Taking a New Contour: A Novel View on Unit Root Test [1]

Yoosoon Chang
Department of Economics
Rice University

and

Joon Y. Park
Department of Economics
Rice University and Sungkyunkwan University

Abstract

In this paper we introduce a new view on the distributions of unit root tests. Taking a contour given by the fixed sum of squares instead of the fixed sample size, we show that the null distributions of most commonly used unit root tests such as the ones by Dickey-Fuller (1979, 1981) and Phillips (1987) are normal in large samples. The normal asymptotics along the new contour continue to hold under the local-to-unity alternatives, in which case the tests have normal limit distributions with mean given by the product of the square root of the level of the contour and the locality parameter. Our results are derived for the general unit root models with innovations satisfying the functional central limit theory that is routinely employed to obtain the unit root asymptotics. Moreover, the new asymptotics are shown to be applicable also for the models with deterministic components, as long as they are removed recursively by using only the past information.

First Draft: August 10, 2002
This version: December 30, 2004

JEL Classification: C12, C16, C22.

Key words and phrases: unit root test, Dickey-Fuller distribution, normal asymptotics, Brownian motion, continuous martingale, time change

[1] We are grateful to Don Andrews, Jushan Bai, Xiahong Chen, Peter Hansen, Byungsoo So, Jim Stock and Donggyu Sul for useful comments. Correspondence address to: Yoosoon Chang, Department of Economics - MS22, Rice University, 6100 Main Street, Houston, TX 77005-1892, Tel: 713-348-2796, Fax: 713-348-5278, Email: yoosoon@rice.edu.

1. Introduction

It is well known that the distributional theories for many of the commonly used unit root tests are nonstandard. For instance, the Dickey-Fuller and Phillips tests both have non-normal distributions, which are usually represented by the functionals of Brownian motion under the null hypothesis of unit root. The characteristics of their null distributions, which are often referred to as the Dickey-Fuller distributions named after who first tabulate them, have been studied by several authors including Evans and Savin (1981, 1984) and Abadir (1993). In particular, the Dickey-Fuller distributions are known to be asymmetric and skewed to the left, as demonstrated in, e.g., Fuller (1996). The nonstandard nature of the limit distribution theory for the unit root tests extends to the case of the near unit root, which is often considered as the local alternative to the unit root null hypothesis. Under the local-to-unity alternative hypothesis, the limit distributions of the unit root tests are generally given by the functionals of Ornstein-Uhlenbeck process.

This paper introduces a novel view on these and other related distributions. The sampling distribution of a statistic is usually obtained for a given sample size. Using the conventional sampling distribution of the statistic for the purpose of statistical inference thus implies that we evaluate the likelihood of a realized value of a statistic against other possible realizations along the contour given by the fixed sample size. In this paper, we consider taking a different contour in obtaining the sampling distribution of the statistic, i.e., the contour that is given by the fixed sum of squares. In order to assess the likelihood of the statistic, we therefore look for other possible realizations with their sum of squares, rather than their sample sizes, holding fixed. As we show in the paper, the distribution theory of a unit root test can be very different, depending upon whether we take the conventional contour of the fixed sample size or the new contour of the fixed sum of squares to evaluate the likelihood of the test statistic.

The distinction and choice between the two contours introduced above matter only for the nonstationary models including the unit root or near unit root. In particular, they become unimportant for the stationary models. For the observations from stationary time series, the sum of squares becomes a constant multiple of the sample size in the limit. The contours of the equi-sample-size and the equi-squared-sum to evaluate the likelihood of a realized sample are thus virtually identical if the size of the sample is large enough. The limit distribution theories are therefore identical regardless of the choice of the contours. This is not so for the samples from unit root processes. If normalized as necessary, the sum of squares of the samples from unit root processes remain to be random even in the limit. For the samples from the unit root and near unit root models, it would thus yield a new asymptotic theory different from the conventional one to evaluate the likelihood of a given realization against all other possible realizations with the same sum of squares.

As an illustration, we provide ten simulated sample paths with equal sample size, and another ten with equal sum of squares, respectively, in Figures 1 and 2.[2] For the equi-sample-size paths provided in Figure 1, the one with largest sum of squares are presented

[2]More precisely, ten sample paths at 5%, . . . , 95% percentiles were chosen out of ten thousand realizations. The sample size is fixed at 100 for Figure 1, while we set the sum of squares to be 0.23 times 100 squared for Figure 2. The setting yields the most comparable results for the two contours considered here.

