime}\right)^{-1}$. However,

$$\frac{1}{l}R_{2,l} = \frac{1}{k}\sum_{m=0}^{k-1}\widehat{U}_{1,i_m}\frac{1}{b}\sum_{s=1}^{b} X_{2,mb+s}^{*} = O_P(1),$$

since

$$\frac{1}{\sqrt{bl}}\sum_{s=1}^{b} X_{2,mb+s}^{*} = \Omega^{*1/2}\frac{1}{\sqrt{b}}\sum_{s=1}^{b}\Omega^{*-1/2}\frac{1}{\sqrt{l}}\sum_{j=1}^{mb+s} U_{2,j}^{*} = \Omega^{*1/2}\frac{1}{\sqrt{b}}\sum_{s=1}^{b} S_l^{*}((mb+s)/l) = O_P(1),$$

see also the proof of Lemma 5.2(i), Section 5. This together with Lemma 5.1(iii), implies that $\widetilde{R}_{2,l} = O_P(1)$. Furthermore, since

$$\frac{1}{l}R_{1,l} = \frac{1}{l}\sum_{m=0}^{k-2}(\widehat{U}_{1,i_m+b} - \widehat{U}_{i_m})\sum_{t=(m+1)b+1}^{l} X_{2,t}^{*}$$

and $l^{-3/2}\sum_{t=r}^{l} X_{2,t}^{*} = \Omega^{*1/2} l^{-1}\sum_{t=r}^{l} S_l^{*}(t/l)$, see the proof of Lemma 5.2(i), we get that $l^{-1}R_{1,l} = O_P(\sqrt{lk})$, and thus $\widetilde{R}_{1,l}$ does not vanish as $n \to \infty$. The above considerations and expression (3.5) lead to the conclusion that the stochastic behavior of $n(\widehat{B} - B)$ can not be successfully approximated by that of the CBB analogue $l(\widehat{B}^{*} - B)$.

3.2. **Residual-based Block Bootstrap (RBB).**

Since the CBB applied to a cointegrated process fails due to the fact that the generated pseudo-time series $X_1^{*}, X_2^{*}, \ldots, X_n^{*}$ mimics the behavior of a full rank random walk, we show in this section that a residual-based block bootstrap (RBB) scheme succeeds in approximating correctly the distribution of the least squares estimator $\widehat{B}$ under the cointegration model (3.1). Using this residual-based block bootstrap, we generate pseudo-time series data $X_1^{+}, X_2^{+}, \ldots, X_n^{+}$ that retains the cointegration structure of the underlying process. This is achieved using the following block bootstrap algorithm. Notice that the notation X_t^{+} instead of X_t^{*} is used in order to distinguish between the pseudo-time series generated by the CBB and by the following RBB algorithm.

Step 1. Compute $\widehat{B}$ and $\widetilde{U}_t = (\widetilde{U}_{1t}^{\prime}, \widetilde{U}_{2t}^{\prime})^{\prime}$ for $t = 1, \ldots, n$ according to

$$\begin{aligned}\widetilde{U}_{1t} &= X_{1t} - \widehat{B}X_{2t},\\ U_{2t} &= X_{2t} - X_{2,t-1}.\end{aligned}$$

Step 2. Choose a block length $b < n$ and let k be the smallest number of blocks needed to get a bootstrap sample of length $l = kb$, such that $l \geq n$. Let $i_0, \ldots, i_{k-1}$ be i.i.d. random variables uniformly distributed on the set $\{0, 1, 2, \ldots, n-b\}$.

Step 3. Let $U_1^{+}, \ldots, U_l^{*}$ with $U_t^{+} = (U_{1t}^{+\prime}, U_{2t}^{+\prime})^{\prime}$ be a moving block bootstrap sample, where for $j = 1, 2, \ldots, b$ and $m = 0, 1, 2, \ldots, k-1$,

$$\begin{aligned}\begin{pmatrix} U_{1,mb+j}^{+} \\ U_{2,mb+j}^{+} \end{pmatrix} &= \begin{pmatrix} \widehat{U}_{1,i_m+j} \\ \widehat{U}_{2,i_m+j} \end{pmatrix} = \begin{pmatrix} \widetilde{U}_{1,i_m+j} - E^{+}(\widetilde{U}_{1,i_m+j}) \\ U_{2,i_m+j} - E^{+}(U_{2,i_m+j}) \end{pmatrix}\\ &= \begin{pmatrix} \widetilde{U}_{1,i_m+j} \\ U_{2,i_m+j} \end{pmatrix} - \frac{1}{n-b+1}\sum_{\tau=0}^{n-b}\begin{pmatrix} \widetilde{U}_{1,\tau+j} \\ U_{2,\tau+j} \end{pmatrix}. \qquad (3.6)\end{aligned}$$

