REM WORKING PAPER SERIES

Quasi-Maximum Likelihood and the Kernel Block Bootstrap for Nonlinear Dynamic Models

Paulo M.D.C. Parente, Richard J. Smith

REM Working Paper 059-2018

November 2018

REM – Research in Economics and Mathematics
Rua Miguel Lúpi 20,
1249-078 Lisboa,
Portugal

ISSN 2184-108X

Quasi-Maximum Likelihood and The Kernel Block Bootstrap for Nonlinear Dynamic Models*

Paulo M.D.C. Parente
ISEG- Lisbon School of Economics & Management, Universidade de Lisboa
REM - Research in Economics and Mathematics;
CEMAPRE- Centro de Matemática Aplicada à Previsão e Decisão Económica.
pparente@iseg.ulisboa.pt

Richard J. Smith
*c*emmap, U.C.L and I.F.S.
Faculty of Economics, University of Cambridge
Department of Economics, University of Melbourne
ONS Economic Statistics Centre of Excellence
rjs27@econ.cam.ac.uk

This version: October 2018

Abstract

This paper applies a novel bootstrap method, the *kernel block bootstrap*, to quasi-maximum likelihood estimation of dynamic models with stationary strong mixing data. The method first kernel weights the components comprising the quasi-log likelihood function in an appropriate way and then samples the resultant transformed components using the standard "m out of n" bootstrap. We investigate the first order asymptotic properties of the KBB method for quasi-maximum likelihood demonstrating, in particular, its consistency and the first-order asymptotic validity of the bootstrap approximation to the distribution of the quasi-maximum likelihood estimator. A set of simulation experiments for the mean regression model illustrates the efficacy of the kernel block bootstrap for quasi-maximum likelihood estimation.

JEL Classification: C14, C15, C22

*Address for correspondence: Richard J. Smith, Faculty of Economics, University of Cambridge, Austin Robinson Building, Sidgwick Avenue, Cambridge CB3 9DD, UK

Keywords: Bootstrap; heteroskedastic and autocorrelation consistent inference; quasi-maximum likelihood estimation.

1 Introduction

This paper applies the kernel block bootstrap (KBB), proposed in Parente and Smith (2018), PS henceforth, to quasi-maximum likelihood estimation with stationary and weakly dependent data. The basic idea underpinning KBB arises from earlier papers, see, e.g., Kitamura and Stutzer (1997) and Smith (1997, 2011), which recognise that a suitable kernel function-based weighted transformation of the observational sample with weakly dependent data preserves the large sample efficiency for randomly sampled data of (generalised) empirical likelihood, (G)EL, methods. In particular, the mean of and, moreover, the standard random sample variance formula applied to the transformed sample are respectively consistent for the population mean [Smith (2011, Lemma A.1, p.1217)] and a heteroskedastic and autocorrelation (HAC) consistent and automatically positive semidefinite estimator for the variance of the standardized mean of the original sample [Smith (2005, Section 2, pp.161-165, and 2011, Lemma A.3, p.1219)].

In a similar spirit, KBB applies the standard "m out of n" nonparametric bootstrap, originally proposed in Bickel and Freedman (1981), to the transformed kernel-weighted data. PS demonstrate, under appropriate conditions, the large sample validity of the KBB estimator of the distribution of the sample mean [PS Theorem 3.1] and the higher order asymptotic bias and variance of the KBB variance estimator [PS Theorem 3.2]. Moreover, [PS Corollaries 3.1 and 3.2], the KBB variance estimator possesses a favourable higher order bias property, a property noted elsewhere for consistent variance estimators using tapered data [Brillinger (1981, p.151)], and, for a particular choice of kernel function weighting and choice of bandwidth, is optimal being asymptotically close to one based on the optimal quadratic spectral kernel [Andrews (1991, p.821)] or Bartlett-Priestley-Epanechnikov kernel [Priestley (1962, 1981, pp. 567-571), Epanechnikov (1969) and Sacks and Yvisacker (1981)]. Here, though, rather than being applied to the original data as in PS, the KBB kernel function weighting is applied to the individual observational components of the quasi-log likelihood criterion function itself.

Myriad variants for dependent data of the bootstrap method proposed in the landmark article Efron (1979) also make use of the standard "m out of n" nonparametric bootstrap, but, in contrast to KBB, applied to "blocks" of the original data. See, *inter alia*, the moving blocks bootstrap (MBB) [Künsch (1989), Liu and Singh (1992)], the circular block bootstrap [Politis and Romano (1992a)], the stationary bootstrap [Politis and Romano (1994)], the external bootstrap for m-dependent data [Shi and Shao (1988)], the frequency domain bootstrap [Hurvich and Zeger (1987), see also Hidalgo (2003)], and its generalization the transformation-based bootstrap [Lahiri (2003)], and the autoregressive sieve bootstrap [Buhlmann (1997)]; for further details on these methods, see, e.g., the monographs Shao and Tu (1995) and Lahiri (2003). Whereas the block length of these other methods is typically a declining fraction of sample size, the implicit KBB block length is dictated by the support of the kernel function and, thus, with unbounded support as in the optimal case, would be the sample size itself.

KBB bears comparison with the tapered block bootstrap (TBB) of Paparoditis and Politis (2001); see also Paparoditis and Politis (2002) . Indeed KBB may be regarded as a generalisation and extension of TBB. TBB is also based on a reweighted sample of the observations but with weight function with bounded support and, so, whereas each KBB data point is in general a transformation of all original sample data, those of TBB use a fixed block size and, implicitly thereby, a fixed number of data points. More generally then, the TBB weight function class is a special case of that of KBB but is more restrictive; a detailed comparison of KBB and TBB is provided in PS Section 4.1.

The paper is organized as follows. After outlining some preliminaries Section 2 introduces KBB and reviews the results in PS. Section 3 demonstrates how KBB can be applied in the quasi-maximum likelihood framework and, in particular, details the consistency of the KBB estimator and its asymptotic validity for quasi-maximum likelihood. Section 4 reports a Monte Carlo study on the performance of KBB for the mean regression model. Finally section 5 concludes. Proofs of the results in the main text are provided in Appendix B with intermediate results required for their proofs given in Appendix A.

2 Kernel Block Bootstrap

To introduce the kernel block bootstrap (KBB) method, consider a sample of T observations, $z_1, ..., z_T$, on the scalar strictly stationary real valued sequence $\{z_t, t \in \mathbb{Z}\}$ with unknown mean $\mu = \mathrm{E}[z_t]$ and autocovariance sequence $R(s) = \mathrm{E}[(z_t - \mu)(z_{t+s} - \mu)]$, $(s = 0, \pm 1, ...)$. Under suitable conditions, see Ibragimov and Linnik (1971, Theorem 18.5.3, pp. 346, 347), the limiting distribution of the sample mean $\bar{z} = \sum_{t=1}^{T} z_t/T$ is described by $T^{1/2}(\bar{z} - \mu) \xrightarrow{d} N(0, \sigma_\infty^2).$, where $\sigma_\infty^2 = \lim_{T\to\infty} \mathrm{var}[T^{1/2}\bar{z}] = \sum_{s=-\infty}^{\infty} R(s)$.

The KBB approximation to the distribution of the sample mean $\bar{z}$ randomly samples the kernel-weighted centred observations

$$z_{tT} = \frac{1}{(\hat{k}_2 S_T)^{1/2}} \sum_{r=t-T}^{t-1} k(\frac{r}{S_T})(z_{t-r} - \bar{z}), t = 1, ..., T, \tag{2.1}$$

where S_T is a bandwidth parameter, $(T = 1, 2, ...)$, $k(\cdot)$ a kernel function and $\hat{k}_j = \sum_{s=1-T}^{T-1} k(s/S_T)^j/S_T$, $(j = 1, 2)$. Let $\bar{z}_T = T^{-1}\sum_{t=1}^{T} z_{tT}$ denote the sample mean of z_{tT}, $(t = 1, ..., T)$. Under appropriate conditions, $\bar{z}_T \xrightarrow{p} 0$ and $(T/S_T)^{1/2}\bar{z}_T/\sigma_\infty \xrightarrow{d} N(0, 1)$; see, e.g., Smith (2011, Lemmas A.1 and A.2, pp.1217-19). Moreover, the KBB variance estimator, defined in standard random sampling outer product form,

$$\hat{\sigma}^2_{\text{KBB}} = T^{-1}\sum_{t=1}^{T}(z_{tT} - \bar{z}_T)^2 \xrightarrow{p} \sigma^2_\infty; \tag{2.2}$$

and is thus an automatically positive semidefinite heteroskedastic and autocorrelation consistent (HAC) variance estimator; see Smith (2011, Lemma A.3, p.1219).

KBB applies the standard "m out of n" non-parametric bootstrap method to the index set $\mathcal{T}_T = \{1, ..., T\}$; see Bickel and Freedman (1981). That is, the indices t^*_s and, thereby, $z_{t^*_s}$, $(s = 1, ..., m_T)$, are a random sample of size m_T drawn from, respectively, $\mathcal{T}_T$ and $\{z_{tT}\}_{t=1}^{T}$, where $m_T = [T/S_T]$, the integer part of T/S_T. The KBB sample mean $\bar{z}^*_{m_T} = \sum_{s=1}^{m_T} z_{t^*_s T}/m_T$ may be regarded as that from a random sample of size m_T taken from the blocks $\mathcal{B}_t = \{k\{(t-r)/S_T\}(z_r - \bar{z})/(\hat{k}_2 S_T)^{1/2}\}_{r=1}^{T}$, $(t = 1, ..., T)$. See PS Remark 2.2, p.3. Note that the blocks $\{\mathcal{B}_t\}_{t=1}^{T}$ are overlapping and, if the kernel function $k(\cdot)$ has unbounded support, the block length is T.

Let $\mathcal{P}^*_\omega$ denote the bootstrap probability measure conditional on $\{z_{tT}\}_{t=1}^{T}$ (or, equivalently, the observational data $\{z_t\}_{t=1}^{T}$) with E^* and var^* the corresponding conditional expectation and variance respectively. Under suitable regularity conditions, see PS Assumptions 3.1-3.3, pp.3-4, the bootstrap distribution of the scaled and centred KBB sample mean $m_T^{1/2}(\bar{z}^*_{m_T} - \bar{z}_T)$ converges uniformly to that of $T^{1/2}(\bar{z} - \mu)$, i.e.,

$$\sup_{x\in\mathcal{R}}\left|\mathcal{P}^*_\omega\{m_T^{1/2}(\bar{z}^*_{m_T} - \bar{z}_T) \le x\} - \mathcal{P}\{T^{1/2}(\bar{z} - \mu) \le x\}\right| \xrightarrow{p} 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P}; \tag{2.3}$$

see PS (Theorem 3.1, p.4).

Given stricter requirements, PS Theorem 3.2, p.5, provides higher order results on moments of the KBB variance estimator $\hat{\sigma}^2_{\text{KBB}}$ (2.2). Let $k^*_{(q)} = \lim_{y\to 0}\{1 - k^*(y)\}/|y|^q$, where the induced self-convolution kernel $k^*(y) = \int_{-\infty}^{\infty} k(x-y)k(x)dx/k_2$, and $MSE(T/S_T, \hat{\sigma}^2_{\text{KBB}}) = (T/S_T)E((\hat{\sigma}^2_{\text{KBB}} - J_T)^2)$, where $J_T = \sum_{s=1-T}^{T-1}(1 - |s|/T)R(s)$. Bias: $E[\hat{\sigma}^2_{\text{KBB}}] = J_T + S_T^{-2}(\Gamma_{k^*} + o(1)) + U_T$, where $\Gamma_{k^*} = -k^*_{(2)}\sum_{s=-\infty}^{\infty}|s|^2 R(s)$ and $U_T = O((S_T/T)^{b-1/2}) + o(S_T^{-2}) + O(S_T^{b-2}T^{-b}) + O(S_T/T) + O(S_T^2/T^2)$ with $b > 1$. Variance: if $S_T^5/T \to \gamma \in (0, \infty]$, then $(T/S_T)var[\hat{\sigma}^2_{\text{KBB}}] = \Delta_{k^*} + o(1)$, where $\Delta_{k^*} = 2\sigma^4_\infty\int_{-\infty}^{\infty} k^*(y)^2 dy$. Mean squared error: if $S_T^5/T \to \gamma \in (0, \infty)$, then $MSE(T/S_T, \hat{\sigma}^2_{\text{KBB}}) = \Delta_{k^*} + \Gamma^2_{k^*}/\gamma + o(1)$. The bias and variance results are similar to Parzen (1957, Theorems 5A and 5B, pp.339-340) and Andrews (1991, Proposition 1, p.825), when the Parzen exponent q equals 2. The KBB bias, cf. the tapered block bootstrap (TBB), is $O(1/S_T^2)$, an improvement on $O(1/S_T)$ for the moving block bootstrap (MBB). The expression $MSE(T/S_T, \hat{\sigma}^2_{\text{KBB}}(S_T))$ is identical to that for the mean squared error of the Parzen (1957) estimator based on the induced self-convolution kernel $k^*(y)$.

