Taking a New Contour: A Novel Approach to Panel Unit Root Tests

Yoosoon Chang
and
Joon Y. Park
Department of Economics
Rice University [1]

Abstract

The paper introduces a novel approach to statistical inference in the nonstationary panels. The approach takes a new contour on which the nonstationary models yield standard normal asymptotics free of nuisance parameters. The new contour is drawn along the line given by the equi-squared-sum, instead of the traditional one given by the equi-sample-size. Along the new contour, we show that the distributions of commonly used unit root tests are normal in large samples. The normal asymptotics hold under both the null of a unit root and the local-to-unity alternative, and for the IV t-ratio as well as the usual t-ratio. Moreover, they are applicable also for the models with intercept or linear time trend, as long as is used the demeaning or detrending method relying only on the past information. Subsequently, we demonstrate that this startling finding may be exploited to invent tools and methodologies for the effective inferences in nonstationary panel context. In particular, our theory implies that the individual tests may be viewed as asymptotically normal samples if they are computed using the samples which have the same sum of squares across all cross-sectional units. Consequently, we may use various functionals of those individual tests to do valid inferences in nonstationary panels.

First Draft: August 10, 2002
This version: March 11, 2004

JEL Classification: C12, C15, C32, C33.

Key words and phrases: Equi-squared-sum contour, dependent panels, heterogeneity, unit root test, Dickey-Fuller distribution, Brownian motion, continuous martingale, time change, recursive demeaning

---

[1] The paper was prepared for the presentation at Panel Workshop held at Cass Business School, London, UK, on March 15, 2004. We are very grateful to the workshop organizer Giovanni Urga for the invitation, and to Peter Pedroni and Donggyu Sul for helpful discussions and comments. Correspondence address to: Yoosoon Chang, Department of Economics - MS 22, Rice University, 6100 Main Street, Houston, TX 77005-1892, Tel: 713-348-2796, Fax: 713-348-5278, Email: yoosoon@rice.edu.

1. Introduction

Inference in nonstationary panels has recently drawn much attention as more panel datasets covering long time spans became available. In particular, the unit root tests in panels have been frequently used by numerous authors to investigate many important economic interrelationships among countries or regions, which may imply convergence, divergence or parity of some sets of economic variables. The growth convergence and the purchasing power parity of exchange rates are prime examples, and routinely examined by the unit root tests applied to the time series observations across the countries or regions under investigation. The unit root tests that are applicable for the panel data have been developed by many authors including Levin, Lin and Chu (2002), Quah (1994), Im, Pesaran and Shin (2003), Maddala and Wu (1999), Choi (2001a, 2001b), Chang (1999, 2002), Chang and Song (2002), Moon and Perron (2001), Phillips and Sul (2001), Bai and Ng (2002) and Pesaran (2003). See, e.g., Phillips and Moon (2000) and Baltagi and Kao (2000) for surveys on the recent development of the unit root tests in panels.

All of the existing panel unit root tests combine some transformations of the individual unit root statistics obtained across cross-sectional units. Quite naturally, under the assumption of cross-sectional independence, the appropriately transformed and standardized combinations of individual tests are well expected to be approximately normal as the number of cross sections N gets large. Indeed, virtually all the panel unit root tests rely on normal N-limit theory, whose validity requires the appropriate normalizations of individual tests, as well as the asymptotic cross-sectional independence. It is therefore fair to say that a good panel test is the one which efficiently normalizes the individual tests and effectively deals with cross-sectional dependency. In this paper, we concentrate more on efficient normalization of the individual tests. The problem of cross-sectional dependency was first addressed in Chang (2002), and dealt with the methodology based on nonlinear instrumental variable estimation. Subsequent authors introduced common factors in their models to allow for cross-sectional dependency.

The efficient normalization of the individual tests is much more difficult than one might think. As is well known, the individual unit root tests have null distributions that are nonstandard and nonnormal. Their time T-asymptotics yield distributions commonly represented by various functionals of Brownian motions, and in particular, known to be asymmetric and skewed. See, e.g., Fuller (1996) for the tabulations of them. Consequently, the standardization through the mean and variance adjustment or the p-value transformation, which are two most frequently used methods for normalization, often works poorly even if T is relatively large. Worse, the errors made in the normalizations for individual tests are accumulated as N of them are combined to compute the panel unit root test. Obviously, the problem gets worse as N increases. We require, however, N tend to infinity to obtain the normal N-asymptotics. This is a serious dilemma. To overcome this difficulty, we take a totally different method in this paper: A novel approach based on a new contour to achieve efficient normalization of the individual tests.

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 along

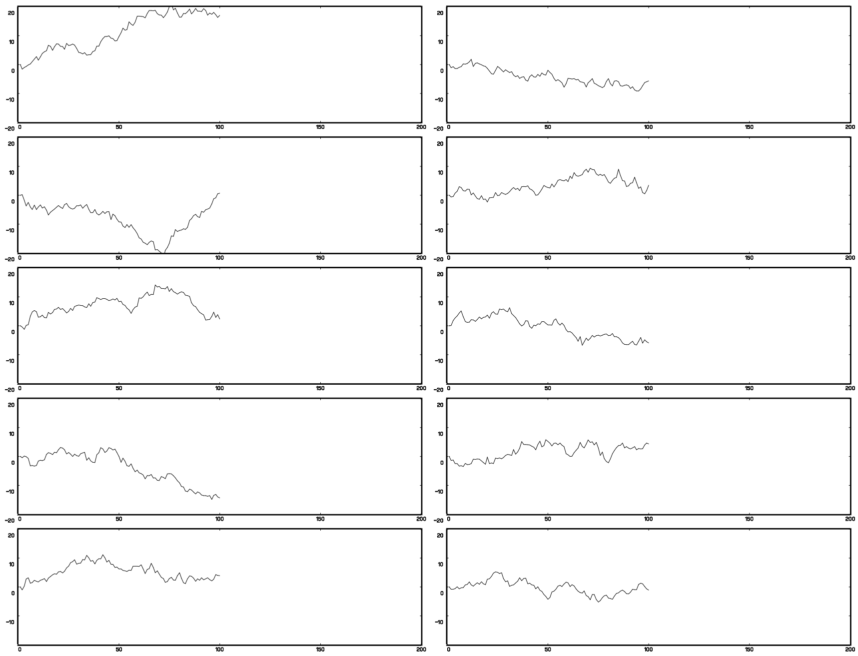

Figure 1: Sample Paths with Equal Sample Size

the contour given by the fixed sample size. In this paper, we suggest to take 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. For the observations from stationary time series, the sum of squares becomes a constant multiple of the sample size for large samples. 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. 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. For the unit root samples, it would thus yield a new meaning 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 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

[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.