Step 4. Generate then the bootstrap pseudo-time series $(X_{1t}^{+\prime}, X_{2t}^{+\prime})'$, $t = 1, \ldots, l$ from

$$X_{2t}^{+} = \begin{cases} X_{20}, & t = 0 \\ X_{2,t-1}^{+} + U_{2t}^{+}, & t \in \{1, \ldots, l\} \end{cases}$$

and

$$X_{1t}^{+} = \widehat{B} X_{2t}^{+} + U_{1t}^{+}, \quad t \in \{1, \ldots, l\}.$$

This leads to the estimator

$$\widehat{B}^{+} = \left(\sum_{t=1}^{l} X_{1t}^{+} X_{2t}^{+\prime} \right) \left(\sum_{t=1}^{l} X_{2t}^{+} X_{2t}^{+\prime} \right)^{-1}.$$

As (3.2) holds under Assumptions 2.1 or 2.2, the following theorem shows that the RBB procedure described above succeeds in mimicking correctly the random behavior of the LS-estimator $\widehat{B}$ in the cointegrated case (3.1).

Theorem 3.1. *Let $\{X_t, t \in \mathbb{N}_0\}$ be an m-dimensional stochastic process following* (3.1) *and assume that the process $\{U_t, t \in \mathbb{N}\}$ satisfies Assumption 2.1 or Assumption 2.2. If $b \to \infty$ as $n \to \infty$ such that $b/\sqrt{n} \to 0$ then*

$$l(\widehat{B}^{+} - \widehat{B}) \Rightarrow \left[\Omega^{1/2} \int_0^1 W(t) dW'(t) \Omega^{1/2} + \Omega_0 + \Omega_1 \right]_{12} \left[\left[\Omega^{1/2} \int_0^1 W(t) W'(t) dt \Omega^{1/2} \right]_{22} \right]^{-1}$$

in probability, where Ω_1, Ω and Ω_0 are defined in (2.2). *Here, $[C]_{12}$ and $[C]_{22}$ denote the $(m_1 \times m_2)$ upper-right and the $(m_2 \times m_2)$ lower-right part of an $(m \times m)$ matrix C with $m = m_1 + m_2$.*

4. Conclusions

In this paper, some asymptotic theory for block bootstrap resampling schemes has been developed when such procedures are applied to multivariate integrated and/or cointegrated time series. We proved a functional central limit theorem for the partial sum process based on pseudo-time series generated using a multivariate continuous-path block bootstrap (CBB) procedure. The pseudo-time series generated in this context is based on integrating resampled blocks of the centered differences of the observed multivariate time series. This basic result is used to establish asymptotic validity of the CBB procedure for estimating the distribution of least squares estimators, both, in the full rank regression and in the spurious regression case. We showed further, that the CBB procedure fails in the case of cointegrated processes since it does not capture appropriately the cointegration structure of the underlying time series. For this kind of integrated multivariate processes, a block resampling scheme based on regression residuals (RBB) is more appropriate. It is shown that the RBB procedure is valid for estimating the distribution of the least square estimator of the cointegration matrix.

5. Proofs

Proof of Theorem 2.1: First, note that

$$\sum_{j=1}^{t} U_t^{*} = \sum_{m=0}^{M_t} \sum_{s=1}^{B_{t,m}} U_{mb+s}^{*} = \sum_{m=0}^{M_t} \sum_{s=1}^{B_{t,m}} \widehat{U}_{i_m+s},$$

where $M_t = \lceil \frac{t}{b} \rceil - 1$, $B_{t,m} = min(b, t - mb)$ and

$$\widehat{U}_{i_m+s} := U_{i_m+s} - E^{*}\left(U_{i_m+s}\right) = U_{i_m+s} - \frac{1}{n-b+1} \sum_{\tau=0}^{n-b} U_{\tau+s}.$$

Here, $\lceil x \rceil$ ($\lfloor x \rfloor$) denotes the smallest (largest) integer that is larger (smaller) or equal to x. Similarly, we set $M_\nu = \lfloor \frac{\lfloor l\nu \rfloor - 1}{b} \rfloor$ and $B_{\nu,m} = min(b, \lfloor l\nu \rfloor - mb)$ and together with the definition of $S_l^*(\nu)$ in (2.5), this leads to

$$\begin{aligned} S_l^*(\nu) &= \frac{1}{\sqrt{l}} (\Omega_l^*)^{-1/2} \sum_{m=0}^{M_\nu} \sum_{s=1}^{B_{\nu,m}} \widehat{U}_{i_m+s} \\ &= \frac{1}{\sqrt{l}} (\Omega_l^*)^{-1/2} \sum_{m=0}^{M_\nu} \sum_{s=1}^{b} \widehat{U}_{i_m+s} - \frac{1}{\sqrt{l}} (\Omega_l^*)^{-1/2} \sum_{s=B_{\nu,M_\nu}+1}^{b} \widehat{U}_{i_{M_\nu}+s} \end{aligned}$$