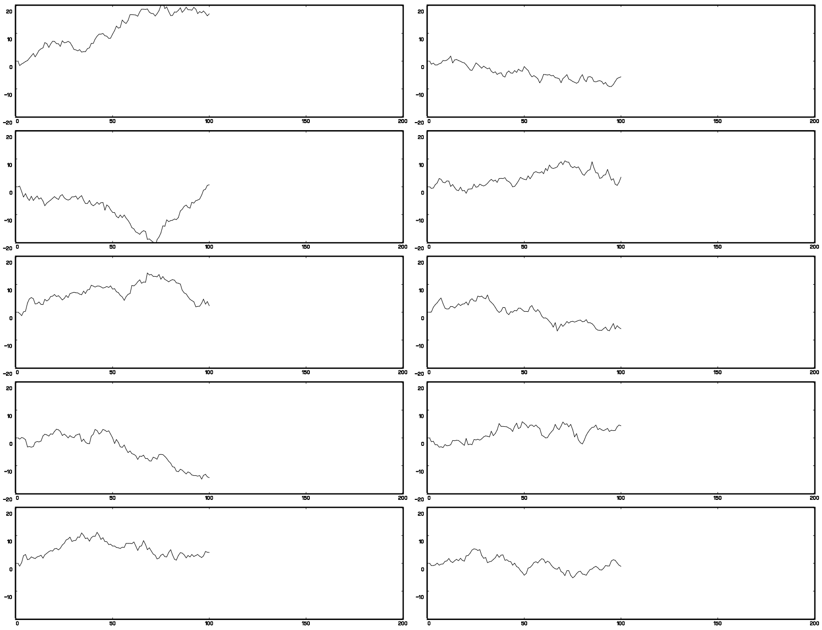

Figure 1: Sample Paths with Equal Sample Size

in the top-left corner and the one with the smallest sum of squares in the bottom-right corner. For the sample paths with equi-squared-sum in Figure 2, the one with the smallest sample size to attain the required squared sum is presented in the top-left corner, and the one with the largest sample size in the bottom-right corner. Figures 1 and 2 represent two different contours we may take to obtain the sampling distributions of the statistics involving unit root processes. The former shows the contour of the samples of fixed sample size (with varying sums of squares as required to have the same sample size), while the latter represents the contour of fixed sum of squares (with varying sample sizes as required to have the same sum of squares).

The new contour is particularly relevant in testing for a unit root. For the unit root test, the information content of the sample is effectively measured, not by the sample size, but by the sum of squares. The most conspicuous characteristic of the unit root processes (that is constrasted with the stationary process) is the presence of stochastic trend. The degree of conspicuousness of their stochastic trend would essentially lead us to believe, or not to believe, the presence of a unit root in the underlying time series. How conspicuous should they be for us to reject, or not to reject, the unit root hypothesis? Here comes the point that we need a formal statistical test. The remaining question is whether to evaluate the likelihood of a given observation against other possible realizations either of the same sample size exhibiting different degrees of stochastic trends, or having similar degrees of stochastic trends with varying sample sizes. At least in this regard, it appears to be more reasonable to consider the contour of the equi-squared-sum rather than that of the equi-sample-size.

Our asymptotic results are rather surprising. Along the new contour of equi-squared-sum, we have the usual normal asymptotics for the unit root and near unit root models.

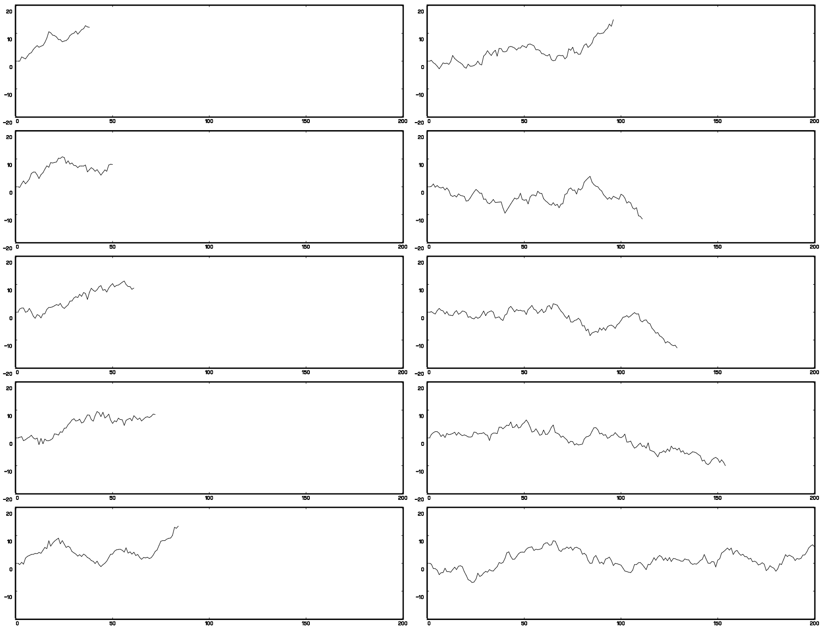

Figure 2: Sample Paths with Equal Sum of Squares

The limit distribution of the t-ratio is standard normal under the null of unit root, and it is simply shifted with the location parameter given by the product of the square root of the level of the contour and the locality parameter under the alternative of local-to-unity. The nonstandard nature of the conventional limit distribution theories of the unit root and near unit root models therefore completely disappears once the contour of equi-squared-sum is taken. Perhaps it may merely suggest that the usual contour of equi-sample-size is inappropriate for the unit root and near unit root models. The new asymptotics are developed under very weak conditions, which only require the innovations to satisfy the usual functional central limit theory. Our results are also applicable for models with deterministic components, as long as they are removed recursively by using only the past information.

The rest of the paper is organized as follows. The main results of the paper are given in Section 2. There we consider the prototype unit root model and the test statistic, and develop a new asymptotics along the contour of the equi-squared-sum. The asymptotics are shown to be normal under both the null of unit root and the alternative of local-to-unity. Section 3 extends our main results. In particular, it is shown that the normal asymptotics continue to apply for the models with intercept if the recursive demeaning is used. The tests based on more general unit root models are also investigated and shown to yield the normal asymptotics along the new contour. The concluding remarks are in Section 4, and the mathematical proofs are given in Appendix. A word on notation. As usual, $\to_d$, $\to_p$ and $\to_{a.s.}$ are used to signify respectively the convergence in distribution, the convergence in probability and the almost sure convergence, and $\sim$ denotes the equivalence in distribution. The standard Brownian motion is denoted by W throughout the paper.

