AARHUS
UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

The wild tapered block bootstrap

Ulrich Hounyo

CREATES Research Paper 2014-32

Department of Economics and Business
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

The wild tapered block bootstrap*

Ulrich Hounyo †
Oxford-Man Institute, University of Oxford,
CREATES, Aarhus University,

September 24, 2014

Abstract

In this paper, a new resampling procedure, called the wild tapered block bootstrap, is introduced as a means of calculating standard errors of estimators and constructing confidence regions for parameters based on dependent heterogeneous data. The method consists in tapering each overlapping block of the series first, then applying the standard wild bootstrap for independent and heteroscedastic distributed observations to overlapping tapered blocks in an appropriate way. It preserves the favorable bias and mean squared error properties of the tapered block bootstrap, which is the state-of-the-art block-based method in terms of asymptotic accuracy of variance estimation and distribution approximation. For stationary time series, the asymptotic validity, and the favorable bias properties of the new bootstrap method are shown in two important cases: smooth functions of means, and M-estimators. The first-order asymptotic validity of the tapered block bootstrap as well as the wild tapered block bootstrap approximation to the actual distribution of the sample mean is also established when data are assumed to satisfy a near epoch dependent condition. The consistency of the bootstrap variance estimator for the sample mean is shown to be robust against heteroskedasticity and dependence of unknown form. Simulation studies illustrate the finite-sample performance of the wild tapered block bootstrap. This easy to implement alternative bootstrap method works very well even for moderate sample sizes.

JEL Classification: C15, C22
Keywords: Block bootstrap, Near epoch dependence, Tapering, Variance estimation.

1 Introduction

The bootstrap of Efron (1979) is a powerful nonparametric method to approximate the sampling distribution and the variance of complicated statistics based on i.i.d. observations. The failure of the i.i.d. resampling scheme to give a consistent approximation to the true limiting distribution of a statistic when observations are not independent has motivated the development of alternative bootstrap methods in the context of dependent data. As an extension of Efron's i.i.d. bootstrap to dependent observations, the moving block bootstrap (MBB) of Künsch (1989) and Liu and Singh (1992) can be used to approximate the sampling distributions and variances of statistics in time series. In order

*I acknowledge support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation, as well as support from the Oxford-Man Institute of Quantitative Finance.

†Department of Economics and Business, Aarhus University, 8210 Aarhus V., Denmark. Email: ulrich.hounyo@oxford-man.ox.ac.uk.

to capture temporal dependence nonparametrically, the MBB samples the overlapping blocks with replacement and then pastes the resampled blocks together to form a bootstrap sample. Based on the idea of resampling blocks, a few variants of the MBB have been developed, such as the nonoverlapping block bootstrap (NBB) (Carlstein (1986)), and the stationary bootstrap (SB) (Politis and Romano (1994)), among others.

For variance estimation in the smooth function model, the MBB and its variants (the so-called “first generation” block bootstrap methods) yield the same convergence rate of the mean squared error (MSE), albeit with a different constant in the leading term of the bias and variance expansions; see, e.g., Lahiri (1999, 2003) and Nordman (2009). In an attempt to reduce the bias and MSE, Carlstein et al. (1998) proposed the matched block bootstrap whereas Paparoditis and Politis (2001, 2002) proposed the tapered block bootstrap (TBB) (one of the so-called “second generation” bootstrap methods). The TBB involves tapering each overlapping block of the series first, then a resampling of those tapered blocks. The TBB offers a superior convergence rate in the bias and MSE compared to the “first generation” block bootstrap methods. The data tapering of the blocks used in the TBB is designed to decrease the bootstrap bias, and has, as a result, an increased accuracy of estimation of sampling characteristics for linear and approximately linear statistics. See also Shao (2010a, 2010b) who developed the extended tapered block bootstrap (ETBB) and the dependent wild bootstrap (DWB) for stationary time series. The ETBB and DWB can preserve the favorable bias and mean squared error properties of the TBB.

The performance of these bootstrap methods in the presence of nonstationarity is not well understood in the literature. Recently, Nordman and Lahiri (2012) have investigated the properties of some block bootstrap methods under a specific form of nonstationarity, with data generated by a linear regression model with weakly dependent errors and non stochastic regressors. In contrast to the stationarity case, Nordman and Lahiri (2012) show that the MBB, SB, and TBB variance estimators often turn out to be invalid with general nonrandom regressors. As a remedy, they propose an additional block randomization step in order to balance out the effects of nonuniform regression weights.

In this paper, we introduce a new resampling method, called the wild tapered block bootstrap (WTBB), that is generally applicable for dependent heterogeneous arrays. As in Gonçalves and White (2002), the data are assumed to satisfy a near epoch dependent (NED) condition, which includes the more restrictive mixing assumption as a special case. NED processes also allow for considerable heterogeneity.

In the case of the sample mean, we found that the WTBB is robust against heteroskedasticity and dependence of unknown form. We also show that Paparoditis and Politis's TBB enjoys this robustness property to heteroskedasticity in this heterogeneous NED context. To the best of our knowledge, the validity of the TBB method has not yet been studied in heterogeneous context, and with the degree of dependence considered here. Our results broaden considerably the scope for application of the new WTBB as well as the TBB in economics and finance, where the homogeneity of data and the mixing assumption are often a concern. For instance, as shown in Hounyo, Gonçalves and

Meddahi (2013) in the context of noisy diffusion models, due to the heterogeneity of high-frequency financial data, a direct application of the "blocks of blocks" bootstrap method suggested by Politis and Romano (1992) and further studied by Bühlmann and Künsch (1995) fails. To handle both the dependence and heterogeneity of the data (most often in the form of heteroskedasticity), Hounyo, Gonçalves and Meddahi (2013) propose the wild blocks of blocks bootstrap (WBBB), which combine the wild bootstrap with the blocks of blocks bootstrap. This procedure relies on the fact that the heteroskedasticity can be handled elegantly by use of the wild bootstrap, and a block-based bootstrap can be used to treat the serial correlation in the data. In this paper we used a similar approach. The WTBB combine the wild bootstrap with the TBB. The WBBB split a pre-specified blocks of observations into non-overlapping blocks with no tapering. The WTBB differs by using overlapping blocks and tapering. Our bootstrap method constitutes an alternative to the existing methods. Similar to the TBB, the WTBB method involves tapering each overlapping block of the demeaned data first, then a resampling of those tapered blocks. Unlike the TBB, the WTBB does not resample overlapping tapered blocks independently with replacement, but apply the standard wild bootstrap to overlapping tapered blocks in an appropriate way. Our WTBB is intimately related to Paparoditis and Politis's (2001) TBB in the same way that Wu's (1986) wild bootstrap is intimately related to Efron's (1979) bootstrap. The favorable bias and mean squared error properties of the TBB over the MBB are also well preserved by the WTBB. There are two different interpretations of the WTBB method, both valid. One is that the WTBB can be view as a simple variant of the traditional wild bootstrap. The main difference from the traditional wild bootstrap is that the data are first tapered in the blocks in an appropriate way before applying the traditional wild bootstrap on the transformed data. The other interpretation is that the WTBB method is akin to the DWB of Shao (2010b). As the DWB, the WTBB extends the traditional wild bootstrap of Wu (1986) to the time series setting by allowing a transformation of the auxilliary variables involved in the wild bootstrap to be dependent, hence, the WTBB is capable of mimicking the dependence in the original series nonparametrically.

We also generalize the WTBB methodology to cover the case of approximately linear statistics, and M-estimators. The first order asymptotic validity and the favorable asymptotic properties of the WTBB are established in these cases for stationary and weakly dependent time series, as in Paparoditis and Politis (2002).

The remainder of this paper is organized as follows. Section 2 describes the WTBB and its connection to various block-based methods in the context of variance estimation as well as distribution estimation, and states the consistency of this method under the framework of a smooth function model, and M-estimators. Section 3 establishes the consistency of the TBB as well as the WTBB for both variance estimation and distribution approximation of the sample mean when data are assumed to satisfy a NED condition. The results from simulation studies are reported in Section 4. Section 5 concludes. Technical details are relegated to the Appendix.

A word on notation. In this paper, and as usual in the bootstrap literature, P^* (E^* and Var^*)

denotes the probability measure (expected value and variance) induced by the bootstrap resampling, conditional on a realization of the original time series. In addition, let "$\to^{d}$" and "$\to^{P}$" denote convergence in distribution and in probability, respectively, and let $O_P(1)$ and $o_P(1)$ denote being bounded in probability and convergence to zero in probability, respectively. Finally, for $\alpha = (\alpha_1, \ldots, \alpha_d)' \in \mathbb{N}^d$, let D^α denote the differentiable operator $D^\alpha = \frac{\partial^{\alpha_1+\ldots+\alpha_d}}{\partial x_1^{\alpha_1}, \ldots, \partial x_d^{\alpha_d}}$ on $\mathbb{R}^d$.

2 The wild tapered block bootstrap

In this section, to facilitate a comparison between the WTBB and other block-based methods, we restrict our attention to stationary (not heterogeneous) and weakly dependent time series. The more general setting, which allows for dependent heterogeneous arrays is adopted in Section 3. Suppose $X_1, \ldots, X_N$ are observations from the strictly stationary real-valued sequence $\{X_t\}_{t\in Z}$ taking value in $\mathbb{R}^m$ and having mean $\mu = E(X_t)$. Let F denote the marginal distribution of X_t. Suppose the quantity of interest is $T(F)$. Given the observations $X_1, \ldots, X_N$, the goal is to make inferences about $T(F)$ based on some estimator $T_N = T_N(X_1, \ldots, X_N)$. In particular, we are interested in constructing a confidence region for $T(F)$ or constructing an estimate of the variance $\sigma_N^2 = Var\left(\sqrt{N}T_N\right)$, or its asymptotic limit $\sigma_\infty^2 = \lim_{N\to\infty} \sigma_N^2$. Typically, an estimate of the sampling distribution of T_N is required, and the WTBB method proposed here is developed for this purpose.

To define the WTBB, we follow substantially Paparoditis and Politis (2001, 2002). We need to introduce a sequence of data-tapering windows $w_n(\cdot)$ for $n = 1, 2, \ldots$; the weights $w_n(t)$ are value in $[0,1]$, with $w_n(t) = 0$ for $t \notin \{1, 2, \ldots, n\}$. From the above, it is immediate that $\|w_n\|_1 \leq n$ and $\|w_n\|_2 \leq n^{1/2}$, where $\|w_n\|_1 = \sum_{t=1}^{n} |w_n(t)|$ and $\|w_n\|_2 = \left(\sum_{t=1}^{n} w_n^2(t)\right)^{1/2}$. The idea behind the (multiplicative) application of a tapering window to data is to give reduced weight to data near the end-points of the window. The notion of tapering for time series especially in connection to spectral estimation is well-studied; see, for example, Brillinger (1975), Priestley (1981) and Künsch (1989). It is customary to obtain the sequence of data-tapering windows $w_n(\cdot)$ by means of dilations of a single function $w : \mathbb{R} \to [0,1]$, so that

$$w_n(t) = w\left(\frac{t-0.5}{n}\right). \tag{1}$$

We will generally follow Paparoditis and Politis (2001, 2002) and assume that the function $w_n(\cdot)$ satisfies the following assumptions.

Assumption 1. We have $w_n(t) \in [0,1]$ for all $t \in \mathbb{R}$, $w_n(t) = 0$ if $t \notin [0,1]$, and $w_n(t) > 0$ for t in a neighbourhood of $\frac{1}{2}$.

Assumption 2. The function $w_n(t)$ is symmetric about $t = \frac{1}{2}$ and nondecreasing for $t \in \left[0, \frac{1}{2}\right]$.

Assumption 3. The self-convolution is twice continously differentiable at the point $t = 0$, where $(w * w)(t) = \int_{-1}^{1} w(x)\, w(x + |t|)\, dx$.

The WTBB algorithm is defined as follows.

Step 1. First, set a block size l, s.t. $l = l_N \in \mathbb{N}$ and $1 \leq l < N$. Let denote by

$$\bar{X}_{l,w} = \frac{1}{Q}\sum_{j=1}^{Q}\sum_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_1}X_{i+j-1} = \sum_{t=1}^{N}\underbrace{\sum_{j=1}^{Q}\frac{w_l(t-j+1)}{Q\|w_l\|_1}}_{\equiv a_N(t)}X_t = \sum_{t=1}^{N}a_N(t)X_t,$$

the tapered moving (overlapping) block sample mean, where $Q \equiv N - l + 1$. Note that $\sum_{t=1}^{N} a_N(t) = 1$. For $j = 1, \ldots, Q$, let

$$B_{j,l,w} = \left\{\frac{w_l(1)}{\|w_l\|_2}\left(X_j - \bar{X}_{l,w}\right), \frac{w_l(2)}{\|w_l\|_2}\left(X_{j+1} - \bar{X}_{l,w}\right), \ldots, \frac{w_l(l)}{\|w_l\|_2}\left(X_{j+l-1} - \bar{X}_{l,w}\right)\right\}$$

denote the jth centered tapered block of l consecutive observations starting at $\frac{w_l(1)}{\|w_l\|_2}\left(X_j - \bar{X}_{l,w}\right)$.

Step 2. Generate Q independent and identically distributed random variables whose distribution is independent of the original sample $u_1, \ldots, u_Q$ with $E(u_1) = 0$ and $E(u_1)^2 = 1$. For $j = 1, \ldots, Q$, multiply all observations within a given block $B_{j,l}$ by the same external random variable u_j.

