Kernel block bootstrap

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

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP48/18

An ESRC Research Centre

Kernel Block Bootstrap*

Paulo M.D.C. Parente
ISEG- Lisbon School of Economics & Management, Universidade de Lisboa
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 Draft: July 2018

Abstract

Abstract

This article introduces and investigates the properties of a new bootstrap method for time-series data, the kernel block bootstrap. The bootstrap method, although akin to, offers an improvement over the tapered block bootstrap of Paparoditis and Politis (2001), admitting kernels with unbounded support. Given a suitable choice of kernel, a kernel block bootstrap estimator of the spectrum at zero asymptotically close to the optimal Parzen (1957) estimator is possible. The paper shows the large sample validity of the kernel block bootstrap and derives the higher order bias and variance of the kernel block bootstrap variance estimator. Like the tapered block bootstrap variance estimator, the kernel block bootstrap estimator has a favourable higher order bias property. Simulations based on the designs of Paparoditis and Politis (2001) indicate that the kernel block bootstrap may be efficacious in practice.

JEL Classification: C14, C15, C22
Keywords: Bias; Confidence Interval; Resampling; Kernel function; Spectral density estimation; Time Series; Variance estimation.

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

1 Introduction

Since its introduction in the landmark article Efron (1979), the bootstrap has become a standard tool for empirical research in statistics. This paper proposes and investigates the properties of a novel bootstrap method, the kernel block bootstrap, appropriate for stationary and weakly dependent data which, importantly, admits an optimal bootstrap method.

The kernel block bootstrap applies the standard non-parametric bootstrap for randomly sampled observations to a kernel function-based weighted transformation of the original data. It generalises the tapered block bootstrap of Paparoditis and Politis (2001) relaxing their requirements on the taper function, in particular, that of bounded support. Critically, therefore, a bootstrap variance estimator asymptotically close to one based on the optimal quadratic spectral (Andrews, 1991, p.821) or Bartlett-Priestley-Epanechnikov kernel (Priestley 1962, 1981, pp. 567-571, Epanechnikov, 1969, and Sacks and Yvisacker, 1981) is possible. The kernel block bootstrap variance estimator also possesses a favourable higher order bias property similar to that for the tapered block bootstrap, a property noted elsewhere for consistent variance estimators using tapered data (Brillinger, 1981, p.151).

Additionally, the paper links some of the extant results on bootstrap variance estimation. Politis and Romano (1994) show that the moving blocks bootstrap variance estimator (K\"{u}nch, 1989, and Liu and Singh, 1992) is approximately equivalent to the Bartlett kernel variance estimator in large samples. Similarly, Paparoditis and Politis (2001) show the tapered block bootstrap variance estimator is asymptotically close to a Parzen (1957) variance estimator constructed using a particular kernel function which is the self-convolution of a unimodal, non-negative taper function with bounded support. Because the kernel function whose self-convolution is the quadratic spectral kernel is admissible, a kernel block bootstrap variance estimator that closely approximates the optimal Parzen (1957) estimator is possible; see, for example, Andrews (1991).

After outlining some preliminaries Section 2 introduces the kernel block bootstrap. Section 3 details the theoretical results, describing the assumptions, the large sample validity of the kernel block bootstrap estimator of the distribution of the sample mean and the higher order asymptotic bias and variance of the kernel block bootstrap variance estimator. Optimality issues relating to the choice of kernel function and bandwidth parameter are also examined. Section 4 contrasts and compares the kernel and tapered block bootstraps and briefly discusses connections with the literature on consistent variance matrix estimation. A simulation study reported in Section 5 compares the kernel and tapered block bootstraps using designs employed in Paparoditis and Politis (2001) and indicates that the kernel block bootstrap may be efficacious in practice. Proofs of the results in the text are provided in the Supplementary Material.

2 Kernel Block Bootstrap

2.1 Some Preliminaries

The set-up closely follows that of Paparoditis and Politis (2001) in which a sample of T observations, $X_1, ..., X_T$, is available on the scalar strictly stationary real valued sequence $\{X_t, t \in \mathbb{Z}\}$ having unknown mean $\mu = E(X_t)$ and autocovariance sequence $R(s) = E((X_t - \mu)(X_{t+s} - \mu))$. The principal objective of the paper is the provision of an efficacious interval estimate for μ.

The approach here uses a kernel block bootstrap approximation to the distribution of the sample mean $\bar{X} = T^{-1} \sum_{t=1}^{T} X_t$. Recall that the asymptotic distribution of the centred and scaled statistic $T^{1/2} \left(\bar{X} - \mu\right)$ is normal with mean 0 and variance $\sigma_\infty^2 = \lim_{T\to\infty} var(T^{1/2}\bar{X}) = \sum_{s=-\infty}^{\infty} R(s)$ if $E(|X_t|^{2+\delta}) < \infty$ and $\sum_{k=1}^{\infty} \alpha_X^{\delta/(2+\delta)}(k) < \infty$ for some $\delta > 0$, (Ibragimov and Linnik, 1971, Theorem 18.5.3, pp. 346, 347). Hence, estimation of σ_∞^2 is of critical importance. The strong mixing coefficients $\alpha_X(k) = \sup_{A,B} |\text{pr}(A \cap B) - \text{pr}(A)\,\text{pr}(B)|$ with A and B events in the σ-algebras $\mathcal{F}_{-\infty}^{0}$ and $\mathcal{F}_{k}^{\infty}$ are the σ-algebras generated by $\{X_t, t \leq 0\}$ and $\{X_t, t \geq k\}$ respectively; see Rosenblatt (1985, pp. 62, 73).

The kernel block bootstrap samples the kernel-weighted centred observations

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

where S_T is a bandwidth parameter, $T = 1, 2, ...$, $k(x)$ a kernel function and $\hat{k}_2 = \sum_{s=1-T}^{T-1} k(s/S_T)^2/S_T$; see also Paparoditis and Politis (2001, Step 2, p.1107).

Remark 2.1. The definition of X_{tT} (2.1) rescales that in Kitamura and Stutzer (1997) and Smith (1997, 2011) by $S_T^{1/2}$ with k_2 replaced without loss by $\hat{k}_2$, see Corollary K.2 in the Supplementary Material, where $k_j = \int_{-\infty}^{\infty} k(x)^j dx$, $j = 1, 2$. The scale normalisation $k_1 = 1$ is imposed throughout.

The sample mean of X_{tT}, $t = 1, ..., T$, is denoted $\bar{X}_T = T^{-1} \sum_{t=1}^{T} X_{tT}$. Under assumptions stated in Section 3, $\bar{X}_T$ is weakly consistent for 0 and $(T/S_T)^{1/2}\bar{X}_T/\sigma_\infty$ converges in distribution to a standard normal variate; see, for example, Smith (2011, Lemmas A.1 and A.2, pp.1217-19). Moreover, the kernel block bootstrap variance estimator, defined in standard random sampling outer product form,

$$\hat{\sigma}_{\text{KBB}}^2 = T^{-1} \sum_{t=1}^{T} (X_{tT} - \bar{X}_T)^2 \tag{2.2}$$

is weakly consistent for σ_∞^2 (Smith, 2011, Lemma A.3, p.1219) and is automatically non-negative.

2.2 Kernel Block Bootstrap

The kernel block bootstrap 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^*, $s = 1, ..., m_T$, are a random sample of size m_T drawn from $\mathcal{T}_T$, where $m_T = [T/S_T]$, the integer part of T/S_T.

The kernel block bootstrap sample mean is

$$\bar{X}^*_{m_T} = \frac{1}{m_T}\sum_{s=1}^{m_T} X_{t_s^* T}$$
$$= \frac{1}{m_T}\sum_{s=1}^{m_T}\frac{1}{(\hat{k}_2 S_T)^{1/2}}\sum_{r=t_s^*-T}^{t_s^*-1} k(\frac{r}{S_T})(X_{t_s^*-r} - \bar{X}). \quad (2.3)$$

Remark 2.2. The reformulation (2.3) emphasises the block-wise nature of the kernel block bootstrap since $\bar{X}^*_{m_T}$ is the sample mean from a random sample of size m_T taken from the blocks $\mathcal{B}_t = \{k\{(t-r)/S_T\}(X_r - \bar{X})/(\hat{k}_2 S_T)^{1/2}\}_{r=1}^T$, $t = 1, ..., T$. Note that the blocks $\{\mathcal{B}_t\}_{t=1}^T$ are overlapping and, if $k(x)$ has unbounded support, the block length is T.

The probability measure conditional on X_{tT}, $t = 1, ..., T$, or, equivalently, the data $X_1, ..., X_T$, is denoted by pr^* with E^* and var^* the corresponding conditional expectation and variance respectively. Therefore, since the bootstrap sample $X_{t_s^* T}$, $s = 1, ..., m_T$, is a random sample of size m_T drawn from the sample space $\{X_{tT}\}_{t=1}^T$ with each sample point X_{tT}, $t = 1, ..., T$, having equal probability $1/T$, it is immediate that $E^*(\bar{X}^*_{m_T}) = \bar{X}_T$ and $m_T var^*(\bar{X}^*_{m_T}) = \hat{\sigma}^2_{\text{KBB}}$.

3 Theoretical Results

3.1 Assumptions

Assumptions 3.1-3.3 below are adaptations of Smith (2011, Assumptions 2.1, 2.2 and 2.3(d)(e), pp.1199-1200) and are sufficient for the uniform convergence of the kernel block bootstrap distribution to its asymptotic counterpart.

Assumption 3.1. The process $\{X_t, t \in Z\}$ is a scalar stationary and strong mixing process with mixing coefficients satisfying $\sum_{k=1}^\infty k^2 \alpha_X(k)^{\delta/(\delta+4)} < \infty$ for some $\delta > 0$.

Assumption 3.1 is weaker than that in Paparoditis and Politis (2001, Theorem 2, p.1108). Noting $\alpha_X(k) \in [0, 1/4]$ (Doukhan, 1994, Remark 1, p. 4), the condition $\sum_{k=1}^\infty k^2 \alpha_X(k)^{\delta/(\delta+1)} < \infty$ holds and is sufficient for the asymptotic validity of the kernel block bootstrap given in Theorem 3.1 below. The stricter Assumption 3.1 is required for the higher order results of Theorem 3.2.

Let $\mathbb{I}(A)$ be the indicator function, that is, $\mathbb{I}(A) = 1$ if A true and 0 otherwise.

Assumption 3.2. **(a)** $S_T \to \infty$ and $S_T = O(T^{\frac{1}{2}-\eta})$ with $0 < \eta < \frac{1}{2}$; **(b)** $k : \mathcal{R} \longmapsto [-k_{\max}, k_{\max}]$, $k_{\max} < \infty$, $k(0) \neq 0$, $k_2 \neq 0$, 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)** $K(\lambda) \neq 0$ for all $\lambda \in \mathcal{R}$ where $K(\lambda) = (2\pi)^{-1} \int_{-\infty}^{\infty} k(x) \exp(-ix\lambda) dx$.

Assumption 3.3. **(a)** $E(|X_t|^{\alpha}) < \Delta < \infty$ for some $\alpha > \max(4(\delta+1), 1/\eta)$; **(b)** σ_{∞}^2 is positive and finite.

Assumptions 3.2(b)(c) ensure that $k_2 > 0$ and, moreover, guarantee that the induced self-convolution kernel $k^*(y) = \int_{-\infty}^{\infty} k(x-y)k(x)dx/k_2$ is a member of the positive semi-definite class $\mathcal{K}_2$ of symmetric kernels (Andrews, 1991, p.822) used for consistent covariance matrix estimation (Smith, 2011, Lemma C.3, p.1234), that is, $\sum_{s=1-T}^{T-1} k^*(s/S_T)\hat{R}_T(s)$ where the sample autocovariance $\hat{R}_T(s) = \sum_{r=\max[1,1-s]}^{\min[T,T-s]} (X_{r+s} - \bar{X})(X_r - \bar{X})/T$. Assumption 3.2(c) ensures that certain normalised sums defined in terms of the kernel $k(x)$ converge appropriately to their integral representation counterparts; see Jansson (2002). Assumption 3.1 together with Assumption 3.3(a) ensures $\{X_t - \mu, t \in \mathbb{Z}\}$ satisfies the hypotheses of Andrews (1991, Lemma 1, p.824). Assumptions 3.1 and 3.3 also guarantee that a central limit theorem of Ibragimov and Linnik (1971) holds.

Let $k^{(j)}(x) = d^j k(x)/dx^j$, $j = 1, 2$. The following assumption on $k(x)$ is needed for results on the higher order bias and variance of $\hat{\sigma}_{\text{KBB}}^2$ (2.2).

Assumption 3.4. **(a)** $k(x)$ is twice continuously differentiable; **(b)** $k^{(j)}(x) \in L^1(\mathcal{R})$ and $\sup_x \left|k^{(j)}(x)\right| < \infty$, $j = 1, 2$; **(c)** $\lim_{|x| \to \infty} k^{(j)}(x) = 0$, $j = 1, 2$; **(d)** $|k(x)| \leq C_k |x|^{-b}$ for some $b > 1$ and some $C_k < \infty$.

Remark 3.1. To allow the kernel functions to have unbounded support, Assumption 3.4(d) imposes a rate of decay on the tails of $k(x)$ implying, in particular, that $\lim_{|x| \to \infty} k(x) = 0$. This extension is important because, as described in Section 3.4, a kernel function with unbounded support is optimal in a particular sense.

3.2 Large Sample Validity

Theorem 3.1 details the uniform convergence of the bootstrap distribution of the scaled and centred kernel block bootstrap sample mean $m_T^{1/2}(\bar{X}_{m_T}^* - \bar{X}_T)$ to the limiting distribution of $T^{1/2}\left(\bar{X} - \mu\right)$.

Theorem 3.1. *Let Assumptions 3.1-3.3 hold. Then, in probability,*

$$\sup_{x \in \mathcal{R}} \left| \text{pr}^*\{m_T^{1/2}(\bar{X}_{m_T}^* - \bar{X}_T) \leq x\} - \text{pr}\{T^{1/2}(\bar{X} - \mu) \leq x\} \right| \to 0. \tag{3.1}$$

Theorem 3.1 mirrors Paparoditis and Politis (2001, Theorem 3, (4), p.1108) for the tapered block bootstrap. Alternatively, (Paparoditis and Politis, 2001, (5), p.1108), $\sup_{x \in \mathcal{R}} |\text{pr}^*\{m_T^{1/2}\bar{X}_{m_T}^* \leq x\} - \text{pr}\{T^{1/2}(\bar{X} -$

$\mu) \leq x\}| \to 0$ in probability but, being only approximately zero in mean, is likely to be less accurate than (3.1); see Paparoditis and Politis (2001, p.1108).

3.3 Asymptotic Bias and Variance

Define $k^*_{(q)} = \lim_{y\to 0} \{1 - k^*(y)\} / |y|^q$ and let $MSE(T/S_T, \hat{\sigma}^2_{\text{KBB}}) = (T/S_T)E((\hat{\sigma}^2_{\text{KBB}} - J_T)^2)$, where $J_T = \sum_{s=1-T}^{T-1}(1 - |s|/T)R(s)$. The following theorem provides the higher order bias, variance and mean-squared error of the kernel block bootstrap variance estimator $\hat{\sigma}^2_{\text{KBB}}$ (2.2).

Theorem 3.2. *Let Assumptions 3.1-3.4 hold and* $k^*_{(2)} \in [0,\infty)$, $\sum_{s=-\infty}^{\infty} |s|^2 R(s) \in [0,\infty)$. *Then* **(a)** $E[\hat{\sigma}^2_{\text{KBB}}] = J_T + S_T^{-2}(\Gamma_{k^*} + o(1)) + U_T$, $\Gamma_{k^*} = -k^*_{(2)} \sum_{s=-\infty}^{\infty} |s|^2 R(s)$, $U_T = O((S_T/T)^{b-1/2}) + o(S_T^{-2}) + O(S_T^{b-2}T^{-b}) + O(S_T/T) + O(S_T^2/T^2)$; **(b)** *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* $\Delta_{k^*} = 2\sigma^4_\infty \int_{-\infty}^{\infty} k^*(y)^2 dy$; **(c)** *if* $S_T^5/T \to \gamma \in (0,\infty)$, *then* $MSE(T/S_T, \hat{\sigma}^2_{\text{KBB}}) = \Delta_{k^*} + \Gamma^2_{k^*}/\gamma + o(1)$.

Theorem 3.2(a)(b) are results 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; see also Paparoditis and Politis (2001, Theorems 1 and 2, pp.1107, 1108). The bandwidth parameter S_T corresponds to the block length B for the moving and tapered block bootstraps; see Section 4.1 below. Hence, the bias of the kernel block bootstrap estimator is of order $O(1/S_T^2)$, a rate identical to that of the tapered block bootstrap but faster than $O(1/S_T)$ for the moving block bootstrap. The variance of the kernel block bootstrap variance estimator is $O(T/S_T)$ coinciding with that for both methods.

3.4 Optimality

To indicate explicitly its dependence on the bandwidth parameter S_T, the kernel block bootstrap variance estimator $\hat{\sigma}^2_{\text{KBB}}$ is now written as $\hat{\sigma}^2_{\text{KBB}}(S_T)$. Theorem 3.2(c) shows that 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 $k^*(y) = \int_{-\infty}^{\infty} k(x-y)k(x)dx/k_2$; see also Andrews (1991, Proposition 1, p.825). The optimality results presented here are an immediate consequence of Theorem 3.2(c) and the theoretical results of Andrews (1991) for the Parzen (1957) estimator. The first result concerns the choice of optimal kernel, while the second discusses the optimal bandwidth parameter.

