12-31-2015

Block bootstrap consistency under weak assumptions

Gray Calhoun
Iowa State University, gcalhoun@iastate.edu

Block bootstrap consistency under weak assumptions

Abstract
This paper weakens the size and moment conditions needed for typical block bootstrap methods (i.e. the moving blocks, circular blocks, and stationary bootstraps) to be valid for the sample mean of Near-Epoch-Dependent functions of mixing processes; they are consistent under the weakest conditions that ensure the original process obeys a Central Limit Theorem (those of de Jong, 1997, Econometric Theory). In doing so, this paper extends de Jong's method of proof, a blocking argument, to hold with random and unequal block lengths. This paper also proves that bootstrapped partial sums satisfy a Functional CLT under the same conditions.

Keywords
resampling, time series, near epoch dependence, functional central limit theorem

Disciplines
Economics

Block bootstrap consistency under weak assumptions

Gray Calhoun*

2015-12-31, commit 5032b34

Abstract

This paper weakens the size and moment conditions needed for typical block bootstrap methods (i.e. the Moving Blocks, Circular Blocks, and Stationary Bootstraps) to be valid for the sample mean of Near-Epoch-Dependent (NED) functions of mixing processes; they are consistent under the weakest conditions that ensure the original NED process obeys a Central Limit Theorem (those of De Jong, 1997, *Econometric Theory*). In doing so, this paper extends De Jong's method of proof, a blocking argument, to hold with random and unequal block lengths. This paper also proves that bootstrapped partial sums satisfy a Functional CLT under the same conditions.

JEL Classification: C12, C15

Keywords: Resampling, Time Series, Near Epoch Dependence, Functional Central Limit Theorem

*Economics Department, Iowa State University, Ames, IA 50011. Telephone: (515) 294-6271. Email: gcalhoun@iastate.edu. Web: http://gray.clhn.org. I would like to thank Helle Bunzel, Dimitris Politis, Robert Taylor, and three anonymous referees for their comments and feedback on earlier versions of this paper.

Block bootstraps, e.g. the Moving Blocks (Kunsch, 1989, and Liu and Singh, 1992), Circular Block (Politis and Romano, 1992), and Stationary Bootstraps (Politis and Romano, 1994), have become popular in Economics, partly because they do not require the researcher to make parametric assumptions about the data generating process. They are valid under general weak dependence and moment conditions. Some recent papers (Gonçalves and White, 2002, and Gonçalves and De Jong, 2003) relax the dependence and moment conditions of the original papers to fit with the Near-Epoch-Dependence (NED) assumptions commonly used in econometrics.[1,2] But these conditions are still stronger than required for a CLT to hold; De Jong (1997) has established the CLT under L_2-NED with smaller size and moment restrictions.[3] This paper shows that these block bootstrap methods consistently estimate the distribution of the sample mean under De Jong's (1997) assumptions and show that an FCLT holds as well.[4] It also relaxes Gonçalves and White's (2002) and Gonçalves and De Jong's (2003) requirement that the expected block length be $o(n^{1/2})$ to the original papers' requirement that it be $o(n)$.

The proof exploits the conditional independence of the blocks in each bootstrap. Each bootstrap proceeds by drawing blocks of M consecutive observations from the original time series, and then pasting these blocks together to create the new bootstrap time series. The Moving Blocks bootstrap does exactly that; the Circular Block bootstrap "wraps" the observations, so that (X_{n-1}, X_n, X_1, X_2), for example, is a possible block of length four (letting X_t denote the original time series). The Stationary Bootstrap wraps the observations and also draws M at random for each block; Politis and Romano (1994) suggest drawing M from the geometric distribution. As the name suggests, the series produced by the Stationary Bootstrap are strictly stationary, while those produced by the other methods are not. Although the Stationary Bootstrap was believed to be much less efficient than other block bootstrap methods due to results of Lahiri (1999), Nordman (2009) has shown that it is only slightly less efficient than the other block-bootstrap methods discussed in this paper, and

[1] Gonçalves and White (2002) show that these bootstrap methods can be applied to heterogeneous $L_{2+\delta}$-NED processes of size $-2(r-1)/(r-2)$ on a strong mixing sequence of size $-r(2+\delta)/(r-2)$, where $r > 2$ and $\delta > 0$, when the original series has uniformly bounded $3r$-moments. Gonçalves and De Jong (2003) relax these conditions to $L_{2+\delta}$-NED of size -1 and $r+\delta$ moments for the original series, and size $-(2+\delta)(r+\delta)/(r-2)$ for the underlying mixing series. Both papers require that the expected block length grow with n and be $o(n^{1/2})$. Gonçalves and Politis (2011) discuss these issues further.

[2] An array $\{X_{nt}\}$ is an L_ρ-NED process on a mixing array $\{V_{nt}\}$ if

$$\|X_{nt} - \mathrm{E}(X_{nt} \mid V_{n,t-m}, \dots, V_{n,t+m})\|_\rho \leq d_{nt} v_m \tag{1}$$

with $v_m \to 0$ as $m \to \infty$ and $\{d_{nt}\}$ an array of positive constants. It is of size $-\gamma$ if $v_m = O(m^{-\gamma-\delta})$ for all $\delta > 0$. Dropping the index "n" gives the series definition. Note that the underlying strong and uniform mixing arrays are not required to be stationary.

[3] De Jong (1997) proves that the CLT holds for averages of L_2-NED processes of size $-1/2$ on a strong mixing series of size $-r/(r-2)$, $r > 2$ and the original series having bounded r-moments.

[4] Radulović (1996) proves consistency for the Moving Blocks Bootstrap for any stationary strong mixing sequence that satisfies the CLT. This paper uses a similar method of proof to his, but also accommodates nonstationary sequences and the Stationary Bootstrap.

has efficiency identical to that of the non-overlapping block bootstrap. Consequently, there has been renewed interest in the Stationary Bootstrap since stationarity of the bootstrap samples can be a useful property in theoretical research. Kreiss and Paparoditis (2011a) provides a recent review of the bootstrap for time-series processes[5] and Gonçalves and Politis (2011) further discuss recent developments in block-bootstraps.

Theorem 1 presents the main result, asymptotic normality of the distribution of bootstrapped sums. This paper adopts the standard notation that E^*, var^*, etc. are the usual operators with respect to the probability measure induced by the bootstrap and will use explicit stochastic array notation for precision. Also note that all results are presented for the scalar case but generalize immediately to random vectors. All of the proofs are presented in the appendix; only proofs for the Stationary Bootstrap are presented, since proofs for the other methods are similar and easier to construct. All limits are taken as $n \to \infty$ unless otherwise noted and $\|\cdot\|_r$ denotes the L_r-norm.

Theorem 1. *Suppose the following conditions hold.*

1. *X_{nt} is L_2-NED of size $-1/2$ on an array $\{V_{nt}\}$ that is either strong mixing of size $-r/(r-2)$ or uniform mixing of size $-r/2(r-1)$, with $r > 2$. The NED magnitude indices are denoted $\{d_{nt}\}$.*

2. *The array $\mu_{nt} - \bar{\mu}_n$ is uniformly bounded and $\sum_{t=1}^{n}(\mu_{nt} - \bar{\mu}_n)^2 \to 0$, where $\mathrm{E}\, X_{nt} = \mu_{nt}$ and $\bar{\mu}_n = n^{-1}\sum_{t=1}^{n}\mu_{nt}$. Moreover, $\sqrt{n}\|\bar{X}_n - \bar{\mu}_n\|_2 \to \sigma > 0$, with $\bar{X}_n = (1/n)\sum_{t=1}^{n} X_{nt}$.*

3. *There exists an array of positive real numbers $\{c_{nt}\}$ such that $(X_{nt} - \mu_{nt})/c_{nt}$ is uniformly L_r-bounded and c_{nt} and d_{nt}/c_{nt} are uniformly bounded in n and t.*

4. *X^*_{nt} is generated by the Stationary Bootstrap with geometric block lengths with success probability p_n, $p_n = cn^{-a}$ and $a, c \in (0,1)$, or by the Moving or Circular Block bootstrap with block length M_n such that $M_n \sim n^a$ for $a \in (0,1)$. Let M_{ni} be the block length of the ith block, $i = 1, \ldots, J_n$, and define $K_{n0} = 0$ and $K_{nj} = \sum_{i=1}^{j} M_{ni}$. The last block, M_{n,J_n}, is defined as $K_{n,J_n} - K_{n,J_n-1}$, so $K_{n,J_n} = n$ a.s.*

Then $\hat{\sigma}^{*2} \to^p \sigma^2$,

$$\sup_x \left| \Pr^*\left[\sqrt{n}(\bar{X}^*_n - \mathrm{E}^*\,\bar{X}^*_n) \le x\right] - \Pr\left[\sqrt{n}(\bar{X}_n - \mathrm{E}\,\bar{X}_n) \le x\right]\right| \to^p 0, \tag{2}$$

and

$$\sup_x \left| \Pr^*\left[\sqrt{n}(\bar{X}^*_n - \mathrm{E}^*\,\bar{X}^*_n)/\hat{\sigma}^*_n \le x\right] - \Phi(x)\right| \to^p 0, \tag{3}$$

