The Size-Power Tradeoff in HAR Inference: Online Supplement

August 14, 2020

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

This supplement provides additional figures, Monte Carlo results, and supplemental proofs.

Contents

S1. Discussion of Additional Figures

Figure S1 plots the implied mean kernel for the Fourier, cosine, and SS basis functions for $B = 8$ and $T = 200$. The Fourier transforms of these implied mean kernels, that is, the frequency-domain implied mean kernel, are plotted in Figure S2 at the frequencies $2\pi j/T$, $j = 0, \ldots, 35$. The EWP (Fourier) estimator is the only one of these four that has an exact kernel representation, and its frequency-domain kernel is the familiar flat (Daniell) kernel that gives equal weight to the first $B/2$ periodogram ordinates. The remaining three implied mean kernels in the frequency domain also concentrate their mass at low frequencies. In the frequency domain the SS implied mean kernel has more leakage than the Fourier and cosine kernels.

Figure S3 shows the power difference, as a function of the local alternative index δ, between the EWP and QS test, for $B = 8$ for EWP and b for QS chosen so that the two tests have the same size. This curve is computed using the expression in Theorem 3 and Remark 5.

S2. Monte Carlo Study

The purpose of this Monte Carlo analysis is twofold. First, we assess the quality of the small-b 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.

S2.1 Estimators and Design

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

We specifically examine the following HAR tests:

1. NW: Kernel estimator with Bartlett/Newey-West kernel, $k(x) = (1-|x|)\mathbf{1}(|x|\leq 1)$
2. KVB: The Kiefer-Vogelsang-Bunzel (2000) test, which is NW with $S = T$ (so $v = 3/2$).
3. QS implemented in the covariance domain: $k(x) = 3[\sin(\pi u)/(\pi u) - \cos(\pi u)]/(\pi u)^2$ for $u = 6x/5$.
4. EWP: Equal-weighted orthonormal series 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 orthonormal series 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 orthonormal series 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 according to $y_t = x_t'\beta + u_t$, with x_{it}, $i = 1,\ldots, m$ and u_t being 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 $\Sigma_{XX} = T^{-1}\sum_{t=1}^{T} x_t x_t'$ (note that in 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 $\delta_2 = [\delta \quad 0]'$.

S2.2 Monte Carlo Results

This section presents a number of representative Monte Carlo results; additional results are contained in Lazarus, Lewis, Stock and Watson (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. We conduct 100,000 replications for each simulation design considered here.

Location model. Figure S4 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.

Figure S5 presents results for $m = 2$ with AR(1) errors, $\rho = 0.5$, and $T = 200$. Compared to the results for $m = 1$, the $m = 2$ frontier fits the simulations slightly better for QS and EWP, but somewhat worse for NW.

Figures S6-S9 provide additional results for the location model for other AR(1) parameters, other sample sizes, ARMA(2,1) disturbances, and other kernels and orthonormal series. In particular, Figure S6 shows additional Monte Carlo results for different values of T for 5 tests: QS, EWP, cosine (type II cosine basis function), NW, and SS. Figure S7 repeats this analysis for $m = 2$, but varies ρ instead of T. Figure S8 shows the spectral density for the ARMA(2,1) process. The parameters are calibrated so that $\omega^{(2)} = 4$ (the same as an AR(1) with $\rho = 0.5$) and with a spectral density approximately symmetric around $\pi/2$, with a minimum at $\pi/2$ (the coefficients are $\rho_1 = 0.048$, $\rho_2 = 0.248$, $\theta = -0.064$). Figure S9 shows results for ARMA(2,1) disturbances with parameters underlying Figure S8, $m = 1$. These results indicate that the 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 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 S10 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 approximations 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 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. These findings reflect results for other designs, kernels, and $m = 2$, the latter of which are reported in Figure S11. Further, additional simulations show that the approximation improves for higher values of T.

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, consistent with the theory, the performance of $q = 2$ kernels is superior to that of $q = 1$ kernels, at least for this design. 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]

S3. Supplemental Proofs

S3.1 Additional Proofs of Main Results

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

$$\begin{aligned} E\hat{\Omega}_j^{OS} &= E\left(\sqrt{\frac{1}{T}}\sum_{t=1}^{T}\phi_j(t/T)\hat{z}_t\right)\left(\sqrt{\frac{1}{T}}\sum_{t=1}^{T}\phi_j(t/T)\hat{z}_t\right)' \\ &= \frac{1}{T}\sum_{t=1}^{T}\sum_{s=1}^{T}\phi_j(t/T)\phi_j(s/T)\Gamma_{s-t} + O(1/T) \\ &= \sum_{u=-(T-1)}^{T-1}\frac{1}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\Gamma_u + O(1/T), \end{aligned} \tag{S.1}$$

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,

$$E\hat{\Omega}^{WOS} - \Omega = \sum_{u=-(T-1)}^{T-1}\left\{\left[\sum_{j=1}^{B}w_j\frac{1}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\right]-1\right\}\Gamma_u - \sum_{|u|\geq T}\Gamma_u + O\left(\frac{1}{T}\right). \tag{S.2}$$

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

$$\left|\sum_{|u|\geq T}\Gamma_u\right| \leq \sum_{|u|\geq T}\left|\Gamma_u\right| \leq \frac{1}{T^2}\sum_{|u|\geq T}|u|^2\left|\Gamma_u\right| = o\left(T^{-2}\right) = o\left((B/T)^q\right), \tag{S.3}$$

so we may focus on the first summation in (S.2). 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

$$\begin{aligned} E\hat{\Omega}^{WOS} - \Omega = &\sum_{u=-L_T}^{L_T}\left\{\left[\sum_{j=1}^{B}\frac{w_j}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\right]-1\right\}\Gamma_u \\ &+ \sum_{L_T<|u|<T}\left\{\left[\sum_{j=1}^{B}\frac{w_j}{T}\sum_{t=\max(1,u)}^{\min(T,T+u)}\phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{t}{T}-\frac{u}{T}\right)\right]-1\right\}\Gamma_u + o\left(\left(\frac{B}{T}\right)^q\right), \end{aligned} \tag{S.4}$$

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

---

[1] In unreported results, we also examined the performance of tests based on plug-in higher-order corrected critical values given by equation (20) of the main 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. Results are given in a previous working paper version of this paper (Lazarus, Lewis, and Stock (2017)).

$$\frac{T^q}{L_T^{q+\zeta} B^q} + \frac{L_T B}{T} \to 0, \tag{S.5}$$

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

$$\begin{aligned}
&\left| \sum_{L_T < |u| < T} \left\{ \left[\sum_{j=1}^{B} \frac{w_j}{T} \sum_{t=\max(1,u)}^{\min(T,T+u)} \phi_j\left(\frac{t}{T}\right) \phi_j\left(\frac{t}{T} - \frac{u}{T}\right) \right] - 1 \right\} \Gamma_u \right| \\
&\qquad \le \sum_{j=1}^{B} w_j \sum_{L_T < |u| < T} \left| \left[\frac{1}{T} \sum_{t=\max(1,u)}^{\min(T,T+u)} \phi_j\left(\frac{t}{T}\right) \phi_j\left(\frac{t}{T} - \frac{u}{T}\right) \right] - 1 \right| \left| \Gamma_u \right| \\
&\qquad \le 2 \sum_{L_T < |u| < T} \left| \Gamma_u \right| \le 2 L_T^{-q-\zeta} \sum_{L_T < |u| < T} |u|^{q+\zeta} \left| \Gamma_u \right| = O\left(L_T^{-q-\zeta}\right) = o\left(\left(\frac{B}{T} \right)^q \right),
\end{aligned} \tag{S.6}$$

where the first inequality applies the triangle inequality, the second inequality uses that $\left| \sum_{t=\max(1,u)}^{\min(T,T+u)} \phi_j(t/T) \phi_j((t-u)/T)/T \right| \le \left| \sum_{t=1}^{T} (\phi_j(t/T))^2 / T \right| = 1$ by Cauchy-Schwarz (and therefore that $\left| \left[\sum_{t=\max(1,u)}^{\min(T,T+u)} \phi_j(t/T) \phi_j((t-u)/T)/T \right] - 1 \right| \le 2$), and where the first equality in the last line applies Assumption 1(b). (Note that setting $T^{-1} \sum_{t=1}^{T} (\phi_j(t/T))^2 = 1$ in finite samples for orthonormal ϕ_j is without loss of generality; see after (S.14) below for further discussion.)

