A Sieve Bootstrap-based Test for Multiple Unit Roots

Xiao Zhong, V. A. Samaranayake
Department of Mathematics & Statistics, Missouri University of Science & Technology, Rolla, MO 65401, USA

Abstract:
We introduce a sieve bootstrap based test for determining if an autoregressive process has two unit roots. This is in contrast to the standard procedure of determining the number of unit roots by first conducting a unit root test, then differencing the series if the null hypothesis of a unit root process is not rejected and repeating the unit root test on the differenced series. It's shown how the proposed procedure can easily be extended to the case of three or more unit roots. In addition, we develop a bootstrap version of a sequential test proposed by Dickey and Pantula in 1987. A Monte Carlo simulation study is carried out to investigate the properties of the two tests. Results show that the sieve bootstrap based test has reasonable small sample properties when the sample size is large and that the Dickey-Pantula test has good power.

Key Words:
Integrated processes; Unit root tests; Multiple unit roots; Sieve Bootstrap; ARIMA

1. Introduction

When modeling empirical time series, it is sometimes necessary to perform unit root tests. One reason for carrying out such tests is to determine if the time series needs differencing to obtain stationarity. More importantly, unit root tests have been applied in the investigation of certain economic hypotheses. For example, Altonji and Ashenfelter (1980) used unit root tests to test an equilibrium hypothesis for wage movements; Nelson and Plosser (1982) applied unit root tests to describe the effect of monetary disturbances on macroeconomic series; Meese and Singleton (1982) explained the importance of unit root testing in the theory of linearized expectations by applying unit root tests to exchange rates. Also, over the last three decades, the unit root tests have drawn more and more attention in many research fields related to economics. In particular, such tests can imply whether or not the shocks (ε_t) to an economic system have a permanent effect on the future econometric pattern. Specifically, if at least one unit root exists, then each shock does have a permanent impact on the future forecasts; otherwise, the impact could be negligible. For more details, see J. Franke et al. (2011, p. 244).

Practically, the existence of unit roots is often suspected by visual inspection of the autocorrelation function (ACF) and data plots. As long as the ACF decays slowly, the time series should be considered having at least one unit root and the operation of differencing the time series may be performed repeatedly to obtain a stationary time series. Many statistical tests for unit roots, including ours, are based on auto-regression tests of linear dependence. On one hand, they simply mitigate the subjectivity of visual inspection of autoregression plots; on the other hand, compared to visual inspection, these tests are more helpful in deciding close-call situations. The most commonly used unit root tests were developed by Dickey and Fuller (1979, 1981) and sometimes referred to as Dickey-Fuller tests. Dickey-Fuller tests assumed that the series have at most one unit root. If there are more than one unit roots, a sequence of Dickey-Fuller tests may be applied to the raw series and the differenced series repeatedly. Intuitively, we expect that if there are more than one unit root, the test for one unit root will strongly indicate that the process needs to be differenced. Hence we expect that the hypothesis of one unit root will be rejected (and the hypothesis of no unit root will be favored) less than 5% of the time when there are more than one unit root present. However, a simulation study done by Dickey and Pantula (1987) doesn't support that intuition. Moreover, Sen (1985) showed that if there are actually two unit roots, then the method of applying Dickey-Fuller tests on the raw and the differenced series repeatedly is not valid. As matter of fact, since the Dickey-Fuller test is based on the assumption of at most one unit root, at least the first few tests in this sequence cannot be theoretically justified. In order to mitigate these problems and perform tests based on a sound theoretical foundation, Dickey and Pantula (1987) proposed a strategy of performing the sequence tests in a different order. In their paper, they propose a method for sequential testing of unit roots. These tests compare a null hypothesis of d unit roots with an alternative of d-1 unit roots. Specifically, one starts with the largest d to test and work down if the null hypothesis of having d unit roots is rejected. The sequential testing procedure stops when a null hypothesis cannot be rejected. This test is recognized for its simplicity (it uses existing t tables) and high power.