The quadratic spectral or Bartlett-Priestley-Epanechnikov kernel is

$$k^*_{\text{QS}}(y) = \frac{3}{(ay)^2}\left(\frac{\sin ay}{ay} - \cos ay\right) \tag{3.2}$$

where $a = 6\pi/5$. The kernel (3.2) is well-known to possess optimality properties for the estimation of spectral densities (Priestley, 1962; 1981, pp. 567-571) and probability densities (Epanechnikov, 1969, Sacks and Yvisacker, 1981). The induced self-convolution kernel $k^*(y) = k^*_{\text{QS}}(y)$ if

$$k(x) = (\frac{5\pi}{8})^{1/2}\frac{1}{x}J_1(\frac{6\pi x}{5}) \text{ if } x \neq 0 \text{ and } (\frac{5\pi}{8})^{1/2}\frac{3\pi}{5} \text{ if } x = 0, \tag{3.3}$$

(Smith, 2011, Example 2.3, p.1204), where $J_v(z) = \sum_{k=0}^{\infty}(-1)^k\,(z/2)^{2k+v}\,/\,\{\Gamma(k+1)\Gamma(k+2)\}$, a Bessel function of the first kind (Gradshteyn and Ryzhik, 1980, 8.402, p.951) with $\Gamma(\cdot)$ the gamma function.

Let $\tilde{\sigma}^2_{\text{KBB}}(S_T)$ denote the kernel block bootstrap variance estimator computed with the kernel function (3.3). Lemma K.5 in the Supplementary Material verifies that (3.3) satisfies Assumptions 3.2 and 3.4.

Since kernel functions are not subject to any normalisation, in any comparison dissimilar results will be obtained for identical kernel functions with arguments scaled differently. Hence, to provide a valid comparison of estimators of σ^2_∞, bandwidth parameters are chosen to ensure that the respective asymptotic variances scaled by T/S_T of Theorem 3.2(b) coincide; see Andrews (1991, p.829). Hence, for $k^*(y)$, the requisite bandwidth parameter is $S_{Tk^*} = S_T/\int_{-\infty}^{\infty} k^*(y)^2 dy$; see Andrews (1991, (4.1), p.829). Note that $\int_{-\infty}^{\infty} k^*_{\text{QS}}(y)^2 dy = 1$.

Corollary 3.1 follows from Theorem 3.2(c) and Andrews (1991, Theorem 2, p.829).

Corollary 3.1. *Let Assumptions 3.1-3.4, $k^*_{(2)} \in [0,\infty)$ and $\sum_{s=-\infty}^{\infty} |s|^2 R(s) \in (0,\infty)$ hold. Then, for any sequence of bandwidth parameters $\{S_T\}$* such that $S_T \to \infty$ and $S_T^5/T \to \gamma \in (0,\infty)$, *the kernel* (3.3) *is preferred to any other kernel function satisfying Assumptions 3.2 and 3.4 in the sense that* $\lim_{T\to\infty} MSE(T/S_T, \hat{\sigma}^2_{\text{KBB}}(S_{Tk^*})) - MSE(T/S_T, \tilde{\sigma}^2_{\text{KBB}}(S_T)) \geq 0$. *The inequality is strict if $k^*(y) \neq k^*_{\text{QS}}(y)$ with positive Lebesgue measure.*

Let $S_T^* = (4\Gamma^2_{k^*}/\Delta_{k^*})^{1/5}T^{1/5}$. Corollary 3.2 is an immediate consequence of Theorem 3.2(c) and Andrews (1991, Corollary 1, p.830); see also Paparoditis and Politis (2001, Section 3.1, p.1110).

Corollary 3.2. If *Assumptions 3.1-3.4 are satisfied, $k^*_{(2)} \in (0,\infty)$, $\sum_{s=-\infty}^{\infty} |s|^2 R(s) < \infty$ and $\Delta_{k^*} \in (0,\infty)$, then, for any sequence of bandwidth parameters $\{S_T\}$* such that $S_T \to \infty$ and $S_T^5/T \to \gamma \in (0,\infty)$, *the sequence $\{S_T^*\}$ is preferred to $\{S_T\}$ in the sense that* $\lim_{T\to\infty} MSE(T^{4/5}, \hat{\sigma}^2_{\text{KBB}}(S_T)) - MSE(T^{4/5}, \hat{\sigma}^2_{\text{KBB}}(S_T^*)) \geq 0$. *The inequality is strict unless $S_T = S_T^* + o(1/T^{1/5})$.*

4 Comparisons

4.1 Tapered Block Bootstrap

It is helpful to re-call the kernel block bootstrap sample mean (2.3)

$$\bar{X}^*_{m_T} = \frac{1}{m_T}\sum_{s=1}^{m_T}\frac{1}{(\hat{k}_2 S_T)^{1/2}}\sum_{r=t_s^*-T}^{t_s^*-1} k(\frac{r}{S_T})(X_{t_s^*-r} - \bar{X}). \tag{4.1}$$

The tapered block bootstrap employs a non-negative taper $w(x)$ 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 $[0,1/2]$ (Paparoditis and Politis, 2001, Assumptions 1 and 2, p.1107). Hence, $w(x)$ is centred and unimodal at $1/2$. Given a positive integer bandwidth parameter S_T, the tapered variates are $Y_{rT}^t = w_{S_T}(r)(X_{t+r-1}-\bar{X})S_T^{1/2}/\left\|w_{S_T}\right\|_2$, $r=1,...,S_T$, where $w_{S_T}(r)=w\{(r-1/2)/S_T\}$ and $\left\|w_{S_T}\right\|_2 = (\sum_{r=1}^{S_T} w_{S_T}(r)^2)^{1/2}$; see Paparoditis and Politis (2001, (3), p.1106, and Step 2, p.1107). Each block $\mathcal{B}_t$ then has equal length S_T, that is, $\mathcal{B}_t = \{Y_{rT}^t\}_{r=1}^{S_T}$, $t=1,...,T-S_T+1$. Thus, with the bootstrap sample $\mathcal{B}_{t_s^*}$, $s=1,...,m_T$, the tapered block bootstrap sample mean (Paparoditis and Politis (2001, Step 3, p.1107) is

$$\begin{aligned}\bar{X}_{m_T}^* &= \frac{1}{S_T m_T}\sum_{s=1}^{m_T}\sum_{r=1}^{S_T} S_T^{1/2} w_{S_T}(r)\frac{S_T^{1/2}}{\left\|w_{S_T}\right\|_2}(X_{t_s^*+r-1}-\bar{X}) \\ &= \frac{1}{m_T}\sum_{s=1}^{m_T}\frac{1}{(\left\|w_{S_T}\right\|_2/S_T^{1/2})S_T^{1/2}}\sum_{r=1}^{S_T} w_{S_T}(r)(X_{t_s^*+r-1}-\bar{X}), \qquad (4.2)\end{aligned}$$

after scaling by $S_T^{1/2}$ to ensure comparability with the kernel block bootstrap sample mean (4.1); see Paparoditis and Politis (2001, Theorem 3, p.1108) and Theorem 3.1. Hence, comparing (4.1) with (4.2), the implicit transformed variates (2.1) are

$$\begin{aligned}X_{tT}^{\text{TBB}} &= \frac{1}{(\left\|w_{S_T}\right\|_2/S_T^{1/2})S_T^{1/2}}\sum_{r=1}^{S_T} w_{S_T}(r)(X_{t+r-1}-\bar{X}),\ t=1,...,T-S_T+1, \\ &= \frac{1}{(\left\|w_{S_T}\right\|_2/S_T^{1/2})S_T^{1/2}}\sum_{r=-[S_T/2]}^{S_T-[S_T/2]-1} w_{S_T}(r+[S_T/2]+1)(X_{t+r}-\bar{X}), \qquad (4.3)\end{aligned}$$

$t=[S_T/2]+1,...,T-S_T+[S_T/2]$, with, when expressed in terms of a kernel function, the taper $w(x)=k(x-1/2)$, $x\in[0,1]$, and the divisor $\left\|w_{S_T}\right\|_2/S_T^{1/2}$ in (4.3) replaced by $\hat{k}_2^{1/2}$ in (4.1), where $\left\|w_{S_T}\right\|_2 = \{\sum_{r=1}^{S_T} k((r-1/2-S_T/2)/S_T)^2\}^{1/2}$.

Important differences between the tapered and kernel block bootstraps, apart from those noted above, are immediately apparent. First, and most crucially, kernel block bootstrap variates (4.1) may be defined using kernel functions with unbounded support, in particular, the optimal kernel function (3.3). Secondly, the tapered block bootstrap, by employing the same block length S_T, omits blocks of length less than S_T formed from the data points $X_{tT} = \sum_{\max[t-T,-[S_T/2]]}^{\min[t-1,S_T-[S_T/2]-1]} w_{S_T}(r+[S_T/2]+1)(X_{t+r}-\bar{X})/(\left\|w_{S_T}\right\|_2/S_T^{1/2})S_T^{1/2}$, $t=1,...,[S_T/2]$ and $t=T-S_T+[S_T/2]+1,...,T$, at the beginning and end of the kernel block bootstrap sample space with kernel $k(x)=w(x+1/2)$.

As a consequence, tapered block bootstrap variance estimators are deficient relative to the kernel block bootstrap variance estimator using the optimal kernel (3.3). To see this, first, Corollary 3.1 implies that the limits of $MSE(T/S_T,\hat{\sigma}_{\text{KBB}}^2(S_{Tk^*}))$ and $MSE(T/S_T,\hat{\sigma}_{k^*}^2(S_{Tk^*}))$ are identical where $\hat{\sigma}_{k^*}^2(S_T) = \sum_{s=1-T}^{T-1} k^*(s/S_T)\hat{R}_T(s)$. In particular, the limits of $MSE(T/S_T,\tilde{\sigma}_{\text{KBB}}^2(S_T))$ and $MSE(T/S_T,\hat{\sigma}_{k_{\text{QS}}^*}^2(S_T))$

coincide. Secondly, let $\hat{\sigma}^2_{\text{TBB}}(S_T) = (T - S_T)^{-1} \sum_{t=[S_T/2]+1}^{T-S_T+[S_T/2]} (X^{\text{TBB}}_{tT} - \bar{X}^{\text{TBB}}_T)^2$ denote the tapered block bootstrap variance estimator, where $\bar{X}^{\text{TBB}}_T = \sum_{t=[S_T/2]+1}^{T-S_T+[S_T/2]} X^{\text{TBB}}_{tT}/(T - S_T)$. Theorems 1, p.1107, and 2, p.1108, of Paparoditis and Politis (2001) establish that, under their Assumptions 1, 2 and 3, p.1107, $MSE(T/S_T, \hat{\sigma}^2_{\text{TBB}}(S_{Tk^*}))$ and $MSE(T/S_T, \hat{\sigma}^2_{k^*}(S_{Tk^*}))$ have the same limit for $k^*(y) = \int_{-1/2}^{1/2} w(x - y + 1/2)w(x + 1/2)dx/w_2$ where $w_2 = \int_{-1/2}^{1/2} w(x + 1/2)^2 dx$. Hence, the resultant induced kernel $k^*_{\text{TBB}}(y)$ differs from the quadratic spectral kernel $k^*_{\text{QS}}(y)$ with positive Lebesgue measure and, therefore, the limit of $MSE(T/S_T, \hat{\sigma}^2_{\text{TBB}}(S_{Tk^*}))$ is strictly greater than that of $MSE(T/S_T, \tilde{\sigma}^2_{\text{KBB}}(S_T))$. Since the limit of $MSE(T/S_T, \hat{\sigma}^2_{\text{KBB}}(S_{Tk^*}))$ is an increasing function of $k^*_{(2)}(\int k^*(x)dx)^2$, see Section S.4.3, pp.S.5-S.6, of the Supplementary Material, to gauge the deficiency of the optimal tapered block bootstrap variance estimator (Paparoditis and Politis, 2001, p.1111) $k^*_{(2)\text{TBB}}(\int k^*_{\text{TBB}}(x)^2 dx)^2 = 1.6456$, noting $\int k^*_{\text{TBB}}(x)^2 dx = 0.5495$, $k^*_{(2)\text{TBB}} = 5.45$, whereas $k^*_{(2)\text{QS}}(\int k^*_{\text{QS}}(x)^2 dx)^2 = 1.4212$ as $\int k^*_{\text{QS}}(x)^2 dx = 1$.

4.2 Consistent Variance Estimation

Politis and Romano (1994) note that the block and stationary bootstrap variance estimators are approximately equivalent to the Bartlett kernel variance estimator; see, e.g., Newey and West (1987). Smith (2005, Section 2, pp.161-165) demonstrates a more general result under Assumptions 3.1-3.3 for kernel-based variance estimators (2.2); see also Smith (2011, Sections 2.4 and 2.5, pp.1199-1202, and Proof of Lemma A.3, pp.1219-1221). In particular, Smith (2011, Example 2.3, p.1204) shows that, in probability, the difference between the variance estimator (2.2) with kernel (3.3) and the optimal quadratic spectral or Bartlett-Priestley-Epanechnikov estimator with kernel (3.2) is negligible asymptotically; see Andrews (1991, (2.7), p.821) and Priestley (1981, (6.2.86), p.444). Theorem 3.2 establishes that, in addition, higher order properties of these variance estimators coincide.

5 Finite Sample Performance

This section investigates the finite sample performance of the kernel block bootstrap based on the optimal kernel (3.3). To provide a basis for comparison with the optimal tapered block bootstrap method using the trapezoidal taper (Paparoditis and Politis, 2001, Sections 3.1 and 3.2, pp. 1110-1112), identical simulation designs are employed.

The first set of simulation experiments investigates the sensitivity of the methods to the choice of the bandwidth while the second set evaluates their performance when the estimated optimal block size/ bandwidth is used.

5.1 Design I

Model 1. Nonlinear autorgressive model.

$$X_t = 0.6\sin(X_{t-1}) + Z_t. \ t \in \mathbb{Z}.$$

Model 2. Exponential autoregressive model.

$$X_t = \{0.8 - 1.1\exp(-50X_{t-1}^2)\}X_{t-1} + 0.1Z_t.\ t \in \mathbb{Z}.$$

Samples $t = 1, ..., T$ are generated from both models with the initialisation $X_{-50} = 0$ and $\{Z_t\}$ independent and identically distributed $N(0, 1)$.

All simulations for the kernel block bootstrap use $\hat{k}_2$ in (2.1) rather than k_2; see Remark 1. This approximation, since it also depends on the bandwidth, appears to compensate for situations in which the values of the bandwidth are too large or too low relative to the optimal bandwidth S_T^*.

The mean μ and variance σ_∞^2 for each process are approximated as follows: 5000 independent sample means $\bar{X}(i) = \sum_{t=1}^{10000} X_t(i)/10000$, $i = 1, ..., 5000$, are computed, each based on a sample $X_t(i)$, $t = 1, ..., 10000$. The population mean μ is then approximated by $\bar{X}_{5000} = \sum_{i=1}^{5000} \bar{X}(i)/5000$ and σ_∞^2 by $10000 \sum_{i=1}^{5000} \left(\bar{X}(i) - \bar{X}_{5000}\right)^2 /5000$. The simulation samples are drawn independently of these samples.

Samples $\{X_t\}_{t=1}^T$ are generated for the sample sizes $T = 200$ and $T = 1000$ with 5000 replications. Within each replication, both kernel and tapered block bootstraps were used to compute 95% equal tailed bootstrap confidence intervals for the population mean μ and the empirical mean squared error of the bootstrap estimators of the long run variance σ_∞^2. The bootstraps were computed with fixed bandwidth/block sizes S_T in the range from 2 to 40. The number of bootstrap replications in all cases was 1000.

Figures 1 and 2 about here

Empirical coverage rates are presented in Figure 1, while the results concerning empirical mean squared errors are displayed in Figure 2. Both Figures 1 and 2 indicate that the performance of the kernel block bootstrap method is sensitive to the choice of the bandwidth/block size parameter S_T. Although it is superior to the tapered block bootstrap for low values of S_T, relative empirical coverage deteriorates quite sharply as S_T increases for Model 1 in Figure 1(a) for the smaller sample size $T = 200$ at moderate and large block sizes. This facet of the kernel block bootstrap is less evident for Model 2 in Figure 1(c) for $T = 200$. For the larger sample size $T = 1000$ given in Figures 1(b) and 1(d) for Models 1 and 2 respectively, the superiority of the kernel block bootstrap occurs over a larger range for S_T with the performance of both bootstraps then broadly similar for moderate and large block sizes. The tapered block bootstrap method tends to be more robust to the choice of block size with empirical coverage initially increasing and then becoming relatively stable over a large range of S_T particularly for the smaller sample size $T = 200$. Similar conclusions may also be drawn from Figure 2 where again, in terms of empirical mean squared error of the respective estimators of σ_∞^2, the kernel block bootstrap is superior initially for small values of S_T with a deterioration in relative performance for moderate and large block sizes for $T = 200$ for Model 1 but this finding is somewhat less pronounced for Model 2.