and since

$$\sup_\nu \left| \frac{1}{\sqrt{l}} (\Omega_l^*)^{-1/2} \sum_{s=B_{\nu,M_\nu}+1}^{b} \widehat{U}_{i_{M_\nu}+s} \right| = O_P(k^{-1/2}),$$

it remains to show

$$\left(\frac{1}{\sqrt{l}} (\Omega_l^*)^{-1/2} \sum_{m=0}^{M_\nu} \sum_{s=1}^{b} \widehat{U}_{i_m+s} \right)_{\nu \in [0,1]} \Rightarrow (W(\nu), \nu \in [0,1]) \, .$$

in probability. We can consider instead the asymptotically equivalent statistic

$$\widetilde{S}_l^*(\nu) = \frac{1}{\sqrt{l}} (\Omega_l^*)^{-1/2} \sum_{m=0}^{\lfloor k\nu \rfloor} \sum_{s=1}^{b} \widehat{U}_{i_m+s} = \frac{1}{\sqrt{k}} \sum_{m=0}^{\lfloor k\nu \rfloor} V_m^*, \tag{5.1}$$

where $\{V_m^*, m = 0, 1, 2, \ldots, \lfloor k\nu \rfloor\}$ with

$$V_m^* = \frac{1}{\sqrt{b}} (\Omega_l^*)^{-1/2} \sum_{s=1}^{b} \widehat{U}_{i_m+s} \tag{5.2}$$

forms a triangular array of random variables that are independently and identically distributed conditionally on $X_1, \ldots, X_n$. Now, to prove $\widetilde{S}_l^* \Rightarrow W$, we have to show

(a) Convergence of all finite dimensional distributions
(b) Tightness

For the first part (a), let $q \in \mathbb{N}$ and set $\underline{\nu} = (\nu_1, \ldots, \nu_q)'$, where $0 \le \nu_1 \le \ldots \le \nu_q \le 1$. Further, set $\lambda = (\lambda_1', \ldots, \lambda_q')'$, where $\lambda_p \in \mathbb{R}^m$ for $p = 1, \ldots, q$, and let $\widetilde{S}_l^*(\underline{\nu})$ be the vector that stacks $\widetilde{S}_l^*(\nu_1), \ldots, \widetilde{S}_l^*(\nu_q)$ and $\Delta\widetilde{S}_l^*(\underline{\nu})$ be the vector that stacks the differences $\widetilde{S}_l^*(\nu_1) - \widetilde{S}_l^*(\nu_0), \ldots, \widetilde{S}_l^*(\nu_q) - \widetilde{S}_l^*(\nu_{q-1})$, where we set $\nu_0 = 0$. Note that the qm-dimensional process $\{\Delta\widetilde{S}_l^*(\underline{\nu}), \nu \in [0,1]\}$ consists of q conditionally independent m-dimensional processes due to the independence of $\{V_m^*, m = 0, 1, 2, \ldots, \lfloor k\nu \rfloor\}$. This allows us to consider them separately. We have

$$\widetilde{S}_l^*(\nu_p) - \widetilde{S}_l^*(\nu_{p-1}) = \frac{1}{\sqrt{k}} \sum_{m=\lfloor k\nu_{p-1} \rfloor + 1}^{\lfloor k\nu_p \rfloor} V_m^* = \frac{1}{\sqrt{l}} (\Sigma_l^*)^{-1/2} \sum_{m=\lfloor k\nu_{p-1} \rfloor + 1}^{\lfloor k\nu_p \rfloor} \sum_{s=1}^{b} \widehat{U}_{i_m+s},$$

and due to $E^*(V_m^*)=0$ and Lemma 5.1(i), we obtain

$$
\begin{aligned}
& Var^*\left(\lambda_p'\left\{\widetilde{S}_l^*(\nu_p)-\widetilde{S}_l^*(\nu_{p-1})\right\}\right)\\
= \;& \frac{1}{k}\lambda_p'\left\{(\Omega_l^*)^{-1/2}\,E^*\left[\sum_{m_1=\lfloor k\nu_{p-1}\rfloor+1}^{\lfloor k\nu_p\rfloor}\sum_{s_1=1}^{b}\widehat{U}_{i_{m_1}+s_1}\sum_{m_2=\lfloor k\nu_{p-1}\rfloor+1}^{\lfloor k\nu_p\rfloor}\sum_{s_2=1}^{b}\widehat{U}_{i_{m_2}+s_2}'\right](\Omega_l^*)^{-1/2}\right\}\lambda_p\\
= \;& \frac{\lfloor k\nu_p\rfloor-\lfloor k\nu_{p-1}\rfloor}{k}\lambda_p'\left\{(\Omega_l^*)^{-1/2}\left(\Omega^*_{\lfloor k\nu_p\rfloor-\lfloor k\nu_{p-1}\rfloor}\right)(\Omega_l^*)^{-1/2}\right\}\lambda_p\\
\underset{k\to\infty}{\longrightarrow} \;& \lambda_p'\left\{(\nu_p-\nu_{p-1})\,I_m\right\}\lambda_p.
\end{aligned}
$$

