Econometric Reviews

ISSN: 0747-4938 (Print) 1532-4168 (Online) Journal homepage: https://www.tandfonline.com/loi/lecr20

Wavelet energy ratio unit root tests

Mirza Trokić

To cite this article: Mirza Trokić (2016): Wavelet energy ratio unit root tests, Econometric Reviews, DOI: 10.1080/07474938.2016.1222232

To link to this article: https://doi.org/10.1080/07474938.2016.1222232

Accepted author version posted online: 22 Aug 2016.
Published online: 26 Oct 2016.

ECONOMETRIC REVIEWS
http://dx.doi.org/10.1080/07474938.2016.1222232

Taylor & Francis
Taylor & Francis Group

Wavelet energy ratio unit root tests

Mirza Trokić

Department of Economics, Bilkent University, Ankara, Turkey

ABSTRACT
This article uses wavelet theory to propose a frequency domain nonparametric and tuning parameter-free family of unit root tests. The proposed test exploits the wavelet power spectrum of the observed series and its fractional partial sum to construct a test of the unit root based on the ratio of the resulting scaling energies. The proposed statistic enjoys good power properties and is robust to severe size distortions even in the presence of serially correlated MA(1) errors with a highly negative moving average (MA) parameter, as well as in the presence of random additive outliers. Any remaining size distortions are effectively eliminated using a novel wavestrapping algorithm.

KEYWORDS
Fractional Brownian motion; fractional integration; hypothesis test; size distortion; statistical power; time series; unit root; variance ratio statistic; wavelet energy ratio; wavestrapping; wavelets

JEL CLASSIFICATION
C01; C12; C15; C21; C40; C46; C63

1. Introduction

Testing for the presence of a unit root is an important empirical exercise, and early seminal works of Dickey and Fuller (1979), Phillips (1987b), and Phillips and Perron (1988) have inspired a pleiad of unit root tests. The lot of these tests however, are plagued by poor statistical power, severe size distortions, and tuning parameter (e.g., lag length, bandwidth, kernel choice.) selection. These are well-recognized issues in unit root models with a linear trend and serially correlated moving average (MA) errors, particularly when the MA root is highly negative. While power and size suffer due to the dissolution of the unit root framework as the MA root approaches negative unity (see Campbell and Perron, 1991), tuning parameter selection renders finite sample performance dependent on tuning parameter specifications without reflecting this specification in the limiting distribution of the statistic. Significant efforts have been made to improve these shortcomings. Elliott et al. (1996) address both size and power issues through point optimal tests, power envelopes, and generalized least squares (GLS) detrending of augmented Dickey-Fuller (ADF) tests. Ng and Perron (2001) and Perron and Qu (2007) address low size and power through optimized truncation lag selection. Similarly, issues concerning tuning parameter selection prompted the development of tuning parameter-free unit root tests as in Park and Choi (1988), Park (1990), Breitung (2002), and Nielsen (2009).

Some thirteen years before the first time domain unit root test of Dickey and Fuller (1979), Granger (1966) had observed that the majority of economic series exhibit power spectra that are characterized by the “overpowering importance of the low frequency components” which are amplified by the presence of trends in mean. Still, the majority of unit root tests, and in fact all those mentioned above, are constructed in the time domain. There are two important exceptions: Choi and Phillips (1993) and Fan and Gençay (2010). Whereas the former relies on Fourier spectral analysis, the latter exploits wavelet theory. This distinction is an important one. Fourier transforms lack a time resolution and are localized only in frequency. This renders Fourier analysis an excellent tool for studying stationary time series. Wavelet transforms, however, are localized both in frequency and time. Accordingly, wavelets are ideally adapted for the study of nonstationary series. Since economic and financial data often exhibit nonstationary patterns over time such as trends, jumps, kinks, volatility clustering, etc., this renders wavelet transforms

CONTACT Mirza Trokić trokic@bilkent.edu.tr Department of Economics, Bilkent University, Ankara 06800, Turkey.
Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/lecr.

a *de facto* natural platform for the construction of frequency domain unit root tests. See Gençay et al. (2001) for an exposition on the contrasts between Fourier and wavelet transforms.

Whereas the primary advantage of the Fan and Gençay (2010) (henceforth FG) test is high statistical power, like many tests in the literature, it is subject to violent size distortions, particularly in the presence of deterministic dynamics and MA serial correlations with a high negative root. Although FG do not consider MA errors at all, simulation evidence in this article cautions against the illusion of power gains in the presence of severe size distortions as size and power are positively related. Furthermore, since the FG test uses the Newey and West (1987) estimator of the long run error variance which requires a suitably chosen kernel bandwidth parameter q, the FG test is not considered tuning parameter-free since q is not reflected asymptotically. In contrast, the Nielsen (2009) test (henceforth NVR) enjoys good power, at times much better than the FG test, is also subject to severe size distortions (albeit less than the FG test), but is tuning parameter-free by design. Moreover, Nielsen (2009) handles size distortions through a sieve bootstrap algorithm of Chang and Park (2003), albeit at the cost of sacrificing the tuning parameter-free property of the statistic.

In light of the above, this article constructs a family of nonparametric, tuning parameter-free, wavelet-based tests for the autoregressive unit root hypothesis. These tests possess good asymptotic power, consistently discriminate the null and alternative hypotheses (see Müller, 2008), and are significantly more robust to size distortions in the presence of errors with highly negative MA roots than either the NVR or FG tests. This is particularly desirable in empirical work on nonstationary economic time series. As shown in Schwert (1987, 1989) and Dods and Giles (1995) for instance, various macroeconomic time series (e.g., inflation rates, stock market volatility) are known to exhibit serial correlation with highly negative MA roots. Similarly, in microeconomic time series (e.g., union strikes, consumer hoarding behavior in face of tax incentives), Franses and Haldrup (1994) demonstrate that large and frequently occurring additive outliers in the levels of nonstationary time series mimic the behavior of highly negative MA roots. Since the proposed tests are designed to filter the frequency range characterizing MA processes with roots approaching negative unity, they are less affected by their presence. This renders the proposed test particularly well suited to the analysis of the aforementioned class of problems. Finally, any size distortions are addressed using a novel *wavestrapping* algorithm. The latter proves more effective than sieve bootstrapping in reducing severe size distortions and leaves the statistic tuning parameter-free. All proofs are contained in the Appendix.

2. Wavelet power spectrum

Wavelet techniques differ from classical spectral tools in that the former can extract not only frequency but also temporal information from an input signal.[1] It is precisely this feature which makes wavelets an ideal tool for *multiresolution analysis* (MRA) — the analysis of signals at different frequencies with varying resolutions.[2] Moving along the time domain, MRA allows one to *zoom* to a desired level of detail such that high (low) frequencies yield good (poor) time resolutions and poor (good) frequency resolutions. Since economic time series often exhibit multiscale features, wavelet techniques can effectively decompose these series into constituent processes associated with different time scales. For instance, since nonstationary series have dominating lower frequency components relative to stationary series, one can exploit this distinction to identify series as $I(1)$ or $I(0)$. This distinction was recognized in FG and will also be exploited in the construction of the new test.

Formally, a wavelet is a real valued function $\psi(\cdot)$ satisfying $\int_{-\infty}^{\infty} \psi(t)\, dt = 0$ and $\int_{-\infty}^{\infty} \psi^2(t)\, dt = 1$. In other words, wavelets integrate to zero and have unit energy[3] and nonzero range. The continuous

[1]Borrowed terminology will be referenced throughout the article. The term *signal* refers to a data source, e.g., a time series.
[2]MRA was introduced in Mallat (1989).
[3]The term *energy* originates from the signal processing literature. It is formalized as $\int_{\infty}^{\infty} |f(t)|^2 dt$, for some function $f(t)$. Restricting $f(t)$ to the real plane, energy and variance are effectively synonymous and will henceforth be used interchangeably.

wavelet transform (CWT) of a time series $y(t)$ is then defined as

$$W(a,b) = \int_{-\infty}^{\infty} y(t)\, \psi_{a,b}^{*}(t)\, dt,$$

where $\psi_{a,b} = \frac{1}{\sqrt{a}} \psi_{a,b}\left(\frac{t-b}{a}\right)$, and $*$ denotes the complex conjugate. See Percival and Walden (2006) for a detailed exposition.

Since continuous functions are rarely observed, the CWT is empirically impractical and a discretized analogue known as the discrete wavelet transform (DWT) is used. Characterizing the DWT are $\mathbf{h} = (h_1, \ldots, h_l)$ and $\mathbf{g} = (g_1, \ldots, g_l)$ — the wavelet (high pass) and scaling (low pass) filters of dyadic length l, respectively. Formally, $\mathbf{h}$ and $\mathbf{g}$ are related through the *quadrature mirror relationship*.[4] Since the DWT is also an orthonormal transform, high and low pass filters exhibit additional orthogonality conditions.[5] The DWT of an input then ensues by filtering the observed series $\mathbf{y} = \{y_t\}_{t=0}^{T}$ with both high and low pass filters, where $y_0 = 0$. This yields two series as follows: the first extracting high frequency behavior of y_t, and the second extracting its low frequency behavior.

In practice, DWT coefficients are derived through the Mallat (1989) *pyramid algorithm*. In this regard, for $T = 2^M$, define the level m matrix of DWT coefficients as $[\mathbf{W}_1, \ldots, \mathbf{W}_m, \mathbf{V}_m]^{\top}$, for all $1 \leq m \leq M$.[6] Here, $\mathbf{W}_m$ and $\mathbf{V}_m$ are $(2^{-m}T \times 1)$ vectors of wavelet and scaling coefficients,[7] respectively, and are associated with changes and averages, respectively, on scales of length $\lambda_m = 2^{m-1}$. The algorithm can now be formalized as a sequence of m iterative convolutions of the input signal with filters $\mathbf{h}$ and $\mathbf{g}$, respectively, to render $[\mathbf{W}_1, \ldots, \mathbf{W}_m, \mathbf{V}_m]^{\top}$. These convolutions are formalized as

$$\mathbf{e}_{m,t}^{\top} \mathbf{W}_m = \sum_{i=1}^{l} h_i \mathbf{e}_{m-1,2t-(i-1)(\text{mod}T)}^{\top} \mathbf{V}_{m-1} \quad \mathbf{e}_{m,t}^{\top} \mathbf{V}_m = \sum_{i=1}^{l} g_i \mathbf{e}_{m-1,2t-(i-1)(\text{mod}T)}{}^{\top} \mathbf{V}_{m-1},$$

where $\mathbf{e}_{m,t} = (0, \ldots, 0, 1, 0, \ldots, 0)^{\top}$ is the canonical basis vector in $\mathbb{R}^{2^{-m}T}$, and $\mathbf{V}_0 = \mathbf{y}$. Each iteration, therefore, convolves the scaling coefficients from the preceding iteration with both the high and low pass filters. The entire algorithm continues until $m = M$, although it can be stopped earlier.

An important property of the DWT transform above is the conservation of energy. It follows from the orthonormality of the DWT generating matrix $\mathcal{W}$ satisfying $[\mathbf{W}_1, \mathbf{V}_1]^{\top} = \mathcal{W}_{\mathbf{y}}$. Here, orthonormality of $\mathcal{W}$ implies $\mathcal{W}^{\top}\mathcal{W} = \mathcal{W}\mathcal{W}^{\top} = I_T$ is an identity matrix of dimension T, and therefore $||\mathbf{y}||^2 = ||\mathcal{W}^{\top} [\mathbf{W}_1, \mathbf{V}_1]^{\top}||^2 = ||\mathbf{W}_1||^2 + ||\mathbf{V}_1||^2$, where $||\cdot||$ denotes the Euclidean norm. In fact, the result can be extended to demonstrate decomposition of energy on a scale-by-scale basis. The latter formalizes as

$$||\mathbf{y}||^2 = \sum_{m=1}^{M} ||\mathbf{W}_m||^2 + ||\mathbf{V}_m||^2. \tag{1}$$

Thus, $||\mathbf{W}_m||^2$ quantifies the energy of $\mathbf{y}$ accounted for at scale λ_m. Moreover, $||\mathbf{W}_m||^2/T$ is the contribution to the sample variance of $\mathbf{y}$ at scale λ_m. This decomposition is known as the *wavelet power spectrum* (WPS) and is arguably the most insightful of the properties of the DWT.

[4]This relationship states that $h_i = (-1)^i g_{l-1-i}, g_i = (-1)^{i+1} h_{l-1-i}$ for $i = 0, \ldots, l-1$.

[5]Both filters exhibit orthogonality to even shifts. Formally, $\sum_{i=0}^{l-1} h_i h_{i+2n} = \sum_{i=0}^{l-1} g_i g_{i+2n} = \sum_{i=0}^{l-1} g_i h_{i+2n} = 0, \forall n \in \mathbb{Z}_{+}$.

[6]Limiting series to dyadic lengths is restrictive. Methods such as the *maximum overlap discrete wavelet transform* (MODWT) otherwise known as the *non-decimated* DWT overcome this shortcoming.

[7]While $\mathbf{W}_m$ and $\mathbf{V}_m$ implicitly depend on l, the notation is suppressed for notational brevity.

3. Wavelet energy ratio tests

Recall that the FG unit root test relativizes the energy of the scaling coefficients to that of total energy. Specifically, their statistic and limiting distributions are formalized as

$$\tau^{FG} = \frac{4T^2\hat{\omega}^2}{\hat{\gamma}_0}\left(\frac{||\mathbf{W}_1||^2}{||\mathbf{W}_1||^2 + ||\mathbf{V}_1||^2}\right),$$

where $\hat{\gamma}_0^2$ consistently estimates $E\left\{u_{2t}^2\right\}$, and $\hat{\omega}^2$ consistently estimates the long-run variance of $\{u\}_T^{t=1}$ using a Bartlett kernel with bandwidth q. Since q is not reflected in the limiting distribution, τ^{FG} is therefore not tuning parameter-free.[8] It bears noticing, however, that one can exploit the WPS to construct an alternative unit root test that is entirely nonparametric and tuning parameter-free. Specifically, the new test relativizes the energy of the scaling coefficients to that of its fractionally differenced transform. The result is a family of nonparametric and tuning parameter-free tests indexed by the fractional parameter d. They will henceforth be referred to as *wavelet scaling ratio* (WSR) tests.

To motivate the new construction, consider a simple AR(1) (near) unit root model augmented with possibly time varying deterministic components. Specifically, consider the model.

$$y_t = \boldsymbol{\gamma}\delta_t + x_t, \tag{2}$$

$$x_t = \phi x_{t-1} + u_t, \tag{3}$$

$$\phi = 1 - c_\phi/T, \tag{4}$$

$$u_t = \psi(L)\epsilon_t, \tag{5}$$

where T is the sample size, L is the lag operator, $\psi(L) = \sum_{j=0}^{\infty}\psi_j L^j$ introduces serial correlation, and $c_\phi/T \in [0,2)$ is the localization constant which interprets x_t as the possibly near unit root process of Phillips (1987a). Moreover, when $\boldsymbol{\gamma} = (\gamma_0, \gamma_1)$ and $\boldsymbol{\delta}_t = (1,t)^\top$, the model is augmented with common deterministic specifications $\boldsymbol{\gamma}\boldsymbol{\delta}_t$. For instance, when $\boldsymbol{\gamma} = \mathbf{0}$, y_t reduces to x_t, and when $\boldsymbol{\gamma} \neq \mathbf{0}$, y_t models an integrated process with nonzero mean and/or linear trend. In this regard, let $\widehat{\boldsymbol{\gamma}}$ denote the ordinary least squares (OLS) estimator of $\boldsymbol{\gamma}$ from regression (2), rendering the residuals $\widehat{y}_t = y_t - \widehat{\boldsymbol{\gamma}}\boldsymbol{\delta}_t$ the OLS detrended analogues of y_t. Unless otherwise specified, subsequent analyses are conducted over $\widehat{\mathbf{y}} = \{\widehat{y}_t\}_{t=0}^{T}$.