[5] Also see the discussion papers by Dahlhaus (2011), Gonçalves and Politis (2011), Horowitz (2011), Jentsch and Mammen (2011), and Kreiss and Paparoditis (2011b).

where Φ is the CDF of the Standard Normal distribution and

$$\hat{\sigma}_n^{*2} = \tfrac{1}{n} \sum_{j=1}^{J_n} \Big\{ \sum_{t=K_{n,j-1}+1}^{K_{nj}} (X_{nt}^* - \bar{X}_n^*) \Big\}^2. \tag{4}$$

As mentioned earlier, these are the same size and mixing conditions used by De Jong (1997). Note that De Jong does allow a little bit more flexibility in the conditions on the array $\{c_{nt}\}$ (see also Davidson, 1993); essentially, he allows for a single set of blocks with the maximal $\{c_{nt}\}$ over each block well-behaved, while this paper requires this condition to hold for every possible partition of blocks. This additional restriction is required because the Stationary Bootstrap will select the blocks randomly and is similar to De Jong and Davidson's (2000) requirement for the FCLT. Similarly, our assumption on the dispersion of the individual means, $(\mu_{nt} - \bar{\mu}_n)^2$, is slightly stronger than Gonçalves and White's (2002) and Gonçalves and De Jong's (2003) to accommodate larger block sizes.[6]

Theorem 1 relies on a general insight about the variance of the sample mean under the bootstrap-induced distribution. It is well-known that a key step in proving the CLT for arbitrary dependent processes is demonstrating that the squared elements converge to a positive and finite limit; i.e. if $\{Z_{nj};\ j = 1, \ldots, J_n\}$ is a representative stochastic array, $\sum_j Z_{nj}^2 \to^p \sigma^2$ is an important necessary condition for $\sum_j Z_{nj} \to^d N(0, \sigma^2)$. (See Section 3.2 of Hall and Heyde, 1980, for further discussion.) For martingale difference arrays, each Z_{nj} is one of the original random variables X_{nt} (typically normalized by $1/\sqrt{n}$), but for other forms of dependence (NED or mixingale arrays, for example, as in De Jong, 1997) each Z_{nj} is a contiguous block of the original random variables,

$$Z_{nj} = \tfrac{1}{\sqrt{n}} \sum_{t=(j-1)M_n+1}^{jM_n} (X_{nt} - \mu_{nt}),$$

that adds up to the original summation (plus potentially a negligible residual) so $\sum_j Z_{nj} = (1/\sqrt{n}) \sum_t (X_{nt} - \mu_{nt}) + o_p(1)$. In De Jong (1997), for example, the CLT for mixingale arrays assumes that there exists a sequence of blocks such that $\sum_j Z_{nj}^2$ converges i.p., and the NED CLT establishes conditions under which such a Z_{nj} exists.

Our insight is that the expectation of squared blocks of the bootstrap process can be expressed as a sequence of *contiguous* blocks of the original process, so the arguments that establish convergence of the original squared blocks can be applied with only minor changes to the bootstrapped blocks. Consider the Moving Blocks Bootstrap,[7] for example, and let

$$Z_{nj}^* = \tfrac{1}{\sqrt{n}} \sum_{t=(j-1)M_n+1}^{jM_n} (X_{nt}^* - \bar{X}_n).$$

[6] Note that this is one reason that we are presenting our results using general triangular-array notation: to allow for the means to change over time for comparison with Gonçalves and White's (2002) and Gonçalves and De Jong's (2003) results.

[7] To make this presentation as simple as possible, assume for now that $n = M_n J_n$ exactly.

Then, conditional on the data, the Z_{nj}^{*2} are independent can be expected to obey an LLN, so $\sum_j (Z_{nj}^{*2} - \mathrm{E}^* Z_{nj}^{*2}) \to^p 0$ and the CLT for the bootstrapped array requires $\sum_j \mathrm{E}^* Z_{nj}^{*2}$ to converge to a positive and finite limit. But, since E^* only averages over the starting point of each block, we have

$$\begin{aligned}\mathrm{E}^* Z_{nj}^{*2} &= \tfrac{1}{n} \sum_{\tau=0}^{n-M_n} \Big(\tfrac{1}{\sqrt{n}} \sum_{t=\tau+1}^{\tau+M_n} (X_{nt} - \bar{X}_n) \Big)^2 \\ &= \tfrac{1}{n} \sum_{\tau_0=0}^{M_n-1} \sum_{j=1}^{J_n} \Big(\tfrac{1}{\sqrt{n}} \sum_{t=(j-1)M_n+\tau_0+1}^{jM_n+\tau_0} (X_{nt} - \bar{X}_n) \Big)^2\end{aligned}$$

after grouping blocks separated by M_n periods. For each τ_0, the summation can be expected to converge in probability through the same arguments that were used to establish the CLT for the original array.[8,9] A similar representation is available for the Circular and Stationary bootstraps.

In short, the basic approach that we use to prove Theorem 1 is based on a fundamental connection between the second moments of the bootstrap process and the sum of squared blocks of the original array. Even though the details of our proof rely on specific techniques for NED arrays, this connection implies that block bootstraps are typically consistent when the original dependent array obeys the CLT and the connection should be useful for proving consistency of the bootstrap under other dependence conditions.

Theorem 1 can also be extended to give an FCLT using arguments from De Jong and Davidson (2000). We show in Theorem 2 that the partial sum of the bootstrapped process obeys an FCLT and can be used to derive critical values for other test statistics under the same assumptions as Theorem 1. For this result, define the following partial sums,

$$W_n(\gamma) = \tfrac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor \gamma n \rfloor} (X_{nt} - \bar{\mu}_n) \qquad \text{and} \qquad W_n^*(\gamma) = \tfrac{1}{\sqrt{n}} \sum_{t=1}^{\lfloor \gamma n \rfloor} (X_{nt}^* - \mu_n^*).$$

Also let W denote standard Brownian Motion and σW denote Brownian Motion scaled by the constant σ.

Theorem 2. *Suppose that the conditions of Theorem 1 hold and let* d *be any distance function that metricizes weak convergence. Then*

$$\Pr{}^*[\mathrm{d}(W_n^*, \sigma W) > \delta] \to^p 0 \tag{5}$$

for any positive δ. *If, in addition,* $\sup_{t=1,\dots,n} |\mu_{nt} - \bar{\mu}_n| = o(1/\sqrt{n})$ *and*

$$n^{-1} \sum_{s,t=1}^{\lfloor \gamma n \rfloor} \mathrm{cov}(X_{ns}, X_{nt}) \to \sigma^2 \gamma \tag{6}$$

[8] Some of the details of the argument will typically need to change because the original CLT only requires convergence for $\tau_0 = 0$, but these details are often incidental to the original argument.

[9] Lemmas 3 and 5 are particularly strong demonstrations of this argument.

for all $\gamma \in [0,1]$*, then*

$$\Pr[\mathrm{d}(W_n, \sigma W) > \delta] \to 0 \tag{7}$$

for any positive δ*.*

If both (5) and (7) hold then the bootstrap can be used to approximate the distribution of partial sums. Note that Theorem 2 imposes stronger assumptions for the original process than for the bootstrapped process. Without (10), the partial sum of the original observations converges to a transformed Brownian Motion with a different covariance process. The bootstrapped partial sum, on the other hand, always converges to standard (but potentially scaled) Brownian Motion because the resampling strategies ensure that the bootstrapped process is globally covariance stationary.

If the original process does not satisfy (10), it would be necessary to normalize W_n with a uniformly consistent estimator of the true covariance process of the series to make use of these results. This would be the case if the variance permanently changes partway through the series, for example. Other methods, such as the Local Block Bootstrap (Dowla et al., 2003; Paparoditis and Politis, 2002), may be able to capture this additional heterogeneity with the bootstrap alone, but we do not pursue that possibility further.

The rest of the paper presents the mathematical proofs in detail.

A Proof of main results

For both results, we only present a proof for the stationary bootstrap. The moving blocks and circular block bootstrap follow the same general argument but are simpler. We will define some additional notation here before presenting the proofs.

