STATISTICA, anno LXIII, n. 3, 2003

IMPROVING THE POWER OF UNIT ROOT TESTS AGAINST FRACTIONAL ALTERNATIVES USING BOOTSTRAP

L. Bisaglia, I. Procidano

1. INTRODUCTION

In the empirical applications of Box-Jenkins ARIMA*(p,d,q)* models, the number of times that the observed time series must be differenced (parameter *d*) is usually determined by intuitive (but informal) methods, for example the analysis of the empirical autocorrelation function of the differenced series. However, for many observed time series, taking the first or the second difference may be too strong. Another approach, is to use models based on fractional integration that permit the difference parameter *d* to assume non-integer values. Since the seminal papers of Granger and Joyeux (1980) and Hosking (1981), this kind of models has received increasing attention because of their ability to capture the persistent temporal dependence that many empirical time series exhibit. At the same time, the unit root hypothesis is included as a special case of this kind of models. Typical features of such data are that, even if stationary, the sample autocorrelations decrease to zero like a power function rather than exponentially and also that the spectral density diverges as the frequencies tend to zero. Thus, they can present spurious local trends and cycles that disappear after a while. So, specially for short time series, it could be almost impossible distinguish a stationary process with long-memory from a unit root process. Simple persisting trends can be distinguished from stationary behaviours only if we have a long enough time series.

Sowell (1990) derives the asymptotic distributions for the Dickey-Fuller test under the hypothesis that the data generating process is a non-stationary fractional integrated process, that is $X_t \sim I(d)$ with $d \in (1/2, 3/2)$. He shows that the limiting distributions of fractionally integrated series are radically different from limiting distributions of series integrated of order zero or one and also that testing for unit roots is complicated because the *t*-statistic in this model only converges when $d = 1$.

Diebold and Rudebush (1991) examine the properties of Dickey-Fuller test under fractionally integrated alternatives and show by Monte Carlo simulations that this test has quite low power and can lead to the incorrect conclusion that a time series has a unit root also when this is not true.

Hassler and Wolters (1994) investigate the probability of rejecting the I(1) hypothesis when unit root tests are applied to fractionally integrated time series. They find out that especially the Augmented Dickey-Fuller test (ADF in the following) performs poorly. In particular, they find that the ADF test loses considerably power when augmented terms are added. They support their theoretical arguments with some Monte Carlo experiments.

Kramer (1998), in contrast with Hassler and Wolters (1994), shows that the ADF test is consistent against fractional alternatives if the order of the autoregression does not tend to infinity too fast. His results are not supported by simulation experiments in finite samples.

In this work we want firstly to clarify, via Monte Carlo study, this apparent contradiction, and, secondarily, we want to see if a bootstrap approach can help to improve the power of ADF test when the alternatives are long-memory. Moreover, we will consider a different class of tests, usually used to determine if a series have long-memory, to test the null hypothesis of unit root against fractional alternatives. Since these tests are asymptotic, they often exhibit non negligible size distortion in small samples. To improve inference, in this work we propose a boostrap procedure to adjust critical values.

The plan of the paper is the following. In section 2 we introduce the problem and recall some known results about fractionally ARIMA models. Section 3 introduces the Augmented Dickey-Fuller test and some fractional unit root tests. In section 4 we consider the problems of bootstrap when the data generating process has unit roots and/or long-memory, and we propose our bootstrap technique. Section 5 presents the Monte Carlo study and the results about the size and the power of the tests. Conclusions are offered in Section 6.

2. ARFIMA PROCESSES AND UNIT ROOTS

The Autoregressive Fractionally Integrated Moving-Average (ARFIMA) process generalizes the usual ARIMA(p,d,q) process by assuming d to be fractional. This generalization provides a more flexible framework to study empirical time series data. In fact, this class of processes can be used to model data dependence that is stronger than allowed in stationary ARMA processes and weaker than implied by non-stationary unit root processes.

