Unit Root Testing with Slowly Varying Trends

Sven Otto*

April 5, 2019

Abstract: A unit root test is proposed for time series with a nonparametric deterministic trend component. It is shown that the pooled OLS estimator of overlapping blocks filters out the trend component asymptotically. Under both fixed-b and small-b block asymptotics, the limiting distribution of the t-statistic for the unit root hypothesis is derived. Furthermore, a nuisance parameter correction provides 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: 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 were first studied by Perron (1989) and Rappoport and Reichlin (1989). Under a structural change with a known breakpoint, these authors' tests reject the unit root hypothesis for many macroeconomic series. Christiano (1992) has 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), propose unit root tests for this framework. Structural changes in innovation variances have been studied by Hamori and Tokihisa (1997), Kim et al. (2002), and Cavaliere (2005), while Cavaliere et al. (2011) consider unit root testing under broken trends together with nonstationary volatility. Leybourne et al. (1998), Kapetanios et al. (2003), and Kılıç (2011) allow for exponential smooth transitions from one trend regime to another. Bierens (1997) approximates a nonlinear mean function with Chebyshev polynomials. Furthermore, Enders and Lee (2012) propose a Fourier series approximation of the trend, which is an approach that can be used when the exact form and date of structural changes are unknown.

Dickey-Fuller 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) estimate 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 approximation, in a small block, a moderately varying trend can be approximated quite closely by a constant. To exploit this fact, 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 (2001) 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 nonparametric 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 Brownian motions 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 homoskedastic errors is analyzed together with the asymptotic behavior of the pooled least squares estimator in the presence of a general nonparametric trend model. For both fixed-b and small-b block asymptotics, the limiting distributions are derived under both the unit root hypothesis and under local alternatives. Section 3 considers pseudo t-statistics for unit root testing, while Section 4 demonstrates that, under heteroskedasticity, nuisance parameters appear in the limiting distributions. The estimation of these parameters is discussed, and heteroskedasticity-robust test statistics are provided. In Section 5, a pre-whitening procedure is proposed in order to account for short-run dynamics, while Section 6 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. In Section 7, these tests are applied to the monthly inflation rates of 25 countries. The results provide some evidence in favor of inflation rates being trend-stationary around a slowly varying deterministic component. Finally, Section 8 presents the conclusion.

2. The pooled estimator

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

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

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 simplified assumptions on the error term, which are relaxed in the subsequent sections:

Assumption 2 (i.i.d. errors)**.** *The process* $\{u_t\}_{t\in\mathbb{N}}$ *is independently distributed, where* $E[u_t] = 0$*,* $E[u_t^2] = \sigma^2$ *and* $E[u_t^4] < \infty$*.*

The principal approach to dealing with a nonparametric, 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 approximated by the first observation in each block. Thereafter, by pooling the $T - B$ individual block regressions, we obtain the following regression equation:

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

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 for a zero trend component 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 trend component can be ignored asymptotically. All asymptotic results are jointly derived for B and $T \to \infty$. 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, $|\mathcal{Y}_{1,T} - \mathcal{X}_{1,T}| = o_P(1)$, and $|\mathcal{Y}_{2,T} - \mathcal{X}_{2,T}| = o_P(1)$ as $B, T \to \infty$.*

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 the following 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 $B, T \to \infty$, it follows that*

(a) $\mathcal{X}_{1,T} = \sum_{j=1}^{T} q_{j,T} + \mathcal{C}_{1,T}$, where $\{q_{j,T}\}_{j \leq T, T \in \mathbb{N}}$ is a martingale difference array with $Var[\sum_{j=1}^{T} q_{j,T}] = \sigma^4 \frac{(B-1)((T-B)(2B-1)-2(B-2))}{6B^2T}$ and $\mathcal{C}_{1,T} = -c \cdot (\frac{\sigma^2}{2} + o_P(1))$, and

(b) $E[\mathcal{X}_{2,T}] = \sigma^2 \frac{(T-B)(B-1)}{2BT} + c \cdot O(B^{1/2}T^{-1/2})$ and $Var[\mathcal{X}_{2,T}] = O(BT^{-1})$.

Lemmas 1 and 2 imply that $\mathcal{Y}_{2,T} = O_P(1)$ if $B = O(T)$, whereas $\mathcal{Y}_{2,T} = o_P(1)$ if $B = o(T)$. This suggests distinguishing between these 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. From Lemmas 1 and 2, it follows that $Var[\mathcal{Y}_{1,T}] \to \sigma^4/3$ and $E[\mathcal{Y}_{2,T}] \to \sigma^2/2$ as $B/T \to 0$ and $B, T \to \infty$. Together with the central limit theorem for martingale difference arrays, the following asymptotic result can be stated:

Theorem 1 (small-b asymptotics)**.** *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, $\mathcal{Y}_{1,T} \xrightarrow{d} \mathcal{N}(-\frac{c\sigma^2}{2}, \frac{\sigma^4}{3})$ and $\mathcal{Y}_{2,T} \xrightarrow{p} \frac{\sigma^2}{2}$.*

[1] Note that the terminology "fixed-b and small-b asymptotics" has also been used in the context of long-run variance estimation. Whereas Kiefer and Vogelsang (2005) use 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.

As a direct consequence, the pooled estimator is asymptotically normally distributed under small-b asymptotics. Together with Slutsky's theorem, it follows that

$$\sqrt{BT}(\hat{\rho}-1) \xrightarrow{d} \mathcal{N}(0, 4/3)$$

under the unit root hypothesis, which is given by $\rho = 1$ or equivalently by $c = 0$. Under fixed-b asymptotics, the numerator and denominator statistics can be represented as a partial sum process of the innovations. The functional central limit theorem then yields the following asymptotic result:

Theorem 2 (fixed-b asymptotics). *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*

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

where $J_c(r)$ is an Ornstein-Uhlenbeck process.

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

$$\sqrt{BT}(\hat{\rho}-1) \xrightarrow{d} \frac{b^{1/2} \int_0^{1-b} \left(W(b+r) - W(r)\right)^2 \, \mathrm{d}r + b^{3/2}(1-b)}{2 \int_0^{1-b} \int_r^{b+r} \left(W(s) - W(r)\right)^2 \, \mathrm{d}s \, \mathrm{d}r}$$

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.

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 numerator of the t-statistic can be represented as $\hat{\rho} - 1 = \mathcal{Y}_{1,T}(\mathcal{Y}_{2,T}\sqrt{BT})^{-1}$. The standard error is obtained from the conditional variance of $\hat{\rho}$. Let

$$s_{\hat{\rho}}^2 = \hat{\sigma}^2 \Big(\sum_{j=1}^{T-B} \sum_{t=2}^{B} (y_{t+j-1} - y_j)^2 \Big)^{-1} = \frac{\hat{\sigma}^2}{\mathcal{Y}_{2,T} B^2 T},$$

where $\hat{\sigma}^2$ denotes some consistent estimator for the error variance σ^2. The conventional t-statistic is then represented as $(\hat{\rho}-1)/s_{\hat{\rho}} = \sqrt{B}\mathcal{Y}_{1,T}/\sqrt{\hat{\sigma}^2\mathcal{Y}_{2,T}}$ and diverges in probability under H_0. Accordingly, we consider a pseudo t-statistic of the form

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

which is $O_P(1)$ under both small-b and fixed-b asymptotics. We consider the residuals $\hat{u}_t = y_t - \hat{\rho}y_{t-1}$ for $t = 2, \ldots, T$ and their sample mean $\overline{\hat{u}} = T^{-1}\sum_{j=1}^{T}\hat{u}_j$. For the error variance estimation, we distinguish between fixed-b and small-b block asymptotics and define

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

The following consistency result can then be obtained:

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) Let $B/T \to 0$ as $B, T \to \infty$. Then, $\hat{\sigma}^2_{sb} \xrightarrow{p} \sigma^2$.

(b) Let $0 < b < 1$, and let $B/T \to b$ as $B, T \to \infty$. Then, $\hat{\sigma}^2_{fb} \xrightarrow{p} \sigma^2$.

In what follows, the pseudo t-tests are defined. For the small-b pseudo t-statistic, we scale $\mathcal{Y}_{1,T}$ and $\mathcal{Y}_{2,T}$ by their finite sample variance and their expectation from Lemma 2 in order to avoid small-sample size distortions. Let

$$v_T^2 = \frac{(T-B)(2B-1) - 2(B-2)}{3B(T-B)},$$

which is equal to $\sigma^{-2}Var[\mathcal{X}_{1,T}]/E[\mathcal{X}_{2,T}]$ under the unit root hypothesis. The small-b pseudo t-statistic is then defined as

$$\tau\text{-SB} = \frac{\mathcal{Y}_{1,T}}{v_T\hat{\sigma}_{\text{sb}}\sqrt{\mathcal{Y}_{2,T}}} = \frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}\Delta y_{t+j}(y_{t+j-1} - y_j)}{\hat{\sigma}_{\text{sb}}\sqrt{\frac{(T-B)(2B-1)-2(B-2)}{3(T-B)}\sum_{j=1}^{T-B}\sum_{t=2}^{B}(y_{t+j-1} - y_j)^2}}.$$

For the fixed-b statistic, we consider the unscaled versions and define

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

The unit root hypothesis is rejected in favor of stationarity if the test statistic falls below the α-quantile of the limiting distribution under H_0, where α is the significance level. From Theorem 1 and 2 and Lemma 3, together with the continuous mapping theorem and Slutsky's theorem, the following limiting result can be stated:

Corollary 1. *Let $\rho = 1$, 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, τ-SB $\xrightarrow{d} \mathcal{N}(0,1)$.

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

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

where $W(r)$ is a standard Brownian motion.

For τ-SB we can rely on standard normal quantiles as critical values. However, the limiting distribution of τ-FB is nonstandard. 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 τ-FB 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 Theorems 1 and 2, it follows that the tests have power against alternatives of the form $\rho = c/\sqrt{BT}$, where $c > 0$.

4. Testing under heteroskedasticity

While stationary time-varying conditional variances such as ARCH and GARCH processes do not affect unit root testing, Hamori and Tokihisa (1997) show that permanent changes in volatility, in contrast, dramatically alter the limiting distributions of unit root tests. Kim et al. (2002) report that an abrupt break in the innovation variance can produce spurious rejections, while Cavaliere (2005) shows that nonstationary volatility induces 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 $\eta(s) = (\int_0^1 \sigma(r)^2 dr)^{-1} \int_0^s \sigma(r)^2 dr$, the transformed process is defined as $W_\eta(r) = W(\eta(r))$, where $0 \leq r \leq 1$. In what follows, we relax Assumption 2 and allow for heteroskedastic errors.

Assumption 3 (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.*

Notice that the approximation result of Lemma 1 is not affected by Assumption 3 and can be formulated under heteroskedasticity as follows:

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,* $|\mathcal{Y}_{1,T} - \mathcal{X}_{1,T}| = o_P(1)$, *and* $|\mathcal{Y}_{2,T} - \mathcal{X}_{2,T}| = o_P(1)$ *as* $B, T \to \infty$.

However, nuisance parameters then appear in the limiting distributions of the numerator and denominator statistics.

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

(a) Let $B/T \to 0$ *as* $B, T \to \infty$. *Then,* $\mathcal{Y}_{1,T} \xrightarrow{d} \mathcal{N}(-\frac{c}{2}\int_0^1 \sigma^2(r)\,\mathrm{d}r, \frac{1}{3}\int_0^1 \sigma^4(r)\,\mathrm{d}r)$, *and* $\mathcal{Y}_{2,T} \xrightarrow{p} \frac{1}{2}\int_0^1 \sigma^2(r)\,\mathrm{d}r$.

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

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

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

In order to correct for the additional nuisance terms, we consider the following estimators. Let

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

and let

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

where $s \in [0,1]$. The robust small-b statistic is then defined as

$$\tau\text{-SB}^{\mathrm{H}} = \frac{\mathcal{Y}_{1,T}}{\upsilon_T \hat{\kappa}\hat{\sigma}_{\mathrm{sb}}^{-1}\sqrt{\mathcal{Y}_{2,T}}} = \frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}\Delta y_{t+j}(y_{t+j-1} - y_j)}{\hat{\kappa}\sigma_{\mathrm{sb}}^{-1}\sqrt{\frac{(T-B)(2B-1)-2(B-2)}{3(T-B)}\sum_{j=1}^{T-B}\sum_{t=2}^{B}(y_{t+j-1} - y_j)^2}}.$$

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) show that the time-transformation in the Gaussian limiting processes can be inverted. Accordingly, we fix some auxiliary sample size $\widetilde{T} \geq T$ and consider the time-transformed series $\tilde{y}_t = y_{\lfloor \hat{\eta}^{-1}(t/\widetilde{T})\widetilde{T} \rfloor}$ for $t = 1, \ldots, \widetilde{T}$, where $\hat{\eta}^{-1}(s)$ is the inverse function of $\hat{\eta}(s)$. In practice, the observed time series is transformed in such a way that copies of adjacent observations between the sample points are inserted in highly volatilie periods. Since $\widetilde{T}$ can be set arbitrarily high, we do not need to discard any observations. We replace the original series in the test statistic by the time-transformed series and define

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

which yields the fixed-b heteroskedasticity-robust statistic

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

Theorem 4. *Let $\rho = 1$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 3.*

(a) Let $B/T \to 0$ as $B, T \to \infty$. Then, $\hat{\sigma}_{sb}^2 \xrightarrow{p} \int_0^1 \sigma^2(r)\,\mathrm{d}r$, $\hat{\kappa}^2 \xrightarrow{p} \int_0^1 \sigma^4(r)\,\mathrm{d}r$, and τ-$SB^H \xrightarrow{d} \mathcal{N}(0,1)$.

(b) Let $0 < b < 1$, and let $B/\widetilde{T} \to b$ as $B, \widetilde{T} \to \infty$. Then, $\hat{\sigma}_{fb}^2 \xrightarrow{p} \int_0^1 \sigma^2(r)\,\mathrm{d}r$, $\hat{\eta}(s) \xrightarrow{p} \eta(s)$ uniformly for all $s \in [0,1]$, and

$$\tau\text{-}FB^H \xrightarrow{d} \frac{\int_0^{1-b}(W(b+r) - W(r))^2\,\mathrm{d}r + b(1-b)}{2\sqrt{b\int_0^{1-b}\int_r^{b+r}(W(s) - W(r))^2\,\mathrm{d}s\,\mathrm{d}r}}.$$

The limiting distributions under the unit root hypothesis of the heteroskedasticity-robust test statistics coincide with those obtained in Section 3 under homoskedasticity. Hence, the critical values from those tests can be retained. For τ-SB$^{\mathrm{H}}$, we consider standard normal quantiles, and, for τ-FB$^{\mathrm{H}}$, we can apply the values from Table 1.

5. 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 4 (serially correlated errors)**.** *The process $\{u_t\}_{t\in\mathbb{N}}$ possesses the stationary AR(p) representation $u_t = \sum_{i=1}^{p}\theta_i u_{t-i} + \epsilon_t$. 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. The lag order p satisfies $T^{-1/4}p\to 0$.*

In the fashion of Said and Dickey (1984), we allow the lag order to grow with the sample size. Asymptotically, this allows for fairly general forms of serially correlated errors, such as stationary and invertible ARMA processes. In order to correct for the effect of short-run dynamics, we follow Breitung and Das (2005) and consider the pre-whitened series $y_t^* = y_t - \sum_{i=1}^{p}\theta_i y_{t-i}$. The series decomposes into $y_t^* = d_t^* + x_t^*$, where the deterministic and the stochastic parts are given by $d_t^* = d_t - \sum_{i=1}^{p}\theta_i d_{t-i}$ and $x_t^* = x_t - \sum_{i=1}^{p}\theta_i x_{t-i}$ respectively. Note that $x_t^* - \rho x_{t-1}^* = \epsilon_t$, where ϵ_t satisfies the same conditions as u_t under Assumption 3. 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 then coincide with those presented in the previous sections.

