

2001s-19

The Bootstrap of the Mean for Dependent Heterogeneous Arrays

Sílvia Gonçalves, Halbert White

Série Scientifique
Scientific Series



CIRANO
Centre interuniversitaire de recherche
en analyse des organisations

Montréal
Mars 2001

CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du ministère de la Recherche, de la Science et de la Technologie, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère de la Recherche, de la Science et de la Technologie, and grants and research mandates obtained by its research teams.

Les organisations-partenaires / The Partner Organizations

- École des Hautes Études Commerciales
- École Polytechnique
- Université Concordia
- Université de Montréal
- Université du Québec à Montréal
- Université Laval
- Université McGill
- MEQ
- MRST
- Alcan inc.
- AXA Canada
- Banque du Canada
- Banque Laurentienne du Canada
- Banque Nationale du Canada
- Banque Royale du Canada
- Bell Québec
- Bombardier
- Bourse de Montréal
- Développement des ressources humaines Canada (DRHC)
- Fédération des caisses populaires Desjardins de Montréal et de l'Ouest-du-Québec
- Hydro-Québec
- Imasco
- Industrie Canada
- Pratt & Whitney Canada Inc.
- Raymond Chabot Grant Thornton
- Ville de Montréal

© 2001 Sílvia Gonçalves et Halbert White. Tous droits réservés. All rights reserved.

Reproduction partielle permise avec citation du document source, incluant la notice ©.

Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.

Ce document est publié dans l'intention de rendre accessibles les résultats préliminaires de la recherche effectuée au CIRANO, afin de susciter des échanges et des suggestions. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs, et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.

This paper presents preliminary research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.

The Bootstrap of the Mean for Dependent Heterogeneous Arrays*

Sílvia Gonçalves[†], Halbert White[‡]

Résumé / Abstract

Actuellement, les conditions assurant la validité des méthodes de bootstrap pour la moyenne d'échantillon des (possiblement hétérogènes) fonctions de dépendance d'époque proche (DEP) des processus de mixage sont inconnues. Ainsi, un des objectifs principaux de cet article est d'établir la validité du bootstrap dans ce contexte, en élargissant l'applicabilité des méthodes de bootstrap à une classe de processus largement adéquats pour les applications en économie et en finance. Les résultats se rapportent au bootstrap de blocs mouvants de Künsch (1989) et Liu et Singh (1992), de même qu'au bootstrap stationnaire de Politis et Romano (1994). Plus particulièrement, nous démontrons que la consistance de l'estimateur de variance du bootstrap pour la moyenne d'échantillon résiste à l'hétéroscédasticité et à la dépendance de forme inconnue. La validité asymptotique de premier ordre de l'approximation du bootstrap à la distribution actuelle de la moyenne d'échantillon est également démontrée dans ce contexte DEP hétérogène.

Presently, conditions ensuring the validity of bootstrap methods for the sample mean of (possibly heterogeneous) near epoch dependent (NED) functions of mixing processes are unknown. A main purpose of this paper is thus to establish the validity of the bootstrap in this context, extending the applicability of bootstrap methods to a class of processes broadly relevant for applications in economics and finance. The results apply to the moving blocks bootstrap of Künsch (1989) and Liu and Singh (1992) as well as to the stationary bootstrap of Politis and Romano (1994). In particular, the consistency of the bootstrap variance estimator for the sample mean is shown to be robust against heteroskedasticity and dependence of unknown form. The first order asymptotic validity of the bootstrap approximation to the actual distribution of the sample mean is also established in this heterogeneous NED context.

Mots Clés : Bootstrap en bloc, dépendance d'époque proche, moyenne d'échantillon

Keywords: Block bootstrap, near epoch dependence, sample mean

* Corresponding Author: Sílvia Gonçalves, CIRANO, 2020 University Street, 25th floor, Montréal, Qc, Canada H3A 2A5 Tel.: (514) 985-4011 Fax: (514) 985-4039 email: silvia.goncalves@umontreal.ca

[†] Université de Montréal and CIRANO

[‡] University of California, San Diego

1. Introduction

The bootstrap proposed by Efron (1979) provides an attractive method for generating statistics useful for inference and the construction of confidence intervals. It amounts to treating the empirical distribution of the data as if it were the population distribution when evaluating a statistic of interest. Given its straightforward applicability, even in situations where standard methods are intractable, bootstrap methods have become increasingly popular among applied econometricians. Under mild regularity conditions, the bootstrap approximation to the distribution of a statistic can be shown to be at least as accurate as the asymptotic normal approximation. Moreover, when asymptotically pivotal statistics are the basis for inference or confidence intervals, bootstrap methods can deliver results more accurate than standard asymptotics.

Bootstrap methods have been most intensively studied for the case of independent identically distributed (i.i.d.) observations (e.g., Bickel and Freedman (1981), Singh (1981)). However, the failure of the i.i.d. resampling scheme to give a consistent approximation to the true distribution of a statistic when observations are not independent (e.g. as remarked in Singh (1981)) has motivated several attempts in the literature to modify and extend Efron's idea to dependent data. Most of the extensions so far apply only to the stationary case. Bootstrap methods appropriate for stationary mixing processes have been proposed and studied by Künsch (1989) and Liu and Singh (1992) (the "moving blocks" bootstrap) and by Politis and Romano (1994) (the "stationary bootstrap"), among others. As it turns out, the moving blocks bootstrap is robust to heteroskedasticity. Lahiri (1992) gives conditions ensuring the second order correctness of Künsch's bootstrap for the normalized sample mean of observations that are not necessarily stationary. More recently, Fitzenberger (1997) has shown that the moving blocks method can be validly applied to heterogeneous mixing processes in the context of linear regressions and quantile regressions. Similarly, Politis et al. (1997) have shown the validity of certain subsampling methods for heterogeneous mixing processes.

For applications in economics, mixing is too strong a dependence condition to be broadly

applicable. Andrews (1984) gives an example of a simple AR(1) process that fails to be strong mixing. The need to accommodate such time series motivates the use of functions of mixing processes, the so-called processes near epoch dependent (NED) on an underlying mixing process (Billingsley, 1968; McLeish, 1975; Gallant and White, 1988). NED processes allow for considerable heterogeneity as well as dependence and include the mixing processes as a special case. An important example of the usefulness of near epoch dependence in economics concerns the standard ARCH (Engle, 1982) and GARCH (Bollerslev, 1986) processes widely used in economics and finance, for which the mixing properties are currently known only under certain restrictive assumptions (Carrasco and Chen, 1999). As Hansen (1991a) and Sin and White (1996) have shown, ARCH and GARCH processes are nevertheless, processes NED on an underlying mixing process, under mild regularity conditions. The NED concept thus makes possible a convenient theory of inference for these models that would otherwise be unavailable.

Presently, conditions ensuring the validity of bootstrap methods for the sample mean of (possibly heterogeneous) NED functions of mixing processes are unknown. A main purpose of this paper is thus to establish the validity of the bootstrap in this context, extending the applicability of bootstrap methods to a class of processes broadly relevant for applications in economics and finance. As is usual in the bootstrap literature, establishing the validity of the bootstrap for the sample mean is an important step towards establishing its validity for more complicated statistics. In Gonçalves and White (2000) we build on the results given here to prove the validity of the bootstrap for general extremum estimators such as quasi-maximum likelihood and generalized method of moments estimators.

Our results apply to the stationary bootstrap (SB) of Politis and Romano (1994), which has not yet been studied in a heterogeneous context, and to the moving blocks bootstrap (MBB) scheme of Künsch (1989) and Liu and Singh (1992), which has not been studied with the degree of dependence considered here. In particular, the consistency of the bootstrap variance estimator for the sample mean is shown to be robust against heteroskedasticity and dependence of unknown form. We also establish the first order asymptotic validity of the bootstrap approximation to the

actual distribution of the sample mean in this heterogeneous, near epoch dependent context.

The plan of the paper is as follows. The main theoretical results for the univariate mean are given in Section 2. Section 3 extends these results to the multivariate mean and to smooth functions of means. Section 4 presents a simulation study to examine the finite sample properties of the methods, and Section 5 concludes.

2. Main Results

Suppose $\{X_{nt}, n, t = 1, 2, \dots\}$ is a double array of not necessarily stationary (heterogeneous) random variables defined on a given probability space (Ω, \mathcal{F}, P) . By assuming that $\{X_{nt}\}$ is near epoch dependent on a mixing process, we permit a considerable degree of dependence and heterogeneity and include mixing processes as a special case. We define $\{X_{nt}\}$ to be NED on a mixing process $\{V_t\}$ provided $E(X_{nt}^2) < \infty$ and $v_k \equiv \sup_{n,t} \|X_{nt} - E_{t-k}^{t+k}(X_{nt})\|_2$ tends to zero at an appropriate rate. Here and in what follows, $\|X_{nt}\|_p \equiv (E|X_{nt}|^p)^{1/p}$ denotes the L_p norm and $E_{t-k}^{t+k}(\cdot) \equiv E(\cdot | \mathcal{F}_{t-k}^{t+k})$, where $\mathcal{F}_{t-k}^{t+k} \equiv \sigma(V_{t-k}, \dots, V_{t+k})$ is the σ -field generated by V_{t-k}, \dots, V_{t+k} . In particular, if $v_k = O(k^{-a-\delta})$ for some $\delta > 0$ we say $\{X_{nt}\}$ is NED of size $-a$. The sequence $\{V_t\}$ is assumed to be strong mixing although analogous results could be derived under the assumption of uniform mixing. We define the strong or α -mixing coefficients as usual, i.e. $\alpha_k \equiv \sup_m \sup_{\{A \in \mathcal{F}_{-\infty}^m, B \in \mathcal{F}_{m+k}^\infty\}} |P(A \cap B) - P(A)P(B)|$, and we require $\alpha_k \rightarrow 0$ as $k \rightarrow \infty$ at an appropriate rate.

Let $\mu_{nt} \equiv E(X_{nt})$ for $t = 1, 2, \dots, n$, $n = 1, 2, \dots$ and let $\bar{\mu}_n \equiv n^{-1} \sum_{t=1}^n \mu_{nt}$ be the average population mean. We regard $\bar{\mu}_n$ as the parameter of interest to be estimated by the sample mean $\bar{X}_n \equiv n^{-1} \sum_{t=1}^n X_{nt}$. Related studies such as Fitzenberger (1997) and Politis et al. (1997) have assumed a common mean, $\mu_{nt} = \mu$. If $\{X_{nt}\}$ is assumed to have a common mean μ , then $\bar{\mu}_n$ is just μ . Although we don't impose this restriction, we shall find it convenient to assume that the population means μ_{nt} satisfy a certain homogeneity condition in order to establish our main consistency result.

Our goal is to conduct inference on $\bar{\mu}_n$ based on a realization of $\{X_{nt}\}$. Alternatively, we

may be interested in constructing a confidence interval for $\bar{\mu}_n$ or in computing an estimate of the standard error of its estimator, the sample mean, $\bar{X}_n \equiv n^{-1} \sum_{t=1}^n X_{nt}$. The bootstrap can be used for these purposes.

We follow Lahiri (1999) in describing the block bootstrap methods. Let $\ell = \ell_n \in \mathbb{N}$ ($1 < \ell < n$) denote the (expected) length of the blocks and let $B_{t,\ell} = \{X_{nt}, X_{n,t+1}, \dots, X_{n,t+\ell-1}\}$ be the block of ℓ consecutive observations starting at X_{nt} ; $\ell = 1$ corresponds to the standard bootstrap. Assume for simplicity that $n = k\ell$.¹ The MBB resamples $k = n/\ell$ blocks randomly with replacement from the set of $n - \ell + 1$ overlapping blocks $\{B_{1,\ell}, \dots, B_{n-\ell+1,\ell}\}$. Thus, if we let I_{n1}, \dots, I_{nk} be i.i.d. random variables uniformly distributed on $\{1, \dots, n - \ell + 1\}$, the MBB pseudo-time series $\{X_{nt}^{*(1)}, t = 1, \dots, n\}$ is the result of arranging the elements of the k resampled blocks $B_{I_{n1},\ell}, \dots, B_{I_{nk},\ell}$ in a sequence: $X_{n1}^{*(1)} = X_{n,I_{n1}}, X_{n2}^{*(1)} = X_{n,I_{n1}+1}, \dots, X_{n\ell}^{*(1)} = X_{n,I_{n1}+\ell-1}, X_{n,\ell+1}^{*(1)} = X_{n,I_{n2}}, \dots, X_{n,kl}^{*(1)} = X_{n,I_{nk}+\ell-1}$. Here and throughout, we use the superscript (1) in $X_{n,t}^{*(1)}$ to denote the bootstrap samples obtained by the MBB. Similarly, we will use the superscript (2) to denote bootstrap samples obtained by the SB resampling scheme.

Unlike the MBB, the stationary bootstrap resamples blocks of random size. Let $p = \ell^{-1}$ be a given number in $(0, 1)$; $p = 1$ corresponds to the standard bootstrap. Let L_{n1}, L_{n2}, \dots be conditionally i.i.d. random variables having the geometric distribution with parameter p so that the probability of the event $\{L_{n1} = k\}$ is $(1-p)^{k-1}p$ for $k = 1, 2, \dots$. Independent of $\{X_{nt}\}$ and of $\{L_{nt}\}$, let I_{n1}, I_{n2}, \dots be i.i.d. random variables having the uniform distribution on $\{1, \dots, n\}$. The SB pseudo-time series $\{X_{nt}^{*(2)}\}$ can be obtained by joining the resampled blocks $B_{I_{n1},L_{n1}}, B_{I_{n2},L_{n2}}, \dots, B_{I_{nK},L_{nK}}$, where $K = \inf\{k \geq 1 : L_{n1} + \dots + L_{nk} \geq n\}$. Thus, the stationary bootstrap amounts to resampling blocks of observations of random length, where each block size has a geometric distribution with parameter p and expected length equal to $\frac{1}{p} = \ell$. We shall require $\ell = \ell_n$ to tend to infinity at an appropriate rate, which is equivalent to letting $p = p_n$ tend to zero. Hence, on average the lengths of the SB blocks tend to infinity with n as

¹If n is not divisible by ℓ , we can let k be the smallest integer satisfying $k\ell > n$ and delete the observations X_{nt}^* for $t > n$ (Politis and Romano, 1994); alternatively, we can let $k = \lfloor n/\ell \rfloor$ be the integer part of n/ℓ and let $n' = \lfloor n/\ell \rfloor \ell$ replace n , as in Andrews, 1999, and references therein.

happens with the (fixed) MBB blocks lengths.

As Politis and Romano (1994) remark, an alternative way to generate the SB resamples is as follows: let $X_{n1}^{*(2)}$ be picked at random from the original n observations, i.e. $X_{n1}^{*(2)} = X_{n,I_n}$, where I_n is an i.i.d. random variable with uniform distribution on $\{1, \dots, n\}$. With probability p , let $X_{n2}^{*(2)}$ be randomly drawn from the original n observations, while with probability $1 - p$, $X_{n2}^{*(2)}$ is the “next” observation in the original time series following X_{n,I_n} , i.e. $X_{n2}^{*(2)} = X_{n,I_n+1}$, and so on until n observations in the resample have been generated. If $t > n$, $X_{nt}^{*(2)}$ is defined to be X_{ns} , where $s = t \pmod{n}$.

In contrast to the MBB resampling method, the stationary bootstrap resample is a strictly stationary process (Politis and Romano, 1994), conditional on the original data.

Given the bootstrap resample $\{X_{n1}^{*(j)}, \dots, X_{nn}^{*(j)}\}$, $j = 1, 2$, one can compute the bootstrap version of the statistic of interest, $\bar{X}_n^{*(j)} \equiv n^{-1} \sum_{t=1}^n X_{nt}^{*(j)}$. For stationary α -mixing processes, Künsch (1989) and Politis and Romano (1994) show that their block bootstrap “works”. As a consequence, by repeating this procedure a large number B of times, one can approximate the true distribution of $\sqrt{n}(\bar{X}_n - \bar{\mu}_n)$ by the approximate sampling distribution of $\sqrt{n}(\bar{X}_n^{*(j)} - \bar{X}_n)$, conditional on the original data, given by the empirical distribution of the B draws of $\sqrt{n}(\bar{X}_n^{*(j)} - \bar{X}_n)$. Likewise, an estimate of the variance of the scaled sample mean $\sigma_n^2 \equiv \text{var}(\sqrt{n}\bar{X}_n)$ is easily obtained by using the bootstrap variance $\hat{\sigma}_{n,j}^2 = \text{var}^*\left(\sqrt{n}\bar{X}_n^{*(j)}\right)$. Here and in the following, the star denotes expectation with respect to $X_{n1}^{*(j)}, \dots, X_{nn}^{*(j)}$ conditional on the data X_{n1}, \dots, X_{nn} . The goal of this section is to extend these results to the heterogeneous NED case.