Step 3. Finally, the centered WTBB pseudo-time series $\left\{X_t^* - \bar{X}_N, t = 1, 2, \ldots, N\right\}$ is the result of taking the sum of elements of the Q overlapping blocks $B_{t,l}$ of size l that have the same indices. This amounts to generate the WTBB pseudo-time series as follows, for $t = 1, 2, \ldots, N$, let

$$X_t^* - \bar{X}_N = \sum_{j=1}^{Q}\left(\sum_{i=1}^{l}\left(\frac{w_l(i)}{\|w_l\|_2}\left(X_{i+j-1} - \bar{X}_{l,w}\right)\right)1_{\{t\}}(i+j-1)\right)u_j \quad (2)$$

$$= \left(X_t - \bar{X}_{l,w}\right)\underbrace{\left(\sum_{j=1}^{Q}\frac{w_l(t-j+1)}{\|w_l\|_2}u_j\right)}_{\equiv \eta_t} \quad (3)$$

$$= \left(X_t - \bar{X}_{l,w}\right)\eta_t, \quad (4)$$

where $1_{\{\cdot\}}$ is the indicator function.

The WTBB algorithm's describes above with a general data-tapering function $w_n(\cdot)$ is quite compact. It is helpful to focus on some particular cases of this algorithm in order to gain further understanding.

Remark 1. If we let $w(t) = 1_{[0,1]}$ (i.e. no tapering) and $l = 1$, then the WTBB boils down to the wild bootstrap of Wu (1986) exactly as the MBB method of Künsch (1989) coincides with Efron's bootstrap when the bootstrap block size $l = 1$. However we will let l tend to infinity as $N \to \infty$, since in this way we will asymptotically able to mimick the (weak) dependence in the original series nonparametrically. The WTBB is intimately related to Paparoditis and Politis's (2001, 2002) TBB in the same way that Efron's bootstrap is intimately related to Wu's (1986) wild bootstrap. When $w(t) = 1_{[0,1]}$ with $1 \leq l < N$, given (4) the centered wild untapered block bootstrap pseudo-time series are generated as follows, for $t = 1, 2, \ldots, N$,

$$X_t^* - \bar{X}_N = \begin{cases} \left(X_t - \bar{X}_{l,w}\right) \frac{1}{\sqrt{l}} \sum_{j=1}^{t} u_j, & \text{if} \quad t \in \{1, \ldots, l\}, \\ \left(X_t - \bar{X}_{l,w}\right) \frac{1}{\sqrt{l}} \sum_{j=1}^{l} u_{t-l+j}, & \text{if} \quad t \in \{l+1, \ldots, Q\}, \\ \left(X_t - \bar{X}_{l,w}\right) \frac{1}{\sqrt{l}} \sum_{j=1}^{N-t+1} u_{Q-j+1}, & \text{if} \quad t \in \{Q+1, \ldots, N\}, \end{cases} \tag{5}$$

where here $\bar{X}_{l,w} = \frac{1}{Ql} \sum_{j=1}^{Q} \sum_{i=1}^{l} X_{i+j-1}$, since $w(t) = 1_{[0,1]}$.

Remark 2. Obviously we could also use nonoverlapping subseries as in Carlstein (1986). This approach will correspond to nonoverlapping WTBB. For the convenience of presentation, we assume here that $N = kl$. Consequently, in step 1 we will consider only k centered tapered nonoverlapping block of l consecutive observations, with the main difference that observations inside the blocks are not centered around $\bar{X}_{l,w}$ (the tapered moving overlapping block sample mean) but centered around the tapered nonoverlapping block sample mean $\tilde{X}_{l,w}$, where $\tilde{X}_{l,w} = \sum_{j=1}^{k} \sum_{i=1}^{l} \frac{w_l(i)}{\|w_l\|_1} X_{i+(j-1)l}$. Whereas in step 2, we only need to generate k i.i.d. random variables $u_1, \ldots, u_k$ with $E(u_1) = 0$ and $E(u_1)^2 = 1$. Then for $j = 1, \ldots, k$, we multiply all observations within the jth centered tapered nonoverlapping block by the same external random variable u_j. This preserves the dependence within each block. Finally, for step 3, the centered nonoverlapping WTBB pseudo-time series are generated as follows, for $j = 1, 2, \ldots, k$ and $i = 1, 2, \ldots, l$,

$$X_{i+(j-1)l}^* - \bar{X}_N = w_l(i) \frac{l^{1/2}}{\|w_l\|_2} \left(X_{i+(j-1)l} - \tilde{X}_{l,w}\right) u_j. \tag{6}$$

Note that in (6) the "inflation" factor $\frac{l^{1/2}}{\|w_l\|_2}$ is necessary to compensate for the decrease of the variance of the nonoverlapping WTBB observations $X_{i+(j-1)l}^*$'s effected by the shrinking caused by the window w_l (for further details see Paparoditis and Politis (2001)).

Also note that if we let $w(t) = 1_{[0,1]}$ (i.e. no tapering), the nonoverlapping WTBB is equivalent to the blockwise wild bootstrap method studied by Shao (2011) in the context of approximation of the sampling distribution of the Cramer-von Mises test statistic. Recently, Hounyo, Gonçalves, and Meddahi (2013) have proposed a wild blocks of blocks bootstrap method, in the context of noisy diffusion models. In their setting, observations are not stationary, they are heterogeneous. As a result, due to some "mean heterogeneity problem" of high-frequency financial data, they propose to center observations not around the sample mean, but around the blocks sample mean. In particular, their bootstrap method amounts to resample as follows, for $j = 1, \ldots, k$, and $i = 1, \ldots, l$,

$$X_{i+(j-1)l}^* = \begin{cases} \bar{X}_{j+1,l} + \left(X_{i+(j-1)l} - \bar{X}_{j+1,l}\right) u_j, & \text{if } 1 \leq j \leq k-1, \\ \bar{X}_{j,l} + \left(X_{i+(j-1)l} - \bar{X}_{j,l}\right) u_j, & \text{if } j = k, \end{cases} \tag{7}$$

where $\bar{X}_{j,l} = l^{-1} \sum_{i=1}^{l} X_{i+(j-1)l}$.

It is well-known that the nonoverlapping blocks based-method is less efficient than the full-overlap block. In the sequel, we will focus on the (overlapping) WTBB method.

Because in the next subsection we discuss and also link the DWB of Shao (2010b) to the WTBB method, here we briefly introduce the DWB procedure. Given the observations $\{X_t\}_{t=1}^N$, the DWB generates the bootstrap observations according to the equation

$$X_t^{*(DWB)} - \bar{X}_N = \left(X_t - \bar{X}_N\right) \eta_t^{(DWB)}, \quad t = 1, 2, \ldots, N, \tag{8}$$

where the random variables $\left\{\eta_t^{(DWB)}\right\}_{t=1}^N$ are independent of $\{X_t\}_{t=1}^N$ with $E\left(\eta_t^{(DWB)}\right) = 0$ and $Var\left(\eta_t^{(DWB)}\right) = 1$ for $t = 1, 2, \ldots, N$. In addition, $\eta_t^{(DWB)}$ is a stationary process such that $cov\left(\eta_t^{(DWB)}, \eta_{t'}^{(DWB)}\right) = \gamma\left(\left(t - t'\right)/l\right)$, where $\gamma(\cdot)$ is a kernel function with $\int_{-\infty}^{\infty} \gamma(u)\, e^{-iux} dx \geq 0$ for $x \in \mathbb{R}$, and l is a bandwidth parameter. Here and throughout, we use the superscript (DWB) in $X_t^{*(DWB)}$ and $\eta_t^{(DWB)}$ to denote the bootstrap samples and the random variable, respectively obtained by the DWB.

2.1 The sample mean

In this subsection, to elucidate the connection between the WTBB and other block-based methods, we investigate the properties of our bootstrap method for the sample mean first. This corresponds to $T_N = \bar{X}_N$ and the bootstrap estimator T_N^* analogue of T_N is given by $T_N^* = \bar{X}_N^* = N^{-1} \sum_{t=1}^N X_i^*$. A closer inspection of $\bar{X}_N^*$ suggests that, we can also write the centered bootstrap sample mean as

$$\bar{X}_N^* - \bar{X}_N = \frac{1}{Q} \sum_{j=1}^{Q} Z_j u_j = \frac{1}{Q} \sum_{j=1}^{Q} Z_j^* \equiv \bar{Z}_Q^*, \tag{9}$$

where $Z_j = \frac{Q}{N}\left(\sum_{i=1}^{l} \frac{w_l(i)}{\|w_l\|_2} X_{i+j-1} - \bar{X}_{l,w} \frac{\|w_l\|_1}{\|w_l\|_2}\right)$, or as

$$\bar{X}_N^* - \bar{X}_N = \frac{1}{N} \sum_{t=1}^{N} \left(X_t - \bar{X}_{l,w}\right) \eta_t, \tag{10}$$

see Lemma 5.1 in the Appendix for further details. Thus, there are two interpretations of the bootstrap sample mean, both valid. One is that the bootstrap sample mean $\bar{X}_N^*$ is an average of Q independent but not necessarily identically distributed components (see equation (9)). According to this viewpoint, the WTBB is a simple variant of the traditional wild bootstrap (Wu (1986), Liu (1988), Mammen (1993)), which was originally proposed in the context of cross-section linear regression models subject to unconditional heteroskedasticity in the error term. The main difference from the traditional wild bootstrap is that the data are first tapering in the blocks in an appropriate way before applying the traditional wild bootstrap on the transformed data. $\{Z_j\}_{j=1}^Q$ are not independent because they rely on many common observations of the original data $\{X_t\}_{t=1}^N$. However, each observation Z_j is a particular linear combination of all of the original data, as we show, it contains all the relevant information on data dependency required for inference on $\bar{X}_N$. The advantages of tapering were pointed out in detail in Künsch (1989) in connection with his proposal of a tapered block jackknife. Paparoditis and Politis (2001) introduce the TBB method in the bootstrap literature. As in Paparoditis and Politis (2001), the values towards the block endpoints are downweighted in the WTBB procedure. The defined above incorporates the same notion of tapering. Also note that when $w(t) = 1_{[0,1]}$, then the centered bootstrap sample mean as defined in (9) shares with Inoue's (2001) simulation based-method, the fact that the same draws of the random variables $\{u_t\}_{t=1}^{Q}$ are used for all observations within each overlapping block of size l. This preserves the dependence within each block, in order to properly mimic the long-run variance.

The other interpretation is that the centered bootstrap sample mean is the average of N dependent and heteroscedastic distributed components (see equation (10)). According to this viewpoint, the WTBB is akin to the DWB, which is recently proposed by Shao (2010b) for stationary time series. As the DWB, the WTBB extends the traditional wild bootstrap of Wu (1986) to the time series setting by allowing the auxilliary variables $\{\eta_t\}_{t=1}^{N}$ (which are a transformation of $\{u_t\}_{t=1}^{Q}$) involved in the wild bootstrap (see equation (4)) to be dependent, hence, the WTBB is capable of mimicking the dependence in the original series nonparametrically. Similar to the DWB, the dependence between neighboring observations X_t and $X_{t'}$ are not only preserved when the indices t and t' are in the same block as the block-based methods. Whenever $\left|t - t'\right| < l$, X_t^* and $X_{t'}^*$ are conditionally dependent. A common undesirable feature of block-based bootstrap methods is that if the sample size N is not a multiple of the block size l, then one must either take a shorter bootstrap sample or use a fraction of the last resampled block. This could lead to some inaccuracy when the block size is large. In contrast, for the WTBB, the size of the bootstrap sample is always the same as the original sample size. It is worth emphasising that despite the fact that the DWB shares some appealing features with the WTBB, the latter is not a particular case of the DWB method. For instance, unlike the DWB, the random variables $\{\eta_t\}_{t=1}^{N}$, here are not stationary even in the simple case of no tapering (i.e. $w(t) = 1_{[0,1]}$), and observations are centered around $\bar{X}_{l,w}$, and not around $\bar{X}_N$, (see the RHS of (4) and (8)). The WTBB is very easy to implement, and require only as external random variable a simple draw from an i.i.d. distribution as for the plain wild bootstrap.

Let $\hat{\sigma}^2_{l,WTBB}$ denote the WTBB estimate of the asymptotic variance σ^2_∞ based on block size l. A straightforward analytical calculation (see Lemma 5.1 in the Appendix for details) shows that

$$\hat{\sigma}^2_{l,WTBB} = \frac{Q}{N}\hat{\sigma}^2_{l,TBB}, \tag{11}$$

where $\hat{\sigma}^2_{l,TBB} = \frac{1}{Q}\frac{1}{\|w_l\|_2^2}\sum_{j=1}^{n-l+1}\left(\sum_{i=1}^{l} w_l(i)\, X_{i+j-1} - \|w_l\|_1 \bar{X}_{l,w}\right)^2$ is the TBB estimate of the asymptotic variance σ^2_∞ given by Paparoditis and Politis (2001). This implies that the WTBB method preserves the favorable bias and mean squared error properties of the TBB. These favorable asymptotic properties of the WTBB are quantified in the following subsection of a large class of approximately linear statistics.