as $n\to\infty$ in probability for any p. To prove asymptotic normality, we show a CLT based on the Liapunov condition for triangular arrays $\{V_m^*, m=0,1,2,\ldots,\lfloor k\nu\rfloor\}$ [see e.g. Serfling (1980)]. More precisely, we consider

$$
\frac{\left(\frac{1}{\sqrt{k}}\right)^{2+\kappa}\sum_{m=\lfloor k\nu_{p-1}\rfloor+1}^{\lfloor k\nu_p\rfloor}E^*\left(\left|\lambda_p'V_m^*\right|^{2+\kappa}\right)}{\left(\frac{1}{k}Var^*\left(\sum_{m=\lfloor k\nu_{p-1}\rfloor+1}^{\lfloor k\nu_p\rfloor}\lambda_p'V_m^*\right)\right)^{(2+\kappa)/2}}.
$$

The denominator is bounded from below and from above with probability tending to one and the nominator vanishes due to

$$
\begin{aligned}
\frac{1}{k^{(2+\kappa)/2}}\sum_{m=\lfloor k\nu_{p-1}\rfloor+1}^{\lfloor k\nu_p\rfloor}E^*\left(\left|\lambda_p'V_m^*\right|^{2+\kappa}\right) &= \frac{(\lfloor k\nu_p\rfloor-\lfloor k\nu_{p-1}\rfloor)}{k^{(2+\kappa)/2}}\frac{1}{n-b+1}\sum_{t=0}^{n-b}\left|\lambda_p'\frac{1}{\sqrt{b}}(\Omega_l^*)^{-1/2}\sum_{s=1}^{b}\widehat{U}_{t+s}\right|^{2+\kappa}\\
&= O_P\left(\frac{1}{k^{\kappa/2}}\right),
\end{aligned}
$$

because of

$$
\frac{1}{n-b+1}\sum_{t=0}^{n-b}\left|\lambda_p'\frac{1}{\sqrt{b}}(\Omega_l^*)^{-1/2}\sum_{s=1}^{b}\widehat{U}_{t+s}\right|^{2+\kappa}=\int_0^1\left|\lambda_p'\widetilde{S}_b(\nu)\right|^{2+\kappa}d\nu\Rightarrow\int_0^1\left|\lambda_p'W(\nu)\right|^{2+\kappa}d\nu=O_P(1).
$$

Here, the partial sum process $\{\widetilde{S}_b(\nu), 0\le\nu\le1\}$ is similarly defined as $\{S_l^*(\nu), 0\le\nu\le1\}$ in (2.5) and fulfills a FLT under the imposed conditions. Now, we can conclude that $S_l^*(\nu_p)-S_l^*(\nu_{p-1})\Rightarrow W(\nu_p-\nu_{p-1})$ in probability. Together, we have that $\Delta\widetilde{S}_l^*(\underline{\nu})$ converges weakly to $(W'(\nu_1-\nu_0),\ldots,W'(\nu_q-\nu_{q-1}))'$. Finally, multiplying $\Delta\widetilde{S}_l^*(\underline{\nu})$ with a suitable matrix gives $\widetilde{S}_l^*(\underline{\nu})$ and by $W(0)=0$ and $W(a-b)=W(a)-W(b)$ for $a,b\in[0,1]$, we get the desired convergence of the finite dimensional distributions. For the tightness of part (b) recall first that in probability measures on the product space are tight if and oly if the marginal probability measures are tight; cf. Billingsley (1999), page 65. Now, for $i\in\{1,2,\ldots,m\}$ consider

$$
\widetilde{S}_{i,l}^*(\nu)=\underline{e}_i'\widetilde{S}_l^*(\nu),
$$

where $\underline{e}_i=(0,\ldots,0,1,0,\ldots,0)$ is the i-th unit vector. Denote by $P_{i,l}^*$ the probability measure of $\widetilde{S}_{i,l}^*$ and by P_l^* the probability measure of $\widetilde{S}_l^*$. Then, by a version of Donsker's Theorem for triangular arrays of row-wise independent random variables, Billingsley (1999), pages 147-148, we immediately get that $P_{i,l}^*$ is tight from which the tightness of P_l^* follows. □

Proof of Theorem 2.2: It holds

$$
l(\widehat{A}^*-I_m)=\left(\frac{1}{l}\sum_{t=2}^{l}U_t^*X_{t-1}^{*'}\right)\left(\frac{1}{l^2}\sum_{t=2}^{l}X_{t-1}^*X_{t-1}^{*'}\right)^{-1}
$$

and the claimed result follows from Lemma 5.1 and the positive definiteness of Ω. □

Lemma 5.1. *Under the assumptions of Theorem 2.2, it holds*

$$(i) \quad \Omega_l^* \to \Omega \quad \textit{in probability},$$

$$(ii) \quad \frac{1}{l}\sum_{t=2}^{l} U_t^* X_{t-1}^{*\prime} \Rightarrow \Omega^{1/2}\int_0^1 W(t)dW'(t)\Omega^{1/2} + \Omega_1,$$

$$(iii) \quad \frac{1}{l^2}\sum_{t=2}^{l} X_{t-1}^* X_{t-1}^{*\prime} \Rightarrow \Omega^{1/2}\int_0^1 W(t)W(t)'dt\Omega^{1/2}$$

in probability, respectively, where joint convergence (of (ii) and (iii)) also applies.