Under the assumption of heterogeneous dependent observations, σ_n^2 is given by

$$\sigma_n^2 = n^{-1} \sum_{t=1}^n E \left[(X_{nt} - \mu_{nt})^2 \right] + 2 \sum_{\tau=1}^{n-1} n^{-1} \sum_{t=1}^{n-\tau} E \left[(X_{nt} - \mu_{nt}) (X_{n,t+\tau} - \mu_{n,t+\tau}) \right].$$

Our first result establishes the consistency of the block bootstrap variance estimators for the sample mean, the MBB variance estimator $\hat{\sigma}_{n,1}^2$ and the SB variance estimator $\hat{\sigma}_{n,2}^2$, when the observations are near epoch dependent on a mixing process. As it turns out, neither of these bootstrap variance estimators require resampling the observations. Indeed, the following formula

for $\hat{\sigma}_{n,1}^2$ is available (Künsch, 1989, Theorems 3.1 and 3.4):

$$\hat{\sigma}_{n,1}^2 = \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \tilde{R}_n(\tau), \quad (2.1)$$

where

$$\tilde{R}_n(\tau) = \sum_{t=1}^{n-|\tau|} \beta_n(t, \tau) (X_{nt} - \bar{X}_{\alpha,n}) (X_{n,t+|\tau|} - \bar{X}_{\alpha,n}), \quad \text{and}$$

$$\bar{X}_{\alpha,n} = \sum_{t=1}^n \alpha_n(t) X_{nt}.$$

The weights $\alpha_n(t)$ and $\beta_n(t, \tau)$ are given as follows (cf. Künsch, 1989, expressions (3.2) and (3.7)):

$$\alpha_n(t) = \begin{cases} \frac{t}{\ell(n-\ell+1)}, & \text{if } t \in \{1, \dots, \ell-1\} \\ \frac{1}{n-\ell+1}, & \text{if } t \in \{\ell, \dots, n-\ell+1\} \\ \frac{n-t+1}{\ell(n-\ell+1)}, & \text{if } t \in \{n-\ell+2, \dots, n\}, \end{cases} \quad (2.2)$$

and

$$\beta_n(t, \tau) = \begin{cases} \frac{t}{(\ell-|\tau|)(n-\ell+1)}, & \text{if } t \in \{1, \dots, \ell-|\tau|-1\} \\ \frac{1}{n-\ell+1}, & \text{if } t \in \{\ell-|\tau|, \dots, n-\ell+1\} \\ \frac{n-t-|\tau|+1}{(\ell-|\tau|)(n-\ell+1)}, & \text{if } t \in \{n-\ell+2, \dots, n-|\tau|\}, \end{cases} \quad (2.3)$$

where $\sum_{t=1}^n \alpha_n(t) = 1$ and $\sum_{t=1}^{n-|\tau|} \beta_n(t, \tau) = 1$. Thus, as Künsch (1989) remarks, $\bar{X}_{\alpha,n}$ is asymptotically equivalent to the sample mean \bar{X}_n if $\ell = o(n)$, and $\tilde{R}_n(\tau)$ is an estimate of the autocovariance at lag τ similar to the usual sample covariance, but with a smaller bias.

Similarly, the SB variance estimator can be calculated with the following formula (Politis and Romano, 1994, Lemma 1):

$$\hat{\sigma}_{n,2}^2 = \hat{R}_n(0) + 2 \sum_{\tau=1}^{n-1} b_n(\tau) \hat{R}_n(\tau), \quad (2.4)$$

where

$$\hat{R}_n(\tau) = n^{-1} \sum_{t=1}^{n-\tau} (X_{nt} - \bar{X}_n) (X_{n,t+\tau} - \bar{X}_n)$$

is the usual autocovariance estimator and

$$b_n(\tau) = \left(1 - \frac{\tau}{n}\right) (1-p)^\tau + \frac{\tau}{n} (1-p)^{n-\tau}$$

is the Politis and Romano (1994) weight, with smoothing parameter $p \equiv \ell^{-1}$.

As is evident from (2.1) and (2.4), the MBB and the SB variance estimators for the sample mean are closely related to a lag window estimator of the spectral density at frequency zero. In particular, as remarked by Fitzenberger (1997) and by Politis and Romano (1994), the MBB variance estimator $\hat{\sigma}_{n,1}^2$ is approximately equivalent to the Bartlett kernel variance estimator considered by Newey and West (1987). Politis and Romano (1994) also discuss the relation between $\hat{\sigma}_{n,1}^2$ and $\hat{\sigma}_{n,2}^2$. They offer an interpretation for the SB variance estimator as a weighted average over ℓ of MBB variance estimators with fixed length ℓ , which suggests that $\hat{\sigma}_{n,2}^2$ should be less sensitive to the choice of p than $\hat{\sigma}_{n,1}^2$ is to the choice of ℓ . In a recent theoretical study, Lahiri (1999) compares several block bootstrap variance estimators, including the MBB and the SB. He concludes that while $\hat{\sigma}_{n,1}^2$ and $\hat{\sigma}_{n,2}^2$ have the same asymptotic bias, the variance of $\hat{\sigma}_{n,2}^2$ is larger than that of $\hat{\sigma}_{n,1}^2$, suggesting that the SB method is asymptotically less efficient than the MBB. In Section 4, we compare the finite sample properties of both methods via some Monte Carlo simulation experiments.

Assumption 2.1 is used to establish our main consistency theorem.

Assumption 2.1

2.1.a) For some $r > 2$, $\|X_{nt}\|_{3r} \leq \Delta < \infty$ for all $n, t = 1, 2, \dots$.

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

Theorem 2.1. Assume $\{X_{nt}\}$ satisfies Assumption 2.1. Then, if $\ell_n \rightarrow \infty$ and $\ell_n = o(n^{1/2})$, for $j = 1, 2$,

$$\hat{\sigma}_{n,j}^2 - (\sigma_n^2 + u_{n,j}^2) \xrightarrow{P} 0, \quad \text{where } u_{n,j}^2 \equiv \text{var}^* \left(n^{-1/2} \sum_{t=1}^n \mu_{nt}^{*(j)} \right).$$

The proof of this and of all our results can be found in the Appendix.

Under arbitrary heterogeneity in $\{X_{nt}\}$ the block bootstrap variance estimators $\hat{\sigma}_{n,j}^2$, $j = 1, 2$, are not consistent for σ_n^2 , but for $\sigma_n^2 + u_{n,j}^2$. The term $u_{n,j}^2$ is related to the heterogeneity in the means $\{\mu_{nt}\}$ and amounts to a block bootstrap variance of the scaled bootstrap population mean

$\sqrt{n}\bar{\mu}_n^*$. Theorem 2.1 makes clear that a necessary condition for the consistency of $\hat{\sigma}_{n,j}^2$ for σ_n^2 is that $u_{n,j}^2 \rightarrow 0$ as $n \rightarrow \infty$. A sufficient condition for $u_{n,j}^2$ to vanish is first order stationarity of $\{X_{nt}\}$: if $\mu_{nt} \equiv E(X_{nt}) = \mu$ for all n, t , then $u_{n,j}^2 = 0$. We have the following lemma.

Lemma 2.1. *If we let $\bar{\mu}_{\alpha,n} = \sum_{t=1}^n \alpha_n(t) \mu_{nt}$,*

$$u_{n,1}^2 = \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_n(t, \tau) (\mu_{nt} - \bar{\mu}_{\alpha,n}) (\mu_{n,t+|\tau|} - \bar{\mu}_{\alpha,n}), \text{ and}$$

$$u_{n,2}^2 = n^{-1} \sum_{t=1}^n (\mu_{nt} - \bar{\mu}_n)^2 + 2 \sum_{\tau=1}^{n-1} b_n(\tau) n^{-1} \sum_{t=1}^{n-\tau} (\mu_{nt} - \bar{\mu}_n) (\mu_{n,t+\tau} - \bar{\mu}_n).$$

In fact, the condition $\lim_{n \rightarrow \infty} u_{n,j}^2 = 0$, $j = 1, 2$, should be interpreted as an homogeneity condition on the means. The following assumption ensures $\lim_{n \rightarrow \infty} u_{n,j}^2 = 0$, $j = 1, 2$.

Assumption 2.2 $n^{-1} \sum_{t=1}^n (\mu_{nt} - \bar{\mu}_n)^2 = o(\ell_n^{-1})$, where $\ell_n \rightarrow \infty$ and $\ell_n = o(n)$.

The following consistency result holds under Assumptions 2.1 and 2.2 and is an immediate consequence of the previous remark.

Corollary 2.1. *Assume $\{X_{nt}\}$ satisfies Assumptions 2.1 and 2.2. Then, if $\ell_n \rightarrow \infty$ and $\ell_n = o(n^{1/2})$,*

$$\hat{\sigma}_{n,j}^2 - \sigma_n^2 \xrightarrow{P} 0, \quad j = 1, 2.$$

This result extends the previous consistency results by Künsch (1989) and Politis and Romano (1994) 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. Notice in particular that we do not assume asymptotic covariance stationarity.

In particular, for $j = 2$, Corollary 2.1 contains a version of Politis and Romano's (1994) Theorem 1 as a special case. Consider a strictly stationary α -mixing sequence $\{X_1, \dots, X_n\}$ satisfying Assumption 2.1. Because a mixing process is trivially near epoch dependent on itself, the NED requirement is automatically satisfied. Corollary 2.1 achieves the same conclusion as

Politis and Romano's (1994) Theorem 1 under the same moment conditions but weaker α -mixing size conditions ($\alpha_k = O(k^{-\lambda})$ for some $\lambda > \frac{2r}{r-2}$ and $r > 2$ here in contrast to $\alpha_k = O(k^{-\lambda})$ for some $\lambda > \frac{3(6+\varepsilon)}{\varepsilon}$ and $\varepsilon > 0$ there). We allow more dependence here, with the familiar trade-off between moment and memory conditions. Nevertheless, we require the stronger condition that $\ell_n = o(n^{1/2})$, i.e. $n^{1/2}p_n \rightarrow \infty$ (with $p_n = \ell_n^{-1} \rightarrow 0$), while Politis and Romano (1994) only require $\ell_n = o(n)$, i.e. $np_n \rightarrow \infty$. Imposing stationarity in our framework will ensure that $\sigma_n^2 \rightarrow \sigma_\infty^2$ as $n \rightarrow \infty$, where $\sigma_\infty^2 = \text{var}(X_1) + 2 \sum_{\tau=1}^{\infty} \text{var}(X_1, X_{1+\tau})$; hence, $\hat{\sigma}_{n,2}^2 \rightarrow \sigma_\infty^2$ in probability, as Politis and Romano (1994) conclude.

Similarly, for $j = 1$, our Corollary 2.1 specializes to Künsch's (1989) Corollary 3.1 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 $\ell_n = o(n^{1/2})$ instead of $\ell_n = o(n)$.

The next theorem establishes the first order asymptotic equivalence between the moving blocks and stationary bootstrap distributions and the normal distribution for the sample mean. A slightly stronger dependence assumption is imposed to achieve this result. Specifically, we require $\{X_{nt}\}$ to be $L_{2+\delta}$ -NED on a mixing process (see Andrews (1988)), for small $\delta > 0$, i.e. $v_k \equiv \sup_{n,t} \left\| X_{nt} - E_{t-k}^{t+k}(X_{nt}) \right\|_{2+\delta} \rightarrow 0$ as $n \rightarrow \infty$. We strengthen Assumption 2.1.b) slightly:

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

Theorem 2.2. Assume $\{X_{nt}\}$ satisfies Assumptions 2.1 and 2.2 strengthened by 2.1.b'). Then $\sigma_n^2 \leq K < \infty$ and if $\sigma_n^2 \geq \kappa > 0$ for all n sufficiently large,

(i) $\sup_{x \in \mathbb{R}} |P[\sigma_n^{-1} \sqrt{n}(\bar{X}_n - \bar{\mu}_n) \leq x] - \Phi(x)| \rightarrow 0$, where Φ is the standard normal cumulative distribution function.

(ii) Assume further that $\ell_n \rightarrow \infty$ and $\ell_n = o(n^{1/2})$. Then, for all $\varepsilon > 0$,

$$P \left\{ \sup_{x \in \mathbb{R}} \left| P^* \left[\sqrt{n} (\bar{X}_n^{*(j)} - \bar{X}_n) \leq x \right] - P \left[\sqrt{n} (\bar{X}_n - \bar{\mu}_n) \leq x \right] \right| > \varepsilon \right\} \rightarrow 0, \quad j = 1, 2,$$

where P^* is the probability measure induced by the bootstrap, conditional on $\{X_{nt}\}_{t=1}^n$.

For $j = 2$, this is an extension of Theorem 2 of Politis and Romano (1994) for stationary mixing observations to the case of NED functions of a mixing process. For $j = 1$, Theorem 2.2 states a weaker conclusion than does Theorem 3.5 of Künsch (1989), since we prove convergence in probability, but not almost sure convergence. On the other hand, we permit heterogeneity and greater dependence.

In part (i) we state the usual asymptotic normality result for the sample mean. We apply a central limit theorem for arrays NED on mixing processes given by Wooldridge and White (see e.g. Gallant and White, 1988, Theorem 5.3). In that result, it suffices that $\|X_{nt}\|_r \leq \Delta < \infty$ for some $r > 2$ and $\{X_{nt}\}$ is NED on $\{\varepsilon_t\}$ of size -1 , as is immediately guaranteed by the stronger conditions we impose to derive our bootstrap results.

In part (ii) we prove the uniform convergence to zero (in probability) of the discrepancy between the actual distribution of $\sqrt{n}(\bar{X}_n - \bar{\mu}_n)$ and the block bootstrap approximation to it. This result follows from the fact that under our assumptions the distribution of $\sigma_n^{-1}\sqrt{n}(\bar{X}_n^{*(j)} - \bar{X}_n)$, conditional on X_{n1}, \dots, X_{nn} , converges weakly to the standard normal distribution for all double arrays $\{X_{nt}\}$ in a set with probability tending to one. In particular, we apply a central limit theorem for triangular arrays and use Assumption 2.1.b') to ensure that Lyapounov's condition is satisfied under our heterogeneous NED context. Assumption 2.1.b) might well be sufficient to verify the weaker Lindeberg condition, as in Künsch (1989, Theorem 3.5), although we have not verified this.

Fitzenberger (1997) has recently proven the consistency of the moving blocks bootstrap approximation to the true sampling distribution of the sample mean for heterogeneous α -mixing processes. Our result extends his by allowing for near epoch dependence on mixing processes. However, in the purely strong mixing case, our moment and memory conditions are more stringent than his. In particular, his Theorem 3.1 only requires $E|X_t|^{2p+\delta} < C$, for small $\delta > 0$ and $p > 2$, and $\{X_t\}$ strong mixing of size $-\frac{p}{p-2}$. Politis et al. (1997, Theorem 3.1) have also established the robustness of the subsampling method for consistent sampling distribution estimation for heterogeneous and dependent data under mild moment conditions ($E|X_t|^{2+2\varepsilon} \leq \Delta < \infty$ for

some $\varepsilon > 0$). Nevertheless, they also assumed a strong mixing process, asymptotic covariance stationarity, and slightly stronger size requirements than ours on the mixing coefficients (α_k of size $-\frac{3(4+\varepsilon)}{\varepsilon}$).

