Econometric Theory, **19**, 2003, 32–48. Printed in the United States of America.
DOI: 10.1017/S0266466603191025

A NOTE ON THE POWER OF BOOTSTRAP UNIT ROOT TESTS

ANDERS RYGH SWENSEN
University of Oslo

In this note we consider the asymptotic power functions of some bootstrap unit root tests under local alternatives and show that they are in fact the same as for ordinary unit root tests. This is regardless of whether the differences of the observations, i.e., the so-called restricted residuals, or the ordinary least squares residuals are used to construct the resampled observations. We also consider models containing a constant and a linear trend and the DF-GLS tests proposed by Elliott, Rothenberg, and Stock (1996, *Econometrica* 64, 813–836). A small Monte Carlo experiment is included.

1. INTRODUCTION

The recent surveys on bootstrap methods in time series by Li and Maddala (1996) and Berkowitz and Kilian (2000) both contain special sections on how to deal with nonstationary data and a discussion of bootstrap unit root tests. Bootstrap procedures offer an opportunity to take into account such factors as sample size, various specifications of the initial condition, and the distribution of the errors. They may therefore have more accurate finite-sample properties than procedures making use of asymptotic approximations, where such elements typically do not enter. Tabulations based on Monte Carlo simulations will necessarily have to specify these factors, and the assumptions may not always be well suited in particular applications. The bootstrap procedures present an interesting possibility making more extensive use of data and may therefore represent an alternative and supplement to tabulations and asymptotic approximations.

Most studies of bootstrap methods for unit root tests have been concerned with type I errors and, in particular, the question of whether the sizes of the tests correspond to the nominal levels. Ferretti and Romo (1996) consider a unit root bootstrap test for AR(1) models and demonstrate the asymptotic validity for independent and autoregressive errors. For a generalization where the order of the autoregressive error series is allowed to increase, see Swensen

I thank the associate editor, Bruce E. Hansen, and three anonymous referees for very constructive comments on the previous versions of the manuscript. Address correspondence to: Anders Rygh Swensen, Department of Mathematics, University of Oslo, P.O. Box 1053, N-0316 Blindern, Oslo, Norway; e-mail: swensen@math.uio.no.

 © 2003 Cambridge University Press 0266-4666/03 $9.50

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

(2002). Another way to develop bootstrap unit root tests for general error processes is to use block-bootstrap methods. In Swensen (2002) it is also shown that the stationary bootstrap of Politis and Romano (1994) can be used to construct bootstrap-based tests that asymptotically have correct level. Paparoditis and Politis (2000, 2001) also develop unit root tests based on the block bootstrap.

In this paper we shall address errors of type II or power. This is done in a situation where there are no nuisance parameters due to a general time series structure of the error process. We therefore assume a simple generating process where the observations $Y_1,\ldots,Y_T$ satisfy the recursion

$$Y_t = \alpha Y_{t-1} + \epsilon_t, \qquad t = 1,\ldots,T, \qquad Y_0 = 0, \tag{1}$$

where the random variables ϵ_t are independent and identically distributed with mean zero and variance σ^2.

The unit root case, i.e, when $\alpha = 1$, is of particular interest, and tests for this hypothesis versus $|\alpha| < 1$, which corresponds to stationary alternatives, have been much studied. The main conclusion of the present paper is that under local alternatives the asymptotic power of the bootstrap-based tests is the same as the asymptotic power of the usual tests for the presence of a unit root.

A closely related paper is a recent contribution by Hansen (1999) about grid bootstrap methods in autoregressive models. However, Hansen mainly discusses confidence intervals. To show that the asymptotic coverage is correct, he studies the distribution of the relevant statistics when both the parameter of the data generating process and the procedure used to construct the bootstrap samples are local to unity. In the present paper the focus is on power properties. Therefore the bootstrap samples are generated satisfying the null hypothesis, whereas the observations are assumed to follow an autoregressive recursion that is local to unity. Another relevant paper is Heimann and Kreiss (1996), where a hypothesis of the form $H_1: |\alpha| \geq 1$ against alternatives $K_1: |\alpha| < 1$ is studied. They show that a bootstrap procedure yields a test that is asymptotically unbiased and consistent. However, they do not derive the local power properties.

The paper is organized as follows. In the next section we introduce the setup and provide some background material. In section 3 we consider the situation where the model does not contain an intercept or a linear trend. It is shown that regardless of whether we use the differences of the observations or the ordinary least squares (OLS) residuals to construct the synthetic data, the local asymptotic power of the bootstrap unit root tests is the same as for the usual ones in this situation. In Section 4 we extend the results to cover models containing a time trend and/or an intercept. We also consider how the bootstrap procedures perform for the powerful generalized least squares Dickey–Fuller (DF-GLS) tests suggested by Elliott, Rothenberg, and Stock (1996). A small Monte Carlo experiment illustrating the results is provided. Some technical proofs are collected in Section 5.

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

2. BACKGROUND

For simplicity, we assume the observations are generated according to (1). However, one can verify that all the arguments that follow remain valid under the more general assumption that the initial variable, Y_0, has a fixed distribution not depending on T.

For the hypothesis that $\alpha = 1$ Dickey and Fuller (1979) propose a test based on the OLS estimator $\hat{\alpha}_T = \sum_{t=2}^{T} Y_t Y_{t-1} / \sum_{t=2}^{T} Y_{t-1}^2$. Simulated critical values for various levels and sample sizes can be found in Fuller (1996) under the assumption that the errors ϵ_t are normally distributed.

It is well known that when $\alpha = 1$ and $T \to \infty$, the distribution of $T(\hat{\alpha}_T - 1)$ converges toward the distribution of the random variable $\frac{1}{2}(W(1)^2 - 1)/\int_0^1 W(s)^2\, ds$, where $W = \{W(s)\}$ is a standard Brownian motion defined on the interval $[0,1]$. The t-statistic based on $\hat{\alpha}_T$ is defined as $t_{\alpha,T} = (\sum_{t=2}^{T} Y_{t-1}^2)^{1/2}(\hat{\alpha} - 1)/S$ where $S^2 = \sum_{t=2}^{T}(Y_t - \hat{\alpha}_T Y_{t-1})^2/(T-2)$. The asymptotic distribution is $\frac{1}{2}(W(1)^2 - 1)/(\int_0^1 W(s)^2\, ds)^{1/2}$.

For local alternatives of the form $\alpha_T = 1 + c/T$ Chan and Wei (1987) and Phillips (1987) show that the asymptotic distribution of $T(\hat{\alpha} - 1)$ is $\frac{1}{2}(J_c(1)^2 - 1)/\int_0^1 J_c(s)^2\, ds$, where J_c is the Ornstein–Uhlenbeck process defined by the stochastic differential equation

$$dJ_c(s) = cJ_c(s)ds + dW(s),$$

with initial condition $J_c(0) = 0$. Here values $c < 0$ correspond to stationary alternatives and $c > 0$ to the explosive ones.

