*Econometric Theory*, **35**, 2019, 142–166.
doi:10.1017/S0266466618000038

# BOOTSTRAP-ASSISTED UNIT ROOT TESTING WITH PIECEWISE LOCALLY STATIONARY ERRORS

YEONWOO RHO
*Michigan Technological University*

XIAOFENG SHAO
*University of Illinois at Urbana-Champaign*

In unit root testing, a piecewise locally stationary process is adopted to accommodate nonstationary errors that can have both smooth and abrupt changes in second- or higher-order properties. Under this framework, the limiting null distributions of the conventional unit root test statistics are derived and shown to contain a number of unknown parameters. To circumvent the difficulty of direct consistent estimation, we propose to use the dependent wild bootstrap to approximate the nonpivotal limiting null distributions and provide a rigorous theoretical justification for bootstrap consistency. The proposed method is compared through finite sample simulations with the recolored wild bootstrap procedure, which was developed for errors that follow a heteroscedastic linear process. Furthermore, a combination of autoregressive sieve recoloring with the dependent wild bootstrap is shown to perform well. The validity of the dependent wild bootstrap in a nonstationary setting is demonstrated for the first time, showing the possibility of extensions to other inference problems associated with locally stationary processes.

## 1. INTRODUCTION

Unit root testing has received a lot of attention in econometrics since the seminal work by Dickey and Fuller (1979, 1981). In their articles, unit root tests were developed under the assumption of independent and identically distributed (i.i.d.) Gaussian errors. Many variants of the Dickey–Fuller test have been proposed, when the error processes are stationary, weakly dependent, and free of the Gaussian assumption. Most of the variants rely on two fundamental approaches to accommodate weak dependence in the error. One approach is the Phillips–Perron test (Phillips, 1987a; Phillips and Perron, 1988), where the longrun

This research was partially supported by NSF grant DMS-1104545. We are grateful to the co-editor and the three referees for their constructive comments and suggestions that led to a substantial improvement of the article. In particular, we are most grateful to Peter C.B. Phillips, who has gone beyond the call of duty for an editor in carefully correcting our English. We also thank Fabrizio Zanello, Mark Gockenbach, Benjamin Ong, and Meghan Campbell for proofreading. Superior, a high performance computing cluster at Michigan Technological University, was used in obtaining results presented in this publication. Address correspondence to Yeonwoo Rho, Department of Mathematical Sciences, Michigan Technological University, Houghton, MI 49931, USA; e-mail: yrho@mtu.edu.


© Cambridge University Press 2018

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

variance of the error process is consistently estimated in a nonparametric way, using heteroscedasticity and autocorrelation consistent estimators (Newey and West, 1987; Andrews, 1991). The other approach is the augmented Dickey–Fuller test (Said and Dickey, 1984), which approximates dependence structure in error processes with an AR($p$) model, where $p$ can grow with respect to the sample size. In addition to these two popular methods and their variants, bootstrap-based tests were also proposed by Paparoditis and Politis (2002, 2003), Chang and Park (2003), Parker, Paparoditis, and Politis (2006), and Cavaliere and Taylor (2009), among others. For reviews and comparisons of some of these bootstrap-based methods, we refer to Paparoditis and Politis (2005) and Palm, Smeekes, and Urbain (2008).

Recently, it has been argued that many macroeconomic and financial series exhibit nonstationary behavior in the error. In particular, heteroscedastic behavior in the error is well known in the unit root testing literature. For instance, the U.S. gross domestic product series was observed to have less variability since the 1980s; see Kim and Nelson (1999), McConnell and Perez-Quiros (2000), Busetti and Taylor (2003), and references therein. Also the majority of macroeconomic data in Stock and Watson (1999) exhibit heteroscedasticity in unconditional variances, as pointed out by Sensier and van Dijk (2004). If there are breaks in the error structure, it is known that traditional unit root tests are biased towards rejecting the unit root assumption (Busetti and Taylor, 2003). For this reason, a number of unit root tests that are robust to heteroscedasticity have been developed in the literature, such as in Busetti and Taylor (2003), Cavaliere and Taylor (2007, 2008a, 2008b, 2009), and Smeekes and Urbain (2014). Such tests allow for smooth and abrupt changes in the unconditional or conditional variance in error processes. However, changes in underlying dynamics do not have to be limited to heteroscedastic behavior. For example, in financial econometrics, it has been argued that a slow decay in sample autocorrelation in squared or absolute stock returns might be due to a smooth change in its dynamics rather than a long-memory behavior. Accordingly, Stărică and Granger (2005) and Fryzlewicz, Sapatinas, and Subba Rao (2008), among others, proposed to model stock returns series with locally stationary models. A unit root test that is robust to changes in general error dynamics would be useful in identifying a long-memory behavior in such financial series.

In a series of articles by Taylor and coauthors, heteroscedasticity in the error is accommodated by assuming that the error is generated from linear processes with heteroscedastic innovations, i.e.,

$$u_t = \sum_{j=0}^{\infty} c_j e_{t-j}, \qquad e_t = \omega_t \varepsilon_t. \tag{1}$$

Here, $\varepsilon_t$ is assumed to be i.i.d. or to be a martingale difference sequence, and $\omega_t$ is a sequence of deterministic numbers that account for heteroscedasticity. The process (1) can be considered as a generalized linear process. Although this

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

generalization allows for some departures from stationarity, it is still restrictive in the following three aspects. First, this kind of linear process cannot accommodate nonlinearity in the error process and does not include nonlinear models that are popular in time series analysis, such as threshold, bilinear, and nonlinear moving average models. Second, this error structure is somewhat special in that temporal dependence and heteroscedasticity can be separated, and it seems that most methods developed to account for heteroscedasticity in the error take advantage of this special error structure. Third, changes in second- or higher-order properties in the error process $u_t$ are not as flexibly accommodated as those in $e_t$. Recently, Smeekes and Urbain (2014) proposed a unit root test for a piecewise modulated stationary process $u_t = \omega_t v_t$, where $v_t$ is a weakly stationary process and $\omega_t$ is a sequence of deterministic numbers that accounts for heteroscedasticity. While a modulated stationary processes is more flexible in handling heteroscedasticity in $u_t$, it is still restrictive in the sense that time dependence and heteroscedasticity can be separated, similarly to (1). For more discussion on the separability of (1) and modulated stationary processes in the context of linear regression models with fixed regressors and nonstationary errors, see Rho and Shao (2015).

In this article, a general framework of nonstationarity is adapted to capture both smooth and abrupt changes in second- or higher-order properties of the error process. Specifically, the error process is assumed to follow a piecewise locally stationary (PLS) process, which was recently proposed by Zhou (2013), as a generalization of a locally stationary process. Locally stationary processes have received a lot of attention since the seminal works of Priestley (1965) and Dahlhaus (1997). Local stationarity naturally expands the notion of stationarity by allowing a change of the second-order properties of a time series; see Dahlhaus (1997), Mallat, Papanicolaou, and Zhang (1998), Giurcanu and Spokoiny (2004), and Zhou and Wu (2009), among others, for more related work. However, locally stationary processes exclude abrupt changes in second- or higher-order properties, which is often observed in real data. To accommodate abrupt changes, PLS processes were proposed to allow for a finite number of breaks in addition to smooth changes. For example, Adak (1998) proposed a PLS model in frequency domain, generalizing the local stationary model due to Dahlhaus (1997). Zhou (2013) proposed another PLS model in time domain as an extension of the framework of Zhou and Wu (2009) and Draghicescu, Guillas, and Wu (2009). This PLS process allows for both nonlinearity and piecewise local stationarity and covers a wide range of processes; see Wu (2005), Zhou and Wu (2009), and Zhou (2013) for more details.

Under the general PLS framework for the error, the limiting null distributions of the conventional unit root test statistics are not pivotal; they depend on the local longrun variance of the PLS error and some other nuisance parameters. A direct consistent estimation of the unknown parameters in the limiting null distributions is very involved, unlike the case of stationary errors (Phillips, 1987a). To overcome this difficulty, we propose to apply the dependent wild bootstrap (DWB) proposed in Shao (2010) to approximate the limiting null distributions. We also

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

provide a rigorous theoretical justification by establishing the functional central limit theorem for the standardized partial sum process of the bootstrapped residuals. This seems to be the first time DWB is justified for PLS processes and in the unit root setting. It suggests the ability of DWB to accommodate both piecewise local stationarity and weak dependence, which can be potentially used for other inference problems related to locally stationary processes.

The rest of the article is organized as follows. Section 2 presents the model, along with the test statistics and their limiting distributions under the null and local alternatives. In Section 3, DWB is described and its consistency is justified. The power behavior under local alternatives is also presented. Section 4 presents some simulation results. Section 5 summarizes the article. Technical details are relegated to the supplementary material.

We now set some standard notation. Throughout the article, $\xrightarrow{\mathcal{D}}$ is used for convergence in distribution and $\Rightarrow$ signifies weak convergence in $D[0,1]$, the space of functions on [0,1] which are right continuous and have left limits, endowed with the Skorohod metric (Billingsley, 1968). Let $a_n \asymp c_n$ indicate $a_n/c_n \to 1$ as $n \to \infty$. For $a \in \mathbb{R}$, $\lfloor a \rfloor$ denotes the largest integer smaller than or equal to $a$. $B(\cdot)$ denotes a standard Brownian motion, and $N(\mu, \Sigma)$ the (multivariate) normal distribution with mean $\mu$ and covariance matrix $\Sigma$. Set $||X||_p = (E|X|^p)^{1/p}$. Let $\mathbf{1}(\mathcal{E})$ be the indicator function, being 1 if the event $\mathcal{E}$ occurs and 0 otherwise.

## 2. UNIT ROOT TESTING UNDER PIECEWISE LOCALLY STATIONARY ERRORS

Following the framework in Phillips and Xiao (1998), consider data $\{y_{1,n}, \ldots, y_{n,n}\}$ generated from

$$y_{t,n} = X_{t,n} + \beta' z_{t,n}, \qquad t = 0, 1, \ldots, n, \tag{2}$$

and

$$X_{t,n} = \rho X_{t-1,n} + u_{t,n}, \quad t = 1, 2, \ldots, n. \tag{3}$$

Here, $\beta$ is a $p \times 1$ vector of coefficients and $z_{t,n}$ is a $p \times 1$ vector of deterministic trend functions, which satisfies the following conditions:

(Z1) There exists a sequence of scaling matrices $D_n$ and a piecewise continuous function $Z(r)$ such that $D_n^{-1} z_{\lfloor nr \rfloor,n} \Rightarrow Z(r)$ as $n \to \infty$.

(Z2) $\int_0^1 Z(r)Z(r)'dr$ is positive definite.

These assumptions include some popular trend functions, such as $(p-1)$st order polynomial trends, and are quite standard in the literature; see Section 2.1 of Phillips and Xiao (1998) and Section 2 of Cavaliere and Taylor (2007). The initial

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

condition, $X_{0,n} = 0$ is assumed in order to simplify the argument. This assumption can be relaxed to allow, e.g., $X_{0,n}$ to be bounded in probability, which does not alter our asymptotic results.