More specifically, ARFIMA(p,d,q) processes are stationary and invertible processes of the form:

$$\Phi(B)\Delta(B)(X_t - \mu) = \Theta(B)\varepsilon_t$$

where the ε_t are white noise with zero mean and variance σ^2, $\Phi(\cdot)$ and $\Theta(\cdot)$ are polynomials in the backward operator B of degrees p and q respectively, $\Delta(B) = (1 - B)^d$ and $d \in (-1/2, 1/2)$ is the fractional parameter. Formally, the fractional difference operator $(1 - B)^d$ can be expressed by a binomial expansion, that is:

$$(1-B)^d = \sum_{j=0}^{\infty} \pi_j B^j \quad with \quad \pi_j = \frac{\Gamma(j-d)}{\Gamma(j+1)\Gamma(-d)}$$

where $\Gamma(\cdot)$ is the Gamma function. We can see that if the parameter d is restricted to the set of integers we obtain the usual ARIMA processes. When $d < 0$ the process X_t is said antipersistent, while when $d > 0$ X_t is simply said to have long-range dependence or long-memory. If $d > 1/2$ the process becomes non stationary, but it is always possible to transform it into a stationary process by taking a suitable number of integer differences (for details about ARFIMA process, see Granger and Joyeux, 1980 and Hosking, 1981).

Now, let x_t, $t = 1,2,\dots,T$, be a real-valued time series whose behaviour is governed by the following model:

$$X_t = X_{t-1} + \eta_t \tag{1}$$

$$(1 - B)^d \eta_t = \varepsilon_t$$

where $d \in (-1/2,1/2)$ and the ε_t are white noise with zero mean and variance σ^2, that is η_t is an ARFIMA($0, d, 0$) process or a fractionally integrated noise. Since it is possible to write (1) as

$$(1 - B)\, X_t = \eta_t\,,$$

that is, X_t is a random walk with fractionally integrated innovations, it follows that

$$(1 - B)^{\delta} X_t = \varepsilon_t$$

where $\delta = (1 + d) \in (1/2, 3/2)$. Even if any fractional process can be written as one that includes an autoregressive unit root, following Diebold and Rudebush (1991) and subsequent papers, only I(1) processes, i.e. $\delta = 1$ (or equivalently $d = 0$), will be said to contain a unit root. Then, our problem is to test the null hypothesis H_0: $\delta = 1$ against fractional alternatives. In the next section we will introduce some tests with this aim.

3. FRACTIONAL UNIT ROOT TESTS

Firstly, we consider one of the most popular unit root test: the Augmented Dickey-Fuller, or ADF test (Said and Dickey, 1984). This test is implemented in many software packages, it would be nice if it worked well also to distinguish unit root by fractional unit root.

We consider the generating process $X_t = \rho X_{t-1} + u_t$, where $u_t \sim ARMA(p,q)$, Fuller (1976) demonstrates that, under the null hypothesis of a unit root ($\rho = 1$), X_t reduces to

$$\Delta X_t = \rho X_{t-1} + \sum_{i=1}^{\infty} \beta_i \Delta X_{t-i} + \varepsilon_t$$

where the ε_t are *i.i.d.*$(0, \sigma^2)$, and the ADF test is simply based on the regression t-statistic for the hypothesis $\rho = 0$. The asymptotic distribution of the original Dickey-Fuller statistic is not altered.

Another approach is to apply some tests for fractional integration to the first differences of X_t. In fact, if X_t is a unit root process ($\delta = 1$) its first differences will be a white noise process and $d = 0$. If X_t is a fractionally integrated process then $\delta \in (1/2, 3/2)$ and its first differences will be a stationary fractionally process with fractional parameter $d \in (-1/2, 1/2)$. Thus the null hypothesis of unit root becomes $H_0: d=0$ in the first differences.

The tests we consider are the t-test based on the log-periodogram regression of Geweke and Porter-Hudack (1993), GPH, the modified rescaled range statistic derived by Lo (1991), MRS, and a LM type test recently developed by Lobato and Robinson (1998).

The first test is based on a semi-parametric procedure to test for fractional integration. It is motivated by the log-spectral density of the ARFIMA process, and amounts to estimating the least squares regression

$$\ln I(w_j) = c - d \ln\left\{4 \sin^2\left(\frac{w_j}{2}\right)\right\} + \zeta_j \qquad j = 1, 2, \ldots, m$$

where $I(\omega_j)$ is the periodogram of X_t at the Fourier frequencies $w_j = 2\pi j/T$ and m is a positive integer, usually chosen as $m = [(T)^{1/2}]$ with $[\cdot]$ denoting the integer part. There is evidence of fractional integration if $\hat{d}$, the least square estimate of the long memory parameter, is significantly different from zero.

