The Continuous-Path Block-Bootstrap

E. PAPARODITIS and D. N. POLITIS
University of Cyprus and University of California, San Diego

Abstract - The situation where the available data arise from a general linear process with a unit root is discussed. We propose a modification of the Block Bootstrap which generates replicates of the original data and which correctly imitates the unit root behavior and the weak dependence structure of the observed series. Validity of the proposed method for estimating the unit root distribution is shown.

Research supported by NSF Grant DMS-97-03964 and by a University of Cyprus Research Grant.

1. INTRODUCTION

Consider time series data $\{X(1), X(2), \ldots, X(n)\}$ arising from the process

$$X(t) \ = \ \rho X(t-1) + U(t), \tag{1}$$

for $t = 1, 2, \ldots$, where $X(1) = 0$, ρ is a constant in $[-1, 1]$, and

Assumption A $\{U(t), t \geq 1\}$ is a stochastic process satisfying

$$U(t) \ = \ \sum_{j=0}^{\infty} \psi_j \varepsilon(t-j) \tag{2}$$

where $\psi_0 = 1$, $\sum_{j=0}^{\infty} j|\psi_j| < \infty$, $\sum_{j=0}^{\infty} \psi_j \neq 0$ and $\{\varepsilon(t), t \in \mathbb{Z}\}$ is a sequence of independent identically distributed random variables with mean zero and $0 < E(\varepsilon^2(1)) < \infty$.

We will be especially concerned with the nonstationary (integrated) case where $\rho = 1$. Note that if $\psi_j = 0$ for $j > 1$ we are in the case of a random walk, i.e., (2) allows for a wide range of weak dependence of the differenced process $X(t) - X(t-1)$.

A number of papers in the econometrics literature has dealt with model (1); see e.g. Hamilton (1994) or Fuller (1996) and the references therein. The traditional approach so far has been based on the Dickey and Fuller (1979) pioneering work and consists of conducting a test of the null hypothesis that there is a unit root; in this connection see also Phillips and Perron (1988), and Ferretti and Romo (1996).

Recently however, there has been some interest in the attempt to go beyond the simple unit root test. Stock (1991) managed to develop confidence intervals for ρ in the equation $X(t) = \rho X(t-1) + U(t)$ based on 'local-to-unity' asymptotics. Hansen (1997) proposed the 'grid-bootstrap' to adress this situation, and reports improved performance. Finally, Romano and Wolf (1998) applied the general subsampling methodology of Politis and Romano (1994) to the AR(1) model with good results; see Politis et al. (1999) for more details.

In the paper at hand, we present a different approach towards inference under the presence of a unit root; our approach is based on a modification of the Block-Bootstrap (BB) of Künsch (1989), and —for reasons to be apparent shortly— is termed "Continuous-Path Block-Bootstrap (CBB)". To motivate the CBB, let us give an illustration demonstrating the failure of the BB under the presence of a unit root.

Figure 1(a) shows a plot of (the natural logarithm of) the S&P 500 stock series index recorded annually from year 1871 to year 1988, while Figure 1(b) shows a realization of a BB pseudo replication of the S&P 500 series using block size 20. It is obvious visually that the bootstrap series is quite dissimilar to the original series, the most striking difference being the presence of strong discontinuities (of the 'jump' type) in the bootstrap series that –not surprisingly- occur every 20 time units, i.e., where the independent bootstrap blocks join.

Figure 1(c) suggests a way to fix this problem by forcing the bootstrap sample path to be continuous. A simple way to do this is to *shift* each of the bootstrap blocks up or down with the goal of ensuring (i) the bootstrap series starts off at the same point as the original series, and that (ii) the bootstrap sample path is continuous. Notably, the bootstrap blocks used in Figure 1(c) are the exact same blocks featuring in Figure 1(b).

At least as far as visual inspection of the plot can discern, the series in Figure 1(c) could just as well have been generated by the same probability mechanism that generated the original S&P 500 series. In other words, it is plausible that a bootstrap algorithm generating series such as the one in Figure 1(c) would be successful in mimicking important features of the original process; thus, the "Continuous-Path Block-Bootstrap" of Figure 1(c) is expected to 'work' in this case.

Of course, the actual yearly S&P 500 data are in discrete time, and talking about continuity is -strictly speaking- inappropriate. Nevertheless, an underlying continuous-time model may always be thought to exist, and the idea of continuity of sample paths is powerful and intuitive; hence the name "Continuous-Path Block-Bootstrap" (CBB for short) for our discrete-time methodology as well. The CBB is described in detail in the next Section, and some of its key properties are proven.

