The Size-Power Tradeoff in HAR Inference: Online Supplement

Eben Lazarus	Daniel J. Lewis	James H. Stock
MIT Sloan	Federal Reserve Bank of NY	Harvard University

CONTENTS

S1. MONTE CARLO STUDY

The purpose of this Monte Carlo analysis is twofold. First, we assess the quality of the approximations to the size/power tradeoffs in the Gaussian location model. Second, we investigate the extent to which the theory derived for the Gaussian multivariate location model generalizes to time series regression with stochastic regressors.

S1.1 ESTIMATORS AND DESIGN

For a given kernel or WOS estimator, we use four values of b, chosen so that $v = 8$, 16, 32, and 64. The tests are labeled accordingly, e.g., NW16 is the Newey-West (Bartlett) test with $v = 16$ equivalent degrees of freedom. For reference, for $T = 200$, NW32 has a truncation parameter of $(3/2)T/v$, which rounds up to 10. For the WOS estimators, we consider tests with equal weights $w_j = 1/B$, so $v = B$. Tests use fixed-b critical values unless stated otherwise. We specifically examine the following HAR tests:

1. NW: Kernel estimator with Bartlett/Newey-West kernel, $k(x) = (1-|x|)\mathbf{1}(|x|\leq1)$.

2. KVB: The Kiefer-Vogelsang-Bunzel (2000) test, i.e., NW with $S = T$ ($v = 3/2$).

3. QS: $k(x) = 3[\sin(\pi x)/(\pi x) - \cos(\pi x)] /(\pi x)^2$.

4. EWP: Equal-weighted WOS estimator using the Fourier basis, $\{\phi_{2j-1}(s), \phi_{2j}(s)\} = \left\{\sqrt{2}\cos(2\pi js), \sqrt{2}\sin(2\pi js)\right\}$, $j = 1, \ldots, B/2$.

5. cos: Equal-weighted WOS estimator using the Type II cosine basis, $\{\phi_j(s)\} = \left\{\sqrt{2}\cos\left[\pi j\left(s - \frac{1/2}{T}\right)\right]\right\}$, $j = 1,\ldots, B$.

6. SS-basis: Equal-weighted split-sample WOS estimator (see Section S3.2 below).

In the location model, the data are generated according to equation (3) in the main text, where u_{it}, $i = 1,.., m$ are independent and follow either a Gaussian AR(1) or an ARMA(2,1), with all m disturbances having the same parameter values. For the regression model, the data are generated following $y_t = x_t'\beta + u_t$, where x_{it}, $i = 1,\ldots, m$ and u_t are independent Gaussian AR(1) processes. Under the null, $\beta = 0$. Under the local alternative, $\beta = T^{-1/2}\Sigma_{XX}^{-1}\Omega^{1/2}\delta$ for $m=1$, where δ is the local alternative index value and $\Sigma_{XX} = T^{-1}\sum_{t=1}^{T}x_t x_t'$ (for the location model, $\Sigma_{XX} = I$, as in the text). For $m = 2$, we set $\beta = T^{-1/2}\Sigma_{XX}^{-1}\Omega^{1/2}\delta_2$, with $\delta_2 = [\delta \quad 0]'$. We conduct 100,000 replications for each design.

S2.2 MONTE CARLO RESULTS

This section presents some representative results; additional results are contained in the working paper version of this paper (Lazarus, Lewis, and Stock (LLS, 2017)) and in Lazarus, Lewis, Stock and Watson (LLSW, 2018). All results are displayed in finite-sample counterparts of Figure 1. For these figures, the axes are not scaled, so that the units are the size distortion and the power loss. The theoretical tradeoffs from Theorem 4(ii) are shown as lines, and the Monte Carlo results are presented as scatter points.
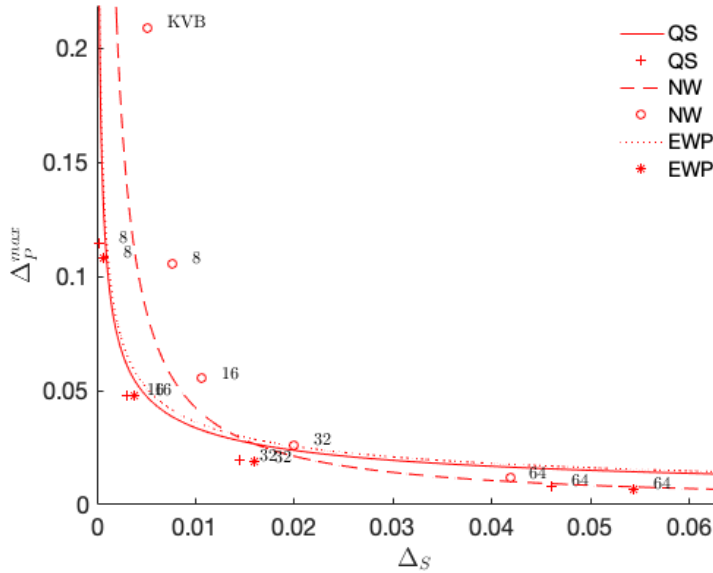
FIGURE S1. Theoretical (lines) and Monte Carlo (symbols) size distortion/power loss for QS, Newey-West, and EWP tests. Location model, $m = 1$, AR(1), $\rho = 0.5$, and $T = 200$.

Location model. Figure S1 presents results for QS, EWP, and NW tests in the location model with Gaussian AR(1) disturbances in the $m = 1$ case with AR parameter $\rho = 0.5$ and $T = 200$. The Monte Carlo results for QS and EWP are close to their theoretical curves. The small-b approximation is less good for Newey-West: the NW Monte Carlo scatter appears to follow a curve that has the same shape as the theoretical curve, but is shifted out. KVB is a limiting case of Newey-West with $S_T = T$ (so $b = 1$ and $\nu = 1.5$), that is, KVB is NW1.5, so KVB lies on the NW Monte Carlo curve.

LLS (2017) contains figures with additional results for the location model, which we discuss briefly here. For $m = 2$ with AR(1) errors, $\rho = 0.5$, and $T = 200$, the frontier fits the simulations slightly better for QS and EWP than in the $m = 1$ case, but somewhat worse for NW. We find in addition that, for $m = 1$ and 2 and with either AR(1) or ARMA(2,1) disturbances, fit (distance from the scatter points to their theoretical tradeoff) improves with T, deteriorates as $\omega^{(2)}$ increases, is better for $q = 2$ kernels than $q = 1$, and does not appreciably deteriorate as process parameters are changed holding $\omega^{(2)}$ constant. The first two results are unsurprising. Our interpretation of the third finding is that the
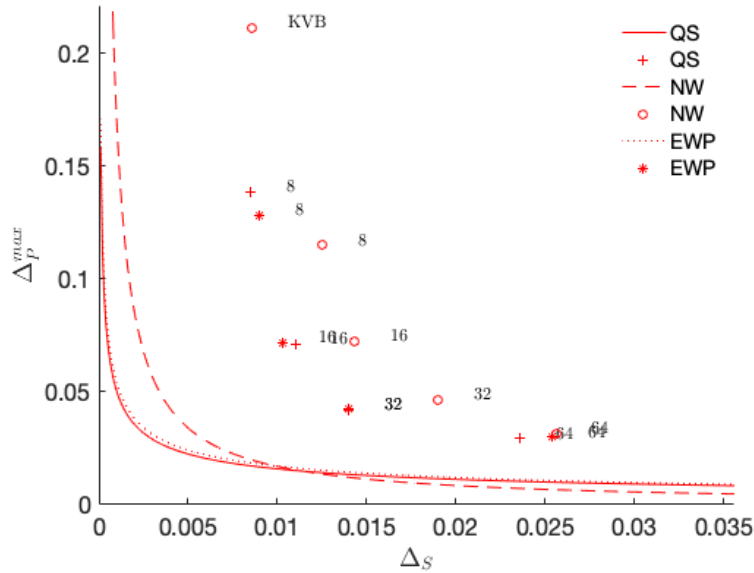
FIGURE S2. Theoretical (lines) and Monte Carlo (symbols) size distortion/power loss for QS, Newey-West, and EWP tests. Stochastic regressor, $m = 1$, AR(1), $\rho = 0.5$, $T = 200$. Theoretical curves are for the Gaussian location model.

order of approximation of the expansions is $o((bT)^{-q})$, so the remainder is of a smaller order for $q = 2$ than for $q = 1$ kernels. The larger values of b used with the NW kernel for a given ν may also play a role. Overall, the simulation results accord with the theory.