Now we turn to the bootstrap test and consider the simplest case first. Let $\epsilon_1^*, \ldots, \epsilon_T^*$ be sampled with replacement from the centered differences

$$\hat{\epsilon}_{t,T} = \Delta Y_t - \overline{\Delta Y_T}, \qquad t = 2, \ldots, T,$$

where $\overline{\Delta Y_T} = [1/(T-1)]\sum_{t=2}^{T} \Delta Y_t$. Alternatively we may use the OLS residuals $Y_t - \hat{\alpha}_T Y_{t-1}$ after centering. Li and Maddala (1996) denote the two resampling schemes S_3 and S_2, respectively.

The resampled observations are computed by the recursion

$$Y_t^* = Y_{t-1}^* + \epsilon_t^*, \qquad t = 1, \ldots, T, \tag{2}$$

where $Y_0^* = 0$. Another possibility, which yields the same asymptotic distributions as long as the distribution of Y_0 is fixed, is to start the recursion at $Y_1^* = Y_1$, the first observation, and generate $Y_2^*, \ldots, Y_T^*$ according to (2). This seems a natural choice when the statistical inference is done conditional on the initial distribution.

Next, define $\hat{\alpha}_T^* = \sum_{t=2}^{T} Y_t^* Y_{t-1}^* / \sum_{t=2}^{T} Y_{t-1}^{*2}$ and $t_{\alpha,T}^* = (\sum_{t=2}^{T} Y_{t-1}^{*2})^{1/2} \times (\hat{\alpha}^* - 1)/S$. If $\alpha = 1$ and P^* denotes the conditional distribution of $Y_1^*, \ldots, Y_T^*$

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

given $Y_1,\ldots,Y_T$ Basawa, Mallik, McCormic, Reeves, and Taylor (1991b) show that for all x, if the sampling scheme S_3 is used,

$$P^*(T(\hat{\alpha}_T^* - 1) \leq x) \to P\left(\frac{1}{2}(W(1)^2 - 1)\Big/\int_0^1 W(s)^2\,ds \leq x\right)$$

a.s. In practice a bootstrap unit root test based on $\hat{\alpha}_T$, say, for the null hypothesis of a unit root, $H_0\colon \alpha = 1$ versus $K_0\colon \alpha < 1$, with asymptotic level γ is therefore obtained by rejecting the null hypothesis if the fraction $(1/B)\sum_{b=1}^{B} 1(\hat{\alpha}_T^* \leq \hat{\alpha}_T)$ is less than the level γ. Here B represents the number of times the recursion (2) and the statistic $\hat{\alpha}_T^* = \hat{\alpha}_{T,b}^*, b = 1,\ldots,B$ are recalculated based on new samples, and $1(\cdot)$ is the indicator function. Ferretti and Romo (1996) and Heimann and Kreiss (1996) show that the sampling scheme S_2 also has the same asymptotic properties.

It is important that (2) is used when the resampled observations Y_t^* are constructed. In fact, Basawa, Mallik, McCormic, Reeves, and Taylor (1991a) show that using $Y_t^* = \hat{\alpha}_T Y_{t-1}^* + \epsilon_t^*$ implies that the asymptotic distribution is random and depends on the realization $\{Y_t\}_{t=0}^{\infty}$.

In the arguments that follow we shall need some results from the theory of weak convergence of probability measures. We consider the space $D[0,1]$ of functions defined on the interval $[0,1]$ that are right continuous functions and have left limits. The limit processes will be continuous with probability 1 and hence, as explained in Pollard (1984), the uniform metric may be used. Also let E^* and Var^* refer to the conditional distribution P^*, so that, e.g., $E^* = E_{P^*}$.

3. MODELS WITHOUT CONSTANT AND LINEAR TREND

Assume that the parameter α has the form $\alpha_T = 1 + c/T$. From (1) the differences, or restricted residuals, can be written as

$$\Delta Y_t = (\alpha_T - 1)Y_{t-1} + \epsilon_t, \qquad t = 2,\ldots,T,$$

and therefore the centered differences $\hat{\epsilon}_t = \hat{\epsilon}_{t,T}$ are equal to

$$\hat{\epsilon}_t = (\alpha_T - 1)(Y_{t-1} - \overline{Y}_{T-1}) + (\epsilon_t - \bar{\epsilon}_T), \qquad t = 2,\ldots,T,$$

where $\overline{Y}_{T-1} = \sum_{t=1}^{T-1} Y_t/(T-1)$ and $\bar{\epsilon}_T = \sum_{t=2}^{T} \epsilon_t/(T-1)$.

Similarly the OLS residuals can be expressed as

$$Y_t - \hat{\alpha}_T Y_{t-1} = (\alpha_T - \hat{\alpha}_T)Y_{t-1} + \epsilon_t, \qquad t = 2,\ldots,T,$$

and the centered OLS residuals, $\hat{\hat{\epsilon}}_t = \hat{\hat{\epsilon}}_{t,T}$, as

$$\hat{\hat{\epsilon}}_t = (\alpha_T - \hat{\alpha}_T)(Y_{t-1} - \overline{Y}_{T-1}) + (\epsilon_t - \bar{\epsilon}_T), \qquad t = 2,\ldots,T.$$

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

From the definition of $\hat{\alpha}_T^*$,

$$T(\hat{\alpha}_T^* - 1) = T \frac{\sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \epsilon_s^* \right) \epsilon_t^*}{\sum_{t=2}^{T} \left(\sum_{s=1}^{t-1} \epsilon_s^* \right)^2}.$$

To find the asymptotic power function of the bootstrap test based on $T(\hat{\alpha}_T^* - 1)$ we shall need the following lemmas, which are valid for both sampling schemes S_3 and S_2, i.e., both centered differences and centered OLS residuals.

LEMMA 1. *Denote the distribution of the observations* $(Y_1, \ldots, Y_T)$ *corresponding to a local parameter* $\alpha_T = 1 + c/T$ *by* $P_{c,T}$. *Then the conditional moments of* $\epsilon_t^*, t = 1, \ldots, T$ *satisfy*

(i) $E^*(\epsilon_t^*) = 0$,
(ii) $\mathrm{Var}^*(\epsilon_t^*) \to \sigma^2$,

in $P_{c,T}$ *probability as* $T \to \infty$ *whether the* ϵ_t^**'s are sampled from the centered differences* $\hat{\epsilon}_t, t = 2, \ldots, T$ *or the centered* OLS *residuals* $\hat{\hat{\epsilon}}_t, t = 2, \ldots, T$.

Consider the random function $S_T^*(s) = \left(1/\sqrt{T}\right) \sum_{j=1}^{[Ts]} \epsilon_j^*, 0 \le s \le 1$ in $D[0,1]$. The conditional distribution of $\{S_T^*(s)\}$ given the observations converges weakly toward the distribution of $\{\sigma W(s)\}$ in $D[0,1]$ in $P_{c,T}$ probability. In other words, we have the following result.

LEMMA 2. *For all bounded continuous functions f defined on* $D[0,1]$

$$E^*[\, f(S_T^*)] \to E[\, f(\sigma W)]$$

in $P_{c,T}$ *probability as* $T \to \infty$.

LEMMA 3: *Let* F_T^* *be the conditional cumulative distribution function of* $T(\hat{\alpha}_T^* - 1)$ *given the observations, i.e.,* $F_T^*(x) = P^*(T(\hat{\alpha}_T^* - 1) \le x)$. *Then for all x*

$$F_T^*(x) \to P\left(\frac{1}{2} (W(1)^2 - 1) \Big/ \int_0^1 W(s)^2 \, ds \le x \right)$$

in $P_{c,T}$ *probability as* $T \to \infty$.