2.2 Smooth function model

Our aim in this subsection is to show the asymptotic properties of the WTBB under the framework of smooth function model. In particular we derived the favorable bias and mean squared error properties of the WTBB, and establish its consistency distribution approximation. Recall that the applicability of the block-based bootstrap methods is limited to linear or approximately linear statistics that are root-N consistent and asymptotically normal (see for e.g. Shao (2010a)). $T_N = T_N(X_1, \ldots, X_N)$ is said to be approximately linear in a neighborhood of F, if it admits the expansion

$$T_N = T(F) + N^{-1} \sum_{t=1}^{N} IF(X_t, F) + R_N,$$

where the remainder term R_N is appropriately small and $IF(X_t, F)$ is the influence function defined as follows

$$IF(x, F) = \lim_{\epsilon \to 0} \frac{T((1-\epsilon)F + \epsilon \delta_x) - T(F)}{\epsilon},$$

where δ_x representing a unit mass on point x (see e.g. Künsch (1989)). In practice, $IF(X_t, F)$ is unknown, but can be replaced by its empirical counterpart $IF(X_t, \rho_N)$, where $\rho_N = \sum_{t=1}^{N} \delta_{X_t}$ is the empirical measure. Under suitable conditions, we have

$$Var\left(\sqrt{N} T_N\right) = N^{-1} Var\left(\sum_{t=1}^{N} IF(X_t, \rho_N)\right) + o(1).$$

Then we can estimate $Var\left(\sqrt{N} T_N\right)$ by applying the WTBB procedure to $IF(X_t, \rho_N)$. In fact, Paparoditis and Politis (2002) proposed to apply the TBB to $IF(X_t, \rho_N)$. However, as pointed out by Shao (2010a), this implicitly assumed that $IF(X_t, \rho_N)$ is known once we observe the data. This is not necessarily the case in practice. As a remedy, Shao (2010a) propose to taper the random weights in the bootstrap empirical measure. Here, to see whether the WTBB is applicable to the approximately linear statistics, we follow Hall and Mammen (1994) and interpret the WTBB in terms of the generation of random measures. The bootstrapped measure ρ_N^* (corresponding to the WTBB) can be considered as a random distribution with weights at the points $X_1, \ldots, X_N$. Specifically, we can write

$$\rho_N^* = \frac{1}{N} \sum_{t=1}^{N} (\eta_t + 1 - a_N(t) \bar{\eta}_N) \delta_{X_t},$$

where $\bar{\eta}_N = N^{-1} \sum_{t=1}^{N} \eta_t$, with $\{\eta_t\}_{t=1}^{N}$ are random variables satisfying equation (3), $a_N(t) = \sum_{j=1}^{Q} \frac{w_l(t-j+1)}{Q\|w_l\|_1}$, and note that $\sum_{t=1}^{N} a_N(t) = 1$. Hence in the case where $T(F) = \int x dF$, the foregoing formulation amounts to $T_N = \bar{X}_N$ and $T(\rho_N^*) = N^{-1} \sum_{t=1}^{N} (\eta_t + 1 - a_N(t) \bar{\eta}_N) X_t = \bar{X}_N + N^{-1} \sum_{t=1}^{N} \left(X_t - \bar{X}_{l,w}\right) \eta_t$, which coincides with the bootstrapped sample mean under the definition in (10). Note that, for more general nonlinear statistics, it may be difficult to obtain bootstrap samples, because ρ_N^* is not a valid probability measure. However, for a large class of statistics, such as smooth functions of means, the empirical influence function is known. In this case we can apply the WTBB to $IF(X_t, \rho_N)$. For this reason, we follow Hall (1992) and Lahiri (2003) and we restrict our attention to the "smooth function model". This framework is sufficiently general to include many statistics of practical interest, such as autocovariance, autocorrelation, the generalized M-estimators of Bustos (1982), the Yule-Walker estimator, and other interesting statistics in time series. Consider the general class of statistics obtained by functions of linear statistics, i.e. let

$$T_N = f\left(N^{-1}\sum_{t=1}^{N}\phi\left(X_t\right)\right), \tag{12}$$

for some functions $f : \mathbb{R}^d \to \mathbb{R}$, and $\phi : \mathbb{R}^m \to \mathbb{R}^d$. Let $\nabla(x) = \{\partial f(x)/\partial x_1, \partial f(x)/\partial x_2, \ldots, \partial f(x)/\partial x_d\}'$ be the vector of first-order partial derivatives of f at x. Note that the empirical influence function $IF(X_t, \rho_N) = \nabla\left(N^{-1}\sum_{t=1}^{N}\phi(X_t)\right)'\left(\phi(X_t) - N^{-1}\sum_{t=1}^{N}\phi(X_t)\right)$, and is known once the data are observed. Consider now the new series

$$Y_t \equiv IF(X_t, \rho_N) \text{ for } t = 1, 2, \ldots, N,$$

note that a more correct notation for Y_t would be $Y_{t,N}$ but no confusion arises with the simpler notation Y_t. Our proposal is to apply the WTBB algorithm to Y_t. Let denote by $\{Y_t^*, t = 1, 2, \ldots, N\}$ and $\bar{Y}_N^* = N^{-1}\sum_{t=1}^{N} Y_i^*$ the corresponding WTBB pseudo-time series and WTBB sample mean, respectively. Recall that under some suitable conditions, we have $\sqrt{N}(T_N - T(F)) \to^d N\left(0, \sigma_\infty^2\right)$. The sampling distribution of $\sqrt{N}(T_N - T(F))$ can be approximated by using, $N^{1/2}\left(\bar{Y}_N^* - E^*\left(\bar{Y}_N^*\right)\right)$.

To state our results we need a smoothness assumption on the function f.

Assumption 4. The function f is differentiable in a neighborhood of $E(\phi(X_t))$ that is, $N_f = \left\{x \in \mathbb{R}^d : \|x - E(\phi(X_t))\|_2 \leq \epsilon\right\}$ for some $\epsilon > 0$, $\sum_{|\alpha|=1}|D^\alpha f(E(\phi(X_t)))| \neq 0$, and the first partial derivatives of f satisfy a Lipschitz condition of order $s > 0$ on N_f.

As usual, we let $\alpha_X(k) \equiv \sup_{\{A \in \mathcal{F}_{-\infty}^0, B \in \mathcal{F}_k^\infty\}}|P(A \cap B) - P(A)P(B)|$, be the strong mixing coefficients, where $\mathcal{F}_{-\infty}^0$, and $\mathcal{F}_k^\infty$ are the σ-algebras generated by $\{X_n, n \leq 0\}$ and $\{X_n, n \geq k\}$, respectively (see e.g. Rosemblatt (1985)).

Theorem 2.1. *Assume that the function f satisfies the smoothness Assumption 4. Also assume that equation (1), Assumptions 1-3 hold and for some $\delta > 0$, $E\left(|IF(X_t, \rho_N)|^{6+\delta}\right) < \infty$, $\sum_{k=1}^{\infty} k^2 \alpha_X(k)^{\delta/(6+\delta)} < \infty$, $E\left(|\phi_j(X_t)|^{2+\delta}\right) < \infty$ for $j = 1, 2, \ldots, d$, and $\mu_{2+\delta}^* = E^*\left(|u_1|^{2+\delta}\right) < \infty$; recall that $\phi(x) = (\phi_1(x), \ldots, \phi_d(x))'$. If $l_N \to \infty$ as $N \to \infty$ such that $l_N = o\left(N^{1/3}\right)$, then,*

a) *The bias and the variance of $\hat{\sigma}_{l,WTBB}^2$ are respectively given by*

$$E\left(\hat{\sigma}_{l,WTBB}^2\right) - \sigma_\infty^2 = \Gamma/l^2 + o\left(1/l^2\right), \text{ and} \tag{13}$$

$$Var\left(\hat{\sigma}_{l,WTBB}^2\right) = \left(\frac{Q}{N}\right)^2 Var\left(\hat{\sigma}_{l,TBB}^2\right) = \Delta\frac{l}{N} + o(l/N), \tag{14}$$

where the asymptotic variance of T_N *is given by*

$$\sigma_\infty^2 = \sum_{k=-\infty}^{+\infty} R_{IF}(k),$$

and $R_{IF}(k) = cov(IF(X_0, F), IF(X_k, F))$. *The bias and variance constants are calculated to be*

$$\Gamma = \frac{(w*w)''(0)}{2(w*w)(0)} \sum_{k=-\infty}^{\infty} k^2 R_{IF}(k), \text{ and } \Delta = 2\sigma_\infty^4 \int_{-1}^{1} \frac{(w*w)^2(x)}{(w*w)^2(0)} dx.$$

b) *In addition if* $\sigma_\infty^2 > 0$, *we have that*

$$\sup_{x\in\mathbb{R}} \left| P^*\left(\sqrt{N}\left(\bar{Y}_N^* - E^*\left(\bar{Y}_N^*\right)\right) \le x\right) - P\left(\sqrt{N}\left(T_N - T(F)\right) \le x\right)\right| \to^P 0. \tag{15}$$

Part a) of Theorem 2.1 shows that the WTBB method shares with the TBB its favorable bias and mean squared error properties. The bias of $\hat{\sigma}_{l,WTBB}^2$ is of order $O\left(1/l^2\right)$, whereas the untapered block bootstrap results is an estimator of σ_∞^2 of bias $O(1/l)$. It also follows that the MSE of estimator $\hat{\sigma}_{l,WTBB}^2$ is of order $O\left(1/l^4\right) + O(l/N)$. To minimize it, one should pick l proportional to $N^{1/5}$, in which case $MSE\left(\hat{\sigma}_{l,WTBB}^2\right) = O\left(N^{-4/5}\right)$ which is a significant improvement over the $O\left(N^{-2/3}\right)$ rate of the "first generation" block bootstrap methods. Part b) provides a theoretical justification for using the bootstrap distribution of $\sqrt{N}\left(\bar{Y}_N^* - E^*\left(\bar{Y}_N^*\right)\right)$ to estimate the distribution of $\sqrt{N}\left(T_N - T(F)\right)$.

2.3 M-Estimator

M-estimation is a widely used technique for statistical inference. In econometrics, M-estimators are a broad class of estimators, which are obtained as the minima of sums of functions of the data. The aim of this subsection is to show the asymptotic validity, the favorable bias and mean squared error properties of the WTBB for M-estimators. These statistics are often approximately linear and are defined implicitly as solutions of an equation such as

$$\sum_{t=1}^{N} \psi\left(X_{t;}T_N\right) = 0, \tag{16}$$

where the function ψ satisfies conditions strong enough to ensure that

$$\sqrt{N}\left(T_N - T(F)\right) \to^d N\left(0, \sigma_\infty^2\right), \tag{17}$$

a prime example of an M-estimator is the maximum likelihood estimator (MLE). To get some results on the order of magnitude of the bias and variance of the WTBB in the case of M-estimators, we require some technical conditions; to state them, let K be a positive constant, and U a (fixed) open neighborhood of $T(F)$.

$$\psi(x;u) \le K \text{ and } \left|\psi'(x;u)\right| \le K \text{ for all } x\in\mathbb{R} \text{ and } u\in U, \tag{18}$$

where $\psi^{'}(x;u) = \frac{d}{du}\psi(x;u)$.

$$\left|\psi^{'}(x;u_1) - \psi^{'}(x;u_2)\right| \leq K|u_1 - u_2| \text{ for all } x \in \mathbb{R}^m \text{ and } u_1, u_2 \in U, \tag{19}$$

$$E\left(\psi^{'}(X_t; T(F))\right) \neq 0. \tag{20}$$

Under the above conditions as in Paparoditis and Politis (2002) the following theorem holds true.

Theorem 2.2. *Assume the set-up of equation (16) where the function ψ is such that equations (17)-(20) are satisfied. Also assume that the function f satisfies the smoothness assumption 4. Also assume that equation (1), Assumptions 1-3 hold and for some $\delta > 0$, $E\left(|IF(X_t, \rho_N)|^{6+\delta}\right) < \infty$, $\sum_{k=1}^{\infty} k^2 \alpha_X(k)^{\delta/(6+\delta)} < \infty$ and $\mu_3^* = E^*\left(|u_1|^3\right) < \infty$. If $l_N \to \infty$ as $N \to \infty$ but with $l = o\left(N^{1/3}\right)$, then, equations (13)–(15) hold true.*

3 Dependent heterogeneous arrays

In practice, econometricians used data that are typically quite complicated, mixing is too strong a dependence condition to be broadly applicable (see, e.g., Andrews (1984) for an example of a simple AR(1) process that fails to be strong mixing). In this section we adopt the framework of Gonçalves and White (2002), which allows for general dependence conditions and also heterogeneity in data. Suppose $\{X_{Nt}, N, t = 1, 2\ldots\}$ is a double array of not necessarily stationary (can be heterogeneous) random variables defined on a given probability space $(\Omega, \mathcal{F}, P)$ and NED on a mixing process $\{V_t\}$. Let $\mu_{Nt} \equiv E(X_{Nt})$ for $t = 1, 2, \cdots, N$, and let $\bar{\mu}_N = N^{-1}\sum_{t=1}^{N} \mu_{Nt}$ be the parameter of interest to be estimated using the sample mean $\bar{X}_N$ (in the sequel, we will focus on the mean). Following Gonçalves and White (2002) we have established the conditions ensuring the validity of the TBB as well as the WTBB for the sample mean of (possibly heterogeneous) NED functions of mixing processes. We define $\{X_{Nt}\}$ to be NED on a mixing process $\{V_t\}$ if $E\left(X_{Nt}^2\right) < \infty$ and $\upsilon_k \equiv \sup_{N,t}\left\|X_{Nt} - E_{t-k}^{t+k}(X_{Nt})\right\|_2 \to 0$ as $k \to \infty$. Here, $\|X_{Nt}\|_p \equiv (E|X_{Nt}|^p)^{1/p}$ is the L_p norm and $E_{t-k}^{t+k}(\cdot) \equiv E\left(\cdot|\mathcal{F}_{t-k}^{t+k}\right)$, where $\mathcal{F}_{t-k}^{t+k} \equiv \sigma(V_{t-k}, \ldots, V_{t+k})$ is the σ-field generated by $V_{t-k}, \ldots, V_{t+k}$. If $\upsilon_k = O\left(k^{-a-\delta}\right)$ for some $\delta > 0$, we say $\{X_{Nt}\}$ is NED of size $-a$. We assume $\{V_t\}$ is strong mixing. The strong mixing coefficients are $\alpha_k \equiv \sup_m \sup_{\{A\in\mathcal{F}_{-\infty}^m, B\in\mathcal{F}_{m+k}^{\infty}\}} |P(A\cap B) - P(A)P(B)|$, and we require $\alpha_k \to 0$ as $k \to \infty$ at an appropriate rate.