Proof: (i) By (2.6) and rewriting the involved sums, we get

$$\begin{aligned}
\Omega_l^* &= Var^*\left(\frac{1}{\sqrt{l}}\sum_{t=1}^{l} U_t^*\right) \\
&= \frac{1}{b}\sum_{s_1,s_2=1}^{b}\frac{1}{n-b+1}\sum_{t=0}^{n-b}\widehat{U}_{t+s_1}\widehat{U}_{t+s_2}' \\
&= \sum_{h=-b-1}^{b-1}\frac{1}{b}\sum_{r=max(1,1-h)}^{min(b,b-h)}\frac{1}{n-b+1}\sum_{t=0}^{n-b}\widehat{U}_{t+r+h}\widehat{U}_{t+r}' \\
&= \sum_{h=-(b-1)}^{b-1}\left\{\frac{1}{n-b+1}\sum_{t=1}^{n}\frac{b-|h|}{b}\widehat{U}_{t+h}\widehat{U}_t'\right\} + O\left(\frac{b^2}{n}\right) \qquad (5.3) \\
&= A_q + (A - A_q),
\end{aligned}$$

where A denotes the first term in (5.3) and A_q the same sum with $\sum_{h=-(b-1)}^{b-1}$ replaced by $\sum_{h=-(q-1)}^{q-1}$ for some fixed q. Now, we use Proposition 6.3.9 in Brockwell and Davis (1991). Under the assumptions, we have for any fixed $h \in \mathbb{Z}$ that $\frac{1}{n-b+1}\sum_{t=1}^{n}\frac{b-|h|}{b}\widehat{U}_{t+h}\widehat{U}_t' \to \Gamma(h)$ holds in probability as $n \to \infty$ and we get

$$\sum_{h=-(q-1)}^{q-1}\left\{\frac{1}{n-b+1}\sum_{t=1}^{n}\frac{b-|h|}{b}\widehat{U}_{t+h}\widehat{U}_t'\right\} + O\left(\frac{b^2}{n}\right) \underset{n\to\infty}{\to} \sum_{h=-q}^{q}\Gamma(h) \underset{q\to\infty}{\to} \sum_{h=-\infty}^{\infty}\Gamma(h)$$

in probability. It remains to show for all $\delta > 0$ that $\lim_{q\to\infty}\limsup_{n\to\infty} P^*(\|A - A_q\| \geq \delta) = 0$ in probability, which can be proved by standard arguments.
(ii) Plugging-in for $X_{t-1}^{*\prime}$ leads to

$$\begin{aligned}
\frac{1}{l}\sum_{t=2}^{l} U_t^* X_{t-1}^{*\prime} &= \frac{1}{l}\sum_{t=1}^{l}\left\{U_t^*\sum_{j=1}^{t-1}U_j^{*\prime}\right\} + \left(\frac{1}{l}\sum_{t=1}^{l}U_t^*\right)X_0' - \frac{1}{l}U_1^* X_0 \\
&= \frac{1}{l}\sum_{p=0}^{k-1}\sum_{r=1}^{b}\left\{U_{pb+r}^*\left(\sum_{m=0}^{M_{pb+r-1}}\sum_{s=1}^{B_{pb+r-1,m}}U_{mb+s}^{*\prime}\right)\right\} + O_{P^*}\left(\frac{1}{\sqrt{l}}\right) \\
&= \frac{1}{l}\sum_{p=0}^{k-1}\sum_{r=1}^{b}\left\{\widehat{U}_{i_p+r}\left(\sum_{m=0}^{p-1}\sum_{s=1}^{b}\widehat{U}_{i_m+s}'\right)\right\} + \frac{1}{l}\sum_{p=0}^{k-1}\sum_{r=1}^{b}\left\{\widehat{U}_{i_p+r}\left(\sum_{s=1}^{r-1}\widehat{U}_{i_p+s}\right)\right\} + O_{P^*}\left(\frac{1}{\sqrt{l}}\right) \\
&= A_1^* + A_2^* + O_{P^*}\left(\frac{1}{\sqrt{l}}\right).
\end{aligned}$$

with an obvious notation for A_1^* and A_2^*. The first term can be expressed as

$$A_1^* = (\Omega_l^*)^{1/2} \left\{ \frac{1}{k} \sum_{p=0}^{k-1} V_p^* \left(\sum_{m=0}^{p-1} V_m^* \right) \right\} (\Omega_l^*)^{1/2},$$