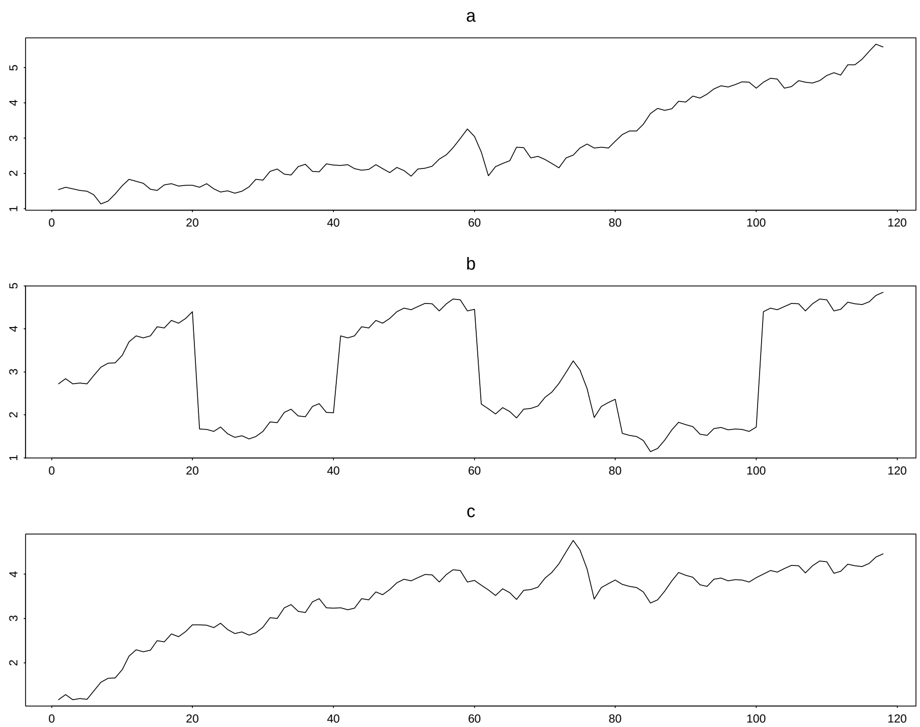

Figure 1: Plot of the natural logarithm of the S&P500 stock index series (a), of a BB realization (b) and of a CBB realization (c) with blocksize 20.

2. THE CONTINUOUS-PATH BLOCK-BOOTSTRAP (CBB)

Before introducing the Continuous-Path Block-Bootstrap (CBB) Method we review Künsch's (1989) Block-Bootstrap (BB). The BB algorithm is carried out conditionally on the original data $\{X(1), X(2), \ldots, X(n)\}$, and thus implicitly defines a bootstrap probability mechanism denoted by $P^\star$ that is capable of generating bootstrap pseudo-series of the type $\{X^\star(t), t = 1, 2, \ldots\}$.

Block-Bootstrap (BB) algorithm:

1. First chose a positive integer $b(< n)$, and let $i_0, i_1, \ldots, i_{k-1}$ be drawn i.i.d. with distribution uniform on the set $\{1, 2, \ldots, n-b+1\}$; here we take $k = [n/b]$, where $[\cdot]$ denotes the integer part, although different choices for k are also possible. The BB constructs a bootstrap pseudo-series $X^\star(1), X^\star(2), \ldots, X^\star(l)$, where $l = kb$, as follows.

2. For $m = 0, 1, \ldots, k-1$, let

$$X^\star(mb + j) := X(i_m + j - 1) \qquad \text{for } j = 1, 2, \ldots, b.$$

The Continuous-Path Block-Bootstrap (CBB) algorithm is now defined in the following three steps below. As before, the algorithm is carried out conditionally on the original data $\{X(1), X(2), \ldots, X(n)\}$, and implicitly defines a bootstrap probability mechanism denoted by P^* that is capable of generating bootstrap pseudo-series of the type $\{X^*(t), t = 1, 2, \ldots\}$. In the following we denote quantities with respect to P^* with an asterisk *.

Continuous-Path block-bootstrap (CBB) algorithm:

1. First calculate the centered residuals

$$\widehat{U}(t) = X(t) - X(t-1) \; - \; \frac{1}{n-1} \sum_{t=2}^{n} (X(t) - X(t-1))$$

   for $t = 2, 3, \ldots, n$. Attention now focuses on the new variables $\widetilde{X}(t)$ defined as follows:

$$\widetilde{X}(t) = \begin{cases} X(1) & \text{for } t = 1 \\ X(1) + \sum_{j=2}^{t} \widehat{U}(j) & \text{for } t = 2, 3, \ldots, n. \end{cases}$$

2. Chose a positive integer $b (< n)$, and let $i_0, i_1, \ldots, i_{k-1}$ be drawn i.i.d. with distribution uniform on the set $\{1, 2, \ldots, n-b\}$; here, we take $k = [n/b]$ as before. The CBB constructs a bootstrap pseudo-series $X^*(1), \ldots, X^*(l)$, where $l = kb$, as follows.

3. Construction of the first bootstrap block. Let

$$X^*(j) := X(1) + [\widetilde{X}(i_0 + j - 1) - \widetilde{X}(i_0)]$$

   for $j = 1, \ldots, b$. To elaborate:

$$X^*(1) := X(1)$$

$$X^*(2) := X^*(1) + [\widetilde{X}(i_0 + 1) - \widetilde{X}(i_0)]$$

$$X^*(3) := X^*(1) + [\widetilde{X}(i_0 + 2) - \widetilde{X}(i_0)]$$

$$\vdots$$

$$X^*(b) := X^*(1) + [\widetilde{X}(i_0 + b - 1) - \widetilde{X}(i_0)].$$

4. Construction of the $(m+1)$-th bootstrap block from the m-th block for $m = 1, \ldots, k-1$. Let

$$X^*(mb + j) := X^*(mb) + [\widetilde{X}(i_m + j) - \widetilde{X}(i_m)]$$

for $j = 1, \ldots, b$. To elaborate:

