Journal of the American Statistical Association

ISSN: 0162-1459 (Print) 1537-274X (Online) Journal homepage: www.tandfonline.com/journals/uasa20

Block Bootstraps for Time Series With Fixed Regressors

Daniel J. Nordman & Soumendra N. Lahiri

To cite this article: Daniel J. Nordman & Soumendra N. Lahiri (2012) Block Bootstraps for Time Series With Fixed Regressors, Journal of the American Statistical Association, 107:497, 233-246, DOI: 10.1080/01621459.2011.646929

To link to this article: https://doi.org/10.1080/01621459.2011.646929

Supplementary materials for this article are available online. Please go to www.tandfonline.com/r/JASA

Block Bootstraps for Time Series With Fixed Regressors

Daniel J. NORDMAN and Soumendra N. LAHIRI

This article examines block bootstrap methods in linear regression models with weakly dependent error variables and nonstochastic regressors. Contrary to intuition, the tapered block bootstrap (TBB) with a smooth taper not only loses its superior bias properties but may also fail to be consistent in the regression problem. A similar problem, albeit at a smaller scale, is shown to exist for the moving and the circular block bootstrap (MBB and CBB, respectively). As a remedy, an additional block randomization step is introduced that balances out the effects of nonuniform regression weights, and restores the superiority of the (modified) TBB. The randomization step also improves the MBB or CBB. Interestingly, the stationary bootstrap (SB) automatically balances out regression weights through its probabilistic blocking mechanism, without requiring any modification, and enjoys a kind of robustness. Optimal block sizes are explicitly determined for block bootstrap variance estimators under regression. Finite sample performance and practical uses of the methods are illustrated through a simulation study and two data examples, respectively. Supplementary materials are available online.

KEY WORDS: MSE expansions; Optimal blocks; Simultaneous confidence bands; Tapers; Variance estimation.

1. INTRODUCTION

Block bootstraps recreate time series by resampling blocks of data and, in recent years, many variants of this method have been proposed and studied in the literature, primarily for stationary time series data. However, their performance in the presence of nonstationarity is not well understood. In this article, we investigate properties of block bootstrap methods under a specific form of nonstationarity, with data generated by a linear regression model with time-dependent errors and nonstochastic regressors. To put the findings of this article in the right perspective, we loosely classify types of block bootstrap methods into two groups—namely, *first- and second-generation* block bootstraps. Examples of first-generation block bootstraps include the moving block bootstrap (MBB) of Künsch (1989) and Liu and Singh (1992), the circular block bootstrap (CBB) of Politis and Romano (1992), and the stationary bootstrap (SB) of Politis and Romano (1994). The first-generation methods are characterized by the property that in variance estimation problems for the sample mean (and similar estimators) based on stationary observations, the bias of the block bootstrap estimators is $O(b^{-1})$, where b is the block size. It is known (cf. Hall, Horowitz, and Jing 1995; Lahiri 1999; Nordman 2009) that, asymptotically, the MBB and the CBB variance estimators are equivalent in terms of mean squared error (MSE) and both outperform the SB for a wide range of block sizes. In turn, a further improvement over all these methods is given by the tapered block bootstrap (TBB) of Paparoditis and Politis (2001, 2002). This bootstrap smooths data blocks with a spectral window and the resulting TBB variance estimator for the sample mean has a smaller bias, namely $O(b^{-2})$, compared with the first-generation block bootstraps. In this sense, the TBB is a *second-generation* block bootstrap. The objective of this article is to explore the extent to which such optimality properties of the two classes of methods carry over to the nonstationary case under the time series regression model.

Nonstationarity of the regression model interacts with certain inherent features of blocking mechanisms and leads to nonstandard behavior of the block bootstrap methods for estimation of various process parameters. To highlight the major issues with the blocking mechanisms of various first- and second-generation block bootstraps in the regression setup and to present the findings in a transparent manner, we restrict attention to the prototypical problem of variance estimation of the ordinary least squares (OLS) estimator of the unknown regression parameter vector. Specifically, we consider the performance of the CBB, MBB, SB, and TBB for variance estimation in the regression model. In contrast to the stationary case, the TBB variance estimator often turns out to be *invalid* with general nonrandom regressors (i.e., outside sample means). We introduce a modified TBB (MTBB) based on additional block randomization. Under mild dependence assumptions, the MTBB produces consistent variance estimators and, more importantly, retains desirable bias properties for virtually *any* nonstochastic regressors considered. While the CBB/MBB methods can exhibit unusual behavior in their biases and even inconsistency, the SB stands as the one block bootstrap that, in its original formulation, is robust under regression and is shown to be consistent, regardless of the regressors. Findings obtained here have ramifications on the optimality of block bootstrap methods for other functionals as well.

We end this section by briefly recalling resampling schemes for linear regression models. Section 2 describes the block bootstraps under the regression model and the variance estimation problem. Bias and variance properties, along with optimal block sizes, are presented in Section 3. Simulation studies in Section 4 then illustrate the performance of block bootstrap estimators for variance estimation and also consider block bootstraps for

Daniel J. Nordman, Associate Professor, Department of Statistics, Iowa State University, Ames, IA 50011 (E-mail: *dnordman@iastate.edu*). Soumendra N. Lahiri, Professor, Department of Statistics, Texas A&M University, College Station, TX 77843 (E-mail: *snlahiri@stat.tam.edu*). The authors thank Professor Len Stefanski and an associate editor for their important contributions, as well as two referees for thoughtful comments, which significantly improved the article. This research is partially supported by NSF grants DMS 0707139, 1007703, and 0906588. **AMS (2000) subject classification:** Primary 62G09; Secondary 62G20; 62J05; 62M10.

Journal of the American Statistical Association
March 2012, Vol. 107, No. 497, Theory and Methods
DOI: 10.1080/01621459.2011.646929

approximating the distribution of studentized OLS estimators. Section 5 provides data examples to illustrate the methods, and Section 6 gives concluding remarks. A supplementary appendix contains all proofs and additional summaries of simulation results.

Two resampling schemes are well known for bootstrapping linear regression models with independent data (cf. Freedman 1981; Efron 1982). As described by Efron (1982, pp. 35–36), one resampling plan draws from estimated residuals randomly and forms a bootstrap sample by treating regressors as fixed, while a second approach resamples vectors of response and regressor variables simultaneously. Our treatment of time series regression is based on the first resampling technique, which has not been well studied for the block bootstrap. In contrast, Fitzenberger (1997) had extended the MBB for inference in linear regression models under the second resampling plan. See Lahiri (1992, 1996), Politis, Romano, and Wolf (1999, sec. 4), and Heagerty and Lumley (2000) for similar results for the MBB and subsampling with sample averages. For models with stochastic regressors, the MBB from Fitzenberger (1997) and the subsampling approach of Politis, Romano, and Wolf (1997) present alternatives to more well-known heteroscedasticity and autocorrelation consistent (HAC) standard errors for OLS (cf. Andrews 1991); see Fitzenberger (1997) and Robinson and Velasco (1997) for a survey of HAC literature. Gonçalves and White (2005) had established the consistency of the MBB variance estimator for OLS estimators in dynamic linear models. In regression models with variables described by linear processes, Hidalgo (2003) proposed a bootstrap based on resampling in the frequency domain. For the types of regression models examined here, the block bootstraps have the advantage of providing consistent inference for various features of the sampling distribution of OLS estimators, not limited to just variance estimation, under very weak process assumptions.

2. BLOCK BOOTSTRAP AND LINEAR REGRESSION

After describing the variance estimation scenario in Section 2.1, Sections 2.2 and 2.3 describe the first-generation block bootstraps (CBB, MBB, and SB), and Section 2.4 introduces an MTBB for regression models.

2.1 Variance Estimation Problem

Consider a length n time series stretch $Y_1, \ldots, Y_n$, where

$$Y_i = \boldsymbol{x}_{i,n}'\boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{1}$$

for some $\boldsymbol{\beta} \in \mathbb{R}^q$. We assume that $\{\varepsilon_t : t \in \mathbb{Z}\}$ is a real-valued, strictly stationary error process with mean zero $E(\varepsilon_t) = 0$ and covariances $r(k) = \text{cov}(\varepsilon_t, \varepsilon_{t+k})$, $k \in \mathbb{Z}$, and that $\{\{\boldsymbol{x}_{i,n}\}_{i=1}^n : n \geq 1\}$ represents a triangular array of nonstochastic regressor variables in $\mathbb{R}^q$ with positive definite $\boldsymbol{A}_n \equiv \sum_{i=1}^n \boldsymbol{x}_{i,n}\boldsymbol{x}_{i,n}'$. Denote the OLS estimator of $\boldsymbol{\beta}$ as $\hat{\boldsymbol{\beta}}_n = \boldsymbol{A}_n^{-1}\sum_{i=1}^n \boldsymbol{x}_{i,n}Y_i$. We shall investigate block bootstrap estimators of the scaled variance $\boldsymbol{\Sigma}_n \equiv \text{var}(\boldsymbol{A}_n^{1/2}\hat{\boldsymbol{\beta}}_n)$ of $\hat{\boldsymbol{\beta}}_n$, where scaling $\boldsymbol{A}_n^{1/2}$ ensures that this covariance matrix is bounded under weak time dependence (i.e., if $\sum_{k=0}^\infty |r(k)| < \infty$). To motivate the block bootstrap estimators, it will be helpful to rewrite $\boldsymbol{\Sigma}_n$ as

$$\boldsymbol{\Sigma}_n = \text{var}\left(\sum_{i=1}^n \boldsymbol{A}_n^{-1/2}\boldsymbol{x}_{i,n}\varepsilon_i\right) \equiv \sum_{k=0}^{n-1} \boldsymbol{Q}_{k,n} r(k), \tag{2}$$

using a system of weighting matrices $\{\boldsymbol{Q}_{k,n}\}_{k=0}^{n-1}$ based on regressor cross-product sums. Here $\boldsymbol{Q}_{0,n} = I_{q\times q}$ is the $q \times q$ identity matrix, and for $k \geq 1$, we define

$$\boldsymbol{Q}_{k,n} \equiv \sum_{i=1}^{n-k} \boldsymbol{M}_n(i, i+k), \tag{3}$$

for $\boldsymbol{M}_n(i,j) \equiv \boldsymbol{A}_n^{-1/2}(\boldsymbol{x}_{i,n}\boldsymbol{x}_{j,n}' + \boldsymbol{x}_{j,n}\boldsymbol{x}_{i,n}')\boldsymbol{A}_n^{-1/2}$, $j > i \geq 1$.

In the regression setting, the quality of a block bootstrap estimator depends on how well the block resampling mechanism captures *both* the process covariances $r(k)$ *and* the regressor weights $\boldsymbol{Q}_{k,n}$ in (2) at "small" lags k. In contrast, typical comparisons among block bootstrap variance estimators for a real-valued sample mean $\hat{\boldsymbol{\beta}}_n = \bar{Y}_n$ (i.e., $\boldsymbol{x}_{i,n} = 1$) reduce to quantifying performance in covariance $r(k)$ estimation, since regressor weights $\boldsymbol{Q}_{k,n} \approx 1$ at small lags are negligible in this case.

2.2 Moving (MBB) and Circular Block (CBB) Bootstraps

Block bootstraps require data blocks $\mathcal{B}(i,j) = (e_i, \ldots, e_{i+j-1})$, $i, j \geq 1$, of consecutive OLS residuals given by $e_i \equiv Y_i - \boldsymbol{x}_{i,n}'\hat{\boldsymbol{\beta}}_n$ for $1 \leq i \leq n$ and periodically extended as $e_i = e_{i-n}$, $i > n$. All bootstraps involve an integer block length $b \in [1, n]$, where $b \to \infty$ with $b/n \to 0$ as $n \to \infty$ to ensure small blocks relative to the sample size.