The second test is an extension of the *range over standard deviation*, or R/S statistic proposed by Hurst (1951). This statistic, Q_T, is defined by

$$Q_T = \frac{1}{\hat{\sigma}_T(q)}\left[\underset{1 \le k \le T}{Max} \sum_{j=1}^{k} (X_j - \bar{X}) - \underset{1 \le k \le T}{Min} \sum_{j=1}^{k} (X_j - \bar{X})\right]$$

where $\bar{X}$ is the sample mean over the T observations, $\hat{\sigma}_T^2(q)$ is given by

$$\hat{\sigma}_T^2(q) = \hat{\sigma}_x^2 + 2\sum_{j=1}^{q}\left(1 - \frac{j}{q+1}\right)\hat{\gamma}_j \qquad q < T$$

and $\hat{\sigma}_x^2$ and $\hat{\gamma}_j$ are the usual sample variance and autocovariance estimators of the data. Usually q is set equal to the integer part of $(3T/2)^{1/3}(2\hat{\rho}/(1-\hat{\rho}^2))^{2/3}$, $\hat{\rho}$ being the sample first-order autocorrelation coefficient. Extreme values of Q_T

are regarded as signs of fractional integration. Under a set of conditions satisfied by a large class of short memory processes, Lo (1991) derives the asymptotic distribution of the statistic $V_T(q) = \frac{Q_T(q)}{\sqrt{T}}$ and gives the critical values.

Finally, the third test we consider is a non parametric one and makes no assumptions on spectral behaviour away from zero frequency. As far as our problem is concerned, the test statistic is given by

$$\lambda = -\sqrt{m}\left[\frac{\sum_{j=1}^{m} \upsilon_j I(\omega_j)}{\sum_{j=1}^{m} I(\omega_j)}\right]$$

where $\upsilon_j = \ln j - \frac{1}{m}\sum_{j=1}^{m} \ln j$ and $m \in (0, T/2)$ is a cutting parameter that we have set equal to the integer part of $(T-1)/2$. Under the null hypothesis (H_0: $d=0$) the statistic λ have an asymptotic standard normal distribution.

Since for some of these tests, Cheung (1993) reports evidence of serious size distortions in small samples, our idea is to correct the distortions using the bootstrap method.

4. BOOTSTRAP UNIT ROOT TESTS

In this section we firstly recall all the problems that can arise by using the bootstrap in presence of unit root processes and then we present our bootstrap scheme.

4.1. *Bootstrap and unit roots*

The consistency of the bootstrap estimator of the distribution of the least square estimator of the autoregressive parameter in a simple first order autoregressive process, which may or may not have a stationary solution, has been investigated by several authors recently.

Basawa *et al.* (1991a) show that, when the data generating process is $X_t = \rho X_{t-1} + u_t$, where the u_t are i.i.d. with zero mean and unit variance, then when $\rho = 1$ the bootstrap least square estimate is asymptotically invalid even if the error distribution is assumed to be normal. In this case, in fact, the consistency of the bootstrap estimator is much more sensitive to how the bootstrap sample is drawn than when it is known that $|\rho < 1|$. One way to overcome this problem is by resampling the restricted residuals under the null hypothesis of a unit root, that is by specifying that the true value of ρ is 1 (see Basawa *et al.* 1991b for details).

Datta (1996) considers a first order autoregressive process where the autoregressive parameter ρ may vary over the entire real line and shows that one way of fixing the problem when $\rho = 1$ is to reduce the bootstrap sample size m. In particular, he demonstrates that if one increases m to infinity in such a way that $m/T \to 0$, where T is the length of the time series, then the bootstrap will approximate the sampling distribution in probability. If $[m(\log \log T)^2]/T \to 0$ as $T \to \infty$, then the bootstrap will work almost surely. It remains open the problem of an optimal choice of m. A different modification of the standard bootstrap which retains the original sample size for the bootstrap is proposed by Datta and Sriram (1997). Their procedure employs a data dependent shrinkage of the least squares estimator towards the critical value $\rho = 1$ and it can be extended to the AR(p) model, $p > 1$.