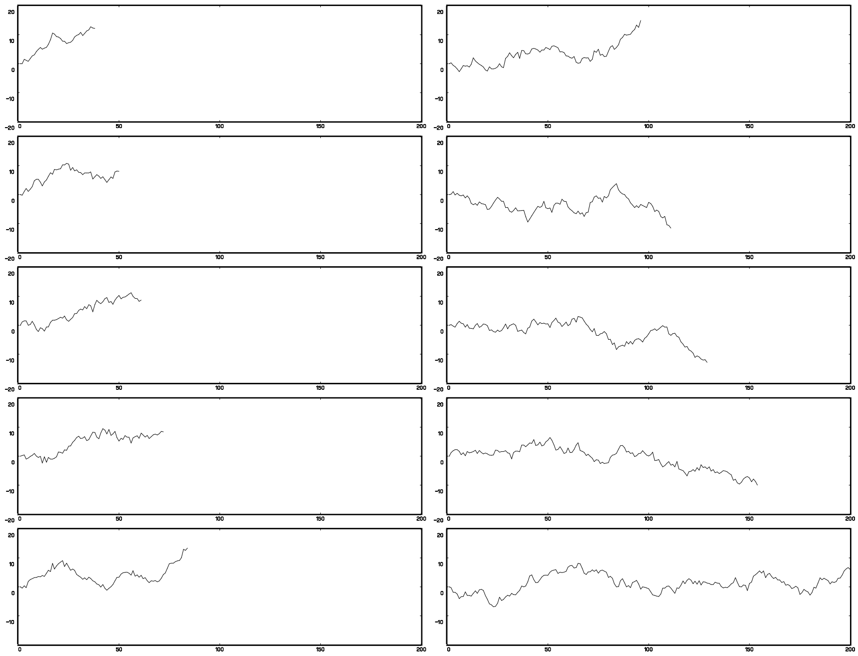

Figure 2: Sample Paths with Equal Sum of Squares

contours we may take to obtain the sampling distributions of the statistics involving unit root processes. Here the choice is whether to look at other possible realizations along the contour of the samples *either* of fixed sample size (with varying sums of squares as required to have the same sample size) as in Figure 1, *or* of fixed sum of squares (with varying sample sizes as required to have the same sum of squares) as in Figure 2.

Both the size and the sum of squares represent the information contents in the sample on population. Needless to say, the larger dataset and the dataset with larger sampling variations would help us perform more precise inference on the underlying data generating mechanism. The asymptotics, however, become completely different depending upon which contour we take. In sharp contrast with the conventional contour of the equi-sample-size, the contour of the equi-squared-sum yields normal asymptotics and conventional statistical theories for the unit root tests. Indeed, we show in the paper that the asymptotics obtained for large sum of squares, along the contour of the equi-squared-sum, are normal. The critical values for the t-ratio can therefore be obtained from the standard normal table, and all other relevant statistical theories both under the null of a unit root and under the alternative of local-to-unity just follow exactly as in the standard regression model given in the elementary econometrics textbook. This is true for models with fitted mean as well, as long as they are removed properly using only the past information for each observation. Our normal asymptotics also well extend to the nonlinear IV approach by Phillips, Park and Chang (2004).

The new asymptotics along the contour of the equi-squared-sum we derive have very important and far-reaching implications for inference in nonstationary panels. For the independent panels, it would imply that the individual unit root tests behave asymptotically as if they were independent and identically distributed standard normals, if we observe their values along the contour of the equi-squared-sum. That is, if we set the sum of squares to be the same across all cross-sectional units, the resulting individual unit root tests can be regarded as standard normal samples. This rather startling result would certainly give us a great opportunity to do effective inferences in nonstationary panels. For instance, we may now use the order statistics such as the maximum or minimum to test the null hypothesis that all (some) cross-sections have unit roots against the alternative that some (all) do not. Under the equi-squared-sum scheme, the individual tests have normal T-asymptotics for each N, regardless of being small or large, or letting it be fixed or tend to infinity. We may indeed rely on standard normal distribution theory for both T, N or their joint asymptotics if we simply aggregate the individual test statistics along the new contour.

There is one special case where two contours become identical. This is when the usual sign function is used as the instrument generating function. The resulting estimator, which is often called the Cauchy estimator, has the fixed sum of squares given by the sample size for all its realizations. The Cauchy t-ratio therefore has the standard normal T-asymptotics, along the contours of both the equi-squared-sum and the equi-sample-size. This is rather important, since usually the data are collected along the contour of the equi-sample-size. When the data are given for a fixed sample size, to do inference along the contour of the equi-squared-sum necessarily implies that we do not use some part of our data. For the standard t-ratio, for instance, we must discard some of our observations to take new contour. This, however, is unnecessary for the Cauchy t-ratio, and we may utilize full samples. Therefore, under the usual circumstances that the individual cross-sectional units have the equal number of time series observations, the use of the Cauchy t-ratio is preferred. In this case, the panel unit root test based on the Cauchy t-ratio may perform better than that based on the standard t-ratio, though individually the OLS estimator is known to be more efficient than the Cauchy estimator. In fact, this is what we found in our simulations.

The rest of the paper is organized as follows. In Section 2, we derive the main results of the paper for the simple unit root test. There we introduce the unit root model and the test statistic, and develop new asymptotics for the unit root test along the contour of the equi-squared-sum. The asymptotics are shown to be normal. Section 3 extends our main results into several directions. In particular, it is shown that our main results continue to hold under the local alternatives and for the models with intercept and linear trend. The Cauchy and other nonlinear instrumental variable estimators are also considered. Section 4 demonstrates that our normal asymptotics along the new equi-squared-sum contour may be exploited to construct tools and methodologies for effective inferences in nonstationary panels. Also discussed are various issues on the inference in nonstationary panels such as cross-sectional dependencies, heterogeneities, and formulations of hypotheses. Order statistics for testing more flexible forms of hypotheses are also considered. Finite sample performance of the newly proposed tests are evaluated via a set of simulations in Section 5. The concluding remarks are given in Section 6, and the mathematical proofs are given in Appendix.

A word on notation. As usual, $\to_d$ and $\to_{a.s.}$ are used to signify respectively the convergence in distribution 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 for the Simple Unit Root Test

We consider the simple AR(1) model

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

and the test of the unit root hypothesis

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

We assume that (ε_t) are white noise with zero mean and unit (known) second moment. The assumptions are far from being necessary. They are introduced here simply to avoid unnecessary complications and focus on the main issue of the paper. The unknown second moment can easily be estimated consistently from the fitted residuals. Moreover, we may consider more general models, i.e., the 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). They all have the same large sample distributions as the test we consider explicitly in the paper, and therefore, for them our subsequent discussions are also applicable. See, e.g., Stock (1994) for the test of a unit root in general models.

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

$$T_n = \frac{\hat{\alpha}_n - 1}{s(\hat{\alpha}_n)} \tag{3}$$

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 $s(\hat{\alpha}_n) = (\sum_{t=1}^n y_{t-1}^2)^{-1/2}$. It is well known that under the null hypothesis of unit root

$$T_n \to_d \left(\int_0^1 W(r)^2 dr \right)^{-1/2} \int_0^1 W(r)\, dW(r) \tag{4}$$

as $n \to \infty$. The limiting distribution appeared in (4), 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. For any given $x > 0$, let $m \geq 1$ be such that

$$m = \inf_{k \geq 1} \left\{ \sum_{t=1}^k y_{t-1}^2 \geq x \right\} \tag{5}$$

and consider the t-ratio T_m for the sample of size m. Here the sample size m is determined by the squared sum of (y_t) achieving a certain level. Note that m is a function of (y_t) as well as x.