Equation (S.4) can thus be written, using $k_{B,T}^{WOS}(u/S)$ as defined after equation (9) in the main text, as

$$\begin{aligned}
E\hat{\Omega}^{WOS} - \Omega &= \sum_{u=-L_T}^{L_T} \left[k_{B,T}^{WOS}\left(\frac{u}{S}\right) - 1 \right] \Gamma_u + o\left(\left(\frac{B}{T} \right)^q \right) \\
&= \sum_{u=-L_T}^{L_T} \left\{ \left[k_B^{WOS}\left(\frac{u}{S}\right) - 1 \right] \left(1 + o\left(\frac{1}{T}\right) \right) \right\} \Gamma_u + o\left(\left(\frac{B}{T} \right)^q \right) \\
&= \left\{ \sum_{u=-L_T}^{L_T} \left[k_B^{WOS}\left(\frac{u}{S}\right) - 1 \right] \Gamma_u \right\} \left(1 + o(1) \right) + o\left(\left(\frac{B}{T} \right)^q \right),
\end{aligned} \tag{S.7}$$

where the second line uses (10), along with the fact that that $k_{B,T}^{WOS}(x)$ and $k_{B,T}^{WOS\prime}(x)$ are uniformly bounded for fixed B (since $\left|\phi_j(s)\right|, \left|\phi_j'(s)\right| \le CB^{5/2}$ 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.5).

Now note that under Assumption 3, in addition to $\left| k_B^{WOS}(x) \right| \le 1$, 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 $k_B^{WOS}(x) = 1 + k_{B,+}^{WOS\prime}(0)x + \frac{1}{2} k_{B,+}^{WOS\prime\prime}(0)x^2 + o\left(x^2\right)$ 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 $k_B^{WOS}(x) = 1 + k_{B,+}^{WOS\prime}(0)|x| + \frac{1}{2}k_{B,+}^{WOS\prime\prime}(0)x^2 + o\left(x^2\right)$ 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

$$1 - k_B^{WOS}(x) = g_{1,B}|x| + g_{2,B}|x|^2 + o\left(|x|^2\right) \tag{S.8}$$

as $x \to 0$, from which it is clear from (11) that $k_B^{(q)}(0) = g_{q,B}$. Using this with (S.7) and the fact that $S = T/B$, we can follow Priestley (1981, p. 459) and write

$$\begin{aligned}
E\hat{\Omega}^{WOS} - \Omega &= \left\{-\left(\frac{B}{T}\right)^q \sum_{u=-L_T}^{L_T} \frac{1-k_B^{WOS}(u/S)}{|u/S|^q}|u|^q\,\Gamma_u\right\}(1+o(1)) + o\left(\left(\frac{B}{T}\right)^q\right) \\
&= -2\pi\left(\frac{B}{T}\right)^q \left(k_B^{(q)}(0)s_z^{(q)}(0)\right)(1+o(1)) + o\left(\left(\frac{B}{T}\right)^q\right) \\
&= -2\pi\left(\frac{B}{T}\right)^q k_B^{(q)}(0)s_z^{(q)}(0) + o\left(\left(\frac{B}{T}\right)^q\right) \\
&= -2\pi\left(\frac{B}{T}\right)^q \left\{\left(\lim_{B\to\infty} k_B^{(q)}(0)\right)(1+o(1))\right\} s_z^{(q)}(0) + o\left(\left(\frac{B}{T}\right)^q\right) \\
&= -2\pi\left(\frac{B}{T}\right)^q \left(\lim_{B\to\infty} k_B^{(q)}(0)\right) s_z^{(q)}(0) + o\left(\left(\frac{B}{T}\right)^q\right).
\end{aligned} \tag{S.9}$$

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

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

$$k_B^{WOS\prime}(x) = B^{-1}\sum_{j=1}^{B} w_j\left(-\int_{B^{-1}x}^{1} \phi_j(s)\phi_j{}'(s - B^{-1}x)ds - \phi_j(B^{-1}x)\phi_j(0)\right), \tag{S.10}$$

so from part (i),

$$k_B^{(q)}(0) = g_{1,B} = -k_{B,+}^{WOS\prime}(0) = B^{-1}\sum_{j=1}^{B} w_j\left(\int_0^1 \phi_j(s)\phi_j{}'(s)ds + \left(\phi_j(0)\right)^2\right). \tag{S.11}$$

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

Similarly, for $x > 0$,

$$k_B^{WOS\prime\prime}(x) = B^{-2}\sum_{j=1}^{B} w_j\left(\int_{B^{-1}x}^{1} \phi_j(s)\phi_j{}''(s - B^{-1}x)ds\right). \tag{S.12}$$

Using that $k_B^{(2)}(0) = -k_{B,+}^{WOS\prime\prime}(0)/2$ 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 the fact that $q = 1$ follows immediately from (S.8) and the definition of q stated after (11). If $k^{(1)}(0) = 0$, note from Lemma A1 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.8), 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.3)).

(iii) For kernel estimators, (17) is a restatement of Andrews (1991, Proposition 1(a)), as noted in the Appendix. For WOS estimators, Sun (2011, p. 361) gives, when generalized to the case of arbitrary WOS weights w_j, that

$$\begin{aligned}\operatorname{var}\left(\operatorname{vec}\hat{\Omega}\right) &= \Omega \otimes \Omega\left(I_{m^2} + K_{mm}\right)\frac{1}{T^2}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\left[\sum_{j=1}^{B} w_j \phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{\tau}{T}\right)\right]^2 + O\left(\frac{1}{T}\right) \\ &= \left(I_{m^2} + K_{mm}\right)\Omega \otimes \Omega\frac{1}{T^2}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\left[\sum_{j=1}^{B} w_j \phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{\tau}{T}\right)\right]^2 + o(b), \end{aligned} \qquad \text{(S.13)}$$

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 $\{\phi_j\}$,

$$\begin{aligned}\frac{1}{T^2}\sum_{t=1}^{T}\sum_{\tau=1}^{T}\left[\sum_{j=1}^{B} w_j \phi_j\left(\frac{t}{T}\right)\phi_j\left(\frac{\tau}{T}\right)\right]^2 &= \sum_{j=1}^{B}\sum_{k=1}^{B} w_j w_k \left[\frac{1}{T}\sum_{j=1}^{B}\phi_j\left(\frac{t}{T}\right)\phi_k\left(\frac{t}{T}\right)\right]^2 \\ &= \sum_{j=1}^{B} w_j^2\left[\frac{1}{T}\sum_{j=1}^{B}\left(\phi_j\left(\frac{t}{T}\right)\right)^2\right]^2 = \sum_{j=1}^{B} w_j^2 \, . \end{aligned} \qquad \text{(S.14)}$$

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 (Lazarus, Lewis, and Stock (2017)). If, for a given T, $\Phi = [\Phi_1 \ldots \Phi_B]$, 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) 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.14) would include an error of order o(b). Finally, we have from (14) that $\nu^{-1} = B^{-1}B\sum_{j=1}^{B} w_j^2 = \sum_{j=1}^{B} w_j^2$. Thus, along with (S.13) and (S.14), we have that $\operatorname{var}(\operatorname{vec}\hat{\Omega}) = \nu^{-1}(I_{m^2} + K_{mm})\Omega \otimes \Omega + o(b)$, 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.23) 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 discussed 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

$$mF_T \xrightarrow{d} \eta'\left(\sum_{j=1}^{B} w_j \Xi_j\right)^{-1} \eta \equiv mF_{\infty,m,B}, \tag{S.15}$$