Following the framework introduced by Zhou (2013), the error process $\{u_{t,n}\}_{t=1}^n$ is assumed to be mean-zero piecewise locally stationary (PLS) with a finite number of break points. Let the break points be denoted by $b_1, b_2, \ldots, b_\tau$, where $0 = b_0 < b_1 < \cdots < b_\tau < b_{\tau+1} = 1$. The process $\{u_{t,n}\}_{t=1}^n$ is considered as a concatenation of $\tau + 1$ measurable functions $G_j(s, \mathcal{F}_t) : [0,1] \times \mathbb{R}^\infty \to \mathbb{R}$, $j = 0, 1, \ldots, \tau$, where

$$u_{t,n} = G_j(s_t, \mathcal{F}_t) \quad \text{if } b_j \le s_t < b_{j+1},$$

$s_t = t/n$, $\mathcal{F}_t = (\ldots, \varepsilon_0, \ldots, \varepsilon_{t-1}, \varepsilon_t)$, and the $\varepsilon_t$ are i.i.d. random variables with mean 0 and variance 1. The following is further assumed:

(A1) For each $j = 0, 1, \ldots, \tau$, the function $G_j(s, \mathcal{F}_t)$ is stochastically Lipschitz continuous with respect to $s$. That is, there exists a finite constant $C$ such that, for $s, s' \in [b_j, b_{j+1}]$, $s \neq s'$,

$$\frac{||G_j(s, \mathcal{F}_0) - G_j(s', \mathcal{F}_0)||_2}{|s - s'|} \le C.$$

(A2) $$\max_{j \in \{0,1,\ldots,\tau\}} \sup_{s \in [b_j, b_{j+1}]} ||G_j(s, \mathcal{F}_0)||_4 < \infty.$$

(A3) $\delta_4(k) = O(\chi^k)$ for some $\chi \in (0,1)$, where $\delta_p(k)$ is the physical dependence measure defined as

$$\delta_p(k) = \max_{j \in \{0,1,\ldots,\tau\}} \sup_{s \in [b_j, b_{j+1}]} ||G_j(s, \mathcal{F}_k) - G_j\{s, (\mathcal{F}_{-1}, \varepsilon_0', \varepsilon_1, \ldots, \varepsilon_k)\}||_p$$

if $k \ge 0$, and $\delta_p(k) = 0$ if $k < 0$. Here, $\varepsilon_0'$ is an i.i.d. copy of $\varepsilon_0$.

(A4) $\inf_{s \in [0,1]} \sigma^2(s) > 0$, where $\sigma^2(s) = \sum_{h=-\infty}^{\infty} c_j(s; h)$ is the longrun variance function, $c_j(s; h) = \mathrm{cov}\{G_j(s, \mathcal{F}_0), G_j(s, \mathcal{F}_h)\}$ for $s \in [b_j, b_{j+1})$, and $\sigma^2(1) = \lim_{s \uparrow 1} \sigma^2(s)$.

If there is no break point and the function $G$ does not depend on its first argument, then the PLS process reduces to a nonlinear causal process $G(\mathcal{F}_t)$, which can accommodate a wide range of stationary processes. A special example is when $G$ is a linear function, in which case $G(\mathcal{F}_t) = \sum_{j=0}^{\infty} c_j \varepsilon_{t-j}$ is the commonly used linear process. For nonlinear time series models that fall into the above framework of a nonlinear causal process, see Shao and Wu (2007).

By introducing the dependence on the relative location $t/n$, the PLS series naturally extends this stationary causal process $G(\mathcal{F}_t)$ to a locally stationary one; see Zhou and Wu (2009). In particular, assumption (A1) states that if $t/n$ and $t'/n$

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

are close and there is no break point in between, then $u_{t,n}$ and $u_{t',n}$ are expected to be stochastically close. In other words, the second- or higher-order property of $u_{t,n}$ should be smoothly changing, except for a finite number of break points. This ensures the local stationarity between break points. As stated in Zhou (2013), the goal of extending from locally stationary processes to PLS processes is to allow for abrupt changes in both second- and high-order properties, and to accommodate both nonlinearity and nonstationarity in a broad fashion.

The physical dependence measure in (A3) was introduced in Zhou and Wu (2009) as an extension of its stationary counterpart first introduced by Wu (2005). Assumption (A3) implies that $u_{t,n}$ is locally short-range dependent and that the dependence decays exponentially fast. When the location $s \in [0, 1]$ is fixed, the process $\{G_j(s, \mathcal{F}_t)\}_{t\in\mathbb{Z}}$ is stationary for each $j$, and (A4) introduces the time-varying longrun variance parameter $\sigma^2(s)$. Assumptions (A1)–(A4) are similar to those of Zhou and Wu (2009), Wu and Zhou (2011), and Zhou (2013). These assumptions are not the weakest possible for our theoretical results to hold but are satisfied by a wide class of time series models. See the following examples:

**Example 2.1**
The PLS framework includes a time-varying linear process as a special case. Suppose $s_t = t/n$ lies in the $j$th segment of [0,1], i.e., $s_t \in [b_j, b_{j+1})$. Let the measurable functions $G_j(s_t, \mathcal{F}_t)$ be linear functions with respect to $\mathcal{F}_t = (\ldots, \varepsilon_{t-1}, \varepsilon_t)$. Then we can write

$$u_{t,n} = G_j(s_t, \mathcal{F}_t) = \sum_{i=0}^{\infty} \psi_{i,j}(s_t)\varepsilon_{t-i}, \qquad s_t \in [b_j, b_{j+1}), \tag{4}$$

where the $\varepsilon_t$ are i.i.d. (0,1). If the following assumptions (LP1)–(LP4) are satisfied, the process in (4) is a PLS process.

(LP1) There exists a finite constant $C$ such that

$$\max_{j=0,\cdots,\tau} \sup_{s\neq s'\in[b_j,b_{j+1}]} |\psi_{i,j}(s) - \psi_{i,j}(s')|/|s - s'| \le C \quad \text{for } i = 0, 1, \ldots.$$

(LP2) $||\varepsilon_t||_4 < \infty$.

(LP3) $\max_{j=0,\cdots,\tau} \sup_{s\neq s'\in[b_j,b_{j+1}]} |\psi_{k,j}(s)| = O(\chi^k)$ for some $\chi \in (0, 1)$.

(LP4) $\inf_{s\in[0,1]} \sigma^2(s) = \inf_{j=0,\cdots,\tau} \inf_{s\in[b_j,b_{j+1}]} \sum_{h=-\infty}^{\infty} \sum_{i=0}^{\infty} \psi_{i,j}(s)\psi_{i+h,j}(s) > 0.$

Under (4), it is not difficult to show that (LP1) implies (A1); (LP2) and (LP3) imply (A2) and (A3); and (LP4) implies (A4). It is worth noting that the heteroscedastic linear process in (1) can be expressed as a special example of our framework in (4) by letting

$$\psi_{i,j}(s_t) = c_i\omega(s_t),$$

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

where $\omega(s_t)$ is a nonstochastic and strictly positive function on [0,1], representing the standard deviation of the error term of the linear process at the relative location $s_t = t/n$. Cavaliere and Taylor (2008a) further assumed

(CT1) $\sum_{i=0}^{\infty} i|c_i| < \infty$.
(CT2) $C(z) = \sum_{i=0}^{\infty} c_i z^i \neq 0$ for all $|z| \leq 1$.

As long as $\omega(s)$ is bounded for $s \in [0, 1]$, (CT1) can be equivalently written as

$$\max_{0\leq j\leq \tau} \sup_{b_j\leq s\leq b_{j+1}} \sum_{i=0}^{\infty} i|\psi_{i,j}(s)| < \infty,$$

which is weaker than (LP3). Note that our sufficient assumptions (A1)–(A4) are not necessary and can be relaxed at the expense of lengthy technical details. The PLS framework allows for locally stationary nonlinear processes as detailed below, and our technical argument is considerably different from that in Cavaliere and Taylor (2008a). The latter relies on theoretical results in Chang and Park (2002) and the Beveridge–Nelson Decomposition (Phillips and Solo, 1992), which are tailored to linear processes.

**Example 2.2**
The PLS process accommodates time-varying nonlinear models. For example, the autoregressive conditional heteroscedasticity (ARCH(1)) model with time-varying coefficients (Dahlhaus and Subba Rao, 2006) can be represented in the PLS form. For simplicity, assume there are no break points. Define $G(s, \mathcal{F}_t) = H\{s, G(s, \mathcal{F}_{t-1}), \varepsilon_t\}$, where $H : [0, 1] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ satisfies

$$H\{s, G(s, \mathcal{F}_{t-1}), \varepsilon_t\} = \{w(s) + \alpha(s)G(s, \mathcal{F}_{t-1})^2\}^{1/2}\varepsilon_t$$

with $\varepsilon_t \overset{i.i.d.}{\sim} (0, 1)$. Under standard assumptions on the smoothness and boundedness of $w(\cdot)$ and $\alpha(\cdot)$, and moment assumptions on $\varepsilon_t$, it can be shown that the time-varying ARCH(1) satisfies (A1)–(A4); see Proposition 5.1 in Shao and Wu (2007) and Assumption 2 in Dahlhaus and Subba Rao (2006). Additionally, as mentioned in Zhou (2013), many stationary nonlinear time series models naturally fall into the framework of $G(\mathcal{F}_t)$ and can be extended to piecewise stationary nonlinear models by introducing break points and allowing different nonlinear models within each segment. This flexibility of modeling complex dynamics of time series is automatically built into the PLS process.

**Remark 2.1.** The PLS framework is based on the physical dependence measure in (A3), whereas the mixing conditions can be understood as being based on a more abstract probabilistic dependence measure, and neither is inclusive of the other. On the one hand, special cases of PLS processes can be shown to satisfy a mixing condition. For example, the time-varying ARCH process introduced in Example 2.2 can be shown to be $\alpha$-mixing under some appropriate conditions on $\omega(s)$ and $\alpha(s)$ and the smoothness of the density of the $\varepsilon_t$, using Theorem 3.1 of Fryzlewicz and Subba Rao (2011). On the other hand, examples that are PLS but

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

not strong-mixing can also be found. For instance, an autoregressive (AR) process of order 1 with AR parameter 0.5 and Bernoulli innovations is known to violate the strong-mixing conditions but can easily fit into the physical dependence measure framework (Andrews, 1984; Wu, 2005).

Given the observations $\{y_{t,n}, z_{t,n}\}_{t=1}^{n}$, consider testing the unit root hypothesis

$$H_0: \rho = 1 \quad \text{vs} \quad H_1: |\rho| < 1.$$

The ordinary least squares (OLS) estimator $\widehat{\rho}_n = \left(\sum_{t=1}^{n} \widehat{X}_{t,n}\widehat{X}_{t-1,n}\right) / \left(\sum_{t=1}^{n} \widehat{X}_{t-1,n}^2\right)$ of $\rho$ is considered, where $\widehat{X}_{t,n} = y_{t,n} - \widehat{\beta}_n' z_{t,n}$ are the OLS residuals of $y_{t,n}$ regressed on $z_{t,n}$.[1] We proceed to define two test statistics that are popular in the literature, namely

$$\mathbf{T}_n = n(\widehat{\rho}_n - 1) \quad \text{and} \quad \mathbf{t}_n = \frac{\left(\sum_{t=1}^{n} \widehat{X}_{t-1,n}^2\right)^{1/2}(\widehat{\rho}_n - 1)}{(s_n^2)^{1/2}},$$

where $s_n^2 = (n-2)^{-1}\sum_{t=1}^{n}\left(\widehat{X}_{t,n} - \widehat{\rho}_n \widehat{X}_{t-1,n}\right)^2$.

In the unit root testing literature, local alternatives, i.e., $\rho_n = 1 + c/n$, $c < 0$, are often considered to examine the behavior of the test when the true $\rho$ is close to the unity. The Ornstein–Uhlenbeck process, $J_c(r) = \int_0^r e^{(r-s)c} dB(s)$, is usually involved in the limiting distributions of $\mathbf{T}_n$ and $\mathbf{t}_n$ for this near-integrated case. Under our error assumptions, define a similar process, $J_{c,\sigma}(r) := \int_0^r e^{(r-s)c}\sigma(s)dB(s)$. Theorem 2.1 below states the limiting distributions of the two test statistics under the null hypothesis, $\rho = 1$, and under local alternatives, $\rho_n = 1 + c/n$, $c < 0$.