To resample data blocks in the MBB, let $\ell \equiv \lfloor n/b \rfloor$ using the floor function $\lfloor\cdot\rfloor$ and generate iid random variables $\{I_1, \ldots, I_{\ell+1}\}$ with a uniform distribution on the index set $\{1, \ldots, n-b+1\}$ of all length b blocks $\{\mathcal{B}(i,b) : i = 1, \ldots, n-b+1\}$ among the residuals $e_1, \ldots, e_n$. The MBB error variables $\varepsilon_1^*, \ldots, \varepsilon_n^*$ are defined by concatenating resampled blocks $\{\mathcal{B}(I_j, b) : j = 1, \ldots, \ell+1\}$, where the last block is truncated to have length $R = n - \ell b < b$ when $n \neq \ell b$. Define

$$Y_i^* = \boldsymbol{x}_i'\hat{\boldsymbol{\beta}}_n + (\varepsilon_i^* - E_*\varepsilon_i^*), \quad i = 1, \ldots, n, \tag{4}$$

with the OLS estimator $\hat{\boldsymbol{\beta}}_n$ playing the role of $\boldsymbol{\beta}$ in the bootstrap world and using E_* and var_* to denote bootstrap expectation and variance conditioned on the data. Letting $\hat{\boldsymbol{\beta}}_n^*$ denote the bootstrap version of $\hat{\boldsymbol{\beta}}_n$ (defined by Y_i^*'s in place of Y_i's), the MBB estimator of the OLS variance $\boldsymbol{\Sigma}_n$ in (2), denoted by $\hat{\boldsymbol{\Sigma}}_{n,\text{MBB}}(b)$, is then

$$E_*\big(\big[\boldsymbol{A}_n^{1/2}(\hat{\boldsymbol{\beta}}_n^* - \hat{\boldsymbol{\beta}}_n)\big]\big[\boldsymbol{A}_n^{1/2}(\hat{\boldsymbol{\beta}}_n^* - \hat{\boldsymbol{\beta}}_n)\big]'\big) = \text{var}_*\left(\boldsymbol{A}_n^{-1/2}\sum_{i=1}^n \boldsymbol{x}_{i,n}(\varepsilon_i^* - E_*\varepsilon_i^*)\right). \tag{5}$$

The CBB of Politis and Romano (1992) has analogous block resampling, but with respect to the circular block collection $\{\mathcal{B}(i,b) : i = 1, \ldots, n\}$ of periodically extended residuals $e_i \equiv e_{i-n}$, $i > n$. Then, taking $\{I_1, \ldots, I_{\ell+1}\}$ as iid and uniform on $\{1, \ldots, n\}$ similarly defines the CBB innovations $\varepsilon_1^*, \ldots, \varepsilon_n^*$. The CBB variance estimator $\hat{\boldsymbol{\Sigma}}_{n,\text{CBB}}(b)$ from (5) can be concisely expressed as

$$\hat{\boldsymbol{\Sigma}}_{n,\text{CBB}}(b) = \sum_{k=0}^{b-1} \boldsymbol{Q}^{\text{CBB}}_{k,b,n} \hat{r}^{\text{CBB}}_n(k), \qquad \hat{r}^{\text{CBB}}_n(k) \equiv \hat{r}_n(k) + \hat{r}_n(n-k), \tag{6}$$

with sample covariances $\hat{r}_n(k) = n^{-1}\sum_{i=1}^{n-k}(e_i - \bar{e}_n)(e_{i+k} - \bar{e}_n)$, $0 \le k < n$, and $\bar{e}_n = n^{-1}\sum_{i=1}^{n} e_i$, which follows because CBB covariances $\text{cov}_*(\varepsilon^*_j, \varepsilon^*_{j+k}) = \hat{r}^{\text{CBB}}_n(k)$ do not depend on the index j, where $1 \le j \le j+k \le b$. The CBB variance (6) involves a bootstrap analog $\boldsymbol{Q}^{\text{CBB}}_{k,b,n} = \boldsymbol{Q}_{k,n} - \boldsymbol{D}^{\text{CBB}}_{k,b,n}$ of covariance weights $\boldsymbol{Q}_{k,n}$ in the OLS variance (2), where

$$\boldsymbol{D}^{\text{CBB}}_{k,b,n} \equiv \left(\sum_{i=1}^{\ell-1}\sum_{j=1}^{k} \boldsymbol{M}_n(bi+j-k, bi+j)\right) + \boldsymbol{\Delta}^{\text{CBB}}_{k,b,n}, \quad 1 \le k < b, \tag{7}$$

and $\boldsymbol{\Delta}^{\text{CBB}}_{k,b,n} \equiv \sum_{j=1}^{\min\{k,R\}} \boldsymbol{M}_n(b\ell + j - k, b\ell + j)$ is taken to be zero whenever $R = n - b\ell$ is zero (i.e., b evenly divides n). Note that, unlike $\boldsymbol{Q}_{k,n}$, terms $\boldsymbol{Q}^{\text{CBB}}_{k,b,n}$ depend on the block size, except for $\boldsymbol{Q}_{0,n} = \boldsymbol{Q}^{\text{CBB}}_{0,n}$ when $k = 0$. The important point is that CBB versions $\boldsymbol{Q}^{\text{CBB}}_{k,b,n}$ fail to match regressor weights $\boldsymbol{Q}_{k,n}$ because independent block resampling causes the CBB to miss regressor cross-products $\boldsymbol{A}_n^{-1/2}\boldsymbol{x}_{i,n}\boldsymbol{x}'_{i+k,n}\boldsymbol{A}_n^{-1/2}$ at a lag $1 \le k < b$ whenever bootstrap responses Y^*_i, Y^*_{i+k} lie in different blocks and are consequently independent. Missed cross-products at the edges of concatenated blocks are collected in the differences $\boldsymbol{D}^{\text{CBB}}_{k,b,n}$ from (7), and Section 3 summarizes how these impact the bias of CBB/MBB estimators.

2.3 Stationary Bootstrap (SB)

The SB of Politis and Romano (1994) generates innovations $\varepsilon^*_1, \dots, \varepsilon^*_n$ in the bootstrap regression model (4) by resampling data blocks of random length. Let $W_1, \dots, W_n$ be iid geometric random variables, independent of the data, where the event $\{W_1 = k\}$, $k \ge 1$, has probability pq^{k-1} for $p = b^{-1} \in (0,1)$, $q = 1 - p$. For $K = \min\{k : \sum_{i=1}^{k} W_i \ge n\}$, let $I_1, \dots, I_K$ be iid uniform variables on $\{1, \dots, n\}$. The SB innovations $\varepsilon^*_1, \dots, \varepsilon^*_n$ are then given by the first n observations in $\mathcal{B}(I_1, W_1), \dots, \mathcal{B}(I_K, W_K)$, and $b = p^{-1}$ represents the expected length of a resampled block. The resulting SB innovations exhibit stationarity and the SB variance estimator from (5) may be written as

$$\begin{aligned}\hat{\boldsymbol{\Sigma}}_{n,\text{SB}}(b) &= \sum_{k=0}^{n-1} \boldsymbol{Q}_{k,n}\hat{r}^{\text{SB}}_n(k)\\ &= \sum_{k=0}^{n-1}\left(q^k \boldsymbol{Q}_{k,n} + (1-q)^{n-k}\boldsymbol{Q}_{n-k,n}\right)\hat{r}_n(k),\end{aligned} \tag{8}$$

using SB covariances $\text{cov}_*(\varepsilon^*_1, \varepsilon^*_{1+k}) \equiv \hat{r}^{\text{SB}}_n(k) = q^k[\hat{r}_n(k) + \hat{r}_n(n-k)]$ and $\boldsymbol{Q}_{n,n} = \boldsymbol{0}_{q\times q}$ (see Politis and Romano 1994). Hence, by the induced stationarity, the estimator (8) correctly captures the regressor weights $\boldsymbol{Q}_{k,n}$ in the OLS variance (2) at all lags.

2.4 Tapered Block Bootstrap (TBB)

2.4.1 The Original Formulation. The TBB requires a sequence of tapering windows $w_b(t) = w(\frac{t-0.5}{b})$, $t \in \mathbb{R}$, $b = 1, 2, \dots$, similar to those used in frequency-domain estimation with time series (cf. Brillinger 1981; Dahlhaus 1983), where the kernel $w : \mathbb{R} \to [0,1]$ is positive in a neighborhood of $1/2$, symmetric around $1/2$, and nondecreasing on $[0, 1/2]$ with $w(t) = 0$ if $t \notin [0,1]$. Following Künsch (1989) and Paparoditis and Politis (2001), we consider the so-called "smooth tapers" having self-convolutions $w * w(t)$, $t \in \mathbb{R}$ satisfying

$$w * w(t) = \int_{-1}^{1} w(x)w(x + |t|)dx$$
$$\text{is twice continuously differentiable at } t = 0. \tag{9}$$

Letting $\mathbb{I}(\cdot)$ denote the indicator function, a second (unsmooth) taper is given by $w(t) = \mathbb{I}(t \in [0,1])$ for producing untapered data blocks in the following.

To motivate a TBB method in the regression problem, we recall the TBB of Paparoditis and Politis (2001, 2002) that uses the block resampling scheme of the MBB method, but each resampled MBB block $\mathcal{B}(I_j, b) = (e_{I_j}, \dots, e_{I_j+b-1})$ is replaced by a tapered version

$$\mathcal{B}_{\text{taper}}(I_j, b) \equiv \frac{\sqrt{b}}{\sqrt{v_b(0)}}\left(w_b(1)e_{I_j}, \dots, w_b(b)e_{I_j+b-1}\right),$$

scaling the ith observation in $B(I_j, b)$ by $w_b(i)\sqrt{b/v_b(0)}$, where $v_b(0) = \sum_{i=1}^{b} w_b^2(i)$. A smooth taper aims to give reduced weight to observations near the ends of a data block, and the TBB reduces to the MBB when $w(t) = \mathbb{I}(t \in [0,1])$. The resulting TBB errors $\varepsilon^*_1, \dots, \varepsilon^*_n$ produce an estimator of $\boldsymbol{\Sigma}_n$, say $\hat{\boldsymbol{\Sigma}}_{n,\text{TBB}}(b)$, analogously to (5). Recall that this variance estimator was originally formulated for the sample mean $\hat{\boldsymbol{\beta}}_n \equiv n^{-1}\sum_{i=1}^{n} Y_i$ of a stretch $Y_i = \boldsymbol{\beta} + \varepsilon_i$, $i = 1, \dots, n$ (i.e., $\boldsymbol{\beta} \in \mathbb{R}$, $\boldsymbol{x}_{i,n} = 1$, $\boldsymbol{A}_n^{-1/2}\boldsymbol{x}_{i,n} = 1/\sqrt{n}$), and in this case, the bias $O(1/b^2)$ of the TBB is known to beat the bias $O(1/b)$ of the MBB (cf. Paparoditis and Politis 2001).

2.4.2 Inconsistency of the TBB in Regression. The estimator $\hat{\boldsymbol{\Sigma}}_{n,\text{TBB}}(b)$ appears natural under the general regression model (4), but its superior bias properties can collapse outside of the sample mean case described earlier. The problem lies in matching TBB innovations ε^*_i to regressors $\boldsymbol{x}_{i,n}$, which inherently assigns all regressors of form $\boldsymbol{x}_{(i-1)b+j}$ to a taper weight $w_b(j)$, $j = 1, \dots, b$, $i \ge 1$. Such "matching" issues do not arise with either constant regressors $\boldsymbol{x}_{i,n} = 1$ (i.e., sample means) or constant weights $w_b(j) = 1$ (i.e., untapered blocks). But when the sequence of regressors varies and the weights $w_b(\cdot)$ are nonconstant as for a smooth taper, the TBB can be asymptotically *biased* and inconsistent. Assuming a smooth taper $w(\cdot)$ with $n = b\ell$, $b^{-1} + b^3/n \to 0$ and sufficient moment conditions (e.g., Theorem 1, Section 3.1), the bias of TBB variance estimator can be expanded as $E\hat{\boldsymbol{\Sigma}}_{n,\text{TBB}}(b) - \boldsymbol{\Sigma}_n = O(1/b^2) - \boldsymbol{C}_{n,\text{TBB}}(w, b, \boldsymbol{x})$, where $\boldsymbol{C}_{n,\text{TBB}}(w, b, \boldsymbol{x})$ defined as

$$\sum_{k=0}^{b-1}\frac{r(k)}{v_b(0)}\sum_{i=1}^{b-k} w_b(i)w_b(i+k)$$
$$\times\left[\boldsymbol{Q}_{k,n} - b\sum_{j=1}^{\ell}\boldsymbol{M}_n((j-1)b+i, (j-1)b+i+k)\right]$$

depends on the taper, block size, and regressor cross-products $\boldsymbol{M}_n(\cdot,\cdot)$ from (3). Since $\lim_{n\to\infty}\text{var}(\hat{\boldsymbol{\Sigma}}_{n,\text{TBB}}(b)) = \boldsymbol{0}$ under mild

Table 1. $\boldsymbol{C}_{n,\text{TBB}}(w, b, \boldsymbol{x})$ for two regressor types $\boldsymbol{x}_{i,n}$, increasing block/sample sizes, MA(1) errors $\{\varepsilon_t\}$ ($r(0) = 2$, $r(1) = 1$), and taper $w(t) = \{1 - \cos(2\pi t)\}/2$, $t \in [0, 1]$

b	$n = b^4$	$\boldsymbol{x}_{i,n} = 1/i$	$\boldsymbol{x}_{i,n} = 1$
10	10,000	1.9181	−1.8724e-04
20	160,000	2.4974	−1.2296e-05
30	810,000	2.7199	−2.4511e-06
40	2,560,000	2.8373	−7.7804e-07

conditions, the behavior of $\boldsymbol{C}_{n,\text{TBB}}(w, b, \boldsymbol{x})$ determines the consistency of the TBB estimator. While not straightforward to analyze directly, Table 1 illustrates $\boldsymbol{C}_{n,\text{TBB}}(w, b, \boldsymbol{x})$ values over two sets of regressors $\boldsymbol{x}_{i,n} = i^{-1}$ and $\boldsymbol{x}_i = 1$ under a smooth taper. Although the TBB has no regressor matching problems in the sample mean case $\boldsymbol{x}_{i,n} = 1$ for which $\boldsymbol{C}_{n,\text{TBB}}(w, b, \boldsymbol{x}) \to 0$ as $n \to \infty$, $\boldsymbol{C}_{n,\text{TBB}}(w, b, \boldsymbol{x})$ fails to converge to zero for the nonconstant regressors $\boldsymbol{x}_{i,n} = i^{-1}$. The TBB is then inconsistent while any *first-generation* block bootstrap variance estimator would be consistent with these regressors (Section 3).

We also note that when the TBB is inconsistent for variance estimation, the method will generally *not* be valid for distribution estimation either. To provide a formal statement of this, let $\mathcal{D}_n$ and $\hat{\mathcal{D}}_{n,\text{TBB}}(b)$, respectively, denote the distributions of the OLS estimator $\boldsymbol{A}_n^{1/2}(\hat{\boldsymbol{\beta}}_n - \boldsymbol{\beta})$ and the TBB version $\boldsymbol{A}_n^{1/2}(\hat{\boldsymbol{\beta}}^*_{n,\text{TBB}} - \hat{\boldsymbol{\beta}}_n)$. Under suitable summability conditions on the cumulants of the innovation process (1) along with $b^{-1} + b/n \to 0$ as $n \to \infty$, if $\boldsymbol{\Sigma}_n - \hat{\boldsymbol{\Sigma}}_{n,\text{TBB}}(b) \xrightarrow{p} \boldsymbol{0}_{q\times q}$ fails, then $\varrho_q(\mathcal{D}_n, \hat{\mathcal{D}}_{n,\text{TBB}}(b)) \xrightarrow{p} 0$ is impossible and so TBB distribution estimation breaks down, where ϱ_q denotes any metric on probability distributions in $\mathbb{R}^q$, metricizing the topology of convergence in distribution. The result requires, for the regressors, only that $\boldsymbol{A}_n = \sum_{i=1}^n \boldsymbol{x}_{i,n}\boldsymbol{x}'_{i,n}$, $n \geq 1$ is positive definite; see the supplementary materials for a proof.

