# Bootstrap union tests for unit roots in the presence of nonstationary volatility

*Econometric Theory*, **28**, 2012, 422–456.
doi:10.1017/S0266466611000387

# BOOTSTRAP UNION TESTS FOR UNIT ROOTS IN THE PRESENCE OF NONSTATIONARY VOLATILITY

STEPHAN SMEEKES
*Maastricht University*

A.M. ROBERT TAYLOR
*University of Nottingham*

Three important issues surround testing for a unit root in practice: uncertainty as to whether or not a linear deterministic trend is present in the data; uncertainty as to whether the initial condition of the process is (asymptotically) negligible or not, and the possible presence of nonstationary volatility in the data. Assuming homoskedasticity, Harvey, Leybourne, and Taylor (2011, *Journal of Econometrics*, forthcoming) propose decision rules based on a four-way union of rejections of quasi-differenced (QD) and ordinary least squares (OLS) detrended tests, both with and without a linear trend, to deal with the first two problems. In this paper we first discuss, again under homoskedasticity, how these union tests may be validly bootstrapped using the sieve bootstrap principle combined with either the independent and identically distributed (i.i.d.) or wild bootstrap resampling schemes. This serves to highlight the complications that arise when attempting to bootstrap the union tests. We then demonstrate that in the presence of nonstationary volatility the union test statistics have limit distributions that depend on the form of the volatility process, making tests based on the standard asymptotic critical values or, indeed, the i.i.d. bootstrap principle invalid. We show that wild bootstrap union tests are, however, asymptotically valid in the presence of nonstationary volatility. The wild bootstrap union tests therefore allow for a joint treatment of all three of the aforementioned issues in practice.

## 1. INTRODUCTION

It is well known that the performance of unit root tests depends on a number of factors not observable by the practitioner applying the tests. Two such factors that have a profound impact on the power of these unit root tests, and in particular the popular (augmented) Dickey-Fuller (DF) tests, are deterministic trends and the initial condition.

In many economic applications it is important to allow for the presence of a linear trend. If a linear trend is present in the data but not accounted for in the

We thank Peter Phillips, three anonymous referees, Christoph Hanck, and seminar participants at Maastricht University, the NAKE Research Day 2010, and the NESG annual conference 2010 for helpful comments on previous versions of this paper. Address correspondence to Robert Taylor, School of Economics, University of Nottingham, Nottingham, NG7 2RD, U.K. e-mail: robert.taylor@nottingham.ac.uk.

test, power will decrease dramatically, such that even for a relatively small (local) trend the unit root tests will never reject, even asymptotically. On the other hand, if a trend is absent from the data but is accounted for in computing the test, then power also drops relative to the test that does not allow for a trend; see Marsh (2007). To deal with this issue, Harvey, Leybourne, and Taylor (hereafter HLT) (2009a) construct a new test formed as a union of rejections of unit root tests with and without a deterministic linear trend. They show that this union test can maintain high power irrespective of the true value of the trend. Moreover, by adjusting the critical values used to determine the individual rejections, the test maintains correct asymptotic size.

Similarly, the initial condition (defined as the deviation of the initial observation from the deterministic components) is also known to have a major impact on the power of unit root tests; see, for example, Phillips and Magdalinos (2009). As investigated by Müller and Elliott (2003), among others, the DF test with ordinary least squares (OLS) detrending, denoted here as *DF-OLS*, suffers from low power relative to the DF test with quasi-differenced (QD) or generalized least squares (GLS) detrending, denoted as *DF-QD*, if the initial condition is small; the opposite occurs if the initial condition is large. Moreover, as with the deterministic trend, the initial condition is not observed, thus leaving the practitioner without proper knowledge of which test to apply. HLT (2009a) again propose a test based on a union of rejections, this time from the *DF-OLS* and *DF-QD* tests, to deal with this situation and show that this test additionally maintains good size and power across different values of the initial condition.

In practice these two factors cannot realistically be viewed in isolation from each other. This motivated HLT (2011) to extend the analysis of HLT (2009a) by considering the impact of both factors simultaneously. They propose a four-way union of rejections of *DF-QD* and *DF-OLS* tests, both with and without trend. A modified version that involves (inconsistent) pretesting for both the initial condition and the linear trend is also proposed and shown to improve the power of the basic four-way union in certain cases.

In this paper we discuss how the union tests of HLT (2011) can be extended to a bootstrap setting, outlining our approach for both the independent and identically distributed (i.i.d.) and wild bootstrap resampling schemes, which can both be combined with a sieve regression. We show that two major complications arise in doing this, however. First, simply applying the bootstrap to the individual tests underlying the union does not control size. We therefore show how to combine the individual tests in a union in a valid way using the bootstrap, an idea that is not dissimilar in spirit to the use of the bootstrap in multiple testing problems (cf. White, 2000; Romano and Wolf, 2005). Bayer and Hanck (2009) also use the bootstrap for combining tests for co-integration, although they use a different approach from that which we take here. The second complication arises because of the uncertainty that exists regarding the deterministic trend. We show that incorporating an estimate of the local trend into the bootstrap data generating process (DGP) presents both some interesting problems and opportunities.

A third factor surrounds the possible presence of nonstationary volatility in the innovations. Applied researchers have recently focused attention on the question of whether or not the variability in the shocks driving macroeconomic time series has changed over time; see, e.g., the literature review in Busetti and Taylor (2003). The empirical evidence has suggested that time-varying behavior, in particular a general decline, in unconditional volatility in the shocks driving macroeconomic time series over the past 20 years or so is a relatively common phenomenon; see, inter alia, Kim and Nelson (1999), McConnell and Perez Quiros (2000), Van Dijk, Osborn, and Sensier (2002), Sensier and Van Dijk (2004) and references therein. Sensier and Van Dijk, for example, report that over 80% of the real and price variables in the Stock and Watson (1999) data set reject the null of constant unconditional innovation variance against the alternative of a one-time change in variance. Watson (1999) also argues that multiple changes in volatility are commonly observed in interest rate data.[1] Consonant with these empirical findings we will allow the volatility process in our time series model to satisfy the conditions of Cavaliere and Taylor (2008); the precise conditions are given in Assumption $1'$ in Section 3. This setup allows for a wide variety of volatility processes, requiring the innovation variance only to be nonstochastic, bounded and to display a finite number of jumps. As discussed in detail by Cavaliere and Taylor (2008), this class includes both single and multiple abrupt breaks in variance, polynomially trending volatility, piecewise trending volatility, and smooth transition variance breaks, so that it should be able to reasonably well approximate the type of volatility paths found in the aforementioned empirical studies.

Nonstationary volatility effects can greatly influence the size of standard unit root tests, even asymptotically, as has been shown by Cavaliere and Taylor (2007, 2008), among others. A solution to this problem is analyzed by Cavaliere and Taylor (2008, 2009a), who employ the wild bootstrap to capture the nonstationary volatility within the resampled data. They show that the wild bootstrap correctly reproduces the first-order limiting null distribution under nonstationary volatility, thereby allowing for the construction of asymptotically valid (pivotal) bootstrap tests. As a consequence, just as one cannot in practice view uncertainty over the deterministic trend and the initial condition in isolation from one another, it is also very difficult to argue, given empirical evidence, that uncertainty over these two factors can be analyzed separately from nonstationary volatility. Indeed, as we will show in this paper, the union tests of HLT (2011) suffer in just the same way that any "standard" asymptotic test does in the presence of nonstationary volatility. The key contribution of our paper is to show that a joint treatment that simultaneously deals with uncertainty regarding the deterministic trend, the initial condition, and the possibility of nonstationary volatility in the data is attainable using our proposed wild bootstrap variant of the union tests of HLT (2011). Computer code that allows practitioners to run the bootstrap union tests is available from http://www.personeel.unimaas.nl/s.smeekes/research.htm.

The structure of the paper is as follows. In Section 2 the bootstrap union tests are introduced under the assumption of homoskedasticity, allowing us to focus initially on the complications that arise from introducing the bootstrap in this setting. Here bootstrap tests based on the sieve combined with both i.i.d. and wild bootstrap resampling schemes are discussed. The (wild) bootstrap union tests in the presence of nonstationary volatility are subsequently investigated in Section 3. Section 4 offers some conclusions. All proofs are contained in the Appendix.

In the following, $\lfloor x \rfloor$ denotes the largest integer less than or equal to $x$; $x := y$ ($x =: y$) indicates that $x$ is defined by $y$ ($y$ is defined by $x$); convergence in distribution (probability) is denoted by $\xrightarrow{d}$ ($\xrightarrow{p}$); $W(r)$ denotes a univariate standard Brownian motion; and $\mathcal{D} = D[0,1]$ denotes the space of right continuous with left limit (càdlàg) processes. As is usual, bootstrap quantities (conditional on the original sample) are indicated by appending a superscript $*$ to the standard notation. Convergence in distribution (probability) of bootstrap statistics is denoted $\xrightarrow{d^*}$ ($\xrightarrow{p^*}$), where the bootstrap convergence holds in probability (cf. Giné and Zinn, 1990).

## 2. BOOTSTRAP UNION TESTS IN HOMOSKEDASTIC MODELS

In this section we introduce bootstrap union tests under the assumption of homoskedasticity. This allows us to focus on the complications arising from the bootstrap, while being able to work with asymptotically pivotal statistics.

### 2.1. The Model

We consider the DGP,

$$\begin{aligned} y_t &= x_t + \mu + \beta_T t, \qquad t = 0, 1, \ldots, T, \\ x_t &= \rho_T x_{t-1} + u_t, \qquad t = 1, \ldots, T, \\ u_t &= \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} =: \psi(L)\varepsilon_t, \qquad (\psi_0 = 1), \end{aligned} \tag{1}$$

where $\rho_T := 1 - c/T$. We wish to test whether or not $y_t$ contains a unit root; that is, our interest focuses on testing $H_0 : c = 0$ against $H_1 : c > 0$. The stochastic process $\psi(L)\varepsilon_t$ is assumed to satisfy the following (standard) linear process condition.

**Assumption 1.** (i) Let $\varepsilon_t$ be i.i.d. with $\mathrm{E}\varepsilon_t = 0$, $\mathrm{E}\varepsilon_t^2 = \sigma^2$, and $\mathrm{E}|\varepsilon_t|^{4+a} < \infty$ for some $a > 0$; (ii) $\psi(z) \neq 0$ for all $|z| \leq 1$, and $\sum_{j=0}^{\infty} j|\psi_j| < \infty$. Also define $\omega_u^2 := \lim_{T\to\infty} T^{-1}\mathrm{E}(\sum_{t=1}^{T} u_t)^2 = \sigma^2\psi(1)^2$.

Assumption 1(i) will be relaxed subsequently when we allow for nonstationary volatility. The next assumption specifies the behavior of the coefficient on the linear trend term in (1), providing an appropriate Pitman (local) drift for our subsequent asymptotic analyses. This assumption coincides with that employed in HLT (2009a, 2011).

**Assumption 2.** The trend coefficient $\beta_T$ in (1) satisfies $\beta_T := T^{-1/2}\omega_u\kappa$.

As argued in HLT (2009a), it is appropriate to consider a local trend model in order that the subsequent asymptotic analysis reflects the uncertainty that exists in finite samples over whether a linear trend is present in the data or not. This is because asymptotically a local trend cannot be consistently detected using pretests, while a fixed trend can; hence in the latter case there is no uncertainty asymptotically.

Finally, for the second source of uncertainty, the initial condition, we again follow HLT (2009a, 2011) and assume the following.

**Assumption 3.** The initial condition, $x_0$, is generated as $x_0 := \alpha\sqrt{\omega_u^2(1-\rho_T)^{-1}}$, for $\rho_T = 1-c/T$, $c > 0$, and for $c = 0$ we set $x_0 = 0$ without loss of generality.

In Assumption 3, $\alpha$ controls the magnitude of $x_0$ relative to the magnitude of the standard deviation of a stationary $AR(1)$ process with parameter $\rho_T$ and innovation long-run variance $\omega_u^2$. The form given for $x_0$ is also consistent with the analysis of Müller and Elliott (2003) and Elliott and Müller (2006).

## 2.2. Unit Root Tests and Their Bootstrap Analogues

As in HLT (2009a, 2011), we consider the OLS- and QD-detrended DF unit root tests. Both tests involve an initial step of detrending to obtain the detrended series $\hat{x}_{t,\gamma}^{\delta}$ as

$$\hat{x}_{t,\gamma}^{\delta} := y_t - \hat{\theta}_\gamma' z_t^{\delta}, \tag{2}$$

$$\hat{\theta}_\gamma := \left(\sum_{t=0}^{T} z_{\bar{c},\gamma,t}^{\delta} z_{\bar{c},\gamma,t}^{\delta\prime}\right)^{-1}\left(\sum_{t=0}^{T} z_{\bar{c},\gamma,t}^{\delta} y_{\bar{c},\gamma,t}\right). \tag{3}$$

Here $\gamma = QD$ and $\gamma = OLS$ for QD and OLS detrending, respectively; hence,

$$z_{\bar{c},\gamma,t}^{\delta} := \begin{cases} z_t^{\delta} - (1-\bar{c}T^{-1})z_{t-1}^{\delta} & \text{if } \gamma = QD, \\ z_t^{\delta} & \text{if } \gamma = OLS, \end{cases} \qquad t \geq 1,$$

while $z_{\bar{c},\gamma,0}^{\delta} := z_0^{\delta}$. The series $y_{\bar{c},\gamma,t}$ is defined analogously to $z_{\bar{c},\gamma,t}^{\delta}$. Furthermore, $\delta = \mu, \tau$, where $z_t^{\mu} = 1$ if just an intercept is included in the regression and $z_t^{\tau} = (1,t)'$ if both an intercept and a linear trend are included. For QD detrending we take $\bar{c} = 7$ and $\bar{c} = 13.5$ if $\delta = \mu$ and $\delta = \tau$, respectively, as recommended by Elliott, Rothenberg, and Stock (1996).

The DF $t$-statistic, denoted $DF\text{-}\gamma^{\delta}$, is then the usual regression $t$-statistic of significance on $\lambda$ in the augmented DF regression

$$\Delta\hat{x}_{t,\gamma}^{\delta} = \lambda\hat{x}_{t-1,\gamma}^{\delta} + \sum_{j=1}^{p}\phi_{p,j}\Delta\hat{x}_{t-j,\gamma}^{\delta} + \varepsilon_{p,t}, \qquad t = p+1,\ldots,T. \tag{4}$$

We make the following (standard) assumption concerning the lag length, $p$ in (4).

