«ΣΠΟΥΔΑΙ», Τόμος 49, Τεύχος 1ο-4ο, Πανεπιστήμιο Πειραιώς / «SPOUDAI»,Vol. 49, No 1-4, University of Piraeus

A SELECTIVE REVIEW ON THE ISSUE OF TESTING FOR A UNIT AUTOREGRESSIVE ROOT

By

Christos Agiakloglou
Department of Economics, University of Piraeus, Greece

Abstract

In the recent few years an increasing effort has been made to establish reliable testing procedures to determine whether or not an observed time series is generated by a unit autoregressive root process. This paper presents in a selective manner some of the most common and widely used test statistics for testing for a unit autoregressive root and evaluates the performance of these test statistics in moderately large samples. (JEL Clas.: C120, C150, C220)

1. Introduction

Time series analysis has been extensively used by several sciences to explain the behavior of many physical, engineering and economic phenomena generated by random processes. In this framework, an objective is to correctly specify the process through which an observed time series is generated so that, for example, inference about future behavior can be made.

According to the methodology of Box and Jenkins (1976), autoregressive integrated moving average models, known as ARIMA models, are fitted to the data following the identification, estimation and diagnostic check stages of model estimation. As part of the identification stage an effort is made to determine whether the observed time series is generated by a stationary or non-stationary process.

Consider for example the ARIMA (p, d, q,) process

$$\Phi(B)\ (1-B)^d X_t = \Theta(B)\varepsilon_t \tag{1.1}$$

where B is the backshift operator, p and q are the orders of the polynomials $\Phi(B)$

and $\Theta(B)$ respectively and the roots of the polynomial equations in ζ, $\Phi(\zeta)=0$, and $\Theta(\zeta) = 0$, are outside the unit root circle $|z|= 1$. An important issue in building models of the form (1.1) is determining an appropriate value of d, where d according to Box and Jenkink, methodology is a non-negative integer number.

The process (1.1) is stationary if d is zero. However, if the process is non-stationary, then, in the presence of a unit autoregressive root, it can be converted to a stationary and invertible process by first differencing assuming that there is only one unit autoregressive root. By transforming a non-stationary process in this way other characteristics of the generating model are more easily uncovered. This determination as to whether or not a series should be differenced is known as the unit root test and d is called the number of unit autoregressive roots.

It should be emphasized however that to determine whether an observed time series is stationary or non-stationary is not always the main concern in time series analysis. Think for example a stationary AR(1) process with a value of the autoregressive parameter close to one and a non-stationary random walk process. If the objective is to select one of the above models in order to use it only for making inference about future behavior then, in this case, it does not make any difference which one of these two models will be chosen since both models will generate very similar, if not identical, forecasts.

To the contrary, the unit root issue is really very important in time series analysis if the objective is to estimate or to built a particular econometric mode. As it is known, the presence of a unit autoregressive root effects the estimates and therefore the test statistics (see for example Park and Phillips (1988, 1989) and Sims, Stock and Watson (1990) as well as the structure of the model in terms of applying the economic theory through time series analysis (see for example Granger and Newbold (1974, 1986), Harvey (1985, 1989), Engle and Granger (1987) and Johansen (1991)).

The objective of this paper is to present in a selective manner, although the literature on unit root tests is vast, and to evaluate the performance of the most common and widely used test statistics for testing for a unit autoregressive root in moderately large samples. Section two introduces the Dickey-Fuller unit root test for the simple random walk process and for the AR(p) process. Sections three and four present the Dickey-Fuller type tests for ARIMA (p, 1, q) models based on the least squares estimation when the orders ρ and q are either unknown or known respectively. Section five displays the unit root test based on the sample autocorrelations and, lastly, section six examines the class of fractionally integrated ARMA models, known as ARFIMA models, that nests the unit root phenomenon as a special case and in which the value of d is assumed to be a real number.

2. The Dickey-Fuller unit root test

Consider the simple first-order autoregressive process

$$X_t = \varphi X_{t-1} + \varepsilon_t \tag{2.1}$$

where e_t is white noise and the mean of the process is taken to be zero. If $|\varphi)<1$ then the process (2.1) is stationary whereas if $|\varphi|$ then the process (2.1) is non-stationary. For the special case where $\varphi=1$ the process (2.1) becomes a random walk, i.e.,

$$\Delta X_t = X_t - X_{t-1} = \varepsilon_t \tag{2.2}$$

where Δ is the first difference operator.

Getting an estimate of φ from model (2.1) is very simple. For example, given T observations, the least squares estimator of φ is

$$\hat{\varphi} = \frac{\sum_{t=2}^{T} X_t X_{t-1}}{\sum_{t=2}^{T} X_{t-1}^2} \tag{2.3}$$

and the distributional behavior of $\hat{\varphi}$ depends on the value of the true φ.

As it is known, for $|\varphi|<1$, $T^{1/2}(\hat{\varphi}-\varphi)$ is asymptotically normal and for $|\varphi|>1$ the limiting distribution of $|\varphi|^T(\hat{\varphi}-\varphi)(\varphi^2-1)^{-1}$ is Caushy. For $|\varphi| = 1$, $T(\hat{\varphi}-1)$ has a non-degenerate asympototic distribution and critical values for the usual t-statistic for testing the null hypothesis that $\varphi=1$ have been tabulated by Fuller (1976).

