Working Paper 93-22
Statistics and Econometrics Series 17
October, 1993

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

BOOTSTRAP TESTS FOR UNIT ROOT AR(1) MODELS

Nélida Ferretti and Juan Romo*

Abstract

In this paper, we propose bootstrap tests for unit roots in first-order autoregressive models. We provide the bootstrap functional limit theory needed to prove the asymptotic validity of these tests both for independent and autoregressive errors; in this case, the usual corrections due to innovations dependence can be avoided. We also present a power empirical study comparing these tests with existing alternative methods.

Key Words

Autoregressive processes, bootstrapping least squares estimator, unit root, bootstrap invariance principle.

*Ferretti, Universidad de la Plata, CONICET and Universidad Carlos III de Madrid; Romo, Universidad Carlos III de Madrid. Research partially supported by Cátedra ARGENTARIA (Universidad Carlos III de Madrid) and DGICYT PB90-0266 (Spain).

BOOTSTRAP TESTS FOR UNIT ROOT AR(1) MODELS

NÉLIDA FERRETTI[1]
Universidad de la Plata, CONICET and Universidad Carlos III de Madrid

and

JUAN ROMO[2]
Universidad Carlos III de Madrid

Abstract

In this paper, we propose bootstrap tests for unit roots in first-order autoregressive models. We provide the bootstrap functional limit theory needed to prove the asymptotic validity of these tests both for independent and autoregressive errors; in this case, the usual corrections due to innovations dependence can be avoided. We also present a power empirical study comparing these tests with existing alternative methods.

Key words and phrases. Autoregressive processes, bootstrapping least squares estimator, unit root, bootstrap invariance principle.

[1]Research partially supported by Cátedra ARGENTARIA (Universidad Carlos III de Madrid, Spain)
[2]Research partially supported by DGICYT PB90-0266 (Spain)

1 INTRODUCTION

Our aim in this paper is to develop bootstrap tests for unit roots in autoregressive models and to establish its asymptotic validity. The bootstrap is a powerful and versatile resampling methodology to approximate the distribution of a statistic. It was introduced by Efron (1979) and since then, it has drawn a lot of attention from both the theoretical and practical points of view. It has been applied in many areas of statistics and it is of great potential use in econometrics.

A problem arising in many time series applications is the question of whether a series should be differenced; this is related to asking if the time series has a unit root.

Let $\{X_t\}$, $t = 1, 2, \ldots$ be a first-order autoregressive process defined by

$$X_t = \beta X_{t-1} + u_t, \qquad X_0 = 0, \tag{1.1}$$

where $\{u_t\}$ is a sequence of independent and identically distributed random variables with $E(u_t) = 0$, and $V(u_t) = \sigma_u^2 < \infty$. We are interested in testing the null hypothesis

$$H_0 : \beta = 1.$$

Let

$$\hat{\beta}_n = \left(\sum_{t=1}^{n} X_{t-1}^2\right)^{-1} \sum_{t=1}^{n} X_t X_{t-1}$$

be the least squares estimator of β, based on a sample of n observations $(X_1, \ldots, X_n)$.

The consistency of $\hat{\beta}_n$ was established by Rubin (1950) for all values of β. The limit distribution of $\hat{\beta}_n$ is, however, different for the three possible situations: stationary ($\beta < 1$), explosive ($\beta > 1$) and unstable ($\beta = 1$). The limit distribution of $\hat{\beta}_n$ is normal for the stationary case and nonnormal for the two nonstationary behaviors. For instance, in the unstable case $\beta = 1$ it is known (see White (1958), Rao (1978, 1980)) that

$$Z_n \to_w Z = \frac{1}{2}(W^2(1) - 1)\left(\int_0^1 W^2(t)dt\right)^{-1/2} \text{ as } n \to \infty, \tag{1.2}$$

where

$$Z_n = \frac{1}{\sigma_u}\left(\sum_{t=1}^{n} X_{t-1}^2\right)^{1/2} (\hat{\beta}_n - \beta)$$

and $\{W(t)\}$ is a standard Wiener process. So, it is interesting to study the bootstrap approximation for the distribution of $\hat{\beta}_n$ in the unstable model.

A way of deciding if a time series should be differenced is visual inspection of the autocorrelation function of the deviations from the fitted model. One of the best known statistics for testing the adequacy of a time series model is the Box-Pierce statistic (Box and Pierce (1970))

$$n \sum_{k=1}^{m} r_k^2,$$

where r_k is the usual lag k residual autocorrelation. If the innovations $\{u_t\}$ are normal, this statistic is asymptotically chi-squared distributed with $m-p$ degrees of freedom for large n, where p is the number of estimated parameters. If $\{X_t\}$ satisfies (1.1) then $p = 0$ under H_0 and the residuals are $X_t - X_{t-1}$. A modified test based on

$$Q_{n,m} = n(n+2)\sum_{k=1}^{m}(n-k)^{-1} r_k^2,$$

was recommended by Ljung and Box (1978). It was shown that it provides a substantially improved chi-square approximation.

If the innovations $\{u_t\}$ in the model are normally distributed, the likelihood ratio test for the hypothesis H_0 is a function of

$$Z_n' = \frac{1}{\sigma_n'}\left(\sum_{t=2}^{n} X_{t-1}^2\right)^{1/2} (\hat{\beta}_n - 1) \tag{1.3}$$

where

$$\sigma_n'^2 = \frac{\sum_{t=2}^{n}(X_t - \hat{\beta}_n X_{t-1})^2}{n-2}.$$

Dickey and Fuller (1979) derived representations for the limiting distributions of $n(\hat{\beta}_n - 1)$ and Z_n' under the assumption $\beta = 1$ and they compared by Monte Carlo techniques the power of the tests based on $n(\hat{\beta}_n - 1)$ and Z_n' with that of Box-Pierce (1970) test statistic. Tables for the percentiles of the distributions can be seen in Fuller (1976, pp. 371, 373) and a slightly more accurate one is given in Dickey (1976).

More related work has been done by Saïd (1982), Saïd and Dickey (1984, 1985) and Saïd (1991) who extended the Dickey-Fuller test procedure to general $ARIMA$ models. Evans and Savin (1981,1984) considered the $AR(1)$ model and calculated the percentiles of the limiting distribution of the least squares estimator of the parameter β in the case $\beta = 1$. Solo (1984) derived a Lagrange-multiplier test for unit roots in $ARIMA$ models. Hasza (1977) and Hasza and Fuller (1979) showed that the limiting distribution of the unit-roots test is invariant with respect to the error distribution for autoregressive models with two unit roots.

The above mentioned papers considered either normally distributed independent innovations or independent and identically distributed errors with mean zero and variance σ_u^2. Dickey and Fuller (1981) showed that the limit distribution of the likelihood-ratio test associated with $\hat{\beta}_n$ is invariant when $\{u_t\}$ in (1.1) is replaced by a stationary AR process.

Phillips (1987) and Phillips and Perron (1988) proposed an alternative procedure for testing the presence of a unit root in a general time series with weakly dependent and heterogeneously distributed innovations. This approach requires introducing adjustments to the Dickey-Fuller test statistics $n(\hat{\beta}_n - 1)$ and Z'_n. They also showed that their modified test statistics have asymptotic distribution tabulated by Dickey and Fuller.

Next, we present the contents of the paper. An introduction to bootstrap methodology is given in Section 2. We describe in Section 3 the bootstrap resampling scheme and we establish the asymptotic validity of the bootstrap test statistic for testing unit roots in $AR(1)$ models when the innovations are independent. Section 4 contains a Monte Carlo study of the power of the proposed test compared to the Ljung–Box test, the Dickey–Fuller test and the test based on the bootstrap analogue of the statistic $n(\hat{\beta}_n - 1)$. In Section 5 we extend the bootstrap test and we establish its asymptotic validity for testing H_0 when the innovations are a stationary autoregressive process. Finally, the proofs can be seen in an appendix.

2 THE BOOTSTRAP METHODOLOGY

The bootstrap, a resampling method introduced by Efron (1979), aims to reproduce from the sample the mechanism generating the data and to use it in the statistic of interest, replacing everywhere the unknown populational model. A formal description of the bootstrap in its simplest and original form is as follows. Let $\mathbf{Y} = (Y_1, \ldots, Y_n)$ be a random sample of size n from a population with distribution G and let $T(Y_1, \ldots, Y_n; G)$ be the specified random variable of interest, possibly depending of the unkonwn distribution G. Let G_n be the empirical distribution function of $Y_1, \ldots, Y_n$, i.e., the distribution giving mass $\frac{1}{n}$ to each of the observations $Y_1, \ldots, Y_n$. The bootstrap method aims to approximate the distribution of $T(Y_1, \ldots, Y_n; G)$ under G by that of $T(Y_1^*, \ldots, Y_n^*; G_n)$ under G_n where $Y_1^*, \ldots, Y_n^*$ denotes a random sample of size n from G_n; although this later distribution cannot usually be explicitly calculated, it is always possible to approximate it very easily by Monte Carlo simulation since G_n is available from the sample. So, the bootstrap technique follows the next steps:

(i) simulate artificially a random sample $\mathbf{Y}^* = (Y_1^*, \ldots, Y_m^*)$ —m not necessarily equal to n— from the empirical probability G_n.