Since the autoregressive parameters of the error process are unknown, they need to be estimated. We augment the regression equation with lagged values of the differenced series, such that

$$\Delta y_t = \varphi y_{t-1} + \sum_{i=1}^{p}\beta_i \Delta y_{t-i} + e_t, \tag{1}$$

for $t=p+1,\ldots,T$, where e_t is a mean-zero error term. Let $\hat{\varphi}$ and $\hat{\beta}_1,\ldots,\hat{\beta}_p$ denote the OLS estimators of the parameters. In the following, we show that $(\hat{\beta}_1, \dots, \hat{\beta}_p)'$ is consistent for $(\theta_1, \dots, \theta_p)'$ under the unit root hypothesis:

Lemma 5. *Let $\rho = 1$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 4. Then $\max_{1 \le i \le p} p|\hat{\beta}_i - \theta_i| = o_P(1)$ as $B, T \to \infty$.*

The estimated pre-whitened series is defined as $\hat{y}_t^* = y_t - \sum_{i=1}^p \hat{\beta}_i y_{t-i}$, 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 6. *Let $\rho = 1$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 4. Then, $|\hat{\mathcal{Y}}_{1,T}^* - \mathcal{X}_{1,T}^*| = o_P(1)$, and $|\hat{\mathcal{Y}}_{2,T}^* - \mathcal{X}_{2,T}^*| = o_P(1)$ as $B, T \to \infty$.*

The estimators $\hat{\sigma}_{\mathrm{sb}}^{*2}$, $\hat{\sigma}_{\mathrm{fb}}^{*2}$, $\hat{\kappa}^{*2}$, and $\hat{\eta}^*(s)$ are defined as their counterparts in Sections 3 and 4, except that the residuals are now defined as $\hat{u}_t = \hat{y}_t^* - \hat{\rho}^* \hat{y}_{t-1}^*$, where $\hat{\rho}^*$ is given by $\sqrt{BT}(\hat{\rho}^* - 1) = \hat{\mathcal{Y}}_{1,T}^* / \hat{\mathcal{Y}}_{2,T}^*$. Analogously to Section 4, we consider the time-transformed pre-whitened series $\tilde{y}_t^* = \hat{y}^*_{\lfloor \hat{\eta}^{-1}(t/\widetilde{T})\widetilde{T} \rfloor}$ for $t = 1, \dots, \widetilde{T}$, where $\widetilde{T} \ge T$, and we define

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

The pre-whitened versions of the test statistics are then given by

$$\tau\text{-SB}^{\mathrm{PW}} = \frac{\hat{\mathcal{Y}}_{1,T}^*}{v_T \hat{\sigma}_{\mathrm{sb}}^* \sqrt{\hat{\mathcal{Y}}_{2,T}^*}}, \quad \tau\text{-SB}^{\text{H-PW}} = \frac{\hat{\mathcal{Y}}_{1,T}^*}{v_T \hat{\kappa}^* \hat{\sigma}_{\mathrm{sb}}^{*-1} \sqrt{\hat{\mathcal{Y}}_{2,T}^*}},$$
$$\tau\text{-FB}^{\mathrm{PW}} = \frac{\hat{\mathcal{Y}}_{1,T}^*}{\hat{\sigma}_{\mathrm{fb}}^* \sqrt{\hat{\mathcal{Y}}_{2,T}^*}}, \quad \tau\text{-FB}^{\text{H-PW}} = \frac{\widetilde{\mathcal{Y}}_{1,T}^*}{\hat{\sigma}_{\mathrm{fb}}^* \sqrt{\widetilde{\mathcal{Y}}_{2,T}^*}}.$$

Theorem 5. *Let $\rho = 1$, let d_t satisfy Assumption 1, and let u_t satisfy Assumption 4.*

(a) Let $B/T \to 0$ as $B, T \to \infty$. Then, $\hat{\sigma}_{sb}^{*2} \xrightarrow{p} \int_0^1 \sigma^2(r)\,\mathrm{d}r$, $\hat{\kappa}^{*2} \xrightarrow{p} \int_0^1 \sigma^4(r)\,\mathrm{d}r$, and τ-SB$^{H\text{-}PW} \xrightarrow{d} \mathcal{N}(0,1)$. Furthermore, τ-SB$^{PW} \xrightarrow{d} \mathcal{N}(0,1)$ if $\sigma_t^2 = \sigma^2$ for all t.

(b) Let $0 < b < 1$, and let $B/\widetilde{T} \to b$ as $B, \widetilde{T} \to \infty$. Then, $\hat{\sigma}_{fb}^{*2} \xrightarrow{p} \int_0^1 \sigma^2(r)\,\mathrm{d}r$,

$\hat{\eta}(s) \xrightarrow{p} \eta(s)$ *uniformly for all* $s \in [0,1]$*, and*

$$\tau\text{-}FB^{H\text{-}PW} \xrightarrow{d} \frac{\int_0^{1-b} \left(W(b+r) - W(r)\right)^2 \,\mathrm{d}r + b(1-b)}{2\sqrt{b \int_0^{1-b} \int_r^{b+r} \left(W(s) - W(r)\right)^2 \,\mathrm{d}s\,\mathrm{d}r}}.$$

Furthermore, if $\sigma_t^2 = \sigma^2$ *for all* $t \in \mathbb{N}$*, then*

$$\tau\text{-}FB^{PW} \xrightarrow{d} \frac{\int_0^{1-b} \left(W(b+r) - W(r)\right)^2 \,\mathrm{d}r + b(1-b)}{2\sqrt{b \int_0^{1-b} \int_r^{b+r} \left(W(s) - W(r)\right)^2 \,\mathrm{d}s\,\mathrm{d}r}}.$$

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). The maximum lag order is inspired by the rule of thumb provided by Schwert (1989) and is fixed as $p^* = \lfloor 4 \cdot (T/100)^{1/5} \rfloor$ or as $p^* = \lfloor 12 \cdot (T/100)^{1/5} \rfloor$.

6. Simulations

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

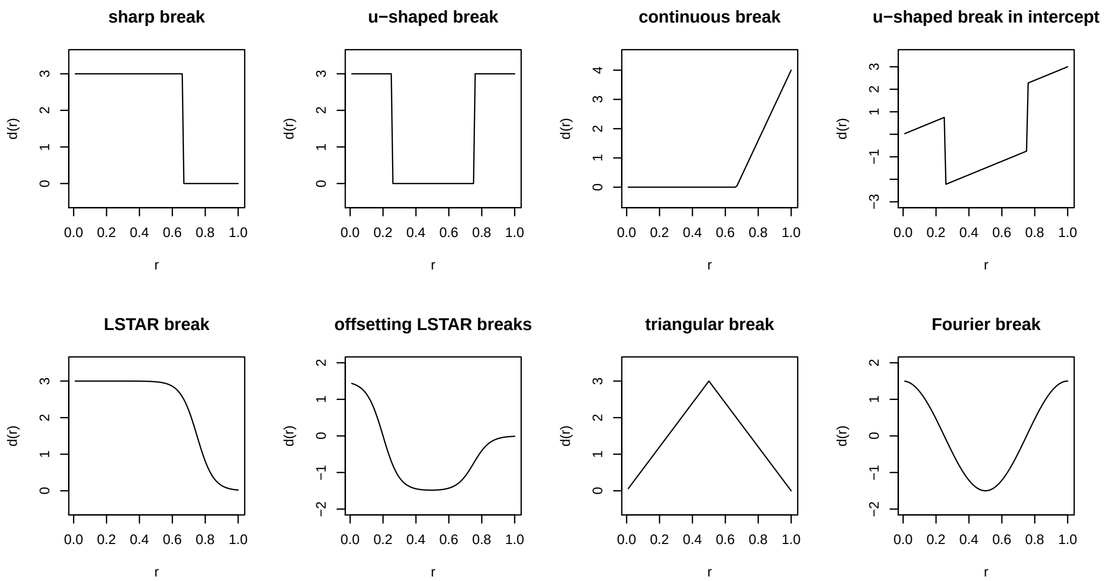
While the zero-trend $d_t = 0$ is the main benchmark, we consider several other trends including sharp breaks and smooth changes of different shapes. The trend specifications are presented in Table 2 and Figure 1. The parameter λ determines the size of the break. Similar trend functions are also considered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

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, which are considered in the Monte Carlo simulations, are presented. The parameter λ determines the size of the trend.

Figure 1: Plots of the trend functions from Table 2



Note: The plots of the of the trend functions from Table 2, which are considered in the Monte Carlo simulations, 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 that follow the variance function $\sigma^2(r) = \lambda \cdot 1_{\{r \leq 2/3\}}$ are presented. As noted by Müller and Elliott (2003), the power of a unit root test depends on the initial condition. Thus, we consider both the zero initial condition $x_0 = 0$ as well as a random initial condition, where $x_0 = \sum_{k=1}^{T} \rho^{T-k} \tilde{\epsilon}_k$ is simulated from i.i.d. standard normal innovations $\tilde{\epsilon}_k$.

For the small-b tests, we consider blocklengths of the form $B = T^\gamma$ with parameters $\gamma \in \{0.5, 0.6, 0.7, 0.8\}$, and, for the fixed-$b$ versions, we implement $B = b \cdot T$ with relative blocklengths $b \in \{0.2, 0.3, 0.4, 0.5\}$. Size and power results are presented for τ-SB and τ-FB as well as for their pre-whitened and heteroskedasticity-robust versions. Both fixed lag augmentation as well as a flexible lag augmentation determined by the BIC are implemented. All empirical size levels are presented for a significance level of 5%, while the power results are size-adjusted. The models are simulated with 100,000 repetitions for sample sizes of $T = 100$ and $T = 300$.

Table 3: Size and size-adjusted power for the zero-trend specification

	zero initial condition				random initial condition			
	$T = 100$		$T = 300$		$T = 100$		$T = 300$	
	$\rho = 1$	$\rho = 0.9$	$\rho = 1$	$\rho = 0.9$	$\rho = 1$	$\rho = 0.9$	$\rho = 1$	$\rho = 0.9$
i.i.d. errors – no lag augmentation (p=0)								
τ-SB, $B = T^{0.5}$	0.057	0.294	0.056	0.845	0.056	0.281	0.058	0.838
τ-SB, $B = T^{0.6}$	0.057	0.351	0.057	0.952	0.057	0.334	0.058	0.948
τ-SB, $B = T^{0.7}$	0.054	0.409	0.056	0.989	0.052	0.394	0.057	0.987
τ-SB, $B = T^{0.8}$	0.040	0.445	0.045	0.997	0.039	0.420	0.046	0.996
τ-FB, $B = 0.2T$	0.049	0.390	0.049	0.991	0.047	0.378	0.051	0.989
τ-FB, $B = 0.3T$	0.051	0.428	0.049	0.996	0.050	0.410	0.050	0.995
τ-FB, $B = 0.4T$	0.053	0.443	0.050	0.997	0.051	0.421	0.051	0.996
τ-FB, $B = 0.5T$	0.053	0.452	0.050	0.997	0.052	0.422	0.051	0.997
ADF	0.054	0.310	0.052	0.995	0.053	0.334	0.052	0.996
DF-GLS	0.078	0.661	0.058	1.000	0.076	0.490	0.059	0.931
DF-GLS-trend	0.069	0.294	0.053	0.993	0.069	0.255	0.054	0.945
EL	0.061	0.117	0.054	0.761	0.061	0.113	0.054	0.736
AR(1) errors – fixed lag augmentation (p=1)								
τ-SB^{PW}, $B = T^{0.5}$	0.010	0.309	0.020	0.805	0.008	0.331	0.018	0.819
τ-SB^{PW}, $B = T^{0.6}$	0.021	0.335	0.036	0.908	0.017	0.357	0.033	0.913
τ-SB^{PW}, $B = T^{0.7}$	0.031	0.368	0.044	0.963	0.027	0.387	0.043	0.964
τ-SB^{PW}, $B = T^{0.8}$	0.026	0.381	0.040	0.981	0.021	0.408	0.038	0.982
τ-FB^{PW}, $B = 0.2T$	0.024	0.355	0.040	0.968	0.020	0.374	0.038	0.969
τ-FB^{PW}, $B = 0.3T$	0.032	0.375	0.043	0.980	0.028	0.393	0.042	0.981
τ-FB^{PW}, $B = 0.4T$	0.037	0.379	0.045	0.983	0.031	0.403	0.044	0.985
τ-FB^{PW}, $B = 0.5T$	0.039	0.379	0.046	0.985	0.033	0.407	0.044	0.986
ADF	0.056	0.242	0.051	0.969	0.055	0.244	0.052	0.969
DF-GLS	0.077	0.589	0.058	1.000	0.079	0.522	0.059	0.990
DF-GLS-trend	0.071	0.239	0.052	0.968	0.069	0.229	0.053	0.951
EL	0.067	0.095	0.056	0.609	0.067	0.094	0.057	0.595
AR(1) errors – flexible lag augmentation where p is determined by BIC								
τ-SB^{PW}, $B = T^{0.5}$	0.004	0.357	0.015	0.837	0.003	0.397	0.013	0.850
τ-SB^{PW}, $B = T^{0.6}$	0.015	0.366	0.032	0.913	0.011	0.395	0.029	0.921
τ-SB^{PW}, $B = T^{0.7}$	0.026	0.387	0.042	0.960	0.020	0.414	0.040	0.962
τ-SB^{PW}, $B = T^{0.8}$	0.023	0.394	0.039	0.978	0.017	0.425	0.037	0.979
τ-FB^{PW}, $B = 0.2T$	0.019	0.377	0.039	0.964	0.014	0.406	0.036	0.967
τ-FB^{PW}, $B = 0.3T$	0.028	0.389	0.043	0.976	0.023	0.415	0.040	0.978
τ-FB^{PW}, $B = 0.4T$	0.034	0.389	0.044	0.980	0.027	0.423	0.043	0.982
τ-FB^{PW}, $B = 0.5T$	0.036	0.385	0.046	0.981	0.029	0.421	0.043	0.983
ADF	0.058	0.240	0.051	0.967	0.057	0.242	0.052	0.967
DF-GLS	0.085	0.542	0.060	0.999	0.086	0.480	0.061	0.986
DF-GLS-trend	0.082	0.219	0.054	0.955	0.080	0.206	0.056	0.932
EL	0.106	0.089	0.066	0.573	0.107	0.087	0.066	0.559

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 is equal to the augmentation order. The random initial condition is simulated from T lagged innovations. For $\rho = 1$, the rejection frequencies are based on the asymptotic critical values for a significance level of 5%, while, for $\rho = 0.9$, the values are size-adjusted.