In essence, Fuller (1976) and Dickey & Fuller (1979) derive the distributions of the unit root test statistic for the first-order autoregressive process (2.1) when the generating model is a random walk. The unit root test for an observed time series X_t, t=1, 2, ..., T. is based on the regression of X_t on its one period lagged value X_{t-1} in three different ways:

a) with no mean

$$X_t = \beta X_{t-1} + \varepsilon_t \tag{2.4}$$

b) with mean

$$X_t = \alpha + \beta X_{t-1} + \varepsilon_t \tag{2.5}$$

c) with mean and a linear time trend

$$X_t = \alpha + \beta X_{t-1} + \delta t + \varepsilon_t \tag{2.6}$$

where e_t are identically and independently distributed as $N(0,\sigma^2)$ random variables.

The test for testing the null hypothesis that $\beta=1$ is conducted either by using the so-called $\hat{\rho}$ test or by using the $\hat{\tau}$ test. The $\hat{\rho}$ test is based on the statistic $T(\hat{\beta}-1)$ where $\hat{\beta}$ is the ordinary least squares estimator of β and the $\hat{\tau}$ test is based on the usual t-statistic for testing $\beta=1$, i.e., $\hat{\tau}= (\hat{\beta}-1)/ \text{se}(\hat{\beta})$ where se $(\hat{\beta})$ is the least squares standard error of $\hat{\beta}$. Critical values for these tests are reported by Fuller (1976).

The Dickey-Fuller tests presented so far are based on the usual t-statistic for testing only for a unit autoregressive root, i.e., testing the null hypothesis that $\beta=1$. However in the case where the objective is to test joint hypotheses the resulting test is an F-test. Dickey and Fuller (1981) derive the distribution of a class of tests based on equations (2.4) through (2.6) for testing joint hypotheses for the following three cases: i) H_0: $\beta=0$ and $\alpha=0$, ii) H_0 : $\beta=0$, $\alpha=0$ and $\delta=0$ and iii) H_0: $\beta=0$ and $\delta=0$ and critical values for these F-tests are tabulated and presented in Dickey and Fuller (1981).

Furthermore the Dickey-Fuller test can also be applied to high-order autoregressive processes as long as the autoregressive order is known. As shown by Fuller (1976) the test for a unit autoregressive root can be extended to a test of an ARIMA (p, 1, 0) process against an ARIMA (p+1, 0, 0) process when the order of the autoiregressive parameter p is known. In order words, for an AR(p) process the Dickey-Fuller test is based on the following regression

$$\Delta X_t = \alpha + \gamma X_{t-1} + \sum_{j=1}^{k} \zeta_j \Delta X_{t-j} + \varepsilon_t \tag{2.7}$$

where ε_t is taken to be white noise, k=p and Δ is the first difference operator. Testing for a unit autoregressive root is equivalent to testing the null hypothesis that γ=0. The $\hat{\rho}$ test will then based on the statistic T($\hat{\gamma}$), where $\hat{\gamma}$ is the least squares estimator of γ, and the $\hat{\tau}$ test will based on the usual t-statistic for testing the null hypothesis that γ=0. A linear time trend term, as in the first-order autoregressive process, is sometimes included in equation (2.7).

The test for a unit autoregressive root based on equation (2.7) is known as the "augmented" Dickey-Fuller test and critical values for this test statistic are presented in Fuller (1976).

Two very important points must be clarified at this stage. The first point has to do with the structure of the error term ε_t in all of the above equations. According to the unit root test the error term ε_t in all equations (2.4) through (2.7) must be a white noise otherwise the test is not valid. The second point has to do with the characterization of the test as a unit root test when this test is specifically applied to models (2.4) through (2.6). For these particular models the unit root test is not just a test for a unit autoregressive root but it is a test for a particular model structure that contains a unit autoregressive root, i.e., a random walk process with or without a drift or a linear time trend, and, therefore, it should not be refer to as the unit root test in the general manner.

Finally, Evans and Savin (1981, 1984) and Nankervis and Savin (1985) examine in detail the power of tests of a random walk process against a first-order autoregressive process.

3. The unit root test based on least squares estimation for ARIMA (p, 1, q) models of unknown orders

The distribution of the Dickey-Fuller tests, as previously indicated, is based on the assumption that the structure of the error term is white noise. However, since most macroeconomic variables are generated by processes having moving average components, the assumption of the error term being a white noise process is very restrictive. Therefore if this assumption is violated the unit root test is not any longer valid.

To eliminate the problem of time series generating models with moving average components Said and Dickey (1984) have initially derived the limit distribution of the test for testing for a unit autoregressive root for ARMA (p, q) models when the orders ρ and q are unknown. According to Said and Dickey

(1984) any ARIMA (p, 1, q) process can be approximated by a sufficient long autoregressive process, ARIMA (k, 0, 0), and the unit root test is based on the "augmented" Dickey-Fuller regression (2.7), where the order of the approximated autoregression k is defined as a function of the sample size T, i.e., $k=o\,(T^{1/3})$.