Consider further the fractional partial sum process

$$\widetilde{z}_t = \Delta_+^{-d} z_t = (1-L)_+^{-d} z_t = \sum_{k=0}^{t-1}\frac{\Gamma(d+k)}{\Gamma(d)\Gamma(k+1)} z_{t-k} = \sum_{k=0}^{t-1}\pi_k(d) z_{t-k},$$

where $d \in \mathbb{R}$ and Δ_+^{-d} is the truncated version of the binomial expansion in L. Next, let $B(t)$ represent a standard Brownian motion process, and denote by $J_{c_\phi}(t)$ and $\widetilde{J}_{c_\phi}(t,d)$ the standard and fractional variants

[8]Since test consistency requires bandwidth parameters to expand at specific rates relative to sample size, both finite sample and asymptotic performance are highly dependent on the tuning parameter choice while the latter is not reflected in the asymptotic distribution. In this regard, the bandwidth choice q is considered a tuning parameter; cf. Nielsen (2009). Moreover, since the bandwidth determines the proportion of information in the covariance structure that is used in the estimate of the long run variance, failing to select the right bandwidth for covariances that dissipate slowly will result in imprecise estimates; cf. Andrews (1991) and Newey and West (1994). Accordingly, the Newey and West (1987) rule of thumb choice $q = 4(T/100)^{2/9}$ used in Fan and Gençay (2010) may not always be appropriate when the underlying process is highly persistent. As pointed out in Kiefer and Vogelsang (2000), "traditional asymptotics requires the bandwidth to increase with sample size but the fraction of sample autocovariances used goes to zero. Thus, information contained in sample autocovariances must be ignored for the asymptotics to work." This partly explains why simulation results in this article show that the FG test adapt poorly to MA serial correlations when the MA parameter is highly negative.

of the Ornstein-Uhlenbeck (O-U) process with parameter c_ϕ, respectively, defined as

$$J_{c_\phi}(t) = B(t) - c_\phi \int_0^t e^{-c_\phi(t-r)} B(r) dr, \tag{6}$$

$$\widetilde{J}_{c_\phi}(t,d) = \frac{1}{\Gamma(d+1)} \int_0^t (t-r)^d dJ_{c_\phi}(r). \tag{7}$$

Moreover, let $\widehat{J}_{c_\phi}(t)$ and $\widetilde{\widetilde{J}}_{c_\phi}(t,d)$ denote the OLS detrended variants of $J_{c_\phi}(t)$ and $\widetilde{J}_{c_\phi}(t,d)$, respectively, defined as

$$\widehat{J}_{c_\phi}(t) = J_\phi(t) - \left(\int_0^1 J_{c_\phi}(r)\mathbf{D}(r)^\top dr\right)\left(\int_0^1 \mathbf{D}(r)\mathbf{D}(r)^\top dr\right)^{-1} \mathbf{D}(t), \tag{8}$$

$$\widetilde{\widetilde{J}}_{c_\phi}(t,d) = \widetilde{J}_{c_\phi}(t,d) - \left(\int_0^1 J_{c_\phi}(r)\mathbf{D}(r)^\top dr\right)\left(\int_0^1 \mathbf{D}(r)\mathbf{D}(r)^\top dr\right)^{-1} \int_0^t \frac{(t-r)^{d-1}}{\Gamma(d)} \mathbf{D}(r) dr, \tag{9}$$

where $\boldsymbol{D}(r) = \left(\mathbb{1}_{\{\gamma_0 \neq 0\}}, r\mathbb{1}_{\{\gamma_1 \neq 0\}}\right)^\top$ and the indicator function $\mathbb{1}_{\{\bullet\}}$ equals 1 when $\bullet$ is true, and zero otherwise. Finally, complete the setup with the following assumptions.

Assumption 1. *(a) $\{\epsilon_t, \mathcal{F}_t\}$ is a MDS with respect to some filtration $\mathcal{F}_t$ and $E\{\epsilon_t^2|\mathcal{F}_t\} = \sigma^2 < \infty$.*
(b) $\sup_{t\in\mathbb{Z}} E\left\{|\epsilon_t|^p\right\} < \infty$ for $p > \max\{2, 2/(2d+1)\}$ and $d > -1/2$.
(c) $\sum_{j=0}^\infty |\psi_j| < \infty$, $\sum_{j=0}^\infty j|\psi_j| < \infty$, and $b_\psi = \sum_{j=0}^\infty \psi_j \neq 0$.

The regularity conditions (a) through (c) are primarily required to invoke (fractional) functional central limit theorems (FCLTs) and allow for a relatively flexible dependence structure in u_t which includes stationary and invertible Autoregressive Moving Average (ARMA) processes. As shown in Johansen and Nielsen (2012), Assumption (b) is necessary and has a long standing tradition in the literature since Davydov (1970). Although the assumption can be very strong when d is close to $-1/2$, several important FCLTs for fractional processes such as Marinucci and Robinson (2000), Davidson and De Jong (2000), Tanaka (1999), Wang et al. (2003), and Lee and Shie (2004), rest on it. Assumption (c) is also salient when defining u_t as a linear process of ϵ_t. An alternative specification is also possible with $\sum_{j=0}^\infty j^{1/2-d}|\psi_j| < \infty$; see Phillips and Solo (1992) for a discussion when $d = 0$, and Wang et al. (2003) when $d > -1/2$.

Next, let $\widetilde{\widetilde{y}}_t = \Delta_+^{-d}\widehat{y}_t$ and $\widetilde{\widetilde{\mathbf{y}}} = \left\{\widetilde{\widetilde{y}}_t\right\}_{t=0}^T$, let $\longrightarrow_d$ denote convergence in distribution, and consider the battery of local to unity hypotheses $H : c_\phi/T \in [0,2)$. Provided $d > 0$ and assumption 1 hold, recall that the limiting distribution of the Nielsen (2009) variance ratio statistic is characterized as

$$\tau^N(d) = T^{2d}\frac{||\widehat{\mathbf{y}}||^2}{||\widetilde{\widetilde{\mathbf{y}}}||^2} \longrightarrow_d \frac{\int_0^1 \widehat{J}_{c_\phi}(s)^2 ds}{\int_0^1 \widetilde{\widetilde{J}}_{c_\phi}(s,d)^2 ds}.$$

In particular, it follows that under the unit root hypothesis $H_0 : \phi = 1$ when $\boldsymbol{\gamma} = \mathbf{0}$, $J_{c_\phi}(t)$ and $\widetilde{J}_{c_\phi}(t,d)$ respectively reduce to $B(t)$ and $B_{d+1}(t)$, where $B_{d+1}(t)$ denotes the type II fractional Brownian motion[9]

$$B_{d+1}(t) = \frac{1}{\Gamma(d+1)} \int_0^t (t-s)^d dB(s). \tag{10}$$

[9]See Davidson and Hashimzade (2009) for a discussion on type I and type II fractional Brownian motions.

The Nielsen (2009) test is in fact a generalization of the classical variance ratio test. It is entirely nonparametric and requires neither estimation of the long-run variance of y_t nor the short term dynamics when u_t exhibits serial correlation. Furthermore, it consistently discriminates the unit root null from alternative hypotheses of stationarity, see Müller (2008). Lastly, since both $\tau^N(d)$ and its asymptotic distributions are indexed by d, the latter is not considered a tuning parameter.

3.1. WSR tests

Recall that the proportion of total energy in y_t, associated with the scaling coefficients at level m, is $||\mathbf{V}_m||^2/||\mathbf{y}||^2$. Consider also the fractional ratio $||\mathbf{V}_m||^2/||\widetilde{\mathbf{V}}_m||^2$ where $\widetilde{\mathbf{V}}_m = \Delta_+^{-d}\mathbf{V}_m$. Figure 1 illustrates these relative energies and fractional ratios when scaling coefficients are generated from a Gaussian white noise process z_t, and a random walk process $y_t = y_{t-1} + z_t$. In particular, Fig. 1 demonstrates that $||\mathbf{V}_m||^2/||\widetilde{\mathbf{V}}_m||^2$ is uniformly (across m) smaller when the associated process is a random walk in contrast to white noise. This polarity derives from the inverse proportionality of frequency length to wavelet scales λ_m. In other words, lowering λ_m stretches (renders less precise) the frequency resolution but compresses (renders more precise) the time resolution. Accordingly, the low pass filters which render $\mathbf{V}_m$ are well adapted to capturing persistent effects and one expects $||\mathbf{V}_m||^2$ to be larger and $||\mathbf{V}_m||^2/||\widetilde{\mathbf{V}}_m||^2$ to be smaller when y_t is a random walk as opposed to white noise.[10] It stands to reason, therefore, that one can use this polarity to test for unit roots. The idea is formalized as the WSR test as follows:

$$\tau_m^{WSR}(d) = \left(2^{-m}T\right)^{2d} \frac{||\mathbf{V}_m||^2}{||\widetilde{\mathbf{V}}_m||^2}. \tag{11}$$

To render the subsequent analysis statistically tractable, it is necessary to analytically characterize the wavelet functions. Here, the analysis is adapted to Daubechies wavelets — an important class of orthogonal wavelet functions indexed by the maximal number of vanishing moments for a given support. Specifically, a wavelet has p vanishing moments if and only if the associated scaling function can recover polynomials of degree $k \leq p - 1$. It is worth noting here that although p is an appropriate index, the nomenclative hierarchy of the Daubechies class is typically indexed by the wavelet length $l = p/2$. For instance, the well-known Haar wavelet belongs to the class with $l = 2$ or $p = 1$. In other words, the Haar scaling function has length $l = 2$ and generates constants. Since the objects of primary interest

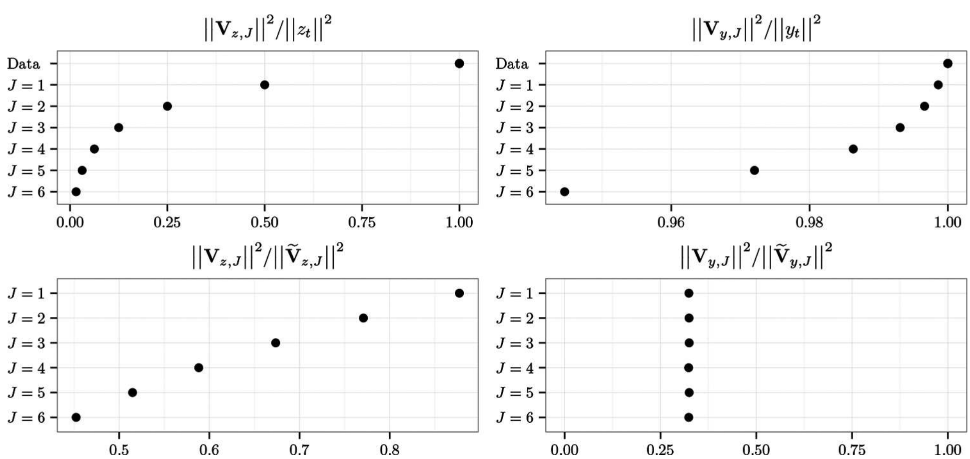

Figure 1. Haar filter Level 6 DWT relative scaling energy decomposition of z_t and $y_t = y_{t-1} + z_t$. Results are derived over 5000 MC replications with $T = 2^{10}$ and fractional parameter $d = 0.10$. *Data* represents the total energy of the input signal.

[10]Granger (1966) first noticed that it is the ill behaved frequencies near the origin which indicate the presence of a unit root.

are the vectors of scaling coefficients, a tedious application of backward substitution on Mallat's pyramid algorithm reduces the tth element of $\widehat{\mathbf{V}}_m$, the level m Daubechies length l scaling vector, to

$$\mathbf{e}_{m,t}^{\top}\widehat{\mathbf{V}}_m = \eta_m(L)\widehat{y}_{2^m t+(2^m-1)(l-2) \pmod T}, \tag{12}$$

$$\eta_m(L) = \prod_{j=1}^{m}\left(g_1 L^{(l-1)2^{j-1}} + g_2 L^{(l-2)2^{j-1}} + \cdots + g_{l-1}L^{2^{j-1}} + g_l\right), \tag{13}$$

where L is the usual lag operator, $\sum_{i=1}^{l} g_i = \sqrt{2}$ and $\sum_{i=1}^{l} g_i^2 = 1, t = 1, \ldots, 2^{-m}T$, and the $\widehat{\bullet}$ notation indicates that the DWT is taken over detrended series $\widehat{\mathbf{y}}$. Furthermore, for $r \in [0,1]$, define the partial sum processes

$$\widehat{\mathbf{V}}_{m,T}(r) = \left(2^{-m}T\right)^{-1/2}\mathbf{e}_{m,\lfloor 2^{-m}Tr\rfloor}^{\top}\widehat{\mathbf{V}}_m, \tag{14}$$

$$\widehat{\widetilde{\mathbf{V}}}_{m,T}(r) = \left(2^{-m}T\right)^{-(d+1/2)}\Delta_{+}^{-d}\widehat{\mathbf{V}}_{m,T}(r). \tag{15}$$

The following lemma, which is of independent interest, characterizes the limiting distributions of $\widehat{\mathbf{V}}_{m,T}(r)$ and $\widehat{\widetilde{\mathbf{V}}}_{m,T}(r)$.

Lemma 1. *Provided assumption 1 hold and y_t is generated by Eqs. (2) to (5), under the battery of hypotheses $H : c_\phi/T \in [0,2)$, for any $m \in \mathbb{Z}_+$, $l < \infty$, $d > -1/2$, and $T \longrightarrow \infty$, we have the following situations:*
1. When $c_\phi = 0$,

$$\widehat{\mathbf{V}}_{m,T}(r) \longrightarrow_d 2^{2m}\psi(1)\sigma_\epsilon\widehat{J}_{c_\phi}(r),$$
$$\widehat{\widetilde{\mathbf{V}}}_{m,T}(r) \longrightarrow_d 2^{2m}\psi(1)\sigma_\epsilon\widehat{\widetilde{J}}c_\phi(r,d);$$

2. When $c_\phi/T \in (0,2)$,

$$\widehat{\mathbf{V}}_{m,T}(r) \longrightarrow_d 2^m\left(\frac{1-\phi^{2^m}}{1-\phi}\right)\psi(1)\sigma_\epsilon\widehat{J}_{c_\phi}(r),$$
$$\widehat{\widetilde{\mathbf{V}}}_{m,T}(r) \longrightarrow_d 2^m\left(\frac{1-\phi^{2^m}}{1-\phi}\right)\psi(1)\sigma_\epsilon\widehat{\widetilde{J}}c_\phi(r,d).$$

The result of lemma 1 is particularly important as it states that for any fixed l and m, under the null of unit root, the scaling vectors follow a Brownian motion, while under the alternative, they follow an O-U process. Turning now to the limiting distribution of the WSR statistic, the following result holds.

Theorem 1. *Provided assumption 1 hold and y_t is generated by Eqs. (2) to (5), under the battery of hypotheses $H : c_\phi/T \in [0,2)$, for any $m \geq 1$, $l < \infty$, $d > -1/2$, and $T \longrightarrow \infty$*

$$\tau_m^{WSR}(d) = \left(2^{-m}T\right)^{2d}\frac{\widehat{||\mathbf{V}||_m}^2}{||\widehat{\widetilde{\mathbf{V}}}_m||^2} \longrightarrow_d \frac{\int_0^1 \widehat{J}_{c_\phi}(s)^2 ds}{\int_0^1 \widehat{\widetilde{J}}_{c_\phi}(s,d)^2 ds}. \tag{16}$$

Theorem 1 establishes that the WSR and NVR tests are asymptotically equivalent. Moreover, in contrast to FG, where τ^{FG} depends on a kernel bandwidth choice q for consistent estimation of the long run variance (although q is not reflected in the asymptotic distribution), $\tau_m^{WSR}(d)$ is by design nuisance parameter-free as d is reflected in the asymptotic distribution. Furthermore, since Daubechies wavelet filters approach the ideal high-pass filter as l grows, FG have suggested that power gains may be achieved by increasing l. While the conclusion is plausible in finite samples, Theorem 1 clearly demonstrates that l is not reflected in the asymptotic distribution of $\tau_m^{WSR}(d)$. Moreover, since d indexes the WSR family of statistics, it is natural to ask whether there exists a d which maximizes local asymptotic power for said family? Simulation analysis in Nielsen (2009) suggests $d = 0.1$. Although this is not a global optimum as choices of $d < 0.1$ yield uniformly (in ρ_c) higher asymptotic local power, the choice is justified since choosing d too small results in severe size distortions. Similar conclusions hold in the case of the WSR statistic. Finally, unlike the FG and NVR tests, it is worth noting that the WSR test is by design, relatively inert to the presence of highly negative MA serial correlation roots. This is a consequence of using scaling energies which filter the frequency band $[0, 1/2]$ which corresponds to the frequency band characterizing MA processes with roots approaching negative unity.

3.2. Wavestrapped WSR statistic

Although the WSR test is particularly effective at reducing size distortions, it leaves much to be desired; see Section 4 for details. While further reductions are possible with the sieve bootstrap of Bühlmann (1997) or Chang and Park (2003), here, a novel *wavestrapping* algorithm is applied to do the same. This has two important advantages. First, unlike sieve bootstrap algorithms, wavestrapping does not require regression fitting as an algorithmic step. Second, whereas the sieve bootstrap depends on lag length specifications for the AR sieve, wavestrapping requires no tuning parameter specifications. This is particularly important as bootstrapping the NVR statistic in Nielsen (2009) forces dependence on the sieve length tuning parameter, thereby rendering $\tau^{NVR}(d)$ no longer tuning parameter-free. This is clearly not a concern with wavestrapping, and both the original and wavestrapped WSR statistics remain tuning parameter-free.

Wavestrapping, first developed in Percival et al. (2000) to resample statistics derived from the spectral density function, is a bootstrap-like procedure applied to wavelet transforms of a time series. The governing principle is, as shown in Flandrin (1992), that a DWT approximately decorrelates long memory processes. This approximate decorrelation lends itself to the application of bootstrap procedures by rendering approximately independent replicates of the wavelet coefficients. These can then be used to reconstruct independent replicates of the underlying time series process through DWT inversion. Nevertheless, Percival et al. (2000) claim the procedure works poorly for short memory processes such as MA(1) since the DWT of such series may not produce adequately decorrelated wavelet coefficients. Instead, they suggest using the *discrete wavelet packet transform* (DWPT) as the underlying decorrelating transform in a top-down search for a collection of least correlated wavelet coefficients based on adaptive white-noise tests.