THEOREM 2.1. *Assume* (A1)–(A4) and (Z1)–(Z2). When $\rho = \rho_n = 1 + c/n$, $c \le 0$,

$$n^{-1/2}\widehat{X}_{\lfloor nr \rfloor, n} \Rightarrow J_{c,\sigma|Z}(r), \tag{5}$$

where $J_{c,\sigma|Z}(r) = J_{c,\sigma}(r) - \{\int_0^1 J_{c,\sigma}(s)Z(s)'ds\}\{\int_0^1 Z(s)Z(s)'ds\}^{-1}Z(r)$ is the Hilbert projection of $J_{c,\sigma}(\cdot)$ onto the space orthogonal to $Z(\cdot)$. The limiting distributions of the two statistics are

$$\mathbf{T}_n \xrightarrow{\mathcal{D}} \mathcal{L}_{\mathbf{T},c} = \frac{\int_0^1 J_{c,\sigma|Z}(r)\sigma(r)dB(r) + 2^{-1}\{\int_0^1 \sigma^2(r)dr - \sigma_u^2\}}{\int_0^1 J_{c,\sigma|Z}^2(r)dr} + c \tag{6}$$

and

$$\mathbf{t}_n \xrightarrow{\mathcal{D}} \mathcal{L}_{\mathbf{t},c} = \frac{\int_0^1 J_{c,\sigma|Z}(r)\sigma(r)dB(r) + 2^{-1}\{\int_0^1 \sigma^2(r)dr - \sigma_u^2\}}{\{\sigma_u^2 \int_0^1 J_{c,\sigma|Z}^2(r)dr\}^{1/2}} + \frac{c\{\int_0^1 J_{c,\sigma|Z}^2(r)dr\}^{1/2}}{\sigma_u}, \tag{7}$$

where $\sigma_u^2 = \lim_{n\to\infty} n^{-1}\sum_{t=1}^{n} E(u_{t,n}^2)$.

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

When $\rho$ is far from the unity, with $c < 0$ and large in absolute value, the limiting distributions are far to the left compared to the limiting null distributions. In this case, the unit root null hypothesis would be rejected with high probability. This implies that the unit root tests based on $\mathbf{T}_n$ and $\mathbf{t}_n$ have nontrivial powers under local alternatives. In the special case $\beta \equiv 0$ and $\sigma(s) = \sigma$, i.e., when there is no deterministic trend and the error is stationary, the two limiting distributions $\mathcal{L}_{\mathbf{T},c}$ and $\mathcal{L}_{\mathbf{t},c}$ reduce to those found in Theorem 1 of Phillips (1987b).

Notice that when $c = 0$, i.e., under the null hypothesis, $J_{c,\sigma|Z}(r) = B_{\sigma|Z}(r)$, which implies $\int_0^1 B_{\sigma|Z}(r)\sigma(r)dB(r) = 2^{-1}\{B^2_{\sigma|Z}(1) - \int_0^1 \sigma^2(r)dr\}$, by Itô's formula.[2] In this case, the limiting null distributions can be written as

$$\mathcal{L}_{\mathbf{T}} := \mathcal{L}_{\mathbf{T},0} = \frac{\{B_{\sigma|Z}(1)^2 - \sigma_u^2\}}{2\int_0^1 B_{\sigma|Z}(r)^2 dr} \quad \text{and} \quad \mathcal{L}_{\mathbf{t}} := \mathcal{L}_{\mathbf{t},0} = \frac{\{B_{\sigma|Z}(1)^2 - \sigma_u^2\}}{2\{\sigma_u^2 \int_0^1 B_{\sigma|Z}(r)^2 dr\}^{1/2}},$$

where $B_\sigma(r) = \int_0^r \sigma(s)dB(s)$ and

$$B_{\sigma|Z}(r) = B_\sigma(r) - \left\{\int_0^1 B_\sigma(s)Z(s)'ds\right\}\left\{\int_0^1 Z(s)Z(s)'ds\right\}^{-1} Z(r)$$

is the Hilbert projection of $B_\sigma(\cdot)$ onto the space orthogonal to $Z(\cdot)$.

Since the limiting null distributions contain a number of unknown parameters, one may try to directly estimate them for inferential purpose. Consistent estimation of the limit of the average marginal variance, $\sigma_u^2$, is not difficult. This can be done by noting that $\sigma_u^2 = \int_0^1 c(s;0)ds$, with $c(s;h)$ defined in (A4) and in the first paragraph of the technical appendix. As a special case of Lemma A.5, $\sigma_u^2$ can be consistently estimated by $n^{-1}\sum_{t=1}^n \widehat{u}_{t,n}^2$, where $\widehat{u}_{t,n} = \widehat{X}_{t,n} - \widehat{\rho}_n \widehat{X}_{t-1,n}$ is the OLS residual. However, consistent estimation of the local longrun variance $\sigma^2(s)$ is not as simple for a PLS process. In the case of a stationary error process, $c(s;h)$ does not depend on $s$, and $\sigma^2(s) = \sum_{h=-\infty}^{\infty} \mathrm{cov}(u_t, u_{t+h}) = \sigma^2$. If it is further assumed that there is no deterministic trend function, i.e., $\beta \equiv 0$, then the limiting null distributions $\mathcal{L}_{\mathbf{T}}$ and $\mathcal{L}_{\mathbf{t}}$ reduce to

$$\mathcal{L}_{\mathbf{T},\sigma(s)=\sigma} = \frac{\{B(1)^2 - \sigma_u^2/\sigma^2\}}{2\int_0^1 B(r)^2 dr} \quad \text{and} \quad \mathcal{L}_{\mathbf{t},\sigma(s)=\sigma} = \frac{\sigma/\sigma_u\{B(1)^2 - \sigma_u^2/\sigma^2\}}{2\{\int_0^1 B(r)^2 dr\}^{1/2}}.$$

These limiting null distributions contain only a couple of unknown parameters and coincide with those in Phillips (1987a). To make inference possible in this stationary error case, as Phillips (1987a) and Phillips and Perron (1998) suggested, the longrun variance $\sigma^2$ of the error process may be consistently estimated using heteroscedasticity and autocorrelation consistent (HAC) estimators. The new statistics (see page 287 of Phillips, 1987a), adjusted by using consistent estimates of $\sigma$ and $\sigma_u$, have pivotal limiting null distributions. However, in the piecewise locally stationary error case, the usual Phillips–Perron adjustment may not lead

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

to pivotal limiting null distributions. Specifically, the HAC-based estimator of the nuisance parameter $\sigma^2(s)$, $s \in [0,1]$, is not known to be consistent in the PLS framework. The parameter $\sigma^2(s)$ is unknown at infinitely many points, and the integral of $\sigma(s)$ over a Brownian motion needs to be estimated as well as $\sigma^2(s)$ itself. This makes the direct estimation of the unknown parameters in the limiting null distributions difficult.

## 3. BOOTSTRAP-ASSISTED UNIT ROOT TEST

To implement the (asymptotic) level $\alpha$ test, the $\alpha$-quantiles of the limiting null distributions $\mathcal{L}_{\mathbf{T}}$ and $\mathcal{L}_{\mathbf{t}}$ need to be identified and estimated. However, it is difficult to consistently estimate the unknown parameters $\sigma(s)$ for all $s \in [0,1]$. As an alternative, we shall use a bootstrap method to approximate the limiting null distributions. When the errors are stationary, it is well known that the Phillips–Perron test or the augmented Dickey–Fuller test have size distortions in finite samples, even though they have been proven to work asymptotically. Bootstrap-based methods have been proposed to improve the finite sample performance. Psaradakis (2001), Chang and Park (2003), and Palm et al. (2008) used the sieve bootstrap (Kreiss, 1988) assuming an infinite order AR structure for the error process. Paparoditis and Politis (2003) applied the block bootstrap (Künsch, 1989), which randomly samples from overlapping blocks of residuals. Swensen (2003) and Parker et al. (2006) extended the stationary bootstrap (Politis and Romano, 1994) to unit root testing, where not only are overlapping blocks randomly chosen, but also the block size is chosen from a geometric distribution. Cavaliere and Taylor (2009) applied the wild bootstrap (Wu, 1986) for unit root M tests (Perron and Ng, 1996), which are modifications of the Phillips–Perron test.

To accommodate both heteroscedasticity and temporal dependence in the error, bootstrap-based methods have been developed by Cavaliere and Taylor (2008a) and Smeekes and Taylor (2012). In their articles, the error is assumed to be a linear process with heteroscedastic innovations as in equation (1), and the wild bootstrap (Wu, 1986) was used along with an AR sieve procedure to filter out the dependence in the error. The combination of an autoregressive sieve (or recolored) filter and wild bootstrap handles heteroscedasticity and serial correlation simultaneously, but its theoretical validity strongly depends upon the linear process assumption on the error. We speculate that this recolored wild bootstrap (RWB) will not work for PLS errors because the naive wild bootstrap can account for heteroscedasticity, but not for weak temporal dependence that is not completely filtered out after applying the AR sieve. The wild bootstrap works in the framework of (1), since temporal dependence is removed by the Phillips–Perron adjustment (Cavaliere and Taylor, 2008a) or the augmented Dickey–Fuller adjustment (i.e., the AR sieve) (Cavaliere and Taylor, 2009; Smeekes and Taylor, 2012). This type of removal is only possible under the assumption that the error is a heteroscedastic linear process, in which case the longrun variance $\sigma^2(s)$ can be factored into two parts; one part is due to heteroscedasticity in the innovations $e_t$, and the other part

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

is due to temporal dependence (i.e., $\sum_{j=0}^{\infty} c_j$) in the error, as shown recently by Rho and Shao (2015). As a result, limiting null distributions after the two popular adjustments for temporal dependence depend only on heteroscedasticity, which can be handled by the wild bootstrap. However, for PLS error processes, even after the Phillips–Perron or the augmented Dickey–Fuller adjustment, the limiting null distributions are still affected by temporal dependence in the error. Therefore, RWB is not expected to work in our setting.

To accommodate nonstationarity and temporal dependence in the error, we propose to adopt the so-called dependent wild bootstrap (DWB), which was first introduced by Shao (2010) in the context of stationary time series. It turns out that DWB is capable of mimicking local weak dependence in the error process and provides a consistent approximation of the limiting null distributions of $\mathbf{T}_n$ and $\mathbf{t}_n$. Note that DWB was developed for stationary time series and its applicability was only proved for smooth function models. Smeekes and Urbain (2014) recently proved the validity of several modified wild bootstrap methods, including DWB, for modulated stationary errors in a multivariate setting. However, as discussed in the introduction, the modulated stationary process is somewhat restrictive due to its separable structure of temporal dependence and heteroscedasticity in its longrun variance. Instead, our PLS framework is considerably more general in allowing both abrupt and smooth change in second- and higher-order properties. From a technical viewpoint, our proofs seem more involved than theirs due to the general error framework we adopt.

In the implementation of DWB, pseudo-residuals are generated by perturbing the original (OLS) residuals using a set $\{W_{t,n}\}_{t=1}^n$ of external variables. The difference between DWB and the original wild bootstrap is that $\{W_{t,n}\}_{t=1}^n$ is made to be dependent in DWB, whereas $\{W_{t,n}\}_{t=1}^n$ is assumed to be independent in the usual wild bootstrap. The following assumptions on $\{W_{t,n}\}_{t=1}^n$ are from Shao (2010):

(B1) $\{W_{t,n}\}_{t=1}^n$ is a realization from a stationary time series with $E(W_{t,n}) = 0$ and $\mathrm{var}(W_{t,n}) = 1$. $\{W_{t,n}\}_{t=1}^n$ are independent of the data, $\mathrm{cov}(W_{t,n}, W_{t',n}) = a\{(t-t')/l\}$, where $a(\cdot)$ is a kernel function and $l = l_n$ is a bandwidth parameter that satisfies $l \asymp Cn^{\kappa}$ for some $0 < \kappa < 1/3$. Assume that $W_{t,n}$ is $l$-dependent and $E(W_{1,n}^4) < \infty$.