To see how the methodology of Said and Dickey works consider the simple model

$$X_t = \varphi X_{t-1} + z_t \tag{3.1}$$

$$z_t = \alpha z_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \tag{3.2}$$

where $|\alpha|<1$, $|\theta|<1$ and $X_0=0$. If $|\varphi|<1$ then the process X_t, generated by equations (3.1) and (3.2), is stationary and the model is an ARMA (2, 1) process. If on the other hand $\varphi=1$ then the model is an ARIMA (1, 1, 1) process. Therefore, testing the null hypothesis that $\varphi=1$ is equivalent to testing for a unit autoregressive root.

Notice that ε_t can be written as

$$\varepsilon_t = \sum_{j=0}^{\infty} \theta^j (z_{t-j} - \alpha z_{t-j-1})$$

from which it follows that

$$X_t - X_{t-1} = (\varphi - 1)\, X_{t-1} + (\alpha - \theta)\, (z_{t-1} + \theta z_{t-2} + \theta^2 z_{t-3} + \dots) + \varepsilon_t \tag{3.3}$$

and, therefore, under the null hypothesis that $\varphi=1$ (3.1) implies that $z_t = X_t - X_{t-1} = \Delta X_t$.

Hence equation (3.3) can be estimated by regressing ΔX_t on X_{t-1}, ΔX_{t-1}, ..., ΔX_{t-k}, where k is an integer number that specifies the number of autoregressive parameters. Notice that equation (3.3) is identical to equation (2.7), and according to Said and Dickey (1984), the test statistic for testing the unit root hypothesis, known as the τ test, will have the same distribution as that originally tabulated by Fuller (1979), but not the ρ test. As shown by Said and Dickey (1984) the limit distribution of the ρ test depends on the unknown moving average and autoregressive parameters and, therefore, the ρ test cannot be used in this case. Furthermore, this approach can be extended to any ARIMA (p, 1, q) process.

Phillips (1987), on the other hand, has derived tests for testing for a unit autoregressive root that do not require any specifications of the orders of the autoregressive or the moving average parameter. Think for example a random walk process in which the error term has an ARMA (p, q) presentation. The unit root test is based on estimating a non-augmented Dickey-Fuller regression, i.e., for the zero mean case, equation (2.4), and the τ test is defined as follows

$$\hat{\tau} = T(\hat{\beta} - 1) - \frac{1}{2}(S^2_{Tk} - S^2_{\varepsilon})\, T^2 \left[\sum_{t=2}^{T} (X_{t-1} - \overline{X}_{-1})\right]^{-1} \tag{3.4}$$

where s_{ε}^2 is the sample variance of the residuals ε_t and

$$S^2_{Tk} = T^{-1} \sum_{t=2}^{T} \varepsilon_t^2 + 2T^{-1} \sum_{j=1}^{k} \omega_{jk} \sum_{t=j+1}^{T} \varepsilon_t \varepsilon_{t-j} \tag{3.5}$$

where the weights $\omega_{jk} = 1 - j/(1+k)$ ensure that the estimate of the variance of s^2_{Tk} is positive (Newey and West (1987)). Critical values for this test can be obtained from Fuller (1976).

The Phillips procedure has also been extended by Phillips and Perron (1988) to allow trend under the alternative. The test statistic is conducted in a similar fashion as in (3.4) and as indicated by Phillips and Perron (1988) the modified test statistic for testing for a unit autoregressive root in the presence of a time trend component has the same distribution as that originally tabulated by Fuller (1979).

Unfortunately, as shown by Stock and Watson (1988), Schwert (1989) and Agiakloglou and Newbold (1992) the performance of these test statistics in moderately large samples is not satisfactory. In particularly, Agiakloglou and Newbold (1992) examined in depth the performance of the "augmented" Dickey-Fuller test not only because it is more attractive and easy to implement it but also because as indicated by Stock and Watson (1988) and Schwert (1989) the performance of the Phillips (1987) and Phillips and Perron (1988) tests was more inconsistent. Moreover, a recent paper by Leybourne and Newbold (1999) encountered that the non-parametric Phillips-Perron test has serious problems of oversized tests even when the true values of the short —and long— run variances are used in place of the sample estimates even for the case of the simple ARIMA (0, 1, 1) model. The paper also indicated that the t ratio variant of the test performs rather more poorly than an implementable version of the Dickey-Fuller test.

To view this consider the simplest possible model with one moving average parameter, the ARIMA (0, 1, 1) process, i.e.,

$$X_t - X_{t-1} = \varepsilon_t - \theta\varepsilon_{t-1} \tag{3.6}$$

in which the unit root hypothesis is true for all values of θ strictly less than one. If θ=1 then the process (3.6) is a white noise process.

As shown by Agiakloglou and Newbold (1992) the performance of the "augmented" Dickey-Fuller test that was applied to processes generated by model (3.6) for various values of θ using samples of 100 observations is strongly affected by the order of the approximating autoregression, recalling that Agiakloglou and Newbold (1992) tried values of k from 0 to 10, and by the magnitude of the moving average parameter.

