This is the accepted version of the following article:
Palm, F.C., S. Smeekes and J.-P. Urbain (2008). Bootstrap Unit-Root Tests: Comparison and Extensions. *Journal of Time Series Analysis* 29 (2), 371–401,
which has been published in final form at http://dx.doi.org/10.1111/j.1467-9892.2007.00565.x

Bootstrap Unit Root Tests: Comparison and Extensions

Franz C. Palm*, Stephan Smeekes, Jean-Pierre Urbain

Department of Quantitative Economics
Universiteit Maastricht
September 27, 2007

Abstract

In this paper we study and compare the properties of several bootstrap unit root tests recently proposed in the literature. The tests are Dickey-Fuller or Augmented DF-tests, either based on residuals from an autoregression and the use of the block bootstrap or on first differenced data and the use of the stationary bootstrap or sieve bootstrap. We extend the analysis by interchanging the data transformations (differences versus residuals), the types of bootstrap and the presence or absence of a correction for autocorrelation in the tests.

We show that two sieve bootstrap tests based on residuals remain asymptotically valid. In contrast to the literature which focuses on a comparison of the bootstrap tests with an asymptotic test, we compare the bootstrap tests among them using response surfaces for their size and power in a simulation study.

This study leads to the following conclusions: (i) augmented DF-tests are always preferred to standard DF-tests; (ii) the sieve bootstrap performs better than the block bootstrap; (iii) difference-based tests appear to have slightly better size properties but residual-based tests appear more powerful.

Keywords: bootstrap unit root tests, monte carlo, response surface.

JEL Codes: C15, C22

*Corresponding author: Department of Quantitative Economics, Universiteit Maastricht, P.O. Box 616, 6200 MD Maastricht, The Netherlands. Email: f.palm@ke.unimaas.nl. We thank Jeroen van den Berg for his assistance with our simulations and Robert Taylor, an anonymous referee and participants at the Econometric Society World Congress, London, August 2005, for helpful comments and suggestions. The usual disclaimer applies.

1 Introduction

Due to the good performance of the bootstrap in finite samples for stationary processes, its application to nonstationary series has recently become increasingly popular. In this paper we study and compare the properties of some bootstrap unit root tests that have recently been proposed in the literature. We also introduce some new tests, show their first order asymptotic validity and compare them to existing tests. The tests considered are Dickey-Fuller (DF) or Augmented Dickey-Fuller (ADF) tests, either based on residuals from an autoregression and the use of the block bootstrap (Paparoditis and Politis, 2003) or on first differenced data and the use of the stationary bootstrap (Swensen, 2003a) or sieve bootstrap (Psaradakis, 2001; Chang and Park, 2003). As mentioned, these papers differ in the way the bootstrap unit root tests. Besides showing the asymptotic validity,[1] all these papers compare the finite sample performance of their test(s) to the asymptotic counterpart(s), and the results are overall encouraging. It is however less clear how these tests perform compared to *each other*. The goal of this paper is to find out which tests perform best under circumstances to be given, and which aspects of the tests determine their finite sample performance. We will analyse and compare the asymptotic properties of these tests, and we will also consider Monte Carlo simulations.

We distinguish three main features of the tests. The first feature is the actual test statistic. Some tests use the DF test, others the ADF. As the ADF statistic is asymptotically pivotal, whereas the DF is not, we might expect a bootstrap ADF test to offer asymptotic refinements over the bootstrap DF test and asymptotic tests (Horowitz, 2001).[2] The second feature is which series exactly should be resampled. Bootstrapping a nonstationary series directly is not valid (Basawa et al., 1991). Therefore a stationary series has to be constructed first. Some tests use residuals from a first-order autoregression of the series, others use first-differences of the series. Swensen (2003b) shows that power functions are the same for both cases if the innovations are iid. However as shown by Paparoditis and Politis (2003, 2005), the use of differences leads to poor behaviour of the bootstrap tests under the alternative. The third feature is the time series bootstrap method that is employed. Some tests that we consider use some form of the *block bootstrap*, in which blocks of (restricted) residuals are resampled. Other tests use the *sieve bootstrap*, that fits an AR model to the (restricted) residuals and resamples the residuals of this AR model. The sieve bootstrap is somewhat easier to use and performs better when valid, but the block bootstrap is valid for more general processes.

Currently, to our knowledge no tests that use the sieve bootstrap based on residuals have been shown to be asymptotically valid for the Data Generating Processes (DGPs) considered in this paper. We adapt the sieve bootstrap tests by Psaradakis (2001) and Chang and Park

[1] We call a test *asymptotically valid* if the bootstrap distribution under the null converges to the asymptotic null distribution.

[2] Park (2003) shows that bootstrap ADF tests offer asymptotic refinements under the assumption the errors are a finite AR process with known order.

(2003) by constructing them using residuals instead of differences and show that these new tests are asymptotically valid. As residual-based tests may have better properties under the alternative than difference-based tests, this is an important extension. With these results, all the tests considered in this paper have been shown to be asymptotically valid.

A word on notation. We denote weak convergence by '$\xrightarrow{d}$', convergence in probability by '$\xrightarrow{p}$' and almost sure convergence by '$\xrightarrow{a.s.}$'. $W(r)$ indicates a standard Brownian motion. As usual, we use the superscript '$*$' to denote bootstrap quantities, both for bootstrap samples and statistics calculated for bootstrap samples. Similarly, '$\xrightarrow{d^*}$' indicates weak convergence of a bootstrap statistic conditional on the original series.

The structure of the paper is as follows. In Section 2, we discuss the bootstrap unit root tests, highlight several features of these tests and prove the asymptotic validity of the new tests proposed. Section 3 contains an extensive Monte Carlo simulation analysis of the various bootstrap unit root tests. The results are summarised using response surfaces. Section 4 concludes.

2 The tests

In this section we discuss several bootstrap unit root tests from a theoretical point of view.

2.1 DF sieve bootstrap test

2.1.1 Difference-based DF sieve bootstrap test: Psaradakis (2001)

Psaradakis (2001) considers the following DGP for the time series $y_t, t = 1, \ldots, n$:

$$y_t = d_t + v_t, \qquad v_t = \rho v_{t-1} + u_t, \tag{1}$$

where d_t consists of deterministic components. Three cases for the deterministic components are considered: the first case is without deterministics, $d_t = 0$, the second case is with only a constant term, $d_t^\mu = \delta_0$, and the third case is with constant term and linear time trend, $d_t^\tau = \delta_0 + \delta_1 t$. The process u_t is assumed to satisfy the following condition with $r = 4$ and $s = 1$:

Assumption 1

(i) the process u_t is generated by $u_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$, with ε_t a sequence of iid random variables with $\mathrm{E}[\varepsilon_t] = 0$, $\mathrm{E}[\varepsilon_t^2] = \sigma_\varepsilon^2 > 0$ and $\mathrm{E}[\varepsilon_t^r] < \infty$.

(ii) (A) Let $\psi_0 = 1$, $\sum_{j=1}^{\infty} j^s |\psi_j| < \infty$ and $\sum_{j=0}^{\infty} \psi_j \neq 0$.

(B) $\sum_{j=0}^{\infty} \psi_j z^j$ is bounded, and bounded away from zero for $\{z \in \mathbb{C} : |z| \leq 1\}$.

Note that this assumption implies that u_t is an invertible linear process; see Phillips and Solo (1992) for more details. We can rewrite the model (1) into the following form

$$y_t = \rho y_{t-1} + d_t^\dagger + u_t, \tag{2}$$

where $d_t^\dagger = \gamma_0 + \gamma_1 t := (1-\rho)\delta_0 + \rho\delta_1 + (1-\rho)\delta_1 t$ (in the first case $\delta_0 = \delta_1 = 0$, in the second case $\delta_1 = 0$). Psaradakis considers the DF coefficient test $n(\hat{\rho} - 1)$ and t-test in equation (2) for testing $\rho = 1$. As stated above, the assumptions on the innovations allow for a sieve bootstrap.

Psaradakis (2001) furthermore needs the following assumption on the order of the autoregression:

Assumption 2 *The order p of the autoregressive approximation is such that $p = p(n) \to \infty$ as $n \to \infty$ with $p(n) = o((n/\ln n)^{1/4})$.*

The exact bootstrap procedure can be described as follows.

Bootstrap Test 1 (Psaradakis, 2001)

1. *Fit an AR(p) model to $\hat{u}_t$, where $\hat{u}_t = \Delta y_t$ if the deterministic part consists of at most a constant term, and $\hat{u}_t = \Delta y_t - n^{-1}\sum_{t=1}^n \Delta y_t$ if the deterministic part contains both a constant term and a linear time trend, to obtain estimates $\hat{\phi}_{j,n}$ and*

$$\hat{\varepsilon}_{t,n} = \hat{u}_t - \sum_{j=1}^p \hat{\phi}_{j,n}\hat{u}_{t-j}, \quad t = 1+p, \dots, n.$$

2. *Generate an iid sample $\varepsilon_{t,n}^*$ by drawing randomly with replacement from $\hat{\varepsilon}_{t,n} - (n-p)^{-1}\sum_{t=1+p}^n \hat{\varepsilon}_{t,n}$.*

3. *Construct bootstrap errors by the recursion*

$$u_{t,n}^* = \sum_{j=1}^p \hat{\phi}_{j,n} u_{t-j,n}^* + \varepsilon_{t,n}^*. \tag{3}$$

4. *The bootstrap sample $y_{t,n}^*$ is generated recursively by*

$$y_{t,n}^* = y_{t-1,n}^* + u_{t,n}^*$$

in case of no deterministic components or an intercept only, and by

$$y_{t,n}^* = n^{-1}\sum_{t=1}^n \Delta y_t + y_{t-1,n}^* + u_{t,n}^*$$

in case of a constant term and a linear trend.

5. *Calculate the DF coefficient test and t-test using the bootstrap sample for the previously specified deterministic specification.*

6. *Repeat steps 2 to 5 B times to find the bootstrap distributions where B denotes the number of bootstrap replications.*

Psaradakis (2001) suggests to estimate the AR(p) model in step 1 using the Yule-Walker equations to ensure that the generated innovations $u^*_{t,n}$ admit a one-sided MA(∞) representation. The asymptotic distribution of the bootstrap statistics under the null is shown to be the same as the asymptotic distribution of the original DF statistics. Note that although the limiting distributions contain nuisance parameters, this does not matter for the bootstrap approach as the critical values for testing are based on the (empirical) distributions of the bootstrap tests that can be approximated by simulation with any accuracy desired.

2.1.2 Residual-based DF sieve bootstrap test: Psaradakis modified

The test we propose here is very similar to the test by Psaradakis (2001), except that it is based on residuals. Paparoditis and Politis (2005) have proposed an ADF coefficient test, and we construct our test in the same way as they do. We will show that our test is asymptotically valid when considering the assumptions made by Psaradakis (2001).

The new algorithm differs from that for the tests by Psaradakis (2001) only in step 1:

Bootstrap Test 2 (Residual-based DF sieve bootstrap procedure)
Replace step 1 from Bootstrap Test 1 by calculating the residuals from the regression

$$\hat{\varepsilon}_{t,n} = \tilde{y}_t - \hat{\rho}_n \tilde{y}_{t-1} - \sum_{j=1}^{p} \hat{\phi}_{j,n} \Delta \tilde{y}_{t-j}, \quad t = 1+p, \ldots, n, \tag{4}$$

where $\tilde{y}_t = y_t$ in the case of a (possibly zero) intercept and $\tilde{y}_t = y_t - \hat{\gamma}_0 - \hat{\gamma}_1 t$ in the case of a linear trend, and $\hat{\gamma}_0$ and $\hat{\gamma}_1$ are the corresponding OLS estimates.