Theorem 2.1 Assume (1) and (2). If we let m_n be defined as in (5) with $x = n^2 c$ for each $n \geq 1$ and some fixed constant $c > 0$, then

$$T_{m_n} \to_d \mathbf{N}(0,1) \tag{6}$$

as $n \to \infty$.

Unlike the conventional result in (4), our approach here yields the normal asymptotics given in (6). A few remarks are now in order.

Remark 2.2 (a) For the choice of

$$c = \frac{1}{n^2} \sum_{t=1}^{n} y_{t-1}^2 \tag{7}$$

we have $m_n = n$ and $T_{m_n} = T_n$. The statistics T_n and T_{m_n} would then have identical values, and may thus be regarded as the same statistic. The large sample distributions in (4) and (6) are derived just by taking two different contours in evaluating the likelihood of a realized value for the statistic. The distribution in (4) is obtained by the conventional approach holding the sample size constant. On the other hand, our new approach yields the distribution in (6) assuming the sum of squares to be constant. The likelihood of a realized value for the statistic is evaluated against other possible realizations from the samples of the same size (with varying sums of squares) and of the same sum of squares (with varying sample sizes), respectively in (4) and (6).

(b) The samples from stationary time series would produce the same sampling distributions for the two different contours considered above. For the stationary samples, $\sum_{t=1}^{n} y_{t-1}^2/n$ 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.

(c) Depending upon which contour we choose to evaluate the likelihood of a realized value for the statistic, the relevant null 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 given by the samples of the same sum of squares, the standard normal critical value should be used.

(d) 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

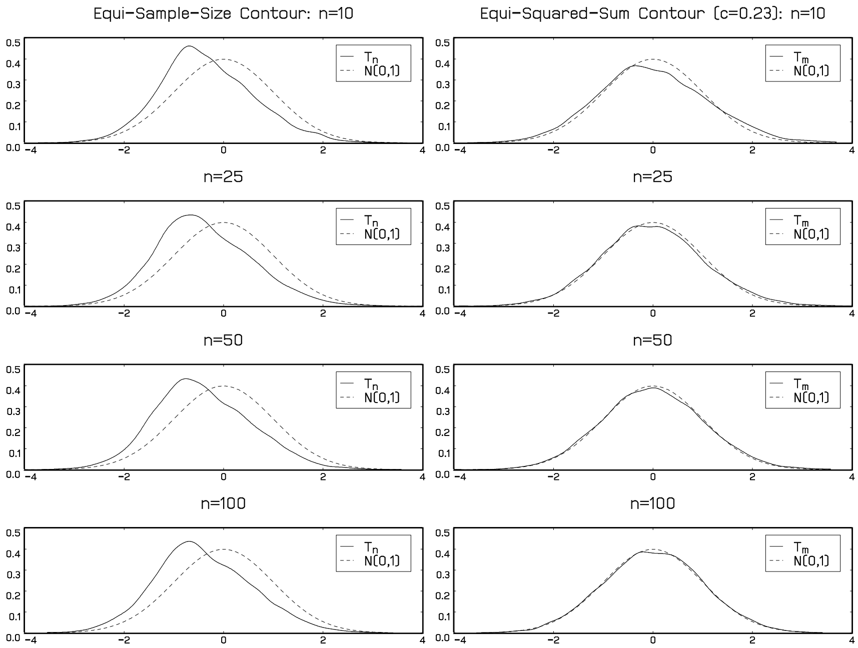

Figure 3: Densities of t-ratios from Equi-Sample-Size and Equi-Squared-Sum Contours

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 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 magnitudes. This seems quite reasonable.

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

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 of 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 moderate size samples, 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.

3. Extensions to More General Models

3.1 Distributions under Local Alternatives

We now consider the local alternative

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

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

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

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

Corollary 3.1 Assume (8). If we let m_n be defined as in (5) with $x = n^2 c$ for each $n \geq 1$ and some fixed constant $c > 0$, then

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

as $n \to \infty$.

If we take the contour of the fixed sum of squares, we would thus get the standard normal limiting distribution theory under both the null and alternative hypotheses. The unit root t-ratio is distributed as standard normal under the null in large samples. Moreover, it is also normal in large samples under the local alternative, with mean shifted by a constant multiple of the locality parameter. Note that the constant c in (10) is given by (7) for the sample of size n. Against the local alternatives, the unit root test is expected to have more powers for the samples with large sums of squares. As is evidently seen from (10) and (7), the large sum of squares has a magnifying effect on the locality parameter.

3.2 Models with Intercept and Time Trend

Our normal asymptotics on the contour of the equi-squared-sum extend well to the models with intercept and linear time trend, if it is removed effectively by using only the past information. In this case, the unit root along the new contour can be based on the regression

$$\triangle \bar{y}_t = (\alpha - 1)\bar{y}_{t-1} + \varepsilon_t \tag{11}$$

where $(\triangle\bar{y}_t)$ and $(\bar{y}_{t-1})$ are demeaned or detrended $(\triangle y_t)$ and (y_{t-1}) that are defined more precisely below. More dynamics can be introduced and AR(p), instead of AR(1) in (11), can be used as we explain later.

First, we look at the model with intercept. To test for the unit root in (y_t) generated as

$$y_t = \mu + y_t^\circ$$

where (y_t°) follows the autoregressive process given in (1), we use (y_t^μ) given by

$$y_t^\mu = y_t - y_0 \tag{12}$$

or

$$y_t^\mu = y_t - \frac{1}{t-1}\sum_{k=1}^{t-1} y_k \tag{13}$$

which is defined recursively for each $t = 1, \ldots, n$. This recursive demeaning was first proposed by So and Shin (1999a) to demean positively correlated stationary AR processes,[3] and later used in Chang (2002) and Phillips, Park and Chang (2004) for the test of the unit root using the nonlinear instrumental variable methodology.

The test for the unit root in (y_t°) can be based on the regression (11) with $\bar{y}_{t-1} = y_{t-1}^\mu$ and $\triangle\bar{y}_t = \triangle y_t$. The conventional limit distribution of the t-ration T_n^μ for the unit root hypothesis in (11) is dependent upon the actual demeaning procedure that we introduce in (12) and (13). If (y_t^μ) given in (12) is used, then the limit distribution of T_n^μ is precisely the same as T_n without intercept given in (4). On the other hand, if (y_t^μ) in (13) 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]

Now we consider the model with linear time trend, which we write

$$y_t = \mu + \nu t + y_t^\circ$$

[3]They found that the recursive demeaning reduces the biases of the parameter estimators.

[4]It follows from the well known Brownian law of iterated logarithm [see, for example, Revuz and Yor (1994, p.53)] that

$$\frac{1}{r}\int_0^r W(s)ds = O(r^{3/2}(\log\log(1/r))^{1/2}) \text{ a.s.}$$

and therefore

$$\frac{1}{r}\int_0^r W(s)ds \to 0 \text{ as } r \to 0$$

The process $W^\mu(r)$ then becomes a continuous stochastic processes defined on $[0,\infty)$, if we let $W^\mu(0) = 0$.