For small and moderate values of θ and taking k3 yields a test with approximately correct significance levels whereas for large values of θ the empirical significance levels are grossly infalted. On the other hand, for all values of the moving average parameter the test improves its performance, in the sense that the empirical significance levels approach the nominal level tests, as long as a high value of k is employed for the implementation of the "augmented" Dickey-Fuller test. Hence, it is less likely to reject the unit root hypothesis using a high value of k but, as indicated by Agiakloglou and Newbold (1992), the cost of using a high value of k is that the test losses its power.

The proposed by Agiakloglou and Newbold (1992) AIC criterion —Akaike Information Criterion— which has the tendency to yield more heavily parameterized models as apposed to SBC criterion —Scwarz Bayesian Criterion— failed to satisfactory select the right order of the approximating autoregression. Recall that the order of the approximating autoregression k is arbitrary defined by Schwert (1989) as proportional to the fourth root of the number of observations T, divided by one hundred, i.e., l_k=int{c(T/100)$^{1/4}$ where c=4 or 12. Thus for series of 100 observations Schwert (1989) tried values of k equal to 4 and 12 indicating that it is less likely to reject the unit root hypothesis for k=12. Said and Dickey (1984), on the other hand, obtained the final value of k by using a sequence of F — tests for various values of k. Unfortunately, such an approach is quite ambiguous and depends greatly on the significance level of the test. Moreover, it can be shown that the model selected by the AIC criterion will not be rejected against more elaborate models by an F — test for series of 100 observations using a 5% significance level.

To the contrary, Hall (1994) and Ng and Perron (1995) made an effort to determine the order of the autoregressive equation (2.7) by using sequential tests, a method in which the value of k is chosen according to the significance of the coefficient of the last lagged dependent variable. In other words, a series of tests are applied to the coefficient of the last lagged dependent variable using either the specific-to-general or the general-to-specific approach and the final value of k is chosen when the coefficient of the last lagged dependent variable is statistical significant from zero.

Hall (1994) examined these testing procedures to the pure autoregressive processes whereas Ng and Perron (1995) applied these procedures to processes that include moving average components. Both studies acknowledge the fact that the specific-to-general approach has the same asymptotic properties as a selection rule based on the information criteria, AIC or SBC, and therefore the chosen value of k will always be small whereas using the general-to-specific approach, starting at some predefined large value of k, it is more likely to select a high value of k than that selected by an information criteria with some loss of the power of the test. However, it should not be ignored the fact that the general-to-specific approach similar to the sequence of F — tests approach proposed by Said and Dickey (1984) is sensitive to the critical values used for the determination of the significance of the last lagged dependent variable. Furthermore, a study by Agiakloglou and Newbold (1996) discusses the issue of the trade-off between size distortions and power loss when the ADF test is applied to processes like AR(1) and ARIMA(0, 1, 1) regard less of the method used to select the order of the approximating regression.

Finally, Perron (1989) has presented a unit root test based on the "augmented" Dickey-Fuller regression (2.7) augmented by a time trend term and a set of three dummy variables to allow for a deterministic change at a given point of time. Critical values for this test can be obtained from Perron (1989). Apart from all the problems that may typically arise in terms of applying a unit root test and had already discussed, Newbold and Agiakloglou (1992) have shown that this particular unit root test is sensitive to the choice of the break point. Thus, if a single point exists, the model selected by the usual criteria for the series of the logarithm of the US common stock prices results in very weak evidence against the unit root hypothesis whereas the unit root hypothesis for the model chosen by Perron (1989) is rejected at the nominal 2.5% level test. Recall that Perron (1989) was able to conclude that for some of the series originally studied by Nelson and Plosser (1982) the unit root hypothesis can be rejected if a structural change is incorporated to the unit root test imposed to the series.

In the vein, Leybourne *et al.* (1998) following the initial study of Perron (1989) explored in some detail the performance of the Dickey-Fuller tests in the presence of a break under the null hypothesis. According to the result applying the Dickey-Fuller tests when the true generating process is integrated of order one, but with a break, can lead to a severe problem of spurious rejections of the unit root hypothesis especially when the break is early in the series.

4. The Unit Root Test based on least squares estimation of ARIMA (p, 1, q) models when the orders ρ and q are known

The unit root test is extended by Said and Dickey (1985) in the case where the orders ρ and q of an ARIMA (p, 1, q) process are known. As shown by Said and Dickey (1985) the unit root test for testing an ARIMA (p, 1, q) process against a stationary ARIMA (p+1, 0, q) process will follow the same asymptotic distribution originally tabulated by Fuller (1976). The estimation procedure is executed using the one-step Gauss-Newton least squares estimation and it is also discussed in Solo (1984). The difference is that the Solo (1984) test is a Lagrange Multiplier test whereas the Said and Dickey (1985) test is a Wald type test.

Consider for example the time series X_t for t= 1, 2, ... T satisfying the following process

$$X_t = c + \varphi X_{t-1} + \varepsilon_t - \theta\varepsilon_{t-1} \qquad (4.1)$$

where $X_0 = 0$, $|\theta| < 1$ and $\varepsilon_t \sim iidN(0, \sigma^2)$. If the null hypothesis is true, i.e., $\varphi=1$, then model (4.1) is simply an invertible, moving average process for ΔX_t since $|\theta|$ is assumed to be less than one.