A well known property of the MBB statistic $\sqrt{n} \left(\bar{X}_n^{*(1)} - \bar{X}_n \right)$ is that its (conditional) expected value is not zero. Indeed, under the MBB resampling scheme $E^* \left(\bar{X}_n^{*(1)} \right) = \sum_{t=1}^n \alpha_n(t) X_{nt}$, where the weights $\alpha_n(t)$ are defined as in (2.2). If $\ell_n = o(n)$, one can show that $E^* \left(\bar{X}_n^{*(1)} \right) = \bar{X}_n + O_P \left(\frac{\ell_n}{n} \right)$ (see e.g. Lemma A.1 of Fitzenberger, 1997). Thus, the MBB distribution has a random bias $\sqrt{n} \left(E^* \left(\bar{X}_n^{*(1)} \right) - \bar{X}_n \right)$, which is of order $O_P \left(\frac{\ell_n}{n^{1/2}} \right)$. (For the SB no such problem exists since $E^* \left(\bar{X}_n^{*(2)} \right) = \bar{X}_n$.) As pointed out by Lahiri (1992), this random bias becomes predominant for second order analysis and prevents the MBB from providing second order improvements over the standard normal approximation. To correct for this bias, he suggests recentering the MBB distribution around the bootstrap mean, that is, to consider the bootstrap distribution of $\sqrt{n} \left(\bar{X}_n^{*(1)} - E^* \left(\bar{X}_n^{*(1)} \right) \right)$. The following result shows that under the assumptions of Theorem 2.2 recentering the MBB distribution around the MBB bootstrap mean is asymptotically valid (to first order) in this heterogeneous NED context.

Corollary 2.2. *Under the assumptions of Theorem 2.2, for all $\varepsilon > 0$,*

$$P \left\{ \sup_{x \in \mathbb{R}} \left| P^* \left[\sqrt{n} \left(\bar{X}_n^{*(1)} - E^* \left(\bar{X}_n^{*(1)} \right) \right) \leq x \right] - P \left[\sqrt{n} \left(\bar{X}_n - \bar{\mu}_n \right) \leq x \right] \right| > \varepsilon \right\} \rightarrow 0,$$

where P^* is the probability measure induced by the MBB bootstrap, conditional on $\{X_{nt}\}_{t=1}^n$.

Theorem 2.2 and Corollary 2.2 justify the use of the MBB and SB distributions to obtain an asymptotically valid confidence interval for $\bar{\mu}_n$ instead of using a consistent estimate of the variance along with the normal approximation. For example, an equal tailed $(1 - \alpha)$ 100% stationary bootstrap confidence interval for $\bar{\mu}_n$ would be $[\bar{X}_n - q_n^* \left(1 - \frac{\alpha}{2} \right), \bar{X}_n - q_n^* \left(\frac{\alpha}{2} \right)]$, where $q_n^* \left(\frac{\alpha}{2} \right)$ and $q_n^* \left(1 - \frac{\alpha}{2} \right)$ are the $\frac{\alpha}{2}$ and $1 - \frac{\alpha}{2}$ quantiles of the SB bootstrap distribution of $\bar{X}_n^{*(2)} - \bar{X}_n$.

3. Extensions

Our main results can be readily extended to the multivariate mean and to smooth functions of sample means.

The first extension we study is the multivariate sample mean. Suppose each X_{nt} takes values in \mathbb{R}^d with i^{th} element given by X_{nti} . The parameter of interest is now the vector of average means $\bar{\mu}_n$, with typical element $\bar{\mu}_{ni} = n^{-1} \sum_{t=1}^n E(X_{nti})$ for $n = 1, 2, \dots$, and $i = 1, \dots, d$. By analogy with Assumption 2.1, the double array of random vectors $\{X_{nt}\}$ is assumed to be near epoch dependent on the strong mixing sequence $\{V_i\}$. For $j = 1, 2$, a multivariate pseudo-time series $\{X_{nt}^{*(j)}\}$ is generated by the MBB and the SB resampling schemes described before in exactly the same way.

We let $V_n \equiv \text{var}(\sqrt{n}\bar{X}_n)$ be the $d \times d$ covariance matrix with typical element $\sigma_{n,il}^2 \equiv \text{var}(\sqrt{n}\bar{X}_{ni}, \sqrt{n}\bar{X}_{nl})$. The MBB and SB covariance estimators denoted by $\hat{V}_{n,j} \equiv \text{var}^*(\sqrt{n}\bar{X}_n^{*(j)})$, for $j = 1, 2$, respectively, can be computed without resampling as before. In fact, following Künsch (1989, Theorems 3.1 and 3.4), the following formula for a typical element $\hat{\sigma}_{n,1,il}^2$ of $\hat{V}_{n,1}$ can be obtained:

$$\hat{\sigma}_{n,1,il}^2 = \tilde{R}_n^{il}(0) + \sum_{\tau=1}^{\ell-1} \left(1 - \frac{\tau}{\ell}\right) \left(\tilde{R}_n^{il}(\tau) + \tilde{R}_n^{li}(\tau)\right),$$

where

$$\tilde{R}_n^{il}(\tau) = \sum_{t=1}^{n-\tau} \beta_n(t, \tau) (X_{nti} - \bar{X}_{\alpha,n,i})(X_{n,t+\tau,l} - \bar{X}_{\alpha,n,l}),$$

and $\bar{X}_{\alpha,n,i} = \sum_{t=1}^n \alpha_n(t) X_{nti}$ for $i = 1, \dots, d$, and $\alpha_n(t)$ and $\beta_n(t, \tau)$ are given in (2.2) and (2.3), respectively. Similarly, for a typical element $\hat{\sigma}_{n,2,il}^2$ of $\hat{V}_{n,2}$ (cf. Lemma 1 in Politis and Romano (1994)), we have

$$\hat{\sigma}_{n,2,il}^2 = \hat{R}_n^{il}(0) + \sum_{\tau=1}^{n-1} b_n(\tau) \left(\hat{R}_n^{il}(\tau) + \hat{R}_n^{li}(\tau)\right),$$

where $\hat{R}_n^{il}(\tau)$ is the usual cross covariance estimator at lag τ , *i.e.* $\hat{R}_n^{il}(\tau) = n^{-1} \sum_{t=1}^{n-\tau} (X_{nti} - \bar{X}_{ni})(X_{n,t+\tau,l} - \bar{X}_{nl})$, $b_n(\tau)$ is the SB weight and $p = p_n \equiv \ell_n^{-1}$.

The following results are the direct counterparts of Theorems 2.1 and 2.2 for the multivariate

sample mean.

Theorem 3.1. *Let $\{X_{nt}\}$ be a double array of $d \times 1$ random vectors with typical element $\{X_{nti}\}$ satisfying Assumption 2.1. Let $V_n \equiv \text{var}(\sqrt{n}\bar{X}_n)$ and $\hat{V}_{n,j} \equiv \text{var}^*(\sqrt{n}\bar{X}_n^{*(j)})$, $j = 1, 2$. If $\ell_n \rightarrow \infty$ and $\ell_n = o(n^{1/2})$, then for $j = 1, 2$,*

$$\hat{V}_{n,j} - (V_n + U_{n,j}) \xrightarrow{P} 0, \quad \text{where } U_{n,j} \equiv \text{var}^*(\sqrt{n}\bar{\mu}_n^{*(j)}).$$

The bias term $U_{n,j}$ can be interpreted as the block bootstrap variance-covariance matrix of $\sqrt{n}\bar{\mu}_n^{*(j)}$ that would result if we could resample the vector time series $\{\mu_{nt}\}$. A multivariate extension of Lemma 2.1 gives the appropriate expressions for $U_{n,j}$ for $j = 1, 2$. If the elements of $\{X_{nt}\}$ satisfy Assumptions 2.1 and 2.2, then $U_{n,j} \rightarrow 0$ as $n \rightarrow \infty$ and we obtain the consistency result $\hat{V}_{n,j} - V_n \xrightarrow{P} 0$.

Theorem 3.2. *Let $\{X_{nt}\}$ be a double array of $d \times 1$ random vectors with typical element $\{X_{nti}\}$ satisfying Assumptions 2.1 and 2.2 strengthened by 2.1.b'). Let $V_n \equiv \text{var}(\sqrt{n}\bar{X}_n)$ be positive definite uniformly in n , i.e. V_n is positive semidefinite for all n and $\det V_n \geq \kappa > 0$ for all n sufficiently large. Then $V_n = O(1)$ and*

(i) $V_n^{-1/2}\sqrt{n}(\bar{X}_n - \bar{\mu}_n) \Rightarrow N(0, I_d)$ under P , and

$$\sup_{x \in \mathbb{R}^d} \left| P \left[\sqrt{n} V_n^{-1/2} (\bar{X}_n - \bar{\mu}_n) \leq x \right] - \Phi(x) \right| \rightarrow 0,$$

where Φ is the standard multivariate normal cumulative distribution function, “ \leq ” applies to each component of the relevant vector and “ \Rightarrow ” denotes convergence in distribution.

Moreover, if $\ell_n \rightarrow \infty$ and $\ell_n = o(n^{1/2})$, then for any $\varepsilon > 0$, and for $j = 1, 2$,

(ii) $V_n^{-1/2}\sqrt{n}(\bar{X}_n^{*(j)} - \bar{X}_n) \Rightarrow N(0, I_d)$ under P^* with probability P approaching one, and

$$P \left\{ \sup_{x \in \mathbb{R}^d} \left| P^* \left[\sqrt{n} (\bar{X}_n^{*(j)} - \bar{X}_n) \leq x \right] - P \left[\sqrt{n} (\bar{X}_n - \bar{\mu}_n) \leq x \right] \right| > \varepsilon \right\} \rightarrow 0,$$

where P^* is the probability measure induced by the bootstrap, conditional on $\{X_{nt}\}_{t=1}^n$.

For $j = 2$, Theorem 3.2 is an extension of Theorem 3 of Politis and Romano (1994) to the heterogeneous near epoch dependent case.

We now consider the validity of the bootstrap for smooth functions of sample means. We follow Horowitz (1999) in describing the “smooth function model”. Specifically, we let $f(\bar{Z}_{n1}, \dots, \bar{Z}_{np})$ denote the statistic of interest, where $f = (f_1, \dots, f_q)$ is a smooth function from \mathbb{R}^p to \mathbb{R}^q in the sense defined below, $\bar{Z}_{ni} = n^{-1} \sum_{t=1}^n h_i(X_{nt})$ for $i = 1, \dots, p$ and each h_i is a known nonstochastic function from \mathbb{R}^d to \mathbb{R} . The parameter of interest is $f(\mu_{Z_{n1}}, \dots, \mu_{Z_{np}})$, where $\bar{\mu}_{Z_{ni}} = n^{-1} \sum_{t=1}^n E[h_i(X_{nt})]$. We will assume that each real-valued function f_i is continuously differentiable in \mathbb{R}^p (or at least in a neighborhood of $\bar{\mu}_{Z_n} \equiv (\bar{\mu}_{Z_{n1}}, \dots, \bar{\mu}_{Z_{np}})'$) and that the gradient of f at $\bar{\mu}_{Z_n}$, the $p \times q$ matrix $\nabla f(\bar{\mu}_{Z_n})$, has full column rank uniformly in n (which implies that each f_i has a nonzero differential at $\bar{\mu}_{Z_n}$). If we further impose the moment and memory conditions on $\{h_i(X_{nt})\}$ as given by Assumption 2.1, the following theorem is true.

Theorem 3.3. *For $h : \mathbb{R}^d \rightarrow \mathbb{R}^p$, let $\{h(X_{nt})\}$ be a double array of $p \times 1$ random vectors with typical element $\{h_i(X_{nt})\}$ satisfying Assumptions 2.1 and 2.2 strengthened by 2.1.b'). For $i = 1, \dots, p$, let $\bar{Z}_{ni} = n^{-1} \sum_{t=1}^n h_i(X_{nt})$, $\bar{\mu}_{Z_{ni}} = n^{-1} \sum_{t=1}^n E[h_i(X_{nt})]$ and write $\bar{Z}_n = (\bar{Z}_{n1}, \dots, \bar{Z}_{np})'$, $\bar{\mu}_{Z_n} = (\bar{\mu}_{Z_{n1}}, \dots, \bar{\mu}_{Z_{np}})'$. If $\Omega_n = \text{var}(\sqrt{n}\bar{Z}_n)$ is positive definite uniformly in n , then $\Omega_n = O(1)$ and*

$$(i) \sup_{x \in \mathbb{R}^p} \left| P \left[\sqrt{n}\Omega_n^{-1/2} (\bar{Z}_n - \bar{\mu}_{Z_n}) \leq x \right] - \Phi(x) \right| \rightarrow 0.$$

If $\ell_n \rightarrow \infty$ and $\ell_n = o(n^{1/2})$, then for $j = 1, 2$,

$$(ii) \sup_{x \in \mathbb{R}^p} \left| P^* \left[\sqrt{n} (\bar{Z}_n^{*(j)} - \bar{Z}_n) \leq x \right] - P \left[\sqrt{n} (\bar{Z}_n - \bar{\mu}_{Z_n}) \leq x \right] \right| \xrightarrow{P} 0.$$

(iii) *If we assume further that $f : \mathbb{R}^p \rightarrow \mathbb{R}^q$ is continuously differentiable and that $\nabla f(\bar{\mu}_{Z_n})$ has full column rank uniformly in n , then for $j = 1, 2$,*

$$\sup_{x \in \mathbb{R}^p} \left| P^* \left[\sqrt{n} \left(f(\bar{Z}_n^{*(j)}) - f(\bar{Z}_n) \right) \leq x \right] - P \left[\sqrt{n} \left(f(\bar{Z}_n) - f(\bar{\mu}_{Z_n}) \right) \leq x \right] \right| \xrightarrow{P} 0.$$

For $j = 2$, Theorem 3.3 is an extension of Theorem 4 of Politis and Romano (1994) to the NED heterogeneous case.

Parts (i) and (ii) follow trivially by an application of Theorem 3.2 to the double array $\{h(X_{nt})\}$. To establish part (iii) we apply a mean value expansion to $f(\bar{Z}_n)$ around $\bar{\mu}_{Z_n}$ to show

that $\sqrt{n}\Gamma_n^{-1/2} (f(\bar{Z}_n) - f(\bar{\mu}_{Z_n})) \Rightarrow N(0, I_q)$, where $\Gamma_n = \nabla' f(\bar{\mu}_{Z_n}) \Omega_n \nabla f(\bar{\mu}_{Z_n})$. Similarly, we show that for $j = 1, 2$, $\sqrt{n}\Gamma_n^{-1/2} (f(\bar{Z}_n^{*(j)}) - f(\bar{Z}_n)) \Rightarrow N(0, I_q)$ under P^* with probability P approaching one, which delivers the result.

4. Simulation results

An immediate application of Theorem 3.3 is the construction of confidence intervals for parameters of linear models that are estimated by least squares or instrumental variables estimators since these estimators are smooth functions of sample means.

In this section we present the results of a simulation study aimed at assessing the finite sample properties of a variety of bootstrap methods for inference in the context of a multivariate linear regression. We consider the problem of building a confidence interval for a single regression parameter and use the finite sample coverage probability of symmetric 95% confidence intervals as our performance criterion. Our study is analogous to the simulation studies by Fitzenberger (1997) (who considers the moving blocks bootstrap) and Politis et al. (1997) (who consider the subsampling and the moving blocks bootstrap). Here we consider moving blocks bootstrap (MBB) and stationary bootstrap (SB) procedures. For comparison, we also consider two other standard kernel covariance estimators, the Bartlett kernel (BT) estimator and Andrews' (1991) Quadratic Spectral estimator (QS).

We follow the basic setup of Andrews (1991) and Andrews and Monahan (1992) who consider the linear regression model $y_t = x_t' \beta + \varepsilon_t$, where $x_t' = (1, \tilde{x}_t')$ contains five regressors, the first one being a constant. The data generating process is one of the processes proposed by Andrews (1991), namely AR(1)-HET 1, in which the errors follow some form of conditional heteroskedasticity. In particular, we let $x_{it} = \rho x_{i,t-1} + v_{it}$, $i = 2, \dots, 5$; $\tilde{\varepsilon}_t = \rho \tilde{\varepsilon}_{t-1} + u_t$ and $\varepsilon_t = |x_{2t}| \tilde{\varepsilon}_t$. Two different distributions are used to generate the innovations v_{it} and u_t : the standard normal distribution and the (centered) exponential distribution with parameter 1. We set the true parameter β equal to zero (without loss of generality) and consider the following values for the AR parameter ρ : $-0.5, 0.2, 0.5$ and 0.9 . In the simulations, 10,000 random samples are generated for the sample

size $n = 128$, which is one of the sample sizes considered in the previous literature.