(B2) $a : \mathbb{R} \to [0,1]$ is symmetric and has compact support on [-1,1], $a(0) = 1$, $\lim_{x\to 0}\{1 - a(x)\}/|x|^q = k_q \neq 0$ for some $q \in (0,2]$, and $\int_{-\infty}^{\infty} a(u)e^{-iux}du \geq 0$ for $x \in \mathbb{R}$.

In practice, $\{W_{t,n}\}_{t=1}^n$ can be sampled from a multivariate normal distribution with mean zero and covariance function $\mathrm{cov}(W_{t,n}, W_{t',n}) = a\{(t-t')/l\}$. There are two user-determined parameters: a kernel function $a(\cdot)$ and a bandwidth parameter $l$. The kernel function affects the performance to a lesser degree than the bandwidth parameter $l$, and the choice of $l$ will be discussed in Section 4. For the kernel function, some commonly used kernels, such as the Bartlett kernel, satisfy (B2).

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

The DWB algorithm in unit root testing is as follows:

ALGORITHM 3.1 [The dependent wild bootstrap (DWB)].

1. Calculate the OLS estimate $\widehat{\beta}_n$ of $\beta$ by fitting $y_{t,n}$ on $z_{t,n}$, and let $\widehat{X}_{t,n} = y_{t,n} - \widehat{\beta}_n' z_{t,n}$.
2. Let $\widehat{\rho}_n$ be the OLS estimate of $\widehat{X}_{t,n}$ on $\widehat{X}_{t-1,n}$. Calculate the statistics $\mathbf{T}_n = n(\widehat{\rho}_n - 1)$ and $\mathbf{t}_n = (\sum_{t=1}^n \widehat{X}_{t-1,n}^2)^{1/2}(\widehat{\rho}_n - 1)/s_n$.
3. Calculate the residuals $\widehat{u}_{t,n} = \widehat{X}_{t,n} - \widehat{\rho}_n \widehat{X}_{t-1,n}$ for all $t = 1, \ldots, n$.
4. Randomly generate the $l$-dependent mean-zero stationary series $\{W_{t,n}\}_{t=1}^n$ satisfying conditions (B1)–(B2) and generate the perturbed residuals $u_{t,n}^* = \widehat{u}_{t,n} W_{t,n}$.
5. Construct the bootstrapped sample $y_{t,n}^*$ using $u_{t,n}^*$ as if $\rho = 1$ is true:

   $$(y_{t,n}^* - \widehat{\beta}_n' z_{t,n}) = (y_{t-1,n}^* - \widehat{\beta}_n' z_{t-1,n}) + u_{t,n}^*, \qquad t = 2, \ldots, n,$$

   and $y_{1,n}^* = \widehat{\beta}_n' z_{1,n} + u_{1,n}^*$.
6. Calculate $\widehat{\beta}_n^*$ by refitting $y_{t,n}^*$ on $z_{t,n}$, and let $\widehat{X}_{t,n}^* = y_{t,n}^* - (\widehat{\beta}_n^*)' z_{t,n}$.
7. Calculate bootstrapped versions of $\widehat{\rho}_n$ and $s_n^2$, i.e., $\widehat{\rho}_n^*$ and $s_n^{*2}$, based on $\{\widehat{X}_{t,n}^*\}_{t=1}^n$, and the bootstrapped test statistics $\mathbf{T}_n^* = n(\widehat{\rho}_n^* - 1)$ and $\mathbf{t}_n^* = \{\sum_{t=1}^n (\widehat{X}_{t-1,n}^*)^2\}^{1/2}(\widehat{\rho}_n^* - 1)/s_n^*$.
8. Repeat steps 2–7 B times, and record the bootstrapped test statistics $\{\mathbf{T}_n^{*(1)}, \ldots, \mathbf{T}_n^{*(B)}\}$ and $\{\mathbf{t}_n^{*(1)}, \ldots, \mathbf{t}_n^{*(B)}\}$. The p-values are

   $$\frac{\sum_{b=1}^B \mathbf{1}\{\mathbf{T}_n^{*(b)} < \mathbf{T}_n\}}{B} \quad \text{and} \quad \frac{\sum_{b=1}^B \mathbf{1}\{\mathbf{t}_n^{*(b)} < \mathbf{t}_n\}}{B}.$$

**Remark 3.1.** Notice that the null hypothesis is not enforced in step 3 of Algorithm 3.1, i.e., unrestricted residuals are used in the construction of the bootstrap samples. Another approach constructing the bootstrap sample is discussed in Paparoditis and Politis (2003), where the null hypothesis is imposed in step 3. Both procedures are consistent under the null hypothesis, but as observed in Paparoditis and Politis (2003) for their residual block bootstrap, unrestricted residuals deliver higher power. The same phenomenon was also observed for DWB-based tests in our (unreported) simulations, so we shall not consider the restricted residual case in detail.

The following theorem provides the core result in the proof of the consistency of DWB in Theorem 3.2 and may be of independent interest.

THEOREM 3.1. *Assume* (A1)–(A4), (Z1)–(Z2), and (B1)–(B2). For any $\rho = 1 + c/n$, $c \le 0$,

$$n^{-1/2} \sum_{t=1}^{\lfloor nr \rfloor} u_{t,n}^* \Rightarrow B_{\sigma|Z}(r) \qquad \textit{in probability}.$$

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

Note that Theorem 3.1 holds not only under the null hypothesis $\rho = 1$ but also under local alternatives. This property makes the DWB method powerful because the bootstrapped distributions correctly mimic the limiting null distributions under both the null and local alternatives. The DWB method can still correctly approximate the limiting null distribution under local alternatives, mainly because $y^*_{t,n}$ are constructed assuming $\rho = 1$ in step 5.

THEOREM 3.2 (Bootstrap consistency and power). *Assume* (A1)–(A4), (Z1)–(Z2), and (B1)–(B2). For any $c \leq 0$,

$$P(\mathbf{T}_n \leq \mathbf{T}^*_{n,\alpha} | \rho = 1 + c/n) \xrightarrow{\mathcal{P}} P(L_{T,c} \leq \mathcal{L}^{\alpha}_{\mathbf{T}}),$$

$$P(\mathbf{t}_n \leq \mathbf{t}^*_{n,\alpha} | \rho = 1 + c/n) \xrightarrow{\mathcal{P}} P(L_{t,c} \leq \mathcal{L}^{\alpha}_{\mathbf{t}}),$$

where $L_{T,c}$ and $L_{t,c}$ are random variables with distribution $\mathcal{L}_{\mathbf{T},c}$ and $\mathcal{L}_{\mathbf{t},c}$, respectively, which are defined in Theorem 2.1. $\mathcal{L}^{\alpha}_{\mathbf{T}}$ and $\mathcal{L}^{\alpha}_{\mathbf{t}}$ are the $\alpha$-quantiles of the limiting null distributions, $\mathcal{L}_{\mathbf{T}}$ and $\mathcal{L}_{\mathbf{t}}$, respectively. $\mathbf{T}^*_{n,\alpha}$ and $\mathbf{t}^*_{n,\alpha}$ are the $\alpha$-quantiles of the distributions of $\mathbf{T}^*_n$ and $\mathbf{t}^*_n$ conditional on the data, respectively.

Under the null hypothesis, i.e., when $c = 0$, Theorem 3.2 establishes the consistency of DWB in approximating the limiting null distributions. Since the bootstrap statistics (asymptotically) replicate the exact null distribution when $\rho = 1$, the (asymptotic) size of our unit root test would be exactly the same as the level of the test. On the other hand, if $c$ is negative and far from 0, the probability of rejecting the null, or the asymptotic power of the test, will be close to 1. If $c$ is not 0 but not too far from 0, Theorem 3.2 states that the probability of rejecting the null is somewhere between the level of the test and 1. This means that the DWB-based unit root tests have nontrivial power under local alternatives.

**Remark 3.2.** The DWB method was originally developed for stationary time series (Shao, 2010). In the construction of DWB samples, $\{W_{t,n}\}_{t=1}^n$ is generated as $l$-dependent stationary time series, so it is natural to expect that DWB would work for stationary time series. However, it does not seem straightforward that this simple form of bootstrap would work in the case of a locally stationary process with unknown breaks. What Theorem 3.1 suggests is that DWB is capable of capturing nonstationary behaviors, without the need to specify any parametric forms of error structures or to know the specific form of nonstationarity such as the location of breaks.

## 4. SIMULATIONS

In this section, the DWB method is compared with the recolored (sieve) wild bootstrap (RWB) method, which was proposed in Cavaliere and Taylor (2009, Sect. 3.3). We also propose to combine the AR sieve idea in RWB with DWB and present this method as the recolored dependent wild bootstrap (RDWB). The

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

RDWB statistics are based on the RWB statistics using DWB to determine the critical values of the tests, so RDWB can be considered as a generalization of RWB.

Before introducing our simulation setting and results, we first present some details about (i) the RDWB algorithm and (ii) the size-corrected power calculation similar to the one in Domínguez and Lobato (2001).

First, the RDWB procedure is described below. Rewrite equation (3) as

$$\Delta X_{t,n} = \pi_0 X_{t-1,n} + \sum_{i=1}^{k} \pi_i \Delta X_{t-i,n} + u_{t,n,k}, \tag{8}$$

where $\Delta$ represents the difference operator.

ALGORITHM 4.1 [The recolored dependent wild bootstrap (RDWB)].

1. Calculate the OLS estimate $\widehat{\beta}_n$ of $\beta$ by fitting $y_{t,n}$ on $z_{t,n}$, and let $\widehat{X}_{t,n} = y_{t,n} - \widehat{\beta}_n' z_{t,n}$.
2. Choose the number $k$ of lags using, for example, the modified Akaike information criterion (MAIC) in Ng and Perron (2001, p. 1529). That is, $\hat{k} = \text{argmin}_{0\le k\le k_{max}} \text{MAIC}(k)$, where $\text{MAIC}(k) = \ln(\hat{\sigma}_k^2) + 2\{\tau_n(k)+k\}/(n-k_{max})$, $k_{max} = \lfloor 12(n/100)^{1/4} \rfloor$, $\tau_n(k) = (\hat{\sigma}_k^2)^{-1}\hat{\pi}_0^2 \sum_{t=k_{max}+1}^{n} \widehat{X}_{t-1,n}^2$, and $\hat{\sigma}_k^2 = (n-k_{max})^{-1}\sum_{t=k_{max}+1}^{n} \hat{u}_{t,n,k}^2$. Here, $\hat{\pi}_0$ and $\hat{u}_{t,n,k}$ are the OLS estimators and residuals from (8) with $k = 0, \ldots, k_{max}$. Find the OLS estimators and residuals, i.e., $\hat{\pi}_i$ and $\hat{u}_{t,n,\hat{k}}$ from (8) with $k = \hat{k}$.
3. Let $\widehat{\rho}_n$ be the OLS estimate of $\widehat{X}_{t,n}$ on $\widehat{X}_{t-1,n}$. Calculate the statistics $\mathbf{T}_n = n(\widehat{\rho}_n - 1)$ and $\mathbf{t}_n = (\sum_{t=1}^{n} \widehat{X}_{t-1,n}^2)^{1/2}(\widehat{\rho}_n - 1)/s_n$.
4. Generate the $l$-dependent mean-zero stationary series $\{W_{t,n}\}_{t=1}^{n}$ satisfying conditions (B1)–(B2) and generate the perturbed residuals $u_{t,n,\hat{k}}^* = \hat{u}_{t,n,\hat{k}} W_{t,n}$.
5. Construct the bootstrapped sample $y_{t,n}^*$ using $\{u_{t,n}^*\}$ under the unit root null hypothesis in (8), i.e., $\pi_0 = 0$, and recolor the bootstrapped residuals:

