Unit Root Testing with Slowly Varying Trends

Sven Otto*

December 10, 2019

Abstract: A unit root test is proposed for time series with a general nonlinear deterministic trend component. It is shown that asymptotically the pooled OLS estimator of overlapping blocks filters out any trend component that satisfies some Lipschitz condition. Under both fixed-b and small-b block asymptotics, the limiting distribution of the t-statistic for the unit root hypothesis is derived. Nuisance parameter corrections provide heteroskedasticity-robust tests, and serial correlation is accounted for by pre-whitening. A Monte Carlo study that considers slowly varying trends yields both good size and improved power results for the proposed tests when compared to conventional unit root tests.

Keywords: unit root tests, nonlinear trends, heteroskedasticity.

JEL Classification: C12, C14, C22

*Corresponding address: Institute of Econometrics and Statistics, University of Cologne, Institute of Econometrics and Statistics, Albertus-Magnus-Platz, D-50923 Köln, Germany. Tel.: +49-221-470-6186. Mail: sven.otto@statistik.uni-koeln.de.

1. Introduction

It is widely debated in the time series literature whether macroeconomic variables such as GDP, inflation, and interest rates are $I(1)$ or $I(0)$ around a deterministic trend. Dickey-Fuller-type unit root tests often fail to reject the null hypothesis for these time series. The trend component of a time series y_t is typically treated as known up to some parameter vector. The most commonly applied unit root tests, such as those developed by Dickey and Fuller (1979), Said and Dickey (1984), Phillips (1987), Phillips and Perron (1988), and Elliott et al. (1996), impose either a constant or a linear trend model. If, however, the deterministic trend component is nonlinear, highly persistent trend-stationary processes can be hardly distinguishable from unit root processes.

It is not only a misspecified trend model that may lead to high power losses, as an overparameterized model can also reduce the power of unit root tests. Therefore, many authors have suggested applying trend models that seem more suitable for macro data. Broken trend models with one-time changes in mean or slope with known breakpoint were first studied by Perron (1989) and Rappoport and Reichlin (1989). Christiano (1992) demonstrated that a broken trend model with an unknown breakpoint is more adequate, and Zivot and Andrews (1992), as well as Banerjee et al. (1992), proposed unit root tests for this framework. Structural changes in innovation variances were studied by Hamori and Tokihisa (1997), Kim et al. (2002), and Cavaliere (2005), while Cavaliere et al. (2011) considered unit root testing under broken trends together with nonstationary volatility. Leybourne et al. (1998), Kapetanios et al. (2003), and Kılıç (2011) allowed for exponential smooth transitions from one trend regime to another. Bierens (1997) approximated a nonlinear mean function with Chebyshev polynomials, and Enders and Lee (2012) proposed a Fourier series approximation of the trend, which are approaches that can be used when the exact form and date of structural changes are unknown. For a comprehensive review on the research on unit root testing see Choi (2015).

Dickey-Fuller-type tests are based on the t-statistic of the first-order autoregressive parameter. In case of a constant trend, the estimator is derived from a regression of Δy_t on $(y_{t-1} - \overline{y})$, where $\overline{y}$ is the sample mean. Schmidt and Phillips (1992) estimated the constant by the initial observation, which results in a regression of Δy_t on $(y_{t-1} - y_1)$. Whereas a constant is often not a good global approximation, in a small block, a smoothly varying trend can be approximated quite closely by a constant. To exploit this fact, we propose a block procedure to filter out the unknown trend component. Blocking was also used in Rooch et al. (2019) to estimate the fractional integration parameter in a similar situation. We divide the series into $T - B$ overlapping blocks of length B. As the blocks can be considered as units of a panel, we follow the panel unit root tests proposed by Breitung (2000) and Levin et al. (2002) and consider a pooled regression of Δy_{j+t} on $(y_{j+t-1} - y_j)$ for $2 \leq t \leq T$ and $1 \leq j \leq T - B$. The deterministic function is approximated locally by a constant. Under a general class of piecewise continuous trend functions, the resulting pooled estimator is consistent as $B, T \to \infty$. The limiting null distribution of the t-statistic is a functional of a Brownian motion under fixed-b asymptotics. Under small-b asymptotics, a normal distribution is obtained.

The paper is organized as follows: In Section 2 the autoregressive model with independent and heteroskedastic errors is analyzed together with the asymptotic behavior of the pooled least squares estimator in the presence of a general nonlinear trend component. For both fixed-b and small-b block asymptotics, the limiting distributions are derived under both the unit root hypothesis and under local alternatives. In the presence of heteroskedastic errors, nuisance parameters appear in the limiting distributions, and the estimation of these parameters is discussed. Section 3 considers pseudo t-tests for the unit root hypothesis, and heteroskedasticity-robust test statistics are provided. In Section 4, a pre-whitening procedure is proposed in order to account for short-run dynamics, while Section 5 reports on Monte Carlo simulations. The tests are found to have only minor size distortions in small samples and are sized correctly in larger samples. It is shown that in the presence of slowly varying trends, pooled tests tend to yield higher power than conventional unit root tests. Finally, Section 6 presents the conclusion.

All proofs are placed in the Appendix. In the following, $W(r)$ denotes a standard Brownian motion and "$\Rightarrow$" stands for weak convergence on the càdlàg space $D[0,1]$ together with a suitable norm. $\Theta(\cdot)$ denotes the exact order Landau symbol, that is, $a_T = \Theta(b_T)$ if and only if $a_T = O(y_T)$ and $a_T \neq o(b_T)$, as $T \to \infty$. Moreover, $\lfloor \cdot \rfloor$ is the integer part of its argument, and Δy_t stands for the differenced series $y_t - y_{t-1}$. Finally, $\xrightarrow{\mathcal{D}}$ and $\xrightarrow{p}$ denote convergence in distribution and convergence in probability.

2. The pooled estimator

We are interested in inference concerning the autoregressive parameter ρ in the model

$$y_t = d_t + x_t, \quad x_t = \rho x_{t-1} + u_t, \quad t = 1, \ldots, T, \tag{1}$$

where ρ is close or equal to one. The deterministic trend component d_t is treated as nonstochastic and fixed in repeated samples, where its functional form is nonparametric and unknown.

Assumption 1 (trend component)**.** *The trend component is given by $d_t = d(t/T)$, where $d(r)$ is a piecewise Lipschitz continuous function.*

Note that any continuously differentiable function is Lipschitz continuous. Lipschitz functions are locally close to a constant value in the sense that there exists some $C < \infty$ such that $|d(r) - d(s)| \leq C|r - s|$ for all $r, s \in \mathbb{R}$. The piecewise Lipschitz condition allows for a partition with a finite number of intervals, such that $d(r)$ is Lipschitz continuous on each interval. This includes both smooth changes as well as abrupt breaks in the trend function. For the initial value, it is assumed that $E[x_0^2] < \infty$. We introduce the pooled estimator and the unit root test statistics under the following assumptions on the error term:

Assumption 2 (heteroskedastic errors)**.** *The process $\{u_t\}_{t \in \mathbb{N}}$ is independently distributed with $E[u_t] = 0$, $E[u_t^2] = \sigma_t^2$ and $E[u_t^4] < \infty$, where $\sigma_t = \sigma(t/T)$. The function $\sigma(r)$ is càdlàg, non-stochastic, strictly positive, and bounded.*

The principal approach to dealing with a general, slowly varying trend is to approximate the unknown trend locally by a constant. Let B be some blocklength that satisfies $2 \leq B < T$. We divide the time series into $T - B$ overlapping blocks of length B and then block-wise estimate ρ via OLS under a constant trend specification. In the fashion of Schmidt and Phillips (1992), as well as Breitung and Meyer (1994), the constant trend is estimated by the first observation in each block, which corresponds to the maximum likelihood estimator under the unit root hypothesis $\rho = 0$. Thereafter, by pooling the $T - B$ individual block regressions, we obtain the regression equation

$$\Delta y_{t+j} = \phi(y_{t+j-1} - y_j) + u_{t+j}, \qquad t = 2, \dots, B, \quad j = 1, \dots, T - B,$$

where $\phi = \rho - 1$. The pooled OLS estimator is formulated as

$$\hat{\phi} = \hat{\rho} - 1 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta y_{t+j}(y_{t+j-1} - y_j)}{\sum_{j=1}^{T-B} \sum_{t=2}^{B} (y_{t+j-1} - y_j)^2}.$$

In the following, we derive the asymptotic properties for the numerator and the denominator separately. The numerator and denominator statistics are defined as

$$\mathcal{Y}_{1,T} = \frac{1}{B^{3/2}T^{1/2}} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta y_{t+j}(y_{t+j-1} - y_j), \quad \mathcal{Y}_{2,T} = \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{t=2}^{B} (y_{t+j-1} - y_j)^2,$$

such that $\sqrt{BT}(\hat{\rho} - 1) = \mathcal{Y}_{1,T}/\mathcal{Y}_{2,T}$. Their counterparts without deterministics are given by

$$\mathcal{X}_{1,T} = \frac{1}{B^{3/2}T^{1/2}} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta x_{t+j}(x_{t+j-1} - x_j), \quad \mathcal{X}_{2,T} = \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{t=2}^{B} (x_{t+j-1} - x_j)^2.$$

In what follows, we show that, under the block procedure, the deterministic component can be ignored asymptotically. All asymptotic results are jointly derived for $B, T \to \infty$. While the statistics $\mathcal{X}_{1,T}$ and $\mathcal{X}_{2,T}$ are infeasible if d_t is unknown, they can be well approximated by $\mathcal{Y}_{1,T}$ and $\mathcal{Y}_{2,T}$ in the following sense:

Lemma 1. *Let $\rho = 1 - c/\sqrt{BT}$ with $c \geq 0$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 2. Then, as $B, T \to \infty$, $\mathcal{Y}_{1,T} - \mathcal{X}_{1,T} = O_P(B^{-1/2})$, and $\mathcal{Y}_{2,T} - \mathcal{X}_{2,T} = O_P(T^{-1/2})$.*

Accordingly, we obtain $(\mathcal{Y}_{1,T} - \mathcal{X}_{1,T}, \mathcal{Y}_{2,T} - \mathcal{X}_{2,T}) \xrightarrow{p} (0, 0)$ jointly, and the block procedure filters out the trend component in the numerator and the denominator asymptotically. Hence, applying Slutsky's theorem, we can write

$$\sqrt{BT}(\hat{\rho} - 1) = \frac{\mathcal{Y}_{1,T}}{\mathcal{Y}_{2,T}} = \frac{\mathcal{X}_{1,T}}{\mathcal{X}_{2,T}} + o_P(1).$$

In order to obtain the limiting distribution, we formulate some properties for the numerator and denominator statistics.

Lemma 2. *Let $\rho = 1 - c/\sqrt{BT}$ with $c \geq 0$, and let u_t satisfy Assumption 2. Then, as $B, T \to \infty$, the following statements hold true:*

(a) $\mathcal{X}_{1,T} = \sum_{j=1}^{T} q_{j,T} - c \cdot \mathcal{W}_T$, where $\{q_{j,T},\ j \leq T,\ T \in \mathbb{N}\}$ is a martingale difference array with $q_{j,T} = B^{-3/2}T^{-1/2} \sum_{t \in \mathcal{I}_j} \sum_{k=1}^{t-1} u_j u_{j-k}$, $\mathcal{I}_j = \{t \in \mathbb{N} :\ 1 \leq t \leq B,\ j + B - T \leq t \leq j - 1\}$, and $\mathcal{W}_T = 1/2 \int_0^1 \sigma^2(r)\,\mathrm{d}r + O_P(B^{1/2}T^{-1/2})$.

(b) $Var[\mathcal{X}_{1,T}] = \Theta(1)$ *and* $Var[\mathcal{X}_{2,T}] = \Theta(BT^{-1})$.

(c) If $c = 0$ *and* $\sigma_t^2 = \sigma^2$ *for all* $t \in \mathbb{N}$,

$$v_T^2 := \frac{\sigma^2 Var[\mathcal{X}_{1,T}]}{E[\mathcal{X}_{2,T}]} = \frac{(T-B)(2B-1) - 2(B-2)}{3B(T-B)}.$$

The previous results suggest distinguishing between two fundamentally different types of blocklength asymptotics. The fixed-b approach denotes the case where the relative blocklength B/T converges to some value b with $0 < b < 1$, such that B and T grow at the same rate. In the small-b approach, we consider a relative blocklength that converges to zero, while $B, T \to \infty$.[1] As the blocks are overlapping, the error terms in the pooled regression equation are correlated, but, fortunately, the correlation structure is known by construction. Together with the central limit theorem for martingale difference arrays, the following asymptotic result can be established for the small-b case:

Theorem 1. *Let* $\rho = 1 - c/\sqrt{BT}$ *with* $c \geq 0$, *let* d_t *satisfy Assumption 1, and let* u_t *satisfy Assumption 2. Let* $B/T \to 0$ *as* $B, T \to \infty$. *Then,*

$$\mathcal{Y}_{1,T} \xrightarrow{\mathcal{D}} \mathcal{N}\left(-\frac{c}{2}\int_0^1 \sigma^2(r)\,\mathrm{d}r,\ \frac{1}{3}\int_0^1 \sigma^4(r)\,\mathrm{d}r \right), \quad \text{and} \quad \mathcal{Y}_{2,T} \xrightarrow{p} \frac{1}{2}\int_0^1 \sigma^2(r)\,\mathrm{d}r.$$

Since $\mathcal{Y}_{2,T}$ converges in probability to a constant, we have joint convergence of $(\mathcal{Y}_{1,T}, \mathcal{Y}_{2,T})$, and the pooled estimator is asymptotically normally distributed under small-b asymptotics. Under the unit root hypothesis $\rho = 1$, or, equivalently, if $c = 0$, it follows that

$$\sqrt{BT}(\hat{\rho} - 1) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0,\ \frac{4}{3}\frac{\int_0^1 \sigma^4(r)\,\mathrm{d}r}{(\int_0^1 \sigma^2(r)\,\mathrm{d}r)^2}\right).$$

When imposing fixed-b asymptotics, the numerator and denominator statistics can be represented as a partial sum process of the innovations. Cavaliere (2005) showed that permanent changes in volatility induce a time-shift in the right-hand-side process of the functional central limit theorem. A variance-transformed Brownian process $W_\eta(r)$ appears in the limiting distributions of Dickey-Fuller-type unit root tests. Given the variance profile η, where $\eta(s) = (\int_0^1 \sigma^2(r)dr)^{-1}\int_0^s \sigma^2(r)dr$, the transformed process is defined as $W_\eta(r) = W(\eta(r))$, where $W(r)$ is a standard Brownian motion.

Theorem 2. *Let* $\rho = 1 - c/\sqrt{BT}$ *with* $c \geq 0$, *let* d_t *satisfy Assumption 1, and let* u_t *satisfy Assumption 2. Let* $0 < b < 1$, *and let* $B/T \to b$ *as* $B, T \to \infty$. *Then,*

$$\begin{pmatrix} \mathcal{Y}_{1,T} \\ \mathcal{Y}_{2,T} \end{pmatrix} \xrightarrow{\mathcal{D}} \begin{pmatrix} 0.5b^{-3/2}\int_0^1 \sigma^2(r)\,\mathrm{d}r\big(\int_0^{1-b}(J_{c,b,\eta}(b+r) - J_{c,b,\eta}(r))^2 - b(1-b)\big) \\ b^{-2}\int_0^1 \sigma^2(r)\,\mathrm{d}r \int_0^{1-b}\int_r^{b+r}(J_{c,b,\eta}(s) - J_{c,b,\eta}(r))^2\,\mathrm{d}s\,\mathrm{d}r \end{pmatrix},$$

where $J_{c,b,\eta}(r) = \int_0^r e^{-(r-s)c/b} dW_\eta(s)$ *is a variance-transformed Ornstein-Uhlenbeck process.*

[1] Note that the terminology "fixed-b and small-b asymptotics" was also used in the context of long-run variance estimation. Whereas Kiefer and Vogelsang (2005) used this wording for the asymptotics of the ratio of the truncation point to the sample size, we consider the ratio of the blocklength to the sample size.