An implication of Lemma 3 is that a test having asymptotic level γ is obtained by rejecting the null hypotheses $H_0 : \alpha = 1$ if

$$F_T^*(T(\hat{\alpha}_T - 1)) \le \gamma. \qquad \textbf{(3)}$$

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

The power of the bootstrap test is therefore

$$P_{c,T}(F_T^{*}(T(\hat{\alpha}_T - 1)) \leq \gamma) = P_{c,T}(T(\hat{\alpha}_T - 1) \leq F_T^{*-1}(\gamma)),$$

where $F_T^{*-1}(\gamma) = \inf\{x : F_T^{*}(x) \geq \gamma\}$. By Lemma 3 $F_T^{*-1}(\gamma) \to DF_\gamma$, in $P_{c,T}$ probability, where DF_γ denotes the γ quantile of the Dickey–Fuller distribution defined by the random variable $\frac{1}{2}(W(1)^2 - 1)/\int_0^1 W(s)^2\, ds$. Hence, the asymptotic power function is $P(\frac{1}{2}(J_c(1)^2 - 1)/\int_0^1 J_c(s)^2\, ds \leq DF_\gamma)$. Because $S^2 \to \sigma^2$ in $P_{c,T}$ probability, there is an analogous result for the bootstrap test constructed using $t_{\alpha,T}^{*}$ and $t_{\alpha,T}$. If DFt_γ is the γ quantile for the distribution defined by the random variable $\frac{1}{2}(W(1)^2 - 1)/(\int_0^1 W(s)^2\, ds)^{1/2}$ we have the following result, using the fact that $\frac{1}{2}(J_c(1)^2 - 1) = \int_0^1 J_c(s)\, dJ_c$.

PROPOSITION 1. *Suppose the observations $Y_1, \ldots, Y_T$ are generated according to (1) where the ϵ_t^{*}'s are sampled from the centered differences $\hat{\epsilon}_t, t = 2, \ldots, T$ or the centered* OLS *residuals $\hat{\tilde{\epsilon}}_t, t = 2, \ldots, T$. The power function of the bootstrap test (3) and the power function of a bootstrap test based on $t_{\alpha,T}^{*}$ and $t_{\alpha,T}$ then converge toward*

$$P\left(\int_0^1 J_c(s)\, dJ_c \Big/ \int_0^1 J_c(s)^2\, ds \leq DF_\gamma\right) \tag{4}$$

and

$$P\left(\int_0^1 J_c(s)\, dJ_c \Big/ \left(\int_0^1 J_c(s)^2\, ds\right)^{1/2} \leq DFt_\gamma\right), \tag{5}$$

respectively, as $T \to \infty$.

Remark 1. The expressions given by (4) and (5) are, as shown by Phillips (1987), the same as the local power functions of the unit root tests based on $\hat{\alpha}_T$ and $t_{\alpha,T}$.

Remark 2. The important property of the variables $\epsilon_1^{*}, \ldots, \epsilon_T^{*}$ is that the conditional moments converge as described in Lemma 1. As we will discuss in the next section, when the model contains an intercept or a time trend, we can also use other resampling schemes for $\epsilon_1^{*}, \ldots, \epsilon_T^{*}$ that still ensure this property.

Remark 3. The conditional cumulative distribution F_T^{*} is a complicated function of the observations, so it cannot be computed explicitly. However, the fraction

$$\frac{1}{B}\sum_{b=1}^{B} 1(\hat{\alpha}_T^{*} \leq \hat{\alpha}_T), \tag{6}$$

where B bootstrap replications are used to compute the pseudo-observations $Y_1^{*}, \ldots, Y_T^{*}$, will converge toward the left-hand side of (3) as $B \to \infty$. When implementing bootstrap procedures it is important to choose the number of boot-

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

strap replications, B, large enough to approximate the conditional distribution of the bootstrapped statistics given the observations. For a recent paper on this issue, see Davidson and MacKinnon (2000).

Remark 4. Other estimators for the variance σ^2 in the definition of the studentized statistics used in the bootstrap procedure are possible. The important point is to use an estimator $\hat{S}^2$ of σ^2 such that for all $\delta, \epsilon > 0$

$$P_{c,T}(P^*(|\hat{S}^2 - \sigma^2| > \delta) > \epsilon) \to 0$$

as $T \to \infty$. Consider, e.g., $S^{*2} = \sum_{t=2}^{T}(Y_t^* - \hat{\alpha}_T Y_{t-1}^*)^2/(T-2)$. It is necessary to show that for all $\delta > 0$ $P^*(|S^{*2} - \sigma^2| > \delta)$ tends to zero in $P_{c,T}$ probability. By subtracting and adding Y_{t-1}^*, S^{*2} may be written as

$$\begin{aligned}\sum_{t=2}^{T}(Y_t^* - \hat{\alpha}_T Y_{t-1}^*)^2/(T-2) &= \sum_{t=2}^{T}(Y_t^* - Y_{t-1}^*)^2/(T-2) - 2T(\hat{\alpha}_T - 1)\sum_{t=2}^{T}(Y_t^* - Y_{t-1}^*)Y_{t-1}^*/T(T-2) \\ &\quad + T^2(\hat{\alpha}_T - 1)^2 \sum_{t=2}^{T} Y_{t-1}^{*2}/T^2(T-2). \qquad \textbf{(7)}\end{aligned}$$

Using Lemma 2 it follows by the continuous mapping theorem that the conditional distributions of $\sum_{t=2}^{T}(Y_t^* - Y_{t-1}^*)Y_{t-1}^*/T = \sum_{t=2}^{T}(\sum_{s=1}^{t-1}\epsilon_s^*)\epsilon_t^*/T$ and $\sum_{t=2}^{T} Y_{t-1}^{*2}/T^2 = \sum_{t=2}^{T}(\sum_{s=1}^{t-1}\epsilon_s^*)^2/T^2$ converge in distribution toward $\sigma^2(W(1)^2 - 1)/2$ and $\sigma^2\int_0^1 W(s)^2\,ds$, respectively, in $P_{c,T}$ probability. Also $T(\hat{\alpha}_T - 1)$ converges in $P_{c,T}$ distribution, so the two last terms in (7) are negligible. The first term equals $\sum_{t=2}^{T}\epsilon_t^{*2}/(T-2)$. But as explained in the proof of Lemma 3, $P^*(|\sum_{t=2}^{T}\epsilon_t^{*2}/(T-2) - \sigma^2| > \delta)$ tends to zero in $P_{c,T}$ probability, which implies that $P^*(|S^{*2} - \sigma^2| > \delta)$ also tends to zero in $P_{c,T}$ probability.

The same arguments apply to the variance estimator $S^{**2} = \sum_{t=2}^{T}(Y_t^* - \hat{\alpha}_T^* Y_{t-1}^*)^2/(T-2)$. We may proceed as before using the fact that the conditional distribution of $T(\hat{\alpha}_T^* - 1)$ converges in $P_{c,T}$ probability so that the terms playing the role of the two last terms in (7) are still negligible.