**Assumption 4.** Let $p \to \infty$ and $p = o(T^{1/3})$ as $T \to \infty$.

The limit distributions of the DF statistics under local alternatives are well known and documented in HLT (2009a, 2011). For completeness we reproduce these results in Lemma 1.

LEMMA 1. *Let $y_t$ be generated according to* **(1)** *and let Assumptions 1, 2, 3, and 4 hold. Then, as $T \to \infty$,*

$$DF\text{-}\gamma^{\delta} \xrightarrow{d} \frac{K_{c,\gamma}^{\delta}(1,\kappa)^2 - K_{c,\gamma}^{\delta}(0,\kappa)^2 - 1}{2\left(\int_0^1 K_{c,\gamma}^{\delta}(r,\kappa)^2 dr\right)^{1/2}},$$

*where*

$$K_{c,QD}^{\mu}(r,\kappa) := K_c(r) + r\kappa, \quad K_{c,OLS}^{\mu}(r,\kappa) := K_c(r) - \int_0^1 K_c(s)ds + \left(r - \frac{1}{2}\right)\kappa,$$

$$K_{c,QD}^{\tau}(r,\kappa) := K_c(r) - r\left(1 + \bar{c} + \frac{1}{3}\bar{c}^2\right)^{-1}\left[(1+\bar{c})K_c(1) + \bar{c}^2 \int_0^1 sK_c(s)ds\right],$$

$$K_{c,OLS}^{\tau}(r,\kappa) := K_c(r) - (4 - 6r)\int_0^1 K_c(s)ds - (12r - 6)\int_0^1 sK_c(s)ds,$$

*and, defining $W_c(r) := \int_0^r e^{-(r-s)c} dW(s)$,*

$$K_c(r) := \begin{cases} W(r) & \text{if } c = 0, \\ \alpha(e^{-rc} - 1)(2c)^{-1/2} + W_c(r) & \text{if } c > 0. \end{cases}$$

**Remark 1.** Observe that the limiting distributions of $DF\text{-}\gamma^{\tau}$, $\gamma = OLS, QD$, do not depend on the local trend coefficient, $\kappa$, as would be expected given the exact invariance of these statistics to the trend parameter. In contrast, the limiting distributions of $DF\text{-}\gamma^{\mu}$, $\gamma = OLS, QD$, depend on $\kappa$ both under the unit root null hypothesis, $c = 0$, and under the alternative, $c > 0$. The limiting distributions of all four statistics depend on $\alpha$, the initial condition magnitude, when $c > 0$.

Various bootstrap versions of the above tests have been proposed; see Palm, Smeekes, and Urbain (2008) for a selective overview. Here we focus on DF tests using the i.i.d. bootstrap (Park, 2002; Chang and Park, 2003; Smeekes, 2009) and the wild bootstrap (Cavaliere and Taylor 2008, 2009a, 2009b), in each case using a sieve regression to account for stationary serial correlation. It should be clear, however, that the same arguments hold for any other bootstrap method that delivers an asymptotically valid bootstrap test, such as the different forms of the block bootstrap (Paparoditis and Politis, 2003; Swensen, 2003; Parker, Paparoditis, and Politis, 2006). We start by describing our bootstrap algorithm. Some noteworthy aspects of the algorithm are subsequently discussed in Remarks 2 and 3.

**Algorithm 1**

1. Calculate $\hat{x}^{\tau}_{t,\tilde{\gamma}} := y_t - \hat{\theta}'_{\tilde{\gamma}} z^{\tau}_t$, where $\hat{\theta}_{\tilde{\gamma}}$ is defined as in (3) with $\tilde{\gamma} = QD, OLS$. It is not necessary that $\tilde{\gamma}$ is equal to $\gamma$, the detrending method used to obtain the test statistic.

2. Estimate an augmented DF regression of order $q$ for $\hat{x}^{\tau}_{t,\tilde{\gamma}}$ by OLS and calculate the residuals

$$\hat{\varepsilon}^{\tau}_{q,t} := \Delta\hat{x}^{\tau}_{t,\tilde{\gamma}} - \hat{\lambda}\hat{x}^{\tau}_{t-1,\tilde{\gamma}} - \sum_{j=1}^{q} \hat{\phi}_{q,j} \Delta\hat{x}^{\tau}_{t-j,\tilde{\gamma}}, \qquad t = q+1, \ldots, T. \tag{5}$$

3. Construct bootstrap errors $\varepsilon^*_t$ in one of the two following ways:

   (i) (i.i.d. bootstrap) Resample with replacement from the recentered residuals $\left(\hat{\varepsilon}^{\tau}_{q,t} - \bar{\hat{\varepsilon}}^{\tau}_{q,t}\right)$.

   (ii) (wild bootstrap) Let $\varepsilon^*_t = \xi^*_t \hat{\varepsilon}^{\tau}_{q,t}$, where $\xi^*_t$ satisfies $\mathrm{E}^* \xi^*_t = 0$ and $\mathrm{E}^* \xi^{*2}_t = 1$.[2]

4. (a) Build $u^*_t$ recursively as $u^*_t = \sum_{j=1}^{q} \hat{\phi}_{q,j} u^*_{t-j} + \varepsilon^*_t$, using the estimated parameters $\hat{\phi}_{q,j}$ from Step 2, and build $x^*_t$ as $x^*_t = x^*_{t-1} + u^*_t$, initialized at $x^*_0 = 0$. (b) Finally, let

$$y^*_t = x^*_t + \theta^{*\prime} z_t, \qquad t = 0, 1, \ldots, T,$$

where either Scheme A: $\theta^* = 0$ or Scheme B: $\theta^* = \hat{\theta}_{\tilde{\gamma}}$.

5. Using the bootstrap sample $y^*_t$, apply the same method of detrending $\gamma$ as applied to the original sample to obtain the detrended bootstrap series $\hat{x}^{\delta *}_{t,\gamma} := y^*_t - \hat{\theta}^*_{\gamma} z^{\delta}_t$, where $\hat{\theta}^*_{\gamma}$ is defined analogously as in (3), but with the bootstrap data. Calculate the bootstrap augmented DF statistic, denoted $DF\text{-}\gamma^{\delta *}_{\tilde{\gamma}}$, from the regression

$$\Delta\hat{x}^{\delta *}_{t,\gamma} = \lambda^* \hat{x}^{\delta *}_{t-1,\gamma} + \sum_{j=1}^{p^*} \phi_{p^*,j} \Delta\hat{x}^{\delta *}_{t-j,\gamma} + \varepsilon^*_{p^*,t}, \qquad t = p^* + 1, \ldots, T.$$

6. Repeat Steps 3 to 5 $N$ times, obtaining bootstrap test statistics $DF\text{-}\gamma^{\delta *}_b$ for $b = 1, \ldots, N$, and select the bootstrap critical value $cv^{\delta *}_{\gamma,\tilde{\gamma}}(\pi)$ as

$$cv^{\delta *}_{\gamma,\tilde{\gamma}}(\pi) := \max\{x : N^{-1} \sum_{b=1}^{N} I(DF\text{-}\gamma^{\delta *}_{\tilde{\gamma},b} < x) \leq \pi\}$$

or, equivalently, as the $\pi$-quantile of the ordered $DF\text{-}\gamma^{\delta *}_{\tilde{\gamma},b}$ statistics. Reject the null of a unit root if $DF\text{-}\gamma^{\delta}$ is smaller than $cv^{\delta *}_{\gamma,\tilde{\gamma}}(\pi)$, where $\pi$ is the nominal level of the test.

**Remark 2.** A crucial aspect of Algorithm 1 surrounds the choice of $\theta^*$ through either Scheme A or Scheme B. Under Scheme A the estimated deterministic component from Step 1 of the algorithm is not added back to the bootstrap sample, while under Scheme B it is. Both schemes have been used in the bootstrap unit root testing literature: for example, Cavaliere and Taylor (2009a) use Scheme A, while Parker, Paparoditis, and Politis (2006, Sect. 4.2) adopt Scheme B. If $\delta = \tau$, then the bootstrap tests are invariant to the value of $\theta^*$, but just as with the asymptotic tests, they are not if $\delta = \mu$. We will denote by $DF\text{-}\gamma_{\tilde{\gamma}}^{\delta *}(A)$ and $DF\text{-}\gamma_{\tilde{\gamma}}^{\delta *}(B)$ the bootstrap DF statistics calculated using Schemes A and B, respectively, in Step 4 of Algorithm 1. We will subsequently show that when $\delta = \mu$, the statistics based on Schemes A and B differ even asymptotically (owing to the fact that the local trend magnitude, $\kappa$, is not consistently estimated) and that, in contrast to Scheme A, Scheme B does not constitute an asymptotically valid bootstrap in the context of the local trend model considered here.

**Remark 3.** The bootstrap samples are generated by imposing the unit root null on the resampling scheme in Step 4(a) of Algorithm 1, thereby avoiding the difficulties with the use of estimated unit roots discussed by Basawa, Mallik, McCormick, Reeves, and Taylor (1991). Notice also that in Step 3 of Algorithm 1 we have used recentered residuals for the i.i.d. bootstrap but not for the wild bootstrap. This is done because the bootstrap errors should have a zero mean to avoid any unwanted drift in the resulting bootstrap data created in Step 4(a). While the wild bootstrap errors have zero mean by construction, the i.i.d. bootstrap errors have a mean equal to the sample mean of the residuals to be resampled, hence the need to recenter them.

We will require the following assumptions on the lag lengths $p^*$ and $q$ in Algorithm 1.

**Assumption 5.** (i) Let $q \to \infty$ and $q = o((T/\ln T)^{1/3})$ as $n \to \infty$. (ii) Let $p^*$ satisfy Assumption 4 and let $p^*/q \to \nu > 1$ as $T \to \infty$, where $\nu$ may be infinite.

**Remark 4.** The first part of Assumption 5 bounds $q$, while the second part essentially states that, for large $T$, $p^*$ should be at least as large as $q$. Often $p$ and $q$ will be identical (it is important to let $p^*$ differ across bootstrap replications though). However, we do not want to impose this, a priori, for two reasons. First, as can be seen from Assumptions 4 and 5(i), the assumptions on the rates are not the same. This is because the lag polynomial serves a different purpose in the augmented DF regression and the sieve regression; in the former it needs only to eliminate the serial correlation, while in the latter it not only needs to eliminate it but also to correctly replicate it and, hence, consistently estimate it at a specified rate; see, for example, Remark 7 of Palm, Smeekes, and Urbain (2010) for details.[3] Second, and more importantly, the sieve regression in (5) might be based on a different specification than the test regression; in particular, this happens in the first step of Algorithm 1 for the case where the original test is based on demeaned data. Doing so removes the possibility of misspecification arising from the omission of a nonzero trend in the sieve regression. Moreover, later on when we combine these tests in a union, this will become particularly important, as we will need a single bootstrap process; see Remark 11. In such cases it is not clear why the lag length with demeaning and detrending should be identical; hence, from the start we allow for a different lag length in the two regressions.

We now detail the limit distributions of the bootstrap tests from Algorithm 1 in the following theorem.

THEOREM 1. *Let $y_t$ be generated according to* (1) *and let Assumptions 1, 2, 3, 4, and 5 hold. Then the bootstrap augmented DF t-statistics from Algorithm 1 satisfy, as $T \to \infty$,*

(i) *If either Scheme A or Scheme B is used,*

$$DF\text{-}\gamma_{\tilde{\gamma}}^{\tau *}(A,B) \xrightarrow{d^*} \frac{K_{0,\gamma}^{\tau}(1,\kappa)^2 - K_{0,\gamma}^{\tau}(0,\kappa)^2 - 1}{2\left(\int_0^1 K_{0,\gamma}^{\tau}(r,\kappa)^2 dr\right)^{1/2}} \quad \text{in probability.}$$

(ii) *If Scheme A is used,*

$$DF\text{-}\gamma_{\tilde{\gamma}}^{\mu *}(A) \xrightarrow{d^*} \frac{K_{0,\gamma}^{\mu}(1,0)^2 - K_{0,\gamma}^{\mu}(0,0)^2 - 1}{2\left(\int_0^1 K_{0,\gamma}^{\mu}(r,0)^2 dr\right)^{1/2}} \quad \text{in probability.}$$

(iii) *If Scheme B is used,*

$$DF\text{-}\gamma_{\tilde{\gamma}}^{\mu *}(B) \xrightarrow{d^*} \frac{K_{c,\gamma,\tilde{\gamma}}^{\mu *}(1,\kappa)^2 - K_{c,\gamma,\tilde{\gamma}}^{\mu *}(0,\kappa)^2 - 1}{2\left(\int_0^1 K_{c,\gamma,\tilde{\gamma}}^{\mu *}(r,\kappa)^2 dr\right)^{1/2}} \quad \text{in probability,}$$

*where*

$$K_{c,QD,\tilde{\gamma}}^{\mu *}(r,\kappa) := W(r) + r(\kappa + B_{c,\tilde{\gamma}}),$$

$$K_{c,OLS,\tilde{\gamma}}^{\mu *}(r,\kappa) := W(r) - \int_0^1 W(s)ds + \left(r - \frac{1}{2}\right)(\kappa + B_{c,\tilde{\gamma}}),$$

*and*

$$B_{c,QD} := \left(1 + \bar{c} + \frac{1}{3}\bar{c}^2\right)^{-1}\left[(1+\bar{c})K_c(1) + \bar{c}^2 \int_0^1 sK_c(s)ds\right],$$

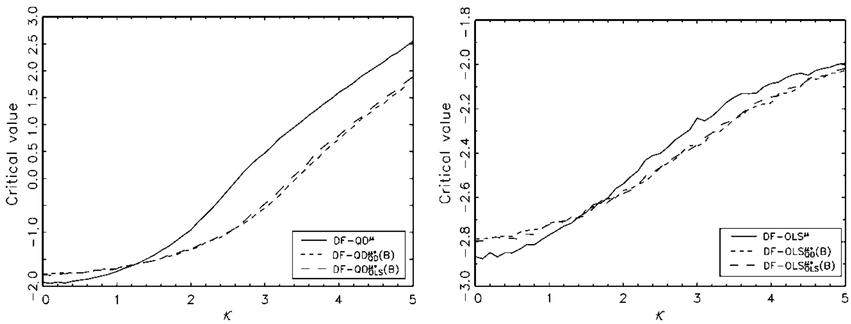
$$B_{c,OLS} := -6\int_0^1 K_c(s)ds + 12\int_0^1 sK_c(s)ds.$$

**Remark 5.** The result in part (i) of Theorem 1 establishes that the bootstrap $DF\text{-}\gamma_{\tilde{\gamma}}^{\tau *}$ statistics attain the same first-order limiting null distribution as the corresponding $DF\text{-}\gamma_{\tilde{\gamma}}^{\tau}$ statistics, regardless of whether Scheme A or B is used in step 4 of Algorithm 1. This result is, of course, expected given the invariance properties of the statistics based on detrended data; cf. Remark 1.