Optimality results for the estimation of σ^2_∞ are an immediate consequence of PS Theorem 3.2, p.5, and the theoretical results of Andrews (1991) for the Parzen (1957) estimator. Smith (2011, Example 2.3, p.1204) shows that the induced self-convolution kernel $k^*(y) = k^*_{\text{QS}}(y)$, where the quadratic spectral (QS) kernel

$$k^*_{\text{QS}}(y) = \frac{3}{(ay)^2}\left(\frac{\sin ay}{ay} - \cos ay\right), a = 6\pi/5, \tag{2.4}$$

if

$$k(x) = (\frac{5\pi}{8})^{1/2}\frac{1}{x}J_1(\frac{6\pi x}{5}) \text{ if } x \neq 0 \text{ and } (\frac{5\pi}{8})^{1/2}\frac{3\pi}{5} \text{ if } x = 0; \tag{2.5}$$

here $J_v(z) = \sum_{k=0}^{\infty}(-1)^k (z/2)^{2k+v} / \{\Gamma(k+1)\Gamma(k+2)\}$, a Bessel function of the first kind (Gradshteyn and Ryzhik, 1980, 8.402, p.951) with $\Gamma(\cdot)$ the gamma function. The QS kernel $k^*_{\text{QS}}(y)$ (2.4) is well-known to possess optimality properties, e.g., for the estimation of spectral densities (Priestley, 1962; 1981, pp. 567-571) and probability densities (Epanechnikov, 1969, Sacks and Yvisacker, 1981). PS Corollary 3.1, p.6, establishes a similar result for the KBB variance estimator $\tilde{\sigma}^2_{\text{KBB}}(S_T)$ computed with the kernel function (2.5). For sensible comparisons, the requisite bandwidth parameter is $S_{Tk^*} = S_T / \int_{-\infty}^{\infty} k^*(y)^2 dy$, see Andrews (1991, (4.1), p.829), if the respective asymptotic variances scaled by T/S_T are to coincide; see Andrews (1991, p.829). Note that $\int_{-\infty}^{\infty} k^*_{\text{QS}}(y)^2 dy = 1$. Then, for any bandwidth sequence $\{S_T\}$ such that $S_T \to \infty$ and $S_T^5/T \to \gamma \in (0, \infty)$, $\lim_{T\to\infty} MSE(T/S_T, \hat{\sigma}^2_{\text{KBB}}(S_{Tk^*})) - MSE(T/S_T, \tilde{\sigma}^2_{\text{KBB}}(S_T)) \geq 0$ with strict inequality if $k^*(y) \neq k^*_{\text{QS}}(y)$ with positive Lebesgue measure; see PS Corollary 3.1, p.6. Also, the bandwidth $S_T^* = (4\Gamma^2_{k^*}/\Delta_{k^*})^{1/5} T^{1/5}$ si optimal in the following sense. For any bandwidth sequence $\{S_T\}$ such that $S_T \to \infty$ and $S_T^5/T \to \gamma \in (0, \infty)$, $\lim_{T\to\infty} MSE(T^{4/5}, \hat{\sigma}^2_{\text{KBB}}(S_T)) - MSE(T^{4/5}, \hat{\sigma}^2_{\text{KBB}}(S_T^*)) \geq 0$ with strict inequality unless $S_T = S_T^* + o(1/T^{1/5})$; see PS Corollary 3.2, p.6.

3 Quasi-Maximum Likelihood

This section applies the KBB method briefly outlined above to parameter estimation in the quasi-maximum likelihood (QML) setting. In particular, under the regularity conditions detailed below, KBB may be used to construct hypothesis tests and confidence intervals. The proofs of the results basically rely on verifying a number of the conditions required for several general lemmata established in Gonçalves and White (2004) on resampling methods for extremum estimators. Indeed, although the focus of Gonçalves and White (2004) is MBB, the results therein also apply to other block bootstrap schemes such as KBB.

To describe the set-up, let the d_z-vectors z_t, $(t = 1, ..., T)$, denote a realisation from the stationary and strong mixing stochastic process $\{z_t\}_{t=1}^{\infty}$. The d_θ-vector θ of parameters is of interest where $\theta \in \Theta$ with the compact parameter space $\Theta \subseteq \mathcal{R}^{d_\theta}$. Consider the log-density $\mathcal{L}_t(\theta) = \log f(z_t; \theta)$ and its expectation $\mathcal{L}(\theta) = \mathrm{E}[\mathcal{L}_t(\theta)]$. The true value θ_0 of θ is defined by

$$\theta_0 = \arg\max_{\theta \in \Theta} \mathcal{L}(\theta)$$

with, correspondingly, the QML estimator $\hat{\theta}$ of θ

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \bar{\mathcal{L}}(\theta),$$

where the sample mean $\bar{\mathcal{L}}(\theta) = \sum_{t=1}^{T} \mathcal{L}_t(\theta)/T$.

To describe the KBB method for QML define the kernel smoothed log density function

$$\mathcal{L}_{tT}(\theta) = \frac{1}{(\hat{k}_2 S_T)^{1/2}} \sum\nolimits_{r=t-T}^{t-1} k(\frac{r}{S_T}) \mathcal{L}_{t-r}(\theta), (t = 1, ..., T),$$

cf. (2.1). As in Section 2, the indices t_s^* and the consequent bootstrap sample $\mathcal{L}_{t_s^* T}(\theta)$, $(s = 1, ..., m_T)$, denote random samples of size m_T drawn with replacement from the index set $\mathcal{T}_T = \{1, ..., T\}$ and the bootstrap sample space $\{\mathcal{L}_{tT}(\theta)\}_{t=1}^{T}$, where $m_T = [T/S_T]$ is the integer part of T/S_T. The bootstrap QML estimator $\hat{\theta}^*$ is then defined by

$$\hat{\theta}^* = \arg\max_{\theta \in \Theta} \bar{\mathcal{L}}_{m_T}^*(\theta)$$

where the bootstrap sample mean $\bar{\mathcal{L}}_{m_T}^*(\theta) = \sum_{s=1}^{m_T} \mathcal{L}_{t_s^* T}(\theta)/m_T$.

Remark 3.1. Note that, because $\mathrm{E}[\partial \mathcal{L}_t(\theta_0)/\partial\theta] = 0$, it is unnecessary to centre $\mathcal{L}_t(\theta)$, $(t = 1, ..., T)$, at $\bar{\mathcal{L}}(\theta)$; cf. (2.1).

The following conditions are imposed to establish the consistency of the bootstrap estimator $\hat{\theta}^*$ for θ_0. Let $f_t(\theta) = f(z_t; \theta)$, $(t = 1, 2, ...)$.

Assumption 3.1 **(a)** *$(\Omega, \mathcal{F}, \mathcal{P})$ is a complete probability space;* **(b)** *the finite d_z-dimensional stochastic process Z_t: $\Omega \longmapsto \mathcal{R}^{d_z}$, $(t = 1, 2, ...)$, is stationary and strong mixing with mixing numbers of size $-v/(v-1)$ for some $v > 1$ and is measurable for all t, $(t = 1, 2, ...)$.*

Assumption 3.2 **(a)** *f: $\mathcal{R}^{d_z} \times \Theta \longmapsto \mathcal{R}^+$ is $\mathcal{F}$-measurable for each $\theta \in \Theta$, Θ a compact subset of $\mathcal{R}^{d_\theta}$;* **(b)** *$f_t(\cdot)$: $\Theta \longmapsto \mathcal{R}^+$ is continuous on Θ a.s.-$\mathcal{P}$;* **(c)** *$\theta_0 \in \Theta$ is the unique maximizer of* $\mathrm{E}[\log f_t(\theta)]$, $\mathrm{E}[\sup_{\theta \in \Theta} |\log f_t(\theta)|^\alpha] < \infty$ *for some $\alpha > v$;* **(d)** *$\log f_t(\theta)$ is global Lipschitz continuous on Θ, i.e., for all $\theta, \theta^0 \in \Theta$, $|\log f_t(\theta) - \log f_t(\theta^0)| \leq L_t \|\theta - \theta^0\|$ a.s.-$\mathcal{P}$ and $\sup_T E[\sum_{t=1}^T L_t/T] < \infty$;*

Let $\mathbb{I}(x \geq 0)$ denote the indicator function, i.e., $\mathbb{I}(A) = 1$ if A true and 0 otherwise.

Assumption 3.3 **(a)** *$S_T \to \infty$ and $S_T = o(T^{\frac{1}{2}})$;* **(b)** *$k(\cdot)$: $\mathcal{R} \longmapsto [-k_{\max}, k_{\max}]$, $k_{\max} < \infty$, $k(0) \neq 0$, $k_1 \neq 0$, and is continuous at 0 and almost everywhere;* **(c)** *$\int_{-\infty}^{\infty} \bar{k}(x)dx < \infty$ where $\bar{k}(x) = \mathbb{I}(x \geq 0) \sup_{y \geq x} |k(y)| + \mathbb{I}(x < 0) \sup_{y \leq x} |k(y)|$;* **(d)** *$|K(\lambda)| \geq 0$ for all $\lambda \in \mathcal{R}$, where $K(\lambda) = (2\pi)^{-1} \int k(x) \exp(-ix\lambda)dx$.*

Theorem 3.1. Let Assumptions 3.1-3.3 hold. Then, **(a)** $\hat{\theta} - \theta_0 \to 0$, prob-$\mathcal{P}$; **(b)** $\hat{\theta}^* - \hat{\theta} \to 0$, prob-$\mathcal{P}^*$, prob-$\mathcal{P}$.

To prove consistency of the KBB distribution requires a strengthening of the above assumptions.

Assumption 3.4 **(a)** *$(\Omega, \mathcal{F}, \mathcal{P})$ is a complete probability space;* **(b)** *the finite d_z-dimensional stochastic process Z_t: $\Omega \longmapsto \mathcal{R}^{d_z}$, $(t = 1, 2, ...)$, is stationary and strong mixing with mixing numbers of size $-3v/(v-1)$ for some $v > 1$ and is measurable for all t, $(t = 1, 2, ...)$.*

Assumption 3.5 **(a)** *f: $\mathcal{R}^{d_z} \times \Theta \longmapsto \mathcal{R}^+$ is $\mathcal{F}$-measurable for each $\theta \in \Theta$, Θ a compact subset of $\mathcal{R}^{d_\theta}$;* **(b)** *$f_t(\cdot)$: $\Theta \longmapsto \mathcal{R}^+$ is continuously differentiable of order 2 on Θ a.s.-P, $(t = 1, 2, ...)$;* **(c)** *$\theta_0 \in \mathrm{int}(\Theta)$ is the unique maximizer of* $\mathrm{E}[\log f_t(\theta)]$.

Define $A(\theta) = \mathrm{E}[\partial^2 \mathcal{L}_t(\theta)/\partial\theta\partial\theta']$ and $B(\theta) = \lim_{T \to \infty} \mathrm{var}[T^{1/2} \partial \bar{\mathcal{L}}(\theta)/\partial\theta]$.

Assumption 3.6 **(a)** *$\partial^2 \mathcal{L}_t(\theta)/\partial\theta\partial\theta'$ is global Lipschitz continuous on Θ;* **(b)** $\mathrm{E}[\sup_{\theta \in \Theta} \|\partial \mathcal{L}_t(\theta)/\partial\theta\|^\alpha] < \infty$ *for some $\alpha > \max[4v, 1/\eta]$, $E[\sup_{\theta \in \Theta} \|\partial^2 \mathcal{L}_t(\theta)/\partial\theta\partial\theta'\|^\beta] < \infty$ for some $\beta > 2v$;* **(c)** *$A_0 = A(\theta_0)$ is non-singular and $B_0 = \lim_{T \to \infty} \mathrm{var}[T^{1/2} \partial \bar{\mathcal{L}}(\theta_0)/\partial\theta]$ is positive definite.*

Under these regularity conditions,

$$B_0^{-1/2} A_0 T^{1/2}(\hat{\theta} - \theta_0) \xrightarrow{d} N(0, I_{d_\theta});$$

see the Proof of Theorem 3.2. Moreover,

THEOREM 3.2. Suppose Assumptions 3.2-3.6 are satisfied. Then, if $S_T \to \infty$ and $S_T = O(T^{\frac{1}{2}-\eta})$ with $0 < \eta < \frac{1}{2}$,

$$\sup_{x \in \mathcal{R}^{d_\theta}} \left| \mathcal{P}^*_\omega \{T^{1/2}(\hat{\theta}^* - \hat{\theta})/k^{1/2} \leq x\} - \mathcal{P}\{T^{1/2}(\hat{\theta} - \theta_0) \leq x\} \right| \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P},$$

where $k = k_2/k_1^2$.