Remark 5. The result in Proposition 1 is of interest also for considering another candidate for bootstrapping. In Li and Maddala (1996, 1997) comparing $\hat{\alpha}_T - 1$ to the bootstrap distribution of $\hat{\alpha}_T^* - \hat{\alpha}_T$ is discussed. Based on evidence from simulation studies they do not recommend this procedure. Actually, the asymptotic distribution of this bootstrap test is easily derived from the earlier result. Remark that $\hat{\alpha}_T^* - \hat{\alpha}_T < \hat{\alpha}_T - 1$ is equivalent to $\hat{\alpha}_T^* - 1 < 2(\hat{\alpha}_T - 1)$. Using the notation introduced earlier, the probability of rejection is $P_{c,T}(2T(\hat{\alpha}_T - 1) \leq F_T^{*-1}(\gamma))$, which, as $T \to \infty$, converges toward $P(\frac{1}{2}(J_c(1)^2 - 1)/\int_0^1 J_c(s)^2\,ds \leq$

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

$(1/2)DF_\gamma$). Thus, taking $c = 0$, we see that the asymptotic size of this test will be too large when $DF_\gamma < 0$, which is the case for the usual nominal levels.

Remark 6. The asymptotic results are consistent with the results reported by Li and Maddala (1996, 1997), who, referring to a preprint of Ferretti and Romo (1996), state that in a simulation study of bootstrap unit root tests they have not found any difference between using differences, as done by Basawa et al. (1991b), and the OLS residuals used by Ferretti and Romo (1996).

4. MODELS WITH CONSTANT AND LINEAR TREND

In practice, models with an intercept or with a constant and linear trend are often estimated. In this section, we derive the asymptotic power of the corresponding bootstrap tests in this context.

Consider first fitting the equation

$$Y_t = \tilde{\mu}_T + \tilde{\alpha}_T Y_{t-1} + \tilde{\epsilon}_t, \qquad t = 2,\ldots,T, \tag{8}$$

where $\tilde{\mu}_T$ and $\tilde{\alpha}_T$ are the OLS estimators. Also in this case there are several possible resampling schemes. In addition to the schemes S_3 and S_2 of Li and Maddala (1996), corresponding to using centered differences and centered residuals, respectively, another possibility is basing the resampling on the OLS residuals from (8).

Because

$$(\tilde{\alpha}_T - 1) = \frac{\displaystyle\sum_{t=2}^{T} Y_{t-1}(Y_t - Y_{t-1}) - (Y_T - Y_1)\frac{\displaystyle\sum_{t=2}^{T} Y_{t-1}}{(T-1)}}{\displaystyle\sum_{t=2}^{T} Y_{t-1}^2 - \frac{\left(\displaystyle\sum_{t=2}^{T} Y_{t-1}\right)^2}{(T-1)}},$$

it follows from Lemma 2 and the continuous mapping theorem that if Y_t^* is constructed by (2) where the ϵ_t^*'s are obtained by resampling from the centered differences $\hat{\epsilon}_t$ based on $\Delta Y_t, t = 2,\ldots,T$ or the centered residuals $\hat{\tilde{\epsilon}}_t$ based on $Y_t - \hat{\alpha}_T Y_{t-1}, t = 2,\ldots,T$, the conditional distribution of $T(\tilde{\alpha}_T^* - 1)$ given the observations $Y_1,\ldots,Y_T$ will converge to the distribution of the random variable

$$\frac{\displaystyle\int_0^1 W^{\mu}(s)\,dW(s)}{\displaystyle\int_0^1 W^{\mu}(s)^2\,ds} = \frac{\displaystyle\int_0^1 \left(W(s) - \int_0^1 W(s)\,ds\right) dW(s)}{\displaystyle\int_0^1 \left(W(s) - \int_0^1 W(s)\,ds\right)^2 ds}. \tag{9}$$

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

As argued in Section 3 this will imply that the asymptotic power function of the bootstrap test constructed from $\tilde{\alpha}_T$ and $\tilde{\alpha}_T^*$ in the same way as $\hat{\alpha}_T$ and $\hat{\alpha}_T^*$ were used previously will be

$$P\left(\frac{\int_0^1 J_c^\mu(s)\,dJ_c(s)}{\int_0^1 J_c^\mu(s)^2\,ds} \leq DF_\gamma^\mu\right), \tag{10}$$

where DF_γ^μ denotes the γ quantile of the random variable defined in (9) and $J_c^\mu = J_c - \int_0^1 J_c(s)\,ds$. Expression (10) has the same form as (4). Similarly, the asymptotic power function of the studentized versions of the tests based on $\tilde{\alpha}_T$ and $\tilde{\alpha}_T^*$, and using resampling schemes S_2 and S_3, has the same form as (5).

The main reason why (10) is true is that that the conditional variance $\mathrm{Var}^*(\epsilon_t^*) \to \sigma^2$ in $P_{c,T}$ probability. Hence, the asymptotic power function based on resampling from the OLS residuals $\tilde{\epsilon}, t = 2,\ldots,T$ from (8) will also be the same if we show that the conditional variance converges when the ϵ_t^* are sampled from these residuals. Because a constant is included in the regression, the sum of the residuals is zero, so that $E^*(\epsilon_t^*) = 0$. The OLS residuals can be written, using (1), as $(\alpha_T - \tilde{\alpha}_T)Y_{t-1} - \tilde{\mu}_T + \epsilon_t, t = 2,\ldots,T$, and because the sum is zero this also equals $(\alpha_T - \tilde{\alpha}_T)(Y_{t-1} - \bar{Y}_{T-1}) + (\epsilon_t - \bar{\epsilon}_T)$, $t = 2,\ldots,T$ where $\bar{Y}_{T-1}$ and $\bar{\epsilon}_T$ are defined in Section 3. The OLS residuals therefore have the same form as the centered residuals from a regression without intercept except that $\tilde{\alpha}_T$ now takes the place of c or $\hat{\alpha}_T$. The convergence in probability of the conditional variance therefore follows by the same argument as in Lemma 1.

After some manipulation expression (10) may be written as

$$P\left(c + \frac{\int_0^1 J_c^\mu(s)\,dW(s)}{\int_0^1 J_c^\mu(s)^2\,ds} \leq DF_\gamma^\mu\right),$$

which shows that the asymptotic power function is identical to the power function of the corresponding unit root test derived in Phillips and Perron (1988).

We now consider fitting the equation

$$Y_t = \breve{\mu}_T + \breve{\beta}_T t + \breve{\alpha}_T Y_{t-1} + \breve{\epsilon}_t, \qquad t = 2,\ldots,T, \tag{11}$$

where $\breve{\mu}_T$, $\breve{\beta}_T$, and $\breve{\alpha}_T$ are the OLS estimators and $\breve{\epsilon}_t$, $t = 2,\ldots,T$ are the corresponding residuals.

As explained by Phillips and Perron (1988), because the model now contains a time trend, the distribution of the coefficient $\breve{\alpha}_T$ remains the same even when the recursion contains a nonzero drift term μ; i.e., the data are generated by the

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

recursion $Y_t = \mu + \alpha Y_{t-1} + \epsilon_t$. We therefore assume, without loss of generality, that $\mu = 0$ in the derivations that follows, i.e., that the data are generated by the recursion (1).