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

$$\sum_{j=1}^{B} w_j \Xi_j = \begin{pmatrix} \varsigma_{11} & \varsigma_{12} \\ \varsigma_{21} & \varsigma_{22} \end{pmatrix}, \tag{S.16}$$

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

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

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

$$E\left[(\varsigma_{11} - \varsigma_{12}\varsigma_{22}^{-1}\varsigma_{21})^2\right] = 1 + 2(2-m)\frac{\psi}{B} + o(b), \tag{S.17}$$

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

$$P(mF_{\infty,m,B} \le z) = E\left[G_m\left(z(\varsigma_{11} - \varsigma_{12}\varsigma_{22}^{-1}\varsigma_{21})\right)\right] = G_m(z) - G_m'(z)z(m-1)\frac{\psi}{B} + \frac{1}{2}G_m''(z)z^2\left(2\frac{\psi}{B}\right) + o(b)$$

$$= G_m(z) + \frac{\psi}{B}\left[G_m''(z)z^2 - G_m'(z)z(m-1)\right] + o(b). \tag{S.18}$$

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

$$1 - \alpha = G_m(\tilde{c}_m^{\alpha}(b)) + \frac{\psi}{B}\left[G_m''(\tilde{c}_m^{\alpha}(b))(\tilde{c}_m^{\alpha}(b))^2 - G_m'(\tilde{c}_m^{\alpha}(b))\tilde{c}_m^{\alpha}(b)(m-1)\right] + o(b). \tag{S.19}$$

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, $V = \mathrm{var}([u_1' \quad u_2' \quad \dots \quad u_T']')$, and $y = [y_1' \quad y_2' \quad \dots \quad y_T']'$, and define $\Omega_{T,GLS} = \mathrm{var}[\mathrm{T}^{1/2}(\hat{\beta}_{GLS} - \beta)]$. 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 finite-sample test statistic. Applying Sun (2011, Lemma 3), Sun (2014b, Lemma 1), this expansion proceeds from the following representation:

$$P\left(mF_T^* \le z\right) = E\left[G_m\left(z\frac{B}{B-m+1}\Theta_T^{-1}\right)\right] + O(1/T), \tag{S.20}$$

where $\Theta_T = e_T'\left[\Omega^{1/2}\hat{\Omega}^{-1}\Omega^{1/2\prime}\right]e_T$, with $e_T = \Omega_{T,GLS}^{-1/2}\sqrt{T}(\hat{\beta}_{GLS}-\beta_0)/\ \|\Omega_{T,GLS}^{-1/2}\sqrt{T}(\hat{\beta}_{GLS}-\beta_0)\ \|$ and where $\|\cdot\|$ 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 is linear in $\hat{\Omega}-\Omega$, $L = ([e_T'\Omega^{-1/2}]\otimes[e_T'\Omega^{-1/2}])\text{vec}(\hat{\Omega}-\Omega)$, and Q is quadratic in that difference, $Q = \text{vec}(\hat{\Omega}-\Omega)'(J_1 - J_2)\text{vec}(\hat{\Omega}-\Omega)/2$, $J_1 = [2\Omega^{-1/2}e_Te_T'\Omega^{-1/2}]\otimes[\Omega^{-1/2}(e_Te_T')\Omega^{-1/2}]$, $J_2 = \Omega^{-1/2}e_Te_T'\Omega^{-1/2}\otimes\Omega^{-1}K_{mm}(I_{m^2}+K_{mm})$ (see Sun (2014b, p. 675)). From part (i) of the theorem, $E\hat{\Omega}-\Omega = -(B/T)^q k^{WOS(q)}(0)\sum_{j=-\infty}^{\infty}|j|^q\Gamma_j + o\left((B/T)^q\right)$, and therefore, again following the steps in Sun (2011, Theorem 4), Sun (2013, Theorem 4.1), Sun (2014b, Theorem 2),

$$E[L] = -(B/T)^q k^{WOS(q)}(0)\omega^{(q)} + o\left((bT)^{-q} + b\right), \tag{S.21}$$

and similarly

$$E[L^2] = 2\frac{\psi}{B} + o\left((bT)^{-q}+b\right) \text{ and } E[Q] = -\frac{\psi}{B}(m-1) + o\left((bT)^{-q}+b\right). \tag{S.22}$$

Then expanding (S.20) as in those theorems,

$$\begin{aligned} P\left(mF_T^* \le z\right) &= G_m\left(z\frac{B}{B-m+1}\right) + G_m'(z)zE[L+Q] + \frac{1}{2}EG_p''(z)z^2E[L^2] + o\left((bT)^{-q}+b\right) + O\left(\frac{1}{T}\right) \\ &= G_m(z) - G_m'(z)z\omega^{(q)}k^{WOS(q)}(0)(bT)^{-q} - G_m'(z)z\frac{\psi}{B}(m-1) + G_p''(z)z^2\frac{\psi}{B} \\ &\quad + o(b) + o\left((bT)^{-q}\right). \end{aligned} \tag{S.23}$$

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.19) 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). We omit the calculations here, since they follow the same steps as in those papers and above, but they are available upon request. Those calculations yield the following expansion:

$$\begin{aligned} \Pr_\delta\left[F_T^* \le c_m^\alpha(b)\right] = &\ G_{m,\delta^2}(\chi_m^\alpha) - G_{m,\delta^2}'(\chi_m^\alpha)\chi_m^\alpha\omega^{(q)}k^{(q)}(0)(B/T)^q \\ &+ \frac{1}{2}\delta^2G_{m+2,\delta^2}'(\chi_m^\alpha)\chi_m^\alpha\frac{\psi}{B} + o(b) + o\left((bT)^{-q}\right). \end{aligned} \tag{S.24}$$

Rearranging gives (19).

(vi) Parts (ii)-(v) apply directly. For part (i), see Proposition S3 below. □

S3.2 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

$$t^{SS} = \sqrt{B+1}\left(\bar{\hat{\beta}} - \beta_0\right)\Big/\sqrt{S_{\hat{\beta}}^2}\text{, where } S_{\hat{\beta}}^2 = \frac{1}{B}\sum_{i=1}^{B+1}\left(\hat{\beta}^{(i)} - \bar{\hat{\beta}}\right)^2, \tag{S.25}$$

where $\hat{\beta}^{(i)}$ is the estimator of β computed using the i^{th} 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

$$\Phi^{SS} = \sqrt{(B+1)}\left(I_{B+1}\otimes\iota_{T/(B+1)}\right)M_\iota^B, \tag{S.26}$$

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

$$\begin{aligned}
\hat{\Omega}^{SS} &= [T/(B+1)]S_{\hat{\beta}}^2 = [T/(B+1)]B^{-1}\breve{\beta}'\left(I_{B+1} - \iota_{B+1}(\iota_{B+1}'\iota_{B+1})^{-1}\iota_{B+1}'\right)\breve{\beta} \\
&= [T/(B+1)]B^{-1}[T/(B+1)]^{-2}y'\left(I_{B+1}\otimes\iota_{T/(B+1)}\right)\left(I_{B+1} - \iota_{B+1}(\iota_{B+1}'\iota_{B+1})^{-1}\iota_{B+1}'\right)\left(I_{B+1}\otimes\iota_{T/(B+1)}'\right)y \\
&= (BT)^{-1}(B+1)\hat{z}'\left(I_{B+1}\otimes\iota_{T/(B+1)}\right)\left(I_{B+1} - \iota_{B+1}(\iota_{B+1}'\iota_{B+1})^{-1}\iota_{B+1}'\right)\left(I_{B+1}\otimes\iota_{T/(B+1)}'\right)\hat{z} \\
&= (BT)^{-1}(B+1)\hat{z}'\left(I_{B+1}\otimes\iota_{T/(B+1)}\right)M_\iota^B M_\iota^{B\prime}\left(I_{B+1}\otimes\iota_{T/(B+1)}'\right)\hat{z} \\
&= \hat{z}'\Phi^{SS}\Phi^{SS\prime}\hat{z}/BT,
\end{aligned} \tag{S.27}$$

where the first equality uses the definition of $\hat{\Omega}^{SS}$, the second applies the expression for $S_{\hat{\beta}}^2$, the third uses $\breve{\beta} = [T/(B+1)]^{-1}\left[I_{B+1}\otimes\iota_{T/(B+1)}'\right]y$, the fourth uses that $\hat{z} = y - \bar{\hat{\beta}}$ 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.27).