The DWPT generalizes the DWT and involves filtering both wavelet and scaling coefficients. At each level m, this produces 2^m wavelet coefficients: 2^{m-1} coefficients corresponding to the low-pass filtering of the $(m-1)$th level wavelet coefficients, and another 2^{m-1} coefficients resulting from the low-pass filtering of the $(m-1)$th level scaling coefficients. The result is a *wavelet packet* (WP) table shown in Fig. 2, which nests the original DWT as $\mathbf{W}_1 = \mathbf{W}_{1,1}, \mathbf{W}_2 = \mathbf{W}_{2,1}, \ldots, \mathbf{W}_m = \mathbf{W}_{m,1}, \mathbf{V}_m = \mathbf{W}_{m,0}$. The idea is to perform a white noise test on the coefficients in each row of the WP table. If the null hypothesis that said coefficients are a sample from a white noise process is rejected, the row is discarded; otherwise it is retained. Resampling then proceeds on the retained rows which are inverted to obtain a wavestrapped version of the original input. The algorithm is formalized in what follows.

1. Fix the Monte Carlo replications MC and the nominal size α.
2. Given $\mathbf{u} = \{u_t\}_{t=1}^{T}$ of length $T = 2^M$, compute a level $M_0 = M-2$, DWPT. (Enter Step 4 with starting values $j = n = 0$ and $W_{0,0} = \mathbf{u}$.)

<table>
<tr><td>$m = 0$</td><td colspan="8">$\boldsymbol{W}_{0,0}$</td></tr>
<tr><td>$m = 1$</td><td colspan="4">$\boldsymbol{W}_{1,0}$</td><td colspan="4">$\boldsymbol{W}_{1,1}$</td></tr>
<tr><td>$m = 2$</td><td colspan="2">$\boldsymbol{W}_{2,0}$</td><td colspan="2">$\boldsymbol{W}_{2,1}$</td><td colspan="2">$\boldsymbol{W}_{2,2}$</td><td colspan="2">$\boldsymbol{W}_{2,3}$</td></tr>
<tr><td>$m = 3$</td><td>$\boldsymbol{W}_{3,0}$</td><td>$\boldsymbol{W}_{3,1}$</td><td>$\boldsymbol{W}_{3,2}$</td><td>$\boldsymbol{W}_{3,3}$</td><td>$\boldsymbol{W}_{3,4}$</td><td>$\boldsymbol{W}_{3,5}$</td><td>$\boldsymbol{W}_{3,6}$</td><td>$\boldsymbol{W}_{3,7}$</td></tr>
</table>

Figure 2. Wavelet packet table.

3. Use $\mathbf{u}$ to compute the statistic of interest τ using. For the WSR statistic, extract the DWT coefficients from the WP table in Step 2 and use them to compute $\tau \equiv \tau_m^{WSR}(d)$.
4. If $j = M_0$, retain $W_{j,n}$; if $j < M_0$, do a white noise test on $W_{j,n}$. If the null hypothesis is not rejected, retain $W_{j,n}$. If it is rejected, transform $W_{j,n}$ into $W_{j+1,2n}$ and $W_{j+1,2n+1}$ and discard $W_{j,n}$. Repeat this step on $W_{j+1,2n}$ and on $W_{j+1,2n+1}$.
5. Set B to some large number such that $\alpha(B+1)$ is an integer[11] and resample with replacement B times from each retained subvector from Step 4.
6. Apply the inverse DWPT to each resampled vector in Step 5 and obtain a wavestrapped series $\mathbf{u}^\star$, $b = 1, \ldots, B$ in the time domain. Use $\mathbf{u}^\star$ to compute a wavestrapped statistics $\tau^\star$ in the same way $\mathbf{u}$ was used to compute τ in Step 3.
7. Let $\mathbf{1}\{\cdot\}$ denote the indicator function and compute the wavestrap p-value as

$$p^\star = \frac{1}{B}\sum_{b=1}^{B} \mathbf{1}\left\{\tau < \tau^\star\right\}$$

8. Repeat Steps 1 through 7 MC times and obtain $p_i^\star$, $i = 1, \ldots, MC$.
9. Compute the wavestrap size distortion as

$$RP^\star = \frac{1}{MC}\sum_{i=1}^{MC} \mathbf{1}\left\{p_i^\star < \alpha\right\}.$$

This algorithm requires the computation of $MC(B+1)$ statistics. This is essentially a double bootstrap procedure and can be expensive to compute even by today's standards. Fortunately, the *fast double bootstrap* (FDB) procedure of Davidson and MacKinnon (2007) reduces the number of computations to $2MC$. The idea is to set $B = 1$ and estimate $RP^\star$ as

$$RP^\star \simeq RP^\star_{FDW} = \frac{1}{MC}\sum_{i=1}^{MC} \mathbf{1}\left\{\tau < Q^\star(\alpha)\right\},$$

where $Q^\star(\alpha)$ is the empirical α-quantile of $\tau^\star$ and the subscript FDW reflects the *fast double wavestrap* context. Size distortion can now be computed as $RP^\star - \alpha$. Like all bootstrap algorithms, the result of Basawa et al. (1991) suggests that wavestrapping should be performed under the null hypothesis. Simulation exercises below demonstrate that wavestrapping effectively eliminates most size distortions exhibited by the WSR statistic and therefore proves to be an effective alternative to classical bootstrap algorithms.

4. Simulation analysis

Finite sample reliability is the ultimate benchmark of test performance and the WSR test is especially attractive in this regard. The test is decidedly effective at reducing severe size distortions in the presence of negative MA serial correlations parameters, linear trends, and random additive outliers. Generally, simulations indicate that the WSR test has the smallest size distortion among the FG and NVR tests, while wavestrapping routines for the WSR test all but eliminate size distortions for even the most problematic scenarios. Moreover, size-adjusted local asymptotic power simulations show that the WSR test dominates the FG test, in some cases even uniformly, for many important scenarios.

The simulations under consideration focus on three empirical designs that typically test the limits of unit root tests. In particular, the unit root hypothesis is tested under a typical AR(1) framework with 1) MA serially correlated errors; 2) linear deterministic dynamics; and 3) random additive outlier dynamics. These paradigms are only natural considering that they arise in many macroeconomic time series configurations and generate a platform where many unit root tests are known to suffer severe size distortion and power loss; see Schwert (1987), see Evans (1991), Franses and Haldrup (1994), Dods and

[11] See Davidson and MacKinnon (2000) for details.

Giles (1995), Ng and Perron (2001), and Nielsen (2008). In general, all three designs are readily nested in the following DGP:

$$\begin{aligned} y_t &= \boldsymbol{\gamma}\boldsymbol{\delta}_t + \pi_t\lambda + x_t, \\ x_t &= \left(1 - c_\phi/T\right) x_{t-1} + u_t, \\ u_t &= \psi(L)\epsilon_t, \end{aligned}$$

where λ is the magnitude of the additive outlier and π_t is a Bernoulli random variable with support $\{0, 1\}$ with probability $p \in (0, 1)$, and zero otherwise. In other words, $P(\pi_t = 1) = p$ and $P(\pi_t = 0) = 1 - p$. The particular appeal of the setup is that Franses and Haldrup (1994) show that additive outliers can induce effects resembling highly negative MA roots in the innovation process, which are known to induce severe size distortions in most unit root tests.

To formalize matters, each simulation compares the level $m \in \{1, 2, 3\}$ WSR test to the NVR and FG tests, over 10,000 Monte Carlo replications with significance level $\alpha = 0.05$, MA(1) serial correlation $\psi(L) = 1 + \psi_1 L$, and sample sizes $T = \{64, 128, 256\}$.[12] All size distortion exercises also include $\tau_m^{WSR\star}$ and $\tau_m^{WSR\star\star}$, respectively the DWT and DWPT versions of the wavestrapped WSR statistic, while local asymptotic power simulations are all adjusted for size and derived over $\phi = 1 - c_\phi/T \in [0.8, 1]$.[13] Finally, simulations for the NVR and WSR tests are computed with $d = 0.1$ and $d = 0.05$, respectively. To mitigate d exhibiting inverse proportionality to T, Nielsen (2009) argues that d should not be lowered too much since the test degenerates as $d \longrightarrow 0$. In finite samples, the effect is reflected through increased size distortion. Nevertheless, as the exercises in Tables 1 to 3 clearly show that the WSR is generally least size distorted, lowering d to 0.05 seems justified. Finally, the FG and WSR tests are both computed using the Haar filter.

Consider size distortion in the model without additive outliers first. In this regard, Table 1 lists rejection frequencies for the classical and detrended variants of tests under consideration. Specifically, while size distortions are evidently problematic for all three tests, they are clearly most troublesome for the FG test, particularly for larger sample sizes. In fact, problems are only exacerbated with the inclusion of linear trends with the FG test exhibiting both violent oversizing and undersizing when ψ_1 approaches negative and positive unity, respectively. Meanwhile, whereas the standard and detrended NVR test performs passably well when $\psi_1 \in [0, 1]$, it too suffers severe size distortion when ψ_1 is near negative unity. On the other hand, while the WSR test clearly dominates both the FG and NVR tests when sample size is large and $\psi_1 \in [-1, 0)$, the test is prone to severe undersizing when $m > 1$ and ψ_1 is close to positive unity. This is particularly troublesome for the detrended statistic with higher wavelet orders. In this regard, increasing m is not particularly advised. Fortunately, the DWT and DWPT wavestrapping algorithm prove rather efficient with the wavestrapped variants of the WSR test generally exhibiting near nominal size.

On the other hand, several patterns emerge in the model with additive outliers. In particular, Table 2 shows that size distortions increase with larger outlier magnitudes, exhibit parabolic patterns in outlier frequency with peaks near $p = 0.4$, and for a fixed (λ, p) pair, generally decrease as sample size increases. Although these patterns pervade all three tests, it is clear that size distortions are again most problematic for the FG test. Similarly, the NVR test, while reasonably sized for very small and very large values of p, is nonetheless highly unattractive otherwise. In contrast, the WSR test stands out as being most resilient to drastic size distortions with rejection frequencies never exceeding 23%, although it can be undersized when p is very large. Moreover, while increasing m can reduce size distortions further still, as in the

[12]Simulations could also have been conducted using the MODWT. Since the MODWT does not suffer the decimation at each scale like the DWT, it may produce further finite sample improvements over the DWT, particularly in terms of power since the MODWT produces series of the same length as the input signal. This is not pursued, however, since the DWT is significantly quicker to compute, requiring $O(T)$ computations vs. $O(T \log_2 T)$ for the MODWT.

[13]Due to excessive size distortion differentials among the NVR, FG, and WSR tests, size adjusted power uses empirical size rather than nominal size α to deflate (inflate) oversized (undersized) tests and calibrate power curves to a common reference point. While this renders different tests directly comparable, the exercise requires Monte Carlo simulations and, therefore, as argued in Horowitz and Savin (2000), is "irrelevant for empirical research."

Table 1. Rejection frequencies without additive outliers: $\lambda = 0$.

					$m=1$			$m=2$			$m=3$		
$\gamma\delta_t$	T	θ	τ^{N}	τ^{FG}	τ_m^{WSR}	$\tau_m^{WSR\star}$	$\tau_m^{WSR\star\star}$	τ_m^{WSR}	$\tau_m^{WSR\star}$	$\tau_m^{WSR\star\star}$	τ_m^{WSR}	$\tau_m^{WSR\star}$	$\tau_m^{WSR\star\star}$
$\gamma\delta_t = 0$													
	64												
		−0.875	0.6862	0.7595	0.4293	0.1607	0.0522	0.1474	0.0581	0.0275	0.0061	0.0273	0.0282
		−0.75	0.3702	0.5619	0.1759	0.0601	0.0323	0.0450	0.0267	0.0248	0.0015	0.0237	0.0309
		−0.625	0.2113	0.3941	0.0915	0.0374	0.0313	0.0240	0.0280	0.0287	0.0006	0.0247	0.0364
		−0.5	0.1288	0.2601	0.0573	0.0351	0.0349	0.0142	0.0285	0.0348	0.0004	0.0339	0.0424
		−0.25	0.0645	0.0808	0.0321	0.0427	0.0444	0.0130	0.0405	0.0459	0.0007	0.0443	0.0479
		0	0.0429	0.0140	0.0250	0.0490	0.0502	0.0084	0.0448	0.0500	0.0003	0.0507	0.0498
		0.5	0.0320	0.0004	0.0193	0.0553	0.0544	0.0105	0.0526	0.0522	0.0004	0.0527	0.0509
		0.875	0.0309	0.0003	0.0210	0.0535	0.0551	0.0086	0.0576	0.0525	0.0004	0.0595	0.0516
	128												
		−0.875	0.6212	0.8906	0.4165	0.1323	0.0368	0.2058	0.0426	0.0178	0.0514	0.0189	0.0204
		−0.75	0.3010	0.6235	0.1634	0.0392	0.0253	0.0685	0.0191	0.0213	0.0194	0.0154	0.0286
		−0.625	0.1678	0.4288	0.0857	0.0224	0.0279	0.0441	0.0208	0.0300	0.0136	0.0230	0.0388
		−0.5	0.1064	0.2868	0.0587	0.0273	0.0355	0.0284	0.0243	0.0393	0.0119	0.0327	0.0446
		−0.25	0.0610	0.1082	0.0429	0.0369	0.0463	0.0238	0.0378	0.0476	0.0100	0.0404	0.0485
		0	0.0449	0.0349	0.0358	0.0460	0.0508	0.0209	0.0446	0.0500	0.0102	0.0505	0.0500
		0.5	0.0379	0.0075	0.0289	0.0492	0.0525	0.0226	0.0537	0.0531	0.0087	0.0497	0.0506
		0.875	0.0399	0.0061	0.0339	0.0566	0.0531	0.0210	0.0485	0.0523	0.0106	0.0577	0.0506
	256												
		−0.875	0.5170	0.9038	0.3589	0.0805	0.0235	0.1898	0.0288	0.0138	0.0771	0.0095	0.0170
		−0.75	0.2318	0.6132	0.1441	0.0214	0.0238	0.0697	0.0105	0.0215	0.0366	0.0112	0.0337
		−0.625	0.1288	0.4061	0.0788	0.0179	0.0304	0.0491	0.0171	0.0342	0.0281	0.0207	0.0420
		−0.5	0.0860	0.2561	0.0618	0.0283	0.0380	0.0403	0.0289	0.0418	0.0225	0.0274	0.0454
		−0.25	0.0554	0.0951	0.0418	0.0391	0.0471	0.0345	0.0417	0.0482	0.0212	0.0399	0.0487
		0	0.0509	0.0479	0.0424	0.0437	0.0505	0.0346	0.0481	0.0498	0.0250	0.0494	0.0506
		0.5	0.0450	0.0259	0.0389	0.0490	0.0516	0.0346	0.0501	0.0514	0.0231	0.0533	0.0507
		0.875	0.0458	0.0254	0.0402	0.0552	0.0519	0.0320	0.0498	0.0513	0.0244	0.0517	0.0505
$\gamma\delta_t \neq 0$													
	64												
		−0.875	0.9956	0.5072	0.5506	0.4561	0.1262	0.0000	0.1512	0.0468	0.0000	0.0629	0.0396
		−0.75	0.8600	0.7478	0.1942	0.1822	0.0654	0.0000	0.0554	0.0281	0.0000	0.0323	0.0299
		−0.625	0.5602	0.7689	0.0618	0.0753	0.0398	0.0000	0.0300	0.0276	0.0000	0.0234	0.0317
		−0.5	0.3070	0.5500	0.0185	0.0338	0.0321	0.0000	0.0194	0.0283	0.0000	0.0256	0.0380
		−0.25	0.0766	0.0834	0.0033	0.0198	0.0385	0.0000	0.0200	0.0396	0.0000	0.0319	0.0436
		0	0.0225	0.0031	0.0006	0.0187	0.0484	0.0000	0.0220	0.0514	0.0000	0.0330	0.0502
		0.5	0.0070	0.0000	0.0008	0.0222	0.0627	0.0000	0.0275	0.0551	0.0000	0.0388	0.0507
		0.875	0.0040	0.0000	0.0003	0.0224	0.0654	0.0000	0.0258	0.0582	0.0000	0.0382	0.0515
	128												
		−0.875	0.9990	0.9975	0.8998	0.5264	0.0984	0.2309	0.1693	0.0300	0.0000	0.0508	0.0246
		−0.75	0.8630	0.9998	0.4180	0.1298	0.0325	0.0346	0.0338	0.0158	0.0000	0.0170	0.0230
		−0.625	0.5344	0.9933	0.1568	0.0378	0.0243	0.0071	0.0156	0.0198	0.0000	0.0143	0.0298
		−0.5	0.2779	0.9120	0.0597	0.0164	0.0240	0.0023	0.0102	0.0262	0.0000	0.0136	0.0370
		−0.25	0.0867	0.3190	0.0177	0.0161	0.0380	0.0006	0.0138	0.0407	0.0000	0.0194	0.0464
		0	0.0370	0.0200	0.0111	0.0206	0.0491	0.0011	0.0210	0.0499	0.0000	0.0250	0.0510
		0.5	0.0165	0.0000	0.0058	0.0205	0.0622	0.0007	0.0233	0.0562	0.0000	0.0264	0.0516
		0.875	0.0166	0.0000	0.0059	0.0226	0.0615	0.0008	0.0242	0.0556	0.0000	0.0240	0.0518
	256												
		−0.875	0.9953	1.0000	0.9122	0.3954	0.0446	0.4689	0.1107	0.0110	0.0396	0.0247	0.0113
		−0.75	0.7593	1.0000	0.4037	0.0508	0.0167	0.0934	0.0108	0.0105	0.0039	0.0049	0.0180
		−0.625	0.4238	0.9871	0.1688	0.0180	0.0176	0.0333	0.0070	0.0179	0.0013	0.0077	0.0305
		−0.5	0.2203	0.8749	0.0799	0.0073	0.0229	0.0141	0.0045	0.0263	0.0005	0.0066	0.0382
		−0.25	0.0768	0.3481	0.0298	0.0115	0.0400	0.0088	0.0118	0.0454	0.0005	0.0143	0.0481
		0	0.0423	0.0570	0.0224	0.0178	0.0489	0.0074	0.0191	0.0502	0.0003	0.0215	0.0497
		0.5	0.0300	0.0018	0.0186	0.0226	0.0578	0.0070	0.0228	0.0547	0.0006	0.0226	0.0512
		0.875	0.0298	0.0006	0.0201	0.0241	0.0601	0.0080	0.0248	0.0560	0.0003	0.0232	0.0519

model without additive outliers, higher order variants of the WSR test are generally undersized, and for small sample sizes can effectively be zero if m/T is not small enough. Accordingly, using them is not advised. Alternatively, wavestrapped variants of the WSR test are much more attractive albeit somewhat undersized.