 Stochastic regressor. Figure S2 shows the QS, EWP, and NW tests on the coefficient on a single stochastic regressor, where both the regressor and dependent variable have AR(1) disturbances with $\rho = 0.5$ and $T = 200$ (intercept included in the regression but not tested). In this DGP, z_t is AR(1) but non-Gaussian. For reference, the theoretical tradeoff curves are shown for the Gaussian location model. It appears that this departure from Gaussianity results in poor performance of the Gaussian small-b asymptotic approximation and that there are missing terms in the expansion as suggested by the calculations in Velasco and Robinson (2001). This said, several key qualitative results in the theory continue to apply to the single stochastic regressor. First, for a given estimator, the Monte Carlo results map out a size-power tradeoff that has a shape similar to the Gaussian theoretical shape, just shifted out. Second, the tradeoffs for the QS and EWP
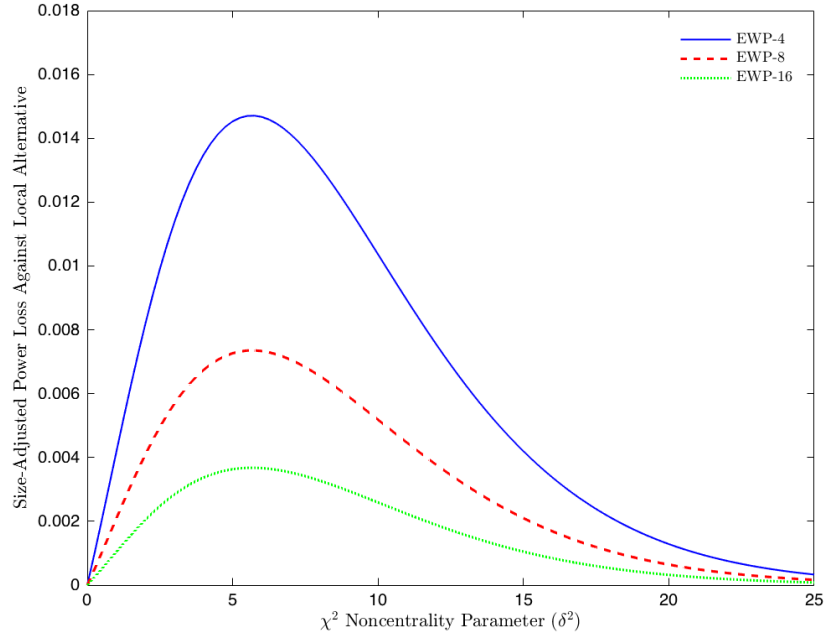
FIGURE S3. Small-*b* approximation to power loss for EWP test compared to same-sized QS test, for different values of *B* in the EWP test. The figure plots the expression in (30) as a function of δ for $m = 1$, $\alpha = 0.05$ (see Remark 5).

estimators are very close to each other. Third, the ranking across estimators is the same as suggested by the theory and confirmed in the Monte Carlo analysis of the location model, that is, the $q = 1$ tests are outperformed by the $q = 2$ tests. We find similar results for other designs, kernels, and values of *m*, and further that the approximation improves for higher values of *T*; again see LLS (2017) and LLSW (2018).

Overall, we can draw three conclusions. First, the theoretical frontiers provide a good description of estimator performance in the Gaussian location model. The fit is better for $q = 2$ kernels than $q = 1$. Second, Monte Carlo performance is consistent with the theory. In particular, the performance of $q = 2$ kernels is superior to that of $q = 1$ kernels in this design, and the cost of using EWP relative to QS is low. As further illustration of the latter theoretical result, Figure S3 plots the theoretical higher-order power loss from using EWP relative to QS as a function of δ for various values of *B*, as discussed in Remark 5. Third, the qualitative results for stochastic regressors are consistent with the theory for

the location model, however the Monte Carlo points no longer lie on the tradeoff derived for the Gaussian location model. We attribute this divergence of the theory and Monte Carlo results to the non-Gaussianity of z_t in the stochastic regressor case.[1]

S2. SUPPLEMENTAL PROOFS

S2.1 ADDITIONAL PROOFS OF MAIN RESULTS

We first provide two preliminary results needed for Theorems 1 and 5, and then prove Theorem 1. Assume for the remainder of the Supplement that Assumptions 1-4 hold.

Lemma S1. For any weights $\{w_j\}$, the Fourier basis minimizes $\left|\sum_{j=1}^{B} w_j \int_0^1 \phi_j(s)\phi_j''(s)ds\right|$ across all WOS estimators up to an error of order $o(1/T)$.

Proof of Lemma S1: The complex Fourier expansion of any ϕ_j in a given basis is again

$$\phi_j(s) = \sum_{l=-\infty}^{\infty} a_{jl}e^{-i2\pi ls} , \tag{S.1}$$

where $\{a_{jl}\}_l$ are the (inverse) Fourier coefficients of $\phi_j(s)$. For any orthonormal series,

$$1 = \int_0^1 |\phi_j(s)|^2 ds = \sum_{l,l'=-\infty}^{\infty} a_{jl}\overline{a}_{jl'} \int_0^1 e^{-i2\pi ls}e^{i2\pi l's}ds = \sum_l |a_{jl}|^2 , \tag{S.2}$$

$$\text{and } 0 = \int_0^1 \phi_j(s)\overline{\phi_{j'\neq j}(s)}ds = \sum_l a_{jl}\overline{a}_{j'\neq j,l} , \tag{S.3}$$

where $\overline{a}_{jl}$ is the complex conjugate of a_{jl}. The minimization problem for real ϕ_j is then

---

[1] In results available in LLS (2017), we also examined the performance of tests based on plug-in higher-order corrected critical values based on equation (20) of the text, using an estimated value of $\omega^{(q)}$. HAR tests using these plug-in critical values generally worked poorly compared to tests using standard fixed-b critical values.

$$\min_{\{a_{jl}\}}\left|\sum_{j=1}^{B}w_j\int_0^1\phi_j(s)\overline{\phi_j''(s)}ds\right| \Leftrightarrow \min_{\{a_{jl}\}}\left|\sum_{j=1}^{B}w_j\sum_{l,l'}a_{jl}\bar{a}_{jl'}4\pi^2l^2\int_0^1 e^{-i2\pi ls}e^{i2\pi l's}ds\right|$$

$$\Leftrightarrow \min_{\{a_{jl}\}}\sum_{j=1}^{B}w_j\sum_l\left|a_{jl}\right|^2 l^2 \tag{S.4}$$

subject to the constraints (S.2) and (S.3), along with $a_{j0}=\int_0^1\phi_j(s)ds=0$.

For now set the summation limits in (S.1) to $\pm\overline{T}$ for some $\overline{T}>T$, so (S.4) can be written $\min_A \mathrm{tr}((AW)^*D(AW)) \Leftrightarrow \min_A \mathrm{tr}(W^2A^*DA)$ s.t. $A^*A=I_B$, where $A=[A_1\ A_{-1}\ A_2\ A_{-2}\ \dots\ A_{\overline{T}}\ A_{-\overline{T}}]'$, $A_l=[a_{1l}\ a_{2l}\ \dots\ a_{Bl}]'$, A^* is the conjugate transpose of A, $W=\mathrm{diag}([\sqrt{w_1}\ \sqrt{w_2}\ \dots\ \sqrt{w_B}])$, and $D=\mathrm{diag}([1\ 1\ 4\ 4\ \dots\ \overline{T}^2\ \overline{T}^2])$. From (S.4), the objective is linear in the entries of $\mathbf{A}_2=A\circ\overline{A}$, where $\circ$ is the Hadamard product and $\overline{A}$ is the elementwise complex conjugate of A.

It will be convenient to transform this problem to $\min_{\tilde{A}}\mathrm{tr}(\tilde{W}^2\tilde{A}^*D\tilde{A})$ s.t. $\tilde{A}^*\tilde{A}=I_{2\overline{T}}$, where $\tilde{W}=[I_B\ 0_{B\times(2\overline{T}-B)}]'W[I_B\ 0_{B\times(2\overline{T}-B)}]$ is padded with zeros relative to W, and $\tilde{A}=[A\ \ H]$ for some $2\overline{T}\times(2\overline{T}-B)$ matrix H, so $\tilde{A}$, $\tilde{W}$, and D are $2\overline{T}\times 2\overline{T}$. The objective is again linear in the entries of $\widetilde{\mathbf{A}}_2=\tilde{A}\circ(\overline{\tilde{A}})$, which is doubly stochastic since $\tilde{A}^*\tilde{A}=I_{2\overline{T}}$ implies $\tilde{A}\tilde{A}^*=I_{2\overline{T}}$. Thus

$$\min_{\tilde{A}\ \text{s.t.}\ \tilde{A}^*\tilde{A}=I_{2\overline{T}}}\mathrm{tr}\left(\tilde{W}^2\tilde{A}^*D\tilde{A}\right)\geq\min_{\Upsilon}\sum_{j,l}w_jl^2\gamma_{jl}, \tag{S.5}$$

where Υ is a doubly stochastic matrix containing the values $\{\gamma_{jl}\}$. The right side of (S.5) is linear in the entries of Υ, and the set of doubly stochastic matrices $\{\Upsilon\}$ is compact and convex. Thus the minimum of the right side is obtained at an extreme point of this set. By Birkhoff's Theorem (e.g., Bhatia (1997, p. 37)), the extreme points of $\{\Upsilon\}$ are the permutation matrices. Any permutation matrix P is unitary, so (S.5) in fact holds with equality, and we can select $\tilde{A}=\arg\min_P\mathrm{tr}(\tilde{W}^2P'DP)$.

Note that D and $\tilde{W}^2$ are psd and diagonal, and D has its diagonal terms (eigenvalues) in ascending order. Given weights $\{w_j\}$, assume first that the weights are ordered

descendingly, $w_1 \geq w_2 \geq \ldots \geq w_B$, and therefore that $\tilde{W}^2$ has its diagonal terms (eigenvalues) in descending order. In this case, the minimum of the objective is achieved trivially by $\tilde{A} = P = I_{2\bar{T}}$, so that the minimizing A is given by the first B columns of $I_{2\bar{T}}$; equivalently, $a_{2j'-1,j'} = a_{2j',-j'} = 1$, $j' = 1,\ldots,B/2$, $a_{jl} = 0$ otherwise. Thus from (S.1),

$$\{\phi_{2j'-1}(s),\phi_{2j'}(s)\} = \{e^{-i2\pi j's},e^{i2\pi j's}\} = \{\sqrt{2}\cos(2\pi j's),\sqrt{2}\sin(2\pi j's)\}, j' = 1,\ldots,B/2, \text{ so}$$

we have in fact selected the Fourier basis. This applies for any $\bar{T}$ used in the finite truncation of (S.1), so we can set arbitrary $\bar{T} \gg T$ so as to apply Jackson's inequality to obtain that the statement holds to $o(1/T)$.