The next theorem shows that, under the assumptions given above, the bootstrap distributions converge to the same limit distribution as the standard test statistics:

Theorem 1 *Let $\tau^*_n = n(\hat{\rho}^*_n - 1)$ and t^*_n be the coefficient and t-statistic, respectively, that follow from Bootstrap Procedure 2. Let $\sigma^2_u = \mathrm{E}[u^2_t]$ and $\sigma^2 = \lim_{n\to\infty} n^{-1}\,\mathrm{E}[(\sum_{t=1}^{n} u_t)^2]$. Under Assumptions 1 with $r = 4$ and $s = 1$ and 2, we have that*

$$\tau^*_n \xrightarrow{d^*} \frac{\int_0^1 W(r)dW(r) + (\sigma^2 - \sigma^2_u)/2\sigma^2}{\int_0^1 W(r)^2 dr} \quad \text{in probability}$$

$$t^*_n \xrightarrow{d^*} \frac{\int_0^1 W(r)dW(r) + (\sigma^2 - \sigma^2_u)/2\sigma^2}{\left((\sigma^2_u/\sigma^2)\int_0^1 W(r)^2 dr\right)^{1/2}} \quad \text{in probability.}$$

where $W(r)$ is a standard Brownian motion on $[0,1]$.

Proof: See Appendix A.

We have shown that the DF sieve test as constructed by Psaradakis (2001) remains asymptotically valid if it is based on residuals instead of differences.

2.2 ADF sieve bootstrap test

2.2.1 Difference-based ADF sieve bootstrap test: Chang and Park (2003)

Chang and Park (2003) consider the DGP

$$y_t = \rho y_{t-1} + u_t, \tag{5}$$

where $u_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j}$. They employ Assumption 1 with $r \geq 4$ and $s \geq 1$. For the order of the autoregressive approximation, Chang and Park (2003) consider two different assumptions:

Assumption 3 *Let $p(n) \to \infty$ and $p(n) = o(n^{\kappa})$ with $\kappa < \frac{1}{2}$ as $n \to \infty$.*

The following assumption is stronger.

Assumption 4 *Let $p(n) = cn^{\kappa}$ for some constant c and $1/rs < \kappa < \frac{1}{2}$.*

The bootstrap procedures of Chang and Park (2003) and Psaradakis (2001) are very similar:

Bootstrap Test 3 (Chang and Park, 2003)
Follow the same steps as in Bootstrap Test 1, but only for the deterministic specification $d_t = 0$. Replace step 5 by

5. *Calculate the ADF coefficient statistic $(1 - \sum_{j=1}^{p} \hat{\phi}_{j,n})^{-1} n(\hat{\rho}_n^* - 1)$ and the corresponding t-statistic*[3] *from the ADF regression*

$$y_{t,n}^* = \rho^* y_{t-1,n}^* + \sum_{j=1}^{p} \phi_j^* \Delta y_{t-j,n}^* + \varepsilon_t^*.$$

Chang and Park (2003) show that their bootstrap tests converge to the same asymptotic distributions under the null as the asymptotic tests. The convergence is shown to hold almost surely under the strong assumptions, and in probability under the weaker assumptions. They claim that their tests are also valid when applied to demeaned or detrended data, but they do not provide any further analysis.

[3]Chang and Park suggest using $\hat{\sigma}_{\varepsilon,n}^2$ (calculated from the original sample) for the t-test instead of $\hat{\sigma}_{\varepsilon,n}^{*2}$ (calculated from the bootstrap sample), although both are appropriate. Similarly, it is possible to use $1 - \sum_{j=1}^{p} \hat{\phi}_{j,n}^*$ for the coefficient test.

2.2.2 Residual-based ADF sieve bootstrap test: Chang and Park modified

Similar to the previous section, we construct a residual-based test that is based on the test by Chang and Park (2003) and resembles the residual-based ADF test of Paparoditis and Politis (2005) strongly.

Bootstrap Test 4 (Residual-based ADF sieve bootstrap test)
Replace step 1 from Bootstrap Test 3 by calculating the residuals from an ADF regression as in the following equation

$$\hat{\varepsilon}_{t,n} = y_t - \tilde{\rho}_n y_{t-1} \sum_{j=1}^{p} \hat{\phi}_{j,n} \Delta y_{t-j}, \quad t = 1 + p, \ldots, n. \tag{6}$$

The next theorem shows that, under the assumptions given above, the bootstrap distributions converge to the same limit distributions as the asymptotic test statistics.

Theorem 2 *Let τ_n^* and t_n^* be the bootstrap coefficient statistic and t-statistic, respectively, that follow from Bootstrap Test 4. Let Assumptions 1 with $r \geq 4$ and $s \geq 1$ and 3 hold. Then*

$$\tau_n^* \xrightarrow{d^*} \frac{\int_0^1 W(r)dW(r)}{\int_0^1 W(r)^2 dr} \quad \textit{in probability},$$

$$t_n^* \xrightarrow{d^*} \frac{\int_0^1 W(r)dW(r)}{\left(\int_0^1 W(r)^2 dr\right)^{1/2}} \quad \textit{in probability}.$$

Proof: See Appendix A.

We have shown that the ADF sieve test as constructed by Chang and Park (2003) is also asymptotically valid if it is based on residuals. In Theorem 2 we have obtained convergence in probability whereas Chang and Park (2003) proved a.s. convergence for their strong assumptions. By imposing the unit root restriction difference-based tests rely on stationary series for which a.s. convergence holds. Although not imposing the unit root when applying the sieve bootstrap is certainly a drawback, our result is worthwhile as it does provide justification for using a residual-based sieve bootstrap, even if it is not the same justification as Chang and Park (2003) provide for their tests.

For finite order AR(p) processes, Paparoditis and Politis (2005) show that under fixed alternatives the bootstrap distribution of the residual-based sieve bootstrap coefficient test is the same as that under the null. For the difference-based sieve bootstrap the distribution under the null differs from that under the alternative for the coefficient test, but not for the t-test. This results in a loss of power for the difference-based sieve bootstrap coefficient test. For the t-tests, both methods are asymptotically equivalent. For AR(∞) processes, Paparoditis and Politis (2005) do not discuss the residual-based sieve test, but they state that the difference-based sieve bootstrap is inappropriate as the differenced process is not invertible if the alternative is true.

2.3 (A)DF block bootstrap test

2.3.1 Residual-based (A)DF block bootstrap test: Paparoditis and Politis (2003)

Paparoditis and Politis (2003) propose a block bootstrap method to test for unit roots. Their method, the residual-based block bootstrap (RBB), is a block bootstrap method applied to residuals of a regression of the series y_t on its first lag. We first state the assumptions under which the RBB is appropriate. Two sets of assumptions are considered, such that one of these should be satisfied by the process y_t to validate the use of the RBB. Paparoditis and Politis (2003) consider the process $y_t = \alpha + \rho y_{t-1} + u_t$ where if $\alpha \neq 0$ there is a drift under the null of $\rho = 1$.

Paparoditis and Politis (2003) consider two sets for u_t. The first is that Assumption 1 (i) and (ii)(A) hold with $r = 4$ and $s = 1$ under the null. Under the alternative these assumptions should hold for y_t as well. Under the additional assumption (ii)(B) the process is invertible as well. This assumption is similar to those Psaradakis and Chang and Park employ.

The second assumption that Paparoditis and Politis (2003) use, is that u_t is strong mixing:

Assumption 5 *For each value of ρ, the series u_t is strong mixing and satisfies the following conditions:* $\mathrm{E}[u_t] = 0$, $\mathrm{E}\,|u_t|^r < \infty$ *for some* $r > 2$, $f_u(0) > 0$, *where f_u denotes the spectral density of u_t, i.e.,* $f_u(\lambda) = \sum_{h=-\infty}^{\infty} \gamma_u(h) \exp(i\lambda h)$ *and* $\gamma_u(h) = \mathrm{E}[u_t u_{t+h}]$. *Furthermore,* $\sum_{k=0}^{\infty} \alpha(k)^{1-2/r} < \infty$, *where $\alpha(\cdot)$ denotes the strong mixing coefficient of u_t.*

In contrast to the condition needed for the sieve bootstrap, one should note that the generating process of u_t does not have to be belong to the class of linear processes to satisfy this condition. Hence, we see here a class of processes (possibly non-linear) for which the block bootstrap is valid but the sieve bootstrap is not.

The procedure proposed by Paparoditis and Politis (2003) can be described as follows.

Bootstrap Test 5 (Paparoditis and Politis, 2003)

1. *Calculate the centred residuals*

$$\tilde{u}_{t,n} = \hat{u}_{t,n} - \frac{1}{n-1}\sum_{j=2}^{n} \hat{u}_{t,n} = (y_t - \tilde{\rho}_n y_{t-1}) - \frac{1}{n-1}\sum_{j=2}^{n}(y_t - \tilde{\rho}_n y_{t-1}),$$

   where $\tilde{\rho}_n$ is a consistent estimator of ρ.

2. *Choose the block length b, and draw points $i_0, i_1, \ldots, i_{k-1}$, where $k = \lfloor (n-1)/b \rfloor$,*[4] *from the uniform distribution on the set $\{1, 2, \ldots, n-b\}$. These points will serve as the*

[4] The bootstrap sample y_t^* will have total length $l = kb + 1$.

beginning points of the blocks of centred residuals:

$$y^*_{t,n} = \begin{cases} y_1 & \text{for } t = 1 \\ \alpha^* + y^*_{t-1,n} + \tilde{u}_{i_m+s,n} & \text{for } t = 2, 3, \ldots, n \end{cases}, \tag{7}$$

where $m = \lfloor (t-2)/b \rfloor$, $s = t - mb - 1$, *and* α^* *is a drift parameter that is either set equal to zero or it is a consistent estimator of* α.

3. *From the bootstrap series* $y^*_{t,n}$ *compute the desired statistics.*

4. *Repeat steps 2 to 3 B times to find the bootstrap distribution.*

Although most bootstrap unit root tests are based on differences, Paparoditis and Politis (2003) formally show that the residual-based block bootstrap coefficient test performs well asymptotically both under the null and under contiguous alternatives whereas the asymptotic distribution of the difference-based block bootstrap (DBB) statistic differs from that of the RBB statistic under the alternative, leading to a loss of power of the DBB test. Moreover, the convergence rate for this DBB bootstrap test is slower under the alternative than for the RBB test. For fixed alternatives, the slower rate of convergence leads to a loss of power of the DBB test compared to the RBB test. For sequences of n^{-1} local alternatives, the two tests have the same power.

In step 1 of the bootstrap procedure, $\tilde{\rho}_n$ should be a consistent estimator of ρ. Furthermore, it is required that $\tilde{\rho}_n$ is $o_p(1)$ if $\rho \neq 1$, $O_p(n^{-1})$ if $\rho = 1$ and $\alpha = 0$, and $O_p(n^{-3/2})$ if $\rho = 1$ and $\alpha \neq 0$. Many estimators satisfy these conditions. Paparoditis and Politis (2003) focus on what they call the least squares (LS) estimator, which is just the DF estimator, and the ADF estimator (they call this the DF estimator). For the validity of the ADF estimator the additional condition (ii)(B) is needed to ensure invertibility.