The recursive detrending of (y_t) can be done to obtain

$$y_t^\tau = y_t - y_0 - \sum_{k=1}^t \frac{1}{k}(y_k - y_0) \tag{14}$$

or

$$y_t^\tau = y_t + \frac{2}{t}\sum_{k=1}^t y_k - \frac{6}{t(t+1)}\sum_{k=1}^t k y_k \tag{15}$$

There can be many other alternatives. The regression (11) may now be fitted with $\bar{y}_{t-1} = y_{t-1}^\tau$ and $\triangle\bar{y}_t = \triangle y_t - (y_n - y_0)/n$.

If we denote by T_n^τ the t-ratio for the unit root hypothesis in regression (11), then we have under the conventional asymptotics

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

as $n \to \infty$, where W^τ is given by

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

or

$$W^\tau(r) = W(r) + \frac{2}{r}\int_0^r W(s)\, ds - \frac{6}{r^2}\int_0^r sW(s)\, ds$$

respectively for (y_t^τ) given in (14) or (15).[5]

We now define a new contour

$$m = \inf_{k\geq 1}\left\{\sum_{t=1}^k \bar{y}_{t-1}^2 \geq x\right\} \tag{16}$$

similarly as in (5), where $\bar{y}_{t-1} = y_{t-1}^\mu$ or y_{t-1}^τ respectively for the models with intercept and linear time trend. Then we have

Corollary 3.2 Assume (1) and (2). If we let m_n be defined as in (16) with $x = n^2 c$ for each $n \geq 1$ and some fixed constant $c > 0$, then

$$T_{m_n}^\mu, T_{m_n}^\tau \to_d \mathbf{N}(0,1)$$

as $n \to \infty$.

Our previous results therefore also apply for the models with intercept and linear time trend. To obtain the normal asymptotics for the models with fitted mean, however, it is important to use only the past information. The normal asymptotics on the contour of the equi-squared-sum do not follow if the usual demeaning or detrending is used.

[5]Exactly as in the previous footnote, the processes introduced here are well defined to be continuous processes if we set them zero at the origin.

3.3 Tests Using Nonlinear IV Estimators

The unit root hypothesis may also be tested using an IV estimator. Recently, Phillips, Park and Chang (2004) consider the IV estimator $\hat{\alpha}_n^{IV} = \left(\sum_{t=1}^n y_{t-1}F(y_{t-1})\right)^{-1} \sum_{t=1}^n F(y_{t-1})y_t$ of α, which uses $F(y_{t-1})$ as an instrument for some instrument generating function $F : \mathbb{R} \to \mathbb{R}$. If we denote by $s(\alpha_n^{IV}) = \left(\sum_{t=1}^n y_{t-1}F(y_{t-1})\right)^{-1} \left(\sum_{t=1}^n F(y_{t-1})^2\right)^{1/2}$ the standard error of the IV estimator, the IV t-ratio T_n^{IV} is given by

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

which reduces to

$$\left(\sum_{t=1}^n F(y_{t-1})^2\right)^{-1/2} \sum_{t=1}^n F(y_{t-1})\varepsilon_t$$

under the null hypothesis of a unit root.

As shown earlier by Phillips, Park and Chang (2004), the asymptotic behavior of the IV t-ratio T_n^{IV} depends crucially on the type of the instrument generating function F, in particular whether it is integrable or asymptotically homogeneous. The IV t-ratio T_n^{IV} with integrable F is thoroughly explored in Chang (2002) for the test of unit roots in panels under cross-sectional dependency. Therefore, here we only consider T_n^{IV} with an asymptotically homogeneous function F. An asymptotically homogeneous function F may be written as

$$F(\lambda x) = \nu(\lambda)H(x) + o(\nu(\lambda))$$

for large λ uniformly in $x \in \mathbb{R}$ over any compact interval, where we call ν the asymptotic order and H the limit homogeneous function of F.[6] It is shown in Phillips, Park and Chang (2004) that

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

along the conventional contour. The limit distribution of T_n^{IV} with an asymptotically homogeneous instrument generating function F is generally nonnormal.

Of course, the OLS estimator $\hat{\alpha}_n$ belongs to the class of the IV estimators considered here. It is easy to see that the OLS estimator is an IV estimator with the instrument generating function $F(x) = x$, i.e., the identity function. Phillips, Park and Chang (2004) show that the OLS estimator is indeed the most efficient IV estimator. Naturally, the usual t-ratio T_n based on the OLS estimator is therefore expected to be most powerful. However, under the local alternatives introduced in (8), we have for any IV t-ratio with asymptotically homogeneous F

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

[6] The reader is referred to Park and Phillips (2004) for the details.

Therefore, as long as $xH(x) \geq 0$ for all $x \in \mathbb{R}$ and $xH(x) > 0$ for all x in a neighborhood of the origin, T_n^{IV} would have nontrivial powers under the local alternatives shrinking to unity at the rate of n^{-1}.[7] Note that

$$\int_0^1 W_\delta(r)H(W_\delta(r))\,dr = \int_{-\infty}^{\infty} xH(x)L_\delta(1,x)\,dx$$

where $L_\delta(r,x)$ is the local time of W_δ at time r and spatial point x. This follows directly from the application of occupation times formula.

If we redefine

$$m = \inf_{k\geq 1}\left\{\sum_{t=1}^{k} F(y_{t-1})^2 \geq x\right\} \tag{17}$$

analogously as in (5), then we have

Corollary 3.3 Assume (1) and (2). If we let m_n be defined as in (17) with $x = n\nu(\sqrt{n})^2 c$ for each $n \geq 1$ and some fixed constant $c > 0$, then

$$T_{m_n}^{IV} \to_d \mathbf{N}(0,1)$$

as $n \to \infty$.

Our previous results for the normal asymptotics therefore continue to hold for the unit root nonlinear IV t-ratio. The only modification required here is to redefine m and m_n so that we may effectively explore the contour given by the sum of squares of $F(y_{t-1})$, in place of y_{t-1}. Of course, the usual t-ratio can be regarded as a special case of the more general nonlinear IV t-ratio. Therefore, our earlier result in Theorem 2.1 follows from Corollary 3.3 as a special case.

A special case of particular interest arises if we set

$$F(x) = \operatorname{sgn}(x) \tag{18}$$

The resulting IV estimator, called the Cauchy estimator, was investigated earlier by So and Shin (1999b). This is an interesting example, for which the two contours coincide. Obviously, we have $F(y_{t-1})^2 \equiv 1$ and $\nu(\lambda) = 1$ for F given in (18). With the only conformable value $c = 1$, we thus have in this case that $m_n = n$. The contours of the equi-sample-size and the equi-squared-sum therefore become identical for any realization of the samples. Consequently, both the conventional approach and our new approach here yield the same normal asymptotics for the Cauchy estimator. Needless to say, we would have the same results for any IV estimator with asymptotically homogeneous F having the sign function as the limit homogeneous function. Note that for all this class of IV estimators we have $x\mathrm{sgn}(x) = |x|$, and therefore, they have nontrivial powers against the local alternatives shrinking towards unity at the rate n^{-1}.

[7] The IV t-ratio with integrable F has nontrivial powers only against the local alternative which shrink at the rate $n^{-1/2}$.