2.4.3 A Modified TBB. We next describe an MTBB method, which involves two rounds of block randomization. In the first step, we resample $\ell_1 = \lceil (n+b)/b \rceil$ blocks (possibly one block more than the CBB/MBB). Let $\{I_j : j = 1, \ldots, \ell_1\}$ be iid uniform on $\{1, \ldots, n-b+1\}$ and concatenate the resampled, tapered blocks $\mathcal{B}_{\text{taper}}(I_j, b)$ to produce an initial sequence of length $b\ell_1$, denoted as $\acute{\varepsilon}^*_1, \ldots, \acute{\varepsilon}^*_{b\ell_1}$. Center these by their bootstrap expectation, say $\tilde{\varepsilon}^*_i = \acute{\varepsilon}^*_i - E_*\acute{\varepsilon}^*_i$, $i = 1, \ldots, b\ell_1$, and wrap as $\tilde{\varepsilon}^*_i = \tilde{\varepsilon}^*_{i-b\ell_1}$, $i > b\ell_1$. In the second randomization step, independent from data $\{Y_i\}_{i=1}^n$ and resampled blocks $\{I_j\}_{j=1}^{b\ell_1}$, generate a random variable I with a uniform distribution on $\{1, \ldots, b\ell_1\}$ and define the MTBB errors $\varepsilon^*_1, \ldots, \varepsilon^*_n$ in (4) as the first n observations from $\tilde{\varepsilon}^*_I, \ldots, \tilde{\varepsilon}^*_{I+b\ell_1-1}$.

The extra randomization step avoids alignment problems between regressors and tapering weights $w_b(\cdot)$. From (5), we obtain an MTBB variance estimator $\hat{\boldsymbol{\Sigma}}_{n,\text{MTBB}}(b)$, which interestingly has a closed form

$$\hat{\boldsymbol{\Sigma}}_{n,\text{MTBB}}(b) = \sum_{k=0}^{b-1} \boldsymbol{Q}_{k,n} \frac{v_b(k)}{v_b(0)} \hat{r}_n^{\text{MTBB}}(k), \tag{10}$$

involving the correct cross-product sums $\boldsymbol{Q}_{k,n}$ from the OLS variance (2), lag weights $v_b(k)/v_b(0)$ for $v_b(j) = \sum_{i=1}^{b-j} w_b(i)w_b(i+j)$, $0 \leq j < b$, and tapered estimators

$$\hat{r}_n^{\text{MTBB}}(k) = \frac{1}{n-b+1} \sum_{i=0}^{n-b} \sum_{j=1}^{b-k} w_b(j)w_b(j+k)(e_{i+j} - \bar{e}_j) \times (e_{i+j+k} - \bar{e}_{j+k})/v_n(k)$$

of covariances based on residual means $\bar{e}_j = \sum_{i=0}^{n-b} e_{i+j}/(n - b + 1)$, $1 \leq j \leq b$.

In the following, we will refer to (10) as the *MTBB variance estimator* and refer to (10) when $w(t) = \mathbb{I}(t \in [0, 1])$ as the *modified MBB (MMBB) variance estimator*.

Remark 1. The MTBB variance estimator has a close correspondence (in fact, an asymptotic equivalence) to a type of kernel estimator $\sum_{k=0}^{b-1} \boldsymbol{Q}_{k,n} u(k/b)\hat{r}_n(k)$ for the OLS variance (2), where $u(k/b)$ are lag weights defined by a kernel $u : [0, 1] \to [0, 1]$ with $u(0) = 1$, b is a bandwidth, and $\hat{r}_n(k)$ are sample covariances of residuals. For sample means, the same connection between MBB/TBB variance estimators and kernel estimators has been noted by Künsch (1989, thm. 3), Paparoditis and Politis (2001), and Politis (2003). There, the MBB is similar to a first-order kernel estimator using Bartlett's window $u(t) = 1 - |t|$, while the TBB variance estimator for the sample mean resembles a second-order kernel estimator with $u(t) = 1 + O(t^2)$ as $t \to 0$ (cf. Parzen 1957), which improves the bias of the TBB. The same equivalences to kernel variance estimators now carry over to the block bootstraps in the regression setting. However, the advantage of the MTBB over kernel variance estimators is that the modified block resampling mechanism can be applied to estimate other features of the sampling distribution of OLS estimators, such as quantiles for calibrating confidence intervals (CIs). We illustrate this through simulation in Section 4.2.

3. RESULTS ON MSE EXPANSIONS

3.1 Bias Expansions

The block bootstrap estimators differ largely in their bias expressions. To state these, we require Assumptions A(i), B(i) (stated as functions of an integer $i \geq 1$), and C:

A(i): $|r(0)| + \sum_{k=1}^{\infty} k^i |r(k)| < \infty$.
B(i): As $n \to \infty$, $b^{-1} + b^i/n \to 0$.
C: For smooth tapers, (9) holds denoting this second derivative at $t = 0$ as $w * w''(0)$.

We make no assumptions on the nonstochastic regressors $\{\boldsymbol{x}_{i,n}\}_{i=1}^n$ outside of $\boldsymbol{A}_n = \sum_{i=1}^n \boldsymbol{x}_{i,n}\boldsymbol{x}'_{i,n}$ being positive definite. However, we may define *standard regressors* as satisfying $\max_{1\leq i\leq n} \|\boldsymbol{A}_n^{-1/2}\boldsymbol{x}_{i,n}\|^2 n = O(1)$, which often allows limit theorems for OLS estimators under mixing conditions (cf. Lahiri 2003b). This includes cases where $\boldsymbol{A}_n/n$ has a positive definite limit, as in the sample mean case. We shall refer to *nonstandard regressors* when this condition is not satisfied.

The following result gives the leading terms in the bias expansions for block bootstrap variance estimators. The conditions on the range of block sizes in each case are inclusive of the order of the optimal block sizes (cf. Section 3.3). Let $\|\boldsymbol{A}\|$ denote the spectral norm of matrix $\boldsymbol{A}$.

Theorem 1. Suppose that $\boldsymbol{A}_n$ is positive definite for $n \geq 1$ and that Assumption A(1) holds. Also, suppose that for the CBB and the MBB, Assumption B(1) holds; for the SB and MMBB, Assumption B(2) holds; and for the MTBB, Assumptions A(2), B(3), and C hold. Then, the bias of a bootstrap variance estimator $\hat{\boldsymbol{\Sigma}}(b)$ of $\boldsymbol{\Sigma}_n$ is given by

$$
\begin{aligned}
&E[\hat{\boldsymbol{\Sigma}}(b) - \boldsymbol{\Sigma}_n] \\
&= \begin{cases}
-\displaystyle\sum_{k=1}^{b-1} \boldsymbol{D}_{k,b,n}^{\text{CBB}} r(k) & \\
\quad + O\left(\frac{\sum_{k=0}^{b-1} \|\boldsymbol{Q}_{k,b,n}^{\text{CBB}}\|}{n} + \left\|\sum_{k=b}^{n-1} \boldsymbol{Q}_{k,n} r(k)\right\|\right) & \text{CBB}, \\
-\displaystyle\sum_{k=1}^{b-1} \boldsymbol{D}_{k,b,n}^{\text{CBB}} r(k) & \\
\quad + O\left(\frac{\sum_{k=0}^{b-1} \|\boldsymbol{Q}_{k,b,n}^{\text{MBB}}\|}{n} + \left\|\sum_{k=b}^{n-1} \boldsymbol{Q}_{k,n} r(k)\right\|\right) & \text{MBB}, \\
-\dfrac{1}{b}\displaystyle\sum_{k=1}^{b-1} \boldsymbol{Q}_{k,n} k r(k) + o(b^{-1}) & \text{SB}, \\
-\dfrac{1}{b}\displaystyle\sum_{k=1}^{b-1} \boldsymbol{Q}_{k,n} k r(k) + o(b^{-1}) & \text{MMBB}, \\
-\dfrac{1}{b^2}\dfrac{w * w''(0)}{2w * w(0)}\displaystyle\sum_{k=1}^{b-1} & \text{MTBB}. \\
\quad \times \boldsymbol{Q}_{k,n} k^2 r(k) + o(b^{-2}) &
\end{cases}
\end{aligned}
$$

Remark 2. To interpret the bias expansions, note that the weight matrices $\boldsymbol{Q}_{k,n}$ from (3), $\boldsymbol{Q}_{k,b,n}^{\text{CBB}}$ from (6), and $\boldsymbol{D}_{k,b,n}^{\text{CBB}}$ from (7) are *uniformly bounded* for any regressors (e.g., by the Cauchy–Schwarz inequality, $\sup_n \max_{1\leq k<n} \|\boldsymbol{Q}_{k,n}\| < C$). MBB weights $\boldsymbol{Q}_{k,b,n}^{\text{MBB}} \equiv \sum_{i=0}^{\ell}\sum_{j=1}^{b-k} \mathbb{I}(ib + j + k \leq n)\boldsymbol{M}_n(ib + j, ib + j + k)c_{j,k,n}$, $0 \leq k < b$, involve terms $\{c_{j,k,n}\}_{j=1}^{b-k}$ that depend on the covariance structure and are uniformly bounded ($\sup_n \max_{j,k} |c_{j,k,n}| < \infty$). For *standard* regressors, we may replace block condition B(1) with B(2) and reformulate the bias of CBB/MBB estimators in Theorem 1 as

$$
\begin{aligned}
E\hat{\boldsymbol{\Sigma}}_n(b) - \boldsymbol{\Sigma}_n &= -\frac{1}{b}\sum_{k=1}^{b-1} \boldsymbol{U}_{k,b,n} k r(k) + o(1/b), \\
\max_{1\leq k\leq b-1} \|\boldsymbol{U}_{k,b,n}\| &= O(1),
\end{aligned}
$$

for $\boldsymbol{U}_{k,b,n} \equiv b\boldsymbol{D}_{k,b,n}^{\text{CBB}}/k$, which gives a bias form resembling that of the SB and the MMBB. That is, the order of the bias for CBB/MBB methods is $O(1/b)$ for standard regressors. For nonstandard regressors, biases of unusual order can result for the CBB/MBB; see Example 1.

Remark 3. Theorem 1 has several implications. We list some of these next.

3.1. Across a large variety of regressor sequences, the SB and the MMBB have the same bias $O(1/b)$ while the MTBB with a smooth taper has bias $O(1/b^2)$. Thus, the higher accuracy associated with the TBB under smooth tapering in the sample mean case is extended to the regression setting by the MTBB.

3.2. While SB, MMBB, and MTBB estimators are guaranteed to be asymptotically unbiased for any regressors with positive definite $\boldsymbol{A}_n$, this is not generally true for CBB/MBB estimators outside of standard regressors. As an example of nonstandard regressors for which the CBB/MBB estimators are asymptotically *biased* and inconsistent, consider regressors that are zero except for consecutive observations lying at the end/beginning of two separate blocks $\boldsymbol{x}_{ib,n} = \boldsymbol{x}_{ib+1,n} = 1$, $i = 1, \ldots, \lfloor n/b \rfloor$. Then, in the OLS target variance (2), the only nonzero regressor weight $\boldsymbol{Q}_{1,n} = 2$ is determined by regressors in differing blocks, which the CBB reconstruction $\boldsymbol{Q}_{1,b,n}^{\text{CBB}} = 0$ misses. Consequently, the CBB bias $-2r(1) + O(1/n)$ is nonnegligible, following from Theorem 1 with $\boldsymbol{D}_{1,b,n}^{\text{CBB}} = 2$ and $\boldsymbol{D}_{k,b,n}^{\text{CBB}} = 0$, $k > 1$, in (7); the MBB bias has this same form.

3.3. Note that the SB has the same leading term in the bias expansion as the MMBB. Thus, the SB works for the same large class of regressors as does the MMBB with the same level of accuracy, but without any modification for the regression setting. The SB has a kind of automatic robustness built into its blocking mechanism that the other block bootstraps lack in their original formulation.

In summary, the MTBB has the best bias of all methods considered in the standard case. First-generation block bootstraps based on nonrandom blocks (e.g., MBB) have accuracies that depend heavily on the underlying regressors. The additional block randomization step in resampling helps to stabilize their bias for the large class of standard regressors (e.g., MMBB).

Example 1. (real-valued polynomial regressors). In the following, $\to$ denotes convergence as $n \to \infty$ and $s_n \sim t_n$ denotes $s_n/t_n \to 1$. Suppose for some $c \in \mathbb{R}$ that $\boldsymbol{x}_{i,n} = i^c$ for $i = 1, \ldots, n$. For the SB/MMBB/MTBB bias, note in (3) that $\boldsymbol{Q}_{k,n} \to 2$ for each fixed $k \geq 1$ if $c \geq -1/2$ and $\boldsymbol{Q}_{k,n} \to 2\sum_{i=1}^{\infty} i^c(i + k)^c / \sum_{i=1}^{\infty} i^{2c}$ if $c < -1/2$. For the CBB/MBB bias, note in (7) that $\boldsymbol{D}_{k,b,n}^{\text{CBB}} \sim 2k/b$ if $c \geq -1/2$ and $\boldsymbol{D}_{k,b,n}^{\text{CBB}} \sim 2k/b^{|2c|}$ if $c < -1/2$. Hence, for standard regressors ($c \geq -1/2$) and a block satisfying $B(2)$ ($b^2/n \to 0$), the CBB, MBB, SB, and MMBB have the same bias

$$-\frac{2}{b}\sum_{k=1}^{\infty} k r(k) + o(1/b),$$

as in the sample mean case (i.e., $c = 0$) by Theorem 1. However, when regressors are nonstandard ($c < -1/2$) and $b^{1+|2c|}/n \to 0$, then CBB/MBB bias becomes

$$-\frac{2}{b^{|2c|}}\sum_{k=1}^{\infty} k r(k) + o(1/b^{|2c|}),$$

which is better than the $O(1/b)$ bias of the SB/MMBB and also, if $c < -1$, the $O(1/b^2)$ bias of the MTBB.