where V_m^* is defined in (5.2) which forms a triangular array of random variables that are independently distributed conditionally on $X_1, \dots, X_n$. By adopting the proof of Theorem 2.4(ii) in Chan and Wei (1988) for our purposes and the FLT in Theorem 2.1, we can show that

$$\frac{1}{k} \sum_{p=0}^{k-1} V_p^* \left(\sum_{m=0}^{p-1} V_m^* \right) \Rightarrow \int_0^1 W(t) dW'(t)$$

in probability. This is possible, because their proof is based exclusively on the FLT and, therefore, the fact that $\{V_m^*\}$ is a triangular array does not alter the proof. For A_2^*, we get

$$E^*(A_2^*) = \frac{1}{l} \sum_{p=0}^{k-1} \sum_{r=1}^{b} \sum_{s=1}^{r-1} E^* \left(\widehat{U}_{i_p+r} \widehat{U}'_{i_p+s} \right) = \frac{1}{b} \sum_{r=1}^{b} \sum_{s=1}^{r-1} \frac{1}{n-b+1} \sum_{t=0}^{n-b} \widehat{U}_{t+r} \widehat{U}'_{t+s},$$

which converges to Ω_1 by Proposition 6.3.9 in Brockwell and Davis (1991). Similar arguments show $Var^*(A_2) = o_P(1)$.

(iii) Since $X_t^* = X_0 + \sum_{j=1}^t Y_j^* = X_0 + \sum_{j=1}^t U_j^*$, we get

$$\begin{aligned}
\frac{1}{l^2} \sum_{t=2}^{l} X_{t-1}^* X_{t-1}^{*\prime} &= \frac{1}{l} \sum_{t=2}^{l} \left(\frac{1}{\sqrt{l}} \sum_{j=1}^{t-1} U_j^* \right) \left(\frac{1}{\sqrt{l}} \sum_{j=1}^{t-1} U_j^{*\prime} \right) + O_{P^*} \left(\frac{1}{\sqrt{l}} \right) \\
&= (\Omega_l^*)^{1/2} \frac{1}{l} \sum_{t=2}^{l} \left(\frac{1}{\sqrt{l}} (\Omega_l^*)^{-1/2} \sum_{j=1}^{t-1} U_j^* \right) \left(\frac{1}{\sqrt{l}} \sum_{j=1}^{t-1} U_j^{*\prime} (\Omega_l^*)^{-1/2} \right) (\Omega_l^*)^{1/2} + O_{P^*} \left(\frac{1}{\sqrt{l}} \right) \\
&= (\Omega_l^*)^{1/2} \frac{1}{l} \sum_{t=2}^{l} S_l^* \left(\frac{t-1}{l} \right) S_l^{*\prime} \left(\frac{t-1}{l} \right) (\Omega_l^*)^{1/2} + O_{P^*} \left(\frac{1}{\sqrt{l}} \right) \\
&= (\Omega_l^*)^{1/2} \int_0^1 S_l^*([l\nu]) S_l^{*\prime}([l\nu]) d\nu (\Omega_l^*)^{1/2} + O_{P^*} \left(\frac{1}{\sqrt{l}} \right) \\
&\Rightarrow \Omega^{1/2} \int_0^1 W(\nu) W'(\nu) d\nu \Omega^{1/2},
\end{aligned}$$

in probability by Theorem 2.1, part (i) of this lemma and the continuous mapping theorem. □

Proof of Theorem 2.3: Due to (2.13), the claimed result follows immediately from Lemma 5.2 below. □

Lemma 5.2. *Under the assumptions of Theorem 2.3, it holds*

(*i*) $l^{-3/2} \sum_{t=1}^{l} X_t^* \Rightarrow \Omega^{1/2} \int_0^1 W(\nu) d\nu,$

(*ii*) $l^{-2} \sum_{t=1}^{l} X_t^* X_t^{*\prime} \Rightarrow \Omega^{1/2} \int_0^1 W(\nu) W'(\nu) d\nu \Omega^{1/2},$

(*iii*) $l^{-2} \sum_{t=1}^{l} (X_t^* - \overline{X}^*)(X_t^* - \overline{X}^*)' \Rightarrow \Omega^{1/2} \left\{ \int_0^1 W(t) W'(t) dt - \int_0^1 W(t) dt \int_0^1 W'(t) dt \right\} \Omega^{1/2}.$