Table 4: Size and size-adjusted power for trends 1-4 from Table 2 and i.i.d. errors

	$T=100,\ \rho=1$		$T=100,\ \rho=0.9$			$T=300,\ \rho=1$		$T=300,\ \rho=0.9$		
	$\lambda=3$	$\lambda=9$	$\lambda=3$	$\lambda=6$	$\lambda=9$	$\lambda=3$	$\lambda=9$	$\lambda=3$	$\lambda=6$	$\lambda=9$
sharp break										
τ-SB, $B=T^{0.5}$	0.056	0.039	0.248	0.162	0.104	0.056	0.053	0.816	0.725	0.586
τ-SB, $B=T^{0.6}$	0.056	0.039	0.281	0.159	0.085	0.058	0.054	0.928	0.834	0.656
τ-SB, $B=T^{0.7}$	0.053	0.038	0.296	0.124	0.043	0.056	0.054	0.970	0.856	0.566
τ-SB, $B=T^{0.8}$	0.044	0.048	0.281	0.075	0.014	0.047	0.052	0.969	0.688	0.173
τ-FB, $B=0.2T$	0.051	0.042	0.293	0.142	0.061	0.051	0.054	0.972	0.845	0.511
τ-FB, $B=0.3T$	0.056	0.056	0.284	0.093	0.024	0.053	0.063	0.970	0.720	0.224
τ-FB, $B=0.4T$	0.062	0.081	0.280	0.077	0.013	0.054	0.074	0.968	0.666	0.146
τ-FB, $B=0.5T$	0.062	0.085	0.292	0.088	0.016	0.053	0.073	0.972	0.690	0.162
ADF	0.050	0.023	0.158	0.027	0.002	0.051	0.040	0.887	0.255	0.006
DF-GLS	0.078	0.065	0.364	0.056	0.003	0.058	0.058	0.983	0.630	0.063
DF-GLS-trend	0.069	0.057	0.238	0.134	0.060	0.053	0.053	0.965	0.793	0.414
EL	0.060	0.044	0.110	0.088	0.071	0.053	0.050	0.710	0.569	0.397
u-shaped break										
τ-SB, $B=T^{0.5}$	0.055	0.022	0.213	0.113	0.069	0.057	0.046	0.785	0.615	0.413
τ-SB, $B=T^{0.6}$	0.056	0.022	0.231	0.098	0.049	0.058	0.047	0.898	0.694	0.397
τ-SB, $B=T^{0.7}$	0.057	0.025	0.209	0.052	0.016	0.055	0.048	0.945	0.630	0.197
τ-SB, $B=T^{0.8}$	0.040	0.018	0.252	0.065	0.017	0.046	0.044	0.941	0.443	0.046
τ-FB, $B=0.2T$	0.053	0.026	0.232	0.080	0.031	0.050	0.050	0.946	0.587	0.147
τ-FB, $B=0.3T$	0.056	0.033	0.231	0.059	0.017	0.051	0.054	0.940	0.442	0.048
τ-FB, $B=0.4T$	0.052	0.025	0.252	0.066	0.018	0.050	0.046	0.929	0.418	0.040
τ-FB, $B=0.5T$	0.045	0.010	0.243	0.061	0.018	0.047	0.031	0.922	0.377	0.032
ADF	0.046	0.011	0.178	0.056	0.019	0.049	0.030	0.911	0.381	0.045
DF-GLS	0.077	0.049	0.374	0.092	0.021	0.058	0.056	0.985	0.658	0.105
DF-GLS-trend	0.063	0.016	0.126	0.018	0.003	0.050	0.036	0.819	0.133	0.002
EL	0.064	0.053	0.108	0.090	0.076	0.055	0.057	0.703	0.547	0.377
continuous break										
τ-SB, $B=T^{0.5}$	0.049	0.015	0.264	0.188	0.115	0.053	0.037	0.839	0.818	0.782
τ-SB, $B=T^{0.6}$	0.049	0.014	0.299	0.185	0.093	0.054	0.036	0.944	0.916	0.853
τ-SB, $B=T^{0.7}$	0.046	0.012	0.327	0.165	0.063	0.053	0.035	0.981	0.937	0.783
τ-SB, $B=T^{0.8}$	0.035	0.011	0.355	0.175	0.059	0.043	0.029	0.989	0.919	0.640
τ-FB, $B=0.2T$	0.042	0.011	0.315	0.172	0.074	0.046	0.030	0.983	0.932	0.745
τ-FB, $B=0.3T$	0.043	0.012	0.331	0.160	0.055	0.046	0.031	0.988	0.914	0.631
τ-FB, $B=0.4T$	0.045	0.015	0.351	0.174	0.058	0.047	0.032	0.991	0.934	0.685
τ-FB, $B=0.5T$	0.046	0.016	0.360	0.184	0.062	0.047	0.033	0.992	0.941	0.704
ADF	0.046	0.011	0.149	0.019	0.001	0.047	0.029	0.891	0.273	0.007
DF-GLS	0.064	0.016	0.381	0.064	0.004	0.055	0.035	0.984	0.663	0.088
DF-GLS-trend	0.061	0.022	0.215	0.091	0.026	0.050	0.036	0.956	0.712	0.260
EL	0.059	0.048	0.116	0.111	0.104	0.053	0.049	0.754	0.725	0.678
u-shaped break in intercept										
τ-SB, $B=T^{0.5}$	0.053	0.018	0.209	0.108	0.063	0.056	0.043	0.784	0.607	0.403
τ-SB, $B=T^{0.6}$	0.054	0.018	0.221	0.088	0.042	0.057	0.043	0.897	0.680	0.373
τ-SB, $B=T^{0.7}$	0.055	0.019	0.194	0.042	0.012	0.054	0.043	0.940	0.592	0.158
τ-SB, $B=T^{0.8}$	0.039	0.015	0.238	0.054	0.013	0.045	0.040	0.933	0.381	0.027
τ-FB, $B=0.2T$	0.052	0.021	0.214	0.065	0.025	0.049	0.045	0.941	0.538	0.108
τ-FB, $B=0.3T$	0.054	0.026	0.214	0.046	0.012	0.051	0.050	0.931	0.382	0.027
τ-FB, $B=0.4T$	0.051	0.020	0.238	0.054	0.012	0.049	0.042	0.923	0.364	0.024
τ-FB, $B=0.5T$	0.043	0.008	0.237	0.057	0.014	0.047	0.029	0.919	0.340	0.022
ADF	0.043	0.007	0.118	0.012	0.001	0.047	0.026	0.785	0.075	0.000
DF-GLS	0.075	0.037	0.345	0.067	0.011	0.058	0.049	0.990	0.681	0.098
DF-GLS-trend	0.063	0.016	0.126	0.018	0.003	0.050	0.036	0.819	0.133	0.002
EL	0.064	0.053	0.108	0.090	0.076	0.055	0.057	0.703	0.547	0.377

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. For $\rho=1$, the rejection frequencies are based on the asymptotic critical values for a significance level of 5%, while, for $\rho=0.9$, the values are size-adjusted.

Table 5: Size and size-adjusted power for trends 5-8 from Table 2 and i.i.d. errors

	$T=100,\ \rho=1$		$T=100,\ \rho=0.9$			$T=300,\ \rho=1$		$T=300,\ \rho=0.9$		
	$\lambda=3$	$\lambda=9$	$\lambda=3$	$\lambda=6$	$\lambda=9$	$\lambda=3$	$\lambda=9$	$\lambda=3$	$\lambda=6$	$\lambda=9$
LSTAR break										
τ-SB, $B=T^{0.5}$	0.052	0.022	0.269	0.211	0.145	0.055	0.043	0.840	0.826	0.800
τ-SB, $B=T^{0.6}$	0.051	0.020	0.308	0.209	0.115	0.056	0.042	0.945	0.926	0.883
τ-SB, $B=T^{0.7}$	0.047	0.017	0.332	0.171	0.063	0.053	0.038	0.983	0.949	0.835
τ-SB, $B=T^{0.8}$	0.036	0.014	0.348	0.161	0.048	0.044	0.032	0.988	0.907	0.598
τ-FB, $B=0.2T$	0.043	0.016	0.325	0.189	0.086	0.048	0.034	0.984	0.944	0.795
τ-FB, $B=0.3T$	0.045	0.016	0.334	0.160	0.053	0.047	0.034	0.987	0.908	0.612
τ-FB, $B=0.4T$	0.047	0.018	0.346	0.163	0.048	0.048	0.036	0.988	0.904	0.571
τ-FB, $B=0.5T$	0.048	0.020	0.355	0.171	0.051	0.049	0.037	0.989	0.907	0.574
ADF	0.049	0.019	0.179	0.037	0.004	0.050	0.037	0.926	0.416	0.030
DF-GLS	0.070	0.029	0.425	0.104	0.010	0.055	0.042	0.990	0.797	0.215
DF-GLS-trend	0.063	0.033	0.248	0.147	0.068	0.052	0.040	0.972	0.854	0.549
EL	0.059	0.046	0.115	0.112	0.106	0.053	0.050	0.751	0.716	0.666
offsetting LSTAR break										
τ-SB, $B=T^{0.5}$	0.050	0.018	0.267	0.196	0.125	0.053	0.040	0.841	0.822	0.786
τ-SB, $B=T^{0.6}$	0.050	0.016	0.297	0.188	0.097	0.055	0.038	0.945	0.919	0.859
τ-SB, $B=T^{0.7}$	0.046	0.014	0.324	0.165	0.064	0.053	0.035	0.981	0.935	0.770
τ-SB, $B=T^{0.8}$	0.035	0.011	0.338	0.153	0.053	0.043	0.029	0.985	0.877	0.506
τ-FB, $B=0.2T$	0.042	0.013	0.319	0.179	0.078	0.046	0.031	0.983	0.928	0.727
τ-FB, $B=0.3T$	0.044	0.014	0.334	0.159	0.059	0.047	0.031	0.985	0.882	0.530
τ-FB, $B=0.4T$	0.045	0.014	0.336	0.154	0.054	0.047	0.032	0.980	0.832	0.421
τ-FB, $B=0.5T$	0.046	0.014	0.326	0.137	0.043	0.047	0.032	0.979	0.798	0.336
ADF	0.052	0.048	0.239	0.121	0.046	0.050	0.044	0.978	0.838	0.476
DF-GLS	0.068	0.023	0.385	0.079	0.008	0.055	0.039	0.969	0.532	0.049
DF-GLS-trend	0.061	0.018	0.191	0.061	0.011	0.049	0.033	0.932	0.532	0.088
EL	0.060	0.050	0.116	0.114	0.109	0.053	0.050	0.755	0.736	0.704
triangular break										
τ-SB, $B=T^{0.5}$	0.051	0.020	0.267	0.204	0.136	0.055	0.042	0.840	0.822	0.793
τ-SB, $B=T^{0.6}$	0.050	0.019	0.308	0.205	0.114	0.056	0.041	0.945	0.924	0.879
τ-SB, $B=T^{0.7}$	0.047	0.016	0.339	0.191	0.083	0.054	0.039	0.983	0.951	0.829
τ-SB, $B=T^{0.8}$	0.036	0.012	0.346	0.172	0.065	0.045	0.032	0.983	0.871	0.526
τ-FB, $B=0.2T$	0.043	0.015	0.331	0.200	0.097	0.048	0.034	0.984	0.945	0.788
τ-FB, $B=0.3T$	0.046	0.016	0.343	0.184	0.075	0.048	0.035	0.983	0.880	0.555
τ-FB, $B=0.4T$	0.046	0.016	0.347	0.171	0.065	0.048	0.034	0.977	0.817	0.418
τ-FB, $B=0.5T$	0.047	0.016	0.349	0.170	0.063	0.048	0.034	0.982	0.838	0.441
ADF	0.052	0.043	0.231	0.098	0.025	0.051	0.046	0.971	0.768	0.331
DF-GLS	0.069	0.023	0.470	0.179	0.051	0.056	0.039	0.990	0.823	0.326
DF-GLS-trend	0.059	0.018	0.193	0.053	0.009	0.051	0.033	0.919	0.467	0.057
EL	0.060	0.055	0.116	0.113	0.113	0.053	0.050	0.760	0.755	0.746
Fourier break										
τ-SB, $B=T^{0.5}$	0.048	0.014	0.258	0.176	0.098	0.053	0.037	0.841	0.820	0.779
τ-SB, $B=T^{0.6}$	0.048	0.012	0.284	0.160	0.071	0.054	0.035	0.944	0.914	0.839
τ-SB, $B=T^{0.7}$	0.044	0.010	0.301	0.130	0.043	0.052	0.032	0.980	0.914	0.658
τ-SB, $B=T^{0.8}$	0.033	0.007	0.310	0.115	0.033	0.042	0.026	0.974	0.743	0.250
τ-FB, $B=0.2T$	0.041	0.009	0.299	0.144	0.054	0.046	0.028	0.981	0.895	0.578
τ-FB, $B=0.3T$	0.042	0.009	0.304	0.120	0.037	0.046	0.027	0.974	0.762	0.279
τ-FB, $B=0.4T$	0.043	0.009	0.309	0.114	0.033	0.046	0.028	0.963	0.673	0.185
τ-FB, $B=0.5T$	0.044	0.009	0.308	0.113	0.032	0.046	0.028	0.968	0.690	0.187
ADF	0.049	0.027	0.203	0.065	0.010	0.049	0.037	0.955	0.623	0.140
DF-GLS	0.064	0.015	0.430	0.140	0.041	0.054	0.034	0.989	0.784	0.235
DF-GLS-trend	0.058	0.011	0.160	0.032	0.004	0.049	0.028	0.891	0.318	0.016
EL	0.061	0.061	0.117	0.117	0.117	0.054	0.054	0.761	0.761	0.761

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. For $\rho=1$, the rejection frequencies are based on the asymptotic critical values for a significance level of 5%, while, for $\rho=0.9$, the values are size-adjusted.

Table 6: Size and size-adjusted power for different trends and AR(1) errors

	$T=100$, $\rho=1$		$T=100$, $\rho=0.9$		$T=300$, $\rho=1$		$T=300$, $\rho=0.9$	
	$\lambda=3$	$\lambda=9$	$\lambda=3$	$\lambda=9$	$\lambda=3$	$\lambda=9$	$\lambda=3$	$\lambda=9$
sharp break								
τ-SB$^{\text{PW}}$, $B=T^{0.5}$	0.004	0.004	0.296	0.150	0.015	0.011	0.797	0.643
τ-SB$^{\text{PW}}$, $B=T^{0.6}$	0.015	0.013	0.310	0.152	0.031	0.028	0.888	0.759
τ-SB$^{\text{PW}}$, $B=T^{0.7}$	0.027	0.027	0.323	0.131	0.042	0.041	0.943	0.798
τ-SB$^{\text{PW}}$, $B=T^{0.8}$	0.026	0.038	0.320	0.092	0.040	0.046	0.958	0.674
τ-FB$^{\text{PW}}$, $B=0.2T$	0.021	0.024	0.320	0.143	0.039	0.043	0.947	0.787
τ-FB$^{\text{PW}}$, $B=0.3T$	0.035	0.051	0.322	0.109	0.045	0.059	0.956	0.692
τ-FB$^{\text{PW}}$, $B=0.4T$	0.042	0.075	0.319	0.094	0.047	0.064	0.958	0.657
τ-FB$^{\text{PW}}$, $B=0.5T$	0.043	0.075	0.323	0.109	0.048	0.061	0.960	0.672
u-shaped break								
τ-SB$^{\text{PW}}$, $B=T^{0.5}$	0.005	0.004	0.304	0.149	0.014	0.010	0.801	0.600
τ-SB$^{\text{PW}}$, $B=T^{0.6}$	0.015	0.013	0.310	0.134	0.031	0.026	0.886	0.676
τ-SB$^{\text{PW}}$, $B=T^{0.7}$	0.030	0.033	0.308	0.084	0.042	0.041	0.936	0.646
τ-SB$^{\text{PW}}$, $B=T^{0.8}$	0.022	0.017	0.331	0.112	0.039	0.038	0.948	0.521
τ-FB$^{\text{PW}}$, $B=0.2T$	0.023	0.026	0.313	0.116	0.039	0.046	0.939	0.616
τ-FB$^{\text{PW}}$, $B=0.3T$	0.032	0.036	0.321	0.099	0.043	0.050	0.947	0.517
τ-FB$^{\text{PW}}$, $B=0.4T$	0.032	0.025	0.329	0.113	0.043	0.039	0.946	0.504
τ-FB$^{\text{PW}}$, $B=0.5T$	0.029	0.011	0.326	0.105	0.042	0.031	0.947	0.474
continuous break								
τ-SB$^{\text{PW}}$, $B=T^{0.5}$	0.004	0.003	0.332	0.180	0.015	0.013	0.825	0.737
τ-SB$^{\text{PW}}$, $B=T^{0.6}$	0.014	0.010	0.340	0.182	0.031	0.029	0.904	0.830
τ-SB$^{\text{PW}}$, $B=T^{0.7}$	0.025	0.018	0.350	0.170	0.042	0.038	0.952	0.862
τ-SB$^{\text{PW}}$, $B=T^{0.8}$	0.022	0.016	0.359	0.174	0.038	0.036	0.970	0.847
τ-FB$^{\text{PW}}$, $B=0.2T$	0.018	0.013	0.348	0.176	0.038	0.034	0.956	0.858
τ-FB$^{\text{PW}}$, $B=0.3T$	0.027	0.020	0.355	0.164	0.042	0.039	0.968	0.841
τ-FB$^{\text{PW}}$, $B=0.4T$	0.032	0.025	0.357	0.174	0.044	0.041	0.972	0.859
τ-FB$^{\text{PW}}$, $B=0.5T$	0.034	0.027	0.357	0.181	0.046	0.042	0.973	0.868
LSTAR break								
τ-SB$^{\text{PW}}$, $B=T^{0.5}$	0.004	0.002	0.292	0.145	0.015	0.013	0.797	0.669
τ-SB$^{\text{PW}}$, $B=T^{0.6}$	0.013	0.007	0.305	0.157	0.031	0.028	0.893	0.803
τ-SB$^{\text{PW}}$, $B=T^{0.7}$	0.024	0.014	0.323	0.142	0.041	0.037	0.948	0.854
τ-SB$^{\text{PW}}$, $B=T^{0.8}$	0.020	0.012	0.333	0.136	0.039	0.034	0.965	0.816
τ-FB$^{\text{PW}}$, $B=0.2T$	0.017	0.010	0.320	0.153	0.038	0.033	0.952	0.847
τ-FB$^{\text{PW}}$, $B=0.3T$	0.026	0.015	0.328	0.136	0.042	0.037	0.964	0.817
τ-FB$^{\text{PW}}$, $B=0.4T$	0.031	0.019	0.331	0.138	0.044	0.040	0.966	0.813
τ-FB$^{\text{PW}}$, $B=0.5T$	0.034	0.021	0.331	0.144	0.045	0.041	0.968	0.810
Fourier break								
τ-SB$^{\text{PW}}$, $B=T^{0.5}$	0.004	0.003	0.335	0.217	0.015	0.013	0.826	0.755
τ-SB$^{\text{PW}}$, $B=T^{0.6}$	0.013	0.009	0.342	0.205	0.032	0.028	0.905	0.839
τ-SB$^{\text{PW}}$, $B=T^{0.7}$	0.025	0.015	0.349	0.180	0.041	0.036	0.951	0.846
τ-SB$^{\text{PW}}$, $B=T^{0.8}$	0.021	0.013	0.353	0.171	0.038	0.034	0.961	0.729
τ-FB$^{\text{PW}}$, $B=0.2T$	0.018	0.011	0.348	0.192	0.038	0.033	0.954	0.830
τ-FB$^{\text{PW}}$, $B=0.3T$	0.027	0.017	0.352	0.173	0.042	0.037	0.960	0.738
τ-FB$^{\text{PW}}$, $B=0.4T$	0.032	0.020	0.349	0.170	0.044	0.039	0.959	0.681
τ-FB$^{\text{PW}}$, $B=0.5T$	0.034	0.021	0.348	0.167	0.045	0.039	0.962	0.695

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 that is determined from the BIC. For $\rho = 1$, the rejection frequencies are based on the asymptotic critical values for a significance level of 5%, while, for $\rho = 0.9$, the values are size-adjusted.