Ferretti and Romo (1996) develop a bootstrap resampling scheme to test H_0: $\rho = 1$ for the first order autoregressive models and establish the asymptotic validity of the bootstrap test statistic both for independent and for AR errors. In the former case, the bootstrap methodology approaches directly the asymptotic distribution, making unnecessary the usual corrections due to the dependence of innovations.

In any case, it is not currently known if the bootstrap can provide improvements in accuracy when the DGP has unit roots and long-memory. In these cases, bootstrap methods based on sieve or other non parametric approximations to the data generating process may work better even if much further research is needed to determine if such approximations work well both in theory and in practice.

The sieve bootstrap was first introduced by Kreiss (1992) and then developed by Bühlmann (1997). This method is based on the idea of sieve approximation: it approximates a general linear, invertible process by a finite autoregressive model with order increasing with the sample size and resampling from the approximated autoregressions. By viewing such autoregressive approximations as a sieve for the underlying infinite-order process, the bootstrap procedure may still be regarded as a non parametric one. Chang and Park (1999) consider a sieve bootstrap for the test of a unit root in models driven by general linear process. The resulting tests are shown to be consistent under very general conditions. Simulations support their theoretical arguments. Psaradakis (2001) uses sieve bootstrap tests of the null hypothesis of an autoregressive unit root in models which are driven by innovations that belong to the class of stationary and invertible linear processes. He shows that the sieve approach provides asymptotically valid tests and supports his results in small samples by simulations.

The wild bootstrap was developed by Liu (1988) following a suggestion of Wu (1986). In his work, Liu shows that the wild bootstrap provides refinements for the linear regression model with heteroskedastic errors. Procidano and Rigatti Luchini (2001) propose this kind of bootstrap to test the hypothesis of unit roots in time series with heteroskedastic innovations. In a related work Pizzi *et al.* (2001) show that wild bootstrap applied to test the null hypothesis of unit root is also robust in the presence of outliers and non linearities.

Our idea is to combine sieve and wild bootstrap to see if this approach works better in the presence both of unit root and long-memory. The bootstrap scheme is presented in the next section.

4.2. *The bootstrap scheme*

Our procedure is motivated by the fact that, if $d \in (-1/2,1/2)$ the fractionally integrated process, η_t, admits the infinite autoregressive representation

$$(1-B)^d \eta_t = \sum_{j=0}^{\infty} \pi_j \eta_{t-j} = \varepsilon_t$$

where $\pi_0 = 1$ and

$$\pi_j = \pi_{j-1} \frac{j-1-d}{j}, \qquad j = 1, 2, \dots,$$

(see, for instance, Brockwell and Davis, 1991),

Having observed the sample $x_1, x_2, ..., x_T$, the algorithm we followed is defined in the following block diagram[1]:

1. Calculate the estimated residuals, $\eta_t = x_t - x_{t-1}$, under the null hypothesis H_0: $\rho = 1$.

↓

2. Fit an autoregressive process to the estimated residuals
$\Delta x_t = \beta_1 \Delta x_{t-1} + \beta_2 \Delta x_{t-2} + \dots + \beta_{p(T)} \Delta x_{t-p(T)}$
and estimate the coefficients $\hat{\beta}_{1,T}, \hat{\beta}_{2,T}, \dots, \hat{\beta}_{p,T}$.

↓

3. Construct the residuals
$$\varepsilon_{t,T} = \Delta x_t - \hat{\beta}_{1,T} \Delta x_{t-1} - \hat{\beta}_{2,T} \Delta x_{t-2} - \dots - \hat{\beta}_{p,T} \Delta x_{t-p}, \qquad t = p+1, \dots, T.$$

↓

[1] The routine to perform these bootstrap tests are written in R and are available upon request by the authors.

4. Extract, with reintroduction, $(T-p-1)$ elements of a random variable, e^*, such that $E(e^*) = 0$, $E(e^{*2}) = 1$.

↓

5. Generate an i.i.d. sample $\varepsilon_{t,T}^*$, $t = p+1, \ldots, T$ in the following way

$$\varepsilon_{t,T}^* = (1\text{-}diag(\boldsymbol{P}))^{-1} \left| \varepsilon_{t,T} \right| \cdot e_t^*$$

where $\mathbf{P} = \mathbf{X}(\mathbf{X'X})^{-1}\mathbf{X'}$ and $\mathbf{X} = [\Delta x_{t-1}, \Delta x_{t-p}, \ldots, \Delta x_{t-p}]$.