Consequently, the pooled estimator is asymptotically represented as a functional of a standard Brownian motion. If $\rho = 1$, then Theorem 2, together with the continuous mapping theorem, implies that

$$\sqrt{BT}(\hat{\rho} - 1) \xrightarrow{\mathcal{D}} \frac{b^{1/2} \int_0^{1-b} (W_\eta(b+r) - W_\eta(r))^2 \, \mathrm{d}r + b^{3/2}(1-b)}{2 \int_0^{1-b} \int_r^{b+r} (W_\eta(s) - W_\eta(r))^2 \, \mathrm{d}s \, \mathrm{d}r}$$

under fixed-b asymptotics. In comparison to the limiting distribution of the ρ-statistic in the Dickey-Fuller framework, the functional includes an additional integral, which results from pooling the block regressions.

In order to estimate the unknown parameters in the limiting distributions, we consider the residuals $\hat{u}_t = y_t - \hat{\rho} y_{t-1}$ for $t = 2, \ldots, T$ and their sample mean $\overline{\hat{u}} = (T-1)^{-1} \sum_{j=2}^T \hat{u}_j$. Let, for notational convenience, $\hat{u}_1 = 0$, and let

$$\hat{\sigma}^2 = \frac{1}{T-2} \sum_{j=2}^{T} (\hat{u}_j - \overline{\hat{u}})^2, \quad \hat{\kappa}^2 = \frac{\sum_{j=1}^{T-B} \sum_{t=1}^{B} \left(\hat{u}_{j+1} - \overline{\hat{u}}\right)^2 \left(\hat{u}_{j+t} - \frac{1}{B} \sum_{k=1}^{B} \hat{u}_{j+k}\right)^2}{\sum_{j=1}^{T-B} \sum_{t=1}^{B} \left(\hat{u}_{j+t} - \frac{1}{B} \sum_{k=1}^{B} \hat{u}_{j+k}\right)^2},$$

$$\hat{\eta}(s) = \frac{\sum_{j=2}^{\lfloor sT \rfloor} \left(\hat{u}_j - \frac{1}{\lfloor sT \rfloor - 1} \sum_{k=2}^{\lfloor sT \rfloor} \hat{u}_k\right)^2 + (sT - \lfloor sT \rfloor) \left(\hat{u}_{\lfloor sT \rfloor + 1} - \frac{1}{\lfloor sT \rfloor} \sum_{k=2}^{\lfloor sT \rfloor + 1} \hat{u}_k\right)^2}{\sum_{j=2}^{T} (\hat{u}_j - \overline{\hat{u}})^2},$$

where $s \in [0, 1]$. We obtain the following consistency results:

Lemma 3. *Let $\rho = 1 - c/\sqrt{BT}$ with $c \geq 0$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 2.*

(a) $\hat{\sigma}^2 \xrightarrow{p} \int_0^1 \sigma^2(r) \, \mathrm{d}r$, *as $B, T \to \infty$.*

(b) $\sup_{s \in [0,1]} |\hat{\eta}(s) - \eta(s)| \xrightarrow{p} 0$, *as $B, T \to \infty$.*

(c) $\hat{\kappa}^2 \xrightarrow{p} \int_0^1 \sigma^4(r) \, \mathrm{d}r / \int_0^1 \sigma^2(r) \, \mathrm{d}r$, *as $B, T \to \infty$ and $B/T \to 0$.*

3. Pseudo t-statistics for unit root testing

The principal concept of Dickey-Fuller-type unit root tests is to consider a t-test for the null hypothesis $H_0 : \rho = 1$. Following this approach in the pooled regression framework, the usual standard error is given by $s_{\hat{\rho}} = \hat{\sigma}(\sum_{j=1}^{T-B} \sum_{t=2}^{B} (y_{t+j-1} - y_j)^2)^{-1/2} = \hat{\sigma}(\mathcal{Y}_{2,T} B^2 T)^{-1/2}$ and the conventional t-statistic is represented as $(\hat{\rho} - 1)/s_{\hat{\rho}} = \sqrt{B} \mathcal{Y}_{1,T} / \sqrt{\hat{\sigma}^2 \mathcal{Y}_{2,T}}$, which diverges in probability under H_0. Accordingly, we consider a scaled pseudo t-statistic of the form

$$\tau = \frac{\hat{\rho} - 1}{s_{\hat{\rho}} \sqrt{B}} = \frac{\mathcal{Y}_{1,T}}{\hat{\sigma} \sqrt{\mathcal{Y}_{2,T}}}, \tag{2}$$

which is $O_P(1)$, as $B, T \to \infty$.

In what follows, pseudo t-tests are defined for both small-b and fixed-b block asymptotics. In order to get a nuisance-parameter-free limiting distribution under small-b asymptotics, we replace $\hat{\sigma}$ by $\hat{\kappa}$ in equation (2). Furthermore, in order to avoid small-sample size distortions, we scale the t-statistic by v_T, which is defined in Lemma 2. The small-b pseudo t-statistic is then given as

$$\tau\text{-SB} = \frac{\mathcal{Y}_{1,T}}{\hat{\kappa} v_T \sqrt{\mathcal{Y}_{2,T}}} = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta y_{t+j}(y_{t+j-1} - y_j)}{\hat{\kappa} v_T \sqrt{B \sum_{j=1}^{T-B} \sum_{t=2}^{B} (y_{t+j-1} - y_j)^2}}.$$

Under fixed-b asymptotics, the nuisance term appears in the Gaussian process itself. By means of transforming the data with its inverse variance profile, Cavaliere and Taylor (2007) showed that the time-transformation in the Gaussian limiting processes can be inverted. Accordingly, we consider the time-transformed series $\tilde{y}_t = y_{\lfloor \hat{\eta}^{-1}(t/T)T \rfloor}$ for $t = 1, \ldots, T$, where $\hat{\eta}^{-1}(s)$ is the inverse function of $\hat{\eta}(s)$. We replace the original series in the test statistic by $\tilde{y}_t$ and define

$$\widetilde{\mathcal{Y}}_{1,T} = \frac{1}{B^{3/2} T^{1/2}} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta \tilde{y}_{t+j}(\tilde{y}_{t+j-1} - \tilde{y}_j), \quad \widetilde{\mathcal{Y}}_{2,T} = \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{t=2}^{B} (\tilde{y}_{t+j-1} - \tilde{y}_j)^2,$$

which yields the fixed-b statistic

$$\tau\text{-FB} = \frac{\widetilde{\mathcal{Y}}_{1,T}}{\hat{\sigma} \sqrt{\widetilde{\mathcal{Y}}_{2,T}}} = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta \tilde{y}_{t+j}(\tilde{y}_{t+j-1} - \tilde{y}_j)}{\hat{\sigma} \sqrt{B \sum_{j=1}^{T-B} \sum_{t=2}^{B} (\tilde{y}_{t+j-1} - \tilde{y}_j)^2}}.$$

In practice, we do not need to discard any observations when transforming the data. The sample size of the auxiliary time series can set arbitrarily high. An auxiliary sample size $\widetilde{T} \geq T$ can be chosen in such a way that $\hat{\eta}^{-1}(t/\widetilde{T}) - \hat{\eta}^{-1}((t-1)/\widetilde{T}) \geq \widetilde{T}^{-1}$ for all $t = 1, \ldots, \widetilde{T}$. Then, the fixed-$b$ statistic is applied to the series $\tilde{y}_t = y_{\lfloor \hat{\eta}^{-1}(t/\widetilde{T})\widetilde{T} \rfloor}$ for $t = 1, \ldots, \widetilde{T}$.

Theorem 3. *Let $\rho = 1 - c/\sqrt{BT}$ with $c \geq 0$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 2.*

(a) Let $B/T \to 0$ as $B, T \to \infty$. Then,

$$\tau\text{-}SB \xrightarrow{\mathcal{D}} \mathcal{N}\left(-\frac{c\sqrt{3}}{2} \frac{\int_0^1 \sigma^2(r)\,\mathrm{d}r}{\sqrt{\int_0^1 \sigma^4(r)\,\mathrm{d}r}},\ 1 \right).$$

(b) Let $0 < b < 1$, and let $B/T \to b$ as $B, T \to \infty$. Then,

$$\tau\text{-}FB \xrightarrow{\mathcal{D}} \frac{\int_0^{1-b} \left(J_{c,b}(b+r) - J_{c,b}(r)\right)^2 \mathrm{d}r - b(1-b)}{2\sqrt{b \int_0^{1-b} \int_r^{b+r} \left(J_{c,b}(s) - J_{c,b}(r)\right)^2 \mathrm{d}s\,\mathrm{d}r}},$$

where $J_{c,b}(r) = \int_0^r e^{-(r-s)c/b} dW(s)$ is a standard Ornstein-Uhlenbeck process.

The unit root hypothesis is rejected in favor of stationarity if the test statistic is smaller than the α-quantile of the limiting distribution for the case $c = 0$, where α is the significance level. For τ-SB we can rely on standard normal quantiles as critical values. The limiting distribution of τ-FB is nonstandard. Note that $J_c(r) = W(r)$ if $c = 0$. Table 1 presents simulated left-tailed quantiles of the null distribution for various relative blocklengths B/T and significance levels.

Table 1: Asymptotic critical values for the fixed-b test

α	B/T								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
0.2	-0.788	-0.812	-0.815	-0.799	-0.761	-0.701	-0.623	-0.520	-0.377
0.1	-1.126	-1.128	-1.104	-1.055	-0.987	-0.903	-0.798	-0.664	-0.486
0.05	-1.403	-1.375	-1.327	-1.257	-1.169	-1.067	-0.939	-0.781	-0.573
0.04	-1.486	-1.446	-1.391	-1.318	-1.222	-1.113	-0.978	-0.814	-0.600
0.03	-1.582	-1.534	-1.471	-1.394	-1.291	-1.169	-1.025	-0.855	-0.630
0.02	-1.709	-1.650	-1.579	-1.489	-1.374	-1.246	-1.094	-0.909	-0.669
0.01	-1.904	-1.830	-1.745	-1.639	-1.511	-1.361	-1.191	-0.995	-0.729
0.001	-2.431	-2.320	-2.203	-2.042	-1.882	-1.692	-1.480	-1.226	-0.905

Note: The sample paths of the standard Brownian motions contained in the asymptotic null distribution of τ-FB are simulated by a discretized version of $W(r)$ on a grid of 50,000 equidistant points. The empirical quantiles are obtained from 100,000 Monte Carlo repetitions.

From the point of view of a practitioner, the τ-SB test has a number of advantages: the distribution is standard normal; thus, there is no need to resort to new tables, and p-values are easy to implement. Furthermore, the unit root test is robust to heteroskedasticity without using any data modification method such as those in Cavaliere and Taylor (2007) and Beare (2018) or wild bootstrap implementations (see Cavaliere and Taylor 2008a).

4. Testing under short-run dynamics

A more realistic scenario for macroeconomic variables is that error terms are serially correlated. We impose the following assumption on the error process:

Assumption 3 (serially correlated errors)**.** *The process $\{u_t\}_{t\in\mathbb{Z}}$ possesses the moving average representation $u_t = \psi(L)\epsilon_t = \sum_{i=0}^{\infty}\psi_i\epsilon_{t-i}$, where L is the usual lag operator, $\psi(z)\neq 0$ for all $|z|\leq 1$, and $\sum_{i=0}^{\infty}|\psi_i|<\infty$. The process $\{\epsilon_t\}_{t\in\mathbb{Z}}$ is independently distributed with $E[\epsilon_t]=0$, $E[\epsilon_t^2]=\sigma_t^2$ and $E[\epsilon_t^4]<\infty$, where $\sigma_t=\sigma(t/T)$. The function $\sigma(r)$ is càdlàg, non-stochastic, strictly positive, and bounded.*

Assumption 3 implies that the moving average representation of u_t is invertible, and we may write $\theta(L)u_t = u_t - \sum_{i=1}^{\infty}\theta_i u_{t-i} = \epsilon_t$, where $\theta(z) = 1-\sum_{i=1}^{\infty}\theta_i z^i$, and $\sum_{i=1}^{\infty}|\theta_i|<\infty$. In order to correct for the effect of short-run dynamics, we follow Breitung and Das (2005) and consider the pre-whitened series $x_t^* = \theta(L)x_t$. By equation (1), it follows that

$$x_t^* = \theta(L)\rho x_{t-1} + \theta(L)u_t = \rho x_{t-1}^* + \epsilon_t,$$

where ϵ_t satisfies the same conditions as u_t under Assumption 2. Consequently, if the unit root statistics are defined in terms of

$$\mathcal{X}_{1,T}^* = \frac{1}{B^{3/2}T^{1/2}}\sum_{j=1}^{T-B}\sum_{t=2}^{B}\Delta x_{t+j}^*(x_{t+j-1}^* - x_j^*), \quad \mathcal{X}_{2,T}^* = \frac{1}{B^2T}\sum_{j=1}^{T-B}\sum_{t=2}^{B}(x_{t+j-1}^* - x_j^*)^2$$

instead of $\mathcal{X}_{1,T}$ and $\mathcal{X}_{2,T}$, their limiting distributions coincide with those presented in the previous sections.

Since the autoregressive parameters of the error process are unknown, they need to be estimated. In the fashion of Said and Dickey (1984) and Chang and Park (2002), we fix some lag order p and consider the AR(p) error representation $u_t = \sum_{i=1}^{p} \theta_i u_{t-i} + \epsilon_{p,t}$ with $\epsilon_{p,t} = \sum_{i=p+1}^{\infty} \theta_i u_{t-i} + \epsilon_t$. Then,

$$\Delta x_t = \phi x_{t-1} + \sum_{i=1}^{p} \theta_i u_{t-i} + \epsilon_{p,T}, \tag{3}$$

which is equal to $\sum_{i=1}^{p} \theta_i \Delta x_{t-i} + \epsilon_{p,T}$ under the unit root hypothesis. The lag order p is allowed to grow with the sample size T. Since the differenced deterministic terms are asymptotically negligible, we may replace Δx_{t-i} by Δy_{t-i} for all $i \geq 0$ in the augmented regression equation. Accordingly, let $(\hat{\varphi}, \hat{\theta}_1, \ldots, \hat{\theta}_p)'$ be the the least squares coefficient vector from the regression of Δy_t on $y_{t-1}, \Delta y_{t-1}, \ldots, \Delta y_{t-p}$, for $t = p+1 \ldots, T$.