(ii) evaluate T at the bootstrap sample to obtain the bootstrap version of the statistic $T^* = T(Y_1^*, \ldots, Y_m^*; G_n)$.

(iii) replicate (i) and (ii) a large number B of times, in order to get B values of T^*, $T_i^* = T(\mathbf{Y}^{*\mathbf{i}}; G_n), i = 1, \ldots, B$.

Finally, a histogram (or, in general, any other estimate of the distribution of T^*) is obtained from T_i^*, $i = 1, \ldots, B$; this is an approximation to the distribution of T^* which in turn is the bootstrap estimation of the unknown distribution of T.

Recently, the bootstrap has been adapted and studied for regression and autoregression models. More information on bootstrap for regression models can be seen, e.g., in González Manteiga, Prada Sánchez and Romo (1992).

The study of the bootstrap for time series and dynamic regression models was started by Freedman (1984). Bose (1988) has shown that, under some regularity conditions, the bootstrap approximation to the distribution of the least-squares estimator in stationary autoregressive models is of order $o(n^{-\frac{1}{2}})$ a.s., improving on the normal approximation (which is $O(n^{-\frac{1}{2}})$); Thombs and Schucany (1990) give bootstrap prediction intervals in this case. The validity of the bootstrap for the least squares estimator in explosive $AR(1)$ models has been established by Basawa *et al.* (1989) and Stute and Gründer (1990) have obtained bootstrap approximations to prediction intervals in this case. Basawa *et al.* (1991-a) prove that the bootstrapped least squares estimator has a random limit distribution for the unstable first-order autoregressive model ($\beta = 1$); however, Basawa *et al.* (1991-b) present a modified sequential bootstrap which works in this situation. The bootstrap for state-space models has been considered by Stoffer and Wall (1991). Under no model assumptions, Künsch (1989) investigates blockwise bootstrap for stationary observations.

Finally, we introduce some terminology that we will need in our work. As we have seen, a general goal of bootstrap resampling is to approximate the distribution $P_G\{T(\mathbf{Y}; G) \leq x\}$ of the statistic $T(\mathbf{Y}; G)$ by using the distribution

$$P_{G_n}\{T(\mathbf{Y}^*; G_n) \leq x\} = P^*\{T(\mathbf{Y}^*; G_n) \leq x\}$$

of $T(\mathbf{Y}^*; G_n)$. This can be expressed in several ways. If $T(\mathbf{Y}; G)$ converges weakly to a distribution $S(G)$, it suffices to show that $T(\mathbf{Y}^*; G_n)$ converges weakly to $S(G)$ for almost all samples $Y_1, \ldots Y_n \ldots$ ($T(\mathbf{Y}^*; G_n) \rightarrow_w S(G)$ *a.s.*) or to establish that the distance between the law of $T(\mathbf{Y}^*; G_n)$ and the law of $S(G)$ tends to zero in probability for any distance metrizing weak convergence ($T(\mathbf{Y}^*; G_n) \rightarrow_w S(G)$ *in probability*).

3 BOOTSTRAP UNIT ROOT TESTS (INDEPENDENT INNOVATIONS)

For the model defined in (1.1) and to test $H_0 : \beta = 1$, our bootstrap resampling scheme can be described in the following way. Let $\epsilon_t = X_t - \hat{\beta}_n X_{t-1}$, $t = 1, \ldots, n$, and define $\hat{\epsilon}_t = \epsilon_t - n^{-1} \sum_{j=1}^{n} \epsilon_j$, the centered residuals. Denote by $\hat{F}_n$ the empirical distribution function based on $\{\hat{\epsilon}_t : t = 1, \ldots, n\}$ and take a random sample $\{\epsilon^*_{n,t} : t = 1, \ldots, n\}$ from $\hat{F}_n$. So, the random variables $\{\epsilon^*_{n,t} : t = 1, \ldots, n\}$ are independent and identically distributed with distribution function $\hat{F}_n$, conditionally on $X_1, \ldots, X_n$. Then, the bootstrap sample $\{X^*_{n,t} : t = 1, \ldots, n\}$ is recursively obtained from the model for $\beta = 1$,

$$X^*_{n,t} = X^*_{n,t-1} + \epsilon^*_{n,t}, \qquad t = 1, \ldots, n \tag{3.4}$$

with $X^*_{n,0} = 0$. The bootstrap least squares estimate is then given by

$$\hat{\beta}^*_n = \left(\sum_{t=1}^{n} X^{*2}_{n,t-1} \right)^{-1} \sum_{t=1}^{n} X^*_{n,t} X^*_{n,t-1}.$$

In the stationary case, $|\beta| < 1$, Bose (1988) showed the asymptotic validity of the bootstrap estimators corresponding to $\hat{\beta}_n$ and in the explosive case, $|\beta| > 1$, this has been established by Basawa, Mallik, McCormick and Taylor (1989). If $|\beta| = 1$, the unstable case, $\hat{\beta}_n$ has the limit distribution given by (1.2), so it is interesting to study the bootstrap approximation in this situation. Basawa, Mallik, McCormick, Reeves and Taylor (1991-a) took i.i.d. $\{u_t^*\}$ with distribution $N(0,1)$ and they obtained $\{\tilde{X}_t^*\}$ from

$$\tilde{X}_t^* = \tilde{\beta}_n \tilde{X}_{t-1}^* + u_t^*, \qquad \tilde{X}_0^* = 0,$$

where $\tilde{\beta}_n$ is the least squares estimate for the $AR(1)$ model; they show that for $\tilde{\beta}_n^* = \left(\sum_{t=1}^n \tilde{X}_{t-1}^{*2}\right)^{-1} \sum_{t=1}^n \tilde{X}_t^* \tilde{X}_{t-1}^*$, the sequence

$$\tilde{Z}_n^* = \left(\sum_{t=1}^n \tilde{X}_{t-1}^{*2}\right)^{1/2} (\tilde{\beta}_n^* - \tilde{\beta}_n)$$

converges to a random distribution not approaching the asymptotic correct one. Thus, this is an example of incorrect behavior of the standard bootstrap. Basawa, Mallik, McCormick, Reeves and Taylor (1991-b) present a modified sequential bootstrap which correctly approaches the limit distribution of the least squares estimator of β; also, they established the asymptotic validity of a bootstrap test statistic for unit roots based on residuals of the form $X_t - X_{t-1}$.

In related work, Rayner (1990) performed a Monte Carlo study to examine the small-sample behavior of the bootstrap and the Student-t approximations to the true distribution of the classical test statistic used for testing hypotheses on the slope parameter in the stationary first-order autoregressive model. Bertail (1991) investigated using Monte Carlo techniques the bootstrap test for linear models and unstable autoregressive models.

In this section, we prove that this resampling algorithm is asymptotically correct under H_0, in the sense that it converges weakly to the limit distribution given by (1.2) for almost all sample $(X_1, \ldots, X_n)$.

3.1 A BOOTSTRAP INVARIANCE PRINCIPLE

The study of the asymptotic behavior of the bootstrap least squares estimate relies on a ***bootstrap invariance principle***, a functional central limit theorem for a stochastic process built from the sequence of partial sums corresponding to the bootstrap resamples.

Consider the sequence of partial sums $S^*_{n,0} = 0$, $S^*_{n,k} = \sum_{j=1}^{k} \epsilon^*_{n,j}$, $k = 1, \ldots, n$, $n \in \mathcal{N}$. A sequence of continuous-time processes $\{Y^*_n(s) : s \in [0,1]\}_{n=1}^{\infty}$ can be obtained from

$$\left\{S^*_{n,k} : k = 1, \ldots, n\right\}_{n=1}^{\infty}$$

by linear interpolation, i.e.,

$$Y^*_n(s) = \frac{1}{\hat{\sigma}_n\sqrt{n}} S^*_{n,[ns]} + (ns - [ns]) \frac{1}{\hat{\sigma}_n\sqrt{n}} \epsilon^*_{n,[ns]+1}, \; s \in [0,1], \; n \in \mathcal{N} \quad (3.5)$$

where $\hat{\sigma}^2_n = V^*(\epsilon^*_{n,t}) = \frac{\sum_{j=1}^{n} \hat{\epsilon}^2_j}{n}$ and $[m]$ denotes the greatest integer less than or equal to m.

Hereafter, P^*, E^*, V^* will denote, respectively, the bootstrap probability, expectation and variance conditionally on the sample $X_1, \ldots, X_n$.

The sample paths of the process $Y^*_n(s)$ are in the space $C[0,1]$ of real continuous functions on $[0,1]$. We endow it with the supremum norm

$$\| f - g \|_\infty = \sup_{x \in [0,1]} | f(x) - g(x) |, \qquad f, g \in C[0,1].$$

To obtain weak convergence on this space is enough to show weak convergence of finite dimensional distribution and tightness (see, e.g., Billingsley (1968) or Pollard (1984) for a detailed treatment).