↓

6. Generate bootstrap replicates, $x_{t,T}^*$, $t = p+1, \ldots, T$, according to

$$\Delta x_{t,T}^* = \hat{\beta}_{1,T}\Delta x_{t-1,T}^* + \hat{\beta}_{2,T}\Delta x_{t-2,T}^* + \ldots + \hat{\beta}_{p,T}\Delta x_{t-p,T}^* + \varepsilon_{t,T}^*$$

$$x_{t,T}^* = \sum_{j=1}^{t} \Delta x_{j,T}^* .$$

↓

7. Obtain the bootstrap version of the ADF test (hereafter ADF*) running the bootstrap ADF regression

Steps 4 – 7 are repeated a large number of times yielding an empirical distribution for ADF*.

Remarks. With regard to step 1, note that the residuals correspond to the first differences of the series, that is $\eta_t = \Delta x_t$, where $\Delta = (1 - B)$,

With regard to step 2, *i)* to estimate the residuals, we prefer to use the Yule-Walker method since it always yields an invertible autoregression, but other estimation methods, like OLS or likelihood methods, are equivalent; *ii)* we fit the autoregressive process with increasing order $p(T)$ as the sample size T increases. Let $p = p(T) \to \infty$ $(T \to \infty)$ with $p(T) = o(T^{1/2+d})$ (Kramer, 1998). In practice we fixed $p_{\max} = 10\log_{10}(T)$ and then, following a suggestion of Bühlmann (1997), we chose the optimal p using the AIC criterion, $p = p_{AIC}$ so *iii)* in contrast with Hassler and Wolters (1994), who fix in their experiments the order p of the ADF autoregression, our procedure selects for each series the suitable value for the order p.

With regard to step 4, in the literature, the further condition that $E(\varepsilon_t^3) = 1$ is often added. In our experiments we chose like auxiliary, or external, distribution,

$e^* \sim U(-\sqrt{3}, \sqrt{3})$. In practice, the choice of the external distribution is not very important as Davidson and Flachaire (2001) show.

With regard to step 6, *i)* it is important to base the bootstrap sampling on this regression with the unit root restriction $\rho = 0$. The samples generated by the same regression without the unit root restriction do not behave like unit root processes, and this will render the subsequent bootstrap procedure inconsistent as shown in Basawa *et al.* (1991a); *ii)* we have to chose appropriately p initial values of $\Delta x^*_{t,T}$. An obvious choice for the initial values, would be to use the initial value x_0 for $x^*_{0,T}$, and generate the bootstrap samples $\Delta x^*_{t,T}$ conditioned on x_0. Even if the choice of initial value may affect the finite sample performance of the bootstrap, however the effect of the initial condition would disappear asymptotically. We have therefore set $x_0 = 0$.

Finally, to compare the performance of the bootstrap test against ADF test, we do not use the critical values tabulated by Fuller (1976) but those more recently tabulated by Mackinnon (1991).

5. MONTE CARLO SIMULATION STUDY

In this section we use some Monte Carlo experiments to investigate the size and power properties of the described tests in finite samples. In particular we want: 1) to compare the performances of the ADF test against Mackinnon's critical values and those obtained with our bootstrap test; 2) to investigate the size and power properties of the other fractional tests in finite samples.

Simulation 1. The properties of the ADF and the ADF* procedure for testing H_0: $\delta = 1$ against H_1: $\delta < 1$ are assessed. We did not consider values of $\delta > 1$ because for these values the non-stationarity of the process is so strong that the ADF test always accept the null of a unit root. Thus, in this case, the test procedure always says that it needs to differentiate the data. The Monte Carlo experiments we conducted is based on the following data generating process:

$$(1 - B)^{\delta} x_t = \varepsilon_t, \quad t = 1, 2, \ldots, T$$

with $\delta = 0.35, 0.45, 0.6, 0.7, 0.8, 0.9, 1$ and $\varepsilon_t \sim i.i.d.\ N(0,1)$. Our results are based on 1000 independent replications with T = 50, 100, 250 and 500 observations and for each simulated series we considered 500 bootstrap replications. The stationary fractional noise are generated using the recursive Durbin-Levinson algorithm (Brockwell and Davis, 1991), and the non-stationary processes are generated autoregressively, $x_t = x_{t-1} + \eta_t$. All series are generated with 100 additional values in order to obtain random starting values. All tests are one-sided with level 0.01, 0.05 and 0.1.