They prove the consistency of the RBB for the DF coefficient test and the ADF coefficient test. For models where $\alpha = 0$, they recommend to use the OLS estimator of ρ in

$$y_t = \alpha + \rho y_{t-1} + u_t \tag{8}$$

or the ADF equivalent[5] as $\tilde{\rho}_n$, which is used to construct the residuals. In the second step α^* is set to zero, as there should be no drift. They also recommend for the RBB ADF test to use the block bootstrap[6] of $y_t - y_{t-1}$ directly as lagged differences instead of $y^*_{t,n} - y^*_{t-1,n}$. For both the tests without deterministic components and the tests with a constant included, the consistency of the DF and ADF RBB tests is proved.

For the case of $\alpha \neq 0$, Paparoditis and Politis (2003) recommend using the same estimator for $\tilde{\rho}_n$ as before but setting $\alpha^* = \tilde{\alpha}_n$ where $\tilde{\alpha}_n$ is the estimator of α in (8). They prove the

[5]Depending on which unit root test is performed.
[6]Using the same blocks as for the residuals.

consistency of the DF coefficient test with constant and trend and claim the consistency of the corresponding ADF test can be established similarly.

2.3.2 Difference-based (A)DF block bootstrap test: Paparoditis and Politis (2003)

Again we consider an alternative version of the tests by Paparoditis and Politis (2003). As the original tests are based on residuals, the modified tests will be based on differences. The new procedure simply replaces $\tilde{\rho}_n$ by 1. Paparoditis and Politis (2003) already showed the asymptotic validity of these alternative tests (also see the discussion of power above).

2.4 DF stationary bootstrap test

2.4.1 Difference-based DF stationary bootstrap test: Swensen (2003a)

Swensen (2003a) considers a unit root test without deterministic components based on the stationary bootstrap of Politis and Romano (1994). He assumes the DGP $y_t = \rho y_{t-1} + u_t$ with the following assumptions on u_t.

Assumption 6

(i) The process u_t is strictly stationary with $\mathrm{E}[u_t] = 0$ for all t.

(ii) If $\gamma(k) = \mathrm{E}[u_t u_{t+k}]$, then $\gamma_0 + \sum_{r=0}^{\infty} |r\gamma(r)| < \infty$

(iii) $\sum_{r,s,t} \kappa_4(r, s, t) = K < \infty$ where $\kappa_4(r, s, t)$ is the fourth cumulant of the distribution of $(u_j, u_{j+r}, u_{j+r+s}, u_{j+r+s+t})$.

Assumption (iii) is used to ensure that the variance of $\frac{1}{n}\sum_t u_t^2$ tends to zero and implies that σ^2 can be consistently estimated. Under these conditions Swensen (2003a) proves the consistency of the DF tests without deterministic components based on the stationary bootstrap. The conditions needed are significantly weaker than those needed for the sieve bootstrap.

The algorithm can be described as follows:

Bootstrap Test 6 (Swensen, 2003)

1. *Compute centred differences*

$$\tilde{u}_t = \Delta y_t - (n-1)^{-1} \sum_{j=2}^{n} \Delta y_j.$$

2. *Apply the stationary bootstrap of Politis and Romano (1994) to the centred residuals to obtain bootstrap errors $u_{t,n}^*$:*

(a) *Draw the index of the starting points of the blocks,* $i_1, i_2, \ldots$, *from the uniform distribution* $P(t_1 = t) = \frac{1}{n}, t = 1, \ldots, n$. *Let* p_L *be a fixed number between 0 and 1. Draw the length of the blocks* $b_1, b_2, \ldots$ *from the geometric distribution* $P(b_1 = l) = (1 - p_L)^{l-1} p_L$. *The expected block length is* $1/p_L$.

(b) *Form blocks using the drawn starting points and block lengths. For block* $m + 1$ *we have*

$$u^*_{t,n} = \tilde{u}_{i_{m+1}+t-b(m)-1} \tag{9}$$

where $t = b(m) + 1, \ldots, b(m) + b_{m+1}$ *and* $b(m) = \sum_{j=1}^{m} b_j$.

(c) *Stop after generating* B *blocks if* $l_B = \sum_{j=1}^{B} b_j \geq n$. *Lay the blocks end-to-end in the order sampled, and cut off the resulting series* $u^*_{1,n}, \ldots, u^*_{l_B,n}$ *at* $u^*_{n,n}$ *if* $l_B > n$.

3. *Construct the bootstrap sample* $y^*_{t,n}$ *with the recursion* $y^*_{t,n} = y^*_{t-1,n} + u^*_{t,n}$.

4. *Compute the bootstrap DF coefficient and t-statistic.*

5. *Repeat steps 2 to 4 B times to find the bootstrap distribution.*

2.4.2 Residual-based DF stationary bootstrap test: Parker, Paparoditis, and Politis (2006)

Again, we also consider a modified version of these tests. Here we base the modified version on residuals instead of differences. Instead of the centred differences we calculate in step 1 centred residuals as in the bootstrap procedure of Paparoditis and Politis (2003). This test has been proposed by Parker et al. (2006) who also show its asymptotic validity.

2.5 Summary of tests considered

When comparing these tests, we will mainly focus on three aspects: whether differences or residuals are used, the bootstrap method and the test statistic.

Table 1 summarises all the test statistics and their main features. A note on the notation: we use τ for a coefficient test and t for a t-test. The first subscript indicates the bootstrap method: S stands for sieve bootstrap, B for block bootstrap, and St for stationary bootstrap; the second subscript indicates whether a test is based on differences (d) or residuals (r). A superscript a states that the test is an augmented DF test.

[Table 1 about here.]

3 Finite sample performance: Monte Carlo results

We analyse and compare the finite sample behaviour of the tests by Monte Carlo simulations.

3.1 Monte Carlo setup

We generate a series $y_t, t = 1, \dots, n$, according to the recursion

$$y_t = \rho y_{t-1} + u_t, \qquad y_0 = 0, \tag{10}$$

where different values for ρ are used: 1, 0.99, 0.95, 0.9 and 0.8. We let u_t be generated by an ARMA(1,1) process:

$$u_t = \phi u_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}, \tag{11}$$

where $\varepsilon_t \sim IN(0,1)$. The values used for ϕ and θ vary from -0.8 to 0.8. [7]

As sample sizes we consider $n = 50$, 100 and 250. We use three different significance levels: 0.01, 0.05 and 0.10. All experiments will be based on 5000 simulations and 999 bootstrap replications. All simulations are performed using GAUSS 6.0.

AIC is used to select the lag length for the sieve bootstrap. We estimate the AR(p) models by OLS.[8] For the lag length in the ADF tests we use the modified AIC by Ng and Perron (2001), both outside and inside the bootstrap procedures. For the block length we choose fixed numbers: 5 for $n = 50$, 8 for $n = 100$ and 15 for $n = 250$. The fact that there is no easy way to estimate block lengths remains a problem.

We perform two sets of simulations with these models. The first set considers the tests based on models without deterministic components. In the second set of simulations the DGPs remain unchanged but the tests are based on models with a constant and a trend. These extensions are not discussed in all papers, so that not all tests we consider have been shown to be theoretically valid. For the ADF test of Paparoditis and Politis (2003), we follow their instructions on how to handle the test allowing for a trend. Chang and Park (2003) indicate that their tests can be applied for the model with trend by applying the bootstrap test to the detrended data. We detrend both the original series (by OLS) and the bootstrap series.[9] For the test proposed by Swensen (2003a), deterministic components are added in the same way as in Psaradakis (2001).[10]

Note however that this is in fact not necessary, as the tests applied are actually invariant to the deterministic components present in the DGP, provided sufficient deterministics are included in the test regression. Therefore, the bootstrap test statistics are also invariant to the deterministics in the bootstrap DGP as long they are correctly specified in the bootstrap test regression.

[7] Specific values used for (ϕ, θ) are: $(0,0)$, $(-0.8,0)$, $(-0.4,0)$, $(0.4,0)$, $(0.8,0)$, $(0,-0.8)$, $(0,-0.4)$, $(0,0.4)$, $(0,0.8)$, $(0.4,0.4)$, $(-0.4,-0.4)$.

[8] The estimated AR(p) model may not be invertible. A solution could be to impose a root bound as in Burridge and Taylor (2004). This is however mainly important for empirical work, as in a large simulation study as ours the number of cases in which the estimated process is not invertible, is very small.

[9] It is crucial to detrend the bootstrap series as well, otherwise the bootstrap distribution will not converge to the correct asymptotic distribution.

[10] More efficient detrending methods are not considered in this paper.

The large number of DGP's and tests statistics in our simulations leads to a huge number of results that is rather hard to analyse in standard tables. We circumvent this problem by estimating different response surfaces for the rejection frequencies observed in our simulations for each of the test statistics. Because the empirical rejection frequency $\hat{P}$ lies between 0 and 1, we use the following transformation:

$$L(\hat{P}) = \ln\left(\frac{\hat{P}}{1-\hat{P}}\right). \tag{12}$$

The dependent variable is $L(\hat{P})$. As explanatory variables we consider several functions of the nominal level and the parameters in the underlying DGP. We will provide more details below. The specific form of the response surfaces is test specific. To avoid lengthy specification searches, we rely on *PcGets* (Hendry and Krolzig, 2001) to select the most appropriate specification from a large set of possible variables. The reported standard errors are White's heteroscedasticity consistent standard errors. Apart from the coefficient estimates and their standard errors, the adjusted R^2 of the regression is also reported.

3.2 Results

In this section we will give the main findings of our simulation study. We focus here on the results for the tests without deterministic trends. The results for the tests allowing for deterministic trends will be briefly discussed below.[11]

Size Tables 2 give a summary of the response surfaces for the size. We consider the following response surface for the size:

$$L(\hat{P}_i) = \beta_1 L(P_{a,i}) + \beta_2' f(L(P_{a,i}), \phi_i, \theta_i, n_i) + \nu_i, \quad i = 1, \ldots, M, \tag{13}$$

where P_a is the nominal size of the test, $f(L(P_{a,i}), \phi_i, \theta_i, n_i)$ is a vector of functions (all of order $O(n^{-1/2})$ and $O(n^{-1/2})$) of $L(P_a)$, the ARMA parameters ϕ_i and θ_i and the sample size n_i and ν_i denotes a disturbance. The number of simulation experiments M is 99.

The term $\beta_2' f(\cdot)$ captures the deviations of the actual size from the nominal size as a function of the parameters of the DGP and sample size. $\beta_1 L(P_{a,i})$ gives an indication of the asymptotic size of the tests. When β_1 is equal to 1, the empirical size of the test is equal to the nominal size for large n. The table gives the estimate of β_1 and its standard error, as well a measure of the fit.

[Table 2 about here.]

[11] The collection of all simulation results, response surfaces and graphical analyses is available on the Internet: `www.personeel.unimaas.nl/s.smeekes/outputreport.pdf`

Several things can be seen from the tables. We see that for some tests β_1 is significantly different from 1, although for most it is close to it. The estimates for the residual-based sieve tests are all not significantly different from one. Most of the other estimates are different from one, where especially the estimates for the DF test are far away from one. Note that these are all DF tests, as opposed to the ADF tests for which β_1 is much closer to 1. The coefficient and t-tests appear to have similar size in most cases. For the block tests, the value for β_1 is higher for the difference-based version than the residual-based version, which indicates that in general the residual-based block tests give higher rejection frequencies than difference-based block tests.

We also see for all tests that the fit increases when we include variables of higher order. Especially the increase in the fit from the first setting to the second is noticeable. This shows that all tests suffer from finite-sample distortions, although some more than others. As can be clearly seen, the adjusted R^2 for the regression on only the nominal size is much higher for the sieve tests than for the block tests. This shows that the (especially ADF) sieve tests suffer less from finite sample distortions than the other tests.