Similarly to the results on empirical coverage, the negative aspects of these findings are ameliorated for the larger sample size $T = 1000$.

5.2 Design II

Given the sensitivity of the performance of the kernel block bootstrap to bandwidth/block size S_T, and because S_T is not fixed but depends on the sample size T, the behaviours of both bootstrap methods are further investigated when implemented using the estimated optimal bandwidth/block size as described in Paparoditis and Politis (2001, section 3.2, pp. 1111-1112) and Politis and White (2004, ftn. c, p. 59). See also Politis and Romano (1995).

The simulation experiments were based on samples $t = 1, ..., 200$ from the MA(2) model $X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$ with $\{Z_t\}$ $N(0,1)$ distributed as above and the initialisations $Z_{-50} = Z_{-49} = 0$ (Paparoditis and Politis 2001, pp. 1113-1114). The MA parameters θ_1 and θ_2 take values in the set $\{-1.0, -0.6, -0.3, 0.1, 0.4, 0.7, 1.0\}$.

Table 1 about here

The empirical mean squared errors of the kernel block bootstrap variance estimator $\hat{\sigma}^2_{\text{KBB}}$ divided by σ^4_∞ are displayed in Table 1. These results are similar to those given in Paparoditis and Politis (2001, Table 1, p.1115) for the tapered block bootstrap. Both methods behave relatively poorly in the presence of negative dependence with similar problematic cases. See, for example, $\theta_1 = -0.6$ $\theta_2 = -0.3$, $\theta_1 = 0.1$ $\theta_2 = -1.0$, $\theta_1 = -0.6$ $\theta_2 = 0.1$ and $\theta_1 = -0.3$ $\theta_2 = 0.1$.

Table 2 about here

The ratio of the empirical mean squared error for the kernel block bootstrap variance estimator $\hat{\sigma}^2_{\text{KBB}}$ divided by that of the tapered block bootstrap is reported in Table 2. These results confirm the conclusion from Table 1 that the kernel block bootstrap estimates of σ^2_∞ are generally better in the lower right triangular part of the table, that is, for positive θ_1 and θ_2, with the opposite result in the upper left triangular portion of the table. Nevertheless, overall, the kernel block bootstrap provides an improvement in 57% of the cases considered. Although not reported here, this proportion increases to 76% if the infeasible optimal bandwidth S^*_T is used. Hence, although the kernel block bootstrap offers theoretical advantages over the tapered block bootstrap, given the increased sensitivity to the choice of the bandwidth, these advantages are somewhat diluted in practice.

Table 3 about here

To examine the location of the estimator of the optimal bandwidth/block size S^*_T for the quadratic spectral kernel block bootstrap, Table 3 displays the empirical mean of the estimated optimal bandwidth/block size together with its standard deviation and the true S^*_T. Again, similar to the results for the tapered block bootstrap reported in Paparoditis and Politis (2001, Table 3, p. 1116), there are substantial deviations between the empirical mean and true S_T^* for the problematic cases. Further investigation of these cases indicated substantial outliers and highly positively skewed distributions of the estimators for S_T^*. Apart from these designs, there is a relatively close correspondence between the empirical mean and true S_T^*.

Acknowledgement

Paulo Parente is grateful for the financial support of Fundação para a Ciência e a Tecnologia (Project PTDC/IIM-ECO/3097/2014)

References

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* **59**, 817-858.

Bickel, P. & Freedman, D. (1981). Some asymptotic theory of the bootstrap. *Ann. Statist.* **9**, 1196-1217.

Brillinger, D. R. (1981). *Time Series: Data Analysis and Theory.* New York: McGraw-Hill.

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. *Ann. Statist.* **7**, 1-26.

Epanechnikov, V. A. (1969). Non-parametric estimation of a multivariate probability density. *Th. Prob. Applic.* **14**, 153-158.

Gradshteyn, I. S. & Ryzhik, I. M. (1980). *Tables of Integrals, Series and Products.* New York: Academic Press.

Ibragimov, I. A. & Linnik, Y. V. (1971). *Independent and Stationary Sequences of Random Variables.* Groningen : Wolters-Noordhoff.

Kitamura, Y. & 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. *Ann. Statist.* **17**, 1217–1241.

Liu, R. & Singh, K. (1992). Moving blocks jackknife and bootstrap capture weak dependence. In *Exploring the Limits of Bootstrap*, R. LePage & L. Billard, eds. New York: Wiley, pp. 224-248.

Newey, W. K. & West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55**, 703-708.

Paparoditis, E. & Politis, D. N. (2001). Tapered block bootstrap. *Biometrika* **88**, 1105-1119.

Parzen, E. (1957). On consistent estimates of the spectrum of a stationary time series. *Ann. Math. Statist.* **28**, 329-348.

Politis, D. N. & Romano, J. P. (1994). The stationary bootstrap. *J. Am. Statist. Assoc.* **89**, 1303-1313.

Politis, D. N. & Romano, J. P. (1995). Bias-corrected nonparametric spectral estimation. *Journal of Time Series Analysis* **16**, 67-103.

Politis, D. N. & White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews* **23**, 53-70.

Priestley, M. B. (1962). Basic considerations in the estimation of spectra. *Technometrics* **4**, 551-5154,

Priestley, M. B. (1981). *Spectral Analysis and Time Series.* Volumes I and 2. New York: Academic Press.

Rosenblatt, M. (1985). *Stationary Sequences and Random Fields.* Boston: Birkhàuser.

Sacks, J. & Ylvisacker, D. (1981). Asymptotically optimum kernels for density estimation at a point. *Ann. Statist.* **9**, 334-346.

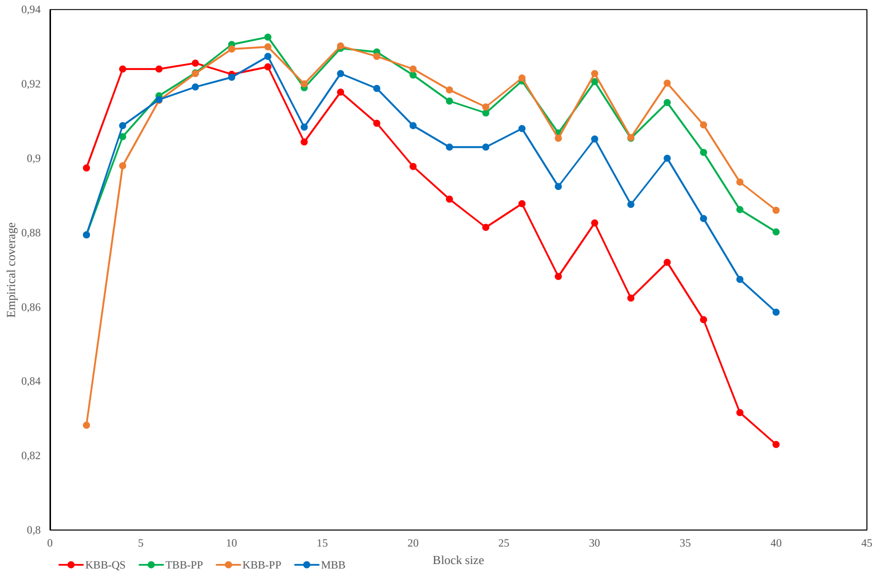
Singh, K. N. (1981). On the asymptotic accuracy of Efron's bootstrap. *Ann. Statist.* **9**, 1187-1195.

Smith, R. J. (1997). Alternative semi-parametric likelihood approaches to generalised method of moments estimation. *Econ. J.* **107**, 503-19.

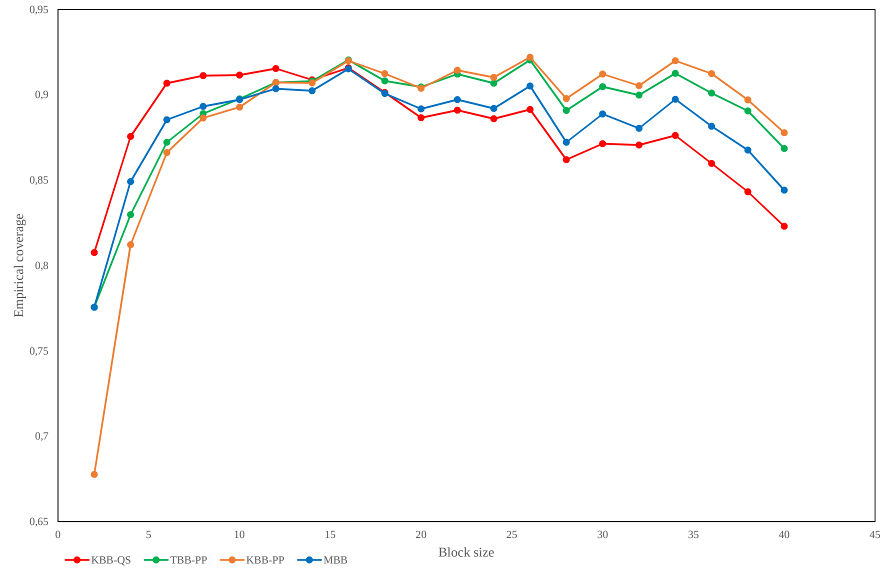
Smith, R. J. (2005). Automatic positive semidefinite HAC covariance matrix and GMM estimation. *Econometric Th.* **21**, 158-170.

Smith, R. J. (2011). GEL criteria for moment condition models. *Econometric Th.* **27**, 1192-1235.

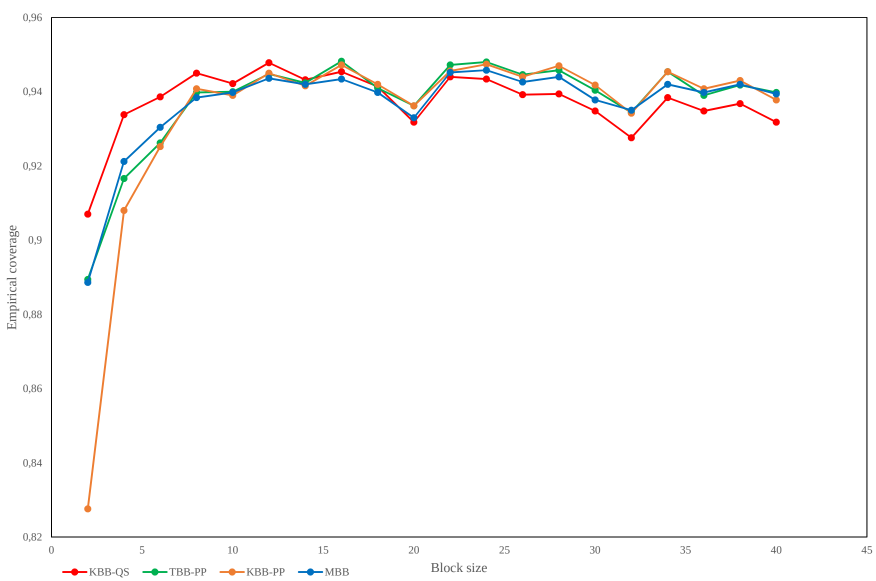
Figure 1: **Empirical coverage, as a function of the block size S_T, of 95% equal-tailed confidence intervals obtained with block bootstraps for Models 1 and 2.**



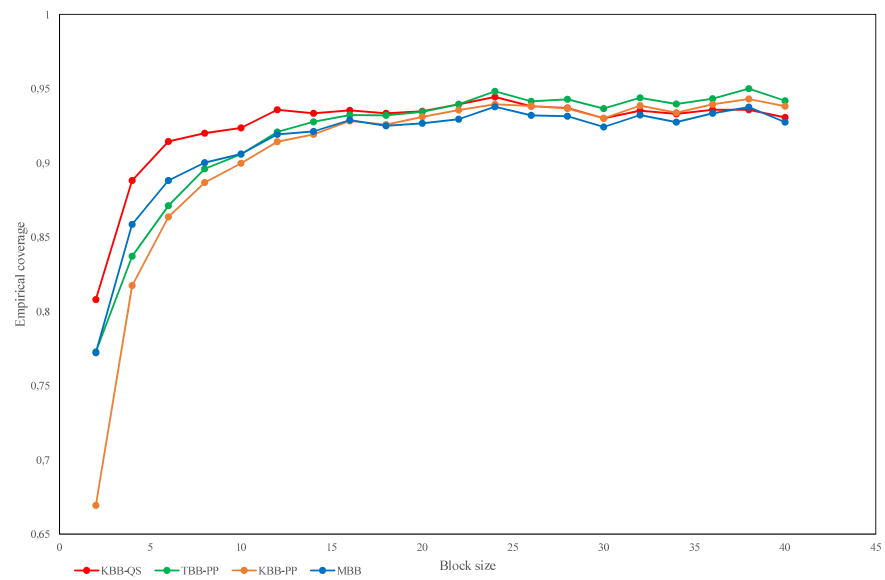
(a) **Model 1:** $T = 200$



(c) **Model 2:** $T = 200$



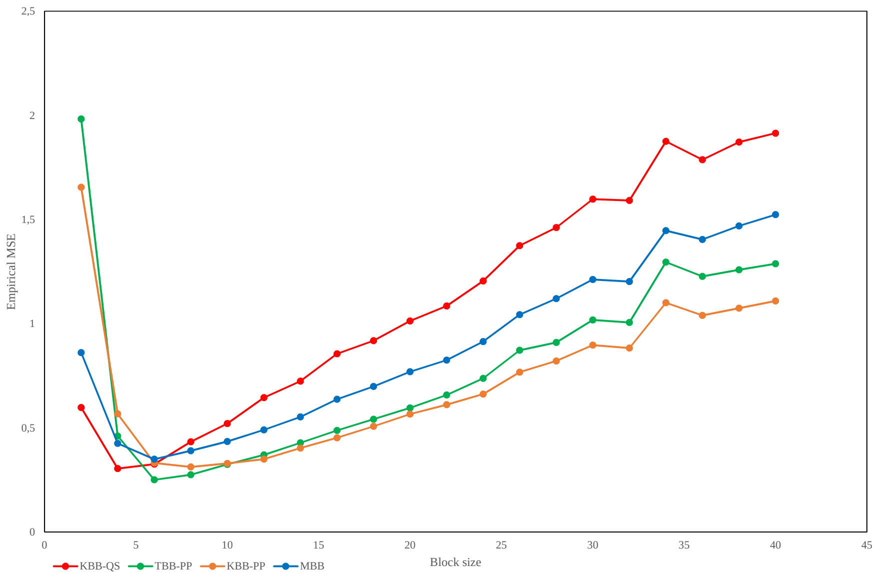
(b) **Model 1:** $T = 1000$



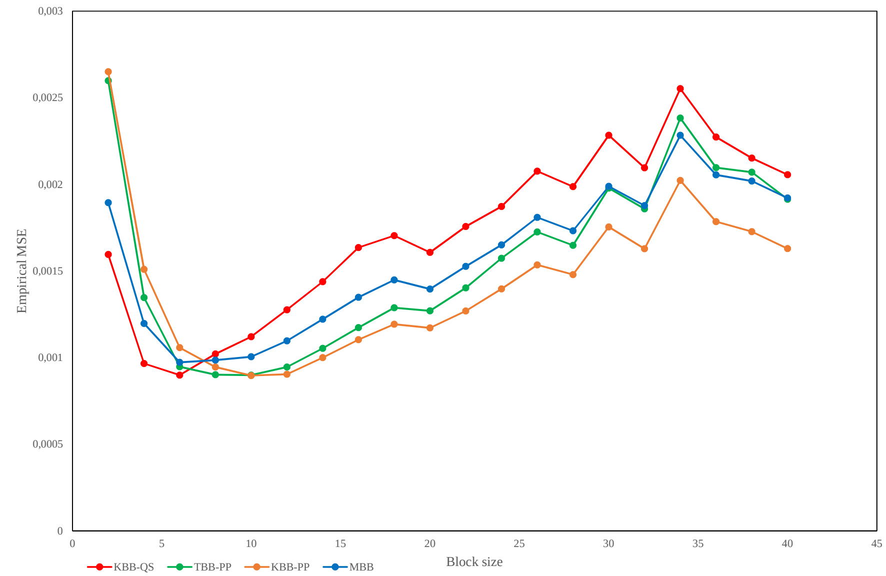
(d) **Model 2:** $T = 1000$

Note: KBB-QS: optimal kernel block bootstrap; TBB-PP: optimal tapered block bootstrap; KBB-PP: kernel block bootstrap with optimal tapered block bootstrap kernel; MBB: moving blocks bootstrap.