$$\Delta(y_{t,n}^* - \widehat{\beta}_n' z_{t,n}) = \sum_{i=1}^{\hat{k}} \hat{\pi}_i \Delta(y_{t-i,n}^* - \widehat{\beta}_n' z_{t-i,n}) + u_{t,n,\hat{k}}^*, \qquad t = \hat{k}, \ldots, n,$$

   and $y_{t,n}^* = \widehat{\beta}_n' z_{t,n} + u_{t,n,\hat{k}}^*$ for $t = 1, \ldots, \hat{k}-1$.
6. Calculate $\widehat{\beta}_n^*$ by refitting $y_{t,n}^*$ on $z_{t,n}$, and let $\widehat{X}_{t,n}^* = y_{t,n}^* - (\widehat{\beta}_n^*)' z_{t,n}$.
7. Calculate the bootstrapped versions of $\widehat{\rho}_n$ and $s_n^2$, i.e., $\widehat{\rho}_n^*$ and $s_n^{*2}$, based on $\{\widehat{X}_{t,n}^*\}_{t=1}^{n}$, and the bootstrapped test statistics $\mathbf{T}_n^* = n(\widehat{\rho}_n^* - 1)$ and $\mathbf{t}_n^* = \{\sum_{t=1}^{n}(\widehat{X}_{t-1,n}^*)^2\}^{1/2}(\widehat{\rho}_n^* - 1)/s_n^*$.

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

8. Repeat steps 2–7 B times, and record the bootstrapped test statistics $\{\mathbf{T}_n^{*(1)}, \ldots, \mathbf{T}_n^{*(B)}\}$ and $\{\mathbf{t}_n^{*(1)}, \ldots, \mathbf{t}_n^{*(B)}\}$. The p-values are
$$\frac{\sum_{b=1}^{B} \mathbf{1}\{\mathbf{T}_n^{*(b)} < \mathbf{T}_n\}}{B} \quad \text{and} \quad \frac{\sum_{b=1}^{B} \mathbf{1}\{\mathbf{t}_n^{*(b)} < \mathbf{t}_n\}}{B}.$$

**Remark 4.1.** If $l = 1$, or equivalently, if an i.i.d. sequence $W_t$ is used in RDWB step 4, then the above-described procedure coincides with RWB in Cavaliere and Taylor (2009).

**Remark 4.2.** In the above procedure, the number $\hat{k}$ of lags is optimized for the original data $X_{t,n}$, and the same $\hat{k}$ is used for the bootstrapped data $X_{t,n}^*$. In general, the number of lags for the original data and that for the bootstrapped data do not have to be the same. For example, for the bootstrap, it can be chosen to be optimized for each bootstrapped sample; that is,

$$k_b^* = \text{argmin}_{0 \le k \le k_{max}} \text{MAIC}^*(k),$$

where $\text{MAIC}^*(k) = \ln(\hat{\sigma}_k^{2*}) + \frac{2(\tau_n^*(k)+k)}{n-k_{max}}$, $\tau_n^*(k) = (\hat{\sigma}_k^{2*})^{-1}\hat{\pi}_0^{2*}\sum_{t=k_{max}+1}^{n}(\widehat{X}_{t-1,n}^*)^2$, $\hat{\sigma}_k^{2*} = (n - k_{max})^{-1}\sum_{t=k_{max}+1}^{n}\hat{u}_{t,n,k}^{2*}$. However, we shall keep the same $k$ for both the original and the bootstrapped data based on the finite sample findings reported in Remark 3 of Cavaliere and Taylor (2009).

For a fair comparison of power, the following size-corrected power procedure similar to Domínguez and Lobato (2001) is adapted.

ALGORITHM 4.2 [Size-corrected power of a bootstrap test]. Consider a level $\alpha$ test with $\mathbf{T}_n$ for a simple exposition. The unit root null hypothesis is rejected if $\mathbf{T}_n < \mathcal{L}_{\mathbf{T},0,\alpha}$, where $\mathcal{L}_{\mathbf{T},0,\alpha}$ indicates the $\alpha$-quantile of $\mathcal{L}_{\mathbf{T},0}$ in Theorem 2.1.

1. Estimate the finite sample counterpart $\widehat{\mathcal{L}}_{\mathbf{T},0,\alpha}$ of $\mathcal{L}_{\mathbf{T},0,\alpha}$ based on $N$ Monte-Carlo replications. Let $N$ be large enough so that $N\alpha$ is an integer. That is, if $\{\mathbf{T}_n^{(1)}, \mathbf{T}_n^{(2)}, \ldots, \mathbf{T}_n^{(N)}\}$ indicates the set of test statistics of $N$ Monte-Carlo replications and $\{\mathbf{T}_n^{[1]}, \mathbf{T}_n^{[2]}, \ldots, \mathbf{T}_n^{[N]}\}$ indicates its ordered version from the smallest to largest, $\widehat{\mathcal{L}}_{\mathbf{T},0,\alpha} = \mathbf{T}_n^{[N\alpha]}$. Note that using this infeasible critical value $\widehat{\mathcal{L}}_{\mathbf{T},0,\alpha}$, the empirical size should be similar to the nominal level $\alpha$.
2. For each Monte-Carlo replication under the null hypothesis, generate $B$ bootstrap samples and calculate the corresponding bootstrap test statistics $\{\mathbf{T}_n^{*(b,i)}\}_{b=1}^{B}$, $i = 1, \ldots, N$. Calculate the empirical size of the bootstrap test $\alpha^{(i)}$ of the $i$th Monte-Carlo replication using the infeasible critical value $\widehat{\mathcal{L}}_{\mathbf{T},0,\alpha}$, i.e.,
$$\alpha^{(i)} = B^{-1}\sum_{b=1}^{B} \mathbf{1}\left(\mathbf{T}_n^{*(b,i)} < \widehat{\mathcal{L}}_{\mathbf{T},0,\alpha}\right).$$

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

The size-corrected level $\alpha^c$ is the average of the $\alpha^{(i)}$, that is, $\alpha^c = N^{-1}\sum_{i=1}^{N}\alpha^{(i)}$.

3. For another set of statistics of $N$ Monte-Carlo replications under the (local) alternative, the size-corrected power is calculated replacing $\alpha$ with its size-corrected version, $\alpha^c$. That is,

$$N^{-1}\sum_{i=1}^{N}\left(\mathbf{T}_n^{(i)} < \mathbf{T}_{n,\alpha^c}^{*(i)}\right),$$

where $\mathbf{T}_{n,\alpha^c}^{*(i)}$ is the $\alpha^c$-quantile of the bootstrapped statistics $\{\mathbf{T}_n^{*(b,i)}\}_{b=1}^{B}$ for the $i$th Monte-Carlo replication.

The following data generating processes (DGPs) are used for comparison of DWB, RWB, and RDWB in finite samples. For simplicity, set $\beta \equiv 0$ so that $\widehat{X}_{t,n} = X_{t,n}$. Consider (3) and $u_{t,n}$ generated from time-varying moving average (MA) and autoregressive (AR) models with lag 1,

$$(\mathrm{MA}_{i,j})\ u_{t,n} = e_{j,t,n} + \phi_i(t/n)e_{j,t-1,n}, \quad (\mathrm{AR}_{i,j})\ u_{t,n} = e_{j,t,n} + \phi_i(t/n)u_{t-1,n}$$

for $t = 1,\ldots,n$, where $e_{j,t,n} = \omega_j(t/n)\varepsilon_t$, $\varepsilon_t \overset{i.i.d.}{\sim} N(0,1)$. The MA or AR coefficient $\phi_i(s)$ is possibly time-varying with the following six choices: for $s \in [0,1]$,

$$\phi_1(s) = 0.8,\ \phi_2(s) = -0.8,\ \phi_3(s) = 0.2 + 0.6\mathbf{1}(s > 0.2),$$

$$\phi_4(s) = 0.2 + 0.6\mathbf{1}(s > 0.8),\ \phi_5(s) = 0.8 - 1.6s,\ \text{and}\ \phi_6(s) = 0.6s - 0.8.$$

The function $\omega_j(s)$ governs possible heteroscedastic behavior in $u_{t,n}$ with the following five choices: for $s \in [0,1]$,

$$\omega_1(s) = 0.5,\ \omega_2(s) = 0.1 + 0.5\mathbf{1}(s > 0.1),\ \omega_3(s) = 0.1 + 0.5\mathbf{1}(s > 0.9),$$

$$\omega_4(s) = 0.1 + 0.5\mathbf{1}(0.4 < s < 0.6),\ \text{and}\ \omega_5(s) = 0.5s + 0.1.$$

Combinations of $\phi_i(s)$ and $\omega_j(s)$ along with the choice of MA or AR lead to 60 DGPs that satisfy the PLS assumption in (A1)–(A4). In particular, if $i = 1$ or 2, $\phi_i(s)$ is constant over $s \in [0,1]$. The corresponding $\{u_{t,n}\}$ processes fall into the category of linear processes with heteroscedastic error in (1), making RWB consistent for any choices of $\omega_j(s)$, $j = 1,\ldots,5$. These settings are to mirror the setup of the Cavaliere–Taylor articles. For all other settings, the asymptotic consistency of RWB is not guaranteed, whereas DWB and RDWB are expected to work asymptotically. Sudden increases and smooth changes in MA or AR coefficients are presented in the cases with $i = 3,4$ and $i = 5,6$, respectively. The variance of $e_{j,t,n}$ is a constant ($j = 1$), a step function with a sudden increase in the beginning ($j = 2$) and end ($j = 3$) of the series, a step function with a sudden

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

increase and decrease in the middle ($j = 4$), or a smoothly increasing sequence ($j = 5$).

The sample sizes $n = 100$ and 400 are considered. The number of Monte-Carlo replications is 2,000, and the number of bootstrap replications is $B = 1{,}000$ for all bootstrap methods. For local alternatives, $c = 0, -5, -10, -15, -20, -25, -30$ are considered. In particular, for DWB and RDWB, in each replication, pseudoseries $(W_{1,n}, \ldots, W_{n,n})'$ are generated from i.i.d. $N(\mathbf{0}_n, \Sigma)$, where $\Sigma$ is an $n$ by $n$ matrix with its $(i, j)$th element being $a\{(i - j)/l\}$. Here the Bartlett kernel is used, i.e., $a(s) = (1 - |s|)\mathbf{1}(|s| \leq 1)$. For DWB and RDWB, the bandwidth parameter $l$ is chosen as $l = \lfloor 6(n/100)^{1/4} \rfloor$. That is, $l = 6$ if $n = 100$, and $l = 8$ if $n = 400$. In Section B of the supplementary material, (i) full details on the effect of different choices of $l$ for selected DGPs are presented and (ii) a data-driven choice of $l$, the minimum volatility method, is proposed. It seems that the empirical sizes are not overly sensitive to the choice of $l$, as long as $l$ is not too small, and the finite sample size comparison with the MV method in Section B of supplementary material supports the above deterministic choice of $l$.