2. Main Results

We consider the autoregressive model

$$y_t = \alpha y_{t-1} + u_t, \tag{1}$$

and the test of the unit root hypothesis

$$\alpha = 1. \tag{2}$$

We assume that

Assumption 2.1 Let $(u_t, \mathcal{F}_t)$ be a martingale difference sequence, with some filtration $(\mathcal{F}_t)$, satisfying invariance principle.

The martingale difference condition in Assumption 2.1 is not necessary and will be relaxed later. It is introduced here simply to avoid unnecessary complications and focus on the main issue of the paper. We may consider more general unit root models driven by linear processes or weakly dependent innovations without any difficulty. For such general models, the unit root test may be based on the regression augmented with the lagged differences as for the tests by Dickey and Fuller (1979, 1981), or can be done using the statistic modified nonparametrically as in the tests by Phillips (1987). These tests will be considered explicitly in a later section. They all have the same limit distributions as the test based on the simple model considered here. We define

$$W_n(r) = \frac{1}{\sigma\sqrt{n}} \sum_{t=1}^{[nr]} u_t \tag{3}$$

for $r \geq 0$, where $[z]$ signifies the largest integer which does not exceed z and σ^2 is the asymptotic variance of (u_t), i.e., the probability limit of $(1/n)\sum_{t=1}^n u_t^2$. Under Assumption 2.1, we have $W_n \to_d W$ as $n \to \infty$, where W is the standard Brownian motion.

Let $y_1, \ldots, y_n$ be the random sample of size n. The unit root hypothesis is routinely tested by the t-ratio on the autoregressive coefficient α, which is given by

$$T_n = \frac{\hat{\alpha}_n - 1}{s(\hat{\alpha}_n)}, \tag{4}$$

where $\hat{\alpha}_n = (\sum_{t=1}^n y_{t-1}^2)^{-1}\sum_{t=1}^n y_{t-1}y_t$ is the least squares estimator of α with the standard error $s(\hat{\alpha}_n) = \hat{\sigma}(\sum_{t=1}^n y_{t-1}^2)^{-1/2}$, and $\hat{\sigma}^2$ denotes the usual error variance estimate obtained from the fitted residual $(\hat{u}_t)$ in regression (1). It is well known that under the null hypothesis of unit root

$$T_n \to_d \left(\int_0^1 W(r)^2 dr\right)^{-1/2} \int_0^1 W(r)\, dW(r) \tag{5}$$

as $n \to \infty$. The limiting distribution appeared in (5), often called the Dickey-Fuller distribution, is nonnormal and skewed to the left. The unit root hypothesis is rejected if T_n takes a large negative value.

We now introduce a new asymptotics. Let $c > 0$ be a fixed number chosen arbitrarily. For each $n = 1, 2, \ldots$, let m_n be defined by

$$m_n = \inf_{k \geq 1} \left\{ \frac{1}{n^2} \sum_{t=1}^{k} y_{t-1}^2 \geq c \right\}, \tag{6}$$

and consider the t-ratio T_{m_n} for the sample of size m_n. Here the sample size m_n is determined by the squared sum of (y_t) achieving a certain level. Note that n no long denotes the sample size. Here it is used simply to define a sequence m_n.

Theorem 2.1 Suppose that (y_t) is generated as in (1) and (2) under Assumption 2.1, and that (m_n) is defined by (6). Then we have

$$T_{m_n} \to_d \mathbb{N}(0, 1) \tag{7}$$

as $n \to \infty$.

Unlike the conventional result in (5), our approach here yields the normal asymptotics given in (7). A few important remarks are now in order.

Remarks (a) The asymptotics in (5) and (7) are derived by taking two different contours: The former holds the sample size fixed with varying sums of squares, while the latter holds the sum of squares fixed with varying sizes of samples. They are useful in different contexts. The asymptotics in (5) is more relevant to the conventional approach, which evaluates the likelihood of a realized value of the statistic against other possible realizations from the samples of the same size. In contrast, our new approach suggests the evaluation of the likelihood of a realized value of the statistic against other possible realizations for the samples of the same sum of squares, and therefore, requires the asymptotics in (7). We may deduce from (5) that, across the samples of the same size fixed at some large n, the t-ratio has distribution close to the Dickey-Fuller distribution. On the other hand, (7) implies that the t-ratio is approximately standard normal if we look at all realizations of the samples of size m_n, where for arbitrarily given $c > 0$

$$\sum_{t=1}^{m_n} y_{t-1}^2 \approx n^2 c$$

with some large fixed n.

(b) For a given sample $y_1, \ldots, y_n$, we may set

$$c = \frac{1}{n^2} \sum_{t=1}^{n} y_{t-1}^2.$$

Obviously, we have in this case $m_n = n$ and $T_{m_n} = T_n$. The statistics T_n and T_{m_n} would then have identical value. Now the choice of the critical values is the choice of the contour, along which we would like to evaluate the likelihood of the given realization. Depending upon which contour we choose to evaluate the likelihood of a realized value for the statistic, the relevant null limit distribution and thus the critical value of the test would be different. If the realized value of the statistic is to be compared with all of its possible values obtained from the samples of the same size, the critical value from the Dickey-Fuller distribution should be used. If, on the other hand, the realized value of the statistic is to be compared with all possible values from the samples of the same sum of squares, the standard normal critical value should be used.