Lemma 4. *Let $\rho = 1 - c/\sqrt{BT}$ with $c \geq 0$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 3. Then, $\max_{1 \leq i \leq p} |\hat{\theta}_i - \theta_i| = O_P(B^{-1/2})$, as $p, B, T \to \infty$.*

The estimated pre-whitened series is defined as $\hat{y}_t^* = y_t - \sum_{i=1}^{p} \hat{\theta}_i y_{t-i}$, and the corresponding numerator and denominator statistics are given by

$$\hat{\mathcal{Y}}_{1,T}^* = \frac{1}{B^{3/2}T^{1/2}} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta \hat{y}_{t+j}^* (\hat{y}_{t+j-1}^* - \hat{y}_j^*), \quad \hat{\mathcal{Y}}_{2,T}^* = \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{t=2}^{B} (\hat{y}_{t+j-1}^* - \hat{y}_j^*)^2.$$

Lemma 5. *Let $\rho = 1 - c/\sqrt{BT}$ with $c \geq 0$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 3. Then, as $p, B, T \to \infty$, $\hat{\mathcal{Y}}_{1,T}^* - \mathcal{X}_{1,T}^* = O_P(pB^{-1/2})$, and $\hat{\mathcal{Y}}_{2,T}^* - \mathcal{X}_{2,T}^* = O_P(pT^{-1/2})$.*

As a direct consequence, $(\hat{\mathcal{Y}}_{1,T}^* - \mathcal{X}_{1,T}^*, \hat{\mathcal{Y}}_{2,T}^* - \mathcal{X}_{2,T}^*) \xrightarrow{p} (0,0)$ if $p = o(B^{1/2})$. Let $\hat{\rho}^*$ be given by $\sqrt{BT}(\hat{\rho}^* - 1) = \hat{\mathcal{Y}}_{1,T}^* / \hat{\mathcal{Y}}_{2,T}^*$ and let the pre-whitened residuals be defined as $\hat{u}_t^* = \hat{y}_t^* - \hat{\rho}^* \hat{y}_{t-1}^*$, for $t = p+1, \ldots, T$. For notational convenience, let $\hat{u}_1^* = \ldots = \hat{u}_p^* = 0$. The pre-whitened counterparts of the estimators from Lemma 3 are defined as

$$\hat{\sigma}^{*2} = \frac{1}{T-2} \sum_{j=2}^{T} (\hat{u}_j^* - \overline{\hat{u}^*})^2, \quad \hat{\kappa}^{*2} = \frac{\sum_{j=1}^{T-B} \sum_{t=1}^{B} \left(\hat{u}_{j+1}^* - \overline{\hat{u}^*}\right)^2 \left(\hat{u}_{j+t}^* - \frac{1}{B}\sum_{k=1}^{B} \hat{u}_{j+k}^*\right)^2}{\sum_{j=1}^{T-B} \sum_{t=1}^{B} \left(\hat{u}_{j+t}^* - \frac{1}{B}\sum_{k=1}^{B} \hat{u}_{j+k}^*\right)^2},$$

$$\hat{\eta}(s)^* = \frac{\sum_{j=2}^{\lfloor sT \rfloor} \left(\hat{u}_j^* - \frac{1}{\lfloor sT \rfloor - 1}\sum_{k=2}^{\lfloor sT \rfloor} \hat{u}_k^*\right)^2 + (sT - \lfloor sT \rfloor)\left(\hat{u}_{\lfloor sT \rfloor + 1}^* - \frac{1}{\lfloor sT \rfloor}\sum_{k=2}^{\lfloor sT \rfloor + 1} \hat{u}_k^*\right)^2}{\sum_{j=2}^{T} (\hat{u}_j^* - \overline{\hat{u}^*})^2}.$$

Analogously, we consider the time-transformed pre-whitened series $\tilde{y}_t^* = \hat{y}_{\lfloor \hat{\eta}^{-1}(t/T)T \rfloor}^*$ for all $t = 1, \ldots, T$, and we define

$$\widetilde{\mathcal{Y}}_{1,T}^* = \frac{1}{B^{3/2}T^{1/2}} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta \tilde{y}_{t+j}^* (\tilde{y}_{t+j-1}^* - \tilde{y}_j^*), \quad \widetilde{\mathcal{Y}}_{2,T}^* = \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{t=2}^{B} (\tilde{y}_{t+j-1}^* - \tilde{y}_j^*)^2.$$

For any lag order $p \geq 0$, the pre-whitened versions of the test statistics are given by

$$\tau\text{-SB}_p = \frac{\hat{\mathcal{Y}}_{1,T}^*}{\hat{\kappa}^* v_T \sqrt{\hat{\mathcal{Y}}_{2,T}^*}}, \quad \tau\text{-FB}_p = \frac{\hat{\mathcal{Y}}_{1,T}^*}{\hat{\sigma}^* \sqrt{\hat{\mathcal{Y}}_{2,T}^*}}.$$

Table 2: Trend functions

	type of the trend	functional form
1	sharp break	$d(r) = \lambda \cdot 1_{\{r \leq 2/3\}}$
2	u-shaped break	$d(r) = \lambda \cdot 1_{\{r \leq 1/4\}} + \lambda \cdot 1_{\{r > 3/4\}}$
3	continuous break	$d(r) = \lambda \cdot (4r \cdot 1_{\{r > 2/3\}} - 8/3)$
4	u-shaped break in intercept	$d(r) = \lambda \cdot (r1_{\{r \leq 1/4\}} + (r-1)1_{\{1/4 < r \leq 3/4\}} + r1_{\{t > 3/4\}})$
5	LSTAR break	$d(r) = \lambda \cdot (1 + \exp(20(r - 0.75)))^{-1}$
6	offsetting LSTAR break	$d(r) = \lambda/(1 + \exp(20(r - 0.2))) - 0.5\lambda/(1 + \exp(20(r - 0.75)))$
7	triangular break	$d(r) = \lambda \cdot (2r1_{\{r \leq 1/2\}} + 2(1-r)1_{\{r > 1/2\}})$
8	Fourier break	$d(r) = \lambda \cdot 0.5 \cos(2\pi r)$

Note: The functional form of the trend functions for the simulations are presented. The parameter λ determines the size of the trend.

Note that τ-SB$_0$ = τ-SB and τ-FB$_0$ = τ-FB. To summarize, we obtain the following limiting distributions:

Theorem 4. *Let $\rho = 1 - c/\sqrt{BT}$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 3. Furthermore, let $p = o(B^{1/2})$.*

(a) Let $B/T \to 0$ as $B, T \to \infty$. Then, $\hat{\kappa}^{*2} \xrightarrow{p} \int_0^1 \sigma^4(r)\,\mathrm{d}r / \int_0^1 \sigma^2(r)\,\mathrm{d}r$, and

$$\tau\text{-}SB_p \xrightarrow{\mathcal{D}} \mathcal{N}\Bigg(-\frac{c\sqrt{3}}{2} \frac{\int_0^1 \sigma^2(r)\,\mathrm{d}r}{\sqrt{\int_0^1 \sigma^4(r)\,\mathrm{d}r}},\ 1 \Bigg).$$

(b) Let $0 < b < 1$, and let $B/T \to b$ as $B, T \to \infty$. Then, $\sup_{r \in [0,1]} |\hat{\eta}(s) - \eta(s)| \xrightarrow{p} 0$, $\hat{\sigma}^{*2} \xrightarrow{p} \int_0^1 \sigma^2(r)\,\mathrm{d}r$, and

$$\tau\text{-}FB_p \xrightarrow{\mathcal{D}} \frac{\int_0^{1-b} \left(J_{c,b}(b+r) - J_{c,b}(r)\right)^2\,\mathrm{d}r - b(1-b)}{2\sqrt{b \int_0^{1-b} \int_r^{b+r} \left(J_{c,b}(s) - J_{c,b}(r)\right)^2\,\mathrm{d}s\,\mathrm{d}r}},$$

where $J_{c,b}(r) = \int_0^r e^{-(r-s)c/b} dW(s)$.

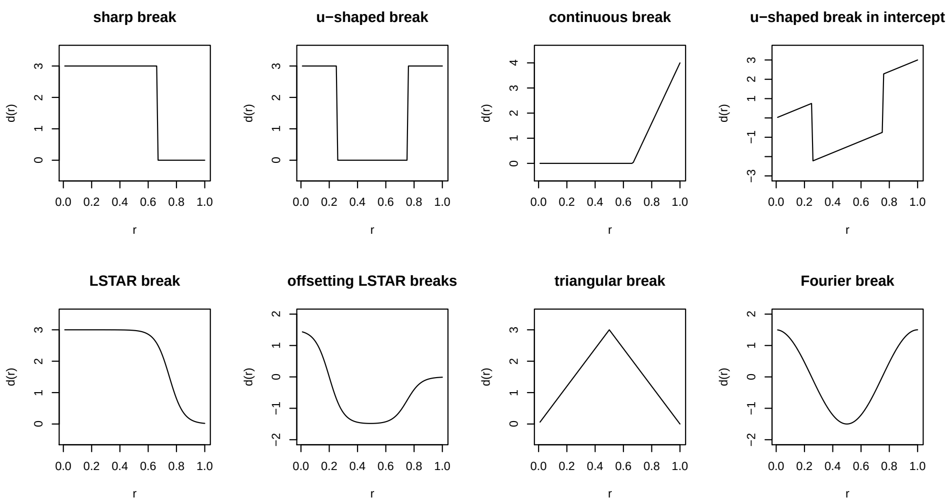
The lag order p is typically unknown in practice and can be chosen using conventional lag order selection methods, such as the Bayesian information criterion (BIC) or by the general-to-specific methodology in the fashion of Ng and Perron (1995). For the maximum lag order p_{max} we can use for instance the rule of thumb provided by Schwert (1989).

5. Simulations

In this section, the finite sample performance of the unit root tests is evaluated by means of Monte Carlo simulations. The analysis includes different specifications for both the deterministic part d_t and the stochastic part x_t.

While the zero-trend $d_t = 0$ is the main benchmark, we consider several other trends including sharp breaks and smooth changes of different shapes. The trend specifications are presented in Table 2 and Figure 1. The parameter λ determines the size of the break. Similar trend functions are also considered in Jones and Enders (2014) in order to evaluate the performance of the unit root test by Enders and Lee (2012).

Figure 1: Plots of the trend functions



Note: The plots of the of the trend functions from Table 2 are presented. The trend size is $\lambda = 3$.

The stochastic part x_t is simulated both under the null hypothesis $\rho = 1$ and the alternative hypothesis $\rho = 0.9$. For the errors u_t, we consider an independent process as well as the AR(1) process $u_t = 0.5u_{t-1} + \epsilon_t$ with standard normal innovations. Furthermore, results with heteroskedastic innovations using the variance function $\sigma^2(r) = 1 + \lambda \cdot 1_{\{r \leq 2/3\}}$ are presented.

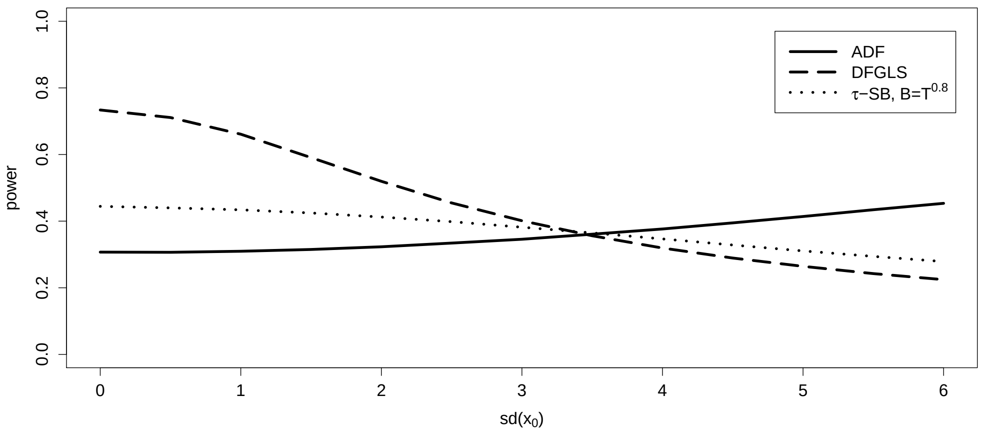
The small-b tests are implemented using blocklengths of the form $B = T^\gamma$ with parameters $\gamma \in \{0.5, 0.6, 0.7, 0.8\}$. For the fixed-$b$ versions, we consider $B = b \cdot T$ with relative blocklengths $b \in \{0.2, 0.4, 0.6\}$. For all tests, the lag augmentation order p is either fixed or flexibly determined by the BIC with a maximum lag order of $p_{max} = 5$. All empirical size levels are presented for a significance level of 5%, and the models are simulated with 100,000 repetitions for sample sizes of $T = 100$ and $T = 300$. As noted by Müller and Elliott (2003), the power of a unit root test depends on the initial condition, and the initial value is simulated as $x_0 \sim \mathcal{N}(0, \sigma_0^2)$ for $\sigma_0^2 \in \{0, 5, 10\}$.

In order to demonstrate the advantage of the fixed-b and small-b unit root tests, their finite sample results are compared to those obtained by conventional unit root tests. As the main benchmark, we consider the augmented Dickey-Fuller test by Said and Dickey (1984) with constant trend specification (ADF henceforth), which is the t-test for the hypothesis $\varphi = 0$ in the regression $\Delta y_t = \varphi y_{t-1} + \beta_0 + \sum_{i=1}^{p} \xi_i \Delta y_{t-i} + e_t$.

Elliott et al. (1996) proposed a feasible point-optimal test with local-to-unity GLS demeaning in the ADF regression. Let the deterministic trend function be given by the vector z_t, and let $\alpha^* = 1 - \bar{c}/T$, where $\bar{c} \in \mathbb{R}$. Furthermore, let $y_{\bar{c},t} = y_t - \alpha^* y_{t-1}$ and $Z_{\bar{c},t} = z_t - \alpha^* z_{t-1}$ for $t \geq 2$, and let $y_{\bar{c},1} = y_1$ and $Z_{\bar{c},1} = z_1$. The Dickey-Fuller GLS test is then the t-test for the hypothesis $\varphi = 0$ in the regression $\Delta y_t^d = \varphi y_{t-1}^d + \sum_{i=1}^{p} \xi_i \Delta y_{t-i}^d + e_t$, where $y_t^d = y_t - \hat{\beta}' z_t$ and where $\hat{\beta}$ is the OLS estimator from a regression of $y_{\bar{c},t}$ on $Z_{\bar{c},t}$. For the constant trend specification (DF-GLS henceforth), we set $z_t = 1$ and $\bar{c} = 7$, and, for the linear trend specification (DF-GLS-trend henceforth), $z_t = (1, t)'$ and $\bar{c} = 13.5$ are considered. Note that the point-optimal test with GLS demeaning is asymptotically equivalent with the Dickey-Fuller test for $d_t = 0$ computed using the series with initial value subtraction (see Elliott et al. 1996)

An approach that does not assume a precise model for the trend component is that developed

Figure 2: Effect of the initial condition on the finite-sample power



Size-adjusted power results for different tests are presented. The initial condition is simulated from a normal distribution with mean zero and different values for $\sigma_0^2 = Var[x_0]$, where σ_0 is shown on the x-axis. The simulation results are reported for for a nominal size level of 5%, for 100,000 replications with $T = 100$, $\rho = 0.9$, the zero trend specification $d_t = 0$, and independent standard normal innovations u_t.

by Enders and Lee (2012) (EL henceforth). A flexible Fourier form is used to approximate smooth breaks in the trend function. Structural changes can be captured by the low frequency components of a series. In its simplest form, Enders and Lee (2012) considered the parametric trend model $d(r) = \alpha_0 + \gamma r + \alpha_1 \sin(2\pi r) + \beta_1 \cos(2\pi r)$. More frequencies could be included, but doing so could lead to an over-fitting problem. The test works as follows: First, the auxiliary regression $\Delta y_t = \delta_0 + \delta_1 \Delta \sin(2\pi t/T) + \delta_2 \Delta \cos(2\pi t/T) + v_t$ is considered with OLS estimates $\widehat{\delta}_0$, $\widehat{\delta}_1$, and $\widehat{\delta}_2$. Let $\widetilde{D}_t = \hat{\delta}_0 t + \hat{\delta}_1 \sin(2\pi t/T) + \hat{\delta}_2 \cos(2\pi t/T)$, which yields the detrended series $\widetilde{S}_t = y_t - \widetilde{D}_t - (y_1 - \widetilde{D}_1)$. Finally, the test statistic is given by the t-statistic for the null hypothesis $\varphi = 0$ in the regression $\Delta y_t = \varphi \widetilde{S}_{t-1} + \beta_0 + \beta_1 \Delta \sin(2\pi t/T) + \beta_2 \Delta \cos(2\pi t/T) + \sum_{i=1}^{p} \xi_i \Delta \widetilde{S}_{t-i} + e_t$.

Harvey and Leybourne (2005, 2006) showed that, if $x_0 \sim \mathcal{N}(0, \sigma_\alpha^2/(1-\rho^2))$ for $\rho = 1-c/T$ with $c > 0$ and some $\sigma_\alpha > 0$, the limiting distributions of the ADF and the DF-GLS test depend on the additional nuisance parameter σ_α. The DF-GLS test is optimal for the zero initial condition $x_0 = 0$, but its power decreases monotonically in σ_α, while the power of the ADF test increases. Figure 2 indicates that the pooled tests are less sensitive to this effect across different values of σ_α. Furthermore, there is no test that outperforms the other tests uniformly across σ_α for this situation in terms of size-adjusted power.

Tables 3–7 present size and actual power results under different model specifications. For smaller sample sizes, the pooled tests have small size distortions, which become larger as the break gets larger. However, for larger sample sizes, the size distortions decline. Overall, the size levels are similar to those obtained from using the conventional unit root tests.