[Figure 1 about here.]

[Figure 2 about here.]

Figure 1 and 2 show graphs of the fitted size plotted against the autoregressive and moving-average parameters ϕ and θ. As nominal level we take 0.05 and as sample size we take 100. The green area indicates a size between 0.03 and 0.07, the blue area indicates a size below that range and the red area above that range.

The fitted transformed sizes are calculated from the response surfaces (13) for specific values ϕ_0, θ_0, n_0 and $P_{a,0}$, substituting estimates $\hat{\beta}_1$ and $\hat{\beta}_2$ for β_1 and β_2 respectively and dropping the disturbance ν_i. Next we apply the inverse of the $L(\cdot)$ transformation to the fitted values to obtain the fitted size.

As well as confirming what the tables tell us, the graphs show how the AR and MA parameter influence the empirical size. For all the tests, we see the well-known size distortions for a large negative MA parameter. The extent of this size distortion differs however. The stationary bootstrap tests and the DF block tests have massive size distortions, that also increase when the AR parameter becomes large and negative. We see that these tests are much more sensitive to the values of ϕ and θ, as they also exhibit a large undersize for large positive values. The ADF block tests mainly exhibit large undersize, especially for large absolute values of ϕ and θ. The sieve tests can be seen to perform quite well; especially the ADF sieve tests, for which the graphs are quite flat and in the correct range. We can also see that in general residual-based tests have higher rejection frequencies than the difference-based tests, except for the ADF sieve tests where both perform equally well.

Power Tables 3 and 4 give summaries of the response surfaces for the power. We choose to report only unadjusted power as we feel this is the most relevant, because this is what matters in practice. We now estimate the following response surface:

$$L(\hat{P}_i) = \beta_0 + \beta_1 L(P_{a,i}) + \sum_{k=1}^{3} \beta_{2,k}(\rho_i - 1)^k + \beta_3' f(\rho_i, L(P_{a,i}), \phi_i, \theta_i, n_i) + \nu_i, \quad i = 1, \ldots, M. \quad (14)$$

The number of simulation experiments M is 396. Again all variables in $f(\rho, L(P_{a,i}), \phi_i, \theta_i, n_i)$ are either of order $O(n^{-1/2})$ or $O(n^{-1})$. So in this case the asymptotic behaviour can be deduced from $\beta^a = (\beta_0, \beta_1, \beta_2')'$. The tables give the estimates plus standard error for the $O(1)$ variables and a measure of the fit. Again we see that the fit increases when we add higher order terms.

[Table 3 about here.]

[Table 4 about here.]

In Figures 3 to 4 we give power curves for varying sample sizes. These plots are derived from the response surfaces in the same way as the surface graphs for the size. For all cases, we have taken $\phi = \theta = 0$.[12] Most of the graphs show that the residual-based tests have higher power than the difference-based tests. However, as we also found that the residual-based tests have larger size distortions than the difference-based tests in general, the higher power will partly be caused by the size distortions. In that respect, we see that the power difference between residual-based and difference-based ADF sieve tests is quite small, while for these tests the behaviour under the null of residual-based tests and difference-based tests was also comparable. Hence, if there is a power advantage for residual-based tests, it is only small.

[Figure 3 about here.]

[Figure 4 about here.]

Deterministic trends The tests allowing for deterministic trends give qualitatively similar results as the ones described above. All tests perform worse, however the effect of including the deterministic trends where in fact none are needed is similar for all tests. Power becomes lower in general, and size seems to fluctuate more for different AR and MA parameters.

4 Conclusion

We have analysed the behaviour of a set of bootstrap unit root tests in finite samples. Moreover, we have shown the validity of two procedures that turn out to work well in finite samples.

[12]Unreported results show the dependence of the empirical power on ϕ and θ is similar as in the case of the size.

From our simulation study we can draw several conclusions. First, ADF tests clearly perform better than DF tests, which is what we expected from our discussion about asymptotically pivotal statistics. We do not observe a clear difference between the coefficient tests and t-tests.

Second, it seems that sieve tests perform better in terms of size than block tests for ARMA models, which is in line with the results for stationary series. We also see that the stationary bootstrap test performs worse in terms of size than the block bootstrap. Added to this, there is also a practical reason to use the sieve bootstrap. The selection of the lag length can be done quite easily, and appears to work if based on an information criterion like AIC or modified AIC. On the other hand, choosing the block length on the basis of intuition is difficult, and there exist no satisfactory methods for it. Taking all this into account, for our set of models the sieve bootstrap is preferable over the block bootstrap.

Third, the choice between difference-based tests and residual-based tests is less obvious. While the residual-based tests have higher power than the difference-based tests, these tests also have higher size distortions. However, when we consider ADF sieve bootstrap test, the residual-based tests perform similarly as the difference-based tests both in terms of size and in terms of power.

These findings are in line with the simulation results reported in the previous studies mentioned in the introduction, in the way the tests perform for different ARMA parameters. Our findings however allowed us to systematically compare existing and newly proposed tests. On the basis of previous studies only, it was not clear how the various tests compared.

Summarising, for the type of processes considered, the ADF sieve tests perform best in our simulation study. Therefore, for settings comparable to ours, we can recommend to use either the tests by Chang and Park (2003) or the ADF sieve tests based on residuals that we proposed. For other types of processes, allowing for broken trends, heteroskedasticity, etc., further research is needed.

References

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

Bühlmann, P. (1995). Moving-average representations of autoregressive approximations. *Stochastic Processes and their Applications 60*, 331–342.

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

Burridge, P. and A. M. R. Taylor (2004). Bootstrapping the HEGY seasonal unit root tests. *Journal of Econometrics 123*, 67–87.

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

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

Hendry, D. F. and H.-M. Krolzig (2001). *Automatic Econometric Model Selection*. London: Timberlake Consultants Press.

Horowitz, J. L. (2001). The bootstrap. In J. J. Heckman and E. E. Leamer (Eds.), *Handbook of Econometrics*, Volume 5, Chapter 52, pp. 3159–3228. Amsterdam: North Holland Publishing.

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

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

Paparoditis, E. and 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.

Park, J. Y. (2003). Bootstrap unit root tests. *Econometrica 71*, 1845–1895.

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

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

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

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

Swensen, A. R. (2003a). Bootstrapping unit root tests for integrated processes. *Journal of Time Series Analysis 24*, 99–126.

Swensen, A. R. (2003b). A note on the power of bootstrap unit root tests. *Econometric Theory 19*, 32–48.

A Proofs

Our proofs are adaptations of the proofs of Psaradakis (2001) and Chang and Park (2003) (which in turn depends on Park (2002)). We only elaborate where our proofs differ from theirs due to the use of residuals instead of differences. To be specific, the residuals to be resampled in our tests are constructed as

$$\hat{\varepsilon}_{t,n} = y_t - \hat{\rho}_n y_{t-1} - \sum_{j=1}^{p} \hat{\phi}_{j,n} \Delta y_{t-j}, \tag{15}$$

where $\hat{\rho}_n, \hat{\phi}_{1,n}, \ldots, \hat{\phi}_{p,n}$ are the OLS estimates from the (augmented Dickey-Fuller) regression of y_t on $y_{t-1}, \Delta y_{t-1}, \ldots, \Delta y_{t-p}$.

The residuals to be resampled in the tests of Psaradakis (2001) and Chang and Park (2003) are constructed as

$$\tilde{\varepsilon}_{t,n} = \Delta y_t - \sum_{j=1}^{p} \tilde{\phi}_{j,n} \Delta y_{t-j}, \tag{16}$$

where $\tilde{\phi}_{1,n}, \ldots, \tilde{\phi}_{p,n}$ are the OLS (or Yule-Walker) estimates from the regression of Δy_t on $\Delta y_{t-1}, \ldots, \Delta y_{t-p}$.

Let $\tilde{\phi}_n = (\tilde{\phi}_{1,n}, \ldots, \tilde{\phi}_{p,n})'$, $\hat{\phi}_n = (\hat{\phi}_{1,n}, \ldots, \hat{\phi}_{p,n})'$ and $x_{p,t} = (\Delta y_{t-1}, \ldots, \Delta y_{t-p})'$. Then $\hat{\phi}_n$ and $\tilde{\phi}_n$ are related by

$$\hat{\phi}_n = \tilde{\phi}_n + (\hat{\rho}_n - 1) \left(\sum_{t=1}^{n} x_{p,t} x_{p,t}' \right)^{-1} \left(\sum_{t=1}^{n} x_{p,t} y_{t-1} \right) \tag{17}$$

as in Chang and Park (2002, Proof of Lemma 3.5). From this we can deduce that

$$\hat{\phi}_{j,n} = \tilde{\phi}_{j,n} + O_p(n^{-1}) \tag{18}$$

under the null hypothesis of a unit root.

One consequence of not imposing the unit root restriction is that ρ has to be estimated so that we are only able to show some results in terms of convergence in probability instead of almost sure convergence.

Note that we only focus on the bootstrap distributions under the null, in line with most of the literature and specifically the papers that we base the new tests on. To analyse power properties, one needs to look at the bootstrap distribution under alternatives as well. In the main text we discuss the findings of Paparoditis and Politis (2005), who consider the power of these type of tests.

A.1 Proof of Theorem 1

In order to prove this theorem we need the following lemmas:

Lemma 1 *Suppose Assumptions 1 (with $r = 4$ and $s = 1$) and 2 hold. Then*

$$\mathrm{E}^*[(\varepsilon_{t,n}^*)^{2w}] = \mathrm{E}[(\varepsilon_t)^{2w}] + o_p(1) \qquad \textit{for } w = 1, 2. \tag{19}$$

Proof of Lemma 1 We adapt the proof of Bühlmann (1997, Proof of Lemma 5.3). First note that

$$\mathrm{E}^*[(\varepsilon_{t,n}^*)^{2w}] = (n-p)^{-1} \sum_{t=p+1}^{n} (\hat{\varepsilon}_{t,n} - \hat{\varepsilon}_n^{(\cdot)})^{2w}, \tag{20}$$

where $\hat{\varepsilon}_n^{(\cdot)} = (n-p)^{-1} \sum_{t=p+1}^{n} \hat{\varepsilon}_{t,n}$.

We first show that

$$\hat{\varepsilon}_n^{(\cdot)} = o_p(1). \tag{21}$$

Note that under the null

$$\varepsilon_t = \Delta y_t - \sum_{j=1}^{\infty} \phi_j \Delta y_{t-j} \tag{22a}$$

$$\hat{\varepsilon}_{t,n} = y_t - \hat{\rho}_n y_{t-1} - \sum_{j=1}^{p} \hat{\phi}_{j,n} \Delta y_{t-j}. \tag{22b}$$

Then we write

$$\begin{aligned}
\hat{\varepsilon}_n^{(\cdot)} &= (n-p)^{-1} \sum_{t=p+1}^{n} (\varepsilon_t - \varepsilon_t + \hat{\varepsilon}_{t,n}) \\
&= (n-p)^{-1} \sum_{t=p+1}^{n} \left[\varepsilon_t - (\Delta y_t - \sum_{j=1}^{\infty} \phi_j \Delta y_{t-j}) + (y_t - \hat{\rho}_n y_{t-1} - \sum_{j=1}^{p} \hat{\phi}_{j,n} \Delta y_{t-j}) \right] \\
&= (n-p)^{-1} \sum_{t=p+1}^{n} \Bigg[\varepsilon_t - (\Delta y_t - (y_t - \hat{\rho}_n y_{t-1})) \\
&\qquad - \sum_{j=1}^{p} (\hat{\phi}_{j,n} - \phi_j) \Delta y_{t-j} + \sum_{j=p+1}^{\infty} \phi_j \Delta y_{t-j} \Bigg] \\
&= (n-p)^{-1} \sum_{t=p+1}^{n} (A_{t,n} + B_{t,n} + C_{t,n} + D_{t,n})
\end{aligned} \tag{23}$$

Hence it has to be shown that the four right hand side components in (23) are $o_p(1)$.