(c) The choice of the contour would ultimately be a subjective matter. However, we may say that it would be more appropriate to choose the contour representing the same amount of information on the hypothesis to be tested. In this regard, the contour of the equi-squared-sum is especially appealing for the test of a unit root. The most important and distinguishing charactersitic of the sample path from the unit root process (in comparison with that from the stationary process) is the presence of stochastic trend, and its magnitude can be effectively measured by the sum of squares. Choosing the contour of the equi-squared-sum for the unit root test thus implies that we assess the likelihood of a realized test value against other possible realizations having the stochastic trends of the same magnitude. This seems quite reasonable.

(d) The distinction and choice of the two contours are unimportant for the samples from stationary time series. The stationary samples yield the same sampling distributions for the two different contours considered here. For the stationary time series (y_t), $\sum_{t=1}^n y_{t-1}^2/n$ converges to a fixed constant as the sample size grows, due to the law of large numbers, making the two contours identical in large samples. However, the two contours can be very different for the samples from the unit root process. Most of all, the first contour is fixed and nonrandom, whereas the second contour is path-dependent. As is well known,

$$\frac{1}{n^2}\sum_{t=1}^n y_{t-1}^2 \to_d \int_0^1 W(r)^2\,dr$$

for the unit root process. The sum of squares, if normalized properly, would thus remain to be random and depend upon a realized value of the underlying process.

(e) Our asymptotics also help to analyze the nonnormality of the Dickey-Fuller distribution. We may clearly see from the proof of Theorem 2.1 that, for a stopping time τ such that $\int_0^\tau W(r)^2dr$ is constant, the distribution of $\int_0^\tau W(r)\,dW(r)/(\int_0^\tau W(r)^2dr)^{1/2}$ is standard normal. The nonnormality of the Dickey-Fuller distribution is due to the evaluation of the integral over the fixed interval $[0,1]$, rather than the random interval $[0,\tau]$, in the limiting t-ratio.

(f) Lai and Siegmund (1983) derive the asymptotics comparable to (7) for the case of iid (u_t).[3] For the Gaussian random walk, we may regard the new asymptotics as those

[3] We were not aware of Lai and Siegmund (1983) until very recently, so we could not mention their work in earlier versions of this paper. We are very grateful to Peter Hansen, Byungsoo So and Jim Stock for bringing their work to our attention. Their result, however, is applicable only for the pure random walk

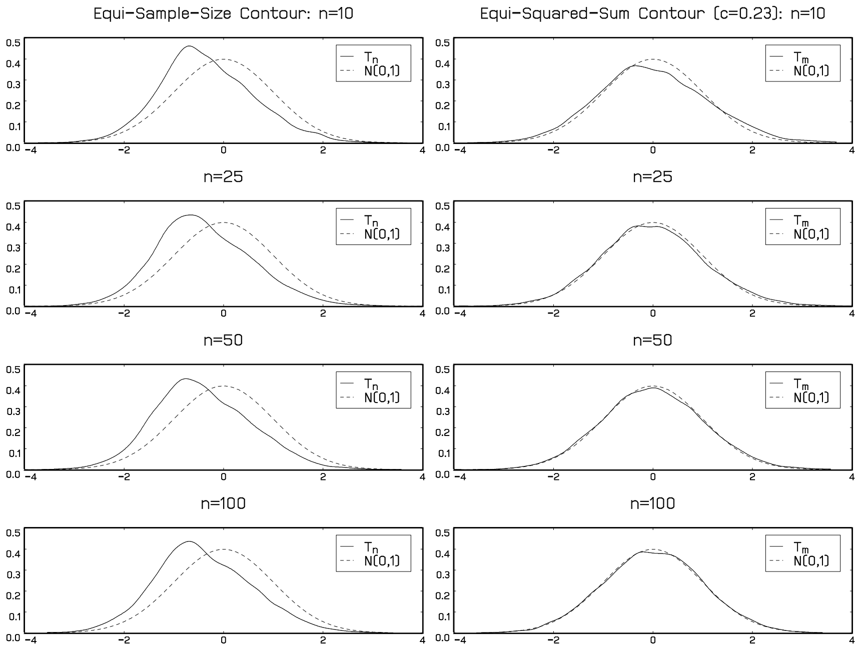

Figure 3: Densities of t-ratios from Equi-Sample-Size and Equi-Squared-Sum Contours

for a sequential sampling scheme which measures time in terms of accumulated Fisher information. Note that Fisher information about α contained in $y_1, \ldots, y_n$ is given by $\sum_{t=1}^{n} y_{t-1}^2$ in this case.