The power of the pooled tests depends on the blocklength. In case of no break, a larger blocklength implies higher power results, which is in line with the theoretical findings that those tests have power in a $1/\sqrt{BT}$ neighborhood of the unit root hypothesis. For blocklengths of $B = T^{0.8}$ in the small-b case and $B = 0.6T$ in the fixed-b case, the power results are similar to those from the ADF test and the Dickey-Fuller GLS test, where the ordering depends on the initial condition (cf. Figure 2). Hence, none of the tests dominates the pooled tests uniformly across these small-sample specifications (although, asymptotically, those tests have power in a $1/T$ neighborhood of the unit root hypothesis). Furthermore, smaller blocklengths, such as $T^{0.6}$

Table 3: Size and power results under the zero-trend specification

initial value	$x_0 = 0$				$x_0 \sim \mathcal{N}(0,5)$				$x_0 \sim \mathcal{N}(0,10)$			
sample size	$T = 100$		$T = 300$		$T = 100$		$T = 300$		$T = 100$		$T = 300$	
ρ	1	0.9	1	0.9	1	0.9	1	0.9	1	0.9	1	0.9
i.i.d. errors – no lag augmentation (p=0)												
τ-SB, $B = T^{0.5}$	0.063	0.346	0.057	0.870	0.064	0.329	0.057	0.864	0.064	0.315	0.057	0.859
τ-SB, $B = T^{0.6}$	0.064	0.407	0.059	0.963	0.064	0.388	0.059	0.961	0.064	0.371	0.059	0.959
τ-SB, $B = T^{0.7}$	0.062	0.459	0.058	0.992	0.061	0.434	0.059	0.991	0.061	0.413	0.059	0.990
τ-SB, $B = T^{0.8}$	0.049	0.428	0.048	0.996	0.049	0.400	0.049	0.995	0.049	0.375	0.049	0.995
τ-FB, $B = 0.2T$	0.042	0.306	0.046	0.973	0.041	0.287	0.046	0.972	0.041	0.270	0.046	0.970
τ-FB, $B = 0.4T$	0.047	0.374	0.047	0.989	0.046	0.346	0.048	0.988	0.047	0.323	0.048	0.987
τ-FB, $B = 0.6T$	0.047	0.386	0.046	0.989	0.047	0.350	0.046	0.988	0.047	0.320	0.046	0.986
ADF	0.054	0.329	0.052	0.996	0.054	0.348	0.050	0.996	0.054	0.367	0.050	0.996
DF-GLS	0.078	0.792	0.058	1.000	0.077	0.617	0.058	0.947	0.077	0.516	0.058	0.858
DF-GLS-trend	0.069	0.371	0.053	0.994	0.069	0.324	0.052	0.955	0.069	0.292	0.052	0.894
EL	0.061	0.140	0.054	0.775	0.061	0.134	0.053	0.755	0.061	0.130	0.053	0.732
AR(1) errors – fixed lag augmentation (p=1)												
τ-SB$_1$, $B = T^{0.5}$	0.012	0.125	0.021	0.679	0.012	0.124	0.022	0.675	0.012	0.121	0.022	0.674
τ-SB$_1$, $B = T^{0.6}$	0.025	0.222	0.038	0.877	0.025	0.220	0.038	0.876	0.025	0.216	0.038	0.875
τ-SB$_1$, $B = T^{0.7}$	0.038	0.305	0.046	0.958	0.037	0.301	0.046	0.957	0.037	0.297	0.046	0.957
τ-SB$_1$, $B = T^{0.8}$	0.034	0.290	0.042	0.972	0.033	0.286	0.042	0.972	0.033	0.281	0.042	0.972
τ-FB$_1$, $B = 0.2T$)	0.025	0.189	0.040	0.922	0.025	0.187	0.040	0.922	0.025	0.184	0.040	0.922
τ-FB$_1$, $B = 0.4T$	0.037	0.270	0.044	0.960	0.037	0.268	0.045	0.961	0.037	0.263	0.045	0.960
τ-FB$_1$, $B = 0.6T$	0.039	0.281	0.044	0.962	0.038	0.276	0.044	0.962	0.037	0.272	0.044	0.961
ADF	0.056	0.263	0.051	0.970	0.056	0.267	0.051	0.971	0.056	0.271	0.051	0.972
DF-GLS	0.077	0.722	0.058	1.000	0.077	0.656	0.058	0.993	0.077	0.602	0.058	0.973
DF-GLS-trend	0.071	0.309	0.052	0.970	0.071	0.297	0.052	0.956	0.071	0.285	0.052	0.937
EL	0.067	0.125	0.056	0.636	0.068	0.125	0.056	0.628	0.068	0.123	0.056	0.620
AR(1) errors – flexible lag augmentation (p determined by BIC)												
τ-SB$_p$, $B = T^{0.5}$	0.006	0.093	0.016	0.680	0.006	0.093	0.016	0.676	0.006	0.091	0.016	0.674
τ-SB$_p$, $B = T^{0.6}$	0.018	0.200	0.033	0.873	0.018	0.198	0.034	0.872	0.018	0.195	0.034	0.871
τ-SB$_p$, $B = T^{0.7}$	0.032	0.296	0.044	0.952	0.032	0.293	0.044	0.953	0.031	0.289	0.044	0.952
τ-SB$_p$, $B = T^{0.8}$	0.032	0.287	0.042	0.968	0.030	0.284	0.041	0.968	0.030	0.280	0.041	0.968
τ-FB$_p$, $B = 0.2T$	0.020	0.171	0.038	0.916	0.020	0.170	0.038	0.917	0.020	0.168	0.038	0.916
τ-FB$_p$, $B = 0.4T$	0.033	0.254	0.043	0.956	0.033	0.254	0.044	0.956	0.033	0.250	0.044	0.956
τ-FB$_p$, $B = 0.6T$	0.035	0.263	0.044	0.957	0.034	0.261	0.043	0.957	0.033	0.258	0.043	0.956
ADF	0.058	0.269	0.051	0.969	0.059	0.272	0.052	0.970	0.059	0.276	0.052	0.971
DF-GLS	0.085	0.703	0.060	0.999	0.084	0.637	0.059	0.991	0.084	0.584	0.059	0.967
DF-GLS-trend	0.082	0.317	0.054	0.960	0.081	0.302	0.055	0.943	0.081	0.289	0.055	0.921
EL	0.106	0.175	0.066	0.637	0.106	0.173	0.064	0.628	0.106	0.171	0.064	0.621

Note: Simulation results are reported for 100,000 replications. The zero-trend $d_t = 0$ is considered for all $t = 1, \ldots, T$. The AR(1) process is given by $u_t = 0.5u_{t-1} + \epsilon_t$. All innovations are simulated independently as standard normal random variables. For the small-b and fixed-b tests, the lag order p refers to the pre-whitening scheme, and, for the conventional tests, p represents the augmentation order. The rejection frequencies are based on the asymptotic critical values for a significance level of 5%.

in the small-b context and $0.2T$ in the fixed-b context, still yield reasonably high power. In particular, the EL test performs much worse in all cases. The size and power results obtained under the AR(1) error specification with both fixed and flexible lag augmentation for the pre-whitening scheme are similar to those produced by i.i.d. errors.

As the tests are designed to yield higher power in the presence of slowly varying trends and breaks, we compare the size-adjusted powers of the tests under the trend specifications presented in Table 2 and Figure 1. For large break sizes λ, it is shown that the smaller the blocklength, the greater the power results. In most cases, the pooled tests have greater power than the ADF, the DF-GLS, the DF-GLS-trend, and the EL test. Furthermore, the power results of the pooled tests are quite uniform across different trend specifications when compared to those of the conventional tests.

Table 6 shows that the pooled tests have reasonable size and power properties under the presence of AR(1) errors and different trend specifications. Furthermore, from Table 7, we can conclude that the tests are sized correctly and have good power properties in the presence of a break in the variance and in the trend function.

Table 4: Size and power results under different trends and i.i.d. errors (1/2)

sample size	$T = 100$						$T = 300$					
ρ	$\rho = 1$			$\rho = 0.9$			$\rho = 1$			$\rho = 0.9$		
λ	3	6	9	3	6	9	3	6	9	3	6	9
sharp break												
τ-SB, $B = T^{0.5}$	0.064	0.064	0.063	0.281	0.194	0.129	0.057	0.058	0.058	0.837	0.752	0.623
τ-SB, $B = T^{0.6}$	0.065	0.067	0.068	0.318	0.198	0.114	0.059	0.061	0.062	0.941	0.861	0.705
τ-SB, $B = T^{0.7}$	0.063	0.068	0.072	0.322	0.155	0.069	0.059	0.060	0.063	0.976	0.885	0.638
τ-SB, $B = T^{0.8}$	0.069	0.117	0.153	0.319	0.189	0.108	0.051	0.056	0.063	0.966	0.709	0.241
τ-FB, $B = 0.2T$	0.041	0.043	0.041	0.218	0.129	0.058	0.046	0.044	0.043	0.936	0.758	0.474
τ-FB, $B = 0.4T$	0.044	0.027	0.011	0.220	0.060	0.009	0.048	0.043	0.033	0.940	0.654	0.237
τ-FB, $B = 0.6T$	0.042	0.022	0.006	0.225	0.055	0.004	0.046	0.040	0.027	0.936	0.639	0.205
ADF	0.050	0.038	0.023	0.169	0.021	0.001	0.049	0.045	0.038	0.898	0.247	0.004
DF-GLS	0.078	0.075	0.065	0.402	0.105	0.011	0.059	0.059	0.059	0.885	0.599	0.142
DF-GLS-trend	0.069	0.067	0.055	0.270	0.164	0.074	0.052	0.052	0.051	0.911	0.729	0.415
EL	0.060	0.056	0.044	0.124	0.096	0.062	0.053	0.052	0.051	0.703	0.565	0.383
u-shaped break												
τ-SB, $B = T^{0.5}$	0.065	0.067	0.064	0.247	0.143	0.089	0.057	0.057	0.058	0.810	0.650	0.452
τ-SB, $B = T^{0.6}$	0.066	0.070	0.070	0.271	0.135	0.072	0.059	0.060	0.062	0.918	0.740	0.464
τ-SB, $B = T^{0.7}$	0.079	0.105	0.109	0.290	0.136	0.069	0.059	0.062	0.066	0.954	0.691	0.280
τ-SB, $B = T^{0.8}$	0.055	0.067	0.069	0.253	0.093	0.034	0.053	0.064	0.079	0.937	0.520	0.116
τ-FB, $B = 0.2T$	0.040	0.031	0.025	0.170	0.059	0.018	0.045	0.041	0.036	0.885	0.477	0.149
τ-FB, $B = 0.4T$	0.045	0.037	0.030	0.196	0.047	0.010	0.048	0.046	0.042	0.878	0.364	0.049
τ-FB, $B = 0.6T$	0.043	0.044	0.057	0.183	0.044	0.013	0.046	0.042	0.038	0.852	0.256	0.024
ADF	0.046	0.027	0.011	0.181	0.030	0.002	0.048	0.041	0.030	0.915	0.329	0.018
DF-GLS	0.077	0.068	0.049	0.435	0.163	0.037	0.059	0.059	0.056	0.885	0.634	0.230
DF-GLS-trend	0.063	0.040	0.017	0.148	0.016	0.000	0.051	0.046	0.036	0.743	0.126	0.001
EL	0.066	0.065	0.054	0.132	0.112	0.080	0.055	0.057	0.057	0.702	0.568	0.405
continuous break												
τ-SB, $B = T^{0.5}$	0.055	0.036	0.017	0.266	0.128	0.029	0.055	0.048	0.038	0.852	0.808	0.719
τ-SB, $B = T^{0.6}$	0.055	0.035	0.016	0.300	0.123	0.019	0.056	0.048	0.038	0.950	0.911	0.789
τ-SB, $B = T^{0.7}$	0.051	0.032	0.014	0.314	0.100	0.011	0.055	0.047	0.036	0.983	0.928	0.680
τ-SB, $B = T^{0.8}$	0.042	0.028	0.014	0.287	0.091	0.010	0.046	0.039	0.030	0.983	0.873	0.449
τ-FB, $B = 0.2T$	0.036	0.023	0.011	0.214	0.080	0.012	0.044	0.037	0.029	0.953	0.846	0.525
τ-FB, $B = 0.4T$	0.040	0.027	0.014	0.261	0.097	0.014	0.046	0.040	0.031	0.972	0.855	0.472
τ-FB, $B = 0.6T$	0.041	0.028	0.015	0.269	0.105	0.016	0.044	0.039	0.032	0.970	0.845	0.461
ADF	0.045	0.027	0.010	0.151	0.011	0.000	0.048	0.040	0.029	0.895	0.235	0.003
DF-GLS	0.064	0.039	0.015	0.351	0.045	0.001	0.056	0.046	0.035	0.885	0.541	0.060
DF-GLS-trend	0.061	0.041	0.021	0.230	0.076	0.011	0.050	0.044	0.035	0.891	0.607	0.192
EL	0.059	0.054	0.047	0.129	0.116	0.097	0.053	0.051	0.048	0.744	0.710	0.652
u-shaped break in intercept												
τ-SB, $B = T^{0.5}$	0.064	0.061	0.056	0.236	0.123	0.068	0.056	0.056	0.055	0.807	0.636	0.424
τ-SB, $B = T^{0.6}$	0.065	0.064	0.058	0.254	0.109	0.049	0.059	0.059	0.058	0.915	0.718	0.414
τ-SB, $B = T^{0.7}$	0.077	0.092	0.089	0.262	0.099	0.039	0.058	0.059	0.060	0.950	0.640	0.202
τ-SB, $B = T^{0.8}$	0.053	0.062	0.058	0.230	0.066	0.017	0.052	0.062	0.073	0.929	0.444	0.063
τ-FB, $B = 0.2T$	0.038	0.029	0.025	0.160	0.055	0.028	0.044	0.039	0.032	0.877	0.435	0.128
τ-FB, $B = 0.4T$	0.044	0.038	0.039	0.198	0.073	0.046	0.048	0.045	0.043	0.877	0.408	0.115
τ-FB, $B = 0.6T$	0.042	0.047	0.086	0.201	0.113	0.134	0.045	0.040	0.038	0.881	0.426	0.164
ADF	0.043	0.022	0.007	0.112	0.004	0.000	0.047	0.037	0.025	0.784	0.051	0.000
DF-GLS	0.073	0.060	0.037	0.353	0.066	0.004	0.058	0.055	0.049	0.907	0.578	0.069
DF-GLS-trend	0.063	0.040	0.017	0.148	0.016	0.000	0.051	0.046	0.036	0.743	0.126	0.001
EL	0.066	0.065	0.054	0.132	0.112	0.080	0.055	0.057	0.057	0.702	0.568	0.405

Note: Simulation results are reported for 100,000 replications. The errors u_t are simulated independently as standard normal random variables. The series are not pre-whitened ($p = 0$). The rejection frequencies are based on the asymptotic critical values for a significance level of 5%.