The one-step Gauss-Newton least squares estimation procedure as shown by Said and Dickey (1985) works as follows. First, model (4.1) is employed to obtain estimates for the constant and for the moving average parameter. Then, assuming that $\varphi=1$ and setting $\hat{\varepsilon}_1=\hat{V}_1=\hat{W}_1=0$, the values of ε_t, $\hat{V}_t$ and $\hat{W}_t$ are computed as follows

$$\hat{\varepsilon}_t = X_t - X_{t-1} + \hat{\theta}\hat{\varepsilon}_{t-1} - \hat{c}$$

$$\hat{V}_t = X_{t-1} + \hat{\theta}\hat{V}_{t-1}$$

$$\hat{W}_t = \hat{\theta}\hat{W}_{t-1} - \hat{\varepsilon}_{t-1}$$

where t= 1, 2, ..., T.

Finally, using the ordinary least squares estimation, the following model is estimated

$$\hat{\varepsilon}_t = \alpha + \beta\hat{V}_t + \gamma\hat{W}_t + u_t \tag{4.2}$$

where t= 2, 3, ..., T, $u_t \sim N(0, \sigma^2)$ and $\hat{\beta} = (\hat{\varphi} - 1)$. Thus, for large sample sizes the $\hat{\rho}$ test will be based on the statistic $T(\hat{\beta})$ and the $\hat{\tau}$ test will be based on the t-statistic for testing the null hypothesis that β=0 where critical values for these tests can be obtained form Fuller (1976).

The most critical point in terms of applying this particular unit root test to processes generated by (3.6) is to obtain consistent initial estimates of the moving average parameter. Said and Dickey (1985) proposed the following three methods of establishing consistent estimates: a) using the first-order autocorrelation coefficient, r_1, of an MA(1) process and solve for $|\theta|<1$, recalling that $r_1=-\theta/(1+\theta^2)$, b) using the Durbin's method, i.e., fitting a long autoregression to the data and then regressing these coefficients on their lags to estimate the moving average parameter, and c) using the true θ. Unfortunately, none of the above methods was able to produce correct empirical significance levels for large values of θ. To the contrary, Agiakloglou and Newbold (1922) were able to obtain significance levels that were close to the nominal levels when the initial value of θ was obtained from a maximum likelihood estimation of an ARIMA (0, 1, 1) model with no constant.

However, it should be emphasized that the objective in this case is not to indicate which method produces the most reliable estimates of the moving average parameter, but to explicitly point out that the performance of this particular unit root test even for the simplest possible model with one moving average term is so sensitive to the initial value of a single parameter.

Furthermore, Dolado and Hidalgo-Moreno (1990), contrary to the proposed one-step Gauss-Newton least squares estimation, proved asymptotically that unit root tests can be based on the estimates of an ARMA (1,1) model using one of the widely available packages for fitting ARIMA models. Unfortunately, empirical test results presented by Agiakloglou and Newbold (1992) that are based on generated series of 100 observations of model (4.1) with φ=1 and large values of Θ, using SPSS through maximum likelihood estimation, and using SAS through unconditional least squares, conditional least squares and maximum likelihood estimation in which the estimates are defined as in Anshley and Newbold (1980) do not support the theoretical results presented by Dolado and Hidalgo-Moreno (1990). Perhaps one explanation for such results is that asymptotics do not work for series of one hundred observations.

Lastly, Pantula (1991) provides theoretical treatment of the case where the moving average root of the generating model approaches one.

5. The Unit Root Tests based on the Sample Autocorrelations of ARIMA (p. 1, q) models

Bierens (1993) has developed a test statistic for testing for a unit autoregressive root based on the sample autocorrelations of an observed time series. The test statistic for a given time series X_t, t= 1, 2, ..., T, is defined in the following way

$$R(\mu) = \frac{T\,[r_k - 1]}{k\,|\,\eta(k)\,|^{\mu}} \qquad \text{if} \quad \eta(k) > 0 \tag{5.1}$$
$$= -T^2 \qquad \text{if} \quad \eta(k) \leq 0$$

where

$$\eta(k) = \frac{k\,(r_{k+1} - r_k)}{r_k - 1}$$

r_k is the k^{th} sample autocorrelation defined as

$$r_k = \frac{\sum_{t=\kappa+1}^{T} (X_t - \overline{X})(X_{t-k} - \overline{X})}{\sum_{t=1}^{T} (X_t - \overline{X})^2} \tag{5.2}$$

for k = 1, 2, ..., $\overline{X}$ is the sample mean and critical values for this test can be obtained from Bierens (1993).