The first lemma establishes the weak convergence of the finite-dimensional distributions for almost all samples $(X_1, \ldots, X_n)$.

Lemma 3.1. Conditionally on $(X_1, \ldots, X_n)$ and for almost all sample paths $(X_1, X_2, \ldots)$,

$$(Y^*_1(s_1), \ldots, Y^*_n(s_d)) \rightarrow_w (W(s_1), \ldots, W(s_d)) \quad (3.6)$$

for all $(s_1, \ldots, s_d) \in [0,1]^d$.

The tightness of our sequence of stochastic processes follows from Lemmas 3.2 and 3.3.

Lemma 3.2. For any $\eta > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} \frac{1}{\delta} P^* \left\{ \max_{1 \le j \le [n\delta]+1} |S^*_{n,j}| > \eta \hat{\sigma}_n \sqrt{n} \right\} = 0$$

conditionally on $(X_1, \ldots, X_n)$ and for almost all sample paths $(X_1, X_2, \ldots)$.

Lemma 3.3. For any $\eta > 0$ and $T > 0$,

$$\lim_{\delta \downarrow 0} \limsup_{n\to\infty} P^* \left\{ \max_{\substack{1 \le j \le [n\delta]+1 \\ 0 \le k \le [nT]+1}} |S^*_{n,j+k} - S^*_{n,k}| > \eta \hat{\sigma}_n \sqrt{n} \right\} = 0$$

conditionally on $(X_1, \ldots, X_n)$ for almost all sample paths $(X_1, X_2, \ldots)$.

Now, we establish the bootstrap invariance principle.

Proposition 3.1. Let $\{\epsilon_n\}_{n=1}^{\infty}$ be a sequence of residuals as defined above. Let $\hat{F}_n$ be the empirical distribution associated to $\hat{\epsilon}_i = \epsilon_i - \frac{\sum_{j=1}^{n} \epsilon_j}{n}$, $i = 1, \ldots, n$ and let $\epsilon^*_{n,i}$, $i = 1, \ldots, n$ be independent random variables with distribution $\hat{F}_n$. Define $\{Y^*_n(t) : t \in [0,1]\}_{n=1}^{\infty}$ by (3.5). Then

$$Y^*_n \longrightarrow_w W \qquad a.s.$$

in $C[0,1]$, where W is the standard one-dimensional Brownian motion on $[0,1]$.

3.2 Asymptotic behavior of the bootstrap statistic

Let

$$Z^*_n = \frac{1}{\hat{\sigma}_n} \left(\sum_{t=1}^{n} X^{*2}_{n,t-1} \right)^{1/2} (\hat{\beta}^*_n - 1) \tag{3.7}$$

be the bootstrap version of Z_n under $\beta = 1$.

Now, in Theorem 3.1 we derive the limit distribution of Z_n^*. To prove this theorem we will need the following lemma.

Lemma 3.4. Let

$$r_n^* = \frac{1}{n}\sum_{i=1}^{n} Y_n^{*2}\left(\frac{i}{n}\right) - \int_0^1 Y_n^{*2}(s)ds.$$

Then, conditionally on $(X_1, \ldots, X_n)$ and for almost all sample paths $(X_1, X_2, \ldots)$

$$r_n^* \longrightarrow_{P^*} 0,$$

as $n \to \infty$.

Our goal is to show that $Z_n^* \to_w Z$ almost surely and so this bootstrap resampling approaches properly the correct limiting distribution. Thus, our main result in this section is the following.

Theorem 3.1. For Z_n^* defined in (3.7), under the model (1.1) with $\beta = 1$, we have that

$$Z_n^* \to_w Z$$

conditionally on $(X_1, \ldots, X_n)$ for almost all sample paths $(X_1, X_2, \ldots)$ where Z is defined in (1.2).

4 EMPIRICAL POWER STUDY

We have performed a Monte Carlo study to investigate the power of the test statistics reported in this article when the model is (1.1).

Two thousand samples of size n = 25, 50, 100, 250 were generated for two different innovation distributions (Normal and Student with 3 degrees of freedom) and 5000 bootstrap samples .

Table 1 and Table 2 show empirical powers of two-sided 5 % level tests for the test statistics: $Q_{n,m}(m = 5)$, Z_n', B_n^* (the bootstrap analogue of the statistic $n(\hat{\beta}_n - 1)$) and Z_n^*. Values of β were 0.80, 0.90, 0.95, 0.99, 1.00, 1.02 and 1.05.

Table 3 and Table 4 give empirical powers of one-sided 5 % level tests ($H_0 : \beta = 1$ versus $H_1 : \beta < 1$) for the test statistics: Z_n', B_n^* and Z_n^*. Values of β were 0.80, 0.90, 0.95, 0.99 and 1.00. The results for the one-sided 5 % tests of $H_0 : \beta = 1$ against the alternative $H_1 : \beta > 1$ are not reported here because they are qualitatively similar to those of the alternative $H_1 : \beta < 1$.

Several routines from IMSL Library were used: GGUBS (basic uniform (0,1) pseudo-random number generator), GGNML (normal random deviate generator), GGCHS (chi-squared random deviate generator) and GGNQF (normal random deviate generator-function form of GGNML). The computer programmes where written in FORTRAN and performed in a DECstation 5000/2000 under ULTRIX-32 at the Universidad Carlos III, Madrid.

We observe that the power of the Ljung–Box $Q_{n,m}$ test was significantly lower than the power of the other tests considered in almost all cases. The performances of Z_n^* test, the Dickey–Fuller Z_n' test and the B_n^* test were similar. The bootstrap Z_n^* test power is never lower than the power of their competitors.

TABLE 1
Empirical power of Two-Sided Size 0.05 Tests
when $\beta = 1$ and F is Normal

		β						
n	Test	0.80	0.90	0.95	0.99	1.00	1.02	1.05
25	$Q_{n,m}$	0.08	0.06	0.06	0.06	0.05	0.08	0.17
	Z'_n	0.21	0.08	0.06	0.05	0.05	0.07	0.20
	B^*_n	0.20	0.07	0.06	0.05	0.05	0.06	0.18
	Z^*_n	0.28	0.10	0.07	0.06	0.06	0.12	0.31
50	$Q_{n,m}$	0.11	0.07	0.06	0.06	0.05	0.11	0.53
	Z'_n	0.58	0.20	0.07	0.05	0.05	0.15	0.71
	B^*_n	0.57	0.20	0.07	0.05	0.05	0.14	0.70
	Z^*_n	0.64	0.23	0.09	0.05	0.05	0.24	0.70
100	$Q_{n,m}$	0.15	0.08	0.06	0.06	0.06	0.32	0.95
	Z'_n	0.99	0.57	0.18	0.04	0.05	0.51	0.97
	B^*_n	0.99	0.57	0.18	0.04	0.05	0.51	0.97
	Z^*_n	0.99	0.61	0.19	0.04	0.05	0.56	0.96
250	$Q_{n,m}$	0.46	0.15	0.07	0.05	0.05	0.95	1.00
	Z'_n	1.00	1.00	0.75	0.06	0.05	0.98	1.00
	B^*_n	1.00	1.00	0.75	0.06	0.05	0.98	1.00
	Z^*_n	1.00	1.00	0.76	0.06	0.05	0.97	1.00

TABLE 2
Empirical power of Two-Sided Size 0.05 Tests when $\beta = 1$ and F is Student with three degrees of freedom

		β						
n	Test	0.80	0.90	0.95	0.99	1.00	1.02	1.05
25	$Q_{n,m}$	0.07	0.04	0.04	0.04	0.05	0.05	0.16
	Z'_n	0.19	0.08	0.05	0.05	0.06	0.08	0.19
	B^*_n	0.19	0.08	0.05	0.05	0.05	0.07	0.16
	Z^*_n	0.25	0.10	0.05	0.04	0.05	0.11	0.34
50	$Q_{n,m}$	0.07	0.06	0.04	0.05	0.04	0.10	0.52
	Z'_n	0.57	0.17	0.07	0.04	0.06	0.13	0.70
	B^*_n	0.58	0.17	0.07	0.05	0.06	0.11	0.69
	Z^*_n	0.64	0.20	0.09	0.04	0.05	0.23	0.70
100	$Q_{n,m}$	0.15	0.06	0.05	0.04	0.05	0.34	0.94
	Z'_n	0.99	0.57	0.17	0.05	0.05	0.54	0.98
	B^*_n	0.99	0.58	0.17	0.05	0.05	0.53	0.98
	Z^*_n	0.99	0.60	0.18	0.05	0.05	0.58	0.97
250	$Q_{n,m}$	0.42	0.12	0.06	0.05	0.05	0.94	1.00
	Z'_n	1.00	1.00	0.74	0.07	0.05	0.97	1.00
	B^*_n	1.00	1.00	0.75	0.08	0.05	0.97	1.00
	Z^*_n	1.00	1.00	0.74	0.08	0.05	0.97	1.00