Example 2. Consider nonstandard regressors $\boldsymbol{x}_{i,n} = e^{n-i}$, $i = 1, \ldots, n$. Then, the SB/MMBB and MTBB estimators have biases, respectively, given by

$$-\frac{2}{b}\sum_{k=1}^{\infty} e^{-k} k r(k) + o(b^{-1}),$$

$$-\frac{1}{b^2}\frac{w * w''(0)}{w * w(0)}\sum_{k=1}^{\infty} e^{-k} k^2 r(k) + o(b^{-2}),$$

using that $\boldsymbol{Q}_{k,n} \to 2e^{-k}$ for each $k \geq 1$. For determining the CBB/MBB bias from Theorem 1, it holds that $\max\{|\boldsymbol{Q}^{\mathrm{MBB}}_{k,b,n}|, |\boldsymbol{Q}^{\mathrm{CBB}}_{k,b,n}|\} \leq Ce^{-k}$ for $1 \leq k < n$ (some $C > 0$) and that $\boldsymbol{D}^{\mathrm{CBB}}_{k,b,n} \sim 2(e^k - e^{-k})/e^{2b}$ when $b/n \to 0$. In choosing a block b proportional to $\log n$, the bias of CBB/MBB estimators becomes $O(1/n)$ in comparison.

3.2 Variance Expansions

A general expression is next provided for the variance of block bootstrap estimators. For concreteness, we consider a bootstrap estimator $\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n$ of the variance $\boldsymbol{a}_n'\boldsymbol{\Sigma}_n\boldsymbol{a}_n = \mathrm{var}(\boldsymbol{a}_n'\boldsymbol{A}_n^{1/2}\hat{\boldsymbol{\beta}}_n)$ of a linear combination $\boldsymbol{a}_n'\boldsymbol{A}_n^{-1/2}\hat{\boldsymbol{\beta}}_n$ of OLS estimators, along a bounded sequence $\boldsymbol{a}_n \in \mathbb{R}^q$ with $\|\boldsymbol{a}_n\| \leq C$ for some $C > 0$. To do so, some notation is required. Define a nonnegative array $\{v_{k,n} : k = 1, \ldots, n-1\}$ that depends on the block bootstrap type, regressors $\boldsymbol{x}_{i,n}$, and $\boldsymbol{a}_n \in \mathbb{R}^q$ as follows:

$$v_{k,n} \equiv \begin{cases} \boldsymbol{a}_n'\boldsymbol{Q}^{\mathrm{CBB}}_{k,b,n}\boldsymbol{a}_n\mathbb{I}(1 \leq k < b \text{ or} & \text{CBB}, \\ \quad n-b+1 \leq k < n) & \\ \boldsymbol{a}_n'\boldsymbol{Q}^{\mathrm{CBB}}_{k,b,n}\boldsymbol{a}_n\mathbb{I}(1 \leq k < b) & \text{MBB}, \\ q^k\boldsymbol{a}_n'\boldsymbol{Q}_{k,n}\boldsymbol{a}_n + (1-q)^{n-k}\boldsymbol{a}_n'\boldsymbol{Q}_{n-k,n}\boldsymbol{a}_n & \text{SB}, \\ (1-b^{-1}k)\boldsymbol{a}_n'\boldsymbol{Q}_{k,n}\boldsymbol{a}_n\mathbb{I}(1 \leq k < b) & \text{MMBB}, \\ \dfrac{v_b(k)}{v_b(0)}\boldsymbol{a}_n'\boldsymbol{Q}_{k,n}\boldsymbol{a}_n\mathbb{I}(1 \leq k < b) & \text{MTBB}, \end{cases}$$

where $\mathbb{I}(\cdot)$ denotes the indicator function. For MMBB/MTBB estimators of $\boldsymbol{a}_n'\boldsymbol{\Sigma}_n\boldsymbol{a}_n$, terms $v_{k,n}$ are the weights on covariance estimators $\hat{r}_n^{\mathrm{MTBB}}(k)$ appearing in $\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_{n,\mathrm{MTBB}}(b)\boldsymbol{a}_n$ from (10); for CBB/SB estimators, $v_{k,n}$ are weights on $\hat{r}_n(k)$ after arranging (6) or (8) as sums of sample covariances. The variance expansions require a special sum

$$s_n \equiv \sum_{k=1}^{n-1} v_{k,n}^2\left(1 - \frac{k}{n}\right)^2$$

and an error term

$$\eta_n \equiv \frac{1}{n} + \frac{\|\mathbf{v}_n\|}{n}\left(\sum_{k=1}^{n-1} v_{k,n}^2\frac{k^2}{n^2}\left[1-\frac{k}{n}\right]^2\right)^{1/2} + \frac{b\|\mathbf{v}_n\|^2}{n^2} + \frac{1}{n^2}\left(\sum_{k=1}^{n-1}|v_{k,n}|\right)^2, \tag{11}$$

where $\|\mathbf{v}_n\|^2 \equiv \sum_{k=1}^{n-1} v_{k,n}^2$. Write $H_n(\lambda) = \sum_{k=1}^{n-1}(1 - n^{-1}k)v_{k,n}e^{\imath k\lambda}$, $\lambda \in \Pi = [-\pi, \pi]$, where $\imath = \sqrt{-1}$, and define a nonnegative kernel $\Phi_n(\lambda) \equiv H_n(\lambda)H_n(-\lambda)/\{2\pi s_n\}$, which satisfies $\int_\Pi \Phi_n(\lambda)d\lambda = 1$. Let $f(\lambda) = (2\pi)^{-1}\sum_{k\in\mathbb{Z}} r(k)e^{-\imath k\lambda}$, $\lambda \in \Pi$ denote the spectral density of $\{\varepsilon_t\}_{t\in\mathbb{Z}}$, and denote the fourth-order cumulant as $\chi(t_1, t_2, t_3) = \mathrm{cumulant}(\varepsilon_0, \varepsilon_{t_1}, \varepsilon_{t_2}, \varepsilon_{t_3})$, $t_1, t_2, t_3 \in \mathbb{Z}$.

Theorem 2 provides a general variance result for block bootstrap estimators, which holds under mild moment assumptions, for any regressors with positive definite $\boldsymbol{A}_n = \sum_{i=1}^n \boldsymbol{x}_{i,n}\boldsymbol{x}_{i,n}'$ and for any block size satisfying $b/n \to 0$ (i.e., $b \to \infty$ is not required).

Theorem 2. Suppose that $\boldsymbol{A}_n$ is positive definite for $n \geq 1$ that Assumption A(1) and $\sum_{t_1,t_2,t_3\in\mathbb{Z}}|\chi(t_1,t_2,t_3)| < \infty$ hold, and that $b/n \to 0$. Then, as $n \to \infty$,

(a) the variance of a block bootstrap estimator $\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n$ of $\boldsymbol{a}_n'\boldsymbol{\Sigma}_n\boldsymbol{a}_n$ is given by

$$\mathrm{var}(\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n) = \frac{(2\pi)^2 s_n}{n}\int_{-\pi}^{\pi}\Phi_n(\lambda)f^2(\lambda)d\lambda + \delta_n, \tag{12}$$

where the first right-hand side term is $O(b/n)$ and $\delta_n = o(b/n)$;

(b) except for the MBB, the expansion (12) holds with

$$s_n = O(b), \qquad \left|\int_{-\pi}^{\pi}\Phi_n(\lambda)f^2(\lambda)d\lambda\right| = O(1),$$
$$\delta_n = O\left([s_n\eta_n/n]^{1/2} + \eta_n\right),$$

and η_n from (11) having a bound $\eta_n = O((b/n)^2 + n^{-1})$;

(c) for the MBB, $|\mathrm{var}(\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_{n,\mathrm{MBB}}(b)\boldsymbol{a}_n) - \mathrm{var}(\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_{n,\mathrm{CBB}}(b)\boldsymbol{a}_n)| = O((b/n)^{3/2})$ holds.

(d) If $\displaystyle\lim_{n\to\infty}\frac{|v_{n-1,n}| + \sum_{k=2}^{n-1}|v_{k,n} - v_{k-1,n}|}{s_n^{1/2}} = 0$, then $\displaystyle\lim_{n\to\infty}\int_{-\pi}^{\pi}\Phi_n(\lambda)\ f^2(\lambda)d\lambda = f^2(0)$.

Remark 4. Essentially for any block bootstrap method, the assumption $b/n \to 0$ guarantees that the variance of the bootstrap estimator asymptotically becomes zero and $\mathrm{var}(\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n) = O(b/n)$ becomes a standard upper bound on the variance order. In some instances with nonstandard regressors, smaller orders can be realized for the block bootstrap's variance. For example, outside of the MBB, a block bootstrap estimator will have $O(1/n)$ variance whenever $\sup_n\sum_{k=1}^{n-1}v_{k,n}^2 < \infty$ holds, following from Theorem 2(b) and the forms of s_n and η_n in (11); by Theorem 2(c), this variance order follows for the MBB as well if $b = O(n^{1/3})$ additionally. The regressors in Example 2 of Section 3.1 illustrate this behavior. The MBB has a different treatment in Theorem 2 because its variance estimator, unlike the CBB estimator (6), lacks a concise form in the regression setting (cf. Section 2.2).

Remark 5. Theorem 2(d) entails that the kernel Φ_n act as an approximate identity, that is, for any $\delta > 0$, $\lim_{n\to\infty}\int_{\delta<|\lambda|<\pi}\Phi_n(\lambda)d\lambda = 0$. For this, we typically need $\lim_{n\to\infty}s_n = \infty$, which follows for many regressors when block size $b \to \infty$ as $n \to \infty$.

3.2.1 *Specialized Variance Result.*

A very concrete form for the bootstrap variance is possible for many (typically standard) regressors and is provided in Corollary 1. To state the result, using (3) and (6), let $\boldsymbol{L}_{k,b,n} \equiv \boldsymbol{Q}_{k,n}$, $k \geq 1$, for SB/MMBB/MTBB methods and $\boldsymbol{L}_{k,b,n} \equiv \boldsymbol{Q}^{\mathrm{CBB}}_{k,b,n}$ for CBB/MBB methods. Define Assumption D(i) (stated in terms of integer $i \geq 1$) as

D(i): for any sequence b satisfying Assumption $B(i)$ as $n \to \infty$, $L \geq 0$ exists, where

$$\frac{1}{b}\sum_{k=1}^{b}(\boldsymbol{a}_n' \boldsymbol{L}_{k,b,n}\boldsymbol{a}_n)^2 \to L \quad \text{and} \quad \sum_{k=2}^{b}|\boldsymbol{a}_n'(\boldsymbol{L}_{k,b,n} - \boldsymbol{L}_{k-1,b,n})\boldsymbol{a}_n| = o(b^{1/2}). \tag{13}$$

Corollary 1 requires D(1) in connection to B(1) (i.e., $b^{-1} + b/n \to 0$) as a block condition, which is mild for many regressors and allows the variance of a bootstrap estimator to be characterized by the behavior of weights $\boldsymbol{Q}_{k,n}$ in the OLS variance (2) at low lags k. Terms $\boldsymbol{L}_{k,b,n}$ only depend on the sequence b for CBB/MBB methods (see Remark 7). From (9), let $\tilde{w}(x) \equiv w * w(x)/w * w(0)$, $x \in \mathbb{R}$ for the MTBB.

Corollary 1. For a given block bootstrap, suppose the conditions of Theorem 2(a) hold with Assumption D(1). Then, as $n \to \infty$, the variance of the block bootstrap estimator $\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n$ of $\boldsymbol{a}_n'\boldsymbol{\Sigma}_n\boldsymbol{a}_n$ is given by

$$\mathrm{var}(\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n) = C_0\{2\pi f(0)\}^2\frac{b}{n} + o(b/n), \tag{14}$$

$C_0 := \tilde{L}_0$ *for CBB/MBB*; $\frac{L_0}{2}$ *for SB*; $\frac{L_0}{3}$ *for MMBB*; $L_0 \int_0^1 \tilde{w}^2(x)dx$ *for MTBB*, where $\sum_{k=1}^{b}(\boldsymbol{a}_n'\boldsymbol{Q}_{k,n}\boldsymbol{a}_n)^2/b \to L_0 \geq 0$ and $\sum_{k=1}^{b}(\boldsymbol{a}_n'\boldsymbol{Q}_{k,b,n}^{\text{CBB}}\boldsymbol{a}_n)^2/b \to \tilde{L}_0 \geq 0$ by D(1).

Remark 6. For alternatively determining $L_0, \tilde{L}_0$ in Corollary 1, if measurable functions $Q(x), D(x) : (0,\infty) \to (0,\infty)$ exist such that $\boldsymbol{a}_n'\boldsymbol{Q}_{\lceil xb\rceil,n}\boldsymbol{a}_n \to Q(x)$ and $\boldsymbol{a}_n'\boldsymbol{D}_{\lceil xb\rceil,b,n}^{\text{CBB}}\boldsymbol{a}_n \to D(x)$ as $n \to \infty$ for $x > 0$ and b satisfying Assumption B(1) [see (3) and (7)], then $L_0 = \int_0^1 [Q(x)]^2 dx$ and $\tilde{L}_0 = \int_0^1 [Q(x) - D(x)]^2 dx$.