Unfortunately, as shown by Agiakloglou (1996), the performance of the (5.1) test statistic for testing for a unit autoregressive root applied to processes generated by (3.6) for various values of the moving average parameter and for sample sizes of 100 and 300 observations is satisfactory only for the negative values of the moving average parameter as long as low values of k are used. On the other hand, for all positive values of the moving average parameter including the random walk process the test not only failed to produce empirical significance levels close to the nominal level tests but also in some cases the empirical significance levels were grossly inflated. Moreover, the empirical significance levels were far away form any nominal level tests if the proposed by Bierens (1993) values of μ=4 and k=12 and 14 for series of 100 and 300 observations respectively were used for all values of the moving average parameter. These empirical results, as shown by Agiakloglou (1996), were also supported on theoretical grounds in terms of applying the (5.1) unit root test to the means of the sample autocorrelations of ARIMA (0, 1, 1) models which can be estimated as a ratio of two quadratic forms in normal deviates (see for example Kendall (1954), Marriott and Pope (1954), Wichern (1973) and Kumar (1973)).

In essence, the performance of the (5.1) test statistic is strongly affected not only by the values of k or μ used for the implementation of the test but also by the values of the moving average parameter. The latter result can also been found in previous studies by Schwert (1989) and by Agiakloglou and Newbold (1992) in which they have indicated that the performance of the existing unit root test statistics is strongly affected by the presence or large positive values of the moving average parameter. To see this once more consider the model chosen by Bierens (1993) the ARIMA (1,1,1) process

$$(1 - \tau B)(X_t - X_{t-1}) = (1 - \theta B)\varepsilon_t \qquad (5.3)$$

where ε_t is white noise. Series of 100 observations are then generated by model (5.3) for values of τ= 0.95 and for values of θ= 0.5, 0.8 and 0.9 based in 1,000 replications. In this case, the unit root test is conducted using values of μ=4 and k=12 proposed by Bierens (1993) and the null hypothesis is rejected at the nominal 5% level test 58, 08 and 267 times out of 1,000 trials for values of θ=0.5, 0.8 and 0.9 respectively.

6. Fractionally Integrated ARMA Processes

Time series models have been recently analyzed, contrary to the traditional Box and Jenkins (1976) methodology, by allowing the degree of differencing to take any real value. The class of models that considers fractional integration, known as ARFIMA models, nests the unit root phenomenon as a special case and examines the behavior of a time series in a more general way than the standard ARIMA analysis.

ARFIMA models are also known as long-memory models whereas standard ARIMA models are known as short-memory models. The distinction between long-memory and short-memory processes is based on the degree of dependence between observations widely separated in time. If the degree of dependency between observations a long time apart decays much more slowly than the standard ARIMA process the process is called long-memory. This is the same to say that the autocorrelations of a long-memory process decay very slowly.

For a given time series X_t the ARFIMA model of orders (p, d, q) is written as in (1.1) where now

$$(1 - B)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j B^j = 1 - dB + \frac{d(d-1)}{2} B^2 - \frac{d(d-1)(d-2)}{6} B^3 + \ldots$$

and ε_t white noise. As shown by Granger and Joyeux (1980) and Hosking (1981) the process (1.1) is stationary if $d<0.5$ and all roots of the polynomial $\Phi(B)$ are outside the unit root circle and it is invertible if $d>-0.5$ and all roots of the polynomial $(\Theta(B)$ are outside the unit root circle. Of course, if the process is non-stationary it can be converted to a stationary process by differencing.

Although fractionally integrated ARMA models have applications in both economics and hydrology, the genesis of such models can be actually found in hydrology, where the Hurst phenomenon is often studied (see for example McLeod and Hipel (1978)). Some theoretical and empirical studies for this type of models include Geweke and Porter-Hudak (1988), Li and McLeod (1986), Diebold and Rudebusch (1989, 1991 a, b), Porter-Hudak (1990), Sowell (1986, 1990) and Yajima (1988).

One of the methods that is frequently used, especially in applications to economic time series (see for example Diebold and Rudebush (1989)), to estimate fractionally integrated models is the two-step estimation procedure proposed by Geweke and Porter-Hudak (1988). According to this approach, first an estimate of the fractional difference parameter is obtained. This estimate is then used to transform the series. Recall that in the case where d is a fractional number the $(1-B)^d$ term can be expressed as an infinite binomial series expansion in power of the backshift operator B. Then the traditional ARMA analysis is applied to the transformed seres to obtain estimates of the autoregressive and moving average parameters.

The potentially most attractive feature of this approach is the method by which the estimate of the fractional difference parameter is obtained. As shown by Geweke and Porter-Hudak (1988) an estimate of d is obtained by estimating the following equation

$$\ln(I(\lambda_j)) = \alpha + \beta \ln[4 \mathrm{sim}^2 (\lambda_j/2)] + \qquad e_j \qquad (6.1)$$

where $\beta=-d$, $I(\lambda_j)$ is the periodogram of X_t at ordinate λ_j and the error terms ε_j, $j=1, 2, \ldots, k$, are independent and identically distributed with mean zero and variance $\pi^2/6$. Geweke and Porter-Hudak (1988) established consistency and asymptotic normality of the estimate of β provided that the number of ordinates in (6.1) is a function of sample size. In practice, k is often taken to be $T^{1/2}$.

Estimates of the fractional difference parameter d along with their standard error can therefore be obtained by ordinary least squares estimation of the Geweke and Porter-Hudak regression (6.1). Tests then for the fractional difference parameter d are based on the usual t — statistic, although, the use of the theoretical variance of the error term of the Geweke and Porter-Hudak regression is strongly recommended to increase the efficiency of the test.