**Remark 6.** The results in parts (ii) and (iii) of Theorem 1 show that the choice between Schemes A and B in step 4 makes a difference, even asymptotically, when considering the statistics based on demeaned data. Under Scheme A, where $\theta^* = 0$ (i.e., where no estimated deterministic component is added to the bootstrap sample data), it is seen from part (ii) that the bootstrap $DF\text{-}\gamma_{\tilde{\gamma}}^{\mu *}$ statistics attain the same first-order limiting null ($c = 0$) distributions as the corresponding $DF\text{-}\gamma_{\tilde{\gamma}}^{\mu}$ statistics when $\kappa = 0$. Consequently, bootstrap tests based on $DF\text{-}\gamma_{\tilde{\gamma}}^{\mu *}(A)$ are asymptotically valid owing to the fact that the asymptotic tests based on $DF\text{-}\gamma_{\tilde{\gamma}}^{\mu}$ are based on the asymptotic critical value relevant for $\kappa = 0$. Both the asymptotic tests and the bootstrap analogue tests under Scheme A will therefore be conservative when $\kappa \neq 0$; see HLT (2009a). In contrast, under Scheme B, where the estimated deterministic component is added to the bootstrap data, we see from part (iii) that the bootstrap tests based on the $DF\text{-}\gamma_{\tilde{\gamma}}^{\mu}(B)$ statistics are asymptotically invalid. This is caused by the fact that $\kappa$ cannot be estimated consistently and instead converges (when scaled) to the random limit, $B_{c,\tilde{\gamma}}$ whose form depends on whether QD or OLS demeaning is used and on the value of $c$ but[4] not on $\kappa$. As a consequence the bootstrap statistics do not replicate the limiting null distribution of the demeaned DF statistics.

While the estimate of $\kappa$ is not consistent, we might still expect that it will provide some information about the true value of $\kappa$.[5] This is especially so since in the bootstrap limit distribution $\kappa$ shows up in the same way as in the original limit distribution, while as noted above, the term causing the invalidity, $B_{c,\tilde{\gamma}}$, does not depend on $\kappa$. For this reason we now investigate how large the influence of the term $B_{c,\gamma}$ is on the limit distributions in part (iii) of Theorem 1. To this end we now graph the asymptotic critical values of $DF\text{-}QD^{\mu}$ and $DF\text{-}OLS^{\mu}$ (these will, of course, coincide for $\kappa = 0$ with those for their bootstrap analogues calculated under Scheme A), together with the corresponding asymptotic bootstrap critical values for Scheme B at a 0.05 nominal level for varying $\kappa$ in Figure 1. The asymptotic critical values, as well as all other asymptotic results in the paper, were obtained by direct simulation of the relevant limiting representations, approximating the standard Brownian motion using i.i.d. $N(0, 1)$ random variables, and with the integrals approximated by normalized sums of 1,000 steps, using Gauss 8.0 with 50,000 Monte Carlo replications.

It is seen from Figure 1 that the bootstrap critical values under Scheme B clearly deviate from the asymptotic critical values, demonstrating the invalidity of the bootstrap. However, the deviation is not very large, and more importantly, the bootstrap limit distributions follow the same tendency to shift to the




(a) $DF\text{-}QD^{\mu}$ (b) $DF\text{-}OLS^{\mu}$

**FIGURE 1.** Asymptotic critical values of $DF$ tests for varying $\kappa$.

right as $\kappa$ increases, as is seen in the limiting distributions of the demeaned DF statistics. Therefore, even though the bootstrap tests are invalid, they still mimic the distribution of the demeaned DF statistics to a reasonable degree. Hence, we will not discard these invalid tests at this stage, but still consider them as a potential option in forming union tests. Moreover, because in the union multiple tests are combined, the error made by the bootstrap under Scheme B may be smoothed out.

**Remark 7.** It is straightforward to show that under a fixed trend, i.e., $\beta_T = \omega_u \kappa$, the bootstrap test $DF - \gamma_{\tilde{\gamma}}^{\mu *}(B)$ is asymptotically valid, contrary to the local trend case discussed above (a similar result is found in Parker, Paparoditis, and Politis, 2006). However, one could argue that the framework of a fixed trend is not the most appropriate to analyze trend uncertainty; a fixed trend can be picked up consistently by pretests, thus rendering union tests superfluous.

## 2.3. Bootstrap Union Tests

HLT (2011) extend the work of HLT (2009a) and propose a four-way union of rejections of $DF\text{-}QD^{\mu}$, $DF\text{-}QD^{\tau}$, $DF\text{-}OLS^{\mu}$, and $DF\text{-}OLS^{\tau}$, thereby simultaneously dealing with uncertainty about the trend and the initial condition. They also provide a scaling constant, $\tau_{\pi}$, with which to multiply the critical values of the four individual tests in order to control the asymptotic size of the union test. Let $cv_{\gamma}^{\delta}(\pi)$ denote the asymptotic critical value of $DF\text{-}\gamma^{\delta}$ at nominal level $\pi$. Then we can denote the rejection rule by

$$\text{Reject } H_0 \text{ if } \Big\{ DF\text{-}QD^{\mu} < \tau_{\pi} cv_{QD}^{\mu}(\pi) \text{ or } DF\text{-}QD^{\tau} < \tau_{\pi} cv_{QD}^{\tau}(\pi)$$

$$\text{or } DF\text{-}OLS^{\mu} < \tau_{\pi} cv_{OLS}^{\mu}(\pi) \text{ or } DF\text{-}OLS^{\tau} < \tau_{\pi} cv_{OLS}^{\tau}(\pi) \Big\}.$$

Alternatively, we may write, as in Harvey, Leybourne, and Taylor (2009b),

$$\text{Reject } H_0 \text{ if} \quad \min\left( DF\text{-}QD^{\mu}, \left(\frac{cv^{\mu}_{QD}(\pi)}{cv^{\tau}_{QD}(\pi)}\right) DF\text{-}QD^{\tau}, \left(\frac{cv^{\mu}_{QD}(\pi)}{cv^{\mu}_{OLS}(\pi)}\right) DF\text{-}OLS^{\mu}, \left(\frac{cv^{\mu}_{QD}(\pi)}{cv^{\tau}_{OLS}(\pi)}\right) DF\text{-}OLS^{\tau} \right) < \tau_{\pi}\, cv^{\mu}_{QD},$$

in which case we can denote the test statistic as

$$UR_4(\pi) = \min\left( DF\text{-}QD^{\mu}, \left(\frac{cv^{\mu}_{QD}(\pi)}{cv^{\tau}_{QD}(\pi)}\right) DF\text{-}QD^{\tau}, \left(\frac{cv^{\mu}_{QD}(\pi)}{cv^{\mu}_{OLS}(\pi)}\right) DF\text{-}OLS^{\mu}, \left(\frac{cv^{\mu}_{QD}(\pi)}{cv^{\tau}_{OLS}(\pi)}\right) DF\text{-}OLS^{\tau} \right). \tag{6}$$

This last form proves particularly useful when setting up the bootstrap; the statistic is now in the form of a minimum over four numbers, which can easily be calculated in each bootstrap iteration. It should also be immediately clear that by bootstrapping this statistic, the bootstrap critical value will automatically incorporate the scaling constant $\tau_{\pi}$ needed to achieve the correct size. This form of bootstrap test statistic closely corresponds to the maximum-based bootstrap statistics employed in White (2000) for the purpose of multiple testing.

**Remark 8.** In (6) the statistic has been scaled with respect to the distribution of $DF\text{-}QD^{\mu}$, but this is obviously an arbitrary choice. In fact, we may write the statistic as

$$UR_4(\pi) := \min\left( \left(\frac{x}{cv^{\mu}_{QD}(\pi)}\right) DF\text{-}QD^{\mu}, \left(\frac{x}{cv^{\tau}_{QD}(\pi)}\right) DF\text{-}QD^{\tau}, \left(\frac{x}{cv^{\mu}_{OLS}(\pi)}\right) DF\text{-}OLS^{\mu}, \left(\frac{x}{cv^{\tau}_{OLS}(\pi)}\right) DF\text{-}OLS^{\tau} \right),$$

for any $x < 0$. In that case the criterion for rejection would be "Reject $H_0$ if $UR_4(\pi) < \tau_{\pi} x$". It is clear then that (6) follows by setting $x = cv^{\mu}_{QD}(\pi)$.

**Remark 9.** HLT (2011) also consider a modified union test, which in some situations has higher power than $UR_4$. It consists of performing "pretests" for a large initial condition and a deterministic trend. These are not true pretests, as the initial condition and local trend cannot be consistently estimated but do provide information on the magnitude of the initial condition and trend. If the initial condition test rejects, only the union of $DF\text{-}OLS^{\mu}$ and $DF\text{-}OLS^{\tau}$ is used. If the test for a trend is rejected, only the union of $DF\text{-}QD^{\tau}$ and $DF\text{-}OLS^{\tau}$ is used. If both are rejected, only $DF\text{-}OLS^{\tau}$ is used. The exact procedure with scaling constants can be found in HLT (2011). We refer to this procedure as $UR_4^m$.

**Remark 10.** Following the recommendations of HLT (2009a), we have chosen to focus in this paper on unions of augmented DF-type unit root statistics. One could, however, perform the same analysis using other unit root statistics such as, for example, the analogous statistics from Phillips and Perron (1988) or the *M* statistics of, inter alia, Ng and Perron (2001). In these two examples we anticipate the effects of the local trend (and its estimation in the bootstrap), and the initial condition would be very similar to those reported here. One might also consider the possibility of including other unit root statistics in the union *together* with the DF-type test statistics. While possible, this is highly unlikely to lead to any gains. The more test statistics one includes in the union, the stricter the size of the individual tests need to be so as to fix the overall size of the union test at a prespecified level. Hence, the more statistics are included, the more power is lost in the resulting union test, other things being equal, unless the included test provides power in a part of the parameter space where the tests so far considered do not. As it would be difficult to find any other unit root test that provides significant power in a region where the four DF-type tests considered do not, this approach will not be considered further.

We now proceed with the bootstrap version of HLT's (2011) $UR_4(\pi)$ test. To that end, consider the following slightly modified union test statistic.

$$UR_4(\pi) := \min\Bigg( DF\text{-}QD^{\mu}, \left(\frac{cv_{QD}^{\mu*}(\pi)}{cv_{QD}^{\tau*}(\pi)}\right) DF\text{-}QD^{\tau}, \left(\frac{cv_{QD}^{\mu*}(\pi)}{cv_{OLS}^{\mu*}(\pi)}\right) DF\text{-}OLS^{\mu}, \left(\frac{cv_{QD}^{\mu*}(\pi)}{cv_{OLS}^{\tau*}(\pi)}\right) DF\text{-}OLS^{\tau} \Bigg).$$

The difference with the previous definition of $UR_4$ is that the critical values $cv_{\gamma}^{\delta}(\pi)$ have been replaced by the values $cv_{\gamma}^{\delta*}(\pi)$, which may come from a bootstrap procedure (but do not have to, as discussed in Remark 12 below). We now give the bootstrap algorithm.

**Algorithm 2**
Perform Steps 1 to 4 of Algorithm 1 to obtain a bootstrap sample $y_t^*$.

5. Calculate $DF\text{-}QD_{\tilde{\gamma}}^{\mu*}$, $DF\text{-}QD_{\tilde{\gamma}}^{\tau*}$, $DF\text{-}OLS_{\tilde{\gamma}}^{\mu*}$, and $DF\text{-}OLS_{\tilde{\gamma}}^{\tau*}$ using the bootstrap sample $y_t^*$. Next calculate $UR_{4,\tilde{\gamma}}^*(\pi)$ as

$$UR^*_{4,\tilde{\gamma}}(\pi) = \min\left( DF\text{-}QD^{\mu*}_{\tilde{\gamma}}, \left(\frac{cv^{\mu*}_{QD}(\pi)}{cv^{\tau*}_{QD}(\pi)}\right) DF\text{-}QD^{\tau*}_{\tilde{\gamma}}, \left(\frac{cv^{\mu*}_{QD}(\pi)}{cv^{\mu*}_{OLS}(\pi)}\right) DF\text{-}OLS^{\mu*}_{\tilde{\gamma}}, \left(\frac{cv^{\mu*}_{QD}(\pi)}{cv^{\tau*}_{OLS}(\pi)}\right) DF\text{-}OLS^{\tau*}_{\tilde{\gamma}} \right).$$

6. Repeat Steps 3–5 $N$ times, obtaining bootstrap test statistics $UR^*_{4,\tilde{\gamma},b}(\pi)$ for $b = 1, \ldots, N$, and select the bootstrap critical value as

$$cv^*_{UR,\tilde{\gamma}}(\pi) = \max\left\{ x : N^{-1} \sum_{b=1}^{N} I(UR^*_{4,\tilde{\gamma},b}(\pi) < x) \le \pi \right\},$$

or, equivalently, as the $\pi$-quantile of the ordered $UR^*_{4,\tilde{\gamma},b}(\pi)$ statistics. Reject the null of a unit root if $UR_4(\pi)$ is smaller than $cv^*_{UR,\tilde{\gamma}}(\pi)$.

We distinguish between $UR^*_{4,\tilde{\gamma},A}$, constructed using $DF\text{-}\gamma^{\mu*}_{\tilde{\gamma}}(A)$ with $\theta^* = 0$, and $UR^*_{4,\tilde{\gamma},B}$, constructed using $DF\text{-}\gamma^{\mu*}_{\tilde{\gamma}}(B)$ with $\theta^* = \hat{\theta}_{\tilde{\gamma}}$.

**Remark 11.** It is important to construct just one bootstrap process from which to calculate all four statistics, and not to construct four different bootstrap processes, in order to correctly replicate the distribution of the union; the original union statistic is also based on just one sample.

From now on we will ease notation by no longer indexing the $UR^*$ tests with respect to $\tilde{\gamma}$. As can be seen in Figure 1, it matters only very slightly whether $\tilde{\gamma} = QD$ or $\tilde{\gamma} = OLS$ is used. This remains the same for the union tests we consider; therefore in the following we will always take $\tilde{\gamma} = OLS$ and simply refer to the bootstrap union tests as $UR^*_{4,A}$ or $UR^*_{4,B}$.