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) propose detrending the series locally in the ADF regression. Let the deterministic trend function be given by the vector z_t, and let $\overline{c} \in \mathbb{R}$. Furthermore, let $y_{\overline{c},t} = y_t - \overline{c} y_{t-1}$ and $Z_{\overline{c},t} = z_t - \overline{c} z_{t-1}$ for $t \geq 2$, and let $y_{\overline{c},1} = y_1$ and $Z_{\overline{c},1} = z_1$. 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 $y_{\overline{c},t}$ on $Z_{\overline{c},t}$. For the constant trend specification (DF-GLS henceforth), let $z_t = 1$ and $\overline{c} = 7$, and, for the linear trend specification (DF-GLS-trend henceforth), $z_t = (1, t)'$ and $\overline{c} = 13.5$ are considered. Elliott et al. (1996) have shown that the Dickey-Fuller GLS test is optimal for the zero initial condition $x_0 = 0$.

An approach that does not assume a precise model for the trend component is that developed 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) consider 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 $\widehat{\delta}_2$. This yields the detrended series

$$\widetilde{S}_t = y_t - (y_1 - \widehat{\delta}_0 - \widehat{\delta}_1 \sin(\tfrac{2\pi}{T}) - \widehat{\delta}_2 \cos(\tfrac{2\pi}{T})) - \widehat{\delta}_0 t - (\widehat{\delta}_1 \sin(\tfrac{2\pi t}{T}) + \widehat{\delta}_2 \cos(\tfrac{2\pi t}{T})).$$

Finally, the test statistic is given by the t-statistic for the null hypothesis $\varphi = 0$ in the

Table 7: Simulation results for robust tests under constant trend and variance

	iid errors, $p=0$				AR(1) errors, p is chosen by BIC			
	$T=100$		$T=300$		$T=100$		$T=300$	
	$\rho=1$	$\rho=0.9$	$\rho=1$	$\rho=0.9$	$\rho=1$	$\rho=0.9$	$\rho=1$	$\rho=0.9$
τ-SB$^{\mathrm{H}}$, $B=T^{0.5}$	0.063	0.298	0.057	0.849	0.006	0.356	0.016	0.838
τ-SB$^{\mathrm{H}}$, $B=T^{0.6}$	0.064	0.355	0.059	0.953	0.018	0.365	0.033	0.913
τ-SB$^{\mathrm{H}}$, $B=T^{0.7}$	0.062	0.408	0.058	0.988	0.032	0.381	0.044	0.960
τ-SB$^{\mathrm{H}}$, $B=T^{0.8}$	0.049	0.434	0.048	0.996	0.032	0.380	0.042	0.976
τ-FB$^{\mathrm{H}}$, $B=0.2T$	0.044	0.348	0.046	0.979	0.020	0.315	0.037	0.942
τ-FB$^{\mathrm{H}}$, $B=0.3T$	0.046	0.386	0.047	0.989	0.028	0.331	0.042	0.960
τ-FB$^{\mathrm{H}}$, $B=0.4T$	0.047	0.400	0.048	0.992	0.033	0.337	0.043	0.966
τ-FB$^{\mathrm{H}}$, $B=0.5T$	0.049	0.402	0.048	0.993	0.034	0.337	0.044	0.970

Note: Simulation results are reported for 100,000 replications. All innovations are simulated independently as standard normal random variables, and the initial condition is $x_0 = 0$. The AR(1) process is given by $u_t = 0.5u_{t-1} + \epsilon_t$. For $\rho = 1$, the rejection frequencies are based on the asymptotic critical values for a significance level of 5%, while, for $\rho = 0.9$, the values are size-adjusted.

regression

$$\Delta y_t = \varphi \widetilde{S}_{t-1} + \beta_0 + \beta_1 \Delta \sin(2\pi t/T) + \beta_2 \Delta \cos(2\pi t/T) + \sum_{i=1}^{p} \xi_i \Delta \widetilde{S}_{t-i} + e_t.$$

For all tests, the lag augmentation order p is either fixed or flexibly determined by the BIC with a maximum lag order of $p^* = \lfloor 4 \cdot (T/100)^{1/5} \rfloor$.

The results presented in Tables 3 – 5 indicate that the pooled tests are slightly undersized for smaller sample sizes, where the size distortions 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 $B = 0.5T$ in the fixed-b case, the power results are higher than for the ADF test and also larger than those obtained when performing the Dickey-Fuller GLS test under a random initial condition. Hence, we do not lose power under these small-sample specifications (although, asymptotically, those tests have power in a $1/T$ neighborhood of the unit root hypothesis). Furthermore, smaller sample sizes, such as $T^{0.6}$ in the small-b context and $0.3T$ 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

Table 8: Simulation results for robust tests under breaks in the trend and variance

	$T=100$ $\rho=1$		$T=100$ $\rho=0.9$		$T=300$ $\rho=1$		$T=300$ $\rho=0.9$	
	$\lambda=2$	$\lambda=3$	$\lambda=2$	$\lambda=3$	$\lambda=2$	$\lambda=3$	$\lambda=2$	$\lambda=3$
sharp break in variance								
τ-SB$^{\mathrm{H}}$, $B=T^{0.5}$	0.066	0.068	0.301	0.291	0.058	0.058	0.829	0.788
τ-SB$^{\mathrm{H}}$, $B=T^{0.6}$	0.072	0.075	0.358	0.347	0.062	0.062	0.941	0.916
τ-SB$^{\mathrm{H}}$, $B=T^{0.7}$	0.082	0.096	0.414	0.406	0.068	0.073	0.986	0.977
τ-SB$^{\mathrm{H}}$, $B=T^{0.8}$	0.085	0.124	0.443	0.433	0.082	0.116	0.997	0.996
τ-FB$^{\mathrm{H}}$, $B=0.2T$	0.043	0.042	0.346	0.318	0.045	0.045	0.958	0.900
τ-FB$^{\mathrm{H}}$, $B=0.3T$	0.044	0.043	0.384	0.350	0.047	0.045	0.977	0.930
τ-FB$^{\mathrm{H}}$, $B=0.4T$	0.044	0.043	0.408	0.366	0.048	0.046	0.987	0.944
τ-FB$^{\mathrm{H}}$, $B=0.5T$	0.043	0.042	0.417	0.379	0.046	0.047	0.993	0.968
sharp break in trend								
τ-SB$^{\mathrm{H}}$, $B=T^{0.5}$	0.063	0.063	0.275	0.251	0.058	0.058	0.837	0.820
τ-SB$^{\mathrm{H}}$, $B=T^{0.6}$	0.063	0.064	0.322	0.281	0.060	0.060	0.942	0.928
τ-SB$^{\mathrm{H}}$, $B=T^{0.7}$	0.061	0.061	0.353	0.294	0.058	0.058	0.981	0.969
τ-SB$^{\mathrm{H}}$, $B=T^{0.8}$	0.058	0.068	0.347	0.267	0.049	0.050	0.988	0.967
τ-FB$^{\mathrm{H}}$, $B=0.2T$	0.043	0.043	0.308	0.270	0.046	0.046	0.967	0.950
τ-FB$^{\mathrm{H}}$, $B=0.3T$	0.045	0.045	0.320	0.264	0.048	0.048	0.974	0.949
τ-FB$^{\mathrm{H}}$, $B=0.4T$	0.047	0.045	0.332	0.270	0.048	0.048	0.978	0.951
τ-FB$^{\mathrm{H}}$, $B=0.5T$	0.047	0.044	0.341	0.286	0.048	0.048	0.981	0.957
sharp break in trend and variance								
τ-SB$^{\mathrm{H}}$, $B=T^{0.5}$	0.067	0.068	0.283	0.256	0.059	0.058	0.816	0.767
τ-SB$^{\mathrm{H}}$, $B=T^{0.6}$	0.070	0.073	0.331	0.297	0.062	0.062	0.933	0.897
τ-SB$^{\mathrm{H}}$, $B=T^{0.7}$	0.080	0.091	0.372	0.334	0.068	0.072	0.979	0.959
τ-SB$^{\mathrm{H}}$, $B=T^{0.8}$	0.095	0.144	0.368	0.315	0.081	0.112	0.989	0.970
τ-FB$^{\mathrm{H}}$, $B=0.2T$	0.043	0.042	0.316	0.278	0.046	0.045	0.947	0.875
τ-FB$^{\mathrm{H}}$, $B=0.3T$	0.044	0.042	0.341	0.287	0.048	0.046	0.962	0.893
τ-FB$^{\mathrm{H}}$, $B=0.4T$	0.044	0.042	0.350	0.290	0.048	0.047	0.967	0.889
τ-FB$^{\mathrm{H}}$, $B=0.5T$	0.044	0.044	0.351	0.281	0.048	0.048	0.964	0.886

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. The sharp break specification is defined by a break in the variance at 2/3 of the sample. For $\rho = 1$, the rejection frequencies are based on the asymptotic critical values for a significance level of 5%, while, for $\rho = 0.9$, the values are size-adjusted.

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 Tables 7 and 8, we can conclude that the heteroskedasticity-robust tests are sized correctly and have good power properties in the presence of a break in the variance and in the trend function.

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 impossible 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. Moreover, the empirical power of τ-SB does not exceed the empirical power of its heteroskedasticity-robust counterpart in the case of homoskedasticity. Hence, the $\tau\text{-SB}^{\text{H}}$ statistic can always be used in favor of τ-SB. The fixed-b statistic τ-FB has slightly better power results than $\tau\text{-FB}^{H}$ in the simulations.

7. Empirical illustrations

In order to illustrate the application of the test procedures, we apply the unit root tests to monthly annualized growth rates of the consumer price index. For monetary policy it is crucial to know whether shocks to the inflation rate will have a permanent or transitory effect. From an econometric point of view, the integrational properties of inflation rates affect the choice of an appropriate model. It has been widely debated in the literature whether inflation rates are $I(1)$ or $I(0)$. Early studies, such as that of MacDonald and Murphy (1989), have shown that conventional Dickey-Fuller tests are often unable to reject the unit root hypothesis for quarterly inflation rates. Evans and Lewis (1995), as well as Ng and Perron (2001), also find strong evidence that inflation rates are nonstationary. The work of Hassler and Wolters (1995), in which the authors used ARFIMA models for monthly data, produces mixed results. In contrast, Rose (1988) finds that for 18 countries, quarterly inflation rates are stationary. Using panel unit root tests, Lee and Wu (2001) provide evidence that the inflation rates of 13 OECD countries do not contain a unit root. Allowing for multiple breaks, Narayan and Narayan (2010) also find strong evidence for stationarity.

Our dataset includes 25 countries with 576 observations covering the period from 1971:1 to 2018:12.[2] The series is pre-whitened, where p is determined by the BIC using a maximal lag order of $p^* = \lfloor 4 \cdot (T/100)^{1/5} \rfloor$. The small-$b$ and fixed-b tests are then applied using different blocklengths. The results of both the pooled tests and some benchmark tests are presented in Table 9. For the conventional statistics, the augmented versions are applied with the same value for p.

At a 10% significance level, the τ-SB test with $B = T^{0.7}$ rejects the unit root hypothesis for 17 of 25 countries, and the τ-FB test with $B = 0.2T$ does so for 14 countries, whereas DF-GLS rejects H_0 for 12 countries, and ADF rejects H_0 only for 5 countries in the

[2]Source: https://data.oecd.org/

Table 9: Unit root tests applied to inflation rates (1971:1 - 2018:12)

	τ-SB^{PW} $B = T^{0.6}$	τ-SB^{PW} $B = T^{0.7}$	τ-FB^{PW} $B = 0.2T$	τ-FB^{PW} $B = 0.3T$	ADF	DF-GLS	EL
AUT	−0.37	−1.36*	−1.12	−0.76	−1.81	−1.60	−2.84
BEL	−1.62*	−2.52***	−1.95***	−1.51**	−2.67*	−2.67***	−3.67
CAN	−0.13	−0.96	−0.58	−0.55	−1.72	−1.24	−2.41
CHE	−0.23	−1.48*	−1.37*	−0.96	−2.38	−1.14	−3.19
DEU	0.13	−0.93	−1.05	−0.82	−2.03	−1.55	−2.80
DNK	−2.31**	−2.12**	−1.15*	−0.73	−1.95	−1.56	−3.92*
ESP	−0.12	−0.37	−0.06	0.07	−1.19	−1.15	−2.45
FIN	0.13	−1.14	−0.57	−0.33	−1.46	−1.47	−2.44
FRA	0.04	−0.37	−0.16	−0.15	−1.24	−1.20	−2.31
GBR	−1.53*	−1.92**	−1.46**	−1.03	−2.32	−1.96**	−3.82*
GRC	−2.12**	−1.96**	−1.15*	−0.78	−2.13	−1.90*	−4.44**
IDN	−3.19***	−3.12***	−2.18***	−1.99***	−5.49***	−4.75***	−5.68***
IND	−3.11***	−3.47***	−2.73***	−2.46***	−6.03***	−4.76***	−6.05***
ITA	−0.95	−1.13	−0.60	−0.38	−1.56	−1.57	−3.24
JPN	−1.13	−1.31*	−0.73	−0.58	−2.13	−1.46	−3.20
KOR	−2.02**	−2.11**	−1.57**	−1.31*	−2.89**	−1.78*	−4.56**
LUX	1.42	−0.99	−0.71	−0.52	−1.85	−1.80*	−2.60
MEX	−2.05**	−1.93**	−1.76**	−1.71**	−3.14**	−2.75***	−4.21**
NLD	0.66	−0.26	−0.24	−0.02	−1.55	−0.76	−2.28
NOR	−1.39*	−1.99**	−1.27*	−0.96	−2.10	−1.53	−4.07**
PRT	−1.70**	−2.04**	−0.93	−0.47	−1.83	−1.82*	−3.75
SWE	−1.61*	−1.74**	−1.40**	−1.04	−2.04	−1.18	−4.38**
TUR	−1.76**	−1.45*	−1.29*	−1.11*	−2.34	−1.71*	−3.74
USA	−1.06	−1.57*	−1.38**	−1.24*	−2.37	−2.13**	−3.42
ZAF	−2.37***	−2.25**	−1.52**	−1.18*	−2.26	−2.11**	−5.03***

Note: Test statistics for 25 countries are reported. ADF is the augmented Dickey-Fuller test with constant trend specification, DFGLS is the test developed by Elliott et al. (1996) with constant trend specification, and EL is the test by Enders and Lee (2012). The asterisks *, **, and *** denote the significance at the 10%, 5%, and 1% level, respectively. The lag order p is determined by the BIC.

dataset. Tests that are more robust to changes in the deterministic component reject the hypothesis of a unit root more frequently than tests that assume a constant trend. Hence, inflation rates might be stationary around a slowly varying trend that reflects different regimes of monetary policy rather than a constant trend.