If the weight values are not in descending order, use that $\tilde{W}^2$ is psd and diagonal to write its singular value decomposition as $\tilde{W}^2 = V\tilde{W}^2_{desc}V'$, where $\tilde{W}^2_{desc}$ is the diagonal matrix containing the eigenvalues (diagonal terms) of $\tilde{W}^2$ ordered descendingly. Then the problem can be rewritten as $\min_{\tilde{A}_V} \text{tr}(\tilde{W}^2_{desc}\tilde{A}^*_V D\tilde{A}_V)$ subject to $\tilde{A}^*_V\tilde{A}_V = I_{2\bar{T}}$, where $\tilde{A}_V = \tilde{A}V$, so that V has been absorbed into the argument to be minimized. But this is the same problem as in the case above, with $\tilde{W}^2$ having its values in descending order. Thus the minimum achieved for (S.5) is equivalent for any set of weights regardless of their ordering. Thus for any set of weights, it is without loss to set them in descending order, in which case the Fourier basis again achieves the minimum, completing the proof. $\square$

Lemma S2. For any WOS test using the Fourier basis, the value $\ell^{(2)}(k)$ is minimized with respect to $\{w_j\}$ by the use of QS weights: $w^*_j = \bar{w}_{QS}[1-(j/B)^2]$, where $\bar{w}_{QS} = \frac{6B}{(B-1)(4B+1)}$.

Proof of Lemma S2: Note first that $\bar{w}_{QS} = 6B/[(B-1)(4B+1)]$ is set so that $\sum_{j=1}^B w^*_j = 1$. From (44), given the use of the Fourier basis, minimizing $\ell^{(2)}(k)$ is equivalent to minimizing $\left(\sum_{j=1}^B w_j j^2\right)^{1/2}\left(\sum_{j=1}^B w^2_j\right)$. Write $B_{QS} = B$. For any alternative set of

weights w_j, write $w_j = w_j^* + \epsilon_j$. We allow for the sequence B_{alt} for to this set of weights to differ from B_{QS} (both must meet Assumption 4). If $B_{alt} > B_{QS}$, then set $w_j^* = 0$ for $j > B_{QS}$, so that $w_j = \epsilon_j \geq 0$ for $B_{QS} < j \leq B_{alt}$. If $B_{alt} < B_{QS}$, then correspondingly $w_j = 0$ for $B_{alt} < j \leq B_{QS}$. Write $\bar{B} = \max(B_{alt}, B_{QS})$. Since $\sum_{j=1}^{\bar{B}} w_j = 1$, we have $\sum_{j=1}^{\bar{B}} \epsilon_j = 0$.

We then equate higher-order size for the two estimators and show that QS dominates with respect to power. Equating higher-order size requires $\sum_{j=1}^{\bar{B}} w_j j^2 = \sum_{j=1}^{\bar{B}} w_j^* j^2$, so

$\sum_{j=1}^{\bar{B}} \epsilon_j j^2 = 0$. Further, $\sum_{j=1}^{\bar{B}} w_j^2 = \sum_{j=1}^{\bar{B}} (w_j^*)^2 + 2\sum_{j=1}^{\bar{B}} \epsilon_j w_j^* + \sum_{j=1}^{\bar{B}} \epsilon_j^2$, and

$$\sum_{j=1}^{\bar{B}} \epsilon_j w_j^* = \bar{w}_{QS} \sum_{j=1}^{B_{QS}} \left[1 - \left(\frac{j}{B_{QS}} \right)^2 \right] \epsilon_j = \bar{w}_{QS} \left\{ \sum_{j=1}^{\bar{B}} \left[1 - \left(\frac{j}{B_{QS}} \right)^2 \right] \epsilon_j + \sum_{j=B_{QS}+1}^{B_{alt}} \left[\left(\frac{j}{B_{QS}} \right)^2 - 1 \right] \epsilon_j \right\}. \quad (S.6)$$

The first term in (S.6) is zero given the steps above. For the second term, if $B_{alt} > B_{QS}$, then as above $w_j = \epsilon_j \geq 0$ for $B_{QS} < j \leq B_{alt}$, and therefore $\sum_{j=B_{QS}+1}^{B_{alt}} [(j/B_{QS})^2 - 1]\epsilon_j \geq 0$

(with equality if $B_{alt} < B_{QS}$). Thus $\sum_{j=1}^{\bar{B}} \epsilon_j w_j^* \geq 0$. It is further trivially the case that $\sum_{j=1}^{\bar{B}} \epsilon_j^2 \geq 0$.

We conclude that $\sum_{j=1}^{\bar{B}} w_j^2 \geq \sum_{j=1}^{\bar{B}} (w_j^*)^2$, with equality if and only if $\epsilon_j = 0$ for all j. Therefore

QS attains greater higher-order power for equivalent higher-order size, and thus

minimizes $\ell^{(2)}(k)$, relative to all alternative WOS estimators using the Fourier basis. $\quad \square$

Proof of Theorem 1:

(i) For kernel estimators, under the equivalent of our Assumptions 1, 2, and 4, Sun (2014b, p. 675) gives equation (15) directly. For WOS estimators, write

$$E\hat{\Omega}_j^{OS} = E\left(\sqrt{\frac{1}{T}} \sum_{t=1}^{T} \phi_j(t/T)\hat{z}_t \right)\left(\sqrt{\frac{1}{T}} \sum_{t=1}^{T} \phi_j(t/T)\hat{z}_t \right)'$$

$$= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\phi_j(t/T)\phi_j(s/T)\Gamma_{s-t} + O(1/T)$$

$$= \sum_{u=-(T-1)}^{T-1}\frac{1}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\Gamma_u + O(1/T), \qquad \text{(S.7)}$$

where the $O(1/T)$ term in the second line arises due to the approximation of $\hat{z}_t$ with z_t under Assumption 1 (see, for example, the proof of Theorem 2 in Sun (2011)). Thus,

$$E\hat{\Omega}^{WOS} - \Omega = \sum_{u=-(T-1)}^{T-1}\left\{\left[\sum_{j=1}^{B}w_j\frac{1}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\right]-1\right\}\Gamma_u - \sum_{|u|\geq T}\Gamma_u + O\left(\frac{1}{T}\right). \quad \text{(S.8)}$$

For $q \leq 2$ (shown later to be without loss of generality), by Assumptions 1(b) and 4,

$$\left|\sum_{|u|\geq T}\Gamma_u\right| \leq \sum_{|u|\geq T}|\Gamma_u| \leq \frac{1}{T^2}\sum_{|u|\geq T}|u|^2|\Gamma_u| = o\left(T^{-2}\right) = o\left((B/T)^q\right), \qquad \text{(S.9)}$$

so we may focus on the first summation in (S.8). Further, $T^{-1} = b^q T^{q-1}(bT)^{-q} = o((bT)^{-q})$ by Assumption 4, so that $O(1/T) = o((bT)^{-q})$.

We may then, following the device in Theorem 1(i) of Phillips (2005), write

$$E\hat{\Omega}^{WOS} - \Omega = \sum_{u=-L_T}^{L_T}\left\{\left[\sum_{j=1}^{B}\frac{w_j}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\right]-1\right\}\Gamma_u$$

$$+ \sum_{L_T<|u|<T}\left\{\left[\sum_{j=1}^{B}\frac{w_j}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\right]-1\right\}\Gamma_u + o\left(\left(\frac{B}{T}\right)^q\right), \quad \text{(S.10)}$$

where $L_T < T$ is a positive integer sequence chosen such that

$$\frac{T^q}{L_T^{q+\zeta}B^q} + \frac{L_T B}{T} \to 0, \qquad \text{(S.11)}$$

where ζ is as in Assumption 1(b). We have

$$\left|\sum_{L_T<|u|<T}\left\{\left[\sum_{j=1}^{B}\frac{w_j}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\right]-1\right\}\Gamma_u\right|$$

$$\leq \sum_{j=1}^{B}w_j\sum_{L_T<|u|<T}\left\|\left[\frac{1}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\right]-1\right\||\Gamma_u|$$

$$\leq 2 \sum_{L_T < |u| < T} |\Gamma_u| \leq 2 L_T^{-q-\zeta} \sum_{L_T < |u| < T} |u|^{q+\zeta} |\Gamma_u| = O\left(L_T^{-q-\zeta}\right) = o\left(\left(\frac{B}{T}\right)^q\right), \quad \text{(S.12)}$$

where the first inequality applies the triangle inequality, the second inequality uses that

$\left| \sum_{t=\max(1,u)}^{\min(T,T+u)} \phi_j(t/T)\phi_j((t-u)/T)/T \right| \leq \left| \sum_{t=1}^{T} (\phi_j(t/T))^2 /T \right| = 1$ by Cauchy-Schwarz

(and therefore that $\left| \left[\sum_{t=\max(1,u)}^{\min(T,T+u)} \phi_j(t/T)\phi_j((t-u)/T)/T \right] - 1 \right| \leq 2$), and where the first

equality in the last line applies Assumption 1(b). (Setting $T^{-1}\sum_{t=1}^{T}(\phi_j(t/T))^2 = 1$ in finite

samples for orthonormal ϕ_j is without loss; see after (S.20) below for further discussion.)