The goal is to build a confidence interval for β_2 . The traditional approach relies on the asymptotic normal approximation and uses a consistent covariance estimator to compute the t-statistic. The kernel methods implement this approach. We use an automatic bandwidth selection procedure to compute the kernel covariances (see Andrews, 1991). For the bootstrap methods, one of several possible approaches is to form confidence intervals based directly on the bootstrap distribution. For example, a symmetric 95% bootstrap confidence interval is given by $\left[\hat{\beta}_{2n} - q_{n,0.95}^*, \hat{\beta}_{2n} + q_{n,0.95}^* \right]$, where $q_{n,0.95}^*$ is the 0.95 quantile of the bootstrap distribution of $\left| \hat{\beta}_{2n}^* - \hat{\beta}_{2n} \right|$, and where $\hat{\beta}_{2n}^*$ is the resampled OLS estimator of β_2 . This is only one of many ways to construct a confidence interval based on bootstrap quantiles. Following Politis (1998), we will call this approach the “root method” (cf. Politis, 1998). Because it does not rely on the normal approximation, this method might be expected to capture the potential skewness of the finite sample distribution of the estimator. To implement the “root method”, we use 500 bootstrap replications for each sample in the Monte Carlo experiment. For the moving blocks bootstrap we consider fixed block sizes that range from $\ell = 1$ to $\ell = 32$; for the stationary bootstrap, we consider similar average block sizes, i.e. we take $p = 1$ to $p = 1/32$.

The bootstrap is also used as a means to obtain a consistent estimator of the asymptotic variance of $\hat{\beta}_{2n}$. In particular, we use the bootstrap to consistently estimate $B_n = \text{var}(n^{-1/2} \sum_{t=1}^n x_t \varepsilon_t)$. To define the bootstrap variance estimator, we let $\{x_t \hat{\varepsilon}_t\}$ replace the unobservable process $\{x_t \varepsilon_t\}$, where $\{\hat{\varepsilon}_t\}$ are the OLS residuals. That is, for each $j = 1, 2$, we consider $\hat{B}_n^{(j)} = \text{var}^* \left(n^{-1/2} \sum_{t=1}^n x_t^* \hat{\varepsilon}_t^{*(j)} \right)$, where $\{x_t^* \hat{\varepsilon}_t^{*(j)}\}$ is a resample from $\{x_t \hat{\varepsilon}_t\}$. Following Andrews (1991), we introduce a small sample degrees of freedom adjustment factor equal to $n/(n - 5)$ in order to take into account the estimation of β when computing $\hat{\varepsilon}_t$; the kernel methods also use this adjustment factor. To compute the element (2, 2) of $\hat{B}_n^{(j)}$ corresponding to $\hat{\beta}_{2n}$ we use the closed form expressions given by (2.1) and (2.4), which amounts to considering an infinite number of bootstrap resamples.

Theorem 3.1 justifies the use of the bootstrap to provide a consistent estimator of the co-

variance matrix B_n . As it turns out, the feasible estimator $\hat{B}_n^{(j)}$ based on $\{x_t^* \varepsilon_t^{*(j)}\}$ is consistent under the same conditions of Theorem 3.1 (see Gonçalves and White, 2000, Theorem 3.1, for a more general result that applies to general extremum estimators). In particular, note that $\{x_t\}$ and $\{\tilde{\varepsilon}_t\}$ are AR(1) processes and thus they are mixing with geometrically decaying memory. This implies that the products $\{x_t \varepsilon_t\}$ are mixing with the same size and thus they are trivially NED on a mixing process. Thus, Assumption 2.1.b) is satisfied. The moment condition, Assumption 2.1.a), is also satisfied given that the innovations are generated as standard normal or as exponential random variables. Since $\{x_t\}$ and $\{\varepsilon_t\}$ are independent with mean zero, it follows trivially that $E(x_t \varepsilon_t) = 0$, so that Assumption 2.2 is also satisfied. An application of Theorem 3.3 justifies using the bootstrap to approximate the quantiles of the finite sampling distribution of $\hat{\beta}_{2n}$.

We now comment briefly on the results. Table 4.1 contains the results of our simulations with Gaussian innovations and Table 4.2 refers to the case where the innovations are exponential. (In the tables, ℓ is used to denote the fixed block size for the MBB; for the SB, ℓ corresponds to $1/p$ so that in this case ℓ denotes the average block size). Our results replicate the findings of previous studies for the kernel methods.

The first feature of note is that all methods consistently undercover. This feature tends to worsen with increasing dependence in the data. This is especially true for $\rho = 0.9$, where all methods have coverage probabilities well below the true nominal level. However, for sufficiently large block sizes, both the stationary bootstrap and the moving blocks bootstrap seem to better handle the increased dependence in the data, especially when the “root method” is used. In particular, for Gaussian errors and $\rho = 0.9$, the confidence interval based on the moving blocks bootstrap quantiles (MBB-root) with $\ell = 16$ achieves a coverage probability of 77.2% while the best kernel method (QS) only covers the true parameter 69.1% of the time. The best stationary bootstrap confidence interval for this scenario (SB-root, with $p = 1/16$) has a coverage probability of 73.8%, ranking below the moving blocks method.

The better performance of the moving blocks bootstrap in comparison to the other methods is

a general pattern in our simulation study. The stationary bootstrap performs similarly, although it ranks slightly below the moving blocks bootstrap. This is in line with recent work of Lahiri (1999), who shows that blocking bootstrap methods that use a random block size tend to have higher mean squared errors than methods that use a fixed block length.

Note also that using the bootstrap to approximate the quantiles of the true sampling distribution of the estimator (MBB-root and SB-root) instead of relying on the normal distribution along with a bootstrap variance estimator (MBB-var and SB-var) typically induces better coverage probabilities.

All methods perform worse under the exponential distribution as compared to the standard normal distribution. Nevertheless, the coverage probability shortfall is largest for the kernel methods and for the MBB-var and SB-var methods. By not relying on the asymptotic normal distribution, the “root method” is able to capture the potential skewness of the finite sample distribution of the estimator.

5. Conclusion

In this paper we establish the first order validity of block bootstrap methods for the sample mean of dependent heterogeneous data. Our results apply to the moving blocks bootstrap of Künsch (1989) and Liu and Singh (1992) as well as to the stationary bootstrap of Politis and Romano (1994). In particular, we show that the MBB and the SB 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. We also prove the first order asymptotic equivalence between the block bootstrap distributions and the normal distribution in this heterogeneous near epoch dependent context.

We perform a simulation study for the case of a multivariate linear regression. The block bootstrap methods perform fairly well compared to more standard kernel methods, and can deliver better inferences, especially under high serial correlation in the data.

Table 4.1: Empirical coverage probabilities with Gaussian Innovations
95% level confidence intervals, $n = 128$

Parameter	Method	$\ell = 1$	$\ell = 5$	$\ell = 6$	$\ell = 7$	$\ell = 8$	$\ell = 12$	$\ell = 16$	$\ell = 32$
$\rho = 0.2$	BT	92.2							
	QS	92.4							
	MBB-var	92.2	91.7	91.4	91.2	91.0	89.6	88.4	82.0
	MBB-root	91.8	91.4	91.2	91.0	90.9	90.2	89.5	85.4
	SB-var	92.2	91.0	90.6	90.3	89.9	88.4	86.9	81.2
	SB-root	91.8	90.5	90.3	90.1	89.7	88.4	86.7	81.9
$\rho = 0.5$	BT	88.8							
	QS	89.4							
	MBB-var	85.2	89.0	88.8	88.6	88.2	87.0	85.8	79.1
	MBB-root	84.5	88.9	88.8	88.7	88.5	88.5	87.9	83.9
	SB-var	85.2	88.3	88.0	87.6	87.2	85.7	84.0	78.5
	SB-root	84.5	87.9	88.0	87.8	87.5	86.3	84.7	79.9
$\rho = 0.9$	BT	67.0							
	QS	69.1							
	MBB-var	43.1	63.3	64.6	65.5	66.1	66.5	65.5	59.6
	MBB-root	42.7	67.1	69.4	71.3	72.5	76.3	77.2	76.8
	SB-var	43.1	63.9	64.5	64.9	65.1	64.4	63.1	58.4
	SB-root	42.7	67.8	69.4	70.5	71.5	73.7	73.8	71.1
$\rho = -0.5$	BT	89.1							
	QS	89.7							
	MBB-var	85.0	89.2	89.1	89.0	88.8	88.0	86.6	80.5
	MBB-root	84.5	88.9	89.0	88.8	88.8	88.6	88.1	84.5
	SB-var	85.0	88.6	88.5	88.2	87.8	86.4	84.9	79.5
	SB-root	84.5	88.2	88.1	87.9	87.7	86.4	84.9	80.4

Table 4.2: Empirical coverage probabilities with Exponential Innovations
95% level confidence intervals, $n = 128$

Parameter	Method	$\ell = 1$	$\ell = 5$	$\ell = 6$	$\ell = 7$	$\ell = 8$	$\ell = 12$	$\ell = 16$	$\ell = 32$
$\rho = 0.2$	BT	85.1							
	QS	85.1							
	MBB-var	85.2	84.3	84.0	83.7	83.4	82.3	81.0	74.6
	MBB-root	85.8	85.4	85.0	84.8	84.7	84.6	83.9	80.0
	SB-var	85.2	84.0	83.7	83.2	82.7	81.1	79.7	74.0
	SB-root	85.8	84.6	84.6	84.2	83.8	82.9	81.5	76.5
$\rho = 0.5$	BT	84.0							
	QS	84.4							
	MBB-var	80.9	83.9	83.6	83.4	83.3	82.1	80.8	73.9
	MBB-root	80.6	85.0	85.2	85.1	85.1	84.9	84.5	80.9
	SB-var	80.9	83.1	82.8	82.4	82.1	80.7	79.0	73.2
	SB-root	80.6	84.5	84.3	84.3	84.1	83.4	81.8	77.0
$\rho = 0.9$	BT	65.6							
	QS	67.5							
	MBB-var	42.5	62.0	63.3	64.3	64.9	65.4	64.4	58.5
	MBB-root	42.2	66.5	68.6	70.2	71.5	75.7	76.5	76.5
	SB-var	42.5	62.7	63.3	63.7	63.8	63.3	62.2	57.0
	SB-root	42.2	67.0	68.8	70.1	71.0	73.1	73.3	71.2
$\rho = -0.5$	BT	86.6							
	QS	87.1							
	MBB-var	83.0	86.5	86.3	86.2	85.9	84.8	83.3	76.5
	MBB-root	82.8	87.3	87.2	87.1	87.0	86.9	86.0	82.3
	SB-var	83.0	85.8	85.5	85.0	84.7	83.1	81.5	75.9
	SB-root	82.8	86.4	86.3	86.1	85.9	84.9	83.1	78.6

A. Appendix

Throughout the Appendix, K_i ($i = 0, 1, 2, \dots$) will denote generic constants that may change from one usage to another.

The following lemma generalizes Lemma 6.7 (a) in Gallant and White (1988, pp. 99-100). It will be used in the proof of our subsequent results.

Lemma A.1. *Assume $\{X_{nt}\}$ is a double array of random $p \times 1$ vectors, $p \in \mathbb{N}$, such that $E(X_{nt}) = 0$ and $\|X_{ntj}\|_{3r} \leq \Delta < \infty$ for some $r > 2$ and for all $n, t = 1, 2, \dots$, $j = 1, \dots, p$; assume further that $\{X_{nt}\}$ is NED on $\{\varepsilon_t\}$ with $v_k \equiv \sup_{n,t} \left\| X_{ntj} - E_{t-k}^{t+k}(X_{ntj}) \right\|_2$ for all $j = 1, \dots, p$, where $E_{t-k}^{t+k}(\cdot) = E(\cdot | \mathcal{F}_{t-k}^{t+k})$, $\mathcal{F}_{t-k}^{t+k} = \sigma(\varepsilon_{t-k}, \dots, \varepsilon_{t+k})$, and $\{\varepsilon_t\}$ is an α -mixing sequence with mixing coefficients denoted by α_k . Then, for fixed $\tau > 0$ and all $t < s \leq t + \tau$,*

$$|\text{cov}(X_{nti}X_{n,t+\tau,j}, X_{nsi}X_{n,s+\tau,j})| \leq K_1 \left(\alpha_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{1}{2} - \frac{1}{r}} + v_{\lfloor \frac{s-t}{4} \rfloor} \right) + K_2 v_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{r-2}{2(r-1)}} + K_3 \left(\alpha_{\lfloor \frac{\tau}{4} \rfloor}^{\left(\frac{1}{2} - \frac{1}{r}\right)} + v_{\lfloor \frac{\tau}{4} \rfloor} \right)^2,$$

for some finite constants K_1, K_2 and K_3 .

Proof. Let $t < s \leq t + \tau$, for any fixed $\tau > 0$, and notice

$$\begin{aligned} & |\text{cov}(X_{nti}X_{n,t+\tau,j}, X_{nsi}X_{n,s+\tau,j})| \\ & \leq |E(X_{nti}X_{n,t+\tau,j}X_{nsi}X_{n,s+\tau,j})| + |E(X_{nti}X_{n,t+\tau,j})| |E(X_{nsi}X_{n,s+\tau,j})|. \end{aligned} \quad (\text{A.1})$$

Under our assumptions, $|E(X_{nti}X_{n,t+\tau,j})| \leq \Delta \left(5\Delta \alpha_{\lfloor \frac{\tau}{4} \rfloor}^{1/2-1/r} + 2v_{\lfloor \frac{\tau}{4} \rfloor} \right)$ (see Gallant and White, 1988, pp. 109-110). Hence,

$$|E(X_{nti}X_{n,t+\tau,j})| |E(X_{nsi}X_{n,s+\tau,j})| \leq K_3 \left(\alpha_{\lfloor \frac{\tau}{4} \rfloor}^{1/2-1/r} + v_{\lfloor \frac{\tau}{4} \rfloor} \right)^2. \quad (\text{A.2})$$

To bound $|E(X_{nti}X_{n,t+\tau,j}X_{nsi}X_{n,s+\tau,j})|$, let $\hat{Y}_{nt\tau}^{ij} \equiv E_{t-\tau-\lfloor \frac{s-t}{2} \rfloor}^{t+\tau+\lfloor \frac{s-t}{2} \rfloor}(X_{nti}X_{nsi}X_{n,t+\tau,j})$ and write

$$\begin{aligned} |E(X_{nti}X_{n,t+\tau,j}X_{nsi}X_{n,s+\tau,j})| & \leq \left| E\left(\hat{Y}_{nt\tau}^{ij} X_{n,s+\tau,j}\right) \right| \\ & \quad + \left| E\left(X_{n,s+\tau,j} \left(X_{nti}X_{nsi}X_{n,t+\tau,j} - \hat{Y}_{nt\tau}^{ij}\right)\right) \right|. \end{aligned} \quad (\text{A.3})$$

Because $\hat{Y}_{nt\tau}^{ij}$ is measurable- $\mathcal{F}^{t+\tau+\lfloor \frac{s-t}{2} \rfloor}$, Hölder's inequality implies

$$\left| E\left(\hat{Y}_{nt\tau}^{ij} X_{n,s+\tau,j}\right) \right| \leq \|X_{nti}X_{nsi}X_{n,t+\tau,j}\|_2 \left\| E^{t+\tau+\lfloor \frac{s-t}{2} \rfloor}(X_{n,s+\tau,j}) \right\|_2 \leq \Delta^3 \left\| E^{t+\tau+\lfloor \frac{s-t}{2} \rfloor}(X_{n,s+\tau,j}) \right\|_2,$$

where we used Hölder's inequality twice in writing $\|X_{nti}X_{nsi}X_{n,t+\tau,j}\|_2 \leq \|X_{nti}\|_{3r} \|X_{nsi}\|_{3r} \|X_{n,t+\tau,j}\|_{3r} \leq \Delta^3 < \infty$. McLeish's inequality and the NED property can be used to bound $\left\| E^{t+\tau+\lfloor \frac{s-t}{2} \rfloor}(X_{n,s+\tau,j}) \right\|_2$

appropriately, delivering

$$\left| E \left(\hat{Y}_{nt\tau}^{ij} X_{n,s+\tau,j} \right) \right| \leq \Delta^3 \left(5\Delta \alpha_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{1}{2}-\frac{1}{r}} + v_{\lfloor \frac{s-t}{4} \rfloor} \right). \quad (\text{A.4})$$