TABLE 3
Empirical power of One-Sided Size 0.05 Tests
when $\beta = 1$ and F is Normal

		β				
n	Test	0.80	0.90	0.95	0.99	1.00
25	Z_n'	0.33	0.17	0.09	0.06	0.05
	B_n^*	0.32	0.16	0.08	0.05	0.05
	Z_n^*	0.40	0.21	0.12	0.07	0.06
50	Z_n'	0.78	0.33	0.15	0.05	0.05
	B_n^*	0.78	0.33	0.15	0.05	0.05
	Z_n^*	0.81	0.37	0.17	0.06	0.06
100	Z_n'	1.00	0.78	0.32	0.08	0.05
	B_n^*	1.00	0.78	0.32	0.08	0.05
	Z_n^*	1.00	0.80	0.35	0.09	0.05
250	Z_n'	1.00	1.00	0.89	0.15	0.04
	B_n^*	1.00	1.00	0.89	0.15	0.05
	Z_n^*	1.00	1.00	0.89	0.16	0.05

TABLE 4
Empirical power of Two-Sided Size 0.05 Tests
when $\beta = 1$ and F is Student with three degrees of freedom

		β				
n	Test	0.80	0.90	0.95	0.99	1.00
25	Z_n'	0.34	0.13	0.07	0.05	0.04
	B_n^*	0.35	0.13	0.08	0.05	0.04
	Z_n^*	0.42	0.16	0.10	0.07	0.05
50	Z_n'	0.78	0.32	0.14	0.06	0.05
	B_n^*	0.79	0.33	0.14	0.06	0.05
	Z_n^*	0.83	0.36	0.16	0.07	0.06
100	Z_n'	1.00	0.76	0.31	0.07	0.05
	B_n^*	1.00	0.79	0.32	0.07	0.05
	Z_n^*	1.00	0.79	0.33	0.07	0.05
250	Z_n'	1.00	1.00	0.90	0.13	0.05
	B_n^*	1.00	1.00	0.91	0.13	0.05
	Z_n^*	1.00	1.00	0.91	0.13	0.05

5 BOOTSTRAP TESTS FOR A UNIT ROOT (AUTOREGRESSIVE INNOVATIONS)

In this section, we extend Theorem 3.1 on the asymptotic behavior of the bootstrap least squares estimator to models with autoregressive errors. For the process $\{X_t\}$, $t = 1, 2, \ldots$ given by

$$X_t = \beta X_{t-1} + u_t, \qquad X_0 = 0, \tag{5.8}$$

consider now that the innovations $\{u_t\}, t = 1, 2, \ldots$ are a stationary autoregressive sequence defined by

$$u_t = \rho u_{t-1} + v_t, \qquad u_0 = 0, \qquad |\rho| < 1, \tag{5.9}$$

where $\{v_t\}$ are independent and identically distributed random variables with $E(v_t) = 0$ and $0 < V(v_t) = \sigma_v^2 < \infty$. We will make also the following assumptions on the $\{v_t\}$:

A1. $\int | \phi(r) | \, dr < \infty$ where ϕ is the characteristic function of each v_t.

A2. $\sup_t E(| v_t |^{\gamma+\eta}) < \infty$, for some $\gamma > 2$ and $\eta > 0$.

A3. $M_0 = \sup\limits_{m,s,k \geq 1} \sup\limits_{\alpha,\tau,\nu} \max\limits_t \left| \frac{\partial}{\partial \nu_t} P(\tilde{U} + \nu \in \bigcup\limits_{j=1}^{s} D_j) \right| < \infty,$
where $D_j = \mathop{\times}\limits_{t=k}^{k+m-1} (\alpha_{jt}, \tau_{jt}), \nu = (\nu_k, \ldots, \nu_{k+m-1}), \tilde{U} = (u_k, \ldots, u_{k+m-1})$.

By Minkowski's inequality, assumption A2 implies that

$$\sup_t E(| u_t |^{\gamma+\eta}) < \infty.$$

So, under A1-A3, by Corollary 3 of Withers (1981), $\{u_t\}$ is strong mixing with mixing coefficients α_m that satisfy

$$\sum_{m=1}^{\infty} \alpha_m^{1-2/\gamma} < \infty.$$

In particular, Gaussian AR(1) processes satisfy A1-A3.

Our goal is again to test the hypothesis $H_0 : \beta = 1$ (under H_0, the model defined by (5.8) and (5.9) is an ARIMA (1,1,0) process). Let $\hat{\beta}_n$ be the least squares estimator of β based on a sample of n observations $(X_1, \ldots, X_n)$. For the statistic

$$T_n = \frac{1}{s_n}\left(\sum_{t=1}^{n} X_{t-1}^2\right)^{\frac{1}{2}} (\hat{\beta}_n - 1),$$

where

$$s_n^2 = \frac{1}{n}\sum_{t=1}^{n}(X_t - \hat{\beta}_n X_{t-1})^2,$$

Phillips (1987) showed that under H_0,

$$\hat{\beta}_n \longrightarrow_P 1 \text{ as } n \to \infty \tag{5.10}$$

and

$$T_n \to_w T = \frac{\sigma}{2\sigma_u}(W^2(1) - \frac{\sigma_u^2}{\sigma^2})\left(\int_0^1 W^2(t)dt\right)^{-1/2} \text{ as } n \to \infty, \tag{5.11}$$

with

$$\sigma_u^2 = \lim_{n\to\infty} \frac{1}{n}\sum_{t=1}^{n} E(u_t^2) \tag{5.12}$$

and

$$\sigma^2 = \lim_{n\to\infty} E(\frac{1}{n}(\sum_{t=1}^{n} u_t)^2). \tag{5.13}$$

Note that for the sequence $\{u_t\}$ given by (5.9), $\sigma_u^2 = \frac{\sigma_v^2}{1-\rho^2}$. On the other hand, since $0 < \sigma_v^2 < \infty$ we have that the limit given in (5.13) exists and $\sigma^2 = \frac{\sigma_v^2}{(1-\rho)^2} > 0$. Finally, observe that for independent innovations, $\sigma^2 = \sigma_u^2$ and this common value is the variance of the innovations in model (1.1); so, in this case, the limit distributions in (1.2) and (5.11) coincide.

For the bootstrap test in this situation, we propose the following resampling strategy. Let $\tilde{u}_t = X_t - \hat{\beta}_n X_{t-1}, t = 1, \ldots, n$ with $\tilde{u}_0 = 0$ and let $\hat{\rho}_n$ be the least squares estimate of ρ obtained from $\tilde{u}_1, \ldots, \tilde{u}_n$. Consider $\tilde{v}_t = \tilde{u}_t - \hat{\rho}_n \tilde{u}_{t-1}$ and define $\hat{v}_t = \tilde{v}_t - n^{-1}\sum_{j=1}^{n} \tilde{v}_j$, the centered residuals.

Let $\hat{F}_n$ be the empirical distribution function based on $\{\hat{v}_t : t = 1, \ldots, n\}$ and take a random sample $\{v^*_{n,t} : t = 1, \ldots, n\}$ from $\hat{F}_n$. Then, construct $\{u^*_{n,t} : t = 1, \ldots, n\}$ from

$$u^*_{n,t} = \hat{\rho}_n u^*_{n,t-1} + v^*_{n,t} \tag{5.14}$$

and also the sequence of bootstrap pseudo-data under H_0,

$$X^*_{n,t} = X^*_{n,t-1} + u^*_{n,t}, \qquad t = 1, \ldots, n \tag{5.15}$$

with $X^*_{n,0} = 0$. The bootstrap least squares estimate of β is

$$\hat{\beta}^*_n = \left(\sum_{t=1}^{n} X^{*2}_{n,t-1} \right)^{-1} \sum_{t=1}^{n} X^*_{n,t} X^*_{n,t-1}.$$

Our first step is to establish a bootstrap invariance principle; for the proof, we will need the following lemma.

Lemma 5.1. Under H_0 and assumptions A1, A2 and A3, we have

(i) $\hat{\rho}_n \longrightarrow_P \rho$.

(ii) $V^*(v^*_{n,t}) = \frac{1}{n} \sum_{j=1}^{n} \hat{v}_j^2 \longrightarrow_{P^*} \sigma_v^2$,

(iii) $V^*(u^*_{n,t}) \longrightarrow_{P^*} \sigma_u^2$,

as $n \to \infty$, conditionally on $(X_1, \ldots, X_n)$ for almost all sample paths $(X_1, X_2, \ldots)$.

Define now the sequence of partial sums $R^*_{n,0} = 0$, $R^*_{n,k} = \sum_{j=1}^{k} u^*_{n,j}$, $k = 1, \ldots, n$, $n \in \mathcal{N}$. From

$$\{R^*_{n,k} : k = 1, \ldots, n\}_{n=1}^{\infty},$$

we obtain the corresponding sequence of continuous-time processes

$$\{U^*_n(s) : s \in [0,1]\}_{n=1}^{\infty},$$

given by

$$U^*_n(s) = \frac{1}{\sigma\sqrt{n}} R^*_{n,[ns]} + (ns - [ns]) \frac{1}{\sigma\sqrt{n}} u^*_{n,[ns]+1}, \; s \in [0,1], \; n \in \mathcal{N}, \tag{5.16}$$

where now

$$\sigma^2 = \frac{\sigma_v^2}{(1-\rho)^2}. \tag{5.17}$$

Herrndorf (1984) obtained functional central limit theorems for sequences of partial sums of dependent random variables. Applying his Corollary 1, we can prove the bootstrap invariance principle in probability that we need for our theorem.