Equation (S.10) can thus be written, with $k_{B,T}^{WOS}(u/S)$ defined after (9) in the main text,

$$E\hat{\Omega}^{WOS} - \Omega = \sum_{u=-L_T}^{L_T} \left[k_{B,T}^{WOS}\left(\frac{u}{S}\right) - 1 \right] \Gamma_u + o\left(\left(\frac{B}{T}\right)^q\right)$$

$$= \sum_{u=-L_T}^{L_T} \left\{ \left[k_B^{WOS}\left(\frac{u}{S}\right) - 1 \right]\left(1 + o\left(\frac{1}{T}\right)\right) \right\} \Gamma_u + o\left(\left(\frac{B}{T}\right)^q\right)$$

$$= \left\{ \sum_{u=-L_T}^{L_T} \left[k_B^{WOS}\left(\frac{u}{S}\right) - 1 \right] \Gamma_u \right\}(1 + o(1)) + o\left(\left(\frac{B}{T}\right)^q\right), \quad \text{(S.13)}$$

where the second line uses (10), along with the fact that $k_{B,T}^{WOS}(x)$ and $k_{B,T}^{WOS\prime}(x)$ are

uniformly bounded for fixed B (since $|\phi_j(s)|, |\phi_j'(s)| \leq CB^{5/2}$ for some $C < \infty$ by

Assumption 3), to obtain that $k_{B,T}^{WOS} = k_B^{WOS} + o(1/T)$ by Riemann approximation, and the

third line uses that $o(L_T/T) = o(1)$ by (S.11).

Now note that under Assumption 3, in addition to $|k_B^{WOS}(x)| \leq 1$, we have $k_B^{WOS}(0) = 1$,

$k_B^{WOS}(x) = k_B^{WOS}(-x)$, and $k_B^{WOS}(x)$ is continuous since $\phi(u/T)$ is continuous for

$u/T \in [0,1]$. And since ϕ is twice continuously differentiable, $k_B^{WOS}(x)$ is twice

continuously differentiable on $[-B,0) \cup (0,B]$. Thus defining $k_{B,+}^{WOS\prime}(x)$ and $k_{B,+}^{WOS\prime\prime}(x)$ as

the first and second right derivatives, respectively, of $k_B^{WOS}(x)$, we have

$k_B^{WOS}(x) = 1 + k_{B,+}^{WOS\prime}(0)x + \frac{1}{2}k_{B,+}^{WOS\prime\prime}(0)x^2 + o(x^2)$ as $x \to 0^+$. Since $k_B^{WOS}(x)$ is even, the first

and second left derivatives satisfy $k_{B,-}^{WOS\prime}(x) = -k_{B,+}^{WOS\prime}(x)$ and $k_{B,-}^{WOS\prime\prime}(x) = k_{B,+}^{WOS\prime\prime}(x)$,

respectively, and thus $k_B^{WOS}(x) = 1 + k_{B,+}^{WOS\prime}(0)|x| + \frac{1}{2}k_{B,+}^{WOS\prime\prime}(0)x^2 + o(x^2)$ as $x \to 0^-$. Thus,

defining $g_{1,B} = -k_{B,+}^{WOS\prime}(0)$ and $g_{2,B} = -k_{B,+}^{WOS\prime\prime}(0)/2$, we can write

$$1 - k_B^{WOS}(x) = g_{1,B}|x| + g_{2,B}|x|^2 + o(|x|^2) \tag{S.14}$$

as $x \to 0$, from which it is clear from (11) that $k_B^{(q)}(0) = g_{q,B}$. Using this with (S.13) and

the fact that $S = T/B$, we can follow Priestley (1981, p. 459) and write

$$E\hat{\Omega}^{WOS} - \Omega = \left\{ -\left(\frac{B}{T}\right)^q \sum_{u=-L_T}^{L_T} \frac{1 - k_B^{WOS}(u/S)}{|u/S|^q} |u|^q \Gamma_u \right\}(1 + o(1)) + o\left(\left(\frac{B}{T}\right)^q\right)$$

$$= -2\pi\left(\frac{B}{T}\right)^q \left(k_B^{(q)}(0)s_z^{(q)}(0)\right)(1 + o(1)) + o\left(\left(\frac{B}{T}\right)^q\right)$$

$$= -2\pi\left(\frac{B}{T}\right)^q \left\{\left(\lim_{B\to\infty} k_B^{(q)}(0)\right)(1 + o(1))\right\} s_z^{(q)}(0) + o\left(\left(\frac{B}{T}\right)^q\right)$$

$$= -2\pi\left(\frac{B}{T}\right)^q \left(\lim_{B\to\infty} k_B^{(q)}(0)\right) s_z^{(q)}(0) + o\left(\left(\frac{B}{T}\right)^q\right). \tag{S.15}$$

Using that $\mu = 0$ for WOS estimators, (15) follows, with $k^{(q)}(0) = \lim_{B\to\infty} k_B^{(q)}(0)$.

(ii) Using equation (10), we have for $x > 0$ that

$$k_B^{WOS\prime}(x) = B^{-1}\sum_{j=1}^B w_j\left(-\int_{B^{-1}x}^1 \phi_j(s)\phi_j'(s - B^{-1}x)ds - \phi_j(B^{-1}x)\phi_j(0)\right), \tag{S.16}$$

so from part (i),

$$k_B^{(q)}(0) = g_{1,B} = -k_{B,+}^{WOS\prime}(0) = B^{-1}\sum_{j=1}^B w_j\left(\int_0^1 \phi_j(s)\phi_j'(s)ds + \left(\phi_j(0)\right)^2\right). \tag{S.17}$$

Integrating by parts, $\int_0^1 \phi_j(s)\phi_j'(s)ds = \phi_j(1)^2 - \phi_j(0)^2 - \int_0^1 \phi_j(s)\phi_j'(s)ds = \frac{\phi_j(1)^2 - \phi_j(0)^2}{2}$. Thus,

with $k^{(q)}(0) = \lim_{B\to\infty} k_B^{(q)}(0)$ and Assumption 3 guaranteeing the existence of the limit since

$w_j = O(B^{-1})$ and $\sum_j \phi_j(s)^2 = O(B^2)$, the first part of (16) follows. Similarly, for $x > 0$,

$$k_B^{WOS''}(x) = B^{-2}\sum_{j=1}^{B} w_j \left(\int_{B^{-1}x}^1 \phi_j(s)\phi_j''(s-B^{-1}x)ds\right). \tag{S.18}$$

Using that $k_B^{(2)}(0) = -k_{B,+}^{WOS''}(0)/2$ if $q = 2$, and taking $B \to \infty$, then delivers the second

part of (16), as the existence of this limit is again guaranteed under Assumption 3.

For the final statement, if $k^{(1)}(0) \neq 0$, then $q = 1$ follows immediately from (S.14) and

the definition of q after (11). If $k^{(1)}(0) = 0$, note from Lemma S1 and (16) that the Fourier

basis minimizes $k^{(2)}(0)$ across WOS estimators, so $k^{(2)}(0) > 0$ for all WOS estimators.

Thus (11) gives that $q \leq 2$, so from (S.14), if $k^{(1)}(0) = 0$, then $q = 2$, as stated. This

extends the classic result that psd kernel estimators have $q \leq 2$ to the implied mean

kernels of WOS estimators (and justifies the notational use of some $q \leq 2$ above (S.9)).

(iii) For kernel estimators, (17) restates Andrews (1991, Proposition 1(a)). For WOS

estimators, generalizing Sun (2011, p. 361) to the case of arbitrary WOS weights w_j,

$$\mathrm{var}\left(\mathrm{vec}\,\hat{\Omega}\right) = \Omega \otimes \Omega\left(I_{m^2} + K_{mm}\right)\frac{1}{T^2}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\left[\sum_{j=1}^{B} w_j \phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{\tau}{T}\right)\right]^2 + O\left(\frac{1}{T}\right)$$

$$= \left(I_{m^2} + K_{mm}\right)\Omega \otimes \Omega\frac{1}{T^2}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\left[\sum_{j=1}^{B} w_j \phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{\tau}{T}\right)\right]^2 + o(b), \tag{S.19}$$

where the second line follows from Magnus and Neudecker (1979, Theorem 3.1(ix)) and

the fact that $T^{-1} = o(b)$ from Assumption 4. Further, by the orthonormality of $\{\phi_j\}$,

$$\frac{1}{T^2}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\left[\sum_{j=1}^{B} w_j \phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{\tau}{T}\right)\right]^2 = \sum_{j=1}^{B}\sum_{k=1}^{B} w_j w_k \left[\frac{1}{T}\sum_{j=1}^{B}\phi_j\left(\frac{t}{T}\right)\phi_k\left(\frac{t}{T}\right)\right]^2$$

$$= \sum_{j=1}^{B} w_j^2 \left[\frac{1}{T}\sum_{j=1}^{B}\left(\phi_j\left(\frac{t}{T}\right)\right)^2\right]^2 = \sum_{j=1}^{B} w_j^2. \tag{S.20}$$

Note that these steps assume the orthonormality of $\{\phi_j\}$ applies for the finite-sample inner product for all T, as in the working paper version of this paper (LLS (2017)). If $\Phi = [\Phi_1 \ldots \Phi_B]$, where $\Phi_j = [\phi_j(1/T)\ \phi_j(2/T)\ \ldots\ \phi_j(1)]'$, does not satisfy $\iota_T'\Phi_j = 0$ and $\Phi'\Phi/T = I_B$, the finite-sample Φ can be constructed as the orthonormalization of the demeaned $\{\Phi_j\}$. Lemma A of Phillips (2005) then shows that for the unadjusted series, $\Phi'\Phi/T = I_B + O(1/T)$, which implies that the finite-sample orthonormalization adjustment introduces an error of at most order $O(1/T) = o(b)$; equivalently, without the adjustment, (S.20) would include an error of order $o(b)$. Finally, we have from (14) that $v^{-1} = B^{-1}B\sum_{j=1}^{B} w_j^2 = \sum_{j=1}^{B} w_j^2$. Thus, along with (S.19) and (S.20), we have that

$$\mathrm{var}(\mathrm{vec}\,\hat{\Omega}) = v^{-1}(I_{m^2} + K_{mm})\Omega \otimes \Omega + o(b),$$ as stated.