First, the relevant probability infrastructure. Let $(S, \Omega, \Pr)$ be a probability space and define the sequence of sub-sigma-fields $\Omega_n \subset \Omega$. Assume that each X_{nt}, V_{nt}, M_{nj}, and u_{nj} is Ω_n-measurable, with u_{nj} the uniform$(1, \dots, n)$ random variables that designate the start period of each bootstrap "block." Also define the σ-field generated by the stationary bootstrap's block lengths alone,

$$\mathcal{M}_n = \sigma(J_n, M_{n1}, \dots, M_{nJ_n}) \tag{8}$$

and the conditional probability $\Pr_{\mathcal{M}}(\cdot) = \Pr[\cdot \mid \mathcal{M}_n]$. (And define $\Pr^*_{\mathcal{M}}(\cdot) = \Pr^*[\cdot \mid \mathcal{M}_n]$, $\mathrm{E}_{\mathcal{M}}(\cdot) = \mathrm{E}^*(\cdot \mid \mathcal{M}_n)$, etc.) An important property is that $\mathcal{M}_n$ is *independent* of the X_{nt}'s, V_{nt}'s, and u_{nj}'s, so we can treat any M_{nj} and J_n terms as constants within $\mathrm{E}_{\mathcal{M}}(\cdot)$ and $\Pr_{\mathcal{M}}[\cdot]$ and integrate over the unconditional distributions of the other random variables. This property is especially important because it allows us to use maximal inequalities and other moment inequalities for mixingale arrays with only small modifications to construct almost sure bounds on the conditional moments $\mathrm{E}_{\mathcal{M}}$.

Also define

$$I_n(\tau,m) = \begin{cases} \{\tau+1,\dots,\tau+m\} & \text{if } 0 \le \tau \le n-m \text{ and } 1 \le m \\ \{1,\dots,m-n+\tau\} \cup \{\tau+1,\dots,n\} & \text{if } n-m < \tau \le n \text{ and } 1 \le m \\ \emptyset & \text{if } m \le 0, \end{cases} \tag{9}$$

so each $I_n(\tau,m)$ defines a potential block of length m of the original observations that could be chosen by the bootstrap.[10] By convention, summations over empty index sets will be considered equal to zero. Note that the I_n satisfy

$$\begin{aligned} I_n(K_{n0}, M_{n1}) &= \{1,\dots,K_{n1}\} \\ I_n(K_{n1}, M_{n2}) &= \{K_{n1}+1,\dots,K_{n2}\} \\ &\vdots \\ I_n(K_{n,J_n-1}, M_{nJ_n}) &= \{K_{n,J_{n-1}}+1,\dots,n\}, \end{aligned} \tag{10}$$

so $I_n(K_{n0},M_{n1}),\dots,I_n(K_{n,J_n-1},M_{nJ_n})$ exactly partition the set $\{1,\dots,n\}$ into consecutive blocks with lengths determined by the bootstrap.

Let

$$Z_n(\tau,m) = \tfrac{1}{\sqrt{n}} \sum_{t \in I_n(\tau,m)} (X_{nt} - \bar{X}_n) \tag{11}$$

$$Z_n^*(\tau,m) = \tfrac{1}{\sqrt{n}} \sum_{t \in I_n(\tau,m)} (X_{nt}^* - \bar{X}_n) \tag{12}$$

and

$$Z_{nj}^* = Z_n^*(K_{n,j-1}, M_{nj}) = \tfrac{1}{\sqrt{n}} \sum_{t=K_{n,j-1}+1}^{K_{n,j}} (X_{nt}^* - \bar{X}_n) \tag{13}$$

and define the corresponding demeaned terms

$$Z_n'(\tau,m) = \tfrac{1}{\sqrt{n}} \sum_{t \in I_n(\tau,m)} (X_{nt} - \mu_{nt}) \tag{14}$$

$$Z_n'^*(\tau,m) = \tfrac{1}{\sqrt{n}} \sum_{t \in I_n(\tau,m)} (X_{nt}^* - \mu_{nt}^*) \tag{15}$$

and

$$Z_{nj}'^* = Z_n'^*(K_{n,j-1}, M_{nj}), \tag{16}$$

[10] The index sets $I_n(\tau,m)$ are designed to "wrap around" and use the first observations when $\tau+m>n$, matching the defining aspect of the stationary and circular block bootstraps.

where μ_{nt}^* is the expected value of the observation in the original dataset corresponding to the tth observation in the bootstrapped dataset. Further, define the filtration

$$\mathcal{G}_{nj} = \sigma(Z_{n1}^*, \dots, Z_{nj}^*,\ X_{n1}, \dots, X_{nn},\ \mathcal{M}_n) \tag{17}$$

so that $\{Z_{nj}^*/\sigma_n^*, \mathcal{G}_{nj}\}$ is a martingale difference array.

By construction, (see Equations (10) and (13))

$$\tfrac{1}{\sqrt{n}} \sum_{t=1}^{n} (X_{nt}^* - \bar{X}_n) = \sum_{j=1}^{J_n} Z_{nj}^* \tag{18}$$

and

$$\mathrm{E}_{\mathcal{M}}^* \, g(Z_{nj}^*, \dots, Z_{nk}^*) = \frac{1}{n^{k-j+1}} \sum_{\tau_1=0}^{n-1} \cdots \sum_{\tau_{k-j+1}=0}^{n-1} g(Z_n(\tau_1, M_{nj}), \dots, Z_n(\tau_{k-j+1}, M_{nk})) \tag{19}$$

almost surely for any function g and any $j \leq k$. Equation (19) conditions on the lengths of each block, but averages over their starting points.

Proof of Theorem 1

First we prove that

$$\sup_x \left| \Pr^* \left[\sqrt{n}(\bar{X}_n^* - \bar{X}_n)/\sigma_n^* \leq x \right] - \Phi(x) \right| \to^p 0 \tag{20}$$

where $\sigma_n^{*2} = n \, \mathrm{E}^*(\bar{X}_n^* - \bar{X}_n)^2$. Rewrite $\sqrt{n}(\bar{X}_n^* - \bar{X}_n)$ as in Equation (18), so

$$\tfrac{1}{\sqrt{n}} \sum_{t=1}^{n} (X_{nt}^* - \bar{X}_n) = \sum_{j=1}^{J_n} Z_{nj}^*$$

and $\{Z_{nj}^*/\sigma_n^*, \mathcal{G}_{nj}\}$ is a martingale difference array. Moreover,

$$\Pr_{\mathcal{M}}^* \Big[\sum_{j=1}^{J_n} Z_{nj}^*/\sigma_n^* \leq x \Big] - \Phi(x) \to^p 0, \tag{21}$$

for all x if $\sigma_n^{*2} \to^p \sigma^2$ (which ensures that σ_n^{*2} is uniformly a.s. positive and holds by Lemma 3) and the following two conditions hold for all positive ϵ:

$$\sum_{j=1}^{J_n} \mathrm{E}_{\mathcal{M}}^* \left(Z_{nj}^{*2} \, 1\{Z_{nj}^{*2} > \epsilon\} \right) \to^p 0 \tag{22}$$

and

$$\Pr^*_{\mathcal{M}}\Big[\,\Big|\sum_{j=1}^{J_n} Z_{nj}^{*2} - \sigma_n^{*2}\Big| > \epsilon\Big] \to^p 0 \tag{23}$$

since (22) and (23) ensure that Z_{nj}^*/σ_n^* obeys a martingale difference CLT (e.g. Hall and Heyde, 1980, Theorem 3.3).[11]

For (23), Z_{nj}^* and Z_{nk}^* (when $k \neq j$) are conditionally uncorrelated given $X_{n1}, \dots, X_{nn}$, and $\mathcal{M}_n$, which implies

$$\sum_{j=1}^{J_n} Z_{nj}^{*2} - \sigma_n^{*2} = \sum_{j=1}^{J_n}\Big(Z_{nj}^{*2} - (1/J_n)\,\mathrm{E}^* \sum_{j=1}^{J_n} Z_{nj}^{*2}\Big)$$

almost surely. But

$$\Big\{\tfrac{n}{M_{nj}}\Big(Z_{nj}^{*2} - (1/J_n)\,\mathrm{E}^*\sum_{j=1}^{J_n} Z_{nj}^{*2}\Big),\ \mathcal{G}_{nj}\Big\}$$

is a uniformly-integrable martingale difference array by Lemma 5 and satisfies the LLN (e.g., Davidson, 1994, Theorem 19.7). So this sum converges to zero in conditional probability, proving (23).