Dickey-Fuller tests are based on finite-order auto-regressions, $AR(p)$, where the parameter p represents the order of the autoregressive time series and is assumed to be known. However, in general, p is unknown. Phillips (1987) and Phillips and Perron (1988) modified the Dickey-Fuller tests for the case where p is unknown; Said and Dickey (1984) used auto-regressions augmented with lagged differences to obtain the Augmented Dickey-Fuller tests (ADF). They showed that these tests are valid for all finite ARMA procedures with unknown orders if we increase the number of lagged differences appropriately as the sample size grows. These unit root tests are more useful than the tests that assume p is known, in practical applications. However, some researchers such as Leybourne and Newbold (1999) found that these unit root tests have serious size distortion and low power issues in finite samples, especially when the model has a moving average component. Subsequently, bootstrap and sieve bootstrap methods have been introduced to such modified Dickey-Fuller tests to improve their finite sample performance.

Basawa et al. (1991a, 1991b) applied a bootstrap process to $AR(1)$ unit root tests and showed that the unit root must be imposed on the generation of bootstrap samples to achieve consistency of the bootstrap unit root tests. Ferretti and Romo (1996) and Datta (1996) also made their contributions to such tests. If the bootstrap procedure is based on a sieve which is an approximation of an infinite dimensional model by a sequence of finite dimensional models, we get the sieve bootstrap procedure introduced by Bühlmann (1997). Specifically, we can approximate any linear process such as AR, MA or $ARMA$ by a finite $AR(\hat{p})$ where $\hat{p}$ increases with the sample size; and resample from the approximated auto-regressions. Chang and Park (2000) considered a sieve bootstrap for the test of a unit root in models driven by general linear processes. Their bootstrapped versions of ADF unit root tests are shown to be consistent under very general conditions and the asymptotic validity of such tests are thus established. Significant improvements on finite sample performance of the tests are also established by Monte Carlo simulations.

In this paper, two different approaches based on the bootstrap for multiple unit root tests are investigated when the underlying model is driven by general linear processes. First, Dickey and Pantula's sequential tests of multiple unit roots are adopted with the modification that the bootstrap is applied to obtain the critical values for multiple unit root tests. Second, a natural test of multiple unit roots based on the DF method and the sieve bootstrap is proposed. The rest of the paper is organized as follows. Section 2 introduces Dickey and Pantulas' tests and presents their asymptotic theories. The bootstrap versions of their tests are described in Section 3. In Section 4, ADF tests and their asymptotic theories are explained, and the sieve bootstrap for ADF multiple unit root tests are explored. In Section 5, the Monte Carlo studies for both methods are presented, and a comparison is given. Section 6 is a summary.

2. Dickey-Pantula's Sequential Tests and Their Limiting Distributions

Assume the time series $\{X_t\}$ satisfy

$$X_t = \sum_{j=1}^{p} \alpha_j X_{t-j} + e_{t}, \quad t = 1, 2, \ldots, \qquad (2.1)$$

Where $\{e_t\}$ is a sequence of iid random variables with mean 0 and variance 1. To make the presentation simple, p is restricted to 3. Extensions for $p > 3$ are easily carried out.

Now, let m_1, m_2 and m_3 represent the roots of the characteristic equation

$$m^3 - \alpha_1 m^2 - \alpha_2 m - \alpha_3 = 0. \qquad (2.2)$$

Assume that $|m_1| \geq |m_2| \geq |m_3|$. Consider the following four hypotheses: H_0: $|m_1| < 1$; H_1: $m_1 = 1, |m_2| < 1$; H_2: $m_1 = 1$, $m_2 = 1, |m_3| < 1$; H_3: $m_1 = m_2 = m_3 = 1$. That is, under H_d, $d = 0, 1, 2, 3$, there are d unit roots. After a re-parameterization of model (2.1), we can write

$$W_t = \theta_1 X_{t-1} + \theta_2 Y_{t-1} + \theta_3 Z_{t-1} + e_t, \tag{2.3}$$

where $Y_t = X_t - X_{t-1}$, $Z_t = Y_t - Y_{t-1}$, $W_t = Z_t - Z_{t-1}$, and the $H_d's$ are transformed into: H_3: $\theta_1 = \theta_2 = \theta_3 = 0$; H_2: $\theta_1 = \theta_2 = 0$, $\theta_3 < 0$; H_1: $\theta_1 = 0$, $\theta_2 < 0$, $\theta_3 < 0$; H_0: $\theta_1 < 0$, $\theta_2 < 0$, $\theta_3 < 0$. The reparameterization is useful because now we can use the usual regression tests for the thetas in (2.3).