To bound the second term in (A.3), by the Cauchy-Schwartz inequality,

$$\left| E \left(X_{n,s+\tau,j} \left(X_{nti} X_{nsi} X_{n,t+\tau,j} - \hat{Y}_{nt\tau}^{ij} \right) \right) \right| \leq \|X_{n,s+\tau,j}\|_2 \left\| X_{nti} X_{nsi} X_{n,t+\tau,j} - \hat{Y}_{nt\tau}^{ij} \right\|_2.$$

Theorem 4.2 of Gallant and White is used to prove that $\{X_{nti} X_{nsi} X_{n,t+\tau,j}\}$ is NED on $\{\varepsilon_t\}$ under our assumptions. In particular, by a reasoning similar to the proof of corollary 4.3.(b) of Gallant and White (1988, p. 61) we obtain

$$\left\| X_{nti} X_{nsi} X_{n,t+\tau,j} - E_{t-m}^{t+m} (X_{nti} X_{nsi} X_{n,t+\tau,j}) \right\|_2 \leq K v_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{r-2}{2(r-1)}}.$$

This implies

$$\left| E \left(X_{n,s+\tau,j} \left(X_{nti} X_{nsi} X_{n,t+\tau,j} - \hat{Y}_{nt\tau}^{ij} \right) \right) \right| \leq \Delta K v_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{r-2}{2(r-1)}}. \quad (\text{A.5})$$

Combining (A.4) and (A.5) with (A.3) yields

$$\left| E (X_{nti} X_{n,t+\tau,j} X_{nsi} X_{n,s+\tau,j}) \right| \leq K_1 \left(\alpha_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{1}{2}-\frac{1}{r}} + v_{\lfloor \frac{s-t}{4} \rfloor} \right) + K_2 v_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{r-2}{2(r-1)}}, \quad (\text{A.6})$$

which implies the desired result once we insert (A.6) and (A.2) into (A.1). ■

Proof of Theorem 2.1. The proof consists of two steps. In particular, for $j = 1, 2$, we prove that

$$\text{Step 1: } \tilde{\sigma}_{n,j}^2 - \sigma_n^2 \xrightarrow{P} 0.$$

$$\text{Step 2: } \hat{\sigma}_{n,j}^2 - \left(\tilde{\sigma}_{n,j}^2 + u_{n,j}^2 \right) \xrightarrow{P} 0.$$

In step 1, we consider an infeasible estimator $\tilde{\sigma}_{n,j}^2$ which is identical to $\hat{\sigma}_{n,j}^2$ except that it replaces $\bar{X}_{\alpha,n}$ and \bar{X}_n in (2.1) and (2.4) for $j = 1, 2$, respectively, with the population means μ_{nt} . So, e.g. for $j = 1$ we consider

$$\tilde{\sigma}_{n,1}^2 = \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell} \right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} (X_{nt} - \mu_{nt}) \left(X_{n,t+|\tau|} - \mu_{n,t+|\tau|} \right),$$

where we let $\beta_{n,t,\tau} \equiv \beta_n(t, \tau)$ throughout. To prove step 1, we show that the bias and the variance of $\tilde{\sigma}_{n,j}^2$ tend to zero as $n \rightarrow \infty$.

Proof of step 1. ($j = 1$): letting $Z_{nt} \equiv X_{nt} - \mu_{nt}$ and $R_{n,t}(\tau) = E(Z_{nt} Z_{n,t+|\tau|})$, it follows that

$$E(\tilde{\sigma}_{n,1}^2) = \sum_{t=1}^n \beta_{n,t,0} R_{n,t}(0) + 2 \sum_{\tau=1}^{\ell-1} \left(1 - \frac{\tau}{\ell} \right) \sum_{t=1}^{n-\tau} \beta_{n,t,\tau} R_{n,t}(\tau).$$

Because $\sigma_n^2 = n^{-1} \sum_{t=1}^n R_{n,t}(0) + 2 \sum_{\tau=1}^{n-1} n^{-1} \sum_{t=1}^{n-|\tau|} R_{n,t}(\tau)$, it follows that

$$\begin{aligned} |E(\tilde{\sigma}_{n,1}^2) - \sigma_n^2| &\leq 2 \sum_{\tau=1}^{\ell-1} \frac{\tau}{\ell} \sum_{t=1}^{n-\tau} |\beta_{n,t,\tau}| |R_{n,t}(\tau)| + 2 \sum_{\ell \leq \tau \leq n-1} n^{-1} \sum_{t=1}^{n-|\tau|} |R_{n,t}(\tau)| \\ &\quad + \sum_{t=1}^n |\beta_{n,t,0} - n^{-1}| |R_{n,t}(0)| + 2 \sum_{\tau=1}^{\ell-1} \sum_{t=1}^{n-|\tau|} |\beta_{n,t,\tau} - n^{-1}| |R_{n,t}(\tau)| \\ &\leq \frac{2n}{n-\ell+1} \sum_{\tau=1}^{n-1} \frac{\tau}{\ell} n^{-1} \sum_{t=1}^{n-|\tau|} |R_{n,t}(\tau)| + O\left(\frac{\ell}{n-\ell+1}\right) + O\left(\frac{\ell^2}{n-\ell+1}\right), \end{aligned}$$

given that $|\beta_{n,t,\tau}| \leq \frac{1}{n-\ell+1}$ for all n, t and where the order of magnitudes of the two last terms in the first inequality were obtained using the expression for $\beta_{n,t,\tau}$ (cf. (2.3)). Given that $\ell = o(n^{1/2})$, these terms are negligible asymptotically. Under the same assumption on ℓ , $\frac{n}{n-\ell+1} \rightarrow 1$ and therefore we only need to show that the limit of $\xi_n \equiv \sum_{\tau=1}^{n-1} \frac{\tau}{\ell} n^{-1} \sum_{t=1}^{n-|\tau|} |R_{n,t}(\tau)|$ is zero to prove that the bias of $\tilde{\sigma}_{n,1}^2$ is zero.

An argument similar to that in Gallant and White (1988, pp.109-110) delivers $|R_{n,t}(\tau)| \leq K \left(\alpha_{[\frac{\tau}{4}]}^{1-\frac{1}{p}-\frac{1}{r}} + v_{[\frac{\tau}{4}]} \right)$, for any $p \geq 2$ and $r > 2$ (Gallant and White's result uses $p = 2$). Thus, for fixed τ , $n^{-1} \sum_{t=1}^{n-\tau} |R_{n,t}(\tau)| \leq K \left(\alpha_{[\frac{\tau}{4}]}^{1-\frac{1}{p}-\frac{1}{r}} + v_{[\frac{\tau}{4}]} \right)$ and $\xi_n \leq K \sum_{\tau=1}^{n-1} \frac{\tau}{\ell} \left(\alpha_{[\frac{\tau}{4}]}^{1-\frac{1}{p}-\frac{1}{r}} + v_{[\frac{\tau}{4}]} \right)$. By taking limits on both sides of this inequality, an application of the Lebesgue dominated convergence theorem as in Gallant and White (1988, p. 111) is used to show that $\lim_{n \rightarrow \infty} \xi_n = 0$. In particular, we note that for each τ , $\frac{\tau}{\ell} \rightarrow 0$ as $n \rightarrow \infty$ given that $\ell_n \rightarrow \infty$. To find a dominating function, we note that the size conditions on α_k imposed by Assumption 2.1 are sufficient to obtain that $\sum_{\tau=1}^{\infty} \tau \left(\alpha_{[\frac{\tau}{4}]}^{1-\frac{1}{p}-\frac{1}{r}} + v_{[\frac{\tau}{4}]} \right) < \infty$ if we let $p = r > 2$. Thus, $\lim_{n \rightarrow \infty} |E(\tilde{\sigma}_{n,1}^2) - \sigma_n^2| = 0$.

Next, we show that $\text{var}(\tilde{\sigma}_{n,1}^2) \rightarrow 0$. Letting $\tilde{R}_{n,0}(\tau) = \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} Z_{nt} Z_{n,t+|\tau|}$, we have that

$$\text{var}(\tilde{\sigma}_{n,1}^2) = \sum_{\tau=-\ell+1}^{\ell-1} \sum_{\lambda=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \left(1 - \frac{|\lambda|}{\ell}\right) \text{cov}\left(\tilde{R}_{n,0}(\tau), \tilde{R}_{n,0}(\lambda)\right). \quad (\text{A.7})$$

By the Cauchy-Schwartz inequality, it suffices to bound $\text{var}\left(\tilde{R}_{n,0}(\tau)\right)$ for any fixed τ :

$$\begin{aligned} \text{var}\left(\tilde{R}_{n,0}(\tau)\right) &= \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau}^2 \text{var}\left(Z_{nt} Z_{n,t+|\tau|}\right) + 2 \sum_{t=1}^{n-|\tau|} \sum_{s=t+1}^{n-|\tau|} \beta_{n,t,\tau} \beta_{n,s,\tau} \left| \text{cov}\left(Z_{nt} Z_{n,t+|\tau|}, Z_{ns} Z_{n,s+|\tau|}\right) \right| \\ &\leq \frac{1}{(n-\ell+1)^2} \sum_{t=1}^{n-|\tau|} \text{var}\left(Z_{nt} Z_{n,t+|\tau|}\right) + \frac{2}{(n-\ell+1)^2} \sum_{t=1}^{n-|\tau|} \sum_{s=t+1}^{n-|\tau|} \left| \text{cov}\left(Z_{nt} Z_{n,t+|\tau|}, Z_{ns} Z_{n,s+|\tau|}\right) \right| \\ &\quad + \frac{2}{(n-\ell+1)^2} \sum_{t=1}^{n-|\tau|} \sum_{s=t+|\tau|+1}^{n-|\tau|} \left| \text{cov}\left(Z_{nt} Z_{n,t+|\tau|}, Z_{ns} Z_{n,s+|\tau|}\right) \right|, \end{aligned}$$

given that $\beta_{n,t,\tau} \leq \frac{1}{n-\ell+1}$ for all t and τ . Given Assumption 2.1 and noticing that Z_{nt} has mean zero by construction, by Lemma A.1 it follows that

$$\begin{aligned} (n-\ell+1)^2 \text{var} \left(\tilde{R}_{n,0}(\tau) \right) &\leq \sum_{t=1}^{n-|\tau|} 2\Delta^2 + K_1 \sum_{t=1}^{n-|\tau|} \sum_{s=t+1}^{t+|\tau|} \left(\alpha_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{1}{2}-\frac{1}{r}} + v_{\lfloor \frac{s-t}{4} \rfloor} \right) + K_2 \sum_{t=1}^{n-|\tau|} \sum_{s=t+1}^{t+|\tau|} v_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{r-2}{2(r-1)}} \\ &\quad + K_3 \sum_{t=1}^{n-|\tau|} \sum_{s=t+1}^{t+|\tau|} \left(\alpha_{\lfloor \frac{|\tau|}{4} \rfloor}^{\left(\frac{1}{2}-\frac{1}{r}\right)} + v_{\lfloor \frac{|\tau|}{4} \rfloor} \right)^2 + K_4 \sum_{t=1}^{n-|\tau|} \sum_{s=t+|\tau|+1}^{n-|\tau|} \left(\alpha_{\lfloor \frac{s-t-|\tau|}{4} \rfloor}^{\frac{1}{2}-\frac{1}{r}} + v_{\lfloor \frac{s-t-|\tau|}{4} \rfloor}^{\frac{r-2}{2(r-1)}} \right). \end{aligned}$$

Notice that Lemma 6.7 (a) in Gallant and White (1988, pp.99-100) is used above to bound the last term when $s > t + |\tau|$.

It follows that

- $\sum_{t=1}^{n-|\tau|} \sum_{s=t+1}^{t+|\tau|} \left(\alpha_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{1}{2}-\frac{1}{r}} + v_{\lfloor \frac{s-t}{4} \rfloor} \right) \leq (n-|\tau|) \sum_{k=1}^{\infty} \left(\alpha_{\lfloor \frac{k}{4} \rfloor}^{\frac{1}{2}-\frac{1}{r}} + v_{\lfloor \frac{k}{4} \rfloor} \right);$
- $\sum_{t=1}^{n-|\tau|} \sum_{s=t+1}^{t+|\tau|} v_{\lfloor \frac{s-t}{4} \rfloor}^{\frac{r-2}{2(r-1)}} \leq (n-|\tau|) \sum_{k=1}^{\infty} v_{\lfloor \frac{k}{4} \rfloor}^{\frac{r-2}{2(r-1)}};$
- $\sum_{t=1}^{n-|\tau|} \sum_{s=t+|\tau|+1}^{n-|\tau|} \left(\alpha_{\lfloor \frac{s-t-|\tau|}{4} \rfloor}^{\frac{1}{2}-\frac{1}{r}} + v_{\lfloor \frac{s-t-|\tau|}{4} \rfloor}^{\frac{r-2}{2(r-1)}} \right) \leq (n-|\tau|) \sum_{k=1}^{\infty} \left(\alpha_{\lfloor \frac{k}{4} \rfloor}^{\frac{1}{2}-\frac{1}{r}} + v_{\lfloor \frac{k}{4} \rfloor}^{\frac{r-2}{2(r-1)}} \right);$
- $\sum_{t=1}^{n-|\tau|} \sum_{s=t+1}^{t+|\tau|} \left(\alpha_{\lfloor \frac{|\tau|}{4} \rfloor}^{\left(\frac{1}{2}-\frac{1}{r}\right)} + v_{\lfloor \frac{|\tau|}{4} \rfloor} \right)^2 = (n-|\tau|) |\tau| \left(\alpha_{\lfloor \frac{|\tau|}{4} \rfloor}^{\left(\frac{1}{2}-\frac{1}{r}\right)} + v_{\lfloor \frac{|\tau|}{4} \rfloor} \right)^2$

$$= (n-|\tau|) |\tau| \left(\alpha_{\lfloor \frac{|\tau|}{4} \rfloor}^{2\left(\frac{1}{2}-\frac{1}{r}\right)} + v_{\lfloor \frac{|\tau|}{4} \rfloor}^2 + 2\alpha_{\lfloor \frac{|\tau|}{4} \rfloor}^{\left(\frac{1}{2}-\frac{1}{r}\right)} v_{\lfloor \frac{|\tau|}{4} \rfloor} \right);$$

For K sufficiently large,

$$\begin{aligned} (n-\ell+1)^2 \text{var} \left(\tilde{R}_{n,0}(\tau) \right) &\leq Kn \left\{ \Delta^2 + \sum_{k=1}^{\infty} \alpha_{\lfloor \frac{k}{4} \rfloor}^{\frac{1}{2}-\frac{1}{r}} + \sum_{k=1}^{\infty} v_{\lfloor \frac{k}{4} \rfloor} + \sum_{k=1}^{\infty} v_{\lfloor \frac{k}{4} \rfloor}^{\frac{r-2}{2(r-1)}} \right\} \\ &\quad + Kn \left(|\tau| \alpha_{\lfloor \frac{|\tau|}{4} \rfloor}^{2\left(\frac{1}{2}-\frac{1}{r}\right)} + |\tau| v_{\lfloor \frac{|\tau|}{4} \rfloor}^2 + 2|\tau| \alpha_{\lfloor \frac{|\tau|}{4} \rfloor}^{\left(\frac{1}{2}-\frac{1}{r}\right)} v_{\lfloor \frac{|\tau|}{4} \rfloor} \right). \end{aligned}$$

Given the size conditions on α_k and v_k , the sums in the curly brackets are finite. Therefore, we can write

$$\begin{aligned} \text{var} \left(\tilde{R}_{n,0}(\tau) \right) &\leq \frac{Kn}{(n-\ell+1)^2} B \\ &\quad + \frac{Kn}{(n-\ell+1)^2} \left(|\tau| \alpha_{\lfloor \frac{|\tau|}{4} \rfloor}^{2\left(\frac{1}{2}-\frac{1}{r}\right)} + |\tau| v_{\lfloor \frac{|\tau|}{4} \rfloor}^2 + 2|\tau| \alpha_{\lfloor \frac{|\tau|}{4} \rfloor}^{\left(\frac{1}{2}-\frac{1}{r}\right)} v_{\lfloor \frac{|\tau|}{4} \rfloor} \right), \quad (\text{A.8}) \end{aligned}$$

for some finite constant B . But the size conditions on α_k and on v_k , their monotonicity and the fact that they converge to 0 as $k \rightarrow \infty$ implies the last term in (A.8) tends to 0 as $|\tau| \rightarrow \infty$. Hence,

there exists some finite constant K independent of n and τ such that $\text{var}\left(\tilde{R}_{n,0}(\tau)\right) \leq \frac{Kn}{(n-\ell+1)^2}$ for every $|\tau| \geq 0$. By the Cauchy-Schwartz inequality we have that $\text{cov}\left(\tilde{R}_{n,0}(\tau), \tilde{R}_{n,0}(\lambda)\right) \leq \frac{Kn}{(n-\ell+1)^2}$ for every τ and λ . Given (A.7) and given that $\ell_n = o(n^{1/2})$ and $\ell_n \rightarrow \infty$, this implies the desired result.