In Figure 3, the densities for the distributions of T_n and T_{m_n} are given and compared with the standard normal distribution. The densities of T_n are obtained for each of the fixed sample sizes $n = 10, 25, 50$ and 100, while the densities of T_{m_n} are computed for the fixed sum of squares given by $n^2 c$ with $n = 10, 25, 50, 100$ and $c = 0.23$. From simulations, we find that the asymptotic expected value of the stopping time τ defined by $\int_0^\tau W(r)^2 dr = c$ is approximately unity with this choice of c. The densities, in all cases, are quite insensitive to the choice for the value of c. Along the contour of the fixed sum of squares, the finite sample distribution of T_{m_n} appears to converge rather rapidly. Our normal asymptotics thus provide very good approximations for the finite sample distributions of T_{m_n}. Even for the samples with moderate sizes, the finite sample distributions are indeed quite close to standard normal. In contrast, the distributions of T_n are quite distinct from standard normal at all sample sizes.

We now consider the local alternative

$$\alpha = 1 - \frac{\delta}{n} \tag{8}$$

model driven by iid innovations. Moreover, their approach is not readily extended to the near unit root model and general unit root models with intercept and serially dependent innovations that we consider in the next section.

for some $\delta > 0$. It is well known that

$$T_n \to_d -\left(\int_0^1 W_\delta(r)^2 dr\right)^{1/2} \delta + \frac{\displaystyle\int_0^1 W_\delta(r)\, dW(r)}{\left(\displaystyle\int_0^1 W_\delta(r)^2 dr\right)^{1/2}} \tag{9}$$

as $n \to \infty$, where W_δ is the Ornstein-Uhlenbeck process given by $W_\delta(r) = \int_0^r \exp[-(r - s)\delta]dW(s)$. In contrast to the conventional asymptotics in (9), our asymptotics yield

Theorem 2.2 Suppose that (y_t) is generated as in (1) and (8) under Assumption 2.1, and that (m_n) is defined by (6). Then we have

$$T_{m_n} \to_d -\left(c/\sigma^2\right)^{1/2} \delta + \mathbb{N}(0,1) \tag{10}$$

as $n \to \infty$.

Our result in (10) establishes the new asymptotics for the t-ratio under the alternative hypothesis of local-to-unity. Along the contour of the fixed sum of squares, we would thus get the normal asymptotics under the local alternative, as well as under the null. Under the alternative of local-to-unity, the t-ratio has limit normal distribution with mean shifted by the product of the locality parameter and the square root of the level of the contour, measured in the unit of the innovation variance. If the new contour is taken, the asymptotic distribution under the local alternative is simply given by a parallel shift of the asymptotic null distribution, exactly as in the standard stationary regression models. This is in sharp contrast with the asymptotics along the conventional contour, which are given by (9).

3. Extensions

Our results in the previous section can be readily extended to more general unit root models. Here we consider models with intercept and models driven by serially dependent innovations. Upon appropriate modifications of the tests and contours, they yield the same new asymptotics as those for the simple unit root model considered earlier.

3.1 Models with Intercept

Now we let (y_t) be generated as

$$y_t = \mu + y_t^0, \tag{11}$$

where (y_t^0) follows the autoregressive process given in (1). The unit root in (y_t^0) may now be tested in the regression

$$y_t^\mu = \alpha y_{t-1}^\mu + u_t \tag{12}$$

using (y_t^μ) given by

$$y_t^\mu = y_t - y_0 \tag{13}$$

or

$$y_t^\mu = y_t - \frac{1}{t-1}\sum_{k=1}^{t-1} y_k, \tag{14}$$

which is defined recursively for each $t = 1, \ldots, n$. This recursive demeaning was first proposed by So and Shin (1999) to demean positively correlated stationary AR processes, and later used in Chang (2002) for the test of the unit root using the nonlinear instrumental variable methodology.

Denote by T_n^μ the t-ratio for the unit root hypothesis (2) in regression (12). The conventional limit distribution of T_n^μ is dependent upon the actual demeaning procedure that we introduce in (13) and (14). If (y_t^μ) given in (13) is used, then the limit distribution of T_n^μ is precisely the same as that of T_n given in (5) from regression (1) without intercept. On the other hand, if (y_t^μ) in (14) is used, then the conventional asymptotics would yield

$$T_n^\mu \to_d \left(\int_0^1 W^\mu(r)^2 dr\right)^{-1/2} \int_0^1 W^\mu(r)\, dW(r),$$

where

$$W^\mu(r) = W(r) - \frac{1}{r}\int_0^r W(s)ds$$

as $n \to \infty$.[4]

We define a new contour

$$m_n = \inf_{k\geq 1}\left\{\frac{1}{n^2}\sum_{t=1}^{k} y_{t-1}^{\mu\,2} \geq c\right\} \tag{15}$$

for each $n = 1, 2, \ldots$ and an arbitrarily given $c > 0$. Note that (m_n) is defined in (15) exactly as in (6), except that it is now based on the demeaned (y_t).

Corollary 3.1 Suppose that (y_t) is generated as in (11) with (y_t^0) given by (1) under Assumption 2.1. Under the unit root hypothesis (2), we have

$$T_{m_n}^\mu \to_d \mathbb{N}(0,1) \tag{16}$$

as $n \to \infty$. Moreover, we have under the alternative hypothesis of local-to-unity (8)

$$T_{m_n}^\mu \to_d -\left(c/\sigma^2\right)^{1/2}\delta + \mathbb{N}(0,1) \tag{17}$$

as $n \to \infty$.