To prove (22), it suffices to show that

$$\mathrm{E}_{\mathcal{M}} \sum_{j=1}^{J_n} \mathrm{E}^*_{\mathcal{M}}\left(Z_{nj}^{*2}\, 1\{Z_{nj}^{*2} > \epsilon\}\right) \to^p 0. \tag{24}$$

For any j and m and for large enough n,

$$\begin{aligned}\mathrm{E}_{\mathcal{M}}\,\mathrm{E}^*_{\mathcal{M}}&\left((Z_n(u_{nj},m)^2 n/m)\, 1\{Z_n(u_{nj},m)^2 n/m > \epsilon n/m\}\right)\\ &= \mathrm{E}\left((Z_n^*(0,m)^2 n/m)\, 1\{Z_n^*(0,m)^2 n/m > \epsilon n/m\}\right)\\ &\leq B(\epsilon n/m)\end{aligned} \tag{25}$$

where B is some finite monotone function such that $B(x) \to 0$ as $x \to \infty$; the existence of a function B that satisfies these conditions (and does not depend on n or m) is a consequence

[11] Conditional on $X_{n1}, \dots, X_{nn}$, J_n, and $M_{n1}, \dots, M_{n,J_n}$, the only stochastic components of $\sum_{j=1}^{J_n} Z_{nj}^*/\sigma_n$ are the start periods of each block, which are discrete uniform$(1, \dots, n)$ and are independent of all of the other random variables in the information set used for conditioning. Consequently, $\Pr^*_{\mathcal{M}}$ is a regular conditional probability and arguments like Hall and Heyde's (1980) Theorem 3.3 apply without modification on this probability measure. See also Section 23.2 of Van der Vaart (2000). Also note that Hall and Heyde's Theorem 3.3 as stated imposes an additional restriction on the sigma-fields. However, as Hall and Heyde discuss on pages 59 and 63–64, that condition is unnecessary here because σ_n^{*2} is measurable with respect to all of the $\mathcal{G}_{nj}$.

of Lemma 5, Equation (38). Since J_n and M_{nj} are both $\mathcal{M}_n$-measureable and independent of the u_{nj}'s and X_{nt}'s, (25) implies that

$$\begin{aligned}
&\mathrm{E}_{\mathcal{M}} \sum_{j=1}^{J_n} \mathrm{E}^*_{\mathcal{M}} \left(Z^{*2}_{nj} 1\{Z^{*2}_{nj} > \epsilon\}\right) \\
&\quad= \sum_{j=1}^{J_n} (M_{nj}/n)\, \mathrm{E}_{\mathcal{M}}\, \mathrm{E}^*_{\mathcal{M}} \left((Z_n(u_{nj}, M_{nj})^2 n/M_{nj})\, 1\{Z_n(u_{nj}, M_{nj})^2 n/M_{nj} > \epsilon n/M_{nj}\}\right) \\
&\quad\leq \sum_{j=1}^{J_n} (M_{nj}/n)\, B(\epsilon n/M_{nj}) \\
&\quad\leq \max_{j=1,\dots,J_n} B(\epsilon n/M_{nj}) \sum_{i=1}^{J_n} M_{ni}/n \\
&\quad= B(\epsilon n/ \max_{j=1,\dots,J_n} M_{nj}) \\
&\quad\to^p 0 \text{ as } n \to \infty.
\end{aligned}$$

The first equality follows from simple algerbra and the definition of the u_{nj}'s. The first inequality is a consequence of (25). The second equality holds by monotonicity of B. And convergence in probability holds by Lemma 1, completing the proof of (22). The Dominated Convergence Theorem and (21) then ensure that

$$\Pr^* \Big[\sum_{j=1}^{J_n} Z^*_{nj}/\sigma^*_n \leq x \Big] - \Phi(x) \to^p 0. \tag{26}$$

(Also see Lemma 2.)

Lemma 3 implies that σ^{*2}_n and $\hat{\sigma}^{*2}_n$ both converge to σ^2 in probability. This convergence then implies that

$$\Pr^* \left[\sqrt{n}(\bar{X}^*_n - \bar{X}_n) \leq x\right] \to^p \Phi(x/\sigma) \tag{27}$$

and

$$\Pr^* \left[\sqrt{n}(\bar{X}^*_n - \bar{X}_n)/\hat{\sigma}^*_n \leq x\right] \to^p \Phi(x) \tag{28}$$

for any x. These results are sufficient for (2) and (3) though an argument attributed to Polyà that proceeds as follows. Let k be a finite integer and define $x_i = \sigma\Phi^{-1}(i/k)$ for $i = 0, \dots, k$ (so $x_0 = -\infty$ and $x_k = +\infty$). For any $x \in [x_i, x_{i+1}]$,

$$\begin{aligned}
\Pr^* \left[\sqrt{n}(\bar{X}^*_n - \bar{X}_n) \leq x\right] - \Phi(x/\sigma) &\leq \Pr^* \left[\sqrt{n}(\bar{X}^*_n - \bar{X}_n) \leq x_{i+1}\right] - \Phi(x_i/\sigma) \\
&= \Pr^* \left[\sqrt{n}(\bar{X}^*_n - \bar{X}_n) \leq x_{i+1}\right] - \Phi(x_{i+1}/\sigma) + 1/k
\end{aligned}$$

and

$$\begin{aligned}\Pr^*\left[\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \le x\right] - \Phi(x/\sigma) &\ge \Pr^*\left[\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \le x_i\right] - \Phi(x_{i+1}/\sigma)\\ &= \Pr^*\left[\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \le x_i\right] - \Phi(x_i/\sigma) - 1/k\end{aligned}$$

almost surely. Then

$$\begin{aligned}\sup_{x\in(-\infty,+\infty)} &\left|\Pr^*\left[\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \le x\right] - \Phi(x/\sigma)\right|\\ &\le \sup_{i=0,\dots,k}\left|\Pr^*\left[\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \le x_i\right] - \Phi(x_i/\sigma)\right| + 1/k\end{aligned}$$

almost surely and (27) ensures that

$$\sup_{i=0,\dots,k}\left|\Pr^*\left[\sqrt{n}(\bar{X}_n^* - \bar{X}_n) \le x_i\right] - \Phi(x_i/\sigma)\right| + 1/k \to^p 1/k$$

for any finite k. Since k is arbitrary, (20) holds. Since Theorem 1's assumptions ensure that the original array obeys the CLT, (2) holds (De Jong, 1997, Theorem 2). A similar argument applies to the asymptotic distribution of $\sqrt{n}(\bar{X}_n^* - \bar{X}_n)/\hat{\sigma}_n^*$, completing the proof. □

Proof of Theorem 2

We will only present proofs for the bootstrap results, since Theorem 3.1 of De Jong and Davidson (2000) establishes the result for the partial sums of the original process. Theorem 1 implies that, for any fixed γ, $W_n^*(\gamma)$ is asymptotically normal with limiting variance $\gamma\sigma^2$, so we can assume that $\sigma^2 = 1$ without loss of generality. Moreover, Lemma 2 shows that it is sufficient to prove unconditional convergence, so we will establish $\Pr^*[\mathrm{d}(W_n^*, W) > \delta)] \to^p 0$. As in De Jong and Davidson (2000), this will hold if we show that W_n^* has asymptotically independent increments and stochastic equicontinuity, namely

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\Pr[\sup_{\gamma\in[0,1]}\sup_{\gamma'\in[\gamma-\delta,\gamma+\delta]}|W_n^*(\gamma) - W_n^*(\gamma')| > \epsilon] = 0 \tag{29}$$

for any positive ϵ.

First observe that we can write W_n^* as

$$W_n^*(\gamma) = \tfrac{1}{\sqrt{n}}\sum_{t=1}^{\lfloor\gamma n\rfloor}(X_{nt}^* - \bar{X}_n) = \sum_{j=1}^{\lfloor\gamma J_n\rfloor} Z_{nj}^* + Z_n^*(K_{n,\lfloor\gamma J_n\rfloor}, \lfloor\gamma n\rfloor - K_{n,\lfloor\gamma J_n\rfloor}).$$

To show that the increments of this process are asymptotically independent, choose $\gamma, \gamma' \in [0,1]$ and $\delta, \delta' > 0$ so that $\delta + \gamma \le \gamma'$. Since the blocks of W_n^* are conditionally uncorrelated given $X_{1n}, \dots, X_{nn}$, and $\mathcal{M}_n$, we have

$$\mathrm{E}\left[(W_n^*(\delta' + \gamma') - W_n^*(\gamma'))(W_n^*(\delta + \gamma) - W_n^*(\gamma))\right] = 0$$

for large enough n if $\gamma' > \gamma + \delta$. If $\gamma' = \gamma + \delta$ then

$$\begin{aligned}&\mathrm{E}\left[(W_n^*(\delta' + \gamma') - W_n^*(\gamma'))(W_n^*(\delta + \gamma) - W_n^*(\gamma))\right] = \\ &\quad \mathrm{E}\,\mathrm{E}_{\mathcal{M}}^*\big\{ - Z_n^*(\lfloor \gamma' n \rfloor, K_{n,\lceil \gamma' J_n \rceil} - \lfloor \gamma' n \rfloor) Z_n^*(K_{n,\lfloor(\gamma+\delta)J_n\rfloor}, \lfloor (\gamma+\delta) n \rfloor - K_{n,\lfloor(\gamma+\delta)J_n\rfloor})\big\}.\end{aligned}$$

But this second quantity can be bounded:

$$\begin{aligned}&\mathrm{E}\,\mathrm{E}_{\mathcal{M}}^*\left\{-Z_n^*(\lfloor \gamma' n \rfloor, K_{n,\lceil \gamma' J_n \rceil} - \lfloor \gamma' n \rfloor) Z_n^*(K_{n,\lfloor(\gamma+\delta)J_n\rfloor}, \lfloor (\gamma+\delta) n \rfloor - K_{n,\lfloor(\gamma+\delta)J_n\rfloor})\right\} \\ &\qquad \le \big\{ \mathrm{E}\,\mathrm{E}_{\mathcal{M}} \left[Z_n^*(\lfloor \gamma' n \rfloor, K_{n,\lceil \gamma' J_n \rceil} - \lfloor \gamma' n \rfloor)\right]^2 \\ &\qquad\quad \times \big\{ \mathrm{E}\,\mathrm{E}_{\mathcal{M}}\, Z_n^*(K_{n,\lfloor(\gamma+\delta)J_n\rfloor}, \lfloor (\gamma+\delta) n \rfloor - K_{n,\lfloor(\gamma+\delta)J_n\rfloor}))^2 \big\}^{1/2} \\ &\qquad \le C\,\mathrm{E}\, M_{n,\lfloor(\gamma+\delta)J_n\rfloor}/n\end{aligned}$$

for some constant C by Lemma 5. This term converges to zero by Lemma 1.