Note that Φ^{SS} is $T\times B$, that $\iota_T'\Phi^{SS} = 0$, and $\Phi^{SS\prime}\Phi^{SS}/T = I_B$ 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 the condition that $\sup_{s\in[0,1]}\left|\phi_j^{(n)}(s)\right| \le C_{n,\phi}\, j^{2n+1/2}$, for all j and n = 0, 1, 2, where $\phi_j^{(n)}(s)$ is the n^{th} derivative of the basis function 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 $\int_0^1 \phi_j(s)\phi_k(s)ds = 1\{j = k\}$ (see, e.g., Abramowitz and Stegun (1965, p. 774)), as required by definition. Note that this implies that $|\phi_j(s)|/\sqrt{j} \le \sqrt{3}$, satisfying the requirement for the 0^{th} 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

$$\frac{d^n}{dx^n} P_j(x) = \frac{d^n}{dx^n} P_j^{(\alpha,\beta)}(x)\bigg|_{\alpha=0,\beta=0} = \frac{\Gamma(j+1+n)}{2^n \Gamma(j+1)} P_{j-n}^{(n,n)}(x), \tag{S.28}$$

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

$$P_j'(\pm 1) = \frac{1}{2}(\pm 1)^{j-1} j(j+1), \quad P_j''(\pm 1) = \frac{1}{8}(\pm 1)^j (j-1)j(j+1)(j+2). \tag{S.29}$$

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 5). For EWP and QS tests with equivalent higher-order size,

$$\Pr_\delta\left[F^*_{QS,T} > c_{QS,\alpha}(b_{QS})\right] - \Pr_\delta\left[F^*_{EWP,T} > c_{EWP,\alpha}(b_{EWP})\right] \approx \frac{1}{2}\delta^2 G'_{m+2,\delta^2}(\chi_m^\alpha)\chi_m^\alpha\left(\nu_{EWP}^{-1} - \nu_{QS}^{-1}\right)$$

$$= \frac{1}{2}\delta^2 G'_{m+2,\delta^2}(\chi_m^\alpha)\chi_m^\alpha\left(1 - \frac{6\sqrt{3}}{5\sqrt{5}}\right)B^{-1}, \tag{S.30}$$

where $\nu_{EWP} = B$.

Proof of Proposition S3: Fix a sequence $B = 1/b_{\text{EWP}}$. To obtain equivalent higher-order size using the QS test, Theorem 1(iv) gives that we must set

$$b_{\text{QS}} = \sqrt{\frac{k^{\text{QS}(2)}(0)}{k^{\text{EWP}(2)}(0)}} b_{\text{EWP}} = \sqrt{\frac{\pi^2/10}{\pi^2/6}} B^{-1} = \sqrt{\frac{3}{5}} B^{-1}, \tag{S.31}$$

where the $k^{(2)}(0)$ values for the two tests are as in the proof of Theorem 5. That proof also uses that $\int_{-\infty}^{\infty} k^2(x)dx = \frac{6}{5}$ for QS, so that given equivalent higher-order size, we have $\nu_{\text{EWP}}^{-1} - \nu_{\text{QS}}^{-1} = B^{-1} - \frac{6}{5}\sqrt{\frac{3}{5}} B^{-1}$. Plugging this into the higher-order power difference in Theorem 3 yields the desired result. □

Proposition S4 (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 S4:

(i) We first consider the SS estimator. Note that the SS basis functions (S.26) do not satisfy the differentiability requirement of Assumption 3. Thus for the SS estimator we calculate $E\hat{\Omega}^{\text{SS}}$ 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 that $\bar{y}_i - \bar{y} = \frac{1}{T}\sum_{t=1}^{T}\left((B+1)1\{t \in T_i\} - 1\right) y_t = \frac{B+1}{T}\sum_{t=1}^{T}\left(1\{t \in T_i\} - \frac{1}{B+1}\right) y_t$. Thus squaring and summing over subsamples, we have

$$\frac{1}{B}\sum_{i=1}^{B+1}\left(\bar{y}_i - \bar{y}\right)^2 = \frac{1}{B}\sum_{i=1}^{B+1}\left(\frac{B+1}{T}\right)^2 \sum_{t=1}^{T}\sum_{s=1}^{T}\left(1\{t \in T_i\} - \frac{1}{B+1}\right)\left(1\{s \in T_i\} - \frac{1}{B+1}\right) y_t y_s . \tag{S.32}$$

Taking the expectation of this value and rearranging,

$$\begin{aligned} E\frac{1}{B}\sum_{i=1}^{B+1}(\bar{y}_i - \bar{y})^2 &= \frac{B+1}{B}\frac{1}{T/B+1}\sum_{u=-(T-1)}^{T-1}\left[\left(1 - \left|\frac{u}{T/B+1}\right|\right)1\left\{|u| \le \frac{T}{B+1}\right\} - \frac{1}{B+1}\left(1 - \left|\frac{u}{T}\right|\right)\right]\Gamma_u \\ &= \frac{B+1}{T}\sum_{u=-(T-1)}^{T-1}\left[\left(\frac{B+1}{B} - \frac{B+1}{B}\left|\frac{u}{T/(B+1)}\right|\right)1\left\{|u| \le \frac{T}{B+1}\right\} - \frac{1}{B} + \frac{1}{B}\left|\frac{u}{T}\right|\right]\Gamma_u . \end{aligned} \tag{S.33}$$

Converting $E\frac{1}{B}\sum_{i=1}^{B+1}(\bar{y}_i - \bar{y})^2$ to $E\hat{\Omega}^{\text{SS}}$ requires multiplying by $T/(B+1)$ given the form of the statistic in (S.25) 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.33)) as

$$k_{\text{B}}^{\text{SS}}(\tilde{v}) = \left(\frac{B+1}{B} - \frac{B+1}{B}|\tilde{v}|\right)1\left\{|\tilde{v}| \le 1\right\} - \frac{1}{B} + \frac{1}{B(B+1)}|\tilde{v}|. \tag{S.34}$$

Thus using the definition of the generalized first derivative in (11), we have $k_{\text{B}}^{\text{SS}(1)}(0) = \frac{B+1}{B} - \frac{1}{B(B+1)} = \frac{B+2}{B+1} \to 1$ as $B \to \infty$. Because $k^{\text{SS}(1)}(0) \neq 0$, $q = 1$ for the SS estimator. Further, comparing $E\hat{\Omega}^{\text{SS}}$ with Ω using (S.33), we obtain that Theorem 1(i) applies for the SS estimator as well, and the tradeoff in Theorem 4(ii) applies. The value $\ell^{(1)}(k^{\text{SS}})$ is equal to $k^{\text{SS}(1)}(0) = 1$ 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 $\ell^{(1)}(k^{\text{NW}}) = k^{(1)}(0)\int_{-\infty}^{\infty} k^2(x)dx = 2/3$, from which we conclude that the Bartlett tradeoff dominates the SS tradeoff, as stated.