Our previous results therefore also apply for the models with intercept, if a recursive demeaning procedure is used. For the normal asymptotics obtained in Corollary 3.1, it is

[4]As noted by Chang (2002), the process W^μ is well defined to be a continuous semimartingale if we set $W^\mu(0) = 0$. This can be readily shown by Brownian law of iterated logarithm [see, e.g., Revuz and Yor (1994, p53)].

important to use the contour in (15) defined with the demeaned (y_t). The normal asymptotics would not follow if the contour (6) based on (y_t) is used. Here we only consider the models with intercept to simplify the exposition. It is, however, clear that we may allow for other deterministic components in the model, as long as they can be removed recursively by using only the past information and the contour is defined appropriately.

3.2 General Unit Root Models

Not surprisingly, the new asymptotics we established in the previous section are also applicable for more general unit root models driven by a serially dependent nonmartingale difference sequence (u_t). For the process (u_t), we may consider two sets of assumptions.

Assumption 3.1 Let (u_t) satisfy invariance principle.

Assumption 3.2 Let $u_t = \pi(L)\varepsilon_t = \sum_{k=0}^{\infty} \pi_k \varepsilon_{t-k}$ and assume that (a) $(\varepsilon_t, \mathcal{F}_t)$ is a martingale difference sequence, with some filtration $(\mathcal{F}_t)$, such that $\mathbb{E}(\varepsilon_t^2) = \sigma^2$, $(1/n)\sum_{t=1}^{n} \varepsilon_t^2 \to_p \sigma^2$ and $\sup_{t \geq 1} \mathbb{E}|\varepsilon_t|^r < K$ for some $r \geq 4$ and constant $K > 0$, and that (b) $\pi(z) \neq 0$ for all $|z| \leq 1$ and $\sum_{k=0}^{\infty} |k|^s |\pi_k| < \infty$ for some $s \geq 1$.

Under Assumption 3.1, we may use the Phillips test to test for the unit root hypothesis in the general unit root model. It is based on the modified t-statistic

$$T_n^a = \frac{\hat{\sigma}}{\hat{\omega}} T_n - \frac{\hat{\omega}^2 - \hat{\sigma}^2}{2\hat{\omega}\left(\dfrac{1}{n^2}\displaystyle\sum_{t=1}^{n} y_{t-1}^2\right)^{1/2}},$$

where T_n is the standard t-ratio and $\hat{\sigma}^2$ and $\hat{\omega}^2$ are respectively any consistent estimates for the shortrun and longrun variances of (u_t). See Phillips (1987) and Stock (1994) for more details. Under Assumption 3.2, the so-called augmented Dickey-Fuller test can be applied. It relies on the unit root regression

$$y_t = \alpha y_{t-1} + \sum_{k=1}^{p-1} \alpha_k \triangle y_{t-k} + \varepsilon_t \tag{18}$$

augmented with the differenced lags, and tests the unit root hypothesis using the standard t-ratio. We denote by T_n^b the t-statistic for the unit root hypothesis (2) in regression (18). Said and Dickey (1984) show that the test is applicable for general invertible ARMA processes of unknown order if we set $p = Kn^r$ with some constant $K > 0$ and $0 < r \leq 1/3$. More recently, Chang and Park (2003) show that it is valid under the conditions in Assumption 3.2 as long as $p \to \infty$ with $p = o(n^{1/2})$.

As is well known, the statistics T_n^a and T_n^b have the same limiting distributions as T_n. For the general unit root model, we have

$$T_n^a, T_n^b \to_d \left(\int_0^1 W(r)^2 dr\right)^{-1/2} \int_0^1 W(r)\, dW(r) \tag{19}$$

as $n \to \infty$, respectively under Assumptions 3.1 and 3.2. This is exactly as in (5). Our new asymptotics extend readily to these statistics, as we show in the following corollary.

Corollary 3.2 Suppose that (y_t) is generated as in (1) and (m_n) is defined by (6). Also, let Assumptions 3.1 and 3.2 hold respectively for $T^a_{m_n}$ and $T^b_{m_n}$. Under the unit root hypothesis (2), we have

$$T^a_{m_n},\, T^b_{m_n} \to_d \mathbb{N}(0,1)$$

as $n \to \infty$. Moreover, we have under the alternative hypothesis of local-to-unity (8)

$$T^a_{m_n},\, T^b_{m_n} \to_d -\left(c/\omega^2\right)^{1/2} \delta + \mathbb{N}(0,1)$$

as $n \to \infty$.

The new asymptotics for the unit root tests in the general unit root models are essentially identical to those in the simple unit root model considered in the previous section. They are the same as the earlier ones under the null hypothesis of unit root. The asymptotics under the alternative hypothesis of local-to-unity are also identical, except that the level of the contour appearing in the shift parameter is now measured in the unit of the longrun variance of the innovations.

4. Conclusion

In this paper, we develop new asymptotics for the unit root tests that are commonly used in practical applications. Our asymptotics take a new contour given by the fixed sum of squares, and contrast with the conventional ones which evaluate the likelihood of a realized value of the test along the contour of the fixed sample size. We show in the paper that if the equi-squared-sum contour is chosen the distribution theories for the tests are normal. They have normal limiting distributions, and we may use the standard normal table for their critical values. As is well known, their conventional asymptotics are nonstandard and nonnormal. Our theories developed in this paper make it clear that we may legitimately use the standard normal table for many of the commonly used unit root tests. It would lead us *not* to making an invalid inference, *but* to exploring a new contour that has never been uncovered.