Table 5: Size and power results under different trends and i.i.d. errors (2/2)

sample size	$T = 100$						$T = 300$					
ρ	$\rho = 1$			$\rho = 0.9$			$\rho = 1$			$\rho = 0.9$		
λ	3	6	9	3	6	9	3	6	9	3	6	9
LSTAR break												
τ-SB, $B = T^{0.5}$	0.057	0.042	0.024	0.282	0.170	0.062	0.055	0.051	0.044	0.856	0.827	0.769
τ-SB, $B = T^{0.6}$	0.057	0.040	0.022	0.318	0.161	0.041	0.057	0.051	0.043	0.954	0.927	0.853
τ-SB, $B = T^{0.7}$	0.054	0.037	0.019	0.327	0.118	0.017	0.056	0.049	0.040	0.985	0.945	0.771
τ-SB, $B = T^{0.8}$	0.044	0.031	0.017	0.287	0.092	0.011	0.047	0.042	0.035	0.983	0.870	0.449
τ-FB, $B = 0.2T$	0.038	0.026	0.014	0.222	0.093	0.019	0.044	0.039	0.032	0.956	0.868	0.599
τ-FB, $B = 0.4T$	0.042	0.030	0.018	0.258	0.098	0.016	0.047	0.042	0.035	0.967	0.821	0.411
τ-FB, $B = 0.6T$	0.042	0.032	0.019	0.262	0.103	0.019	0.045	0.041	0.033	0.964	0.799	0.377
ADF	0.049	0.034	0.019	0.189	0.028	0.001	0.049	0.044	0.036	0.932	0.402	0.019
DF-GLS	0.070	0.051	0.029	0.415	0.101	0.006	0.056	0.051	0.043	0.899	0.671	0.197
DF-GLS-trend	0.063	0.050	0.033	0.265	0.142	0.048	0.051	0.046	0.041	0.916	0.758	0.449
EL	0.059	0.053	0.046	0.129	0.115	0.094	0.053	0.051	0.048	0.741	0.704	0.644
offsetting LSTAR break												
τ-SB, $B = T^{0.5}$	0.056	0.038	0.019	0.276	0.152	0.048	0.056	0.049	0.041	0.854	0.819	0.746
τ-SB, $B = T^{0.6}$	0.055	0.036	0.017	0.307	0.142	0.032	0.057	0.049	0.039	0.952	0.916	0.813
τ-SB, $B = T^{0.7}$	0.052	0.033	0.015	0.320	0.115	0.016	0.056	0.048	0.037	0.983	0.925	0.671
τ-SB, $B = T^{0.8}$	0.042	0.027	0.013	0.281	0.088	0.011	0.047	0.040	0.031	0.978	0.809	0.326
τ-FB, $B = 0.2T$	0.036	0.023	0.011	0.212	0.081	0.014	0.043	0.038	0.029	0.950	0.823	0.471
τ-FB, $B = 0.4T$	0.040	0.026	0.012	0.240	0.077	0.010	0.046	0.039	0.031	0.949	0.691	0.225
τ-FB, $B = 0.6T$	0.039	0.025	0.012	0.229	0.062	0.006	0.045	0.039	0.030	0.930	0.573	0.115
ADF	0.052	0.048	0.048	0.269	0.135	0.059	0.050	0.047	0.045	0.981	0.837	0.452
DF-GLS	0.069	0.045	0.023	0.435	0.136	0.015	0.055	0.048	0.039	0.845	0.511	0.142
DF-GLS-trend	0.060	0.038	0.018	0.211	0.054	0.005	0.049	0.042	0.033	0.854	0.458	0.074
EL	0.060	0.055	0.049	0.131	0.121	0.106	0.053	0.052	0.049	0.747	0.723	0.684
triangular break												
τ-SB, $B = T^{0.5}$	0.055	0.040	0.023	0.282	0.168	0.060	0.055	0.050	0.042	0.855	0.824	0.761
τ-SB, $B = T^{0.6}$	0.056	0.039	0.021	0.318	0.164	0.045	0.057	0.050	0.041	0.954	0.924	0.847
τ-SB, $B = T^{0.7}$	0.054	0.036	0.019	0.335	0.142	0.029	0.056	0.050	0.040	0.985	0.947	0.769
τ-SB, $B = T^{0.8}$	0.042	0.028	0.015	0.290	0.105	0.017	0.046	0.041	0.034	0.977	0.826	0.388
τ-FB, $B = 0.2T$	0.037	0.026	0.014	0.224	0.100	0.024	0.044	0.039	0.031	0.955	0.864	0.579
τ-FB, $B = 0.4T$	0.041	0.028	0.014	0.258	0.098	0.018	0.047	0.041	0.032	0.949	0.715	0.273
τ-FB, $B = 0.6T$	0.041	0.028	0.016	0.262	0.105	0.021	0.045	0.040	0.032	0.957	0.758	0.333
ADF	0.052	0.047	0.042	0.256	0.105	0.027	0.051	0.048	0.045	0.975	0.782	0.331
DF-GLS	0.067	0.045	0.023	0.459	0.175	0.027	0.056	0.049	0.039	0.891	0.682	0.314
DF-GLS-trend	0.059	0.038	0.018	0.202	0.048	0.004	0.050	0.042	0.033	0.841	0.409	0.052
EL	0.060	0.058	0.054	0.133	0.127	0.118	0.053	0.052	0.051	0.752	0.742	0.726
Fourier break												
τ-SB, $B = T^{0.5}$	0.054	0.034	0.015	0.261	0.119	0.025	0.055	0.048	0.038	0.852	0.809	0.718
τ-SB, $B = T^{0.6}$	0.054	0.033	0.014	0.287	0.103	0.013	0.056	0.048	0.037	0.951	0.905	0.762
τ-SB, $B = T^{0.7}$	0.050	0.028	0.011	0.289	0.074	0.006	0.055	0.045	0.034	0.981	0.893	0.496
τ-SB, $B = T^{0.8}$	0.040	0.022	0.009	0.247	0.051	0.003	0.046	0.038	0.028	0.963	0.644	0.113
τ-FB, $B = 0.2T$	0.035	0.020	0.008	0.195	0.056	0.006	0.043	0.036	0.027	0.944	0.756	0.292
τ-FB, $B = 0.4T$	0.038	0.021	0.009	0.218	0.053	0.004	0.045	0.037	0.027	0.923	0.526	0.079
τ-FB, $B = 0.6T$	0.038	0.022	0.009	0.223	0.055	0.004	0.044	0.037	0.027	0.930	0.557	0.095
ADF	0.048	0.037	0.026	0.217	0.054	0.007	0.049	0.044	0.036	0.959	0.594	0.102
DF-GLS	0.066	0.037	0.015	0.427	0.115	0.009	0.055	0.046	0.035	0.885	0.633	0.205
DF-GLS-trend	0.057	0.031	0.011	0.172	0.023	0.001	0.048	0.039	0.028	0.808	0.267	0.010
EL	0.061	0.061	0.061	0.134	0.134	0.134	0.053	0.053	0.053	0.755	0.755	0.755

Note: Simulation results are reported for 100,000 replications. The errors u_t are simulated independently as standard normal random variables. The series are not pre-whitened ($p = 0$). The rejection frequencies are based on the asymptotic critical values for a significance level of 5%.

Table 6: Size and power results under different trends and AR(1) errors

sample size	$T = 100$						$T = 300$					
ρ	$\rho = 1$			$\rho = 0.9$			$\rho = 1$			$\rho = 0.9$		
λ	3	6	9	3	6	9	3	6	9	3	6	9
sharp break												
τ-SB$_p$, $B = T^{0.5}$	0.006	0.007	0.009	0.069	0.042	0.032	0.015	0.014	0.012	0.607	0.477	0.362
τ-SB$_p$, $B = T^{0.6}$	0.019	0.022	0.028	0.160	0.111	0.087	0.033	0.032	0.030	0.837	0.754	0.650
τ-SB$_p$, $B = T^{0.7}$	0.034	0.043	0.060	0.250	0.187	0.147	0.044	0.044	0.046	0.934	0.877	0.781
τ-SB$_p$, $B = T^{0.8}$	0.051	0.108	0.172	0.307	0.325	0.309	0.043	0.047	0.057	0.946	0.860	0.712
τ-FB$_p$, $B = 0.2T$	0.023	0.045	0.069	0.149	0.167	0.156	0.038	0.043	0.058	0.882	0.790	0.692
τ-FB$_p$, $B = 0.4T$	0.029	0.024	0.019	0.187	0.092	0.039	0.042	0.037	0.033	0.920	0.779	0.555
τ-FB$_p$, $B = 0.6T$	0.029	0.016	0.008	0.192	0.072	0.016	0.041	0.033	0.024	0.918	0.767	0.501
u-shaped break												
τ-SB$_p$, $B = T^{0.5}$	0.007	0.011	0.019	0.079	0.066	0.060	0.015	0.014	0.013	0.610	0.471	0.346
τ-SB$_p$, $B = T^{0.6}$	0.020	0.032	0.047	0.175	0.148	0.129	0.032	0.032	0.035	0.833	0.726	0.592
τ-SB$_p$, $B = T^{0.7}$	0.051	0.097	0.140	0.305	0.293	0.253	0.045	0.049	0.058	0.926	0.831	0.680
τ-SB$_p$, $B = T^{0.8}$	0.034	0.053	0.080	0.254	0.213	0.174	0.044	0.055	0.072	0.935	0.816	0.623
τ-FB$_p$, $B = 0.2T$	0.024	0.044	0.056	0.149	0.145	0.117	0.037	0.042	0.054	0.862	0.694	0.515
τ-FB$_p$, $B = 0.4T$	0.034	0.048	0.061	0.211	0.150	0.103	0.043	0.045	0.053	0.902	0.701	0.435
τ-FB$_p$, $B = 0.6T$	0.037	0.060	0.089	0.218	0.164	0.127	0.043	0.048	0.060	0.896	0.662	0.380
continuous break												
τ-SB$_p$, $B = T^{0.5}$	0.006	0.005	0.004	0.079	0.049	0.022	0.016	0.015	0.015	0.654	0.594	0.511
τ-SB$_p$, $B = T^{0.6}$	0.018	0.016	0.013	0.173	0.115	0.060	0.033	0.032	0.030	0.859	0.818	0.746
τ-SB$_p$, $B = T^{0.7}$	0.030	0.027	0.023	0.256	0.173	0.087	0.044	0.043	0.040	0.943	0.906	0.826
τ-SB$_p$, $B = T^{0.8}$	0.029	0.027	0.023	0.250	0.172	0.092	0.041	0.039	0.037	0.956	0.908	0.792
τ-FB$_p$, $B = 0.2T$	0.019	0.017	0.015	0.151	0.107	0.060	0.037	0.036	0.034	0.902	0.851	0.751
τ-FB$_p$, $B = 0.4T$	0.032	0.029	0.025	0.230	0.168	0.100	0.043	0.041	0.040	0.943	0.897	0.796
τ-FB$_p$, $B = 0.6T$	0.033	0.031	0.027	0.237	0.177	0.108	0.043	0.041	0.039	0.943	0.898	0.792
LSTAR break												
τ-SB$_p$, $B = T^{0.5}$	0.005	0.004	0.003	0.063	0.025	0.010	0.016	0.015	0.014	0.613	0.505	0.424
τ-SB$_p$, $B = T^{0.6}$	0.017	0.014	0.010	0.149	0.078	0.035	0.033	0.033	0.030	0.842	0.779	0.706
τ-SB$_p$, $B = T^{0.7}$	0.030	0.025	0.019	0.231	0.126	0.055	0.043	0.042	0.039	0.937	0.890	0.809
τ-SB$_p$, $B = T^{0.8}$	0.028	0.024	0.018	0.222	0.120	0.053	0.041	0.039	0.037	0.951	0.887	0.746
τ-FB$_p$, $B = 0.2T$	0.018	0.015	0.011	0.133	0.074	0.037	0.037	0.036	0.034	0.894	0.829	0.723
τ-FB$_p$, $B = 0.4T$	0.031	0.027	0.021	0.204	0.123	0.063	0.044	0.042	0.040	0.936	0.866	0.729
τ-FB$_p$, $B = 0.6T$	0.032	0.028	0.023	0.215	0.134	0.071	0.042	0.041	0.039	0.934	0.858	0.717
Fourier break												
τ-SB$_p$, $B = T^{0.5}$	0.005	0.005	0.003	0.081	0.055	0.029	0.016	0.015	0.015	0.658	0.608	0.533
τ-SB$_p$, $B = T^{0.6}$	0.017	0.014	0.011	0.176	0.122	0.067	0.033	0.032	0.030	0.860	0.821	0.750
τ-SB$_p$, $B = T^{0.7}$	0.030	0.025	0.019	0.257	0.175	0.090	0.044	0.042	0.039	0.941	0.897	0.796
τ-SB$_p$, $B = T^{0.8}$	0.029	0.024	0.018	0.247	0.162	0.079	0.041	0.039	0.036	0.946	0.853	0.648
τ-FB$_p$, $B = 0.2T$	0.019	0.016	0.012	0.150	0.105	0.058	0.037	0.036	0.033	0.898	0.832	0.693
τ-FB$_p$, $B = 0.4T$	0.031	0.026	0.021	0.224	0.153	0.082	0.043	0.041	0.038	0.925	0.811	0.593
τ-FB$_p$, $B = 0.6T$	0.032	0.028	0.021	0.231	0.158	0.086	0.043	0.041	0.038	0.930	0.828	0.621

Note: Simulation results are reported for 100,000 replications. The errors u_t are simulated from $u_t = 0.5u_{t-1} + \epsilon_t$ with independent standard normal innovations, and the series are pre-whitened with a lag order p that is determined from the BIC. The rejection frequencies are based on the asymptotic critical values for a significance level of 5%.

Table 7: Size and size-adjusted power of robust tests under breaks in trend and variance

sample size	$T = 100$						$T = 300$					
ρ	$\rho = 1$			$\rho = 0.9$			$\rho = 1$			$\rho = 0.9$		
λ	2	3	4	2	3	4	2	3	4	2	3	4
sharp break in variance												
τ-SB, $B = T^{0.5}$	0.067	0.069	0.069	0.344	0.337	0.329	0.057	0.056	0.056	0.847	0.806	0.767
τ-SB, $B = T^{0.6}$	0.071	0.075	0.077	0.420	0.421	0.416	0.062	0.061	0.061	0.954	0.933	0.909
τ-SB, $B = T^{0.7}$	0.081	0.095	0.107	0.526	0.565	0.585	0.068	0.072	0.074	0.992	0.987	0.981
τ-SB, $B = T^{0.8}$	0.085	0.124	0.162	0.569	0.683	0.756	0.082	0.116	0.147	0.999	1.000	1.000
τ-FB, $B = 0.2T$	0.040	0.039	0.040	0.283	0.261	0.238	0.045	0.044	0.042	0.947	0.882	0.812
τ-FB, $B = 0.4T$	0.043	0.042	0.041	0.346	0.308	0.276	0.047	0.046	0.045	0.982	0.935	0.876
τ-FB, $B = 0.6T$	0.042	0.040	0.042	0.349	0.327	0.307	0.045	0.045	0.045	0.989	0.974	0.947
sharp break in trend and variance												
τ-SB, $B = T^{0.5}$	0.066	0.068	0.070	0.324	0.305	0.283	0.057	0.056	0.056	0.836	0.786	0.737
τ-SB, $B = T^{0.6}$	0.071	0.074	0.076	0.391	0.370	0.344	0.062	0.061	0.061	0.946	0.916	0.881
τ-SB, $B = T^{0.7}$	0.080	0.091	0.099	0.474	0.470	0.444	0.067	0.071	0.073	0.988	0.976	0.959
τ-SB, $B = T^{0.8}$	0.095	0.145	0.194	0.526	0.595	0.627	0.081	0.111	0.136	0.997	0.996	0.993
τ-FB, $B = 0.2T$	0.040	0.039	0.038	0.260	0.228	0.197	0.046	0.044	0.043	0.935	0.854	0.765
τ-FB, $B = 0.4T$	0.044	0.042	0.043	0.295	0.240	0.200	0.047	0.046	0.046	0.960	0.872	0.772
τ-FB, $B = 0.6T$	0.042	0.043	0.047	0.292	0.240	0.205	0.046	0.045	0.046	0.954	0.866	0.766

Note: Simulation results are reported for 100,000 replications. The errors u_t are simulated independently as standard normal random variables, and the series are not pre-whitened ($p = 0$). The sharp break specification is defined by a break in the variance at 2/3 of the sample. The rejection frequencies are based on the asymptotic critical values for a significance level of 5%.

The blocklength B is a tuning parameter that needs to be chosen carefully, and any optimality result would depend on the actual trend model. In practice, however, the trend model is unknown, which makes it hard to derive an optimal blocklength. Although theoretical recommendations cannot be formulated based on the current analysis, the small-b tests with $B = T^{0.7}$ and the fixed-b tests with $T = 0.2B$ yield very promising results for all trend functions studied in this paper and are therefore recommended as the default settings.

6. Conclusion

We have presented two variants of a unit root test under an unknown trend specification that are robust under both heteroskedasticity and autocorrelation. When applied to finite samples, the tests show good size properties. The fixed-b pooled test statistic converges to a functional of a Brownian motion under the unit root hypothesis, while the small-b variant shows a standard normal distribution in the limit. Autocorrelation-robust versions of the tests were introduced using a pre-whitening scheme. Monte Carlo simulations indicate that, while under the zero-trend specification, the fixed-b and small-b tests perform similar to the conventional tests in terms of size and power, under sharp breaks as well as smooth changes in the trend, their power is much higher. Furthermore, the powers of the tests are less sensitive to the initial value when compared to the augmented Dickey-Fuller test and the Dickey-Fuller GLS test.