The limit distributions of the $UR_4$ and $UR^*_4$ statistics follow directly from the continuous mapping theorem and the limit distributions of the individual (bootstrap) DF statistics (cf. White, 2000, Prop. 2.2). It is therefore clear that $UR^*_{4,A}$ is asymptotically valid, having the same first-order limit null distribution as $UR_4$ when $\kappa = 0$, while $UR^*_{4,B}$ is invalid because the underlying tests are invalid; cf. Remark 6.

**Remark 12.** There are two options for the choice of $cv^{\delta*}_{\gamma}(\pi)$: One can take the asymptotic critical value, or one can take the bootstrap critical value from bootstrapping the individual DF statistics. Asymptotically these are equivalent. However, in finite samples the bootstrap critical value may be preferable as it will usually be a better approximation of the true critical value than the asymptotic one. While it seems that using the bootstrap critical value might involve an additional bootstrap step to determine it, it can in fact be determined in the same bootstrap procedure as the calculation of $UR_4^*$, as the individual DF statistics must be calculated anyway; hence no additional bootstrap iterations are necessary. One should further note that, if bootstrap critical values are used, $cv_\gamma^{\mu *}$ should be based on the demeaned statistics from Scheme A of Step 4 of the algorithm; that is, with $\theta^* = 0$. If they are based on Scheme B, with $\theta^* = \hat{\theta}_{\tilde{y}}$, too much weight is given to the $DF\text{-}\gamma^\mu$ statistics for large $\kappa$, thus having a detrimental effect on power.

We will next analyze the asymptotic properties of the union test. We focus on the comparison of the asymptotic $UR_4$ test (and its bootstrap equivalent $UR_{4,A}^*$) with the bootstrap $UR_{4,B}^*$ test. We also add the $UR_4^m$ test of HLT (2011).

Figure 2 gives the asymptotic size for varying $\kappa$. The invalidity of the $UR_{4,B}^*$ test can be seen as the asymptotic size is above the nominal level of 0.05 for small $\kappa$. However, surprisingly, the size does not rise above 0.06, making the size distortion rather modest. For large $\kappa$ the $UR_{4,B}^*$ test is not as conservative as the $UR_4$ and $UR_4^m$ tests, and its size appears to converge to the nominal level, which is as expected given Remark 7. This might lead to the test having higher power for larger $\kappa$.

For the power analysis we add the $DF\text{-}OLS^\tau$ test, as this is the only individual test of the four tests considered in the union that does not have trivial power in any part of the parameter space. Figure 3 shows the asymptotic (uncorrected) power curves of the tests. For small $\kappa$ the power of $UR_{4,B}^*$ is very close to that

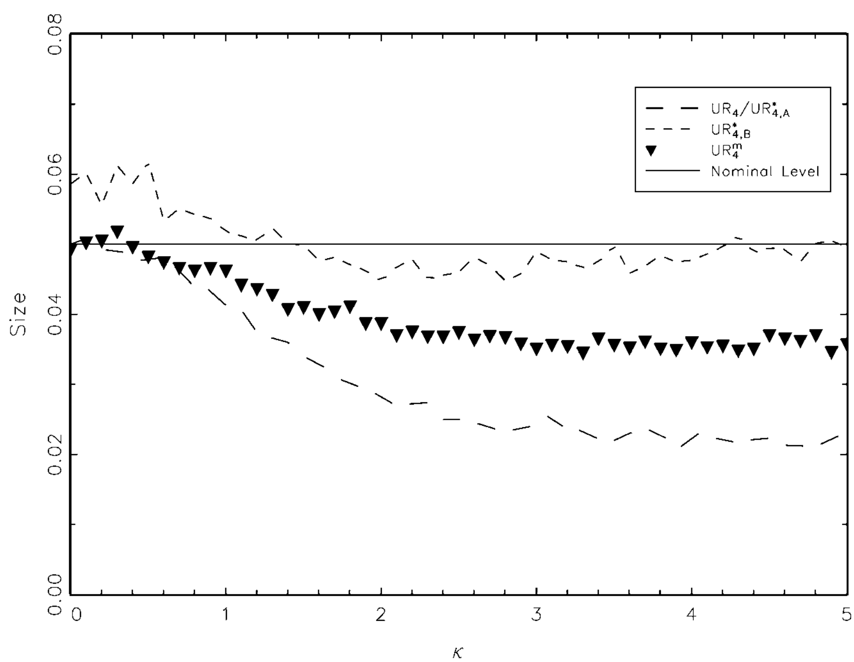



**Figure 2.** Asymptotic size of *UR* tests.

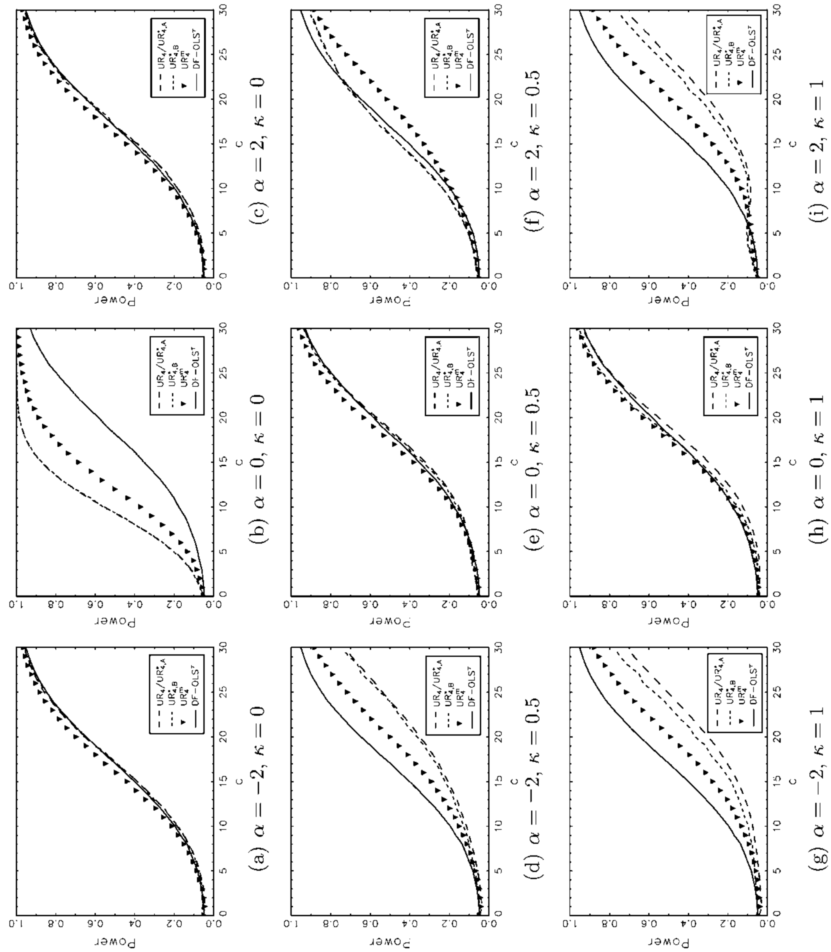



**FIGURE 3.** Asymptotic power of *UR* tests.

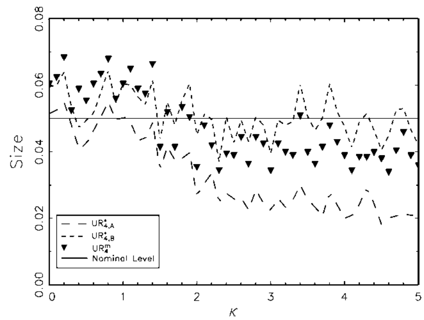
of $UR_4$ and, hence, also to $UR^*_{4,A}$. From $\kappa = 1$ on the power difference between $UR^*_{4,B}$ and $UR_4$ is noticeable, however the power of $UR^m_4$ is still higher for $\kappa = 1$. Unreported curves (which can be found in Smeekes and Taylor, 2010) show that $\kappa$ has to increase to 2 such that the power of $UR^*_{4,B}$ is higher than that of $UR^m_4$ (depending on the initial condition). For even larger $\kappa$, the power advantage of the $UR^*_{4,B}$ becomes greater, by virtue of the convergence of the size toward the nominal level. However, such a large $\kappa$ can effectively be considered as a fixed trend, and therefore this is arguably not the most relevant range for applying the union test.

The impact of the magnitude of the initial condition on the power of the tests is somewhat more varied. In general the $UR_4$ and $UR^*_{4,B}$ tests tend to be relatively more powerful in comparison with the $UR^m_4$ test for a small initial condition (in absolute sense), while the opposite occurs for a large initial condition. Notice that the effect of the initial condition is not symmetric around zero, as can clearly be seen from Figure 3(d) and 3(f). Moreover, there seems to be an interaction with $\kappa$; intermediate values of $\kappa$ show different patterns across the values of the initial condition than small and large values of $\kappa$. As expected, the $DF\text{-}OLS^\tau$ test dominates the union tests in terms of power for large trend and initial condition, but can be considerably less powerful than the union tests if the trend and initial condition are small; this is most clearly seen in Figure 3(b), where $\alpha = \kappa = 0$, but is also seen in the additional power curves reported in Smeekes and Taylor (2010) for intermediate cases, such as $\kappa = 0.25$ and $|\alpha| \leq 1$.

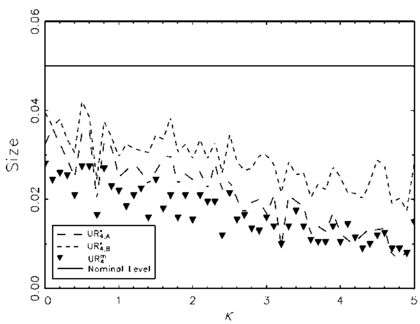
It is also interesting to observe from Figure 3 that all of the union-based tests, including the bootstrap union tests, display a lack of asymptotic unbiasedness (i.e., each displays asymptotic local power below the nominal asymptotic level of the test for some values of $c$ under $H_1$) for some values of $\kappa$ and $\alpha$. In contrast, the $DF\text{-}OLS^\tau$ test is seen to be asymptotically unbiased with local power in excess of the nominal level under $H_1$, regardless of $\kappa$ and $\alpha$.

To investigate if the finite sample performance of the individual bootstrap unit root tests discussed earlier carries over to the bootstrap union test, we perform a short Monte Carlo experiment for a small sample size. We only consider size here; we use DGP **(1)** with $c = 0$ and $u_t = \varphi u_{t-1} + \varepsilon_t + \vartheta \varepsilon_{t-1}$, with $\varepsilon_t \sim N(0,1)$. We take $T = 50$ and consider the $UR^*_{4,A}$, $UR^*_{4,B}$, and $UR^m_4$ tests, where we apply the i.i.d. scheme for the bootstrap tests.[6] Lag length selection (also within the bootstrap) is done by modified Akaike information criterion (MAIC) (Ng and Perron, 2001) with a maximum lag length of $12(T/100)^{1/4}$; as recommended by Perron and Qu (2007) we apply MAIC only to the OLS demeaned and detrended series. The sieve bootstrap regression is based on OLS detrended series (with lag length again selected by MAIC). We consider five combinations of $\varphi = \{-0.4, 0, 0.4\}$ and $\vartheta = \{-0.4, 0, 0.4\}$. We take 0.05 as nominal level of the tests. The results are based on 2,000 Monte Carlo replications and 499 bootstrap replications. Simulations are again programmed in Gauss 8.0.

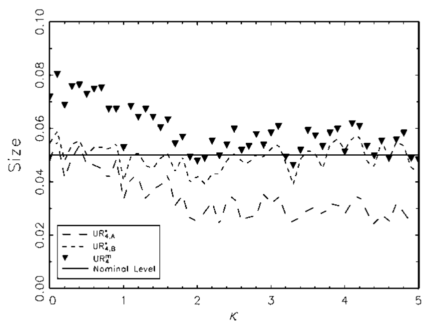
The results are given in Figure 4. The results for the model without serial correlation closely resemble the asymptotic results. If there is serial correlation,



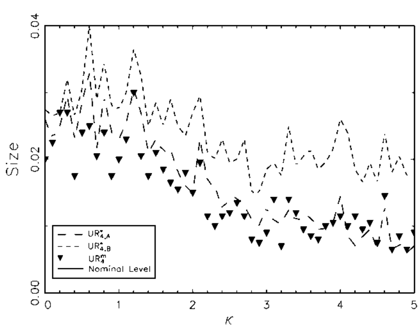

(a) $\varphi = 0$ and $\vartheta = 0$



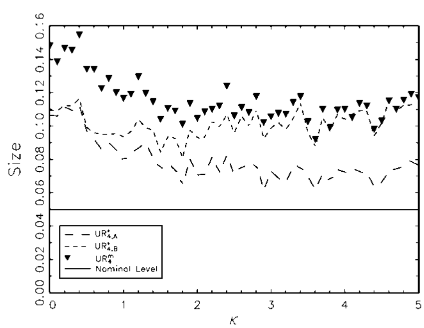

(b) $\varphi = 0.4$ and $\vartheta = 0$




(c) $\varphi = -0.4$ and $\vartheta = 0$




(d) $\varphi = 0$ and $\vartheta = 0.4$




(e) $\varphi = 0$ and $\vartheta = -0.4$

**FIGURE 4.** Empirical size of $UR$ tests for $T = 50$.

the bootstrap tests, as expected, have size closer to the nominal level than the asymptotic $UR_4^m$ test. Remarkably, the size correction not only occurs if the asymptotic test is oversized, but also if it is undersized. These results are in line with the results for the individual bootstrap unit root tests (cf. Chang and Park, 2003; Smeekes, 2009). It is also noticeable that the $UR^*_{4,A}$ test in general has somewhat better size properties than $UR^*_{4,B}$.

Concluding, the bootstrap $UR^*_{4,B}$ test can indeed mimic the effect of the local trend, although at the cost of invalidity. As expected, its power is higher than that of the conservative $UR_4$ test for larger values of $\kappa$. However the power difference only becomes noticeable for quite large values of $\kappa$, in particular in comparison to HLT's (2011) $UR_4^m$ test. One can therefore raise the question of how much, at least asymptotically, the bootstrap $UR_{4,B}^*$ test improves on the asymptotic tests, in particular as the price of invalidity has to be paid. In finite samples it may still have all the benefits over the asymptotic tests that all bootstrap tests have, but the same holds for the $UR_{4,A}^*$ test.

We will therefore now move on to a setting where the bootstrap test does have a large asymptotic advantage: the setting of nonstationary volatility.