Because in this section we also establish the validity of the TBB method for the sample mean when data are assumed to satisfy a NED condition, here we briefly introduce the TBB procedure of Paparoditis and Politis (2001). For a fixed block size l, s.t. $l = l_N \in \mathbb{N}$ and $1 \leq l < N$, let denote by $\mathcal{B}_{j,l} = \{X_{Nj}, \ldots, X_{N,j+l-1}\}$ be the jth block, $j = 1, \ldots, Q = N - l + 1$. The number of blocks in the bootstrap sample is denoted by $k = \lfloor N/l \rfloor$. For the convenience of presentation, we assume that $N = lk$. The TBB consists of two steps: (1) let $I_0, \ldots, I_{k-1}$ be i.i.d. random variables uniformly distributed on the set $\{1, 2, \dots, Q\}$; and (2) for $m = 0, 1, \dots, k-1$, let

$$X^{*(TBB)}_{N,ml+i} = w_l(i) \frac{l^{1/2}}{\|w_l\|_2} \left(X_{N,I_m+i-1} - \bar{X}_N \right), \ i = 1, \dots, l. \tag{21}$$

Here and throughout, we use the superscript (TBB) in $X^{*(TBB)}_{N,ml+i}$ to denote the bootstrap samples obtained by the TBB. When $w(t) = 1_{[0,1]}$, the TBB reduces to the MBB. Note that the TBB uses the same block resampling scheme as for the MBB method, but each resampled MBB block is replaced by a tapered version. In order to state our results, we follow Gonçalves and White (2002) and make the following assumption to establish the validity of the TBB and the WTBB methods in this heterogeneous NED context:

Assumption 5.

a) For some $r > 0$, $\|X_{Nt}\|_{3r} \leq \Delta < \infty$ for all $N, t = 1, 2, \dots$

b) $\{X_{Nt}\}$ is near epoch dependent (NED) on $\{V_t\}$ with NED coefficients α_k of size $-\frac{2(r-1)}{(r-2)}$; $\{V_t\}$ is an α-mixing sequence with α_k of size $-\frac{2r}{r-2}$.

As Gonçalves and White (2002) pointed out, we also found in Theorem 3.1 below that under arbitrary heterogeneity in $\{X_{Nt}\}$ the TBB variance estimator $\hat{\sigma}^2_{l,TBB}$ is not consistent for σ^2_N, but for $\sigma^2_N + U_N$. The bias term U_N is related to the heterogeneity in the means $\{\mu_{Nt}\}$ and can be interpreted as the TBB variance estimate of the scaled sample mean $\sqrt{N}\bar{\mu}^{*(TBB)}_N = N^{-1/2}\sum_{t=1}^N \mu^{*(TBB)}_{Nt}$ that would result if we could resample the vector time series $\{\mu_{Nt}\}$. We follow Gonçalves and White (2002) and call $\left\{\mu^{*(TBB)}_{Nt}\right\}$ the "resampled version" of $\{\mu_{Nt}\}$. The variance σ^2_N can be easily obtained by using the TBB variance $\hat{\sigma}^2_{l,TBB}$ under some homogeneity condition. The following Lemma and its corollary provide the theoretical justification.

Theorem 3.1. *Assume $\{X_{Nt}\}$ satisfies Assumptions 1-3 and Assumption 5. If $l_N \to \infty$ as $N \to \infty$ such that $l_N = o\left(N^{1/2}\right)$, then,*

a) $\hat{\sigma}^2_{l,TBB} - \left(\sigma^2_N + U_N\right) \xrightarrow{P} 0$, *where* $U_N \equiv Var^*\left(N^{-1/2}\sum_{t=1}^N \mu^{*(TBB)}_{Nt}\right)$.

b) $U_N = \sum\limits_{\tau=-l+1}^{l-1} \frac{v_l(\tau)}{v_l(0)} \sum\limits_{t=1}^{N-|\tau|} \beta_{N,t,\tau}\left(\mu_{Nt} - \bar{\mu}_{l,w}\right)\left(\mu_{N,t+|\tau|} - \bar{\mu}_{l,w}\right)$,

where $v_l(\tau) = \sum_{i=1}^{l-|\tau|} w_l(i)\, w_l(i+|\tau|)$, $\bar{\mu}_{l,w} = \sum_{t=1}^N a_N(t)\,\mu_{Nt}$, *and*

$\beta_{N,t,\tau} = \frac{1}{v_l(\tau)}\frac{1}{Q}\sum_{j=1}^Q w_l(t-j+1)\, w_l(t-j+1+|\tau|)$ *with* $\tau < j$.

c) $\hat{\sigma}^2_{l,TBB} - \sigma^2_N \xrightarrow{P} 0$, *as* $\lim_{N\to\infty} U_N = 0$.

Thus, the condition $\lim_{N\to\infty} U_N = 0$ is the homogeneity condition on the mean, analogous conditions is given by Liu (1988) and by Gonçalves and White (2002). To ensure this condition, one can for example suppose that

Assumption 6 $N^{-1}\sum_{t=1}^{N}(\mu_{Nt}-\bar{\mu}_N)^2 = o\left(l_N^{-1}\right)$ where $l_N = o\left(N^{1/2}\right)$.

Assumption 6 amounts to Assumption 2.2 in Gonçalves and White (2002). As they explain, this assumption is rather general allowing for breaks in mean. See Gonçalves and White (2002) for particular examples of processes that satisfy Assumption 6.

The following consistency result holds under Assumptions 1-3, Assumptions 5-6 and is an immediate consequence of the previous Theorem 3.1.

Corollary 4.1. Assume $\{X_{Nt}\}$ satisfies Assumptions 1-3, Assumptions 5-6. If $l_N \to \infty$ as $N \to \infty$ such that $l_N = o\left(N^{1/2}\right)$, then,

a) $\hat{\sigma}^2_{l,TBB} - \sigma^2_N \xrightarrow{P} 0$.

b) $\hat{\sigma}^2_{l,WTBB} - \sigma^2_N \xrightarrow{P} 0$; recall that $\hat{\sigma}^2_{l,WTBB} = \frac{Q}{N}\hat{\sigma}^2_{l,TBB}$ and $\frac{Q}{N} \to 1$ as $N \to \infty$.

This result extends the previous consistency results on $\hat{\sigma}^2_{l,TBB}$ by Paparoditis and Politis (2001) as well as our new estimator $\hat{\sigma}^2_{l,WTBB}$ (when the statistics of interest is the sample mean), for stationary mixing observations to the case of NED functions of a mixing process. In particular, Corollary 4.1 contains a version of Theorem 1 and Theorem 2 of Paparoditis and Politis (2001) and our Theorem 2.1 as a special case, when $\{X_t\}$ is a stationary α-mixing sequence, under the same moment conditions and weaker α-mixing conditions, but under the stronger requirement that $l_N = o\left(N^{1/2}\right)$ instead of $l_N = o(N)$. Here we show that the variance of $\hat{\sigma}^2_{l,TBB}$ and $\hat{\sigma}^2_{l,WTBB}$ are $O\left(\frac{l^2}{N}\right)$, instead of the previous sharper result $O\left(\frac{l}{N}\right)$ when data are stationary, which explains the loss of $l_N = o(N)$.

The next theorem establishes the first order asymptotic validity for the TBB and the WTBB under general dependence conditions. As in Gonçalves and White (2002), we require a slightly stronger dependence condition than Assumption 5.b). Specifcally, we impose:

Assumption 5.b') For some small $\delta > 0, \{X_{Nt}\}$ is $L_{2+\delta}$-NED on $\{V_t\}$ with NED coefficients υ_k of size $-\frac{2(r-1)}{r-2}$; $\{V_t\}$ is an α-mixing sequence with α_k of size $-\frac{(2+\delta)r}{r-2}$.

The next theorem states the consistency results for the TBB as well as the WTBB.

Theorem 3.2. *Assume $\{X_{Nt}\}$ satisfies Assumption 5-6, strengthened by Assumption 5.b'). Also assume equation (1), and Assumptions 1-3. If $l_N \to \infty$ as $N \to \infty$ such that $l_N = o\left(N^{1/2}\right)$, then*

a) $\sup_{x\in\mathbb{R}}\left|P^*\left(N^{1/2}\left(\bar{X}_N^{*(TBB)} - E^*\left(\bar{X}_N^{*(TBB)}\right)\right) \le x\right) - P\left(N^{1/2}\left(\bar{X}_N - \bar{\mu}_N\right) \le x\right)\right| = o_P(1)$.

b) $\sup_{x\in\mathbb{R}}\left|P^*\left(N^{1/2}\left(\bar{X}_N^{*} - E^*\left(\bar{X}_N^{*}\right)\right) \le x\right) - P\left(N^{1/2}\left(\bar{X}_N - \bar{\mu}_N\right) \le x\right)\right| = o_P(1)$, *if for any $\delta > 0$, $E^*\left|u_j\right|^{2+\delta} < \infty$.*

Theorem 3.2 justifies using the TBB as well as the WTBB to build asymptotically valid confidence intervals for (or test hypotheses about) $\bar{\mu}_N$, even though there may be considerable heterogeneity.

Part a1) of Theorem 3.2 is an extension of Theorem 3 of Paparoditis and Politis (2001) to the case of dependent heterogeneous double arrays of random variables, where the stationary mixing assumption is replaced by the more general assumption of a (possibly heterogeneous) double array near epoch dependent on a mixing process. Thus here, we allow for more dependence and heterogeneity in the data. Even if part a) of Theorem 3.2 states results under the same assumptions as Gonçalves and White (2002), note that this result also can be seen as a generalisation of Gonçalves and White's (2002) results for the MBB method. Since, the MBB is a particular case of the TBB method.

Up to this point, we have justified the consistency of the WTBB for distribution and variance approximation under the framework of the smooth function model for stationary (not heterogeneous) and weakly dependent time series. Whereas for the sample mean we show the consistency of the TBB as well as the WTBB for distribution and variance approximation under a wide class of data generating processes, the processes near epoch dependent on a mixing process.

A natural question is whether the WTBB distribution can offer the second-order correctness, that is better than normal approximation. If the external random variables $\{u_t\}_{t=1}^{Q}$, in addition to having mean 0 and variance 1, also has its third central moment equal to 1, we conjectured that the WTBB would share with the Wu's wild bootstrap and block-based bootstrap methods the property of higher-order accuracy after studentization/ standardisation and under some additional regularity conditions, although a rigorous proof is well beyond the scope of this paper. The proof of this claim requires the development of valid Edgeworth expansions for the WTBB distribution (see for example Lahiri (1991) or Gotze and Künsch (1996)). Here we follow Paparoditis and Politis (2001, 2002) and merely give an informal justification of the superiority of the unstudentised WTBB distribution estimator over its block bootstrap counterpart.

Note that the Berry-Essen bound (25) given in the proof of part b) of Theorem 2.1 reveals that not only equation (15) is true, but in addition, choosing l proportional to $N^{1/5}$, it follows that

$$\sup_{x\in\mathbb{R}}\left|P^*\left(\sqrt{N}\left(\bar{Y}_N^* - E^*\left(\bar{Y}_N^*\right)\right)\leq x\right) - P\left(\sqrt{N}\left(T_N - T\left(F\right)\right)\leq x\right)\right| = O_P\left(N^{-1/2}\right). \qquad (22)$$

Recall that the untapered block bootstrap analog of (22) would have a RHS of order $O_P\left(N^{-1/3}\right)$ which is must worse. More interessing the TBB analog of (22) (cf. equation (16) of Paparoditis and Politis (2002)) would have a RHS of order $O_P\left(N^{-2/5}\right)$ which is must worse than $O_P\left(N^{-1/2}\right)$ for the new WTBB method.

4 Simulation studies

In this section, we study via simulations the finite-sample performance of the WTBB compared to the MBB, TBB, and DWB methods for the sample mean. Performance is measured in terms of coverage probability of two-sided 95% level intervals. In the simulation studies, we considered two different models generating the observations, namely:

Model 1. Nonlinear autoregressive model, NAR,

$$X_t = \rho \sin(X_{t-1}) + \upsilon_t,$$

for $t \in \mathbb{Z}$, where $\{\upsilon_t\}$ i.i.d. $N(0,1)$, with $\rho \in \{0.2, 0.6\}$.

Model 2. Heteroskedastic autoregressive AR(1),

$$X_t = \rho X_{t-1} + \upsilon_t, \quad \text{and} \quad \upsilon_t = s_t \widetilde{\upsilon}_t,$$

for $t \in \mathbb{Z}$, where $\{\widetilde{\upsilon}_t\}$ i.i.d. $N(0,1)$, with $\rho \in \{0.2, 0.8\}$. Here $\{s_t\}$ denotes a sequence of real numbers that might be regarded as seasonal effects. Throughout, we choose $\{s_t\}$ to be the infinite repetition of the sequence $\{1,1,1,2,3,1,1,1,1,2,4,6\}$.

Note that, among the block-based bootstrap methods, the theoretical advantage of the TBB over the MBB has been confirmed for model 1, (in particular, with $\rho = 0.6$) through simulation studies by Paparoditis and Politis (2001). For this reason, it seems natural to study the new WTBB method in this case. We also consider model 2, in order to investigate the performance of the WTBB when there are dependent "strongly" heterogeneous data. This model is used by Politis Romano and Wolf (1997) in another context for heteroskedastic times series. Note that in this model, the innovations are independent but heteroskedastic. Then model 2 generates a weakly dependent, heteroskedastic time series.

We generate repeated trials of length $N = 200$ from these processes. The block sizes range from $l = 1$ to $l = 40$. In order to generate the TBB as well as the WTBB observations we need a data-tapering window function $w(\cdot)$. We define the following family of trapezoidal functions as

$$w_c^{trap}(t) = \begin{cases} \frac{t}{c}, & \text{if} \quad t \in [0,c], \\ 1, & \text{if} \quad t \in [c, 1-c], \\ \frac{1-t}{c}, & \text{if} \quad t \in [1-c, 1], \\ 0, & \text{if} \quad t \notin [0,1], \end{cases} \tag{23}$$

where c is some fixed constant in $(0, 1/2]$. To make the comparison fair, in our simulation, we took $c = 0.43$, since it was found in Paparoditis and Politis (2001) that $w(t) = w_{0.43}^{trap}(t)$ offers the optimal (theoretical) MSE provided we fix the covariance structure of a time series. We also use $\gamma(t) = \left(w_{0.43}^{trap} * w_{0.43}^{trap}\right)(t) / \left(w_{0.43}^{trap} * w_{0.43}^{trap}\right)(0)$, where $\gamma(\cdot)$ is the covariance function of the external random variable $\eta^{(DWB)}$ used to generate the DWB observations. With this choice of the kernel function for the DWB, the favorable bias and MSE properties of the TBB variance estimator over other block-based counterparts in the mean case automatically carries over to the DWB. We used $\left\{\eta_t^{(DWB)}\right\}_{t=1}^{N}$ multivariate normal as in Shao (2010b), whereas to generate the WTBB data we use three different external random variables.