3.4 Tests in General Unit Root Models

All our previous results may be easily and naturally extended to more general unit root models. For the test of a unit root in the AR(p) model, we may consider the regression

$$y_t = \alpha y_{t-1} + \sum_{k=1}^{p-1} \alpha_k \triangle y_{t-k} + \varepsilon_t \tag{19}$$

and test whether $\alpha = 1$ using the t-ratio defined similarly as in (3). This is well known. In this case, we let

$$x_t = (\triangle y_{t-1}, \ldots, \triangle y_{t-p+1})'$$

and define

$$y_{p,t} = y_t - \left(\sum_{t=1}^n y_t x_t' \right) \left(\sum_{t=1}^n x_t x_t' \right)^{-1} x_t$$

Then the new contour for the t-ratio defined similarly as in (3) for the regression (19) is given by

$$m = \inf_{k \geq 1} \left\{ \sum_{t=1}^k y_{p,t-1}^2 \geq x \right\} \tag{20}$$

in place of (5). If we denote by $T_{m_n}^p$ the t-ratio based on regression (19) using the sample of size m given in (20), then it can be readily shown that our earlier result continues to apply.

Corollary 3.4 Assume (19) and (2). If we let m_n be defined as in (20) with $x = n^2 c$ for each $n \geq 1$ and some fixed constant $c > 0$, then

$$T_{m_n}^p \to_d \mathbf{N}(0, 1)$$

as $n \to \infty$.

We may show that the statistic $T_{m_n}^p$ has the same distribution as given in Corollary 3.1 of Section 3.1 under the local alternative (8). Moreover, the fitted mean can be allowed and treated exactly as in Section 3.2, and the result in Corollary 3.2 holds also for the statistic $T_{m_n}^p$. The nonlinear IV approach introduced in Section 3.3 can also be exploited in this case.

As is well known, the unit root test based on the AR(p) model (19) is valid for more general underlying processes if we let the order p of the AR model increase as the sample size gets large. This was first noted by Said and Dickey (1984), who show that the test based on the standard t-ratio is valid for general invertible ARMA processes of unknown order if we set $p = c\, n^\kappa$ with some constant $c > 0$ and $0 < \kappa \leq 1/3$. More recently, Chang and Park (2003) show that the procedure is indeed valid for more general linear processes with minimum summability condition on their coefficients and under much weaker condition $p = o(n^{1/2})$ on the rate of increase for the fitted AR orders. It can be shown that the result by Chang and Park (2003) continue to hold if we take the new contour. Therefore, the proposed procedure exploiting the new contour is applicable for a broad range of time series models under very mild conditions.

4. Panel Unit Root Tests

We now consider a panel unit root model given by

$$y_{it} = \alpha_i y_{i,t-1} + \varepsilon_{it}, \quad i = 1, \ldots, N; \; t = 1, \ldots, T. \tag{21}$$

with

$$\alpha_i = 1 \tag{22}$$

As usual, the index i denotes individual cross-sectional units, such as individuals, households, industries or countries, and the index t denotes time periods. The cross-sectional dimension N is not restricted and allowed to take large or small values. For the moment, we assume that the error terms ε_{it} are white noise with zero mean. Moreover, we let (ε_{it}) be uncorrelated across cross-sectional units. We make these simplifying assumptions temporarily to concentrate on more important aspects of our new methodology. They are not crucial and can be relaxed. This will be discussed later in more detail with other issues in panel unit root tests.

Panel unit root tests have been one of the most active research area for the past several years. This is largely due to the availability of panel data with long time span, and the growing use of cross-country and cross-region data over time to test for many important economic inter-relationships, especially those involving convergence/divergence of various economic variables. The notable contributers in theoretical research on the subject include Levin, Lin and Chu (2002), Im, Pesaran and Shin (2003), Maddala and Wu (1999), Choi (2001a, 2001b), Chang (1999, 2002), Phillips and Sul (2001), Moon and Perron (2001), and Bai and Ng (2002). There have been numerous related empirical researches as well. Examples include MacDonald (1996), Frankel and Rose (1996), Oh (1996) and Papell (1997), just to name a few. The papers by Banerjee (1999), Phillips and Moon (1999) and Baltagi and Kao (2000) provide extensive surveys on the recent developments on the testing for unit roots in panels.

To introduce the approach relying on the new contour, we let, as in (5), $m_i \geq 1$ be such that

$$m_i = \inf_{k \geq 1} \left\{ \sum_{t=1}^{k} y_{i,t-1}^2 \geq x \right\} \tag{23}$$

for any given $x > 0$ and for each $i = 1, \ldots, N$, and consider the t-ratio T_{m_i} for the sample of size m_i. Here the sample size m_i is determined by the squared sum of (y_{it}) achieving a certain level. Note that m_i is also given as a function of T, but its dependency on T is suppressed for expositional brevity. From the results established in earlier sections, we may easily deduce that

Theorem 4.1 Assume (21) and (22). If we let m_i be defined as in (23) with $x = T^2 c$ for each $T \geq 1$ and some fixed constant $c > 0$, then

$$T_{m_i} \to_d \mathbf{N}(0, 1) \tag{24}$$

as $T \to \infty$ for all $i = 1, \ldots, N$, and become independent across $i = 1, \ldots, N$.

The asymptotics for each individual t-ratio T_{m_i} derived by taking the contour of equi-squared-sum follow exactly in the same manner as for the univariate case established in Theorem 2.1, under the same set of conditions modified for our panel setting here. The standard normal limit theory of the univariate t-ratio given in Theorem 2.1 therefore continues to apply for each individual t-ratio T_{m_i} for $i = 1, \ldots, N$. Moreover, the t-ratios from different cross-sections are asymptotically independent.

Now suppose we are interested in testing whether the series (y_{it}) generated as in (21) has a unit root in all cross-sections $i = 1, \ldots, N$, against the alternative that (y_{it}) are stationary in all cross-section i. The null hypothesis is therefore formulated as $\mathrm{H}_0 : \alpha_i = 1$ for all i, and tested against the stationarity alternative $\mathrm{H}_1 : |\alpha_i| < 1$ for all i. The test statistic we first consider for testing the panel unit root hypothesis is a simple average of the individual t-ratio statistics for testing the unity of the AR coefficient computed from each cross-sectional unit. The test is defined as

$$S = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} T_{m_i} \tag{25}$$

The limit theory for S follows immediately from Theorem 4.1 as

Theorem 4.2 Under the assumptions of Theorem 4.1, we have

$$S \ \to_d \ \mathbf{N}(0,1)$$

as $T \to \infty$ for each $N \geq 1$.

Our limit theory here is derived using T-asymptotics only, and the factor $N^{-1/2}$ in the definition of the test statistic S in (25) is used just as a normalization factor, since S is based on the sum of N asymptotically independent random variables. This implies that the dimension of the cross-sectional units N may take any value, small as well as large.

Our methodology has some important advantages over the existing tests. In our approach, the individual t-ratios are normal and independent across individual units as long as T tends to infinity. This gives us the flexibility to fully investigate the presence of the unit root test in panels. For instance, we may use the order statistics, such as the minimum and maximum of T_{m_i} across i, to test the null hypothesis $\mathrm{H}_0 : \alpha_i = 1$ for all i against $\mathrm{H}_1 : |\alpha_i| < 1$ for some i, or the null hypothesis $\mathrm{H}_0 : \alpha_i = 1$ for some i against $\mathrm{H}_1 : |\alpha_i| < 1$ for all i. The limit theory for these order statistics can easily be obtained, since T_{m_i} are in the limit nothing but independent standard normal random variates. Moreover, our normal asymptotics make it much easier to deal with the cross-sectional dependencies, which are known to be extremely difficult to control using conventional approaches relying on nonnormal T-asymptotics.