$$X^*(mb+1) := X^*(mb) + [\widetilde{X}(i_m+1) - \widetilde{X}(i_m)]$$

$$X^*(mb+2) := X^*(mb) + [\widetilde{X}(i_m+2) - \widetilde{X}(i_m)]$$

$$\vdots$$

$$X^*(mb+b) := X^*(mb) + [\widetilde{X}(i_m+b) - \widetilde{X}(i_m)].$$

An intuitive way to understand the CBB construction is based on the discussion regarding Figure 1(c) in the Introduction and goes as follows: (i) construct a BB pseudo-series $\{X^*(t), t = 1, 2, \ldots\}$ based on blocks of size equal to $b+1$ from the series $\widetilde{X}(t)$; (ii) shift the first block (of size $b+1$) by an amount selected such that the bootstrap series starts off at the same point as the original series; (iii) shift the second BB block (of size $b+1$) by another amount selected such that the first observation of this new bootstrap block matches exactly the last observation of the previous bootstrap block; (iv) join the two blocks but delete the last observation of the previous bootstrap block from the bootstrap series; (v) repeat parts (iii) and (iv) until all the generated BB blocks are used up.

Note that the CBB is applied to $\{\widetilde{X}(t)\}$ and not to $\{X(t)\}$. The reason is that although the $X(t)$ series is produced via the zero mean innovations $U(t)$, the observed finite-sample realization of the innovations will likely have nonzero (sample) mean; this discrepancy has an important effect on the bootstrap distribution effectively leading to a random walk with *drift* in the bootstrap world. Fortunately, there is an easy fix-up by recentering the innovations; a similar necessity for residual centering has been recommended early on even in regular linear regression —see Freedman (1981).

Note that a CBB series using block size b is associated to a BB construction with block size $b+1$. This phenomenon is only due to the fact that we are dealing with discrete-time processes; it would not occur in a continuous-time setting. The reason for this is our step (iv) above: although we are effecting the matching of the first observation of a new bootstrap block to the last observation of the previous bootstrap block, it does not seem advisable to leave both occurrences of this common (matched) value to exist side-by-side; one of the two must be deleted as step (iv) suggests.

3. ESTIMATION OF THE UNIT ROOT DISTRIBUTION

In this section the properties of the CBB in estimating the distribution of the first order autoregressive coefficient in the presence of a unit root are considered. Recall that based on the observations $\{X(1), X(2), \ldots, X(n)\}$ a common estimator of the first order autoregressive coefficient ρ in (1) is given by

$$\hat{\rho} \;=\; \frac{\sum_{t=2}^{T} X(t)X(t-1)}{\sum_{t=2}^{n} X^2(t-1)}. \qquad (3)$$

The CBB version of ρ is given by

$$\hat{\rho}^* = \frac{\sum_{t=2}^{l} X^*(t)X^*(t-1)}{\sum_{t=2}^{l} X^{*^2}(t-1)} \qquad (4)$$

and the distribution of the statistic $l(\hat{\rho}^*-1)$ is used to estimate the distribution of $n(\hat{\rho}-1)$

Consider first the basic random walk case. For this case the following result can be established.

Theorem 1 *Let $X(t) = X(t-1) + \varepsilon(t)$, $t = 1, 2, \ldots$ where $X(0) = 0$, $\varepsilon(t) \sim IID(0, \sigma^2)$ and $E(\varepsilon^4(1)) < \infty$. If $b \to \infty$ as $n \to \infty$ such that $b/\sqrt{n} \to 0$ then*

$$\sup_{x \in \mathbf{R}} \left| P^*\Big(l(\hat{\rho}^* - 1) \leq x\Big) - P\Big(n(\hat{\rho} - 1) \leq x\Big)\right| \to 0$$

in probability.

The asymptotic validity of the CBB for the general case where the stationary process $\{U(t)\}$ satisfies Assumption A is established in the following theorem which is our main result.

Theorem 2 *Let $X(t) = X(t-1) + U(t)$, $t = 1, 2, \ldots$ where $\{U(t), t \in \mathbb{Z}\}$ satisfies Assumption A and $E(\varepsilon^4(1)) < \infty$. If $b \to \infty$ as $n \to \infty$ such that $b/\sqrt{n} \to 0$ then*

$$\sup_{x \in \mathbf{R}} \left| P^*\Big(l(\hat{\rho}^* - 1) \leq x\Big) - P\Big(n(\hat{\rho} - 1) \leq x\Big)\right| \to 0$$

in probability.

It is noteworthy that, under the assumptions of Theorem 2, the asymptotic distribution of $\hat{\rho}$ has a complicated form, depending on many unknown parameters such as the infinite sum $\sum_{j=0}^{\infty}\psi_j$; see Hamilton (1994) or Fuller (1996). The CBB effortlessly achieves the required distribution estimation, and provides an attractive alternative as compared to the asymptotic distribution with estimated parameters. Notably, estimation of the sum $\sum_{j=0}^{\infty}\psi_j$ is tantamount to estimating the spectral density of the differenced series (evaluated at the origin) which is a highly nontrivial problem.

4. PROOFS

Proof of Theorem 1: Recall that $\widehat{U}(i_m+s)=\widetilde{X}(i_m+s)-\widetilde{X}(i_m+s-1)$. It is easily seen that for $t=2,3,\ldots,l$

$$X^*(t)=X(1)\;+\sum_{m=0}^{[(t-1)/b]}\sum_{s=1}^{B}\widehat{U}(i_m+s) \tag{5}$$

where $B=\min\{b-\delta_{0,m},t-mb-\delta_{0,m}\}$, $\delta_{i,j}$ is Kronecker's delta, i.e., $\delta_{i,j}=1$ if $i=j$ and zero else. Alternatively, we can write

$$X^*(t)=\begin{cases} X(1) & \text{for}\quad t=1\\ X^*(t-1)\;+\;\widehat{U}(i_m+s) & \text{for}\quad t=2,3,\ldots,l\end{cases} \tag{6}$$

where $m=[(t-1)/b]$ and $s=t-mb-\delta_{0,m}$.