Remark 7. In Corollary 1, the proportionality constant $\tilde{L}_0$ for the variance of CBB/MBB estimators has a different characterization than other bootstrap methods, and requires a type of convergence based on the CBB regressor weights $\boldsymbol{Q}_{k,b,n}^{\text{CBB}} = \boldsymbol{Q}_{k,n} - \boldsymbol{D}_{k,b,n}^{\text{CBB}}$ from (6). Recall that difference terms $\boldsymbol{D}_{k,b,n}^{\text{CBB}}$ in (7) depend on the block size b as well as the regressors, which impacts the variance of CBB/MBB estimators in Corollary 1 as well as their bias in Theorem 1.

Corollary 1 shows that many known variance properties of block bootstrap estimators in the stationary sample mean case extend directly to the regression setting. Note that condition $b^{-1}\sum_{k=1}^{b}(\boldsymbol{a}_n'\boldsymbol{Q}_{k,n}\boldsymbol{a}_n)^2 \to L_0$, assumed for SB, MMBB, and MTBB estimators in (13), is essentially a property of the regressors, not the bootstrap method. However, the block resampling mechanism of each bootstrap creates a "characteristic smoothing" of the regressor terms $\boldsymbol{Q}_{k,n}$, apparent in (8), (10), and the variance terms $v_{k,n}$ arising in Theorem 2. Hence, the *relative sizes* of these bootstrap variances in Corollary 1 are directly attributed to resampling mechanisms, not the underlying regressors. Consequently, the *same* variance relationships among block bootstraps in the sample mean case can continue to hold for broader regression models, even for the MMBB and MTBB. For example, the variance of the MTBB is a natural extension of the TBB's variance in the sample mean case (cf. Paparoditis and Politis 2001), and the asymptotic variance of the MMBB estimator is 2/3 that of the SB estimator, which mimics the variance relationship for MBB and SB estimators with sample means (cf. Nordman 2009). The following example provides an illustration.

Example 1 (continued). Consider the variance $\mathrm{var}(\hat{\boldsymbol{\Sigma}}_n(b))$ of a bootstrap estimator of the OLS variance $\boldsymbol{\Sigma}_n$ (2), supposing for some $c \in \mathbb{R}$ that $\boldsymbol{x}_{i,n} = i^c$ for $i = 1,\ldots,n$. In which case, $\sum_{k=2}^{b}|\boldsymbol{L}_{k,b,n} - \boldsymbol{L}_{k-1,b,n}| = O(1) = o(b^{1/2})$ holds in (13). For standard regressors (i.e., $c \geq -1/2$), $\boldsymbol{Q}_{\lceil xb\rceil,n} \to Q(x) \equiv 2$, and $\boldsymbol{D}_{\lceil xb\rceil,n} \to D(x) \equiv 2x$ holds for $x > 0$ and any $b^{-1} + b/n \to 0$. Corollary 1 (Remark 6) gives that

$$C_0 =: 2 \text{ for SB}, \quad 4/3 \text{ for CBB/MBB/MMBB}, \quad 4\int_0^1 \tilde{w}^2(t)dt \text{ for MTBB},$$

matching variance results known for the sample mean (i.e., $c = 0$) (cf. Lahiri 1999).

But for nonstandard regressors (i.e., $c < -1/2$), $\boldsymbol{Q}_{\lceil xb\rceil,n} \to 0$ and $\boldsymbol{D}_{\lceil xb\rceil,b,n} \to 0$ hold for $x > 0$ and any $b^{-1} + b/n \to 0$. Since $\tilde{L}_0 = \tilde{L}_0 = 0$, the variance of any bootstrap estimator $\hat{\boldsymbol{\Sigma}}_n(b)$ is $o(b/n)$ by Corollary 1. More precisely, from Theorem 2(b),

$$\mathrm{var}(\hat{\boldsymbol{\Sigma}}_n(b)) = \begin{cases} O(b^{2(1+c)}/n), & \text{if } c \in [-1,-1/2)\,, \\ & \text{and } b^{-1} + b/n \to 0 \\ O(1/n), & \text{if } c < -1 \text{ and } b/n \to 0, \end{cases}$$

so that, as a function of c, there is a gradual progression in the order of the variance of a block bootstrap estimator from $O(b/n)$, when $c \geq -1/2$, to $O(1/n)$, when $c < -1$.

3.3 Optimal Block Size

We may consider the optimal block size b^{opt} for minimizing the MSE of a block bootstrap estimator of $\boldsymbol{a}_n'\boldsymbol{\Sigma}_n\boldsymbol{a}_n$, $\boldsymbol{a}_n \in \mathbb{R}^q$ (bounded $\|\boldsymbol{a}_n\|$), given by

$$\mathrm{MSE}(\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n) = \{E\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n - \boldsymbol{a}_n'\boldsymbol{\Sigma}_n\boldsymbol{a}_n\}^2 + \mathrm{var}(\boldsymbol{a}_n'\hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n).$$

To simplify the MSE expansions and the formulas for the optimal block sizes, we shall assume that there exists a sequence $\{h_k\}$ such that

$$\lim_{n\to\infty}\boldsymbol{a}_n'\boldsymbol{Q}_{k,n}\boldsymbol{a}_n = h_k, \quad k \geq 1. \tag{15}$$

using (3). For the CBB/MBB, we require an additional condition [see (7)] that

$$\lim_{n\to\infty} b\cdot\boldsymbol{a}_n'\boldsymbol{D}_{k,b,n}^{\text{CBB}}\boldsymbol{a}_n = h_k^{\text{CBB}} \text{ for any } b \text{ satisfying B(2)}, \quad k \geq 1, \tag{16}$$

which is related to assuming standard regressors; see Remark 2 and Example 1 of Section 3.1. A different treatment of CBB/MBB methods, under condition (16), is necessary because the asymptotic bias and the variance of these estimators depend on $\boldsymbol{Q}_{k,b,n}^{\text{CBB}}$ in (6), not $\boldsymbol{Q}_{k,n}$. From (9), (15), and (16), let

$H(i) \equiv \sum_{k=1}^{\infty} h_k k^i r(k)$, $i = 1, 2$, $\tilde{H}(1) \equiv \sum_{k=1}^{\infty} h_k^{\text{CBB}} r(k)$, and $\|\tilde{w}\|^2 = \int_0^1 \tilde{w}^2(x)dx$ for $\tilde{w}(x) = w * w(x)/w * w(0)$.

Theorem 3. Suppose that $\boldsymbol{A}_n$ is positive definite and (15) holds, that Assumptions A(1), $B(2)$, and $\sum_{t_1,t_2,t_3 \in \mathbb{Z}} |\chi(t_1, t_2, t_3)| < \infty$ hold, and that terms $H(1), H(2), \tilde{H}(1)$ are nonzero. For the MTBB, suppose Assumptions A(2), B(3), and C hold, and for the CBB/MBB, suppose (16) holds.

(a) If Assumption D(2) holds with $L > 0$ in (13), then, as $n \to \infty$, the MSE of a block bootstrap estimator of $\boldsymbol{a}_n' \boldsymbol{\Sigma}_n \boldsymbol{a}_n$ is given by

$$\mathrm{MSE}(\boldsymbol{a}_n' \hat{\boldsymbol{\Sigma}}_n(b)\boldsymbol{a}_n) = \begin{cases} \left(\dfrac{[\tilde{H}(1)]^2}{b^2} + \tilde{L}_0[2\pi f(0)]^2 \dfrac{b}{n}\right)(1+o(1)) & \text{CBB/MBB,} \\ \left(\dfrac{[H(1)]^2}{b^2} + \dfrac{L_0}{2}[2\pi f(0)]^2 \dfrac{b}{n}\right)(1+o(1)) & \text{SB,} \\ \left(\dfrac{[H(1)]^2}{b^2} + \dfrac{L_0}{3}[2\pi f(0)]^2 \dfrac{b}{n}\right)(1+o(1)) & \text{MMBB,} \\ \left(\dfrac{[\tilde{w}''(0)H(2)]^2}{4b^4} + \|\tilde{w}\|^2 L_0 [2\pi f(0)]^2 \dfrac{b}{n}\right) & \text{MTBB,} \\ \times (1+o(1)) & \end{cases}$$

where $\sum_{k=1}^{b} (\boldsymbol{a}_n' \boldsymbol{Q}_{k,n} \boldsymbol{a}_n)^2/b \to L_0 > 0$ and $\sum_{k=1}^{b} (\boldsymbol{a}_n' \boldsymbol{Q}_{k,b,n}^{\text{CBB}} \boldsymbol{a}_n)^2/b \to \tilde{L}_0 > 0$ by $D(2)$.

(b) If the following conditions hold as $n \to \infty$,

$$\frac{1}{\sqrt{b}} \sum_{k=1}^{b} |\boldsymbol{a}_n' \boldsymbol{Q}_{k,n} \boldsymbol{a}_n - h_k| \to 0, \quad \frac{1}{\sqrt{b}} \sum_{k=2}^{b} |h_k - h_{k-1}| \to 0,$$
$$\frac{1}{b} \sum_{k=1}^{b} h_k^2 \to h^2 > 0,$$

then the MSE expansion in (a) is valid with $L_0 = h^2$ for the SB, MMBB, and MTBB.

(c) If conditions in (b) hold upon replacing $\boldsymbol{Q}_{k,n}$, h_k, and h^2 with $\boldsymbol{Q}_{k,b,n}^{\text{CBB}}$, $(h_k - b^{-1} h_k^{\text{CBB}})$ and $\tilde{h}^2$, then the MSE expansion in (a) is valid for the CBB and MBB with $\tilde{L}_0 = \tilde{h}^2$.

Remark 8. Theorem 3(b) and (c) provides a device for connecting the behavior of regressor terms $\boldsymbol{Q}_{k,n}$ appearing in the bootstrap bias and variance contributions to the bootstrap MSE; these conditions imply that Assumption D(2) holds in Theorem 3(a) and can be used for determining the variance of block bootstrap estimators in Corollary 1 under block condition B(2) (i.e., these conditions and B(2) can replace Assumptions B(1) and D(1) in Corollary 1 and $h^2, \tilde{h}^2$ can replace $L_0, \tilde{L}_0$ there). For illustration, consider the real-valued regressors $\boldsymbol{x}_{i,n} = i^c$, $c \geq -1/2$, in the example of Section 3.1. For such regressors, both (15) and (16) hold for sequences $h_k = 2$ and $h_k^{\text{CBB}} = 2k$, $k \geq 1$, and it follows that $|\boldsymbol{Q}_{k,n} - h_k|$, $|\boldsymbol{Q}_{k,n}^{\text{CBB}} - (h_k - b^{-1} h_k^{\text{CBB}})| \leq C(k+b)/n$ for $C > 0$ independent of b, n. The conditions in parts (b) and (c) of Theorem 3 are then easily checked with $h^2 = 4$ and $\tilde{h}^2 = 4/3$ and the resulting MSE expressions in Theorem 3 will match those known for the same mean case ($c = 0$).

The following is a direct consequence of Theorem 3.

Corollary 2. Suppose the conditions of Theorem 3. Then, as $n \to \infty$, the optimal block size for a block bootstrap variance estimator $\boldsymbol{a}_n' \hat{\boldsymbol{\Sigma}}_n(b) \boldsymbol{a}_n$ of $\boldsymbol{a}_n' \boldsymbol{\Sigma}_n \boldsymbol{a}_n$ is given by

$$b_{\text{opt}} = \begin{cases} \left(\dfrac{2[\tilde{H}(1)]^2}{\tilde{L}_0[2\pi f(0)]^2}\right)^{1/3} n^{1/3}(1+o(1)) & \text{CBB/MBB,} \\ \left(\dfrac{4[H(1)]^2}{L_0[2\pi f(0)]^2}\right)^{1/3} n^{1/3}(1+o(1)) & \text{SB,} \\ \left(\dfrac{6[H(1)]^2}{L_0[2\pi f(0)]^2}\right)^{1/3} n^{1/3}(1+o(1)) & \text{MMBB,} \\ \left(\dfrac{[\tilde{w}''(0)H(2)]^2}{\|\tilde{w}\|^2 L_0[2\pi f(0)]^2}\right)^{1/5} n^{1/5}(1+o(1)) & \text{MTBB.} \end{cases}$$

Remark 9. A proposal of Paparoditis and Politis (2001, 2002) is helpful for choosing a tapering window $w(\cdot)$ for the MTBB. At its optimal block, the asymptotic MSE of the MTBB is proportional to $[\|\tilde{w}\|^4 |\tilde{w}''(0)|]^{2/5}$ [see (9)], which can be minimized for $w(\cdot)$. Paparoditis and Politis (2001) showed that $c = 0.43$ minimizes this out of a class of trapezoidal tapers $w_c^{\text{trap}}(t) = c^{-1} t \mathbb{I}(t \in [0, c]) + \mathbb{I}(t \in [c, 1-c]) + c^{-1}(1-t)\mathbb{I}(t \in [1-c, 1])$, $t \in \mathbb{R}$, $c \in [0, 1]$, and that $(\tilde{w}_{0.43}^{\text{trap}})''(0) = -10.9$ and $\|\tilde{w}_{0.43}^{\text{trap}}\| = 0.27475$. Other smooth tapers (cf. Paparoditis and Politis 2002, sec. 3) often have self-convolutions $\tilde{w}$ resembling $\tilde{w}_{0.43}^{\text{trap}}$ and induce similar large-sample properties for the MTBB.