Unfortunately, as shown by Agiakloglou *et* al.(1993) for two simple ARMA processes, the AR(1) and the MA(1) process, the Geweke and Porter-Hudak estimation procedure of obtaining consistent estimates of the fractional difference parameter is not reliable. Estimates of d may some times be seriously biased even for large sample sizes. Therefore, using the Geweke and Porter-Hudak approach as a test procedure to test the null hypothesis that d is zero can often lead to inconsistent conclusions.

Evidence of the bias of the Geweke and Porter-Hudak (1988) estimation procedure is also reported by Hurvich and Ray (1995) for non-stationary and non-invertible fractionally integrated processes.

Finally, Sowell (1992) has developed a maximum likelihood estimation procedure of obtaining simultaneously estimates of the autoregressive and moving average parameters and the fractional difference parameter d. An approximate maximum likelihood algorithm can also be found in Hipel and McLeod (1978), However, it should be taken into account the fact that, although for the traditional ARIMA analysis the presence or the absence of the unknown mean does not affect the estimates or the standard errors, for these particular ARFIMA processes evidence of bias in the presence of the sample mean as apposed to the unknown population mean is reported by Newbold and Agiakloglou (1993).

7. Summary

Although the unit root issue is really very important in time series analysis the existing testing procedures have failed so far for moderately large samples to successfully determine whether or not an observed time series is generated by a unit autoregressive root process. Perhaps one explanation to this sad phenomenon is the fact that asymptotics do not work very well for small samples although in economics we typically do not have large samples.

Broadly speaking the decision as to whether or not to accept the unit root hypothesis is not merely a matter of simply applying say the "augmented" Dickey-Fuller test. In fact it is very difficult to believe that by choosing any arbitrary value of k to implement the "augmented" Dickey-Fuller test, based on the test statistic for testing the unit root hypothesis, can successfully determine whether the series is stationary or non-stationary. As discussed, the performance of the "augmented" Dickey-Fuller test is so significantly affected by the order of the approximating autoregression that an analyst will get different test results for different values of k (see for example Newbold *et al* (1993)).

Another way that an analyst can approach this problem is by simply applying the traditional ARIMA analysis to the series. The benefit of using this approach is that on the one hand the ARIMA analysis does not depend on the number of the autoregressive parameters used to determine whether the series is stationary or not and on the other hand the traditional ARIMA analysis is independent of *prior* model specifications.

References

Agiakloglou, C. (1996) A note on the unit root test based on the sample autocorrelations. SPOUDAI, 46, 97-106.

Agiakloglou, C. and Newbold, P. (1996), The balance between size and power in Dickey-Fuller tests with data-dependent rules for the choice of truncation lag. Economic Letters, 52, 229-234.

Agiakloglou, C. and Newbold, P. (1992), Empirical Evidence on Dickey — Fuller — type tests. Journal of Time Series Analysis, 13, 471-483.

Agiakloglou, C, Newbold, P. and Wohar, M. (1993). Bias in an estimator of the fractional difference parameter. Journal of Time Series Analysis, 14, 235-246.

Anshley, C. F. and Newbold, P. (1980), Finite sample properties of estimators for autoregressive moving average models. Journal of Econometrics, 13, 159-183.

Bierens, J. H. (1993). Higher-order sample autocorrelations and the unit root test. Journal of Econometrics, 57, 137-160.

Box, G.E.P. and Jenkins, G. M. (1976). "Time Series: Analysis: Forecasting and Control", San Francisco: Holden-Day.

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.

Diebold, F. X. and Rudebusch, G. D. (1989) Long memory and persistence in aggregate output. Journal of Monetary Economics, 24, 189-209.

Diebold, F. X. and Rudebusch, G. D. (1991a). Is consumption too smooth? Long memory and the Deaton paradox. The review of Economics and Statistics, 73, 1-9.

Diebold, F. X. and Rudebusch, G. D. (1991b). On the power of the Dickey-Fuller tests against fractional alternatives. Economic Letters, 35, 155-160.

Dolado, J. J. and Hidalgo-Moreno, J. (1990). The asymptotic distribution of the iterated Gauss-Newton estimators of an ARIMA process. Econometric Theory, 6, 490-494.

Engle, R. F. and Granger. C. W.J. (1987). Cointegration and error correction: representation, estimation and testing. Econometrica, 55, 251-276.

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

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

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

Geweke, C. W. G. and Porter-Hudak, S. (1988). The estimation and application of long memory time series models. Journal of Time Series Analysis, 4, 221-238.

Granger, C. W.J. and Joyeux, R. (1980). An introduction to long memory time series models and fractional differencing. Journal of Time Series Analysis, 1, 15-29.

Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in econometrics. Journal of Ecoinometrics, 2, 111-120.

Granger, C. W.J. and Newbold, P. (1986). "Forecasting Economic Time series" second edition. New York: Academic Press.

Hall, A. (1994). Testing for a unit root in time series with pretest data based model selection. Journal of Business and Economic Statistics, 12, 461-470.