(iv)-(v) For kernel estimators, given Assumptions 1 and 2, equation (18) follows from Sun (2014b) equation (16), along with $mc_m^{\alpha}(b) = \chi_m^{\alpha} + O(b)$, as shown below after equation (S.29) in proving the expansions for WOS estimators. Equation (19) follows from the proof of Sun (2014b) Theorem 5 for the case of the Gaussian location model.

For WOS estimators, first note that Assumption 1 directly implies that a multivariate martingale functional central limit theorem holds for the partial sums of z_t (see, e.g., Helland (1982)): for $\lambda \in [0,1]$, we have that $T^{-1/2}\sum_{t=1}^{\lfloor T\lambda \rfloor} z_t \xrightarrow{d} \Omega^{1/2}W_m(\lambda)$, where $\lfloor \cdot \rfloor$ is the greatest lesser integer function and W_m is an m-dimensional standard Brownian motion on the unit interval. (This verifies an assumption by Sun (2013, 2014b), whose results we apply.) We thus have (extending the result after (14)) that $\hat{\Omega} \xrightarrow{d} \Omega^{1/2}\left(\sum_{j=1}^{B} w_j\Xi_j\right)\Omega^{1/2\prime}$, where Ξ_j are i.i.d. standard m-dimensional Wishart with one degree of freedom.

Therefore, as in Sun (2014b, eq. (8)-(9)), we have in this case that

$$mF_T \xrightarrow{d} \eta'\left(\sum_{j=1}^{B} w_j\Xi_j\right)^{-1}\eta \equiv mF_{\infty,m,B}, \tag{S.21}$$

where $\eta \sim N(0,I_m)$ and η is independent of Ξ_j for all j. Write

$$\sum_{j=1}^{B} w_j \Xi_j = \begin{pmatrix} \varsigma_{11} & \varsigma_{12} \\ \varsigma_{21} & \varsigma_{22} \end{pmatrix}, \tag{S.22}$$

where $\varsigma_{11} \in \mathbb{R}$, $\varsigma_{22} \in \mathbb{R}^{(m-1)\times(m-1)}$, and so on. Then using Sun (2014b, equation (10)), we

have equivalently that $mF_{\infty,m,B} \sim \|\eta\|^2 / (\varsigma_{11} - \varsigma_{12}\varsigma_{22}^{-1}\varsigma_{21})$. We then proceed to take a Taylor

expansion of $G_m(z \times (\varsigma_{11} - \varsigma_{12}\varsigma_{22}^{-1}\varsigma_{21}))$ around $G_m(z)$ for arbitrary argument z. Note first

that it can be shown quickly (as in Lemma 3 of Sun (2014b)) that

$$E(\varsigma_{11}) = \sum_{j=1}^{B} w_j = 1,$$

$$E(\varsigma_{11} - \varsigma_{12}\varsigma_{22}^{-1}\varsigma_{21}) = 1 - (m-1)\left(\sum_{j=1}^{B} w_j^2\right)(1 + o(1)) = 1 - \frac{\psi}{B}(m-1) + o(b),$$

$$E\left[(\varsigma_{11} - \varsigma_{12}\varsigma_{22}^{-1}\varsigma_{21})^2\right] = 1 + 2(2-m)\frac{\psi}{B} + o(b), \tag{S.23}$$

where again $B = b^{-1}$. Thus a Taylor expansion gives that

$$P(mF_{\infty,m,B}z) = E\left[G_m\left(z(\varsigma_{11} - \varsigma_{12}\varsigma_{22}^{-1}\varsigma_{21})\right)\right]$$

$$= G_m(z) - G'_m(z)z(m-1)\frac{\psi}{B} + \frac{1}{2}G''_m(z)z^2\left(2\frac{\psi}{B}\right) + o(b)$$

$$= G_m(z) + \frac{\psi}{B}\left[G''_m(z)z^2 - G'_m(z)z(m-1)\right] + o(b). \tag{S.24}$$

Using this and denoting by $\tilde{c}^\alpha_{m,B}$ the $1-\alpha$ quantile of the distribution $mF_{\infty,m,B}$, we have

$$1 - \alpha = G_m(\tilde{c}^\alpha_m(b)) + \frac{\psi}{B}\left[G''_m(\tilde{c}^\alpha_m(b))(\tilde{c}^\alpha_m(b))^2 - G'_m(\tilde{c}^\alpha_m(b))\tilde{c}^\alpha_m(b)(m-1)\right] + o(b). \tag{S.25}$$

Moving to F_T^*, following Sun (2011, Lemma 3) and Sun (2014b, Lemma 1), first

define the GLS estimator of β as $\hat{\beta}_{GLS} = [(\iota_T \otimes I_m)'V^{-1}(\iota_T \otimes I_m)]^{-1}(\iota_T \otimes I_m)'V^{-1}y$, where

ι_T is a $T \times 1$ vector of ones, $V = \text{var}([u'_1 \quad u'_2 \quad \ldots \quad u'_T]')$, and $y = [y'_1 \quad y'_2 \quad \ldots \quad y'_T]'$, and

define $\Omega_{T,GLS} = \text{var}[T^{1/2}(\hat{\beta}_{GLS} - \beta)]$. The independence of the GLS estimator from $\hat{\Omega}$

(which is in general not satisfied for the OLS estimator $\hat{\beta}$ given autocorrelation in u_t)

allows for a more convenient expansion of the test statistic. Applying Sun (2011, Lemma

3), Sun (2014b, Lemma 1), this expansion proceeds from the following representation:

$$P\left(mF_T^* \le z\right) = E\left[G_m\left(z\frac{B}{B-m+1}\Theta_T^{-1}\right)\right] + O(1/T),\qquad\text{(S.26)}$$

where $\Theta_T = e_T'\left[\Omega^{1/2}\hat{\Omega}^{-1}\Omega^{1/2\prime}\right]e_T$, $e_T = \Omega_{T,GLS}^{-1/2}\sqrt{T}(\hat{\beta}_{GLS}-\beta_0)/\ \|\Omega_{T,GLS}^{-1/2}\sqrt{T}(\hat{\beta}_{GLS}-\beta_0)\|$ and

where $\|\bullet\|$ is the Frobenius norm. Then applying Sun (2011, Theorem 4), Sun (2013,

Theorem 4.1), Sun (2014b, Theorem 2), we can expand $\Theta^{-1} = 1 + L + Q + o_p\left((bT)^{-q}+b\right)$,

where $L = ([e_T'\Omega^{-1/2}]\otimes[e_T'\Omega^{-1/2}])\text{vec}(\hat{\Omega}-\Omega)$, $Q = \text{vec}(\hat{\Omega}-\Omega)'(J_1 - J_2)\text{vec}(\hat{\Omega}-\Omega)/2$,

$J_1 = [2\Omega^{-1/2}e_T e_T'\Omega^{-1/2}]\otimes[\Omega^{-1/2}(e_T e_T')\Omega^{-1/2}]$, $J_2 = \Omega^{-1/2}e_T e_T'\Omega^{-1/2}\otimes\Omega^{-1}K_{mm}(I_{m^2}+K_{mm})$

(see Sun (2014b, p. 675)). From part (i) of the theorem, we have that

$E\hat{\Omega}-\Omega = -(B/T)^q k^{WOS(q)}(0)\sum_{j=-\infty}^{\infty}|j|^q\Gamma_j + o\left((B/T)^q\right)$, and therefore, again following

the steps in Sun (2011, Theorem 4), Sun (2013, Theorem 4.1), Sun (2014b, Theorem 2),

$$E[L] = -(B/T)^q k^{WOS(q)}(0)\omega^{(q)} + o\left((bT)^{-q}+b\right),\qquad\text{(S.27)}$$

$$E[L^2] = 2\frac{\psi}{B} + o\left((bT)^{-q}+b\right)\ \text{and}\ E[Q] = -\frac{\psi}{B}(m-1) + o\left((bT)^{-q}+b\right).\qquad\text{(S.28)}$$

Then expanding (S.26) as in those theorems,

$$P\left(mF_T^* \le z\right) = G_m\left(z\frac{B}{B-m+1}\right) + G_m'(z)zE[L+Q] + \frac{1}{2}EG_p''(z)z^2 E[L^2] + o\left((bT)^{-q}+b\right) + O\left(\frac{1}{T}\right)$$

$$= G_m(z) - G_m'(z)z\omega^{(q)}k^{WOS(q)}(0)(bT)^{-q} - G_m'(z)z\frac{\psi}{B}(m-1) + G_p''(z)z^2\frac{\psi}{B}$$

$$+o(b) + o\left((bT)^{-q}\right).\qquad\text{(S.29)}$$

Set $z = \tilde{c}_m^{\alpha}(b)$ in this equation, and note that (i) $\tilde{c}_m^{\alpha}(b) = \chi_m^{\alpha} + O(b)$ as in Sun (2014b, p.