## 3. WILD BOOTSTRAP UNION TESTS WITH NONSTATIONARY VOLATILITY

### 3.1. Unit Root Testing in Models with Nonstationary Volatility

Cavaliere and Taylor (2008, 2009b) consider testing for unit roots in settings where the volatility exhibits nonstationary behavior. They show that standard unit root tests are asymptotically not correctly sized for such volatility processes. Therefore, they propose wild bootstrap tests that are robust to nonstationary volatility; not only are these tests asymptotically valid, they are also shown to perform very well in finite samples.

The issues of uncertainty about the presence of a trend and the initial condition cannot realistically be seen in isolation from the possible presence of nonstationary volatility. However, the asymptotic tests developed in HLT (2011) can no longer be applied in the presence of nonstationary volatility, as the asymptotic unit root tests underlying the union are no longer correctly sized, even asymptotically. Moreover, it is impossible to find asymptotic scaling constants to control the asymptotic size of the union tests, as these will depend on the form of the nonstationary volatility.

The bootstrap tests discussed in the previous section, however, do not suffer from these problems and retain their validity, provided the wild bootstrap variant is used. The i.i.d. bootstrap is not valid in this setting. Hence, the bootstrap tests we consider here are robust to nonstationary volatility, trend uncertainty, and uncertainty about the initial condition. It is important to note that the scaling constants $cv_\gamma^{\delta*}(\pi)$ must now be chosen using the bootstrap critical values. The asymptotic critical values are no longer valid, and, moreover, the value of the critical values depends on the form of the volatility. It still holds that the bootstrap critical values should be based on the bootstrap samples generated with $\theta^* = 0$.

As discussed in the Introduction, we follow the framework of Cavaliere and Taylor (2008), and use the following assumption concerning the form of nonstationary volatility allowed in the innovations.

**Assumption 1′.** (i) Let $u_t = \psi(L)\varepsilon_t$, where $\varepsilon_t = \sigma_t e_t$, and let $\psi(z)$ and $e_t$ satisfy Assumption 1 (with $\mathrm{E}e_t^2 = 1$); (ii) The volatility term $\sigma_t$ satisfies $\sigma_t =$

$\omega(t/T)$ for all $r \in [0,1]$, where $\omega(\cdot) \in \mathcal{D}$ is nonstochastic and strictly positive. For $t < 0$, $\sigma_t \leq \bar{\sigma} < \infty$.

**Remark 13.** Given the results in Cavaliere and Taylor (2009b), it would be possible to extend the framework to allow for a wider class of volatility processes, including nonstationary stochastic volatility and certain generalized autoregressive conditional heteroskedasticity (GARCH) processes. Nothing would change in the setup of the tests; only the theory would become more involved. We do not consider this further here for expositional simplicity.

As in Cavaliere and Taylor (2008), we define the variance profile, $\eta(r)$, as

$$\eta(r) := \left(\int_0^1 \omega(s)^2 ds\right)^{-1} \left(\int_0^r \omega(s)^2 ds\right).$$

Furthermore, we define $\bar{\omega}^2 := \int_0^1 \omega(s)^2 ds$, which equals the limit of $T^{-1}\sum_{t=1}^T \sigma_t^2$ and may therefore be interpreted as the asymptotic average variance. Note that Assumptions 2 and 3 remain unchanged, although now $\omega_u^2 = \bar{\omega}^2 \psi(1)^2$.

We now state the limiting distributions of the $DF\text{-}\gamma^{\delta}$ statistics, the proof of which is a simple adaptation of the proof of Theorem 1 of Cavaliere and Taylor (2007).

LEMMA 2. *Let $y_t$ be generated according to* (1) *and let Assumptions 1′, 2, 3, and 4 hold. Then, as $T \to \infty$, we have that*

$$DF\text{-}\gamma^{\delta} \xrightarrow{d} \frac{K^{\delta}_{\eta,c,\gamma}(1,\kappa)^2 - K^{\delta}_{\eta,c,\gamma}(0,\kappa)^2 - 1}{2\left(\int_0^1 K^{\delta}_{\eta,c,\gamma}(r,\kappa)^2 dr\right)^{1/2}},$$

*where*

$$K^{\mu}_{\eta,c,QD}(r,\kappa) := K_{\eta,c}(r) + r\kappa,$$

$$K^{\mu}_{\eta,c,OLS}(r,\kappa) := K_{\eta,c}(r) - \int_0^1 K_{\eta,c}(s)ds + \left(r - \frac{1}{2}\right)\kappa,$$

$$K^{\tau}_{\eta,c,QD}(r,\kappa) := K_{\eta,c}(r) - r\left(1 + \bar{c} + \frac{1}{3}\bar{c}^2\right)^{-1}\left[(1+\bar{c})K_{\eta,c}(1) + \bar{c}^2\int_0^1 sK_{\eta,c}(s)ds\right],$$

$$K^{\tau}_{\eta,c,OLS}(r,\kappa) := K_{\eta,c}(r) - (4-6r)\int_0^1 K_{\eta,c}(s)ds - (12r-6)\int_0^1 sK_{\eta,c}(s)ds,$$

*and*

$$K_{\eta,c}(r) := \begin{cases} W_{\eta,0}(r) & \text{if } c = 0, \\ \alpha(e^{-rc}-1)(2c)^{-1/2} + W_{\eta,c}(r) & \text{if } c > 0, \end{cases}$$

*where* $W_{\eta,c}(r) := \int_0^r e^{-(r-s)c} dW(\eta(s))$.

We next present the limiting distributions of the bootstrap DF test statistics. As noted previously, the bootstrap tests as described in the previous section remain valid, provided the wild bootstrap is used in Step 3 of bootstrap Algorithm 1. In what follows we therefore make reference only to the wild bootstrap version of Algorithm 1.

THEOREM 2. *Let $y_t$ be generated according to* (1) *and let Assumptions 1′, 2, 3, 4, and 5 hold. Let $DF\text{-}\gamma_{\tilde{\gamma}}^{\delta *}$ denote the bootstrap augmented DF t-statistics from Algorithm 1 with $\gamma, \tilde{\gamma} = QD, OLS$, and $\delta = \mu, \tau$. Then, as $T \to \infty$, we have that*

*(i) If either Scheme A or Scheme B is used,*

$$DF\text{-}\gamma_{\tilde{\gamma}}^{\tau *}(A,B) \xrightarrow{d^*} \frac{K_{\eta,0,\gamma}^{\tau}(1,\kappa)^2 - K_{\eta,0,\gamma}^{\tau}(0,\kappa)^2 - 1}{2\left(\int_0^1 K_{\eta,0,\gamma}^{\tau}(r,\kappa)^2 dr\right)^{1/2}} \quad \textit{in probability.}$$

*(ii) If Scheme A is used,*

$$DF\text{-}\gamma_{\tilde{\gamma}}^{\mu *}(A) \xrightarrow{d^*} \frac{K_{\eta,0,\gamma}^{\mu}(1,0)^2 - K_{\eta,0,\gamma}^{\mu}(0,0)^2 - 1}{2\left(\int_0^1 K_{\eta,0,\gamma}^{\mu}(r,0)^2 dr\right)^{1/2}} \quad \textit{in probability.}$$

*(iii) If Scheme B is used*

$$DF\text{-}\gamma_{\tilde{\gamma}}^{\mu *}(B) \xrightarrow{d^*} \frac{K_{\eta,c,\gamma,\tilde{\gamma}}^{\mu *}(1,\kappa)^2 - K_{\eta,c,\gamma,\tilde{\gamma}}^{\mu *}(0,\kappa)^2 - 1}{2\left(\int_0^1 K_{\eta,c,\gamma,\tilde{\gamma}}^{\mu *}(r,\kappa)^2 dr\right)^{1/2}} \quad \textit{in probability,}$$

*where*

$$K_{\eta,c,QD,\tilde{\gamma}}^{\mu *}(r,\kappa) := W_{\eta,0}(r) + r(\kappa + B_{\eta,c,\tilde{\gamma}}),$$

$$K_{\eta,c,OLS,\tilde{\gamma}}^{\mu *}(r,\kappa) := W_{\eta,0}(r) - \int_0^1 W_{\eta,0}(s)ds + \left(r - \frac{1}{2}\right)(\kappa + B_{\eta,c,\tilde{\gamma}}),$$

*and*

$$B_{\eta,c,QD} := \left(1 + \bar{c} + \frac{1}{3}\bar{c}^2\right)^{-1} \left[(1+\bar{c})K_{\eta,c}(1) + \bar{c}^2 \int_0^1 s K_{\eta,c}(s)ds\right],$$

$$B_{\eta,c,OLS} := -6\int_0^1 K_{\eta,c}(s)ds + 12\int_0^1 s K_{\eta,c}(s)ds.$$

**Remark 14.** The implications of the results in Theorem 2 are qualitatively similar to those from the results in Theorem 1 for the constant volatility case. Principally, the detrended wild bootstrap DF statistics attain the same first-order limit null distribution as the corresponding detrended DF statistics. This result has already been established in Cavaliere and Taylor (2008). For the bootstrap demeaned DF statistics, again the choice between Schemes A and B in Step 4 of Algorithm 1 is crucial. Asymptotically valid bootstrap tests are again obtained under Scheme A, but not under Scheme B. Notice that under Scheme B, the additional random term in the limit distribution, $B_{\eta,c,\tilde{\gamma}}$, now also depends on the form of the nonstationary volatility.

**Remark 15.** As demonstrated in Cavaliere and Taylor (2008), usually one does not need Assumption 5 when applying the wild bootstrap, and it suffices to assume that $q \leq p^*$, where neither is required to increase with the sample size. This is also true in our setting for deriving the limit distributions of $DF\text{-}\gamma^{\tau*}$ and $DF\text{-}\gamma^{\mu*}(A)$, but it is not true for the limit distributions of $DF\text{-}\gamma^{\mu*}(B)$. For this test the nuisance parameters arising from the estimation of $\theta^*$ (i.e., $\omega_u$) imply that these must also be correctly reproduced within the bootstrap for them to cancel out in the limiting distribution.

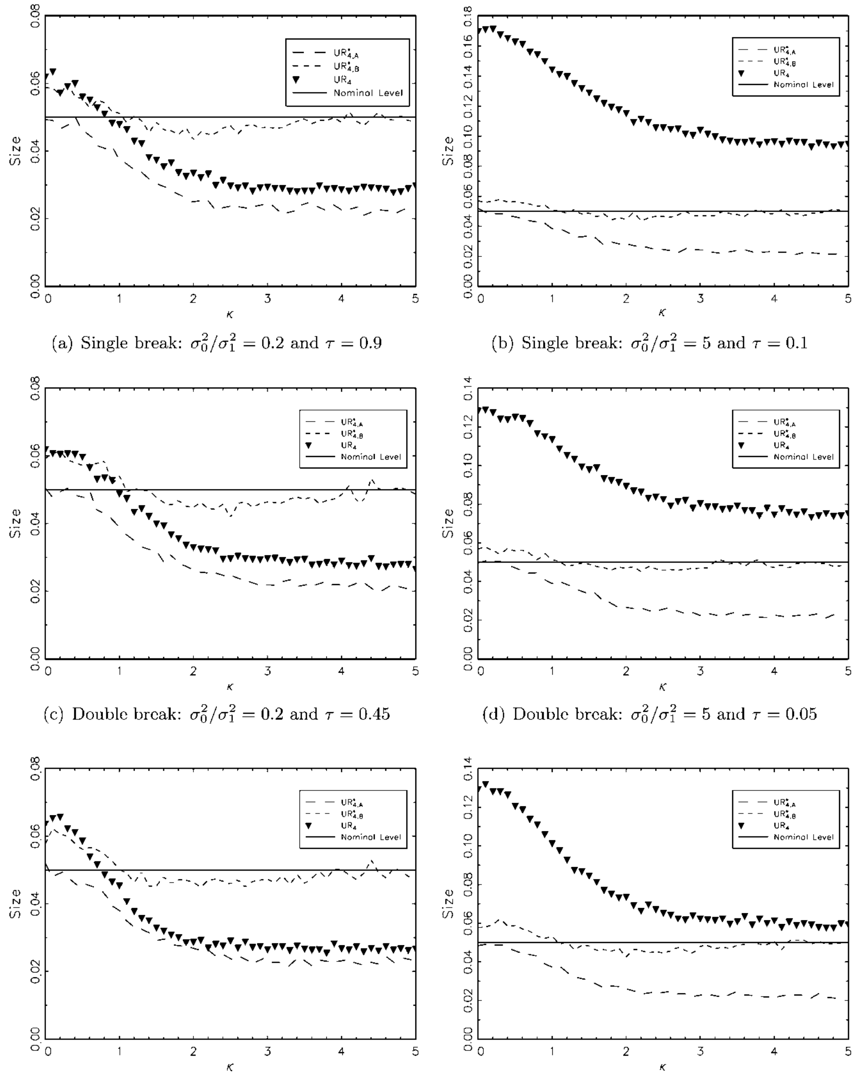
## 3.2. Wild Bootstrap Union Tests

The asymptotic and bootstrap distributions of the $UR_4^*$ tests follow directly from the continuous mapping theorem. Therefore, we expect the bootstrap union tests to be able to reproduce the impact of the volatility on the asymptotic distribution, unlike the asymptotic union tests. To investigate this, we simulate the asymptotic distributions of $UR_4$, $UR_{4,A}^*$, and $UR_{4,B}^*$ for a number of different volatility models. In particular, we consider the following settings that correspond to the simulation models used by Cavaliere and Taylor (2008).

1. Single break in volatility: $\sigma_t^2 = \sigma_0^2 + (\sigma_1^2 - \sigma_0^2) I(t > \lfloor \tau T \rfloor)$.
2. Double break in volatility: $\sigma_t^2 = \sigma_0^2 + (\sigma_1^2 - \sigma_0^2) I(\lfloor \tau T \rfloor < t < \lfloor (1-\tau) T \rfloor)$.
3. Trending volatility: $\sigma_t = \sigma_0 + (\sigma_1 - \sigma_0) t$.

Figure 5 gives the results for size of these models. The asymptotic union test $UR_4$ is, as expected, not correctly sized, and in some situations it is quite severely oversized. The bootstrap tests behave in exactly the same way as in the i.i.d. case; $UR_{4,A}^*$ is conservative, its size decreases as $\kappa$ increases, while $UR_{4,B}^*$ is slightly oversized for small $\kappa$ but correctly sized when $\kappa$ increases. The fact that the behavior of the bootstrap tests is the same over all combinations considered here is very encouraging, as it indicates that the bootstrap tests are highly robust to nonstationary volatility, and we may therefore generalize the conclusions drawn from the i.i.d. case.