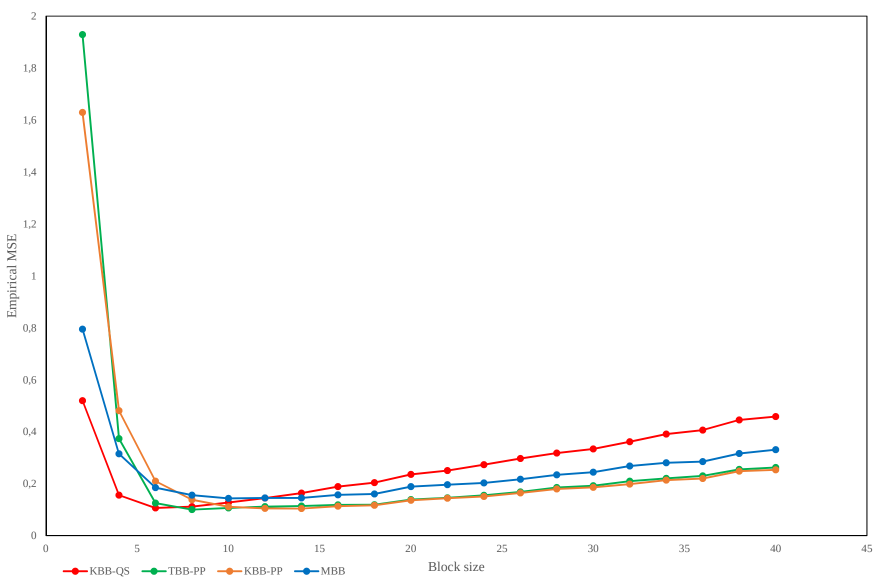
Figure 2: **Empirical mean squared error, as a function of the block size S_T, of block bootstrap estimators for σ^2_∞ for Models 1 and 2.**



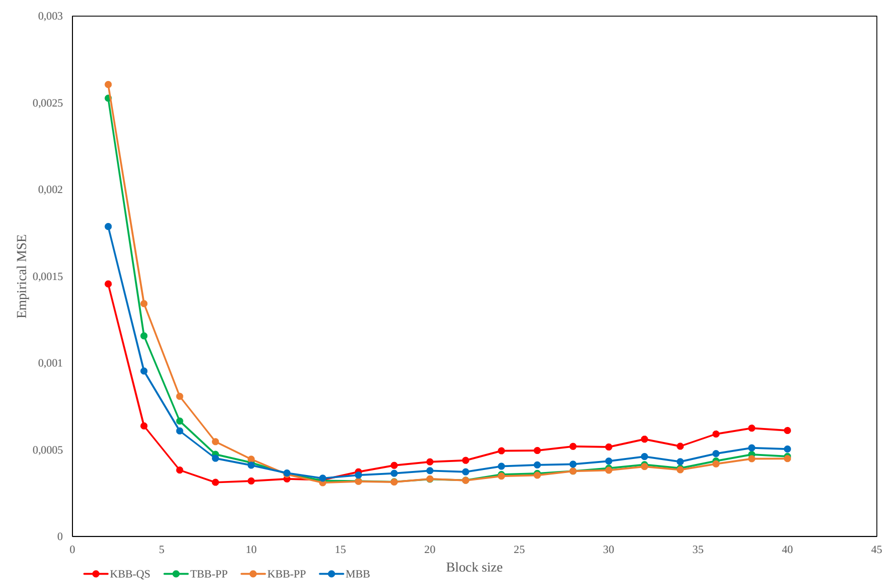
(a) **Model 1:** $T = 200$



(c) **Model 2:** $T = 200$



(b) **Model 1:** $T = 1000$



(d) **Model 2:** $T = 1000$

Note: KBB-QS: optimal kernel block bootstrap; TBB-PP: optimal tapered block bootstrap; KBB-PP: kernel block bootstrap with optimal tapered block bootstrap kernel; MBB: moving blocks bootstrap.

Table 1. Ratio of empirical mean squared error for the kernel block bootstrap variance estimator $\hat{\sigma}^2_{\text{KBB}}$ divided by σ^4_∞ with estimated block size, for different cases of $\text{MA}(2)$ models and $T = 200$.

θ_1	$\theta_2 = -1.0$	$\theta_2 = -0.6$	$\theta_2 = -0.3$	$\theta_2 = 0.1$	$\theta_2 = 0.4$	$\theta_2 = 0.7$	$\theta_2 = 1.0$
-1.0	0.221	1.368	5.232	16.686	0.882	0.112	0.097
-0.6	0.266	7.825	276.073	0.145	0.096	0.080	0.079
-0.3	0.893	14.064	1.656	0.047	0.069	0.077	0.074
0.1	34.325	0.242	0.179	0.061	0.066	0.071	0.075
0.4	0.396	0.146	0.077	0.053	0.072	0.077	0.076
0.7	0.226	0.114	0.056	0.056	0.074	0.083	0.082
1.0	0.211	0.097	0.045	0.056	0.078	0.079	0.079

Table 2. Ratio of empirical mean squared error for the kernel block bootstrap divided by the corresponding empirical mean squared error for tapered block bootstrap with estimated block size, for different cases of $\text{MA}(2)$ models and $T = 200$.

θ_1	$\theta_2 = -1.0$	$\theta_2 = -0.6$	$\theta_2 = -0.3$	$\theta_2 = 0.1$	$\theta_2 = 0.4$	$\theta_2 = 0.7$	$\theta_2 = 1.0$
-1.0	1.091	1.504	1.525	1.082	5.216	1.090	0.969
-0.6	1.075	1.567	1.540	0.955	0.951	0.910	0.897
-0.3	1.388	1.680	1.502	0.840	0.935	0.929	0.919
0.1	1.644	1.042	1.132	1.121	0.922	0.910	0.936
0.4	1.136	1.043	0.961	0.989	0.937	0.913	0.906
0.7	1.038	1.023	0.950	0.954	0.928	0.900	0.940
1.0	1.091	0.995	0.934	0.931	0.910	0.942	0.916

Table 3. The empirical mean of the optimal bandwdith estimator for the kernel block bootstrap, with its standard deviation in parentheses and the true value of S_T^* in square brackets, for different cases of $\mathrm{MA}(2)$ models and $T = 200$.

θ_1	$\theta_2 = -1.0$	$\theta_2 = -0.6$	$\theta_2 = -0.3$	$\theta_2 = 0.1$	$\theta_2 = 0.4$	$\theta_2 = 0.7$	$\theta_2 = 1.0$
-1.0	9.51	12.22	8.72	12.82	7.64	6.95	6.93
	[8.76] (4.66)	[11.43] (11.18)	[17.05] (10.08)	[27.54] (8.94)	[5.50] (5.10)	[6.96] (3.02)	[6.64] (2.49)
-0.6	16.62	21.85	13.29	13.20	8.08	6.15	6.05
	[13.19] (15.86)	[26.90] (17.61)	[38.52] (17.36)	[5.11] (10.11)	[5.39] (9.10)	[5.87] (2.15)	[5.81] (2.16)
-0.3	21.92	22.85	12.73	4.61	5.01	5.65	5.59
	[22.96] (19.05)	[45.97] (17.48)	[12.02] (10.73)	[2.08] (1.37)	[4.98] (1.56)	[5.36] (2.96)	[5.37] (2.05)
0.1	23.31	15.42	6.08	1.98	4.52	5.11	5.26
	[55.30] (17.13)	[12.36] (13.84)	[6.32] (2.93)	[3.32] (1.34)	[4.54] (1.73)	[4.86] (2.47)	[4.94] (3.07)
0.4	21.37	9.19	4.57	2.96	4.66	5.05	5.11
	[18.24] (19.71)	[8.31] (4.66)	[4.51] (2.46)	[3.39] (1.85)	[4.28] (8.89)	[4.58] (3.43)	[4.68] (3.41)
0.7	14.08	6.58	3.78	3.28	4.35	4.99	5.19
	[11.66] (10.66)	[6.30] (2.50)	[3.35] (1.99)	[3.35] (2.11)	[4.06] (3.44)	[4.35] (4.13)	[4.46] (4.69)
1.0	9.53	5.08	3.32	3.42	4.14	4.80	5.10
	[8.76] (4.53)	[5.07] (2.47)	[2.50] (1.86)	[3.27] (2.53)	[3.88] (3.23)	[4.15] (4.23)	[4.28] (4.91)

Supplement to "Kernel Block bootstrap"

Paulo M.D.C. Parente
Instituto Universitário de Lisboa (ISCTE-IUL)
Business Research Unit (BRU-IUL)

Richard J. Smith
*c*emmap, U.C.L and I.F.S.
Faculty of Economics, University of Cambridge
Department of Economics, University of Melbourne
O.N.S. Economic Statistics Centre of Excellence

This Draft: July 2018

S.1 Introduction

This Supplement to the paper Kernel Block Bootstrap details the proofs of Theorems 1 and 2 together with a number of subsidiary results used in establishing Theorems 1 and 2.

S.2 Preliminaries

Throughout the Supplement, C denotes a generic positive constants that may be different in different uses with CS, M, and T the Cauchy-Schwarz, Markov and triangle inequalities respectively.

To simplify the analysis, the Supplement considers the transformed centred observations

$$X_{tT} = \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=t-1}^{t-T} k(\frac{s}{S_T})(X_{t-s} - \bar{X})$$

with k_2 substituting for $\hat{k}_2 = \sum_{t=1-T}^{T-1} k(t/S_T)^2/S_T$ in the main text. Corollary K.2 establishes that the results given below also apply for X_{tT} as defined in the main text.

Without loss of generality, since

$$X_{tT} = \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=t-1}^{t-T} k(\frac{s}{S_T})((X_{t-s} - \mu) - (\bar{X} - \mu))$$

and

$$\bar{X}_T = \frac{1}{T} \sum_{t=1}^{T} \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=t-1}^{t-T} k(\frac{s}{S_T})((X_{t-s} - \mu) - (\bar{X} - \mu)),$$

the transformed and original samples are regarded below as having been drawn from a zero mean process, i.e., $\mu = 0$.

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

Recall

$$m_T var^*(\bar{X}^*_{m_T}) = \sum_{t=1}^{T} (X_{tT} - \bar{X}_T)^2/T.$$

S.3 Some Notation

For ease of reference some notation used in the following is collected here.

Let

$$k_{tT}(\frac{s}{S_T}) = \frac{1}{k_2 S_T} k(\frac{t-s}{S_T}) k(\frac{t}{S_T}), k_T(\frac{s}{S_T}) = \sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T})$$

and

$$k_a(x) = \frac{1}{k_2} k(x+a)k(x),$$

with $k^{(j)}(x) = d^j k(x)/dx^j$ and $k_a^{(j)}(x) = d^j k_a(x)/dx^j$, $j = 1, 2$.

Also define

$$\hat{R}_{tT}(s) = \frac{1}{T} \sum_{r=\max[1,1-s,1-t]}^{\min[T-1,T-s,T-t]} X_{r+s} X_r, \hat{R}_T(s) = \frac{1}{T} \sum_{r=\max[1,1-s]}^{\min[T-1,T-s]} X_{r+s} X_r.$$

Recall

$$J_T = \sum_{s=1-T}^{T-1} (1 - \frac{|s|}{T}) R(s);$$

$$k_q^* = \lim_{y \to 0} \frac{1 - k^*(y)}{|y|^q};$$

$$\Gamma_{k^*} = -\frac{1}{S_T^2} k_q^* \sum_{s=-\infty}^{\infty} |s|^2 R(s), \Delta_{k_*} = 2\sigma_\infty^4 \int_{-\infty}^{\infty} k^*(x)^2 dx.$$

S.4 Proofs of Results

S.4.1 Large Sample Validity

Proof of Theorem 1. The result is proven in Steps 1-5 below; cf. Politis and Romano (1992, Proof of Theorem 2, pp. 1993-5). For simplicity, let $m_T = T/S_T$ be integer.

Step 1: $\bar{X} \to 0$ in probability. Follows by White (1984, Theorem 3.47, p.46).

Step 2: $\text{pr}\{T^{1/2}\bar{X}/\sigma_\infty \leq x\} \to \Phi(x)$, where $\Phi(\cdot)$ is the standard normal distribution function. Follows by White (1984, Theorem 5.19, p.124).

Step 3: $\sup_x \left|\text{pr}\{T^{1/2}\bar{X}/\sigma_\infty \leq x\} - \Phi(x)\right| \to 0$. Follows by Pólya's Theorem (Serfling, 1980, Theorem 1.5.3, p.18) from Step 2 and the continuity of $\Phi(\cdot)$.

STEP 4: $var^*((T/S_T)^{1/2}\bar{X}^*_{m_T}) \to \sigma^2_\infty$ in probability. Note $E^*(\bar{X}^*_{m_T}) = \bar{X}_T$. Thus,

$$\begin{aligned} var^*((T/S_T)^{1/2}\bar{X}^*_{m_T}) &= var^*(X_{t^*T}) \\ &= \frac{1}{T}\sum_{t=1}^{T}(X_{tT} - \bar{X}_T)^2 \\ &= \frac{1}{T}\sum_{t=1}^{T}X^2_{tT} - \bar{X}^2_T. \end{aligned}$$

The result follows since $\bar{X}^2_T = O_p(S_T/T)$ (Smith, 2011, Lemma A.2, p.1219), $S_T/T = o(1)$, Assumption 2(a), and $\sum_{t=1}^{T} X^2_{tT}/T \xrightarrow{p} \sigma^2_\infty$ (Smith, 2011, Lemma A.3, p.1219).

STEP 5:

$$\lim_{T\to\infty} \text{pr}\left\{\sup_x \left|\text{pr}^*\{\frac{\bar{X}^*_{m_T} - E^*(\bar{X}^*_{m_T})}{var^*(\bar{X}^*_{m_T})^{1/2}} \le x\} - \Phi(x)\right| \ge \varepsilon\right\} = 0.$$

Applying the Berry-Esséen inequality, Serfling (1980, Theorem 1.9.5, p.33), noting the bootstrap sample observations $\{X_{t^*_sT}\}_{s=1}^{m_T}$ are independent and identically distributed,

$$\sup_x \left|\text{pr}^*\{\frac{(T/S_T)^{1/2}(\bar{X}^*_{m_T} - \bar{X}_T)}{var^*((T/S_T)^{1/2}\bar{X}^*_{m_T})^{1/2}} \le x\} - \Phi(x)\right| \le \frac{C}{m_T^{1/2}} var^*(X_{t^*T})^{-3/2}E^*(\left|X_{t^*T} - \bar{X}_T\right|^3).$$

Now $var^*(X_{t^*T}) \xrightarrow{p} \sigma^2_\infty > 0$; see the Proof of Step 4 above. Furthermore, $E^*(\left|X_{t^*T} - \bar{X}_T\right|^3) = T^{-1}\sum_{t=1}^{T}\left|X_{tT} - \bar{X}_T\right|^3$ and

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T}\left|X_{tT} - \bar{X}_T\right|^3 &\le \max_t \left|X_{tT} - \bar{X}_T\right| \frac{1}{T}\sum_{t=1}^{T}(X_{tT} - \bar{X}_T)^2 \\ &= O_p(S_T^{1/2}T^{1/\alpha}). \end{aligned}$$

The equality follows since $\max_t \left|X_{tT} - \bar{X}_T\right| \le \max_t \left|X_{tT}\right| + \left|\bar{X}_T\right| = O_p(S_T^{1/2}T^{1/\alpha}) + O_p((S_T/T)^{1/2}) = O_p(S_T^{1/2}T^{1/\alpha})$ by M, Assumption 3(a), (Newey and Smith, 2004, Proof of Lemma A1, p.239), and $\sum_{t=1}^{T}(X_{tT} - \bar{X}_T)^2/T = O_p(1)$, see the Proof of Step 4 above. Therefore

$$\begin{aligned} \sup_x \left|\text{pr}^*\{\frac{(T/S_T)^{1/2}(\bar{X}^*_{m_T} - \bar{X}_T)}{var^*((T/S_T)^{1/2}\bar{X}^*_{m_T})^{1/2}} \le x\} - \Phi(x)\right| &\le \frac{1}{m_T^{1/2}}O_p(1)O_p(S_T^{1/2}T^{1/\alpha}) \\ &= \frac{S_T^{1/2}}{m_T^{1/2}}O_p(T^{1/\alpha}) = o_p(1), \end{aligned}$$

by Assumption 2(a), yielding the required conclusion.■

S.4.2 ASYMPTOTIC BIAS, VARIANCE AND MEAN SQUARED ERROR

PROOF OF THEOREM 2. **(a)**

$$\begin{aligned} E(\hat{\sigma}^2_{\text{KBB}}) &= \frac{1}{T}\sum_{t=1}^{T}E((X_{tT} - \bar{X}_T)^2) \\ &= \frac{1}{T}\sum_{t=1}^{T}E(X^2_{tT}) - E(\bar{X}^2_T). \end{aligned}$$

Now

$$\begin{aligned}\frac{1}{T}\sum_{t=1}^{T} X_{tT}^2 &= \sum_{s=1-T}^{T-1}\sum_{t=\max[1-T,1-s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T})\hat{R}_{tT}(s) \\ &= \sum_{s=1-T}^{T-1} k_T(\frac{s}{S_T})\hat{R}_T(s) \\ &\quad + \sum_{s=1-T}^{T-1}\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T})(\hat{R}_{tT}(s)-\hat{R}_T(s)).\end{aligned}$$

From Lemma B.1,

$$\sum_{s=1-T}^{T-1}(k^*(\frac{s}{S_T})-k_T(\frac{s}{S_T}))E(\hat{R}_T(s)) = (O((\frac{S_T}{T})^{b-1/2})+o(\frac{1}{S_T^2}))J_T + O(\frac{S_T^{b-2}}{T^b}),$$

and, from Lemma B.2,

$$\sum_{s=1-T}^{T-1}\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T})E(\hat{R}_{tT}(s)-\hat{R}_T(s)) = o(\frac{S_T}{T}).$$