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

$$k_B^{\text{EWP}(2)}(0) = -\frac{1}{2}\frac{1}{B}\sum_{j=1}^{B/2}\frac{1}{B^2}2\left(-4\pi^2 j^2\right) = \frac{\pi^2}{6}\frac{(B+1)(B+2)}{B^2} \xrightarrow{B\to\infty} \frac{\pi^2}{6}. \tag{S.35}$$

Similarly, for cosine basis functions, using their limiting implied mean kernel form, we have $\phi''_j(s) = -\sqrt{2}\pi^2 j^2 \cos(\pi js)$ and $\int_0^1 \phi_{2,j-1}(s)\phi''_{2,j-1}(s)ds = -\pi^2 j^2$. Summing over j yields

$$k_B^{\cos(2)} = -\frac{1}{2}\frac{1}{B}\sum_{j=1}^{B}\frac{1}{B^2}\left(-\pi^2 j^2\right) = \frac{\pi^2}{6}\frac{(B+1)(B+1/2)}{B^2} \xrightarrow{B\to\infty} \frac{\pi^2}{6}. \tag{S.36}$$

Results (S.35) and (S.36) 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 S5 (Section 4.3). Assume that the condition provided in Footnote 18 holds. Then:

(i) The maximum weighted average power (WAP) test solving equation (31) features

$$b^{WAP} = q^{\frac{1}{1+q}}\tilde{d}_{m,\alpha,q}\left(\frac{k^{(q)}(0)}{\psi}\right)^{\frac{1}{1+q}}\left(\tilde{\omega}^{(q)}\right)^{\frac{1}{1+q}}T^{\frac{-q}{1+q}}, \tag{S.37}$$

as stated in equation (32), with

$$\tilde{\omega}^{(q)} = \int_{|\rho|\le\bar{\rho}}\left[\bar{\omega}^{(q)}(\rho) - \omega^{(q)}(\rho)\right]d\Pi(\rho), \tag{S.38}$$

$$\tilde{d}_{m,\alpha,q} = \left\{ \int_\delta G'_{m,\delta^2}(\chi_m^\alpha)\chi_m^\alpha d\Pi_\delta(\delta) \Bigg/ \left[\tfrac{1}{2} \int_\delta \delta^2 G'_{m+2,\delta^2}(\chi_m^\alpha)\chi_m^\alpha d\Pi_\delta(\delta) \right] \right\}^{\frac{1}{1+q}}. \quad \text{(S.39)}$$

(ii) The power loss of the test using the WAP-maximizing sequence in (32) depends on k only through $\ell^{(q)}(k)$, so that Grenander's (1951) uncertainty principle provided in Remark 3 applies for WAP-maximizing tests.

(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 S5:

(i) Let $\overline{c}_{m,T}^\alpha(b)$ be the size-adjusted critical value (20) based on the boundary value of $\overline{\omega}^{(q)}$: $\overline{c}_{m,T}^\alpha(b) = \left[1+\overline{\omega}^{(q)}k^{(q)}(0)(bT)^{-q}\right]c_m^\alpha(b)$. It follows from (18) that the null rejection rate of the test using this size-adjusted critical value, evaluated at the true value of $\omega^{(q)}$, is

$$\Pr_0\left[F_T^* > \overline{c}_{m,T}^\alpha(b)\right] = \alpha + G'_m(\chi_m^\alpha)\chi_m^\alpha\left(\omega^{(q)} - \overline{\omega}^{(q)}\right)k^{(q)}(0)(bT)^{-q} + o(b) + o\left((bT)^{-q}\right), \quad \text{(S.40)}$$

from which it follows that, for a given sequence b and under the condition in Footnote 18,

$$\sup_{\omega^{(q)} \le \overline{\omega}^{(q)}} \Pr_0\left[F_T^* > \overline{c}_{m,T}^\alpha(b)\right] \le \alpha + o(b) + o\left((bT)^{-q}\right). \quad \text{(S.41)}$$

The expression for $\Delta_P\left(\omega^{(q)}(\rho),\delta\right)$ in (31) then follows directly from (41) 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

$$\Delta_P^{WAP} = \left(q^{-q/(1+q)} + q^{1/(1+q)}\right)\tilde{a}_{m.\alpha.q}\left[\left(k^{(q)}(0)\right)^{1/q}\psi\right]^{\frac{q}{1+q}}\left(\tilde{\omega}^{(q)}\right)^{\frac{1}{1+q}}T^{\frac{-q}{1+q}}, \quad \text{(S.42)}$$

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

Supplemental Figures

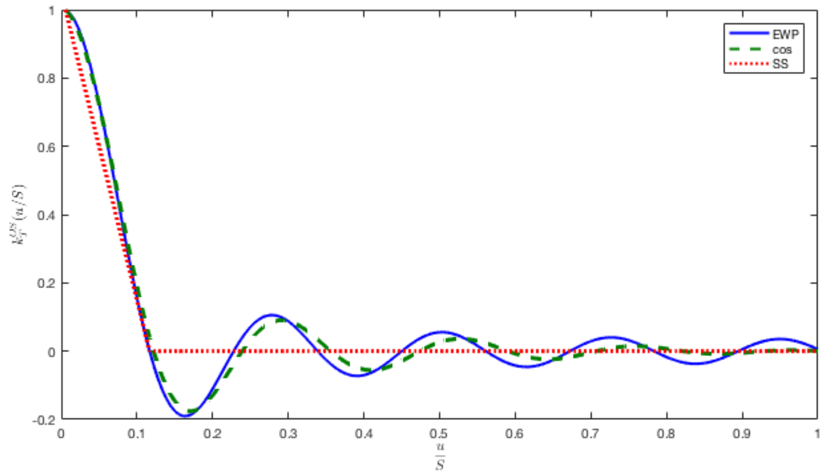

Figure S1. Implied mean kernel of basis function estimators with B = 8, time domain: Fourier/EWP (blue, solid), cosine (green, dash), and split-sample (red, dot).

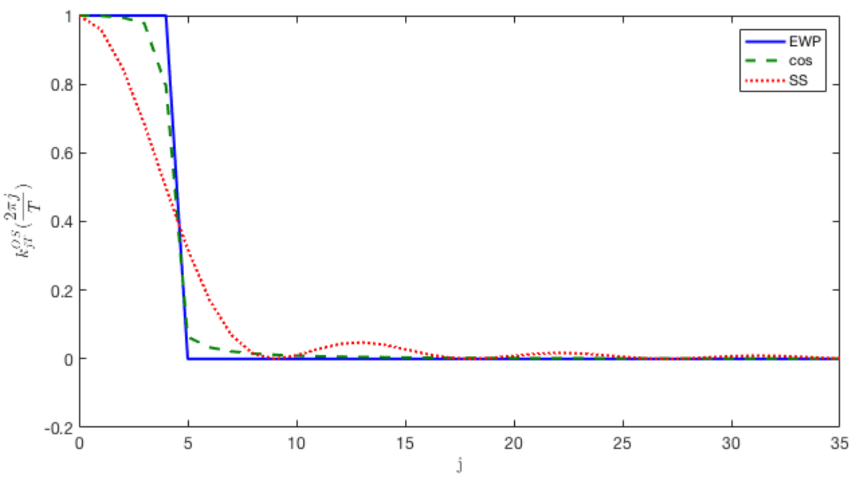

Figure S2. Implied mean kernel of basis function estimators with B = 8, frequency domain: Fourier/EWP (blue, solid), cosine (green, dash), and split-sample (red, dot). The frequency domain kernel is normalized to 1 at $\omega = 0$ and computed over the periodogram ordinates (so the horizontal axis value j corresponds to $2\pi j/T$).