WTBB1 $u_j \sim$ i.i.d. $N(0,1)$, implying that $E^*(u_j) = 0$, $E^*(u_j^2) = 1$, $E^*(u_j^3) = 0$ and $E^*(u_j^4) = 3$.

WTBB2 A two point distribution $u_j \sim$ i.i.d. suggested by Mammen (1993) such that:

$$u_j = \begin{cases} \frac{1+\sqrt{5}}{2}, & \text{with} \quad \text{prob} \quad p = \frac{\sqrt{5}-1}{2\sqrt{5}} \\ \frac{1-\sqrt{5}}{2}, & \text{with} \quad \text{prob} \quad 1-p \end{cases},$$

for which $E^*(u_j) = 0$, $E^*(u_j^2) = E^*(u_j^3) = 1$ and $E^*(u_j^4) = 2$.

WTBB3 The so-called Rademacher, i.e. the two point distribution $u_j \sim$ i.i.d. proposed by Liu (1988) such that:

$$u_j = \begin{cases} 1, & \text{with} \quad \text{prob} \quad p = \frac{1}{2} \\ -1, & \text{with} \quad \text{prob} \quad 1-p \end{cases},$$

for which we have $E^*(u_j) = 0$, $E^*(u_j^2) = 1$, $E^*(u_j^3) = 0$ and $E^*(u_j^4) = 1$.

Note that all three choices of u_j are asymptotically valid when used to construct the unstudentized bootstrap intervals or to estimate σ_N^2, since the conditions $E^*(u_j) = 0$ and $E^*(u_j^2) = 1$ are satisfied. In the case of independent but not necessarily identically distributed observations, the further condition $E^*(u_j^3) = 1$ (satisfied by WTBB2) is often added as a necessary condition for refinement for the traditional wild bootstrap. The Rademacher distribution (WTBB3) also satisfies the necessary conditions for refinements in the case of unskewed disturbances. Davidson and Flachaire (2007) advocated the use of the Rademacher distribution.

For each time series and each block size, we generated 999 MBB, TBB, DWB and WTBB pseudo-series to obtain the bootstrap-based critical values. Then we repeated this procedure 1000 times and plotted the empirical coverage of nominal 95% symmetric confidence intervals as a function of block size in Figure 1. For all bootstrap methods, finite sample performance is far from perfect (especially for model 2) and gets worse as the degree of dependence in the data increases. Model 2 exhibits overall larger coverage distortions than model 1. For the WTBB method, in our simulations, none of the three resampling schemes (i.e., WTBB1, WTBB2 and WTBB3) clearly dominates the others.

Starting with model 1, as a first observation (cf. Figure 1 (a) and (b)), it is striking how close all bootstrap methods (MBB, TBB, DWB and WTBB) analyzed here are in terms of empirical coverage rate for small block size (for $N = 200$, with $l \leq 8$). As l increases, the difference may be considerable between the WTBB and the MBB. But the two methods DWB and WTBB are still close, the difference is less than 2.5 percentage point in most cases, with the WTBB noticeably superior to the TBB. For $\rho = 0.6$, the largest coverage rate of the WTBB is 93.2% given by the block size $l = 9$, whereas it is 91.3% for the MBB with the block size $l = 11$, instead of the desired nominal 95%. The empirical coverage distortions seem to increase with increases in l for $l \geq 15$.

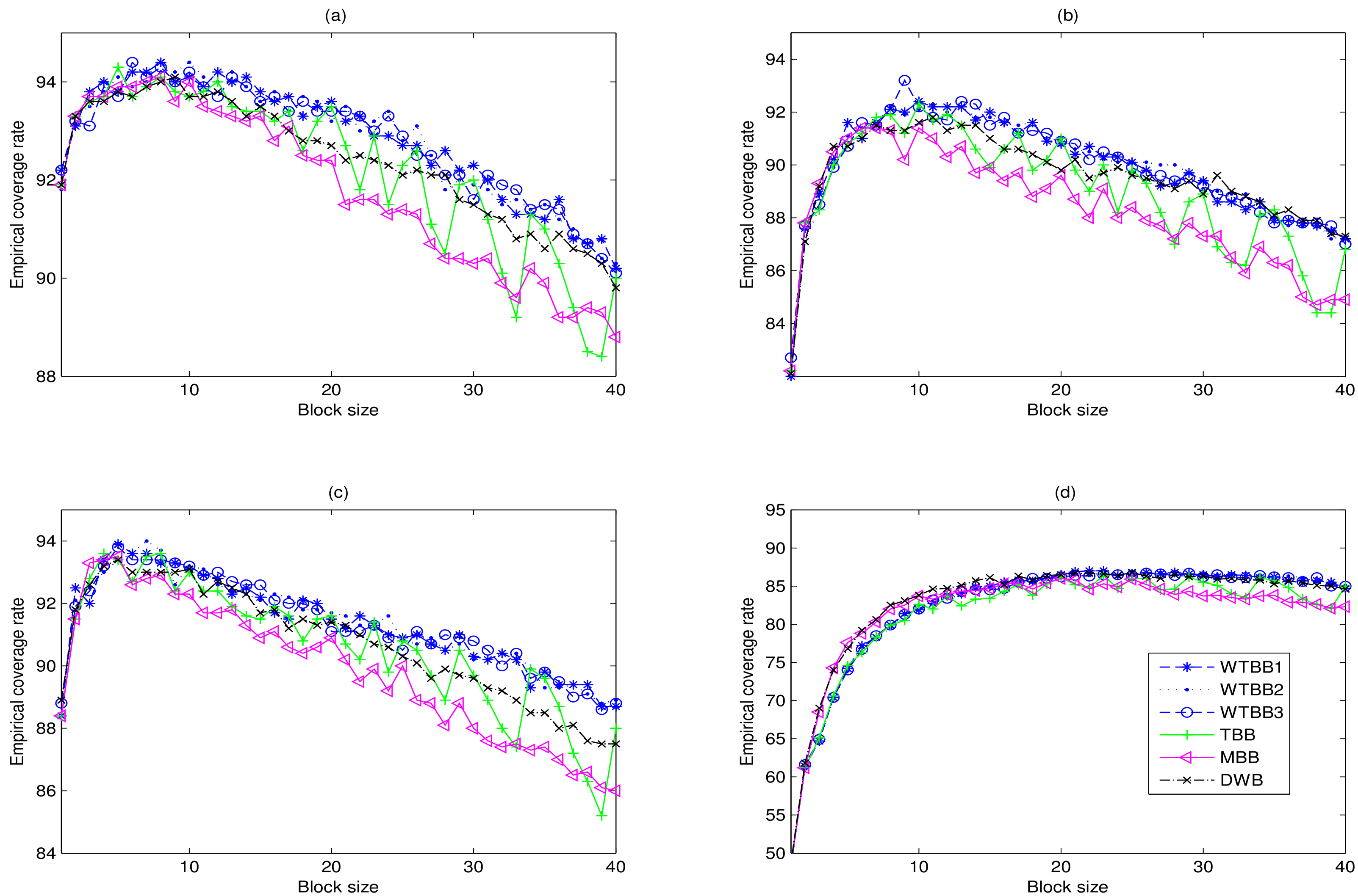

Figure 1: Empirical coverage, as a function of the block size l, of 95% symmetric confidence intervals of the mean, obtained for the nonlinear autoregressive model, Model 1, and the heteroskedastic autoregressive AR model, Model 2, for a sample size $N = 200$. (a) Model 1 with $\rho = 0.2$. (b) Model 1 with $\rho = 0.6$. (c) Model 2 with $\rho = 0.2$. (d) Model 2 with $\rho = 0.8$.

Turning now to the analysis of model 2, Figure 1 (c) and (d) shows that finite-sample coverage distortions are slightly larger. It appears that the advantage of "tapering" over "no-tapering" methods is noticeable for moderately large block sizes. Indeed, the MBB seems to perform very poorly compared to bootstrap schemes using tapering (i.e., TBB, DWB, and WTBB). In particular, for $\rho = 0.8$, the MBB-based intervals undercover consistently for large l. It turns out that this kind of heteroskedasticity, generated by model 2, had moderate impact on the performance of bootstrap methods (i.e., MBB, TBB, DWB, and WTBB) studied here. The performance of the WTBB is slightly better than that of the DWB, for large block size. Based on the foregoing simulations results, the WTBB, DWB and the TBB are the three best bootstrap methods we would recommend in the case of "strongly" heteroskedastic times series. No formal theoretical results exist that may justify the use of DWB in this context.

In the foregoing simulation studies, we do not consider the issue of bandwidth selection, which is very important in practice. In view of the connection between the TBB and WTBB, in particular $\hat{\sigma}^2_{l,WTBB} = \frac{Q}{N}\hat{\sigma}^2_{l,TBB}$, for MSE-optimal block size, the practical block size choice suggested by Paparoditis and Politis (2002) is expected to work for the WTBB. However, the optimal block size for MSE may be suboptimal for the purpose of distribution estimation; see Hall, Horowitz and Jing (1995). We will not pursue this approach further here. We leave this analysis to future work.

5 Concluding remarks

This paper proposes a new bootstrap method for time series, the WTBB, that is generally applicable to variance estimation and sampling distribution approximation for the smooth function model. Within the framework of the smooth function model, we show that the WTBB is asymptotically equivalent to the TBB, which outperforms all other block-based methods in terms of the bias and MSE. Computationally, it is very convenient to implement the new WTBB method. In particular, the choice of the external random variable is very flexible, as for the plain wild bootstrap.

In the case of the sample mean of dependent heterogeneous data, we establish the first order asymptotic validity of the WTBB as well as the TBB. In particular, we show that the WTBB and the TBB variance estimators for the sample mean are consistent under a wide class of data generating processes, the processes near epoch dependent on a mixing process. Finally, simulation studies demonstrate that the WTBB performs well even for moderate sample sizes and in most cases outperforms other bootstrap procedures that take autocorrelation into account. It merits considerable further study.

Simulation evidence also indicates that the DWB seems to be valid for dependent heterogeneous data. We did not attempt to show the theoretical validity of the DWB for dependent heterogeneous arrays. We plan on investigating this issue in future work. Another promising extension is to study the higher-order accuracies of the TBB, DWB, and WTBB methods.

Appendix

Lemma 5.1. *Let $\{X_t^*, t = 1, 2, \ldots, N\}$ be a sequence of the WTBB pseudo-time series, we have that*

a) $\bar{X}_N^* - \bar{X}_N = \frac{1}{N}\sum_{t=1}^{N}\left(X_t - \bar{X}_{l,w}\right)\eta_t$, *where* $\eta_t = \sum_{j=1}^{Q}\frac{w_l(t-j+1)}{\|w_l\|_2}u_j$.

b) $\bar{X}_N^* - \bar{X}_N = \frac{1}{Q}\sum\limits_{j=1}^{Q} Z_j u_j = \frac{1}{Q}\sum\limits_{j=1}^{Q} Z_j^* \equiv \bar{Z}_Q^*$, *where* $Z_j = \frac{Q}{N}\left(\sum\limits_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_2}X_{i+j-1} - \bar{X}_{l,w}\frac{\|w_l\|_1}{\|w_l\|_2}\right)$.

c) $Var^*\left(\sqrt{N}\bar{X}_N^*\right) = \frac{Q}{N}\left(N\sigma^2_{Jack}\right) = \frac{Q}{N}\hat{\sigma}^2_{l,TBB}$,

where $\hat{\sigma}^2_{l,TBB} = N\sigma^2_{Jack} = \frac{1}{Q}\frac{1}{\|w_l\|_2^2}\sum\limits_{j=1}^{n-l+1}\left(\sum\limits_{i=1}^{l} w_l(i)\, X_{i+j-1} - \|w_l\|_1\,\bar{X}_{l,w}\right)^2$.