Proof of step 2. ($j = 1$): Let $s_{n,1}^2 = \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} X_{nt} X_{n,t+|\tau|}$. We have that

$$\begin{aligned} \hat{\sigma}_{n,1}^2 &= s_{n,1}^2 + \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} \left(-\bar{X}_{\alpha,n} X_{nt} - \bar{X}_{\alpha,n} X_{n,t+|\tau|} + \bar{X}_{\alpha,n}^2\right) \quad \text{and} \\ \tilde{\sigma}_{n,1}^2 &= s_{n,1}^2 + \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} \left(-\mu_{n,t+|\tau|} X_{nt} - \mu_{nt} X_{n,t+|\tau|} + \mu_{nt} \mu_{n,t+|\tau|}\right), \end{aligned}$$

where $\bar{X}_{\alpha,n} = \sum_{t=1}^n \alpha_{n,t} X_{nt}$. We let $\alpha_{n,t} \equiv \alpha_n(t)$ throughout. If we let $\bar{\mu}_{\alpha,n} = \sum_{t=1}^n \alpha_{n,t} \mu_{nt}$ and note that $\sum_{t=1}^n \alpha_{n,t} = 1$ and $\sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} = 1$ for each τ , after some manipulations we can write

$$\begin{aligned} \hat{\sigma}_{n,1}^2 - \tilde{\sigma}_{n,1}^2 &= A_{n1} + A_{n2} + A_{n3} + A_{n4}, \quad \text{where} \\ A_{n1} &= -(\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n}) \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} (Z_{nt} + Z_{n,t+|\tau|}), \\ A_{n2} &= \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} (\mu_{nt} - \bar{\mu}_{\alpha,n}) Z_{n,t+|\tau|}, \\ A_{n3} &= \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} (\mu_{n,t+|\tau|} - \bar{\mu}_{\alpha,n}) Z_{nt}, \quad \text{and} \\ A_{n4} &= - \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} \left(\bar{X}_{\alpha,n}^2 - (\mu_{nt} + \mu_{n,t+|\tau|}) \bar{X}_{\alpha,n} + \mu_{nt} \mu_{n,t+|\tau|}\right). \end{aligned}$$

Notice that if we assume $\mu_{nt} = \mu$ for all t , $\bar{\mu}_{\alpha,n} = \mu$ since $\sum_s \alpha_{n,s} = 1$, which implies $A_{n2} = A_{n3} = 0$ and $A_{n4} = -(\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n})^2 \ell$, because $\sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} = 1$ for every τ . Thus, if $\mu_{nt} = \mu$ for all t , the result follows by showing that $A_{n1} = o_P(1)$ and $(\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n})^2 = o_P(\ell^{-1})$. If μ_{nt} is not constrained to be equal to μ for every t , the expression for A_{n4} is more complicated. In particular, we obtain

$$\begin{aligned} A_{n4} &= A'_{n4} + u_{n,1}^2, \quad \text{where} \\ A'_{n4} &= (\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n})^2 \ell + 2(\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n}) \bar{\mu}_{\alpha,n} \ell \\ &\quad - (\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n}) \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} (\mu_{nt} + \mu_{n,t+|\tau|}), \quad \text{and} \\ u_{n,1}^2 &= \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} (\mu_{nt} - \bar{\mu}_{\alpha,n}) (\mu_{n,t+|\tau|} - \bar{\mu}_{\alpha,n}). \end{aligned}$$

That $u_{n,1}^2 = \text{var}^* \left(n^{-1/2} \sum_{t=1}^n \mu_{nt}^{*(1)} \right)$ follows by Theorems 3.1 and 3.4 of Künsch (1989), when we apply the MBB resampling scheme to the time series $\{\mu_{nt}\}$. Thus, to complete the proof of step 2 we need to show that A_{n1}, A_{n2}, A_{n3} and A'_{n4} converge to zero in probability.

We first show that $(\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n}) \ell = o_P(1)$. By definition, $\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n} = \sum_{t=1}^n \alpha_{n,t} Z_{nt}$ and $0 < \alpha_{n,t} \leq \frac{1}{n-\ell+1}$ for every t . Thus, arguing as in Gallant and White (1988, p. 103), by an application of Chebyshev's inequality and McLeish's inequality (McLeish, 1975), we obtain for any $\varepsilon > 0$ and $\ell \equiv \ell_n = o(n^{1/2})$,

$$P \left[\left| \ell \sum_{t=1}^n \alpha_{n,t} Z_{nt} \right| \geq \varepsilon \right] \leq \frac{\ell^2}{(n-\ell+1)^2 \varepsilon^2} E \left[\max_{1 \leq j \leq n} \left(\sum_{t=1}^j |Z_{nt}| \right)^2 \right] \leq \frac{\ell^2}{n} \left(\frac{n}{n-\ell+1} \right)^2 K \Delta^2 \rightarrow 0,$$

since under our assumptions, it follows that $\{|Z_{nt}|, \mathcal{F}_n^t\}$ is a mixingale of size $-\frac{1}{2}$ with mixingale constants $c_{nt} \leq \Delta < \infty$ for all n, t (apply e.g. Theorem 17.12 and Corollary 17.6 in Davidson, 1994). Thus, $(\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n}) \ell = o_P(1)$. This immediately implies that $A'_{n4} = o_P(1)$ since under our assumptions $|\mu_{nt}| \leq K$ for all n, t and $\sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} = 1$ for any fixed τ .

Next, we consider A_{n1} . Because $\bar{X}_{\alpha,n} - \bar{\mu}_{\alpha,n} = o_P(1)$ it suffices to prove that $\sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} Y_{nt} = o_P(1)$, where $Y_{nt} \equiv Z_{nt} + Z_{n,t+|\tau|}$. An application of Markov's inequality, Cauchy-Schwartz inequality, and McLeish's inequality yields for any $\varepsilon > 0$,

$$\begin{aligned} P \left[\left| \sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} \beta_{n,t,\tau} Y_{nt} \right| \geq \varepsilon \right] &\leq \frac{1}{(n-\ell+1) \varepsilon} E \left[\sum_{\tau=-\ell+1}^{\ell-1} \left(1 - \frac{|\tau|}{\ell}\right) \sum_{t=1}^{n-|\tau|} |Y_{nt}| \right] \\ &\leq \frac{K}{\varepsilon} \frac{\ell}{n^{1/2}} \left(\frac{n}{n-\ell+1} \right) \rightarrow 0, \end{aligned}$$

given our assumptions on the block size ℓ . We can show that A_{n2} and A_{n3} converge to zero in probability by a similar argument as the one used for A_{n1} , by letting $Y_{nt} \equiv (\mu_{nt} - \bar{\mu}_{\alpha,n}) Z_{n,t+|\tau|}$ and $Y_{nt} \equiv (\mu_{n,t+|\tau|} - \bar{\mu}_{\alpha,n}) Z_{nt}$, respectively (and noting that $|\mu_{nt} - \bar{\mu}_{\alpha,n}| \leq 2K < \infty$ for all t, n). This completes the proof of step 2 for $j = 1$.

The proof of the theorem for the SB follows closely that for the MBB and we omit the details.

Proof of step 1. ($j = 2$):

Let $\tilde{\sigma}_{n,2}^2 = \hat{R}_{n,0}(0) + 2 \sum_{\tau=1}^{n-1} b_n(\tau) \hat{R}_{n,0}(\tau)$, where $\hat{R}_{n,0}(\tau) = n^{-1} \sum_{t=1}^{n-\tau} Z_{nt} Z_{n,t+\tau}$ and $Z_{nt} \equiv X_{nt} - \mu_{nt}$ as before. Given that $E(\hat{R}_{n,0}(\tau)) = n^{-1} \sum_{t=1}^{n-\tau} R_{n,t}(\tau)$, where $R_{n,t}(\tau) = E(Z_{nt} Z_{n,t+\tau})$, it follows that $E(\tilde{\sigma}_{n,2}^2) - \sigma_n^2 = 2 \sum_{\tau=1}^{n-1} (b_n(\tau) - 1) n^{-1} \sum_{t=1}^{n-\tau} R_{n,t}(\tau)$. We use exactly the same argument as before to prove that $E(\tilde{\sigma}_{n,2}^2) - \sigma_n^2 \rightarrow 0$ (in particular, we remark that $\lim_{n \rightarrow \infty} b_n(\tau) = 1$ for each τ and that $|b_n(\tau)| \leq 2$ for all n and τ). We then apply the Lebesgue dominated con-

vergence theorem and use the size conditions on α_τ and v_τ to find a dominating function, which yields the desired result).

To show that the variance of $\tilde{\sigma}_{n,2}^2$ tends to 0, we follow the same steps as for the MBB and show that $\text{var}\left(\hat{R}_{n,0}(\tau)\right)$ is bounded by $\frac{S}{n}$, where S is a constant independent of τ and n . In particular, we can show that

$$\text{var}\left(\tilde{\sigma}_{n,2}^2\right) \leq \frac{S}{n} + 4\frac{S}{n} \sum_{\tau=1}^{n-1} b_n(\tau) + 4\frac{S}{n} \sum_{\tau=1}^{n-1} \sum_{\lambda=1}^{n-1} b_n(\tau) b_n(\lambda) = O\left((np_n^2)^{-1}\right).$$

Given that $p_n \equiv \ell_n^{-1}$, the assumption that $\ell_n = o(n^{1/2})$ implies that $np_n^2 \rightarrow \infty$. Thus, $\text{var}(\tilde{\sigma}_{n,2}^2) \rightarrow 0$, which completes the proof of step 1.

Proof of step 2. ($j = 2$): Let $s_{n,2}^2 = n^{-1} \sum_{t=1}^n X_{nt}^2 + 2 \sum_{\tau=1}^{n-1} b_n(\tau) n^{-1} \sum_{t=1}^{n-\tau} X_{nt} X_{n,t+\tau}$. We obtain

$$\hat{\sigma}_{n,2}^2 = s_{n,2}^2 - \bar{X}_n^2 - 2\bar{X}_n^2 \sum_{\tau=1}^{n-1} \left(1 - \frac{\tau}{n}\right) (1-p)^\tau, \quad \text{and} \quad (\text{A.9})$$

$$\begin{aligned} \tilde{\sigma}_{n,2}^2 &= s_{n,2}^2 - 2n^{-1} \sum_{t=1}^n X_{nt} \mu_{nt} + n^{-1} \sum_{t=1}^n \mu_{nt}^2 \\ &\quad - 2 \sum_{\tau=1}^{n-1} \left(1 - \frac{\tau}{n}\right) (1-p)^\tau \left(n^{-1} \sum_{t=1}^{n-\tau} X_{nt} \mu_{n,t+\tau} + n^{-1} \sum_{t=1}^{n-\tau} X_{n,t+\tau} \mu_{nt} \right) \\ &\quad + 2 \sum_{\tau=1}^{n-1} \left(1 - \frac{\tau}{n}\right) (1-p)^\tau \left(n^{-1} \sum_{t=1}^{n-\tau} \mu_{nt} \mu_{n,t+\tau} \right). \end{aligned} \quad (\text{A.10})$$

If we let $\bar{\mu}_n \equiv n^{-1} \sum_{t=1}^n \mu_{nt}$, it follows that the difference between (A.9) and (A.10) can be simplified to

$$\hat{\sigma}_{n,2}^2 - \tilde{\sigma}_{n,2}^2 = \zeta_{n1} + \zeta_{n2} + \zeta_{n3} + \zeta_{n4} + u_{n,2}^2, \quad \text{where}$$

$$\zeta_{n1} = -(\bar{X}_n - \bar{\mu}_n)^2; \quad \zeta_{n2} = -2(\bar{X}_n - \bar{\mu}_n)^2 \sum_{\tau=1}^{n-1} \left(1 - \frac{\tau}{n}\right) (1-p)^\tau;$$

$$\zeta_{n3} = -n^{-1} \sum_{t=1}^n (X_{nt} - \mu_{nt})(\bar{\mu}_n - \mu_{nt});$$

$$\zeta_{n4} = -2 \sum_{\tau=1}^{n-1} b_n(\tau) \left[n^{-1} \sum_{t=1}^{n-\tau} (X_{nt} - \mu_{nt})(\bar{\mu}_n - \mu_{n,t+\tau}) + n^{-1} \sum_{t=1}^{n-\tau} (X_{n,t+\tau} - \mu_{n,t+\tau})(\bar{\mu}_n - \mu_{nt}) \right];$$

$$u_{n,2}^2 = \left\{ n^{-1} \sum_{t=1}^n (\mu_{nt} - \bar{\mu}_n)^2 + 2 \sum_{\tau=1}^{n-1} b_n(\tau) n^{-1} \sum_{t=1}^{n-\tau} (\mu_{nt} - \bar{\mu}_n)(\mu_{n,t+\tau} - \bar{\mu}_n) \right\}.$$

By a CLT for the sample mean, $\sqrt{n}(\bar{X}_n - \bar{\mu}_n) = O_P(1)$ and it follows that $\zeta_{n1} \equiv -(\bar{X}_n - \bar{\mu}_n)^2 = O_P(n^{-1}) = o_P(1)$. This also implies that $\zeta_{n2} = O_P\left(\frac{\ell_n}{n}\right) = o_P(1)$ because $\sum_{\tau=1}^{n-1} \left(1 - \frac{\tau}{n}\right) (1-p)^\tau \leq$

$\frac{1}{p_n}$ and because $p_n \equiv \ell_n^{-1}$ and $\ell_n = o(n^{1/2})$ by assumption. The remaining terms are $o_P(1)$ by an argument similar to the one used for the MBB to show that $A_{ni} = o_P(1)$. In particular, we can show that $\zeta_{n3} = O_P(n^{-1})$ and $\zeta_{n4} = O_P\left(\frac{1}{n^{1/2}p_n}\right)$.

By Lemma 1 in Politis and Romano (1994), $u_{n,2}^2 \equiv \text{var}^*\left(\sqrt{n}\bar{\mu}_n^{*(2)}\right)$. Hence, $\hat{\sigma}_{n,2}^2 - (\hat{\sigma}_{n,2}^2 + u_{n,2}^2) \rightarrow 0$ in probability, which completes the proof of step 2 for $j = 2$, and completes the proof of Theorem 2.1. ■

Proof of Lemma 2.1. For $j = 1$, apply Theorems 3.1 and 3.4 of Künsch (1989) to $\{\mu_{nt}\}$, and for $j = 2$ apply Lemma 1 of Politis and Romano (1994) to $\{\mu_{nt}\}$. ■

Proof of Corollary 2.1. Immediate from Theorem 2.1 and the remark that follows it. ■

Proof of Theorem 2.2. By Theorem 5.3 in Gallant and White (1988) and under our assumptions, $\sigma_n^{-1}\sqrt{n}(\bar{X}_n - \bar{\mu}_n) \Rightarrow N(0, 1)$. (i) follows then by an application of Polya's theorem (e.g. Serfling, 1980) given that the normal distribution is continuous.