It is trivial that $(n-p)^{-1} \sum_{t=p+1}^{n} A_{t,n}$ and $(n-p)^{-1} \sum_{t=p+1}^{n} D_{t,n}$ are $o_p(1)$. Next we turn to $B_{t,n}$:

$$\begin{aligned}
B_{t,n} &= -\Delta y_t + y_t - \hat{\rho}_n y_{t-1} \\
&= (1 - \hat{\rho}_n) y_{t-1}
\end{aligned} \tag{24}$$

Under the null $1 - \hat{\rho}_n = O_p(n^{-1})$ (Chang and Park, 2002), so that

$$(n-p)^{-1} \sum_{t=p+1}^{n} (1 - \hat{\rho}_n) y_{t-1} = (1 - \hat{\rho}_n)(n-p)^{-1} \sum_{t=p+1}^{n} y_{t-1} = o_p(1). \tag{25}$$

Finally, we consider $C_{t,n}$. By the Cauchy-Schwartz inequality,

$$\begin{aligned}
&\left| (n-p)^{-1} \sum_{t=p+1}^{n} \sum_{j=1}^{p} (\hat{\phi}_{j,n} - \phi_j) \Delta y_{t-j} \right| \\
&\qquad \leq \left[\sum_{j=1}^{p} (\hat{\phi}_{j,n} - \phi_j)^2 \right]^{1/2} \left[(n-p)^{-1} \sum_{t=p+1}^{n} \sum_{j=1}^{p} (\Delta y_{t-j})^2 \right]^{1/2}.
\end{aligned} \tag{26}$$

As $\hat{\phi}_{j,n} - \phi_j = O_p\left((\ln n/n)^{1/2}\right) + o\left(p^{-1}\right)$ (Chang and Park, 2002, Lemma 3.5) and $p(n) = o\left((n/\ln n)^{1/4}\right)$, we have

$$\begin{aligned}\left[\sum_{j=1}^{p}(\hat{\phi}_{j,n}-\phi_j)^2\right]^{1/2} &= \left[\sum_{j=1}^{p}\left\{O_p\left((\ln n/n)^{1/2}\right)+o\left(p^{-1}\right)\right\}^2\right]^{1/2}\\ &= \left[\sum_{j=1}^{p}\left\{O_p\left((\ln n/n)^{1/4}\right)+o_p\left(p^{-1}(\ln n/n)^{1/2}\right)+o\left(p^{-2}\right)\right\}\right]^{1/2} \\ &= \left[O_p\left(p(\ln n/n)^{1/4}\right)+o_p\left((\ln n/n)^{1/2}\right)+o\left(p^{-1}\right)\right]^{1/2}\\ &= \left[o_p\left((\ln n/n)^{1/2}\right)+o(p^{-1})\right]^{1/2}.\end{aligned} \tag{27}$$

Therefore,

$$(n-p)^{-1}\sum_{t=p+1}^{n}C_t = \left[o_p\left((\ln n/n)^{1/2}\right)+o(p^{-1})\right]^{1/2}O_p(p^{1/2}) = o_p(1). \tag{28}$$

Having shown (21), we now need to show that

$$(n-p)^{-1}\sum_{t=p+1}^{n}(\hat{\varepsilon}_{t,n})^{2w} = \mathrm{E}[(\varepsilon_t)^{2w}]+o_p(1). \tag{29}$$

As in (23), write

$$\begin{aligned}\hat{\varepsilon}_{t,n} &= \varepsilon_t-(\Delta y_t-(y_t-\hat{\rho}_n y_{t-1}))-\sum_{j=1}^{p}(\hat{\phi}_{j,n}-\phi_j)\Delta y_{t-j}+\sum_{j=p+1}^{\infty}\phi_j\Delta y_{t-j}\\ &= A_{t,n}+B_{t,n}+C_{t,n}+D_{t,n}.\end{aligned} \tag{30}$$

Using the arguments in (24) to (28), we have

$$(n-p)^{-1}\sum_{t=p+1}^{n}|B_{t,n}|^{2w} = o_p(1), \tag{31a}$$

$$(n-p)^{-1}\sum_{t=p+1}^{n}|C_{t,n}|^{2w} = o_p(1), \tag{31b}$$

$$(n-p)^{-1}\sum_{t=p+1}^{n}|D_{t,n}|^{2w} = o_p(1). \tag{31c}$$

Then

$$(n-p)^{-1}\sum_{t=p+1}^{n}(\hat{\varepsilon}_{t,n})^{2w} = (n-p)^{-1}\sum_{t=p+1}^{n}(A_{t,n}+B_{t,n}+C_{t,n}+D_{t,n})^{2w}.$$

If we expand the right-hand side of this last equation, we will get a sum of terms of the form

$$(n-p)^{-1}\sum_{t=p+1}^{n}A_{t,n}^{a}B_{t,n}^{b}C_{t,n}^{c}D_{t,n}^{d}$$

where $a,b,c,d\geq 0$ and $a+b+c+d=2w$. Next we apply Hölder's inequality to each of these terms.

Then, because of (31), all these terms are $o_p(1)$ apart from the term $(n-p)^{-1}\sum_{t=p+1}^{n} A_{t,n}^{2w}$. Hence,

$$(n-p)^{-1}\sum_{t=p+1}^{n}(\hat{\varepsilon}_{t,n})^{2w} = (n-p)^{-1}\sum_{t=p+1}^{n}(\varepsilon_{t,n})^{2w} + o_p(1).$$

We can then establish (29) by applying the weak law of large numbers.

Finally, we expand the right-hand side of (20) and again apply Hölder's inequality to the cross-terms. The proof is then completed using (29) and (21). □

Lemma 2 *Suppose Assumptions 1 (with $r = 4$ and $s = 1$) and 2 hold. Then as $n \to \infty$:*

(a) there exists a random variable n_0 such that $\sup_{n \geq n_0} \sum_{j=0}^{\infty} j|\hat{\psi}_{j,n}| < \infty$ in probability;

(b) $\sup_{0 \leq j \leq \infty} |\hat{\psi}_{j,n} - \psi_j| = o_p(1)$;

(c) $\mathrm{Var}^*[\varepsilon_{t,n}^*] - \sigma_\varepsilon^2 = o_p(1)$

(d) $\mathrm{Var}^*[n^{-1/2}\sum_{t=1}^{n}\varepsilon_{t,n}^*] - \sigma^2 = o_p(1)$.

Proof of Lemma 2 (a) As shown in Bühlmann (1995, Lemma 2.2) it is sufficient to prove that

$$\sum_{j=0}^{\infty} j|\hat{\phi}_{j,n} - \phi_j| = o_p(1). \tag{32}$$

From the triangular inequality, we have

$$\sum_{j=0}^{\infty} j|\hat{\phi}_{j,n} - \phi_j| \leq \sum_{j=0}^{\infty} j|\hat{\phi}_{j,n} - \tilde{\phi}_{j,n}| + \sum_{j=0}^{\infty} j|\tilde{\phi}_{j,n} - \phi_j|. \tag{33}$$

Bühlmann (1995, Proof of Lemma 3.1) shows that $\sum_{j=0}^{\infty} j|\tilde{\phi}_{j,n} - \phi_j| = o(1)$ a.s. and furthermore we have that

$$\sum_{j=0}^{\infty} j|\hat{\phi}_{j,n} - \tilde{\phi}_{j,n}| \leq p^2 \max_{1 \leq j \leq p} |\hat{\phi}_{j,n} - \tilde{\phi}_{j,n}| = o_p(1). \tag{34}$$

Hence, $\sum_{j=0}^{\infty} j|\hat{\phi}_{j,n} - \phi_j| = o_p(1)$ and the proof of part (a) is completed. □

Proof of Lemma 2 (b) See Bühlmann (1995, Proof of Theorem 3.2) for the proof. □

Proof of Lemma 2 (c) and (d) Using Lemma 1 and parts (a) and (b) these results follow as in Psaradakis (2001, Proof of Lemma 2). □

Lemma 3 *Let $S_n(r) = n^{-1/2}\sum_{t=1}^{\lfloor nr \rfloor} u_{t,n}^*$ and suppose Assumptions 1 (with $r = 4$ and $s = 1$) and 2 hold. Then*

$$S_n(r) \Rightarrow \sigma W(r). \tag{35}$$

Proof of Lemma 3 See Psaradakis (2001, Proof of Lemma 3). □

Lemma 4 *Let $\xi_{t,n}^* = \sum_{i=1}^{t} u_{t,n}^*$ and let Assumptions 1 (with $r = 4$ and $s = 1$) and 2 hold. Then* a $n^{-3/2}\sum_{t=1}^{n}\xi^*_{t-1,n} \Rightarrow \sigma\int_0^1 W(r)dr$,

b $n^{-2}\sum_{t=1}^{n}\xi^{*2}_{t-1,n} \Rightarrow \sigma^2\int_0^1 W(r)^2dr$,

c $n^{-5/2}\sum_{t=1}^{n}t\xi^*_{t-1,n} \Rightarrow \sigma\int_0^1 rW(r)dr$,

d $n^{-1}\sum_{t=1}^{n}\xi^*_{t-1,n}u^*_{t,n} \Rightarrow \sigma^2\int_0^1 W(r)dW(r)+\frac{1}{2}(\sigma^2-\sigma_u^2)$,

e $n^{-3/2}\sum_{t=1}^{n}tu^*_{t,n} \Rightarrow \sigma\int_0^1 rdW(r)$.

Proof of Lemma 4 See Psaradakis (2001, Proof of Lemma A.1). □

Proof of Theorem 1 See Psaradakis (2001, Proof of Theorem 1). □

A.2 Proof of Theorem 2

Again we first need several lemmas.

Lemma 5 *Let Assumption 1 (with $r \geq 4$ and $s \geq 1$) hold and let $p(n) = o((n/\ln n)^{1/2})$. Then it follows that*

(a) $\max_{1\leq j\leq p}|\hat{\phi}_{j,n}-\phi_j| = O_p((\ln n/n)^{1/2})+o(p^{-s})$

(b) $\hat{\sigma}_n^2 = \sigma^2+O_p((\ln n/n)^{1/2})+o(p^{-s})$

(c) $\sum_{j=1}^{p}\hat{\phi}_{j,n} = \sum_{j=1}^{\infty}\phi_j+O_p(p(\ln n/n)^{1/2})+o(p^{-s})$.