Appendix: Proofs of Theorems

Proof of Theorem 2.1 Let W_n be defined as in (3). It is well known that $W_n \to_d W$ in the space $D(\mathbb{R})$ of cadlag functions endowed with the supremum norm. Moreover, by extending the underlying probability space if necessary, we may assume that W_n and W are defined in the same probability space and that $W_n \to_{a.s} W$ uniformly. Such a construction is possible for instance by the Skorohod embedding. See Hall and Heyde (1980) for details.

For any fixed constant $c > 0$, we let $\tau_n(c)$ be given by

$$\sigma^2 \int_0^{\tau_n(c)} W_n(r)^2 dr = c, \tag{20}$$

and define a stopping time $\tau(c)$ to be such that

$$\sigma^2 \int_0^{\tau(c)} W(r)^2 dr = c. \tag{21}$$

Since $W_n \to_{a.s.} W$ uniformly, it follows from (20) and (21) that

$$\tau_n(c) \to_{a.s.} \tau(c)$$

as $n \to \infty$. Moreover, upon noticing $\tau_n(c) = m_n/n + O(n^{-1})$ a.s., we may further deduce that

$$\frac{m_n}{n} \to_{a.s.} \tau(c)$$

as $n \to \infty$.

Under the null hypothesis of unit root (2), we have

$$\begin{aligned} T_{m_n} &= (1/\hat{\sigma}) \left(\sum_{t=1}^{m_n} y_{t-1}^2 \right)^{-1/2} \sum_{t=1}^{m_n} y_{t-1} u_t \\ &= \left(\int_0^{m_n/n} W_n(r)^2 dr \right)^{-1/2} \int_0^{m_n/n} W_n(r)\, dW_n(r) + o_p(1) \\ &= \left(\int_0^{\tau(c)} W(r)^2 dr \right)^{-1/2} \int_0^{\tau(c)} W(r)\, dW(r) + o_p(1) \end{aligned} \tag{22}$$

as $n \to \infty$, since $W_n \to_{a.s.} W$ uniformly, $m_n/n \to_{a.s.} \tau(c)$ and $\hat{\sigma}^2 \to_p \sigma^2$ as $n \to \infty$. Therefore, it now suffices to show that

$$T = \left(\int_0^{\tau(c)} W(r)^2 dr \right)^{-1/2} \int_0^{\tau(c)} W(r)\, dW(r) \sim \mathbb{N}(0,1), \tag{23}$$

due to (22).

To establish (23), we first define a continuous martingale

$$M(s) = \sigma \int_0^s W(r)\, dW(r)$$

and note that it has the quadratic variation given by

$$[M](s) = \sigma^2 \int_0^s W(r)^2 dr.$$

Therefore, we have

$$\tau(s) = \inf_{r>0} \{[M]_r \geq s\},$$

from which it follows that

$$V(s) = M(\tau(s)) = \sigma \int_0^{\tau(s)} W(r)\, dW(r) \tag{24}$$

is the DDS Brownian motion of the martingale M. The reader is referred to, e.g., Revuz and Yor (1994) for the DDS Brownian motion. Due to (21), (23) and (24), we now have

$$T = \left(\frac{c}{\sigma^2}\right)^{-1/2} \left(\frac{V(c)}{\sigma}\right) \sim V(1),$$

from which (23) follows immediately. The proof is therefore complete. □

Proof of Theorem 2.2 The proof is analogous to that of Theorem 2.1. We define a stochastic process $W_{n\delta}$ by

$$W_{n\delta}(r) = \frac{1}{\sigma\sqrt{n}}\, y_{[nr]}$$

for $r \geq 0$. Under Assumption 2.1 and the alternative of local-to-unity (8), we have $W_{n\delta} \to_d W_\delta$ uniformly in $D(\mathbf{R})$, where $D(\mathbb{R})$ is defined as in the proof of Theorem 2.1. This is well known. If we define $\tau_\delta(c)$ by

$$\sigma^2 \int_0^{\tau_\delta(c)} W_\delta(r)^2 dr = c \tag{25}$$

for any fixed constant $c > 0$, then $m_n/n \to_{a.s.} \tau_\delta(c)$ exactly as in the proof of Theorem 2.1.

Under the alternative of local-to-unity (8), we have

$$\begin{aligned} T_{m_n} &= -(1/\hat{\sigma}) \left(\sum_{t=1}^{m_n} y_{t-1}^2\right)^{1/2} \frac{\delta}{n} + (1/\hat{\sigma}) \left(\sum_{t=1}^{m_n} y_{t-1}^2\right)^{-1/2} \sum_{t=1}^{m_n} y_{t-1} u_t \\ &= -\left(\int_0^{m_n/n} W_{n\delta}(r)^2 dr\right)^{1/2} \delta + \left(\int_0^{m_n/n} W_{n\delta}(r)^2 dr\right)^{-1/2} \int_0^{m_n/n} W_{n\delta}(r)\, dW_n(r) \\ &= -\left(\int_0^{\tau_\delta(c)} W_\delta(r)^2 dr\right)^{1/2} \delta + \left(\int_0^{\tau_\delta(c)} W_\delta(r)^2 dr\right)^{-1/2} \int_0^{\tau_\delta(c)} W_\delta(r)\, dW(r) + o_p(1) \end{aligned} \tag{26}$$

as $n \to \infty$. We now consider the DDS Brownian motion

$$V_\delta(s) = \sigma \int_0^{\tau_\delta(s)} W_\delta(r)\, dW(r)$$

of the continuous martingale

$$M_\delta(s) = \sigma \int_0^s W_\delta(r)\, dW(r),$$

from which the stated result follows immediately, due to (25) and (26). □

Proof of Corollary 3.1 The proof is entirely analogous to those of Theorems 2.1 and 2.2. The details are therefore omitted. □