4 Simulation Results

In this section we report the results of a set of Monte Carlo experiments comparing the finite sample performance of different methods for the construction of confidence intervals for the parameters of the mean regression model when there is autocorrelation in the data. We investigate KBB, MBB and confidence intervals based on HAC covariance matrix estimators.

4.1 Design

We consider the same simulation design as that of Andrews (1991, Section 9, pp.840-849) and Andrews and Monaham (1992, Section 3, pp.956-964), i.e., linear regression with an intercept and four regressor variables. The model studied is:

$$y_t = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \beta_3 x_{3,t} + \beta_4 x_{4,t} + \sigma_t u_t, \tag{4.1}$$

where σ_t is a function of the regressors $x_{i,t}$, $(i = 1, ..., 4)$, to be specified below. The interest concerns 95% confidence interval estimators for the coefficient β_1 of the first non-constant regressor.

The regressors and error term u_t are generated as follows. First,

$$u_t = \rho u_{t-1} + \varepsilon_{0,t},$$

with initial condition $u_{-49} = \varepsilon_{0,-49}$. Let

$$\tilde{x}_{i,t} = \rho \tilde{x}_{i,t-1} + \varepsilon_{i,t}, (i = 1, ..., 4),$$

with initial conditions $\tilde{x}_{i,-49} = \varepsilon_{i,-49}$, $(i = 1, ..., 4)$. As in Andrews (1991), the innovations ε_{it}, $(i = 0, ..., 4)$, $(t = -49, ..., T)$, are independent standard normal random variates. Define $\tilde{x}_t = (\tilde{x}_{1,t}, ..., \tilde{x}_{4,t})'$ and $\bar{x}_t = \tilde{x}_t - \sum_{s=1}^{T} \tilde{x}_s/T$. The regressors $x_{i,t}$, $(i = 1, ..., 4)$, are then constructed as in

$$\begin{aligned} x_t &= (x_{1,t}, ..., x_{4,t})' \\ &= [\sum_{s=1}^{T} \bar{x}_s \bar{x}_s'/T]^{-1/2} \bar{x}_t, (t = 1, ..., T). \end{aligned}$$

The observations on the dependent variable y_t are obtained from the linear regression model (4.1) using the true parameter values $\beta_i = 0$, $(i = 0, ..., 4)$.

The values of ρ are 0, 0.2, 0.5, 0.7 and 0.9. Homoskedastic, $\sigma_t = 1$, and heteroskedastic, $\sigma_t = |x_{1t}|$, regression errors are examined. Sample sizes $T = 64$, 128 and 256 are considered.

The number of bootstrap replications for each experiment was 1000 with 5000 random samples generated. The bootstrap sample size or block size m_T was defined as $\max\{[T/S_T], 1\}$.

4.2 Bootstrap Methods

Confidence intervals based on KBB are compared with those obtained for MBB [Fitzenberger (1997), Gonçalves and White (2004)] and TBB [Paparoditis and Politis (2002)].

Bootstrap confidence intervals are commonly computed using the standard percentile [Efron (1979)], the symmetric percentile and the equal tailed [Hall (1992, p.12)] methods.[1] For succinctness only the best results are reported for each of the bootstrap methods, i.e., the standard percentile KBB and MBB methods and the equal-tailed TBB method.

To describe the standard percentile KBB method, let $\hat{\beta}_1$ denote the LS estimator of β_1 and $\hat{\beta}_1^*$ its bootstrap counterpart.Because the asymptotic distribution of the LS estimator $\hat{\beta}_1$ is normal and hence symmetric about β_1, in large samples the distributions of $\hat{\beta}_1 - \beta_1$ and $\beta_1 - \hat{\beta}_1$ are the same. From the uniform consistency of the bootstrap, Theorem 3.2, the distribution of $\beta_1 - \hat{\beta}_1$ is well approximated by the distribution of $\hat{\beta}_1^* - \hat{\beta}_1$. Therefore, the bootstrap percentile confidence interval for β_1 is given by

$$\left([1 - \frac{1}{\hat{k}^{1/2}}]\hat{\beta}_1 + \frac{\hat{\beta}_{1,0.025}^*}{\hat{k}^{1/2}}, [1 - \frac{1}{\hat{k}^{1/2}}]\hat{\beta}_1 + \frac{\hat{\beta}_{1,0.975}^*}{\hat{k}^{1/2}} \right),$$

[1]The standard percentile method is valid here as the asymptotic distribution of the least squares estimator is symmetric; see Politis (1998, p.45).

where $\hat{\beta}^*_{1,\alpha}$ is the 100α percentile of the distribution of $\hat{\beta}^*_1$ and $\hat{k} = \hat{k}_2/\hat{k}_1^2$.[2] For MBB, $\hat{k} = 1$.

TBB is applied to the sample components $[\sum_{t=1}^{T}(1, x'_t)'(1, x'_t)/T]^{-1}(1, x'_t)'\hat{\varepsilon}_t$, $(t = 1, ..., T)$, of the LS influence function, where $\hat{\varepsilon}_t$ are the LS regression residuals; see Paparoditis and Politis (2002).[3] The equal-tailed TBB confidence interval does not require symmetry of the distribution of $\hat{\beta}_1$. Thus, because the distribution of $\hat{\beta}_1 - \beta_1$ is uniformly close to that of $(\hat{\beta}^*_1 - \hat{\beta}_1)/\hat{k}$ for sample sizes large enough, the equal-tailed TBB confidence interval is given by

$$\left([1+\frac{1}{\hat{k}}]\hat{\beta}_1 - \frac{\hat{\beta}^*_{1,0.975}}{\hat{k}}, [1+\frac{1}{\hat{k}}]\hat{\beta}_1 - \frac{\hat{\beta}^*_{1,0.025}}{\hat{k}}\right).$$

KBB confidence intervals are constructed with the following choices of kernel function: truncated [TR], Bartlett [BT], (2.5) [QS] kernel functions, the last with the optimal quadratic spectral kernel (2.4) as the associated convolution, and the kernel function based on the optimal trapezoidal taper of Paparoditis and Politis (2001) [PP], see Paparoditis and Politis (2001, p1111). The respective confidence interval estimators are denoted by KBB$_{\text{J}}$, where I = TR, BT, QS and PP. TBB confidence intervals are computed using the optimal Paparoditis and Politis (2001) trapezoidal taper.

Standard t-statistic confidence intervals using heteroskedastic autocorrelation consistent (HAC) estimators for the asymptotic variance matrix are also considered based on the Bartlett, see Newey and West (1987), and quadratic spectral, see Andrews (1991), kernel functions. The respective HAC confidence intervals are denoted by BT and QS.

---

[2] Alternatively, $\tilde{k}_j$ can be replaced by k_j, where $k_j = \int_{-\infty}^{\infty} k(x)^j dx$, $(j = 1, 2)$.

[3] TBB employs a non-negative taper $w(\cdot)$ with unit interval support and range which is strictly positive in a neighbourhood of and symmetric about 1/2 and is non-decreasing on the interval $[0, 1/2]$, see Paparoditis and Politis (2001, Assumptions 1 and 2, p.1107). Hence, $w(\cdot)$ is centred and unimodal at 1/2. Given a positive integer bandwidth parameter S_T, the TBB sample space is $[\sum_{t=1}^{T}(1, x'_t)'(1, x'_t)/T]^{-1}\{S_T^{1/2}\sum_{j=1}^{S_T} w_{S_T}(j)(1, x'_{t+j-1})'\hat{\varepsilon}_{t+j-1}/\left\|w_{S_T}\right\|_2\}_{t=1}^{T-S_T+1}$, where $w_{S_T}(j) = w\{(j-1/2)/S_T\}$ and $\left\|w_{S_T}\right\|_2 = (\sum_{j=1}^{S_T} w_{S_T}(j)^2)^{1/2}$; cf. Paparoditis and Politis (2001, (3), p.1106, and Step 2, p.1107). Because $\sum_{t=1}^{T}(1, x'_t)'(1, x'_t)/T$ is the identity matrix in the Andrews (1991) design adopted here, TBB draws a random sample of size $m_T = T/S_T$ with replacement from the TBB sample space $\{S_T^{1/2}\sum_{j=1}^{S_T} w_{S_T}(j)(1, x'_{t+j-1})'\hat{\varepsilon}_{t+j-1}/\left\|w_{S_T}\right\|_2\}_{t=1}^{T-S_T+1}$. Denote the TBB sample mean $\bar{z}^*_T = \sum_{s=1}^{m_T} S_T^{1/2}\sum_{j=1}^{S_T} w_{S_T}(j)(1, x'_{t^*_s+j-1})'\hat{\varepsilon}_{t^*_s+j-1}/\left\|w_{S_T}\right\|_2 S_T m_T$ and sample mean $\bar{z}_T = \sum_{t=1}^{T-S_T+1} S_T^{1/2}\sum_{j=1}^{S_T} w_{S_T}(j)(1, x'_{t+j-1})'\hat{\varepsilon}_{t+j-1}/\left\|w_{S_T}\right\|_2 (T - S_T + 1)$.Then, from Paparoditis and Politis (2002, Theorem 2.2, p.135), $\sup_{x\in\mathcal{R}^{d_\theta}}\left|\mathcal{P}^*_\omega\{T^{1/2}(\bar{z}^*_T - \bar{z}_T) \leq x\} - \mathcal{P}\{T^{1/2}(\hat{\theta} - \theta_0) \leq x\}\right| \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. See Parente and Smith (2018, Section 4.1, pp.6-8) for a detailed comparison of TBB and KBB.

4.3 Bandwidth Choice

The accuracy of the bootstrap approximation in practice is particularly sensitive to the choice of the bandwidth or block size. Gonçalves and White (2004) suggest basing the choice of MBB block size on the automatic bandwidth obtained in Andrews (1991) for the Bartlett kernel, noting that the MBB bootstrap variance estimator is asymptotically equivalent to the Bartlett kernel variance estimator. Smith (2011, Lemma A.3, p.1219) obtained a similar equivalence between the KBB variance estimator and the corresponding HAC estimator based on the implied kernel function $k^*(\cdot)$; see also Smith (2005, Lemma 2.1, p.164). We therefore adopt a similar approach to that of Gonçalves and White (2004) to the choice of the bandwidth for the KBB confidence interval estimators, in particular, the (integer part) of the automatic bandwidth of Andrews (1991) for the implied kernel function $k^*(\cdot)$. Despite lacking a theoretical justification, the results discussed below indicate that this procedure fares well for the simulation designs studied here.

The optimal bandwidth for HAC variance matrix estimation based on the kernel $k^*(\cdot)$ is given by

$$S_T^* = \left(q k_q^2 \alpha(q) T \Big/ \int_{-\infty}^{\infty} k^*(x)^2 dx \right)^{1/(2q+1)},$$

where $\alpha(q)$ is a function of the unknown spectral density matrix and $k_q = \lim_{x \to 0}[1 - k(x)]/|x|^q$, $q \in [0, \infty)$; see Andrews (1991, Section 5, pp.830-832). Note that $q = 1$ for the Bartlett kernel and $q = 2$ for the Parzen, quadratic spectral kernels and the optimal Paparoditis and Politis (2001) taper.

The optimal bandwidth S_T^* requires the estimation of the parameters $\alpha(1)$ and $\alpha(2)$. We use the semi-parametric method recommended in Andrews (1991, (6.4), p.835) based on AR(1) approximations and using the same unit weighting scheme there. Let $z_{it} = x_{it}(y_t - x_t'\hat{\beta})$, $(i = 1, ..., 4)$. The estimators for $\alpha(1)$ and $\alpha(2)$ are given by

$$\hat{\alpha}(1) = \frac{\sum_{i=1}^{4} \frac{4\hat{\rho}_i^2 \, \hat{\sigma}_i^4}{(1-\hat{\rho}_i)^6(1+\hat{\rho}_i)^2}}{\sum_{i=1}^{4} \frac{\hat{\sigma}_i^4}{(1-\hat{\rho}_i)^4}}, \hat{\alpha}(2) = \frac{\sum_{i=1}^{4} \frac{4\hat{\rho}_i^2 \, \hat{\sigma}_i^4}{(1-\hat{\rho}_i)^8}}{\sum_{i=1}^{4} \frac{\hat{\sigma}_i^4}{(1-\hat{\rho}_i)^4}},$$

where $\hat{\rho}_i$ and $\hat{\sigma}_i^2$ are the estimators of the AR(1) coefficient and the innovation variance in a first order autoregression for z_{it}, $(i = 1, ..., 4)$. To avoid extremely large values of the bandwidth due to erroneously large values of $\hat{\rho}_i$, which tended to occur for large values of the autocorrelation coefficient ρ, we replaced $\hat{\rho}_i$ by the truncated version $\max[\min[\hat{\rho}_i, 0.97], -0.97]$.