The normal limit theory is also obtained for the existing panel unit root tests, such as the pooled OLS test by Levin, Lin and Chu (2002) and the group mean t-bar statistic by Im, Pesaran and Shin (2003); however, their tests involve the mean and variance adjustments for the individual tests, and is applicable only for large N. Furthermore, their theories require cross-sectional independence. More recently, several authors have made serious attempt to allow for cross-sectional dependencies. Chang (1999) allows for dependencies of unrestricted form, but her bootstrap procedure requires the dimension of time series T to be substantially larger than that of the cross-section N, which is restrictive for many practical applications. On the other hand, the procedures by Choi (2001b), Phillips and Sul (2001), Moon and Perron (2001) and Bai and Ng (2002) allow for cross-sectional dependencies, but for those in some specific forms. Finally, Chang (2002) and Chang and Song (2002) uses the nonlinear IV approach to invent the tests for panel unit roots that are valid in the presence of arbitrary cross-sectional correlations. Their limit theories are also normal.

Our framework is flexible enough to accommodate virtually all ingredients of previous researches. As we demonstrate in Section 3.4, the nonlinear IV approach is possible for our new methodology. The nonlinear IV methods developed in Chang can therefore be implemented here to deal with cross-sectional correlations in arbitrary forms. Moreover, the factor models for the cross-sectional dependencies employed in Choi (2001b), Phillips and Sul (2001), Moon and Perron (2001), and Bai and Ng (2002) can also be used together with our novel approach here. Our tests yield normal asymptotics and do not need the mean and variance adjustment for the individual tests. Finally, it is obvious that our results also extend to the panels with heterogeneous deterministic components such as individual fixed effects. Moreover, we may consider more general AR(p_i) models

$$y_{it} = \alpha_i y_{i,t-1} + \sum_{k=1}^{p_i} \alpha_{ik} \triangle y_{i,t-k} + \varepsilon_{it}, \quad i = 1, \ldots, N;\ t = 1, \ldots, T$$

for each individual unit i. The required extension is trivial given our results in Section 3.4.

5. Simulations

We evaluate the performance of the panel unit root test S defined in (25), which is the average t-ratio test along the new contour. We specify the simulation model as

$$y_{it} = \mu_i + y_{it}^{\circ}$$

and let the stochastic component y_{it}° be generated as

$$y_{it}^{\circ} = \alpha_i y_{i,t-1}^{\circ} + u_{it} \tag{26}$$

where

$$\begin{aligned} u_{it} &= \rho_i u_{i,t-1} + v_{it} \\ v_{it} &= \pi_i w_t + \varepsilon_{it} \end{aligned}$$

and $|\rho_i| < 1$, (w_t) and (ε_{it}) are independent and identically distributed random sequences. Under our specification, the stochastic component (y_{it}°) of (y_{it}) has a unit root when $\alpha_i = 1$.

Our simulation model is simple, yet it is general enough to consider various important aspects of panel models that would affect the finite sample performance of the unit root tests. Our model allows for the individual fixed effects parametrized as (μ_i), the heterogeneous serial correlation structures given by the autoregressive coefficients (ρ_i), and the presence of common factors (w_t) generating cross-sectional dependency with the heterogeneous factor loading coefficients (π_i). The factor structure has been routinely used to generate cross-sectional dependency, see Bai and Ng (2002), Moon and Perron (2001) and Phillips and Sul (2001). We consider three cases, which are specified below.

	Cases	Specifications
(a)	Prototype Case	$\rho_i = \pi_i = 0$
(b)	Independent Case	$\pi_i = 0$
(c)	Dependent Case	

For the prototype case, we consider the simplest model. This model is used only to demonstrate that our methodology really works as the theory predicts, i.e., that along the new contour the individual tests can be regarded as independent standard normals and our test based on S is distributed as standard normal approximately. The independent case looks at the independent panels with no common factors, and the dependent case explores the most general case.

The simulation results are reported in Tables 1–4. All the reported simulation results are based on 10,000 iterations. Panels with dimensions $N = 10, 50, 100, 200$ and $T = (25), 50, 100, 200$ are considered in Tables (1),3 and 4, while in Table 2 $N = 1, 2, 3$ and $T = 1000$ are looked at. In all cases, (ε_{it}) and (w_t) are generated as independent standard normals. For the prototype case, we fix the contour at the sum of squares given by cT^2 with $c = 0.23$, and the t-ratios based on the least squares estimators are computed for each individual units along the contour. The data are generated for each cross-section until its data reach the contour of the given sum of squares. For the independent and dependent cases, the autoregresive coefficients (ρ_i) are randomly drawn from the uniform distribution with support $[0.2, 0.4]$. The factor loading coefficients (π_i) on the common factor are also randomly drawn for the depedent case from the uniform distribution with support $[1, 4]$, providing ample heterogeneity among the individual responsiveness to the common factor.

For the unit root null hypothesis, we set the coefficients $\alpha_i = 1$ for all $i = 1, \ldots, N$, and investigate the finite sample sizes relative to the nominal 1%, 5% and 10% test sizes. To examine the rejection probabilities and the size-adjusted powers, we consider the stationary alternative with (α_i) generated randomly from the uniform distribution with support [0.95,1]. For the independent models, we consider the t-ratios based on the OLS and the Cauchy estimators along the new contour, which we call respectively the CP and Cauchy tests. For comparison purpose, we also consider the usual average t-ratio test along the conventional contour with the mean and variance corrections as done by Im, Pesaran and Shin (2003), which we denote by IPS. To implement the OLS based t-ratios, we use the new contour given by the minimum sum of squares across the cross-sections. On the other hand, for the dependent panels, we first defactor the data[8] and then consider the averaged OLS and Cauchy t-tests along the new contour. They are then compared with the Fisher-type

[8]We used the defactoring procedure suggested by Phillips and Sul (2001).

test suggested by Phillips and Sul (2001) along the conventional contour, which we denote by the PS test. Table 1 presents the simulation results for the prototype case. There it is clearly seen that the average t-test based on OLS perform well for all T and N. In particular, the normal approximation theory seems to be quite precise even for the cases with small T and N.

Table 2 compares the average t-tests based on the OLS and Cauchy estimators. To compare their discriminatory powers when T is large, we consider the local alternatives with the choices of local parameters $\delta = 1, 5, 10$.[9] Since the OLS is more efficient than the Cauchy estimator, one may well expect that the OLS based test be more powerful for $N = 1$. However, the comparsion may not hold as N increases, since imposing the new contour implies loosing some observations if the data set has the fixed sample size as in our simulations. This is indeed what we observe from our simulations. The average t-test based on the OLS estimator has more discriminatory powers for $N = 1$ and many cases for $N = 2$. The superiority of the OLS based t-test, however, quickly disappears as N becomes 3 or larger. It appears that the relative efficiency of the OLS estimator over the Cauchy estimator is not large enough to pay off the cost of adjusting the OLS estimator to the new contour even for N as small as 3. The average t-test based on the Cauchy estimator has largely the correct sizes, as is well predicted since it has standard normal limit distribution for large T and any N. In contrast, the test based on the OLS estimator is not expected to behave like standard normal when N is very small.