665), and (ii) $mF_T = (B/(B-m+1))mF_T^* = mF_T^*(1+O(b))$, so that

$mc_m^{\alpha}(b) = \tilde{c}_m^{\alpha}(b) + O(b) = \chi_m^{\alpha} + O(b)$, where $c_m^{\alpha}(b)$ is the fixed-b critical value as in the

text (i.e., the $1-\alpha$ quantile of the limiting distribution for F_T^* with B fixed). Combining

this with (S.25) then gives the null expansion (18) for WOS tests.

The expansion under the local alternative uses the calculations above (extended to incorporate the local alternative) to apply the results of Sun (2011, Theorem 5(b)) and Sun (2014b, Theorem 5). Those calculations (omitted here, since they follow the same steps as in those papers and above, but are available upon request) yield the expansion

$$\Pr_{\delta}\left[F_T^* \le c_m^{\alpha}(b)\right] = G_{m,\delta^2}(\chi_m^{\alpha}) - G'_{m,\delta^2}(\chi_m^{\alpha})\chi_m^{\alpha}\omega^{(q)}k^{(q)}(0)(B/T)^q$$

$$+\frac{1}{2}\delta^2 G'_{m+2,\delta^2}(\chi_m^{\alpha})\chi_m^{\alpha}\frac{\psi}{B} + o(b) + o\left((bT)^{-q}\right). \tag{S.30}$$

Rearranging gives (19).

(vi) Parts (ii)-(v) apply directly. For part (i), see Proposition S4 below. □

S2.2 PROOFS FOR REMARKS AND ADDITIONAL STATEMENTS

Proposition S1 (Section 2.2). The Ibragimov-Müller (2010) LRV estimator, which is the sample variance of subsample estimators of β on $B+1$ equal-sized subsamples, can be expressed as an equal-weighted WOS estimator.

Proof of Proposition S1:   For convenience, suppose $T/(B+1)$ is an integer and $m = 1$, though the derivation below applies straightforwardly to the more general cases. The Ibragimov-Müller (2010) split-sample (SS) test statistic is then

$$t^{SS} = \sqrt{B+1}\left(\bar{\hat{\beta}} - \beta_0\right)\Big/\sqrt{S_{\hat{\beta}}^2}, \text{ where } S_{\hat{\beta}}^2 = \frac{1}{B}\sum_{i=1}^{B+1}\left(\hat{\beta}^{(i)} - \bar{\hat{\beta}}\right)^2, \tag{S.31}$$

where $\hat{\beta}^{(i)}$ is the estimator of β computed using the ith subsample and $\bar{\hat{\beta}} = \frac{1}{B+1}\sum_{i=1}^{B+1}\hat{\beta}^{(i)}$.

Note that $\bar{\hat{\beta}} - \beta_0 = \bar{z}_0$, and define $\hat{\Omega}^{SS} = [T/(B+1)]S_{\hat{\beta}}^2$. Let $\breve{\beta}$ be the $B+1$ vector with i^{th} element $\breve{\beta}_i = \hat{\beta}^{(i)}$, so that $S_{\hat{\beta}}^2 = B^{-1}\breve{\beta}'\left(I_{B+1} - \iota_{B+1}(\iota_{B+1}'\iota_{B+1})^{-1}\iota_{B+1}'\right)\breve{\beta}$, where I_{B+1} is the $(B+1)\times(B+1)$ identity matrix and ι_{B+1} is the $(B+1)$-vector of 1's. Define

$$\Phi^{SS} = \sqrt{(B+1)}\left(I_{B+1}\otimes\iota_{T/(B+1)}\right)M_t^B, \tag{S.32}$$

where M_ι^B is the $(B+1)\times B$ matrix of eigenvectors of $I_{B+1} - \iota_{B+1}(\iota_{B+1}{}'\iota_{B+1})^{-1}\iota_{B+1}'$ associated with its B unit eigenvalues. Then

$$\hat{\Omega}^{SS} = [T/(B+1)]S_{\breve{\beta}}^2 = [T/(B+1)]B^{-1}\breve{\beta}'\left(I_{B+1} - \iota_{B+1}(\iota_{B+1}{}'\iota_{B+1})^{-1}\iota_{B+1}'\right)\breve{\beta}$$

$$= [T/(B+1)]B^{-1}[T/(B+1)]^{-2}y'\left(I_{B+1}\otimes\iota_{T/(B+1)}\right)\left(I_{B+1} - \iota_{B+1}(\iota_{B+1}{}'\iota_{B+1})^{-1}\iota_{B+1}'\right)\left(I_{B+1}\otimes\iota_{T/(B+1)}'\right)y$$

$$= (BT)^{-1}(B+1)\hat{z}'\left(I_{B+1}\otimes\iota_{T/(B+1)}\right)\left(I_{B+1} - \iota_{B+1}(\iota_{B+1}{}'\iota_{B+1})^{-1}\iota_{B+1}'\right)\left(I_{B+1}\otimes\iota_{T/(B+1)}'\right)\hat{z}$$

$$= (BT)^{-1}(B+1)\hat{z}'\left(I_{B+1}\otimes\iota_{T/(B+1)}\right)M_\iota^B M_\iota^{B'}\left(I_{B+1}\otimes\iota_{T/(B+1)}'\right)\hat{z} = \hat{z}'\Phi^{SS}\Phi^{SS'}\hat{z}/BT, \qquad \text{(S.33)}$$

where the first equality uses the definition of $\hat{\Omega}^{SS}$, the second applies (S.31), the third

uses $\breve{\beta} = [T/(B+1)]^{-1}\left[I_{B+1}\otimes\iota_{T/(B+1)}'\right]y$, the fourth uses $\hat{z} = y - \overline{\beta}$ and the properties of

$I_{B+1} - \iota_{B+1}(\iota_{B+1}{}'\iota_{B+1})^{-1}\iota_{B+1}'$, the fifth uses the idempotence of $I_{B+1} - \iota_{B+1}(\iota_{B+1}{}'\iota_{B+1})^{-1}\iota_{B+1}'$ and the

definition of M_ι^B, and the final equality uses the definition of Φ^{SS} in (S.32).

Note that Φ^{SS} is $T\times B$, that $\iota_T'\Phi^{SS} = 0$, and $\Phi^{SS'}\Phi^{SS}/T = I_B$ as required for series

estimators (for which $\Phi = [\Phi_1 \ldots \Phi_B]$, where $\Phi_j = [\phi_j(1/T)\ \phi_j(2/T)\ \ldots\ \phi_j(1)]'$). Thus $\hat{\Omega}^{SS}$

is an equal-weighted WOS estimator as defined in (8) with basis matrix Φ^{SS}. □

Proposition S2 (Section 3.3). The Fourier, Type II cosine, and Legendre polynomial

bases satisfy $\sup_{s\in[0,1]}\left|\phi_j^{(n)}(s)\right| \le C_{n,\phi}j^{2n+1/2}$ for all j and $n = 0, 1, 2$, where $\phi_j^{(n)}(s)$ is the nth

derivative of ϕ_j and the constant $C_{n,\phi}$ does not depend on j, as required for Assumption 3.

Proof of Proposition S2:   The Fourier and cosine basis functions satisfy $|\phi_j(x)| \le 1$ for

all j. For the Fourier basis, we have $\phi_{2j-1}'(s) = -2\sqrt{2}\pi j\sin(2\pi js)$,

$\phi_{2j}'(s) = 2\sqrt{2}\pi j\cos(2\pi js)$, $\phi_{2j-1}''(s) = -4\sqrt{2}\pi^2 j^2\cos(2\pi js)$,

$\phi_{2j}''(s) = -4\sqrt{2}\pi^2 j^2\sin(2\pi js)$, and thus $|\phi_k'(s)|/j \le 2\sqrt{2}\pi$, $|\phi_k''(s)|/j^2 \le 4\sqrt{2}\pi^2$, with k

= 2j – 1, 2j, for all j, so that the condition is satisfied. Similarly, for the cosine basis, $|\phi_j'(s)|/j \le \sqrt{2}\pi$, $|\phi_j''(s)|/j^2 \le \sqrt{2}\pi^2$, so that the condition is satisfied.

For the Legendre case, first denote the Legendre polynomial of degree j by $P_j(x)$, $x \in [-1,1]$. The Legendre basis functions are then defined as $\phi_j(s) = P_j(x)\sqrt{2j+1}$, for $s = (x+1)/2$, so that the basis functions are shifted to $s \in [0,1]$ and normalized such that $\int_0^1 \phi_j(s)\phi_k(s)ds = 1\{j = k\}$ (e.g., Abramowitz and Stegun (AS, 1965, p. 774)), as required by definition. Thus $|\phi_j(s)|/\sqrt{j} \le \sqrt{3}$, satisfying the requirement for the 0th derivative, as $|P_j(x)| \le 1$ (Abramowitz and Stegun (1965, eq. 22.14.7)) and $\sqrt{2j+1}/\sqrt{j} \le \sqrt{3}$.