A non-parametric version of Andrews (1991) bandwidth estimator based on flattop lag-window of Politis and Romano (1995) is also considered given by

$$\hat{\alpha}(q) = \frac{\sum_{i=1}^{4} [\sum_{j=-M_i}^{M_i} |j|^q \lambda(\frac{j}{M_i}) \hat{R}_i(j)]^2}{\sum_{i=1}^{4} [\sum_{j=-M_i}^{M_i} \lambda(\frac{j}{M_i}) \hat{R}_i(j)] \sum_{i=-M_j}^{M_j} \lambda\left(\frac{i}{M_j}\right) \hat{R}_j(i)^2},$$

where $\lambda(t) = \mathbb{I}(|t| \in [0, 1/2]) + 2(1 - |t|)\mathbb{I}(|t| \in (1/2, 1])$, $\hat{R}_i(j)$ is the sample jth autocovariance estimator for z_{it}, $(i = 1, ..., 4)$, and M_i is computed using the method described in Politis and White (2004, ftn c, p.59).

The MBB and TBB block sizes are given by $\min[\lceil \hat{S}_T^* \rceil, T]$, where $\lceil \cdot \rceil$ is the ceiling function and S_T^* the optimal bandwidth estimator for the Bartlett kernel $k^*(\cdot)$ for MBB and for the kernel $k^*(\cdot)$ induced by the optimal Paparoditis and Politis (2001) trapezoidal taper for TBB.

4.4 Results

Tables 1 and 2 provide the empirical coverage rates for 95% confidence interval estimates obtained using the methods described above for the homoskedastic and heteroskedastic cases respectively.

Tables 1 and 2 around here

Overall, to a lesser or greater degree, all confidence interval estimates display undercoverage for the true value $\beta_1 = 0$ but especially for high values of ρ, a feature found in previous studies of MBB, see, e.g., Gonçalves and White (2004), and confidence intervals based on t-statistics with HAC variance matrix estimators, see Andrews (1991). As should be expected from the theoretical results of Section 3, as T increases, empirical coverage rates approach the nominal rate of 95%.

A closer analysis of the results in Tables 1 and 2 reveals that the performance of the various methods depends critically on how the bandwidth or block size is computed. While, for low values of ρ, both the methods of Andrews (1991) and Politis and Romano (1995) produce very similar results, the Andrews (1991) automatic bandwidth yields results closer to the nominal 95% coverage for higher values of ρ. However, this is not particularly surprising since the Andrews (1991) method is based on the correct model.

A comparison of the various KBB confidence interval estimates with those using MBB reveals that generally the coverage rates for MBB are closer to the nominal 95% than those of KBB_{TR} although both are based on the truncated kernel. However, MBB usually produces coverage rates lower than those of KBB_{BT}, KBB_{QS} and KBB_{PP} especially for higher values of ρ, apart from the homoskedastic case with $T = 64$, see Table 1, when the covergae rates for MBB are very similar to those obtained for KBB_{BT} and KBB_{QS}.

The results with homoskedastic innovations in Table 1 indicate that the TBB coverage is poorer than that for KBB and MBB. In contradistinction, for heteroskedastic innovations, Table 2 indicates that TBB displays reasonable coverage properties compared with KBB_{BT}, KBB_{QS} and KBB_{PP} for the larger sample sizes $T = 128$ and 256 for all values of ρ except $\rho = 0.9$.

All bootstrap confidence interval estimates outperform those based on HAC t-statistics for higher values of ρ whereas for lower values both bootstrap and HAC t-statistic methods produce similarly satisfactory results.

Generally, one of the KBB class of confidence interval estimates, KBB_{BT}, KBB_{QS} and KBB_{PP}, outperforms any of the other methods. With homoskedastic innovations, see Table 1, the coverage rates for KBB_{PP} confidence interval estimates are closest to the nominal 95%, no matter which optimal bandwidth parameters estimation method is used; a similar finding of the robustness of KBB_{PP} to bandwidth choice is illustrated in the simulation study of Parente and Smith (2018) for the case of the mean. The results in Table 2 for heteroskedastic innovations are far more varied. As noted above, for low values of ρ, the coverage rates of KBB, TBB and HAC t-statistic confidence interval estimates are broadly similar. Indeed, the best results are obtained by KBB_{BT} for moderate values of ρ and by KBB_{PP} for high values of ρ, especially $\rho = 0.9$. Note, though, for $T = 64$, that with the Politis and Romano (1995) bandwidth estimate KBB_{QS} appears best method for low values of ρ.

4.5 Summary

In general, confidence interval estimates based on KBB_{PP} provide the best coverage rates for all sample sizes and especially for larger values of the autoregression parameter ρ.

5 Conclusion

This paper applies the kernel block bootstrap method to quasi-maximum likelihood estimation of dynamic models under stationarity and weak dependence. The proposed bootstrap method is simple to implement by first kernel-weighting the components comprising the quasi-log likelihood function in an appropriate way and then sampling the resultant transformed components using the standard "m out of n" bootstrap for independent and identically distributed observations.

We investigate the first order asymptotic properties of the kernel block bootstrap for quasi-maximum likelihood demonstrating, in particular, its consistency and the first-order asymptotic validity of the bootstrap approximation to the distribution of the quasi-maximum likelihood estimator. A set of simulation experiments for the mean regression model illustrates the efficacy of the kernel block bootstrap for quasi-maximum likelihood estimation. Indeed, in these experiments, it outperforms other bootstrap methods for the sample sizes considered, especially if the kernel function is chosen as the optimal taper suggested by Paparoditis and Politis (2001).

Appendix

Throughout the Appendices, C and Δ denote generic positive constants that may be different in different uses with C, M, and T the Chebyshev, Markov, and triangle inequalities respectively. UWL is a uniform weak law of large numbers such as Newey and McFadden (1994, Lemma 2.4, p.2129) for stationary and mixing (and, thus, ergodic) processes.

A similar notation is adopted to that in Gonçalves and White (2004). For any bootstrap statistic $T^*(\cdot,\omega)$, $T^*(\cdot,\omega) \to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$ if, for any $\delta > 0$ and any $\xi > 0$, $\lim_{T\to\infty} \mathcal{P}\{\omega : \mathcal{P}^*_\omega\{\lambda : |T^*(\lambda,\omega)| > \delta\} > \xi\} = 0$.

To simplify the analysis, the appendices consider the transformed centred observations

$$\mathcal{L}_{tT}(\theta) = \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=t-1}^{t-T} k(\frac{s}{S_T})\mathcal{L}_{t-s}(\theta)$$

with k_2 substituting for $\hat{k}_2 = S_T^{-1} \sum_{t=1-T}^{T-1} k(t/S_T)^2$ in the main text since $\hat{k}_2 - k_2 = o(1)$, cf. PS Supplement Corollary K.2, p.S.21.

For simplicity, where required, it is assumed T/S_T is integer.

Appendix A: Preliminary Lemmas

Assumption A.1. (Bootstrap Pointwise WLLN.) For each $\theta \in \Theta \subset \mathcal{R}^{d_\theta}$, Θ a compact set, $S_T \to \infty$ and $S_T = o(T^{-1/2})$

$$(k_2/S_T)^{1/2}[\bar{\mathcal{L}}^*_{m_T}(\theta) - \bar{\mathcal{L}}_T(\theta)] \to 0, \text{ prob-}\mathcal{P}^*_\omega, \text{ prob-}\mathcal{P}.$$

Remark A.1. See Lemma A.2 below.

Assumption A.2. (Uniform Convergence.)

$$\sup_{\theta \in \Theta} \left| (k_2/S_T)^{1/2} \bar{\mathcal{L}}_T(\theta) - k_1 \bar{\mathcal{L}}(\theta) \right| \rightarrow 0 \text{ prob-}\mathcal{P}.$$

Remark A.2. The hypotheses of the UWLs Smith (2011, Lemma A.1, p.1217) and Newey and McFadden (1994, Lemma 2.4, p.2129) for stationary and mixing (and, thus, ergodic) processes are satisfied under Assumptions 3.1-3.3. Hence, $\sup_{\theta \in \Theta} \left\| (k_2/S_T)^{1/2} \bar{\mathcal{L}}_T(\theta) - k_1 \mathcal{L}(\theta)] \right\| \rightarrow 0$ prob-$\mathcal{P}$ noting $\sup_{\theta \in \Theta} \left\| \bar{\mathcal{L}}(\theta) - \mathcal{L}(\theta)] \right\| \rightarrow 0$ prob-$\mathcal{P}$ where $\mathcal{L}(\theta) = \text{E}[\mathcal{L}_t(\theta)]$. Thus, Assumption A.2 follows by T and k_1, $k_2 = O(1)$.

Assumption A.3. (Global Lipschitz Continuity.) For all $\theta, \theta^0 \in \Theta$, $|\mathcal{L}_t(\theta) - \mathcal{L}_t(\theta^0)| \leq L_t \|\theta - \theta^0\|$ a.s.$\mathcal{P}$ where $\sup_T E[\sum_{t=1}^T L_t/T] < \infty$.

Remark A.3. Assumption A.3 is Assumption 3.2(c).

Lemma A.1. (Bootstrap UWL.) Suppose Assumptions A.1-A.3 hold. Then, for $S_T \rightarrow \infty$ and $S_T = o(T^{1/2})$, for any $\varepsilon > 0$ and $\delta > 0$,

$$\lim_{T \rightarrow \infty} \mathcal{P}\{\mathcal{P}^*_\omega\{\sup_{\theta \in \Theta} \left| (k_2/S_T)^{1/2} \bar{\mathcal{L}}^*_{m_T}(\theta) - k_1 \bar{\mathcal{L}}(\theta) \right| > \delta\} > \xi\} = 0.$$

Proof. From Assumption A.2 the result is proven if

$$\lim_{T \rightarrow \infty} \mathcal{P}\{\mathcal{P}^*_\omega\{\sup_{\theta \in \Theta} (k_2/S_T)^{1/2} \left| \bar{\mathcal{L}}^*_{m_T}(\theta) - \bar{\mathcal{L}}_T(\theta) \right| > \delta\} > \xi\} = 0.$$

The following preliminary results are useful in the later analysis. By global Lipschitz continuity of $\mathcal{L}_t(\cdot)$ and by T, for T large enough,

$$\begin{aligned} (k_2/S_T)^{1/2} \left| \bar{\mathcal{L}}_T(\theta)) - \bar{\mathcal{L}}_T(\theta^0) \right| &\leq \frac{1}{T} \sum_{t=1}^T \frac{1}{S_T} \sum_{s=t-T}^{t-1} \left| k\left(\frac{s}{S_T}\right) \right| \left| \mathcal{L}_{t-s}(\theta) - \mathcal{L}_{t-s}(\theta^0) \right| \quad \text{(A.1)} \\ &= \frac{1}{T} \sum_{t=1}^T \left| \mathcal{L}_t(\theta) - \mathcal{L}_t(\theta^0) \right| \left| \frac{1}{S_T} \sum_{s=1-t}^{T-t} k\left(\frac{s}{S_T}\right) \right| \\ &\leq C \left\| \theta - \theta^0 \right\| \frac{1}{T} \sum_{t=1}^T L_t \end{aligned}$$

since for some $0 < C < \infty$

$$\left| \frac{1}{S_T} \sum_{s=1-t}^{T-t} k\left(\frac{s}{S_T}\right) \right| \leq O(1) < C$$

uniformly t for large enough T, see Smith (2011, eq. (A.5), p.1218). Next, for some $0 < C^* < \infty$,

$$\begin{aligned}(k_2/S_T)^{1/2}E^*[\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}^*_{m_T}(\theta^0)\right|] &= \frac{1}{m_T}\sum_{s=1}^{m_T}\frac{1}{S_T}E^*[\sum_{r=t_s^*-T}^{t_s^*-1}\left|k\left(\frac{r}{S_T}\right)\right|\left|\mathcal{L}_{t_s^*-r}(\theta)-\mathcal{L}_{t_s^*-r}(\theta^0)\right|] \\ &= \frac{1}{T}\sum_{s=1}^{T}\left|\mathcal{L}_t(\theta)-\mathcal{L}_t(\theta^0)\right|\frac{1}{S_T}\sum_{r=t-T}^{t-1}\left|k\left(\frac{r}{S_T}\right)\right| \\ &\leq C^*\left\|\theta-\theta^0\right\|\frac{1}{T}\sum_{t=1}^{T}L_t.\end{aligned}$$

Hence, by M, for some $0 < C^* < \infty$ uniformly t for large enough T,

$$\mathcal{P}^*_\omega\{(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}^*_{m_T}(\theta^0)\right|>\epsilon\}\leq\frac{C^*}{\epsilon}\left\|\theta-\theta^0\right\|\frac{1}{T}\sum_{t=1}^{T}L_t. \tag{A.2}$$