Phillips and Perron (1988) also show that when $\alpha = 1$ in (1), $T(\breve{\alpha}_T - 1)$ converges in distribution toward the distribution of the random variable of the form

$$\int_0^1 W^\tau(s)\, dW(s) \Big/ \int_0^1 W^\tau(s)^2\, ds, \tag{12}$$

where $W^\tau(s) = W(s) - (4 - 6s)\int_0^1 W(s)\, ds + (6 - 12s)\int_0^1 sW(s)\, ds$ (see Phillips and Perron, 1988, p. 339). Now, again from Lemma 1 in Phillips (1987) it follows that the process $Y_{[Ts]}/\sqrt{T}, 0 \le s \le 1$ converges weakly toward the Ornstein–Uhlenbeck process $\sigma J_c(s)$, $0 \le s \le 1$ as elements of $D[0,1]$ when $\alpha_T = 1 + c/T$. Building on this result Phillips and Perron show that $T(\breve{\alpha}_T - 1)$ converges in $P_{c,T}$ distribution toward a random variable defined in (12) but with J_c playing the role of W.

Proceeding as before, the residuals ϵ_t^* can be sampled from the centered differences $\hat{\epsilon}_t$, from the centered residuals $\hat{\tilde{\epsilon}}_t$, or from the residuals $\tilde{\epsilon}_t$. In fact, residuals from fitting (11) also may be used. This is done by Hansen (1999), who also demonstrates that it results in procedures having correct asymptotic properties. To see that in the present setting remark that these residuals may be written, using that $\sum_{t=2}^T \breve{\epsilon}_t = 0$,

$$\begin{aligned}
\breve{\epsilon}_t &= Y_t - \breve{\mu}_T - \breve{\beta}_T t - \breve{\alpha}_T Y_{t-1} \\
&= (\alpha_T - \breve{\alpha}_T) Y_{t-1} - \breve{\mu}_T - \breve{\beta}_T t + \epsilon_t \\
&= (\alpha_T - \breve{\alpha}_T)(Y_{t-1} - \overline{Y}_T) - \breve{\beta}_T \left(t - \frac{T^2 + T - 2}{2(T-1)} \right) + (\epsilon_t - \bar{\epsilon}_T), \\
& \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad t = 2, \ldots, T.
\end{aligned}$$

Using that $\breve{\beta}_T = O_P(T^{-3/2})$ (see, e.g., Banerjee, Dolado, Galbraith, and Hendry, 1993, p. 125), it follows by a similar argument as used in Section 5 in the proof of Lemma 1 that $\mathrm{Var}^*(\epsilon_t^*) \to \sigma^2$ in $P_{c,T}$ probability if $\epsilon_t^*, t = 1 \ldots, T$ are sampled from the OLS residuals $\breve{\epsilon}_t, t = 2, \ldots, T$. The restricted residuals will in this case be the residuals from imposing $\alpha = 1$ in estimating (11). From the preceding expression we see that these residuals can also be used without changing the asymptotic properties of the bootstrap tests in local neighborhoods of $\alpha = 1$.

The resampled observations can be generated by (2) or possibly (2) augmented with the estimated intercept $\tilde{\mu}_T$. We let $\breve{\mu}_T^*$, $\breve{\beta}_T^*$, and $\breve{\alpha}_T^*$ denote the estimators of a regression of the form (11) based on these resampled observations. From Lemma 2 it follows by an application of the continuous mapping theorem that the conditional distribution of $T(\breve{\alpha}_T^* - 1)$ given the observations $Y_1, \ldots, Y_T$

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

will converge to the distribution of the random variable (12) in $P_{c,T}$ probability. Thus, the power function of the bootstrap test based on $\breve{\alpha}_T$ and $\breve{\alpha}_T^*$ converges toward

$$P\left(\int_0^1 J_c^\tau(s)\,dJ_c(s)\Big/\int_0^1 J_c^\tau(s)^2\,ds \le DF_\gamma^\tau\right) \tag{13}$$

as $T \to \infty$, where DF_γ^τ is the γ quantile of the random variable given by (12) and $J_c^\tau(s) = J_c(s) - (4 - 6s)\int_0^1 J_c(s)\,ds + (6 - 12s)\int_0^1 sJ_c(s)\,ds$. Again the asymptotic power function has the form (4), and again the asymptotic power function for the studentized version of the bootstrap test may be expressed in a form similar to (5).

In models containing a constant or a linear trend there are redundant regressors when the observations are generated by the autoregressive scheme (1). This property is exploited by Elliott et al. (1996) to construct more powerful tests. Because this paper is about power properties, it may be of interest to consider the bootstrap versions of these tests.

They are defined as follows. Let $\bar{c}$ be a negative scalar and $Y = (Y_1,\ldots,Y_T)'$. Compute the so-called quasi-detrended observations and regressors $Y_{\bar{c}} = (Y_1, \Delta Y_2 - \bar{c}Y_1/T,\ldots,\Delta Y_T - \bar{c}Y_{T-1}/T)'$, $z_1 = (1,-\bar{c}/T,\ldots,-\bar{c}/T)'$, and $z_2 = (1,1 - \bar{c}/T,\ldots,1 - \bar{c}(T-1)/T)'$. Let $\hat{\beta}_0$ and $(\hat{\beta}_1,\hat{\beta}_2)$ be the coefficients from regressing $Y_{\bar{c}}$ on z_1 and (z_1,z_2), respectively. Using these estimates the demeaned observations are $Y_t^d = Y_t - \hat{\beta}_0$ when the model contains a constant, and the detrended observations are $Y_t^d = Y_t - \hat{\beta}_1 - \hat{\beta}_2 t$ when the model contains a constant and a linear trend. We can then compute the statistics $\hat{\alpha}_T^d = \sum_{t=2}^T Y_{t-1}^d Y_t^d / \sum_{t=2}^T (Y_{t-1}^d)^2$. The DF-GLS statistics are now defined as the t-statistics $[\sum_{t=2}^T (Y_{t-1}^d)^2]^{1/2}(\hat{\alpha}_T^d - 1)/\hat{\sigma}$. The estimator $\hat{\sigma}$ is based on the original observations and may be taken as $S_\mu^2 = [1/(T-3)]\sum_{t=2}^T (Y_t - \tilde{\mu}_T - \tilde{\alpha}_T Y_{t-1})^2$ when the model contains a constant and $S_\tau^2 = [1/(T-4)]\sum_{t=2}^T (Y_t - \breve{\mu}_T - \breve{\beta}t - \breve{\alpha}_T Y_{t-1})^2$ when the model contains a linear trend.