Proposition 5.1. Let $\{u_{n,t}^*\}$ and $\{U_n^*(s) : s \in [0,1]\}_{n=1}^{\infty}$ be as defined in (5.14) and (5.16), respectively. Under H_0 and assumptions A1, A2 and A3,

$$U_n^* \longrightarrow_w W \qquad \textit{in probability}$$

in $C[0,1]$, where W is the standard one-dimensional Brownian motion on $[0,1]$.

Consider now the bootstrap version of the statistic T_n,

$$T_n^* = \frac{1}{s_n}\left(\sum_{t=1}^{n} X_{n,t-1}^{*2}\right)^{\frac{1}{2}} (\hat{\beta}_n^* - 1). \tag{5.18}$$

The main result in this section is contained in Theorem 5.1 which gives the limiting distribution of the statistic T_n^*.

Theorem 5.1. Let T_n^* be as defined in (5.18), under the model (5.8) with $\beta = 1$ and $\{u_t\}$ following (5.9) and assumptions A1, A2 and A3. Then

$$T_n^* \longrightarrow_w T \qquad \textit{in probability}$$

where T is defined in (5.11).

Remarks. (i) The asymptotic distribution in (5.11) depends on unknown parameters σ^2 and σ_u^2. To overcome this difficulty, Phillips (1987) introduced modified statistics with the asymptotic distribution given in (1.2) which is independent of σ^2 and σ_u^2. However, our bootstrap technique does not require any modified statistic because it directly approaches the unknown limiting distribution.

(ii) When $\rho = \rho_0$ is known in model (5.9), the bootstrap convergence in probability of Theorem 5.1, could be strength to almost sure convergence.
(iii) It is straightforward to extend this bootstrap results to model (5.8) with $AR(p)$ innovations, under conditions analogous to assumptions A1-A3.

6 CONCLUSION

We have established the asymptotic correctness of bootstrap tests for unit roots in first-order autoregressive models, both for independent and autoregressive errors. The Monte Carlo comparisons lead us to conclude that, in the first-order autoregressive process with independent and identically distributed innovations, the power of the bootstrap analogue of the Z_n test was similar and never lower to that of the Z_n' test and the bootstrap analogue of $n(\hat{\beta}_n - 1)$ test. Moreover, in the first-order autoregressive process with stationary autoregressive innovations, we have shown that the modifications to the Dickey-Fuller statistic Z_n introduced by Phillips (1987) are not needed when using the bootstrap test.

Acknowledgments. The authors wish to thank Ana Justel for some comments and helping with the simulation work.

7 MATHEMATICAL APPENDIX

Proof of Lemma 3.1. It is enough to show that, for all $r, s \in [0,1]$,

$$(Y_n^*(r), Y_n^*(s)) \rightarrow_w (W(r), W(s)) \qquad a.s.$$

Now, conditionally on $(X_1, \ldots, X_n)$, since

$$|Y_n^*(s) - \frac{1}{\hat{\sigma}_n\sqrt{n}} \sum_{j=1}^{[ns]} \epsilon_{n,j}^*| \leq \frac{1}{\hat{\sigma}_n\sqrt{n}} |\epsilon_{n,[ns]+1}^*|,$$

we obtain by the Čebišev inequality that

$$P^* \left\{ |Y_n^*(s) - \frac{1}{\hat{\sigma}_n\sqrt{n}} \sum_{j=1}^{[ns]} \epsilon_{n,j}^*| > \delta \right\} \leq \frac{1}{\delta^2 n}. \qquad (7.19)$$

Therefore,

$$\|(Y_n^*(r), Y_n^*(s)) - \frac{1}{\hat{\sigma}_n\sqrt{n}}(\sum_{j=1}^{[nr]} \epsilon_{n,j}^*, \sum_{j=1}^{[ns]} \epsilon_{n,j}^*)\| \longrightarrow_{P^*} 0 \qquad a.s.$$

and it suffices to prove that

$$\frac{1}{\hat{\sigma}_n\sqrt{n}}(\sum_{j=1}^{[nr]} \epsilon_{n,j}^*, \sum_{j=1}^{[ns]} \epsilon_{n,j}^*) \longrightarrow_w (W(r), W(s)) \qquad a.s.$$

This is equivalent to show that, conditionally on $(X_1, \ldots, X_n)$,

$$\frac{1}{\hat{\sigma}_n\sqrt{n}}(\sum_{j=1}^{[nr]} \epsilon_{n,j}^*, \sum_{j=[nr]+1}^{[ns]} \epsilon_{n,j}^*) \longrightarrow_w (W(r), W(s) - W(r)) \qquad a.s.; \tag{7.20}$$

but, since the components on the left hand side are conditionally independent random variables with zero mean and variance one, (7.20) follows by the the bootstrap central limit theorem for triangular arrays obtained by using Lindeberg condition. □

Proof of Lemma 3.2. By the bootstrap central limit theorem for triangular arrays , we have that $(1/\hat{\sigma}_n\sqrt{[n\delta]+1})S_{n,[n\delta]+1}^*$ converges weakly almost surely to a standard normal random variable V. Fix $\lambda > 0$ and let $\{\varphi_k\}_{k=1}^{\infty}$ be a sequence of bounded, continuous functions on $\Re$ with $\varphi_k \downarrow \mathbf{1}_{(-\infty,\lambda]\cup[\lambda,\infty)}$. We have for each k,

$$\limsup_{n\to\infty} P^*\left\{|S_{n,[n,\delta]+1}^*| \geq \lambda\hat{\sigma}_n\sqrt{n\delta}\right\}$$
$$\leq \lim_{n\to\infty} E^*\left(\varphi_k\left(\frac{1}{\hat{\sigma}_n\sqrt{n\delta}}S_{n,[n\delta]+1}^*\right)\right) = E^*(\varphi_k(V)).$$

Then, if $k \to \infty$ we obtain

$$\limsup_{n\to\infty} P^*\left\{|S_{n,[n\delta]+1}^*| \geq \lambda\hat{\sigma}_n\sqrt{n\delta}\right\} \leq P(|V| > \lambda) \leq \frac{1}{\lambda^3}E(|V|^3). \tag{7.21}$$

We now define $\tau_n^* = \min\{j \geq 1 : |S_{n,j}^*| > \eta\hat{\sigma}_n\sqrt{n}\}$. If $0 < \delta < \frac{\eta^2}{2}$, we have

$$P^*\left\{\max_{i\le j\le [n\delta]+1}|S^*_{n,j}|>\eta\hat\sigma_n\sqrt{n}\right\}\le P^*\left\{|S^*_{n,[n\delta]+1}|\ge\hat\sigma_n\sqrt{n}(\eta-\sqrt{2\delta})\right\}$$

$$+\sum_{j=1}^{[n\delta]}P^*\left\{|S^*_{n,[n\delta]+1}|<\hat\sigma_n\sqrt{n}(\eta-\sqrt{2\delta})|\tau^*_n=j\right\}P^*\{\tau^*_n=j\}. \qquad (7.22)$$

But for $\tau^*_n=j$,

$$|S^*_{n,[n\delta]+1}|<\hat\sigma_n\sqrt{n}(\eta-\sqrt{2\delta})$$

implies $|S^*_{n,j}-S^*_{n,[n\delta]+1}|>\hat\sigma_n\sqrt{2n\delta}$, and by Čebišev inequality it follows that

$$P^*\left\{|S^*_{n,[n\delta]+1}|<\hat\sigma_n\sqrt{n}(\eta-\sqrt{2\delta})\mid\tau^*_n=j\right\}$$
$$\le\frac{1}{2n\delta\hat\sigma_n^2}V^*\left(\sum_{i=j+1}^{[n\delta]+1}\epsilon^*_{n,i}\right),\quad 1\le j\le[n\delta]. \qquad (7.23)$$

Moreover, the right hand side in (7.23) is bounded above by 1/2.
Therefore, going back to (7.22)

$$P^*\left\{\max_{0\le j\le[n\delta]+1}|S^*_{n,j}|>\eta\hat\sigma_n\sqrt{n}\right\}$$
$$\le P^*\left\{|S^*_{n,[n\delta]+1}|\ge\hat\sigma_n\sqrt{n}(\eta-\sqrt{2\delta})\right\}+\frac{1}{2}P^*\left\{\tau^*_n\le[n\delta]\right\}$$
$$\le P^*\left\{|S^*_{n,[n\delta]+1}|\ge\hat\sigma_n\sqrt{n}(\eta-\sqrt{2\delta})\right\}+\frac{1}{2}P^*\left\{\max_{0\le j\le[n\delta]+1}|S^*_{n,j}|>\eta\hat\sigma_n\sqrt{n}\right\}.$$