For the first and second derivatives, note first that the Legendre polynomial $P_j(x)$ is equivalent to the Jacobi polynomial $P_j^{(\alpha,\beta)}(x)$ with $\alpha = \beta = 0$ (Abramowitz and Stegun (1965, eq. 22.5.24)). Thus applying a well-known property of Jacobi polynomial derivatives (e.g., Shen, Tang, and Wang (2011, eq. (3.101))), we have that

$$\frac{d^n}{dx^n}P_j(x) = \frac{d^n}{dx^n}P_j^{(\alpha,\beta)}(x)\Bigg|_{\alpha=0,\beta=0} = \frac{\Gamma(j+1+n)}{2^n\Gamma(j+1)}P_{j-n}^{(n,n)}(x), \qquad (\text{S.34})$$

$j \ge n$, where $\Gamma(\cdot)$ is the gamma function. (Boundedness for the case $j = 1$, $n = 2$ is immediate, as $P_1''(x) = 0$.) And $\max_{x\in[-1,1]}\left|P_{j-n}^{(n,n)}(x)\right| = \max\left|P_{j-n}^{(n,n)}(\pm 1)\right|$ (Shen, Tang, and Wang (2011, eq. (3.125))), so $P_j'(x)$ and $P_j''(x)$ are maximized at a boundary point $x = \pm 1$. From Shen, Tang, and Wang (2011, eq. (3.177a)-(3.177b)), at these points,

$$P_j'(\pm 1) = \frac{1}{2}(\pm 1)^{j-1}j(j+1), \quad P_j''(\pm 1) = \frac{1}{8}(\pm 1)^j(j-1)j(j+1)(j+2). \qquad (\text{S.35})$$

The uniform boundedness of $|P_j'(x)|/j^2$ and $|P_j''(x)|/j^4$ follows immediately. Then using that $\phi_j(s) = P_j(x)\sqrt{2j+1}$ as above, we have that $|\phi_j'(s)|/j^{2+1/2}$ and $|\phi_j''(s)|/j^{4+1/2}$ are uniformly bounded as well, as stated. □

Proposition S3 (Remark 3). Statements (a)-(e) in Remark 3 hold.

Proof of Proposition S3:

(a) For kernel estimators, Priestley (1981, eq. (6.2.123)) extends a result of Parzen (1957) to show that given a process with known mean, equation (15) holds without the terms in b if $b^q T^{q-1} \to 0$. Thus

$$\text{bias}\left(\hat{s}_z(0)\right) = E\hat{s}_z(0) - s_z(0) = -(bT)^{-q} k^{(q)}(0) s_z^{(q)}(0) + o\left((bT)^{-q}\right), \qquad \text{(S.36)}$$

and this equation holds as well for WOS estimators (including with unknown mean) by (13) and (15). For variance, (17) holds in both cases, so that

$$\text{var}\left(\hat{s}_z(0)\right) = 2v^{-1}\left(s_z(0)\right)^2 + o(b). \qquad \text{(S.37)}$$

So up to higher-order terms, $MSE\left(\hat{s}_z(0)\right) = (bT)^{-2q}\left(k^{(q)}(0) s_z^{(q)}(0)\right)^2 + 2b\psi\left(s_z(0)\right)^2$, which

is shown by Priestley (1981, eq. (7.5.9)) to satisfy $\min_b MSE\left(\hat{s}_z(0)\right) \propto \left(\ell^{(q)}(k)\right)^{2q/(2q+1)}$.

(b) Using the two equations from (a), the objective function evaluates to

$$a(bT)^{-q} k^{(q)}(0)\left|s_z^{(q)}(0)\right| + 2(1-a)b\psi\left(s_z(0)\right)^2 + o\left((bT)^{-q}\right) + o(b). \qquad \text{(S.38)}$$

The minimizing value of b is invariant, up to a multiplicative constant, to transformations of the objective function of the form

$$a_1(bT)^{-q} k^{(q)}(0)\left|s_z^{(q)}(0)\right| + a_2 b\psi\left(s_z(0)\right)^2 + o\left((bT)^{-q}\right) + o(b), \qquad \text{(S.39)}$$

for $a_1, a_2 > 0$. Sun and Yang (2020, p. 11) show that (i) objective function (e) can be expressed in this form, and (ii) its minimum is achieved for $b \propto \left(k^{(q)}(0)/\psi\right)^{1/(q+1)} T^{-q/(q+1)}$ (see also LLSW (2018, rejoinder eq. (1))), so that the minimized objective function is, to higher order and up to an additive constant, proportional to $\left(\ell^{(q)}(k)\right)^{q/(q+1)}$.

(c) By the proof of Corollary 1, both objectives can be expressed in the form (S.39), so part (b) applies.

(d) See LLSW (2018, eq. (24)-(25)).

(e) See part (b). □

Proposition S4 (Remark 5). For EWP and QS tests with equivalent higher-order size,

equation (30) in the text holds, with $v_{EWP} = B$.

Proof of Proposition S4:   Fix a sequence $B = 1/b_{EWP}$. To obtain equivalent higher-order size using the QS test, Theorem 1(iv) gives that we must set

$$b_{QS} = \sqrt{\frac{k^{QS(2)}(0)}{k^{EWP(2)}(0)}} b_{EWP} = \sqrt{\frac{\pi^2/10}{\pi^2/6}} B^{-1} = \sqrt{\frac{3}{5}} B^{-1}, \qquad (S.40)$$

where the $k^{(2)}(0)$ values for the two tests are as in the proof of Theorem 5. Further,

$$\int_{-\infty}^{\infty} k^2(x) dx = \frac{6}{5}$$ for QS, so that given equivalent higher-order size, we have

$$v_{EWP}^{-1} - v_{QS}^{-1} = B^{-1} - \frac{6}{5}\sqrt{\frac{3}{5}} B^{-1}.$$ Plugging this into Theorem 3 yields the desired result.   □

Proposition S5 (Remark 6).

(i)   The Bartlett kernel and SS estimator both have $q = 1$, and the Bartlett kernel's size-power tradeoff curve is strictly below the SS tradeoff curve.

(ii)   The EWP estimator is asymptotically equivalent to the equal-weighted WOS estimator using type II cosine basis functions, and both have $q = 2$.

Proof of Proposition S5:

(i) We first consider the SS estimator. Note that the SS basis functions (S.32) do not satisfy the differentiability requirement of Assumption 3. Thus for the SS estimator we calculate $E\hat{\Omega}^{SS}$ directly; in doing so, we show that the SS implied mean kernel is similar to the Bartlett kernel for a subsample of $T/(B+1)$ observations (where it is assumed for notational simplicity that this ratio is integer-valued, as the non-integer case follows immediately setting the subsample size to $[T/(B+1)]$).

First, given $\bar{y}_i - \bar{y} \equiv \frac{1}{T_i}\sum_{t \in T_i} y_t - \frac{1}{T}\sum_{t=1}^{T} y_t$ (where, abusing notation, T_i denotes both the number of observations in subsample i, $T_i = T/(B+1)$, and the subsample that t indexes),

we have $\bar{y}_i - \bar{y} = \frac{1}{T}\sum_{t=1}^{T}\left((B+1)1\{t \in T_i\} - 1\right)y_t = \frac{B+1}{T}\sum_{t=1}^{T}\left(1\{t \in T_i\} - \frac{1}{B+1}\right)y_t$. Thus

squaring and summing over subsamples, we have

$$\frac{1}{B}\sum_{i=1}^{B+1}(\bar{y}_i - \bar{y})^2 = \frac{1}{B}\sum_{i=1}^{B+1}\left(\frac{B+1}{T}\right)^2\sum_{t=1}^{T}\sum_{s=1}^{T}\left(1\{t \in T_i\} - \frac{1}{B+1}\right)\left(1\{s \in T_i\} - \frac{1}{B+1}\right)y_t y_s. \quad (S.41)$$

Taking the expectation of this value and rearranging,

$$E\frac{1}{B}\sum_{i=1}^{B+1}(\bar{y}_i - \bar{y})^2 = \frac{B+1}{B}\frac{1}{T/B+1}\sum_{u=-(T-1)}^{T-1}\left[\left(1 - \left|\frac{u}{T/B+1}\right|\right)1\{|u| \le \frac{T}{B+1}\} - \frac{1}{B+1}\left(1 - \left|\frac{u}{T}\right|\right)\right]\Gamma_u$$

$$= \frac{B+1}{T}\sum_{u=-(T-1)}^{T-1}\left[\left(\frac{B+1}{B} - \frac{B+1}{B}\left|\frac{u}{T/(B+1)}\right|\right)1\{|u| \le \frac{T}{B+1}\} - \frac{1}{B} + \frac{1}{B}\left|\frac{u}{T}\right|\right]\Gamma_u. \quad (S.42)$$

Converting $E\frac{1}{B}\sum_{i=1}^{B+1}(\bar{y}_i - \bar{y})^2$ to $E\hat{\Omega}^{SS}$ requires multiplying by $T/(B+1)$ given the form of

the statistic in (S.31) compared to the usual t-statistic. Thus in this case defining S such

that $T = S(B+1)$ given that there are $B+1$ subsamples and setting $\tilde{v} = u/S$, we can write

the SS implied mean kernel (i.e., the expression in brackets in (S.42)) as

$$k_B^{SS}(\tilde{v}) = \left(\frac{B+1}{B} - \frac{B+1}{B}|\tilde{v}|\right)1\{|\tilde{v}| \le 1\} - \frac{1}{B} + \frac{1}{B(B+1)}|\tilde{v}|. \quad (S.43)$$

Thus using the definition of the generalized first derivative in (11), we have

$k_B^{SS(1)}(0) = \frac{B+1}{B} - \frac{1}{B(B+1)} = \frac{B+2}{B+1} \to 1$ as $B \to \infty$. Because $k^{SS(1)}(0) \ne 0$, $q = 1$ for the SS

estimator. Further, comparing $E\hat{\Omega}^{SS}$ with Ω using (S.42), we obtain that Theorem 1(i)

applies for the SS estimator as well, and the tradeoff in Theorem 4(ii) applies. The value

$\ell^{(1)}(k^{SS})$ is equal to $k^{SS(1)}(0) = 1$ given $\psi = 1$ for equal-weighted WOS estimators.