Figure 3 presents the performance of the averaged t-ratios based on OLS and Cauchy estimators along the contour given by the minimum sum of squares, and the average t-test along the conventional contour with the mean and varaince modifications given in Im, Pesaran and Shin (2003). We may summarize our findings here as follows. First, the OLS based t-ratio along the new contour does well for all $T = 50, 100, 200$. Second, the Cauchy based t-ratio along new contour does well for all T and N and more powerful than the OLS based test for the sample sizes considered. The reason the OLS based test turned out to be less powerful is that we are losing too many data to be able to take the new contour given by the minimum sum of squares. Indeed, it is clearly demonstrated that the average t-ratio based on the OLS estimator performs poorly relative to that based on the Cauchy estimator especially when N is large and T is small. This is the case where losing data to fit the new contour can be truly detrimental on the performance of the test.

Figure 4 presents the simulation results for the dependent case. In general, our observations on the average t-ratios based on the OLS and Cauchy estimators for the independent case continue to apply for this case. We note, however, that the OLS based t-ratio along the new contour does well for larger $T = 100, 200$, but has serious size distortions when T small. The average t-ratio based on the Cauchy estimator has good sizes and power for all T and N.

[9] Note that the local alternatives are formulated with the normalizing factor $T\sqrt{N}$. This is because the average t-test uses the normalizing factor $\sqrt{N}$.

6. Conclusion

In this paper, we develop a novel view on the interpretation of the unit root distributions. More explicitly, we show that if we take the new contour given by the equi-squared-sum instead of the usual equi-sample-size, then the limit distributions of the common unit root tests may be viewed as being normal. Subsequently, we demonstrate that this finding may be exploited to invent tools and methodologies for the effective inferences in nonstationary panels. In the panel context, our theory implies that the individual tests behave asymptotically as normal when they are computed simply along the equi-squared-sum contour across the individual units. Consequently, we may use various functionals of those individual tests to do inference in nonstationary panels. Here we only concentrate on the panel unit root models. This is just for expositional simplicity, which was intended to deliver the main contents of the proposed methodology more clearly. Quite obviously, we may use essentially the same approach to develop the corresponding methods of inference that are applicable for cointegrated panels.

Appendix: Proofs of Theorems

Proof of Theorem 2.1 Assume (2). Define $W_n(r) = n^{-1/2}y_{[nr]}$, where $[x]$ denotes the largest integer not exceeding $x \geq 0$. 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.

Recall that m_n is defined so as to satisfy

$$\int_0^{m_n/n} W_n(r)^2 dr = c$$

for some fixed $c > 0$. Therefore, if we define $\tau(c)$ to be such that

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

for the given $c > 0$, we have

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

as $n \to \infty$ since $W_n \to_{a.s.} W$ uniformly.

Under the null hypothesis of unit root, we have

$$\begin{aligned} T_{m_n} &= \left(\sum_{t=1}^{m_n} y_{t-1}^2 \right)^{-1/2} \sum_{t=1}^{m_n} y_{t-1}\varepsilon_t \\ &= \left(\int_0^{m_n/n} W_n(r)^2 dr \right)^{-1/2} \int_0^{m_n/n} W_n(r)\, dW_n(r) \end{aligned}$$

$$= \left(\int_0^{\tau(c)} W(r)^2 dr \right)^{-1/2} \int_0^{\tau(c)} W(r)\, dW(r) + o(1) \text{ a.s.} \tag{27}$$

as $n \to \infty$, since $W_n \to_{a.s.} W$ uniformly and $m_n/n \to_{a.s.} \tau(c)$ as $n \to \infty$. However, the process V defined by

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

is the DDS Brownian motion of the martingale M

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

and therefore,

$$c^{-1/2} V(c) = \left(\int_0^{\tau(c)} W(r)^2 dr \right)^{-1/2} \int_0^{\tau(c)} W(r)\, dW(r) \sim \mathbf{N}(0,1) \tag{28}$$

for any given $c > 0$. The reader is referred to, e.g., Revuz and Yor (1994) for the DDS Brownian motion. The stated result now follows readily from (27) and (28), and the proof is complete. ∎

Proof of Corollary 3.1 We use the same notation as in the proof of Theorem 2.1. Assume (8) and let $W_{n\delta}(r) = n^{-1/2} y_{[nr]}$. It follows that $W_{n\delta} \to_d W_\delta$ uniformly in $D(\mathbb{R})$. This is well known. If we define $\tau_\delta(c)$ by

$$\int_0^{\tau_\delta(c)} W_\delta(r)^2 dr = c \tag{29}$$

for a given fixed $c > 0$, 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, we have

$$\begin{aligned} T_{m_n} &= -\left(\sum_{t=1}^{m_n} y_{t-1}^2 \right)^{1/2} \frac{\delta}{n} + \left(\sum_{t=1}^{m_n} y_{t-1}^2 \right)^{-1/2} \sum_{t=1}^{m_n} y_{t-1} \varepsilon_t \\ &= -\left(\int_0^{m_n/n} W_{n\delta}(r)^2 dr \right)^{1/2} \delta + \left(\int_0^{m_n/n} W_{n\delta}(r)^2 dr \right)^{-1/2} \int_0^{m_n/n} W_{n\delta}(r)\, dW_n(r) \\ &= -\left(\int_0^{\tau_\delta(c)} W_\delta(r)^2 dr \right)^{1/2} \delta + \left(\int_0^{\tau_\delta(c)} W_\delta(r)^2 dr \right)^{-1/2} \int_0^{\tau_\delta(c)} W_\delta(r)\, dW(r) + o(1) \text{ a.s} \end{aligned} \tag{30}$$

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

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

of the martingale

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

from which the stated result follows immediately, due to (29) and (30). ∎

Proof of Corollary 3.2 The proof is straightforward given our earlier results, and therefore, omitted. ∎

Proof of Corollary 3.3 The proof follows immediately from Park and Phillips (1999) and our earlier results. Therefore, it is omitted. ∎

Proof of Corollary 3.4 The proof is obvious, and the details are omitted.

Proof of Theorem 4.1 Obvious from our earlier results.

Proof of Theorem 4.2 The stated result follows immediately from Theorem 4.1.

References

Bai, J. and S. Ng (2002). "A PANIC Attack on Unit Roots and Cointegration," forthcoming in *Econometrica*.

Baltagi, B.H. and C. Kao (2000). "Nonstationary panels, cointegration in panels and dynamic panels: A survey," mimeographed, Department of Economics, Texas A& M University.

Banerjee, A. (1999). "Panel Data Unit Roots and Cointegration: An Overview," *Oxford Bulletin of Economics & Statistics* 61: 607-629.

Chang, Y. (1999). "Bootstrap unit root tests in panels with cross-sectional dependency," forthcoming in *Journal of Econometrics*.

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 (2003). "A Sieve Bootstrap for the Test of a Unit Root," *Journal of Time Series Analysis* 24 (4), 379-400.