It follows that

$$P^*\left\{\max_{0\le j\le[n\delta]+1}|S^*_{n,j}|>\eta\hat\sigma_n\sqrt{n}\right\}\le 2P^*\left\{|S^*_{n,[n\delta]+1}|\ge\hat\sigma_n\sqrt{n}(\eta-\sqrt{2\delta})\right\}.$$

Putting $\lambda=(\eta-\sqrt{2\delta})/\sqrt{\delta}$ in (7.21), we have

$$\limsup_{n\to\infty}\frac{1}{\delta}P^*\left\{\max_{0\le j\le[n\delta]+1}|S^*_{n,j}|>\eta\hat\sigma_n\sqrt{n}\right\}\le\frac{2\sqrt{\delta}}{(\eta-\sqrt{2\delta})^3}E(|V|^3).$$

Now letting $\delta \downarrow 0$ the lemma follows. □

Proof of Lemma 3.3. Once we have Lemma 3.2, the proof follows as in Lemma 4.19 of Karatzas and Shreve (1988), page 69, replacing S_k by $S^*_{n,k}$, $k = 1, \ldots, n, n \in \mathcal{N}$. □

Proof of Proposition 3.1. Let Ω_1, Ω_2 and Ω_3 be the sets where lemmas 3.1, 3.2 and 3.3, respectively, hold. For all the sample paths in $\Omega_1 \cap \Omega_2 \cap \Omega_3$, the proof in page 71 of Karatzas and Shreve (1988) gives the tightness of $\{Y^*_n\}_{n=1}^{\infty}$; this and the finite dimensional convergence in (3.6) imply, by theorem 4.15 in Karatzas and Shreve (1988), the weak convergence in $C[0,1]$. □

Proof of Lemma 3.4. It is straightforward from Proposition 3.1. □

Proof of Theorem 3.1. Observe that

$$\begin{aligned} Z^*_n &= \tfrac{1}{\hat{\sigma}_n} \left(\textstyle\sum_{t=1}^n X^{*2}_{n,t-1}\right)^{1/2} \left(\hat{\beta}^*_n - 1\right) \\ &= \tfrac{1}{\hat{\sigma}_n} \left(\textstyle\sum_{t=1}^n X^{*2}_{n,t-1}\right)^{-1/2} \left(\textstyle\sum_{t=1}^n X^*_{n,t-1} \epsilon^*_{n,t}\right). \end{aligned}$$

Now, by squaring (3.4) and by summing, we obtain

$$\sum_{t=1}^{n} X^*_{n,t-1} \epsilon^*_{n,t} = \frac{1}{2} X^{*2}_{n,n} - \frac{1}{2} \sum_{t=1}^{n} \epsilon^{*2}_{n,t}. \tag{7.24}$$

Then, expressing the quantities $X^*_{n,t}$ in terms of $Y^*_n(t)$, defined in (3.5), we have

$$X^{*2}_{n,n} = n\hat{\sigma}^2_n Y^{*2}_n(1) \tag{7.25}$$

and

$$\sum_{t=1}^{n} X^{*2}_{n,t-1} = n\hat{\sigma}^2_n \sum_{i=1}^{n-1} Y^{*2}_n \left(\frac{i}{n}\right). \tag{7.26}$$

It follows from (7.24), (7.25) and (7.26) that

$$Z^*_n = \frac{1}{2} \left(Y^{*2}_n(1) - \frac{1}{\hat{\sigma}^2_n n} \sum_{t=1}^{n} \epsilon^{*2}_{n,t} \right) \left(\frac{1}{n} \sum_{i=1}^{n-1} Y^{*2}_n \left(\frac{i}{n}\right) \right)^{-1/2}.$$

By the bootstrap weak law of large numbers and Proposition 3.1, the numerator converges, for almost all $(X_1, \ldots, X_n)$, to $\frac{1}{2}(W^2(1) - 1)$. Moreover, from Lemma 3.4, Proposition 3.1 and the continuous mapping theorem, the denominator tends to $\left(\int_0^1 W^2(t)dt\right)^{1/2}$. Since can be easily proved that the bootstrap version of Slutsky's theorem holds, the theorem follows.□

Proof of Lemma 5.1. (i) First, we will prove that

$$\hat{\rho}_n - \rho_n \longrightarrow_P 0, \tag{7.27}$$

where ρ_n is the least squares estimate of ρ obtained from $u_1, \ldots, u_n$.
From Theorem 3.1 of Phillips (1987) for the sequence $\{X_t\}$ and by the law of large numbers of McLeish (1975) we have

$$\begin{aligned} \hat{\rho}_n &= \left(\sum_{t=1}^{n} \tilde{u}_{t-1}^2\right)^{-1} \sum_{t=1}^{n} \tilde{u}_t \tilde{u}_{t-1} \\ &= \rho_n \psi_n + o_P(1), \end{aligned} \tag{7.28}$$

where

$$\rho_n = \left(\sum_{t=1}^{n} u_{t-1}^2\right)^{-1} \sum_{t=1}^{n} u_t u_{t-1}, \tag{7.29}$$

$$\psi_n = \left(\sum_{t=1}^{n} \tilde{u}_{t-1}^2\right)^{-1} \sum_{t=1}^{n} u_{t-1}^2 \tag{7.30}$$

and $o_P(1)$ is a sequence of random variables tending to zero in probability and, so it is enough to see that

$$\psi_n \longrightarrow_P 1. \tag{7.31}$$

From the definition of $\tilde{u}_t$, we obtain

$$\begin{aligned} \frac{1}{n}\sum_{t=1}^{n} \tilde{u}_{t-1}^2 = \frac{1}{n}\sum_{t=1}^{n} u_{t-1}^2 - \frac{2(\hat{\beta}_n - 1)}{n}\sum_{t=1}^{n} u_{t-1}X_{t-2} \\ + \frac{(\hat{\beta}_n - 1)^2}{n}\sum_{t=1}^{n} X_{t-2}^2. \end{aligned} \tag{7.32}$$

So, from Theorem 3.1 of Phillips (1987) for the sequence $\{X_t\}$, we deduce

$$\frac{1}{n}\sum_{t=1}^{n}\tilde{u}_{t-1}^2 = \frac{1}{n}\sum_{t=1}^{n}u_{t-1}^2 + o_P(1). \tag{7.33}$$

By the law of large numbers of McLeish (1975) it follows that

$$\frac{1}{n}\sum_{t=1}^{n}u_{t-1}^2 \longrightarrow_P \sigma_u^2. \tag{7.34}$$

Then (7.31) follows from (7.30), (7.32), (7.33) and (7.34). Hence from (7.28), (7.29) and (7.31) one concludes (7.27). Therefore, from the consistency of ρ_n, (i) follows.

Now we will prove (ii). We have that

$$\begin{aligned} V^*(v_{n,t}^*) &= \frac{1}{n}\sum_{j=1}^{n}\hat{v}_j^2 \\ &= \frac{1}{n}\sum_{j=1}^{n}\tilde{v}_j^2 - \left(\frac{1}{n}\sum_{j=1}^{n}\tilde{v}_j\right)^2. \end{aligned} \tag{7.35}$$

First, we will prove that

$$\frac{1}{n}\sum_{j=1}^{n}\tilde{v}_j^2 \longrightarrow_{P^*} \sigma_v^2. \tag{7.36}$$

From the definition of $\tilde{v}_t$, we have

$$\frac{1}{n}\sum_{j=1}^{n}\tilde{v}_j^2 = \frac{1}{n}\sum_{j=1}^{n}\tilde{u}_j^2 - \frac{2\hat{\rho}_n}{n}\sum_{j=1}^{n}\tilde{u}_j\tilde{u}_{j-1} + \frac{\hat{\rho}_n^2}{n}\sum_{j=1}^{n}\tilde{u}_{j-1}^2. \tag{7.37}$$

Moreover, from the definition of $\tilde{u}_t$ we obtain

$$\begin{aligned} \frac{1}{n}\sum_{j=1}^{n}\tilde{u}_j\tilde{u}_{j-1} &= \frac{1}{n}\sum_{j=1}^{n}u_ju_{j-1} - \frac{2(\hat{\beta}_n-1)}{n}\sum_{j=1}^{n}u_jX_{j-1} \\ &\quad - \frac{\hat{\beta}_n-1}{n}\sum_{j=1}^{n}u_{j-1}X_{j-1} - \frac{\hat{\beta}_n-1}{n}\sum_{j=1}^{n}u_jX_{j-2}. \end{aligned} \tag{7.38}$$

By Theorem 3.1 of Phillips (1987) for the sequence $\{X_t\}$ and the law of large numbers of McLeish (1975), it follows that

$$\frac{1}{n}\sum_{j=1}^{n}\tilde{u}_j\tilde{u}_{j-1} = \frac{1}{n}\sum_{j=1}^{n}u_ju_{j-1} + o_{P^*}(1). \tag{7.39}$$