We now turn to a consideration of the power properties of these tests. We also add a wild bootstrap implementation of the $DF\text{-}OLS^{\tau}$ test, denoted $DF\text{-}OLS^{\tau*}$, to the graphs, as again this test is the only one of the four individual tests considered that does not have trivial power for any parameter combination. Moreover, since it is based on the wild bootstrap, it is also feasible and asymptotically valid. To be of practical value, therefore, the bootstrap union tests should provide power




(a) Single break: $\sigma_0^2/\sigma_1^2 = 0.2$ and $\tau = 0.9$

(b) Single break: $\sigma_0^2/\sigma_1^2 = 5$ and $\tau = 0.1$

(c) Double break: $\sigma_0^2/\sigma_1^2 = 0.2$ and $\tau = 0.45$

(d) Double break: $\sigma_0^2/\sigma_1^2 = 5$ and $\tau = 0.05$

(e) Trending: $\sigma_0/\sigma_1 = 0.2$

(f) Trending: $\sigma_0/\sigma_1 = 5$

**FIGURE 5.** Asymptotic size of $UR$ tests.

advantages over this test for at least a reasonable part of the parameter space. As the size of the asymptotic test $UR_4$ is often far from the nominal level, we size-correct the power at $\kappa = 0$. Hence, the power is corrected for the nonstationary volatility, but not for the trend.

Power graphs are presented in Figures 6 and 7. In order to conserve space we report here a representative selection of a wider set of results that can be found

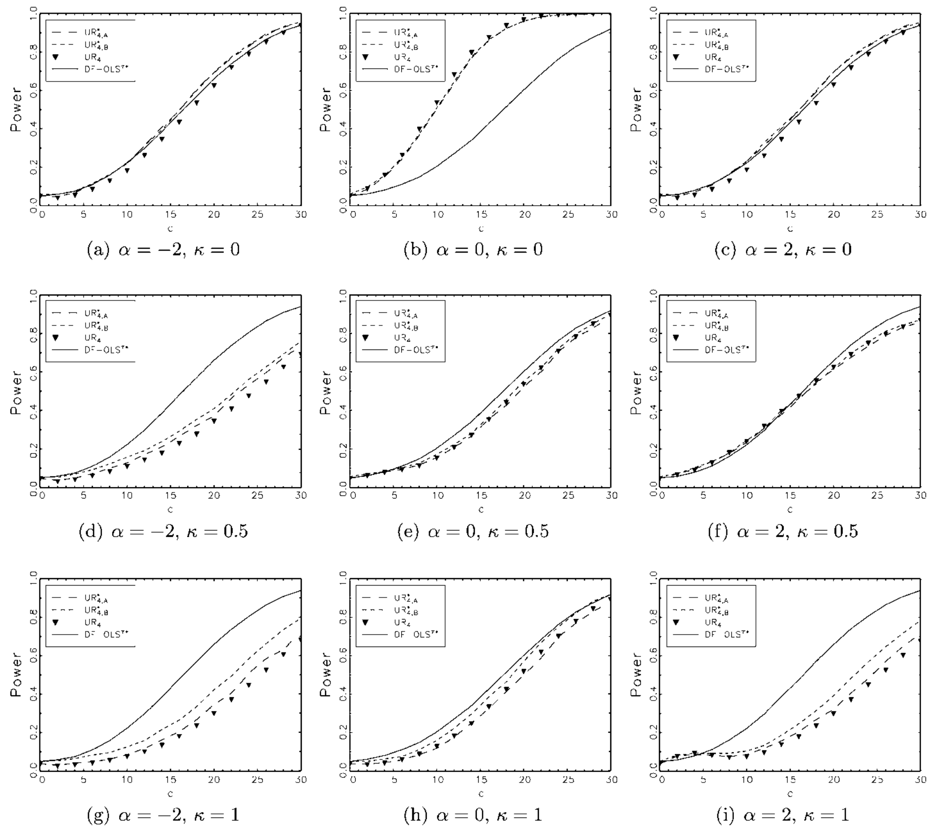
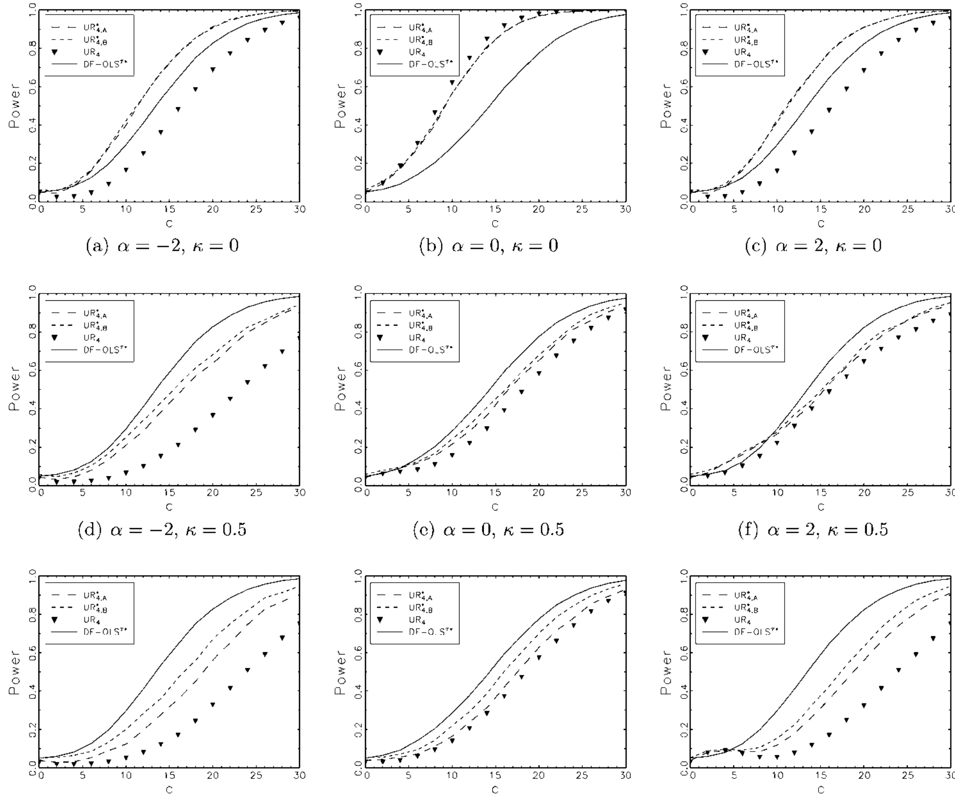



**FIGURE 6.** Asymptotic power of $UR$ tests; single break: $\sigma_0^2/\sigma_1^2 = 0.2$ and $\tau = 0.9$.




(a) $\alpha = -2,\ \kappa = 0$ (b) $\alpha = 0,\ \kappa = 0$ (c) $\alpha = 2,\ \kappa = 0$

(d) $\alpha = -2,\ \kappa = 0.5$ (e) $\alpha = 0,\ \kappa = 0.5$ (f) $\alpha = 2,\ \kappa = 0.5$

(g) $\alpha = -2,\ \kappa = 1$ (h) $\alpha = 0,\ \kappa = 1$ (i) $\alpha = 2,\ \kappa = 1$

**FIGURE 7.** Asymptotic power of $UR$ tests; trending: $\sigma_0/\sigma_1 = 0.2$.

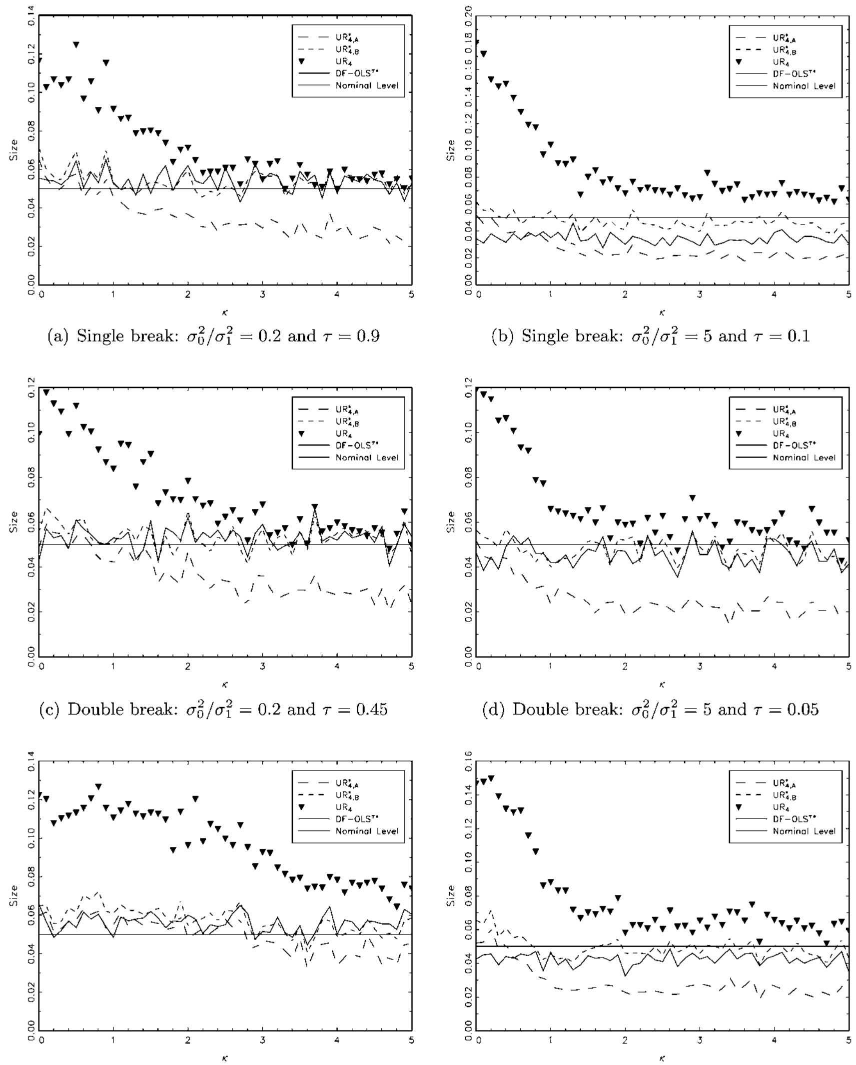
in Smeekes and Taylor (2010). The union tests are again seen to be substantially more powerful than the $DF\text{-}OLS^{\tau *}$ test for small $\kappa$ and small $|\alpha|$, while the power loss with respect to the $DF\text{-}OLS^{\tau *}$ test for the other combinations is again relatively minor in most cases. As in the homoskedastic case discussed previously, the power curves reported in Smeekes and Taylor (2010) show that for intermediate cases the union tests are also more powerful than the $DF\text{-}OLS^{\tau *}$ test. Thus, the union tests are able to deal with the trend and initial condition uncertainty, just as in the homoskedastic case analyzed by HLT (2009a, 2011). Moreover, the bootstrap tests are also highly robust to the presence of nonstationary volatility in the innovations.

It is worth noting that the power properties of the tests can differ considerably in specific models. This can, for example, be seen for the trending volatility model, where the union test also offers (unexpected) power gains on the $DF\text{-}OLS^{\tau *}$ test for large $|\alpha|$. This can be explained by the fact that the nonstationary volatility also has a direct effect on the size-corrected local power function, a point also noted by Cavaliere and Taylor (2008, p. 8). Therefore, it should not be expected that the power curves, and hence also the relations between the tests, are exactly the same as in the i.i.d. case. This can also be seen by the difference in the power curves of the $UR_4$ and $UR_{4,A}^*$ tests; based on the homoskedastic case, one might expect the curves to be equal after the size correction. As this is clearly not the case, it illustrates the direct effect of the nonstationary volatility on the power curves. Notice also from Figures 6 and 7 that the union tests, including the size-corrected $UR_4$ test, again display a lack of asymptotic unbiasedness as in the homoskedastic case. Although there is no evidence of a lack of asymptotic unbiasedness in $DF\text{-}OLS^{\tau *}$ in Figures 6 and 7, $DF\text{-}OLS^{\tau *}$ is also not asymptotically unbiased, as can be seen in, for example, in Figure 18 of Smeekes and Taylor (2010).

Comparing the two bootstrap union tests, we see that as in the homoskedastic case, the $UR_{4,B}^*$ test is somewhat more powerful than the $UR_{4,A}^*$ for $\kappa > 0$, but the power difference only becomes substantial for large $\kappa$. It is therefore doubtful if this relatively minor power gain makes it worthwhile using a test that is invalid.[7]

We next consider the finite sample properties of the bootstrap union tests. We perform a small Monte Carlo experiment again for $T = 50$.[8] We use DGP (1) with $u_t = \sigma_t e_t$, where $e_t \sim N(0, 1)$. For the size results we consider the same specifications for $\sigma_t$ as we did for the asymptotic analysis. Power is investigated only for the single-break model considered above. The tests applied are $UR_{4,A}^*$, $UR_{4,B}^*$, $UR_4$, and $DF\text{-}OLS^{\tau *}$. We apply the wild bootstrap and set all lag lengths to zero. Results are again based on 2,000 Monte Carlo replications, 499 bootstrap replications, and the nominal level set to 0.05.

The results for size are given in Figure 8. While the finite sample size distortions of the $UR_4$ test are significant as the size of $UR_4$ at $\kappa = 0$ is above 0.1 for all models considered here, the size distortion of the bootstrap tests is minimal. The $UR_{4,A}^*$ test displays the same conservative behavior as is found in the asymptotic analysis. It is quite noticeable that the size distortions of the asymptotically invalid $UR_{4,B}^*$ test are also very small; size is rarely more than 0.07. Moreover,




(a) Single break: $\sigma_0^2/\sigma_1^2 = 0.2$ and $\tau = 0.9$

(b) Single break: $\sigma_0^2/\sigma_1^2 = 5$ and $\tau = 0.1$

(c) Double break: $\sigma_0^2/\sigma_1^2 = 0.2$ and $\tau = 0.45$

(d) Double break: $\sigma_0^2/\sigma_1^2 = 5$ and $\tau = 0.05$

(e) Trending: $\sigma_0/\sigma_1 = 0.2$

(f) Trending: $\sigma_0/\sigma_1 = 5$

**FIGURE 8.** Empirical size of $UR$ tests for $T = 50$.

the test does not display the undersize found for the $UR^*_{4,A}$ test. As expected, the $DF\text{-}OLS^{\tau *}$ test has good size properties.