Under some mild conditions on the regressor variables, the large-sample optimal block sizes and MSE results have parallels with those known for block bootstraps with stationary sample means (Hall et al. 1995; Paparoditis and Politis 2001; Lahiri 2003a, chap. 5; Politis and White 2004), though proportionality constants must be adjusted to accommodate the underlying regressors. Additionally, at its optimal block length, the MTBB continues to achieve the best MSE order $O(n^{-4/5})$ among block bootstraps, which have orders $O(n^{-2/3})$ in comparison.

4. NUMERICAL STUDIES

We examine finite sample performance of block bootstraps under a regression model through simulation. In Section 4.1, we compare the block bootstraps for variance estimation, while in Section 4.2, we study the problem of approximating the sampling distributions of OLS estimators for calibrating CIs for regression parameters.

4.1 Variance Estimation

We consider $Y_i = \boldsymbol{x}_{i,n}' \boldsymbol{\beta} + \varepsilon_i$, $i = 1, \ldots, n$ with $\boldsymbol{\beta} = (\beta_0, \beta_1)' = (1, -1)'$ and standard regressors $\boldsymbol{x}_{i,n} = (1, u_i)' \in \mathbb{R}^2$ based on a sequence $u_i = (2 + z_i)\sqrt{i}$, $i \geq 1$ fixed throughout the study, upon generating $\{z_i\}$ as iid normal $N(0, 1)$ variables. Letting $\hat{\beta}_{i,n}$ denote the OLS estimator of β_i, consider block bootstrap estimators of $\sigma_{i,n}^2 \equiv n^{i+1} \mathrm{var}(\hat{\beta}_{i,n}) = \boldsymbol{a}_{i,n}' \boldsymbol{\Sigma}_n \boldsymbol{a}_{i,n}$, $i = 0, 1$, with bounded row vectors $\boldsymbol{a}_{0,n}' = (n^{1/2}, 0)\boldsymbol{A}_n^{-1/2}$ and $\boldsymbol{a}_{1,n}' = (0, n)\boldsymbol{A}_n^{-1/2}$ using notation of Section 3.2. Then, when considering variance estimation of the slope $\hat{\beta}_{1,n}$ with these regressors, it holds that $h_k = 144/169 = h_k^{\text{CBB}}/k$ in (15) and (16) for $k \geq 1$ and that $L_0 = h_1^2 = 3\tilde{L}_0$; for the intercept $\hat{\beta}_{0,n}$,

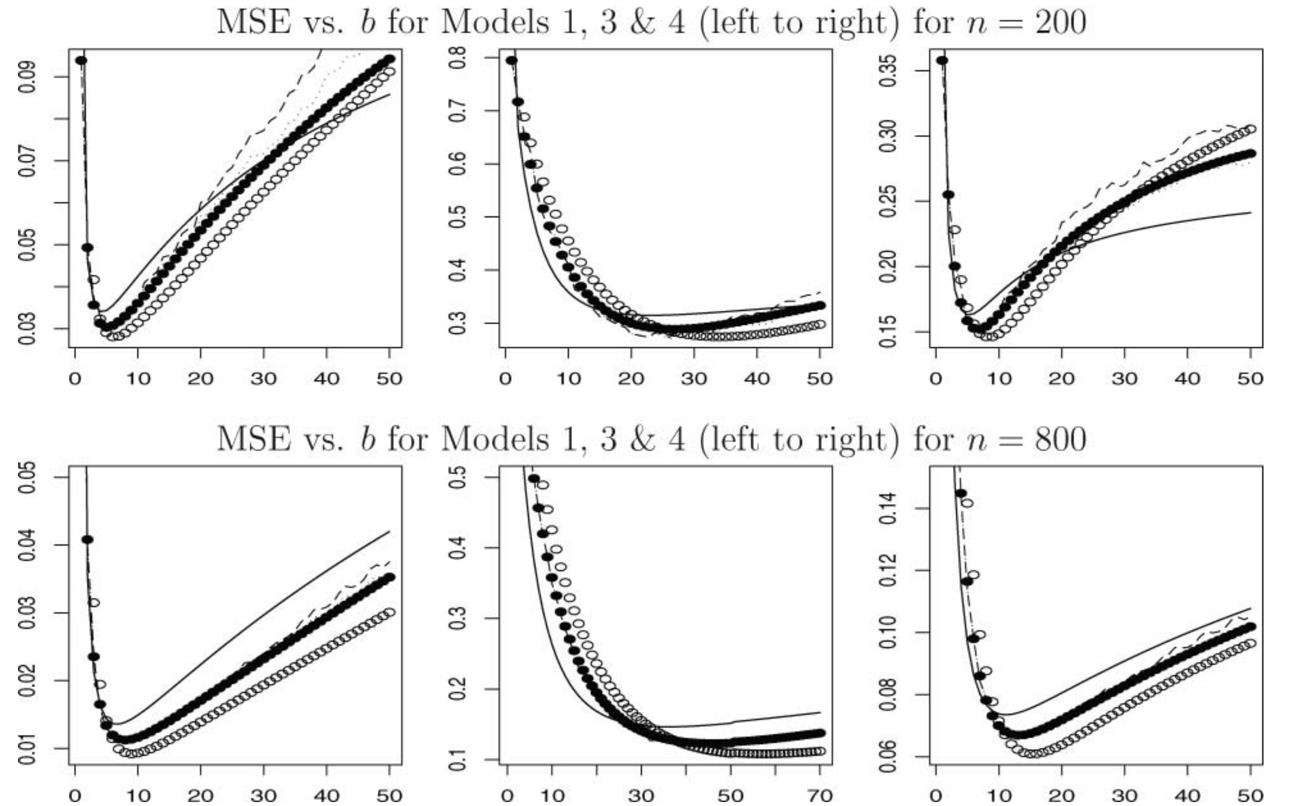

Figure 1. Scaled MSE $E\{\hat{\sigma}^2_{0,n}(b)/\sigma^2_{0,n} - 1\}^2$ versus block b for bootstrap estimators $\hat{\sigma}^2_{0,n}(b)$ of OLS intercept variance $\sigma^2_{0,n}$ (from 10,000 simulations). These include CBB ($\cdots$), MBB (- -), SB (–), MMBB (•), and MTBB (∘) under Model 1 (AR(1), $\theta = 0.3$), Model 3 (AR(1), $\theta = 0.9$), and Model 4 (exponential AR).

$h_k = 594/169 = h_k^{\text{CBB}}/k$ holds instead. Hence, bootstrap variance estimators have different MSEs between slope and intercept problems (Theorem 3), but optimal block sizes (Corollary 2) are the same in both cases (these depend on ratios $[H(j)]^2/L_0 = [\sum_{k=1}^{\infty} k^j r(k)]^2$, $j = 1, 2$, free of h_k values here). CBB, MBB, and MMBB then have matching expressions for their large-sample MSEs and optimal block sizes. The MTBB is implemented using the smooth trapezoidal taper described in Remark 9.

We first consider stationary errors $\{\varepsilon_t\}$ generated through AR(1) models $\varepsilon_t = \theta\varepsilon_{t-1} + Z_t$ for $\theta = 0.3, 0.6, 0.9$ (denoted as Models 1, 2, and 3, respectively) as well as a nonlinear model $\varepsilon_t = \{0.8 - 1.1\exp(50\varepsilon^2_{t-1})\}\varepsilon_{t-1} + 0.1Z_t$ (denoted as Model 4), where $\{Z_t\}$ are iid $N(0, 1)$ innovations. The linear autoregressive models allow increasing positive dependence and the nonlinear version has been considered by Paparoditis and Politis (2001). For sample sizes $n = 200, 800$, and all six block bootstrap methods, Figure 1 displays a collection of scaled MSE curves, as a function of block size, in the intercept problem. (See the online supplementary material for figures of scaled MSE and bias curves in both slope and intercept problems.) As CBB/MBB resampling distorts the regressor sequence, the bias/MSE curves for these methods are less smooth than for the other block bootstraps. But, as according to theory, these curves do begin to match those of the MMBB for larger samples $n = 800$ in both OLS slope and intercept variance estimation. Generally, the MTBB achieves the smallest bias and MSE out of the block bootstraps, which often holds over a range of block values.

For three representative block bootstraps (SB, MMBB, and MTBB) in inference on the OLS slope variance $\sigma^2_{1,n} = n^2\mathrm{var}(\hat{\beta}_{1,n})$, Figure 2 gives a rough impression of MSE convergence over block sizes of form $b = Cn^\tau$, having optimal orders $\tau = 1/5$ for the MTBB and $\tau = 1/3$ otherwise. Also included are four additional innovation distributions as $\varepsilon_t = 0.4\varepsilon_{t-1} + 0.2\varepsilon_{t-2} + 0.1\varepsilon_{t-3} + Z_t$ (Model 5), $\tilde{\varepsilon}_t = 0.2Z_{t-1} + 0.3Z_{t-2} + Z_t$ (Model 6), $\bar{\varepsilon}_t = 0.4\bar{\varepsilon}_{t-1} + 0.2\bar{\varepsilon}_{t-2} + 0.1\bar{\varepsilon}_{t-3} + \tilde{\varepsilon}_t$ (Model 7), and $\varepsilon_t = \bar{\varepsilon}_t^2\,\mathrm{sign}(\bar{\varepsilon}_t)$ (Model 8), where again $\{Z_t\}$ denote iid $N(0, 1)$ variables. For all models and bootstrap methods, MSE convergence appears fastest when $C = 4.0$ and is particularly slow for stronger dependence models (i.e., nonlinear Model 8 and autoregressive Model 3). Repeating these simulations with nonnormal errors (e.g., centered exponentials) produced qualitatively similar results.

For each block bootstrap in Figure 2, we also consider empirical estimates $\hat{b}_{\text{opt}} = \hat{C}n^\tau$ of the optimal block sizes by adjusting a proposal of Paparoditis and Politis (2001) and Politis and White (2004) for the regression setting. That is, in the optimal block expressions of Corollary 2, we replace $2\pi f(0) = \sum_{k=-\infty}^{\infty} r(k)$ and $H(i) = \sum_{k=1}^{\infty} h_k k^i r(k)$, $i = 1, 2$ with estimates

$$\sum_{k=-M}^{M} \lambda(k/M)\hat{r}_n(k), \qquad \sum_{k=1}^{M} \boldsymbol{a}'_{1,n}\boldsymbol{Q}_{k,n}\boldsymbol{a}_{1,n}k^i\lambda(k/M)\hat{r}_n(k),$$

using the sample covariance $\hat{r}_n(k) = n^{-1}\sum_{j=1}^{n-|k|}(e_i - \bar{e}_n)(e_{i+|k|} - \bar{e}_n)$ of residuals and a "flat-top" window $\lambda(t) = \mathbb{I}(|t| \le 1/2) + 2(1 - |t|)\mathbb{I}(|t| > 1/2)$, $t \in [-1, 1]$ (Politis and

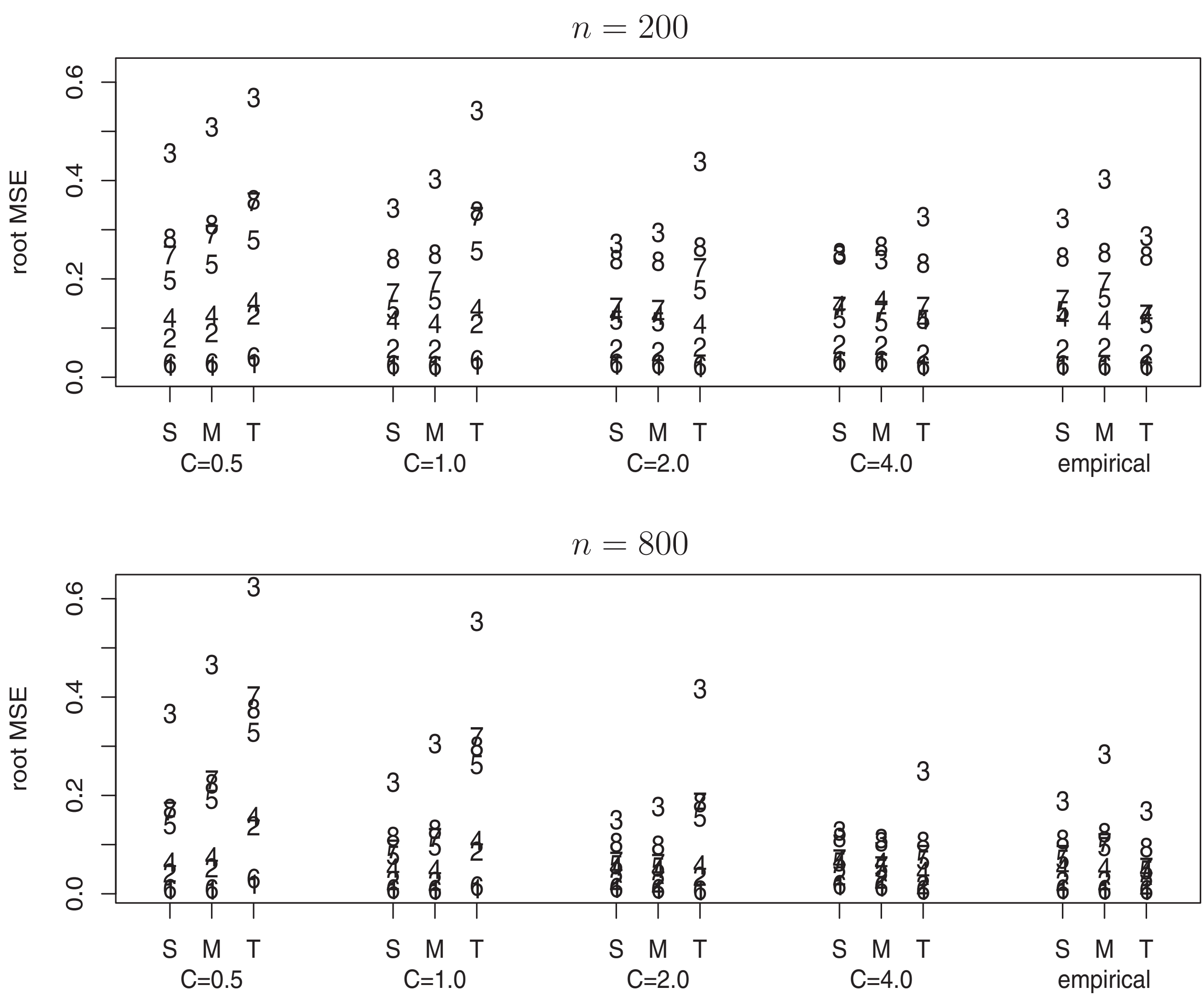

Figure 2. Root MSE $[E\{\hat{\sigma}^2_{1,n}(b)/\sigma^2_{1,n} - 1\}^2]^{1/2}$ for SB, MMBB, and MTBB (T,S,M) estimators $\hat{\sigma}^2_{1,n}(b)$ of OLS slope variance $\sigma^2_{1,n}$ based on blocks $b = Cn^{\tau}$, $C = 0.5, 1, 2, 4$, or empirical blocks $b = \hat{C}n^{\tau}$, with $\tau = 1/5$ for MTBB and $1/3$ for SB or MMBB. For each sample size, block choice, and bootstrap, root MSEs are denoted by model numbers 1–8 (based on 10,000 simulations).