Elliott et al. (1996) showed that under local alternatives $1 + c/T$ the asymptotic distribution of the DF-GLS statistic is $\frac{1}{2}(J_c(1)^2 - 1)/(\int_0^1 J_c(s)^2\,ds)^{1/2}$ in the demeaned case. Thus the asymptotic power function is the same as for the Dickey–Fuller t-statistic in the model without constant and trend. In the detrended case the asymptotic distribution is given by $\frac{1}{2}(V_c(1,\bar{c})^2 - 1)/(\int_0^1 V_c(s,\bar{c})^2\,ds)^{1/2}$ where the process $\{V_c(t,\bar{c})\}$ is defined by

$$V_c(t,\bar{c}) = J_c(t) - t\left[\lambda J_c(1) + 3(1-\lambda)\int_0^1 sJ_c(s)\,ds\right], \qquad 0 \le t \le 1 \tag{14}$$

and $\lambda = (1 - \bar{c})/(1 - \bar{c} + \bar{c}^2/3)$. The parameter $\bar{c}$ in the definition of the DF-GLS statistics must be specified, and based on the previous calculations Elliott et al. (1996) recommend $\bar{c} = -7$ in the demeaned case and $\bar{c} = -13.7$ in the detrended case. For these values they calculate the critical values for various sample sizes by simulation.

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

To compute bootstrap versions of the DF-GLS tests we use resampled observations $Y_1^*,\ldots,Y_T^*$ constructed as described earlier to compute the quasi-detrended resampled observations $Y_{\bar{c}}^* = (Y_1^*, \Delta Y_2^* - \bar{c}Y_1^*/T,\ldots,\Delta Y_T^* - \bar{c}Y_{T-1}^*/T)'$. The coefficients $\hat{\beta}_0^*$ and $(\hat{\beta}_1^*, \hat{\beta}_2^*)$ are found by regressing $Y_{\bar{c}}^*$ on z_1 and (z_1, z_2), respectively. The demeaned and detrended resampled observations $Y_t^{*d}, t = 1,\ldots,T$ are then defined as $Y_t^* - \hat{\beta}_0^*$ and $Y_t^* - \hat{\beta}_0^* - \hat{\beta}_1^* t$, respectively. The vector Y^{*d} is computed for each sample and is used to find $\hat{\alpha}_T^{*d} = \sum_{t=2}^T Y_{t-1}^{*d} Y_t^{*d} / \sum_{t=2}^T (Y_{t-1}^{*d})^2$. The bootstrap version of the DF-GLS statistics is then $[\sum_{t=2}^T (Y_{t-1}^{*d})^2]^{1/2}(\hat{\alpha}_T^{*d} - 1)/\hat{\sigma}$. As discussed earlier, other estimators for σ^2 also can be used.

Let G_T^* and H_T^* be the conditional cumulative distribution functions of the DF-GLS statistic $[\sum_{t=2}^T (Y_{t-1}^{*d})^2]^{1/2}(\hat{\alpha}_T^{*d} - 1)/\hat{\sigma}$ in the demeaned and detrended cases, respectively. Then the asymptotic distribution can be described in the following analogue to Lemma 3.

LEMMA 4. *The conditional cumulative distributions G_T^* and H_T^* satisfy*

$$G_T^*(x) \to P\left(\frac{1}{2}(W(1)^2 - 1) \Big/ \left(\int_0^1 W(s)^2\, ds\right)^{1/2} \le x\right)$$

and

$$H_T^*(x) \to P\left(\frac{1}{2}(V_0(1,\bar{c})^2 - 1) \Big/ \left(\int_0^1 V_0(s,\bar{c})^2\, ds\right)^{1/2} \le x\right)$$

for all x in $P_{c,T}$ probability as $T \to \infty$, where $V_0(s,\bar{c})$ is defined in (14).

Thus, defining $DF - GLS_\gamma$ analogous to DFt_γ, but with $\{V_0(t,\bar{c})\}$ in place of $\{W(t)\}$, we can prove the following proposition by the same arguments used earlier.

PROPOSITION 2. *Suppose the observations $Y_1,\ldots,Y_T$ are generated according to (1). Then the power functions of the* DF-GLS *bootstrap tests converge toward*

$$P\left(\frac{1}{2}(J_c(1)^2 - 1) \Big/ \left(\int_0^1 J_c(s)^2\, ds\right)^{1/2} \le DFt_\gamma\right) \tag{15}$$

in the demeaned case and

$$P\left(\frac{1}{2}(V_c(1,\bar{c})^2 - 1) \Big/ \left(\int_0^1 V_c(s,\bar{c})^2\, ds\right)^{1/2} \le DF - GLS_\gamma\right) \tag{16}$$

in the detrended case, as $T \to \infty$, where $V_c(s,\bar{c})$ is defined in (14).

To illustrate the results in this section we have conducted a small Monte Carlo experiment for the case where the model contains a constant. For $T = 50$, $T = 100$, and $T = 250$, 10,000 samples were generated and the sim-

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

ulated observations calculated by the recursion (1) where $\alpha = 1 + c/T$ with $c = 0, -1, -2, -5, -10, -15$ and $c = -20$. The sequence $\{\epsilon_t\}$ consists of independent standard normal variables generated using the random number generator ran2 in Press, Teukolsky, Vetterling, and Flannery (1992).

For each sample, $Y_1, \ldots, Y_T$, 5,000 bootstrap replications were constructed using (2) where $\epsilon_1^*, \ldots, \epsilon_T^*$ were sampled with replacement. Three sampling schemes were used: centered differences of the simulated observations, centered residuals from a fit using OLS in a model without intercept, and residuals from a fit using OLS in a model with intercept. In each bootstrap sample the following model was fitted by OLS:

$$Y_t^* = \hat{\mu}^* + \hat{\alpha}^* Y_{t-1}^* + error^*, \qquad t = 2, \ldots, T.$$

In addition to $\hat{\alpha}^*$, two t-statistics were computed, one based on the variance estimator $S_\mu^2 = \sum_t (Y_t - \hat{\mu} - \hat{\alpha} Y_{t-1})^2$, which we denote by t_α^*, and another based on $S_\mu^{**2} = \sum_t (Y_t^* - \hat{\mu}^* - \hat{\alpha}^* Y_{t-1}^*)^2$, denoted by $t_{\alpha^*}^*$. The bootstrapped DF-GLS statistic is denoted by gls^*.

Table 1 shows the fraction of the samples for which the various bootstrap tests rejected the null hypothesis of a unit root at the 5% nominal significance level. There are several interesting features worth mentioning. The first is that we did not notice any difference between the three sampling schemes. All resulted in the figures reported in Table 1. This is similar to what Li and Maddala (1996, 1997) report for the case where the model does not contain an intercept. Also, all bootstrap tests have correct size. Concerning the power, the ranking of the bootstrap tests is the same as one can expect from the asymptotic results of Figure 2 in Elliott et al. (1996). The bootstrap DF-GLS test is the most powerful, followed by the test based on $\hat{\alpha}^*$, the coefficient estimated by OLS. The studentized versions have least power. The two versions of the studentized test

TABLE 1. Fraction of rejections of bootstrap unit root tests in model with constant

	$T = 50$				$T = 100$				$T = 250$			
c	$\hat{\alpha}^*$	t_α^*	$t_{\alpha^*}^*$	gls^*	$\hat{\alpha}^*$	t_α^*	$t_{\alpha^*}^*$	gls^*	$\hat{\alpha}^*$	t_α^*	$t_{\alpha^*}^*$	gls^*
0	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
−1	0.07	0.06	0.06	0.08	0.07	0.06	0.06	0.08	0.07	0.06	0.06	0.08
−2	0.09	0.07	0.06	0.11	0.10	0.07	0.07	0.12	0.10	0.07	0.07	0.12
−5	0.18	0.12	0.11	0.26	0.19	0.13	0.12	0.29	0.19	0.12	0.12	0.31
−10	0.48	0.36	0.33	0.61	0.46	0.33	0.31	0.67	0.47	0.32	0.31	0.71
−15	0.80	0.68	0.64	0.89	0.78	0.64	0.61	0.91	0.79	0.62	0.60	0.94
−20	0.96	0.91	0.89	0.98	0.95	0.88	0.87	0.98	0.95	0.86	0.85	0.99

Note: Figures are based on 10,000 Monte Carlo simulations, 5,000 bootstrap replications in each with nominal level 5%.