Proof of Lemma 5 part (a) See Chang and Park (2002, Lemma 3.5). □

Proof of Lemma 5 part (b) See Bühlmann (1995, Proof of Theorem 3.2) □

Proof of Lemma 5 part (c) See Chang and Park (2002, Lemma 3.5). □

Lemma 6 *Let Assumption 1 (with $r \geq 4$ and $s \geq 1$) hold and let $p(n) = o((n/\ln n)^{1/2})$. Then $n^{1-r/2}\,\mathrm{E}^*|\varepsilon^*_{t,n}|^r \xrightarrow{p} 0$ and*

$$W_n^*(i) = \frac{1}{\hat{\sigma}_n\sqrt{n}}\sum_{k=1}^{[ni]}\varepsilon^*_{k,n} \xrightarrow{d*} W(i) \qquad \text{as } n\to\infty.$$

Proof of Lemma 6 As shown in Park (2002, Theorem 2.2) we only need to show

$$n^{1-r/2}\,\mathrm{E}^*|\varepsilon^*_{t,n}|^r = n^{1-r/2}\left(\frac{1}{n}\sum_{t=1}^{n}\left|\hat{\varepsilon}_{t,n}-\frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_{t,n}\right|^r\right) \xrightarrow{p} 0. \tag{36}$$

Our proof will follow the lines of Park (2002, Proof of Lemma 3.2). We have that

$$\frac{1}{n}\sum_{t=1}^{n}\left|\hat{\varepsilon}_{t,n}-\frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_{t,n}\right|^r \leq c(A_n+B_n+C_n+D_n) \tag{37}$$

where

$$A_n = \frac{1}{n}\sum_{t=1}^{n} |\varepsilon_t|^r \qquad B_n = \frac{1}{n}\sum_{t=1}^{n} |\varepsilon_{t,n} - \varepsilon_t|^r$$

$$C_n = \frac{1}{n}\sum_{t=1}^{n} |\hat{\varepsilon}_{t,n} - \varepsilon_{t,n}|^r \qquad D_n = \left|\frac{1}{n}\sum_{t=1}^{n} \hat{\varepsilon}_{t,n}\right|^r$$

and

$$\varepsilon_{t,n} = \Delta y_t - \sum_{j=1}^{p} \phi_j \Delta y_{t-j}. \tag{38}$$

$\hat{\varepsilon}_{t,n}$ is defined in (15). Furthermore, let $\phi_{j,n}$ be defined such that in

$$\Delta y_t = \sum_{j=1}^{p} \phi_{j,n} \Delta y_{t-j} + e_{t,n}, \tag{39}$$

$e_{t,n}$ is uncorrelated with $\Delta y_{t-1}, \ldots, \Delta y_{t-p}$.

Hence we have to show that $n^{1-r/2}A_n$, $n^{1-r/2}B_n$, $n^{1-r/2}C_n$ and $n^{1-r/2}D_n \xrightarrow{p} 0$. The results for A_n and B_n are shown in Park (2002, Proof of Lemma 3.2).

Next we turn to C_n. We write

$$\begin{aligned}\hat{\varepsilon}_{t,n} &= y_t - \hat{\rho}_n y_{t-1} - \sum_{j=1}^{p} \hat{\phi}_{j,n} \Delta y_{t-j} \\ &= (y_t - \hat{\rho}_n y_{t-1} - \Delta y_t) + (\Delta y_t - \sum_{j=1}^{p} \hat{\phi}_{j,n} \Delta y_{t-j}) \\ &= -(\hat{\rho}_n - 1) y_{t-1} + (\varepsilon_{t,n} - \sum_{j=1}^{p} (\hat{\phi}_{j,n} - \phi_{j,n}) \Delta y_{t-j} - \sum_{j=1}^{p} (\phi_{j,n} - \phi_j) \Delta y_{t-j})\end{aligned} \tag{40}$$

It then follows that

$$|\hat{\varepsilon}_{t,n} - \varepsilon_{t,n}|^r \leq c \left(|(\hat{\rho}_n - 1) y_{t-1}|^r + \left|\sum_{j=1}^{p} (\hat{\phi}_{j,n} - \phi_{j,n}) \Delta y_{t-j}\right|^r + \left|\sum_{j=1}^{p} (\phi_{j,n} - \phi_j) \Delta y_{t-j}\right|^r \right). \tag{41}$$

where $c = 3^{r-1}$. We define

$$C_{0n} = \frac{1}{n}\sum_{t=1}^{n} |(\hat{\rho}_n - 1) y_{t-1}|^r \tag{42a}$$

$$C_{1n} = \frac{1}{n}\sum_{t=1}^{n} \left|\sum_{j=1}^{p} (\hat{\phi}_{j,n} - \phi_{j,n}) \Delta y_{t-j}\right|^r \tag{42b}$$

$$C_{2n} = \frac{1}{n}\sum_{t=1}^{n} \left|\sum_{j=1}^{p} (\phi_{j,n} - \phi_j) \Delta y_{t-j}\right|^r \tag{42c}$$

so that it needs to be shown that $n^{1-r/2}C_{in} \xrightarrow{a.s.} 0$ for $i = 0, 1, 2$. The result for C_{2n} follows from Park (2002, Proof of Lemma 3.2).

Again following Park (2002, Proof of Lemma 3.2), C_{1n} is majorised by

$$\begin{aligned}
&\left(\max_{1\le j\le p}|\hat{\phi}_{j,n}-\phi_{j,n}|^r\right)\frac{1}{n}\sum_{t=1}^{n}\sum_{j=1}^{p}|\Delta y_{t-j}|^r \\
&\le \left(\max_{1\le j\le p}|\hat{\phi}_{j,n}-\phi_{j,n}|^r\right)\frac{p}{n}\left(\sum_{t=0}^{n-1}|\Delta y_t|^r+\sum_{t=-1}^{1-p}|\Delta y_t|^r\right) \\
&\le c\left(\max_{1\le j\le p}(|\tilde{\phi}_{j,n}-\phi_{j,n}|^r+|\hat{\phi}_{j,n}-\tilde{\phi}_{j,n}|^r)\right)\frac{p}{n}\left(\sum_{t=0}^{n-1}|\Delta y_t|^r+\sum_{t=-1}^{1-p}|\Delta y_t|^r\right) \\
&= \left[O\left((\ln n/n)^r\right)+O_p\left(n^{-r}\right)\right](p/n)O(n)=O\left(p(\ln n/n)^r\right)=o_p\left((\ln n/n)^{r-1/2}\right).
\end{aligned} \tag{43}$$

As $r\ge 4$, $C_{1n}\xrightarrow{p}0$.

Next we consider C_{0n}. Rewrite the expression in (42a) for C_{0n} as

$$\frac{1}{n}\sum_{t=1}^{n}|(\hat{\rho}_n-1)y_{t-1}|^r=|\hat{\rho}_n-1|^r\frac{1}{n}\sum_{t=1}^{n}|y_{t-1}|^r=o_p(1). \tag{44}$$

This proves that $n^{1-r/2}C_n\xrightarrow{p}0$.

For D_n we need to prove that

$$\frac{1}{n}\sum_{t=1}^{n}\hat{\varepsilon}_{t,n}=\frac{1}{n}\sum_{t=1}^{n}\varepsilon_{t,n}+o_p(1)=\frac{1}{n}\sum_{t=1}^{n}\varepsilon_t+o_p(1), \tag{45}$$

which, by (40) and the result that $\varepsilon_{t,n}=\varepsilon_t+\sum_{j=p+1}^{\infty}\phi_j\Delta y_{t-j}$, holds if

$$\frac{1}{n}\sum_{t=1}^{n}\sum_{j=p+1}^{\infty}\phi_j\Delta y_{t-j}\xrightarrow{p}0, \tag{46}$$

$$\frac{1}{n}\sum_{t=1}^{n}\sum_{j=p+1}^{\infty}(\phi_{j,n}-\phi_j)\Delta y_{t-j}\xrightarrow{p}0, \tag{47}$$

$$\frac{1}{n}\sum_{t=1}^{n}\sum_{j=p+1}^{\infty}(\hat{\phi}_{j,n}-\phi_{j,n})\Delta y_{t-j}\xrightarrow{p}0, \tag{48}$$

$$\frac{1}{n}\sum_{t=1}^{p}(1-\hat{\rho}_n)y_{t-1}\xrightarrow{p}0, \tag{49}$$

where (46) and (47) follow from Park (2002, Proof of Lemma 3.2). For (48) we define

$$N_n=\sum_{j=1}^{p}(\hat{\phi}_{j,n}-\phi_{j,n})\sum_{t=1}^{n}\Delta y_{t-j} \tag{50}$$

and

$$Q_n=\sum_{j=1}^{p}\left|\sum_{t=1}^{n}\Delta y_{t-j}\right|. \tag{51}$$

Then N_n is dominated by

$$Q_n \max_{1 \le j \le p} |\hat{\phi}_{j,n} - \phi_{j,n}|. \tag{52}$$

Park (2002, Proof of Lemma 3.2) shows that $Q_n = o\left(pn^{1/2}(\ln n)^{1/r}(\ln \ln n)^{(1+\delta)/r}\right)$ a.s. for any $\delta > 0$. Furthermore,

$$\max_{1 \le j \le p} |\hat{\phi}_{j,n} - \phi_{j,n}| \le \max_{1 \le j \le p} (|\tilde{\phi}_{j,n} - \phi_{j,n}| + |\hat{\phi}_{j,n} - \tilde{\phi}_{j,n}|) = O\left((\ln n/n)^{1/2}\right) \text{ a.s.} + O_p\left(n^{-1}\right), \tag{53}$$

from which we can conclude that $N_n = o_p(n)$, which proves the result.

Finally, from (25) it is easy to see (49) holds as well. This completes the proof of D_n and hence of Lemma 6. □

Lemma 7 *Let Assumption 1 (with $r \ge 4$ and $s \ge 1$) hold and let $p(n) = O((n/\ln n)^{1/3})$. Then*

$$V_n^*(i) = \frac{1}{\sqrt{n}} \sum_{k=1}^{[ni]} u_{k,n}^* \xrightarrow{d*} \sigma \left(\sum_{j=0}^{\infty} \psi_j \right) W(i) \qquad \text{as } n \to \infty.$$

Proof of Lemma 7 Given Lemma 2, see Park (2002, Proof of Theorem 3.3). □

Lemma 8 *Let $\omega^2 = (1/n) \sum_{t=1}^n \Delta y_t$ and $\omega^{*2} = (1/n) \sum_{t=1}^n \Delta y_t^*$. Furthermore assume that Assumption 1 holds with $r \ge 4$ and $s \ge 1$. and $p(n) = o((n/\ln n)^{1/2})$. Then we have for any $\delta > 0$, $P[|\omega^{*2} - \omega^2| \ge \delta] \xrightarrow{p} 0$.*

Proof of Lemma 8 See Park (2002, Proof of Lemma 4.1). □

Proof of Theorem 2 Given Lemmas 5 to 8, all the relevant lemmas found in Chang and Park (2003) are valid for the test with residuals. The proof then concludes by Chang and Park (2003, Proof of Theorem 2). □

Table 1: Main features of the tests.

Test[a]	Bootstrap Method	Based on	Test Statistic
$\tau_{S,d}$	Sieve	Differences	DF τ
$t_{S,d}$	Sieve	Differences	DF t
$\tau_{S,r}$	Sieve	Residuals	DF τ
$t_{S,r}$	Sieve	Residuals	DF t
$\tau^a_{S,d}$	Sieve	Differences	ADF τ
$t^a_{S,d}$	Sieve	Differences	ADF t
$\tau^a_{S,r}$	Sieve	Residuals	ADF τ
$t^a_{S,r}$	Sieve	Residuals	ADF t
$\tau_{B,r}$	Block	Residuals	DF τ
$\tau_{B,d}$	Block	Differences	DF τ
$\tau^a_{B,r}$	Block	Residuals	ADF τ
$\tau^a_{B,d}$	Block	Differences	ADF τ
$\tau_{St,d}$	Stationary	Differences	DF τ
$t_{St,d}$	Stationary	Differences	DF t
$\tau_{St,r}$	Stationary	Residuals	DF τ
$t_{St,r}$	Stationary	Residuals	DF t

[a] We use τ for a coefficient test and t for a t-test. The first subscript indicates the bootstrap method: S stands for sieve bootstrap, B for block bootstrap, and St for stationary bootstrap; the second subscript indicates whether a test is based on differences (d) or residuals (r). A superscript a states that the test is an augmented DF test.