Table 2. Size distortions with additive outliers: $\lambda \neq 0$.

					$m=1$			$m=2$			$m=3$		
T	λ	p	τ^N	τ^{FG}	τ_m^{WSR}	$\tau_m^{WSR\star}$	$\tau_m^{WSR\star\star}$	τ_m^{WSR}	$\tau_m^{WSR\star}$	$\tau_m^{WSR\star\star}$	τ_m^{WSR}	$\tau_m^{WSR\star}$	$\tau_m^{WSR\star\star}$
64													
	5												
		0.05	0.0970	0.1774	0.0433	0.0340	0.0384	0.0134	0.0305	0.0404	0.0004	0.0375	0.0453
		0.2	0.2018	0.3808	0.0875	0.0358	0.0311	0.0213	0.0249	0.0305	0.0006	0.0247	0.0407
		0.4	0.2374	0.4270	0.1022	0.0398	0.0302	0.0257	0.0238	0.0276	0.0008	0.0226	0.0359
		0.6	0.1966	0.3857	0.0809	0.0355	0.0300	0.0203	0.0223	0.0281	0.0004	0.0224	0.0353
		0.8	0.1209	0.2938	0.0448	0.0220	0.0303	0.0088	0.0158	0.0315	0.0007	0.0188	0.0425
		0.95	0.0413	0.1048	0.0174	0.0254	0.0396	0.0058	0.0278	0.0421	0.0000	0.0335	0.0462
	10												
		0.05	0.2190	0.3888	0.1025	0.0424	0.0310	0.0279	0.0272	0.0293	0.0006	0.0284	0.0393
		0.2	0.4075	0.5924	0.2098	0.0697	0.0327	0.0557	0.0292	0.0238	0.0016	0.0215	0.0299
		0.4	0.3961	0.5450	0.1972	0.0671	0.0310	0.0496	0.0276	0.0212	0.0010	0.0141	0.0229
		0.6	0.2860	0.4194	0.1207	0.0374	0.0242	0.0238	0.0170	0.0204	0.0004	0.0098	0.0254
		0.8	0.1115	0.2484	0.0361	0.0187	0.0254	0.0049	0.0088	0.0234	0.0000	0.0088	0.0345
		0.95	0.0123	0.0799	0.0026	0.0118	0.0393	0.0002	0.0156	0.0419	0.0000	0.0260	0.0481
128													
	5												
		0.05	0.0905	0.2091	0.0568	0.0303	0.0393	0.0281	0.0313	0.0423	0.0097	0.0339	0.0474
		0.2	0.1582	0.4137	0.0856	0.0229	0.0272	0.0357	0.0171	0.0301	0.0105	0.0201	0.0413
		0.4	0.2009	0.4822	0.1109	0.0272	0.0258	0.0471	0.0164	0.0264	0.0122	0.0169	0.0367
		0.6	0.1811	0.4651	0.0934	0.0228	0.0259	0.0354	0.0140	0.0285	0.0109	0.0149	0.0358
		0.8	0.1235	0.3555	0.0636	0.0211	0.0308	0.0283	0.0163	0.0341	0.0080	0.0198	0.0421
		0.95	0.0499	0.1505	0.0295	0.0222	0.0424	0.0165	0.0266	0.0452	0.0063	0.0299	0.0489
	10												
		0.05	0.1874	0.4403	0.1003	0.0246	0.0280	0.0412	0.0175	0.0291	0.0112	0.0197	0.0405
		0.2	0.3510	0.6739	0.2061	0.0470	0.0256	0.0816	0.0201	0.0219	0.0218	0.0136	0.0256
		0.4	0.3788	0.6888	0.2220	0.0475	0.0244	0.0845	0.0169	0.0165	0.0193	0.0107	0.0250
		0.6	0.3055	0.5896	0.1621	0.0367	0.0248	0.0597	0.0144	0.0203	0.0122	0.0087	0.0266
		0.8	0.1549	0.4185	0.0711	0.0138	0.0226	0.0209	0.0047	0.0232	0.0029	0.0049	0.0318
		0.95	0.0374	0.1895	0.0139	0.0085	0.0345	0.0049	0.0073	0.0372	0.0011	0.0110	0.0436
256													
	5												
		0.05	0.0707	0.1822	0.0510	0.0301	0.0418	0.0365	0.0314	0.0448	0.0228	0.0361	0.0476
		0.2	0.1222	0.3898	0.0793	0.0208	0.0304	0.0462	0.0170	0.0350	0.0257	0.0209	0.0414
		0.4	0.1440	0.4664	0.0900	0.0222	0.0297	0.0503	0.0165	0.0315	0.0276	0.0182	0.0401
		0.6	0.1453	0.4464	0.0914	0.0187	0.0266	0.0504	0.0146	0.0288	0.0255	0.0159	0.0382
		0.8	0.1119	0.3617	0.0699	0.0192	0.0323	0.0407	0.0155	0.0352	0.0232	0.0195	0.0419
		0.95	0.0570	0.1553	0.0423	0.0267	0.0446	0.0306	0.0283	0.0463	0.0180	0.0324	0.0486
	10												
		0.05	0.1382	0.4262	0.0883	0.0213	0.0293	0.0514	0.0167	0.0319	0.0289	0.0179	0.0410
		0.2	0.2750	0.6637	0.1694	0.0226	0.0203	0.0820	0.0104	0.0191	0.0367	0.0078	0.0298
		0.4	0.3191	0.7228	0.1964	0.0316	0.0185	0.0961	0.0110	0.0172	0.0375	0.0070	0.0240
		0.6	0.2736	0.6620	0.1639	0.0232	0.0194	0.0769	0.0088	0.0172	0.0333	0.0060	0.0241
		0.8	0.1739	0.5175	0.0940	0.0130	0.0243	0.0419	0.0054	0.0261	0.0182	0.0054	0.0349
		0.95	0.0545	0.2603	0.0310	0.0100	0.0331	0.0169	0.0093	0.0373	0.0099	0.0124	0.0435

Consider size adjusted power next. Table 3 and 4, respectively, illustrate the case of classical and detrended tests for the model without additive outliers. Although the FG test dominates for larger sample sizes when $\psi_1 \geq -0.25$, it is otherwise underpowered with power critically failing (going to zero) for all sample sizes when $\psi_1 < -0.5$. This is particularly evident for detrended statistics. In contrast, the NVR test dominates when $\psi_1 \geq -0.25$, although only marginally in relation to the WSR test. The leverage ensues from the larger effective sample size in the NVR test available for power computation. Specifically, the downsampling mechanism generating the DWT effectively reduces the sample size exploitable in

Table 3. Size adjusted power for classical statistics without additive outliers: $\boldsymbol{\gamma}\delta_t = \mathbf{0}$, $\lambda = 0$.

		$T=64$					$T=128$					$T=256$				
θ	ρ	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$
$\forall\theta$																
	1	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
−0.875																
	0.99	0.0644	0.0554	0.0665	0.0702	0.0678	0.0886	0.0721	0.0875	0.0870	0.0873	0.1379	0.1017	0.1385	0.1382	0.1279
	0.98	0.0881	0.0566	0.0872	0.0898	0.0870	0.1410	0.0823	0.1434	0.1399	0.1341	0.3050	0.1202	0.3075	0.2985	0.2698
	0.97	0.1111	0.0555	0.1117	0.1091	0.1059	0.2140	0.0708	0.2127	0.2045	0.1890	0.5221	0.0884	0.5233	0.4974	0.4381
	0.96	0.1410	0.0509	0.1432	0.1475	0.1421	0.3138	0.0599	0.3114	0.3004	0.2701	0.7363	0.0524	0.7314	0.6971	0.6104
	0.95	0.1787	0.0467	0.1782	0.1790	0.1633	0.4249	0.0392	0.4174	0.3979	0.3596	0.8807	0.0254	0.8731	0.8314	0.7376
	0.9	0.4370	0.0148	0.4159	0.3937	0.3249	0.8963	0.0030	0.8801	0.8290	0.7137	0.9999	0.0001	1.0000	0.9987	0.9835
	0.8	0.8891	0.0007	0.8570	0.7868	0.6042	1.0000	0.0000	0.9998	0.9965	0.9575	1.0000	0.0000	1.0000	1.0000	0.9999
	0.5	1.0000	0.0000	0.9999	0.9927	0.8557	1.0000	0.0000	1.0000	1.0000	0.9986	1.0000	0.0000	1.0000	1.0000	1.0000
−0.75																
	0.99	0.0648	0.0644	0.0644	0.0643	0.0606	0.0876	0.0819	0.0887	0.0880	0.0837	0.1357	0.1407	0.1309	0.1275	0.1190
	0.98	0.0914	0.0756	0.0929	0.0899	0.0831	0.1430	0.1213	0.1409	0.1409	0.1310	0.2928	0.2973	0.2792	0.2610	0.2382
	0.97	0.1096	0.0868	0.1097	0.1034	0.0913	0.2136	0.1608	0.2085	0.2073	0.1853	0.4931	0.4615	0.4616	0.4248	0.3746
	0.96	0.1419	0.1061	0.1418	0.1310	0.1180	0.3048	0.2063	0.2957	0.2843	0.2461	0.7001	0.6010	0.6601	0.5995	0.5249
	0.95	0.1817	0.1035	0.1727	0.1626	0.1351	0.4007	0.2154	0.3886	0.3615	0.3089	0.8439	0.6740	0.8089	0.7408	0.6431
	0.9	0.4303	0.1083	0.4047	0.3560	0.2680	0.8689	0.1827	0.8359	0.7624	0.6239	0.9996	0.5750	0.9988	0.9870	0.9388
	0.8	0.8996	0.0347	0.8510	0.7345	0.5200	1.0000	0.0217	0.9992	0.9859	0.8929	1.0000	0.0588	1.0000	1.0000	0.9970
	0.5	1.0000	0.0002	0.9999	0.9839	0.7886	1.0000	0.0000	1.0000	1.0000	0.9901	1.0000	0.0000	1.0000	1.0000	1.0000
−0.625																
	0.99	0.0618	0.0621	0.0610	0.0618	0.0645	0.0825	0.0899	0.0836	0.0828	0.0807	0.1210	0.1398	0.1210	0.1138	0.1081
	0.98	0.0853	0.0833	0.0870	0.0829	0.0805	0.1369	0.1423	0.1378	0.1328	0.1282	0.2627	0.3015	0.2556	0.2388	0.2231
	0.97	0.1085	0.0967	0.1100	0.1045	0.1032	0.2032	0.2147	0.2010	0.1899	0.1739	0.4430	0.5083	0.4244	0.3886	0.3552
	0.96	0.1370	0.1115	0.1414	0.1341	0.1231	0.2910	0.2839	0.2826	0.2636	0.2349	0.6308	0.7004	0.5940	0.5431	0.4846
	0.95	0.1658	0.1317	0.1649	0.1511	0.1429	0.3833	0.3660	0.3705	0.3339	0.2913	0.7845	0.8441	0.7504	0.6755	0.6009
	0.9	0.3965	0.2190	0.3783	0.3225	0.2698	0.8316	0.6277	0.7898	0.6969	0.5784	0.9957	0.9913	0.9896	0.9648	0.9050
	0.8	0.8678	0.1932	0.8162	0.6780	0.4962	0.9992	0.5365	0.9946	0.9663	0.8511	1.0000	0.9852	1.0000	0.9997	0.9915
	0.5	1.0000	0.0198	0.9996	0.9697	0.7419	1.0000	0.0374	1.0000	0.9998	0.9741	1.0000	0.3320	1.0000	1.0000	1.0000
−0.5																
	0.99	0.0676	0.0620	0.0668	0.0649	0.0626	0.0781	0.0853	0.0780	0.0758	0.0745	0.1287	0.1495	0.1270	0.1244	0.1218
	0.98	0.0897	0.0883	0.0899	0.0883	0.0818	0.1269	0.1411	0.1278	0.1224	0.1156	0.2652	0.3287	0.2558	0.2439	0.2352
	0.97	0.1142	0.1096	0.1121	0.1081	0.0972	0.1881	0.2036	0.1851	0.1705	0.1620	0.4355	0.5451	0.4193	0.3942	0.3649
	0.96	0.1496	0.1357	0.1478	0.1372	0.1242	0.2708	0.2869	0.2620	0.2406	0.2212	0.6048	0.7367	0.5781	0.5392	0.4935
	0.95	0.1828	0.1557	0.1720	0.1587	0.1417	0.3535	0.3790	0.3424	0.3092	0.2781	0.7545	0.8755	0.7238	0.6685	0.6112
	0.9	0.4214	0.3074	0.3875	0.3297	0.2675	0.7762	0.7637	0.7278	0.6390	0.5385	0.9919	0.9988	0.9834	0.9557	0.8994
	0.8	0.8530	0.4520	0.7784	0.6444	0.4593	0.9960	0.9039	0.9851	0.9397	0.8162	1.0000	1.0000	0.9999	0.9996	0.9902
	0.5	1.0000	0.2668	0.9984	0.9463	0.6969	1.0000	0.6825	1.0000	0.9993	0.9545	1.0000	0.9996	1.0000	1.0000	0.9998
−0.25																
	0.99	0.0614	0.0610	0.0604	0.0611	0.0607	0.0808	0.0873	0.0813	0.0784	0.0775	0.1220	0.1340	0.1244	0.1217	0.1158
	0.98	0.0804	0.0831	0.0751	0.0769	0.0733	0.1233	0.1340	0.1243	0.1198	0.1171	0.2537	0.3003	0.2535	0.2447	0.2295
	0.97	0.1016	0.1051	0.0931	0.0935	0.0949	0.1805	0.2076	0.1806	0.1717	0.1647	0.4008	0.5122	0.3999	0.3803	0.3553
	0.96	0.1270	0.1331	0.1214	0.1195	0.1138	0.2495	0.2932	0.2488	0.2291	0.2144	0.5620	0.7152	0.5577	0.5282	0.4851
	0.95	0.1456	0.1578	0.1365	0.1357	0.1323	0.3264	0.3939	0.3218	0.2951	0.2717	0.6891	0.8557	0.6809	0.6473	0.5904
	0.9	0.3350	0.3460	0.3082	0.2801	0.2489	0.7120	0.8313	0.6854	0.6165	0.5373	0.9682	0.9987	0.9630	0.9382	0.8854
	0.8	0.7487	0.6632	0.6747	0.5706	0.4326	0.9764	0.9885	0.9619	0.9063	0.7885	1.0000	1.0000	1.0000	0.9991	0.9836
	0.5	0.9989	0.8303	0.9858	0.9024	0.6424	1.0000	0.9989	0.9999	0.9975	0.9379	1.0000	1.0000	1.0000	1.0000	0.9988
0																
	0.99	0.0642	0.0658	0.0630	0.0625	0.0623	0.0812	0.0880	0.0811	0.0822	0.0785	0.1205	0.1313	0.1198	0.1179	0.1167
	0.98	0.0861	0.0873	0.0850	0.0821	0.0816	0.1296	0.1514	0.1297	0.1307	0.1249	0.2433	0.2877	0.2429	0.2399	0.2345
	0.97	0.1052	0.1103	0.1028	0.1019	0.0978	0.1856	0.2166	0.1846	0.1822	0.1718	0.3847	0.4867	0.3854	0.3747	0.3542
	0.96	0.1276	0.1358	0.1252	0.1237	0.1179	0.2561	0.3099	0.2530	0.2438	0.2272	0.5381	0.6916	0.5390	0.5208	0.4851
	0.95	0.1521	0.1568	0.1481	0.1418	0.1344	0.3259	0.4071	0.3229	0.3123	0.2851	0.6638	0.8331	0.6653	0.6391	0.5946
	0.9	0.3453	0.3540	0.3310	0.3018	0.2586	0.6813	0.8222	0.6698	0.6266	0.5414	0.9545	0.9977	0.9531	0.9311	0.8832
	0.8	0.6991	0.6915	0.6571	0.5702	0.4339	0.9651	0.9906	0.9560	0.9141	0.8009	0.9997	1.0000	0.9995	0.9978	0.9858
	0.5	0.9955	0.9448	0.9773	0.8954	0.6540	1.0000	1.0000	1.0000	0.9978	0.9330	1.0000	1.0000	1.0000	1.0000	0.9987

Table 3. Size adjusted power for classical statistics without additive outliers: $\boldsymbol{\gamma}\boldsymbol{\delta}_t = \mathbf{0}$, $\lambda = 0$.