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

based on t_α^* and $t_{\alpha^*}^*$ have comparable properties, with the one using t_α^* being slightly more powerful.

The power of the bootstrap DF-GLS test increases as T grows, the power of the test based on $\hat{\alpha}^*$ is fairly constant for the values of T considered, and the power of the studentized bootstrap test shows a slight decrease.

We have also computed the values of the ordinary unit root tests based on $\hat{\alpha}$ and t_α using the critical values from Fuller (1996). The results are all within 0.01 of the corresponding figures of Table 1 corresponding to $\hat{\alpha}^*$- and $t_{\alpha^*}^*$-based tests, respectively. The asymptotic values corresponding to $T = \infty$ can be gleaned from Figure 2 in Elliott et al. and are close to those where $T = 250$.

5. PROOFS

Proof of Lemma 1. Consider first the case where $\epsilon_t^* = \epsilon_{t,T}^*, t = 1,\ldots,T$ are sampled with replacement from the centered differences $\hat{\epsilon}_t, t = 2,\ldots,T$. Given $(Y_1,\ldots,Y_T)$, $\epsilon_t^*, t = 1,\ldots,T$ are independent and identically distributed. The first property follows from $E^*(\epsilon_t^*) = [1/(T-1)]\sum_{t=2}^T \hat{\epsilon}_{t,T} = 0$.

Furthermore

$$\begin{aligned}
\mathrm{Var}^*(\epsilon_t^*) &= \sum_{t=2}^T \hat{\epsilon}_{t,T}^2/(T-1) \\
&= \sum_{t=2}^T [(\alpha_T - 1)(Y_{t-1} - \bar{Y}_{T-1}) + (\epsilon_t - \bar{\epsilon}_T)]^2/(T-1) \\
&= \sum_{t=2}^T [c(Y_{t-1} - \bar{Y}_{T-1})/T + (\epsilon_t - \bar{\epsilon}_T)]^2/(T-1) \\
&= c^2 \sum_{t=2}^T (Y_{t-1} - \bar{Y}_{T-1})^2/(T-1)T^2 \\
&\quad + 2c \sum_{t=2}^T (Y_{t-1} - \bar{Y}_{T-1})(\epsilon_t - \bar{\epsilon}_T)/(T-1)T \\
&\quad + \sum_{t=2}^T (\epsilon_t - \bar{\epsilon}_T)^2/(T-1). \qquad \textbf{(17)}
\end{aligned}$$

The process $\{(1/\sqrt{T})Y_{[Ts]}\}_{0\le s\le 1}$, where $[Ts]$ denotes the integer part of Ts, converges toward the Ornstein–Uhlenbeck process $\{\sigma J_c(s)\}_{0\le s\le 1}$ in distribution in $D[0,1]$. This follows from Lemma 1 in Phillips (1987). As explained in Phillips and Perron (1988), therefore $\sum_{t=2}^T (Y_{t-1} - \bar{Y}_{T-1})^2/T^2$ converges in distribution toward $\sigma^2 \int_0^1 J_c^\mu(s)^2\, ds$ where $J_c^\mu = J_c - \int_0^1 J_c(s)\, ds$. Hence, the first term on the right-hand side of (17) tends to zero in $P_{c,T}$ probability. The absolute value of the second term is bounded by $(2c/(T-1)^{1/2})[\sum_{t=2}^T (Y_{t-1} - \bar{Y}_{T-1})^2/T^2]^{1/2}[\sum_{t=2}^T (\epsilon_t - \bar{\epsilon}_T)^2/(T-1)]^{1/2}$. By the weak law of large numbers

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

$\sum_{t=2}^{T}(\epsilon_t - \bar{\epsilon}_T)^2/(T-1)$ tends to σ^2 in $P_{c,T}$ probability, which implies that $\mathrm{Var}^*(\epsilon_t^*) \to \sigma^2$ in $P_{c,T}$ probability.

We now consider the case where $\epsilon_t^* = t = 1,\ldots,T$ are sampled with replacement from the centered OLS residuals $\hat{\hat{\epsilon}}_t, t = 2,\ldots,T$. Because $[1/(T-1)] \times \sum_{t=2}^{T}\hat{\hat{\epsilon}}_{t,T} = 0$, $E^*(\epsilon_t^*) = 0$. Now $\mathrm{Var}^*(\epsilon_t^*)$ is given by a similar expression to (17) with c replaced by $T(\alpha_T - \hat{\alpha}_T)$. But $T(\alpha_T - \hat{\alpha}_T)$ converges in $P_{c,T}$ distribution, as shown by Phillips (1987), and hence is bounded in probability. Again $\mathrm{Var}^*(\epsilon_t^*) \to \sigma^2$ in $P_{c,T}$ probability. ■

Proof of Lemma 2. The result follows from Theorem V.19 in Pollard (1984). The details can be found in Swensen (2002) for the case where $c = 0$. The generalization to convergence in $P_{c,T}$ probability is straightforward using the convergence of the conditional moments proved in Lemma 1. ■

Proof of Lemma 3. The lemma follows from Lemma 2 by an application of the continuous mapping theorem along the lines of Basawa et al. (1991b), (for a similar result, see also Heimann and Kreiss, 1996, p. 97). Here are the details. Using that $2\sum_{t=2}^{T}(\sum_{s=1}^{t-1}\epsilon_s^*)\epsilon_t^* = (\sum_{t=1}^{T}\epsilon_t^*)^2 - \sum_{t=1}^{T}\epsilon_t^{*2}$, we may write $T(\hat{\alpha}_T^* - 1)$ as

$$\frac{\frac{1}{2}\left(S_T^*(1)^2 - \frac{1}{T}\sum_{t=1}^{T}\epsilon_t^{*2}\right)}{\int_0^1 S_T^*(s)^2\,ds} = \frac{\frac{1}{2}\left(S_T^*(1)^2 - \sigma^2\right)}{\int_0^1 S_T^*(s)^2\,ds} - \frac{\frac{1}{2}\left(\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^{*2} - \sigma^2\right)}{\int_0^1 S_T^*(s)^2\,ds}. \tag{18}$$

The function $g \to (g(1)^2 - \sigma^2)/2\int_0^1 g(s)^2\,ds$ is continuous in the uniform metric in $D[0,1]$ if $g \neq 0$, and the continuous mapping theorem can be applied to the first term on the right-hand side of (18). As for the second term, for all $\delta > 0$, $P^*[|(1/T)\sum_{t=2}^{T}\epsilon_t^{*2} - \sigma^2| > \delta]$ converges to zero in $P_{c,T}$ probability. To verify that, let $Z_T^* = \sum_{t=2}^{T}\epsilon_t^{*2}/T$. Then $E^*(\exp(isZ_T^*)) = \{E^*[\exp(i(s/T)\epsilon_1^{*2})]\}^{T-1}$ and $E^*[\exp((is/T)\epsilon_1^{*2})] = 1 + (is/T)E^*(\epsilon_1^{*2}) + o_P(1/T)$. From (17) it follows that $E^*(\epsilon_1^{*2}) = \mathrm{Var}^*(\epsilon_1^*) \to \sigma^2$ in $P_{c,T}$ probability and therefore $E^*(\exp(isZ_T^*)) \to \exp(is\sigma^2)$ in probability. Hence $P^*[|(1/T)\sum_{t=2}^{T}\epsilon_t^{*2} - \sigma^2| > \delta]$ tends to zero in $P_{c,T}$ probability. ■

Proof of Lemma 4. We only consider the detrended case. The proof is essentially an adaptation of the arguments of the proof of Lemma A4 and Theorem 1.c in Elliott et al. (1996).