Tables 1 presents the empirical rejection probabilities for the 5% tests (results for 1% and 10% tests are quite similar and are available upon request by the authors). The most important results are: *i*) for fixed δ, the power of the two tests increases with the sample size (even if it remains low for values of $\delta > 0.5$, non-stationary case, and near one) reflecting thus the consistency of the ADF test according to the theoretical results of Krämer (1998); *ii*) for a fixed sample size, T, the power decreases with increasing δ: result that is not surprising since we are approaching the null hypothesis; *iii*) for all values of fractional parameter δ and for all sample size considered the power of the bootstrap ADF test is (almost) always higher than the power of simple ADF test, while the estimated size ($\delta = 1$) is fairly close to the nominal levels of significance.

Table 2 reports (for sake of completeness) the Monte Carlo means for the estimated autoregressive order p_{AIC} for each value of the fractional parameter and sample size.

TABLE 1

Estimate of power of adf tests

DGP: ARFIMA(0, δ, 0), level α = 0.05

T	*δ (d = δ-1)*	*0.35*	*0.45*	*0.6*	*0.7*	*0.8*	*0.9*	*1.0*
50	ADF	0.474	0.266	0.164	0.108	0.074	0.055	0.053
	ADF*	0.479	0.271	0.170	0.111	0.080	0.057	0.056
100	ADF	0.618	0.356	0.219	0.178	0.086	0.082	0.053
	ADF*	0.623	0.362	0.233	0.183	0.093	0.085	0.051
250	ADF	0.844	0.530	0.386	0.276	0.190	0.102	0.040
	ADF*	0.837	0.539	0.400	0.285	0.198	0.109	0.045
500	ADF	0.939	0.680	0.508	0.389	0.263	0.133	0.053
	ADF*	0.937	0.683	0.511	0.407	0.267	0.134	0.053

TABLE 2

Autoregressive order selection: P_{AIC}

T	***0.35***	***0.45***	***0.6***	***0.7***	***0.8***	***0.9***	***1.0***
50	3.094	2.737	2.215	1.765	1.267	0.92	0.641
100	5.104	4.359	3.517	2.658	1.942	1.217	0.842
250	0.853	7.728	5.835	4.767	3.226	1.769	0.766
500	11.366	11.024	8.436	6.470	4.429	2.432	0.960

Simulation 2. The size and power properties of the second class of fractional unit root tests we considered are explored. In this case the Monte Carlo experiment we conducted is slightly different. We consider, in fact, the following model:

$$(1 - B)^{\delta} x_t = \varepsilon_t , \quad t = 1, 2, \dots, T$$

with $\delta = 0.55, 0.7, 0.9, 1.0, 1.1, 1.3, 1.45$ and $\varepsilon_t \sim$ *i.i.d.* $N(0,1)$. Our results are based on 1000 independent replications with T = 100, 250 and 500 observations and for each simulated series we considered 500 bootstrap replications.

Tables 3 presents the empirical power of the tests at a nominal 5% level of significance. The simulation results suggest that in general, the bootstrap testing procedure is able to improve the tests, in fact: *i)* every bootstrap test has better size properties than any of the original test; *ii)* the bootstrap procedure does not impose distorsions where the original tests are well-sized; *iii)* the power properties suggest that the bootstrap LM test is superior when we are testing for unit root against fractional integration in small samples, in larger samples all the LM tests became equivalent.

Remark. To choose the number of bootstrap samples, in both the simulation studies, so as to minimize experimental randomness, we have followed the so called *p*-value approach, as described, for example in Davidson and MacKinnon (2000).

6. CONCLUSION

In this work we examine the performance of the sieve-wild bootstrap method to test the unit root hypothesis against fractional alternatives. The tests we considered are, firstly, a classical test for unit root, the Augmented Dickey-Fuller, in addition we considered a different class of tests usually used to determine if a series have long-memory. These tests are the modified rescaled range, the Geweke Porter-Hudack and a LM test.