Panel A: Sieve tests								
	$\tau_{S,d}$	$t_{S,d}$	$\tau_{S,r}$	$t_{S,r}$	$\tau^a_{S,d}$	$t^a_{S,d}$	$\tau^a_{S,r}$	$t^a_{S,r}$
Explanatory variables up to order $O(1)$								
$L(P_a)$	1.15	1.15	0.95	0.95	1.01	0.91	0.95	0.93
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.01)
Adj. R^2	0.76	0.77	0.66	0.66	0.82	0.89	0.82	0.91
Explanatory variables up to order $O(n^{-1/2})$								
$L(P_a)$	0.86	0.87	0.96	0.96	0.95	0.99	0.97	1.02
	(0.03)	(0.03)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Adj. R^2	0.95	0.95	0.95	0.95	0.97	0.98	0.97	0.98
Explanatory variables up to order $O(n^{-1})$								
$L(P_a)$	0.88	0.87	0.98	1.00	0.86	0.96	1.00	1.00
	(0.03)	(0.04)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)
Adj. R^2	0.95	0.95	0.95	0.95	0.97	0.98	0.97	0.98
Panel B: Block-type tests								
	$\tau_{B,r}$	$\tau_{B,d}$	$\tau^a_{B,r}$	$\tau^a_{B,d}$	$\tau_{St,d}$	$t_{St,d}$	$\tau_{St,r}$	$t_{St,r}$
Explanatory variables up to order $O(1)$								
$L(P_a)$	0.68	1.05	1.04	1.34	1.20	1.23	0.69	0.64
	(0.08)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)	(0.09)	(0.09)
Adj. R^2	0.12	0.40	0.48	0.45	0.46	0.53	0.11	0.10
Explanatory variables up to order $O(n^{-1/2})$								
$L(P_a)$	0.50	1.03	1.17	1.72	1.15	1.12	0.51	0.53
	(0.12)	(0.06)	(0.06)	(0.12)	(0.06)	(0.06)	(0.14)	(0.13)
Adj. R^2	0.88	0.93	0.85	0.71	0.93	0.93	0.86	0.87
Explanatory variables up to order $O(n^{-1})$								
$L(P_a)$	0.90	1.07	0.92	1.30	1.23	1.19	0.91	0.87
	(0.03)	(0.05)	(0.02)	(0.05)	(0.06)	(0.05)	(0.03)	(0.03)
Adj. R^2	0.95	0.94	0.91	0.93	0.93	0.95	0.95	0.95

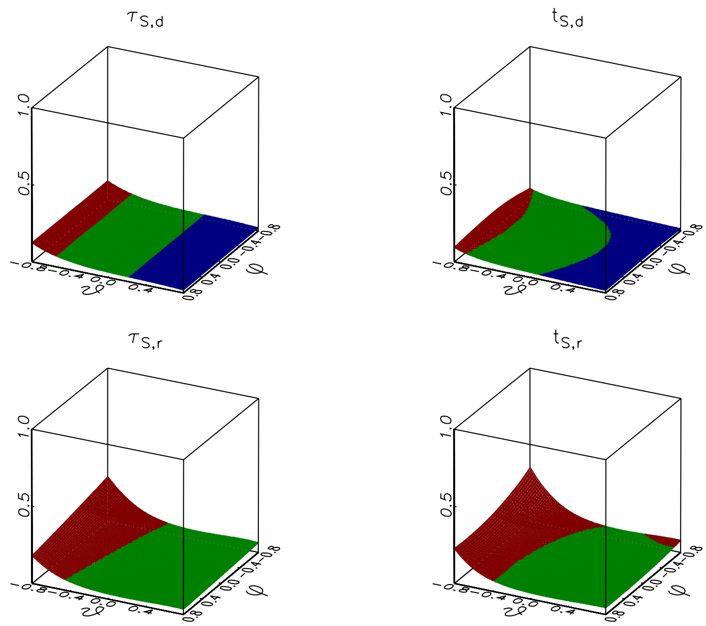
Table 2: Response surfaces of size

	$\tau_{S,d}$	$t_{S,d}$	$\tau_{S,r}$	$t_{S,r}$	$\tau^a_{S,d}$	$t^a_{S,d}$	$\tau^a_{S,r}$	$t^a_{S,r}$
Explanatory variables up to order $O(1)$								
$L(P_a)$	1.54	1.52	0.97	0.95	0.98	0.86	0.90	0.86
	(0.11)	(0.11)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.04)
Constant	1.50	1.47						
	(0.36)	(0.36)						
$(\rho-1)$	-50.58	-50.06	-58.09	-57.19	-54.45	-50.47	-51.80	-48.52
	(5.48)	(5.39)	(6.17)	(6.14)	(6.62)	(5.48)	(6.33)	(5.14)
$(\rho-1)^2$	-116.14	-116.80	-120.31	-117.36	-125.90	-123.84	-111.34	-122.96
	(27.79)	(27.35)	(31.37)	(31.37)	(31.61)	(26.28)	(30.68)	(24.47)
$(\rho-1)^3$								
Adj. R^2	0.55	0.55	0.51	0.50	0.49	0.51	0.49	0.52
Explanatory variables up to order $O(n^{-1/2})$								
$L(P_a)$	1.54	1.52	1.04	1.03	2.19	1.98	2.03	1.98
	(0.04)	(0.04)	(0.02)	(0.02)	(0.17)	(0.15)	(0.15)	(0.14)
Constant	2.75	2.77			4.52	3.92	3.94	3.78
	(0.21)	(0.21)			(0.59)	(0.52)	(0.55)	(0.52)
$(\rho-1)$	-131.44	-128.45	-162.95	-159.51	-140.67	-127.78	-139.64	-139.05
	(6.46)	(6.81)	(5.80)	(6.24)	(9.02)	(7.13)	(8.22)	(10.26)
$(\rho-1)^2$	-335.05	-328.89	-405.38	-393.02	-362.25	-336.56	-340.95	-564.93
	(33.69)	(35.75)	(32.58)	(35.33)	(43.06)	(35.43)	(39.92)	(106.11)
$(\rho-1)^3$								-741.68
								(327.03)
Adj. R^2	0.95	0.94	0.93	0.93	0.89	0.89	0.90	0.89
Explanatory variables up to order $O(n^{-1})$								
$L(P_a)$	3.36	3.39	1.33	1.04	3.40	3.17	3.20	3.16
	(0.39)	(0.40)	(0.13)	(0.02)	(0.61)	(0.51)	(0.55)	(0.50)
Constant	10.49	9.12	1.52		10.89	8.12	9.92	8.14
	(1.37)	(1.48)	(0.51)		(2.09)	(1.84)	(1.93)	(1.78)
$(\rho-1)$	-180.89	-235.58	-272.02	-276.88	-179.23	-242.31	-183.29	-238.07
	(8.11)	(18.61)	(14.28)	(13.88)	(9.43)	(31.97)	(8.73)	(30.72)
$(\rho-1)^2$	-477.17	-888.64	-742.83	-764.57	-362.25	-991.26	-340.95	-1048.42
	(52.85)	(133.58)	(87.88)	(92.06)	(38.60)	(239.34)	(33.83)	(231.53)
$(\rho-1)^3$	-473.46	-1240.03				-1484.91		-1616.62
	(165.86)	(274.24)				(493.75)		(481.09)
Adj. R^2	0.97	0.97	0.96	0.96	0.92	0.93	0.94	0.92

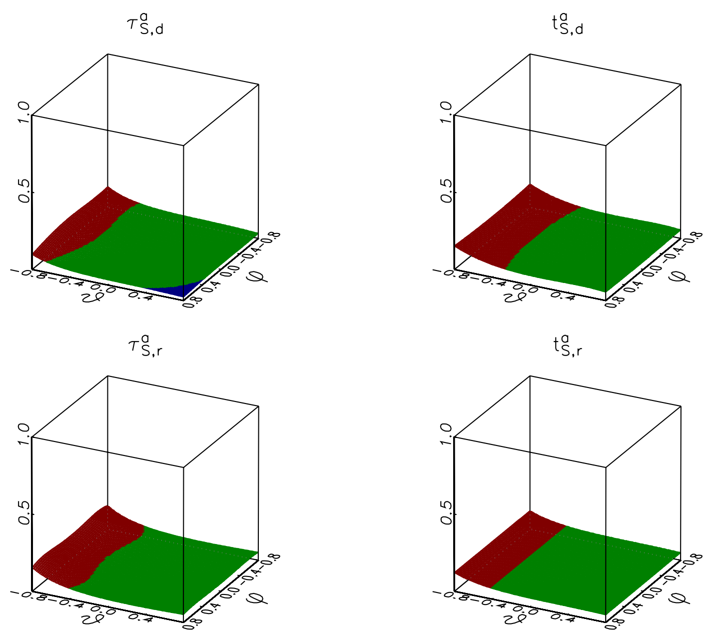
Table 3: Response surfaces of power - part I

	$\tau_{B,r}$	$\tau_{B,d}$	$\tau^a_{B,r}$	$\tau^a_{B,d}$	$\tau_{St,d}$	$t_{St,d}$	$\tau_{St,r}$	$t_{St,r}$
				Explanatory variables up to order $O(1)$				
$L(P_a)$	1.09	1.48	0.96	1.23	1.72	1.87	1.15	1.07
	(0.22)	(0.16)	(0.06)	(0.06)	(0.16)	(0.16)	(0.22)	(0.22)
Constant	1.79	1.48			1.83	2.27	2.02	1.97
	(0.78)	(0.55)			(0.55)	(0.53)	(0.82)	(0.81)
$(\rho-1)$	-71.93	-62.79	-52.52	-46.45	-59.59	-57.91	-70.66	-69.22
	(11.42)	(8.34)	(7.03)	(6.90)	(8.41)	(8.11)	(11.98)	(11.76)
$(\rho-1)^2$	-174.04	-123.95	-97.71	-81.61	-112.72	-112.41	-170.92	-169.64
	(52.80)	(40.98)	(33.92)	(32.77)	(41.34)	(40.01)	(55.16)	(53.93)
$(\rho-1)^3$								
Adj. R^2	0.31	0.47	0.50	0.49	0.48	0.49	0.29	0.28
				Explanatory variables up to order $O(n^{-1/2})$				
$L(P_a)$	0.48	1.48	1.93	2.40	1.72	1.52	0.46	0.68
	(0.21)	(0.06)	(0.18)	(0.30)	(0.06)	(0.12)	(0.21)	(0.24)
Constant	1.02	1.71	2.94	2.56	2.16	1.58	1.26	0.89
	(0.34)	(0.36)	(0.63)	(1.04)	(0.38)	(0.23)	(0.35)	(0.39)
$(\rho-1)$	-178.26	-171.88	-147.07	-133.75	-164.84	-161.46	-170.06	-265.95
	(14.23)	(10.89)	(10.53)	(15.79)	(11.50)	(11.47)	(14.47)	(50.81)
$(\rho-1)^2$	-562.84	-457.14	-334.33	-304.25	-422.11	-416.20	-533.99	-2035.21
	(65.82)	(52.36)	(48.79)	(74.40)	(55.18)	(56.45)	(67.15)	(711.28)
$(\rho-1)^3$								-5177.00
								(2381.08)
Adj. R^2	0.89	0.93	0.90	0.80	0.93	0.92	0.89	0.89
				Explanatory variables up to order $O(n^{-1})$				
$L(P_a)$	0.53	1.48	1.93	2.40	1.72	1.65		
	(0.17)	(0.05)	(0.14)	(0.17)	(0.05)	(0.10)		
Constant		0.70	3.49	3.70	1.06	1.58	1.02	0.94
		(0.20)	(0.52)	(0.65)	(0.20)	(0.20)	(0.34)	(0.33)
$(\rho-1)$	-351.34	-324.64	-204.24	-172.79	-313.36	-287.02	-276.15	-275.72
	(39.12)	(21.19)	(8.50)	(8.93)	(21.05)	(23.15)	(35.27)	(34.28)
$(\rho-1)^2$	-1979.17	-1068.85	-334.33	-304.25	-992.24	-889.42	-2067.15	-2150.97
	(350.62)	(130.46)	(36.81)	(42.35)	(129.22)	(133.88)	(505.94)	(488.18)
$(\rho-1)^3$	-3170.25						-5169.89	-5536.15
	(782.36)						(1732.32)	(1664.08)
Adj. R^2	0.91	0.95	0.94	0.91	0.95	0.95	0.92	0.92