For (29), fix $\delta > 0$ such that $D = 2/\delta$ is a positive integer and let $\gamma_d = d/D$ for $d = 0, \ldots, D$. Mimicking the argument in De Jong and Davidson (2000) gives the bounds

$$\begin{aligned}&\Pr[\sup_{\gamma\in[0,1]} \sup_{\gamma'\in[\gamma-\delta,\gamma+\delta]} |W_n^*(\gamma) - W_n^*(\gamma')| > \epsilon] \\ &\qquad \le \Pr[\sup_{d=1,\ldots,D} \sup_{\gamma\in[0,\delta]} |W_n^*(\gamma + \gamma_d) - W_n^*(\gamma_d)| > \epsilon/2] \\ &\qquad \le (4/\epsilon^2) \sum_{d=1}^{D} \mathrm{E}[\sup_{\gamma\in[0,\delta]} |W_n^*(\gamma + \gamma_d) - W_n^*(\gamma_d)|^2 \\ &\qquad\qquad \times 1\{\sup_{\gamma\in[0,\delta]} |W_n^*(\gamma + \gamma_d) - W_n^*(\gamma_d)|^2 > \epsilon^2/4\}] \\ &\qquad \le (4/\epsilon^2) \max_{d=1,\ldots,D} \mathrm{E}[\sup_{\gamma\in[0,\delta]} (1/\delta)|W_n^*(\gamma + \gamma_d) - W_n^*(\gamma_d)|^2 \\ &\qquad\qquad \times 1\{\sup_{\gamma\in[0,\delta]} (1/\delta)|W_n^*(\gamma + \gamma_d) - W_n^*(\gamma_d)|^2 > \epsilon^2/4\delta\}].\end{aligned}$$

Lemma 5 implies that

$$\sup_{\gamma\in[0,\delta]} (1/\delta)|W_n^*(\gamma + \gamma_d) - W_n^*(\gamma_d)|^2$$

is uniformly integrable, so an argument similar to that used in the proof of Theorem 1 shows that there exists a finite and monotone function B such that $B(x) \to 0$ as $x \to \infty$ and

$$\begin{aligned}&\mathrm{E}[\sup_{\gamma\in[0,\delta]} (1/\delta)|W_n^*(\gamma + \gamma_d) - W_n^*(\gamma_d)|^2 \times \\ &\qquad 1\{\sup_{\gamma\in[0,\delta]} (1/\delta)|W_n^*(\gamma + \gamma_d) - W_n^*(\gamma_d)|^2 > x\}] \le B(x)\end{aligned}$$

for all d and δ and all large enough n.

As a result,

$$\lim_{\delta\to 0}\limsup_{n\to\infty}\Pr[\sup_{\gamma\in[0,1]}\sup_{\gamma'\in[\gamma-\delta,\gamma+\delta]}|W_n^*(\gamma)-W_n^*(\gamma')|>\epsilon]\le\lim_{\delta\to 0}(4/\epsilon^2)B(\epsilon^2/4\delta)$$
$$=0,$$

completing the proof. □

B Supplemental results

Lemma 1. *Suppose that $M_{n1}, M_{n2}, \ldots$ are i.i.d. geometric random variables with success parameter $p_n = cn^{-a}$ with $a, c \in (0,1)$, and that $\ell_n = (p_n \log p_n^{-1})^{-1}$ and define J_n so that $\sum_{i=1}^{J_n-1} M_{ni} < n \le \sum_{i=1}^{J_n} M_{ni}$ Then*

1. $\max_{i=1,\ldots,\lfloor Cnp_n\rfloor} M_{ni}/n \to^p 0$ *for any positive C,*
2. $\max_{i=1,\ldots,J_n} M_{ni}/n \to^p 0$,
3. $\max_{i=1,\ldots,\lfloor Cnp_n\rfloor} M_{ni}/\ell_n^{1+\epsilon} \to^p 0$ *as $n\to\infty$ for any positive ϵ and C,*
4. $\max_{i=1,\ldots,J_n} M_{ni}/\ell_n^{1+\epsilon} \to^p 0$ *as $n\to\infty$ for any positive ϵ, and*
5. $\sum_{i=1}^{J_n} M_{ni}^2/n^2 \to^p 0$.

Proof of Lemma 1. To prove part 1, observe that, for any increasing positive sequence x_n such that $x_n p_n \to \infty$,

$$\Pr\Big[\max_{i=1,\ldots,\lfloor Cnp_n\rfloor} M_{ni} \le x_n\Big] = (1-(1-p_n)^{x_n})^{\lfloor Cnp_n\rfloor} \to \lim\exp(-Cnp_n(1-p_n)^{x_n})$$

and $Cnp_n(1-p_n)^{x_n} \to \lim Cnp_n e^{-x_n p_n}$. Now, let $x_n = nx$ for any positive number x. Then

$$\Pr\Big[\max_{i=1,\ldots,\lfloor Cnp_n\rfloor} M_{ni}/n \le x\Big] \to \lim\exp(-Cnp_n e^{-np_n x}) = \exp(0) = 1.$$

Since x is arbitrary, $\max_{i=1,\ldots,\lfloor Cnp_n\rfloor} M_{ni}/n \to^p 0$.

For part 2, take C to be an arbitrary constant strictly greater than one. For any x,

$$\Pr\Big[\max_{i=1,\ldots,J_n} M_{ni} > x\Big] \le \Pr\Big[\max_{i=1,\ldots,\lfloor Cnp_n\rfloor} M_{ni} > x \text{ or } J_n > \lfloor Cnp_n\rfloor\Big]$$
$$\le \Pr\Big[\max_{i=1,\ldots,\lfloor Cnp_n\rfloor} M_{ni} > x\Big] + \Pr\Big[\sum\nolimits_{i=1}^{\lfloor Cnp_n\rfloor} M_{ni} < n\Big]$$

The first term converges to zero by part 1 and the second term by the LLN.

For part 3, let $x_n = \ell_n^{1+\epsilon} x$ and note that

$$p_n \ell_n^{1+\epsilon} \geq p_n^{-(\epsilon-\delta-\epsilon\delta)} = c^{-(\epsilon-\delta-\epsilon\delta)} n^{a(\epsilon-\delta-\epsilon\delta)} \equiv b n^{a(\epsilon-\delta-\epsilon\delta)}$$

for any $\delta > 0$ and large enough n. Choose δ small enough that $\epsilon > \delta(1+\epsilon)$. Then

$$n p_n \exp(-\ell_n^{1+\epsilon} p_n) \leq n p_n \exp(-b n^{a(\epsilon-\delta-\epsilon\delta)}) = c v_n^{\frac{1-a}{a(\epsilon-\delta-\epsilon\delta)}} \exp(-b v_n) \to 0,$$

with $v_n = n^{a(\epsilon-\delta-\epsilon\delta)}$. Consequently,

$$\Pr[\max_i M_{ni}/\ell_n^{1+\epsilon} \leq x] \to \exp(0) = 1$$

as well.

The proof of part 4 is the same as part 2, making the obvious substitutions. Part 5 holds because $\sum_{i=1}^{J_n} M_{ni}^2/n^2 \leq \max_{i=1,\dots,J_n} M_{ni}/n$ which converges to zero in probability by part 2. □

Lemma 2. *If $\{A_n\}$ is a sequence of events in Ω then the following are equivalent:*

$$\Pr[A_n] \to 0, \qquad \Pr^*[A_n] \to 0 \text{ in } L_1, \qquad \text{and} \qquad \Pr^*_{\mathcal{M}}[A_n] \to 0 \text{ in } L_1.$$

Proof. Since $|\Pr[A_n]| = \mathrm{E}|\Pr^*[A_n]| = \mathrm{E}|\Pr^*_{\mathcal{M}}[A_n]|$ these conditions are equivalent by definition. □

Lemma 3. *Under the conditions of Theorem 1,*

$$\sum_{j=1}^{J_n} (Z_{nj}^{*2} - \mathrm{E}^*_{\mathcal{M}} Z_{nj}^{*2}) \to^p 0, \tag{30}$$

$$\sum_{j=1}^{J_n} (Z_{nj}^{*2} - Z_{nj}^{\prime *2}) \to^p 0, \tag{31}$$

and $\Pr^*[|\sigma_n^{*2} - \sigma^2| > \epsilon] \to^p 0$. *If, in addition,* $\bar{X}_n^* - \bar{X}_n = O_p(1/\sqrt{n})$ *then* $\Pr^*[|\hat{\sigma}_n^{*2} - \sigma^2| > \epsilon] \to^p 0$.