The procedure proceeds as follows: perform a regression of W_t over X_{t-1}, Y_{t-1} and Z_{t-1} to get the least squares estimates $\hat{\theta}_i$ and the corresponding t-statistics $t_{i,n}(3)$, $i = 1, 2, 3$, where $t_{i,n} = \frac{\hat{\theta}_i}{s(\hat{\theta}_i)}$, n denotes the sample size, and $s(\hat{\theta}_i)$ is the standard error of $\hat{\theta}_i$ obtained from the regression.

Now, a sequential testing procedure is considered. We test the null hypothesis H_3 against the alternative hypothesis H_2 first by considering the t-statistic $t^*_{3,n}(3)$ obtained by regression of W_t on Z_{t-1}. Then, we can test the null hypothesis H_2 against the alternative hypothesis H_1 by considering the t-statistic $t^*_{2,n}(3)$ obtained by regression of W_t on Y_{t-1} and Z_{t-1}. Moreover, let $t^*_{1,n}(3) = t_{1,n}(3)$.

Pantula (1987, 1986b) proved that the asymptotic distributions of the $t^*_{d,n}$ statistics under H_d for $d = 1, 2, 3$ are certain functionals of a standard Brownian motion.

In summary, Dickey and Pantula proposed the following sequential procedure for testing the hypotheses:

1) Reject H_3 of three unit roots and go to step 2) if $t^*_{3,n}(3) \leq \hat{\tau}_{n,\alpha}$, where $\hat{\tau}_{n,\alpha}$ was given by Fuller (1976).

2) Reject H_2 of two unit roots and go to step 3) if $t^*_{2,n}(3) \leq \hat{\tau}_{n,\alpha}$, where $\hat{\tau}_{n,\alpha}$ was given by Fuller (1976).

3) Reject H_1 of one unit root in favor of H_0 of no unit roots if $t^*_{1,n}(3) \leq \hat{\tau}_{n,\alpha}$, where $\hat{\tau}_{n,\alpha}$ was given by Fuller (1976).

Note that these critical points are not available for all significance levels and sample sizes. Therefore, the bootstrap-based critical points may be an alternative.

3. The Bootstrap Dickey-Pantula Tests

In this section, we modify the Dickey-Pantula test in two ways. First we obtain the critical points using the bootstrap. Second, we observe that the Dickey-Pantula method requires first testing H_p where p is the order of the autoregressive process, even when it is reasonable to assume that the number of unit roots is less than p. Hence we modify their method to accommodate such cases by starting the sequential testing at a value of $d \leq p$.

Let's assume $p = 3$ for the simplicity of explanation, and the maximum number of unit roots, d, is 2. Extension to other values of p and d can be done easily.

Define $\{W_t\}$ as the third difference of $\{X_t\}$, $\{Z_t\}$ as the second difference of $\{X_t\}$, $\{Y_t\}$ as the first difference of $\{X_t\}$, where $t = 1, 2, \ldots, n$. Then the transformed model is

$$W_t = \theta_1 X_{t-1} + \theta_2 Y_{t-1} + \theta_3 Z_{t-1} + e_t \ . \tag{3.1}$$

The four hypotheses in terms of θs are the same as in Section 2. Now,

1) To get $t^*_{2,n}(3)$, fit the regression model:

$$W_t = \theta_2 Y_{t-1} + \theta_3 Z_{t-1} + e_t \ . \tag{3.2}$$

Then let $t = t^*_{2,n}(3) = \frac{\hat{\theta}_2}{s(\hat{\theta}_2)}$.

2) Now, fit the model (under the null hypothesis):

$$W_t = \theta_3 Z_{t-1} + e_t \ . \tag{3.3}$$

Obtain all the centered residuals:

$$\text{Res}(i) = \hat{e}_i - \bar{\hat{e}},$$

where

$$\hat{e}_i = W_i - \widehat{W}_i \ , \quad i = 1, 2, \ldots, n.$$

3) Sample with replacement from all the centered residuals to get the bootstrap sample of errors, $\{e_t^{bt}\}_{t=1}^{n+50}$.