Therefore,

$$\begin{aligned}\frac{S_T}{T}\sum_{t=1}^{T} E(X_{tT}^2) &= J_T - \frac{1}{S_T^2}(\Gamma_{k^*}+o(1)) \\ &\quad +(O((\frac{S_T}{T})^{b-1/2})+o(\frac{1}{S_T^2}))J_T + O(\frac{S_T^{b-2}}{T^b}) + o(\frac{S_T}{T}).\end{aligned}$$

since, see Parzen (1957) and Andrews (1991, Proposition 1(b), p.825),

$$\sum_{s=1-T}^{T-1} k^*(\frac{s}{S_T})E(\hat{R}_T(s)) = J_T - \frac{1}{S_T^2}(k_{(2)}^*\sum_{s=-\infty}^{\infty}|s|^2 R(s) + o(1)).$$

By Lemma B.3,

$$E(\bar{X}_T^2) \le O(\frac{S_T}{T})J_T + O((\frac{S_T}{T})^2).$$

Therefore, collecting terms,

$$\begin{aligned}E(\hat{\sigma}_{\text{KBB}}^2) &= J_T + \frac{1}{S_T^2}(\Gamma_{k^*}+o(1)) \\ &\quad +(O((\frac{S_T}{T})^{b-1/2})+o(\frac{1}{S_T^2}))J_T + O(\frac{S_T^{b-2}}{T^b}) + o(\frac{S_T}{T}) \\ &\quad +O(\frac{S_T}{T})J_T + O((\frac{S_T}{T})^2) \\ &= J_T + \frac{1}{S_T^2}(\Gamma_{k^*}+o(1)) + O((\frac{S_T}{T})^{b-1/2}) + o(\frac{1}{S_T^2}) + O(\frac{S_T^{b-2}}{T^b}) + O(\frac{S_T}{T}) + O((\frac{S_T}{T})^2).\blacksquare\end{aligned}$$

(b) From Theorems V.1-V.3

$$\begin{aligned}\frac{T}{S_T}var(\hat{\sigma}_{\text{KBB}}^2) &= 2\sigma_\infty^4\int_{-\infty}^{\infty} k^*(y)^2dy + o(1) + O(\frac{S_T}{T}) + O(\frac{1}{T^2}) \\ &= 2\sigma_\infty^4\int_{-\infty}^{\infty} k^*(y)^2dy + o(1).\blacksquare\end{aligned}$$

(c) Let

$$U_T = O((\frac{S_T}{T})^{b-1/2}) + o(\frac{1}{S_T^2}) + O(\frac{S_T^{b-2}}{T^b}) + O(\frac{S_T}{T}) + O((\frac{S_T}{T})^2).$$

From (a) and (b)

$$\begin{aligned} MSE(T/S_T, \hat{\sigma}^2_{\text{KBB}}) &= \frac{T}{S_T} E((\hat{\sigma}^2_{\text{KBB}} - J_T)^2) \\ &= \frac{T}{S_T} var(\hat{\sigma}^2_{\text{KBB}}) + \frac{T}{S_T}(E(\hat{\sigma}^2_{\text{KBB}}) - J_T)^2 \\ &= \frac{T}{S_T} var(\hat{\sigma}^2_{\text{KBB}}) + \frac{T}{S_T}(\frac{1}{S_T^2}(\Gamma_{k^*} + o(1)) + U_T)^2. \end{aligned}$$

Now

$$\frac{T}{S_T}(\frac{1}{S_T^2}(\Gamma_{k^*} + o(1)) + U_T)^2 = ((\frac{T}{S_T^5})^{1/2}(\Gamma_{k^*} + o(1)) + (\frac{T}{S_T})^{1/2} U_T)^2.$$

and, in particular,

$$(\frac{T}{S_T})^{1/2} U_T = O((\frac{S_T}{T})^{b-1}) + o((\frac{T}{S_T^5})^{1/2}) + O(\frac{1}{S_T^2}(\frac{S_T}{T})^{b-1/2}) + O((\frac{S_T}{T})^{1/2}) + O((\frac{S_T}{T})^{3/2})$$

All the terms are $o(1)$ by Assumptions 2(a) and 4(d) and by hypothesis $S_T^5/T \to \gamma \in (0, \infty)$. The result is then immediate.■

S.4.3 Optimality

Let the induced kernel function $k^*(y) = \int_{-\infty}^{\infty} k(x-y)k(x)dx/k_2$ satisfy Assumptions 1, 2(a)(b)(c) and $|K^*(\lambda)| > 0$ for all $\lambda \in \mathbb{R}$ where $K^*(\lambda) = (2\pi)^{-1}\int_{-\infty}^{\infty} k^*(y)\exp(-iy\lambda)dy$. Also let $\hat{\sigma}^2_{k^*}(S_T) = \sum_{s=1-T}^{T-1} k^*(s/S_T)\hat{R}_T(s)$ and $S_{Tk^*} = S_T/\int k^*(y)^2 dy$.

Then, if $S_T^5/T \to \gamma$ for some $\gamma \in (0, \infty)$ and $\sum_{s=-\infty}^{\infty} |s|^2 R(s) \in (0, \infty)$, $\lim_{T\to\infty}(MSE(T/S_T, \hat{\sigma}^2_{k^*}(S_{Tk^*}) - MSE(T/S_T, \hat{\sigma}^2_{k^*_{\text{QS}}}(S_T))) \geq 0$ with strict inequality if $k^*(y)$ and $k^*_{\text{QS}}(y)$ differ by positive Lebesgue measure.

To show this, a proof almost identical to Andrews (1991, Proof of Theorem 2, pp.853-854) for kernel $k^*(y)$, bandwidth sequence S_{Tk^*} and $q = 2$ suffices except that Andrews (1991, Theorem 1(c), p.827) is replaced by Andrews (1991, Proposition 1(c), p.825). Since $S_{Tk^*}^5/T \to \gamma/(\int k^*(y)^2 dy)^5$ and $T/S_T = (1/\int k^*(y)^2 dy)T/S_{Tk^*}$,

$$\lim_{T\to\infty} MSE(T/S_T, \hat{\sigma}^2_{k^*}(S_{Tk^*})) = \frac{1}{\gamma}(\int k^*(y)^2 dy)^4 \Gamma^2_{k^*} + 2(\sum_{s=-\infty}^{\infty} R(s))^2,$$

where $\Gamma_{k^*} = -k^*_{(2)} \sum_{s=-\infty}^{\infty} |s|^2 R(s)$, provided $k^*_{(2)} < \infty$. If $k_2 = \infty$, $\lim_{T\to\infty} MSE(T/S_T, \hat{\sigma}^2_k(S_{Tk^*})) = \infty$ since the bias term is unbounded; see Andrews (1991, Proof of Theorem 2, pp.853-854). In conclusion, $k^*_{(2)}(\int k^*(y)^2 dy)^2 \geq k^*_{(2)\text{QS}}$ with strict inequality if $k^*(y)$ and $k^*_{\text{QS}}(y)$ differ by positive Lebesgue measure; see Andrews (1991, eq. (A.20), p.854).

S.5 Preliminary Results

S.5.1 Asymptotic Bias

S.5.1.1 Preliminaries

Note

$$E(\hat{R}_T(s)) = (1 - \frac{|s|}{T})R(s), s = 0, \pm 1, ...$$

S.5.1.2 Results

Lemma B.1. *Let Assumptions 1, 2, 3(a) and 4(a)(b)(d) hold. If, in addition, Assumption 4(c) holds, that is,* $\int_{-\infty}^{\infty} k_a^{(2)}(x)dx = 0$, *and* $\sum_{s=-\infty}^{\infty} |s|^2 R(s) < \infty$, *then*

$$\sum_{s=1-T}^{T-1} (k^*(\frac{s}{S_T}) - k_T(\frac{s}{S_T}))E(\hat{R}_T(s)) = (O((\frac{S_T}{T})^{b-1/2}) + o(\frac{1}{S_T^2}))J_T + O(\frac{S_T^{b-2}}{T^b}).$$

Proof. From Lemmas K.2 and K.3,

$$k^*(\frac{s}{S_T}) - k_T(\frac{s}{S_T}) = O((\frac{S_T}{T})^{b-1/2}) + o(\frac{1}{S_T^2})$$

uniformly s. Hence

$$\sum_{s=1-T}^{T-1} (k^*(\frac{s}{S_T}) - k_T(\frac{s}{S_T}))E(\hat{R}_T(s)) = (O((\frac{S_T}{T})^{b-1/2}) + o(\frac{1}{S_T^2}))J_T.$$

The result is then immediate since by hypothesis $\sum_{s=-\infty}^{\infty} |s|^2 R(s) = O(1)$.■

Lemma B.2. *Under Assumptions 1, 2, 3(a) and 4,*

$$\sum_{s=1-T}^{T-1} \sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T})E(\hat{R}_{tT}(s) - \hat{R}_T(s)) = o(\frac{S_T}{T}).$$

Proof. Consider the difference

$$\sum_{r=\max[1,1-s,1-t]}^{\min[T-1,T-s,T-t]} E(X_{r+s}X_r) - \sum_{r=\max[1,1-s]}^{\min[T-1,T-s]} E(X_{r+s}X_r).$$

Suppose $t > s$. Then

$$\begin{aligned} \sum_{r=\max[1,1-s,1-t]}^{\min[T-1,T-s,T-t]} E(X_{r+s}X_r) - \sum_{r=\max[1,1-s]}^{\min[T-1,T-s]} E(X_{r+s}X_r) &= \sum_{r=\max[1,1-s]}^{\min[T-1,T-t]} E(X_{r+s}X_r) - \sum_{r=\max[1,1-s]}^{\min[T-1,T-s]} E(X_{r+s}X_r) \\ &= - \sum_{r=\min[T,T-t]+1}^{\min[T-1,T-s]} E[X_{r+s}X_r]. \end{aligned}$$

Using Doukhan (1994, Theorem 3.(1), p.9),

$$
\begin{aligned}
|E(X_{r+s}X_r)| &\leq 8\alpha_X(s)^{\delta/(\delta+1)}E(|X_r|^{2(\delta+1)})^{1/2(\delta+1)}E(|X_{r+s}|^{2(\delta+1)})^{1/2(\delta+1)} \\
&\leq C\alpha_X(s)^{\delta/(\delta+1)}
\end{aligned}
$$

where the last inequality follows from Assumption 3(a). By T

$$
\begin{aligned}
\left|\sum_{r=\min[T,T-t]+1}^{\min[T-1,T-s]} E(X_{r+s}X_r)\right| &\leq \sum_{r=\min[T,T-t]+1}^{\min[T-1,T-s]} |E(X_{r+s}X_r)| \\
&\leq C\alpha_X(s)^{\delta/(\delta+1)}\max[(\min[T,T-s]-\min[T,T-t]),0] \\
&= C\alpha_X(s)^{\delta/(\delta+1)}(t-s) \text{ if } s>0 \text{ or } C\alpha_X(s)^{(v-1)/v}\max[t,0] \text{ if } s\leq 0 \\
&\leq C\alpha_X(s)^{\delta/(\delta+1)}\max[t,0].
\end{aligned}
$$

For $t \leq s$

$$
\sum_{r=\max[1,1-s,1-t]}^{\min[T-1,T-s,T-t]} E(X_{r+s}X_r) - \sum_{r=\max[1,1-s]}^{\min[T-1,T-s]} E(X_{r+s}X_r) = \sum_{r=\max[1,1-s]}^{\max[1,1-t]-1} E(X_{r+s}X_r).
$$

Similarly by T

$$
\begin{aligned}
\left|\sum_{r=\max[1,1-s]}^{\max[1,1-t]-1} E(X_{r+s}X_r)\right| &\leq C\alpha_X(s)^{\delta/(\delta+1)}\max[(\max[1,1-t]-\max[1,1-s]),0] \\
&= C\alpha_X(s)^{\delta/(\delta+1)}\max[-t,0] \text{ if } s>0 \text{ or } C\alpha_X(s)^{\delta/(\delta+1)}(s-t) \text{ if } s<0 \\
&\leq C\alpha_X(s)^{\delta/(\delta+1)}\max[-t,0].
\end{aligned}
$$

Consequently

$$
\left|\sum_{r=\max[1,1-s,1-t]}^{\min[T-1,T-s,T-t]} E(X_{r+s}X_r) - \sum_{r=\max[1,1-s]}^{\min[T-1,T-s]} E(X_{r+s}X_r)\right| \leq C\alpha_X(s)^{\delta/(\delta+1)}|t|.
$$

Hence, also by T,

$$
\begin{aligned}
\left|\sum_{s=1-T}^{T-1}\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(s)E(\hat{C}_{tT}(t)-\hat{C}_T(s))\right| &\leq C\sum_{s=1-T}^{T-1}\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]}|k_{tT}(s)| \\
&\quad \times\alpha_X(s)^{\delta/(\delta+1)}\frac{|t|}{T} \\
&\leq \frac{C}{k_2}\frac{1}{S_T}\sum_{t=1-T}^{T-1}\frac{|t|}{T}\left|k(\frac{t}{S_T})\right| \\
&\quad \times\sum_{s=1-T}^{T-1}\left|k(\frac{t-s}{S_T})\right|\alpha_X(s)^{\delta/(\delta+1)}.
\end{aligned}
$$

By the mean value theorem

$$
k(\frac{t-s}{S_T}) = k(\frac{t}{S_T}) - \frac{1}{S_T}k^{(1)}(\frac{c_{tT}(s)}{S_T})s
$$

where $c_T(s) \in (t-s,t)$. By T $|k((t-s)/S_T)| \leq |k(t/S_T)| + |(s/S_T)| \sup_x \left|k^{(1)}(x)\right|$ and, thus, by Assumption 1

$$\begin{aligned}
\sum_{s=1-T}^{T-1} \left|k(\frac{t-s}{S_T})\right| \alpha(s)^{\delta/(\delta+1)} &\leq \left|k(\frac{t}{S_T})\right| \sum_{s=1-T}^{T-1} \alpha_X(s)^{\delta/(\delta+1)} \\
&\quad + \frac{1}{S_T} \sup_x \left|k^{(1)}(x)\right| \sum_{s=1-T}^{T-1} |s|\, \alpha_X(s)^{\delta/(\delta+1)} \\
&= \left|k(\frac{t}{S_T})\right| o(1) + o(\frac{1}{S_T})
\end{aligned}$$

uniformly t. From Lemma K.4,

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

and, similarly

$$\frac{1}{S_T} \sum_{t=1-T}^{T-1} \frac{|t|}{T} \left|k(\frac{t}{S_T})\right|^2 o(1) \leq o(1) \sup_x |k(x)|\, O(\frac{S_T}{T}) = o(\frac{S_T}{T}).$$

Therefore

$$\left| \sum_{s=1-T}^{T-1} \sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T}) E(\hat{R}_{tT}(s) - \hat{R}_T(s)) \right| \leq o(\frac{S_T}{T}).\blacksquare$$

LEMMA B.3. *Under Assumptions 1, 2. 3(a) and 4,*

$$E(\bar{X}_T^2) \leq O(\frac{S_T}{T}) J_T + O((\frac{S_T}{T})^2).$$

PROOF. Write

$$\begin{aligned}
\bar{X}_T &= \frac{1}{T} \frac{1}{(k_2 S_T)^{1/2}} \sum_{t=1}^{T} \sum_{s=t-1}^{t-T} k(\frac{s}{S_T}) X_{t-s} \\
&= \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=1-T}^{T-1} k(\frac{s}{S_T}) \frac{1}{T} \sum_{t=\max[1,1-s]}^{\min[T,T-s]} X_t \\
&= \bar{X} \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=1-T}^{T-1} k(\frac{s}{S_T}) \\
&\quad + \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=1-T}^{T-1} k(\frac{s}{S_T}) (\frac{1}{T} \sum_{t=\max[1,1-s]}^{\min[T,T-s]} X_t - \bar{X}) \\
&= \mathcal{A}_T + \mathcal{B}_T.
\end{aligned}$$

Using the c_r inequality, White (1984, Proposition 3.8, p.33),

$$E((\mathcal{A}_T + \mathcal{B}_T)^2) \leq 2(E(\mathcal{A}_T^2) + E(\mathcal{B}_T^2)).$$

First,

$$\begin{aligned}
E(\mathcal{A}_T^2) &= \frac{1}{T} \sum_{s=1-T}^{T-1} (1 - \frac{|s|}{T}) R(s) \{ \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=1-T}^{T-1} k(\frac{s}{S_T}) \}^2 \\
&= O(\frac{S_T}{T}) J_T.
\end{aligned}$$

Secondly, by CS,

$$\begin{aligned}
E(\mathcal{B}_T^2) &\leq (\frac{1}{(k_2 S_T)^{1/2}} \sum_{s=1-T}^{T-1} \left| k(\frac{s}{S_T}) \right|) \{ \frac{1}{(k_2 S_T)^{1/2}} \sum_{s=1-T}^{T-1} \left| k(\frac{s}{S_T}) \right| E((\frac{1}{T} \sum\nolimits_{t=\max[1,1-s]}^{\min[T,T-s]} X_t - \bar{X})^2) \} \\
&\leq (\frac{1}{(k_2 S_T)^{1/2}} \sum_{s=1-T}^{T-1} \left| k(\frac{s}{S_T}) \right|) (\frac{1}{(k_2 S_T)^{1/2}} \sum_{s=1-T}^{T-1} |s| \left| k(\frac{s}{S_T}) \right|) O(\frac{1}{T^2}) \\
&= O(\frac{S_T^2}{T^2});
\end{aligned}$$

the second inequality follows from a Doukhan (1994) moment bound, see, for example, Politis et al. (1997, Lemma A.1, eq. (A.4), p.304), noting the $O(T^{-2})$ term is independent of s with the third equality obtained from Lemma K.4.■

S.5.2 Asymptotic Variance

S.5.2.1 Preliminaries

Nordman (2009, Theorem 2, p.365) is used below and is stated here for ease of reference. Let

$$\mathcal{T}_T = \sum_{s=0}^{T-1} a_{s,T} \hat{R}_T(s).$$

Also let

$$\mathcal{A}_T = \sum_{s=1}^{T-1} a_{s,T}^2 (1 - \frac{s}{T})^2 > 0,$$

$$\mathcal{B}_T = \frac{1}{T} + \frac{\log(T)}{T^2} (\sum_{s=0}^{T-1} |a_{s,T}|)^2 + \frac{1}{T^2} \sum_{s=0}^{T-1} |a_{s,T}| \, s (1 - \frac{s}{T})$$

and

$$\mathcal{C}_T = \sum_{s=2}^{T-1} |a_{s,T} - a_{s-1,T}| \, .$$

Define the smoothing window $H_T(\omega) = \sum_{s=0}^{T-1} a_{s,T}^2 (1 - T^{-1} s) \exp(is\omega)$ and the non-negative kernel $\kappa_T(\omega) = H_T(\omega) H_T(-\omega)/(2\pi \mathcal{A}_T)$, $\omega \in \mathbb{R}$. Assumptions A.1 and A.2, p.362, of Nordman (2009) are now restated.