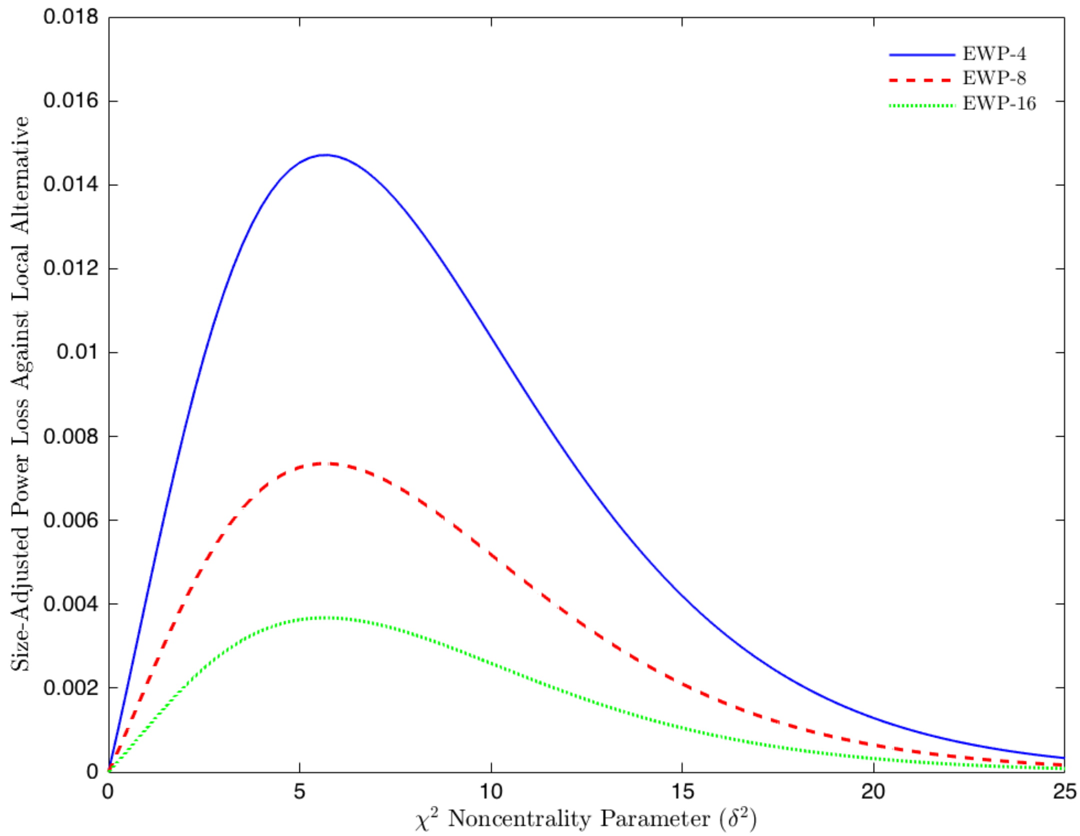

Figure S3. Small-*b* approximation to power loss for EWP test, compared to QS test, for different values of *B* in the EWP test and with *b* for the QS test chosen so that the EWP and QS tests have the same higher-order size when evaluated using fixed-*b* critical values. The figure plots the final expression in (30) as a function of δ. Gaussian location model, *m*=1, 5% significance level.

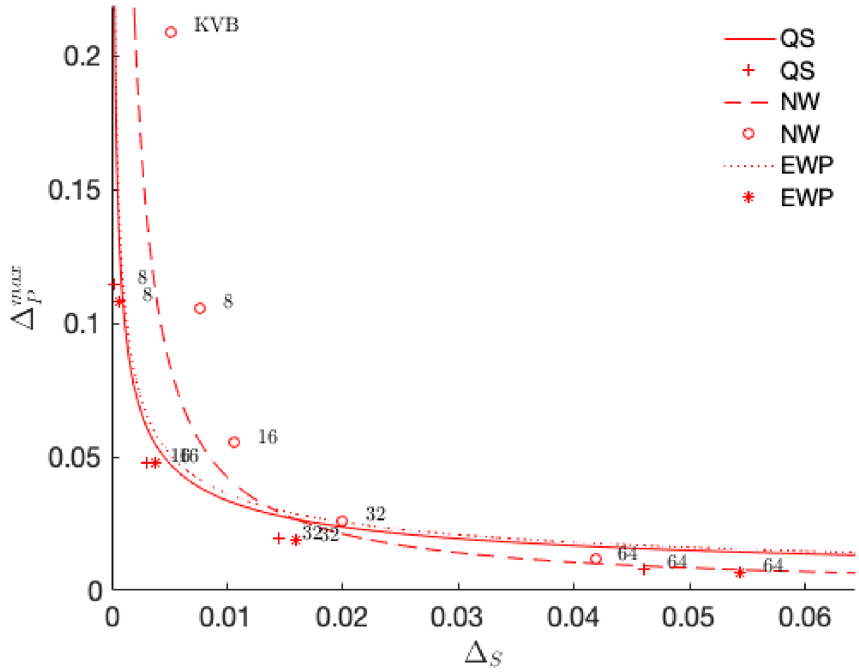

Figure S4. Theoretical (lines) and Monte Carlo (symbols) size distortion/power loss curves for QS, Newey-West, and EWP estimators: Location model, $m = 1$, AR(1), $\rho = 0.5$, and $T = 200$.

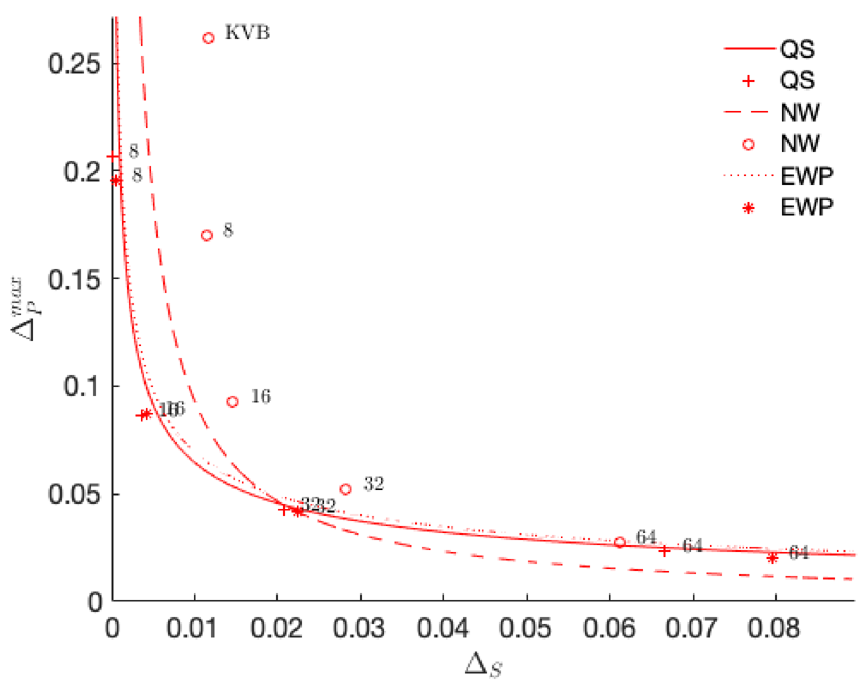

Figure S5. Theoretical (lines) and Monte Carlo (symbols) size distortion/power loss curves for QS, Newey-West, and EWP estimators: Location model, $m = 2$, AR(1), $\rho = 0.5$, $T = 200$.