4) Then we can get the bootstrap samples: $\{Z_t^{bt}\}_{t=1}^{n+50}$, $\{W_t^{bt}\}_{t=1}^{n+50}$, $\{Y_t^{bt}\}_{t=1}^{n+50}$, $\{X_t^{bt}\}_{t=1}^{n+50}$ easily by using the recursive equations

$$Z_t^{bt} = \left(1+\hat{\theta}_3\right)Z_{t-1}^{bt} + e_t^{bt} \quad ,$$

$$Z_0^{bt} = 0,$$

$$W_t^{bt} = Z_t^{bt} - Z_{t-1}^{bt} \quad ,$$

$$Y_t^{bt} = \sum_{j=1}^{t} Z_j^{bt} \quad ,$$

$$X_t^{bt} = \sum_{j=1}^{t} Y_j^{bt} \quad . \tag{3.4}$$

5) Do the regression of (3.2) with the bootstrap samples obtained in 4) and calculate the bootstrap t, t^{bt}, as in 1).

6) Repeat 2) ~ 5) many times (e.g., 2,000 times) and determine the critical value τ^{bt} which is the 5th percentile of the 2,000 t^{bt} values.

7) If the t from 1) is less than τ^{bt}, then reject the null hypothesis of two unit roots; otherwise, don't reject.

To estimate coverage probabilities associated with this method, we repeated Steps 1) to 7) 2,000 times. The results are summarized in Section 6.

Note that the above procedure can be modified to test H_1 as well.

4. The ADF Tests and the Sieve Bootstrap Tests for Two Unit Roots

In this section, we assume the time series $\{X_t\}$ be given by:

$$X_t = \alpha X_{t-1} + u_t \tag{4.1}$$

where $u_t = \pi(L)\varepsilon_t$, is an infinite moving average process, and $\{X_t\}$ can be approximated by a finite auto regression $AR(p)$, with the order p increasing as the sample size n grows. We will first discuss the ADF test to show the parallel between it and the proposed test for two unit roots.

The ADF test of the unit root hypothesis for the time series $\{X_t\}$ given above can now be based on the regression

$$X_t = \alpha X_{t-1} + \sum_{k=1}^{p-2} \alpha_k \nabla X_{t-k} + \varepsilon_{p-2,t} \quad , \tag{4.2}$$

where p is estimated using the AICC criteria.

Under the null hypothesis of 1 unit root, $\alpha = 1$. And the t-statistic considered here is:

$$T_n = \frac{\widehat{\alpha}_n - 1}{s(\widehat{\alpha}_{n})} \quad , \tag{4.3}$$

which is the extension of the t-test given by Dickey and Fuller (1979, 1981) for the $AR(1)$ model.

Under certain assumptions, the asymptotic distribution of the chosen t-test statistic is a functional of Brownian motion. A proof can be found in Chang and Park (2001).

In order to apply the sieve bootstrap to the above test, first, we approximate the time series by an $AR(\hat{p})$ process where $\hat{p}$ is chosen among $\{2, 3, \dots, pmax\}$ to minimize AICC. The value $pmax$ is appropriately chosen so that $pmax \to \infty$ as $n \to \infty$. Now as Chang and Park (2001) showed, the bootstrap version of the ADF test can be constructed by constructing bootstrap samples of $\{X_t\}$ obtained under the null hypothesis.

Now, we will introduce our test for two unit roots. Again, for the simplicity, we assume $\hat{p}$ = 3 and the maximum number of unit roots d is 2. The extensions to other values of $\hat{p}$ and d can be carried out easily. Then, follow the following steps:

1) Fit

$$X_t = \rho_1 X_{t-1} + \rho_2 \nabla X_{t-1} + \sum_{k=1}^{p-2} \alpha_k \nabla^2 X_{t-k} \tag{4.4}$$

to the data and calculate the t-statistic

$$T_{2,n} = \frac{\widehat{\rho}_2 - 1}{s(\widehat{\rho}_2)} \quad . \tag{4.5}$$

2) Fit the model

$$X_t = X_{t-1} + \nabla X_{t-1} + \sum_{k=1}^{p-2} \alpha_k \nabla^2 X_{t-k} + \varepsilon_t \tag{4.6}$$

and compute the estimates $\hat{\alpha}_k, k = 1, 2, \dots, p-2,$ and determine the residuals $\hat{\varepsilon}_t$.