Now, assume without loss of generality that $\sigma^2=1$. Furthermore, set $\widehat{U}(i_m+s)\equiv e(i_m+s)$ and note that in the random walk case considered here we have by the definition of $U(i_m+s)$ that

$$e(i_m+s)\quad=\quad\varepsilon(i_m+s)-\frac{1}{n-1}\sum_{t=2}^{n-1}\varepsilon(t). \tag{7}$$

By the centering of the $\widehat{U}(t)$'s we have

$$\begin{aligned} E^*(e(i_m+s)) &= \frac{1}{n-b}\sum_{t=1}^{n-b}\varepsilon(t+s)+\frac{1}{n-1}\sum_{t=2}^{n}\varepsilon(t)\\ &= O_P(b^{1/2}n^{-1}).\end{aligned}$$

Substituting expression (6) we get

$$l\,(\hat{\rho}^*-1)\quad=\quad\frac{l\displaystyle\sum_{t=2}^{l}\Big(X^*(t)-X^*(t-1)\Big)X^*(t-1)}{\displaystyle\sum_{t=2}^{l}X^{*^2}(t-1)}$$

$$= \Big(l^{-2}\sum_{t=2}^{l} X^{*^2}(t-1)\Big)^{-1}\frac{1}{l}\Big[\sum_{s=1}^{b-1} e(i_0+s)X^*(s) + \sum_{m=1}^{k-1}\sum_{s=1}^{b} e(i_m+s)X^*(mb+s-1)\Big].$$

Applying (5) we get after some simple algebra that

$$\begin{aligned}
&\frac{1}{l}\Big[\sum_{s=1}^{b-1} e(i_0+s)X^*(s) + \sum_{m=1}^{k-1}\sum_{s=1}^{b} e(i_m+s)X^*(mb+s-1)\Big] \\
&= \frac{1}{l}\varepsilon(1)\Big[\sum_{s=1}^{b-1} e(i_0+s) + \sum_{m=1}^{k-1}\sum_{s=1}^{b} e(i_m+s)\Big] \\
&\quad + \frac{1}{2l}\Big[\Big(\sum_{s=1}^{b-1} e(i_0+s) + \sum_{m=1}^{k-1}\sum_{s=1}^{b} e(i_m+s)\Big)^2 \qquad (8)\\
&\quad - \Big(\sum_{s=1}^{b-1} e^2(i_0+s) + \sum_{m=1}^{k-1}\sum_{s=1}^{b} e^2(i_m+s)\Big)\Big] \\
&= \frac{1}{l}\varepsilon(1)\Big[\sum_{s=1}^{b-1} e(i_0+s) + \sum_{m=1}^{k-1}\sum_{s=1}^{b} e(i_m+s)\Big] \\
&\quad - \frac{1}{2}\Big[\frac{1}{l}\Big(\sum_{s=1}^{b-1} e^2(i_0+s) + \sum_{m=1}^{k-1}\sum_{s=1}^{b} e^2(i_m+s)\Big) - 1\Big] \\
&\quad + \frac{1}{2}\Big[\Big(\frac{1}{\sqrt{l}}\Big(\sum_{s=1}^{b-1} e(i_0+s) + \sum_{m=1}^{k-1}\sum_{s=1}^{b} e(i_m+s)\Big)\Big)^2 - 1\Big]. \qquad (9)
\end{aligned}$$

Thus

$$\begin{aligned}
l(\hat{\rho}^*-1) &= \frac{1}{2}\Big(l^{-2}\sum_{t=2}^{l} X^{*^2}(t-1)\Big)^{-1}\Big\{\Big[\frac{1}{\sqrt{l}}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} e(i_m+s)\Big]^2 - 1\Big\} \\
&\quad - \frac{1}{2}\Big(l^{-2}\sum_{t=2}^{l} X^{*^2}(t-1)\Big)^{-1}\Big[\frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} e^2(i_m+s) - 1\Big] \\
&\quad + \frac{1}{2}\Big(l^{-2}\sum_{t=2}^{l} X^{*^2}(t-1)\Big)^{-1}\Big[\frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} e(i_m+s)\varepsilon(1)\Big]. \qquad (10)
\end{aligned}$$

Because of (10) and in order to establish the desired result we have to show that the following three assertions are true:

$$T_{1,n} \equiv \frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} e(i_m+s)\varepsilon(1) = o_{P^*}(1), \qquad (11)$$

$$\Big|T_{2,n}\Big| \equiv \Big|\frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} e^2(i_m+s)-1\Big| = o_{P^*}(1) \qquad (12)$$

and

$$\Big(\frac{1}{l^2}\sum_{t=2}^{l} X^{*^2}(t-1),\ \frac{1}{\sqrt{l}}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} e(i_m+s)\Big) \overset{d^*}{\to} \Big(G_1,\ G_2\Big) \qquad (13)$$

in probability, where

$$G_1 = \sum_{i=1}^{\infty}\gamma_i^2 U_i^2, \qquad G_2 = \sum_{i=1}^{\infty}\sqrt{2}\gamma_i U_i,$$