Proof of Lemma 5.1 part a). Result follows directly given equation (4) and the definition of $\bar{X}_N^*$.
Proof of Lemma 5.1 part b). Given equation (4) and the definition of $\bar{X}_N^*$, we have that

$$\begin{aligned}
\bar{X}_N^* - \bar{X}_N &= \frac{1}{N}\sum_{t=1}^{N}\left(X_t - \bar{X}_{l,w}\right)\eta_t \\
&= \frac{1}{N}\sum_{t=1}^{N}\left(\left(\sum_{j=1}^{Q}\frac{w_l\left(t-j+1\right)}{\|w_l\|_2}u_j\right)\left(X_t - \bar{X}_{l,w}\right)\right) \\
&= \frac{1}{N}\sum_{t=1}^{N}\left(\sum_{j=1}^{Q}\left(\frac{w_l\left(t-j+1\right)}{\|w_l\|_2}\left(X_t - \bar{X}_{l,w}\right)\right)u_j\right).
\end{aligned}$$

Given that $w_l(j) = 0$ if $j \notin \{1, 2, \ldots, l\}$, we can write

$$\begin{aligned}
\bar{X}_N^* - \bar{X}_N &= \frac{1}{N}\sum_{j=1}^{Q}\left(\sum_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_2}\left(X_{i+j-1} - \bar{X}_{l,w}\right)\right)u_j \\
&= \frac{1}{N}\sum_{j=1}^{Q}\left(\sum_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_2}X_{i+j-1} - \bar{X}_{l,w}\sum_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_2}\right)u_j \\
&= \frac{1}{N}\sum_{j=1}^{Q}\left(\sum_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_2}X_{i+j-1} - \bar{X}_{l,w}\frac{\|w_l\|_1}{\|w_l\|_2}\right)u_j,
\end{aligned}$$

it follows that

$$\begin{aligned}
\bar{X}_N^* - \bar{X}_N &= \frac{1}{Q}\sum_{j=1}^{Q}\underbrace{\frac{Q}{N}\left(\sum_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_2}X_{i+j-1} - \bar{X}_{l,w}\frac{\|w_l\|_1}{\|w_l\|_2}\right)}_{\equiv Z_j}u_j \\
&= \frac{1}{Q}\sum_{j=1}^{Q}Z_j u_j = \frac{1}{Q}\sum_{j=1}^{Q}Z_j^* \equiv \bar{Z}_Q^*.
\end{aligned}$$

Proof of Lemma 5.1 part c). Given part b) of Lemma **5.1,** we can write

$$\begin{aligned} Var^*\left(\sqrt{N}\bar{X}_N^*\right) &= Var^*\left(\sqrt{N}\bar{Z}_Q^*\right) \\ &= \frac{N}{Q^2}\sum_{j=1}^{Q} Var^*\left(Z_j u_j\right) = \frac{N}{Q^2}\sum_{j=1}^{Q} Z_j^2 Var^*\left(u_j\right) \\ &= \frac{1}{N}\sum_{j=1}^{Q}\left(\sum_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_2}X_{i+j-1} - \bar{X}_{l,w}\frac{\|w_l\|_1}{\|w_l\|_2}\right)^2 \underbrace{Var\left(u_j\right)}_{=1} \\ &= \frac{Q}{N}\underbrace{\frac{1}{Q}\frac{1}{\|w_l\|_2^2}\sum_{j=1}^{Q}\left(\sum_{i=1}^{l} w_l(i)\, X_{i+j-1} - \|w_l\|_1\bar{X}_{l,w}\right)^2}_{=N\sigma^2_{jack}=\hat{\sigma}^2_{l,TBB}}, \end{aligned} \tag{24}$$

where $\|w_l\|_1 \bar{X}_{l,w} = \frac{1}{Q}\sum_{j=1}^{Q}\sum_{i=1}^{l} w_l(i)\, X_{i+j-1}$, and σ^2_{Jack} is the tapered jackknife variance estimator defined in Künsch (1989, p. 1220).

Proof of Theorem 2.1 Part a). Results follow respectively from Theorem 2.1 of Paparoditis and Politis (2002) in conjunction with part c) of our Lemma 5.1 since $\hat{\sigma}^2_{l,WTBB} = \frac{Q}{N}\hat{\sigma}^2_{l,TBB}$, $\frac{Q}{N} \to 1$ and under our assumed conditions the variance of the linearized statistic $N^{-1}\sum_{t=1}^{N} IF\left(X_t, \rho_N\right)$ approximates well the variance of the nonlinear statistic T_N.

Proof of Theorem 2.1 Part b). The proof of this result follows closely that of equation (10) of Theorem 2.1 of Paparoditis and Politis (2002**).** First note that the assumed conditions are sufficient to ensure that the statistic $N^{-1}\sum_{t=1}^{N}\phi\left(X_t\right)$ is asymptotically normal at rate $\sqrt{N}$. Thus a Taylor expansion of f around $E\left(\phi\left(X_t\right)\right)$ confirms that $\sqrt{N}\left(T_N - T\left(F\right)\right) \to^d N\left(0, \sigma^2_\infty\right)$. Therefore, to prove part b) of Theorem 2.1, we just need to show that the WTBB distribution is approximately close to $\Phi\left(x/\sigma_\infty\right)$, where $\Phi\left(\cdot\right)$ denotes the standard normal distribution function. Note that $N^{1/2}\left(\bar{Y}_N^* - E^*\left(\bar{Y}_N^*\right)\right) = \sum_{j=1}^{Q} z_j^*$, where $z_j^* = \frac{N^{1/2}}{Q}\left(Z_j u_j - E^*\left(Z_j u_j\right)\right)$. Also note that $E^*\left(z_j^*\right) = 0$ and $Var^*\left(\sum_{j=1}^{Q} z_j^*\right) = \hat{\sigma}^2_{l,WTBB} \xrightarrow{P} \sigma^2_\infty$ by part a) of Theorem 2.1. Moreover, since $z_1^*, \ldots, z_Q^*$ are conditionally independent, by the Berry-Esseen bound, for some small $\delta > 0$ and for some constant $K > 0$ (which changes from line to line),

$$\sup_{x\in\mathbb{R}}\left|P^*\left(N_N^{1/2}\left(\bar{Y}_N^* - E^*\left(\bar{Y}_N^*\right)\right) \le x\right) - \Phi\left(x/\sigma_\infty\right)\right| \le K\sum_{j=1}^{Q} E^*\left|z_j^*\right|^{2+\delta}.$$

To check the Berry-Esseen conditions, we can bound $\sum_{j=1}^{Q} E^*\left|z_j^*\right|^3$ which converges to zero in probability as $l \to \infty$, $N \to \infty$ such that $l = o\left(N^{1/3}\right)$. Indeed, we have that

$$\begin{aligned}\sum_{j=1}^{Q} E^* \left|z_j^*\right|^3 &= \sum_{j=1}^{Q} E^* \left|\frac{N^{1/2}}{Q}\left(Z_j u_j - E^*\left(Z_j u_j\right)\right)\right|^3 \\ &\leq 2\frac{N^{3/2}}{Q^3}\sum_{j=1}^{Q} E^* \left|Z_j u_j\right|^3 = 2\frac{N^{3/2}}{Q^3}\sum_{j=1}^{Q} E^* \left|Z_j\right|^3 E^* \left|u_j\right|^3,\end{aligned}$$

where the inequality follows from the C_r and the Jensen inequalities. Given the definition of $Z_j = \frac{Q}{N}\left(\sum_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_2}Y_{i+j-1} - \bar{Y}_{l,w}\frac{\|w_l\|_1}{\|w_l\|_2}\right)$, and the fact that by assumption $E^*\left|u_j\right|^3 < \infty$, we can write

$$\begin{aligned}\sum_{j=1}^{Q} E^* \left|z_j^*\right|^3 &\leq K\frac{N^{3/2}}{Q^3}\sum_{j=1}^{Q}\frac{Q^3}{N^3}\left|\sum_{i=1}^{l}\frac{w_l(i)}{\|w_l\|_2}Y_{i+j-1} - \bar{Y}_{l,w}\frac{\|w_l\|_1}{\|w_l\|_2}\right|^3 \\ &\leq K\frac{Q}{N}\left(\frac{l^3}{N}\right)^{1/2}\left(\frac{1}{Q}\sum_{j=1}^{Q}\left|\frac{1}{l}\sum_{i=1}^{l} w_l(i)\frac{l^{1/2}}{\|w_l\|_2}Y_{i+j-1}\right|^3 + \left|\frac{\|w_l\|_1\bar{Y}_{l,w}}{l^{1/2}\|w_l\|_2}\right|^3\right).\end{aligned}$$

Also note that, we can write

$$\frac{1}{Q}\sum_{j=1}^{Q}\left|\frac{1}{l}\sum_{i=1}^{l} w_l(i)\frac{l^{1/2}}{\|w_l\|_2}Y_{i+j-1}\right|^3 = O\left(\frac{1}{Q}\sum_{j=1}^{Q}\left|\frac{1}{l}\sum_{i=1}^{l} Y_{i+j-1}\right|^3\right).$$

In the above we follow the proof of Theorem 3 of Paparoditis and Politis (2001) and used the facts that $w_l(i) \leq 1$, and $l^{1/2}/\|w_l\|_2 = O(1)$; the latter follows because equation (1) implies that $\|w_l\|_2/l \to \int_0^1 w^2(t)\,dt > 0$ by assumption 1. Thus we have

$$\frac{1}{Q}\sum_{j=1}^{Q}\left|\frac{1}{l}\sum_{i=1}^{l} w_l(i)\frac{l^{1/2}}{\|w_l\|_2}Y_{i+j-1}\right|^3 = \frac{1}{l^{3/2}}O\left(\frac{1}{Q}\sum_{j=1}^{Q}\left|\frac{1}{l^{1/2}}\sum_{i=1}^{l} Y_{i+j-1}\right|^3\right) = O_P\left(\frac{1}{l^{3/2}}\right),$$

where we used the fact that $\frac{1}{Q}\sum_{j=1}^{Q}\left|\frac{1}{l^{1/2}}\sum_{i=1}^{l} Y_{i+j-1}\right|^3 = O_P(1)$ under the assumed conditions, see for instance the proof of Theorem 2 of Paparoditis and Romano (1992, p. 1994). It follows that $\frac{1}{Q}\sum_{j=1}^{Q}\left|\frac{1}{l}\sum_{i=1}^{l} w_l(i)\frac{l^{1/2}}{\|w_l\|_2}Y_{i+j-1}\right|^3 = O_P\left(\frac{1}{l^{3/2}}\right)$. Similarly by using the result of the proof of Theorem 3 of Paparoditis and Politis (2001), we have that

$$\frac{\|w_l\|_1\bar{Y}_{l,w}}{l^{1/2}\|w_l\|_2} = \frac{C}{l^{1/2}\|w_l\|_2} = O_p\left(\frac{l}{N}\right),$$

where we used the fact that $C \equiv \|w_l\|_1 \bar{Y}_{l,w} = O_p\left(\frac{l^2}{N}\right)$ and $\frac{1}{l^{1/2}\|w_l\|_2} = O_p(l)$. Then it follows that

$$\begin{aligned}\sum_{j=1}^{Q} E^*\left|z_j^*\right|^3 &= \frac{Q}{N}\left(\frac{l^3}{N}\right)^{1/2}\left(O_p\left(\frac{1}{l^{3/2}}\right) + O_p\left(\frac{l^3}{N^3}\right)\right) \\ &= \underbrace{\frac{Q}{N}}_{\to 1}\left(\underbrace{\frac{1}{N^{1/2}}}_{=o(1)} O_p(1) + \underbrace{\frac{1}{N^{1/2}}}_{=o(1)}\underbrace{\left(\frac{l^3}{N}\right)^{3/2}}_{=o(1)} O_p(1)\right) = o_p(1). \qquad (25)\end{aligned}$$

Thus $\sup_{x\in\mathbb{R}}\left|P^*\left(T_N^* \leq x\right) - \Phi\left(x/\sigma_\infty\right)\right| = o_p(1)$. Finally, our conclusion follows from the argument in the proof of Theorem 4.1 of Lahiri (2003). We omit the details here.

Proof of Theorem 2.2. The proof is similar to the proof of Theorem 2.1 in conjunction with Corollary 4.1 of Künsch (1989).

Proof of Theorem 3.1 part a). Here, we follow essentially Gonçalves and White (2002) in our proof. Recall that from part c) of Lemma 5.1 we have $\hat{\sigma}_{l,TBB}^2 = N\sigma_{jack}^2$, next using Theorem 3.1 of Künsch (1989), it follows that

$$\hat{\sigma}_{l,TBB}^2 = \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau}\left(X_{Nt} - \bar{X}_{l,w}\right)\left(X_{N,t+|\tau|} - \bar{X}_{l,w}\right). \qquad (26)$$

Given (26), the rest of the proof contains two steps. In (1) we show that $\tilde{\sigma}_N^2 - \sigma_N^2 \xrightarrow{P} 0$, and in (2) we show that $\hat{\sigma}_{l,TBB}^2 - \left(\tilde{\sigma}_N^2 + U_N\right) \xrightarrow{P} 0$, where $\tilde{\sigma}_N^2$ is an infeasible estimator which is identical to $\hat{\sigma}_{l,TBB}^2$ except it replaces $X_{Nt} - \bar{X}_{l,w}$ with $X_{Nt} - \mu_{Nt}$ in (26). In particular, we defined $\tilde{\sigma}_N^2$ as follows

$$\hat{\sigma}_{l,TBB}^2 = \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau}\left(X_{Nt} - \mu_{Nt}\right)\left(X_{N,t+|\tau|} - \mu_{N,t+|\tau|}\right). \qquad (27)$$

For step 1, we also have two steps.

i) We show that $\lim_{N\to\infty}\left|E\left(\tilde{\sigma}_N^2\right) - \sigma_N^2\right| = 0$.

ii) We show that $Var\left(\tilde{\sigma}_N^2\right) \to 0$.

Define $Z_{Nt} \equiv X_{Nt} - \mu_{Nt}$ and $R_{N,t}(\tau) = E\left(Z_{Nt}Z_{N,t+\tau}\right)$. Given the definitions of $\tilde{\sigma}_N^2$ and σ_N^2, we can write

$$E\left(\tilde{\sigma}_N^2\right) = \sum_{t=1}^{N} \beta_{N,t,0} R_{N,t}(0) + 2\sum_{\tau=1}^{l-1}\frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-\tau} \beta_{N,t,\tau} R_{N,t}(\tau), \text{ and}$$

$$\sigma_N^2 = \frac{1}{N}\sum_{t=1}^{N} R_{N,t}(0) + \frac{2}{N}\sum_{\tau=1}^{l-1}\sum_{t=1}^{N-\tau} R_{N,t}(\tau) + \frac{2}{N}\sum_{\tau=l}^{N-1}\sum_{t=1}^{N-\tau} R_{N,t}(\tau).$$

Then using the triangle inequality we have,

$$\begin{aligned}\left|E\left(\tilde{\sigma}_N^2\right)-\sigma_N^2\right| &\leq \sum_{t=1}^{N}\left|\beta_{N,t,0}-N^{-1}\right| R_{Nt}(0)+\sum_{t=1}^{N}\left|\beta_{N,t,0}-N^{-1}\right| R_{Nt}(0)\\ &\quad +2\sum_{\tau=1}^{l-1}\left|\frac{\upsilon_l(\tau)}{\upsilon_l(0)}\sum_{t=1}^{N-\tau}\beta_{N,t,\tau}-N^{-1}\right| R_{N,t}(\tau)+2\sum_{\tau=l}^{n-1}N^{-1}\sum_{t=1}^{n-\tau}\left|R_{N,t}(\tau)\right|\\ &= o(1),\end{aligned}$$

where we used the same argument like Gonçalves and White (2002) to bound the terms in their equation (A.3). Specifically it is due to the assume size conditions on α_k and υ_k and because, $|R_{N,t}(\tau)| \leq \Delta\left(5\alpha_{\left[\frac{\tau}{4}\right]}^{\left(\frac{1}{2}-\frac{1}{r}\right)}+\upsilon_{\left[\frac{\tau}{4}\right]}\right)$ (see Gallant and White, 1988, pp. 109-110).