		$T = 64$					$T = 128$					$T = 256$				
θ	ρ	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$
0.5																
	0.99	0.0643	0.0636	0.0640	0.0641	0.0623	0.0807	0.0879	0.0832	0.0812	0.0815	0.1209	0.1403	0.1232	0.1220	0.1226
	0.98	0.0837	0.0820	0.0827	0.0810	0.0810	0.1192	0.1374	0.1219	0.1181	0.1168	0.2415	0.3046	0.2496	0.2458	0.2409
	0.97	0.0990	0.0990	0.0980	0.0964	0.0937	0.1672	0.1928	0.1725	0.1647	0.1588	0.3752	0.4865	0.3878	0.3764	0.3625
	0.96	0.1180	0.1263	0.1171	0.1128	0.1062	0.2238	0.2715	0.2311	0.2213	0.2111	0.5192	0.6664	0.5375	0.5186	0.4913
	0.95	0.1461	0.1482	0.1471	0.1382	0.1273	0.2971	0.3543	0.3069	0.2884	0.2687	0.6431	0.8116	0.6621	0.6362	0.6004
	0.9	0.3055	0.2991	0.3022	0.2777	0.2379	0.6247	0.7257	0.6334	0.5895	0.5213	0.9394	0.9953	0.9485	0.9265	0.8829
	0.8	0.6435	0.5848	0.6297	0.5491	0.4288	0.9433	0.9675	0.9431	0.8982	0.7932	0.9989	1.0000	0.9992	0.9973	0.9832
	0.5	0.9775	0.9006	0.9639	0.8690	0.6216	1.0000	0.9997	0.9998	0.9940	0.9287	1.0000	1.0000	1.0000	1.0000	0.9988
0.875																
	0.99	0.0660	0.0670	0.0682	0.0642	0.0629	0.0859	0.0895	0.0855	0.0873	0.0846	0.1123	0.1232	0.1155	0.1141	0.1098
	0.98	0.0795	0.0784	0.0804	0.0783	0.0732	0.1249	0.1362	0.1261	0.1265	0.1194	0.2140	0.2602	0.2218	0.2179	0.2101
	0.97	0.0986	0.1001	0.1021	0.0982	0.0928	0.1845	0.2030	0.1845	0.1845	0.1733	0.3487	0.4496	0.3630	0.3538	0.3361
	0.96	0.1197	0.1258	0.1219	0.1161	0.1096	0.2562	0.2759	0.2582	0.2517	0.2326	0.4923	0.6229	0.5105	0.4939	0.4607
	0.95	0.1445	0.1467	0.1478	0.1406	0.1267	0.3178	0.3513	0.3214	0.3128	0.2836	0.6147	0.7631	0.6376	0.6120	0.5678
	0.9	0.2990	0.2836	0.3033	0.2763	0.2370	0.6553	0.7166	0.6624	0.6266	0.5520	0.9271	0.9899	0.9382	0.9156	0.8640
	0.8	0.6328	0.5456	0.6277	0.5487	0.4155	0.9416	0.9593	0.9417	0.9042	0.7988	0.9988	1.0000	0.9989	0.9969	0.9824
	0.5	0.9742	0.8611	0.9620	0.8695	0.6182	0.9997	0.9994	0.9997	0.9949	0.9299	1.0000	1.0000	1.0000	1.0000	0.9983

Table 4. Size adjusted power for detrended statistics without additive outliers: $\boldsymbol{\gamma}\boldsymbol{\delta}_t \neq \mathbf{0}$, $\lambda = 0$.

		$T = 64$					$T = 128$					$T = 256$				
θ	ρ	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$
$\forall\theta$																
	1	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
−0.875																
	0.99	0.0559	0.0462	0.0548	0.0576	0.0533	0.0501	0.0576	0.0540	0.0537	0.0530	0.0612	0.0759	0.0619	0.0605	0.0626
	0.98	0.0534	0.0467	0.0564	0.0568	0.0525	0.0552	0.0489	0.0626	0.0633	0.0579	0.0947	0.0667	0.0919	0.0909	0.0928
	0.97	0.0588	0.0403	0.0571	0.0606	0.0537	0.0711	0.0374	0.0799	0.0781	0.0735	0.1517	0.0432	0.1519	0.1478	0.1449
	0.96	0.0615	0.0365	0.0597	0.0620	0.0567	0.0831	0.0277	0.0945	0.0943	0.0847	0.2318	0.0205	0.2276	0.2175	0.2076
	0.95	0.0663	0.0303	0.0701	0.0693	0.0680	0.1071	0.0150	0.1136	0.1168	0.1035	0.3444	0.0071	0.3301	0.3152	0.2902
	0.9	0.1032	0.0059	0.1051	0.1019	0.0819	0.2851	0.0003	0.3037	0.2903	0.2296	0.8871	0.0000	0.8746	0.8310	0.7266
	0.8	0.2775	0.0000	0.2637	0.2157	0.1397	0.8227	0.0000	0.8242	0.7475	0.5317	1.0000	0.0000	1.0000	0.9996	0.9859
	0.5	0.9527	0.0000	0.8866	0.6137	0.2217	1.0000	0.0000	1.0000	0.9959	0.8314	1.0000	0.0000	1.0000	1.0000	1.0000
−0.75																
	0.99	0.0470	0.0535	0.0457	0.0497	0.0472	0.0542	0.0821	0.0520	0.0519	0.0530	0.0664	0.1267	0.0631	0.0647	0.0635
	0.98	0.0496	0.0624	0.0529	0.0581	0.0506	0.0615	0.1065	0.0588	0.0589	0.0560	0.1050	0.2371	0.0983	0.0992	0.0941
	0.97	0.0495	0.0673	0.0508	0.0541	0.0468	0.0742	0.1230	0.0734	0.0705	0.0694	0.1694	0.3488	0.1602	0.1515	0.1373
	0.96	0.0611	0.0721	0.0622	0.0640	0.0538	0.0933	0.1358	0.0906	0.0874	0.0843	0.2519	0.4306	0.2362	0.2211	0.1941
	0.95	0.0652	0.0742	0.0645	0.0682	0.0608	0.1265	0.1405	0.1187	0.1122	0.1016	0.3585	0.4716	0.3382	0.3129	0.2681
	0.9	0.1113	0.0496	0.1085	0.1120	0.0840	0.3543	0.0837	0.3308	0.2822	0.2313	0.9080	0.3026	0.8697	0.8026	0.6674
	0.8	0.3063	0.0084	0.2838	0.2537	0.1511	0.9191	0.0058	0.8737	0.7468	0.5189	1.0000	0.0145	1.0000	0.9993	0.9623
	0.5	0.9839	0.0000	0.9287	0.7029	0.2806	1.0000	0.0000	1.0000	0.9977	0.8456	1.0000	0.0000	1.0000	1.0000	0.9998
−0.625																
	0.99	0.0487	0.0592	0.0453	0.0455	0.0534	0.0529	0.0770	0.0491	0.0503	0.0490	0.0571	0.1187	0.0587	0.0580	0.0597
	0.98	0.0501	0.0674	0.0453	0.0478	0.0508	0.0640	0.1186	0.0624	0.0615	0.0610	0.0881	0.2517	0.0887	0.0865	0.0845
	0.97	0.0561	0.0808	0.0517	0.0549	0.0587	0.0814	0.1616	0.0791	0.0790	0.0805	0.1446	0.4234	0.1433	0.1363	0.1324
	0.96	0.0625	0.0921	0.0572	0.0604	0.0617	0.0963	0.2048	0.0900	0.0850	0.0804	0.2222	0.5979	0.2166	0.1929	0.1785
	0.95	0.0673	0.1016	0.0628	0.0632	0.0684	0.1303	0.2555	0.1204	0.1146	0.1064	0.3294	0.7553	0.3158	0.2812	0.2515
	0.9	0.1199	0.1245	0.1081	0.1040	0.0955	0.3577	0.3991	0.3237	0.2763	0.2169	0.8806	0.9443	0.8427	0.7490	0.6137
	0.8	0.3356	0.0835	0.2858	0.2337	0.1648	0.9093	0.2748	0.8497	0.7213	0.4934	1.0000	0.8477	0.9999	0.9954	0.9358
	0.5	0.9907	0.0031	0.9360	0.6945	0.3080	1.0000	0.0069	1.0000	0.9947	0.8133	1.0000	0.0754	1.0000	1.0000	0.9981

Table 4. Size adjusted power for detrended statistics without additive outliers: $\gamma\delta_t \neq \mathbf{0}$, $\lambda = 0$.

		$T = 64$					$T = 128$					$T = 256$				
θ	ρ	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$
−0.5																
	1	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
	0.99	0.0499	0.0724	0.0504	0.0475	0.0502	0.0512	0.0748	0.0507	0.0516	0.0527	0.0686	0.1353	0.0669	0.0638	0.0621
	0.98	0.0497	0.0816	0.0508	0.0527	0.0571	0.0646	0.1272	0.0631	0.0624	0.0619	0.1021	0.2815	0.1002	0.0930	0.0906
	0.97	0.0542	0.0996	0.0531	0.0565	0.0572	0.0842	0.1767	0.0792	0.0778	0.0765	0.1623	0.4741	0.1536	0.1396	0.1271
	0.96	0.0631	0.1122	0.0667	0.0638	0.0618	0.1024	0.2454	0.0970	0.0935	0.0868	0.2438	0.6619	0.2325	0.2075	0.1848
	0.95	0.0714	0.1395	0.0719	0.0678	0.0642	0.1292	0.3177	0.1212	0.1116	0.1016	0.3423	0.8157	0.3152	0.2790	0.2465
	0.9	0.1243	0.2076	0.1204	0.1130	0.0962	0.3549	0.6254	0.3195	0.2722	0.2157	0.8756	0.9958	0.8275	0.7322	0.6081
	0.8	0.3528	0.2620	0.3166	0.2399	0.1668	0.8973	0.7524	0.8279	0.6899	0.4782	1.0000	0.9982	0.9994	0.9907	0.9227
	0.5	0.9919	0.1094	0.9359	0.6725	0.2883	1.0000	0.3940	0.9999	0.9907	0.7852	1.0000	0.9654	1.0000	1.0000	0.9962
−0.25																
	0.99	0.0520	0.0592	0.0515	0.0550	0.0494	0.0503	0.0867	0.0510	0.0515	0.0513	0.0689	0.1301	0.0688	0.0647	0.0635
	0.98	0.0533	0.0817	0.0541	0.0530	0.0501	0.0650	0.1309	0.0640	0.0648	0.0610	0.1008	0.2718	0.0969	0.0924	0.0880
	0.97	0.0599	0.0975	0.0566	0.0578	0.0558	0.0769	0.2008	0.0754	0.0745	0.0695	0.1461	0.4648	0.1396	0.1278	0.1212
	0.96	0.0673	0.1209	0.0643	0.0634	0.0630	0.0972	0.2791	0.0938	0.0899	0.0834	0.2199	0.6624	0.2088	0.1910	0.1771
	0.95	0.0699	0.1478	0.0720	0.0728	0.0649	0.1209	0.3694	0.1155	0.1105	0.1006	0.3127	0.8138	0.2935	0.2661	0.2391
	0.9	0.1247	0.2646	0.1158	0.1090	0.0928	0.3254	0.7708	0.2949	0.2658	0.2143	0.8187	0.9977	0.7748	0.6993	0.5959
	0.8	0.3392	0.4981	0.2898	0.2332	0.1585	0.8449	0.9720	0.7664	0.6342	0.4466	0.9994	1.0000	0.9969	0.9821	0.9083
	0.5	0.9805	0.6390	0.9002	0.6460	0.2837	1.0000	0.9863	0.9996	0.9786	0.7491	1.0000	1.0000	1.0000	1.0000	0.9920
0																
	0.99	0.0563	0.0572	0.0550	0.0534	0.0521	0.0558	0.0786	0.0545	0.0527	0.0531	0.0609	0.0991	0.0607	0.0599	0.0613
	0.98	0.0561	0.0740	0.0551	0.0539	0.0536	0.0677	0.1139	0.0673	0.0651	0.0666	0.0977	0.1956	0.0973	0.0965	0.0920
	0.97	0.0571	0.0927	0.0555	0.0534	0.0552	0.0795	0.1693	0.0777	0.0768	0.0761	0.1456	0.3309	0.1443	0.1381	0.1311
	0.96	0.0630	0.1063	0.0618	0.0572	0.0601	0.1016	0.2366	0.0979	0.0927	0.0901	0.2140	0.4918	0.2114	0.2036	0.1876
	0.95	0.0721	0.1198	0.0690	0.0654	0.0665	0.1224	0.3020	0.1168	0.1120	0.1070	0.3056	0.6588	0.2999	0.2824	0.2534
	0.9	0.1230	0.2418	0.1172	0.1080	0.0922	0.3139	0.7027	0.2889	0.2550	0.2185	0.7753	0.9811	0.7530	0.6907	0.5894
	0.8	0.3226	0.4948	0.2786	0.2253	0.1576	0.8121	0.9653	0.7494	0.6347	0.4648	0.9977	0.9999	0.9948	0.9792	0.9102
	0.5	0.9599	0.7965	0.8623	0.6055	0.2790	0.9998	0.9966	0.9992	0.9722	0.7495	1.0000	1.0000	1.0000	1.0000	0.9904
0.5																
	0.99	0.0466	0.0403	0.0457	0.0469	0.0496	0.0636	0.0510	0.0634	0.0640	0.0634	0.0642	0.0715	0.0651	0.0644	0.0640
	0.98	0.0549	0.0392	0.0553	0.0560	0.0526	0.0723	0.0653	0.0725	0.0709	0.0670	0.0927	0.1307	0.0933	0.0952	0.0920
	0.97	0.0557	0.0461	0.0553	0.0552	0.0577	0.0831	0.0875	0.0822	0.0795	0.0768	0.1417	0.2013	0.1414	0.1397	0.1333
	0.96	0.0608	0.0433	0.0598	0.0607	0.0604	0.1052	0.1161	0.1049	0.1040	0.0993	0.2083	0.3072	0.2074	0.1985	0.1854
	0.95	0.0673	0.0506	0.0674	0.0673	0.0661	0.1156	0.1432	0.1127	0.1089	0.1062	0.2770	0.4106	0.2769	0.2614	0.2389
	0.9	0.1075	0.0798	0.1056	0.1019	0.0901	0.2981	0.3650	0.2898	0.2665	0.2247	0.7468	0.8587	0.7404	0.6891	0.5947
	0.8	0.2773	0.2036	0.2596	0.2239	0.1587	0.7550	0.7728	0.7246	0.6345	0.4680	0.9937	0.9952	0.9931	0.9804	0.9077
	0.5	0.8885	0.5745	0.8097	0.5988	0.2751	0.9996	0.9751	0.9988	0.9717	0.7464	1.0000	0.9998	1.0000	1.0000	0.9906
0.875																
	0.99	0.0458	0.0363	0.0447	0.0435	0.0465	0.0549	0.0467	0.0567	0.0563	0.0585	0.0655	0.0662	0.0640	0.0642	0.0634
	0.98	0.0535	0.0377	0.0517	0.0511	0.0533	0.0600	0.0561	0.0605	0.0610	0.0657	0.0988	0.1110	0.0974	0.0958	0.0952
	0.97	0.0572	0.0400	0.0560	0.0561	0.0561	0.0776	0.0715	0.0772	0.0752	0.0799	0.1465	0.1803	0.1432	0.1403	0.1340
	0.96	0.0608	0.0381	0.0600	0.0580	0.0598	0.0930	0.1008	0.0937	0.0911	0.0925	0.2127	0.2633	0.2089	0.2018	0.1860
	0.95	0.0650	0.0414	0.0634	0.0664	0.0650	0.1139	0.1174	0.1148	0.1133	0.1113	0.2792	0.3627	0.2748	0.2636	0.2388
	0.9	0.1108	0.0621	0.1078	0.0989	0.0942	0.2816	0.2944	0.2769	0.2559	0.2249	0.7420	0.8087	0.7320	0.6837	0.5832
	0.8	0.2809	0.1490	0.2643	0.2232	0.1605	0.7435	0.6707	0.7238	0.6280	0.4773	0.9946	0.9888	0.9929	0.9808	0.9094
	0.5	0.8742	0.4987	0.8045	0.5799	0.2791	0.9993	0.9581	0.9976	0.9681	0.7537	1.0000	1.0000	1.0000	0.9999	0.9912

power computations of WSR tests to $2^{-m}T$. Accordingly, the higher order WSR tests are generally underpowered relative to their lower order counterparts.