$\gamma_i = (-1)^{i+1}2/[(2i-1)\pi]$ and $\{U_i\}_{i=1,2,\ldots}$ is a sequence of independent standard Gaussian variables. The assertion of the Theorem follows then because under validity of (11) to (13) and by Slutsky's theorem we get

$$d_K\Big(\mathcal{L}\Big\{l(\hat\rho^*-1)|X_1,X_2,\ldots,X_n\Big\},\mathcal{L}\Big\{(2G_1)^{-1}(G_2^2-1)\Big\}\Big) \to 0$$

in probability, which is the asymptotic distribution of $n(\hat\rho-1)$; cf. Fuller (1996). Here d_K denotes Kolmogorov's distance $d_K(\mathcal{P},\mathcal{Q}) = \sup_{x\in\mathbf{R}}|\mathcal{P}(X\leq x)-\mathcal{Q}(X\leq x)|$ between probability measures $\mathcal{P}$ and $\mathcal{Q}$.

We proceed to show that (11) to (13) are true.

To see (11) note that

$$\begin{aligned} T_{1,n} &= \frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} e(i_m+s)\varepsilon(1) \\ &= \varepsilon(1)\frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} \varepsilon(i_m+s) + O_P(n^{-1/2}) \\ &= \widetilde{T}_{1,n} + O_P(n^{-1/2}). \end{aligned}$$

Note that $\widetilde{T}_{1,n}$ is a block bootstrap estimator of the mean $E(\varepsilon_t)$ based on the i.i.d. sample $\varepsilon(2),\varepsilon(3),\ldots,\varepsilon(n)$. Thus assertion (11) follows because

$$E^*(\widetilde{T}_{1,n}) \to 0 \quad \text{and} \quad Var^*(\widetilde{T}_{1,n}) = O_P(l^{-1}).$$

To establish (12) verify first using

$$E^*\Big(l^{-1}\sum_{m=1}^{k-1}\sum_{s=1}^{b}\varepsilon(i_m+s)\Big) = O_P((n-b)^{-1/2}),$$

and

$$E^*\Big(l^{-1}\sum_{m=1}^{k-1}\sum_{s=1}^{b}\varepsilon(i_m+s)\Big)^2 = O_P((n-b)^{-1})$$

that

$$\frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} e^2(i_m+s) = \frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}} \varepsilon^2(i_m+s) + O_{P^*}(n^{-1/2}(n-b)^{-1/2}).$$

The desired result follows then by recognizing that the first term on the right hand side of the above equation is a block bootstrap estimator of $E(\varepsilon^2(1)) = 1$ based on blocks from the i.i.d. sequence $\varepsilon(2), \varepsilon(2), \ldots, \varepsilon(n)$.

Consider (13). Let $\mathbf{e}_{i_0} = (e(i_0+1), e(i_0+2), \ldots, e(i_0+b-1))$, $\mathbf{e}_{i_m} = (e(i_m+1), e(i_m+2), \ldots, e(i_m+b))$ for $m = 1, 2, \ldots, k-2$ and $\mathbf{e}_{i_{k-1}} = (e(i_{k-1}+1), e(i_{k-1}+2), \ldots, e(i_{k-1}+b-1))$. Denote by $\mathbf{e}$ be the l-dimensional random vector

$$\mathbf{e} = (\varepsilon(1),\ \mathbf{e}_{i_0},\ \mathbf{e}_{i_1},\ \ldots,\ \mathbf{e}_{i_{k-1}})' \tag{14}$$

and by $\mathbf{A}_l$ the $(l-1)\times(l-1)$ matrix given by $\mathbf{A}_l = \sum_{s=1}^{l-1}\mathbf{I}_s$ where $\mathbf{I}_s$ is the $(l-1)\times(l-1)$ matrix with (i,j)th element equal to one if $1 \le i, j \le s$ and zero else. Note that $\mathbf{A}_l = \mathbf{Q}_l\mathbf{\Lambda}_l\mathbf{Q}_l'$ where the (i,j)th element of the orthogonal matrix $\mathbf{Q}_l$ is given by $q_{i,j} = 2(2l-1)^{-1/2}\cos[(4l-2)^{-1}(2j-1)(2i-1)\pi]$ and the i-th element of the diagonal matrix $\mathbf{\Lambda}_l = \mathrm{diag}(\lambda_{1,l}, \lambda_{2,l}, \ldots, \lambda_{l-1,l})$ is given by $\lambda_{i,l} = 0.25\sec^2[(l-i)(2l-1)\pi]$; cf. Fuller (1996). Using this decomposition of the matrix $\mathbf{A}_l$ we have

$$\begin{aligned}\frac{1}{l^2}\sum_{t=2}^{l} X^{*^2}(t-1) &= \frac{1}{l^2}\mathbf{e}'\mathbf{A}_l\,\mathbf{e} \\ &= \frac{1}{l^2}\sum_{i=1}^{l-1}\lambda_{i,l}\,U_i^{*^2}\end{aligned}$$

where the random variable U_i^* is given by

$$U_i^* \;=\; q_{i1}\varepsilon(1) \;+\; \sum_{m=0}^{k-1} V_{i,m}^*,$$

and

$$\begin{aligned}V_{i,m}^* &= \sum_{s=1}^{b-\delta_{0,m}-\delta_{k-1,m}} q_{i,mb+s+\delta_{0,m}}\,e(i_m+s) \\ &= \sum_{s=1}^{b-\delta_{0,m}-\delta_{k-1,m}} q_{i,mb+s+\delta_{0,m}}\,\varepsilon(i_m+s) + O_{P^*}(b^{1/2}k^{-1/2}n^{-1/2}).\end{aligned}$$

Therefore, $E^*(V_{i,m}^*) = O_P(b^{1/2}k^{-1/2}n^{-1/2})$ for $i = 1, 2, \ldots, l-1$ and $m = 0, 1, 2, \ldots, k-1$ and $E^*(U_i^*) = O_p(k^{-1/2})$.