Power curves for the single break model are given in Figure 9. Again, the power curves of $UR_4$ are size-corrected at $\kappa = 0$. The power curves show the same patterns as the asymptotic power curves. The bootstrap union tests are more powerful than the $DF\text{-}OLS^{\tau *}$ test for small $\kappa$ and $|\alpha|$, and somewhat less powerful for

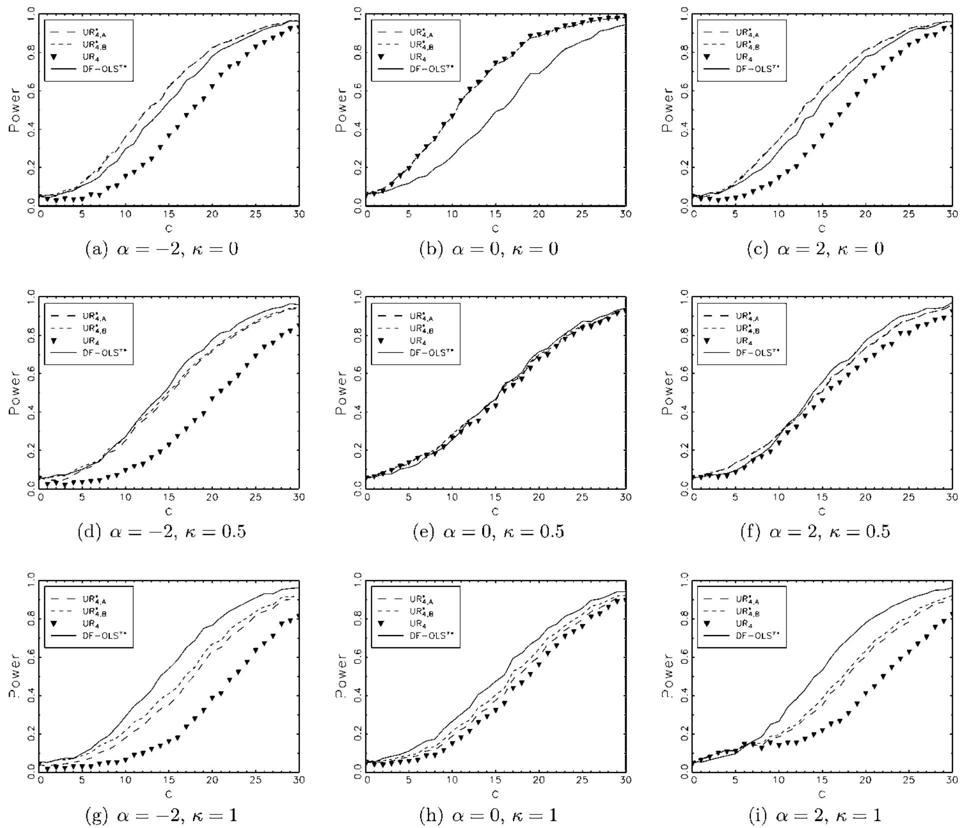



**FIGURE 9.** Empirical power of $UR$ tests for $T = 50$; single break: $\sigma_0^2/\sigma_1^2 0.2$ and $\tau = 0.9$.

large $\kappa$ and $|\alpha|$. The power loss at $\kappa = 1$ of the bootstrap union tests with respect to the $DF\text{-}OLS^{\tau *}$ test is, however, less than that seen in the asymptotic results. The difference between the analysis of $UR_4$ and $UR^*_{4,A}$ seen in the asymptotic power curves is also present.

Given the results presented here, the bootstrap union tests proposed in this paper would appear to constitute a valuable option if one needs to deal simultaneously with uncertainty regarding the trend and the initial condition and to provide results that are simultaneously robust to the possible presence of nonstationary volatility. Extant tests in the literature cannot perform satisfactorily in this situation, with the possible exception of the wild bootstrap $DF\text{-}OLS^{\tau *}$ test. However, as shown in this section, the bootstrap union tests have a clear power advantage over this test for those combinations of $\kappa$ and $\alpha$ that result in uncertainty about their values, while for the other combinations of these parameters the power loss is modest.

## 4. CONCLUSIONS

In this paper we have developed bootstrap tests, based on both the i.i.d. and wild bootstrap approaches, combined with the sieve principle to account for stationary serial correlation, designed to be robust over uncertainty about the presence of a deterministic trend and uncertainty about the initial condition, thereby extending the union tests of HLT (2009a, 2011) to a bootstrap setting. Moreover, provided the wild bootstrap variant is employed, our proposed bootstrap tests are also shown to be robust to nonstationary volatility in the innovations.

We considered two bootstrap union tests, $UR^*_{4,A}$ and $UR^*_{4,B}$; the first is a valid (conservative) test, the second is an invalid test, although its size does not appear to deviate to any great degree from the nominal level, and it is somewhat more powerful. In the setting of homoskedasticity the first test is asymptotically equivalent to the asymptotic $UR_4$ test of HLT (2011), while the second closely resembles it. Simulation evidence demonstrates that the proposed bootstrap tests can deliver an improvement in finite sample performance over the asymptotic union tests.

In cases where the volatility of the innovations is nonstationary, the asymptotic union tests of HLT (2011) fail, just as regular asymptotic tests do. Here the wild bootstrap is used as in Cavaliere and Taylor (2008, 2009b) to make our bootstrap union tests asymptotically robust to nonstationary volatility. The power properties of the bootstrap union test in relation to the trend and initial condition remain similar to those that pertain in the homoskedastic case. Hence, in this setting the bootstrap union tests clearly provide clear advantages over the existing tests both asymptotically and in finite samples, as was demonstrated through simulation evidence. Computer programs, written in Gauss, which enable practitioners to run the bootstrap union tests developed in this paper on real data, are available from http://www.personeel.unimaas.nl/s.smeekes/research.htm.

We conclude by briefly discussing an important limitation of our approach. We have constrained the deterministic trend function to be either a linear trend or a constant. In reality, trend behavior will likely be more complex than this. Cavaliere, Harvey, Leybourne, and Taylor (2011) develop unit root tests that allow for either a linear trend or a single break (of fixed magnitude) in trend and that are robust to nonstationary volatility of the form considered in this paper, but that require an asymptotically negligible assumption on the initial value. However, their approach is based on the use of pretesting (for the presence of a break in trend) rather than the union of rejections approach outlined in this paper. It would be interesting to extend the approach taken in this paper to allow for a more general class of potential trend functions, including the broken trend case. However, this would constitute a far from trivial extension of the results in this paper and is left as a suggestion for further research.

## NOTES

1. Note that these studies predate the recent credit crisis. Including the period after 2008 one would see large increases in the unconditional volatility. If anything, the recent financial turmoil and the apparent corresponding rise in unconditional volatility reinforce the need to allow for the possibility of nonconstancy in unconditional volatility.
2. In this paper we take $\xi_t^*$ to be standard normal. Other choices are also possible, although Cavaliere and Taylor (2008, Rem. 6) mention that this has almost no impact on finite sample behavior.
3. Compare also with the rates required for the DF $t$-test and coefficient test in Chang and Park (2002); stronger assumptions on the lag length are required for the coefficient test for the same reasons as above.
4. The dependence on $c$ will clearly have some impact on power (cf. Paparoditis and Politis, 2003, 2005). However, unreported simulations show that the influence of $c$ on the limiting distributions is very small, and therefore power is hardly compromised.
5. A similar argument is given for the use of inconsistent pretests in HLT (2011).
6. The wild bootstrap gives virtually identical results to those reported here. Results for $T = 100$ are also reported in Smeekes and Taylor (2010); these are very similar to those for $T = 50$.
7. Even though the distribution of $UR_{4,A}^*$ is only identical to the asymptotic distribution of the union test for $\kappa = 0$, it is valid in the sense that it is a size $\pi$ test, whereas $UR_{4,B}^*$ clearly is not; cf. Remarks 6 and 14.
8. Corresponding results for $T = 100$, also reported in Smeekes and Taylor (2010), are very similar to those reported here.

## REFERENCES

Basawa, I.V., A.K. Mallik, W.P. McCormick, J.H. Reeves, & R.L. Taylor (1991) Bootstrapping unstable first-order autoregressive processes. *Annals of Statistics* 19, 1098–1101.

Bayer, C. & C. Hanck (2009) Combining Non-cointegration Tests. METEOR Research Memorandum 09/012, Maastricht University.

Busetti, F. & A.M.R. Taylor (2003) Testing against stochastic trend in the presence of variance shifts. *Journal of Business and Economic Statistics* 21, 510–531.

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

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

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

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

Cavaliere, G. & A.M.R. Taylor (2009b) Heteroskedastic time series with a unit root. *Econometric Theory* 25, 1228–1276.

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, 379–400.

Elliott, G. & U.K. Müller (2006) Minimizing the impact of the initial condition on testing for unit roots. *Journal of Econometrics* 135, 285–310.

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

Giné, E. & J. Zinn (1990) Bootstrapping general empirical measures. *Annals of Probability* 18, 851–869.

Hannan, E.J. & L. Kavalieris (1986) Regression, autoregression models. *Journal of Time Series Analysis* 7, 27–49.

Harvey, D.I., S.J. Leybourne, & A.M.R. Taylor (2009a) Unit root testing in practice: Dealing with uncertainty over the trend and initial condition. *Econometric Theory* 25, 587–636.

Harvey, D.I., S.J. Leybourne, & A.M.R. Taylor (2009b) Unit root testing in practice: Dealing with uncertainty over the trend and initial condition — Rejoinder. *Econometric Theory* 25, 658–667.

Harvey, D.I., S.J. Leybourne, & A.M.R. Taylor (2011) Testing for unit roots in the presence of uncertainty over both the trend and initial condition. *Journal of Econometrics*, forthcoming.

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

Marsh, P. (2007) The available information for invariant tests of a unit root. *Econometric Theory* 23, 686–710.

McConnell, M.M. & G. Perez Quiros (2000) Output fluctuations in the United States: What has changed since the early 1980s? *American Economic Review* 90, 1464–1476.

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

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

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

Palm, F.C., S. Smeekes, & J.-P. Urbain (2010) A sieve bootstrap test for cointegration in a conditional error correction model. *Econometric Theory* 26, 647–681.

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

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

Park, J.Y. (2002) An invariance principle for sieve bootstrap in time series. *Econometric Theory* 18, 469–490.

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

Perron, P. & Z. Qu (2007) A simple modification to improve the finite sample properties of Ng and Perron's unit root tests. *Economics Letters* 94, 12–19.

Phillips, P.C.B. & T. Magdalinos (2009) Unit root and cointegrating limit theory when initialization is in the infinite past. *Econometric Theory* 25, 1682–1715.

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

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

Romano, J.P. & M. Wolf (2005) Stepwise multiple testing as formalized data snooping. *Econometrica* 73, 1237–1282.
Sensier, M. & D. Van Dijk (2004) Testing for volatility changes in U.S. macroeconomic time series. *Review of Economics and Statistics* 86, 833–839.
Smeekes, S. (2009) Detrending Bootstrap Unit Root Tests. METEOR Research Memorandum 09/056, Maastricht University.
Smeekes, S. & A.M.R. Taylor (2010) Bootstrap Union Tests for Unit Roots in the Presence of Nonstationary Volatility. Granger Centre Discussion Paper 10/03, University of Nottingham.
Stock, J.H. (1994) Unit roots, structural breaks and trends. In R.F. Engle & D.L. McFadden (eds.), *Handbook of Econometrics*, vol. 4, Ch. 46, pp. 2739–2841. North Holland.
Stock, J.H. & M.W. Watson (1999) A comparison of linear and nonlinear univariate models for forecasting macroeconomic time series. In R.F. Engle & H.White (eds.), *Cointegration, Causality and Forecasting: A Festschrift in Honour of 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, 99–126.
Van Dijk, D., D.R. Osborn, & M. Sensier (2002) Changes in Variability of the Business Cycle in the G7 Countries. Econometric Institute Report EI 2002-28, Erasmus University Rotterdam.
Watson, M.W. (1999) Explaining the increased variability in long-term interest rates. *Federal Reserve Bank of Richmond Economic Quarterly* 85, 71–96.
White, H. (2000) A reality check for data snooping. *Econometrica* 68, 1097–1126.

# APPENDIX

**Proof of Theorem 1.** We focus here on the proof for the sieve combined with i.i.d. bootstrap test. The proof for the wild bootstrap follows similarly.

It follows from results in, among others, Park (2002), that under Assumptions 1 and 5, $T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor} u_t^* \xrightarrow{d^*} \omega_u W(r)$ in probability. The results in (i) and (ii) then follow as in Smeekes (2009). We now focus on (iii). Let $\theta^* := (\mu^*, \beta^*)'$ and note that for OLS detrending $\hat{x}_{t,OLS}^{\mu*} = y_t^* - T^{-1}\sum_{t=1}^{T} y_t^* = x_t^* - T^{-1}\sum_{t=1}^{T} x_t^* + \beta^* t - \frac{1}{2}(T+1)\beta^*$. Then

$$\begin{aligned}
T^{-1/2}\hat{x}_{\lfloor Tr\rfloor,OLS}^{\mu*} &= T^{-1/2}x_{\lfloor Tr\rfloor}^* - T^{-3/2}\sum_{t=1}^{T} x_t^* \\
&\quad + T^{1/2}\left(\frac{\lfloor Tr\rfloor}{T} - \frac{1}{2}\right)\left[\beta_T + (\beta^* - \beta_T)\right] + o_p^*(1) \\
&\xrightarrow{d^*} \omega_u\left[W(r) - \int_0^1 W(r)dr + \left(r - \frac{1}{2}\right)(\kappa + B_{c,\tilde{\gamma}})\right] \text{ in probability,}
\end{aligned}$$

as $T^{1/2}(\beta^* - \beta_T) = T^{1/2}(\hat{\beta}_{\tilde{\gamma}} - \beta_T) \xrightarrow{d} B_{c,\tilde{\gamma}}$, which can easily be derived from standard results (cf. Stock, 1994; Elliott et al., 1996). For QD detrending we can derive in a similar way that $T^{-1/2}\hat{x}_{\lfloor Tr\rfloor,QD}^{\mu*} = T^{-1/2}x_{\lfloor Tr\rfloor}^* + T^{1/2}\frac{\lfloor Tr\rfloor}{T}\beta + T^{1/2}\frac{\lfloor Tr\rfloor}{T}(\beta^* - \beta) + o_p^*(1) \xrightarrow{d^*}$ $\omega_u\left[W(r) + r(\kappa + B_{c,\tilde{\gamma}})\right]$ in probability. Result (iii) then follows in the same way as (i) and (ii). ■