Proof. For (31), $(Z_{nj}^{*2} - \mathrm{E}^*_{\mathcal{M}} Z_{nj}^{*2}) \cdot (n/M_{nj})$ is a uniformly integrable martingale difference array, by Lemma 5, and satisfies the LLN. (See Davidson, 1994, Theorem 19.7.) For (31), observe that

$$\Big|\sum_{j=1}^{J_n} (Z_{nj}^{*2} - Z_{nj}^{\prime *2})\Big| \leq 2\Big(\sum_{j=1}^{J_n} Z_{nj}^{*2}\Big)^{1/2} \Big(\tfrac{1}{n}\sum_{t=1}^{n} (\mu_{nt} - \bar{X}_n)^2\Big)^{1/2} + \tfrac{1}{n}\sum_{t=1}^{n} (\mu_{nt} - \bar{X}_n)^2$$

after several applications of the Cauchy-Schwarz inequality. Lemma 4, along with (19) and (30), implies that $\sum_j Z_{nj}^{*2} = O_p(1)$; $(1/n)\sum_t (\mu_{nt} - \bar{X})^2$ converges to zero in probability by assumption on $\mu_{nt} - \bar{\mu}_n$ and because $\bar{X}$ itself obeys the LLN.

To show that σ_n^{*2} converges, we can write

$$\sigma_n^{*2} - \sigma^2 = \mathrm{E}^* \Big\{ \sum_{j=1}^{J_n} (Z_{nj}^{*2} - Z_{nj}^{\prime *2}) \Big\} + \tfrac{1}{n} \sum_{\tau=0}^{n-1} \mathrm{E}^* \Big\{ \sum_{j=1}^{J_n} \big(Z_n'(\tau, M_{nj})^2 - \mathrm{E}_{\mathcal{M}}\, Z_n'(\tau, M_{nj})^2 \big) \Big\} + \tfrac{1}{n} \sum_{\tau=0}^{n-1} \mathrm{E}^* \Big\{ \sum_{j=1}^{J_n} \mathrm{E}_{\mathcal{M}}\, Z_n'(\tau, M_{nj})^2 - \sigma^2 \Big\}. \tag{32}$$

Uniform integrability ensures that the convergence in (30) holds in L_1 as well and Lemma 2 then implies that the first term in (32) converges to zero in probability. Lemma 4 proves that the second and third summation converge to zero in probability.

Next,

$$\hat{\sigma}_n^{*2} - \sigma_n^{*2} = \sum_{j=1}^{J_n} \big(Z_{nj}^* + (M_{nj}/\sqrt{n})(\bar{X}_n - \bar{X}_n^*) \big)^2 - \mathrm{E}^* \sum_{j=1}^{J_n} Z_{nj}^{*2}$$

so, in light of the previous arguments, $\hat{\sigma}_n^{*2} \to^p \sigma^2$ if

$$(\bar{X}_n - \bar{X}_n^*)^2 \sum_{j=1}^{J_n} M_{nj}^2 / n \to^p 0, \tag{33}$$

which holds by Lemma 1 and assumption. □

Lemma 4. *If the conditions of Theorem 1 hold then*

$$\Pr \Big[\Big| \tfrac{1}{n} \sum_{\tau=0}^{n-1} \sum_{j=1}^{J_n} \big[Z_n'(\tau, M_{nj})^2 - \mathrm{E}_{\mathcal{M}}\, Z_n'(\tau, M_{nj})^2 \big] \Big| > \epsilon \Big] \to 0 \tag{34}$$

and

$$\Pr \Big[\Big| \tfrac{1}{n} \sum_{\tau=0}^{n-1} \sum_{j=1}^{J_n} \mathrm{E}_{\mathcal{M}}\, Z_n'(\tau, M_{nj})^2 - \sigma^2 \Big| > \epsilon \Big] \to 0. \tag{35}$$

For these two proofs, let $\ell_n = (p_n \log p_n^{-1})^{-1}$ and let $L_{nj} = \lfloor n / M_{nj} \rfloor$; ℓ_n represents a smaller block size that satisfies $\ell_n J_n / n \to^p 0$.

Proof of (34). We can express this summation as

$$
\begin{aligned}
&\tfrac{1}{n}\sum_{\tau=0}^{n-1}\sum_{j=1}^{J_n}\left\{Z_n'(\tau,M_{nj})^2-\mathrm{E}_{\mathcal{M}}\,Z_n'(\tau,M_{nj})^2\right\}\\
&\quad=\tfrac{1}{n}\sum_{j=1}^{J_n}\sum_{\tau=0}^{M_{nj}-1}\sum_{i=0}^{L_{nj}-1}\Big\{\big[Z_n'(\tau+iM_{nj},M_{nj}-\ell_n)+Z_n'(\tau+(i+1)M_{nj}-\ell_n,\ell_n)\big]^2\\
&\qquad\quad-\mathrm{E}_{\mathcal{M}}\big[Z_n'(\tau+iM_{nj},M_{nj}-\ell_n)+Z_n'(\tau+(i+1)M_{nj}-\ell_n,\ell_n)\big]^2\Big\}\\
&\qquad\qquad+\tfrac{1}{n}\sum_{j=1}^{J_n}\sum_{\tau=M_{nj}L_{nj}}^{n-1}\left\{Z_n'(\tau,M_{nj})^2-\mathrm{E}_{\mathcal{M}}\,Z_n'(\tau,M_{nj})^2\right\} \qquad (36)
\end{aligned}
$$

almost surely. By Lemma 6 (Equation 41), for any $\delta > 0$ there exist positive C and ϵ such that

$$
\begin{aligned}
&\Big\|\tfrac{1}{n}\sum_{j=1}^{J_n}\sum_{\tau=0}^{M_{nj}-1}\sum_{i=0}^{L_{nj}-1}\left\{Z_n'(\tau+iM_{nj},M_{nj}-\ell_n)^2-\mathrm{E}_{\mathcal{M}}\left(Z_n'(\tau+iM_{nj},M_{nj}-\ell_n)^2\right)\right\}\Big\|_1\\
&\quad\le \mathrm{E}\,\tfrac{1}{n}\sum_{j=1}^{J_n}\sum_{\tau=0}^{M_{nj}-1}\mathrm{E}_{\mathcal{M}}\Big|\sum_{i=0}^{L_{nj}-1}\left\{Z_n'(\tau+iM_{nj},M_{nj}-\ell_n)^2-\mathrm{E}_{\mathcal{M}}\left(Z_n'(\tau+iM_{nj},M_{nj}-\ell_n)^2\right)\right\}\Big|\\
&\quad\le \mathrm{E}\Big(\tfrac{1}{n}\sum_{j=1}^{J_n}\sum_{\tau=0}^{M_{nj}-1}\big(2\delta+C\,\tfrac{M_{nj}}{n^{1/2}\ell^{1/2+\epsilon}}\big)\Big)
\end{aligned}
$$

for large enough n, which converges to 2δ by Lemma 1. Lemma 7 ensures that there exists a value C (possibly different from the value above) such that

$$
\begin{aligned}
&\mathrm{E}\Big(\tfrac{1}{n}\sum_{j=1}^{J_n}\sum_{\tau=0}^{M_{nj}-1}\sum_{i=0}^{L_{nj}-1}Z_n'(\tau+(i+1)M_{nj}-\ell_n,\ell_n)^2\Big)\\
&\qquad=\mathrm{E}\,\mathrm{E}_{\mathcal{M}}\Big(\tfrac{1}{n}\sum_{j=1}^{J_n}\sum_{\tau=0}^{M_{nj}-1}\sum_{i=0}^{L_{nj}-1}Z_n'(\tau+(i+1)M_{nj}-\ell_n,\ell_n)^2\Big)\\
&\qquad\le C\,\mathrm{E}\Big(\sum_{j=1}^{J_n}L_{nj}\ell_nM_{nj}/n^2\Big)
\end{aligned}
$$

and

$$
\mathrm{E}\Big(\tfrac{1}{n}\sum_{j=1}^{J_n}\sum_{\tau=M_{nj}L_{nj}}^{n-1}Z_n'(\tau,M_{nj})^2\Big)\le C\,\mathrm{E}\Big(\sum_{j=1}^{J_n}M_{nj}^2/n^2\Big)
$$

for large enough n,[12] both of which converge to zero in L_1 as $n \to \infty$ by Lemma 1. These three convergence results imply that the RHS of (36) converges to zero in probability, completing the proof. ☐