To prove (ii), we first remark that for both $j = 1, 2$, we can write

$$\begin{aligned} \sigma_n^{-1}\sqrt{n}(\bar{X}_n^{*(j)} - \bar{X}_n) &= \sigma_n^{-1}\sqrt{n}(\bar{Z}_n^{*(j)} - E^*(\bar{Z}_n^{*(j)})) + \sigma_n^{-1}\sqrt{n}(E^*(\bar{Z}_n^{*(j)}) - \bar{Z}_n) \\ &\quad + \sigma_n^{-1}\sqrt{n}(\bar{\mu}_n^{*(j)} - \bar{\mu}_n) \\ &\equiv A_n^{(j)} + B_n^{(j)} + C_n^{(j)}, \end{aligned}$$

where $Z_{nt} \equiv X_{nt} - \mu_{nt}$ has mean zero by construction and $Z_{nt}^{*(j)} \equiv X_{nt}^{*(j)} - \mu_{nt}^{*(j)}$.

Proof for $j = 1$. For the MBB, $E^*(\bar{Z}_n^{*(1)}) \neq \bar{Z}_n$, hence $B_n^{(1)} \neq 0$. In particular (cf. Künsch, 1989, expression 3.14, p. 1226),

$$E^*(\bar{Z}_n^{*(1)}) = (n - \ell + 1)^{-1} \ell^{-1} \sum_{j=0}^{n-\ell} \sum_{t=1}^{\ell} Z_{n,t+j}. \quad (\text{A.11})$$

Thus, we start by proving that under our assumptions $B_n \xrightarrow{P} 0$. It is easy to show that $E^*(\bar{Z}_n^{*(1)}) = \bar{Z}_n + O_P\left(\frac{\ell}{n}\right)$ (e.g. Lemma A.1 of Fitzenberger (1997, p. 263)). Therefore, $B_n = O_P\left(\frac{\ell}{n^{1/2}}\right)$, given that $\sigma_n^{-1} = O(1)$ by the assumption that $\sigma_n^2 > \kappa > 0$. Since $\ell \equiv \ell_n = o(n^{1/2})$, it follows that $B_n = o_P(1)$.

Next, we show that $A_n \Rightarrow N(0, 1)$, given Z_{n1}, \dots, Z_{nn} in a set with probability converging to one. Following Künsch (1989), we can write $\bar{Z}_n^{*(1)} = k^{-1} \sum_{i=1}^k U_{n,i}$, where $\{U_{n,i}\}$ are i.i.d. with $P^*\left(U_{n,i} = \frac{Z_{n,j+1} + \dots + Z_{n,j+\ell}}{\ell}\right) = \frac{1}{n-\ell+1}$, $j = 0, \dots, n - \ell$. Thus, $E^*(\bar{Z}_n^{*(1)}) = E^*(U_{n,1})$, and it follows that

$$A_n^{(1)} = \sigma_n^{-1} \sqrt{n} \left(k^{-1} \sum_{i=1}^k [U_{n,i} - E^*(U_{n,1})] \right) \equiv \sum_{i=1}^k \tilde{Z}_{ni},$$

where $\tilde{Z}_{ni} \equiv \sigma_n^{-1} n^{1/2} k^{-1} [U_{n,i} - E^*(U_{n,1})] = \sigma_n^{-1} n^{-1/2} \ell [U_{n,i} - E^*(U_{n,1})]$, given that $k = \frac{n}{\ell}$. In particular, $\{\tilde{Z}_{ni}\}$ are i.i.d. with $E^*(\tilde{Z}_{ni}) = 0$ and $\text{var}^*(\tilde{Z}_{n1}) = \sigma_n^{-2} n k^{-2} \text{var}^*(U_{n1}) = k^{-1} \frac{\hat{\sigma}_{n,1}^2}{\sigma_n^2}$. By Katz's (1963) Berry-Esseen Bound, for some small $\delta > 0$,

$$\begin{aligned} \sup_{x \in \mathbb{R}} \left| P^* \left(\frac{\sum_{i=1}^k \tilde{Z}_{ni}}{\sqrt{\text{var}^*(\sum_{i=1}^k \tilde{Z}_{ni})}} \leq x \right) - \Phi(x) \right| &\leq K k^{-\delta/2} E^* \left| \frac{\tilde{Z}_{n1}}{\sqrt{\text{var}^*(\tilde{Z}_{n1})}} \right|^{2+\delta} \\ &\leq K k^{-\delta/2} |\text{var}^*(\tilde{Z}_{n1})|^{-1-\delta/2} E^* |\tilde{Z}_{n1}|^{2+\delta} \leq K \left(\frac{\hat{\sigma}_{n,1}^2}{\sigma_n^2} \right)^{-1-\delta/2} k E^* |\tilde{Z}_{n1}|^{2+\delta}. \end{aligned}$$

But $\frac{\hat{\sigma}_{n,1}^2}{\sigma_n^2} \rightarrow 1$ by Theorem 2.1 for $j = 1$. Thus, it suffices to show that $k E^* |\tilde{Z}_{n1}|^{2+\delta} \xrightarrow{P} 0$. By Markov's inequality, it suffices to show that $E \left| k E^* |\tilde{Z}_{n1}|^{2+\delta} \right| \rightarrow 0$. Now, given the definition of \tilde{Z}_{n1} and the fact that $k = \frac{n}{\ell}$, it follows that

$$\begin{aligned} E \left| k E^* |\tilde{Z}_{n1}|^{2+\delta} \right| &= \frac{n}{\ell(n-\ell+1)} \sum_{j=0}^{n-\ell} E \left| \sigma_n^{-1} n^{-1/2} \left(\sum_{t=1}^{\ell} Z_{n,t+j} - \ell E^*(U_{n,1}) \right) \right|^{2+\delta} \\ &\leq \frac{n}{\ell(n-\ell+1)} \frac{1}{n^{1+\frac{\delta}{2}} \sigma_n^{2+\delta}} \sum_{j=0}^{n-\ell} \left[\left(E \left| \sum_{t=1}^{\ell} Z_{n,t+j} \right|^{2+\delta} \right)^{1/(2+\delta)} + \left(E |\ell E^*(U_{n,1})|^{2+\delta} \right)^{1/(2+\delta)} \right]^{2+\delta} \\ &= \frac{n}{\ell(n-\ell+1)} \frac{1}{n^{1+\frac{\delta}{2}} \sigma_n^{2+\delta}} \sum_{j=0}^{n-\ell} \left[\left\| \sum_{t=1}^{\ell} Z_{n,t+j} \right\|_{2+\delta} + \|\ell E^*(U_{n,1})\|_{2+\delta} \right]^{2+\delta}, \end{aligned} \quad (\text{A.12})$$

where the inequality follows by Minkowski inequality. We now show that

$$\left\| \sum_{t=1}^{\ell} Z_{n,t+j} \right\|_{2+\delta} \leq K \Delta \ell^{1/2}, \quad (\text{A.13})$$

$$\text{and } \|\ell E^*(U_{n,1})\|_{2+\delta} \leq K \Delta \ell^{1/2}. \quad (\text{A.14})$$

To prove (A.13) and (A.14) we use an extension of a maximal inequality for mixingales (Hansen, Lemma 2, 1991) to a double array setting. According to this result, if $\{Z_{nt}, \mathcal{F}_n^t\}$ is a $L_{2+\delta}$ -mixingale for some $\delta > 0$ and the mixingale coefficients $\{\psi_k\}$ satisfy $\sum_{k=1}^{\infty} \psi_k < \infty$, then $\left\| \sum_{t=\tau}^{\tau+r-1} Z_{nt} \right\|_{2+\delta} \leq \left\| \max_{j \leq r} \left| \sum_{t=\tau}^{\tau+j-1} Z_{nt} \right| \right\|_{2+\delta} \leq K \left(\sum_{t=\tau}^{\tau+r-1} c_{nt}^2 \right)^{1/2}$, where c_{nt} are the mixingale constants. Under Assumption 2.1 strengthened by 2.1.b') it follows that $\{Z_{nt}\}$ is a $L_{2+\delta}$ -mixingale with mixingale coefficients $\psi_k = 6\alpha_k^{\frac{1}{2+\delta} - \frac{1}{3r}} + v_k$ and constants $c_{nt} \leq \Delta < \infty$ (see e.g. Davidson, Corollary 17.6, 1994). Hence, it suffices to assume α_k is of size $-\frac{3r(2+\delta)}{3r-(2+\delta)}$ and v_k is of size -1 . But, under

Assumptions 2.1. strengthened by 2.1.b'), it follows that these conditions are satisfied, which proves (A.13). To prove (A.14), notice that $E^*(U_{n,1}) = E^*(\bar{Z}_n^{*(1)})$. Thus, by (A.11),

$$\|lE^*(U_{n,1})\|_{2+\delta} = \left\| (n-\ell+1)^{-1} \ell \ell^{-1} \sum_{j=0}^{n-\ell} \sum_{t=1}^{\ell} Z_{n,t+j} \right\|_{2+\delta} \leq \frac{1}{n-\ell+1} \sum_{j=0}^{n-\ell} \left\| \sum_{t=1}^{\ell} Z_{n,t+j} \right\|_{2+\delta} \leq K\Delta\ell^{1/2},$$

where the first inequality follows by Minkowski's inequality and the second inequality follows by the maximal inequality for $L_{2+\delta}$ -mixingales. Therefore, it follows from (A.12) that

$$E \left| kE^* \left| \tilde{Z}_{n1} \right|^{2+\delta} \right| \leq \frac{n}{\ell(n-\ell+1)} \frac{1}{n^{1+\frac{\delta}{2}} \sigma_n^{2+\delta}} \sum_{j=0}^{n-\ell} (2K\Delta\ell^{1/2})^{2+\delta} = O\left(\left(\frac{\ell}{n}\right)^{\delta/2}\right) \rightarrow 0$$

given our assumption on ℓ_n . Finally, under our assumptions, we can show that $E^*(C_n) = O\left(\frac{\ell}{n}\right) \rightarrow 0$ and $\text{var}^*(C_n) = \sigma_n^{-2} u_{n,1}^2 \rightarrow 0$. Thus, $C_n^{(1)} \xrightarrow{P^*} 0$ and by the Asymptotic Equivalence Lemma (e.g. White, 1984, Lemma 4.7),

$$\sup_{x \in \mathbb{R}} \left| P^* \left[\sigma_n^{-1} \sqrt{n} \left(\bar{X}_n^{*(1)} - \bar{X}_n \right) \leq x \right] - \Phi(x) \right| \xrightarrow{P} 0,$$

given that $B_n^{(1)} \xrightarrow{P} 0$ and that $C_n^{(1)} \xrightarrow{P^*} 0$, where the convergence is uniform by Polyá's theorem given that $\Phi(\cdot)$ is continuous. This completes the proof of Theorem 2.2.(ii) for $j = 1$.

The idea of the proof for the SB is similar to that for the MBB in that the sample mean is also expressed as the sum of i.i.d. blocks of observations and the Berry-Esseen Bound is used to prove the (uniform) convergence of this sum to a standard normal distribution. However, for the SB the number of blocks and the number of observations within each block are random, which renders the proof more complicated than for the MBB. We follow Politis and Romano (1994) and refer to the proof of their Theorem 2 for the relevant details.

First, note that $B_n^{(2)} = 0$ since for the SB $E^*(\bar{Z}_n^{*(1)}) = \bar{Z}_n$. Hence, it suffices to show that with probability approaching one $A_n^{(2)} \Rightarrow N(0,1)$, conditional on the original sample, and $C_n^{(2)} \xrightarrow{P^*} 0$.

To prove that $A_n^{(2)} \Rightarrow N(0,1)$ we verify the conditions (C1), (C2), and (C3) in Politis and Romano's (1994) proof of their Theorem 2. In our more general case, where σ_n^2 is not assumed to have a limit value σ_∞^2 , the appropriate version of these conditions is as follows.

$$\text{(C1)} \quad \frac{n\bar{Z}_n^2}{np_n} \xrightarrow{P} 0;$$

$$\text{(C2)} \quad \hat{C}_n(0) + 2 \sum_{\tau=1}^{\infty} (1-p_n)^\tau \hat{C}_n(\tau) - \sigma_n^2 \xrightarrow{P} 0;$$

$$\text{(C3)} \quad \frac{p_n}{n^{1+\frac{\delta}{2}}} \sum_{r=1}^{\infty} \sum_{\tau=1}^n \left| S_{\tau,r} - \frac{\bar{Z}_n}{p_n} \right|^{2+\delta} (1-p_n)^{r-1} p_n \xrightarrow{P} 0,$$

where, in (C2), $\hat{C}_n(\tau)$ denotes the circular autocovariance of $\{Z_{nt}\}$ at lag τ , i.e. $\hat{C}_n(\tau) = n^{-1} \sum_{t=1}^n Z_{nt} Z_{nt+\tau} - \bar{Z}_n^2$, and in (C3), $S_{\tau,r}$ is defined as the sum of observations in block $B_{\tau,b}$: $B_{\tau,b} = \{Z_{n\tau}, Z_{n,\tau+1}, \dots, Z_{n,\tau+r-1}\}$.

Proof of (C1): this follows trivially because $n\bar{Z}_n^2 \equiv n(\bar{X}_n - \bar{\mu}_n)^2 = O_P(1)$ (by the CLT) and because $np_n \rightarrow \infty$.

Proof of (C2): If we let $\sigma_{n,\infty}^2 = \hat{C}_n(0) + 2 \sum_{\tau=1}^{\infty} (1-p_n)^\tau \hat{C}_n(\tau)$, we can write $\sigma_{n,\infty}^2$ as

$$\begin{aligned} \sigma_{n,\infty}^2 &= \hat{\sigma}_{n,2}^2 \sum_{j=0}^{\infty} (1-p_n)^{nj} + \hat{C}_n(0) \sum_{j=1}^{\infty} (1-p_n)^{nj} \\ &= \sigma_n^2 + (\hat{\sigma}_n^2 - \sigma_n^2) \frac{1}{1 - (1-p_n)^n} + \hat{C}_n(0) \frac{(1-p_n)^n}{1 - (1-p_n)^n}, \end{aligned}$$

where $\hat{\sigma}_n^2 = \hat{C}_n(0) + 2 \sum_{\tau=1}^{n-1} (1-p_n)^\tau \hat{C}_n(\tau)$. By an argument similar to the proof of Theorem 2.1 for $j = 2$, we can show that $\hat{\sigma}_n^2 - \sigma_n^2 \xrightarrow{P} 0$. Hence, $\sigma_{n,\infty}^2 - \sigma_n^2 \rightarrow 0$ in probability given that $np_n \rightarrow \infty$ and $p_n \rightarrow 0$ and that $\hat{C}_n(0) = n^{-1} \sum_{t=1}^n Z_{nt}^2 - \bar{Z}_n^2 = O_P(1)$.

Proof of (C3): this follows provided we ensure $E \left| \sum_{t=\tau}^{\tau+r-1} X_{nt} \right|^{2+\delta} \leq Kr^{1+\frac{\delta}{2}}$, where the constant K only depends on the NED coefficients of $\{X_{nt}\}$. Politis and Romano (1994) use Yokoyama's (1980) moment inequality for mixing processes to achieve this. We use the maximal inequality for mixingales (Hansen, Lemma 2, 1991) referred to above, which holds under our assumptions.