Proof: Considering part (i), Theorem 2.1, Lemma 5.1 and the continuous mapping theorem, gives

$$\begin{aligned}
l^{-3/2}\sum_{t=1}^{l} X_t^* &= l^{-3/2}\sum_{t=1}^{l}\sum_{j=1}^{t-1} U_j^* + l^{-1/2}X_0 + l^{-3/2}\sum_{t=1}^{l} U_t^* \\
&= (\Omega_l^*)^{1/2}\frac{1}{l}\sum_{t=1}^{l}\frac{1}{\sqrt{l}}(\Omega_l^*)^{-1/2}\sum_{j=1}^{t-1} U_j^* + O_{P^*}\left(l^{-1/2}\right) + O_{P^*}\left(l^{-1}\right) \\
&= (\Omega_l^*)^{1/2}\frac{1}{l}\sum_{t=1}^{l} S_l^*\left(\frac{t-1}{l}\right) + O_{P^*}\left(l^{-1/2}\right) \\
&= (\Omega_l^*)^{1/2}\sum_{t=1}^{l}\int_{\frac{t-1}{l}}^{\frac{t}{l}} S_l^*\left(\lfloor l\nu\rfloor\right)d\nu + O_{P^*}\left(l^{-1/2}\right) \\
&= (\Omega_l^*)^{1/2}\int_0^1 S_l^*\left(\lfloor l\nu\rfloor\right)d\nu + O_{P^*}\left(l^{-1/2}\right) \\
&\Rightarrow \Omega^{1/2}\int_0^1 W(\nu)d\nu
\end{aligned}$$

in probability. Part (ii) follows from $l^{-2}\sum_{t=1}^{l} X_t^* X_t^{*\prime} = l^{-2}\sum_{t=2}^{l} X_{t-1}^* X_{t-1}^{*\prime} + O_{P^*}(\frac{1}{l})$ and Lemma 5.1 and (iii) is an immediate consequence of (i) and (ii) of this lemma, due to

$$l^{-2}\sum_{t=1}^{l}(X_t^* - \overline{X}^*)(X_t^* - \overline{X}^*)' = l^{-2}\sum_{t=1}^{l} X_t^* X_t^{*\prime} - \left(l^{-3/2}\sum_{t=1}^{l} X_t^*\right)\left(l^{-3/2}\sum_{t=1}^{l} X_t^*\right)'.$$

□

Proof of Theorem 3.1: It holds

$$l(\widehat{B}^+ - \widehat{B}) = \left(\frac{1}{l}\sum_{t=1}^{l} U_{1t}^+ X_{2t}^{+\prime}\right)\left(\frac{1}{l^2}\sum_{t=1}^{l} X_{2t}^+ X_{2t}^{+\prime}\right)^{-1}$$

and the claimed result follows from Lemma 5.3. □

Lemma 5.3. *Let Ω_l^+ be defined analogue to (2.6), but based on $X_1^+,\ldots,X_n^+$. Under the assumptions of Theorem 3.1, it holds*

$$\begin{aligned}
&(i) \quad \Omega_l^+ \to \Omega \quad \textit{in probability}, \\
&(ii) \quad \frac{1}{l}\sum_{t=1}^{l} U_{1t}^+ X_{2t}^{+\prime} \Rightarrow \left[\Omega^{1/2}\int_0^1 W(t)dW'(t)\Omega^{1/2} + \Omega_0 + \Omega_1\right]_{12}, \\
&(iii) \quad \frac{1}{l^2}\sum_{t=1}^{l} X_{2t}^+ X_{2t}^{+\prime} \Rightarrow \left[\Omega^{1/2}\int_0^1 W(t)W'(t)dt\,\Omega^{1/2}\right]_{22},
\end{aligned}$$

in probability, respectively, where joint convergence (of (ii) and (iii)) also applies.

Proof: First, by (3.1) and (3.6), we get

$$
\begin{aligned}
\begin{pmatrix} U^+_{1,mb+j} \\ U^+_{2,mb+j} \end{pmatrix} &= \begin{pmatrix} \widetilde{U}_{1,i_m+j} \\ U_{2,i_m+j} \end{pmatrix} - \frac{1}{n-b+1}\sum_{\tau=0}^{n-b} \begin{pmatrix} \widetilde{U}_{1,\tau+j} \\ U_{2,\tau+j} \end{pmatrix} \\
&= \begin{pmatrix} U_{1,i_m+j} \\ U_{2,i_m+j} \end{pmatrix} - \frac{1}{n-b+1}\sum_{\tau=0}^{n-b} \begin{pmatrix} U_{1,\tau+j} \\ U_{2,\tau+j} \end{pmatrix} \\
&\quad - \begin{pmatrix} (\widehat{B}-B)X_{2,i_m+j} \\ 0 \end{pmatrix} - \frac{1}{n-b+1}\sum_{\tau=0}^{n-b} \begin{pmatrix} (\widehat{B}-B)X_{2,\tau+j} \\ 0 \end{pmatrix}
\end{aligned}
\tag{5.4}
$$

for $j = 1, \ldots, b$ and $m = 0, \ldots, k-1$.