Proof of (35). After using similar arguments to the previous part of the proof, the conclusion holds if

$$\tfrac{1}{n}\sum_{j=1}^{J_n}\sum_{\tau=0}^{M_{nj}-1}\sum_{i=0}^{L_{nj}-1} \mathrm{E}_{\mathcal{M}}\, Z_n'(\tau + iM_{nj}, M_{nj} - \ell_n)^2 \to^p \sigma^2,$$

which is a direct implication of Lemma 6.[13] ☐

Lemma 5. *Under the conditions of Theorem 1,*

$$\lim_{C\to\infty} \limsup_{n\to\infty} \sup_{m'=1,\dots,n} \mathrm{E}\Big(\Big(\sum_{\tau=0}^{n-1} \max_{m=1,\dots,m'} Z_n(\tau,m)^2/m'\Big) \times 1\Big\{\sum_{\tau=0}^{n-1} \max_{m=1,\dots,m'} Z_n(\tau,m)^2/m' > C\Big\}\Big) = 0. \quad (37)$$

and

$$\lim_{C\to\infty} \limsup_{n\to\infty} \sup_{\substack{\tau=0,\dots,n-1\\ m'=1,\dots,n}} \mathrm{E}\Big(\max_{m=1,\dots,m'} (n\, Z_n^*(\tau,m)^2/m') \times 1\Big\{\max_{m=1,\dots,m'} n\, Z_n^*(\tau,m)^2/m' > C\Big\}\Big) = 0, \quad (38)$$

making the family of random variables $\{\max_{m=1,\dots,m'} Z_n^*(\tau,m)^2 n/m'; \tau, m', n\}$ *uniformly integrable.*

Proof. For (37), take $\epsilon > 0$. The sum

$$\sum_{\tau=0}^{n-1} \max_{m=1,\dots,m'} Z_n(\tau,m)^2/m'$$

is uniformly integrable if and only if it is uniformly L_1-bounded and there exists an ϵ' such that

$$\sup_{n,m'} \mathrm{E}\Big(1(A)\sum_{\tau=0}^{n-1} \max_{m=1,\dots,m'} Z_n(\tau,m)^2/m'\Big) < \epsilon$$

[12] Lemma 6 and 7 are stated for the unconditional expectation, which may cause some confusion here. Note that M_{nj} is measurable with respect to $\mathcal{M}_n$ and is treated the same as the constant m in the statement of these Lemmas. The other random variables in these expressions are independent of $\mathcal{M}_n$. So the only effect of conditioning on $\mathcal{M}_n$ is to prevent integration over the distributions of the M_{nj} and the expectation otherwise behaves exactly like the unconditional expectation.

[13] The discussion in Footnote 12 applies here as well.

for every event $A \in \Omega_n$ satisfying $\Pr(A) < \epsilon'$. (Davidson, 1994, 12.9.) Lemma 7 ensures uniform integrability of $\max_{m=1,\dots,m'} Z_n'(\tau, m)^2 n/m'$, so there exists an ϵ' such that

$$\sup_{n,m',\tau} \mathrm{E}\left(1(A) \max_{m=1,\dots,m'} Z_n'(\tau, m)^2 n/m'\right) \leq \epsilon/9$$

for all $A \in \Omega_n$ with $\Pr(A) \leq \epsilon'$. Then, writing

$$\begin{aligned} Z_n(\tau, m) &= Z_n'(\tau, m) + \tfrac{1}{\sqrt{n}} \sum_{t \in I_n(\tau,m)} \big((\mu_{nt} - \bar{\mu}_n) + (\bar{\mu}_n - \bar{X}_n)\big) \\ &= Z_n'(\tau, m) + \tfrac{1}{\sqrt{n}} \sum_{t \in I_n(\tau,m)} (\mu_{nt} - \bar{\mu}_n) - (m/n) Z_n'(0, n) \end{aligned}$$

gives the relationship

$$\begin{aligned} &\mathrm{E}\left(1(A) \sum_{\tau=0}^{n-1} \max_{m=1,\dots,m'} Z_n(\tau, m)^2/m'\right) \\ &\qquad \leq 3 \sum_{\tau=0}^{n-1} \mathrm{E}\left(1(A) \max_{m=1,\dots,m'} Z_n'(\tau, m)^2/m'\right) \\ &\qquad\quad + (3/nm')\, \mathrm{E}\left(1(A) \sum_{\tau=0}^{n-1} \max_{m=1,\dots,m'} \Big(\sum_{t \in I_n(\tau,m)} (\mu_{nt} - \bar{\mu}_n)\Big)^2\right) \\ &\qquad\quad + (3m'/n)\, \mathrm{E}(1(A) Z_n'(0, n)^2) \\ &\qquad \leq (2\epsilon/3) + (3\epsilon' m'/n) \sum_{t=1}^{n} (\mu_{nt} - \bar{\mu}_n)^2 \end{aligned}$$

where the last inequality holds as a consequence of A's construction. Since $\sum_{t=1}^n (\mu_{nt} - \bar{\mu}_n)^2 \to 0$, choose n large enough that this sum is less than $\epsilon/9\epsilon'$. The proof that that the first moment is bounded is similar, which completes the proof of (37).

For (38), we will first show that, for every $\epsilon > 0$, there exists an $\epsilon' > 0$ with the property that

$$\mathrm{E}_{\mathcal{M}}\left(1(A) \times \max_{m=1,\dots,m'} Z_n^*(\tau, m)^2 n/m'\right) < \epsilon \tag{39}$$

for any event $A \in \mathcal{F}_n \equiv \sigma(X_{n1}, \dots, X_{nn};\ M_{n1}, \dots, M_{n,J_n};\ J_n)$ with $\Pr(A) \leq \epsilon'$. Then we will show that this property implies uniform integrability.[14]

[14] The novelty in this part of the proof is the measurability requirement. The sigma-field $\mathcal{F}_n$ intentionally exlcudes the block start periods $u_{n1}, \dots, u_{n,J_n}$, so $Z_n^*(\tau, m)^2$ is not $\mathcal{F}_n$-measureable. If we instead proved (39) for all $A \in \Omega_n$, as we did for the first part of the proof, then uniform integrebality would be an immediate consequence and would not need to be shown separately. See Davidson (1994, Theorem 12.9).

Take an arbitrary $\epsilon > 0$ and a value of $\epsilon' > 0$ so that

$$\sup_{m'=1,\dots,n} \mathrm{E}_{\mathcal{M}} \Big(1(A) \sum_{\tau=0}^{n-1} \tfrac{1}{m'} \max_{m=1,\dots,m'} Z_n(\tau,m)^2 \Big) < \epsilon/6$$

for any $A \in \mathcal{F}_n$ with $\Pr(A) \leq \epsilon'$, which is shown to exist in the first part of this Lemma. Then, for any x, define $J(x)$ to be the block index such that $K_{n,J(x)-1} < x \leq K_{n,J(x)}$. For any m, we can decompose $Z_n^*(\tau, m)$ as

$$Z_n^*(\tau,m) = Z_n^*(\tau, K_{n,J(\tau)} - \tau) + \sum_{j=J(\tau)+1}^{J(\tau+m)-1} Z_{nj}^* + Z_n^*(K_{n,J(\tau+m)-1}, m - K_{n,J(\tau+m)-1}),$$

giving

$$\begin{aligned} &\mathrm{E}_{\mathcal{M}}\, \mathrm{E}_{\mathcal{M}}^* \Big(1(A) \times \max_{m=1,\dots,m'} n\, Z_n^*(\tau,m)^2/m' \Big) \\ &\qquad \leq 3\, \mathrm{E}_{\mathcal{M}}\, \mathrm{E}_{\mathcal{M}}^* \Big[1(A) \tfrac{n}{m'} \max_{m=1,\dots,m'} \Big(Z_n^*(\tau, K_{n,J(\tau)} - \tau) + \sum_{j=J(\tau)+1}^{J(\tau+m)-1} Z_{nj}^* \Big)^2 \Big] \\ &\qquad\quad + 3\, \mathrm{E}_{\mathcal{M}}\, \mathrm{E}_{\mathcal{M}}^* \big[1(A) \tfrac{n}{m'} \max_{m=1,\dots,m'} Z_n^*(K_{n,J(\tau+m)-1}, m - K_{n,J(\tau+m)-1})^2 \big] \qquad (40) \end{aligned}$$

almost surely.