Acknowledgements

I would like to thank Jörg Breitung and Matei Demetrescu for their extensive advice and support. My thanks also go to Hans Manner, Markus Kösler, Robinson Kruse-Becher, Dominik Wied, Nazarii Salish, Uwe Hassler, and Martin Wagner for their helpful comments. The suggestions made by participants attending the 2015 RMSE meeting in Cologne, the SMYE conference 2017 in Halle (Saale), the SNDE conference 2017 in Paris, and the IAAE conference 2017 in Sapporo are also highly appreciated. Furthermore, the usage of the CHEOPS HPC cluster for parallel computing and a conference grant of the International Association for Applied Econometrics are greatfully acknowledged.

References

Banerjee, A., Lumsdaine, R. L., and Stock, J. H. (1992). Recursive and sequential tests of the unit-root and trend-break hypotheses: theory and international evidence. *Journal of Business & Economic Statistics*, 10:271–287.

Beare, B. K. (2018). Unit root testing with unstable volatility. *Journal of Time Series Analysis*, 39(6):816–835.

Bierens, H. J. (1997). Testing the unit root with drift hypothesis against nonlinear trend stationarity, with an application to the us price level and interest rate. *Journal of Econometrics*, 81:29–64.

Breitung, J. (2000). The local power of some unit root tests for panel data. In *Advances in Econometrics, Vol. 15: Nonstationary Panels, Panel Cointegration, and Dynamic Panels*, pages 161–177. Emerald Group Publishing.

Breitung, J. and Das, S. (2005). Panel unit root tests under cross-sectional dependence. *Statistica Neerlandica*, 59:414–433.

Breitung, J. and Meyer, W. (1994). Testing for unit roots in panel data: are wages on different bargaining levels cointegrated? *Applied Economics*, 26:353–361.

Cavaliere, G. (2005). Unit root tests under time-varying variances. *Econometric Reviews*, 23:259–292.

Cavaliere, G., Harvey, D. I., Leybourne, S. J., and Taylor, A. R. (2011). Testing for unit roots in the presence of a possible break in trend and nonstationary volatility. *Econometric Theory*, 27:957–991.

Cavaliere, G. and Taylor, A. R. (2007). Testing for unit roots in time series models with nonstationary volatility. *Journal of Econometrics*, 140:919–947.

Cavaliere, G. and Taylor, A. R. (2008a). Bootstrap unit root tests for time series with nonstationary volatility. *Econometric Theory*, 24(1):43–71.

Cavaliere, G. and Taylor, A. R. (2008b). Time-transformed unit root tests for models with non-stationary volatility. *Journal of Time Series Analysis*, 29:300–330.

Chang, Y. and Park, J. Y. (2002). On the asymptotics of adf tests for unit roots. *Econometric Reviews*, 21(4):431–447.

Choi, I. (2015). *Almost all about unit roots: Foundations, developments, and applications*. Cambridge University Press.

Christiano, L. J. (1992). Searching for a break in GNP. *Journal of Business & Economic Statistics*, 10:237–250.

Davidson, J. (1994). *Stochastic limit theory: an introduction for econometricians.* Oxford University Press.

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74:427–431.

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

Enders, W. and Lee, J. (2012). A unit root test using a fourier series to approximate smooth breaks. *Oxford Bulletin of Economics and Statistics*, 74:574–599.

Hamori, S. and Tokihisa, A. (1997). Testing for a unit root in the presence of a variance shift. *Economics Letters*, 57:245–253.

Harvey, D. I. and Leybourne, S. J. (2005). On testing for unit roots and the initial observation. *The Econometrics Journal*, 8(1):97–111.

Harvey, D. I. and Leybourne, S. J. (2006). Power of a unit-root test and the initial condition. *Journal of Time Series Analysis*, 27(5):739–752.

Jones, P. M. and Enders, W. (2014). On the use of the flexible fourier form in unit root tests, endogenous breaks, and parameter instability. In *Recent Advances in Estimating Nonlinear Models*, pages 59–83. Springer.

Kapetanios, G., Shin, Y., and Snell, A. (2003). Testing for a unit root in the nonlinear STAR framework. *Journal of Econometrics*, 112:359–379.

Kiefer, N. M. and Vogelsang, T. J. (2005). A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. *Econometric Theory*, 21:1130–1164.

Kılıç, R. (2011). Testing for a unit root in a stationary ESTAR process. *Econometric Reviews*, 30:274–302.

Kim, T.-H., Leybourne, S., and Newbold, P. (2002). Unit root tests with a break in innovation variance. *Journal of Econometrics*, 109:365–387.

Levin, A., Lin, C.-F., and Chu, C.-S. J. (2002). Unit root tests in panel data: asymptotic and finite-sample properties. *Journal of Econometrics*, 108:1–24.

Leybourne, S., Newbold, P., and Vougas, D. (1998). Unit roots and smooth transitions. *Journal of Time Series Analysis*, 19:83–97.

Müller, U. K. and Elliott, G. (2003). Tests for unit roots and the initial condition. *Econometrica*, 71:1269–1286.

Ng, S. and Perron, P. (1995). Unit root tests in ARMA models with data-dependent methods for the selection of the truncation lag. *Journal of the American Statistical Association*, 90:268–281.

Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis. *Econometrica*, 57:1361–1401.

Phillips, P. C. (1987). Time series regression with a unit root. *Econometrica*, 55:277–301.

Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika*, 75:335–346.

Rappoport, P. and Reichlin, L. (1989). Segmented trends and non-stationary time series. *The Economic Journal*, 99:168–177.

Rooch, A., Zelo, I., and Fried, R. (2019). Estimation methods for the lrd parameter under a change in the mean. *Statistical Papers*, 60:313–347.

Said, S. E. and Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika*, 71:599–607.

Schmidt, P. and Phillips, P. C. (1992). Lm tests for a unit root in the presence of deterministic trends. *Oxford Bulletin of Economics and Statistics*, 54:257–287.

Schwert, G. W. (1989). Tests for unit roots: a monte carlo investigation. *Journal of Business & Economic Statistics*, 7:147–159.

Zivot, E. and Andrews, D. W. (1992). Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics*, 10:251–270.

A. Appendix: Proofs

A.1. Auxiliary results

Lemma A.1. *Let $\rho = 1 - c/\sqrt{BT}$ with $c \geq 0$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 2. Furthermore, let $1 \leq s \leq B$. Then,*

(a) $\sum_{r=1}^{B} \big| \sum_{j=1}^{T-B} \Delta d_{r+j} \Delta d_{s+j} \big| = O(1)$

(b) $\sum_{r=1}^{B} \big| \sum_{j=1}^{T-B} \Delta d_{r+j} \Delta x_{s+j} \big| = O_P(T^{1/2})$

Proof. Since $d(r)$ is piecewise Lipschitz continuous on the unit interval, there are a finite number of points where $d(r)$ is not continuous. Let those points be given by $\{\pi_1, \ldots, \pi_L\}$, where $L < \infty$ and $0 < \pi_1 < \ldots < \pi_L < 1$. We can represent $d(r)$ by some function $\delta(r)$ that is Lipschitz continuous on the entire domain. Then, for any $r \in [0,1]$, we obtain $d(r) = \delta(r) + \sum_{l=1}^{L} \lambda_l 1_{\{r \geq \pi_l\}}$, with $\sum_{l=1}^{L} |\lambda_l| < \infty$. Let $p_l = \lfloor \pi_l T \rfloor$ for $l = 1, \ldots, L$, and let $\delta_t = \delta(t/T)$ for $t = 1, \ldots, T$. Then, $d_t = \delta_t + \sum_{l=1}^{L} \lambda_l 1_{\{t \leq p_l\}}$, and consequently, $\Delta d_t = \Delta \delta_t + \sum_{l=1}^{L} \lambda_l 1_{\{t = p_l\}}$. Due to the Lipschitz continuity of $\delta(r)$, there exists a constant $C_1 < \infty$, such that

$$\big|\Delta d_t\big| \leq C_1 T^{-1} + \sum_{l=1}^{L} |\lambda_l| 1_{\{t = p_l\}}, \tag{A.1}$$

for all indices $t = 2, \ldots, T$. Furthermore, $\sum_{l=1}^{L} |\lambda_l| = C_2$ and $\sigma(r) < C_3$ for some constants $C_2, C_3 < \infty$, and $C = \max\{C_1, C_2, C_3, E[|x_0|], c, 1\} < \infty$. For (a), we have

$$\sum_{r=1}^{B} \Big| \sum_{j=1}^{T-B} \Delta d_{r+j} \Delta d_{s+j} \Big| \leq 3C^2 B T^{-1} + \sum_{r=1}^{B} \sum_{j=1}^{T-B} \sum_{l_1, l_2 = 1}^{L} |\lambda_{l_1} \lambda_{l_2}| 1_{\{r+j = p_{l_1}\}} 1_{\{s+j = p_{l_2}\}} = O(1).$$

To show (b), note that $\Delta x_t = (\rho - 1) x_{t-1} + u_t = (\rho - 1)(x_{t-1} - x_0) + (\rho - 1) x_0 + u_t$. We decompose $\sum_{r=1}^{B} \Big| \sum_{j=1}^{T-B} \Delta d_{r+j} \Delta x_{s+j} \Big| \leq A_1 + A_2 + A_3$, where

$$A_1 = \frac{c \sum_{r=1}^{B} | \sum_{j=1}^{T-B} \Delta d_{r+j} (x_{s+j-1} - x_0)|}{B^{1/2} T^{1/2}}, \quad A_2 = \sum_{r=1}^{B} \Big| \sum_{j=1}^{T-B} \Delta d_{r+j} u_{s+j} \Big|,$$

$$A_3 = \frac{c \sum_{r=1}^{B} \sum_{j=1}^{T-B} |\Delta d_{r+j} x_0|}{B^{1/2} T^{1/2}}.$$

From inequality (A.1), Jensen's inequality, and the MA-representation $x_t - x_0 = \sum_{m=0}^{t-1} \rho^m u_{t-m}$, it follows that

$$\begin{aligned} E[|A_1|] &\leq (1-\rho) \sum_{r=1}^{B} \sqrt{E\Big[\Big(\sum_{j=1}^{T-B} \sum_{m=0}^{s+j-2} \rho^m \Delta d_{r+j} u_{s+j-m-1} \Big)^2\Big]} \\ &\leq (1-\rho) \sum_{r=1}^{B} \sqrt{\sum_{j=1}^{T-B} \sum_{m_1, m_2 = 0}^{\infty} C^2 \rho^{m_1 + m_2} |\Delta d_{r+j} \Delta d_{r+j-m_1+m_2}|} \\ &\leq 2C^2 (1-\rho) \sqrt{\sum_{j=1}^{T-B} \sum_{m_1, m_2 = 0}^{\infty} \rho^{m_1 + m_2}} = O(T^{1/2}), \end{aligned}$$

$$E[|A_2|] \leq \sum_{r=1}^{B} \sqrt{E\Big[\Big(\sum_{j=1}^{T-B} \Delta d_{r+j} u_{s+j}\Big)^2\Big]} \leq \sum_{r=1}^{B} \sqrt{C^2 \sum_{j=1}^{T-B} |\Delta d_{r+j}|^2} = \sqrt{4C^4 T} = O(T^{1/2}),$$

and

$$E[|A_3|] = E\Big[(1-\rho) \sum_{r=1}^{B} \Big| \sum_{j=1}^{T-B} \Delta d_{r+j} x_0 \Big|\Big] \leq \frac{cE[|x_0|]}{B^{1/2}T^{1/2}} \sum_{r=1}^{B} \sum_{j=1}^{T-B} |\Delta d_{r+j}| \leq \frac{2C^3 B^{1/2}}{T^{1/2}} = O(1).$$

The assertion follows by Markov's inequality and the triangle inequality. □

A.2. Proof of Lemma 1

First, we reformulate the numerator and denominator statistics. Note that

$$\begin{aligned} &\Delta y_{t+j}(y_{t+j-1} - y_j) - \Delta x_{t+j}(x_{t+j-1} - x_j) \\ &= \Delta d_{t+j}(d_{t+j-1} - d_j) + \Delta d_{t+j}(x_{t+j-1} - x_j) + \Delta x_{t+j}(d_{t+j-1} - d_j), \end{aligned}$$

and

$$(y_{t+j-1} - y_j)^2 - (x_{t+j-1} - x_j)^2 = (d_{t+j-1} - d_j)^2 + 2(x_{t+j-1} - x_j)(d_{t+j-1} - d_j).$$

We decompose $\mathcal{Y}_{1,T} - \mathcal{X}_{1,T} = S_1 + S_2 + S_3$ and $\mathcal{Y}_{2,T} - \mathcal{X}_{2,T} = S_4 + S_5$, where

$$\begin{aligned} S_1 &= \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta d_{t+j}(d_{t+j-1} - d_j)}{B^{3/2}T^{1/2}}, \quad S_2 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta d_{t+j}(x_{t+j-1} - x_j)}{B^{3/2}T^{1/2}}, \\ S_3 &= \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta x_{t+j}(d_{t+j-1} - d_j)}{B^{3/2}T^{1/2}}, \quad S_4 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} (d_{t+j-1} - d_j)^2}{B^2 T}, \\ S_5 &= \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} 2(x_{t+j-1} - x_j)(d_{t+j-1} - d_j)}{B^2 T}. \end{aligned}$$

Lemma A.1 yields $S_1 + S_2 + S_3 = O_P(B^{-1/2})$, and $S_4 + S_5 = O_P(T^{-1/2})$, and the assertion follows by Slutsky's theorem.

A.3. Proof of Lemma 2

(a): From the representation $\Delta x_{t+j} = u_{t+j} + \phi x_{t+j-1}$ with $\phi = -c/\sqrt{BT}$, we decompose the numerator statistic into $\mathcal{X}_{1,T} = S_1 + S_2 + S_3 + S_4$, where

$$\begin{aligned} S_1 &= \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} u_{t+j} u_{k+j}}{B^{3/2}T^{1/2}}, \quad S_2 = \frac{(\rho-1)\sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} u_{k+j} x_{t+j-1}}{B^{3/2}T^{1/2}}, \\ S_3 &= \frac{(\rho-1)\sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} u_{t+j} x_{k+j-1}}{B^{3/2}T^{1/2}}, \quad S_4 = \frac{(\rho-1)^2\sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} x_{t+j-1} x_{k+j-1}}{B^{3/2}T^{1/2}}. \end{aligned}$$

The first term is rearranged as

$$S_1 = \sum_{t=1}^{B} \sum_{j=t+1}^{t+T-B} \sum_{k=1}^{t-1} \frac{u_j u_{k+j-t}}{B^{3/2}T^{1/2}} = \sum_{j=1}^{T} \sum_{t \in \mathcal{I}_j} \sum_{k=1}^{t-1} \frac{u_j u_{j-k}}{B^{3/2}T^{1/2}} = \sum_{j=1}^{T} q_{j,T},$$

which is a sum of elements of a martingale difference array. For the second term, note that $E[\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}u_{k+j}x_0]=O(B^{3/2}T)$, which yields

$$
\begin{aligned}
E[S_2] &= \frac{-c\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sum_{m=0}^{t+j-2}\rho^m E[u_{k+j}u_{t+j-1-m}]}{B^2T}+O(B^{-1/2}) \\
&= \frac{-c\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\rho^{t-k-1}E[u_{k+j}^2]}{B^2T}+O(B^{-1/2}) \\
&= \frac{-c\sum_{j=1}^{T-B}\sum_{k=1}^{t-1}(B-k)E[u_{k+j}^2]}{B^2T}+O(B^{-1/2})=-\frac{c}{2}\int_0^1\sigma^2(r)\,\mathrm{d}r+o(1),
\end{aligned}
$$

and

$$
\begin{aligned}
E[S_2^2] &= \frac{c^2E[(\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sum_{j=1}^{T-B}\sum_{m=0}^{t+j-2}\rho^m u_{k+j}u_{t+j-1-m})^2]}{B^4T^2}+O(B^{-1}) \\
&= \frac{c^2(\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sum_{j=1}^{T-B}\rho^{t-k-1}\sigma_{k+j}^2)^2}{B^4T^2}+O(B^{-1})=\frac{c^2}{4}\Big(\int_0^1\sigma^2(r)\,\mathrm{d}r\Big)^2+o(1).
\end{aligned}
$$

Hence $Var[S_2]=o(1)$, and $S_2=-c/2\int_0^1\sigma^2(r)\,\mathrm{d}r+o_P(1)$. Let $C=\max\{C_1,C_2,1\}<\infty$, where $E[u_t^2]<C_1$, and $E[u_t^4]<C_2^2$, for all $t\in\mathbb{N}$. Furthermore, let

$$
\begin{aligned}
\tilde{S}_3 &= \frac{(\rho-1)\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}u_{t+j}(x_{k+j-1}-x_0)}{B^{3/2}T^{1/2}}, \\
\tilde{S}_4 &= \frac{(\rho-1)^2\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}(x_{t+j-1}-x_0)(x_{k+j-1}-x_0)}{B^{3/2}T^{1/2}}.
\end{aligned}
$$

Then, $E[|S_3-\tilde{S}_3|]=O(T^{-1/2})$, since $E[|T^{-1/2}\sum_{j=1}^{T-B}u_{t+j}x_0|]=O(1)$, and

$$
\begin{aligned}
E[|\tilde{S}_3|] &\le \frac{(1-\rho)\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sqrt{E[(\sum_{j=1}^{T-B}\sum_{m=0}^{k+j-2}\rho^m u_{t+j}u_{k+j-1-m})^2]}}{B^{3/2}T^{1/2}} \\
&\le \frac{(1-\rho)CB^{1/2}\sum_{m=0}^{T}\rho^m}{T^{1/2}}=O(B^{1/2}T^{-1/2}).
\end{aligned}
$$

Analogously, we have $E[|S_4-\tilde{S}_4|]=O(T^{-1/2})$, and $E[|\tilde{S}_4|]=O(B^{1/2}T^{-1/2})$, which implies that $S_3+S_4=O_P(B^{1/2}T^{-1/2})$, and the assertion follows with $\mathcal{W}_T:=-(S_1+S_2+S_3)/c$.