Similar conclusions also hold for the model with additive outliers in Table 5. In particular, the applicability of the FG test is generally only limited to higher sample sizes and outlier frequencies $p < 0.1$, whereas the NVR and WSR are decently sized in all cases except for very large λ and p, although for reasons mentioned earlier, the NVR test performs marginally better. In general, all three tests are highly unreliable for large λ and p, particularly when sample sizes are small. This of course is not very surprising

Table 5. Size adjusted power without a trend and with additive outliers: $\boldsymbol{\gamma}\boldsymbol{\delta}_t = \mathbf{0}$, $\lambda \neq 0$.

			$T = 64$					$T = 128$					$T = 256$				
λ	p	ρ	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$
$\forall\lambda$	$\forall p$																
		1	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500	0.0500
5																	
	0.05																
		0.99	0.0667	0.0702	0.0626	0.0591	0.0623	0.0840	0.0829	0.0822	0.0806	0.0807	0.1364	0.1412	0.1339	0.1312	0.1305
		0.98	0.0871	0.0827	0.0880	0.0793	0.0810	0.1310	0.1315	0.1308	0.1240	0.1238	0.2765	0.3057	0.2703	0.2621	0.2533
		0.97	0.1085	0.1005	0.1052	0.0955	0.0952	0.1965	0.1962	0.1891	0.1790	0.1705	0.4358	0.4995	0.4216	0.4000	0.3749
		0.96	0.1227	0.1148	0.1190	0.1097	0.1080	0.2558	0.2564	0.2449	0.2276	0.2120	0.5922	0.6898	0.5719	0.5392	0.4980
		0.95	0.1652	0.1468	0.1592	0.1431	0.1406	0.3450	0.3383	0.3276	0.3036	0.2772	0.7243	0.8288	0.7059	0.6613	0.6086
		0.9	0.3447	0.2544	0.3232	0.2818	0.2474	0.7170	0.6819	0.6823	0.6127	0.5215	0.9744	0.9955	0.9621	0.9334	0.8760
		0.8	0.7140	0.4058	0.6611	0.5379	0.4105	0.9729	0.9034	0.9468	0.8792	0.7516	0.9999	1.0000	0.9997	0.9947	0.9642
		0.5	0.9921	0.4523	0.9618	0.8151	0.5290	1.0000	0.9282	0.9997	0.9780	0.7832	1.0000	1.0000	1.0000	0.9998	0.9721
	0.1																
		0.99	0.0692	0.0705	0.0680	0.0718	0.0708	0.0833	0.0903	0.0799	0.0734	0.0749	0.1253	0.1341	0.1261	0.1216	0.1210
		0.98	0.0856	0.0804	0.0832	0.0860	0.0844	0.1390	0.1487	0.1341	0.1262	0.1237	0.2517	0.2857	0.2504	0.2342	0.2275
		0.97	0.1049	0.1004	0.1058	0.1040	0.1037	0.1905	0.2108	0.1824	0.1684	0.1595	0.4105	0.4742	0.4043	0.3727	0.3524
		0.96	0.1239	0.1166	0.1242	0.1254	0.1187	0.2658	0.2755	0.2536	0.2382	0.2237	0.5712	0.6482	0.5550	0.5085	0.4715
		0.95	0.1520	0.1338	0.1486	0.1451	0.1381	0.3482	0.3582	0.3279	0.3012	0.2768	0.6982	0.7887	0.6778	0.6169	0.5663
		0.9	0.3184	0.2223	0.3114	0.2834	0.2429	0.6998	0.6404	0.6509	0.5655	0.4821	0.9582	0.9819	0.9367	0.8825	0.8020
		0.8	0.6462	0.2720	0.5960	0.4973	0.3605	0.9559	0.7560	0.9082	0.7876	0.6165	0.9995	0.9984	0.9979	0.9761	0.8698
		0.5	0.9921	0.4523	0.9618	0.8151	0.5290	1.0000	0.9282	0.9997	0.9780	0.7832	1.0000	1.0000	1.0000	0.9998	0.9721
	0.2																
		0.99	0.0568	0.0667	0.0576	0.0583	0.0608	0.0884	0.0863	0.0867	0.0822	0.0807	0.1393	0.1433	0.1363	0.1330	0.1286
		0.98	0.0772	0.0810	0.0753	0.0773	0.0775	0.1356	0.1294	0.1314	0.1299	0.1242	0.2754	0.2866	0.2693	0.2565	0.2360
		0.97	0.0889	0.1006	0.0899	0.0910	0.0902	0.1941	0.1855	0.1853	0.1774	0.1693	0.4279	0.4537	0.4089	0.3826	0.3440
		0.96	0.1176	0.1049	0.1162	0.1152	0.1081	0.2611	0.2318	0.2458	0.2318	0.2166	0.5647	0.6022	0.5351	0.4890	0.4373
		0.95	0.1319	0.1193	0.1313	0.1294	0.1226	0.3255	0.2770	0.3071	0.2767	0.2488	0.6647	0.7012	0.6279	0.5720	0.4998
		0.9	0.2480	0.1523	0.2409	0.2230	0.1877	0.5827	0.3827	0.5314	0.4491	0.3591	0.9004	0.8855	0.8502	0.7329	0.5938
		0.8	0.3850	0.1139	0.3438	0.2690	0.1758	0.8106	0.2594	0.6920	0.4852	0.2888	0.9943	0.8529	0.9659	0.7912	0.4640
		0.5	0.9483	0.1747	0.8483	0.5809	0.2891	1.0000	0.5546	0.9951	0.8619	0.4428	1.0000	0.9950	1.0000	0.9937	0.7244
	0.3																
		0.99	0.0640	0.0549	0.0620	0.0618	0.0623	0.0844	0.0783	0.0841	0.0811	0.0827	0.1299	0.1362	0.1289	0.1243	0.1216
		0.98	0.0792	0.0688	0.0771	0.0753	0.0727	0.1281	0.1123	0.1316	0.1265	0.1208	0.2433	0.2527	0.2376	0.2162	0.2025
		0.97	0.1008	0.0796	0.0965	0.0958	0.0886	0.1720	0.1430	0.1728	0.1606	0.1514	0.3722	0.3869	0.3545	0.3216	0.2928
		0.96	0.1104	0.0864	0.1064	0.1032	0.0973	0.2221	0.1748	0.2182	0.2005	0.1832	0.4654	0.4812	0.4383	0.3883	0.3411
		0.95	0.1310	0.0936	0.1275	0.1238	0.1094	0.2656	0.1976	0.2599	0.2328	0.2053	0.5472	0.5500	0.5075	0.4363	0.3792
		0.9	0.1952	0.0941	0.1854	0.1654	0.1333	0.3867	0.1814	0.3477	0.2725	0.2031	0.7264	0.5837	0.6317	0.4544	0.3077
		0.8	0.1953	0.0327	0.1681	0.1193	0.0682	0.4562	0.0503	0.3478	0.1783	0.0723	0.8996	0.2909	0.7193	0.3264	0.0888
		0.5	0.5330	0.0240	0.3671	0.1565	0.0416	0.9668	0.0424	0.7735	0.2717	0.0401	1.0000	0.3872	0.9974	0.6144	0.0541
10																	
	0.05																
		0.99	0.0607	0.0586	0.0607	0.0588	0.0617	0.0793	0.0879	0.0781	0.0787	0.0787	0.1377	0.1400	0.1339	0.1295	0.1250
		0.98	0.0786	0.0722	0.0804	0.0790	0.0798	0.1245	0.1228	0.1238	0.1234	0.1229	0.2716	0.2780	0.2577	0.2457	0.2310
		0.97	0.0990	0.0857	0.1002	0.0944	0.0944	0.1847	0.1770	0.1822	0.1754	0.1696	0.4399	0.4477	0.4167	0.3872	0.3511
		0.96	0.1244	0.0956	0.1323	0.1280	0.1236	0.2512	0.2344	0.2465	0.2379	0.2205	0.6051	0.6078	0.5723	0.5285	0.4770
		0.95	0.1453	0.1155	0.1485	0.1416	0.1349	0.3279	0.2779	0.3194	0.3040	0.2811	0.7438	0.7219	0.7091	0.6512	0.5767
		0.9	0.2852	0.1426	0.2926	0.2721	0.2435	0.6921	0.4153	0.6672	0.6012	0.5046	0.9825	0.9411	0.9644	0.9115	0.8232
		0.8	0.5883	0.1191	0.5770	0.4868	0.3570	0.9657	0.3438	0.9399	0.8329	0.6551	1.0000	0.9369	0.9996	0.9894	0.9008
		0.5	0.9544	0.0492	0.8970	0.6499	0.3506	1.0000	0.1061	0.9995	0.9487	0.5959	1.0000	0.6331	1.0000	0.9995	0.8880
	0.1																
		0.99	0.0596	0.0566	0.0628	0.0641	0.0630	0.0797	0.0770	0.0813	0.0835	0.0822	0.1406	0.1402	0.1421	0.1362	0.1273
		0.98	0.0734	0.0724	0.0782	0.0815	0.0796	0.1253	0.1159	0.1293	0.1292	0.1224	0.2691	0.2625	0.2646	0.2488	0.2250
		0.97	0.0934	0.0782	0.0988	0.0989	0.0931	0.1803	0.1443	0.1820	0.1733	0.1621	0.4305	0.4083	0.4251	0.3875	0.3465
		0.96	0.1076	0.0883	0.1138	0.1111	0.1048	0.2418	0.1683	0.2467	0.2319	0.2088	0.5855	0.5229	0.5657	0.5086	0.4430
		0.95	0.1309	0.0898	0.1371	0.1305	0.1160	0.3077	0.1904	0.3069	0.2828	0.2491	0.6959	0.5895	0.6663	0.5923	0.5158
		0.9	0.2238	0.0882	0.2313	0.2165	0.1827	0.5789	0.1658	0.5614	0.4786	0.3763	0.9421	0.5785	0.9040	0.7909	0.6291
		0.8	0.3654	0.0463	0.3630	0.2949	0.1981	0.8453	0.0508	0.7815	0.5790	0.3515	0.9999	0.2075	0.9969	0.9074	0.5835
		0.5	0.9544	0.0492	0.8970	0.6499	0.3506	1.0000	0.1061	0.9995	0.9487	0.5959	1.0000	0.6331	1.0000	0.9995	0.8880

Table 5. Size adjusted power without a trend and with additive outliers: $\boldsymbol{\gamma}\delta_t = \mathbf{0}$, $\lambda \neq 0$.

			$T = 64$					$T = 128$					$T = 256$				
λ	p	ρ	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$	τ^N	τ^{FG}	$\tau^{WSR}_{m=1}$	$\tau^{WSR}_{m=2}$	$\tau^{WSR}_{m=3}$
	0.2																
		0.99	0.0620	0.0557	0.0650	0.0660	0.0607	0.0785	0.0757	0.0824	0.0764	0.0788	0.1255	0.1221	0.1229	0.1213	0.1203
		0.98	0.0770	0.0619	0.0793	0.0764	0.0724	0.1096	0.0874	0.1144	0.1050	0.1052	0.2366	0.2024	0.2310	0.2238	0.2122
		0.97	0.0914	0.0621	0.0953	0.0931	0.0851	0.1451	0.1001	0.1527	0.1412	0.1354	0.3371	0.2444	0.3276	0.3006	0.2725
		0.96	0.1026	0.0569	0.1025	0.0996	0.0927	0.1733	0.0921	0.1772	0.1628	0.1516	0.4081	0.2332	0.3880	0.3441	0.2944
		0.95	0.1176	0.0581	0.1175	0.1134	0.1034	0.2010	0.0852	0.2038	0.1836	0.1650	0.4611	0.2026	0.4280	0.3612	0.2925
		0.9	0.1290	0.0293	0.1262	0.1192	0.0916	0.2247	0.0233	0.2216	0.1627	0.1160	0.6206	0.0394	0.5204	0.3405	0.1805
		0.8	0.0872	0.0053	0.0824	0.0591	0.0336	0.2080	0.0008	0.1771	0.0805	0.0301	0.8064	0.0001	0.6088	0.2186	0.0345
		0.5	0.5920	0.0091	0.4826	0.2576	0.0995	0.9927	0.0035	0.9263	0.5452	0.1481	1.0000	0.0107	1.0000	0.9290	0.2577
	0.3																
		0.99	0.0611	0.0528	0.0566	0.0563	0.0589	0.0797	0.0673	0.0787	0.0778	0.0807	0.1218	0.1036	0.1213	0.1185	0.1155
		0.98	0.0646	0.0550	0.0622	0.0635	0.0667	0.0993	0.0745	0.0966	0.0916	0.0924	0.1861	0.1287	0.1807	0.1742	0.1609
		0.97	0.0728	0.0474	0.0685	0.0662	0.0695	0.1180	0.0680	0.1157	0.1064	0.0995	0.2164	0.1135	0.2087	0.1868	0.1621
		0.96	0.0672	0.0427	0.0617	0.0654	0.0666	0.1174	0.0514	0.1144	0.1060	0.1002	0.2258	0.0789	0.2107	0.1794	0.1412
		0.95	0.0709	0.0338	0.0715	0.0683	0.0660	0.1124	0.0377	0.1073	0.0975	0.0840	0.2231	0.0463	0.2022	0.1563	0.1126
		0.9	0.0368	0.0107	0.0352	0.0327	0.0284	0.0581	0.0038	0.0514	0.0346	0.0190	0.1650	0.0004	0.1213	0.0529	0.0156
		0.8	0.0096	0.0010	0.0074	0.0052	0.0031	0.0100	0.0000	0.0063	0.0020	0.0008	0.0739	0.0000	0.0329	0.0032	0.0000
		0.5	0.0354	0.0008	0.0252	0.0106	0.0024	0.1679	0.0000	0.0927	0.0133	0.0007	0.9674	0.0000	0.7068	0.0655	0.0007

considering that under the alternative of stationarity, $\rho < 1$, frequently occurring outliers, particularly those of larger magnitudes, generate trend-like (nonstationary) effects, thereby precluding decent power.

5. Conclusion

The WSR unit root test presented here exploits the wavelet power spectrum of the observed series and its fractional partial sum to construct a test based on the ratio of norms of the unit scale DWT scaling energies. The proposed test is nonparametric, tuning parameter-free, has good size, is robust to size distortions arising from highly negative MA errors, and is constructed entirely in the wavelet spectral domain. This is a direct improvement over the FG test of FG, which requires tuning parameter specifications through estimation and suffers violent size distortions in the presence of a negative MA parameter. Moreover, theoretical results demonstrate that the WSR and NVR statistic of Nielsen (2009) converge to the same limiting distribution. These results are further extended to models with drifts and linear trends in the context of OLS detrending. Simulation exercises demonstrate that both the WSR and NVR tests enjoy similar power properties although power in both is visibly weaker than that exhibited by the FG test. Where the WSR test truly shines, however, is in finite sample performance.

Simulation experiments show that the WSR test exhibits nontrivial size distortion reductions even when the MA parameter is highly negative. Moreover, the test is more robust to size distortions arising from lowering d than the corresponding NVR test. Accordingly, choosing $d = 0.05$ in contrast to $d = 0.10$ as suggested in Nielsen (2009) has little consequence in terms of size distortion but produces noticeable gains in power. Any remaining size distortions are effectively eliminated using a novel wavestrapping algorithm. Simulations demonstrate that wavestrapping is a viable alternative to traditional time series resampling techniques and can effectively reduce size distortions. Furthermore, unlike the sieve bootstrap, wavestrapping requires no tuning parameter specifications and tuning parameter-free statistics retain this property even when wavestrapped.

Finally, it is not difficult to see the potential of the WSR statistic in tests for cointegration rank. One can generalize the τ^{WSR} by forming a ratio of the scaling vectors y_t where the both the numerator and denominator are fractionally differenced. In particular,

$$\tau_m^{WSR}(d_1) = \left(2^{-m}T\right)^{2d_1} \frac{||\Delta_+^d \widehat{\mathbf{V}}_m||^2}{||\Delta_+^{d+d_1} \widehat{\mathbf{V}}_m||^2} \longrightarrow_d \frac{\int_0^1 \widetilde{\widetilde{J}}_{\phi^{2^m}}(s,d)^2 ds}{\int_0^1 \widetilde{\widetilde{J}}_{\phi^{2^m}}(s,d+d_1)^2 ds}.$$

The statistic using y_t instead of $\mathbf{V}_m$ has been used in Nielsen (2010) to test for cointegration rank in a multivariate framework. Benefits to unit root testing accruing from using τ_m^{WSR} are expected to carry over in tests for cointegration rank as well. This work is being researched further.