To show that $Var\left(\tilde{\sigma}_N^2\right)\to 0$, define $\tilde{R}_{N,0}(\tau)=\sum_{t=1}^{N-|\tau|}\beta_{N,t,\tau}Z_{Nt}Z_{N,t+|\tau|}$, and write $Var\left(\tilde{\sigma}_N^2\right)=\sum_{\tau=-l+1}^{l-1}\sum_{\lambda=-l+1}^{l-1}\frac{\upsilon_l(\tau)\upsilon_l(\lambda)}{\upsilon_l^2(0)}Cov\left(\tilde{R}_{N,0}(\tau),\tilde{R}_{N,0}(\lambda)\right)$. We show that $Var\left(\tilde{R}_{N,0}(\tau)\right)=O\left(\frac{1}{N}\right)$, which by Cauchy-Schwarz inequality implies that $Var\left(\tilde{\sigma}_N^2\right)=O\left(\frac{l^2}{N}\right)$, since we have $\sum_{\tau=-l+1}^{l-1}\sum_{\lambda=-l+1}^{l-1}\frac{\upsilon_l(\tau)\upsilon_l(\lambda)}{\upsilon_l^2(0)}=l^2$. Note that we can write,

$$\begin{aligned}Var\left(\tilde{R}_{N,0}(\tau)\right) &= \sum_{t=1}^{N-|\tau|}\beta_{N,t,\tau}^2 Var\left(Z_{Nt}Z_{N,t+|\tau|}\right)\\ &\quad +2\sum_{t=1}^{N-|\tau|}\sum_{s=t+1}^{N-|\tau|}\beta_{N,t,\tau}\beta_{N,s,\tau}Cov\left(Z_{Nt}Z_{N,t+|\tau|},Z_{Ns}Z_{N,s+|\tau|}\right)\\ &\leq \frac{1}{Q^2}\sum_{t=1}^{N-|\tau|}Var\left(Z_{Nt}Z_{N,t+|\tau|}\right)+\frac{2}{Q^2}\sum_{t=1}^{N-|\tau|}\sum_{s=t+1}^{N-|\tau|}Cov\left(Z_{Nt}Z_{N,t+|\tau|},Z_{Ns}Z_{N,s+|\tau|}\right)\\ &\quad +\frac{2}{Q^2}\sum_{t=1}^{N-|\tau|}\sum_{s=t+|\tau|+1}^{N-|\tau|}Cov\left(Z_{Nt}Z_{N,t+|\tau|},Z_{Ns}Z_{N,s+|\tau|}\right)\end{aligned}$$

given that $\beta_{N,t,\tau}\leq\frac{1}{Q}$ for all t and τ.

$$\begin{aligned}Q^2Var\left(\tilde{R}_{N,0}(\tau)\right) &\leq KN\left\{\Delta^2+\sum_{k=1}^{\infty}\alpha_{\left[\frac{k}{4}\right]}^{\frac{1}{2}-\frac{1}{r}}+\sum_{k=1}^{\infty}\upsilon^{\left[\frac{k}{4}\right]}+\sum_{k=1}^{\infty}\upsilon_{\left[\frac{k}{4}\right]}^{\frac{r-2}{2(r-1)}}\right\}\\ &+ KN\left(|\tau|\,\alpha_{\left[\frac{|\tau|}{4}\right]}^{2\left(\frac{1}{2}-\frac{1}{r}\right)}+|\tau|\,\upsilon_{\left[\frac{|\tau|}{4}\right]}^2+2\,|\tau|\,\alpha_{\left[\frac{k}{4}\right]}^{\frac{1}{2}-\frac{1}{r}}\upsilon^{\left[\frac{|\tau|}{4}\right]}\right).\end{aligned}$$

Thus, using argument similar to that of Gonçalves and White (2002) to bound the terms in their equation (A.4), it follows that $Var\left(\tilde{R}_{N,0}(\tau)\right)\leq K\frac{N}{Q^2}$. Hence, $Var\left(\tilde{R}_{N,0}(\tau)\right)=O\left(\frac{1}{N}\right)$.

For step 2, define $S_{N,1}=\sum_{\tau=-l+1}^{l-1}\frac{\upsilon_l(\tau)}{\upsilon_l(0)}\sum_{t=1}^{N-|\tau|}\beta_{N,t,\tau}X_{Nt}X_{N,t+|\tau|}$, thus given (26) and (27), it follows that

$$\hat{\sigma}^2_{l,TBB} = S_{N,1} + \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \left(-\bar{X}_{l,w} X_{Nt} - \bar{X}_{l,w} X_{N,t+|\tau|} + \bar{X}^2_{l,w}\right), \text{ and}$$

$$\tilde{\sigma}^2_N = S_{N,1} + \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \left(-\mu_{N,t+|\tau|} X_{Nt} - \mu_{Nt} X_{N,t+|\tau|} + \mu_{Nt}\mu_{N,t+|\tau|}\right).$$

Then we have $\hat{\sigma}^2_{l,TBB} - \left(\tilde{\sigma}^2_N + U_N\right) = A_{N1} + A_{N2} + A_{N3} + A_{N4}$, where

$$\hat{\sigma}^2_{l,TBB} - \tilde{\sigma}^2_N = \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \begin{pmatrix} -\bar{X}_{l,w} Z_{Nt} - \bar{X}_{l,w}\mu_{Nt} - \bar{X}_{l,w} Z_{N,t+|\tau|} \\ -\bar{X}_{l,w}\mu_{N,t+|\tau|} + \mu_{N,t+|\tau|} Z_{Nt} \\ +\mu_{Nt} Z_{N,t+|\tau|} + \bar{X}^2_{l,w} + \mu_{N,t+|\tau|}\mu_{Nt} \end{pmatrix},$$

by adding and substracting appropriately, we can write

$$\hat{\sigma}^2_{l,TBB} - \tilde{\sigma}^2_N = A_{N1} + A_{N2} + A_{N3} + A_{N4},$$

where

$$A_{N1} = -\left(\bar{X}_{l,w} - \bar{\mu}_{l,w}\right) \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \left(Z_{Nt} + Z_{N,t+|\tau|}\right),$$

$$A_{N2} = \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \left(\mu_{Nt} - \bar{\mu}_{l,w}\right) Z_{N,t+|\tau|},$$

$$A_{N3} = \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \left(\mu_{N,t+|\tau|} - \bar{\mu}_{l,w}\right) Z_{N,t},$$

$$A_{N4} = \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \left(\bar{X}^2_{l,w} - \left(\mu_{Nt} + \mu_{N,t+|\tau|}\right)\bar{X}_{l,w} + \mu_{Nt}\mu_{N,t+|\tau|}\right),$$

with $\bar{\mu}_{l,w} = \sum_{t=1}^{N} a_N(t)\,\mu_{Nt}$. We have that

$$\begin{aligned} A_{N4} &= \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \begin{pmatrix} \left(\bar{X}_{l,w} - \bar{\mu}_{l,w}\right)^2 + 2\left(\bar{X}_{l,w} - \bar{\mu}_{l,w}\right)\bar{\mu}_{l,w} \\ -\left(\mu_{Nt} + \mu_{N,t+|\tau|}\right)\left(\bar{X}_{l,w} - \bar{\mu}_{l,w}\right) + \bar{\mu}^2_{l,w} \\ -\left(\mu_{Nt} + \mu_{N,t+|\tau|}\right)\bar{\mu}_{l,w} + \mu_{Nt}\mu_{N,t+|\tau|} \end{pmatrix} \\ &= U_N + \left(\bar{X}_{l,w} - \bar{\mu}_{l,w}\right)^2 \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \\ &\quad + \left(\bar{X}_{l,w} - \bar{\mu}_{l,w}\right) \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \left(2\bar{\mu}_{l,w} - \left(\mu_{Nt} + \mu_{N,t+|\tau|}\right)\right) \\ &= U_N + A'_{N4}, \end{aligned}$$

where $U_N = \sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau} \left(\mu_{Nt} - \bar{\mu}_{l,w}\right)\left(\mu_{N,t+|\tau|} - \bar{\mu}_{l,w}\right)$.

The rest of the proof follows closely that for the Theorm 2.1 of Gonçalves and White (2002), however for completeness, we present the relevent details. We now show that $\bar{X}_{l,w} - \bar{\mu}_{l,w} = o_P\left(l^{-1}\right)$. Define

$\phi_{Nt}(x) = \omega_{Nt}$, where $\omega_{Nt} \equiv \sum_{j=1}^{Q} \frac{w_l(t-j+1)}{\|w_l\|_1}$, and note that $\phi_{Nt}(\cdot)$ is uniformly Lipschitz continuous. Next, write $\bar{X}_{l,w} - \bar{\mu}_{l,w} = N^{-1}\sum_{t=1}^{N} Y_{Nt}$, where $Y_{Nt} = \phi_{Nt}(Z_{Nt})$ is a mean zero NED array on $\{V_t\}$ of the same size as Z_{Nt} by Theorem 17.12 of Davidson (1994), satisfying the same moment conditions. Hence, results follow by using the same argument as in Gonçalves and White (2002). In particular, by Lemma A.1 of Gonçalves and White (2002) $\{Y_{N,t}, \bar{\mathcal{F}}^t\}$ is a L_2-mixingale of size $-\frac{3r-2}{3(r-2)}$, and thus of size $-1/2$, with uniformly bounded constants, and by Lemma A.2 of Gonçalves and White (2002) $E\left(\max_{1\leq j\leq N}\left(\sum_{t=1}^{j} Y_{Nt}\right)^2\right) = O(N)$. By Chebyshev's inequality, for ϵ, $P\left[l\left(\bar{X}_{l,w} - \bar{\mu}_{l,w}\right) > 0\right] \leq \frac{l^2}{\epsilon^2 Q^2} E\left(\sum_{t=1}^{N} Y_{Nt}\right)^2 = O\left(\frac{l^2 N}{Q^2}\right) = o(1)$, if $l = o\left(N^{1/2}\right)$. This implies $A'_{N4} = o_P(1)$ ans similarly $A_{N1} = o_P(1)$, given that we have $\sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \sum_{t=1}^{N-|\tau|} \beta_{N,t,\tau}\left(Z_{Nt} + Z_{N,t+|\tau|}\right) = O_P(l)$.

To prove that $A_{N3} = o_P(1)$, define $\mathcal{Y}_{Nt\tau} = \omega_{Nt\tau}\left(\mu_{N,t+|\tau|} - \bar{\mu}_{l,w}\right) Z_{N,t} = \phi_{Nt\tau}(Z_{N,t})$, where $\omega_{Nt\tau} \equiv \frac{1}{\upsilon_l(\tau)} \sum_{j=1}^{Q} w_l(t-j+1)\, w_l(t-j+1+|\tau|)$ with $\tau < j$, and $\phi_{Nt\tau}(\cdot)$ is uniformly Lipschitz continous. Arguing as in Gonçalves and White (2002), $\{\mathcal{Y}_{Nt\tau}, \bar{\mathcal{F}}^t\}$ is a L_2-mixingale of size $-1/2$, with uniformly, with mixingale constants $c^{\mathcal{Y}}_{Nt\tau} \leq K \max\{\|w_l\|_{3r}, 1\}$ which are bounded uniformly in N, t, and τ. Thus,

$$
\begin{aligned}
P\left[\left|\sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \frac{1}{Q} \sum_{t=1}^{N-|\tau|} \mathcal{Y}_{Nt\tau}\right| \geq \epsilon\right] &\leq \frac{1}{Q\epsilon}\left[\sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} E\left|\sum_{t=1}^{N-|\tau|} \mathcal{Y}_{Nt\tau}\right|\right] \\
&\leq \frac{1}{Q\epsilon}\left[\sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} E\left(\left(\sum_{t=1}^{N-|\tau|} \mathcal{Y}_{Nt\tau}\right)^2\right)^{1/2}\right] \\
&\leq \frac{1}{Q\epsilon}\left[\sum_{\tau=-l+1}^{l-1} \frac{\upsilon_l(\tau)}{\upsilon_l(0)} \left(K \sum_{t=1}^{N-|\tau|} \left(c^{\mathcal{Y}}_{Nt\tau}\right)^2\right)^{1/2}\right] K\frac{lN^{1/2}}{Q} \\
&= o(1)
\end{aligned}
$$

where the first inequality holds by Markov's inequality, the second inequality holds by Jensen's inequality, the third inequality holds by Lemma A.2 of Gonçalves and White (2002) applied to $\{\mathcal{Y}_{Nt\tau}\}$ for each τ, and the last inequality holds by the uniform boundedness of $c^{\mathcal{Y}}_{Nt\tau}$. The proof of $A_{N2} = o_P(1)$ follows similarly.

Proof of Theorem 3.1 part b) Immediate from the proof of part a) of Theorem 3.1.

Proof of Theorem 3.1 part c) Immediate from the proof of part a) of Theorem 3.1.

Proof of Theorem 3.2 part a) The proof follows exactly the proof of Theorem 2.2 in Gonçalves and white (2002), and therefore we omit the details.