Romano 1995). To be completely empirical, we also replace the limit L_0 in Corollary 2 with an approximation $\sum_{k=1}^{M}(\boldsymbol{a}'_{1,n}\boldsymbol{Q}_{k,n}\boldsymbol{a}_{1,n})^2/M$. The simulation is automated by taking $M = n^{1/5}$ for consistent block estimation (cf. Politis and White 2004, thm. 3.2), though other block estimates may follow by inspecting the correlogram of residuals as detailed by Paparoditis and Politis (2001, sec. 3.2). In Figure 2, empirical block choices generally perform as well as the best deterministic block choice from the collection $b = Cn^{\tau}$, $C = 0.5, 1, 2, 4$. The same essential pattern in Figure 2 also holds for the OLS intercept case and an alternative choice $M = 2n^{1/5}$ did not significantly change results. The supplementary material provides further simulation comparison of SB/MMBB/MTBB methods with empirical blocks, in which the MTBB emerges best in terms of average MSE ranking over all eight models and various n.

4.2 Distributional Approximations and Interval Estimation

The block bootstrap reconstructions $Y_1^*, \ldots, Y_n^*$ of time series data (1) have the advantage that these can be used to approximate functionals related to the sampling distribution of OLS estimators, going beyond variances. The validity of the block bootstraps for variance estimation is an important indication that the block resampling mechanisms can be appropriate for application in broader problems with regression, because bootstraps often fail for more general time series inference when variance estimation (capturing the dependence structure) goes wrong (cf. Lahiri 2003a).

Here, we investigate the block bootstraps for estimating the distribution of OLS estimators of β_0, β_1 under the simulation framework of Section 4.1. In particular, we consider bootstrap estimation of the distribution of studentized OLS estimators $T_{i,n} = n^{(1+i)/2}(\hat{\beta}_{i,n} - \beta_i)/\hat{\tau}_{i,n}$, $i = 0, 1$, where $\hat{\tau}^2_{i,n} = \sum_{k=0}^{M-1}\boldsymbol{a}'_{i,n}\boldsymbol{Q}_{k,n}\boldsymbol{a}_{i,n}u(k/M)\hat{r}_n(k)$ is an estimator of $n^{1+i}\mathrm{var}(\hat{\beta}_{i,n})$ involving residual sample covariances $\hat{r}_n(k)$ and lag weights $u(k/M)$ defined by a bandwidth M and continuous tapering window $u : [0, 1] \to [0, 1]$ with $u(0) = 1$ [values of $\boldsymbol{a}_{0,n}, \boldsymbol{a}_{1,n}$ are from Section 4.1 and the regressor variables there define $\boldsymbol{Q}_{k,n}$ from (3)]. The form of the studentizing factor $\hat{\tau}^2_{i,n}$ belongs to an estimator class considered by Götze and Künsch (1996) in establishing the second-order correctness of the MBB for studentized sample means. Here, we use the Parzen window $u(\cdot)$ with bandwidth $M = n^{1/5}$ of optimal order (cf. Priestley 1981, p. 447); repeating simulations with Tukey–Hanning or Bartlett–Priestly windows or a bandwith $M = 2n^{1/5}$ produced qualitatively similar results. For a given data stretch $Y_1, \ldots, Y_n$, the sampling distributions of interest are approximated by the empirical distribution of $T^*_{i,n} = (\hat{\beta}^*_{i,n} - \hat{\beta}_{i,n})/\hat{\tau}^*_{i,n}$ values generated over several bootstrap resamples (600 used here). These bootstrap distribution estimators are used to calibrate symmetric two-sided, or upper one-sided, 95% CIs for β_i, $i = 0, 1$, by $\hat{\beta}_{i,n} \pm n^{-(1+i)/2}\hat{\tau}_{i,n} q^*_{95}$ or $\hat{\beta}_{i,n} - n^{-(1+i)/2}\hat{\tau}_{i,n}\tilde{q}^*_5$, where q^*_{95} is the 95th percentile of produced $|T^*_{i,n}|$ values and $\tilde{q}^*_5$ is the 5th percentile of $T^*_{i,n}$ values. We refer to these as percentile-t bootstrap intervals. For comparison, we also include normal approximation intervals based on $T_{i,n}$, $i = 0, 1$.

From each of the eight models from Section 4.1, we generated samples of sizes $n = 100, 500$, and computed bootstrap intervals from CBB/MBB/SB/MMBB/MTBB methods with the MTBB using the trapezoidal taper from Remark 9 (a cosine-bell taper also produced similar results). We considered block sizes $b = n^{1/4}, 3n^{1/4}$ for the one-sided case and $b = n^{1/5}, 3n^{1/5}$ for the two-sided case, which are known to be the optimal orders for distribution estimation for the sample mean (cf. Hall et al. 1995). Coverage percentages for Models 1, 4, and 7 appear in Table 2 for the intercept, where these models are chosen because the underlying dependence increases with the model number. Full results for all models and both slope/intercept parameters are available in the supplementary materials.

The coverage accuracies for CIs of the slope are typically better than those for the intercept, though patterns for both are overall similar. Coverage accuracies tend to improve as the dependence weakens and block sizes $b = 3n^{1/4}, 3n^{1/5}$ appear to give marginally better results. All bootstrap intervals hold the nominal 95% coverage level better than the normal approximation. We repeated the simulations with bootstrap intervals without the studentization step, which approximate the distribution of $\hat{\beta}_{i,n} - \beta_i$ with a version $\hat{\beta}^*_{i,n} - \hat{\beta}_{i,n}$; these also performed uniformly better than the normal approximation but not as well as percentile-t bootstrap intervals shown here.

Table 2. Coverage percentages (based on 4,000 simulations) for two-sided and upper one-sided 95% intervals for intercept β_0 based on percentile t-bootstrap (CBB, MBB, SB, MMBB, and MTBB denoted as C, M, S, MM, and T) and normal approximation (Norm)

Two-sided		$b = n^{1/5}$					$b = 3n^{1/5}$				
Model, n	Norm	C	M	S	MM	T	C	M	S	MM	T
1, 100	90	92	93	93	93	93	93	93	94	93	93
4, 100	85	90	90	90	90	90	90	91	91	91	91
7, 100	69	83	80	79	80	81	85	86	86	86	86
1, 500	92	93	93	94	93	93	95	95	95	95	95
4, 500	87	91	91	91	92	91	92	91	91	92	91
7, 500	72	85	83	81	83	83	86	83	82	84	83
One-sided		$b = n^{1/4}$					$b = 3n^{1/4}$				
1, 100	92	93	94	94	94	94	94	94	94	94	94
4, 100	89	92	91	91	91	91	92	92	93	92	92
7, 100	81	88	86	86	86	86	88	89	89	89	89
1, 500	93	93	94	93	94	94	95	95	95	95	95
4, 500	91	93	92	92	92	92	93	92	92	92	92
7, 500	81	89	87	86	87	87	89	87	87	88	87

5. DATA EXAMPLES

5.1 Data Example 1

The Time Series Data Library (Hyndman n.d.) presents counts $C_1, \ldots, C_{120}$ of monthly accidents resulting in serious injury/death on UK roadways between January 1975 and December 1984 (cf. Brockwell and Davis 2002, p. 217, for

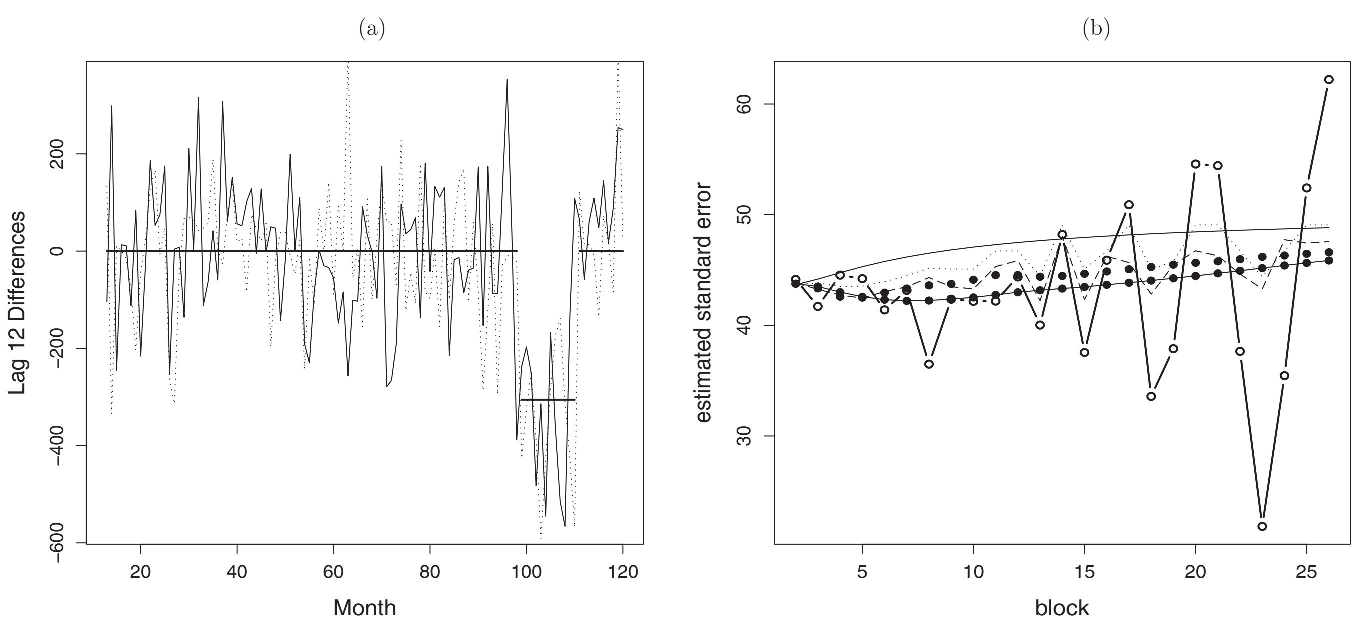

Figure 3. (a) Differenced monthly accident count series (–) and least squares fit (step function). For illustration, one MTBB reconstruction (––) based on a block size $b = 4 \approx n^{1/4}$ is included. (b) Block bootstrap estimates of the standard error of OLS estimator $\hat{\beta}_{1,n}$ over block sizes $b \geq 2$ for CBB (···), MBB (- -), SB (–), MMBB (•), MTBB (– • –), and TBB (– ∘ –). The MTBB/TBB use the taper from Remark 9.