Proof of Corollary 3.2 We first consider the Phillips test. We define

$$W_n(r) = \frac{1}{\omega\sqrt{n}} \sum_{t=1}^{[nr]} u_t,$$

and assume $W_n \to_{a.s.} W$ similarly as in the proof of Theorem 2.1. Moreover, let the stopping time $\tau(c)$ be defined as

$$\omega^2 \int_0^{\tau(c)} W(r)^2 dr = c$$

for any fixed constant $c > 0$. Under the null hypothesis of unit root (2), we have

$$\begin{aligned} T^a_{m_n} &= (1/\hat{\omega}) \left(\frac{1}{n^2} \sum_{t=1}^{m_n} y^2_{t-1} \right)^{-1/2} \left[\frac{1}{n} \sum_{t=1}^{m_n} y_{t-1} u_t - \frac{1}{2} \left(\hat{\omega}^2 - \hat{\sigma}^2 \right) \right] \\ &= \left(\int_0^{m_n/n} W_n(r)^2 dr \right)^{-1/2} \int_0^{m_n/n} W_n(r)\, dW_n(r) + o_p(1) \\ &= \left(\int_0^{\tau(c)} W(r)^2 dr \right)^{-1/2} \int_0^{\tau(c)} W(r)\, dW(r) + o_p(1) \end{aligned}$$

as $n \to \infty$. The stated result now follows exactly as in the proof of Theorem 2.1.

Under the alternative of local-to-unity (8), we define

$$W_{n\delta}(r) = \frac{1}{\omega\sqrt{n}} y_{[nr]}$$

and let $W_{n\delta} \to_{a.s.} W$ as in the proof of Theorem 2.2. Moreover, define the stopping time $\tau_\delta(c)$ by

$$\omega^2 \int_0^{\tau_\delta(c)} W_\delta(r)^2 dr = c.$$

Then it follows that

$$\begin{aligned} T^a_{m_n} &= -(1/\hat{\omega}) \left(\sum_{t=1}^{m_n} y^2_{t-1} \right)^{1/2} \frac{\delta}{n} + (1/\hat{\omega}) \left(\frac{1}{n^2} \sum_{t=1}^{m_n} y^2_{t-1} \right)^{-1/2} \left[\frac{1}{n} \sum_{t=1}^{m_n} y_{t-1} u_t - \frac{1}{2} \left(\hat{\omega}^2 - \hat{\sigma}^2 \right) \right] \\ &= -\left(\int_0^{m_n/n} W_{n\delta}(r)^2 dr \right)^{1/2} \delta + \left(\int_0^{m_n/n} W_{n\delta}(r)^2 dr \right)^{-1/2} \int_0^{m_n/n} W_{n\delta}(r)\, dW_n(r) + o_p(1) \\ &= -\left(\int_0^{\tau_\delta(c)} W_\delta(r)^2 dr \right)^{1/2} \delta + \left(\int_0^{\tau_\delta(c)} W_\delta(r)^2 dr \right)^{-1/2} \int_0^{\tau_\delta(c)} W_\delta(r)\, dW(r) + o_p(1) \end{aligned}$$

as $n \to \infty$. The rest of the proof is exactly identical to the proof of Theorem 2.2. This completes the proof for the Phillips test.

The proof of the stated result for the augmented Dickey-Fuller statistic $T_{m_n}^b$ is virtually identical, given the asymptotic theories provided by Chang and Park (2002). The details are therefore omitted to save the space. $\square$

References

Abadir, K.M. (1993). “The limiting distribution of the autocorrelation coefficient under a unit root,” *Annals of Statistics*, 21, 1058-1070.

Chang, Y. (2002). “Nonlinear IV unit root tests in panels with cross-sectional dependency,” *Journal of Econometrics*, 110, 261-292.

Chang, Y. and J.Y. Park (2002). “On the asymptotics of ADF tests for unit roots,” *Econometric Reviews*, 21, 431-447.

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

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

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

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

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

Hall, P. and C.C. Heyde (1980). *Martingale Limit Theory and Its Application.* Academic Press: New York.

Lai, T.L. and D. Siegmund (1983). “Fixed accuracy estimation of an autoregressive parameter,” *Annals of Statistics* 11, 478-485.

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

Said, S.E. and D.A. Dickey (1984). “Testing for unit roots in autoregressive-moving average models of unknown order,” *Biometrika* 71, 599-608.

So, B.S. and D.W. Shin (1999). “Recursive mean adjustment in time-series inferences,” *Statistics and Probability Letters*, 43, 65-73.

Stock, J.H. (1994). “Unit roots and structural breaks,” In R.F. Engle and D. McFadden, eds., *Handbook of Econometrics*, Vol. 4, 2739-2841, Elsevier: Amsterdam.

Revuz, D. and M. Yor (1994). *Continuous Martingale and Brownian Motion*, 2nd ed. Springer-Verlag: New York.