From the usual weak law of large numbers and the law of large numbers of McLeish (1975), one concludes

$$\frac{1}{n}\sum_{j=1}^{n}u_ju_{j-1} = \frac{1}{2n\rho}\left(\sum_{j=1}^{n}u_j^2 + \rho^2\sum_{j=1}^{n}u_{j-1}^2 - \sum_{j=1}^{n}v_j^2\right)$$
$$\longrightarrow_{P^*} \rho\sigma_u^2. \tag{7.40}$$

Then, from (7.33), (7.34), (7.35), (7.37), (7.38), (7.39), (7.40) and part (i) of the lemma, (7.36) follows.
Moreover, we have

$$\frac{1}{n}\sum_{j=1}^{n}\tilde{v}_j = \frac{1}{n}\sum_{j=1}^{n}(\tilde{u}_j - \hat{\rho}_n\tilde{u}_{j-1}) \tag{7.41}$$

and

$$\frac{1}{n}\sum_{j=1}^{n}\tilde{u}_j = \frac{1}{n}\sum_{j=1}^{n}u_j - \frac{\hat{\beta}_n - 1}{n}\sum_{j=1}^{n}X_{j-1}. \tag{7.42}$$

Hence, from (7.42), the law of large numbers of McLeish (1975) and Theorem 3.1 of Phillips (1987) for the sequence $\{X_t\}$ we obtain

$$\frac{1}{n}\sum_{j=1}^{n}\tilde{u}_j = \frac{1}{n}\sum_{j=1}^{n}u_j + o_{P^*}(1). \tag{7.43}$$

So, from the law of large numbers of McLeish (1975) we have

$$\frac{1}{n}\sum_{j=1}^{n}\tilde{u}_j \longrightarrow_{P^*} 0.$$

Then from (7.41), (7.43) and part (i) of this lemma, we deduce

$$\frac{1}{n}\sum_{j=1}^{n}\tilde{v}_j \longrightarrow_{P^*} 0. \tag{7.44}$$

Hence, from (7.35), (7.36) and (7.44), (ii) follows.

Finally, (iii) is an inmediate consequence of (i) and (ii). □

Proof of Proposition 5.1. We have to show that for any distance d metrizing weak convergence,

$$d(\mathcal{L}(U_n^*), \mathcal{L}(W)) \longrightarrow_P 0,$$

where $\mathcal{L}(X)$ is the distribution of the variable X.
This is equivalent to prove that for any subsequence $\{U_{n_k}^*\}$, there exists a further subsequence $\{U_{n_{k_j}}^*\}$, such that

$$d(\mathcal{L}(U_{n_{k_j}}^*), \mathcal{L}(W)) \longrightarrow_{a.s.} 0$$

(see, e.g., Giné and Zinn (1990)).

From the convergence in probability in Lemma 5.1, it follows that given the subsequence $\{n_k\}$, there exists a subsequence $\{n_{k_j}\}$ of it for which the lemma holds almost surely. Now, we will check that conditions in Corollary 1 of Herrndorf (1984) hold a.s. for the subsequence $\{u_{n_{k_j},t}^* : t = 1, \ldots, n_{k_j}\}$. To simplify the notation we will just write $\{n\}$ for the subsequence $\{n_k\}$.

Obviously,

$$E^*(u_{n,t}^*) = 0 \tag{7.45}$$

and

$$E^*(u_{n,t}^{*2}) < \infty. \tag{7.46}$$

Now, we will first prove that

$$\frac{E^*(R_{n,n}^{*2})}{n} \longrightarrow_{P^*} \sigma^2 = \frac{\sigma_v^2}{(1-\rho)^2} \text{ as } n \to \infty, \tag{7.47}$$

where $R_{n,n}^* = \sum_{j=1}^n u_{n,j}^*$. We have that

$$\frac{E^*(R_{n,n}^{*2})}{n} = E^*(u_{1,1}^{*2}) + \frac{2}{n} \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E^*(u_{n,i}^* u_{n,j}^*).$$

Then, from Lemma 5.1 and the fact that $|\rho| < 1$, we obtain (7.47). So, we have

$$\frac{E^*(R^{*2}_{n_k,n_k})}{n_k} \longrightarrow_{a.s.} \sigma^2 \text{ as } n_k \to \infty, \tag{7.48}$$

for a further subsequence.

Now, we will show that $\{u_n^*\}$ is strong mixing with mixing coefficients α_m^* satisfying

$$\sum_{m=1}^{\infty} \alpha_m^{*1-2/\gamma} < \infty. \tag{7.49}$$

For this, we first establish

A1*. $\int | \phi^*(r) | \, dr < \infty$, where ϕ^* is the characteristic function of each $v_{n,t}^*$.

A2*. $\sup_t E(| v_{n,t}^* |^\theta) \longrightarrow_{P^*} M_1 < \infty$ for some $\theta > 0$.

A3*. $\sup_{m,s,k\geq 1} \sup_{\alpha,\tau,\nu} \max_t \left| \frac{\partial}{\partial \nu_t} P(\tilde{U}^* + \nu \in \bigcup_{j=1}^{s} D_j) \right| \longrightarrow_{P^*} M_2 < \infty,$

where $D_j = \times_{t=k}^{k+m-1} (\alpha_{jt}, \tau_{jt}), \nu = (\nu_k, \ldots, \nu_{k+m-1}), \tilde{U}^* = (u_{n,k}^*, \ldots, u_{n,k+m-1}^*)$

conditionally on $(X_1, \ldots, X_n)$ for almost all $(X_1, \ldots, X_n, \ldots)$ and, in this way, we get the conditions of Corollary 3 in Withers (1981) for $\{v_{n,t}^* : t = 1, ..., n\}$ almost surely along subsequences.

Condition A1* is obvious, since the distribution of $v_{n,t}^*$ is $\hat{F}_n$. For A2*, let $\theta = \gamma + \eta$, where $\gamma > 2$ and $\eta > 0$ are the same as in the assumption A2. We have

$$E(| v_{n,t}^* |^\theta) \leq 2^{\theta-1} \frac{1}{n} \left(\sum_{j=1}^{n} | \tilde{v}_j |^\theta + | \frac{1}{n} \sum_{j=1}^{n} \tilde{v}_j |^\theta \right). \tag{7.50}$$

From the definition of $\tilde{v}_t$, we deduce that

$$\frac{1}{n} \sum_{j=1}^{n} | \tilde{v}_j |^\theta \leq \frac{2^{\theta-1}}{n} \left(\sum_{j=1}^{n} | v_j |^\theta + | \hat{\rho}_n - \rho |^\theta \sum_{j=1}^{n} | \tilde{u}_{j-1} |^\theta \right). \tag{7.51}$$

On the other hand, from the definition of $\tilde{u}_t$ and (5.8) with $\beta = 1$, it follows that

$$\frac{1}{n}\sum_{j=2}^{n} | \tilde{u}_{j-1} |^{\theta} \leq \frac{2^{\theta-1}}{n}\left(\sum_{j=2}^{n} | u_{j-1} |^{\theta} + | \hat{\beta}_n - 1 |^{\theta} \sum_{j=2}^{n} | X_{j-2} |^{\theta}\right). \tag{7.52}$$

From assumption A2, the law of large numbers of McLeish (1975) and Theorem 3.1 of Phillips (1987) for the sequence $\{X_t\}$, we deduce that

$$\frac{1}{n} | \hat{\beta}_n - 1 |^{\theta} \sum_{j=2}^{n} | X_{j-2} |^{\theta} \longrightarrow_{P^*} 0 \tag{7.53}$$

and

$$\frac{1}{n}\sum_{j=2}^{n}(| u_{j-1} |^{\theta} - E(| u_{j-1} |^{\theta})) \longrightarrow_{P^*} 0. \tag{7.54}$$

Thus, from Lemma 5.1 (i), (7.51), (7.52), (7.53) and (7.54), we have

$$\frac{1}{n}\sum_{j=1}^{n} | \tilde{v}_j |^{\theta} \leq \frac{2^{\theta-1}}{n}\sum_{j=1}^{n} | v_j |^{\theta} + o_{P^*}(1). \tag{7.55}$$

Then, from (7.44), (7.50), (7.55), assumption A2 and the weak law of large numbers, condition A2* follows.

Condition A3* follows from condition A3 because the empirical distribution corresponding to $\{\hat{v}_1, \ldots, \hat{v}_n\}$ tends to the distribution of v_t, by (7.44), Lemma 5.1 (i) and Boldin (1982).