Tables 1 and 2 present the empirical sizes of the three methods when the nominal size is 5%. When the model is stationary with positive coefficient $\phi(s) = \phi_1(s) = 0.8$, i.e., $(\mathrm{MA}/\mathrm{AR}_{1,j})$ for $j = 1, \ldots, 5$, all three bootstrap methods produce reasonably accurate sizes, except that the DWB method tends to under-reject for the AR models. This under-rejecting behavior of DWB is observed consistently for most AR models. This might be due to the fact that the DWB method mimics the time dependence in the original data in a manner similar to MA models, so that it does not produce as accurate sizes for AR models as for MA models. The RDWB method nicely compensates this shortcoming by applying an AR-based prewhitening. The prewhitening effect is most noticeable when the model is stationary with negative coefficient $\phi(s) = \phi_2(s) = -0.8$, i.e., $(\mathrm{MA}/\mathrm{AR}_{2,j})$ for $j = 1, \ldots, 5$. For these models with negative autocorrelation, the size-distortion of the DWB method is very large at both sample sizes with slight less distortion for larger sample size. This suggests that although the DWB method should work asymptotically for the negative coefficient case, this convergence could be too slow to be useful in practice. On the other hand, after applying the AR-based prewhitening, similar to RWB, finite sample sizes are brought closer to the nominal level.

A careful examination of RWB shows that it has a fairly accurate size, especially when it is theoretically supported ($i = 1, 2$). However, for some DGPs with changing MA or AR coefficients ($i = 3, 4, 5, 6$), RWB does not seem to be consistent. In particular, in the MA models, the sizes of RWB tend to further deviate from the nominal level as the sample size $n$ increases when there is a sudden increase in the variance in innovations at the latter part of the series ($j = 3$) with changing variance (see $(\mathrm{MA}_{3,3})$, $(\mathrm{MA}_{4,3})$, $(\mathrm{MA}_{5,3})$, and $(\mathrm{MA}_{6,3})$) or when both MA coefficient and variance of innovations change smoothly (see $(\mathrm{MA}_{5,5})$). In the AR models, RWB tends to have heavier size distortion as $n$ increases when the AR coefficient changes drastically from negative to positive

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

**TABLE 1.** Empirical sizes for DWB, RWB, and RDWB for MA models with $\phi_i(s)$ and $\omega_j(s)$ based on 2,000 Monte-Carlo replications and 1,000 bootstrap replications under $\rho = 1$. The nominal level is 5%

| | | $n = 100$ | | | | | | $n = 400$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | DWB | | RWB | | RDWB | | DWB | | RWB | | RDWB | |
| $i$ | $j$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ |
| | 1 | 2.9 | 3.1 | 4.8 | 5.1 | 4.7 | 4.7 | 3.9 | 3.8 | 4.9 | 4.5 | 4.5 | 4.4 |
| | 2 | 3.2 | 3.2 | 4.5 | 4.5 | 5.0 | 4.9 | 3.1 | 3.3 | 4.4 | 4.5 | 4.5 | 4.5 |
| 1 | 3 | 4.3 | 4.8 | 4.2 | 4.9 | 4.9 | 5.3 | 4.5 | 4.3 | 5.3 | 5.5 | 5.9 | 5.9 |
| | 4 | 4.3 | 4.3 | 3.5 | 3.6 | 5.3 | 5.3 | 4.3 | 4.5 | 5.5 | 5.7 | 5.9 | 6.0 |
| | 5 | 2.8 | 2.9 | 3.9 | 4.0 | 4.1 | 4.0 | 3.5 | 3.6 | 4.5 | 5.0 | 4.9 | 5.0 |
| | 1 | 82.2 | 81.9 | 20.5 | 20.4 | 19.8 | 20.0 | 78.5 | 78.4 | 12.8 | 12.8 | 12.2 | 12.0 |
| | 2 | 87.2 | 86.8 | 22.9 | 23.1 | 22.3 | 22.5 | 80.3 | 80.2 | 11.1 | 11.0 | 10.6 | 10.7 |
| 2 | 3 | 95.5 | 94.5 | 21.6 | 20.4 | 21.4 | 20.6 | 96.2 | 95.8 | 13.0 | 13.0 | 12.4 | 12.4 |
| | 4 | 80.9 | 80.7 | 17.4 | 17.4 | 18.4 | 18.3 | 73.5 | 73.3 | 10.8 | 10.8 | 11.1 | 11.2 |
| | 5 | 91.8 | 91.2 | 23.9 | 23.9 | 23.4 | 23.5 | 87.3 | 87.1 | 13.6 | 13.6 | 12.8 | 12.8 |
| | 1 | 3.5 | 3.7 | 5.5 | 5.1 | 5.0 | 5.0 | 3.5 | 3.5 | 4.9 | 4.8 | 4.7 | 4.5 |
| | 2 | 4.0 | 3.8 | 5.1 | 5.0 | 5.2 | 5.0 | 3.7 | 3.6 | 5.3 | 5.5 | 4.8 | 5.1 |
| 3 | 3 | 5.2 | 5.4 | 5.3 | 5.9 | 6.1 | 6.3 | 4.7 | 4.8 | 7.8 | 7.8 | 7.4 | 7.1 |
| | 4 | 4.7 | 5.0 | 5.0 | 5.3 | 5.5 | 5.8 | 4.2 | 4.0 | 6.8 | 6.8 | 6.2 | 6.3 |
| | 5 | 4.2 | 4.3 | 6.2 | 5.9 | 6.0 | 5.8 | 4.1 | 3.8 | 6.3 | 5.8 | 5.2 | 5.2 |
| | 1 | 4.0 | 4.0 | 4.4 | 4.3 | 5.0 | 4.9 | 3.6 | 4.0 | 4.9 | 4.8 | 4.6 | 4.7 |
| | 2 | 4.0 | 4.0 | 5.1 | 4.7 | 5.2 | 5.0 | 3.2 | 3.4 | 4.2 | 4.0 | 4.0 | 4.2 |
| 4 | 3 | 6.0 | 6.7 | 4.9 | 6.4 | 5.3 | 6.6 | 4.9 | 5.2 | 9.5 | 9.0 | 7.7 | 7.5 |
| | 4 | 5.7 | 5.7 | 2.8 | 2.8 | 4.8 | 4.9 | 5.5 | 5.5 | 5.5 | 5.2 | 6.2 | 6.2 |
| | 5 | 3.5 | 3.1 | 4.1 | 3.5 | 4.7 | 4.2 | 3.6 | 3.7 | 5.4 | 5.2 | 4.9 | 4.7 |
| | 1 | 7.4 | 7.6 | 6.6 | 6.5 | 6.2 | 6.4 | 6.2 | 5.9 | 8.3 | 7.9 | 5.7 | 5.2 |
| | 2 | 5.9 | 6.1 | 5.1 | 4.9 | 5.1 | 5.2 | 6.7 | 6.9 | 7.3 | 7.3 | 6.2 | 6.5 |
| 5 | 3 | 8.5 | 8.0 | 7.8 | 10.3 | 7.3 | 9.2 | 5.9 | 6.5 | 16.4 | 15.2 | 10.1 | 10.0 |
| | 4 | 8.7 | 8.6 | 4.1 | 4.3 | 5.5 | 5.5 | 7.4 | 7.5 | 6.6 | 6.3 | 5.7 | 5.9 |
| | 5 | 5.1 | 5.1 | 6.4 | 6.2 | 5.3 | 5.5 | 4.3 | 4.5 | 9.2 | 8.8 | 5.1 | 5.1 |
| | 1 | 33.6 | 33.1 | 11.3 | 11.1 | 10.0 | 9.5 | 25.4 | 24.9 | 10.5 | 10.5 | 8.0 | 8.1 |
| | 2 | 30.7 | 30.8 | 10.7 | 10.2 | 9.8 | 9.9 | 23.4 | 23.2 | 8.2 | 8.1 | 6.8 | 7.0 |
| 6 | 3 | 35.4 | 33.7 | 12.8 | 12.6 | 12.1 | 12.2 | 25.9 | 25.5 | 16.2 | 15.3 | 11.6 | 11.1 |
| | 4 | 33.1 | 33.1 | 7.9 | 7.8 | 8.6 | 8.6 | 26.5 | 26.2 | 6.9 | 6.8 | 6.7 | 6.7 |
| | 5 | 27.3 | 26.7 | 11.4 | 10.5 | 9.8 | 9.7 | 20.0 | 19.7 | 9.2 | 8.8 | 6.7 | 6.7 |

($i = 5$; see ($AR_{5,1}$), ($AR_{5,2}$), ($AR_{5,3}$), and ($AR_{5,5}$)) or when the AR coefficient is negative and changes smoothly and the variance in innovations suddenly increases at the latter part of the series (see ($AR_{6,3}$)). This size distortion might be an indication that the AR prewhitening (RWB) alone does not work in theory, and the dependence in the error is not completely filtered out. By contrast, RDWB tends to have more accurate sizes for these models, although size distortion due to inaccurate prewhitening is still apparent to a lesser degree. On the other hand, as long as the MA or AR coefficients are nonnegative, DWB without prewhitening is always demonstrated to have more accurate size as $n$ increases. In particular, for MA models with changing MA coefficient ($i = 3, 4, 5$), DWB tends to produce the best size with the most consistent behavior among the three bootstrap methods.

Overall, the size for RDWB seems to be the most reliable among the three bootstrap methods if the underlying DGP is not known. In some unreported

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

**TABLE 2.** Empirical sizes for DWB, RWB, and RDWB for AR models with $\phi_i(s)$ and $\omega_j(s)$ based on 2,000 Monte-Carlo replications and 1,000 bootstrap replications under $\rho = 1$. The nominal level is 5%