Let $\mathbf{J}_l$ be the $(l-1)$-dimensional vector $\mathbf{J}_l = (1, 1, \ldots, 1)'$. Using the fact that $(U_1^*, U_2^*, \ldots, U_{l-1}^*)' = \mathbf{Q}_l \mathbf{e}$ we have

$$\begin{aligned} l^{-1/2} \sum_{m=0}^{k-1} \sum_{s=1}^{b-\delta_{0,m}} e(i_m + s) &= l^{-1/2} \mathbf{J}_l' \mathbf{e} \\ &= \sum_{i=1}^{l-1} k_{i,l} U_i^* \end{aligned}$$

where $k_{i,l} = l^{-1/2} \mathbf{J}_l' \mathbf{Q}_l^{-1} \mathbf{1}_i$ and $\mathbf{1}_i$ is the $(l-1) \times 1$ vector with one in the *ith* position and zero elsewhere. Thus we have for the term on the left hand side of (13) that

$$(l^{-2} \sum_{t=2}^{l-1} X^{*^2}(t-1),\ l^{-1/2} \sum_{m=0}^{k-1} \sum_{s=1}^{b-\delta_{0,m}} e(i_m + s)) = (\sum_{i=1}^{l-1} l^{-2} \lambda_{i,l} U_i^*,\ \sum_{i=1}^{l-1} k_{i,l} U_i^*).$$

To establish the desired asymptotic distribution consider first the asymptotic behavior of the bootstrap variable U_i^*. Since

$$\begin{aligned} E^*\Big(\sum_{s=1}^{b} q_{i,mb+s} e(i_m + s)\Big)^2 &= \frac{1}{n-b} \sum_{t=1}^{n-b} \sum_{s_1,s_2=1}^{b} q_{i,mb+s_1} q_{i,mb+s_2} e(t+s_1) e(t+s_2) \\ &= \frac{1}{n-b} \sum_{t=1}^{n-b} \sum_{s_1,s_2=1}^{b} q_{i,mb+s_1} q_{i,mb+s_2} \varepsilon(t+s_1) \varepsilon(t+s_2) \\ &\quad + O_P(bk^{-1}(n-b)^{-1/2} n^{-1/2}) \end{aligned}$$

we get using $E^*(V_{i,m}^*) = O_P(b^{1/2} k^{-1/2} n^{-1/2})$ that

$$\begin{aligned} Var^*(U_i^*) &= \frac{1}{n-b} \sum_{t=1}^{n-b} \sum_{m=0}^{k-1} \sum_{s_1,s_2=1}^{b-\delta_{0,m}-\delta_{k-1,m}} q_{i,mb+s_1+\delta_{0,m}} q_{i,mb+s_2+\delta_{0,m}} \\ &\quad \times \varepsilon(t+s_1) \varepsilon(t+s_2) + O_p(b(n-b)^{-1/2} n^{-1/2}) \\ &= Var(\varepsilon_1) \sum_{r=2}^{l-1} q_{i,r}^2 + o_p(1) \\ &= Var(\varepsilon_1) + o_p(1). \end{aligned}$$

The last equality above follows since $b/n \to 0$, $\sum_{r=1}^{l-1} q_{i,r}^2 = 1$ and $q_{i,r} = O(l^{-1/2})$ uniformly in r. Furthermore, and because $E(U_i^*) = O_P(k^{-1/2})$ we have $Cov^*(U_i^*, U_j^*) = E^*(U_i^* U_j^*) + O_P(k^{-1})$ and by the independence of the

$V_{i,m}^*$

$$
\begin{aligned}
E^*(U_i^* U_j^*) &= \sum_{m=0}^{k-1} E^*(V_{i,m}^* V_{j,m}^*) + o_P(1) \\
&= \frac{1}{n-b} \sum_{t=1}^{n-b} \sum_{m=0}^{k-1} \sum_{s_1,s_2=1}^{b-\delta_{0,m}-\delta_{k-1,m}} q_{i,mb+s_1+\delta_{0,m}} q_{j,mb+s_2+\delta_{0,m}} \\
&\quad \times \varepsilon(t+s_1)\varepsilon(t+s_2) + O_P(b^{3/2} n^{-3/2}).
\end{aligned}
$$

Using

$$
(n-b)^{-1} \sum_{t=1}^{n-b} \varepsilon^2(t+s) \rightarrow Var(\varepsilon(1))
$$

and

$$
(n-b)^{-1} \sum_{t=1}^{n-b} \varepsilon(t+s_1)\varepsilon(t+s_2) = O_P((n-b)^{-1/2})
$$

for $s_1 \neq s_2$ uniformly in s_1 and s_2, we get by the property $\sum_{r=1}^{l-1} q_{i,r} q_{j,r} = 0$ for $i \neq j$ that

$$
\begin{aligned}
E^*(U_i^* U_j^*) &= Var(\varepsilon_1) \sum_{r=2}^{l} q_{i,r} q_{j,r} + O_P(b(n-b)^{-1/2}) \\
&= O(l^{-1}) + O_P(b(n-b)^{-1/2}).
\end{aligned}
$$

Thus, $Cov^*(U_i^*, U_j^*) \rightarrow 0$ in probability for $i \neq j$.