3) Now, sample with replacement $\{\varepsilon_t^{bt}\}$ from $\widetilde{\varepsilon_t} = \hat{\varepsilon}_t - \bar{\hat{\varepsilon}}_t$.

4) Use $\{\varepsilon_t^{bt}\}$ to generate $\{X_t^{bt}\}$ by using (4.6) recursively, with α_k replaced by $\hat{\alpha}_k$ and ε_t replaced by ε_t^{bt}. Note that we set $X_t = 0$ for all $t \le 0$.

5) Fit (4.4) to $\{X_t^{bt}\}$ and calculate the bootstrap t-statistic $T_{2,n}^{bt}$ using (4.5), where the bootstrap estimates $\hat{\rho}_2^{bt}$ and $s(\hat{\rho}_2^{bt})$ replace $\hat{\rho}_2$ and $s(\hat{\rho}_2)$.

6) Repeat Step 2) ~ Step 5) many times (e.g., 2,000 times), and calculate τ^{bt}, the lower 5th percentile of the bootstrap t-statistic $T_{2,n}^{bt}$. If $T_{2,n} < \tau^{bt}$, reject the null hypothesis of 2 unit roots; otherwise, don't reject.

In order to estimate the coverage probabilities, we repeated Step 1) ~ Step 6) 2,000 times and we computed the significance level and the power based on the proportion of rejection of all the Monte Carlo simulations. The results are summarized in Section 5 as well.

5. Simulation Results and Comparison

In order to determine the finite sample properties of these tests we carried out the following setting of experiments: n = 50 and 100; p = 3; d = 2; the number of Monte Carlo simulations = 2,000; the number of bootstrap samples = 2,000; $pmax$ = 10; $\hat{p} \in$ [2, 10]. The simulation results are attached in Table 1 on next page.

As seen from the preliminary results listed in Table 1, the bootstrap ADF type tests for two unit roots have some difficulty maintaining their size and show reasonable power. On the other hand, the bootstrap version of the Dickey-Pantula test is good at maintaining the size and show good power. One reason for the relatively lower power of the ADF type tests for two unit roots may be the fact that the order p is estimated using AICC while the Dickey-Pantula version uses the exact value of p and hence has an advantage. Thus the comparison between the two tests may not be a fair one.

Also note that the ADF type test for two unit roots does not perform well when the sample size is 50 but improves its power when the sample size increases to 100. Even then, it shows very low power when one of the roots is 0.9; while the bootstrap version of Dickey-Pantula test also has low power when a root is close to unity, the power does not dip as low as the other test.

An issue that needs further investigation is the fact that the significance level of the ADF type test for two unit roots decreases when the 3rd root decreases. This is somewhat counterintuitive. It is possible that the estimate of ρ_2 is affected by the 3rd root. Additional investigations on this matter as well as other modifications that may yield additional power are being presently carried out.

Table 1: The comparison of results between sieve bootstrap ADF tests and sieve bootstrap Dickey-Pantula tests

n	r_1	r_2	r_3	sig level	power
50	1	1	0.8	0.0655 (0.0610)	
50	1	1	0.2	0.0225 (0.0630)	
50	1	0.8	0.2		0.0855 (0.2860)
50	1	0.5	0.2		0.4785 (0.8200)
50	1	0.2	0.2		0.8990 (0.9760)
50	1	0.9	0.5		0.0365 (0.1155)
50	1	0.8	0.5		0.0555 (0.2385)
50	0.9	0.9	0.9		0.0490 (0.1205)
50	0.9	0.9	0.5		0.0630 (0.2110)
50	0.9	0.9	0.2		0.0660 (0.2380)
50	0.9	0.5	0.2		0.4945 (0.9200)
50	0.9	0.1	0.2		0.8480 (0.9950)
100	1	1	0.8	0.0715 (0.0595)	
100	1	1	0.2	0.0215 (0.0530)	
100	1	0.8	0.2		0.3295 (0.7705)
100	1	0.5	0.2		0.9630 (1.0000)
100	1	0.2	0.2		0.9600 (1.0000)
100	1	0.9	0.5		0.0640 (0.2780)
100	1	0.8	0.5		0.1855 (0.6630)
100	0.9	0.9	0.9		0.0905 (0.3100)
100	0.9	0.9	0.5		0.2275 (0.6605)
100	0.9	0.9	0.2		0.2495 (0.7405)
100	0.9	0.5	0.2		0.9195 (1.0000)
100	0.9	0.1	0.2		0.9285 (1.0000)