(b): Note that by mathematical induction on n, the identity $\sum_{t=2}^{n}\sum_{k=1}^{t-1}a_k=\sum_{k=1}^{n-1}(n-k)a_k$ holds true for any sequence $(a_t)_{t\in\mathbb{N}}$. The index set $\mathcal{I}_j$ can be expressed as

$$
\mathcal{I}_j=\begin{cases}\{t\in\mathbb{N}:\ 2\le t\le j-1\} & \text{if } j\in[1,B],\\ \{t\in\mathbb{N}:\ 2\le t\le B\} & \text{if } j\in[B+1,T-B],\\ \{t\in\mathbb{N}:\ j+B-T\le t\le B\} & \text{if } j\in[T-B+1,T].\end{cases}
$$

For $j\in[1,B]$, it follows that

$$
B^{3/2}T^{1/2}q_{j,T}=\sum_{t=2}^{j-1}\sum_{k=1}^{t-1}u_ju_{j-k}=\sum_{k=1}^{j-1}(j-1-k)u_ju_{j-k}=\sum_{k=1}^{j-2}ku_ju_{k+1}, \tag{A.2}
$$

and, analogously, if $j \in [B+1, T-B]$, we obtain

$$B^{3/2}T^{1/2}q_{j,T} = \sum_{t=2}^{B}\sum_{k=1}^{t-1} u_j u_{j-k} = \sum_{k=1}^{B}(B-k)u_{j-k} = \sum_{k=1}^{B-1} k u_j u_{j-B+k}. \tag{A.3}$$

Let $i := j + B - T$. If $j \in [T-B+1, T]$, or, equivalently, if $i \in [1, B]$, we have

$$\begin{aligned} B^{3/2}T^{1/2}q_{i,T} &= \sum_{t=i}^{B}\sum_{k=1}^{t-1} u_j u_{j-k} = \sum_{t=2}^{B}\sum_{k=1}^{t-1} u_j u_{j-k} - \sum_{t=2}^{i-1}\sum_{k=1}^{t-1} u_j u_{j-k} \\ &= \sum_{k=1}^{B}(B-k)u_j u_{j-k} - \sum_{k=1}^{i-1}(i-1-k)u_j u_{j-k} = \sum_{k=1}^{B-1} k u_j u_{j-B+k} - \sum_{k=1}^{i-2} k u_j u_{T-B+k+1}. \end{aligned} \tag{A.4}$$

Then,

$$Var\Big[\sum_{j=1}^{T} q_{j,T}\Big] = \sum_{j=B+1}^{T-B} E[q_{j,T}^2] + o(1) = \frac{1}{B^3 T}\sum_{j=B+1}^{T-B}\sum_{k=1}^{B-1} k^2 \sigma_j^2 \sigma_{j-B+k}^2 = \Theta(1),$$

and the first part of (b) has been shown. For the second part, we decompose the denominator statistic into $\mathcal{X}_{2,T} = S_5 + S_6 + S_7$, where

$$\begin{aligned} S_5 &= \frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}(\sum_{k=1}^{t-1} u_{j+k})^2}{B^2 T}, \quad S_6 = \frac{2(\rho-1)\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sum_{l=1}^{t-1} x_{j+k-1}u_{j+l}}{B^2 T}, \\ S_7 &= \frac{(\rho-1)^2\sum_{j=1}^{T-B}\sum_{t=2}^{B}(\sum_{k=1}^{t-1} x_{j+k-1})^2}{B^2 T}. \end{aligned}$$

The first term satisfies

$$E[S_5] = \frac{1}{B^2 T}\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sigma_{j+k}^2, \quad E[S_5^2] = \frac{1}{B^4 T^2}\Big(\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sigma_{j+k}^2\Big)^2 + \Theta(BT^{-1}), \tag{A.5}$$

which yields $Var[S_5] = \Theta(BT^{-1})$. Analogously to S_3 and S_4, $E[|S_6|] + E[|S_7|] = O(B^{1/2}T^{-1/2})$, which implies that $Var[\mathcal{X}_{2,T}] = \Theta(BT^{-1})$.

(c): With a constant error variance, equations (A.2)–(A.4) yield

$$B^3 T \cdot E\big[q_{j,T}^2\big] = \begin{cases} \sigma^4 \sum_{k=1}^{j-2} k^2 & \text{if } j \in [1, B], \\ \sigma^4 \sum_{k=1}^{B-1} k^2 & \text{if } j \in [B+1, T-B], \\ \sigma^4 (\sum_{k=1}^{B-1} k^2 + \sum_{k=1}^{i-2}(k^2 - 2k(B-k))) & \text{if } j \in [T-B+1, T]. \end{cases}$$

Combining all cases and applying the Gaussian summation formulas yields

$$\begin{aligned} Var[\mathcal{X}_{1,T}] &= \sum_{j=1}^{T} E[q_{j,T}^2] = \frac{\sigma^4}{B^3 T}\Big[(T-B)\sum_{k=1}^{B-1} k^2 + \sum_{j=1}^{B}\sum_{k=1}^{j-2}[4k^2 - 2Bk]\Big] \\ &= \sigma^4 \frac{(T-B)(B-1)(2B-1) - 2(B-1)(B-2)}{6B^2 T}, \end{aligned} \tag{A.6}$$

since $c = 0$. For the denominator, we have $S_6 = S_7 = 0$, since $c = 0$. Then,

$$E[\mathcal{X}_{2,T}] = E[S_5] = \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{k=1}^{B-1} (B-k)\sigma^2 = \sigma^2 \frac{(T-B)(B-1)}{2BT}, \tag{A.7}$$

and the assertion follows with equations (A.6) and (A.7).

A.4. Proof of Theorem 1

From Lemma 2, it follows that $\max_{1 \le j \le T} E[q_{j,T}^2] = o(1)$, and Jensen's and Markov's inequalities yield $\max_{1 \le j \le T} |q_{j,T}| \xrightarrow{p} 0$. Furthermore, from equation (A.3), it follows that

$$\begin{aligned}
&Var\Big[\sum_{j=1}^{T} q_{j,T}\Big] = \frac{\sum_{j=B+1}^{T-B} \sum_{k=1}^{B-1} k^2 E[u_j^2] E[u_{j-B+k}^2]}{B^3 T} + o(1) \\
&= \int_{\frac{B}{T}}^{\frac{T-B}{T}} \int_0^1 s^2 \sigma^2(r) \sigma^2(\tfrac{j-\lfloor (1-s)B \rfloor}{T}) \, \mathrm{d}s \, \mathrm{d}r + o(1) = \int_0^1 \int_0^1 s^2 \sigma^4(r) \, \mathrm{d}s \, \mathrm{d}r + o(1) \\
&= \frac{1}{3} \int_0^1 \sigma^4(r) \, \mathrm{d}r + o(1).
\end{aligned}$$

Since $\{q_{j,T}\}$ is a martingale difference array, we can apply the central limit theorem from Theorem 24.3 in Davidson (1994), which yields $\sum_{j=1}^{T} q_{j,T} / \sqrt{Var[\sum_{j=1}^{T} q_{j,T}]} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$, as $T \to \infty$. Furthermore, from Lemma 2, $E[\mathcal{X}_{1,T}] = -c/2 \int_0^1 \sigma^2(r) \, \mathrm{d}r + o(1)$, and the first statement follows from Lemma 1. For the second statement, note that, by equation (A.5),

$$\begin{aligned}
E[S_5] &= \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{k=1}^{B-1} (B-k) \sigma^2(\tfrac{j+k}{T}) = \int_0^{\frac{T-B}{T}} \int_0^1 (1-s) \sigma^2(r + s\tfrac{B}{T}) \, \mathrm{d}s \, \mathrm{d}r + o(1) \\
&= \int_0^1 \int_0^1 (1-s) \sigma^2(r) \, \mathrm{d}s \, \mathrm{d}r + o(1) = \frac{1}{2} \int_0^1 \sigma^2(r) \, \mathrm{d}r + o(1).
\end{aligned}$$

Furthermore, from Lemma 2, $Var[\mathcal{X}_{2,T}] = o(1)$, and the assertion follows by Chebyshev's inequality together with Lemma 1.

A.5. Proof of Theorem 2

Let $X_T(r) = T^{-1/2} \sum_{k=1}^{\lfloor rT \rfloor} u_k$ and $Y_T(r) = T^{-1/2} x_{\lfloor rT \rfloor}$ for $r \ge 0$. From Lemmas 1 and 2 in Cavaliere (2005), it follows that $X_T \Rightarrow \overline{\sigma} W_\eta$, where $\overline{\sigma}^2 = \int_0^1 \sigma^2(r) \, \mathrm{d}r$ denotes the average variance. For notational convenience, we set $u_0 = x_0$. Note that a Taylor expansion around 0 yields $e^{-x} = 1 - x + o(x)$, which implies that $\rho = 1 - c/\sqrt{BT} = \exp(-c/\sqrt{BT}) + o(1/\sqrt{BT})$. Then, with the continuous mapping theorem, we obtain

$$\begin{aligned}
\frac{1}{\overline{\sigma}\sqrt{T}} x_{\lfloor rT \rfloor} &= \sum_{k=0}^{\lfloor rT \rfloor} \rho^{\lfloor rT \rfloor - k} \frac{u_k}{\overline{\sigma}\sqrt{T}} = \sum_{k=0}^{\lfloor rT \rfloor} e^{-(\lfloor rT \rfloor - k)c/\sqrt{BT}} \frac{u_k}{\overline{\sigma}\sqrt{T}} + o_P(1) \\
&= \int_0^r e^{-(r-s)c/b} dX_T(s) + o_P(1) \Rightarrow \int_0^r e^{-(r-s)c/b} dW_\eta(s) = J_{c,b,\eta}(r),
\end{aligned} \tag{A.8}$$

which yields $Y_T \Rightarrow \overline{\sigma} J_{c,b,\eta}$. We rewrite

$$\Delta x_{t+j} x_{t+j-1} = \frac{\Delta x_{t+j}(x_{t+j-1} + x_{t+j} - \Delta x_{t+j})}{2}$$
$$= \frac{(x_{t+j} - x_{t+j-1})(x_{t+j} + x_{t+j-1}) - (\Delta x_{t+j})^2}{2} = \frac{x_{t+j}^2 - x_{t+j-1}^2 - (\Delta x_{t+j})^2}{2}$$

such that

$$\sum_{t=2}^{B} \Delta x_{t+j}(x_{t+j-1} - x_j) = \sum_{t=1}^{B} \frac{x_{t+j}^2 - x_{t+j-1}^2 - (\Delta x_{t+j})^2}{2} - \Delta x_{t+j} x_j$$
$$= \frac{1}{2}(x_{j+B}^2 - x_j^2) - (x_{j+B}x_j - x_j^2) - \frac{1}{2}\sum_{t=1}^{B}(\Delta x_{t+j})^2 = \frac{(x_{j+B} - x_j)^2}{2} - \frac{1}{2}\sum_{t=1}^{B}(\Delta x_{t+j})^2.$$

Then, with Lemma 1,

$$\mathcal{Y}_{1,T} = \mathcal{X}_{1,T} + o_P(1) = \frac{\sum_{j=1}^{T-B}(x_{B+j} - x_j)^2 - \sum_{j=1}^{T-B}\sum_{t=1}^{B}(\Delta x_{t+j})^2}{2B^{3/2}T^{1/2}}$$
$$= \frac{\int_0^{1-b}(Y_T(b+r) - Y_T(r))^2 \,\mathrm{d}r - \frac{1}{T^2}\sum_{j=1}^{T-B}\sum_{t=1}^{B}(\Delta x_{t+j})^2}{2b^{3/2}} + o_P(1).$$

From $\Delta x_t = u_t$, it follows that

$$E\left[\frac{1}{T^2}\sum_{j=1}^{T-B}\sum_{t=1}^{B}(\Delta x_{t+j})^2\right] = \frac{1}{T^2}\sum_{j=1}^{T-B}\sum_{t=1}^{B} E[u_{t+j}^2] = b(1-b)\int_0^1 \sigma^2(r)\,\mathrm{d}r + o(1),$$

which implies that

$$\mathcal{Y}_{1,T} = \frac{\int_0^{1-b}(Y_T(b+r) - Y_T(r))^2 \,\mathrm{d}r - b(1-b)\int_0^1 \sigma^2(r)\,\mathrm{d}r}{2b^{3/2}} + o_P(1). \tag{A.9}$$

Furthermore, Lemma 1 yields

$$\mathcal{Y}_{2,T} = \mathcal{X}_{2,T} + o_P(1) = \frac{1}{b^2}\int_0^{1-b}\int_r^{b+r}(Y_T(s) - Y_T(r))^2 \,\mathrm{d}s\,\mathrm{d}r + o_P(1). \tag{A.10}$$

The assertion follows from equation (A.8), together with the continuous mapping theorem.

A.6. Proof of Lemma 3

Since $(1-\hat{\rho}) = O_P(B^{-1/2}T^{-1/2})$ and $x_t = O_P(T^{1/2})$, the residuals satisfy

$$\hat{u}_t = y_t - \hat{\rho} y_{t-1} = \Delta y_t + (1-\hat{\rho})y_{t-1} = \Delta d_t + u_t + (\rho - 1)x_{t-1} + (1-\hat{\rho})y_{t-1} = u_t + O_P(B^{-1/2})$$

and $\overline{\hat{u}} = O_P(T^{-1/2})$. Then, for any $s \in [0,1]$,

$$\frac{1}{T}\sum_{j=1}^{\lfloor sT \rfloor}(\hat{u}_j - \overline{\hat{u}})^2 = \frac{1}{T}\sum_{j=1}^{\lfloor sT \rfloor} u_j^2 + O_P(B^{-1/2}) = \int_0^s \sigma^2(r)\,\mathrm{d}r + o_P(1), \tag{A.11}$$

and (a) follows with $s = 1$. Furthermore, by Slutsky's theorem, $\hat{\eta}(s) = \eta(s) + o_P(1)$ holds pointwise for all $s \in [0, 1]$. Then, (b) follows by Dini's theorem since both $\hat{\eta}(s)$ and $\eta(s)$ are continuous, monotone, and bounded. For (c), note that

$$\frac{1}{T-B}\sum_{j=1}^{T-B}\Big(\hat{u}_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}_{j+k}\Big)^2 = \frac{1}{T-B}\sum_{j=1}^{T-B} u_{j+t}^2 + O_P(B^{-1/2}), \tag{A.12}$$

for any $t = 1, \ldots, B$. Equations (A.11) and (A.12) yield

$$\frac{1}{(T-B)B}\sum_{j=1}^{T-B}\sum_{t=1}^{B}\Big(\hat{u}_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}_{j+k}\Big)^2 = \int_0^1 \sigma^2(r)\,\mathrm{d}r + o_P(1),$$
$$\frac{1}{(T-B)B}\sum_{j=1}^{T-B}\sum_{t=1}^{B}(\hat{u}_{j+1} - \overline{\hat{u}})^2\Big(\hat{u}_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}_{j+k}\Big)^2 = \int_0^1 \sigma^4(r)\,\mathrm{d}r + o_P(1),$$

as $B, T \to \infty$ and $B/T \to 0$, and the result follows by Slutsky's theorem.