Consider next the asymptotic distribution of the U_i^*'s and recall that $U_i^* = \sum_{m=0}^{k-1} V_{i,m}^* + o_{P^*}(1)$ where the $V_{i,m}^*$ are independent (but not identically distributed) zero mean random variables. Applying a CLT for triangular arrays of independent random variables (see Corollary of Serfling (1981, p. 32)) we can show that $d_K(\mathcal{L}(U_i^*), \mathcal{L}(Z)) \rightarrow 0$ in probability as $n \rightarrow \infty$, where Z denotes a standard Gaussian distributed random variables. To elaborate and because $|\sum_{m=0}^{k-1} Var^*(V_{i,m}^*) - 1| = o_P(1)$ it suffices to show that $\sum_{m=0}^{k-1} E^*|V_{i,m}^*|^{\nu} = o_P(1)$ for some $\nu > 2$. This, however, follows since

$$
\begin{aligned}
E^*|V_{i,m}^*|^{\nu} &= \frac{1}{n-b} \sum_{t=1}^{n-b} \Big| \sum_{s=1}^{b} q_{i,mb+s+\delta_{0,m}} \varepsilon(t+s) \Big|^{\nu} + o(1) \\
&= O_P(b^{\nu/2} l^{-\nu/2})
\end{aligned}
$$

and therefore, $\sum_{m=0}^{k-1} E^*|V_{i,m}^*|^{\nu/2} = O_P(k^{1-\nu/2}) \rightarrow 0$. Therefore, and because $Cov^*(U_i^*, U_j^*) \rightarrow 0$ in probability we get that for $0 < N < l$ fixed

$$
\Big(U_1^*, U_2^*, \ldots, U_N^*\Big)' \stackrel{d^*}{\rightarrow} \Big(U_1, U_2, \ldots, U_N\Big)' \tag{15}
$$

in probability as $n \to \infty$, where $(U_1, U_2, \ldots, U_N)'$ is a random vector having a N-dimensional Gaussian $N(\mathbf{0}, I_N)$ distribution and $\mathbf{I}_N$ is the $N \times N$ unity matrix.

The rest of the proof proceeds along the lines of the proof of Theorem 10.1.1 of Fuller (1996, p. 550). Briefly, since $\lim_{l\to\infty} \sum_{i=1}^{l-1} |l^{-2}\lambda_{i,l} - \gamma_i^2| = 0$ we get that

$$\begin{aligned}(\sum_{i=1}^{l-1} \lambda_{i,l} U_i^{*^2}, \sum_{i=1}^{l-1} k_{i,l} U_i^*) &= (\sum_{i=1}^{N} \gamma_i U_i^{*^2}, \sum_{i=1}^{N} \sqrt{2}\gamma_i U_i^*) \\ &\quad +(\sum_{i=N+1}^{l-1} \gamma_i U_i^{*^2}, \sum_{i=N+1}^{l-1} \sqrt{2}\gamma_i U_i^*) + o_{P^*}(1) \\ &= M_{1,l} + M_{2,l} + o_{P^*}(1)\end{aligned}$$

with an obvious notation for $M_{1,l}$ and $M_{2,l}$. Now, by the summability of the sequence $\{\gamma_i^2\}$ we have that $M_{2,L} = o_P(1)$ as $N \to \infty$ uniformly in l. From this, equation (15) and Lemma 6.3.1 of Fuller (1996) we then get

$$\Big(\sum_{i=1}^{l-1} \lambda_{i,l} U_i^{*^2}, \sum_{i=1}^{l-1} k_{i,l} U_i^*\Big) \xrightarrow{d^*} \Big(G_1, G_2\Big)$$

in probability which concludes the proof of (13) and of the theorem. □

Proof of Theorem 2: We only give a sketch of the proof. We first show that

$$l^{-2} \sum_{t=2}^{l} X^{*^2}(t-1) = \psi^2(1) l^{-2} \sum_{t=2}^{l} \Big(\sum_{m=0}^{[(t-1)/b]} \sum_{s=1}^{M} e(i_m + s)\Big)^2 + o_{P^*}(1) \qquad (16)$$

and

$$\begin{aligned}\frac{1}{l}\sum_{t=2}^{l} X^*(t-1)(X^*(t) - X^*(t-1)) &= \frac{\psi^2(1)}{2}\Big[\Big(\frac{1}{\sqrt{l}} \sum_{m=0}^{k-1} \sum_{s=1}^{b-\delta_{0,m}} e(i_m+s)\Big)^2 \\ &\quad -\sigma^2 \sum_{j=0}^{\infty} \psi_j^2/\psi^2(1)\Big] + o_{P^*}(1), \qquad (17)\end{aligned}$$

where $e(i_m + s)$ is defined in (7). To see this let $C_\psi = \sum_{j=0}^{\infty} \psi_j$ and note that by assumption A and using a Beveridge-Nelson decomposition (cf. Hamilton (1994), Proposition 17.2), $\sum_{t=2}^{n} U(t)$ can be written in the form

$$\sum_{t=2}^{n} U(t) = C_\psi \sum_{t=2}^{n} \varepsilon(t) + \eta(n) - \eta(1)$$

where $\eta(t)=\sum_{j=0}^{\infty}\alpha_j\varepsilon(t-j)$ and $\alpha_j=-\sum_{i=j+1}^{\infty}\psi_i$. Since $\sum_{j=0}^{\infty}|\alpha_j|<\infty$ we have

$$(n-1)^{-1}\sum_{t=2}^{n}U(t)=C_{\psi}(n-1)^{-1}\sum_{t=2}^{n}\varepsilon(t)+O_P(n^{-1}). \tag{18}$$