Table 4: Response surfaces of power - part II

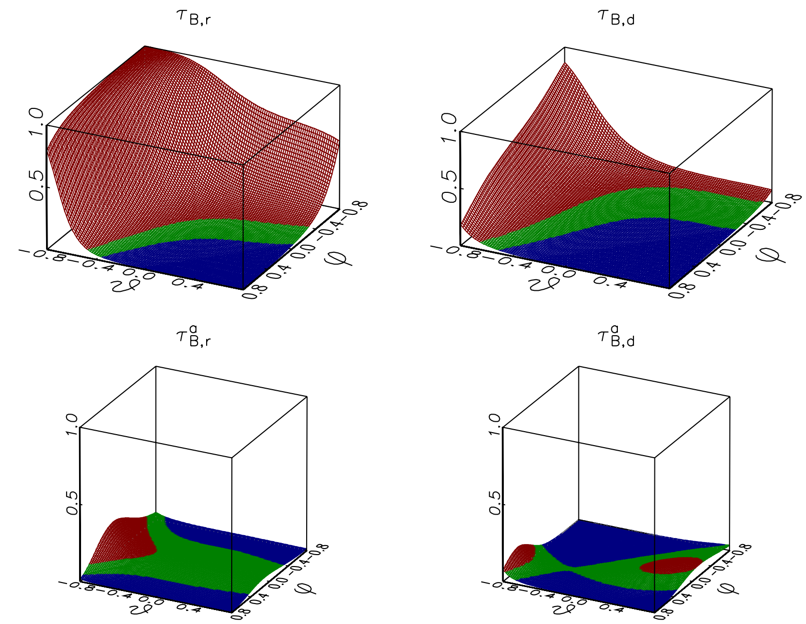


(a) DF tests

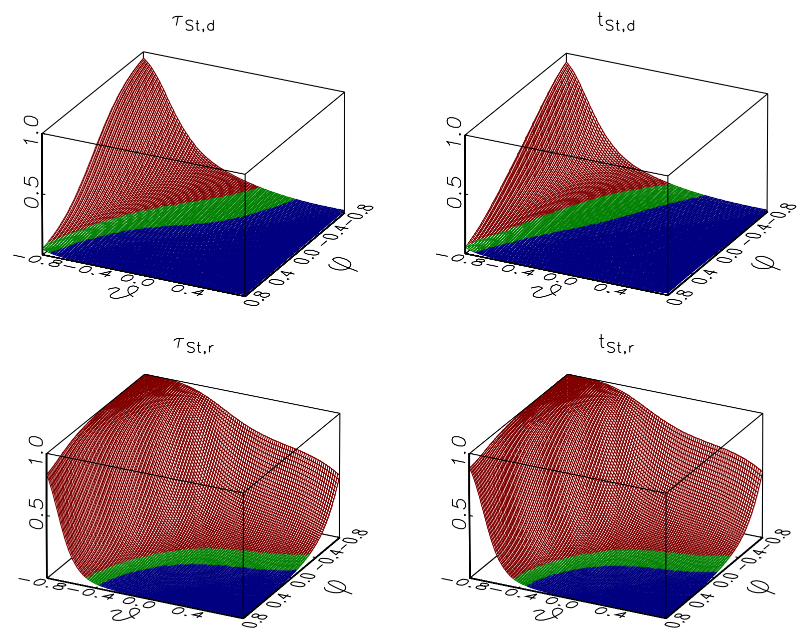


(b) ADF tests

Figure 1: Size as a function of ϕ and θ for sieve tests



(a) Block tests



(b) Stationary tests

Figure 2: Size as a function of ϕ and θ for block-type tests

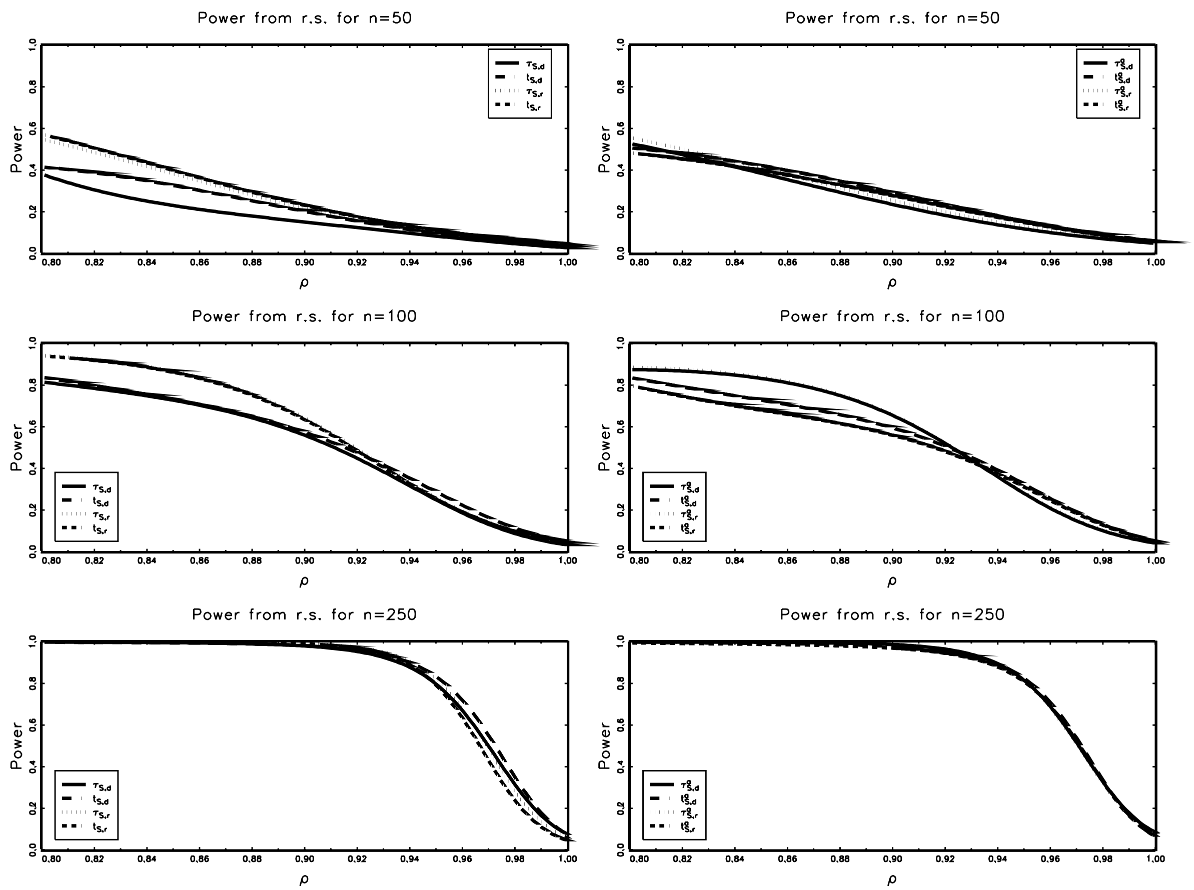


Figure 3: Power curves for sieve tests

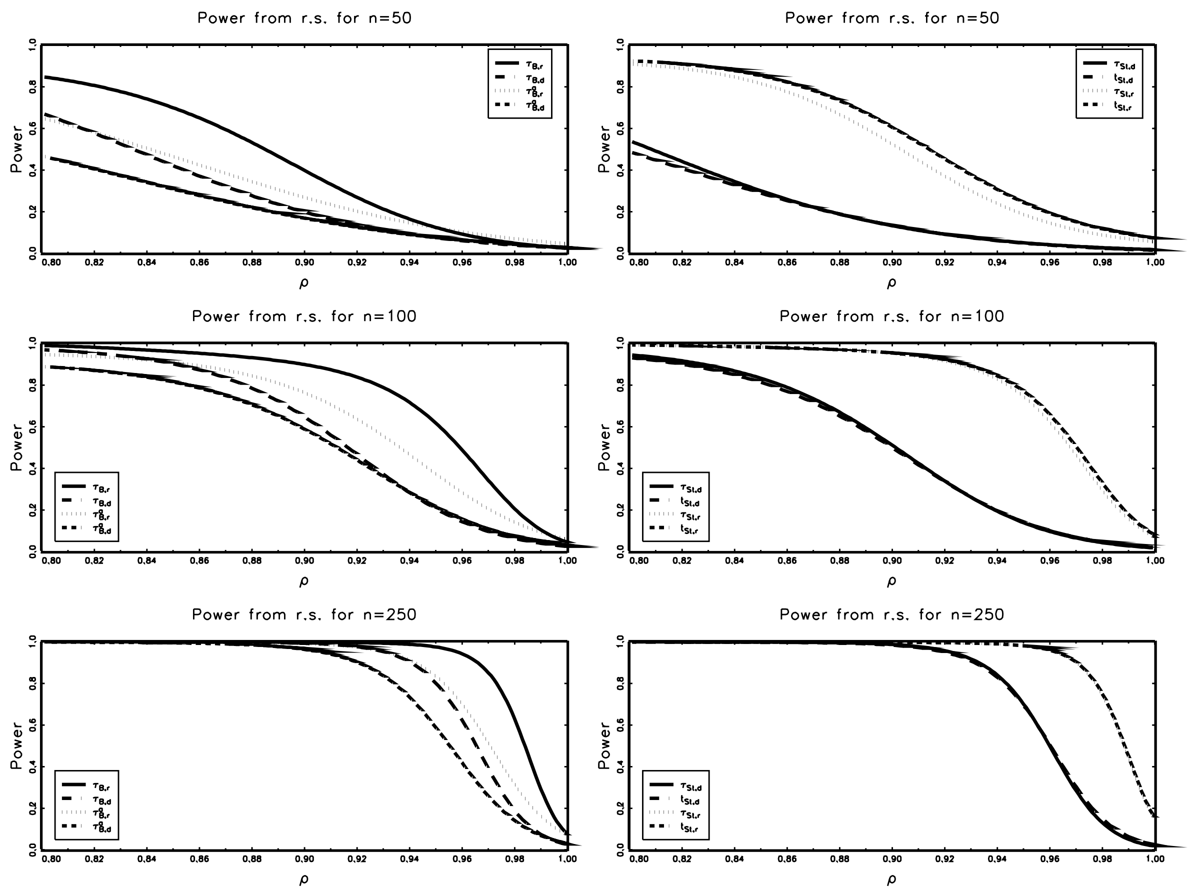


Figure 4: Power curves for block-type tests