The remaining part of the proof is identical to Gonçalves and White (2000, Proof of Lemma A.2, pp.30-31) and is given here for completeness; cf. Hall and Horowitz (1996, Proof of Lemma 8, p.913). Given $\varepsilon > 0$, let $\{\eta(\theta_i,\varepsilon),(i=1,...,I)\}$ denote a finite subcover of Θ where $\eta(\theta_i,\xi)=\{\theta\in\Theta:\|\theta-\theta_i\|<\varepsilon\}$, $(i=1,...,I)$. Now

$$\sup_{\theta\in\Theta}(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}_T(\theta)\right|=\max_{i=1,\dots,I}\sup_{\theta\in\eta(\theta_i,\xi)}(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}_T(\theta)\right|.$$

The argument $\omega\in\Omega$ is omitted for brevity as in Gonçalves and White (2000). It then follows that, for any $\delta > 0$ (and any fixed ω),

$$\mathcal{P}^*_\omega\{\sup_{\theta\in\Theta}(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}_T(\theta)\right|>\delta\}\leq\sum\nolimits_{i=1}^{I}\mathcal{P}^*_\omega\{\sup_{\theta\in\eta(\theta_i,\xi)}(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}_T(\theta)\right|>\delta\}.$$

For any $\theta\in\eta(\theta_i,\xi)$, by T and Assumption A.3,

$$\begin{aligned}(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}_T(\theta)\right| &\leq (k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta_i)-\bar{\mathcal{L}}_T(\theta_i)\right|+(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}^*_{m_T}(\theta_i)\right| \\ &\quad+(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}_T(\theta)-\bar{\mathcal{L}}_T(\theta_i)\right|.\end{aligned}$$

Hence, for any $\delta > 0$ and $\xi > 0$,

$$\begin{aligned}\mathcal{P}\{\mathcal{P}^*_\omega\{\sup_{\theta\in\eta(\theta_i\varepsilon)}(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}_T(\theta)\right|>\delta\}>\xi\} &\leq \mathcal{P}\{\mathcal{P}^*_\omega\{(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta_i)-\bar{\mathcal{L}}_T(\theta_i)\right|>\frac{\delta}{3}\}>\frac{\xi}{3}\} \\ &+\mathcal{P}\{\mathcal{P}^*_\omega\{(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T}(\theta)-\bar{\mathcal{L}}^*_{m_T}(\theta_i)\right|>\frac{\delta}{3}\}>\frac{\xi}{3}\} \\ &+\mathcal{P}\{(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}_T(\theta)-\bar{\mathcal{L}}_T(\theta_i)\right|>\frac{\delta}{3}\}.\end{aligned} \tag{A.3}$$

By Assumption A.1

$$\mathcal{P}\{\mathcal{P}_\omega^*\{(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}_{m_T}^*(\theta_i)-\bar{\mathcal{L}}_T(\theta_i)\right|>\frac{\delta}{3}\}>\frac{\xi}{3}\}<\frac{\xi}{3}$$

for large enough T. Also, by M (for fixed ω) and Assumption A.3, noting $L_t \geq 0$, $(t=1,...,T)$, from eq. (A.2),

$$\begin{aligned}\mathcal{P}_\omega^*\{(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}_{m_T}^*(\theta)-\bar{\mathcal{L}}_{m_T}^*(\theta_i)\right| &> \frac{\delta}{3}\} \leq \frac{3C^*}{\delta}\left\|\theta-\theta_i\right\|\frac{1}{T}\sum\nolimits_{t=1}^T L_t \\ &\leq \frac{3C^*\varepsilon}{\delta}\frac{1}{T}\sum\nolimits_{t=1}^T L_t\end{aligned}$$

as $T^{-1}\sum\nolimits_{t=1}^T L_t$ satisfies a WLLN under the conditions of the theorem. As a consequence, for any $\delta>0$ and $\xi>0$, for T sufficiently large,

$$\begin{aligned}\mathcal{P}\{\mathcal{P}_\omega^*\{(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}_{m_T}^*(\theta)-\bar{\mathcal{L}}_{m_T}^*(\theta_i)\right| &> \frac{\delta}{3\varepsilon}\}>\frac{\xi}{3}\}\leq\mathcal{P}\{\frac{3C^*\varepsilon}{\delta}\frac{1}{T}\sum\nolimits_{t=1}^T L_t>\frac{\xi}{3}\} \\ &= \mathcal{P}\{\frac{1}{T}\sum\nolimits_{t=1}^T L_t>\frac{\xi\delta}{9C^*\varepsilon}\} \\ &\leq \frac{9C^*\varepsilon}{\xi\delta}E[\frac{1}{T}\sum\nolimits_{t=1}^T L_t] \\ &\leq \frac{9C^*\varepsilon\Delta}{\xi\delta}<\frac{\xi}{3}\end{aligned}$$

for the choice $\varepsilon<\xi^2\delta/27C^*\Delta$, where, since, by hypothesis $E[\sum_{t=1}^T L_t/T]=O(1)$, the second and third inequalities follow respectively from M and Δ a sufficiently large but finite constant such that $\sup_T E[\sum_{t=1}^T L_t/T]<\Delta$. Similarly, from eq. (A.1), for any $\delta>0$ and $\xi>0$, by Assumption A.3, $\mathcal{P}\{(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}_T(\theta)-\bar{\mathcal{L}}_T(\theta_i)\right|>\delta/3\}\leq\mathcal{P}\{C\varepsilon\sum_{t=1}^T L_t/T>\delta/3\}\leq 3C\varepsilon\Delta/\delta<\xi/3$ for T sufficiently large for the choice $\varepsilon<\xi\delta/9C\Delta$.

Therefore, from eq. (A.3), the conclusion of the Lemma follows if

$$\varepsilon=\frac{\xi\delta}{9\Delta}\max\left(\frac{1}{C},\frac{\xi}{3C^*}\right).\blacksquare$$

Lemma A.2. (Bootstrap Pointwise WLLN.) Suppose Assumptions 3.1, 3.2(a) and 3.3 are satisfied. Then, if $T^{1/\alpha}/m_T\to 0$ and $\mathrm{E}[\sup_{\theta\in\Theta}\left|\log f_t(\theta)\right|^\alpha]<\infty$ for some $\alpha>v$, for each $\theta\in\Theta\subset\mathcal{R}^{d_\theta}$, Θ a compact set,

$$(k_2/S_T)^{1/2}[\bar{\mathcal{L}}_{m_T}^*(\theta)-\bar{\mathcal{L}}_T(\theta)]\to 0,\text{ prob-}\mathcal{P}_\omega^*,\text{ prob-}\mathcal{P}.$$

Proof: The argument θ is suppressed throughout for brevity. First, cf. Gonçalves and White (2004, Proof of Lemma A.5, p.215),

$$(k_2/S_T)^{1/2}(\bar{\mathcal{L}}_{m_T}^*-\bar{\mathcal{L}}_T)=(k_2/S_T)^{1/2}(\bar{\mathcal{L}}_{m_T}^*-\mathrm{E}^*[\bar{\mathcal{L}}_{m_T}^*])-(k_2/S_T)^{1/2}(\bar{\mathcal{L}}_T-\mathrm{E}^*[\bar{\mathcal{L}}_{m_T}^*]).$$

Since $\mathrm{E}^*[\bar{\mathcal{L}}^*_{m_T}] = \bar{\mathcal{L}}_T$, cf. PS (Section 2.2, pp.2-3), the second term $\bar{\mathcal{L}}_T - \mathrm{E}^*[\bar{\mathcal{L}}^*_{m_T}]$ is zero. Hence, the result follows if, for any $\delta > 0$ and $\xi > 0$ and large enough T, $\mathcal{P}\{\mathcal{P}^*_\omega\{(k_2/S_T)^{1/2}\left|\bar{\mathcal{L}}^*_{m_T} - \mathrm{E}^*[\bar{\mathcal{L}}^*_{m_T}]\right| > \delta\} > \xi\} < \xi$.

Without loss of generality, set $\mathrm{E}^*[\bar{\mathcal{L}}^*_{m_T}] = 0$. Write $\mathcal{K}_{tT} = (k_2/S_T)^{1/2}\mathcal{L}_{tT}$, $(t = 1, ..., T)$. First, note that

$$\begin{aligned}\mathrm{E}^*[\left|\mathcal{K}_{t^*_sT}\right|] &= \frac{1}{T}\sum\nolimits_{t=1}^{T}|\mathcal{K}_{tT}| = \frac{1}{T}\sum\nolimits_{t=1}^{T}\left|\frac{1}{S_T}\sum\nolimits_{s=t-T}^{t-1}k(\frac{s}{S_T})\mathcal{L}_{t-s}\right| \\ &\leq O(1)\frac{1}{T}\sum\nolimits_{t=1}^{T}|\mathcal{L}_t| = O_p(1),\end{aligned}$$

uniformly, $(s = 1, ..., m_T)$, by WLLN and $\mathrm{E}[\sup_{\theta\in\Theta}|\log f_t(\theta)|^\alpha] < \infty$, $\alpha > 1$. Also, for any $\delta > 0$,

$$\begin{aligned}\frac{1}{T}\sum\nolimits_{t=1}^{T}|\mathcal{K}_{tT}| - \frac{1}{T}\sum\nolimits_{t=1}^{T}|\mathcal{K}_{tT}|\,\mathbb{I}(|\mathcal{K}_{tT}| &< m_T\delta) = \frac{1}{T}\sum\nolimits_{t=1}^{T}|\mathcal{K}_{tT}|\,\mathbb{I}(|\mathcal{K}_{tT}| \geq m_T\delta) \\ &\leq \frac{1}{T}\sum\nolimits_{t=1}^{T}|\mathcal{K}_{tT}|\max_t \mathbb{I}(|\mathcal{K}_{tT}| \geq m_T\delta).\end{aligned}$$

Now, by M,

$$\max_t|\mathcal{K}_{tT}| = O(1)\max_t|\mathcal{L}_t| = O_p(T^{1/\alpha});$$

cf. Newey and Smith (2004, Proof of Lemma A1, p.239). Hence, since, by hypothesis, $T^{1/\alpha}/m_T = o(1)$, $\max_t \mathbb{I}(|\mathcal{K}_{tT}| \geq m_T\delta) = o_p(1)$ and $\sum_{t=1}^{T}|\mathcal{K}_{tT}|/T = O_p(1)$,

$$\mathrm{E}^*[\left|\mathcal{K}_{t^*_sT}\right|\mathbb{I}(\left|\mathcal{K}_{t^*_sT}\right| \geq m_T\delta)] = \frac{1}{T}\sum\nolimits_{t=1}^{T}|\mathcal{K}_{tT}|\,\mathbb{I}(|\mathcal{K}_{tT}| \geq m_T\delta) = o_p(1). \tag{A.4}$$

The remaining part of the proof is similar to that for Khinchine's WLLN given in Rao (1973, pp.112-144). For each s define the pair of random variables

$$V_{t^*_sT} = \mathcal{K}_{t^*_sT}\mathbb{I}(\left|\mathcal{K}_{t^*_sT}\right| < m_T\delta), W_{t^*_sT} = \mathcal{K}_{t^*_sT}\mathbb{I}(\left|\mathcal{K}_{t^*_sT}\right| \geq m_T\delta),$$

yielding $\mathcal{K}_{t^*_sT} = V_{t^*_sT} + W_{t^*_sT}$, $(s = 1, ..., m_T)$. Now

$$\mathrm{var}^*[V_{t^*_sT}] \leq \mathrm{E}^*[V^2_{t^*_sT}] \leq m_T\delta\mathrm{E}^*[\left|V_{t^*_sT}\right|]. \tag{A.5}$$

Write $\bar{V}^*_{m_T} = \sum\nolimits_{s=1}^{m_T}V_{t^*_sT}/m_T$. Thus, from eq. (A.5), using C,

$$\begin{aligned}\mathcal{P}^*\{\left|\bar{V}^*_{m_T} - \mathrm{E}^*[V_{t^*_sT}]\right| &\geq \varepsilon\} \leq \frac{\mathrm{var}^*[V_{t^*_sT}]}{m_T\varepsilon^2} \\ &\leq \frac{\delta\mathrm{E}^*[\left|V_{t^*_sT}\right|]}{\varepsilon^2}.\end{aligned}$$