A.7. Proof of Theorem 3

Note that $v_T \to \sqrt{2/3}$, and $\hat{\kappa} v_T \sqrt{\mathcal{Y}_{2,T}} \xrightarrow{p} \sqrt{\int_0^1 \sigma^4(r)\,\mathrm{d}r/3}$, which follows from Theorem 1 and Lemma 3. Then, (a) follows together with Slutsky's theorem. For (b), let $\tilde{x}_{\lfloor rT \rfloor} = x_{\lfloor \hat{\eta}^{-1}(r)T \rfloor}$ and $\tilde{u}_{\lfloor rT \rfloor} = u_{\lfloor \hat{\eta}^{-1}(r)T \rfloor}$. Furthermore, let $\widetilde{X}_T(r) = T^{-1/2}\sum_{k=1}^{\lfloor rT \rfloor}\tilde{u}_k$ and $\widetilde{Y}_T(r) = T^{-1/2}\tilde{x}_{\lfloor rT \rfloor}$. Theorem 1 in Cavaliere and Taylor (2008b) states that $\widetilde{X}_T \Rightarrow \overline{\sigma} W$, where $\overline{\sigma}^2 = \int_0^1 \sigma^2(r)\,\mathrm{d}r$, and, analogously to (A.8), it follows that $\widetilde{Y}_T \Rightarrow J_{c,b}$. Following equations (A.9) and (A.10), we obtain

$$\tau\text{-FB} = \frac{(\int_0^{1-b}(\widetilde{Y}(b+r) - \widetilde{Y}(r))^2\,\mathrm{d}r - b(1-b)\overline{\sigma}^2)/(2b^{3/2})}{\sqrt{\overline{\sigma}\int_0^{1-b}\int_r^{b+r}(\widetilde{Y}(s) - \widetilde{Y}(r))^2\,\mathrm{d}s\,\mathrm{d}r/b^2}} + o_P(1),$$

and the assertion follows with the continuous mapping theorem and Slutsky's theorem.

A.8. Proof of Lemma 4

The coefficients $\hat{\theta}_1, \ldots, \hat{\theta}_p$ can be represented as $\hat{\theta}_i = e_i'\hat{\gamma}$, $i = 1, \ldots, p$, where e_i is the i-th unit vector, and $\hat{\gamma} = (\sum_{t=p+1}^{T} z_t z_t')^{-1}\sum_{t=p+1}^{T} z_t \Delta y_t$ with $z_t = (\Delta y_{t-1}, \ldots, \Delta y_{t-p}, T^{-1/2}y_{t-1})'$. From equation (3), we have

$$\Delta x_t = w_t'\beta + \epsilon_{p,t}, \tag{A.13}$$

where $w_t = (u_{t-1}, \ldots, u_{t-p}, T^{-1/2}x_{t-1})'$, and $\beta = (\theta_1, \ldots, \theta_p, T^{1/2}\phi)'$. Note that $\theta_i = e_i'\beta$ for all $i = 1, \ldots, p$. Furthermore, we have

$$z_t = f_t + \phi g_t + w_t, \tag{A.14}$$

where $f_t = (\Delta d_{t-1}, \dots, \Delta d_{t-p}, T^{-1/2} d_{t-1})'$, and $g_t = (x_{t-1}, \dots, x_{t-p}, 0)'$. Equations (A.13) and (A.14) yield

$$\Delta y_t = \Delta d_t + \Delta x_t = \Delta d_t + w_t' \beta + \epsilon_{p,t} = z_t' \beta + a_t,$$

where $a_t = \epsilon_t + \sum_{i=p+1}^{\infty} \theta_i u_{t-i} + \Delta d_t - f_t' \beta - \phi g_t' \beta$. Then, $\hat{\gamma} = \beta + (\sum_{t=p+1}^{T} z_t z_t')^{-1} \sum_{t=p+1}^{T} z_t a_t$, and

$$\hat{\theta}_i - \theta_i = e_i' \Big(\sum_{t=p+1}^{T} z_t z_t' \Big)^{-1} \sum_{t=p+1}^{T} z_t a_t, \qquad i = 1, \dots, p.$$

Note that $\|e_i'(T^{-1} \sum_{t=p+1}^{T} z_t z_t')^{-1}\| = O_P(1)$ and $\|T^{-1} \sum_{t=p+1}^{T} z_t\| = O_P(1)$, where $\| \cdot \|$ denotes some suitable vector norm. It thus remains to show that

$$\Big\| \frac{1}{T} \sum_{t=p+1}^{T} z_t a_t \Big\| = O_P(B^{-1/2}).$$

From $E[z_t \epsilon_t] = 0$ and the central limit theorem, $\|T^{-1} \sum_{t=p+1}^{T} z_t \epsilon_t\| = O_P(T^{-1/2})$. Then,

$$\begin{aligned} E\big[\|z_t(a_t - \epsilon_t)\|\big] \leq & \sum_{i=p+1}^{\infty} |\theta_i| \cdot E\big[\|z_t u_{t-i}\|\big] + E\big[\|z_t\|\big] \Big(|\Delta d_t| + \sum_{i=1}^{p} |\theta_i| \cdot |\Delta d_{t-i}| + \phi d_{t-1} \Big) \\ & + \phi \sum_{i=1}^{p} |\theta_i| \cdot E\big[\|z_t x_{t-i}\|\big], \end{aligned}$$

which is $O(B^{-1/2})$, and the assertion follows with Markov's inequality.

A.9. Proof of Lemma 5

The proof is split in two parts. In the first part, we show that $\mathcal{Y}_{1,T}^* - \hat{\mathcal{Y}}_{1,T}^* = O_P(pB^{-1/2})$ and $\mathcal{Y}_{2,T}^* - \hat{\mathcal{Y}}_{2,T}^* = O_P(pT^{-1/2})$, where

$$\mathcal{Y}_{1,T}^* = \frac{1}{B^{3/2} T^{1/2}} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \Delta y_{t+j}^* (y_{t+j-1}^* - y_j^*), \quad \mathcal{Y}_{2,T}^* = \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{t=2}^{B} (y_{t+j-1}^* - y_j^*)^2,$$

and $y_t^* = \theta(L) y_t$. We have

$$\begin{aligned} \mathcal{Y}_{1,T}^* - \hat{\mathcal{Y}}_{1,T}^* &= \frac{1}{B^{3/2} T^{1/2}} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \big(\Delta y_{t+j}^* \Delta y_{k+j}^* - \Delta \hat{y}_{t+j}^* \Delta \hat{y}_{k+j}^* \big), \\ \mathcal{Y}_{2,T}^* - \hat{\mathcal{Y}}_{2,T}^* &= \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \Big[\big(\Delta y_{k+j}^*\big)^2 - \big(\Delta \hat{y}_{k+j}\big)^2 \Big]. \end{aligned}$$

Let $\hat{\theta}(z) = 1 - \sum_{i=1}^{p} \hat{\theta}_i z^i$, which yields $\hat{y}_t^* = \hat{\theta}(L) y_t$. Furthermore, let $\theta^*(z) = \sum_{i=1}^{\infty} \theta_i^* z^i$, where

$$\theta_i^* = \begin{cases} \theta_i - \hat{\theta}_i & \text{for } i \le p, \\ \theta_i & \text{for } i > p, \end{cases}$$

which yields $\theta^*(L) y_t = y_t^* - \hat{y}_t^*$. Let, for notational convenience, $\theta_0 = \hat{\theta}_0 = -1$. Then,

$$\begin{aligned} &\Delta y_{t+j}^* \Delta y_{k+j}^* - \Delta \hat{y}_{t+j}^* \Delta \hat{y}_{k+j}^* = (\Delta y_{t+j}^* - \Delta \hat{y}_{t+j}^*) \Delta y_{k+j}^* + \Delta \hat{y}_{t+j}^* (\Delta y_{k+j}^* - \Delta \hat{y}_{k+j}) \\ &= \sum_{i=1}^{\infty} \theta_i^* \left[\Delta y_{t+j-i} \Delta y_{k+j}^* + \Delta \hat{y}_{t+j}^* \Delta y_{k+j-i} \right] \\ &= -\sum_{i=1}^{\infty} \theta_i^* \Big(\sum_{m=0}^{\infty} \theta_m \Delta y_{t+j-i} \Delta y_{k+j-m} + \sum_{m=0}^{p} \hat{\theta}_m \Delta y_{t+j-m} \Delta y_{k+j-i} \Big), \end{aligned}$$

and

$$\begin{aligned} &\big(\Delta y_{k+j}^*\big)^2 - \big(\Delta \hat{y}_{k+j}\big)^2 = \big(\Delta y_{k+j}^* - \Delta \hat{y}_{k+j}^*\big)^2 + 2\big(\Delta y_{k+j}^* - \Delta \hat{y}_{k+j}^*\big) \Delta \hat{y}_{j+j}^* \\ &= \sum_{i,m=1}^{\infty} \theta_i^* \theta_m^* \Delta y_{k+j-i} \Delta y_{k+j-m} - 2 \sum_{i=1}^{\infty} \sum_{m=0}^{p} \theta_i^* \hat{\theta}_m \Delta y_{k+j-i} \Delta y_{k+j-m}. \end{aligned}$$

We decompose $\mathcal{Y}_{1,T}^* - \hat{\mathcal{Y}}_{1,T}^* = S_1 + S_2$ and $\mathcal{Y}_{2,T}^* - \hat{\mathcal{Y}}_{2,T}^* = S_3 + S_4$, where

$$S_1 = -\frac{1}{B^{3/2}} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sum_{i=1}^{\infty} \sum_{m=0}^{\infty} \theta_i^* \theta_m A_{t-i,k-m}, \quad S_2 = -\frac{1}{B^{3/2}} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sum_{i=1}^{\infty} \sum_{m=0}^{p} \theta_i^* \hat{\theta}_m A_{t-m,k-i},$$
$$S_3 = \frac{1}{B^2 T^{1/2}} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sum_{i,m=1}^{\infty} \theta_i^* \theta_m^* A_{k-i,k-m}, \quad S_4 = -\frac{2}{B^2 T^{1/2}} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sum_{i=1}^{\infty} \sum_{m=0}^{p} \theta_i^* \hat{\theta}_m A_{k-i,k-m},$$

with $A_{r,s} = T^{-1/2} \sum_{j=1}^{T-B} \Delta y_{r+j} \Delta y_{s+j}$. Note that

$$\begin{aligned} A_{r,s} &= \frac{1}{T^{1/2}} \sum_{j=1}^{T-B} \big(\Delta d_{r+j} + \phi x_{r+j-1} + u_{r+j}\big)\big(\Delta d_{s+j} + \phi x_{s+j-1} + u_{s+j}\big) \\ &= \frac{1}{T^{1/2}} \sum_{j=1}^{T-B} \big(\phi x_{r+j-1} u_{s+j} + \phi u_{r+j} x_{s+j-1} + \phi^2 x_{r+j-1} x_{s+j-1}\big) + O_P(T^{-1/2}) \\ &= O_P(B^{-1/2}), \end{aligned}$$

and, by Lemma 4, $\sum_{i=1}^{p} \theta_i^* = O_P(pB^{-1/2})$. It then follows that $S_1 + S_2 = O_P(pB^{-1/2})$, and $S_3 + S_4 = O_P(pB^{-1}T^{-1/2})$. Hence, the first part of the proof is completed.

For the second part of the proof, note that $d_t^* = y_t^* - x_t^* = \theta(L) d_t = d^*(t/T)$ is Lipschitz continuous, since

$$|d^*(t/T) - d^*(s/T)| = \theta(L)(d_t - d_s) \le C\Big(1 - \sum_{i=1}^{\infty} |\theta_i|\Big) \Big| \frac{t-s}{T} \Big|,$$

where $C(1 - \sum_{i=1}^{\infty} |\theta_i|) < \infty$. Hence, under Assumption 3, the transformed statistics $\mathcal{Y}_{i,T}^*$ and

$\mathcal{X}^*_{i,T}$, $i = 1, 2$, have the same properties as $\mathcal{Y}_{i,T}$ and $\mathcal{X}_{i,T}$, $i = 1, 2$, under Assumption 2. Thus, Lemma 1 yields $\mathcal{Y}^*_{1,T} - \mathcal{X}^*_{1,T} = O_P(B^{-1/2})$ and $\mathcal{Y}^*_{2,T} - \mathcal{X}^*_{2,T} = O_P(T^{-1/2})$. Finally, the triangle inequality implies that $\hat{\mathcal{Y}}^*_{1,T} - \mathcal{X}^*_{1,T} = O_P(pB^{-1/2})$ and $\hat{\mathcal{Y}}^*_{2,T} - \mathcal{X}^*_{2,T} = O_P(pB^{-1/2})$.

A.10. Proof of Theorem 4

Let $\hat{\theta}(L)$ and $\theta^*(L)$ be defined as in the proof of Lemma 5, and let $\hat{d}^*_t = \hat{\theta}(L)d_t$ and $\hat{x}^*_t = \hat{\theta}(L)x_t$. Analogously to the proof of Lemma 3, we have

$$\hat{u}^*_t = \hat{y}^*_t - \hat{\rho}^* \hat{y}_{t-1} = \Delta \hat{y}^*_t + (1 - \hat{\rho}^*)\hat{y}^*_{t-1} = \Delta \hat{x}^*_t + O(B^{-1/2}) = \epsilon_t + o_P(1).$$

The consistencies of $\hat{\sigma}^{*2}$, $\hat{\kappa}^{*2}$, and $\hat{\eta}(s)^*$ then follow from the fact that

$$\frac{1}{T}\sum_{j=1}^{\lfloor sT \rfloor}(\hat{u}^*_j - \overline{\hat{u}^*})^2 = \frac{1}{T}\sum_{j=1}^{\lfloor sT \rfloor}\epsilon_j^2 + o_P(1) = \int_0^s \sigma^2(r)\,\mathrm{d}r + o_P(1), \quad s \in [0,1],$$

and

$$\begin{aligned}
\frac{1}{T-B}\sum_{j=1}^{T-B}\Big(\hat{u}^*_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}^*_{j+k}\Big)^2 &= \frac{1}{T-B}\sum_{j=1}^{T-B}\epsilon^2_{j+t} + o_P(1),\\
\frac{1}{(T-B)B}\sum_{j=1}^{T-B}\sum_{t=1}^{B}\Big(\hat{u}^*_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}^*_{j+k}\Big)^2 &= \int_0^1 \sigma^2(r)\,\mathrm{d}r + o_P(1),\\
\frac{1}{(T-B)B}\sum_{j=1}^{T-B}\sum_{t=1}^{B}(\hat{u}^*_{j+1} - \overline{\hat{u}^*})^2\Big(\hat{u}^*_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}^*_{j+k}\Big)^2 &= \int_0^1 \sigma^4(r)\,\mathrm{d}r + o_P(1),
\end{aligned}$$

where the last two equations hold true as $B/T \to 0$, analogously to Lemma 3.

Finally, since the pre-whitened numerator and denominator statistics $(\mathcal{X}^*_{1,T}, \mathcal{X}^*_{2,T})$ under Assumption 3 have the same properties as $(\mathcal{X}_{1,T}, \mathcal{X}_{2,T})$ under Assumption 2, the assertion follows with Lemma 5 and the proof of Theorem 3.