**Proof of Lemma 2.** For $c = 0$ it follows from Cavaliere and Taylor (2007, Thm. 1) that $T^{-1/2}x_{\lfloor Tr\rfloor} \xrightarrow{d} \bar{\omega}\psi(1)W_{\eta,0}(r)$, while for $c > 0$ we can write $T^{-1/2}x_{\lfloor Tr\rfloor} =$ $T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}\Delta x_t + T^{-1/2}x_0 = T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}\Delta x_t + T^{-1/2}\alpha\sqrt{\omega_u^2(1-\rho_T)^{-1}} \xrightarrow{d} \bar{\omega}\psi(1)$

$\left[W_{\eta,c}(r)+\alpha(e^{-rc}-1)(2c)^{-1/2}\right]$. Furthermore, by Lemma 2 in Cavaliere and Taylor (2008) we have that the residual variance estimator $\hat{\sigma}^2 \xrightarrow{p} \bar{\omega}^2$. Cavaliere and Taylor (2007, Thm. 1) and Cavaliere and Taylor (2008, Lem. 2) show that $\left\|\hat{\Phi}_p - \Phi_p\right\| = o_p(p^{-1/2})$, where $\hat{\Phi}_p := (\hat{\phi}_{p,1},\ldots,\hat{\phi}_{p,p})'$ and $\Phi_p := (\phi_1,\ldots,\phi_p)'$, by which the lag augmentation can be handled as in Chang and Park (2002, Lems. 3.1 and 3.2) and Smeekes (2009, Lem. 1). ■

**Proof of Theorem 2.** The invariance principle for $\varepsilon_t^*$, $T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}\varepsilon_t^* \xrightarrow{d^*} \bar{\omega}W_{\eta,0}(r)$, in probability, follows directly from the proof of Theorem 2, equation (A.4), of Cavaliere and Taylor (2008). We next show that

$$T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor} u_t^* \xrightarrow{d^*} \bar{\omega}\psi(1)W_{\eta,0}(r) \quad \text{in probability.} \tag{A.1}$$

Letting $\hat{\Phi}(L) := 1-\sum_{j=1}^{q}\hat{\phi}_{q,j}L^j$ and $\hat{\Psi}(1) := \hat{\Phi}(1)^{-1}$, we can write using the Beveridge-Nelson decomposition, $T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}u_t^* = T^{-1/2}\sum_{t=1}^{\lfloor Tr\rfloor}\hat{\Psi}(1)\varepsilon_t^* + T^{-1/2}(\bar{u}_0^* - \bar{u}_{\lfloor Tr\rfloor}^*)$, where $\bar{u}_t^* := \hat{\Psi}(1)\sum_{i=1}^{q}(\sum_{j=i}^{q}\hat{\Phi}_j)u_{t-i+1}^*$. As $\left\|\hat{\Phi}_q - \Phi_q\right\| = o_p(q^{-1/2})$, it follows directly that $\hat{\Psi}(1) \xrightarrow{p} \Psi(1)$. Then (7) follows if we can show that $\mathrm{P}^*\left\{\max_{0\le t\le T}|T^{-1/2}\bar{u}_t^*| > \epsilon\right\} = o_p(1)$. As in Cavaliere and Taylor (2007, eq. (14)) we have that

$$\mathrm{P}^*\left\{\max_{0\le t\le T}\left|T^{-1/2}\bar{u}_t^*\right| > \epsilon\right\} \le \sum_{t=0}^{T}\mathrm{P}^*\left\{\left|\bar{u}_t^*\right| > \epsilon T^{1/2}\right\} \le \sum_{t=0}^{T}\frac{\mathrm{E}^*\left|\bar{u}_t^*\right|^{4+a}}{\epsilon^{4+a}T^{(4+a)/2}}$$

for some $a > 0$, by the Bonferroni and Markov inequality. In similar spirit as Park (2002, eq. (31)), we may write, for large $T$, that $u_t^* = \sum_{j=0}^{\infty}\psi_j^*\varepsilon_{t-j}^*$, using that for large $T$ the estimated lag polynomial will be invertible. Furthermore, $\bar{u}_t^* = \sum_{j=0}^{\infty}\bar{\psi}_j^*\varepsilon_{t-j}^*$, where $\bar{\psi}_j^* := \sum_{i=j+1}^{\infty}\psi_i^*$. Then, as $\varepsilon_t^* = \hat{\varepsilon}_{q,t}^{\tau}\xi_t^*$, by the Marcinkiewicz-Zygmund inequality and Minkowski's inequality we have that

$$\begin{aligned}
\mathrm{E}^*\left|\bar{u}_t^*\right|^{4+a} &= \mathrm{E}^*\left|\sum_{j=0}^{\infty}\bar{\psi}_j^*\varepsilon_{t-j}^*\right|^{4+a} \le c\,\mathrm{E}^*\left(\sum_{j=0}^{\infty}\bar{\psi}_j^{*2}\varepsilon_{t-j}^{*2}\right)^{(4+a)/2} \\
&\le c\left\{\sum_{j=0}^{\infty}\left(\mathrm{E}^*[\bar{\psi}_j^{*2}\varepsilon_{t-j}^{*2}]^{(4+a)/2}\right)^{2/(4+a)}\right\}^{(4+a)/2} \\
&= c\left\{\sum_{j=0}^{\infty}\bar{\psi}_j^{*2}\left[\mathrm{E}^*\left|\xi_t^*\hat{\varepsilon}_{q,t-j}^{\tau}\right|^{4+a}\right]^{2/(4+a)}\right\}^{(4+a)/2} \\
&\le c'\left(\sum_{j=0}^{\infty}\bar{\psi}_j^{*2}\left|\hat{\varepsilon}_{q,t-j}^{\tau}\right|^2\right)^{(4+a)/2},
\end{aligned}$$

where $c$ and $c'$ are constants not depending on $T$. Then

$$\sum_{t=0}^{T}\frac{\mathrm{E}^*\left|\bar{u}_t^*\right|^{4+a}}{\epsilon^{4+a}T^{2+a/2}} \le \sup_t\frac{\mathrm{E}^*\left|\bar{u}_t^*\right|^{4+a}}{\epsilon^{4+a}T^{1+a/2}} \le c''T^{-(2+a)/2}\sup_t\left|\hat{\varepsilon}_{q,t}^{\tau}\right|^{4+a}\left(\sum_{j=0}^{\infty}\bar{\psi}_j^{*2}\right)^{(4+a)/2},$$

where $c'' = c'\epsilon^{-(4+a)}$. It follows from Phillips and Solo (1992, p. 973) that $\sum_{j=0}^{\infty} \bar{\psi}_j^{*2} = O_p(1)$ if $\sum_{j=0}^{\infty} j^{1/2} \left|\psi_j^*\right| = O_p(1)$. This in turn follows if $\sum_{j=1}^{q} j^{1/2} \left|\hat{\phi}_{q,j}\right| = O_p(1)$ (cf. Hannan and Kavalieris, 1986, p. 30). This holds as $\sum_{j=1}^{q} j^{1/2} \left|\hat{\phi}_{q,j}\right| \leq \sum_{j=1}^{q} j^{1/2} \left|\hat{\phi}_{q,j} - \phi_j\right| + \sum_{j=1}^{q} j^{1/2} \left|\phi_j\right|$, while

$$\sum_{j=1}^{q} j^{1/2} \left|\hat{\phi}_{q,j} - \phi_j\right| \leq q^{1/2} \sum_{j=1}^{q} \left|\hat{\phi}_{q,j} - \phi_j\right| \leq q^{1/2} \left\|\hat{\Phi}_q - \Phi_q\right\| = q^{1/2} o_p(q^{-1/2}) = o_p(1)$$

and $\sum_{j=1}^{q} j^{1/2} \left|\phi_j\right| = O(1)$ by Assumption 1(i).

Therefore if we can show that $T^{-1} \sup_t \left|\hat{\varepsilon}_{q,t}^{\tau}\right|^{4+a} = O_p(1)$, the result follows. For expositional simplicity we assume that the residuals $\hat{\varepsilon}_{q,t}^{\tau}$ are obtained while imposing the null hypothesis and in absence of detrending. The argument we present below can straightforwardly, but tediously, be extended to allow for the inclusion of the lagged level in the regression and OLS or QD-detrending; cf. Smeekes (2009) for the homoskedastic case. Our proof is an adaptation of Park (2002, Proof of Lem. 3.2). Note that

$$\begin{aligned} T^{-1} \sup_t \left|\hat{\varepsilon}_{q,t}^{\tau}\right|^{4+a} &\leq T^{-1} \sum_{t=1}^{T} \left|\hat{\varepsilon}_{q,t-j}^{\tau}\right|^{4+a} \\ &\leq 3^{3+a} T^{-1} \sum_{t=1}^{T} \left\{ \left|\hat{\varepsilon}_{q,t}^{\tau} - \varepsilon_{q,t}\right|^{4+a} + \left|\varepsilon_{q,t} - \varepsilon_t\right|^{4+a} + \left|\varepsilon_t\right|^{4+a} \right\} \\ &=: 3^{3+a}(A_T + B_T + C_T), \end{aligned}$$

where $A_T$, $B_T$, and $C_T$ are implicitly defined. Note that $\varepsilon_{q,t}$ is defined as $\varepsilon_{q,t} := u_t - \sum_{j=1}^{q} \phi_j u_{t-j} = \varepsilon_t + \sum_{j=q+1}^{\infty} \phi_j u_{t-j}$.

We first look at $A_T$. We can write $\hat{\varepsilon}_{q,t}^{\tau} - \varepsilon_{q,t} = -\sum_{j=1}^{q} (\hat{\phi}_{q,j} - \phi_j) u_{t-j}$. Then, applying Cauchy's inequality we have

$$\begin{aligned} A_T &= T^{-1} \sum_{t=1}^{T} \left| \sum_{j=1}^{q} (\hat{\phi}_{q,j} - \phi_j) u_{t-j} \right|^{4+a} \\ &\leq T^{-1} \sum_{t=1}^{T} \left[ \sum_{j=1}^{q} \left(\hat{\phi}_{q,j} - \phi_j\right)^2 \right]^{(4+a)/2} \left[ \sum_{j=1}^{q} u_{t-j}^2 \right]^{(4+a)/2} \\ &= \left\|\hat{\Phi}_q - \Phi_q\right\|^{4+a} T^{-1} \sum_{t=1}^{T} \left( \sum_{j=1}^{q} u_{t-j}^2 \right)^{(4+a)/2} \\ &= o_p(q^{-(4+a)/2}) O_p(q^{(4+a)/2}) = o_p(1). \end{aligned}$$

For $B_T$ we can apply the Markov inequality to obtain $\mathrm{P}(B_T > \epsilon) = \mathrm{P}\left(T^{-1} \sum_{t=1}^{T} \left|\varepsilon_{q,t} - \varepsilon_t\right|^{4+a} > \epsilon\right) \leq \epsilon^{-1} T^{-1} \sum_{t=1}^{T} \mathrm{E}\left|\varepsilon_{q,t} - \varepsilon_t\right|^{4+a}$. Note we may write $\varepsilon_{q,t} - \varepsilon_t = \sum_{j=q+1}^{\infty} \phi_j u_{t-j} = \sum_{j=q+1}^{\infty} \phi_j \left(\sum_{k=0}^{\infty} \psi_j \varepsilon_{t-j-k}\right) = \sum_{j=q+1}^{\infty} \pi_{q,j} \varepsilon_{t-j}$, where $\pi_{q,j} :=$

$\sum_{k=q+1}^{j}\phi_k\psi_{j-k}$ and

$$\begin{aligned}\sum_{j=q+1}^{\infty}\left|\pi_{q,j}\right| &= \sum_{j=q+1}^{\infty}\left|\sum_{k=q+1}^{j}\phi_k\psi_{j-k}\right| \le \sum_{j=q+1}^{\infty}\sum_{k=q+1}^{j}|\phi_k|\left|\psi_{j-k}\right| \\ &\le \sum_{l=q+1}^{\infty}\sum_{m=0}^{\infty}|\phi_l||\psi_m| \\ &= \left(\sum_{l=q+1}^{\infty}|\phi_l|\right)\left(\sum_{m=0}^{\infty}|\psi_m|\right) \\ &= o(q^{-1})O(1) = o(q^{-1}),\end{aligned}$$

as Assumption 1 implies that $\sum_{j=q+1}^{\infty}\left|\phi_j\right| = o(q^{-1})$. Then

$$\begin{aligned}\mathrm{E}\left|\varepsilon_{q,t}-\varepsilon_t\right|^{4+a} = \mathrm{E}\left|\sum_{j=q+1}^{\infty}\pi_{q,j}\varepsilon_{t-j}\right|^{4+a} &\le \left(\sum_{j=q+1}^{\infty}\left|\pi_{q,j}\right|\left[\mathrm{E}\left|\varepsilon_{t-j}\right|^{4+a}\right]^{1/(4+a)}\right)^{4+a} \\ \le \max_t|\sigma_t|^{4+a}\,\mathrm{E}\,|e_t|^{4+a}\left(\sum_{j=q+1}^{\infty}\left|\pi_{q,j}\right|\right)^{4+a} &= o(q^{-(4+a)}),\end{aligned}$$

from which it follows that $B_T = o_p(1)$.

It follows straightforwardly that $C_T = O_p(1)$ as $T^{-1}\sum_{t=1}^{T}|\varepsilon_t|^{4+a} \le \max_t|\sigma_t|^{4+a}\,T^{-1}\sum_{t=1}^{T}|e_t|^{4+a} = O_p(1)$, by Assumption 1 and 1′.

Putting these results together we find that $T^{-1}\sup_t\left|\hat{\varepsilon}_{q,t}^{\tau}\right|^{4+a} = O_p(1)$, which in turn shows that $\mathrm{P}^*\left\{\max_{0\le t\le T}|T^{-1/2}\bar{u}_t^*| > \epsilon\right\} = o_p(1)$. Therefore we can conclude that (7) holds.

It then follows straightforwardly for cases (i) and (ii) that $T^{-1/2}\hat{x}_{\lfloor Tr\rfloor,\gamma}^{\delta*} \xrightarrow{d^*} K_{\eta,0,\gamma}^{\delta}(r)$ in probability, while for case (iii) it follows along the same lines as in the proof of Theorem 1 that $T^{-1/2}\hat{x}_{\lfloor Tr\rfloor,\gamma}^{\mu*} \xrightarrow{d^*} K_{\eta,c,\gamma,\tilde{\gamma}}^{\mu*}(r)$ in probability. Finally, the lag augmentation can again be handled as in Chang and Park (2003, Thm. 2) and Smeekes (2009, Lem. 5) to find the limiting distributions of the $DF\text{-}\gamma^{\delta*}$ statistics. ■