| | | $n = 100$ | | | | | | $n = 400$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | DWB | | RWB | | RDWB | | DWB | | RWB | | RDWB | |
| $i$ | $j$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ | $T_n$ | $t_n$ |
| | 1 | 0.2 | 0.5 | 3.8 | 4.2 | 4.0 | 4.2 | 0.8 | 0.8 | 3.5 | 3.8 | 3.7 | 3.9 |
| | 2 | 0.8 | 0.6 | 4.0 | 4.3 | 4.6 | 4.7 | 0.4 | 0.6 | 3.8 | 3.7 | 4.0 | 4.1 |
| 1 | 3 | 0.5 | 0.6 | 5.5 | 6.9 | 5.9 | 7.2 | 0.7 | 1.1 | 3.8 | 4.6 | 4.3 | 5.1 |
| | 4 | 0.3 | 0.4 | 1.8 | 1.9 | 2.8 | 2.9 | 1.0 | 1.1 | 5.1 | 5.0 | 5.6 | 5.6 |
| | 5 | 0.4 | 0.4 | 3.2 | 3.3 | 3.5 | 3.8 | 0.5 | 0.7 | 4.0 | 3.9 | 4.1 | 4.0 |
| | 1 | 45.1 | 45.0 | 7.8 | 7.6 | 7.5 | 7.7 | 32.9 | 33.0 | 5.3 | 5.3 | 5.7 | 5.5 |
| | 2 | 45.1 | 45.1 | 6.2 | 6.0 | 7.3 | 7.2 | 32.9 | 33.0 | 5.2 | 5.2 | 5.5 | 5.5 |
| 2 | 3 | 62.4 | 61.5 | 10.4 | 9.6 | 11.0 | 10.6 | 53.2 | 52.4 | 9.7 | 9.3 | 9.8 | 9.8 |
| | 4 | 44.1 | 44.2 | 5.5 | 5.5 | 6.6 | 6.7 | 34.8 | 34.8 | 6.4 | 6.4 | 6.8 | 6.8 |
| | 5 | 51.7 | 51.5 | 7.6 | 7.4 | 9.0 | 9.0 | 41.4 | 41.2 | 5.5 | 5.7 | 6.0 | 6.0 |
| | 1 | 0.5 | 0.6 | 4.8 | 5.0 | 4.0 | 4.3 | 0.9 | 1.0 | 5.8 | 5.8 | 4.8 | 4.9 |
| | 2 | 0.4 | 0.5 | 4.5 | 4.5 | 4.1 | 4.1 | 0.7 | 0.7 | 4.1 | 3.9 | 3.5 | 3.6 |
| 3 | 3 | 0.7 | 0.8 | 7.3 | 9.8 | 6.8 | 8.9 | 1.1 | 1.1 | 5.2 | 5.9 | 4.9 | 5.7 |
| | 4 | 0.2 | 0.4 | 3.2 | 3.2 | 3.3 | 3.5 | 0.9 | 1.0 | 6.2 | 5.8 | 5.5 | 5.3 |
| | 5 | 0.5 | 0.9 | 3.8 | 4.2 | 3.5 | 4.0 | 0.9 | 0.8 | 5.1 | 5.0 | 4.9 | 4.7 |
| | 1 | 1.7 | 1.7 | 3.5 | 3.5 | 3.6 | 3.7 | 2.1 | 2.0 | 6.5 | 5.7 | 5.1 | 4.8 |
| | 2 | 1.7 | 1.7 | 4.0 | 4.1 | 3.9 | 3.9 | 1.8 | 1.8 | 6.3 | 6.3 | 5.4 | 5.2 |
| 4 | 3 | 1.5 | 1.5 | 6.1 | 11.9 | 6.0 | 11.3 | 1.8 | 2.1 | 9.4 | 10.5 | 7.4 | 8.6 |
| | 4 | 3.4 | 3.3 | 2.1 | 2.0 | 3.6 | 3.6 | 3.7 | 3.8 | 3.9 | 3.8 | 5.4 | 5.1 |
| | 5 | 1.1 | 1.0 | 4.3 | 4.3 | 3.6 | 3.8 | 1.2 | 1.4 | 8.8 | 8.0 | 5.4 | 5.1 |
| | 1 | 3.1 | 3.4 | 6.6 | 6.3 | 5.2 | 5.3 | 2.8 | 3.0 | 9.3 | 8.3 | 5.5 | 5.0 |
| | 2 | 3.1 | 2.9 | 4.2 | 4.2 | 3.9 | 3.7 | 3.5 | 3.8 | 9.2 | 8.7 | 6.0 | 6.3 |
| 5 | 3 | 3.2 | 3.1 | 7.5 | 15.2 | 6.9 | 14.1 | 1.8 | 2.0 | 13.5 | 14.0 | 9.2 | 10.5 |
| | 4 | 6.6 | 6.3 | 3.1 | 3.0 | 4.5 | 4.2 | 5.8 | 5.7 | 5.8 | 5.4 | 5.8 | 5.5 |
| | 5 | 2.1 | 2.2 | 6.4 | 6.3 | 4.2 | 4.3 | 2.3 | 2.2 | 11.2 | 10.3 | 6.2 | 6.4 |
| | 1 | 20.8 | 20.6 | 5.8 | 5.6 | 6.3 | 6.2 | 15.2 | 15.3 | 5.5 | 5.5 | 5.5 | 5.5 |
| | 2 | 18.4 | 18.6 | 6.0 | 5.9 | 6.5 | 6.5 | 13.2 | 13.3 | 5.2 | 5.0 | 4.8 | 4.8 |
| 6 | 3 | 26.1 | 25.4 | 9.4 | 9.3 | 9.3 | 9.3 | 17.3 | 16.8 | 10.9 | 10.1 | 8.8 | 8.6 |
| | 4 | 23.4 | 23.6 | 4.7 | 4.8 | 6.4 | 6.4 | 16.2 | 16.2 | 5.5 | 5.5 | 5.9 | 5.8 |
| | 5 | 18.9 | 18.2 | 5.7 | 5.7 | 6.2 | 6.5 | 12.7 | 12.2 | 5.7 | 5.4 | 5.1 | 5.0 |

simulations, we have observed the following: (i) the large size distortion associated with the DWB method for negative autocorrelation models, $(\mathrm{MA}/\mathrm{AR}_{2,1})$, can be reduced to below the nominal 5% level if we use restricted residuals, at the price of power loss; (ii) a comparison with residual block bootstrap in Paparoditis and Politis (2003) shows that the size for the residual block bootstrap can be quite distorted for some DGPs, e.g., $(\mathrm{MA}_{5,5})$. This indicates the inability of residual block bootstrap to consistently approximate the limiting null distribution when the error process is PLS.

Figures 1 and 2 present the power curves of $\mathbf{t}_n$ for DWB, RWB, and RDWB for selected DGPs with $n = 100$ and 400, respectively. The size-adjusted power curves in the first panel, $(\mathrm{MA}_{4,1})$, are representative for most of the cases where all three bootstrap methods have reasonably accurate sizes, where (i) DWB tends to have the best power, (ii) RDWB tends to have slightly better power than RWB when $n = 100$, and (iii) RWB and RDWB are fairly comparable in terms of sizes

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

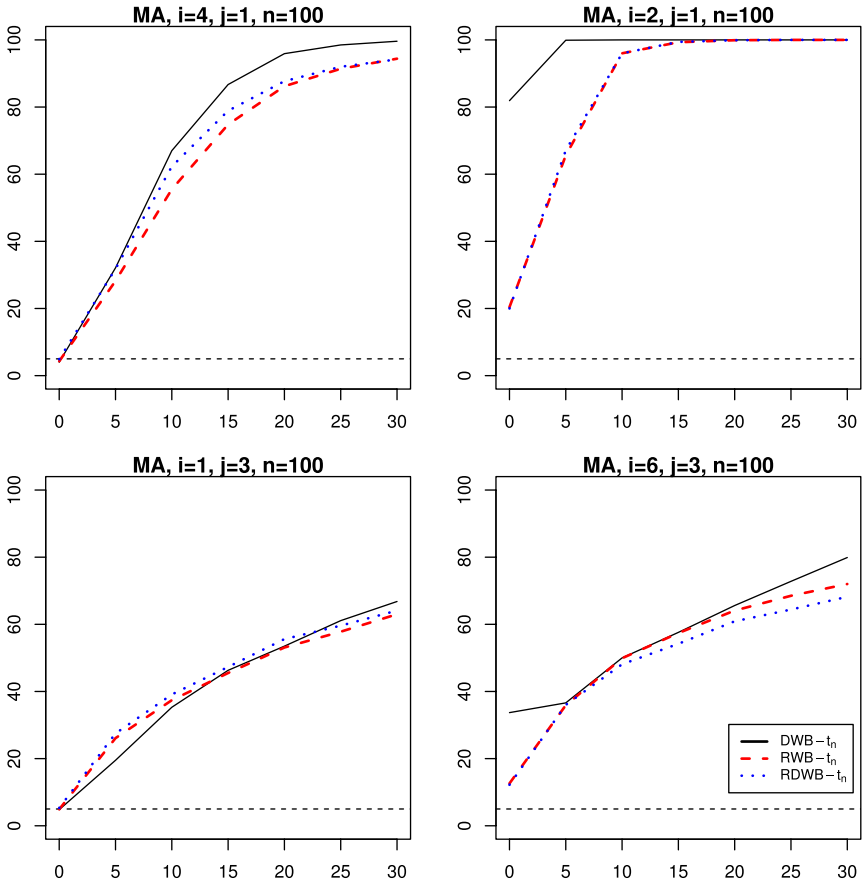



**FIGURE 1.** Rejection frequencies (%) versus $-c$, where $\rho = 1 + c/n$ for DWB, RWB, and RDWB unit root tests when the error $u_{t,n}$ is generated following the MA processes with $\phi_i(s)$ and $\omega_j(s)$ for selected $(i, j)$s. The plots present empirical sizes when $c = 0$ and size-adjusted powers when $c \neq 0$. 2,000 Monte-Carlo replications and 1,000 bootstrap replications are used. The sample size is $n = 100$ and the nominal level is 5%.

and powers in general. Most MA or AR models with positive MA or AR coefficient for at least part of a series ($\phi_i(s)$ with $i = 1, 3, 4, 5$) and constant, early break, or smooth change in error variance ($\omega_j(s)$ with $j = 1, 2, 5$) tend to have a similar shape. The second panel, ($MA_{2,1}$), represents the size-adjusted curves when the MA or AR coefficients are negative at all time points ($i = 2, 6$) so that the finite sample size of DWB is highly distorted. Even though DWB has the best power, it is not recommended due to its big size distortion in this case. It seems that RWB and RDWB do not have much difference in terms of size-adjusted power.

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

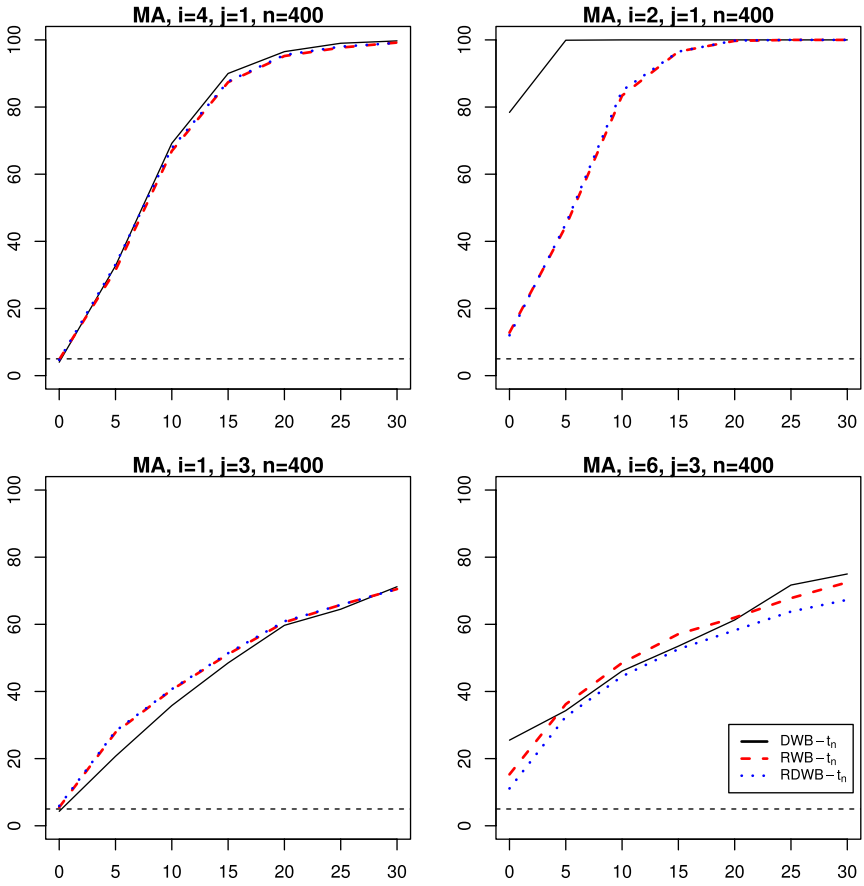



**FIGURE 2.** Rejection frequencies (%) versus $-c$, where $\rho = 1 + c/n$ for DWB, RWB, and RDWB unit root tests when the error $u_{t,n}$ is generated following the MA processes with $\phi_i(s)$ and $\omega_j(s)$ for selected $(i, j)$s. The plots present empirical sizes when $c = 0$ and size-adjusted powers when $c \neq 0$. 2,000 Monte-Carlo replications and 1,000 bootstrap replications are used. The sample size is $n = 400$ and the nominal level is 5%.