Since Z_{nj}^* is a martingale difference array, Doob's maximal inequality for martingales (see Davidson, 1994, Theorem 15.15, for example) along with the construction of Z_{nj}^* gives a bound for the first term on the RHS of (40),

$$\begin{aligned} &\mathrm{E}_{\mathcal{M}}\, \mathrm{E}_{\mathcal{M}}^* \Big[1(A) \tfrac{n}{m'} \max_{m=1,\dots,m'} \Big(Z_n^*(\tau, K_{n,J(\tau)} - \tau) + \sum_{j=J(\tau)+1}^{J(\tau+m)-1} Z_{nj}^* \Big)^2 \Big] \\ &\qquad \leq \tfrac{1}{m'} \sum_{u=0}^{n-1} \Big\{ \mathrm{E}_{\mathcal{M}} \big[1(A) Z_n(u, K_{n,J(\tau)} - \tau)^2 \big] + \sum_{j=J(\tau)+1}^{J(\tau+m')-1} \mathrm{E}_{\mathcal{M}} \big[1(A) Z_n(u, M_{nj})^2 \big] \Big\} \end{aligned}$$

and this term is less than $\epsilon/6$ by construction of A. For the second term on the RHS (40), use the sequence of inequalities

$$\begin{aligned}
&\mathrm{E}_{\mathcal{M}}\,\mathrm{E}^*_{\mathcal{M}}\left[1(A)\tfrac{n}{m'}\max_{m=1,\dots,m'} Z^*_n(K_{n,J(\tau+m)-1}, m-K_{n,J(\tau+m)-1})^2\right]\\
&\qquad\leq \sum_{j=J(\tau)+1}^{J(\tau+m')-1} \tfrac{n}{m'}\,\mathrm{E}_{\mathcal{M}}\,\mathrm{E}^*_{\mathcal{M}}\left[1(A) \max_{m=1,\dots,\min(M_{nj},m'-K_{n,j-1})} Z^*_n(K_{n,j},m)^2\right]\\
&\qquad\leq \sum_{j=J(\tau)+1}^{J(\tau+m')} \tfrac{M_{nj}}{m'}\sum_{u=0}^{n-1}\mathrm{E}_{\mathcal{M}}\left[1(A) \max_{m=1,\dots,\min(M_{nj},m'-K_{n,j-1})} Z_n(u,m)^2/M_{nj}\right]\\
&\qquad\leq \sum_{j=J(\tau)+1}^{J(\tau+m')} \epsilon\cdot\min(M_{nj}, m'-K_{n,j-1})/6m'\\
&\qquad\leq \epsilon/6.
\end{aligned}$$

Together, these bounds imply (39). A similar argument also implies that $nZ^*_n(\tau,m)^2/m'$ has finite first moment.

Now, for uniform integrability, take $\epsilon > 0$, choose ϵ' so that (39) holds for all $A \in \mathcal{F}_n$ s.t. $\Pr(A) \leq \epsilon'$, and choose C_0 large enough that

$$\Pr_{\mathcal{M}}\left[\max_{m=1,\dots,m'} n\, Z^*_n(\tau,m)^2/m' > C_0\right] \leq \epsilon'.$$

Markov's inequality guarantees the existence of this C_0. Then

$$\mathrm{E}_{\mathcal{M}}\left(\max_{m=1,\dots,m'} n\, Z^*_n(\tau,m)^2/m' \times 1\left\{\max_{m=1,\dots,m'} n\, Z^*_n(\tau,m)^2/m' > C\right\}\right) \leq \epsilon$$

for all $C > C_0$. Since ϵ is arbitrary, this completes the proof. □

Lemma 6. *Suppose the conditions of Theorem 1 hold. For any positive δ, there exist positive and finite constants C, n_0, and ϵ such that for all $n > n_0$, $m = 1, \dots, n$, $\tau = 0, \dots, m$, and $\ell = 1, \dots, m-1$:*

$$\mathrm{E}\Big|\sum_{i=0}^{\lfloor n/m\rfloor-1}\left[Z'_n(\tau+im, m-\ell)^2 - \mathrm{E}\left(Z'_n(\tau+im,m-\ell)^2\right)\right]\Big| \leq 2\delta + C\cdot\left(\tfrac{m}{n}\right)^{1/2}\left(\tfrac{m}{\ell^{1+\epsilon}}\right)^{1/2}. \quad (41)$$

Also, there exists a constant C and a finite function $D(x)$ such that $D(x) \to 0$ as $x \to \infty$ and, for large enough n,

$$\mathrm{E}\Big|\sum_{i=0}^{\lfloor n/m\rfloor-1}\mathrm{E}(Z'_n(\tau+im,m-\ell)^2 - \sigma^2\Big| \leq C\,D(\ell). \quad (42)$$

Results (41) and (42) are direct extensions of De Jong's (1997) Lemmas 5 and 4, respectively, replacing De Jong's implicit use of inequalities with explicit inequalities. The supplemental appendix presents a proof of (42) to show the main idea.

Note that these results apply immediately to the conditional expectation $\mathrm{E}_{\mathcal{M}}$ because the block lengths M_{jn} do not appear in (41) and (42) and are independent of all of the random variables used for these bounds.

Lemma 7. *Under the conditions of Theorem 1,*

$$\lim_{C\to 0}\limsup_{n\to\infty}\sup_{\substack{\tau=0,\dots,n-1\\ m'=1,\dots,n}}\mathrm{E}\big(\big(\max_{m=1,\dots,m'} Z_n'(\tau,m)^2 n/m'\big) \times 1\{\max_{m=1,\dots,m'} Z_n'(\tau,m)^2 n/m' > C\}\big)\big) = 0. \quad (43)$$

For proof, see supplemental appendix for the proof of (43), which follows McLeish (1975, Lemma 6.5) and McLeish (1977, Lemma 3.5) almost exactly and is also presented as Theorem 16.13 in Davidson (1994). The same comment that follows Lemma 6 applies here as well: the bounds apply equally well to $\mathrm{E}_{\mathcal{M}}$.

References

Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley, New York.

Dahlhaus, R. (2011). Discussion: Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4):379–381.

Davidson, J. (1993). An L_1-Convergence theorem for heterogeneous mixingale arrays with trending moments. *Statistics & Probability Letters*, 16(4):301–304.

Davidson, J. (1994). *Stochastic Limit Theory: An Introduction for Econometricians*. Advanced Texts in Econometrics. Oxford University Press.

De Jong, R. M. (1997). Central limit theorems for dependent heterogeneous random variables. *Econometric Theory*, 13(3):353–367.

De Jong, R. M. and Davidson, J. (2000). The functional central limit theorem and weak convergence to stochastic integrals I: Weakly dependent processes. *Econometric Theory*, 16(5):621–642.

Dowla, A., Paparoditis, E., and Politis, D. N. (2003). Locally stationary processes and the Local Block Bootstrap. In Akritas, M. G. and Politis, D. N., editors, *Recent Advances and Trends in Nonparametric Statistics*, pages 437–444. Elsevier (North Holland).

Gonçalves, S. and De Jong, R. M. (2003). Consistency of the stationary bootstrap under weak moment conditions. *Economics Letters*, 81(2):273–278.

Gonçalves, S. and Politis, D. (2011). Discussion: Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4):383–386.

Gonçalves, S. and White, H. (2002). The bootstrap of the mean for dependent heterogeneous arrays. *Econometric Theory*, 18(6):1367–1384.

Hall, P. and Heyde, C. (1980). *Martingale limit theory and its application*. Academic Press.

Horowitz, J. L. (2011). Discussion: Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4):387–389.

Jentsch, C. and Mammen, E. (2011). Discussion: Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4):391–392.

Kreiss, J.-P. and Paparoditis, E. (2011a). Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4):357–378.

Kreiss, J.-P. and Paparoditis, E. (2011b). Rejoinder: Bootstrap methods for dependent data: A review. *Journal of the Korean Statistical Society*, 40(4):393–395.

Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, 17(3):1217–1241.

Lahiri, S. (1999). Theoretical comparisons of block bootstrap methods. *The Annals of Statistics*, 27(1):386–404.

Liu, R. Y. and Singh, K. (1992). Moving blocks Jackknife and Bootsrap capture weak dependence. In LePage, R. and Billard, L., editors, *Exploring the Limits of Bootstrap*, pages 225–248. John Wiley, New York.

McLeish, D. (1975). Invariance principles for dependent variables. *Probability Theory and Related Fields*, 32:165–178.

McLeish, D. (1977). On the invariance principle for nonstationary mixingales. *The Annals of Probability*, 5(4):616–621.

Nordman, D. (2009). A note on the stationary bootstrap's variance. *The Annals of Statistics*, 37(1):359–370.

Paparoditis, E. and Politis, D. (2002). Local block bootstrap. *Comptes Rendus Mathematique*, 335(11):959–962.

Politis, D. N. and Romano, J. P. (1992). A circular block resampling procedure for stationary data. In Page, R. and LePage, R., editors, *Exploring the Limits of Bootstrap*, pages 263–270. John Wiley, New York.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89(428):1303–1313.

Radulović, D. (1996). The bootstrap of the mean for strong mixing sequences under minimal conditions. *Statistics & probability letters*, 28(1):65–72.

Van der Vaart, A. W. (2000). *Asymptotic statistics.* Cambridge University Press.