To prove $C_n^{(2)} \xrightarrow{P^*} 0$, note that $E^* \left(n^{-1/2} \sum_{t=1}^n \left(\mu_{nt}^{*(2)} - \mu_{nt} \right) \right) = 0$, given the stationarity of the Politis and Romano resampling scheme. Moreover, $\text{var}^* \left(n^{-1/2} \sum_{t=1}^n \left(\mu_{nt}^{*(2)} - \mu_{nt} \right) \right) \equiv u_{n,2}^2$, and by Assumption 2.2, $u_{n,2}^2 \rightarrow 0$. This completes the proof of Theorem 2.1 (ii) for $j = 2$. ■

Proof of Corollary 2.2. Immediate from the proof of Theorem 2.2 for $j = 1$. ■

Proof of Theorem 3.1. We consider $\{Y_{nt} \equiv \lambda' X_{nt}\}$ for any $\lambda \in \mathbb{R}^d$, and we assume without loss of generality that λ has unit norm, i.e. $\lambda' \lambda = 1$. We now verify Assumption 2.1 for $\{Y_{nt}\}$. By hypothesis, each element of $\{X_{nt}\}$ satisfies Assumption 2.1. Thus, for some $r > 2$, $\|Y_{nt}\|_{3r} = \|\lambda' X_{nt}\|_{3r} \leq \sum_{i=1}^d |\lambda_i| \|X_{nti}\|_{3r} \leq \Delta \sqrt{d} < \infty$ for all n, t . Moreover, $\{Y_{nt} \equiv \lambda' X_{nt}\}$ is NED on $\{\varepsilon_t\}$ of size $-\frac{2(r-1)}{r-2}$. By Theorem 2.1 applied to $\{Y_{nt}\}$ we obtain that for any $\lambda \in \mathbb{R}^d$ such that $\lambda' \lambda = 1$ and for $j = 1, 2$, $\lambda' \hat{V}_{n,j} \lambda - (\lambda' V_n \lambda + \lambda' U_{n,j} \lambda) \xrightarrow{P} 0$, where

$$\begin{aligned} U_{n,1} &\equiv \sum_{t=1}^n \beta_{n,t,0} (\mu_{nt} - \bar{\mu}_{\alpha,n}) (\mu_{nt} - \bar{\mu}_{\alpha,n})' \\ &\quad + \sum_{\tau=1}^{\ell-1} \left(1 - \frac{\tau}{\ell}\right) \sum_{t=1}^{n-\tau} \beta_{n,t,\tau} \left[(\mu_{nt} - \bar{\mu}_{\alpha,n}) (\mu_{n,t+\tau} - \bar{\mu}_{\alpha,n})' + (\mu_{n,t+\tau} - \bar{\mu}_{\alpha,n}) (\mu_{n,t} - \bar{\mu}_{\alpha,n})' \right]; \end{aligned}$$

$$\begin{aligned}
U_{n,2} &\equiv n^{-1} \sum_{t=1}^n (\mu_{nt} - \bar{\mu}_n) (\mu_{nt} - \bar{\mu}_n)' \\
&\quad + \sum_{\tau=1}^{n-1} b_n(\tau) n^{-1} \sum_{t=1}^{n-\tau} \left[(\mu_{nt} - \bar{\mu}_n) (\mu_{n,t+\tau} - \bar{\mu}_n)' + (\mu_{n,t+\tau} - \bar{\mu}_n) (\mu_{nt} - \bar{\mu}_n)' \right].
\end{aligned}$$

Because $U_{n,j} = \text{var}^* \left(\sqrt{n} \bar{\mu}_n^{*(j)} \right)$ for $j = 1, 2$ and λ is arbitrary, the desired result follows. ■

Proof of Theorem 3.2. That $\{V_n\} = O(1)$ follows from Assumption 2.1 (see Gallant and White, 1988, pp. 86 for the details of the proof of a similar result). The remaining results follow from Theorem 2.2 by considering linear combinations of the elements of X_{nt} . In particular, consider $\{Y_{nt} \equiv \lambda' X_{nt}\}$, where $\lambda \in \mathbb{R}^d$ with $\lambda' \lambda = 1$. By an argument similar to the one used in the proof of Theorem 3.1 we can prove that $\{Y_{nt}\}$ satisfies Assumptions 2.1 and 2.2 strengthened by 2.1.b') given that each element of $\{X_{nt}\}$ satisfies these assumptions. Hence, we obtain for any $\lambda \in \mathbb{R}^d$ such that $\lambda' \lambda = 1$, $(\lambda' V_n \lambda)^{-1/2} \lambda' \sqrt{n} (\bar{X}_n - \bar{\mu}_n) \Rightarrow N(0, 1)$, where the assumption that V_n is uniformly positive definite implies that $\lambda' V_n \lambda \geq \kappa > 0$ as required by Theorem 2.2. Moreover, for $j = 1, 2$, $\sup_{x \in \mathbb{R}} \left| P^* \left[\lambda' \sqrt{n} (\bar{X}_n^{*(j)} - \bar{X}_n) \leq x \right] - P \left[\lambda' \sqrt{n} (\bar{X}_n - \bar{\mu}_n) \leq x \right] \right| \xrightarrow{P} 0$, which implies part (ii) since λ is arbitrary. ■

Proof of Theorem 3.3. Parts (i) and (ii) follow trivially by an application of Theorem 3.1 to the double array $\{h(X_{nt})\}$. Part (iii) is established by routine arguments, given part (i). In particular, we apply a mean value expansion to $f(\bar{Z}_n)$ around $\bar{\mu}_{Z_n}$ to show that $\sqrt{n} \Gamma_n^{-1/2} (f(\bar{Z}_n) - f(\bar{\mu}_{Z_n})) \Rightarrow N(0, I_q)$, where $\Gamma_n = \nabla' f(\bar{\mu}_{Z_n}) \Omega_n \nabla f(\bar{\mu}_{Z_n})$. Hence, to prove part (iii), it suffices to show that for $j = 1, 2$, $\sqrt{n} \Gamma_n^{-1/2} (f(\bar{Z}_n^{*(j)}) - f(\bar{Z}_n)) \Rightarrow N(0, I_p)$ under the bootstrap probability measure P^* for all sequences $\{X_{nt}\}$ that lie in a set with probability approaching one.

Consider $\mathcal{X}_n = \{X_{n1}, \dots, X_{nn}\}$ such that

$$\bar{Z}_n - \bar{\mu}_{Z_n} \rightarrow 0 \tag{A.15}$$

$$\sqrt{n} \Omega_n^{-1/2} (\bar{Z}_n^{*(j)} - \bar{Z}_n) \Rightarrow N(0, I_p), \tag{A.16}$$

where \Rightarrow denotes convergence in distribution under P^* .

Claim. $\sqrt{n} \Gamma_n^{-1/2} (f(\bar{Z}_n^{*(j)}) - f(\bar{Z}_n)) \Rightarrow N(0, I_p)$ under P^* for every \mathcal{X}_n .

Proof of Claim. Fix \mathcal{X}_n . For each $i = 1, \dots, p$ and $j = 1, 2$, apply a mean value expansion to $f_i(\bar{Z}_n^{*(j)})$,

$$\begin{aligned}
\sqrt{n} \left(f_i \left(\bar{Z}_n^{*(j)} \right) - f_i \left(\bar{Z}_n \right) \right) &= \left[\nabla' f_i \left(\bar{\mu}_{Z_n} \right) - \nabla' f_i \left(\zeta_n^{*(i,j)} \right) \right] \sqrt{n} \left(\bar{Z}_n^{*(j)} - \bar{Z}_n \right) \\
&\quad + \nabla' f_i \left(\bar{\mu}_{Z_n} \right) \sqrt{n} \left(\bar{Z}_n^{*(j)} - \bar{Z}_n \right) \\
&= r_{ni}^{*(j)} + \nabla' f_i \left(\bar{\mu}_{Z_n} \right) \sqrt{n} \left(\bar{Z}_n^{*(j)} - \bar{Z}_n \right),
\end{aligned}$$

where $\zeta_n^{*(i,j)}$ lies in the segment connecting $\bar{Z}_n^{*(j)}$ and \bar{Z}_n and the superscript (i, j) reflects the fact that the mean value may change for every i and j . We first show that $r_{ni}^{*(j)} \xrightarrow{P^*} 0$. For any \mathcal{X}_n , $\sqrt{n}\Omega_n^{-1/2} \left(\bar{Z}_n^{*(j)} - \bar{Z}_n \right) \Rightarrow N(0, I_p)$, where $\Omega_n^{-1/2} = O(1)$ given that Ω_n is $O(1)$ and is uniformly positive definite. Thus, $\sqrt{n} \left(\bar{Z}_n^{*(j)} - \bar{Z}_n \right) = o_{P^*}(1)$ and $\bar{Z}_n^{*(j)} - \bar{Z}_n = o_{P^*}(1)$. Since $\bar{Z}_n - \bar{\mu}_{Z_n} = o(1)$ for all \mathcal{X}_n , it follows that for such $\{\mathcal{X}_{nt}\}$, $\zeta_n^{*(i,j)} - \bar{\mu}_{Z_n} = o_{P^*}(1)$, where $\bar{\mu}_{Z_n}$ is interior to $C = \times_{i=1}^p [-2\Delta^{3r}, 2\Delta^{3r}]$, a compact subset of \mathbb{R}^p , uniformly in n . Since $\nabla' f_i$ is continuous by assumption, it is uniformly continuous in C . It follows that $\nabla' f_i \left(\zeta_n^{*(i,j)} \right) - \nabla' f_i \left(\bar{\mu}_{Z_n} \right) = o_{P^*}(1)$ by Proposition 2.30 of White (1984). Hence, for all $i = 1, \dots, p$, and $j = 1, 2$, $r_{ni}^{*(j)} = o_{P^*}(1)$, or in vector form, $\sqrt{n} \left(f \left(\bar{Z}_n^{*(j)} \right) - f \left(\bar{Z}_n \right) \right) - \nabla' f \left(\bar{\mu}_{Z_n} \right) \sqrt{n} \left(\bar{Z}_n^{*(j)} - \bar{Z}_n \right) \xrightarrow{P^*} 0$. The proof of the Claim follows by an application of Corollary 4.24 and Lemma 4.7 of White (1984), given condition (A.16).

To complete the proof of Theorem 3.3 by a subsequence argument it suffices to show that the convergences in (A.15) and (A.16) hold in probability for the original sequence $\{\mathcal{X}_{nt}\}$. But this follows trivially by parts (i) and (ii). ■

References

- [1] Andrews, D.W.K. (1984) Non-strong mixing autoregressive processes. *Journal of Applied Probability* 21, 930-934.
- [2] Andrews, D.W.K. (1988) Laws of large numbers for dependent nonidentically distributed random variables. *Econometric Theory* 4, 458-467.
- [3] Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 59, 817-858.
- [4] Andrews, D.W.K. (1999) Higher-order Improvements of a computationally attractive k -step bootstrap for extremum estimators. *Cowles Foundation for Research in Economics*, Yale University.
- [5] Andrews, D.W.K. & J.C. Monahan (1992) An improved heteroskedasticity and autocorrelation consistent covariance matrix estimator. *Econometrica* 60, 953-966.

- [6] Bickel, P.J. & D.A. Freedman (1981) Some asymptotic theory for the bootstrap. *Annals of Statistics* 9, 1196-1217.
- [7] Billingsley, P. (1968) *Convergence of probability measures*. New York: Wiley.
- [8] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307-327.
- [9] Carrasco, M. & X. Chen (1999) Beta-mixing and moment properties of various GARCH, stochastic volatility and ACD models. London School of Economics, manuscript.
- [10] Davidson, J. (1994) *Stochastic limit theory*. Oxford: Oxford University Press.
- [11] Efron, B. (1979) Bootstrap methods: another look at the jackknife. *Annals of Statistics* 7, 1-26.
- [12] Engle, R.F. (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50, 987-1006.
- [13] Fitzenberger, B. (1997) The moving blocks bootstrap and robust inference for linear least squares and quantile regressions. *Journal of Econometrics* 82, 235-287.
- [14] Gallant, A.R. & H. White (1988) *A unified theory of estimation and inference for nonlinear dynamic models*. New York: Basil Blackwell.
- [15] Gonçalves, S. & H. White (2000) Maximum likelihood and the bootstrap for nonlinear dynamic models. University of California, San Diego, manuscript.
- [16] Hansen, B.E. (1991a) GARCH(1,1) processes are near epoch dependent. *Economics Letters* 36, 181-186.
- [17] Hansen, B.E. (1991b) Strong laws for dependent heterogeneous processes. *Econometric Theory* 7, 213-221.
- [18] Horowitz, J. (1999) The bootstrap. Forthcoming in *Handbook of Econometrics*, Vol. 5.
- [19] Katz, M.L. (1963) Note on the Berry-Esseen theorem. *Annals of Mathematical Statistics* 34, 1107-1108.
- [20] Künsch, H.R. (1989) The jackknife and the bootstrap for general stationary observations. *Annals of Statistics* 17, 1217-1241.

- [21] Lahiri, S.N. (1992) Edgeworth correction by ‘moving block’ bootstrap for stationary and nonstationary data. In LePage and Billiard (eds), *Exploring the limits of the bootstrap*, pp. 183-214. New York: Wiley.
- [22] Lahiri, S.N. (1999) Theoretical comparisons of block bootstrap methods. *Annals of Statistics* 27, 386-404.
- [23] Liu, R.Y. & K. Singh (1992) Moving blocks jackknife and bootstrap capture weak dependence. In LePage and Billiard (eds), *Exploring the limits of the bootstrap*, pp. 224-248. New York: Wiley.
- [24] McLeish, D.L. (1975) A maximal inequality and dependent strong laws. *Annals of Probability* 5, 616-621.
- [25] Newey, W.K. & K.D. West (1987) A simple positive semi-definite, heteroskedastic and autocorrelation consistent covariance matrix. *Econometrica* 55, 703-708.
- [26] Politis, D. & J. Romano (1994) The stationary bootstrap. *Journal of American Statistical Association* 89, 1303-1313.
- [27] Politis, D., Romano, J. & M. Wolf (1997) Subsampling for heteroskedastic time series. *Journal of Econometrics* 81, 281-317.
- [28] Politis, D. (1998) Computer-intensive methods in statistical analysis. *IEEE Signal Processing Magazine*, January 1998, 39-55.
- [29] Serfling, R.J. (1980) *Approximation theorems of mathematical statistics*. New York: Wiley.
- [30] Sin, C.-Y. & H. White (1996) Information criteria for selecting possibly misspecified parametric models. *Journal of Econometrics* 71, 207-225.
- [31] Singh, K. (1981) On the asymptotic accuracy of Efron’s bootstrap. *Annals of Statistics* 9, 1187-1195.
- [32] White, H. (1984) *Asymptotic theory for econometricians*. Orlando: Academic Press.
- [33] Yokoyama, R. (1980) Moment bounds for stationary mixing sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 52, 45-57.

Liste des publications au CIRANO *

Cahiers CIRANO / *CIRANO Papers* (ISSN 1198-8169)

- 99c-1 Les Expos, l'OSM, les universités, les hôpitaux : Le coût d'un déficit de 400 000 emplois au Québec — Expos, Montréal Symphony Orchestra, Universities, Hospitals: The Cost of a 400,000-Job Shortfall in Québec / Marcel Boyer
- 96c-1 Peut-on créer des emplois en réglementant le temps de travail? / Robert Lacroix
- 95c-2 Anomalies de marché et sélection des titres au Canada / Richard Guay, Jean-François L'Her et Jean-Marc Suret
- 95c-1 La réglementation incitative / Marcel Boyer
- 94c-3 L'importance relative des gouvernements : causes, conséquences et organisations alternative / Claude Montmarquette
- 94c-2 Commercial Bankruptcy and Financial Reorganization in Canada / Jocelyn Martel
- 94c-1 Faire ou faire faire : La perspective de l'économie des organisations / Michel Patry

Série Scientifique / *Scientific Series* (ISSN 1198-8177)

- 2001s-18 Perspectives on IT Outsourcing Success: Covariance Structure Modelling of a Survey of Outsourcing in Australia / Anne C. Rouse, Brian Corbitt et Benoit A. Aubert
- 2001s-17 A Theory of Environmental Risk Disclosure / Bernard Sinclair-Desgagné et Estelle Gozlan
- 2001s-16 Marriage Market, Divorce Legislation and Household Labor Supply / Pierre-André Chiappori, Bernard Fortin et Guy Lacroix
- 2001s-15 Properties of Estimates of Daily GARCH Parameters Based on Intra-Day Observations / John W. Galbraith et Victoria Zinde-Walsh
- 2001s-14 A Ricardian Model of the Tragedy of the Commons / Pierre Lasserre et Antoine Soubeyran
- 2001s-13 Carbon Credits for Forests and Forest Products / Robert D. Cairns et Pierre Lasserre
- 2001s-12 Estimating Nonseparable Preference Specifications for Asset Market Participants / Kris Jacobs
- 2001s-11 Autoregression-Based Estimators for ARFIMA Models / John W. Galbraith et Victoria Zinde-Walsh
- 2001s-10 Heterogeneous Returns to Human Capital and Dynamic Self-Selection / Christian Belzil et Jörgen Hansen

* Vous pouvez consulter la liste complète des publications du CIRANO et les publications elles-mêmes sur notre site Internet à l'adresse suivante :