(ii) Plugging-in for $X^{+\prime}_{t-1}$ leads to

$$
\begin{aligned}
\frac{1}{l}\sum_{t=1}^{l} U^+_{1t}X^{+\prime}_{2t} &= \frac{1}{l}\sum_{t=1}^{l}\left\{U^+_{1t}\sum_{j=1}^{t-1}U^{+\prime}_{2j}\right\} + \frac{1}{l}\sum_{t=1}^{l}U^+_{1t}U^{+\prime}_{2t} + \left(\frac{1}{l}\sum_{t=1}^{l}U^+_{1t}\right)X'_{20} \\
&= \frac{1}{l}\sum_{t=1}^{l}\left\{U^+_{1t}\sum_{j=1}^{t-1}U^{+\prime}_{2j}\right\} + \frac{1}{l}\sum_{t=1}^{l}U^+_{1t}U^{+\prime}_{2t} + O_{P^+}\left(\frac{1}{\sqrt{l}}\right).
\end{aligned}
$$

The asymptotic behaviour of the first term above follows by using (5.4) immediately from Lemma 5.1(ii), where the second term on the last right-hand side of (5.4) vanishes asymptotically due to $(\widehat{B}-B) = O_P(n^{-1})$. Similarly, for the second term, we have

$$
\begin{aligned}
E^+\left(\frac{1}{l}\sum_{t=1}^{l}U^+_{1t}U^{+\prime}_{2t}\right) &= \frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b}E^+\left(U_{1,i_m+s}U'_{2,i_m+s}\right) + o_P(1) \\
&= \frac{1}{b}\sum_{s=1}^{b}\frac{1}{n-b+1}\sum_{\tau=0}^{n-b}U_{1,\tau+s}U'_{2,\tau+s} + o_P(1),
\end{aligned}
$$

which converges to Ω_0 by Proposition 6.3.9 in Brockwell and Davis (1991). Similar arguments show that its conditional variance vanishes asymptotically in probability. Part (i) follows by similar arguments from (5.4) and Lemma 5.1(i) and part (iii) follows immediately from Lemma 5.1(iii). □

References

[1] Badillo, R., Belaire-Franch, J. and Reverte, C. (2010). Residual-based block bootstrap for cointegration testing. *Applied Economics Letters*, 17, 999–1003.

[2] Billingsley, P. (1999). Convergence of Probability Measures. John Wiley & Sohns, Inc.: New York.

[3] Brockwell, P. J. and Davis, R. A. (1991). Time series: theory and methods. 2nd ed. Springer-Verlag, Berlin.

[4] Chan, N. H. and Wei, C. Z. (1988). Limiting Distributions of Least Squares Estimates of Unstable Autoregressive Processes. *Annals of Statistics*, 16, 367–401.

[5] Künsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *Ann. Statist.*, 17, No. 3, 1217-1241.

[6] Lahiri, S. N. (2003). Resampling methods for dependent data. Springer Series in Statistics. New York, NY: Springer.

[7] Li, H. and Maddala, G. S. (1997). Bootstrapping Cointegrating Regressions. *Journal of Econometrics*, 80, 297–318.

[8] Liu, R. Y. and Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In: Exploring the Limits of Bootstrap (Eds.: LePage, R. and Billard, L.), 225-248. Wiley, New York.

[9] Lütkepohl (2005). New Introduction to Multiple Time Series Analysis. Berlin: Springer.

[10] Hamilton, J. D. (1994). Time Series Analysis. Princeton University Press, Princeton, New Jersey.

[11] Paparoditis, E. and Politis, D. N. (2001) Unit root testing via the continuous-path block bootstrap, Discussion Paper 20012006, Department of Economics, University of California, San Diego, USA.
[12] Paparoditis, E. and Politis, D. N. (2003). Residual-based block bootstrap for unit root testing. *Econometrica*, 71, 813-855.
[13] Parker, C., Paparoditis, E. and Politis, D. N. (2006). Unit root testing via the stationary bootstrap. *Journal of Econometrics*, 133, 601–638.
[14] Phillips, P. C. B. (1986). Understanding spurious regressions in econometrics. *Journal of Econometrics*, 33, 311–340.
[15] Phillips, P. C. B. (1988). Weak Convergence to the Matrix Stochastic Integral $\int_{0^1} BdB'$. *Journal of Mulitvariate Analysis*, 24, 252–264.
[16] Phillips, P. C. B. (2001). Bootstrapping spurious regression. Cowles Foundation Discussion Paper # 1331, Yale University.
[17] Phillips, P. C. B. (2010). Bootstrapping I(1) data. *Journal of Econometrics*, 158, 280–284.
[18] Phillips, P. C. B. and Durlauf, S. N. (1986). Multiple time series regression with integrated processes. *Rev. Econ. Stud.*, 53, 473–495.
[19] Serfling, R. (1980). Approximation Theorems of Mathematical Statistics. New York: John Wiley.

Department of Economics, University of Mannheim, L7, 3-5, 68131 Mannheim, Germany
E-mail address: `cjentsch@mail.uni-mannheim.de`

Department of Mathematics and Statistics, University of Cyprus, P.O.Box 20537, CY-1678 Nicosia, Cyprus
E-mail address: `stathis@ucy.ac.cy`

Department of Mathematics, University of California, San Diego, La Jolla, CA 92093-0112, USA
E-mail address: `politis@math.ucsd.edu`