The last two panels focus on the comparison between RWB and RDWB. The third panel, $(\text{MA}_{1,3})$, is representative when there is a jump in $\omega(s)$ at the end of the series ($j = 3$) and when both RWB and RDWB have reasonable sizes. Models with $i = 1, 2, 6$ and $j = 3, 4$ tend to have a similar pattern if DWB is ignored due to its high size distortion when $i = 2, 6$. In this case, RWB and RDWB have similar powers, as RWB tends to have slightly better power when $i = 5$ or 6, whereas RDWB tends to have slightly higher power when $i = 1$. The last panel, $(\text{MA}_{6,3})$, is representative for the case when RWB is not consistent. $(\text{AR}_{i,j})$ or $(\text{MA}_{i,j})$ with $i = 3, 4, 5$ and $j = 3, 4$ fall into this category. In this case, RDWB seems to present

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

the most reasonable size and power. Even though RWB appears to have the best power, RWB does not seem to be consistent due to the considerable increase in its finite sample size as $n$ increases for some models. Complete power curves for all DGPs are presented in the supplementary material as Figures C.1–C.4. It is worth noting that $\mathbf{t}_n$ tends to produce more accurate sizes with little power loss (or slightly better power) than $\mathbf{T}_n$.

In summary, RDWB, the combination of RWB and DWB, appears to work well in finite samples. It tends to produce reasonably high powers and fairly accurate sizes in all models under examination. In the situation when DWB or RWB have a large size distortion, the size accuracy of RDWB is well maintained and its power appears quite reasonable in all cases. One downside associated with RDWB is that it requires two tuning parameters: the truncation lag in the AR sieve and the bandwidth parameter in DWB. In this article, we choose the number of lags for RWB and RDWB using the MAIC method. As for the bandwidth parameter, it seems that DWB and RDWB are not sensitive to the choice of the bandwidth parameter and the proposed deterministic choice seems to perform reasonably well in finite samples. Given that the DGP is unknown in practice, we shall recommend the use of RDWB.

## 5. CONCLUSION

In this article, we present a new bootstrap-based unit root testing procedure that is robust to changing second- and higher-order properties in the error process. The error is modeled as a piecewise locally stationary (PLS) process, which is general enough to include time-varying nonlinear processes as well as heteroscedastic linear processes as special cases. In particular, the PLS process does not impose a separable structure on its longrun variance as do heteroscedastic linear processes and modulated stationary processes, which have been adopted in the literature to model heteroscedasticity and weak dependence of the error. Under the PLS framework, the limiting null distributions of two popular test statistics are derived and the dependent wild bootstrap (DWB) method is used to approximate these nonpivotal distributions. The functional central limiting theorem has been established for the standardized partial sum process of the DWB residuals, and bootstrap consistency is justified under local alternatives. The DWB-based unit root test has asymptotically nontrivial local power. The DWB method was originally proposed for stationary time series. By showing its consistency in the PLS setting, we broaden its applicability and its use in the locally stationary context is worth further exploration. For finite sample simulations, we propose a recolored DWB (RDWB), combining the AR sieve idea used in the RWB test with DWB to improve the performance of the DWB-based test. In many cases, the RDWB method tends to provide the most accurate sizes and reasonably good power, compared to the use of DWB or RWB alone. In practice, with little knowledge of the error structure, the RDWB-based test seems preferable due to its robustness for a large class of nonstationary error processes.

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

## NOTES

1. It is worth mentioning that the generalized least squares (GLS) detrending can be used instead of the OLS detrending to make the tests more powerful. This is quite straightforward and well established in the literature (Elliott et al., 1996; Müller and Elliott, 2003; Smeekes, 2013), and will not be pursued in this article.

2. Ito's formula can be written as $dB_\sigma(r) = \sigma(r)dB(r)$. Using Itô's formula, we derive $B_\sigma^2(r) = B_\sigma^2(0) - \int_0^r 2\sigma(s)B_\sigma(s)dB(s) + 2^{-1}\int_0^r 2\sigma^2(s)ds$, which leads to $\int_0^r B_\sigma(s)\sigma(s)dB(s) = 2^{-1}\{B_\sigma^2(r) - \int_0^r \sigma^2(s)ds\}$.

## REFERENCES


Adak, S. (1998) Time-dependent spectral analysis of nonstationary time series. *Journal of the American Statistical Association* 93(444), 1488–1501.

Andrews, D.W.K. (1984) Non-strong mixing autoregressive processes. *Journal of Applied Probability* 21(4), 930–934.

Andrews, D.W.K. (1991) Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica* 59(3), 817–858.

Billingsley, P. (1968) *Convergence of Probability Measures*. Wiley.

Busetti, F. & A.M.R. Taylor (2003) Variance shifts, structural breaks, and stationarity tests. *Journal of Business and Economic Statistics* 21(4), 510–531.

Cavaliere, G. & A.M.R. Taylor (2007) Testing for unit roots in time series models with non-stationary volatility. *Journal of Econometrics* 140(2), 919–947.

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

Cavaliere, G. & A.M.R. Taylor (2008b) Time-transformed unit root tests for models with non-stationary volatility. *Journal of Time Series Analysis* 29(2), 300–330.

Cavaliere, G. & A.M.R. Taylor (2009) Bootstrap M unit root tests. *Econometric Reviews* 28(5), 393–421.

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

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

Dahlhaus, R. (1997) Fitting time series models to nonstationary processes. *The Annals of Statistics* 25(1), 1–37.

Dahlhaus, R. & S. Subba Rao (2006) Statistical inference for time-varying arch processes. *The Annals of Statistics* 34(3), 1075–1114.

Dickey, D.A. & W.A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74(366), 427–431.

Dickey, D.A. & W.A. Fuller (1981) Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica* 49(4), 1057–1072.

Domínguez, M.A. & I.N. Lobato (2001) Size Corrected Power for Bootstrap Tests. Working papers 102, Centro de Investigacion Economica, ITAM.

Draghicescu, D., S. Guillas, & W.B. Wu (2009) Quantile curve estimation and visualization for nonstationary time series. *Journal of Computational and Graphical Statistics* 18(1), 1–20.

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

Fryzlewicz, P., T. Sapatinas, & S. Subba Rao (2008) Normalized least-squares estimation in time-varying arch models. *The Annals of Statistics* 36(2), 742–786.

Fryzlewicz, P. & S. Subba Rao (2011) Mixing properties of arch and time-varying arch processes. *Bernoulli* 17(1), 320–346.


https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

Giurcanu, M. & V. Spokoiny (2004) Confidence estimation of the covariance function of stationary and locally stationary processes. *Statistics and Decisions* 22(4), 283–300.

Kim, C.-J. & C.R. Nelson (1999) Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *The Review of Economics and Statistics* 81(4), 608–616.

Kreiss, J.-P. (1988) Asymptotic statistical inference for a class of stochastic processes. Habilitationsschrift, Universität Hamburg.

Künsch, H.R. (1989) The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics* 17(3), 1217–1241.

Mallat, S., G. Papanicolaou, & Z. Zhang (1998) Adaptive covariance estimation of locally stationary processes. *The Annals of Statistics* 26(1), 1–47.

McConnell, M.M. & G. Perez-Quiros (2000) Output fluctuations in the united states: What has changed since the early 1980's. *American Economic Review* 90(5), 1464–1476.

Müller, U.K. & G. Elliott (2003) Tests for unit roots and the initial condition. *Econometrica* 71(4), 1269–1286.

Newey, W. & K.D. West (1987) A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55(3), 703–708.

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

Palm, F.C., S. Smeekes, & J.-P. Urbain (2008) Bootstrap unit-root tests: Comparison and extensions. *Journal of Time Series Analysis* 29(2), 371–401.

Paparoditis, E. & D.N. Politis (2002) Local block bootstrap. *Comptes Rendus Mathematique* 335(11), 959–962.

Paparoditis, E. & D.N. Politis (2003) Residual-based block bootstrap for unit root testing. *Econometrica* 71(3), 813–855.

Paparoditis, E. & D.N. Politis (2005) Bootstrapping unit root tests for autoregressive time series. *Journal of the American Statistical Association* 100(470), 545–553.

Parker, C., E. Paparoditis, & D.N. Politis (2006) Unit root testing via the stationary bootstrap. *Journal of Econometrics* 133(2), 601–638.

Perron, P. & S. Ng (1996) Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies* 63(3), 435–463.

Phillips, P.C.B. (1987a) Time series regression with a unit root. *Econometrica* 55(2), 277–301.

Phillips, P.C.B. (1987b) Towards a unified asymptotic theory for autoregression. *Biometrika* 74(3), 535–547.

Phillips, P.C.B. & P. Perron (1988) Testing for a unit root in time series regression. *Biometrika* 75(2), 335–346.

Phillips, P.C.B. & V. Solo (1992) Asymptotics for linear processes. *The Annals of Statistics* 20(2), 971–1001.

Phillips, P.C.B. & Z. Xiao (1998) A primer on unit root testing. *Journal of Economic Surveys* 12(5), 423–469.

Politis, D.N. & J.P. Romano (1994) The stationary bootstrap. *Journal of the American Statistical Association* 89(428), 1303–1313.

Priestley, M.B. (1965) Evolutionary spectra and non-stationary processes. *Journal of the Royal Statistical Society: Series B* 27(2), 204–237.

Psaradakis, Z. (2001) Bootstrap tests for an autoregressive unit root in the presence of weakly dependent errors. *Journal of Time Series Analysis* 22(5), 577–594.

Rho, Y. & X. Shao (2015) Inference for time series regression models with weakly dependent and heteroscedastic errors. *Journal of Business & Economic Statistics* 33(3), 444–457.

Said, S.E. & D.A. Dickey (1984) Testing for unit roots in autoregressive-moving average models of unknown order. *Biometrika* 71(3), 599–607.

Sensier, M. & D. van Dijk (2004) Testing for volatility changes in U.S. macroeconomic time series. *The Review of Economics and Statistics* 86(3), 833–839.

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press

Shao, X. (2010) The dependent wild bootstrap. *Journal of the American Statistical Association* 105(489), 218–235.
Shao, X. & W.B. Wu (2007) Asymptotic spectral theory for nonlinear time series. *The Annals of Statistics* 35(4), 1773–1801.
Smeekes, S. (2013) Detrending bootstrap unit root tests. *Econometric Reviews* 32(8), 869–891.
Smeekes, S. & A.M.R. Taylor (2012) Bootstrap union tests for unit roots in the presence of nonstationary volatility. *Econometric Theory* 28(2), 422–456.
Smeekes, S. & J.-P. Urbain (2014) A Multivariate Invariance Principle for Modified Wild Bootstrap Methods with an Application to Unit Root Testing. Technical report.
Stărică, C. & C. Granger (2005) Nonstationarities in stock returns. *Review of Economics and Statistics* 87(3), 503–522.
Stock, J. & M. Watson (1999) A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In R. Engle & H. White (eds.), *Cointegration, Causality and Forecasting: A Festschrift for Clive W.J. Granger*, pp. 1–44. Oxford University Press.
Swensen, A.R. (2003) Bootstrapping unit root tests for integrated processes. *Journal of Time Series Analysis* 24(1), 99–126.
Wu, C.F.J. (1986) Jackknife, bootstrap and other resampling methods in regression analysis (with discussion). *The Annals of Statistics* 14(4), 1261–1350.
Wu, W.B. (2005) Nonlinear system theory: Another look at dependence. *Proceedings of the National Academy of Sciences of the United States of America* 102(40), 14150–14154.
Wu, W.B. & Z. Zhou (2011) Gaussian approximations for non-stationary multiple time series. *Statistica Sininca* 21(3), 1397–1413.
Zhou, Z. (2013) Heteroscedasticity and autocorrelation robust structural change detection. *Journal of the American Statistical Association* 108(502), 726–740.
Zhou, Z. & W.B. Wu (2009) Local linear quantile estimation for nonstationary time series. *The Annals of Statistics* 37(5B), 2696–2729.

https://doi.org/10.1017/S0266466618000038 Published online by Cambridge University Press