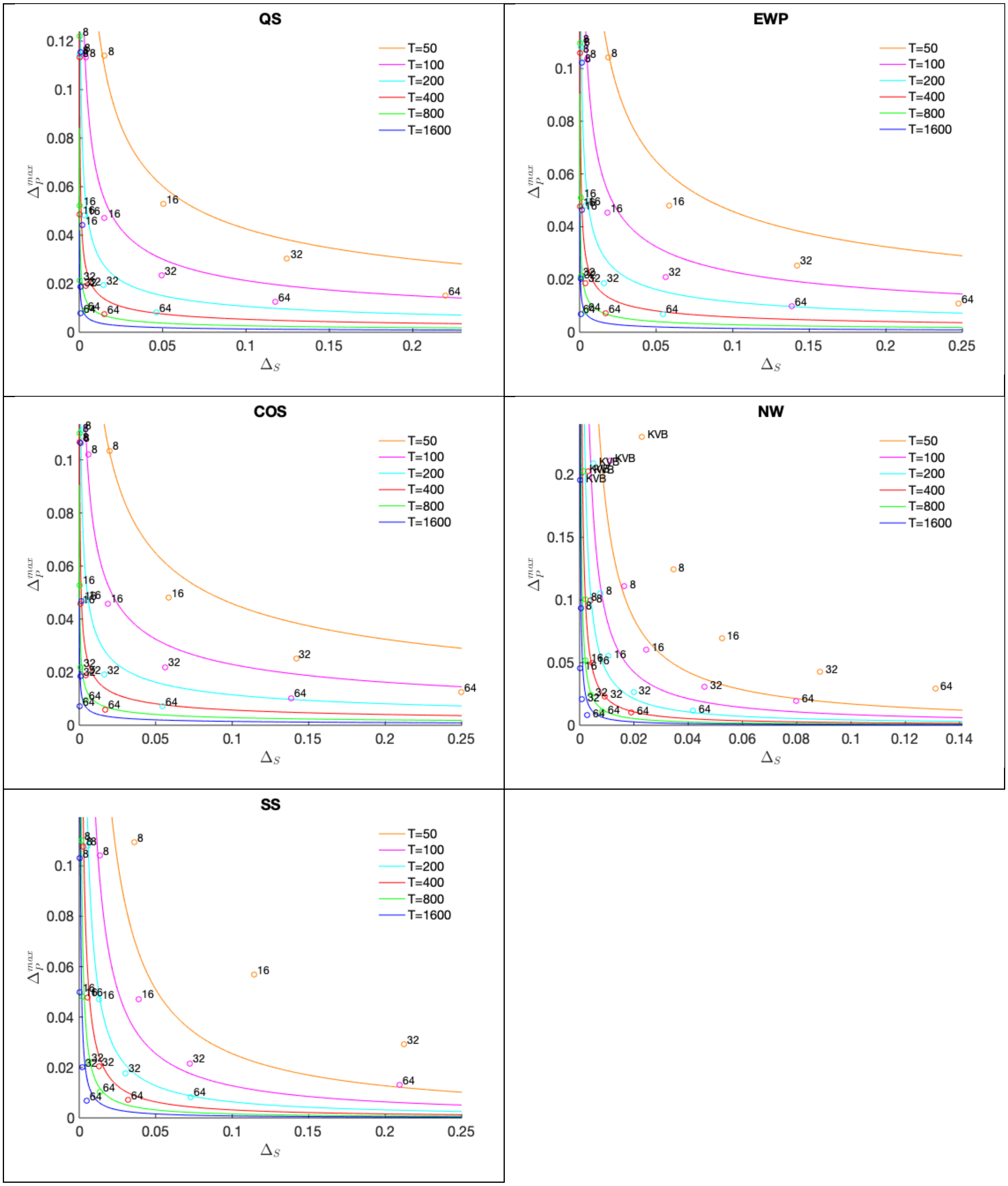

Figure S6. Location model, AR(1), m = 1, ρ = 0.5. Theoretical size distortion/power loss tradeoff curves for each estimator with Monte Carlo results for T ranging from 50 to 1600.

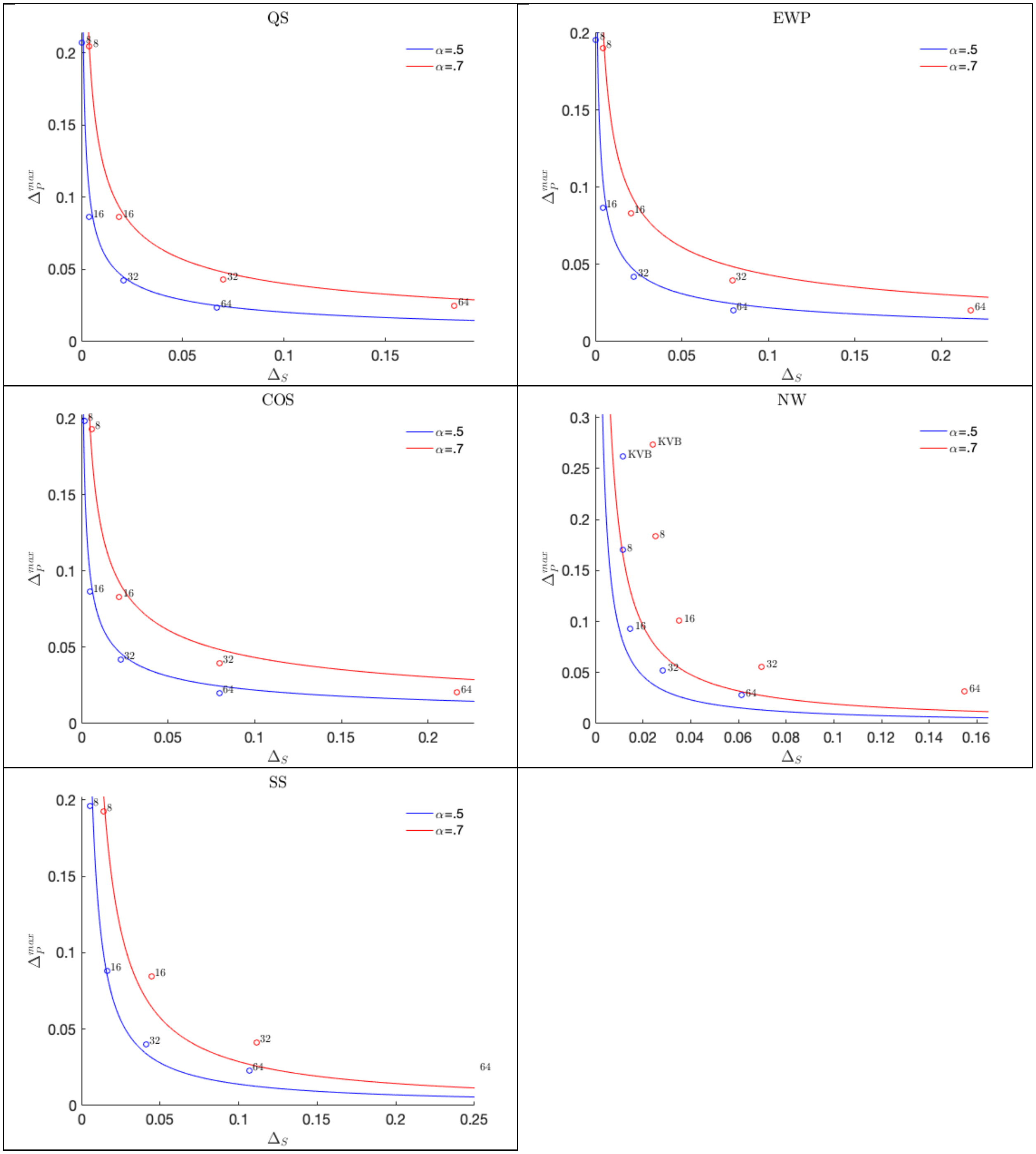

Figure S7. Location model, AR(1), m = 2, ρ = .5 and .7, T = 200. Theoretical size distortion/power loss tradeoff curves for each estimator and Monte Carlo results (dots).

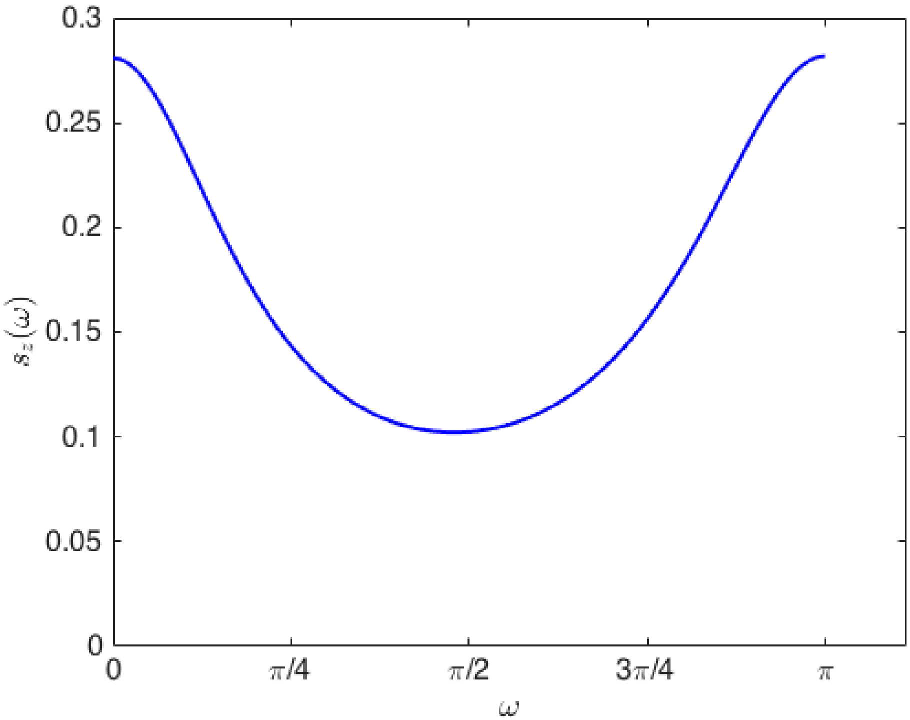

Figure S8. Spectral density of calibrated ARMA(2,1), $\omega^{(2)} = 4$.