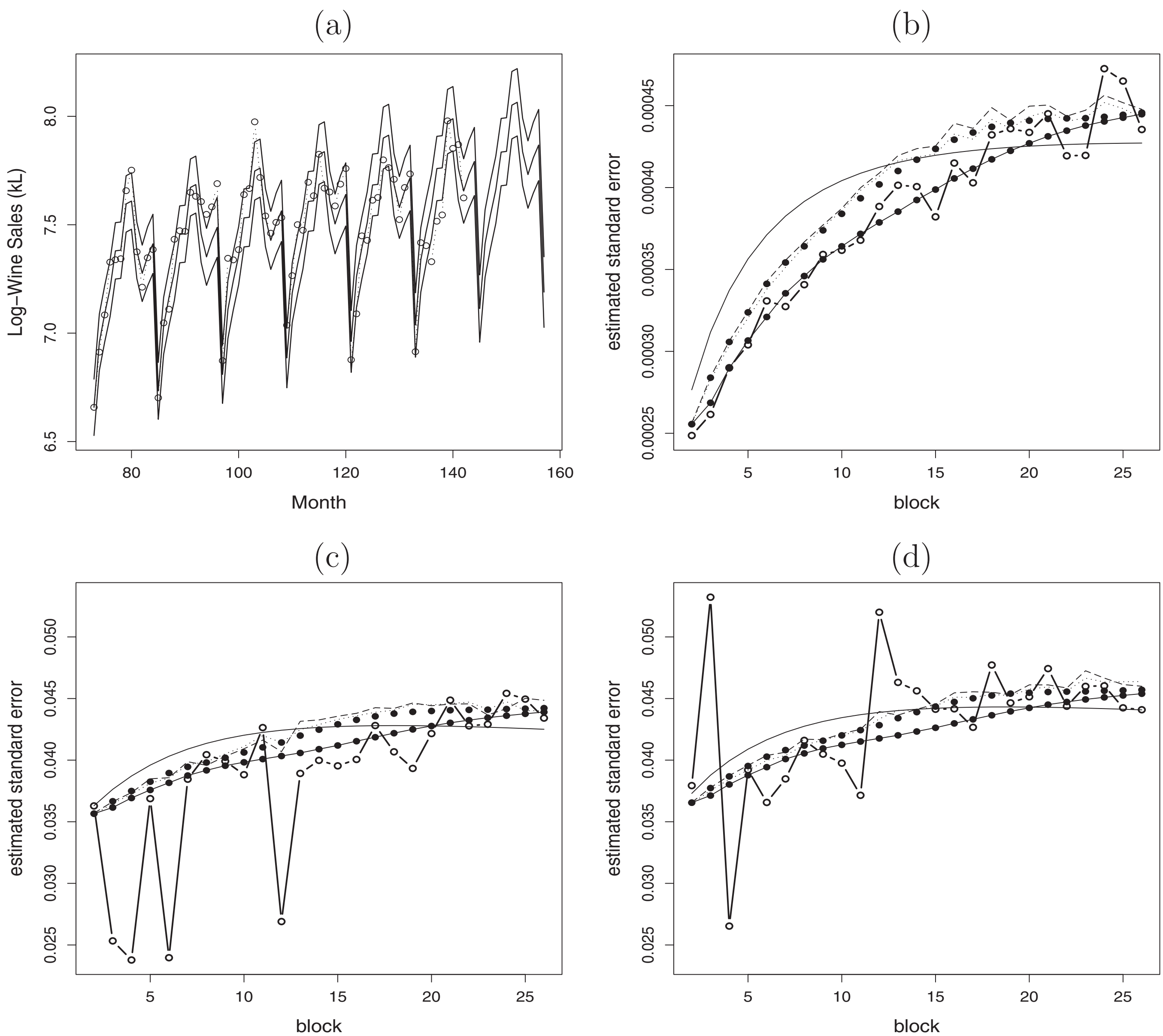

Figure 4. (a) Last values (months 73–142) of log-monthly Australian wine sales (− ○ −) with OLS fit (middle −) and simultaneous 95% confidence bands (upper/lower −) over 1–157 months based on MTBB with percentile-t calibration. (b)–(d) Standard error estimates for OLS estimators of β_1 in (b), β_2 in (c), and β_9 in (d) with CBB (···), MBB (- -), SB (−), MMBB (•), MTBB (− • −), and TBB (− ○ −) (taper from Remark 9).

maximum likelihood estimation with these data). In February 1983 (the 98th time point), seat-belt legislation was passed, and one can examine whether the data suggest a subsequent drop in the mean number of such monthly accidents, corresponding to a negative β_1 parameter in a regression model $EC_i = \beta_0 + \beta_1 \mathbb{I}(i \geq 99), i = 1, \ldots, 120$, where $\mathbb{I}(\cdot)$ denotes the indicator function. The data exhibit seasonality of period 12, and to remove possibly stochastic seasonal effects, we consider a differenced data model $Y_t = C_t - C_{t-12} = \beta_1 \mathbb{I}(99 \leq i \leq 100) + \varepsilon_t$, $t = 13, \ldots, 120$ with stationary ε_t's. The differenced data appear in Figure 3(a) along with the OLS estimate of the piece-wise constant trend EY_t based on $\hat{\beta}_{1,n} = -305.58$; the fit appears reasonable.

This example deviates slightly from a constant trend, for which the TBB is known to be valid. Figure 3(b) shows estimated standard errors for $\hat{\beta}_{1,n}$ from block bootstraps with various block sizes b (MTBB/TBB use the Remark 9 taper). For this nonconstant trend, the TBB exhibits the regressor mismatching described in Section 2.4, as evidenced by the volatility of its estimates as a function of b. The MTBB overcomes this problem. In fact, the estimates from bootstraps involving additional block randomization, namely SB/MMBB/MTBB, vary smoothly with b and are more robust to small block changes. The CBB/MBB methods have less smooth curves in Figure 3(b), as these methods lack additional block randomization. However, compared with the TBB, the CBB/MBB estimates are less erratic, which roughly supports the finding that these methods induce less distortion than the TBB in recreating regression samples to the extent that CBB/MBB variance estimators are generally consistent.

Based on the CBB/MBB/SB/MMBB/MTBB with percentile-t calibration (20,000 resamples) and $b = 3 \approx n^{1/4}$, one-sided upper 99% confidence bounds for β_1 are -196.32, -197.47, -198.16, -198.88, and -199.15, suggesting a drop in average monthly auto injuries following the seat-belt legislation. These bounds also tend to be relatively smooth functions of b; for example, for $b = 10 \approx 3n^{1/4}$, the bounds become -200.32, -199.15, -194.34, -202.36, and -204.66. In contrast, the TBB confidence bounds exhibit volatile behavior as a function of the block size, qualitatively similar to Figure 3(b) (e.g., extreme cases are bounds of -182.44, -250.26 for $b = 23, 25$).

5.2 Data Example 2

We next consider monthly sales of Australian red wine (kiloliters) between January 1980 and October 1991, from Brockwell and Davis (2002). These are log-transformed to stabilize variances and the last half of the series appears in Figure 4(a). For the transformed data, an upward trend with a regular 12-month seasonal pattern is evident, so we consider a linear regression model $Y_t = \beta_1 t + \sum_{j=1}^{12} \beta_{j+1} \mathbb{I}(t \in \text{month } j) + \varepsilon_t, t = 1, \ldots, 142$, with stationary errors. Figure 4(a) shows the trend estimated by OLS, and residual analysis indicates that the model is reasonable. Figure 4(b)–(d) shows estimated standard errors for OLS estimators of β_1 (slope), β_2 (seasonal low), and β_9 (seasonal peak) from block bootstraps with various block sizes b. For the slope $\hat{\beta}_{1,n} = 0.00632$, block estimates of 7 for the MTBB, 4 for the SB, and 5 for all others produce estimated standard errors 3.20e-4, 3.24e-4, 3.35e-4, 3.24e-4, and 3.37e-4 from CBB, MBB, SB, MMBB, and MTBB methods, respectively. In this case, standard errors from the original TBB are the least smooth as a function of b, but often in line with the other bootstraps. However, because of the design matrix here (dummy variables for β_2, β_9), one should expect estimated standard errors for $\hat{\beta}_{2,n}$ and $\hat{\beta}_{9,n}$ to be *roughly equal* for a given bootstrap and given b. Interestingly, the TBB clearly violates this in Figure 4(c) and (d), which again reveals bias problems from regressor mismatching that the MTBB avoids. To reiterate the use of block bootstraps beyond variance estimation, we also applied the MTBB to recreate studentized versions of OLS estimates for the trend function and produce *simultaneous* 95% confidence bands over a time range that included all the data and an additional 15 months, extrapolating the linear growth in (log-scale) wine sales over time (see Figure 4(a)).

6. CONCLUSIONS

In considering the problem of variance estimation, we have provided a study on how optimality properties of block bootstraps for stationary time series transfer to a simple nonstationary setting with linear regression models with dependent errors and nonstochastic regressors. There are some dramatic role reversals. Perhaps unexpectedly, the TBB, as a second-generation bootstrap, breaks down while, for instance, the first-generation SB method translates easily into the regression setting. As a remedy, we introduce the MTBB method that adds an additional block randomization step to provide a valid reconstruction of time series regression models. Compared with the well-studied sample mean case (Hall et al. 1995; Lahiri 1999; Politis and White 2004), the bias and variance properties of block bootstrap variance estimators depend more intricately on the nonstochastic regressors, but there are some similarities. The MTBB turns out to be generally consistent with superior bias $O(1/b^2)$ compared with the typical $O(1/b)$ bias associated with other block bootstraps for many types of regressor variables, and forms of optimal block sizes can carry over to a regression setting. By establishing the validity of block bootstraps for variance estimation, we have suggested that the block resampling mechanisms are appropriate for using these bootstraps to approximate other features of the sampling distributions of least squares estimators, with some numerical evidence of this in Section 4. For the regression models here, the MTBB recreations of time series extend the TBB to inference scenarios where this second-generation bootstrap would otherwise be invalid.

SUPPLEMENTARY MATERIALS

Proofs and simulation summaries: An appendix containing proofs of the results as well as additional tables and figures for the numerical studies of Section 4.

[Received September 2009. Revised November 2011.]

REFERENCES

Andrews, D. W. K. (1991), "Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858. [234]

Brillinger, D. R. (1981), *Time Series: Data Analysis and Theory*, San Francisco, CA: Holden-Day. [235]

Brockwell, P. J., and Davis, R. A. (2002), *Time Series: Theory and Methods* (2nd ed.), New York: Springer. [243,245]

Dahlhaus, R. (1983), "Spectral Analysis with Tapered Data," *Journal of Time Series Analysis*, 4, 163–175. [235]

Efron, B. (1982), *The Jackknife, the Bootstrap and Other Resampling Plans*, Philadelphia, PA: Society for Industrial and Applied Mathematics. [234]

Fitzenberger, B. (1997), "The Moving Blocks Bootstrap and Robust Inference for Linear Least Squares and Quantile Regressions," *Journal of Econometrics*, 85, 235–287. [234]

Freedman, D. A. (1981), "Bootstrapping Regression Models," *The Annals of Statistics*, 12, 121–145. [234]

Gonçalves, S., and White, H. (2005), "Bootstrap Standard Error Estimates for Linear Regression," *Journal of the American Statistical Association*, 100, 970–979. [234]

Götze, F., and Künsch, H. R. (1996), "Second-Order Correctness of the Blockwise Bootstrap for Stationary Obervations," *The Annals of Statistics*, 24, 1914–1933. [242]

Hall, P., Horowitz, J. L., and Jing, B.-Y. (1995), "On Blocking Rules for the Bootstrap With Dependent Data," *Biometrika*, 82, 561–574. [233,240,243,245]

Heagerty, P. J., and Lumley, T. (2000), "Window Subsampling of Estimating Function With Application to Regression Models," *Journal of the American Statistical Association*, 95, 197–211. [234]

Hidalgo, J. (2003), "An Alternative Bootstrap to Moving Blocks for Time Series Regression Models," *Journal of Econometrics*, 117, 367–399. [234]

Hyndman, R. J. (n.d.), *Time Series Data Library* [on-line]. Available at *http://robjhyndman.com/TSDL* (accessed on October 1, 2011). [243]

Künsch, H. R. (1989), "The Jackknife and the Bootstrap for General Stationary Observations," *The Annals of Statistics*, 17, 1217–1241. [233,235,236]

Lahiri, S. N. (1992), "Second Order Optimality of Stationary Bootstrap," in *Exploring the Limits of Bootstrap*, eds. R. Lepage and L. Billard, New York: Wiley, pp. 183–214. [234]

——— (1996), "On Edgeworth Expansion and the Moving Block Bootstrap for Studentized M-Estimators in Multiple Linear Regression Models," *Journal of Multivariate Analysis*, 56, 206–225. [234]

——— (1999), "Theoretical Comparisons of Block Bootstrap Methods," *The Annals of Statistics*, 27, 386–404. [233,239,245]

——— (2003a), *Resampling Methods for Dependent Data*, New York: Springer. [240,242]

——— (2003b), "Central Limit Theorems for Weighted Sums of a Spatial Process Under a Class of Stochastic and Fixed Designs," *Sankhya,* Series A, 65, 356–388. [236]

Liu, R. Y., and Singh, K. (1992), "Moving Blocks Jackknife and Bootstrap Capture Weak Dependence," in *Exploring the Limits of the Bootstrap,* eds. R. Lepage and L. Billard, New York: Wiley, pp. 225–248. [233]

Nordman, D. J. (2009), "A Note on the Stationary Bootstrap's Variance," *The Annals of Statistics*, 37, 359–370. [233,239]

Paparoditis, E., and Politis, D. N. (2001), "Tapered Block Bootstrap," *Biometrika*, 88, 1105–1119. [233,235,236,239,240,241]

——— (2002), "The Tapered Block Bootstrap for General Statistics From Stationary Sequences," *Econometrics Journal*, 5, 131–148. [233,235,240]

Parzen, E. (1957), "On Consistent Estimates of the Spectrum of a Stationary Time Series," *Annals of Mathematical Statistics*, 28, 329–348. [236]

Politis, D. N., and Romano, J. P. (1992), "A Circular Block Resampling Procedure for Stationary Data," in *Exploring the Limits of Bootstrap*, eds. R. Lepage and L. Billard, New York: Wiley, pp. 263–270. [233,234]

——— (1994), "The Stationary Bootstrap," *Journal of the American Statistical Association*, 89, 1303–1313. [233,235]

——— (1995), "Bias-Corrected Nonparametric Spectral Estimation," *Journal of Time Series Analysis*, 16, 67–103. [241]

Politis, D. N., Romano, J. P., and Wolf, M. (1997), "Subsampling for Heteroskedastic Time Series," *Journal of Econometrics*, 81, 281–317. [234]

Politis, D. N., Romano, J. P., and Wolf, M. (1999), *Subsampling*, New York: Springer. [234]

Politis, D. N. (2003), "The Impact of Bootstrap Methods on Time Series Analysis," *Statistical Science*, 18, 219–230. [236]

Politis, D. N., and White, H. (2004), "Automatic Block-Length Selection for the Dependent Bootstrap," *Econometric Reviews*, 23, 53–70. [240,241,245]

Priestley, M. B. (1981), *Spectral Analysis and Time Series*, London: Academic Press. [242]

Robinson, P. M., and Velasco, C. (1997), "Autocorrelation-Robust Inference," in *Handbook of Statistics* (Vol. 15), eds. G. S. Maddala and C. R. Rao, Amsterdam: North-Holland, pp. 267–298. [234]