Let N_T be the diagonal matrix with diagonal elements $\sqrt{T}$ and 1 and $Z_{\bar{c}} = (z_1, z_2)$. Define Z as $Z = Z_{\bar{c}}N_T$. Then $(\hat{\beta}_1^*, \hat{\beta}_2^*)' = (Z_{\bar{c}}'Z_{\bar{c}})^{-1}Z_{\bar{c}}'Y_{\bar{c}}^* = N_T[(1/T)Z'Z]^{-1}[(1/T)Z'Y_{\bar{c}}^*]$. But $[(1/T)Z'Z]^{-1} = D + A_T$ where D is a diagonal matrix with elements 1 and $(\int_0^1(1 - \bar{c}s)^2\,ds)^{-1}$ along the diagonal and $A_T = \{a_{ij,T}\}$ where $a_{ij,T} = o(1), i,j = 1,2$. Then $(\hat{\beta}_1^*, \hat{\beta}_2^*)' = DN_T^2[(1/T)Z_{\bar{c}}'Y_{\bar{c}}^*] + N_TA_TN_T^{-1}N_T^2[(1/T)Z_{\bar{c}}'Y_{\bar{c}}^*]$.

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

But

$$N_T^2\left(\frac{1}{T}\,Z_{\bar{c}}'Y_{\bar{c}}^*\right)=\begin{pmatrix} Y_1^*-\dfrac{\bar{c}}{T}\,(Y_T^*-Y_1^*)+\dfrac{\bar{c}^2}{T^2}\displaystyle\sum_{t=2}^{T}Y_{t-1}^* \\ \dfrac{1}{T}\left[(1-\bar{c})Y_T^*+\dfrac{\bar{c}^2}{T}\displaystyle\sum_{t=2}^{T}\dfrac{(t-1)}{T}\,Y_{t-1}^*+\dfrac{\bar{c}}{T}\,Y_T^*\right]\end{pmatrix},$$

so that because $\left(1/\sqrt{T}\right)Y_{[Ts]}^*=S_T^*(s)$, an application of Lemma 2 and the continuous mapping theorem implies that the conditional distribution of $(\hat{\beta}_1^*/\sqrt{T},\sqrt{T}\hat{\beta}_2^*)'=\sqrt{T}N_T^{-2}(\hat{\beta}_1^*,\hat{\beta}_2^*)$ converges in $P_{c,T}$ probability toward the distribution of $(0,\sigma\beta)'$ where $\beta=[(1-\bar{c})W(1)+\bar{c}^2\int_0^1 sW(s)\,ds]/[\int_0^1(1-\bar{c}s)^2\,ds]$. Here we use that

$$\sqrt{T}N_T^{-1}A_TN_T^{-1}=\begin{pmatrix}\dfrac{a_{11,T}}{\sqrt{T}} & a_{12,T}\\ a_{21,T} & \sqrt{T}a_{22,T}\end{pmatrix}.$$

Because $S_T^{*d}(s)=\left(1/\sqrt{T}\right)Y_{[sT]}^{*d}=S_T^*(s)-\left(1/\sqrt{T}\right)\hat{\beta}_1^*-\hat{\beta}_2^*\left([Ts]/\sqrt{T}\right)$, it follows, again by Lemma 2 and the continuous mapping theorem because $V_0(s,\bar{c})=W(s)-\beta s$, that $E^*[\,f(S_T^{*d})]\to E[\,f(\sigma V_0)]$ in $P_{c,T}$ probability for all continuous f defined on $D[0,1]$. Arguing as in Lemma 3 will finish the proof. ■

REFERENCES

Banerjee, A., J. Dolado, J.W. Galbraith, & D.F. Hendry (1993) *Co-integration, Error-Correction, and the Econometric Analysis of Non-stationary Data*. Oxford: Oxford University Press.

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

Basawa, I.V., A.K. Mallik, W.P. McCormic, J.H. Reeves, & R.L. Taylor (1991b) Bootstrap test of significance and sequential bootstrap estimation for unstable first order autoregressive processes. *Communications in Statistics A* 20, 1015–1026.

Berkowitz, J. & L. Kilian (2000) Recent developments in bootstrapping time series. *Econometric Reviews* 19, 1–48.

Chan, N.H. & C.Z. Wei (1987) Asymptotic inference for nearly nonstationary AR(1) processes. *Annals of Statistics* 15, 1050–1063.

Davidson, R. & J.G. MacKinnon (2000) Bootstrap tests: How many bootstraps? *Econometric Reviews* 19, 55–68.

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

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

Ferretti, N. & J. Romo (1996) Unit root bootstrap tests for AR(1) models. *Biometrika* 83, 849–860.

Fuller, W.A. (1996) *Introduction to Statistical Time Series*, 2nd ed. New York: Wiley.

Hansen, B.E. (1999) The grid bootstrap and the autoregressive model. *Review of Economics and Statistics* 81, 594–607.

Heimann, G. & J.-P. Kreiss (1996) Bootstrapping general first order autoregression. *Statistics and Probability Letters* 30, 87–98.

Li, H. & G.S. Maddala (1996) Bootstrapping time series models. *Econometric Reviews* 15, 115–158.

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press

Li, H. & G.S. Maddala (1997) Bootstrapping cointegrating regressions. *Journal of Econometrics* 80, 297–318.

Paparoditis, E. & D.N. Politis (2000) Unit Root Testing via the Continuous-Path Block Bootstrap. Manuscript, University of California, San Diego.

Paparoditis, E. & D.N. Politis (2001) The continuous-path block-bootstrap. In M. Puri (ed.), *Asymptotics in Statistics and Probability: Papers in honor of George Roussas*, pp. 305–320. Zeist, the Netherlands: VSP Publications.

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

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

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

Pollard, D. (1984) *Convergence of Stochastic Processes*. New York: Springer-Verlag.

Press, W.H., S.A. Teukolsky, W.T. Vetterling, & B.P. Flannery, (1992) *Numerical Recipes in Fortran*, 2nd ed. Cambridge: Cambridge University Press.

Swensen, A.R. (2002) Bootstrapping unit root tests for integrated processes. *Journal of Time Series Analysis*, in press.

https://doi.org/10.1017/S0266466603191025 Published online by Cambridge University Press