For the Bartlett/Newey-West test, Priestley (1981) Table 7.1 gives $k^{(1)}(0) = 1$ and $q =$

1, while Table 6.1 gives that $\int_{-\infty}^{\infty} k^2(x)dx = 2/3$, so that $\ell^{(1)}(k^{NW}) = k^{(1)}(0)\int_{-\infty}^{\infty} k^2(x)dx = 2/3$,

from which we conclude that the Bartlett tradeoff dominates the SS tradeoff.

(ii) For the Fourier basis functions used in the EWP estimator, we have, as in Proposition S2, $\phi''_{2j-1} = -4\sqrt{2}\pi^2 j^2 \cos(2\pi js)$, $\phi''_{2j} = -4\sqrt{2}\pi^2 j^2 \sin(2\pi js)$, which give that

$\int_0^1 \phi_{2j-1}(s)\phi''_{2j-1}(s)ds = \int_0^1 \phi_{2j}(s)\phi''_{2j}(s)ds = -4\pi^2 j^2$. Then applying Theorem 1(ii),

$$k_B^{EWP(2)}(0) = -\frac{1}{2}\frac{1}{B}\sum_{j=1}^{B/2}\frac{1}{B^2}2\left(-4\pi^2 j^2\right) = \frac{\pi^2}{6}\frac{(B+1)(B+2)}{B^2} \xrightarrow{B\to\infty} \frac{\pi^2}{6}. \qquad (S.44)$$

Similarly, for cosine basis functions, using their limiting implied mean kernel form,

$\phi''_j(s) = -\sqrt{2}\pi^2 j^2 \cos(\pi js)$ and $\int_0^1 \phi_{2,j-1}(s)\phi''_{2,j-1}(s)ds = -\pi^2 j^2$. Summing over j,

$$k_B^{\cos(2)} = -\frac{1}{2}\frac{1}{B}\sum_{j=1}^{B}\frac{1}{B^2}\left(-\pi^2 j^2\right) = \frac{\pi^2}{6}\frac{(B+1)(B+1/2)}{B^2} \xrightarrow{B\to\infty} \frac{\pi^2}{6}. \qquad (S.45)$$

Results (S.44) and (S.45) and Theorem 1(ii) give that $q = 2$ for both estimators; these results, along with $\psi = 1$ for equal-weighted WOS estimators, then imply given Theorem 4(ii) that the estimators are asymptotically equivalent. □

Proposition S6 (Section 4.3). Assume that the remainder terms in equation (18) in the text are $o(b) + o\left((bT)^{-q}\right)$ uniformly in $\left|\omega^{(q)}\right| \leq \overline{\omega}^{(q)}$. Then:

(i) The maximum weighted average power (WAP) test solving equation (31) in the text features b^{WAP} as stated in equation (32), with

$$\tilde{\omega}^{(q)} = \int_{|\rho|\leq\overline{\rho}}\left[\overline{\omega}^{(q)}(\rho) - \omega^{(q)}(\rho)\right]d\Pi(\rho) \text{ and } \tilde{d}_{m,\alpha,q} = \left(\frac{\int_\delta G'_{m,\delta^2}(\chi_m^\alpha)\chi_m^\alpha d\Pi_\delta(\delta)}{\frac{1}{2}\int_\delta \delta^2 G'_{m+2,\delta^2}(\chi_m^\alpha)\chi_m^\alpha d\Pi_\delta(\delta)}\right)^{\frac{1}{1+q}}. \qquad (S.46)$$

(ii) The power loss of the test using b^{WAP} in (32) depends on k only through $\ell^{(q)}(k)$.

(iii) The test asymptotically delivering the highest WAP uses the QS kernel, and more generally, $q = 1$ kernels are asymptotically dominated by $q = 2$ kernels.

Proof of Proposition S6:

(i) Let $\overline{c}_{m,T}^\alpha(b)$ be the size-adjusted critical value (20) based on the boundary value of

$\overline{\omega}^{(q)}: \overline{c}_{m,T}^{\alpha}(b) = \left[1+\overline{\omega}^{(q)}k^{(q)}(0)(bT)^{-q}\right]c_m^{\alpha}(b)$. From (18), the null rejection rate of the test using this size-adjusted critical value, evaluated at the true value of $\omega^{(q)}$, is

$$\Pr_0\left[F_T^* > \overline{c}_{m,T}^{\alpha}(b)\right] = \alpha + G_m'(\chi_m^{\alpha})\chi_m^{\alpha}\left(\omega^{(q)} - \overline{\omega}^{(q)}\right)k^{(q)}(0)(bT)^{-q} + o(b) + o\left((bT)^{-q}\right), \text{(S.47)}$$

from which it follows that, for a given sequence b and under the assumed condition,

$$\sup_{\omega^{(q)} \leq \overline{\omega}^{(q)}} \Pr_0\left[F_T^* > \overline{c}_{m,T}^{\alpha}(b)\right] \leq \alpha + o(b) + o\left((bT)^{-q}\right). \tag{S.48}$$

The expression for $\Delta_P\left(\omega^{(q)}(\rho),\delta\right)$ in (31) then follows directly from (33) in the proof of Theorem 2 (omitting higher-order remainder terms). Solving (31) yields (32), with $\tilde{\omega}^{(q)}$ and $\tilde{d}_{m,\alpha,q}$ as stated.

(ii)-(iii) Substituting b^{WAP} in (32) into the expression for $\Delta_P\left(\omega^{(q)}(\rho),\delta\right)$, we obtain that the power loss of the the test using the WAP-maximizing sequence is

$$\Delta_P^{WAP} = \left(q^{-q/(1+q)} + q^{1/(1+q)}\right)\tilde{a}_{m.\alpha.q}\left[\left(k^{(q)}(0)\right)^{1/q}\psi\right]^{\frac{q}{1+q}}\left(\tilde{\omega}^{(q)}\right)^{\frac{1}{1+q}}T^{\frac{-q}{1+q}}, \tag{S.49}$$

where $\tilde{a}_{m.\alpha.q} = \left[\int G_{m,\delta^2}'(\chi_m^{\alpha})\chi_m^{\alpha}d\Pi_\delta(\delta)\right]^{1/(1+q)}\left[\frac{1}{2}\int \delta^2 G_{m+2,\delta^2}'(\chi_m^{\alpha})\chi_m^{\alpha}d\Pi_\delta(\delta)\right]^{q/(1+q)}$. Note that $\ell^{(q)}(k) = \left(k^{(q)}(0)\right)^{1/q}\psi$. The remaining stated results then follow. $\square$

REFERENCES

Abramowitz, M. and I.A. Stegun (1965). *Handbook of Mathematical Functions*. New York: Dover.

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

Bhatia, R. (1997). *Matrix Analysis*. New York: Springer.

Helland, I. S. (1982). "Central Limit Theorems for Martingales with Discrete or Continuous Time," *Scandinavian Journal of Statistics* 9, 79-94.

Ibragimov, R. and Müller, U.K. (2010). "*t*-statistic based correlation and heterogeneity robust inference," *Journal of Business and Economic Statistics* 28, 453-468.

Kiefer, N., T.J. Vogelsang, and H. Bunzel (2000). "Simple Robust Testing of Regression Hypotheses," *Econometrica*, 69, 695-714.

Lazarus, E., D.J. Lewis, and J.H. Stock (2017). "The Size-Power Tradeoff in HAR Inference," manuscript, Harvard University.

Lazarus, E., D.J. Lewis, J.H. Stock, and M.W. Watson (2018). "HAR Inference: Recommendations for Practice" (with discussion and rejoinder). *Journal of Business and Economic Statistics* 36, 541-575.

Magnus, J.R. and H. Neudecker (1979). "The Commutation Matrix: Some Properties and Applications," *Annals of Statistics* 7, 381-394.

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

Phillips, P. C. B. (2005). "HAC Estimation by Automated Regression," *Econometric Theory* 21, 116–142.

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

Shen, J., T. Tang, and L.-L. Wang (2011). *Spectral Methods: Algorithms, Analysis and Applications*. Berlin: Springer-Verlag.

Sun, Y. (2011). "Autocorrelation Robust Trend Inference with Series Variance Estimator and Testing-Optimal Smoothing Parameter," *Journal of Econometrics* 164, 345-366.

Sun, Y. (2013). "Heteroscedasticity and Autocorrelation Robust F Test Using Orthonormal Series Variance Estimator," *The Econometrics Journal*, 16, 1–26.

Sun, Y. (2014b). "Let's Fix It: Fixed-*b* Asymptotics versus Small-*b* Asymptotics in Heteroskedasticity and Autocorrelation Robust Inference," *Journal of Econometrics* 178, 659-677.

Sun, Y. and J. Yang (2020). "Testing-optimal Kernel Choice in HAR Inference," manuscript, University of California San Diego.

Velasco, C. and P.M. Robinson (2001). "Edgeworth Expansions for Spectral Density Estimates and Studentized Sample Mean," *Econometric Theory* 17, 497-539.