Finally, we will show that, for some $\gamma^* > 2$

$$(E^*(| u^*_{n,t} |^{\gamma^*}))^{1/\gamma^*} \longrightarrow_{P^*} M_3 < \infty. \tag{7.56}$$

By Minkowski's inequality, it follows that

$$(E^*(| u^*_{n,t} |^{\gamma^*}))^{1/\gamma^*} \leq (\sum_{j=0}^{\infty} | \hat{\rho}_n |^{j}) \sup_t (E(| v^*_{n,t} |^{\gamma^*}))^{1/\gamma^*}.$$

Hence (7.56) follows the assumption A2* with $\gamma^* = \theta$, Lemma 5.1 (i) and the fact that $| \rho |< 1$. Thus, along a subsequence, we obtain

$$(E^*(| u^*_{n,t} |^{\gamma^*}))^{1/\gamma^*} \longrightarrow_{a.s.} M_3 < \infty. \tag{7.57}$$

So, from (7.45), (7.46), (7.48), (7.49) and (7.57), conditions in Corollary 1 of Herrndorf (1984) are satisfied almost surely along subsequences and the proposition follows. □

Proof of Theorem 5.1. Note that

$$\begin{aligned} T_n^* &= \tfrac{1}{s_n}\left(\textstyle\sum_{t=1}^n X_{n,t-1}^{*2}\right)^{1/2}\left(\hat{\beta}_n^* - 1\right) \\ &= \tfrac{1}{s_n}\left(\textstyle\sum_{t=1}^n X_{n,t-1}^{*2}\right)^{-1/2}\left(\textstyle\sum_{t=1}^n X_{n,t-1}^{*}u_{n,t}^*\right). \end{aligned}$$

Then, by squaring (5.15) and by summing, we have

$$\sum_{t=1}^{n} X_{n,t-1}^* u_{n,t}^* = \frac{1}{2}X_{n,n}^{*2} - \frac{1}{2}\sum_{t=1}^{n} u_{n,t}^{*2}. \tag{7.58}$$

Hence, expressing the quantities $X_{n,t}^*$ in terms of $U_n^*(t)$, defined in (5.16), we obtain

$$X_{n,n}^{*2} = n\sigma^2 U_n^{*2}(1) \tag{7.59}$$

and

$$\sum_{t=1}^{n} X_{n,t-1}^{*2} = n\sigma^2 \sum_{i=1}^{n-1} U_n^{*2}\left(\frac{i}{n}\right). \tag{7.60}$$

From (7.58), (7.59) and (7.60) it follows that

$$T_n^* = \frac{\sigma}{2s_n}\left(U_n^{*2}(1) - \frac{1}{\sigma^2 n}\sum_{t=1}^{n} u_{n,t}^{*2}\right)\left(\frac{1}{n}\sum_{i=1}^{n-1} U_n^{*2}\left(\frac{i}{n}\right)\right)^{-1/2}.$$

By the proof in page 297 of Phillips (1987) we have

$$s_n^2 \longrightarrow_P \sigma_u^2, \qquad as \quad n \to \infty. \tag{7.61}$$

Finally, we have to show that for any distance d metrizing weak convergence,

$$d(\mathcal{L}(T_n^*), \mathcal{L}(T)) \longrightarrow_P 0.$$

This is equivalent to prove that for any subsequence $\{T_{n_k}^*\}$, there exists a subsequence $\{T_{n_{k_j}}^*\}$, such that

$$d(\mathcal{L}(T^*_{n_{k_j}}),\mathcal{L}(T)) \longrightarrow_{a.s.} 0.$$

Given a subsequence $\{n_k\}$, by the bootstrap weak law of large numbers, Proposition 5.1 and (7.61), there exists a further subsequence $\{n_{k_j}\}$ such that the numerator converges, almost surely to $\frac{\sigma}{2\sigma_u}(W^2(1) - \frac{\sigma_u^2}{\sigma^2})$. Moreover, from Lemma 3.4 (with $\{Y_n^*(t) : t \in [0,1]\}_{n=1}^{\infty}$ replaced by $\{U_n^*(t) : t \in [0,1]\}_{n=1}^{\infty}$), Proposition 5.1 and the continuous mapping theorem, the denominator tends to $\left(\int_0^1 W^2(t)dt\right)^{1/2}$, along the subsequence. Then, the proof of the theorem follows from the bootstrap version of Slutsky's theorem almost surely along a subsequence and so, in probability. □

REFERENCES

Basawa, I.V., Mallik, A.K., McCormick, W.P. and Taylor, R.L. (1989). Bootstrapping explosive autoregressive processes. *Annals of Statistics* **17** 1479-1486.

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

Basawa, I.V., Mallik, A.K., McCormick, W.P., Reeves, J.H. and Taylor, R.L. (1991-b). Bootstrap test of significance and sequential bootstrap estimation for unstable first order autoregressive processes. *Communications in Statistcs* **20 (3)** 1015-1026.

Bertail, P. (1991). Test bootstrap: applications au modele lineaire et au modele autoregressif d'ordre 1. Preprint.

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

Boldin, M.N. (1982) Estimation of the distribution of noise in an autorregresion scheme. *Theory Probability Appl.* **27** 866-871.

Bose, A. (1988). Edgeworth correction by bootstrap in autoregressions. *Annals of Statistics* **16** 1709-1722.

Box, G.E.P. and Pierce, D.A. (1970). Distribution of residual autoregressive integrated moving average time series models. *Journal of the American Statistical Association* **65**, 1509-1526.

Dickey, D.A. (1976). Estimation and hypothesis testing in nonstationary time series. *Unpublished Ph.D. thesis*, Iowa University, Ames, IA.

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

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

Efron, B. (1979). Bootstrap methods: another look to the jackknife. *Annals of Statistics* **7**, 1-26.

Evans, G.B.A and Savin, N.E. (1981). Testing for unit roots: 1. *Econometrica* **49** 753-779.

Evans, G.B.A and Savin, N.E. (1984). Testing for unit roots: 2. *Econometrica* **52** 1241-1269.

González Manteiga, W., Prada Sánchez, J. M. and Romo, J. (1992). The bootstrap: a review. Preprint.

Freedman, D.A. (1984). On bootstrapping two-stage least squares estimates in stationary linear models. *Annals of Statistics* **12**, 3, 827-842.

Fuller, W.A. (1976). *Introduction to Statistical Time Series.* Wiley, New York, NY.

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

Hasza, D.P. (1977). Estimation in nonstationary time series. *Unpublished Ph.D. thesis.* Iowa State University, Ames, IA.

Hasza, D.P. and Fuller, W.A. (1979). Estimation for autoregression processes with unit roots. *Annals of Statistics* **7** 1106-1120.

Herrndorf, N. (1984). A functional central limit theorem for weakly dependent sequences of random variables. *Annals of Statistics* **12**, 1, 141-15.

Karatzas, I. and Shreve, S. (1988). *Brownian Motion and Stochastic Calculus.* Springer-Verlag, New York.

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

Ljung, G.M. and Box, G.E.P. (1978). On a measure of lack of fit in time series models. *Biometrika* **65** 297-303.

McLeish, D.L. (1975). A maximal Inequality and Dependent Strong Laws. *Annals of Probability* **3** 829-839.

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

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

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

Rao, M.M. (1978). Asymptotic distribution of an estimator of the boundary parameter of an unstable process. *Annals of Statistics* **6** 185-190.

Rao, M.M. (1980). Correction to "Asymptotic distribution of an estimator of the boundary parameter of an unstable process". *Annals of Statistics* **8** 1403.

Rayner, R. K. (1990). Bootstrapping p-values and power in the first order autoregression: a Monte Carlo investigation. *Journal of Business & Economics Statistics*, **8**, 2, 251-263.

Rubin, H. (1950). Consistency of maximun likelihood estimates in the explosive case. *In Statistical Inference in Dynamic Economic Models* (T. C. Koopmans, ed.) 356-364. Wiley, New York.

Saïd, S.E. (1982). Testing for unit roots in autoregressive moving average models. *Unpublished Ph.D. thesis.* North Carolina State University, Raleigh, NC.

Saïd, S.E. (1991). Unit–roots test for time-series data with a linear time trend. *Journal of Econometrics* **47** 285-303.

Saïd, S.E. and Dickey, D.A. (1984). Testing for unit roots in ARMA(p,q) models with unknown p and q. *Biometrika* **71** 599-607.

Saïd, S.E. and Dickey, D.A. (1985). Hypothesis testing in ARIMA (p,1,q) models. *Journal of the American Statistical Association* **80** 369-374.

Solo, V. (1984). The order of differencing in ARIMA models. *Journal of the American Statistical Association* **79** 916-921.

Stoffer, D.S. and Wall, K.D. (1991). Bootstrapping state space models: Gaussian maximum likelihood estimation and the Kalman filter. *Journal of the American Statistical Association* **86**, 416, 1024-1033.

Stute, W. and Gründer, B.(1990). Prediction intervals for explosive AR(1) processes. *Lecture Notes in Economics and Mathematical Systems* **376**, 121-130.

Thombs, L.A. and Schucany, W.R. (1990). Bootstrap prediction intervals for autoregression. *Journal of the American Statistical Association* **85**, 410, 486-492.

White, J.S. (1958). The limiting distribution of the serial correlation coefficient in the explosive case. *Annals of Mathemathical Statistics* **29** 1188-1197.

Withers, C.S. (1981). Conditions for linear processes to be strong- mixing. *Z. Wahrscheinlichkeitstheorie und verwandte Gebiete* **57** 477-480.

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
28903 Getafe (Madrid)- Spain