Assumption N.1. $\sum_{s=-\infty}^{\infty} |s| \, |R(s)| < \infty$.

Assumption N.2. $\sum_{t_1,t_2,t_3=-\infty}^{\infty} |\kappa(X_0, X_{t_1}, X_{t_2}, X_{t_3})| < \infty$.

Theorem N. (Nordman (2009) Theorem 2, p.365.) *Suppose that Assumptions N.1 and N.2 hold. Then, if in addition* $\sup_T \max_{0 \leq s \leq T-1} |a_{s,T}| < \infty$, *then* **(a)**

$$var(\mathcal{T}_T) = (2\pi)^2 \frac{\mathcal{A}_T}{T} \int_{-\pi}^{\pi} \kappa_T(\omega) f^2(\omega) d\omega + O((\mathcal{A}_T \mathcal{B}_T / T)^{1/2} + \mathcal{B}_T);$$

(**b**)

$$\lim_{T\to\infty}\int_{-\pi}^{\pi}\kappa_T(\omega)f^2(\omega)d\omega = f^2(0)$$

if $\lim_{T\to\infty}\mathcal{A}_T=\infty$ *and* $\sup_T\mathcal{C}_T<\infty$ *also hold.*

Note

$$\frac{1}{T^2}var(\sum_{t=1}^{T}(X_{tT}-\bar{X}_T)^2)=\frac{1}{T^2}var(\sum_{t=1}^{T}X_{tT}^2-T\bar{X}_T^2)$$

and the components of

$$\begin{aligned}\frac{1}{k_2T}\sum_{t=1}^{T}X_{tT}^2 &= \sum_{s=1-T}^{T-1}\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]}k_{tT}(\frac{s}{S_T})\hat{R}_{tT}(s)\\ &= \sum_{s=1-T}^{T-1}k_T(\frac{s}{S_T})\hat{R}_T(s)\\ &\quad+\sum_{s=1-T}^{T-1}\sum_{t=\max[1-T,1-s]}^{\min[T-1,T-1+s]}k_{tT}(\frac{s}{S_T})(\hat{R}_{tT}(s)-\hat{R}_T(s))\end{aligned}$$

which together with $\bar{X}_T^2$ are examined below.

S.5.2.2 Results

Define

$$\begin{aligned}a_{s,T} &= 2k_T(\frac{s}{S_T})\\ &= 2k^*(\frac{s}{S_T})+O(u_T), s=\pm1,...,\end{aligned}$$

and $a_{0,T}=\tilde{k}_T(0)=k^*(0)+O(u_T)$ where, by Lemmas K.2 and K.3, if $b>1$ and $\int_{-\infty}^{\infty}k_a^{(2)}(x)dx=0$,

$$O(u_T)=O((\frac{S_T}{T})^{b-1/2})+o(\frac{1}{S_T^2})$$

uniformly s. Thus

$$\mathcal{A}_T=\sum_{s=1}^{T-1}(2k_T(\frac{s}{S_T}))^2(1-\frac{s}{T})^2,$$

$$\mathcal{B}_T=\frac{1}{T}+4(\sum_{s=1}^{T-1}\left|k_T(\frac{s}{S_T})\right|)^2\frac{\log(T)}{T^2}+\frac{1}{T^2}2\sum_{s=1}^{T-1}\left|k_T(\frac{s}{S_T})\right|s(1-\frac{s}{T}),$$

$$\mathcal{C}_T=2\sum_{s=2}^{T-1}\left|k_T(\frac{s}{S_T})-k_T(\frac{s-1}{S_T})\right|.$$

Lemma V.1. *If Assumption 2 is satisfied then* $\sup_T\max_{0\le s\le T-1}|a_{s,T}|<\infty$.

PROOF. Using Jansson (2002, Lemma 1, p.1451),

$$\begin{aligned}
\sup_T \max_{0\le s\le T-1} |a_{s,T}| &= 2\sup_T \max_{0\le s\le T-1} \left| k_T(\frac{s}{S_T}) \right| \\
&= \frac{2}{k_2}\sup_T \max_{0\le s\le T-1} \left| \frac{1}{S_T} \sum_{t=\max[1-T,1-s]}^{\min[T-1,T-1+s]} k(\frac{t-s}{S_T})k(\frac{t}{S_T}) \right| \\
&\le \frac{2}{k_2} k_{\max} \sup_T \left| \frac{1}{S_T} \sum_{s=1-T}^{T-1} k(\frac{t}{S_T}) \right| < \infty. \blacksquare
\end{aligned}$$

LEMMA V.2. *Let Assumptions 2 and 4 hold. Then,*

$$\frac{\mathcal{A}_T}{S_T} = \frac{2}{S_T}\sum_{s=1-T}^{T-1} k^*(\frac{s}{S_T})^2 + O(u_T) + O(\frac{T}{S_T}u_T^2) + O(\frac{1}{S_T}).$$

PROOF. Now

$$\mathcal{A}_T \le 4\sum_{s=1}^{T-1} k_T(\frac{s}{S_T})^2.$$

By symmetry

$$2\sum_{s=1}^{T-1} k_T(\frac{s}{S_T})^2 = \sum_{s=1-T}^{T-1} k_T(\frac{s}{S_T})^2 - k_T(0)^2.$$

Then, by Lemmas K.2 and K.3,

$$\begin{aligned}
\frac{1}{S_T}\sum_{s=1-T}^{T-1} k_T(\frac{s}{S_T})^2 &= \frac{1}{S_T}\sum_{s=1-T}^{T-1} (k^*(\frac{s}{S_T})^2 + k^*(\frac{s}{S_T})O(u_T) + O(u_T^2)) \\
&= \frac{1}{S_T}\sum_{s=1-T}^{T-1} k^*(\frac{s}{S_T})^2 + O(u_T) + O(\frac{T}{S_T}u_T^2).
\end{aligned}$$

Finally

$$\frac{1}{S_T}k_T(0)^2 = \frac{1}{S_T}k^*(0)^2 + O(\frac{u_T}{S_T}) = O(\frac{1}{S_T}). \blacksquare$$

LEMMA V.3. *If Assumptions 2 and 4 are satisfied, then*

$$\mathcal{B}_T = \frac{1}{T} + O(\frac{\log(T)}{T^2}S_T^2) + O((\frac{S_T}{T})^2).$$

PROOF. By symmetry and Smith (2005, eq. (A.5), p.169),

$$\begin{aligned}
\frac{2}{S_T}\sum_{s=1}^{T-1} \left| k_T(\frac{s}{S_T}) \right| &= \frac{1}{S_T}\sum_{s=1-T}^{T-1} \left| k_T(\frac{s}{S_T}) \right| - \frac{1}{S_T}k_T(0) \\
&= O(1) + O(\frac{1}{S_T}).
\end{aligned}$$

Thus

$$(\sum_{s=1}^{T-1}\left|2k_T(\frac{s}{S_T})\right|)^2\frac{\log(T)}{T^2}=O(\frac{\log(T)}{T^2}S_T^2).$$

Next consider

$$\frac{1}{T^2}\sum\nolimits_{s=1}^{T-1}\left|2k_T(\frac{s}{S_T})\right|s(1-\frac{s}{T})\leq\frac{1}{T^2}\sum_{s=1-T}^{T-1}\left|k_T(\frac{s}{S_T})\right||s|.$$

By T and Lemma K.4,

$$\begin{aligned}\frac{1}{T^2}\sum_{s=1-T}^{T-1}\left|k_T(\frac{s}{S_T})\right||s| &\leq \frac{1}{k_2T^2}\sum_{t=1-T}^{T-1}\left|k(\frac{t}{S_T})\right|\frac{1}{S_T}\sum_{s=1-T}^{T-1}|s|\left|k(\frac{t-s}{S_T})\right| \\ &\leq O(\frac{1}{T^2})\sum_{t=1-T}^{T-1}\left|k(\frac{t}{S_T})\right|O(S_T) \\ &\leq O((\frac{S_T}{T})^2)(\int_{-\infty}^{\infty}\bar{k}(x)dx)=O((\frac{S_T}{T})^2)\end{aligned}$$

by Assumption 2(c).■

Lemma V.4. *Suppose Assumptions 2(b)(c) and 4(a)(b) are satisfied. Then*

$$\mathcal{C}_T<\infty.$$

Proof. Note that, for $s\geq 2$,

$$\begin{aligned}k_T(\frac{s}{S_T})-k_T(\frac{s-1}{S_T}) &= \frac{1}{k_2S_T}\sum_{t=1-T+s}^{T-1}k(\frac{t}{S_T})(k(\frac{t-s}{S_T})-k(\frac{t-(s-1)}{S_T})) \\ &\quad -\frac{1}{k_2S_T}k(\frac{s-T}{S_T})k(\frac{1-T}{S_T}).\end{aligned}$$

Hence, by T

$$\begin{aligned}\mathcal{C}_T &= 2\sum_{s=2}^{T-1}\left|k_T(\frac{s}{S_T})-k_T(\frac{s-1}{S_T})\right| \\ &\leq \frac{2}{k_2S_T}\sum_{s=2}^{T-1}\sum_{t=1-T+s}^{T-1}\left|k(\frac{t}{S_T})\right|\left|k(\frac{t-s}{S_T})-k(\frac{t-(s-1)}{S_T})\right| \\ &\quad +\frac{2}{k_2S_T}\left|k(\frac{1-T}{S_T})\right|\sum_{s=2}^{T-1}\left|k(\frac{s-T}{S_T})\right| \\ &\leq \frac{2}{k_2S_T}\sum_{t=1-T}^{T-1}\left|k(\frac{t}{S_T})\right|\sum_{s=2}^{T-1}\left|\int_{\frac{t-(s-1)}{S_T}}^{\frac{t-s}{S_T}}k^{(1)}(x)dx\right|+c \\ &\leq \frac{2}{k_2S_T}\sum_{t=1-T}^{T-1}\left|k(\frac{t}{S_T})\right|\int_{\frac{t-(T-1)}{S_T}}^{\frac{t-1}{S_T}}\left|k^{(1)}(x)\right|dx+c \\ &\leq \frac{2}{k_2S_T}\sum_{t=1-T}^{T-1}\left|k(\frac{t}{S_T})\right|C+c<\infty\end{aligned}$$

where c and C are positive constants. The second inequality follows from Assumptions 2(c) and 4(a) and Lemma K.4, the fourth inequality from Assumption 4(b) since $\int_{\frac{t-(T-1)}{S_T}}^{\frac{t-1}{S_T}} \left|k^{(1)}(x)\right| dx \leq \int_{-\infty}^{\infty} \left|k^{(1)}(x)\right| dx \leq C$ and the final inequality since $\sum_{t=1-T}^{T-1} \left|k(\frac{t}{S_T})\right| / S_T = O(1)$ by Assumption 2(c) and Lemma K.4. ■

Theorem N is applied to

$$\mathcal{T}_T = \sum_{s=1-T}^{T-1} k_T(\frac{s}{S_T})\hat{R}_T(s).$$

THEOREM V.1. *Let Assumptions 1-4 hold. If* $\sum_{s=-\infty}^{\infty} |s|^2 \mathcal{R}(s) < \infty$ *and* $S_T \geq O(T^{1/5})$ *then*

$$\frac{T}{S_T} var(\mathcal{T}_T) = 2\sigma_\infty^4 \int_{-\infty}^{\infty} k^*(y)^2 dy + o(1).$$

PROOF. Lemma V.1 establishes $\sup_T \max_{0 \leq s \leq T-1} |a_{s,T}| < \infty$. Assumption N.1 follows by hypothesis and Assumption N.2 holds under Assumption 1 by Andrews (1991, Lemma 1, p.824). The additional conditions required for Theorem N(b) are verified by noting $\lim_{T\to\infty} \mathcal{A}_T = \infty$ from Lemma V.2 and $\sup_T \mathcal{C}_T < \infty$ from Lemma V.4.

Therefore, by Theorem N, since $f(0) = \sigma_\infty^2/2\pi$,

$$\frac{T}{S_T} var(\mathcal{T}_T) = \frac{\mathcal{A}_T}{S_T}(\sigma_\infty^4 + o(1)) + \frac{T}{S_T} O((\frac{\mathcal{A}_T \mathcal{B}_T}{T})^{1/2} + \mathcal{B}_T).$$

From Lemma V.2

$$\frac{\mathcal{A}_T}{S_T} = \frac{2}{S_T} \sum_{s=1-T}^{T-1} k^*(\frac{s}{S_T})^2 + O(u_T) + O(\frac{T}{S_T} u_T^2) + O(\frac{1}{S_T}).$$

Now, $O(u_T) = o(1)$,

$$O((\frac{T}{S_T})^{1/2} u_T) = O((\frac{S_T}{T})^{b-1}) + o((\frac{T}{S_T^5})^{1/2}) = o(1)$$

if $S_T \geq O(T^{1/5})$. Thus

$$\frac{\mathcal{A}_T}{S_T} = 2\int_{-\infty}^{\infty} k^*(y)^2 dy + o(1)$$

since

$$\frac{1}{S_T} \sum_{s=1-T}^{T-1} k^*(\frac{s}{S_T})^2 = \int_{-\infty}^{\infty} k^*(y)^2 dy + o(1).$$

Finally, since $\mathcal{B}_T = o(1)$, to establish the order of the remainder, consider $(T/S_T)^2 \mathcal{A}_T \mathcal{B}_T / T = (\mathcal{A}_T/S_T)(T\mathcal{B}_T/S_T)$. From Lemma V.3,

$$\frac{T}{S_T}\mathcal{B}_T = \frac{1}{S_T} + O(\frac{\log(T)}{T} S_T) + O(\frac{S_T}{T}).$$

Hence the result follows.■

Set

$$\mathcal{R}_T = \sum_{s=1-T}^{T-1} \sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T})(\hat{R}_{tT}(s) - \hat{R}_T(s)).$$

THEOREM V.2. *If Assumption 1-4 hold, then*

$$\frac{T}{S_T} var(\mathcal{R}_T) = O(\frac{S_T^2}{T}).$$

PROOF. Let

$$\begin{aligned}
\xi_{st} &= \sum_{r=\max[1,1-s,1-t]}^{\min[T-1,T-s,T-t]} X_{r+s}X_r - \sum_{r=\max[1,1-s]}^{\min[T-1,T-s]} X_{r+s}X_r \\
&= - \sum_{r=\min[T,T-t]+1}^{\min[T-1,T-s]} X_{r+s}X_r \text{ if } t > s \\
&= \sum_{r=\max[1,1-s]}^{\max[1,1-t]-1} X_{r+s}X_r \text{ if } t \leq s.
\end{aligned}$$

Thus

$$var(\mathcal{R}_T) \leq \frac{1}{T^2} E[(\sum_{s=1-T}^{T-1} \sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T})\xi_{st})^2].$$

By T and using the Minkowski inequality (White, 1984, Ex. 3.53(i), p.46),

$$\begin{aligned}
E((\sum_{s=1-T}^{T-1} \sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T})\xi_{st})^2) &\leq E((\sum_{s=1-T}^{T-1} \sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} \left| k_{tT}(\frac{s}{S_T})\xi_{st} \right|)^2) \\
&\leq (\sum_{s=1-T}^{T-1} E((\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} \left| k_{tT}(\frac{s}{S_T})\xi_{st} \right|)^2)^{1/2})^2.
\end{aligned}$$

A further application of the Minkowski inequality

$$\begin{aligned}
E((\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} \left| k_{tT}(\frac{s}{S_T})\xi_{st} \right|)^2) &\leq (\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} E(\left| k_{tT}(\frac{s}{S_T}) \right|^2 \xi_{st}^2)^{1/2})^2 \\
&= (\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} \left| k_{tT}(\frac{s}{S_T}) \right| E(\xi_{st}^2)^{1/2})^2.
\end{aligned}$$

By a Doukhan (1994) moment bound, see, for example, Politis et al. (1997, Lemma A.1, eq. (A.4), p.304), and noting that ξ_{st} consists of no more than $|t|$ terms,

$$E(|\xi_{st}|^2) \leq |t| O(1)$$

uniformly s and t. Thus, uniformly s,

$$\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} \left| k_{tT}(\frac{s}{S_T}) \right| E(\xi_{st}^2)^{1/2} \leq O(1) \frac{1}{k_2 S_T} \sum_{t=1-T}^{T-1} |t|^{1/2} \left| k(\frac{t-s}{S_T}) \right| \left| k(\frac{t}{S_T}) \right|.$$

Therefore, combining the above results,

$$\begin{aligned}
var(\mathcal{R}_T) &\leq O(\frac{1}{T^2})(\frac{1}{k_2 S_T}\sum_{t=1-T}^{T-1}|t|^{1/2}\left|k(\frac{t}{S_T})\right|\sum_{s=1-T}^{T-1}\left|k(\frac{t-s}{S_T})\right|)^2 \\
&= O(\frac{1}{T^2})(O(S_T^{1/2})O(S_T))^2 = O(\frac{S_T^3}{T^2})
\end{aligned}$$

noting $\sum_{s=1-T}^{T-1}|k\{(t-s)/S_T\}| = O(S_T)$ with the inequality following from Lemma K.4.■

THEOREM V.3. *If Assumptions 1-4 hold,*

$$var(\bar{X}_T^2) \leq O(\frac{S_T^2}{T^2}).$$

PROOF. Recall from the Proof of Lemma B.3

$$\bar{X}_T = \frac{1}{T}\sum_{t=1}^{T} k_{tT} X_t$$

where

$$k_{tT} = \frac{1}{(k_2 S_T)^{1/2}}\sum_{s=t-1}^{t-T} k(\frac{s}{S_T}), t = 1, ..., T.$$

Since $var(\bar{X}_T^2) \leq E(\bar{X}_T^4)$

$$var(\bar{X}_T^2) \leq O(\frac{S_T^2}{T^4})E((\sum_{t=1}^{T} X_t)^4)$$

noting $k_{tT}/S_T^{1/2} = O(1)$ uniformly t.