8. 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. Both heteroskedasticity- and autocorrelation-robust versions of the tests were introduced. 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. In terms of power, the small-b tests with a blocklength of $B = T^{0.7}$ and the fixed-b tests with $B = 0.2 \cdot T$ perform well for moderately varying trends.

From the point of view of a practitioner, the τ-SB^H test has a number of advantages: First, the distribution is standard normal; thus, there is no need to resort to new tables, and p-values are easy to implement. Second, for the sample sizes used in the Monte Carlo simulation, the power tends to be higher than for conventional unit root tests under many trend specifications. Finally, the test is robust to heteroskedasticity.

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. Finally, I would like to acknowledge and thank the International Association for Applied Econometrics for a conference grant to travel to the IAAE Annual Conference 2017.

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.

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. (2001). 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 Limited.

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 non-stationary volatility. *Journal of Econometrics*, 140:919–947.

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

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.

Evans, M. D. and Lewis, K. K. (1995). Do expected shifts in inflation affect estimates of the long-run fisher relation? *The Journal of Finance*, 50:225–253.

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

Hassler, U. and Wolters, J. (1995). Long memory in inflation rates: international evidence. *Journal of Business & Economic Statistics*, 13:37–45.

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.

Lee, H.-Y. and Wu, J.-L. (2001). Mean reversion of inflation rates: evidence from 13 oecd countries. *Journal of Macroeconomics*, 23:477–487.

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.

MacDonald, R. and Murphy, P. D. (1989). Testing for the long run relationship between nominal interest rates and inflation using cointegration techniques. *Applied Economics*, 21:439–447.

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

Narayan, P. K. and Narayan, S. (2010). Is there a unit root in the inflation rate? new evidence from panel data models with multiple structural breaks. *Applied Economics*, 42:1661–1670.

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.

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

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.

Rose, A. K. (1988). Is the real interest rate stable? *The Journal of Finance*, 43:1095–1112.

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. Proofs

A.1. Proof of Lemma 1

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 $0 < \pi_1 < \ldots < \pi_L < 1$, and $L < \infty$. 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 $t_l = \lfloor \pi_l T \rfloor$ for $1 \leq l \leq L$, and let $\delta_t = \delta(t/T)$ for $1 \leq t \leq T$. Then, $d_t = \delta_t + \sum_{l=1}^{L} \lambda_l 1_{\{t \leq t_l\}}$, and consequently,

$$\Delta d_{t+j} = \Delta \delta_t + \sum_{l=1}^{L} \lambda_l 1_{\{t+j=t_l\}}, \quad d_{t+j-1} - d_j = \delta_{t+j-1} - \delta_j + \sum_{k=1}^{t-1} \sum_{l=1}^{L} \lambda_l 1_{\{k+j=t_l\}}.$$

Due to the Lipschitz continuity of $\delta(r)$, there exists a constant $C_1 < \infty$, such that

$$|\Delta d_{t+j}| \leq C_1 T^{-1} + \sum_{l=1}^{L} |\lambda_l| 1_{\{t+j=t_l\}}, \quad |d_{t+j-1} - d_j| \leq C_1 B T^{-1} + \sum_{k=1}^{t-1} \sum_{l=1}^{L} |\lambda_l| 1_{\{k+j=t_l\}}.$$

There also exists a constant $C_2 < \infty$, such that $\sum_{l=1}^{L} |\lambda_l| \leq C_2$. Furthermore, the error variance is bounded with $E[u_t^2] = \sigma^2 < \infty$ for all $1 \leq t \leq T$. We then define a common constant $C < \infty$ such that $C = \max\{C_1, C_2, \sigma, 1\}$.

Rearranging the model equation yields $\Delta x_t = (\rho - 1) x_{t-1} + u_t$. For the numerator statistic $\mathcal{Y}_{1,T}$, note that

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

such that $\mathcal{Y}_{1,T} - \mathcal{X}_{1,T} = S_1 + S_2 + S_3$, where

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

In what follows, we show that S_1, S_2, and S_3 converge to zero in probability. For the first term, we obtain

$$
\begin{aligned}
|S_1| &\leq \frac{1}{B^{3/2}T^{1/2}} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \Big(CT^{-1} + \sum_{l=1}^{L} |\lambda_l| 1_{\{t+j=t_l\}}\Big)\Big(CBT^{-1} + \sum_{k=1}^{t-1}\sum_{l=1}^{L} |\lambda_l| 1_{\{k+j=t_l\}}\Big) \\
&\leq \frac{4C^2B^2T^{-1}}{B^{3/2}T^{1/2}} = o(1).
\end{aligned}
$$

For the second term, note that $x_{t+j-1} - x_j = \sum_{k=1}^{t-1} u_{k+j} + (\rho-1)x_{k+j-1}$. With the MA-representation $x_t = \sum_{m=0}^{t-1} \rho^m u_{t-m} + x_0$, we then decompose $S_2 = S_{2,1} + S_{2,2}$, where

$$
\begin{aligned}
S_{2,1} &= \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \Delta d_{t+j} u_{k+j}}{B^{3/2}T^{1/2}}, \\
S_{2,2} &= \frac{(\rho-1)\sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \Big(\sum_{m=0}^{k+j-2} \rho^m u_{k+j-1-m} + x_0\Big)\Delta d_{t+j}}{B^{3/2}T^{1/2}}.
\end{aligned}
$$

Jensen's inequality then implies that

$$
\begin{aligned}
E[|S_{2,1}|] &\leq \frac{\sum_{t=2}^{B} \sum_{k=1}^{t-1} E[|\sum_{j=1}^{T-B} \Delta d_{t+j} u_{k+j}|]}{B^{3/2}T^{1/2}} \leq \frac{\sum_{t=2}^{B} \sum_{k=1}^{t-1} \sqrt{E[(\sum_{j=1}^{T-B} \Delta d_{t+j} u_{k+j})^2]}}{B^{3/2}T^{1/2}} \\
&\leq \frac{C\sum_{t=2}^{B} \sum_{k=1}^{t-1} \sqrt{\sum_{j=1}^{T-B} |\Delta d_{t+j}|^2}}{B^{3/2}T^{1/2}} \leq \frac{C\sum_{t=2}^{B} \sqrt{\sum_{j=1}^{T-B} \Big(CT^{-1} + \sum_{l=1}^{L} |\lambda_l| 1_{\{t+j=t_l\}}\Big)^2}}{B^{1/2}T^{1/2}} \\
&\leq \frac{4C^2}{T^{1/2}} = o(1)
\end{aligned}
$$

and

$$
\begin{aligned}
E[|S_{2,2}|] &\leq \frac{(\rho-1)\sum_{t=2}^{B} \sum_{k=1}^{t-1} E[|\sum_{j=1}^{T-B} \sum_{m=0}^{k+j-2} \rho^m \Delta d_{t+j} u_{k+j-1-m}|]}{B^{3/2}T^{1/2}} + \frac{cC}{T} E[|x_0|] \\
&\leq \frac{C(\rho-1)\sum_{t=2}^{B} \sum_{k=1}^{t-1} \sqrt{\sum_{j=1}^{T-B} (\frac{1}{1-\rho})^2 |\Delta d_{t+j}|^2}}{B^{3/2}T^{1/2}} + o(1) \\
&\leq \frac{C\sum_{t=2}^{B} \sqrt{\sum_{j=1}^{T-B} |\Delta d_{t+j}|^2}}{B^{1/2}T^{1/2}} + o(1) \leq \frac{2C^2}{T^{1/2}} + o(1) = o(1).
\end{aligned}
$$

Analogously, from $\Delta x_{t+j} = u_{t+j} + (\rho - 1)x_{t+j-1}$, we decompose $S_3 = S_{3,1} + S_{3,2}$, where

$$S_{3,1} = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} u_{t+j}(d_{t+j-1} - d_j)}{B^{3/2}T^{1/2}},$$
$$S_{3,2} = \frac{(\rho - 1)\sum_{j=1}^{T-B} \sum_{t=2}^{B} \left(\sum_{m=0}^{t+j-2} \rho^m u_{t+j-1-m} + x_0 \right)(d_{t+j-1} - d_j)}{B^{3/2}T^{1/2}}.$$

Jensen's inequality then yields

$$E[|S_{3,1}|] \leq \frac{\sum_{t=2}^{B} \sum_{k=1}^{t-1} \sqrt{E[(\sum_{j=1}^{T-B} u_{t+j}\Delta d_{k+j})^2]}}{B^{3/2}T^{1/2}} \leq \frac{C \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sum_{j=1}^{T-B} |\Delta d_{k+j}|^2}{B^{3/2}T^{1/2}}$$
$$\leq \frac{C \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sqrt{\sum_{j=1}^{T-B} (CT^{-1} + \sum_{l=1}^{L} |\lambda_l| 1_{\{k+j=t_l\}})^2}}{B^{3/2}T^{1/2}} \leq \frac{4C^2}{T^{1/2}}$$

as well as

$$E[|S_{3,2}|] \leq \frac{(\rho - 1)\sum_{t=2}^{B} E[| \sum_{j=1}^{T-B} \sum_{m=0}^{t+j-2} \rho^m u_{t+j-1-m}(d_{t+j-1} - d_j)|]}{B^{3/2}T^{1/2}} + \frac{4cC^2}{T} E[|x_0|]$$
$$\leq \frac{(\rho - 1)\sum_{t=2}^{B} \sqrt{E[(\sum_{j=1}^{T-B} \sum_{m=0}^{t+j-2} \rho^m u_{t+j-1-m}(d_{t+j-1} - d_j))^2]}}{B^{3/2}T^{1/2}} + o(1)$$
$$\leq \frac{(\rho - 1)\sum_{t=2}^{B} \sqrt{\frac{4C^4 B}{(1-\rho)^2}}}{B^{3/2}T^{1/2}} + o(1) \leq \frac{2C^2}{T^{1/2}} + o(1) = o(1).$$

Consequently, $E[|\mathcal{Y}_{1,T} - \mathcal{X}_{1,T}|] \leq |S_1| + |S_2| + |S_3| = o(1)$, and, by Markov's inequality, it follows that $|\mathcal{Y}_{1,T} - \mathcal{X}_{1,T}| = o_P(1)$. To show the second result, we analogously rewrite

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

and define

$$S_4 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} (d_{t+j-1} - d_j)^2}{B^2 T}, \quad S_5 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^{B} 2(x_{t+j-1} - x_j)(d_{t+j-1} - d_j)}{B^2 T},$$

where $\mathcal{Y}_{2,T} - \mathcal{X}_{2,T} = S_4 + S_5$. For the first term, note that

$$|S_4| \leq \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \left(CBT^{-1} + \sum_{k=1}^{t-1} \sum_{l=1}^{L} |\lambda_l| 1_{\{k+j=t_l\}} \right)^2 \leq 4C^2 T^{-1} = o(1).$$

From $(x_{t+j-1} - x_j) = \sum_{k=1}^{t-1} u_{k+j} + (\rho - 1)x_{k+j-1}$ the second term is then decomposed into $S_5 = S_{5,1} + S_{5,2}$, where

$$S_{5,1} = \frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1} 2u_{k+j}(d_{t+j-1} - d_j)}{B^2T},$$
$$S_{5,2} = \frac{2(\rho-1)\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}(\sum_{m=0}^{k+j-2}\rho^m u_{k+j-1-m} + x_0)(d_{t+j-1} - d_j)}{B^2T}.$$

Jensen's inequality yields

$$\begin{aligned}
E[|S_{5,1}|] &\leq \frac{2\sum_{t=2}^{B}\sum_{k=1}^{t-1} E[|\sum_{j=1}^{T-B} u_{k+j}(d_{t+j-1} - d_j)|]}{B^2T} \\
&\leq \frac{2\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sqrt{E[(\sum_{j=1}^{T-B} u_{k+j}(d_{t+j-1} - d_j))^2]}}{B^2T} \leq \frac{2C\sum_{t=2}^{B}\sqrt{\sum_{j=1}^{T-B}|d_{t+j-1} - d_j|^2}}{BT} \\
&\leq \frac{2C\sum_{t=2}^{B}\sqrt{\sum_{j=1}^{T-B}(CBT^{-1} + \sum_{k=1}^{t-1}\sum_{l=1}^{L}|\lambda_l| 1_{\{k+j=t_l\}})^2}}{BT} \leq \frac{7C^2}{B^{1/2}} = o(1)
\end{aligned}$$

and

$$\begin{aligned}
E[|S_{5,2}|] &\leq \frac{2(\rho-1)}{B^2T}\sum_{t=2}^{B}\sum_{k=1}^{t-1} E\Big[\Big|\sum_{j=1}^{T-B}\sum_{m=0}^{k+j-2}\rho^m u_{k+j-1-m}(d_{t+j-1} - d_j)\Big|\Big] + \frac{4E[|x_0|]cC}{B} \\
&\leq \frac{2(\rho-1)\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sqrt{E[(\sum_{j=1}^{T-B}\sum_{m=0}^{k+j-2}\rho^m u_{k+j-1-m}(d_{t+j-1} - d_j))^2]}}{B^2T} \\
&\leq \frac{2(\rho-1)\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sqrt{\frac{4C^4B}{(1-\rho)^2}}}{B^2T} + o(1) \leq \frac{2C^2}{B^{1/2}} + o(1) = o(1).
\end{aligned}$$

Hence, $E[|\mathcal{Y}_{2,T} - \mathcal{X}_{2,T}|] \leq |S_4| + |S_5| = o(1)$, and, from Markov's inequality, it follows that $|\mathcal{Y}_{2,T} - \mathcal{X}_{2,T}| = o_P(1)$.

A.2. Proof of Lemma 2

From the representation $\Delta x_{t+j} = u_{t+j} + (\rho-1)x_{t+j-1}$, we decompose the numerator statistic into $\mathcal{X}_{1,T} = S_1 + S_2 + S_3 + S_4$, where

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

Note that $S_2 + S_3 + S - 4 = 0$ if $c = 0$. First, we show that S_1 is the sum of the elements of a martingale difference array. We rearrange

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

where $\mathcal{I}_j = \{t \in \mathbb{N} : \ 1 \leq t \leq B, \ j + B - T \leq t \leq j-1\}$. The elements of the sum $q_{j,T} = \sum_{t\in\mathcal{I}_j} \sum_{k=1}^{t-1} \frac{u_j u_{j-k}}{B^{3/2}T^{1/2}}$ form a martingale difference sequence for $j \leq T$ and $T \in \mathbb{N}$. Hence, we have $E[S_1] = 0$ and $Var[S_1] = \sum_{j=1}^{T} E[q_{j,T}^2] = \frac{1}{B^3 T} \sum_{j=1}^{T} E[\tilde{q}_{j,T}^2]$, where $\tilde{q}_{j,T} = B^{3/2}T^{1/2} q_{j,T}$. The index set can be expressed as

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

Note that by mathematical induction on n, the identity

$$\sum_{t=2}^{n} \sum_{k=1}^{t-1} a_k = \sum_{k=1}^{n-1} (n-k) a_k \tag{2}$$

holds true for any sequence $(a_t)_{t\in\mathbb{N}}$. For $j \in [1, B]$, it follows that

$$\tilde{q}_{j,T} = \sum_{t=2}^{j-1} \sum_{k=1}^{t-1} u_j u_{j-k} = \sum_{k=1}^{j-1} (j-1-k) u_{j-k} = \sum_{k=1}^{j-2} k u_j u_{k+1}, \quad E[\tilde{q}_{j,T}^2] = \sigma^4 \sum_{k=1}^{j-2} k^2.$$

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

$$\tilde{q}_{j,T} = \sum_{t=2}^{B} \sum_{k=1}^{t-1} u_j u_{j-k} = \sum_{k=1}^{B} (B-k) u_{j-k} = \sum_{k=1}^{B-1} k u_j u_{j-B+k}, \quad E[\tilde{q}_{j,T}^2] = \sigma^4 \sum_{k=1}^{B-1} k^2. \tag{3}$$