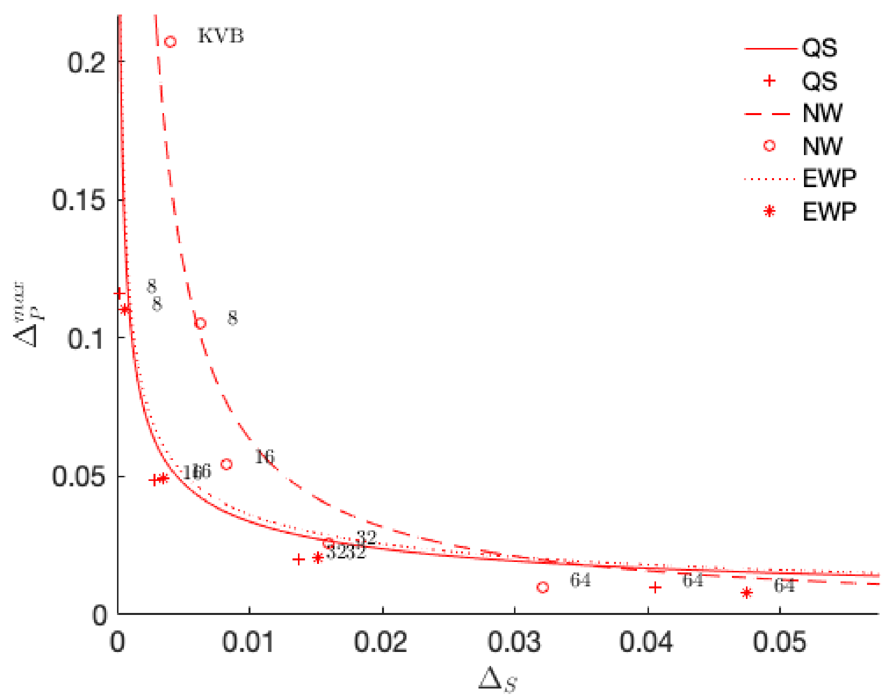

Figure S9. Location model, ARMA(2,1), m = 1, T = 200. Theoretical size distortion/power loss tradeoff curves for QS, Newey-West, and EWP estimators with Monte Carlo results. ARMA(2,1) parameters fixed such that $\omega^{(2)} = 4$, equivalent to AR(1) with $\alpha = 0.5$ (parameter values as in Figure S8).

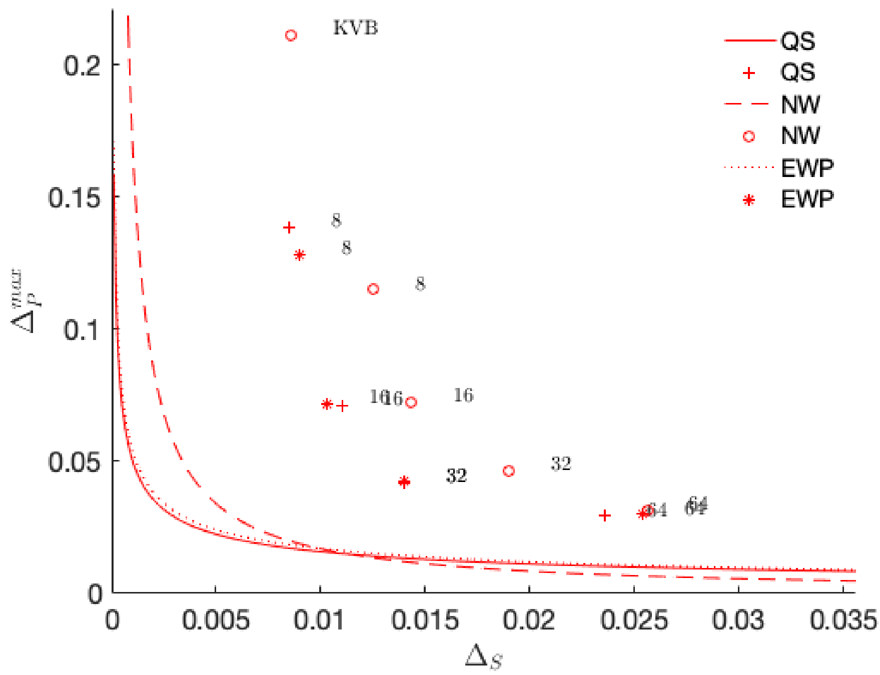

Figure S10. Theoretical (lines) and Monte Carlo (symbols) size distortion/power loss curves for QS, Newey-West, and EWP estimators: Stochastic regressor, m = 1, AR(1), ρ = 0.5, T = 200. Theoretical curves are for the Gaussian location model.

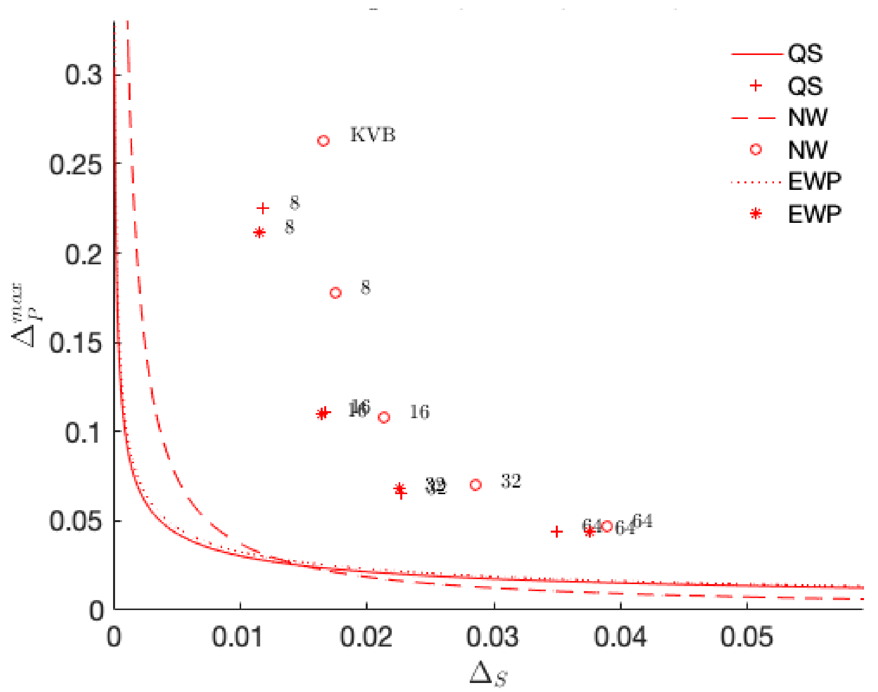

Figure S11. Stochastic regressor, AR(1), $m = 2$, $\rho = 0.5$, $T = 200$. Theoretical size distortion/power loss tradeoff curves for QS, Newey-West, and EWP estimators with Monte Carlo results. Theoretical curves are for the Gaussian location model.