By a Doukhan (1994) moment bound, see, for example, Politis et al. (1997, Lemma A.1, eq. (A.4), p.304),

$$E(\left|\sum_{t=1}^{T} X_t\right|^4) \leq 3024\max[1, \mathcal{C}^2(4,\delta)]\mathcal{D}(4,\delta,T)$$

for each T where

$$\begin{aligned}
\mathcal{C}(4,\delta) &= \sum_{j=0}^{\infty}(j+1)^2\alpha_X(j)^{\delta/(4+\delta)}, \\
\mathcal{D}(4,\delta,T) &= \max[\mathcal{L}(4,\delta,T), [\mathcal{L}(2,\delta,T)]^2], \\
\mathcal{L}(4,\delta,T) &= \sum_{t=1}^{T} E(|X_t|^{4+\delta})^{\frac{4}{4+\delta}} \leq CT, \\
\mathcal{L}(2,\delta,T) &= \sum_{t=1}^{T} E(|X_t|^{2+\delta})^{\frac{2}{2+\delta}} \leq CT
\end{aligned}$$

from Assumption 3(a).

Now $\mathcal{C}(4,\delta)$ is bounded by Assumption 1. Therefore

$$\begin{aligned}
E(\bar{X}_T^4) &\leq O(\frac{S_T^2}{T^4})3024\max[1, \mathcal{C}^2(4,\delta)]\mathcal{D}(4,\delta,T)] \\
&\leq O(\frac{S_T^2}{T^2}).■
\end{aligned}$$

S.5.3 Kernel Functions

S.5.3.1 Notation

Recall

$$k_{tT}(\frac{s}{S_T}) = \frac{1}{k_2 S_T} k(\frac{t-s}{S_T})k(\frac{t}{S_T}), k_T(\frac{s}{S_T}) = \sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]} k_{tT}(\frac{s}{S_T})$$

$$k_a(x) = \frac{1}{k_2} k(x+a)k(x),$$

with $k^{(j)}(x) = d^j k(x)/dx^j$ and $k_a^{(j)}(x) = d^j k_a(x)/dx^j$, $j = 1, 2$.

Let

$$\begin{aligned} Z_T^{c_T}(a) &= \frac{1}{k_2 S_T} \sum_{t=1-c_T}^{c_T-1} k(\frac{t-a}{S_T})k(\frac{t}{S_T}) \\ &\quad + \frac{1}{2k_2 S_T} k(\frac{-c_T-a}{S_T})k(\frac{-c_T}{S_T}) \\ &\quad + \frac{1}{2k_2 S_T} k(\frac{c_T-a}{S_T})k(\frac{c_T}{S_T}). \end{aligned}$$

S.5.3.2 Useful Lemmata

Lemma K.1. *Let $k(\cdot) \in C^{\,2}([-c,c])$ and suppose $k(\cdot)$ satisfies Assumptions 4(a) and (b). Then*

$$Z_T^{c_T}(a) = \frac{1}{k_2} \int_{-c}^{c} k_a(x)dx + \frac{1}{12 k_2 S_T^2} \int_{-c}^{c} k_a^{(2)}(x)dx + o(\frac{1}{S_T^2})$$

uniformly a.

Proof. The proof is an adaptation of Cruz-Uribe and Neugebauer (2002, Proof of Theorem 1.23, pp.20-21).

Consider the interval $[-c_T/S_T, c_T/S_T]$ and define the subintervals $J_i = [x_{i-1}, x_i]$ of equal length $1/S_T$, i.e., $x_i = (i - c_T)/S_T$, $i = 1, ..., 2c_T$, with $x_0 = -c_T/S_T$. Then

$$Z_T^{c_T}(a) = \frac{1}{2k_2 S_T} \sum_{i=1}^{2c_T} (k_a(x_{i-1}) + k_a(x_i)).$$

Define m_i as the mid-point of interval J_i, $(i = 1, ..., 2S_T)$. Hence, using integration by parts,

$$\begin{aligned} \int_{x_{i-1}}^{x_i} (x - m_i) k_a^{(1)}(x)dx &= [(x-m_i)k_a(x)]_{x_{i-1}}^{x_i} - \int_{x_{i-1}}^{x_i} k_a(x)dx \\ &= \frac{1}{2S_T}(k_a(x_{i-1}) + k_a(x_i)) - \int_{x_{i-1}}^{x_i} k_a(x)dx. \end{aligned}$$

Also, again applying integration by parts,

$$\begin{aligned} \frac{1}{2}\int_{x_{i-1}}^{x_i} (\frac{1}{4S_T^2} - (x-m_i)^2) k_a^{(2)}(x)dx &= [\frac{1}{2}(\frac{1}{4S_T^2} - (x-m_i)^2) k_a^{(1)}(x)]_{x_{i-1}}^{x_i} + \frac{1}{2}\int_{x_{i-1}}^{x_i} 2(x-m_i) k_a^{(1)}(x)dx \\ &= \int_{x_{i-1}}^{x_i} (x-m_i) k_a^{(1)}(x)dx. \end{aligned}$$

Now

$$\frac{1}{2}\int_{x_{i-1}}^{x_i}(\frac{1}{4S_T^2}-(x-m_i)^2)dx = \frac{1}{2}[\frac{x}{4S_T^2}-\frac{(x-m_i)^3}{3}]_{x_{i-1}}^{x_i} = \frac{1}{2}[\frac{1}{4S_T^3}-\frac{2}{24S_T^3}] = \frac{1}{12S_T^3}.$$

Therefore

$$\frac{1}{12S_T^3}\inf_{x\in J_i}k_a^{(2)}(x) \le \int_{x_{i-1}}^{x_i}(x-m_i)k_a^{(1)}(x)dx \le \frac{1}{12S_T^3}\sup_{x\in J_i}k_a^{(2)}(x).$$

Multiplying by S_T^2 and summing over $i=1,...,2c_T$,

$$\frac{1}{12k_2S_T}\sum_{i=1}^{2c_T}\inf_{x\in J_i}k_a^{(2)}(x) \le S_T^2[Z_T^{c_T}(a)-\int_{-c_T/S_T}^{c_T/S_T}k_a(x)dx] \le \frac{1}{12k_2S_T}\sum_{i=1}^{2c_T}\sup_{x\in J_i}k_a^{(2)}(x).$$

Now, by hypothesis $k_a^{(2)}(\cdot)$ is continuous from Assumption 4(a) and $c_T=c+o(\frac{1}{S_T})$. Therefore the conclusion holds since both LHS and RHS of the above inequalities converge to $\lim_{S_T\to\infty}\int_{-c_T/S_T}^{c_T/S_T}k_a^{(2)}(x)dx/12k_2$, that is, $\int_{-c}^{c}k_a^{(2)}(x)dx/12k_2$, uniformly a if

$$\lim_{T\to\infty}\sup_{a\in\mathbb{R}}\left|\frac{1}{12k_2S_T}\sum_{i=1}^{2c_T}\sup_{x\in J_i}k_a^{(2)}(x)-\frac{1}{12k_2S_T}\sum_{i=1}^{2c_T}\inf_{x\in J_i}k_a^{(2)}(x)\right|=0.$$

The case $c_T/S_T\to\infty$ and $k(x)$ with unbounded support is considered here; $k(x)$ with bounded support follows straightforwardly. Define $l_T=-MS_T+c_T+1$ and $u_T=MS_T+c_T$ and let

$$D_T(a) = \frac{1}{12k_2S_T}\sum_{i=1}^{2c_T}(\sup_{x\in J_i}-\inf_{x\in J_i})k_a^{(2)}(x) = \frac{1}{12k_2S_T}(\sum_{i=1}^{l_T-1}+\sum_{i=l_T}^{u_T}+\sum_{i=u_T+1}^{2c_T})(\sup_{x\in J_i}-\inf_{x\in J_i})k_a^{(2)}(x).$$

To simplify the notation write $s_{i,T}(a)=\sup_{x\in J_i}k_a^{(2)}(x)-\inf_{x\in J_i}k_a^{(2)}(x)$ and $S_T(a)=\frac{1}{12k_2S_T}\sum_{i=1}^{2c_T}s_{i,T}(a)$ and $k^{(0)}(x)=k(x)$.

For all $M>0$, there exists a T^* such that, for all $T>T^*$, $c_T/S_T>M$ since $c_T/S_T\to\infty$. Now, as J_i is compact, there exist $x_i^{\text{S}}[a]\in J_i$ and $x_i^{\text{I}}[a]\in J_i$ such that $\sup_{x\in J_i}k_a^{(2)}(x)=k_a^{(2)}(x_i^{\text{S}}[a])$ and $\inf_{x\in J_i}k_a^{(2)}(x)=k_a^{(2)}(x_i^{\text{I}}[a])$, $i=1,...,2c_T$. For T large enough, there exists a constant $M>0$ such that $\left|k^{(j)}(x)\right|<\varepsilon$, $j=0,1,2$, for all $|x|\ge M$ as $\lim_{|x|\to\infty}k^{(j)}(x)=0$, $j=0,1,2$, from Assumption 4(c). Hence, noting $k_a^{(2)}(x)=k^{(2)}(x+a)k(x)+2k^{(1)}(x)k^{(1)}(x+a)+k^{(2)}(x)k(x+a)$, $\left|k_a^{(2)}(x)\right|\le\varepsilon g_a(x)$ for $|x|\ge M$ where $g_a(x)=\left|k^{(2)}(x+a)\right|+2\left|k^{(1)}(x+a)\right|+|k(x+a)|$. Now note that for T large enough we have

$$\left|\frac{1}{12k_2S_T}\sum_{i=1}^{l_T-1}(\sup_{x\in J_i}-\inf_{x\in J_i})k_a^{(2)}(x)\right| \le \frac{1}{12k_2S_T}\varepsilon\sum_{i=1}^{l_T-1}[g_a(x_i^{\text{S}}[a])+g_a(x_i^{\text{I}}[a])],$$

$$\left|\frac{1}{12k_2S_T}\sum_{i=u_T+1}^{2c_T}(\sup_{x\in J_i}-\inf_{x\in J_i})k_a^{(2)}(x)\right| \le \frac{1}{12k_2S_T}\varepsilon\sum_{i=u_T+1}^{2c_T}[g_a(x_i^{\text{S}}[a])+g_a(x_i^{\text{I}}[a])].$$

Both sums are $o(1)$ as $\varepsilon > 0$ is arbitrary and, since both sums have $c_T - MS_T$ terms, from Assumption 4(b), $\sum_{i=1}^{l_T-1}[g_a(x_i^{\mathrm{s}}[a]) + g_a(x_i^{\mathrm{I}}[a])]/S_T = O(1)$ and $\sum_{i=u_T+1}^{2c_T}[g_a(x_i^{\mathrm{s}}[a]) + g_a(x_i^{\mathrm{I}}[a])]/S_T = O(1)$.

Now $|x_i^{\mathrm{s}}[a]|, |x_i^{\mathrm{I}}[a]| \leq M$, $i = l_T, ..., u_T$.

For $|a| > 2M$, $g_a(x_i^{\mathrm{s}}[a]), g_a(x_i^{\mathrm{I}}[a]) < \varepsilon$. Consequently,

$$\left|\frac{1}{12k_2S_T}\sum_{i=l_T}^{u_T}(\sup_{x\in J_i} - \inf_{x\in J_i})k_a^{(2)}(x)\right| \leq \frac{1}{12k_2S_T}\varepsilon\sum_{i=l_T}^{u_T}[g_0(x_i^{\mathrm{s}}[a]) + g_0(x_i^{\mathrm{I}}[a])] = o(1),$$

since, as above, ε is arbitrary and $\sum_{i=l_T}^{u_T}[g_0(x_i^{\mathrm{s}}[a]) + g_0(x_i^{\mathrm{I}}[a])]/S_T = O(1)$.

For $|a| \leq 2M$, because the set $a \in [-2M, 2M]$ is compact, only equicontinuity of $\sum_{i=l_T}^{u_T}(\sup_{x\in J_i} - \inf_{x\in J_i})$ $k_a^{(2)}(x)/12k_2S_T$ need be demonstrated as pointwise convergence is uniform on compact sets; see Rudin (1976, Exercise 16, p.168). That is, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$\left|\frac{1}{12k_2S_T}\sum_{i=l_T}^{u_T}(\sup_{x\in J_i} - \inf_{x\in J_i})[k_{a_1}^{(2)}(x) - k_{a_2}^{(2)}(x)]\right| < \varepsilon$$

for all $|a_1 - a_2| < \delta$. By T

$$\left|\frac{1}{12k_2S_T}\sum_{i=l_T}^{u_T}(\sup_{x\in J_i} - \inf_{x\in J_i})[k_{a_1}^{(2)}(x) - k_{a_2}^{(2)}(x)]\right| \leq \frac{1}{12k_2S_T}\sum_{i=l_T}^{u_T}\left|(\sup_{x\in J_i} - \inf_{x\in J_i})[k_{a_1}^{(2)}(x) - k_{a_2}^{(2)}(x)]\right|.$$

For every $\varepsilon > 0$, for T large enough. there exists a $\delta > 0$ such that

$$\left|(\sup_{x\in J_i} - \inf_{x\in J_i})[k_{a_1}^{(2)}(x) - k_{a_2}^{(2)}(x)]\right| < \frac{6k_2\varepsilon}{M}$$

because $k_a^{(2)}(x)$ is continuous in x and a, from the continuity of $k^{(j)}(x)$, $j = 0, 1, 2$, by Assumption 4(a), and, thus, uniformly continuous in x and a as $(x, a) \in [-M, M] \times [-2M, 2M]$ by the compactness of $[-2M, 2M]$.■

Let $c_T = T - 1$ and $s = aS_T$.

Corollary K.1. *Suppose $k(\cdot)$ satisfies Assumptions 4(a) and 4(b) respectively. Then*

$$Z_T^{c_T}(a) = \frac{1}{k_2}\int_{-c_T/s_T}^{c_T/s_T} k_a(x)dx + \frac{1}{12k_2S_T^2}\int_{-\infty}^{\infty} k_a^{(2)}(x)dx + o(\frac{1}{S_T^2})$$

uniformly a.

Proof. The proof follows that of Lemma K.1 above. Given the interval $[-c_T/S_T, c_T/S_T]$, again define the subintervals $J_i = [x_{i-1}, x_i]$ of equal length $1/S_T$, i.e., $x_i = (i - c_T)/S_T$, $(i = 1, ..., 2c_T)$, with $x_0 = -c_T/S_T$. Then, as before,

$$Z_T^{c_T}(a) = \frac{1}{2k_2S_T}\sum_{i=1}^{2c_T}(k_a(x_{i-1}) + k_a(x_i))$$

and

$$\frac{1}{12k_2S_T}\sum_{i=1}^{2c_T}\inf_{x\in J_i}k_a^{(2)}(x)\le S_T^2[Z_T^{c_T}(a)-\int_{-c_T/S_T}^{c_T/S_T}k_a(x)dx]\le\frac{1}{12k_2S_T}\sum_{i=1}^{2c_T}\sup_{x\in J_i}k_a^{(2)}(x).$$

The conclusion holds since both LHS and RHS of the above inequalities converge to $\lim_{S_T\to\infty}\int_{-c_T/S_T}^{c_T/S_T}k_a^{(2)}(x)dx/12$, i.e., $\int_{-\infty}^{\infty}k_a^{(2)}(x)dx/12$.■

Let

$$\tilde{k}_T(\frac{s}{S_T})=\sum_{t=1-T}^{T-1}k_{tT}(\frac{s}{S_T}).$$

LEMMA K.2. *Suppose Assumptions 2 and 4 hold. Then,*

$$\begin{aligned}\tilde{k}_T(\frac{s}{S_T}) &= k_T(\frac{s}{S_T})+|s|\,O(\frac{1}{T^bS_T})\\ &= k_T(\frac{s}{S_T})+O((\frac{S_T}{T})^{b-1/2})\end{aligned}$$

uniformly s.