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

$$\tilde{q}_{i,T}^2 = \Big(\sum_{t=i}^{B}\sum_{k=1}^{t-1} u_j u_{j-k}\Big)^2 = \Big(\sum_{t=2}^{B}\sum_{k=1}^{t-1} u_j u_{j-k} - \sum_{t=2}^{i-1}\sum_{k=1}^{t-1} u_j u_{j-k}\Big)^2$$

$$= \Big(\sum_{k=1}^{B}(B-k) u_j u_{j-k} - \sum_{k=1}^{i-1}(i-1-k) u_j u_{j-k}\Big)^2$$

$$= \Big(\sum_{k=1}^{B-1} k u_j u_{j-B+k}\Big)^2 + \Big(\sum_{k=1}^{i-2} k u_j u_{T-B+k+1}\Big)^2 - 2\sum_{k=1}^{i-1}\sum_{l=1}^{B}(i-1-k)(B-l) u_j^2 u_{j-k} u_{j-l}$$

such that $E[\tilde{q}_{i,T}^2] = \sigma^4[\sum_{k=1}^{B-1} k^2 + \sum_{k=1}^{i-2}[k^2 - 2k(B-k)]]$. Combining all cases yields

$$\sum_{j=1}^{T} E[\tilde{q}_{j,T}^2] = \sigma^4\Big[(T-B)\sum_{k=1}^{B-1} k^2 + \sum_{j=1}^{B}\sum_{k=1}^{j-2}[4k^2 - 2Bk]\Big]$$

$$= \frac{\sigma^4 B(B-1)}{6}[(T-B)(2B-1) - 2(B-2)]$$

by the Gaussian summation formulas and consequently

$$Var[S_1] = \sigma^4 \frac{(T-B)(B-1)(2B-1) - 2(B-1)(B-2)}{6B^2T}.$$

It remains to show that $S_2 + S_3 + S_4 = c \cdot (\sigma^2/2 + o_P(1))$. Let C be a constant such that $C = \sup\{E[u_j^2], E[u_j^4], c, 1\}$. Note that

$$\sum_{m=0}^{T} |\rho|^m = \frac{1 - |\rho|^{T+1}}{1-\rho} = (1 - |\rho|^{T+1})\frac{\sqrt{BT}}{c} = o(\sqrt{BT}).$$

The second term then satisfies

$$E[|S_2|] \leq \frac{(\rho - 1)\sum_{t=2}^{B}\sum_{k=1}^{t-1} E[|\sum_{j=1}^{T-B}\sum_{m=0}^{k+j-2}\rho^m u_{k+j-1-m} u_{t+j}|]}{B^{3/2}T^{1/2}}$$

$$\leq \frac{C^2\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sqrt{\sum_{j=1}^{T-B}\sum_{m=0}^{k+j-2}\rho^{2m}}}{B^2T} = o(1)$$

and, for the fourth term, we obtain

$$
\begin{aligned}
E[|S_4|] &\le \frac{(\rho-1)^2 \sum_{t=2}^{B} \sum_{k=1}^{t-1} E[| \sum_{j=1}^{T-B} \sum_{m=0}^{t+j-2} \sum_{l=0}^{k+j-2} \rho^m \rho^l u_{t+j-1-m} u_{k+j-1-l}|]}{B^{3/2} T^{1/2}} \\
&\le \frac{C^3 B^2 T \sqrt{(\sum_{m=0}^{T} |\rho|^m)^2}}{B^{5/2} T^{3/2}} = o(1).
\end{aligned}
$$

Hence, by Markov's inequality, $S_2 + S_4 = o_P(1)$. For the third term, we obtain

$$
\begin{aligned}
E[S_3] &= -\frac{c \sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sum_{m=0}^{t+j-2} \rho^m E[u_{k+j} u_{t+j-1-m}]}{B^2 T} \\
&= -\frac{c\sigma^2 (T-B) \sum_{t=2}^{B} \sum_{k=1}^{t-1} \rho^{k-1}}{B^2 T} = -\frac{c\sigma^2}{2} + o(1)
\end{aligned}
$$

and

$$
\begin{aligned}
E[S_3^2] &= \frac{c^2 E[(\sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sum_{m=0}^{t+j-2} \rho^m E[u_{k+j} u_{t+j-1-m})^2]}{B^4 T^2} \\
&= \frac{c^2 \sigma^4 (\sum_{t=2}^{B} \sum_{k=1}^{t-1} \rho^{k-1})^2}{B^4} + o(1) = \frac{c^2 \sigma^4}{4} + o(1).
\end{aligned}
$$

Hence $Var[S_3] = o(1)$, and, by Chebyshev's inequality, it follows that $S_3 = c\sigma^2/2 + o_P(1)$.

To show (b), we decompose the denominator statistic into $\mathcal{X}_{2,T} = S_5 + S_6 + S_7$, where

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

The first term satisfies $E[S_5] = \frac{1}{B^2 T} \sum_{j=1}^{T-B} \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sigma^2 = \sigma^2 \frac{(T-B)(B-1)}{2BT}$ and

$$
Var[S_5] = \frac{1}{B^4 T^2} E\Big[\Big(\sum_{j=1}^{T-B} \sum_{t=2}^{B} \big(\sum_{k=1}^{t-1} u_{j+k} \big)^2 \Big)^2 \Big] - E[D_0]^2 = O(BT^{-1}).
$$

Analogously to S_2 and S_4, we obtain

$$
\begin{aligned}
E[|S_6|] &\le \frac{2c \sum_{t=2}^{B} \sum_{k=1}^{t-1} \sum_{l=1}^{t-1} E[| \sum_{j=1}^{T-B} \sum_{m=0}^{j+k-2} \rho^m u_{j+k-1-m} u_{j+l}|]}{B^{5/2} T^{3/2}} \\
&\le \frac{2cB^{1/2}}{T^{3/2}} \cdot O(T) = O(B^{1/2} T^{-1/2})
\end{aligned}
$$

and

$$E[|S_7|] \leq \frac{c^2 \sum_{j=2}^{B} \sum_{k,l=1}^{t-1} E[| \sum_{j=1}^{T-B} \sum_{m=0}^{j+k-2} \sum_{n=0}^{j+l-2} \rho^{m+n} u_{j+k-1-m} u_{j+k-1-n}|]}{B^3 T^2}$$
$$\leq \frac{c^2}{T^2} o(T^{5/4} B^{1/4}) = o(1).$$

Since $S_6 = S_7 = 0$ if $c = 0$, the assertion follows.

A.3. Proof of Theorem 1

From the proof of Lemma 2, it follows that $\max_{1 \leq j \leq T} E[q_{j,T}^2] = O(T^{-1})$. Then, by Jensen's inequality, $\max_{1 \leq j \leq T} E[|q_{j,T}|] = O(T^{-1/2})$ and therefore $\max_{1 \leq j \leq T} |q_{j,T}| \xrightarrow{p} 0$. Furthermore, Lemma 2 yields $E[\mathcal{X}_{1,T}] = -\frac{c\sigma^2}{2} + o(1)$ and

$$V_T = Var\Big[\sum_{j=1}^{T} q_{j,T}\Big] = \frac{\sigma^4}{3} + o(1).$$

The first result then follows from Lemmas 1 and 8. Furthermore, Lemma 2 implies that $E[\mathcal{X}_{2,T}] = \sigma^2/2 + o(1)$ and that $Var[\mathcal{X}_{2,T}] = o(1)$. By Chebyshev's inequality together with Lemma 1, the second result then follows.

A.4. Proof of Theorem 2

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

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

Let $Y_T(r) = T^{-1/2} x_{\lfloor rT \rfloor}$ for $r \geq 0$. Then, with Lemma 1,

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

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

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

which implies that

$$\mathcal{Y}_{1,T} = \frac{\int_0^{1-b}(Y_T(b+r) - Y_T(r))^2 \,\mathrm{d}r - b(1-b)\sigma^2}{2b^{3/2}} + o_P(1). \tag{4}$$

Furthermore, Lemma 1 yields

$$\mathcal{Y}_{2,T} = \mathcal{X}_{2,T} + o_P(1) = \frac{1}{b^2}\int_0^{1-b}\int_r^{b+r}(Y_T(s) - Y_T(r))^2 \,\mathrm{d}s\,\mathrm{d}r + o_P(1). \tag{5}$$

The assertion then follows from Lemma 7, together with the continuous mapping theorem.

A.5. 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

$$\begin{aligned}\hat{u}_t &= y_t - \hat{\rho} y_{t-1} = \Delta y_t + (1-\hat{\rho}) y_{t-1} = \Delta d_t + u_t + (\rho - 1)x_{t-1} + (1-\hat{\rho})y_{t-1} \\ &= u_t + \Delta d_t + \mathcal{O}_P(B^{-1/2}).\end{aligned}$$

Let $\overline{u}_j = \frac{1}{B}\sum_{k=1}^{B} u_{j+k}$ and $\overline{\Delta d}_j = \frac{1}{B}\sum_{k=1}^{B} \Delta d_{j+k}$. Then, for $t = 1, \ldots, B$,

$$\sum_{j=1}^{T-B}\Big(\hat{u}_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}_{j+k}\Big)^2 = \sum_{j=1}^{T-B}(u_{j+t} - \overline{u}_j + \Delta d_{t+j} - \overline{\Delta d}_j)^2 + O_P(TB^{-1/2}).$$

Then $\sum_{j=1}^{T-B}(\hat{u}_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}_{j+k})^2 = \sum_{j=1}^{T-B} u_{j+t}^2 + o_P(T)$, where $\sum_{j=1}^{T}(\Delta d_t)^2 = o(T)$ holds true due to the piecewise Lipschitz continuity of $d(r)$. Consequently,

$$\frac{1}{(T-B)(B-1)}\sum_{j=1}^{T-B}\sum_{t=1}^{B}\left(\hat{u}_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}_{j+k}\right)^2 = \frac{1}{T}\sum_{j=1}^{T-B} u_{j+t}^2 + o_P(1) = \sigma^2 + o_P(1)$$

as $B, T \to \infty$ and $B/T \to 0$. Under fixed-b asymptotics, we obtain

$$\frac{1}{T}\sum_{j=1}^{T}(\hat{u}_j - \overline{\hat{u}})^2 = \frac{1}{T}\sum_{j=1}^{T} u_j^2 + o_P(1) = \sigma^2 + o_P(1),$$

as $B/T \to b$, $0 < b < 1$ and $B, T \to \infty$. The assertion follows from Slutsky's theorem.

A.6. Proof of Lemma 4

We follow the proof of Lemma 1, except that the variance is time-varying and bounded with $E[u_t^2] = \sigma^2(t/T) < C_3^2$ for some constant $C_3 < \infty$. We consider the common constant $C = \max\{C_1, C_2, C_3, 1\}$, and the remaining steps follow analogously to the proof of Lemma 1.

A.7. Proof of Theorem 3

We follow the proof of Lemma 2 and obtain $S_2 + S_4 = o_P(1)$, as well as

$$\begin{aligned}
E[S_3] &= -\frac{c}{B^2 T}\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\rho^{t-k-1}E[u_{k+j}^2] = -\frac{c\int_0^1 \sigma^2(r)\,\mathrm{d}r}{B^2}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\rho^{t-k-1} + o(1)\\
&= -\frac{c}{2}\int_0^1 \sigma^2(r)\,\mathrm{d}r + o(1)
\end{aligned}$$

and $Var[S_3] = o(1)$. Furthermore,

$$\begin{aligned}
Var[S_1] &= \frac{\sum_{j=1}^{T} E[\tilde{q}_{j,T}^2]}{B^3 T} = \frac{\sum_{j=B+1}^{T-B}\sum_{k=1}^{B-1} k^2 E[u_j^2]E[u_{j-B+k}^2]}{B^3 T} + o(1)\\
&= \int_{\frac{B}{T}}^{\frac{T-B}{T}}\int_0^1 s^2\sigma^2(r)\sigma^2\left(\tfrac{j-\lfloor(1-s)B\rfloor}{T}\right)\mathrm{d}s\,\mathrm{d}r + o(1) = \int_0^1\int_0^1 s^2\sigma^4(r)\,\mathrm{d}s\,\mathrm{d}r + o(1)\\
&= \frac{1}{3}\int_0^1 \sigma^4(r)\,\mathrm{d}r + o(1).
\end{aligned}$$

Lemma 8 then yields $\mathcal{X}_{1,T} \xrightarrow{d} \mathcal{N}(-\frac{c}{2}\int_0^1 \sigma^2(r)\,\mathrm{d}r, \frac{1}{3}\int_0^1 \sigma^4(r)\,\mathrm{d}r)$. For the denominator statistic, we obtain $S_6 + S_7 = o_P(1)$ and

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

while $Var[S_5] = o(1)$. Then, following the proof of Theorem 1, result (a) follows with Lemma 4.

To show (b), we consider equation (4) under heteroskedasticity, which is given by

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

where $\overline{\sigma}^2 = \int_0^1 \sigma^2(r)\,\mathrm{d}r$. Together with equation (5) and Lemma 7, the assertion follows according to Slutsky's theorem and the continuous mapping theorem.

A.8. Proof of Theorem 4

From the proof of Lemma 3, we have $\sum_{j=1}^{T-B}(\hat{u}_{j+t} - \frac{1}{B}\sum_{k=1}^{B}\hat{u}_{j+k})^2 = \sum_{j=1}^{T-B}u_{j+t}^2 + o_P(T)$ and $\sum_{j=1}^{\lfloor sT\rfloor}(\hat{u}_j - \overline{\hat{u}})^2 = \sum_{j=1}^{\lfloor sT\rfloor}u_j^2 + o_P(T)$, where $s \in [0,1]$. Then, as $B, T \to \infty$ and $B/T \to 0$, we obtain $\hat{\sigma}^2_{\mathrm{sb}} = \int_0^1 \sigma^2(r)\,\mathrm{d}r + o_P(1)$ and $\hat{\kappa}^2 = \int_0^1 \sigma^4(r)\,\mathrm{d}r + o_P(1)$. Then, (a) follows from the proof of Theorem 1. For (b), note that fixed-b asymptotics yield

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

as $B/T \to b$, $0 < b < 1$ and $B, T \to \infty$. Then, $\hat{\sigma}^2_{\mathrm{fb}} = \int_0^1 \sigma^2(r)\,\mathrm{d}r + o_P(1)$ follows with $s = 1$. Furthermore, Slutsky's theorem implies that $\hat{\eta}(s) = \eta(s) + o_P(1)$ holds pointwise for all $s \in [0,1]$. Uniform convergence then follows by Dini's theorem since both $\hat{\eta}(s)$ and $\eta(s)$ are continuous, monotone, and bounded. Following the notation of Lemma 7, the statistic τ-FB$^{\mathrm{H}}$ is a continuous functional of $\tilde{Y}_T(r)$, and, analogously to the proof of Theorem 2, the assertion follows with the continuous mapping theorem and Slutsky's theorem.