Appendix

Proof of Lemma 1. Begin first by abstracting from deterministic dynamics. In this regard, from Eq. 12, it is readily verified that

$$\left(1-\phi^{2^m}L\right)\mathbf{e}_{m,t}^{\top}\mathbf{V}_m=\eta_m(L)\phi_m(L)u_{2^m t+(2^m-1)(l-2)(\text{Mod}T)},$$

where $t=1,\ldots,2^{-m}T$, $\eta_m(L)$ is defined in Eq. 13, and $\phi_m(L)=\sum_{i=0}^{2^m-1}\phi^i L^i$. Then, for any $r\in[0,1]$, note that

$$\begin{aligned}
\mathbf{e}_{m,\lfloor 2^{-m}Tr\rfloor}^{\top}\mathbf{V}_m&=\eta_m(L)\phi_m(L)\left(1-\phi^{2^m}L\right)^{-1}u_{2^m t+(2^t-1)(l-2)(\text{Mod}T)}\\
&=\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor}\phi^{2^m(\lfloor 2^{-m}Tr\rfloor-t)}u_{2^m t+(2^m-1)(l-2)(\text{Mod }T)}\\
&=\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor}\phi^{2^m(\lfloor 2^{-m}Tr\rfloor-t)}\eta_m(L)\phi_m(L)\psi(L)\epsilon_{2^m t+(2^m-1)(l-2)(\text{Mod }T)},
\end{aligned}$$

where the penultimate line follows from $u_t=\psi(L)\epsilon_t$. Accordingly, the partial sum process in Eq. 14 now derives from

$$\mathbf{V}_{m,T}(r)=2^{m/2}T^{-1/2}\mathbf{e}_{m,\lfloor 2^{-m}Tr\rfloor}^{\top}\mathbf{V}_m.$$

Next, for any $\phi\in[0,1]$ and $l,m<\infty$, since $\sum\limits_{i=1}^{l}g_i=\sqrt{2}$, it implies that $\phi_m(1)=\sum\limits_{k=0}^{2^m-1}\phi^k<\infty$, $\eta_m(1)=\prod\limits_{j=1}^{m}\sum\limits_{i=1}^{l}g_i=2^{m/2}<\infty$, and since ϵ_t and $\psi(L)$ satisfy assumption 1, $\psi(1)<\infty$. Accordingly, $\phi_m(1)\eta_m(1)\psi(1)<\infty$. In this regard, let $v_t=\phi_m(L)\eta_m(L)\psi(L)\epsilon_t=\xi(L)\epsilon_t$ and note that v_t admits the Beveridge-Nelson (BN) decomposition (cf. Phillips and Solo (1992)): $v_t=\xi(1)\epsilon_t+\overline{v}_{t-1}-\overline{v}_t$, where $\overline{v}_t=\phi_m(L)\eta_m(L)\overline{\psi}(L)v_t$ with $\overline{\psi}(L)=\sum\limits_{i=0}^{\infty}\overline{\psi}_jL^j$ and $\overline{\psi}_j=\sum\limits_{i=j+1}^{\infty}\psi_i$. Accordingly, $\mathbf{V}_{m,T}(r)$ admits the representation

$$\begin{aligned}
\mathbf{V}_{m,T}(r)&=2^{m/2}T^{-1/2}\xi(1)\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor}\phi^{2^m(\lfloor 2^{-m}Tr\rfloor-t)}\epsilon_{t+(2^m-1)(l-2)(\text{Mod }T)}\\
&\quad+2^{m/2}T^{-1/2}\left(\phi^{2^m(\lfloor 2^{-m}Tr\rfloor-1)}\overline{v}_{(2^m-1)(l-2)(\text{Mod }T)}-\overline{v}_{\lfloor 2^{-m}Tr\rfloor+(2^m-1)(l-2)(\text{Mod }T)}\right)\\
&\equiv 2^{m/2}T^{-1/2}\left(\xi(1)S_{1,T}(r)+S_{2,T}(r)\right),
\end{aligned}$$

where $T^{-1/2}S_{2,T}(r)$ is readily shown to vanish as $T\longrightarrow\infty$. Moreover, since for any positive integer k the binomial theorem implies that $\left(e^{-c_\phi/T}\right)^k=e^{-kc_\phi/T}=\left(1-c_\phi/T+O(T^2)\right)^k=\left(1-c_\phi/T\right)^k+O(T^2)$, it follows that S_1 (where, due to asymptotic negligibility, terms of order $O(T^{-2})$ and lower have been removed) is further decomposed as

$$
\begin{aligned}
S_{1,T}(r) &= \sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor} \phi^{2^m(\lfloor 2^{-m}Tr\rfloor - t)} \epsilon_{2^m t+(2^m-1)(l-2)(\text{Mod } T)} \\
&= \sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor - t)/T} \epsilon_{2^m t+(2^m-1)(l-2)(\text{Mod } T)} \\
&= \sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor} \epsilon_{2^m t+(2^m-1)(l-2)(\text{Mod } T)} \\
&\quad + \sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor-1} \left(e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor - t)/T} - e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor - t-1)/T}\right) \sum_{q=1}^{t} \epsilon_{2^m q+(2^m-1)(l-2)(\text{Mod } T)} \\
&= \sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor} \epsilon_{2^m t+(2^m-1)(l-2)(\text{Mod } T)} - \frac{2^m c_\phi}{T} \sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor-1} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor - t)/T} \sum_{q=1}^{t} \epsilon_{2^m q+(2^m-1)(l-2)(\text{Mod } T)} \\
&\equiv S_{11,T}(r) - S_{12,T}(r),
\end{aligned}
$$

where the penultimate line follows from summation by parts and the mean value theorem (MVT) expansion of $e^{-2^m c_\phi z/T}$ for any $z \in [\lfloor 2^{-m}Tr\rfloor - 1, \lfloor 2^{-m}Tr\rfloor]$ and $d \in (0,1)$. In particular, the latter states that

$$
\begin{aligned}
e^{-2^m c_\phi \lfloor 2^{-m}Tr\rfloor/T} &= e^{-2^m c_\phi\left(\lfloor 2^{-m}Tr\rfloor - 1\right)/T} - \frac{2^m c_\phi}{T} e^{-2^m c_\phi\left(\lfloor 2^{-m}Tr\rfloor - 1\right)/T} + \frac{1}{2}\left(\frac{2^m c_\phi}{T}\right)^2 e^{-2^m c_\phi\left(\lfloor 2^{-m}Tr\rfloor - d\right)/T} \\
&= e^{-2^m c_\phi\left(\lfloor 2^{-m}Tr\rfloor - 1\right)/T} - \frac{2^m c_\phi}{T} e^{-2^m c_\phi\left(\lfloor 2^{-m}Tr\rfloor - 1\right)/T}\left(1 + O(T^{-1})\right).
\end{aligned}
$$

Next, for $s \in [0,1]$, define $W_T(s) = T^{-1/2}\sum_{q=1}^{\lfloor Ts\rfloor} \epsilon_{q+(2^m-1)(l-2)(\text{Mod } T)}$ and note that a standard application of the FCLT implies that $W_T(s) \longrightarrow_d B(s)$ as $T \longrightarrow \infty$. In this regard, consider $S_{11,T}(r)$ and note that for $s = 2^m t/T$,

$$
\begin{aligned}
T^{-1/2}S_{11,T}(r) &= T^{-1/2}\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor} \epsilon_{2^m t+(2^m-1)(l-2)(\text{Mod } T)} \\
&= T^{-1/2}\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor} \int_{2^m(t-1)/T}^{2^m t/T} T^{1/2} dW_T(s) \\
&= \int_0^r dW_T(s).
\end{aligned}
$$

It now readily follows from the continuous mapping theorem (CMT) and the FCLT that

$$
T^{-1/2}S_{11,T}(r) = W_T(r) \longrightarrow_d \sigma_\epsilon B(r).
$$

Turning to $S_{12,T}(r)$, consider $W_T(s)$ for $s = 2^m t/T$, and note that $S_{12,T}(r)$ admits the following representation:

$$\begin{aligned}
T^{-1/2}S_{12,T}(r) &= T^{-1/2}\frac{2^m c_\phi}{T}\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor-1} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor-t)/T}\sum_{q=1}^{t}\epsilon_{2^m q+(2^m-1)(l-2)(\text{Mod } T)}\\
&= T^{-1/2}\frac{2^m c_\phi}{T}\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor-1} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor-t)/T}\sum_{q=1}^{t}\int_{2^m(q-1)/T}^{2^m q/T} T^{1/2}dW_T(s)\\
&= \frac{2^m c_\phi}{T}\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor-1} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor-t)/T}\int_0^{2^m t/T} dW_T(s)\\
&= \frac{2^m c_\phi}{T}\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor-1} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor-t)/T} W_T(2^m t/T)\\
&= c_\phi\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor-1}\int_{t/T}^{2^m(t+1)/T} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor-\lfloor 2^{-m}Ts\rfloor)/T} W_T(s)ds\\
&= c_\phi\int_0^r e^{-c_\phi(r-s)} W_T(s)ds + R_T(r),
\end{aligned}$$

where $R_T(r)$ is the approximation error

$$\begin{aligned}
R_T(r) = {}& c_\phi\int_0^{1/T} e^{-c_\phi(\lfloor 2^{-m}Tr\rfloor-\lfloor 2^{-m}Ts\rfloor)/T} W_T(s)ds\\
&+ c_\phi\int_{\lfloor 2^{-m}Tr\rfloor/T}^{r} e^{-c_\phi(\lfloor 2^{-m}Tr\rfloor-\lfloor 2^{-m}Ts\rfloor)/T} W_T(s)ds\\
&+ c_\phi\int_0^r \left(e^{-c_\phi(\lfloor 2^{-m}Tr\rfloor-\lfloor 2^{-m}Ts\rfloor)/T} - e^{-c_\phi(r-s)}\right) W_T(s)ds.
\end{aligned}$$

Standard arguments (cf. proof of Theorem 3 in Nielsen, 2009) show that $\sup_{0\le r\le 1} R_T(r) \longrightarrow_p 0$, where $\longrightarrow_p$ denotes convergence in probability. Accordingly, applying the CMT and the FCLT, it follows that

$$\begin{aligned}
T^{-1/2}S_{12,T}(r) &= c_\phi\int_0^r e^{-c_\phi(r-s)} W_T(s)ds + o_p(1)\\
&\longrightarrow_d \sigma_\epsilon c_\phi\int_0^r e^{-c_\phi(r-s)} B(s)ds.
\end{aligned}$$

Putting everything together,

$$\begin{aligned}
\mathbf{V}_{m,T}(r) &= 2^{m/2}T^{-1/2}\xi(1)S_{1,T}(r) + o_p(1) \\
&= 2^{m/2}T^{-1/2}\xi(1)\left(S_{11,T}(r) - S_{12,T}(r)\right) + o_p(1) \\
&\longrightarrow_d 2^{m/2}2^{m/2}\phi_m(1)\psi(1)\sigma_\epsilon\left(B(r) - c_\phi\int_0^r e^{-c_\phi(r-s)}B(s)ds\right) \\
&= 2^m\phi_m(1)\psi(1)\sigma_\epsilon J_{c_\phi}(r).
\end{aligned}$$

To demonstrate the case for $\widetilde{\mathbf{V}}_{m,T}(r)$, note that

$$\begin{aligned}
\widetilde{\mathbf{V}}_{m,T}(r) &= \left(2^{-m}T\right)^{-d}\Delta_+^{-d}\mathbf{V}_m(r) = \left(2^{-m}T\right)^{-(1/2+d)}\sum_{k=0}^{\lfloor 2^{-m}Tr\rfloor-1}\pi_k(d)\mathbf{e}_{m,\lfloor 2^{-m}Tt\rfloor-k}^\top\mathbf{V}_m \\
&= \left(2^{-m}T\right)^{-(1/2+d)}\sum_{k=1}^{\lfloor 2^{-m}Tr\rfloor}\pi_{\lfloor 2^{-m}Tr\rfloor-k}(d)\mathbf{e}_{m,k}^\top\mathbf{V}_m \\
&= \left(2^{-m}T\right)^{-(1/2+d)}\sum_{k=1}^{\lfloor 2^{-m}Tr\rfloor}\pi_{\lfloor 2^{-m}Tr\rfloor-k}(d)\xi(1)\sum_{t=1}^{k}\phi^{2^m(k-t)}\epsilon_{2^mt+(2^m-1)(l-2)(\text{Mod } T)} + o_p(1) \\
&\equiv \left(2^{-m}T\right)^{-(1/2+d)}\xi(1)Q_1 + o_p(1),
\end{aligned}$$

where the penultimate line follows from the BN decomposition of $\mathbf{e}_{m,k}^\top\mathbf{V}_m$. Similarly, using summation by parts and the MVT expansion of $e^{-2^mc_\phi z/T}$, it readily follows that

$$\begin{aligned}
Q_{1.T}(r) &= \sum_{k=1}^{\lfloor 2^{-m}Tr\rfloor}\pi_{\lfloor 2^{-m}Tr\rfloor-k}(d)\sum_{t=1}^{k}\phi^{2^m(k-t)}\epsilon_{2^mt+(2^m-1)(l-2)(\text{Mod } T)} \\
&= \sum_{k=1}^{\lfloor 2^{-m}Tr\rfloor}\pi_{\lfloor 2^{-m}Tr\rfloor-k}(d)\left(\sum_{t=1}^{k}\epsilon_{2^mt+(2^m-1)(l-2)(\text{Mod } T)}\right. \\
&\quad\left.-\frac{2^mc_\phi}{T}\sum_{t=1}^{k-1}e^{-2^mc_\phi(k-t)/T}\sum_{q=1}^{t}\epsilon_{2^mq+(2^m-1)(l-2)(\text{Mod } T)}\right) \\
&= \sum_{k=1}^{\lfloor 2^{-m}Tr\rfloor}\sum_{t=1}^{k}\pi_{\lfloor 2^{-m}Tr\rfloor-k}(d)\epsilon_{2^mt+(2^m-1)(l-2)(\text{Mod } T)} \\
&\quad-\frac{2^mc_\phi}{T}\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor-1}e^{-2^mc_\phi(\lfloor 2^{-m}Tr\rfloor-t)/T}\sum_{k=1}^{t}\pi_{t-k}(d)\sum_{q=1}^{k}\epsilon_{2^mq+(2^m-1)(l-2)(\text{Mod } T)} \\
&= \sum_{k=1}^{\lfloor 2^{-m}Tr\rfloor}\pi_{\lfloor 2^{-m}Tr\rfloor-k}(d+1)\epsilon_{2^mk+(2^m-1)(l-2)(\text{Mod } T)} \\
&\quad-\frac{2^mc_\phi}{T}\sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor-1}e^{-2^mc_\phi(\lfloor 2^{-m}Tr\rfloor-t)/T}\sum_{k=1}^{t}\pi_{t-k}(d+1)\epsilon_{2^mk+(2^m-1)(l-2)(\text{Mod } T)} \\
&\equiv Q_{11,T}(r) - Q_{12,T}(r),
\end{aligned}$$

where the antepenultimate line follows by interchanging the orders of summation as in Nielsen (2009), while the penultimate line follows from Sowell (1990) and Wang et al. (2002).