* *Results for the sieve bootstrap Dickey-Pantula tests are given in parenthesis.*
r_1,r_2, r_3 *are the three roots of the time series*

6. Summary

Testing for two unit roots in a time series has not received as much attention as the case of testing for one unit root. The only procedure that tests for two unit roots using a single test was proposed by Dickey and Pantula in 1987. This test requires taking *p* differences of the time series where *p* is the order of the autoregressive process. We modify this test so that the percentile points are derived using the bootstrap. In addition, a new test for two unit roots, which is an extension of the Augmented Dickey-Fuller test formulation for a single unit root case to two unit roots, was developed. Preliminary results show that this test shows promise but the bootstrap version of the Dickey-Fuller test is superior.

References

Altonji, J., and Ashenfelter, O. (1980), "Wage Movements and the Labor Market Equilibrium Hypothesis," *Economica*, 47, 217-245.

Basawa, I. V., Mallik, A. K., McCormick, W. P., Reeves, J. H. and Taylor, R. L (1991a), "Bootstrapping Unstable First-Order Autoregressive Processes," *Annals of Statistics*, 19, 1098-1101.

Bühlmann, P. (1997), "Sieve Bootstrap for Time Series," *Bernoulli*, 3, 123-148.

Chang, Park. (2000), "A Sieve Bootstrap for the Test of A Unit Root," *Journal of Time Series Analysis*, 24, 379-400.

Dickey, D. A. and Fuller, W. A. (1979), "Distribution of Estimators for Autoregressive Time Series with A Unit Root," *Journal of American Statistics Association*, 74, 427-431.

Datta, S. (1996), "On Asymptotic Properties of Bootstrap for AR(1) Processess," *Journal of Statistical Planning and Inference,* 53, 361-374.

Dickey, Pantula. (1987), "Determining the Order of Differencing in Autoregressive Processes," *Journal of Business and Economic Statistics*, 5, 455-461.

Ferretti, N. and Romo, J. (1996), "Unit Root Bootstrap Tests for AR(1) Models," *Biometrika,* 83, 849-860.

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

J. Franke et al. (2011), *Statistics of Financial Markets: An Introduction*, Universitext, Springer-Verlag Berlin Heidelberg.

Leybourne, S. and Newbold, P. (1999), "On the Size Properties of Phillips-Perron Tests," *Journal of Time Series Analysis,* 20, 51-61.

Meese, R. A., and Singleton, K. J. (1982), "On Unit Roots and the Empirical Modeling of Exchange Rates," *Journal of Finance*, 37, 1029-1034.

Nelson, C. R., and Plosser, C. I. (1982), "Trends and Random Walks in Macroeconomic Time Series," *Journal of Monetary Economics*, 9, 139-162.

Pantula, S. G. (1986b), “Determining the Order of Differencing in Autoregressive Processes,” unpublished manuscript, North Carolina State University, Dept. of Statistics.

Phillips, P. C. B. (1987), “Time Series Regression with a Unit Root,” *Econometrica,* 55, 277-301.

Phillips, P. C. B. and Perron, P. (1988), “Testing for a Unit Root in Time Series Regression,” *Biometrika* 75, 335-346.

Said, S. E. and Dickey, D. A. (1984), “Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order,” *Biometrika* 71, 599-608.

Sen, D. L. (1985), “Robustness of Single Unit Root Test Statistics in the Presence of Multiple Unit Roots,” unpublished Ph.D. dissertation, North Carolina State University, Dept. of Statistics.