A.9. Proof of Lemma 5

Equation (1) can be rewritten as $\Delta y_t = z_t'\beta + e_t$, where $z_t = (\Delta y_{t-1}, \ldots, \Delta y_{t-p}, y_{t-1})'$ and $\beta = (\beta_1, \ldots, \beta_p, \varphi)'$ for $t = p+1, \ldots, T$. The least squares estimator for β is then given by $\hat{\beta} = (\sum_{t=p+1}^T z_t z_t')^{-1} \sum_{t=p+1}^T z_t \Delta y_t$. We then derive a consistent estimator for the coefficients $(\theta_1, \ldots, \theta_p)'$ that is asymptotically equal to $(\hat{\beta}_1, \ldots, \hat{\beta}_p)'$. From Assumption 4, it follows that $\Delta y_t = \Delta d_t + \Delta x_t = \Delta d_t + \phi x_{t-1} + \sum_{i=1}^p \theta_i u_{t-i} + \epsilon_t$. Then, from $\rho = 1$, we obtain $\Delta y_t - \Delta d_t^* = \phi y_{t-1} + \sum_{i=1}^p \theta_i \Delta y_{t-i} + \epsilon_t$, which can be equivalently rewritten as $\Delta y_t - \delta_t = z_t'\vartheta + \epsilon_t$ for $t = p+1, \ldots, T$, where $\vartheta = (\theta_1, \ldots, \theta_p, \phi)'$. The OLS estimator is given by $\hat{\vartheta} = (\sum_{t=p+1}^T z_t z_t')^{-1} \sum_{t=p+1}^T z_t(\Delta y_t - \Delta d_t^*)$. Note that the estimator satisfies $p\|\hat{\vartheta} - \vartheta\|_V = o_P(1)$, where $\|\cdot\|_V$ is an arbitrary vector norm on $\mathbb{R}^{p+1}$. In the following, we show that $p\|\hat{\beta} - \hat{\vartheta}\|_V = o_P(1)$. Note that $\hat{\beta} - \hat{\vartheta} = (\sum_{t=p+1}^T z_t z_t')^{-1} \sum_{t=p+1}^T z_t \Delta d_t^*$. For notational convenience, let $\theta_0 = -1$. Then, $\Delta d_t^* = -\sum_{i=1}^p \theta_i \Delta d_{t-i}$. Following the proof of Lemma 1, there exists a constant $C < \infty$ such that the deterministic part satisfies $|\Delta d_t^*| \leq \sum_{i=1}^p |\theta_i|(CT^{-1} + \sum_{l=1}^L |\lambda_l| 1_{\{t-i=t_l\}})$, where $\sum_{l=1}^L |\lambda_l| < C$, $\sum_{i=0}^p |\theta_i| < pC$ and $0 \leq t_l \leq 1$ for $l = 1, \ldots, L$. Then,

$$\begin{aligned}
\Big\| \frac{p}{T} \sum_{t=p+1}^T z_t \Delta d_t^* \Big\|_V &\leq \frac{p}{T} \sum_{t=p+1}^T |\Delta d_t^*| \|z_t\|_V \\
&\leq \frac{p}{T} \sum_{t=p+1}^T \sum_{i=0}^p |\theta_i| \Big(CT^{-1} + \sum_{l=1}^L |\lambda_l| 1_{\{t-i=t_l\}} \Big) \|z_t\|_V \\
&\leq \frac{p}{T^{1/2}} \Big(\frac{C^2}{T^{3/2}} \sum_{t=p+1}^T \|z_t\|_V + \frac{1}{pT^{1/2}} \sum_{i=0}^p \sum_{l=1}^L |\theta_i||\lambda_l| \|z_{t_l+i}\|_V \Big) \\
&= O_P(p^2 T^{-1/2}) = o_P(1).
\end{aligned}$$

Let $\|\cdot\|_M$ be the matrix norm induced by $\|\cdot\|_V$. Then, $\|(\frac{1}{T}\sum_{t=p+1}^T z_t z_t')^{-1}\|_M = O_P(1)$ and consequently $p\|\hat{\beta} - \hat{\vartheta}\|_V \leq \|(\frac{1}{T}\sum_{t=p+1}^T z_t z_t')^{-1}\|_M \; \|\frac{p}{T}\sum_{t=p+1}^T z_t \Delta d_t^*\|_V = o_P(1)$. The triangle inequality then yields $p\|\hat{\beta} - \vartheta\| \leq p\|\hat{\beta} - \hat{\vartheta}\|_V + \|\hat{\vartheta} - \vartheta\|_V = o_P(1)$, and the assertion follows by setting $\|\cdot\|_V$ equal to the maximum norm.

A.10. Proof of Lemma 6

For notational convenience, we define $\hat{\theta}_i = \hat{\beta}_i$ for $i = 1, \ldots, p$ and $\theta_0 = \hat{\theta}_0 = -1$. The pre-whitened series satisfy $y_t^* = -\sum_{i=0}^p \theta_i y_{t-i}$, $d_t^* = -\sum_{i=0}^p \theta_i d_{t-i}$, $x_t^* = -\sum_{i=0}^p \theta_i x_{t-i}$, and the estimated pre-whitened series are given by $\hat{y}_t^* = -\sum_{i=0}^p \hat{\theta}_i y_{t-i}$, $\hat{d}_t^* = -\sum_{i=0}^p \hat{\theta}_i d_{t-i}$, and $\hat{x}_t^* = -\sum_{i=0}^p \hat{\theta}_i x_{t-i}$. From Lemma 5, it follows that $(\hat{\theta}_0, \ldots, \hat{\theta}_p)' \xrightarrow{p} (\theta_0, \ldots, \theta_p)'$.

Furthermore, the estimated pre-whitened series satisfies $\hat{x}_t^* = x_t^* + \sum_{i=0}^p (\theta_i - \hat{\theta}_i) x_{t-i}$, and $\rho = 1$ yields $\Delta x_t^* = \epsilon_t$, as well as $\Delta \hat{x}_t^* = -\sum_{i=0}^p \hat{\theta}_i u_{t-i} = \epsilon_t + \sum_{i=0}^p (\theta_i - \hat{\theta}_i) u_{t-i}$. Following the proof of Lemma 1, there exists a common constant $1 \leq C < \infty$ such that

$$|\Delta d_{t+j}^*| \leq \sum_{i=0}^p |\theta_i||\Delta d_{t+j-i}| \leq \sum_{i=0}^p |\theta_i| \Big(CT^{-1} + \sum_{l=1}^L |\lambda_l| 1_{\{t+j-i=t_l\}} \Big),$$

$$|d_{t+j-1}^* - d_j^*| \leq \sum_{i=0}^p |\theta_i||d_{t+j-1-i} - d_{j-i}| \leq \sum_{i=0}^p |\theta_i| \Big(CBT^{-1} + \sum_{k=1}^{t-1} \sum_{l=1}^L |\lambda_l| 1_{\{k+j-i=t_l\}} \Big),$$

where the discontinuity points are given by $0 \leq t_l \leq 1$ for $l = 1, \ldots, L$. Furthermore, the constant satisfies $\sum_{l=1}^L |\lambda_l| < C$ and $\sum_{i=0}^p |\theta_i| < pC$. The stationary error term has an MA representation $u_t = \sum_{m=0}^\infty \psi_m \epsilon_{t-m}$, where $\sum_{m=0}^\infty |\psi_m| \leq C$. The variance of ϵ_t is bounded with $\sigma^2(r) \leq C^2$ for all $r \leq 1$. Let

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

In the first part of the proof, we show that $|\hat{\mathcal{Y}}_{1,T}^* - \hat{\mathcal{X}}_{1,T}^*| = o_P(1)$ and $|\hat{\mathcal{Y}}_{2,T}^* - \hat{\mathcal{X}}_{2,T}^*| = o_P(1)$. We decompose $\hat{\mathcal{Y}}_{1,T}^* - \hat{\mathcal{X}}_{1,T}^* = S_1 + S_2 + S_3$ and $\hat{\mathcal{Y}}_{2,T}^* - \hat{\mathcal{X}}_{2,T}^* = S_4 + S_5$, where

$$S_1 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^B \Delta \hat{d}_{t+j}^* (\hat{d}_{t+j-1}^* - \hat{d}_j^*)}{B^{3/2} T^{1/2}}, \quad S_2 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^B \Delta \hat{d}_{t+j}^* (\hat{x}_{t+j-1}^* - \hat{x}_j^*)}{B^{3/2} T^{1/2}},$$
$$S_3 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^B \Delta \hat{x}_{t+j}^* (\hat{d}_{t+j-1}^* - \hat{d}_j^*)}{B^{3/2} T^{1/2}}$$

and

$$S_4 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^B (\hat{d}_{t+j-1}^* - \hat{d}_j^*)^2}{B^2 T}, \quad S_5 = \frac{\sum_{j=1}^{T-B} \sum_{t=2}^B 2(\hat{x}_{t+j-1}^* - \hat{x}_j^*)(\hat{d}_{t+j-1}^* - \hat{d}_j^*)}{B^2 T}.$$

For S_1, note that

$$\sum_{j=1}^{T-B} \sum_{t=2}^B |\Delta \hat{d}_{t+j}^* (\hat{d}_{t+j-1}^* - \hat{d}_j^*)| \leq \sum_{i_1, i_2 = 0}^p \sum_{j=1}^{T-B} \sum_{t=2}^B |\hat{\theta}_{i_1}||\hat{\theta}_{i_2}||A_{i_1,i_2,j,t}|,$$

where $A_{i_1,i_2,j,t} = (CT^{-1} + \sum_{l=1}^L |\lambda_l| 1_{\{t+j-i=t_l\}})(CBT^{-1} + \sum_{k=1}^{t-1} \sum_{l=1}^L |\lambda_l| 1_{\{k+j-i=t_l\}})$. From $\sum_{j=1}^{T-B} \sum_{t=2}^B |A_{i_1,i_2,j,t}| \leq 4C^2 B$, it follows that $|S_1| \leq \frac{4C^2}{B^{1/2} T^{1/2}} (\sum_{i=0}^p |\hat{\theta}_i|)^2$ and that

$$E[|S_1|] \le \frac{4C^2}{B^{1/2}T^{1/2}}\Big(\sum_{i=0}^{p}|\theta_i|\Big)^2 + o(1) = O(p^2B^{-1/2}T^{-1/2}) = o(1).$$

For S_2, note that

$$\begin{aligned}
&\Big|\sum_{j=1}^{T-B}\sum_{t=2}^{B}\Delta\hat{d}^*_{t+j}(\hat{x}^*_{t+j-1}-\hat{x}^*_j)\Big| = \Big|\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\Delta\hat{d}^*_{t+j}\Delta\hat{x}^*_{k+j}\Big|\\
&= \Big|\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sum_{i_1,i_2=0}^{p}\hat{\theta}_{i_1}\hat{\theta}_{i_2}\Delta d_{t+j-i_1}u_{k+j-i_2}\Big|\\
&\le \sum_{t=2}^{B}\sum_{k=1}^{t-1}\sum_{i_1,i_2=0}^{p}\sum_{m=0}^{\infty}\Big|\psi_m\hat{\theta}_{i_1}\hat{\theta}_{i_2}\sum_{j=1}^{T-B}\Delta d_{t+j-i_1}\epsilon_{k+j-i_2-m}\Big|.
\end{aligned}$$

Then,

$$E[|S_2|] \le \frac{1}{B^{3/2}T^{1/2}}\sum_{m=0}^{\infty}\sum_{i_1,i_2=0}^{p}\sum_{t=2}^{B}\sum_{k=1}^{t-1}|\psi_m|E\Big[\Big|\hat{\theta}_{i_1}\hat{\theta}_{i_2}\sum_{j=1}^{T-B}\Delta d_{t+j-i_1}\epsilon_{k+j-i_2-m}\Big|\Big].$$

From the Cauchy-Schwarz inequality, it follows that

$$\begin{aligned}
&\sum_{t=2}^{B}\sum_{k=1}^{t-1}E\Big[\Big|\hat{\theta}_{i_1}\hat{\theta}_{i_2}\sum_{j=1}^{T-B}\Delta d_{t+j-i_1}\epsilon_{k+j-i_2-m}\Big|\Big]\\
&\le \sqrt{E[\hat{\theta}^2_{i_1}\hat{\theta}^2_{i_2}]}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sqrt{E\Big[\Big(\sum_{j=1}^{T-B}\Delta d_{t+j-i_1}\epsilon_{k+j-i_2-m}\Big)^2\Big]}\\
&\le |\theta_{i_1}\theta_{i_2}+o(1)|\sum_{t=2}^{B}\sum_{k=1}^{t-1}\sqrt{\sum_{j=1}^{T-B}|\Delta d_{t+j-i_1}|^2\sigma^2_{k+j-i_2-m}}\\
&\le CB|\theta_{i_1}\theta_{i_2}+o(1)|\sum_{t=2}^{B}\sqrt{\sum_{j=1}^{T-B}\Big(CT^{-1}+\sum_{l=1}^{L}|\lambda_l|1_{\{t+j-i_1=t_l\}}\Big)^2} \le 2C^2B^{3/2}|\theta_{i_1}\theta_{i_2}+o(1)|.
\end{aligned}$$

Consequently, $E[|S_2|] \leq \frac{2C^5p^2}{T^{1/2}} + o(1) = O(p^2T^{-1/2}) = o(1)$. For S_3, note that

$$\Big|\sum_{j=1}^{T-B} \Delta\hat{x}^*_{t+j}(\hat{d}^*_{t+j-1} - \hat{d}^*_j)\Big| = \Big|\sum_{i_1,i_2=0}^{p}\sum_{j=1}^{T-B}\hat{\theta}_{i_1}\hat{\theta}_{i_2}u_{t+j-i_1}(d_{t+j-1-i_2} - d_{j-i_2})\Big|$$
$$\leq \sum_{i_1,i_2=0}^{p}\sum_{m=0}^{\infty}\Big|\psi_m\hat{\theta}_{i_1}\hat{\theta}_{i_2}\sum_{j=1}^{T-B}\epsilon_{t+j-i_1-m}(d_{t+j-1-i_2} - d_{j-i_2})\Big|.$$

Furthermore,

$$E\Big[\Big(\sum_{j=1}^{T-B}\epsilon_{t+j-i_1-m}(d_{t+j-1-i_2} - d_{j-i_2})\Big)^2\Big] = \sum_{j=1}^{T-B}|d_{t+j-1-i_2} - d_{j-i_2}|^2\sigma^2_{t+j-i_1-m}$$
$$\leq C\sum_{j=1}^{T-B}\Big(CBT^{-1} + \sum_{k=1}^{t-1}\sum_{l=1}^{L}|\lambda_l|1_{\{k+j-i_2=t_l\}}\Big)^2 \leq 4C^3B.$$

From the Cauchy-Schwarz inequality, it follows that

$$E\Big[\Big|\hat{\theta}_{i_1}\hat{\theta}_{i_2}\sum_{j=1}^{T-B}\epsilon_{t+j-i_1-m}(d_{t+j-1-i_2} - d_{j-i_2})\Big|\Big]$$
$$\leq \sqrt{E[\hat{\theta}^2_{i_1}\hat{\theta}^2_{i_2}]}\sqrt{E\Big[\Big(\sum_{j=1}^{T-B}\epsilon_{t+j-i_1-m}(d_{t+j-1-i_2} - d_{j-i_2})\Big)^2\Big]} \leq 2C^{3/2}B|\theta_{i_1}\theta_{i_2} + o(1)|.$$

Consequently,

$$E[|S_3|] \leq \frac{1}{B^{3/2}T^{1/2}}\sum_{t=2}^{B}\sum_{i_1,i_2=0}^{p}\sum_{m=0}^{\infty}|\psi_m|E\Big[\Big|\hat{\theta}_{i_1}\hat{\theta}_{i_2}\sum_{j=1}^{T-B}\epsilon_{t+j-i_1-m}(d_{t+j-1-i_2} - d_{j-i_2})\Big|\Big]$$
$$\leq \frac{2C^{3/2}B}{B^{3/2}T^{1/2}}\Big(\sum_{i=0}^{p}|\theta_i + o(1)|\Big)^2\sum_{m=0}^{\infty}|\psi_m| = O(p^2B^{-1/2}T^{-1/2}) = o(1).$$

For S_4, we obtain

$$|S_4| \leq \frac{1}{B^2T}\sum_{t=2}^{B}\sum_{j=1}^{T-B}\Big(\sum_{i=0}^{p}|\hat{\theta}_i|\Big(C^2BT^{-1} + \sum_{k=1}^{t-1}\sum_{l=1}^{L}|\lambda_l|1_{\{k+j-i=t_l\}}\Big)\Big)^2 \leq \frac{4C^4}{T}\Big(\sum_{i=0}^{p}|\hat{\theta}_i|\Big)^2,$$

which implies that $E[|S_4|] \leq \frac{4C^4}{T}\Big(\sum_{i=0}^p |\theta_i|\Big)^2 + o(1) = O(p^2T^{-1}) = o(1)$. For the fifth term, we obtain

$$|S_5| \leq \frac{2}{B^2T}\sum_{i=0}^{p}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\Big|\sum_{j=1}^{T-B}\hat{\theta}_i u_{k+j-i}(\hat{d}^*_{t+j-1} - \hat{d}^*_j)\Big|$$
$$\leq \frac{2}{B^2T}\sum_{i_1,i_2=0}^{p}\sum_{m=0}^{\infty}\sum_{t=2}^{B}\sum_{k=1}^{t-1}\Big|\sum_{j=1}^{T-B}\psi_m\hat{\theta}_{i_1}\hat{\theta}_{i_2}\epsilon_{k+j-m-i_1}(d_{t+j-1-i_2} - d_{j-i_2})\Big|.$$

Note that $E[\hat{\theta}^2_{i_1}] = \theta^2_{i_1} + o(1)$ and $E[\hat{\theta}^2_{i_2}] = \theta^2_{i_2} + o(1)$ and that

$$E\Big[\Big(\sum_{j=1}^{T-B}\epsilon_{k+j-m-i_1}(d_{t+j-1-i_2} - d_{j-i_2})\Big)^2\Big] = \sum_{j=1}^{T-B}\sigma^2_{k+j-m-i_1}|d_{t+j-1-i_2} - d_{j-i_2}|^2$$
$$\leq C^2\sum_{j=1}^{T-B}\Big(CBT^{-1} + \sum_{k=1}^{t-1}\sum_{l=1}^{L}|\lambda_l| 1_{\{k+j-i_2=t_l\}}\Big)^2 \leq 4C^4B.$$