TABLE 3

Estimate of power of unit root tests

DGP: ARFIMA(0, δ, 0), level $\alpha = 0.05$

T	δ ($d = \delta$-1)	***0.55***	***0.7***	***0.9***	*1.0*	***1.1***	***1.3***	***1.45***
100	MRS	0.290	0.242	0.116	0.730	0.040	0.042	0.024
	MRS*	0.132	0.118	0.046	0.060	0.110	0.146	0.070
	GPH	0.230	0.142	0.068	0.048	0.068	0.152	0.276
	GPH*	0.198	0.114	0.050	0.046	0.078	0.196	0.296
	LM	0.898	0.540	0.044	0.027	0.190	0.882	0.988
	LM*	0.980	0.802	0.164	0.050	0.220	0.906	0.986
250	MRS	0.672	0.486	0.166	0.065	0.114	0.232	0.120
	MRS*	0.478	0.314	0.096	0.054	0.176	0.332	0.188
	GPH	0.414	0.258	0.078	0.058	0.076	0.266	0.554
	GPH*	0.380	0.210	0.066	0.053	0.092	0.322	0.564
	LM	1.000	0.994	0.262	0.036	0.504	0.998	1.000
	LM*	1.000	0.998	0.398	0.051	0.526	0.998	1.000
500	MRS	0.906	0.700	0.172	0.058	0.142	0.428	0.384
	MRS*	0.814	0.554	0.116	0.052	0.210	0.502	0.406
	GPH	0.684	0.398	0.078	0.060	0.082	0.434	0.746
	GPH*	0.630	0.358	0.064	0.057	0.104	0.480	0.760
	LM	1.000	1.000	0.600	0.046	0.788	1.000	1.000
	LM*	1.000	1.000	0.692	0.052	0.790	1.000	1.000

We find out that the ADF bootstrap works generally better than the ADF, even if the power of the test is quite low especially if the data generating process is a non stationary fractional integrated process ($\delta > ½$). Anyway the ADF test seems to be consistent.

The MRS*, GPH* and LM* tests have nice size properties, however, because of his higher power, we suggest the use of LM* test.

We conclude that the bootstrap testing procedure provides a practical and effective method to improve tests when we are searching for a unit root against fractional alternatives.

For future research it would be interesting to analyse how this bootstrap approach can improve inference *i)* when the observed time series displays observations that are non-normally distributed and/or conditionally heteroskedastic, situations that can arise frequently in case of analysing financial time series data; *ii)* when we consider models such cointegrating regression and error correction models. Such extensions and applications are currently under way by the authors.

Dipartimento di Scienze Statistiche LUISA BISAGLIA
Università di Padova

Dipartimento di Statistica ISABELLA PROCIDANO
Università di Ca' Foscari, Venezia

REFERENCES

I.V. BASAWA, A.C. MALLIK, W.P. MCCORMIK, J.H. REEVES, R.L. TAYLOR (1991a), *Bootstrapping unstable first-order autoregressive processes*, "The Annals of Statistics", 19, 1098-1101.

I.V. BASAWA, A.C. MALLIK, W.P. MCCORMIK, J.H. REEVES, R.L. TAYLOR (1991b), *Bootstrap test of significance and sequential bootstrap estimation for unstable first order autoregressive process*, "Communications in Statistics. Theory and Methods", 20, 1015-1026.

P.J. BROCKWELL, R.A. DAVIS (1991), *Time series: theory and methods*, Springer, New York.

P. BÜHLMANN (1997), *Sieve bootstrap for time series*, "Bernoulli", 3, 123-148.

Y. CHANG, J.Y. PARK (1999), *A sieve bootstrap for the test of a unit root*, Working Paper No. 1999-14, RICE University.

Y.W. CHEUNG (1993), *Tests for fractional integration: a Monte Carlo investigation*, "Journal of Time Series Analysis", 14, 331-345.

S. DATTA (1996), *On asymptotic properties of bootstrap for AR(1) processes*, "Journal of Statistical Planning and Inference", 53, 361-374.