Also $\left|\bar{\mathcal{K}}_T - \mathrm{E}^*[V_{t_s^*T}]\right| = o_p(1)$, i.e., for any $\varepsilon > 0$, T large enough, $\left|\bar{\mathcal{K}}_T - \mathrm{E}^*[V_{t_s^*T}]\right| \leq \varepsilon$, since by T, noting $\mathrm{E}^*[V_{t_s^*T}] = \sum_{t=1}^{T} \mathcal{K}_{tT}\mathbb{I}(\,|\mathcal{K}_{tT}| < m_T\delta)/T$,

$$\begin{aligned}\left|\bar{\mathcal{K}}_T - \mathrm{E}^*[V_{t_s^*T}]\right| &= \left|\frac{1}{T}\sum_{t=1}^{T}\mathcal{K}_{tT} - \frac{1}{T}\sum_{t=1}^{T}\mathcal{K}_{tT}\mathbb{I}(|\mathcal{K}_{tT}| < m_T\delta)\right| \\ &\leq \frac{1}{T}\sum_{t=1}^{T}|\mathcal{K}_{tT}|\,\mathbb{I}(|\mathcal{K}_{tT}| \geq m_T\delta) = o_p(1)\end{aligned}$$

from eq. (A.4). Hence, for T large enough,

$$\mathcal{P}^*\{\left|\bar{V}^*_{m_T} - \bar{\mathcal{K}}_T\right| \geq 2\varepsilon\} \leq \frac{\delta \mathrm{E}^*[\left|V_{t_s^*T}\right|]}{\varepsilon^2}. \tag{A.6}$$

By M,

$$\begin{aligned}\mathcal{P}^*\{W_{t_s^*T} &\neq 0\} = \mathcal{P}^*\{\left|\mathcal{K}_{t_s^*T}\right| \geq m_T\delta\} \\ &\leq \frac{1}{m_T\delta}\mathrm{E}^*[\left|\mathcal{K}_{t_s^*T}\right|\mathbb{I}(\left|\mathcal{K}_{t_s^*T}\right| \geq m_T\delta)] \leq \frac{\delta}{m_T}.\end{aligned} \tag{A.7}$$

To see this, $\mathrm{E}^*[\left|\mathcal{K}_{t_s^*T}\right|\mathbb{I}(\left|\mathcal{K}_{t_s^*T}\right| \geq m_T\delta)] = o_p(1)$ from eq. (A.4). Thus, for T large enough, $\mathrm{E}^*[\left|\mathcal{K}_{t_s^*T}\right|\mathbb{I}(\left|\mathcal{K}_{t_s^*T}\right| \geq m_T\delta)] \leq \delta^2$ w.p.a.1. Write $\bar{W}^*_{m_T} = \sum_{s=1}^{m_T} W_{t_s^*T}/m_T$. Thus, from eq. (A.7),

$$\mathcal{P}^*\{\bar{W}^*_{m_T} \neq 0\} \leq \sum_{s=1}^{m_T}\mathcal{P}^*\{W_{t_s^*T} \neq 0\} \leq \delta. \tag{A.8}$$

Therefore,

$$\begin{aligned}\mathcal{P}^*\{\left|\bar{\mathcal{K}}^*_{m_T} - \bar{\mathcal{K}}_T\right| &\geq 4\varepsilon\} \leq \mathcal{P}^*\{\left|\bar{V}^*_{m_T} - \bar{\mathcal{K}}_T\right| + \left|\bar{W}^*_{m_T}\right| \geq 4\varepsilon\} \\ &\leq \mathcal{P}^*\{\left|\bar{V}^*_{m_T} - \bar{\mathcal{K}}_T\right| \geq 2\varepsilon\} + \mathcal{P}^*\{\left|\bar{W}^*_{m_T}\right| \geq 2\varepsilon\} \\ &\leq \frac{\delta E^*[\left|\bar{V}_{t_s^*T}\right|]}{\varepsilon^2} + \mathcal{P}^*\{\left|\bar{W}^*_{m_T}\right| \neq 0\} \leq \frac{\delta E^*[\left|V_{t_s^*T}\right|]}{\varepsilon^2} + \delta.\end{aligned}$$

where the first inequality follows from T, the third from eq. (A.6) and the final inequality from eq. (A.8). Since δ may be chosen arbitrarily small enough and $\mathrm{E}^*[\left|V_{t_s^*T}\right|] \leq \mathrm{E}^*[\left|\mathcal{K}_{t_s^*T}\right|] = O_p(1)$, the result follows by M.■

Lemma A.3. Let Assumptions 3.2(a)(b), 3.3, 3.4 and 3.6(b)(c) hold. Then, if $S_T \to \infty$ and $S_T = O(T^{\frac{1}{2}-\eta})$ with $0 < \eta < \frac{1}{2}$,

$$\sup_{x\in\mathcal{R}}\left|\mathcal{P}^*_\omega\{m_T^{1/2}(\frac{\partial\bar{\mathcal{L}}^*_{m_T}(\theta_0)}{\partial\theta} - \frac{\partial\bar{\mathcal{L}}_T(\theta_0)}{\partial\theta}) \leq x\} - \mathcal{P}\{T^{1/2}\frac{\partial\bar{\mathcal{L}}(\theta_0)}{\partial\theta} \leq x\}\right| \to 0, \text{ prob-}\mathcal{P}.$$

PROOF. The result is proven in Steps 1-5 below; cf. Politis and Romano (1992, Proof of Theorem 2, pp. 1993-5). To ease exposition, let $m_T = T/S_T$ be integer and $d_\theta = 1$.

STEP 1. $d\bar{\mathcal{L}}(\theta_0)/d\theta \to 0$ prob-$\mathcal{P}$. Follows by White (1984, Theorem 3.47, p.46) and $\mathrm{E}[\partial\mathcal{L}_t(\theta_0)/\partial\theta] = 0$.

STEP 2. $\mathcal{P}\{B_0^{-1/2}T^{1/2}d\bar{\mathcal{L}}(\theta_0)/d\theta \leq x\} \to \Phi(x)$, where $\Phi(\cdot)$ is the standard normal distribution function. Follows by White (1984, Theorem 5.19, p.124).

STEP 3. $\sup_x \left|\mathcal{P}\{B_0^{-1/2}T^{1/2}d\bar{\mathcal{L}}(\theta_0)/d\theta \leq x\} - \Phi(x)\right| \to 0$. Follows by Pólya's Theorem (Serfling, 1980, Theorem 1.5.3, p.18) from Step 2 and the continuity of $\Phi(\cdot)$.

STEP 4. $\mathrm{var}^*[m_T^{1/2}d\bar{\mathcal{L}}^*_{m_T}(\theta_0)/d\theta] \to B_0$ prob-$\mathcal{P}$. Note $\mathrm{E}^*[d\bar{\mathcal{L}}^*_{m_T}(\theta_0)/d\theta] = d\bar{\mathcal{L}}_T(\theta_0)/d\theta$. Thus,

$$\begin{aligned}
\mathrm{var}^*[m_T^{1/2}\frac{d\bar{\mathcal{L}}^*_{m_T}(\theta_0)}{d\theta}] &= \mathrm{var}^*[\frac{d\mathcal{L}_{t^*T}(\theta_0)}{d\theta}] \\
&= \frac{1}{T}\sum_{t=1}^{T}(\frac{d\mathcal{L}_{tT}(\theta_0)}{d\theta} - \frac{d\bar{\mathcal{L}}_T(\theta_0)}{d\theta})^2 \\
&= \frac{1}{T}\sum_{t=1}^{T}(\frac{d\mathcal{L}_{tT}(\theta_0)}{d\theta})^2 - (\frac{d\bar{\mathcal{L}}_T(\theta_0)}{d\theta})^2.
\end{aligned}$$

the result follows since $(d\bar{\mathcal{L}}_T(\theta_0)/d\theta)^2 = O_p(S_T/T)$ (Smith, 2011, Lemma A.2, p.1219), $S_T = o(T^{1/2})$ by hypothesis and $T^{-1}\sum_{t=1}^{T}(d\mathcal{L}_{tT}(\theta_0)/d\theta)^2 \to B_0$ prob-$\mathcal{P}$ (Smith, 2011, Lemma A.3, p.1219).

STEP 5.

$$\lim_{T\to\infty}\mathcal{P}\left\{\sup_x\left|\mathcal{P}^*_\omega\{\frac{d\bar{\mathcal{L}}^*_{m_T}(\theta_0)/d\theta - \mathrm{E}^*[d\bar{\mathcal{L}}^*_{m_T}(\theta_0)/d\theta]}{\mathrm{var}^*[d\bar{\mathcal{L}}^*_{m_T}(\theta_0)/d\theta]^{1/2}} \leq x\} - \Phi(x)\right| \geq \varepsilon\right\} = 0.$$

Applying the Berry-Esséen inequality, Serfling (1980, Theorem 1.9.5, p.33), noting the bootstrap sample observations $\{d\mathcal{L}_{t_s^*T}(\theta_0)/d\theta\}_{s=1}^{m_T}$ are independent and identically distributed,

$$\begin{aligned}
\sup_x\left|\mathcal{P}^*_\omega\{\frac{m_T^{1/2}(d\bar{\mathcal{L}}^*_{m_T}(\theta_0)/d\theta - d\bar{\mathcal{L}}_T(\theta_0)/d\theta)}{\mathrm{var}^*[m_T^{1/2}d\bar{\mathcal{L}}^*_{m_T}(\theta_0)/d\theta]^{1/2}} \leq x\} - \Phi(x)\right| &\leq \frac{C}{m_T^{1/2}}\mathrm{var}^*[\frac{d\mathcal{L}_{t^*T}(\theta_0)}{d\theta}]^{-3/2} \\
&\quad \times\mathrm{E}^*[\left|\frac{d\mathcal{L}_{t^*T}(\theta_0)}{d\theta} - \frac{d\bar{\mathcal{L}}_T(\theta_0)}{d\theta}\right|^3].
\end{aligned}$$

Now $\mathrm{var}^*[d\mathcal{L}_{t^*T}(\theta_0)/d\theta] \to B_0 > 0$ prob-$\mathcal{P}$; see the Proof of Step 4 above. Furthermore, $\mathrm{E}^*[\left|d\mathcal{L}_{t^*T}(\theta_0)/d\theta - d\bar{\mathcal{L}}_T(\theta_0)/d\theta\right|^3] = T^{-1}\sum_{t=1}^{T}\left|d\mathcal{L}_{tT}(\theta_0)/d\theta - d\bar{\mathcal{L}}_T(\theta_0)/d\theta\right|^3$ and

$$\begin{aligned}
\frac{1}{T}\sum_{t=1}^{T}\left|\frac{d\mathcal{L}_{tT}(\theta_0)}{d\theta} - \frac{d\bar{\mathcal{L}}_T(\theta_0)}{d\theta}\right|^3 &\leq \max_t\left|\frac{d\mathcal{L}_{tT}(\theta_0)}{d\theta} - \frac{d\bar{\mathcal{L}}_T(\theta_0)}{d\theta}\right|\frac{1}{T}\sum_{t=1}^{T}(\frac{d\mathcal{L}_{tT}(\theta_0)}{d\theta} - \frac{d\bar{\mathcal{L}}_T(\theta_0)}{d\theta})^2 \\
&= O_p(S_T^{1/2}T^{1/\alpha}).
\end{aligned}$$

The equality follows since

$$\begin{aligned}\max_t \left|\frac{d\mathcal{L}_{tT}(\theta_0)}{d\theta} - \frac{d\bar{\mathcal{L}}_T(\theta_0)}{d\theta}\right| &\leq \max_t \left|\frac{d\mathcal{L}_{tT}(\theta_0)}{d\theta}\right| + \left|\frac{d\bar{\mathcal{L}}_T(\theta_0)}{d\theta}\right| \\ &= O_p(S_T^{1/2}T^{1/\alpha}) + O_p((S_T/T)^{1/2}) = O_p(S_T^{1/2}T^{1/\alpha})\end{aligned}$$

by M and Assumption 3.6(b), cf. Newey and Smith (2004, Proof of Lemma A1, p.239), and $\sum_{t=1}^{T}(d\mathcal{L}_{tT}(\theta_0)/d\theta - d\bar{\mathcal{L}}_T(\theta_0)/d\theta)^2/T = O_p(1)$, see the Proof of Step 4 above. Therefore

$$\begin{aligned}\sup_x \left|\mathcal{P}_\omega^*\{\frac{(T/S_T)^{1/2}(d\bar{\mathcal{L}}_{m_T}^*(\theta_0)/d\theta - d\bar{\mathcal{L}}_T(\theta_0)/d\theta)}{\text{var}^*[(T/S_T)^{1/2}d\bar{\mathcal{L}}_{m_T}^*(\theta_0)/d\theta]^{1/2}} \leq x\} - \Phi(x)\right| &\leq \frac{1}{m_T^{1/2}}O_p(1)O_p(S_T^{1/2}T^{1/\alpha}) \\ &= \frac{S_T^{1/2}}{m_T^{1/2}}O_p(T^{1/\alpha}) = o_p(1),\end{aligned}$$

by hypothesis, yielding the required conclusion.■

Lemma A.4. Suppose that Assumptions 3.2(a)(b), 3.3, 3.4 and 3.6(b)(c) hold. Then, if $S_T \to \infty$ and $S_T = O(T^{\frac{1}{2}-\eta})$ with $0 < \eta < \frac{1}{2}$,

$$(k_2/S_T)^{1/2}\frac{\partial\bar{\mathcal{L}}_T(\theta_0)}{\partial\theta} = k_1\frac{\partial\bar{\mathcal{L}}(\theta_0)}{\partial\theta} + o_p(T^{-1/2}).$$