PROOF. Consider the difference

$$\begin{aligned}\tilde{k}_T(\frac{s}{S_T})-k_T(\frac{s}{S_T}) &= \frac{1}{k_2S_T}\left(\sum_{t=1-T}^{T-1}-\sum_{t=\max[1-T,1-T+s]}^{\min[T-1,T-1+s]}\right)k\left(\frac{t-s}{S_T}\right)k\left(\frac{t}{S_T}\right)\\ &= \frac{1}{k_2S_T}\left(\sum_{t=\min[T,T+s]}^{T-1}+\sum_{t=1-T}^{\max[-T,-T+s]}\right)k\left(\frac{t-s}{S_T}\right)k\left(\frac{t}{S_T}\right).\end{aligned}$$

Firstly, if $s\ge 0$, $\min[T,T+s]=T$ and $\max[-T,-T+s]=-T+s$. Then, by CS, using Assumption 2(b),

$$\begin{aligned}\frac{1}{S_T}\left|\sum_{t=1-T}^{-T+s}k\left(\frac{t-s}{S_T}\right)k\left(\frac{t}{S_T}\right)\right| &\le \frac{1}{S_T}\sum_{t=1-T}^{-T+s}\left|k\left(\frac{t-s}{S_T}\right)k\left(\frac{t}{S_T}\right)\right|\\ &\le \frac{1}{S_T^{1/2}}(\sum_{t=1-T-s}^{-T}\left|k\left(\frac{t}{S_T}\right)\right|^2)^{1/2}(\frac{1}{S_T}\sum_{t=1-T}^{-T+s}\left|k\left(\frac{t}{S_T}\right)\right|^2)^{1/2}\\ &\le \frac{1}{S_T^{1/2}}(\sum_{t=-\infty}^{-T}\left|k\left(\frac{t}{S_T}\right)\right|^2)^{1/2}(k_2^{1/2}+o(1))\\ &\le C_k\frac{1}{S_T^{1/2-b}}(\sum_{t=T}^{\infty}|t|^{-2b})^{1/2}O(1)\\ &\le O(S_T^{b-1/2})(\int_T^{\infty}|t|^{-2b}\,dt)^{1/2}\\ &= O(S_T^{b-1/2})(T^{1-2b})^{1/2}=O((\frac{S_T}{T})^{b-1/2})\end{aligned}$$

or

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

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

$$\leq k_{\max}\frac{|s|}{T^b S_T} = |s|\,O(\frac{1}{T^b S_T}).$$

uniformly s. Secondly, if $s \leq 0$, $\min[T, T+s] = T+s$ and $\max[-T, -T+s] = -T$. Then, similarly, uniformly s,

$$\frac{1}{S_T}\left|\sum_{t=T+s}^{T-1} k\left(\frac{t-s}{S_T}\right)k\left(\frac{t}{S_T}\right)\right| \leq \frac{1}{S_T^{1/2}}(\sum_{t=T}^{T-1-s}\left|k\left(\frac{t}{S_T}\right)\right|^2)^{1/2}(\frac{1}{S_T}\sum_{t=T+s}^{T-1}\left|k\left(\frac{t}{S_T}\right)\right|^2)^{1/2}$$

$$\leq C_k\frac{1}{S_T^{1/2-b}}(\sum_{t=T}^{\infty} t^{-2b})^{1/2}O(1) = O((\frac{S_T}{T})^{b-1/2})$$

or

$$\frac{1}{S_T}\left|\sum_{t=T+s}^{T-1} k\left(\frac{t-s}{S_T}\right)k\left(\frac{t}{S_T}\right)\right| \leq k_{\max}\frac{1}{S_T}\sum_{t=T}^{T-1-s}\left|k\left(\frac{t}{S_T}\right)\right|$$

$$\leq k_{\max}\frac{|s|}{T^b S_T} = |s|\,O(\frac{1}{T^b S_T}).$$

Therefore,

$$\tilde{k}_T(\frac{s}{S_T}) - k_T(\frac{s}{S_T}) = O((\frac{S_T}{T})^{b-1/2})$$

uniformly s.■

Again let $a = s/S_T$.

LEMMA K.3. *Suppose $S_T \to \infty$, $S_T/T \to 0$ and Assumptions 2(b)(c) and 4(a)-(b) hold. Then, if Assumption 4(d) is satisfied,*

$$k^*(\frac{s}{S_T}) = \tilde{k}_T(\frac{s}{S_T}) + \frac{1}{12k_2S_T^2}\int_{-\infty}^{\infty} k_a^{(2)}(x)dx + O((\frac{S_T}{T})^{b-1/2}) + o(\frac{1}{S_T^2})$$

uniformly s. If $\int_{-\infty}^{\infty} k_a^{(2)}(x)dx = 0$ the remainder is

$$O((\frac{S_T}{T})^{b-1/2}) + o(\frac{1}{S_T^2}).$$

PROOF. Set $c_T = T-1$ and consider

$$\sum_{t=-c_T}^{c_T} k_{tT}(\frac{s}{S_T}) = Z_T^{c_T}(a) + O(\frac{S_T^{b-1}}{T^b})$$

uniformly s, where $Z_T^{c_T}(a)$ is defined above and $a = s/S_T$. To see this, by Assumption 4(e),

$$\begin{aligned} \left| k_{c_T T}(\frac{s}{S_T}) \right| &\leq k_{\max} \frac{1}{k_2 S_T} \left| k(\frac{c_T}{S_T}) \right| \\ &\leq C_k k_{\max} \frac{1}{k_2 S_T} \left| \frac{c_T}{S_T} \right|^{-b} \\ &\leq O(\frac{S_T^{b-1}}{T^b}) \end{aligned}$$

uniformly s. Similarly, uniformly s,

$$\left| k_{-c_T T}(\frac{s}{S_T}) \right| \leq O(\frac{S_T^{b-1}}{T^b}).$$

By Corollary K.1,

$$Z_T^{c_T}(a) - \frac{1}{k_2} \int_{-c_T/S_T}^{c_T/S_T} k_a(x) dx = \frac{1}{12 k_2 S_T^2} \int_{-\infty}^{\infty} k_a^{(2)}(x) dx + o(\frac{1}{S_T^2})$$

uniformly a. Now by CS, also using Assumption 4(d),

$$\begin{aligned} \left| \int_{c_T/S_T}^{\infty} k_a(x) dx \right| &\leq \int_{\infty}^{c_T/S_T} |k(x+a)k(x)| \, dx \\ &\leq (\int_{\infty}^{c_T/S_T} k(x+a)^2 dx)^{1/2} (\int_{\infty}^{c_T/S_T} k(x)^2 dx)^{1/2} \\ &\leq C_k (\int_{-\infty}^{\infty} k(x)^2 dx)^{1/2} (\int_{\infty}^{c_T/S_T} |x|^{-2b} \, dx)^{1/2} \\ &= C_k k_2^{1/2} [|x|^{1-2b}]_{\infty}^{c_T/S_T} \\ &= O((\frac{S_T}{T})^{b-1/2}) \end{aligned}$$

uniformly a. Similarly, uniformly a,

$$\left| \int_{-\infty}^{-c_T/S_T} k_a(x) dx \right| \leq O((\frac{S_T}{T})^{b-1/2}). \blacksquare$$

Corollary K.2. *Suppose $S_T \to \infty$, $S_T/T \to 0$ and Assumptions 2(b)(c) and 4(a)-(c) hold. Then, if Assumption 4(d) is satisfied,*

$$k_2 = \frac{1}{S_T} \sum_{s=1-T}^{T-1} k(\frac{s}{S_T})^2 + O((\frac{S_T}{T})^{b-1/2}) + o(\frac{1}{S_T^2}).$$

Proof. Set $s = 0$ in Lemma K.3.$\blacksquare$

The proof of the following Lemma is based on Smith (2011, Proof of Lemma C.1, p.1231-1232).

LEMMA K.4. *If $S_T \to \infty$, $S_T/T \to 0$ and $k_T(\cdot) \to k(\cdot)$ a.e., then, for $r > 0$, if $\int_{-\infty}^{\infty} |x|^r \bar{k}(x)dx < \infty$,*

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

uniformly s.

PROOF. Note that

$$\frac{1}{S_T}\sum_{s=1-T}^{T-1} |s|^r \left| k(\frac{s}{S_T}) \right| = S_T^r \int_{(1-T)/S_T}^{(T-1)/S_T} |x|^r |k_{T,s}(x))| \, dx$$

writing $k_{T,s}(x) = k((s-1)/S_T)$, $(s-1)/S_T \le x < s/S_T$, if $s \le 0$, $k(s/S_T)$, $(s-1)/S_T < x \le s/S_T$, if $s > 0$. Since $\left| \int_{(1-T)/S_T}^{(T-1)/S_T} |x|^r |k(x)| \, dx \right| \le \int_{-\infty}^{\infty} |x|^r \bar{k}(x)dx$, the result follows from the dominated convergence theorem noting $k_{T,s}(\cdot) \to k(\cdot)$ almost everywhere and $|k(\cdot)| \le \bar{k}(\cdot)$.■

LEMMA K.5. The optimal kernel function

$$k(x) = (\frac{5\pi}{8})^{1/2}\frac{1}{x}J_1(\frac{6\pi x}{5})$$

satisfies the hypotheses of Assumptions 2 and 4.

PROOF. To simplify the exposition $k(x)$ may be re-expressed as $c_k f(y)$, where $f(y) = J_1(y)/y$, $y = c_0 x$, $c_0 = 6\pi/5$ and $c_k = (6\pi/5)(5\pi/8)^{1/2}$. The result of Lemma K.5 for $k(x)$ may be equivalently proved for $f(y)$. To see this $k^{(1)}(x) = c_k f^{(1)}(y)(dy/dx)$, $k^{(2)}(x) = c_k f^{(2)}(y)(dy/dx)^2$ and $|k(x)| \le C|x|^{-3/2}$ if $|f(y)| \le C|y|^{-3/2}$, where $f^{(j)}(y) = d^j f(y)/dy^j$. The function $\bar{k}(x)$ satisfies Assumption 2(c) $\int_{-\infty}^{\infty} \bar{k}(x)dx < \infty$ if $\int_{-\infty}^{\infty} \bar{f}(x)dx < \infty$, where $\bar{f}(y) = \sup_{z \ge y} |f(z)|$ if $y \ge 0$ or $\sup_{z \le y} |f(z)|$ if $y < 0$, since $\bar{k}(x) = c_k \bar{f}(c_0 x)$.

A result used extensively is $\sup_{z \ge 0} |\sqrt{z} J_v(z)| \le C$, $v > 0$; see Olenko (2006, Theorem 2.1, p.456). Hence, for $z > 0$, $|J_v(z)| \le C/\sqrt{z}$ and, thus, for $y > 0$,

$$-C/y^{3/2} \le f(y) \le C/y^{3/2}.$$

Assumption 2(b). By direct inspection $f(y)$ achieves its maximum when $y \to 0$, i.e., $|f(y)| \le \lim_{y \to 0} |f(y)| = 1/2$.

Assumption 2(c). By T and $|f(y)| \le C$, $\left|\bar{f}(y)\right| \le C$. Thus, given symmetry about 0, for some positive constants c and C

$$\begin{aligned}\int_{-\infty}^{\infty} \bar{f}(y)dy &= 2\int_0^{\infty} \bar{f}(y)dy = 2(\int_0^c + \int_c^{\infty})\bar{f}(y)dy \\ &\leq 2C + 2\int_c^{\infty} \bar{f}(y)dy \\ &\leq 2C + 2\int_c^{\infty} \sup_{z\geq y} |J_1(z)/z|\, dy \\ &\leq 2C + 2\int_c^{\infty} \frac{1}{y^{3/2}} \sup_{z\geq y} \left|\sqrt{z}J_1(z)\right| dy \\ &\leq 2C + 2C\int_c^{\infty} \frac{1}{y^{3/2}} dy < C\end{aligned}$$

since $\sup_{z\geq 0} |\sqrt{z}J_v(z)| \leq C$, $v > 0$.

Assumption 4(a). Since $f^{(1)}(y) = -J_2(y)/y$ and $f^{(2)}(y) = -J_2(y)/y^2 + J_3(y)/y$ noting $d(J_v(y)/y^v)/dy = -J_{v+1}(y)/y^v$ (Watson, 1966, section 3.56, p.66), $f(y)$, $f^{(1)}(y)$ and $f^{(2)}(y)$ are continuous on $\mathbb{R}$.

Assumption 4(b). Given symmetry about 0, for some positive constants c and C

$$\begin{aligned}\int_{-\infty}^{\infty} \left|f^{(1)}(y)\right| dy &= 2\int_0^{\infty} \frac{1}{y} |J_2(y)|\, dy = 2(\int_0^c + \int_c^{\infty})\frac{1}{y} |J_2(y)|\, dy \\ &\leq 2C + 2\int_c^{\infty} \frac{1}{y} |J_2(y)|\, dy \leq 2C + 2C\int_c^{\infty} \frac{1}{y^{3/2}} dy) < C.\end{aligned}$$

Similarly

$$\begin{aligned}\int_{-\infty}^{\infty} \left|f^{(2)}(y)\right| dy &= 2(\int_0^c + \int_c^{\infty}) \left|f^{(2)}(y)\right| dy + 2\int_c^{\infty} \left|f^{(2)}(y)\right| dy \\ &\leq C + 2\int_c^{\infty} \left|-\frac{1}{y^2} J_2(y) + \frac{1}{y} J_3(y)\right| dy \\ &\leq C + \int_c^{\infty} \frac{1}{y^2} |J_2(y)|\, dy + \int_c^{\infty} \frac{1}{y} |J_3(y)|\, dy \\ &\leq C + \int_c^{\infty} \frac{1}{y^{5/2}} dy + \int_c^{\infty} \frac{1}{y^{3/2}} dy \leq C.\end{aligned}$$

The first derivative $f^{(1)}(y)$ achieves its maximum at $y^* = 2.29991036426349$ and $y^{**} = -y^*$, that is, $\left|f^{(1)}(y)\right| \leq \left|f^{(1)}(y^*)\right| = \left|f^{(1)}(y^{**})\right| = 0.179962865$. The second derivative $\left|f^{(2)}(y)\right|$ achieves its maximum when $y \to 0$, that is, $\left|f^{(2)}(y)\right| \leq \lim_{y\to 0} \left|-J_2(y)/y^2 + J_3(y)/y\right| = 1/8$.

Assumption 4(c). $f(y)$ is a symmetric function and, thus, $\lim_{y\to\infty} f(y) = \lim_{y\to-\infty} f(y)$. From above, for $y > 0$, $-C/y^{3/2} \leq f(y) \leq C/y^{3/2}$, $\lim_{|y|\to\infty} f(y) = 0$. Now $f^{(1)}(-y) = -f^{(1)}(y)$. Hence $\lim_{y\to-\infty} f^{(1)}(y) = -\lim_{y\to\infty} f^{(1)}(y)$. Similarly, for $y > 0$, since $f^{(1)}(y) = -J_2(y)/y$, $-C/y^{3/2} \leq f^{(1)}(y) \leq C/y^{3/2}$. Hence $\lim_{|y|\to\infty} f^{(1)}(y) = 0$. A similar argument shows $\lim_{|y|\to\infty} f^{(2)}(y) = 0$.

Assumption 4(d). $|J_1(y)/y| \leq C\,|y|^{-3/2}$ as $|J_1(y)| \leq C/\sqrt{y}$. Hence $b = 3/2$.■

Additional References

Cruz-Uribe, D. & C. J. Neugebauer (2002). Sharp error bounds for the trapezoidal rule and Simpson's rule. *J. Inequal. Pure and Appl. Math.* 3(4) Art. 49.

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

GRADSHTEYN, I. S. & I.M. RYZHIK (1980). *Tables of Integrals, Series and Products.* New York: Academic Press.

OLENKO, A. Y. (2006). Upper bound on $\sqrt{x}J_v(x)$ and its applications. *Integral Transforms and Special Functions* 17, 455-67.

POLITIS, D. N. & ROMANO, J. (1992b). A general resampling scheme for triangular arrays of α-mixing random variables with application to the problem of spectral density estimation. *Ann. Statist.* **20**, 1985-2007.

POLITIS, D. N., ROMANO, J.P. & WOLF, M. (1997): Subsampling for heteroskedastic time series. *J. Economet.* 81, 281-317.

RUDIN, W. (1976). *Principles of Mathematical Analysis.* Third edition. New York: McGraw-Hill.

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

WATSON, G. N. (1966). *A Treatise on the Theory of Bessel Functions.* Cambridge: Cambridge University Press.

WHITE, H. (1984). *Asymptotic Theory for Econometricians.* New York: Academic Press.