S. DATTA, T.N. SRIRAM (1997), *A modified bootstrap for autoregression without stationarity*, "Journal of Statistical Planning and Inference", 59, 19-30.

R. DAVIDSON, E. FLACHAIRE (2001), *The wild bootstrap, tamed at last*, Discussion Paper No. DARP 58, London School of Economics.

R. DAVIDSON, J.G. MACKINNON (2000), *Bootstrap tests: how many bootstrap?*, "Econometric Reviews", 19, 55-68.

F.X. DIEBOLD, G.D. RUDEBUSH (1991), *On the power of Dickey-Fuller tests against fractional alternatives*, "Economics Letters", 35, 155-160.

N. FERRETTI, J. ROMO (1996), *Unit root bootstrap tests for AR(1) models*, "Biometrika", 83, 849-860.

W.A. FULLER (1976), *Introduction to statistical time series*, Wiley, New York.

J. GEWEKE, S. PORTER-HUDACK (1983), *The estimation and application of long memory time series models,* "Journal of Time Series Analysis", 4, 221-238.

C.W.J. GRANGER, R. JOYEUX (1980), *An introduction to long-memory time series models and fractional differencing*, "Journal of Time Series Analysis", 1, 15-30.

W. KRÄMER (1998), *Fractional integration and the augmented Dickey-Fuller test*, "Economics Letters", 61, 269-272.

J.P. KREISS (1992), *Bootstrap procedures for AR(∞)-process*, in K.H. Jöckel, G. Rothe and W. Sender (eds.), "Bootstrapping and related techniques", Lecture notes in economics and mathematical systems 376, Springer: Heidelberg.

U. HASSLER, J. WOLTERS (1994), *On the power of unit root test against fractional alternatives*, "Economics Letters", 45, 1-5.

J.R.M. HOSKING (1981), *Fractional differencing*, "Biometrika", 68, 165-176.

H.E. HURST (1951), *Long-term storage capacity of reservoir,* Transactions of the American Society of Civil Engineers, 116, 770-799.

R.Y. LIU (1988), *Bootstrap procedures under some non i.i.d. models*, "Annals of Statistics", 16, 1697-1708.

A.W. LO (1991), *Long-term memory in stock market prices,* "Econometrica", 59, 1279-1313.

I.N. LOBATO, P.M. ROBINSON (1998), *A non parametric test for I(0),* "Review of Economic Studies", 65, 475-495.

J.C. MACKINNON (1991), *Critical values for cointegration tests*, in R.F. Engle and C.W.J. Granger (eds.), "Long run economic relationships", Oxford Univeristy Press, 267-276.

C. PIZZI, I. PROCIDANO, S. RIGATTI LUCHINI, S. (2001), *Un test bootstrap robusto per radici unitarie in presenza di outlier e non linearità*, in "Modelli complessi e metodi computazionali intensivi per la stima e la previsione", a cura di Corrado Provasi, Cleup editrice, 271-276.

I. PROCIDANO, S. RIGATTI LUCHINI (2001), *Testing unit roots by bootstrap*, Metron, 60, 175-189.

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

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

F. SOWELL (1990), *The fractional unit root distribution*, "Econometrica", 58, 495-505.

C.F.J. WU (1986), *Jacknife, bootstrap, and other resampling methods in regression analysis* (with discussion), "Annals of Statistics", 14, 1261-1295.

RIASSUNTO

Un approccio bootstrap per migliorare la potenza dei test di radici unitarie contro alternative frazionarie

In questo lavoro vedremo come utilizzare alcuni test asintotici di integrazione frazionaria per testare l'ipotesi nulla di radici unitarie contro alternative frazionarie. Dal momento che tali test risultano distorti in piccoli campioni, anche sotto l'ipotesi di Gaussianità, utilizzeremo una particolare metodologia *bootstrap* per migliorarne il comportamento. I risultati mostrano che l'approccio *bootstrap* migliora sia la distorsione sia la potenza dei test considerati.

SUMMARY

Improving the power of unit root tests against fractional alternatives using bootstrap

In this paper asymptotic tests for unit root hypothesis against fractional alternatives are considered. Since they are generally badly sized in small samples, even for normally distributed processes, we consider a new bootstrap approach to correct such size distorsion. We find out that the bootstrap approach always improves the size and the power of the considered tests.