Then, by the Cauchy-Schwarz inequality,

$$E[|S_5|] \leq \frac{4C^2B^{1/2}}{B^2T}\sum_{i_1,i_2=0}^{p}\sum_{m=0}^{\infty}\sum_{t=2}^{B}\sum_{k=1}^{t-1}|\psi_m||\theta_{i_1}\theta_{i_2} + o(1)| \leq \frac{4C^5}{B^{1/2}} + o(1) = o(1).$$

By Markov's inequality, it follows that $S_1+S_2+S_3 = o_P(1)$ and that $S_4+S_5 = o_P(1)$. In the second part of the proof, we show that $|\hat{\mathcal{X}}^*_{1,T} - N^*_0| = o_P(1)$ and $|\hat{\mathcal{X}}^*_{2,T} - D^*_0| = o_P(1)$. Note that the estimated pre-whitened series satisfies $\hat{x}^*_t = x^*_t + \sum_{i=0}^p(\theta_i - \hat{\theta}_i)x_{t-i}$ such that $\Delta\hat{x}^*_{t+j} = \Delta x^*_{t+j} + \sum_{i=0}^p(\theta_i - \hat{\theta}_i)\Delta x_{t+j-i}$ as well as

$$\hat{x}^*_{t+j-1} - \hat{x}^*_j = x^*_{t+j-1} - x^*_j + \sum_{i=1}^{p}(\theta_i - \hat{\theta}_i)(x_{t+j-1-i} - x_{j-1}).$$

Then,

$$\Delta\hat{x}^*_{t+j}(\hat{x}^*_{t+j-1} - \hat{x}^*_j) - \Delta x^*_{t+j}(x^*_{t+j-1} - x^*_j)$$
$$= (\Delta\hat{x}^*_{t+j} - \Delta x^*_{t+j})(\hat{x}^*_{t+j-1} - \hat{x}^*_j) + \Delta x^*_{t+j}[(\hat{x}^*_{t+j-1} - \hat{x}^*_j) - (x^*_{t+j-1} - x^*_j)]$$
$$= \sum_{i=0}^{p}(\theta_i - \hat{\theta}_i)\Delta x_{t+j-i}(\hat{x}^*_{t+j-1} - \hat{x}^*_j) + \sum_{i=1}^{p}(\theta_i - \hat{\theta}_i)\Delta x^*_{t+j}(x_{t+j-1-i} - x_{j-1})$$

and

$$(\hat{x}^*_{t+j-1}-\hat{x}^*_j)^2-(x^*_{t+j-1}-x^*_j)^2$$
$$=\Big(\sum_{i=0}^{p}(\theta_i-\hat{\theta}_i)x_{t+j-1-i}-x_{j-i}\Big)^2-2\sum_{i=0}^{p}(\theta_i-\hat{\theta}_i)(x^*_{t+j-1}-x^*_j)(x_{t+j-1-i}-x_{j-i}).$$

Hence, we can decompose $\hat{\mathcal{X}}^*_{1,T}-N^*_0=S_6+S_7$, where

$$S_6=\frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{i=0}^{p}(\theta_i-\hat{\theta}_i)\Delta x_{t+j-i}(\hat{x}^*_{t+j-1}-\hat{x}^*_j)}{B^{3/2}T^{1/2}},$$
$$S_7=\frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{i=1}^{p}(\theta_i-\hat{\theta}_i)\Delta x^*_{t+j}(x_{t+j-1-i}-x_{j-1})}{B^{3/2}T^{1/2}}$$

and $\hat{\mathcal{X}}^*_{2,T}-D^*_0=S_8+S_9$, where

$$S_8=\frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}(\sum_{i=0}^{p}(\theta_i-\hat{\theta}_i)(x_{t+j-1-i}-x_{j-i})^2}{B^2T},$$
$$S_9=\frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{i=0}^{p}2(\theta_i-\hat{\theta}_i)(x^*_{t+j-1}-x^*_j)(x_{t+j-1-i}-x_{j-i})}{B^2T}.$$

Let $A_i=p(\theta_i-\hat{\theta}_i)$. From Lemma 5, it follows that $\max_{1\le i\le p}A_i=o_P(1)$. We then rearrange the terms such that

$$S_6=\frac{1}{p}\sum_{i=1}^{p}A_iS_{6,i},\quad S_7=\frac{1}{p}\sum_{i=1}^{p}A_iS_{7,i},\quad S_8=\frac{1}{p^2}\sum_{i_1,i_2=1}^{p}A_{i_1}A_{i_2}S_{8,i_1,i_2},\quad S_9=\frac{1}{p}\sum_{i=1}^{p}A_iS_{9,i},$$

where

$$S_{6,i}=\frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}(\hat{x}^*_{t+j-1}-\hat{x}^*_j)u_{t+j-i}}{B^{3/2}T^{1/2}},\quad S_{7,i}=\frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}(x_{t+j-1}-x_j)\epsilon_{t+j}}{B^{3/2}T^{1/2}},$$
$$S_{8,i_1,i_2}=\frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k_1,k_2=1}^{t-1}u_{k_1+j-i_1}u_{k_2+j-i_2}}{B^2T},$$
$$S_{9,i}=\frac{\sum_{j=1}^{T-B}\sum_{t=2}^{B}\sum_{k_1,k_2=1}^{t-1}\epsilon_{k_1+j}u_{k_2+j-i}}{B^2T}$$

with $0 \le i, i_1, i_2 \le p$. From the Cauchy-Schwarz inequality, it follows that

$$E[|S_{6,i}|] \le \sum_{i_2=0}^{p} \sum_{m_1,m_2=0}^{\infty} \sum_{t=2}^{B} \psi_{m_1}\psi_{m_2} \frac{|\theta_{i_2} + o(1)|}{B^{3/2}T^{1/2}} \sqrt{\Big(\sum_{j=1}^{T-B}\sum_{k=1}^{t-1} \epsilon_{t+j-i_1-m_1}\epsilon_{k+j-i_2-m_2}\Big)^2}$$
$$\le \frac{1}{B^{3/2}T^{1/2}} C^3 B\sqrt{TBC^4} = O(1)$$

and Jensen's inequality yields

$$E[|S_{7,i}|] \le \frac{1}{B^{3/2}T^{1/2}} \sum_{m=0}^{\infty}\sum_{t=2}^{B} \psi_m \sqrt{E\Big[\Big(\sum_{j=1}^{T-B}\sum_{k=1}^{t-1} \epsilon_{t+j}\epsilon_{k+j-m}\Big)^2\Big]}$$
$$\le \frac{1}{B^{3/2}T^{1/2}} CB\sqrt{TBC^4} = O(1).$$

From $E[(\sum_{j=1}^{T-B}\sum_{k_2=1}^{t-1} \epsilon_{k_1+j-i_1-m_1}\epsilon_{k_2+j-i_2-m_2})^2] \le TBC^4$ and from Jensen's inequality, it follows that

$$E[|S_{8,i_1,i_2}|] \le \frac{1}{B^2T} \sum_{t=2}^{B}\sum_{k_1=1}^{t-1}\sum_{m_1,m_2=0}^{\infty} E\Big[\Big|\sum_{j=1}^{T-B}\sum_{k_2=1}^{t-1} \psi_{m_1}\psi_{m_2}\epsilon_{k_1+j-i_1-m_1}\epsilon_{k_2+j-i_2-m_2}\Big|\Big]$$
$$\le C^4 = O(1).$$

Finally, for the last term, we obtain

$$E[|S_{9,i}|] \le \frac{2}{B^2T} \sum_{t=2}^{B}\sum_{k_1=1}^{t-1}\sum_{m=0}^{\infty} E\Big[\Big|\sum_{j=1}^{T-B}\sum_{k_2=1}^{t-1} \psi_m \epsilon_{k_1+j}\epsilon_{k_2+j-i-m}\Big|\Big]$$
$$\le \frac{2}{B^2T} \sum_{t=2}^{B}\sum_{k_1=1}^{t-1}\sum_{m=0}^{\infty} |\psi_m| \sqrt{\sum_{j=1}^{T-B}\sum_{k_2=1}^{t-1} \sigma^2_{k_1+j}\sigma^2_{k_2+j-i-m}} \le 2C^3 = O(1).$$

Then, for any $0 \le i, i_1, i_2 \le p$, we have $S_{6,i} = O_P(1)$, $S_{7,i} = O_P(1)$, $S_{8,i_1,i_2} = O_P(1)$, and $S_{9,i} = O_P(1)$, which follows from Markov's inequality. By the law of large numbers, it then follows that $S_6 + S_7 = o_P(1)$ and $S_8 + S_9 = o_P(1)$ Consequently, the triangle inequality implies that $|\hat{\mathcal{Y}}^*_{1,T} - N^*_0| \le |\hat{\mathcal{Y}}^*_{1,T} - \hat{\mathcal{X}}^*_{1,T}| + |\hat{\mathcal{X}}^*_{1,T} - N^*_0| = o_P(1)$ as well as $|\hat{\mathcal{Y}}^*_{2,T} - D^*_0| \le |\hat{\mathcal{Y}}^*_{2,T} - \hat{\mathcal{X}}^*_{2,T}| + |\hat{\mathcal{X}}^*_{2,T} - D^*_0| = o_P(1)$.

A.11. Proof of Theorem 5

First, note that

$$\hat{u}_t = \hat{y}_t^* - \hat{\rho}^* \hat{y}_{t-1} = \Delta \hat{y}_t^* + (1-\hat{\rho}^*)\hat{y}_{t-1}^* = \Delta \hat{d}_t^* + \Delta \hat{x}_t^* + (1-\hat{\rho}^*)\hat{y}_{t-1}^*,$$

where $\Delta \hat{x}_t^* = \epsilon_t + (\rho - 1)x_{t-1}^* + \sum_{i=1}^p (\theta_i - \hat{\beta}_i)x_{t-i}$. Then, for all $s \in [0,1]$,

$$\begin{aligned}\hat{u}_{\lfloor sT \rfloor} &= \epsilon_{\lfloor sT \rfloor} + \Delta \hat{d}_{\lfloor sT \rfloor}^* + (\rho - 1)x_{\lfloor sT \rfloor - 1}^* + \sum_{i=1}^p (\theta_i - \hat{\beta}_i)x_{\lfloor sT \rfloor - i} + (1-\hat{\rho}^*)\hat{y}_{\lfloor sT \rfloor - 1}^* \\ &= \epsilon_{\lfloor sT \rfloor} + o_P(1),\end{aligned}$$

which follows from the fact that $p = o(T^{1/4})$, $\Delta d_{\lfloor sT \rfloor} = \mathcal{O}(T)$, $(\rho - 1) = O(B^{-1/2}T^{-1/2})$ as well as $\Delta x_{\lfloor sT \rfloor} = O_P(1)$, $\max_{1 \leq i \leq p} p|\hat{\beta}_i - \theta_i| = o_P(1)$, $(\hat{\rho}^* - 1) = O_P(B^{-1/2}T^{-1/2})$, and $\hat{y}_{\lfloor sT \rfloor} = O(T^{1/2})$. Let $\overline{\epsilon}_j = \frac{1}{B}\sum_{k=1}^B \epsilon_{j+k}$. Then, for any $t = 1, \ldots, B$, we have

$$\sum_{j=1}^B \Big(\hat{u}_{t+j} - \frac{1}{B}\sum_{k=1}^B \hat{u}_{j+k} \Big)^2 = \sum_{j=1}^{T-B} (\epsilon_{t+j} - \overline{\epsilon}_j)^2 + o_P(T) = \sum_{j=1}^{T-B} \epsilon_{t+j}^2 + o_P(T).$$

Analogously, $\sum_{j=1}^{\lfloor sT \rfloor} (\hat{u}_j - \overline{\hat{u}})^2 = \sum_{j=1}^{\lfloor sT \rfloor} \epsilon_j^2 + o_P(T)$ for all $s \in [0,1]$. Then, as $B, T \to \infty$ and $B/T \to 0$, we obtain

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

and the consistency of $\hat{\sigma}_{\text{sb}}^{*2}$ and $\hat{\kappa}^{*2}$ follows by Slutsky's theorem. For (b), note that fixed-b asymptotics yields

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

as $B/T \to b$, $0 < b < 1$ and $B, T \to \infty$. The consistency of $\hat{\sigma}_{\text{fb}}^{*2}$ then follows with $s = 1$. Furthermore, Slutsky's theorem implies that (ii) holds pointwise. The uniform convergence then follows by Dini's theorem since both $\hat{\eta}(s)$ and $\eta(s)$ are continuous, monotone, and bounded.

Finally, since the pre-whitened numerator and denominator statistics $(\mathcal{X}^*_{1,T}, \mathcal{X}^*_{2,T})$ under Assumption 4 have the same properties as $(\mathcal{X}_{1,T}, \mathcal{X}_{2,T})$ under Assumption 3, the assertion follows with Lemma 6.

A.12. Central Limit Theorems

Lemma 7 (FCLTs)**.** *Let $\{u_t\}_{t\in\mathbb{N}}$ be independently distributed with $E[u_t] = 0$, $E[u_t^2] = \sigma_t^2$ and $E[u_t^4] < \infty$ with $\sigma_t = \sigma(t/T)$, where the function $\sigma(r)$ is càdlàg, non-stochastic, strictly positive, and bounded. Let $\rho = 1 - c/\sqrt{BT}$ with $c \geq 0$. Let $\hat{\eta}(r)$ be a consistent estimator for the variance profile $\eta(r)$, and let $\tilde{x}_{\lfloor rT \rfloor} = x_{\lfloor \hat{\eta}^{-1}(r)T \rfloor}$ and $\tilde{u}_{\lfloor rT \rfloor} = u_{\lfloor \hat{\eta}^{-1}(r)T \rfloor}$. Let $W(r)$ be a standard Brownian motion, and let $W_\eta(r) = W(\eta(r))$ be its variance-transformed counterpart (see Davidson 1994, p.486). An Ornstein-Uhlenbeck process is given by $J_c(r) = \int_0^r e^{-(r-s)c} dW(s)$, and the variance-transformed Ornstein-Uhlenbeck process is defined by $J_{c,\eta}(r) = \int_0^r e^{-(r-s)c} dW_\eta(s)$. Furthermore, let*

$$X_T(r) = \sum_{k=1}^{\lfloor rT \rfloor} \frac{u_k}{\sqrt{T}}, \quad Y_T(r) = \frac{x_{\lfloor rT \rfloor}}{\sqrt{T}}, \quad \tilde{X}_T(r) = \sum_{k=1}^{\lfloor rT \rfloor} \frac{\tilde{u}_k}{\sqrt{T}}, \quad \tilde{Y}_T(r) = \frac{\tilde{x}_{\lfloor rT \rfloor}}{\sqrt{T}}.$$

Let $B/T \to b$, where $0 < b < 1$. Then, (a) $X_T \Rightarrow \overline{\sigma} W_\eta$, (b) $Y_T \Rightarrow \overline{\sigma} J_{c/b,\eta}$, (c) $\tilde{X}_T \Rightarrow \overline{\sigma} W$, and (d) $\tilde{Y}_T \Rightarrow \overline{\sigma} J_{c/b}$, as $B, T \to \infty$, where $\overline{\sigma}^2 = \int_0^1 \sigma^2(r)\,\mathrm{d}r$ is the average variance, and where "$\Rightarrow$" denotes weak convergence on the càdlàg space $D[0,1]$ together with a suitable norm.

Proof. Result (a) follows from Lemmas 1 and 2 in Cavaliere (2005). To show (b), we set $u_0 = x_0$ for convenience. 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

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

Result (c) follows by Theorem 1 in Cavaliere and Taylor (2008), and (d) follows analogously to the proof of (b). □

Lemma 8 (CLT for md-arrays)**.** *Let $\{\{q_{j,T}\}_{1\leq j\leq T}\}_{T\in\mathbb{N}}$ be a martingale difference array* *with* $V_T = Var[\sum_{j=1}^{T} q_{j,T}] < \infty$ *and* $\max_{1 \leq j \leq T} |q_{j,T}| \xrightarrow{p} 0$. *Then, as* $T \to \infty$,

$$\frac{1}{\sqrt{V_T}} \sum_{j=1}^{T} q_{j,T} \xrightarrow{d} \mathcal{N}(0,1).$$

Proof. The result follows from Theorem 24.3 in Davidson (1994). □