Focusing on $Q_{11,T}(r)$ first, it follows that

$$\begin{aligned}
\left(2^{-m}T\right)^{-(1/2+d)} Q_{11,T}(r) &= \left(2^{-m}T\right)^{-(1/2+d)} \sum_{k=1}^{\lfloor 2^{-m}Tr \rfloor} \pi_{\lfloor 2^{-m}Tr \rfloor - k}(d+1)\epsilon_{2^m k+(2^m-1)(l-2)(\mathrm{Mod}\ T)} \\
&= \left(2^{-m}T\right)^{-(1/2+d)} \sum_{k=1}^{\lfloor 2^{-m}Tr \rfloor} \frac{\left(\lfloor 2^{-m}Tr \rfloor - k\right)^d}{\Gamma(d+1)} \epsilon_{2^m k+(2^m-1)(l-2)(\mathrm{Mod}\ T)} \\
&= \left(2^{-m}T\right)^{-1/2} \sum_{k=1}^{\lfloor 2^{-m}Tr \rfloor} \frac{(r - 2^m k/T)^d}{\Gamma(d+1)} \epsilon_{2^m k+(2^m-1)(l-2)(\mathrm{Mod}\ T)} \\
&= \left(2^{-m}T\right)^{-1/2} \sum_{k=1}^{\lfloor 2^{-m}Tr \rfloor} \frac{(r - 2^m k/T)^d}{\Gamma(d+1)} \int_{2^m(k-1)/T}^{2^m k/T} dS_{11,T}(s) \\
&= 2^{m/2} \sum_{k=1}^{\lfloor 2^{-m}Tr \rfloor} \int_{2^m(k-1)/T}^{2^m k/T} \frac{(r - 2^m k/T)^d}{\Gamma(d+1)} T^{-1/2} dS_{11,T}(s) \\
&= 2^{m/2} \sum_{k=1}^{\lfloor 2^{-m}Tr \rfloor} \int_{2^m(k-1)/T}^{2^m k/T} \frac{(r-s)^d}{\Gamma(d+1)} dW_T(s) \\
&= 2^{m/2} \int_0^r \frac{(r-s)^d}{\Gamma(d+1)} T^{-1/2} dW_T(s) + \widetilde{R}_{1,1}(r),
\end{aligned}$$

where $s = 2^m k/T$ and the approximation error $\sup_{0 \le r \le 1} \widetilde{R}_{1,1}(r) \longrightarrow_p 0$, see Nielsen (2009) for details. Invoking the CMT and the FCLT to $W_T(s)$, it follows that

$$\begin{aligned}
\left(2^{-m}T\right)^{-(1/2+d)} Q_{11,T}(r) &= 2^{m/2} \int_0^r \frac{(r-s)^d}{\Gamma(d+1)} T^{-1/2} dW_T(s) + o_p(1) \\
&\longrightarrow_d 2^{m/2} \sigma_\epsilon \int_0^r \frac{(r-s)^d}{\Gamma(d+1)} dB(s) \\
&= 2^{m/2} \sigma_\epsilon B_{d+1}(r).
\end{aligned}$$

To handle $Q_{12,T}(r)$, observe first that

$$Q_{11,T}(2^m t/T) = \sum_{k=1}^{t} \pi_{t-k}(d+1)\epsilon_{2^m k+(2^m-1)(l-2)(\mathrm{Mod}\ T)}.$$

Accordingly,

$$\begin{aligned}
Q_{1,2} &= \frac{2^m c_\phi}{T} \sum_{t=1}^{\lfloor 2^{-m}Tr \rfloor - 1} e^{-2^m c_\phi (\lfloor 2^{-m}Tr \rfloor - t)/T} \sum_{k=1}^{t} \pi_{t-k}(d+1)\epsilon_{2^m k+(2^m-1)(l-2)(\mathrm{Mod}\ T)} \\
&= \frac{2^m c_\phi}{T} \sum_{t=1}^{\lfloor 2^{-m}Tr \rfloor - 1} e^{-2^m c_\phi (\lfloor 2^{-m}Tr \rfloor - t)/T} Q_{11,T}(2^m t/T).
\end{aligned}$$

It now follows that for $s = 2^m t/T$

$$\left(2^{-m}T\right)^{-(1/2+d)} Q_{12,T}(r) = \left(2^{-m}T\right)^{-(1/2+d)} \frac{2^m c_\phi}{T} \sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor -1} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor - t)/T} Q_{11,T}(2^m t/T)$$

$$= c_\phi \sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor -1} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor - t)/T} \int_{2^m t/T}^{2^m(t+1)/T} \left(2^{-m}T\right)^{-(1/2+d)} Q_{11,T}(s)ds$$

$$= c_\phi \sum_{t=1}^{\lfloor 2^{-m}Tr\rfloor -1} \int_{2^m t/T}^{2^m(t+1)/T} e^{-2^m c_\phi(\lfloor 2^{-m}Tr\rfloor - \lfloor 2^{-m}Ts\rfloor)/T} \left(2^{-m}T\right)^{-(1/2+d)} Q_{11,T}(s)ds$$

$$= c_\phi \int_0^r e^{-c_\phi(r-s)} \left(2^{-m}T\right)^{-(1/2+d)} Q_{11,T}(s)ds + \widetilde{R}_T(r),$$

where $\sup_{0\le r\le 1} \widetilde{R}_T(r) \longrightarrow_p 0$; see the proof of Theorem 3 in Nielsen (2009). Invoking the CMT and the fractional FCLT yields

$$\left(2^{-m}T\right)^{-(1/2+d)} Q_{1,2} = c_\phi \int_0^r e^{-c_\phi(r-s)} \left(2^{-m}T\right)^{-(1/2+d)} Q_{11,T}(s)ds + o_p(1)$$

$$\longrightarrow_d 2^{m/2}\sigma_\epsilon c_\phi \int_0^r e^{-c_\phi(r-s)} B_{d+1}(s)ds.$$

Putting everything together,

$$\begin{aligned}
\widetilde{\mathbf{V}}_{m,T}(r) &= \left(2^{-m}T\right)^{-(1/2+d)} \xi(1)Q_1 + o_p(1)\\
&= \left(2^{-m}T\right)^{-(1/2+d)} \xi(1)\left(Q_{11_T}(r) - Q_{12_T}(r)\right) + o_p(1)\\
&\longrightarrow_d 2^{m/2}\sigma_\epsilon 2^{m/2}\phi_m(1)\psi(1)\left(B_{d+1}(r) - c_\phi \int_0^r e^{-c_\phi(r-s)} B_{d+1}(s)ds\right)\\
&= 2^m \phi_m(1)\psi(1)\sigma_\epsilon \widetilde{J}_{c_\phi}(t,d).
\end{aligned}$$

At last, addressing the presence of deterministic dynamics, note that $y_t = x_t - (\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma})\boldsymbol{\delta}_t$. Accordingly, denote by $\mathbf{V}_m, \widehat{\mathbf{V}}_m, \mathbf{V}_m^\delta$ the scaling vectors of the DWT of x_t, y_t, and δ_t, respectively. In this regard, note from Eq. 12 and the expansion of $2^m\eta_m(L)t$ that $\mathbf{V}_m^\delta$ admits the following representation:

$$\begin{aligned}
\mathbf{e}_{m,t}^\top \mathbf{V}_m^\delta &= \eta_m(L)\left(1, 2^m t\right)^\top\\
&= \left(\eta_m(1), 2^m\eta_m(L)t\right)^\top\\
&= \left(2^{m/2}, 2^{m/2}2^m t - 2^{\frac{3m-1}{2}} \sum_{j=1}^{l-1}(2^m - 1)g_j\right)^\top\\
&= 2^{m/2}\left(1, 2^m t + O(1)\right)^\top.
\end{aligned}$$

Furthermore, define $N_m(T) = \operatorname{diag}\left(1, 2^m T^{-1}\right)$ and observe that as $T \longrightarrow \infty$,

$$\begin{aligned} \mathbf{V}^{\delta}_{m,T}(r) &= N_m(T)\mathbf{e}^{\top}_{m,\lfloor 2^{-m}Tr \rfloor}\mathbf{V}^{\delta}_m \\ &= 2^{m/2}\left(1, \lfloor Tr \rfloor + o(1)\right)^{\top} \\ &\longrightarrow 2^{m/2}\left(1, r\right)^{\top} \\ &= 2^{m/2}\mathbf{D}(r). \end{aligned}$$

Moreover, for $r = 2^m s/T$, it is not difficult to show that

$$\begin{aligned} 2^{m/2}T^{-1/2}\left(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}\right)\mathbf{e}^{\top}_{m,\lfloor 2^{-m}Tr \rfloor}\mathbf{V}^{\delta}_m &= \left(2^m T^{-1}\sum_{s=1}^{2^{-m}T} N_m(T)\mathbf{e}^{\top}_{m,s}\mathbf{V}^{\delta}_m\mathbf{V}^{\delta^{\top}}_m\mathbf{e}_{m,s}N_m(T)\right)^{-1} \\ &\quad\times\left(2^m T^{-1}\sum_{s=1}^{2^{-m}T} 2^{m/2}T^{-1/2}\mathbf{e}^{\top}_{m,s}\mathbf{V}_m\mathbf{V}^{\delta^{\top}}_m\mathbf{e}_{m,s}N_m(T)\right)N_m(T)\mathbf{e}^{\top}_{m,\lfloor 2^{-m}Tr \rfloor}\mathbf{V}^{\delta}_m \\ &= \left(\int_{2^m/T}^{1+2^m/T}\mathbf{V}^{\delta}_{m,T}(r)\mathbf{V}^{\delta^{\top}}_{m,T}(r)\right)^{-1}\left(\int_{2^m/T}^{1+2^m/T}\mathbf{V}_{m,T}(r)\mathbf{V}^{\delta^{\top}}_{m,T}(r)\right)\mathbf{V}^{\delta}_{m,T}(r). \end{aligned}$$

Therefore, it follows that

$$\begin{aligned} \widehat{\mathbf{V}}_{m,T}(r) &= 2^m T^{-1}\mathbf{e}^{\top}_{m,\lfloor 2^{-m}Tr \rfloor}\mathbf{V}_m - 2^{m/2}T^{-1/2}\left(\widehat{\boldsymbol{\gamma}} - \boldsymbol{\gamma}\right)\mathbf{e}^{\top}_{m,\lfloor 2^{-m}Tr \rfloor}\mathbf{V}^{\delta}_m \\ &= \mathbf{V}_{m,T}(r) - \left(\int_{2^m/T}^{1+2^m/T}\mathbf{V}^{\delta}_{m,T}(r)\mathbf{V}^{\delta^{\top}}_{m,T}(r)\right)^{-1}\left(\int_{2^m/T}^{1+2^m/T}\mathbf{V}_{m,T}(r)\mathbf{V}^{\delta^{\top}}_{m,T}(r)\right)\mathbf{V}^{\delta}_{m,T}(r) \\ &\longrightarrow_d 2^m\phi_m(1)\psi(1)\sigma_\epsilon J_{c_\phi}(r) - 2^m\phi_m(1)\psi(1)\sigma_\epsilon\left(\int_0^1\mathbf{D}(r)\mathbf{D}^{\top}(r)\right)^{-1}\left(\int_0^1 J_{c_\phi}(r)\mathbf{D}^{\top}(r)\right)\mathbf{D}(r) \\ &= 2^m\phi_m(1)\psi(1)\sigma_\epsilon\widehat{J}_{c_\phi}(r). \end{aligned}$$

The proof for the fractionally differenced series $\widehat{\widehat{\mathbf{V}}}_{m,T}(r)$ follows in essentially the same way. The details are found in the proof of Theorem 1 in Nielsen (2009) and are therefore omitted. The lemma now follows by noting that $\phi_m(1) = 2^m$ if $c_\phi = 0$ and $\phi_m(1) = \dfrac{1-\phi^{2^m}}{1-\phi}$ if $c_\phi/T \in (0, 2)$.

Proof of Theorem 1. The proof follows immediately from the results in lemma 1 and similar proofs in Nielsen (2009).

Acknowledgment

I am grateful to Ramazan Gençay, Morten Nielsen, Asger Lunde, Joon Park, Victoria Zinde-Walsh, Russell Davidson, Burak Alparslan Eroğlu, Taner Yigit, and seminar participants at Aarhus University, Bilkent University, Indiana University, and McGill University.

Funding

I also thank the Scientific and Technological Research Council of Turkey (TÜBİTAK) 3501 grant for supporting this research financially (project no.: K-SOBAG-115K557).

References

Andrews, D. W. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica: Journal of the Econometric Society* 59(3):817–858.

Basawa, I., Mallik, A., McCormick, W., Reeves, J., Taylor, R. (1991). Bootstrapping unstable first-order autoregressive processes. *The Annals of Statistics* 19(2):1098–1101.

Breitung, J. (2002). Nonparametric tests for unit roots and cointegration. *Journal of Econometrics* 108(2):343–363.

Bühlmann, P. (1997). Sieve bootstrap for time series. *Bernoulli* 3(2):123–148.

Campbell, J. Y., Perron, P. (1991). Pitfalls and opportunities: What macroeconomists should know about unit roots. *NBER Macroeconomics Annual*, 141–201.

Chang, Y., Park, J. (2003). A sieve bootstrap for the test of a unit root. *Journal of Time Series Analysis* 24(4):379–400.

Choi, I., Phillips, P. (1993). Testing for a unit root by frequency domain regression. *Journal of Econometrics* 59(3):263–286.

Davidson, J., De Jong, R. M. (2000). The functional central limit theorem and weak convergence to stochastic integrals ii. *Econometric Theory* 16(05):643–666.

Davidson, J., Hashimzade, N. (2009). Type i and type ii fractional brownian motions: A reconsideration. *Computational Statistics & Data Analysis* 53(6):2089–2106.

Davidson, R., MacKinnon, J. (2000). Bootstrap tests: How many bootstraps? *Econometric Reviews* 19(1):55–68.

Davidson, R., MacKinnon, J. (2007). Improving the reliability of bootstrap tests with the fast double bootstrap. *Computational Statistics & Data Analysis* 51(7):3259–3281.

Davydov, Y. A. (1970). The invariance principle for stationary processes. *Theory of Probability & Its Applications* 15(3):487–498.

Dickey, D., Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 427–431.

Dods, J. L., Giles, D. E. (1995). Alternative stratetgies for augmenting the dickey-fuller test: size-robustness in the face of pre-testing. *Journal of Statistical Computation and Simulation* 53(3-4):243–258.

Elliott, G., Rothenberg, T., Stock, J. (1996). Efficient tests for an autoregressive unit root. *Econometrica* 64(4):813–836.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *The American Economic Review* 81(4):922–930.

Fan, Y., Gençay, R. (2010). Unit root tests with wavelets. *Econometric Theory* 26(05):1305–1331.

Flandrin, P. (1992). Wavelet analysis and synthesis of fractional brownian motion. *Information Theory, IEEE Transactions on* 38(2):910–917.

Franses, P. H., Haldrup, N. (1994). The effects of additive outliers on tests for unit roots and cointegration. *Journal of Business & Economic Statistics* 12(4):471–478.

Gençay, R., Selçuk, F., Whitcher, B. J. (2001). *An introduction to wavelets and other filtering methods in finance and economics*. Academic Press, An Imprint of Elsevier.

Granger, C. (1966). The typical spectral shape of an economic variable. *Econometrica: Journal of the Econometric Society*:150–161.

Horowitz, J. L., Savin, N. (2000). Empirically relevant critical values for hypothesis tests: A bootstrap approach. *Journal of Econometrics* 95(2):375–389.

Johansen, S., Nielsen, M. o. (2012). A necessary moment condition for the fractional functional central limit theorem. *Econometric Theory* 28(03):671–679.

Kiefer, N. M., Vogelsang, T. J. (2000). A new approach to the asymptotics of heteroskedasticity-autocorrelation robust testing. Technical report, mimeo, Department of Economics, Cornell University.

Lee, C., Shie, F.-S. (2004). Fractional integration and the phillips-perron test. *Academia Economic Papers* 32:273–312.

Mallat, S. (1989). A theory for multiresolution signal decomposition: The wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* 11(7):674–693.

Marinucci, D., Robinson, P. (2000). Weak convergence of multivariate fractional processes. *Stochastic Processes and their applications* 86(1):103–120.

Müller, U. (2008). The impossibility of consistent discrimination between i (0) and i (1) processes. *Econometric Theory* 24(03):616–630.

Newey, W. K., West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *The Review of Economic Studies* 61(4):631–653.

Newey, W., West, K. (1987). A simple positive semidefinite heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55:150–161.

Ng, S., Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power. *Econometrica* 69(6):1519–1554.

Nielsen, M. (2008). A powerful tuning parameter free test of the autoregressive unit root hypothesis. *Working Papers*.

Nielsen, M. (2009). A powerful test of the autoregressive unit root hypothesis based on a tuning parameter free statistic. *Econometric Theory* 25(06):1515–1544.

Nielsen, M. (2010). Nonparametric cointegration analysis of fractional systems with unknown integration orders. *Journal of Econometrics* 155(2):170–187.

Park, J., Choi, B. (1988). A new approach to testing for a unit root. *Center for Analytic Economics, Cornell University, Working Paper* 88:23.
Park, J. (1990). Testing for unit roots and cointegration by variable addition. *Advances in Econometrics* 8(2):107–133.
Percival, D., Sardy, S., Davison, A. (2000). Wavestrapping time series: Adaptive wavelet-based bootstrapping. *Nonlinear and Nonstationary Signal Processing*, 442–471.
Percival, D., Walden, A. (2006). *Wavelet methods for time series analysis*, Volume 4. Cambridge Univ Pr.
Perron, P., Qu, Z. (2007). A simple modification to improve the finite sample properties of ng and perron's unit root tests. *Economics Letters* 94(1):12–19.
Phillips, P. C., Solo, V. (1992). Asymptotics for linear processes. *The Annals of Statistics*, 971–1001.
Phillips, P. C. (1987a). Towards a unified asymptotic theory for autoregression. *Biometrika* 74(3):535–547.
Phillips, P., Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika* 75(2):335–346.
Phillips, P. (1987b). Time series regression with a unit root. *Econometrica: Journal of the Econometric Society*, 277–301.
Schwert, G. W. (1987). Effects of model specification on tests for unit roots in macroeconomic data. *Journal of Monetary Economics* 20(1):73–103.
Schwert, G. W. (1989). Tests for unit roots: A monte carlo investigation. *Journal of Business & Economic Statistics* 7(2):147–59.
Sowell, F. (1990). The fractional unit root distribution. *Econometrica: Journal of the Econometric Society*, 495–505.
Tanaka, K. (1999). The nonstationary fractional unit root. *Econometric theory* 15(04):549–582.
Wang, Q., Lin, Y.-X., Gulati, C. M. (2003). Strong approximation for long memory processes with applications. *Journal of Theoretical Probability* 16(2):377–389.
Wang, Q., Lin, Y., Gulati, C. (2002). Asymptotics for general nonstationary fractionally integrated processes without prehistoric influence. *Journal of Applied Mathematics and Decision Sciences* 6(4):255–269.