Proof. Cf. Smith (2011, Proof of Lemma A.2, p.1219). Recall

$$(k_2/S_T)^{1/2}\frac{\partial\bar{\mathcal{L}}_T(\theta_0)}{\partial\theta} = \frac{1}{S_T}\sum_{r=1-T}^{T-1} k\left(\frac{r}{S_T}\right)\frac{1}{T}\sum_{t=\max[1,1-r]}^{\min[T,T-r]}\frac{\partial\mathcal{L}_t(\theta_0)}{\partial\theta}.$$

The difference between $\sum_{t=\max[1,1-r]}^{\min[T,T-r]}\partial\mathcal{L}_t(\theta_0)/\partial\theta$ and $\sum_{t=1}^{T}\partial\mathcal{L}_t(\theta_0)/\partial\theta$ consists of $|r|$ terms. By C, using White (1984, Lemma 6.19, p.153),

$$\begin{aligned}\mathcal{P}\{\frac{1}{T}\left|\sum_{t=1}^{|r|}\frac{\partial\mathcal{L}_t(\theta_0)}{\partial\theta}\right| &\geq \varepsilon\} \leq \frac{1}{(T\varepsilon)^2}\text{E}[\left|\sum_{t=1}^{|r|}\frac{\partial\mathcal{L}_t(\theta_0)}{\partial\theta}\right|^2] \\ &= |r|\,O(T^{-2})\end{aligned}$$

where the $O(T^{-2})$ term is independent of r. Therefore, using Smith (2011, Lemma C.1, p.1231),

$$
\begin{aligned}
(k_2/S_T)^{1/2}\frac{\partial\bar{\mathcal{L}}_T(\theta_0)}{\partial\theta} &= \frac{1}{S_T}\sum_{r=1-T}^{T-1}k\left(\frac{r}{S_T}\right)(\frac{\partial\bar{\mathcal{L}}(\theta_0)}{\partial\theta}+|r|\,O_p(T^{-2}))\\
&= \frac{1}{S_T}\sum_{s=1-T}^{T-1}k\left(\frac{s}{S_T}\right)\frac{\partial\bar{\mathcal{L}}(\theta_0)}{\partial\theta}+O_p(T^{-1})\\
&= (k_1+o(1))\frac{\partial\bar{\mathcal{L}}(\theta_0)}{\partial\theta}+O_p(T^{-1})\\
&= k_1\frac{\partial\bar{\mathcal{L}}(\theta_0)}{\partial\theta}+o_p(T^{-1/2}).\blacksquare
\end{aligned}
$$

Appendix B: Proofs of Results

Proof of Theorem 3.1. Theorem 3.1 follows from a verification of the hypotheses of Gonçalves and White (2004, Lemma A.2, p.212). To do so, replace n by T, $Q_T(\cdot,\theta)$ by $\bar{\mathcal{L}}(\theta)$ and $Q_T^*(\cdot,\omega,\theta)$ by $\bar{\mathcal{L}}_{m_T}^*(\omega,\theta)$. Conditions (a1)-(a3), which ensure $\hat{\theta}-\theta_0\to 0$, prob-$P$, hold under Assumptions 3.1 and 3.2. To establish $\hat{\theta}^*-\hat{\theta}\to 0$, prob-$\mathcal{P}^*$, prob-$P$, Conditions (b1) and (b2) follow from Assumption 3.1 whereas Condition (b3) is the bootstrap UWL Lemma A.1 which requires Assumption 3.3.$\blacksquare$

Proof of Theorem 3.2. The structure of the proof is identical to that of Gonçalves and White (2004, Theorem 2.2, pp.213-214) for MBB requiring the verification of the hypotheses of Gonçalves and White (2004, Lemma A.3, p.212) which together with Pólya's Theorem (Serfling, 1980, Theorem 1.5.3, p.18) and the continuity of $\Phi(\cdot)$ gives the result.

Assumptions 3.2-3.4 ensure Theorem 3.1, i.e., $\hat{\theta}^*-\hat{\theta}\to 0$, prob-$\mathcal{P}^*$, prob-$\mathcal{P}$, and $\hat{\theta}-\theta_0\to 0$. The assumptions of the complete probability space $(\Omega,\mathcal{F},\mathcal{P})$ and compactness of Θ are stated in Assumptions 3.4(a) and 3.5(a). Conditions (a1) and (a2) follow from Assumptions 3.5(a)(b). Condition (a3) $B_0^{-1/2}T^{1/2}\partial\bar{\mathcal{L}}(\theta_0)/\partial\theta\xrightarrow{d}N(0,I_{d_\theta})$ is satisfied under Assumptions 3.4, 3.5(a)(b) and 3.6(b)(c) using the CLT White (1984, Theorem 5.19, p.124); cf. Step 4 in the Proof of Lemma A.3 above. The continuity of $A(\theta)$ and the UWL Condition (a4) $\sup_{\theta\in\Theta}\left\|\partial^2\bar{\mathcal{L}}(\theta)/\partial\theta\partial\theta'-A(\theta)\right\|\to 0$, prob-$\mathcal{P}$, follow since the hypotheses of the UWL Newey and McFadden (1994, Lemma 2.4, p.2129) for stationary and mixing (and, thus, ergodic) processes are satisfied under Assumptions 3.4-3.6. Hence, invoking Assumption 3.6(c), from a mean value expansion of $\partial\bar{\mathcal{L}}(\hat{\theta})/\partial\theta=0$ around $\theta=\theta_0$ with $\theta_0\in\text{int}(\Theta)$ from Assumption 3.5(c), $T^{1/2}(\hat{\theta}-\theta_0)\xrightarrow{d}N(0,A_0^{-1}B_0A_0^{-1})$.

Conditions (b1) and (b2) are satisfied under Assumptions 3.5(a)(b) as above. To verify Condition (b3),

$$
\begin{aligned}
m_T^{1/2}\frac{\partial\bar{\mathcal{L}}^*_{m_T}(\hat{\theta})}{\partial\theta} &= m_T^{1/2}(\frac{\partial\bar{\mathcal{L}}^*_{m_T}(\theta_0)}{\partial\theta}-\frac{\partial\bar{\mathcal{L}}_T(\theta_0)}{\partial\theta}) \\
&\quad +m_T^{1/2}\frac{\partial\bar{\mathcal{L}}_T(\theta_0)}{\partial\theta}+m_T^{1/2}(\frac{\partial\bar{\mathcal{L}}^*_{m_T}(\hat{\theta})}{\partial\theta}-\frac{\partial\bar{\mathcal{L}}^*_{m_T}(\theta_0)}{\partial\theta}).
\end{aligned}
$$

With Lemma A.3 replacing Gonçalves and White (2002, Theorem 2.2(ii), p.1375), the first term converges in distribution to $N(0,B_0)$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. The sum of the second and third terms converges to 0, prob-$\mathcal{P}^*$, prob-$\mathcal{P}$. To see this, first, using the mean value theorem for the third term, i.e.,

$$
m_T^{1/2}(\frac{\partial\bar{\mathcal{L}}^*_{m_T}(\hat{\theta})}{\partial\theta}-\frac{\partial\bar{\mathcal{L}}^*_{m_T}(\theta_0)}{\partial\theta})=\frac{1}{S_T^{1/2}}\frac{\partial^2\bar{\mathcal{L}}^*_{m_T}(\dot{\theta})}{\partial\theta\partial\theta'}T^{1/2}(\hat{\theta}-\theta_0),
$$

where $\dot{\theta}$ lies on the line segment joining $\hat{\theta}$ and θ_0. Secondly, $(k_2/S_T)^{1/2}\partial^2\bar{\mathcal{L}}^*_{m_T}(\dot{\theta})/\partial\theta\partial\theta'\to k_1A_0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, using the bootstrap UWL $\sup_{\theta\in\Theta}(k_2/S_T)^{1/2}\left\|\partial^2\bar{\mathcal{L}}^*_{m_T}(\theta)/\partial\theta\partial\theta'-\partial^2\bar{\mathcal{L}}_T(\theta)/\partial\theta\partial\theta'\right\|\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, cf. Lemma A.1, and the UWL $\sup_{\theta\in\Theta}\left\|(k_2/S_T)^{1/2}\partial^2\bar{\mathcal{L}}_T(\theta)/\partial\theta\partial\theta'-k_1A(\theta)\right\|\to 0$, prob-$\mathcal{P}$, cf. Remark A.2. Condition (b3) then follows since $T^{1/2}(\hat{\theta}-\theta_0)+A_0^{-1}T^{1/2}\partial\bar{\mathcal{L}}(\theta_0)/\partial\theta\to 0$, prob-$\mathcal{P}$, and $m_T^{1/2}\partial\bar{\mathcal{L}}_T(\theta_0)/\partial\theta-(k_1/k_2^{1/2})T^{1/2}\partial\bar{\mathcal{L}}(\theta_0)/\partial\theta\to 0$, prob-$\mathcal{P}$, cf. Lemma A.4. Finally, Condition (b4) $\sup_{\theta\in\Theta}\left\|(k_2/S_T)^{1/2}[\partial^2\bar{\mathcal{L}}^*_{m_T}(\theta)/\partial\theta\partial\theta'-\partial^2\bar{\mathcal{L}}_T(\theta)/\partial\theta\partial\theta']\right\|\to 0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$, is the bootstrap UWL Lemma A.1 appropriately revised using Assumption 3.6.

Because $\hat{\theta}\in\text{int}(\Theta)$ from Assumption 3.5(c), from a mean value expansion of the first order condition $\partial\bar{\mathcal{L}}^*_{m_T}(\hat{\theta}^*)/\partial\theta=0$ around $\theta=\hat{\theta}$,

$$
T^{1/2}(\hat{\theta}^*-\hat{\theta})=[(k_2/S_T)^{1/2}\frac{\partial^2\bar{\mathcal{L}}^*_{m_T}(\dot{\theta})}{\partial\theta\partial\theta'}]^{-1}m_T^{1/2}\frac{\partial\bar{\mathcal{L}}^*_{m_T}(\hat{\theta})}{\partial\theta},
$$

where $\dot{\theta}$ lies on the line segment joining $\hat{\theta}^*$ and $\hat{\theta}$. Noting $\hat{\theta}^*-\hat{\theta}\to 0$, prob-$\mathcal{P}^*$, prob-$\mathcal{P}$, and $\hat{\theta}-\theta_0\to 0$, prob-$\mathcal{P}$, $(k_2/S_T)^{1/2}\partial^2\bar{\mathcal{L}}^*_{m_T}(\dot{\theta})/\partial\theta\partial\theta'\to k_1A_0$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$. Therefore, $T^{1/2}(\hat{\theta}^*-\hat{\theta})$ converges in distribution to $N(0,(k_2/k_1^2)A_0^{-1}B_0A_0^{-1})$, prob-$\mathcal{P}^*_\omega$, prob-$\mathcal{P}$.■

References

Anatolyev, S. (2005): "GMM, GEL, Serial Correlation, and Asymptotic Bias," *Econometrica*, 73, 983-1002

Andrews, D.W.K. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817-858.

Andrews, D.W.K., and Monahan, J.C. (1992): "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator," *Econometrica*, 60, 1992, 953-966.

Atkinson, K.E. (1989): *An Introduction to Numerical Analysis* (2nd ed.). New York: Wiley.

Beran, R. (1988): "Prepivoting Test Statistics: a Bootstrap View of Asymptotic Refinements," *Journal of the American Statistical Association*, 83, 687-697.

Bickel, P., and Freedman, D. (1981): "Some Asymptotic Theory of the Bootstrap," *Annals of Statistics*, 9, 1196-1217.

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

Bühlmann, P. (1997): "Sieve Bootstrap for Time Series," *Bernoulli*, 3, 123-148.

Cattaneo, M.D., Crump, R.K., and Jansson, M. (2010): "Bootstrapping Density-Weighted Average Derivatives," *CREATES Research Paper* 2010-23, School of Economics and Management, University of Aarhus.

Cohn, D L. (1980): *Measure Theory.* Boston: Birkhäuser.

Davidson, J.E.H. (1994): *Stochastic Limit Theory.* Oxford: Oxford University Press.

Davidson, J., and de Jong, R. (2000): "Consistency of Kernel Estimators of Heteroscedastic and Autocorrelated Covariance Matrices," *Econometrica*, 68, 407-424.

Doukhan, P. (1994): *Mixing: Properties and Examples.* Lecture Notes in Statistics 85. New York: Springer.

Efron, B. (1979): "Bootstrap Methods: Another Look at the Jackknife," *Annals of Statistics*, 7, 1-26.

Epamechnikov, V.A. (1969): "Non-parametric Estimation of a Multivariate Probability Density," Theory of Probability and Its Applications, 14, 153-158.