Furthermore, a same type of decomposition can be applied for the observations within each bootstrap block, i.e.,

$$\sum_{s=1}^{b}U(i_m+s)=C_{\psi}\sum_{s=1}^{b}\varepsilon(i_m+s)+\eta(i_m+b)-\eta(i_m). \tag{19}$$

Now, using (5), (18) and (19) we get

$$X^*(t)=X(1)+\sum_{m=0}^{[(t-1)/b]}\Big(C_{\psi}\sum_{s=1}^{B}e(i_m+s)+(\eta(i_m+B)-\eta(i_m))+O_P(n^{-1})\Big) \tag{20}$$

for $t=2,3,\ldots,l$ where $B=\min\{b-\delta_{0,m},t-mb-\delta_{0,m}\}$. Substituting the above expression for $X^*(t)$ in $l^{-2}\sum_{t=2}^{l}X^{*^2}(t)$, assertion (16) follows after some straightforward calculations.

To see (17) note first that using arguments identical to those in (8) we have

$$\begin{aligned}\frac{1}{l}\sum_{t=2}^{l}X^*(t-1)(X^*(t)-X^*(t-1)) &= \frac{1}{l}U(1)\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}}\widehat{U}(i_m+s)\\ &\quad -\frac{1}{2}\Big(\frac{1}{l}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}}\widehat{U}^2(i_m+s)-\sigma^2\sum_{j=0}^{\infty}\psi_j^2\Big)\\ &\quad +\frac{1}{2}\Big[\Big(\frac{1}{\sqrt{l}}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}}\widehat{U}(i_m+s)\Big)^2-\sigma^2\sum_{j=0}^{\infty}\psi_j^2\Big]\\ &= \frac{1}{2}\Big[\Big(\frac{1}{\sqrt{l}}\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}}\widehat{U}(i_m+s)\Big)^2-\sigma^2\sum_{j=0}^{\infty}\psi_j^2\Big]+o_{P^*}(1).\end{aligned}$$

(17) follows then using (20) and

$$\frac{1}{l}\Big(\sum_{m=0}^{k-1}\sum_{s=1}^{b-\delta_{0,m}}\widehat{U}(i_m+s)\Big)^2 = \frac{1}{l}X^{*^2}(l)+o_p(1).$$

Under validity of (16) and (17) and along the same lines as in the proof of Theorem 1 it follows that

$$\begin{aligned}&\Big(l^{-2}\sum_{t=2}^{l}X^{*^2}(t-1),\ \ l^{-1}\sum_{t=2}^{l}X^*(t-1)(X^*(t)-X^*(t-1))\Big)\\ &\qquad\stackrel{d^*}{\to}\ \Big(\sigma^2\psi^2(1)G_1,\ \ 2^{-1}\sigma^2\psi^2(1)[G_2^2-\sum_{j=0}^{\infty}\psi_j^2/\psi^2(1)]\Big)\end{aligned}$$

in probability. Now, since

$$n(\hat{\rho}-1) \;\to\; \Big(2G_1\Big)^{-1}\Big(G_2^2 - \sum_{j=0}^{\infty}\psi_j^2/\psi^2(1)\Big)$$

in distribution as $n \to \infty$ (cf. Fuller (1996), Hamilton (1994)), the proof of the theorem is concluded by applying Slutsky's theorem. □

REFERENCES

1. D. A. Dickey and W. A. Fuller (1979), Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* **74**, 427-431.

2. N. Ferretti and J. Romo (1996), Unit root bootstrap tests for AR(1) models, *Biometrika* **83**, 849-860.

3. D. A. Freedman (1981), Bootstrapping regression models, *Annals of Statistics* **9**, 1218-1228.

4. W. Fuller (1996), *Introduction to Statistical Time Series*, (2nd Ed.), John Wiley, New York.

5. J. D. Hamilton (1994), *Time Series Analysis*, Princeton University Press, Princeton, New Jersey.

6. H. R. Künsch (1989), The jackknife and the bootstrap for general stationary observations, *Annals of Statistics* **17**, 1217-1241.

7. P. C. B. Phillips and P. Perron (1988), Testing for a unit root in time series regression, *Biometrika*, **75**, 335-346.

8. D. N. Politis and J. P. Romano (1994), Large sample confidence regions based on subsamples under minimal assumptions, *Annals of Statistics* **22**, 2031-2050.

9. D. N. Politis, J. P. Romano and M. Wolf (1999), *Subsampling*, Springer, New York.

10. J. P. Romano and M. Wolf (1998), Subsampling confidence intervals for the autoregressive root, Technical Report, Department of Statistics, Stanford University.

11. R. Serfling (1980), *Approximation Theorems of Mathematical Statistics.* John Wiley, New York.

12. J. H. Stock (1991), Confidence intervals for the largest autoregressive root in U. S. macroeconomic time series, *Journal of Monetary Economics* **28**, 435-459.