Harvey, A. C. (1985). Trends and cycles in macroeconomic time series. Journal of Business and Economic Statistics, 3, 216-227.

Harvey, A. C. (1989). "Forecasting Structural Time Series Models and the Kalman Filter", University Press, Cambridge.

Hipel, K. W. and McLeod, A. I. (1978). Preservation of the rescaled adjusted range 3. Fractional Gaussian noise algorithms. Water Resources Research, 14, 517-518.

Hosking, J. R. M. (1981), Fractional differencing. Biometrika, 68, 165-176.

Hurvich, C. M. and Ray, B. K. (1995). Estimation of the memory parameter for non-stationary or noninvertible fractionally integrated processes. Journal of Time Series Analysis, 16, 17-42.

Johansenm, S. (1991). Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica, 59, 1551-1580.

Kendall, M. G. (1954), Note on bias in the estimation of autocorrelation. Biometrika, 41, 403-404.

Kumar, A. (1973), Expectation of product of quadratic forms. SANKHYA, B, 35, 359-362.

Leybourne, S. and Newbold, P. (1999). On the size properties of Phillips-Perron tests. Journal of Time Series Analysis, 20, 51-61.

Leybourne, S. I., Mills, R., C. and Newbold, P. (1998), Spurious rejections by Dickey-Fuller tests in the presence of a break under the null. Journal of Econometrics, 87, 191-203.

Li, W.K. and McLeod, A.I. (1986), Fractional time series modeling. Biometrika, 73, 217-221.

Marriott, F. H. and Pope, J. A. (1954), Bias in the estimation of autocorrelations, Biometrika, 41, 390-402.

McLeod, A. I. and Hipel, K. W. (1978). Preservation of the rescaled adjusted range 1. A reassessment of the Hurst phenomenon. Water Resources Research, 14, 491-508.

Nankervis, J. C. and Savin, N.E. (1985), Testing the autoregressive parameter with the t statistic. Journal of Econometrics, 27, 143-161.

Nelson, C. R. and Plosser, C. I. (1982). Trends and random walks in macroeconomic time series. Journal of Monetary Economics, 10, 139-162.

Newbold, P. and Agiakloglou, C. (1992). U.S. common stock prices, 1871-1970: Playing with Dummies. Journal of Quantitative Finance and Accounting, 2, 215-220.

Newbold, P. and Agiakloglou, C. (1993). Bias in the sample autocorrelations of fractional noise. Biometrika, 80, 698-702.

Newbold, P., Agiakloglou, C. and Miller, J. (1993). Long-term Inference based on short-term forecasting models. In Developments in Time Series Analysis, In Honor of M. B. Priestley, edited by T.S. Rao, London: Chapman and Hall.

Newey, W.K. and West, K.D. (1987). A simple positive semidefinite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica, 55, 703-708.

Ng, S. and Perron, P. (1995), Unit root tests in ARMA models with data-dependent method, for the selection of the truncation lag. Journal of the American Statistical Association, 90, 268-281.

Pantula, S. G. (1991). Asymptotic distributions of unit-root tests when the process is nearly stationary. Journal of Business and Economic Statistics, 9, 63-71.

Park, J. Y. and Phillips, P.C.B. (1988). Statistical inference in regressions with integrated processes: part 1. Econometric Theory, 4, 468-497.

Park, J. Y. and Phillips, P.C.B. (1989). Statistical inference in regressions with integrated processes part 2. Econometric Theory, 5, 95-131.

Perron, P. (189), The great crash, the oil price stock, and the unit root hypothesis. Econometrica, 57, 1361-1401.

Phillips, P.C.B. (1987), Time series 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.

Porter-Hudak, S. (1990). An application of the seasonal fractionally differenced model to the monetary aggregates. Journal of the American Statistical Association, 85, 338-344.

Said, S.E., and Dickey, D.A. (1984). Testing for unit roots in autoregressive — moving average models of unknown order. Biometrika, 71, 599-607.

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

Schwert, G. W. (1989). Tests for unit roots: A monte carlo investigation. Journal of Business and Economic Statistics, 7, 147-160.

Sims, C.A., Stock, J.H. and Watson, M. W. (1990). Inference in linear time series models with some unit roots. Econometrica, 58, 113-144.

Solo, V. (1984), The order of differencing in ARMA models. Journal of the American Statistical Association, 76, 916-921.

Sowell, F. B. (1986). "Fractionally Integrated Vector Time Series". Ph. D. Thesis, Duke University.

Sowell, F. B. (1990). The fractional unit root distribution. Econometrica, 58, 495-505.

Sowell, F. B. (1992). Maximum likelihood estimation of stationary univariate fractionally integrated time series models. Journal of Econometrics, 53, 165-188.

Stock, J. H. and Watson, M. W. (1988). Testing for common trends. Journal of the American Statistical Association, 83, 1097-1107.

Wichern, D. W. (1973). The behavior of the sample autocorrelation function for integrated moving average process. Biometrika, 60, 235-239.

Yajima, Y. (1988). On estimation of a regression model with long memory stationary errors. The Annals of Statistics, 16, 791-807.