Fitzenberger, B. (1997): "The Moving Blocks Bootstrap and Robust Inference for Linear Least Squares and Quantile Regressions," *Journal of Econometrics*, 82, 235-287.

Gallant, R. (1987): *Nonlinear Statistical Models.* New York: Wiley.

Gonçalves, S., and White, H. (2000): "Maximum Likelihood and the Bootstrap for Nonlinear Dynamic Models," *University of California at San Diego, Economics Working Paper Series 2000-32*, Department of Economics, UC San Diego.

Gonçalves, S., and White, H. (2001): "The Bootstrap of the Mean for Dependent Heterogeneous Arrays," *CIRANO working paper, 2001s-19.*

Gonçalves, S., and White, H. (2002): "The Bootstrap of the Mean for Dependent Heterogeneous Arrays," *Econometric Theory*, 18, 1367-1384.

Gonçalves, S., and White, H. (2004): "Maximum Likelihood and the Bootstrap for Nonlinear Dynamic Models," *Journal of Econometrics*, 119, 199-219.

Hahn, J. (1997): "A Note on Bootstrapping Generalized Method of Moments Estimators," *Econometric Theory*, 12, 187-197.

Hall, P. (1992): *The Bootstrap and Edgeworth Expansion.* New York: Springer.

Hall, P., and Horowitz, J. (1996): "Bootstrap Critical Values for Tests Based on Generalized-Method-of-Moments Estimators," *Econometrica*, 64, 891-916.

Hansen, B. (1992): "Consistent Covariance Matrix Estimation for Dependent Heterogeneous Processes," *Econometrica*, 60, 967-972.

Hidalgo, J. (2003): "An Alternative Bootstrap to Moving Blocks for Time Series Regression Models," *Journal of Econometrics*, 117, 369-399.

Horowitz, J. (2001): "The Bootstrap." In Heckman, J.J., and Leamer, E. (Eds.), *Handbook of Econometrics*, Vol. 5, 3159-3228. Amsterdam:. Elsevier.

Hurvich, C., and Zeger, S. (1987): "Frequency Domain Bootstrap Methods for Time Series," *Statistics and Operation Research Working Paper, New York University*, New York.

Kitamura, Y. (1997): "Empirical Likelihood Methods with Weakly Dependent Processes," *Annals of Statistics*, 25, 2084-2102.

Kitamura, Y., and Stutzer, M. (1997): "An Information-Theoretic Alternative to Generalized Method of Moments Estimation," *Econometrica*, 65, 861-874.

Künsch, H. (1989): "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, 17, 1217–1241.

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

Liu, R., and Singh, K. (1992): “Moving Blocks Jackknife and Bootstrap Capture Weak Dependence.” In LePage, R., and Billard L. (Eds.), *Exploring the Limits of Bootstrap*, 224-248. New York: Wiley.

Machado, J., and Parente, P. (2005): “Bootstrap Estimation of Covariance Matrices via the Percentile Method,” *Econometrics Journal*, 8, 70-78.

Newey, W.K. (1991): “Uniform Convergence in Probability and Stochastic Equicontinuity,” *Econometrica*, 59, 1161-1167.

Newey, W.K. and Smith, R.J. (2004): “Higher Order Properties of GMM and Generalized Empirical Likelihood Estimators,” *Econometrica*, 72, 219-255.

Newey, W.K. and West, K. (1987): “A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix,” *Econometrica*, 55, 703-708.

Owen, A. (1988): “Empirical Likelihood Ratio Confidence Intervals for a Single Functional,” *Biometrika*, 75, 237-249.

Owen, A. (2001): *Empirical Likelihood.* New York: Chapman and Hall.

Paparoditis, E., and Politis, D.N. (2001): “Tapered Block Bootstrap,” *Biometrika*, 88, 1105-1119.

Paparoditis, E., and Politis, D.N. (2002): “The Tapered Block Bootstrap for General Statistics from Stationary Sequences,” *Econometrics Journal*, 5, 131-148.

Parente, P.M.D.C., and Smith, R.J. (2018): “Kernel Block Bootstrap,” CWP 48/18, Centre for Microdata Methods and Practice, U.C.L and I.F.S.

Politis, D.N. (1998): “Computer-Intensive Methods in Statistical Analysis,” *IEEE Signal Processing Magazine*, 15, No. 1, 39-55, January.

Politis, D.N., and Romano, J. (1992a): “A Circular Block-Resampling Procedure for Stationary Data.” In LePage, R., and Billard L. (Eds.), *Exploring the Limits of Bootstrap*, 263-270. New York: Wiley.

Politis, D.N., and Romano, J. (1992b): “A General Resampling Scheme for Triangular Arrays of α-Mixing Random Variables with Application to the Problem of Spectral Density Estimation,” *Annals of Statistics*, 20, 1985-2007.

Politis, D.N., and Romano, J. (1994): "The Stationary Bootstrap," *Journal of the American Statististical Association*, 89, 1303-1313.

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

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

Priestley, M.B. (1962): "Basic Considerations in the Estimation of Spectra," Technometrics, 4, 551-5154,

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

Qin, J., and J. Lawless (1994): "Empirical Likelihood and General Estimating Equations," *Annals of Statistics*, 22, 300-325.

Rao, C.R. (1973): *Linear Statistical Inference and its Applications.* Second Edition. New York: Wiley.

Sacks, J., and Ylvisacker, D. (1981): "Asymptotically Optimum Kernels for Density Estimation at a Point," *Annals of Statistics*, 9, 334-346.

Serfling, R. (2002): *Approximation Theorems of Mathematical Statistics.* New York: Wiley.

Shao, J. and Tu, D. (1995): *The Jackknife and Bootstrap.* New York: Springer.

Shi, X., and Shao, J. (1988): "Resampling Estimation when the Observations are m-Dependent," *Communications in Statistics*, A, 17, 3923-3934.

Singh, K.N. (1981): "On the Asymptotic Accuracy of Efron's Bootstrap," *Annals of Statistics*, 9, 1187-1195.

Singh, K.N., and Xie, M. (2010): "Bootstrap - A Statistical Method." In *International Encyclopedia for Education*, Third Edition, 46-51. Amsterdam: Elsevier.

Smith, R.J. (1997): "Alternative Semi-parametric Likelihood Approaches to Generalised Method of Moments Estimation," *Economic Journal*, 107, 503-19.

Smith, R. J. (2005): "Automatic Positive Semidefinite HAC Covariance Matrix and GMM Estimation," *Econometric Theory*, 21, 158-170.

Smith. R.J. (2011): "GEL Criteria for Moment Condition Models," *Econometric Theory*, 27, 1192-1235.

Tasaki, H. (2009): "Convergence Rates of Approximate Sums of Riemann Integrals," *Journal of Approximation Theory*, 161, 477-490.

White, H. (1999): *Asymptotic Theory for Econometricians.* Academic Press: New York.

Table 1. Empirical Coverage Rates: Nominal 95% Confidence Intervals. Homoskedastic Innovations.

		Andrews (1991)					Politis and Romano (1995)				
T	ρ	0	0.2	0.5	0.7	0.9	0	0.2	0.5	0.7	0.9
64	KBB_{TR}	93.24	91.98	87.60	83.84	71.94	93.12	91.90	87.78	81.72	67.82
	KBB_{BT}	93.76	92.78	89.68	87.90	80.44	93.08	92.18	90.58	85.48	73.46
	KBB_{QS}	93.74	92.68	90.42	89.42	84.92	93.24	92.44	90.22	85.16	74.92
	KBB_{PP}	94.20	93.46	91.80	91.14	87.48	93.90	93.02	91.76	88.24	79.94
	MBB	93.48	92.80	89.40	87.64	79.30	93.24	92.70	90.36	85.52	73.74
	TBB	91.56	90.38	85.96	80.38	59.26	90.74	90.12	86.44	79.00	59.64
	BT	93.30	91.94	86.64	80.26	62.62	92.92	91.46	86.80	78.50	59.68
	QS	93.02	91.86	87.28	81.62	64.16	92.54	91.34	87.84	80.10	61.98
128	KBB_{TR}	93.88	93.14	90.50	87.88	79.84	94.00	92.48	90.10	85.44	76.22
	KBB_{BT}	94.48	94.30	92.12	90.10	85.32	94.46	93.28	91.90	87.54	81.24
	KBB_{QS}	94.00	93.64	92.06	91.14	88.12	94.38	93.14	91.28	88.02	82.42
	KBB_{PP}	94.08	94.12	92.82	92.14	90.10	94.62	93.18	91.76	89.66	85.72
	MBB	94.02	93.78	91.42	89.20	84.76	94.30	92.88	90.64	87.48	80.60
	TBB	92.92	92.68	90.34	86.92	72.50	93.14	91.96	90.10	85.44	71.76
	BT	93.88	93.28	89.40	86.16	72.50	94.20	92.60	89.24	84.04	70.14
	QS	93.82	93.52	90.58	87.32	74.10	93.98	92.60	90.22	85.24	72.38
256	KBB_{TR}	95.04	93.28	91.96	89.66	85.16	94.68	93.86	91.12	89.86	83.72
	KBB_{BT}	95.36	94.14	92.80	91.02	88.12	95.26	94.84	92.18	90.96	86.62
	KBB_{QS}	94.78	93.70	92.68	91.38	89.74	95.06	94.32	92.00	90.94	87.50
	KBB_{PP}	94.84	94.00	92.90	91.90	90.86	94.98	94.28	92.14	91.66	89.36
	MBB	94.76	93.82	92.10	90.26	86.26	95.02	94.12	91.62	90.72	85.50
	TBB	94.50	93.56	92.36	90.04	82.08	94.48	94.00	91.82	90.02	81.44
	BT	94.86	93.50	91.60	88.66	80.70	94.94	93.94	90.84	89.02	79.24
	QS	94.94	93.70	92.28	89.62	82.44	94.88	94.28	91.48	89.88	80.90

Table 2. Empirical Coverage Rates: Nominal 95% Confidence Intervals. Heteroskedastic Innovations.

		Andrews					Politis and Romano				
T	ρ	0	0.2	0.5	0.7	0.9	0	0.2	0.5	0.7	0.9
	KBB_{TR}	91.72	89.94	86.78	80.34	67.70	90.52	89.80	85.36	78.42	63.94
	KBB_{BT}	92.24	90.56	88.98	84.34	74.92	90.58	90.62	87.12	82.28	69.30
	KBB_{QS}	92.20	90.56	88.96	85.32	77.48	90.78	90.90	87.48	82.04	69.76
64	KBB_{PP}	92.02	89.98	89.08	85.52	78.66	90.32	90.52	87.80	82.60	72.56
	MBB	91.72	90.20	88.10	83.32	73.04	90.62	90.40	86.70	81.22	68.54
	TBB	90.88	88.88	85.94	79.54	61.18	88.78	88.44	85.10	77.54	58.78
	BT	91.42	89.62	85.48	77.14	60.46	90.24	89.30	84.40	75.00	56.10
	QS	91.18	89.36	86.42	79.70	62.96	89.52	89.04	85.14	77.10	59.02
	KBB_{TR}	92.36	91.96	88.20	85.18	75.62	92.00	92.06	88.50	83.40	72.96
	KBB_{BT}	92.76	92.80	90.04	87.10	80.38	92.46	92.48	90.72	85.38	77.44
	KBB_{QS}	92.40	92.28	89.96	87.72	82.32	92.08	92.16	90.20	85.20	77.68
128	KBB_{PP}	91.98	92.08	89.76	87.88	83.30	91.90	92.08	90.02	85.74	79.70
	MBB	92.12	92.20	89.20	86.14	78.32	92.14	92.04	89.72	84.80	75.64
	TBB	92.46	92.34	89.88	85.76	72.20	91.90	91.94	90.24	84.48	70.80
	BT	92.34	92.04	87.84	83.06	69.08	92.32	92.08	87.96	81.10	67.20
	QS	92.16	91.98	88.94	85.18	71.70	91.98	92.06	89.28	82.76	69.24
	KBB_{TR}	93.94	93.40	90.30	88.70	82.14	93.72	93.02	90.86	87.84	81.44
	KBB_{BT}	94.56	94.48	91.62	90.10	84.96	94.06	93.82	91.86	89.36	83.64
	KBB_{QS}	94.10	93.94	91.54	90.10	86.02	93.72	93.50	91.62	89.06	84.36
256	KBB_{PP}	94.16	93.70	91.40	90.22	86.70	93.56	93.16	91.74	89.26	84.96
	MBB	94.14	93.64	90.40	89.02	82.84	93.64	93.28	91.28	88.64	82.32
	TBB	94.42	94.00	91.70	90.06	81.54	94.22	93.46	92.24	89.54	81.08
	BT	94.08	93.58	89.88	88.04	77.82	94.02	93.16	90.68	86.78	77.72
	QS	94.20	93.74	90.92	89.32	79.82	93.92	93.36	91.40	88.04	79.74