Proof of Theorem 3.2 part b) First note that the assumed conditions are sufficient to ensure that $\sqrt{N}\left(T_N - T(F)\right) \to^d N\left(0, \sigma^2_\infty\right)$ (see part (i) of Theorem 2.2 of Gonçalves and White (2002)). Therefore, to prove part b) of Theorem 3.2, we just need to show that the WTBB distribution is approximately close to $\Phi\left(x/\sigma_{\infty}\right)$. Note that, we can write

$$N^{1/2}\left(\bar{X}_N^* - E^*\left(\bar{X}_N^*\right)\right) = N^{1/2}\left(\bar{Z}_N^* - E^*\left(\bar{Z}_N^*\right)\right) = \sum_{j=1}^{Q} z_{Nj}^*,$$

where $Z_{Nt}^* \equiv X_{Nt}^* - \mu_{Nt}^*$, and $z_{Nj}^* = \frac{N^{1/2}}{Q}\left(\mathcal{Z}_{Nj}u_j - E^*\left(\mathcal{Z}_{Nj}u_j\right)\right)$, with $\mathcal{Z}_{Nj} \equiv \frac{Q}{N}\left(\sum_{i=1}^{l} \frac{w_l(i)}{\|w_l\|_2} Z_{N,i+j-1} - \bar{Z}_{l,w}\frac{\|w_l\|_1}{\|w_l\|_2}\right)$. Also note that $E^*\left(z_{Nj}^*\right) = 0$ and that

$$Var^*\left(\sum_{j=1}^{Q} z_{Nj}^*\right) = \frac{N}{Q}\hat{\sigma}_{l,TBB}^2 \xrightarrow{P} \sigma_{\infty}^2,$$

by part a2) of Corollary 4.1. Moreover, since $z_{N1}^*, \ldots, z_{NQ}^*$ are conditionally independent, by the Berry-Esseen bound, for some small $\delta > 0$ and for some constant $K > 0$ (which changes from line to line),

$$\sup_{x\in\mathbb{R}}\left|P^*\left(N^{1/2}\left(\bar{Z}_N^* - E^*\left(\bar{Z}_N^*\right)\right) \leq x\right) - \Phi\left(x/\sigma_{\infty}\right)\right| \leq K\sum_{j=1}^{Q} E^*\left|z_{Nj}^*\right|^{2+\delta},$$

which converges to zero in probability as $l \to \infty$, $N \to \infty$ such that $l = o\left(N^{1/2}\right)$. We have

$$\begin{aligned}
\sum_{j=1}^{Q} E^*\left|z_{Nj}^*\right|^{2+\delta} &= \sum_{j=1}^{Q} E^*\left|\frac{N^{1/2}}{Q}\left(\mathcal{Z}_{Nj}u_j - E^*\left(\mathcal{Z}_{Nj}u_j\right)\right)\right|^{2+\delta} \\
&\leq 2\frac{N^{1+\delta/2}}{Q^{2+\delta}}\sum_{j=1}^{Q} E^*\left|\mathcal{Z}_{Nj}u_j\right|^{2+\delta} \\
&= 2\frac{N^{1+\delta/2}}{Q^{2+\delta}}\sum_{j=1}^{Q} E^*\left|\mathcal{Z}_{Nj}\right|^{2+\delta} E^*\left|u_j\right|^{2+\delta} \\
&\leq K\frac{N^{1+\delta/2}}{Q^{2+\delta}}\sum_{j=1}^{Q} E^*\left|\mathcal{Z}_{Nj}\right|^{2+\delta}, \qquad (28)
\end{aligned}$$

where the first inequality follows from the C_r and the Jensen inequalities, whereas the second inequality uses the fact that by assumption $E^*\left|u_j\right|^{2+\delta} < \infty$. Next, note that

$$\begin{aligned}
E\left|\frac{N^{1+\delta/2}}{Q^{2+\delta}}\sum_{j=1}^{Q} E^*\left|\mathcal{Z}_{Nj}\right|^{2+\delta}\right| &\leq \frac{N^{1+\delta/2}}{Q^{2+\delta}}\sum_{j=1}^{Q} E\left|E^*\left|\mathcal{Z}_{Nj}\right|^{2+\delta}\right| \\
&= \frac{N^{1+\delta/2}}{Q^{2+\delta}}\left(\frac{Q}{N}\right)^{2+\delta}\frac{1}{\|w_l\|_2^{2+\delta}}\sum_{j=1}^{Q} E\left|\sum_{i=1}^{l} w_l(i)\, Z_{N,i+j-1} - \|w_l\|_1\bar{Z}_{l,w}\right|^{2+\delta} \\
&= \frac{N^{-(1+\delta/2)}}{\|w_l\|_2^{2+\delta}}\sum_{j=1}^{Q} E\left|\sum_{i=1}^{l} w_l(i)\, Z_{N,i+j-1} - \|w_l\|_1\bar{Z}_{l,w}\right|^{2+\delta} \\
&\leq \frac{N^{-(1+\delta/2)}}{\|w_l\|_2^{2+\delta}}\sum_{j=1}^{Q}\left(\left\|\sum_{i=1}^{l} w_l(i)\, Z_{N,i+j-1}\right\|_{2+\delta} + \left\|\|w_l\|_1\bar{Z}_{l,w}\right\|_{2+\delta}\right)^{2+\delta} \qquad (29)
\end{aligned}$$

where the first inequality follows from the triangle inequality, whereas the second inequality uses the Minkowski inequality. Under our assumptions,

$$\begin{aligned}\left\|\sum_{i=1}^{l} w_l(i)\, Z_{N,i+j-1}\right\|_{2+\delta} &\leq \underbrace{\max_{1\leq i\leq l} w_l(i)}_{\leq 1}\left\|\sum_{i=1}^{l} Z_{N,i+j-1}\right\|_{2+\delta} \\ &\leq \left\|\max_{1\leq t\leq l}\left|\sum_{i=j}^{j+t-1} Z_{N,i}\right|\right\|_{2+\delta} \leq K\left(\sum_{i=j}^{j+l-1} c_{Ni}^{\in}\right)^{1/2} \leq Kl^{1/2},\end{aligned}$$

by Lemmas A.3 and A.4 of Gonçalves and White (2002), given that c_{Ni} are uniformly bounded. Similarly, $\left\|\|w_l\|_1 \bar{Z}_{l,w}\right\|_{2+\delta} = O\left(l^{1/2}\right)$, which from (28) and (29) implies $\sum_{j=1}^{Q} E^*\left|z_{Nj}^*\right|^{2+\delta} = O\left(\frac{1}{N^{\delta/2}}\right) = o(1)$, since $l^{1/2}/\|w_l\|_2 = O(1)$.

References

[1] Andrews, D.W.K. (1984). "Non-strong mixing autoregressive processes," *Journal of Applied Probability* 21, 930-934.

[2] Brillinger, D. R. (1975). "Time Series, Data Analysis and Theory," *New York: Holt, Rinehart and Winston.*

[3] Bühlmann, P. and H. R. Künsch (1995). "The blockwise bootstrap for general parameters of a stationary time series," *Scandinavian Journal of Statistics*, 22(1), 35-54.

[4] Bustos, O. (1982). "General M-estimates for contaminated pth order autoregressive processes: consistency and asymptotic normality," *Z. Wahrsch. Verw. Gebiete* 59 491-504.

[5] Carlstein, E. (1986). "The Use of Subseries Values for Estimating the Variance of a General Statistic From a Stationary Time Series," *The Annals of Statistics*, 14, 1171-1179.

[6] Carlstein, E., Do, K.-A., Hall, P., Hesterberg, T., and Künsch, H. (1998). "Matched-Block Bootstrap for Dependent Data," *Bernoulli, 4, 305-328.*

[7] Davidson, J. (1994). "Stochastic limit theory," *Oxford: Oxford University Press.*

[8] Davidson, R., and E. Flachaire (2007). "Asymptotic and Bootstrap Inference for Inequality and Poverty Measures," *Journal of Econometrics*, 141, 141-166.

[9] Efron, B. (1979). "Bootstrap Methods: Another look at the jackknife," *The Annals of Statistics*, 7, 1-26.

[10] Gonçalves, S., and H. White (2002). "The bootstrap of the mean for dependent heterogeneous arrays," *Econometric Theory*, 18, 1367-1384.

[11] Götze, F. and H. Künsch (1996). "Second order correctness of the blockwise bootstrap for stationary observations," *Annals of Statistics* 24, 1914-33.

[12] Hall, P., (1992). "The bootstrap and Edgeworth expansion," *Springer-Verlag, New York.*

[13] Hall, P., and Mammen, E. (1994). "On General Resampling Algorithms and Their Performance in Distribution Estimation," *The Annals of Statistics,* 22, 2011-2030.

[14] Hall, P., Horowitz, J. L. and Jing, B.-Y. (1995). "On blocking rules for the bootstrap with dependent data," *Biometrika* 82, 561-574.

[15] Hounyo, U. , Gonçalves, S., and N. Meddahi (2013). "Bootstrapping pre-averaged realized volatility under market microstructure noise," Working Paper 28, CREATES, Aarhus University.

[16] Inoue, A. (2001). "Testing for Distributional Changes in Time Series," *Econometric Theory*, 17, 156-187.

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

[18] Lahiri, S. N. (1991). "Second order optimality of stationary bootstrap," *Statistics and Probability Letters* 11, 335-41.

[19] Lahiri, S. N. (1999). "Theoretical comparisons of block bootstrap methods, " *Annals of Statistics*, 27, 386-404.

[20] Lahiri, S. N. (2003). "Resampling Methods for Dependent Data," *New York: Springer.*

[21] Liu, R.Y. (1988). "Bootstrap procedure under some non-i.i.d. models," *Annals of Statistics* 16, 1696-1708.

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

[23] Mammen, E. (1993). "Bootstrap and wild bootstrap for high dimensional linear models," *Annals of Statistics*, 21, 255-285.

[24] Nordman, D. J. (2009). "A note on the stationary bootstrap," *Annals of Statistics*, 37, 359-370. 27, 386-404.

[25] Nordman, D. J. and Lahiri, S. N. (2012). "Block bootstrap methods for time series regression with fixed regressors," *Journal of the American Statistical Association*, 107, 233-246.

[26] Paparoditis, E., and Politis, D. N. (2001). "Tapered block bootstrap," *Biometrika*, 88, 1105-1119.

[27] Paparoditis, E., and Politis, D. N. (2002). "The tapered block bootstrap for general statistics from stationary sequences," *Econometrics Journal*, 5, 131-148.

[28] Politis, D. N. and Romano, J. P. (1992). "A general resampling scheme for triangular arrays of α-mixing random variables," *Annals of Statistics* 20, 1985-2007.

[29] Politis, D. N. and Romano, J. P. (1994). "The Stationary Bootstrap," *Journal of the American Statistical Association*, 89, 1303-1313.

[30] Politis, D. N. Romano, J. P. and M. Wolf (1997). "Subsampling for heteroscedastic time series," *Journal of Econometrics* 81, 281–317.

[31] Priestley, M. B. (1981). "Spectral Analysis and Time Series," *New York: Academic Press.*

[32] Rosenblatt, M. (1985). "Stationary Sequences and Random Fields," *Boston: Birkhäuser.*

[33] Shao, X. (2010a). "Extended tapered block bootstrap," *Statistica Sinica*, 20, 807-821.

[34] Shao, X. (2010b). "The dependent wild bootstrap," *Journal of the American Statistical Association*, 105, 218-235.

[35] Shao, X. (2011). "A bootstrap-assisted spectral test of white noise under unknown dependence," *Journal of Econometrics* 162, 213-224.

[36] Wu, C.F.J., (1986). "Jackknife, bootstrap and other resampling methods in regression analysis," *Annals of Statistics*, 14, 1261-1295.

Research Papers 2013

CREATES
Center for Research in Econometric Analysis of Time Series

2014-15: Federico Carlini and Katarzyna Łasak: On an Estimation Method for an Alternative Fractionally Cointegrated Model

2014-16: Mogens Bladt, Samuel Finch and Michael Sørensen: Simulation of multivariate diffusion bridges

2014-17: Markku Lanne and Henri Nyberg: Generalized Forecast Error Variance Decomposition for Linear and Nonlinear Multivariate Models

2014-18: Dragan Tevdovski: Extreme negative coexceedances in South Eastern European stock markets

2014-19: Niels Haldrup and Robinson Kruse: Discriminating between fractional integration and spurious long memory

2014-20: Martyna Marczak and Tommaso Proietti: Outlier Detection in Structural Time Series Models: the Indicator Saturation Approach

2014-21: Mikkel Bennedsen, Asger Lunde and Mikko S. Pakkanen: Discretization of Lévy semistationary processes with application to estimation

2014-22: Giuseppe Cavaliere, Morten Ørregaard Nielsen and A.M. Robert Taylor: Bootstrap Score Tests for Fractional Integration in Heteroskedastic ARFIMA Models, with an Application to Price Dynamics in Commodity Spot and Futures Markets

2014-23: Maggie E. C. Jones, Morten Ørregaard Nielsen and Michael Ksawery Popiel: A fractionally cointegrated VAR analysis of economic voting and political support

2014-24: Sepideh Dolatabadim, Morten Ørregaard Nielsen and Ke Xu: A fractionally cointegrated VAR analysis of price discovery in commodity futures markets

2014-25: Matias D. Cattaneo and Michael Jansson: Bootstrapping Kernel-Based Semiparametric Estimators

2014-26: Markku Lanne, Jani Luoto and Henri Nyberg: Is the Quantity Theory of Money Useful in Forecasting U.S. Inflation?

2014-27: Massimiliano Caporin, Eduardo Rossi and Paolo Santucci de Magistris: Volatility jumps and their economic determinants

2014-28: Tom Engsted: Fama on bubbles

2014-29: Massimiliano Caporin, Eduardo Rossi and Paolo Santucci de Magistris: Chasing volatility - A persistent multiplicative error model with jumps

2014-30: Michael Creel and Dennis Kristensen: ABC of SV: Limited Information Likelihood Inference in Stochastic Volatility Jump-Diffusion Models

2014-31: Peter Christoffersen, Asger Lunde and Kasper V. Olesen: Factor Structure in Commodity Futures Return and Volatility

2014-32: Ulrich Hounyo: The wild tapered block bootstrap