Chang, Y. and W. Song (2002)."Panel Unit Root Tests in the Presence of Cross-Sectional Dependency and Heterogeneity," mimeographed, Department of Economics, Rice University.

Choi, I. (2001a). "Unit root tests for panel data," *Journal of International Money and Finace* 20: 219-247.

Choi, I. (2001b). "Unit root tests for cross-sectionally correlated panels," mimeographed, Kukmin University.

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.

Frankel, J. and A. Rose (1996). "A panel project on purchasing power parity: Mean reversion within and between countries," *Journal of International Economics* 40: 209-224.

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

Im, K., H. Pesaran and Y. Shin (2003). "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics* 115, pp.53-74.

Kao, C. (1999). "Spurious regression and residual-based tests for cointegration in panel data," *Journal of Econometrics* 90: 1-44.

Levin, A., C. Lin and C.J. Chu (2002). "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties," *Journal of Econometrics* 108, 1-24.

MacDonald, R. (1996). "Panel unit root tests and real exchange rates," *Economics Letters* 50: 7-11.

Maddala, G.S. and S. Wu (1999). "A comparative study of unit root tests with panel data and a new simple test: Evidence from simulations and bootstrap," *Oxford Bulletin of Economics and Statistics* 61, 631-.

Moon, H.R. and B. Perron, (2001). "Testing for a unit root in panels with dynamic factors," forthcoming in *Journal of Econometrics*.

Oh, K.Y. (1996). "Purchasing power parity and unit root tests using panel data," *Journal of International Money and Finance*, 15, 405-418.

Papell, D.H. (1997). "Searching for stationarity: Purchasing power parity under the current float," *Journal of International Economics* 43: 313-323.

Park, J.Y. and P.C.B. Phillips (1999). "Asymptotics for nonlinear transformations of integrated time series," *Econometric Theory*, 15, 269-298.

Pedroni, P. (1997). "Panel cointegration: Asymptotic and finite sample properties of pooled time series tests with and application to the PPP hypothesis," mimeographed, Department of Economics, Inidiana University

Pesaran, M.H. (2003). "A simple panel unit root test in the presence of cross section dependence," mimeographed, Department of Economics, Cambridge University.

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

Phillips, P.C.B. and H.R. Moon (1999). "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica* 67: 1057-1111.

Phillips, P.C.B. and H.R. Moon (2000). "Nonstationary Panel Data Analysis: An Overview of Some Recent Developments," *Econometric Reviews* 19, 263-286.

Phillips, P.C.B., J.Y. Park and Y. Chang (2004). "Nonlinear instrumental variable estimation of an autoregression," *Journal of Econometrics* 118, 219-246.

Phillips, P.C.B. and D. Sul (2001). "Dynamic panel estimation and homogeneity testing under cross section dependence," mimeographed, University of Auckland.

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 (1999a). "Recursive mean adjustment in time-series inferences," *Statistics and Probability Letters*, 43, 65-73.

So, B.S. and D.W. Shin (1999b). "Cauchy estimators for autoregressive processes with applications to unit root tests and confidence intervals," *Econometric Theory*, 15, 165-176.

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.

Table 1: Average t-Test Along the New Contour: Prototype Case

		Sizes			Rejection Probabilities			Size-Adjusted Powers		
T	N	1%	5%	10%	1%	5%	10%	1%	5%	10%
25	10	0.013	0.059	0.112	0.083	0.225	0.340	0.066	0.202	0.314
	50	0.013	0.054	0.101	0.331	0.580	0.700	0.299	0.561	0.698
	100	0.011	0.046	0.088	0.580	0.798	0.878	0.565	0.808	0.893
	200	0.008	0.040	0.077	0.933	0.982	0.992	0.944	0.985	0.994
50	10	0.012	0.055	0.106	0.279	0.531	0.663	0.254	0.510	0.651
	50	0.010	0.049	0.099	0.908	0.975	0.989	0.906	0.975	0.989
	100	0.011	0.054	0.104	0.994	0.999	1.000	0.993	0.999	1.000
	200	0.009	0.048	0.093	1.000	1.000	1.000	1.000	1.000	1.000
100	10	0.011	0.051	0.104	0.973	0.996	0.998	0.970	0.996	0.998
	50	0.011	0.051	0.100	1.000	1.000	1.000	1.000	1.000	1.000
	100	0.011	0.050	0.094	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.011	0.048	0.094	1.000	1.000	1.000	1.000	1.000	1.000
200	10	0.010	0.049	0.100	1.000	1.000	1.000	1.000	1.000	1.000
	50	0.010	0.049	0.100	1.000	1.000	1.000	1.000	1.000	1.000
	100	0.010	0.052	0.103	1.000	1.000	1.000	1.000	1.000	1.000
	200	0.010	0.048	0.098	1.000	1.000	1.000	1.000	1.000	1.000

Table 2: Average t-Tests Based on OLS and Cauchy Estimators Along the New Contour: Independent Case with $T = 100$

($\alpha = 1 - \delta/T$)			Sizes			Rejection Probabilities			Size-Adjusted Powers		
N	δ	Tests	1%	5%	10%	1%	5%	10%	1%	5%	10%
1	0	CP	0.012	0.055	0.107						
		Cauchy	0.010	0.052	0.105						
	5	CP				0.073	0.263	0.410	0.058	0.239	0.392
		Cauchy				0.065	0.242	0.407	0.065	0.235	0.391
	10	CP				0.218	0.567	0.738	0.182	0.533	0.720
		Cauchy				0.175	0.470	0.653	0.176	0.460	0.640
	15	CP				0.363	0.729	0.864	0.313	0.699	0.850
		Cauchy				0.321	0.640	0.795	0.324	0.632	0.784
2	0	CP	0.012	0.054	0.109						
		Cauchy	0.011	0.053	0.106						
	5	CP				0.174	0.419	0.577	0.153	0.404	0.553
		Cauchy				0.169	0.475	0.659	0.163	0.456	0.642
	10	CP				0.583	0.856	0.931	0.545	0.845	0.924
		Cauchy				0.466	0.798	0.909	0.455	0.785	0.900
	15	CP				0.808	0.965	0.989	0.783	0.961	0.986
		Cauchy				0.694	0.918	0.967	0.685	0.911	0.964
4	0	CP	0.011	0.056	0.108						
		Cauchy	0.009	0.053	0.103						
	5	CP				0.384	0.661	0.783	0.370	0.636	0.769
		Cauchy				0.470	0.807	0.912	0.482	0.798	0.907
	10	CP				0.923	0.989	0.997	0.917	0.987	0.996
		Cauchy				0.885	0.983	0.995	0.889	0.982	0.995
	15	CP				0.994	1.000	1.000	0.994	1.000	1.000
		Cauchy				0.977	0.998	0.999	0.978	0.997	0.999

Table 3: Average t-Tests Along the Conventional and New Contours: Independent Case

			Sizes			Rejection Probabilities			Size-Adjusted Powers		
T	N	Tests	1%	5%	10%	1%	5%	10%	1%	5%	10%
25	10	IPS	0.015	0.066	0.132	0.032	0.122	0.216	0.021	0.096	0.170
		CP	0.013	0.058	0.108	0.060	0.204	0.327	0.047	0.178	0.304
		Cauchy	0.012	0.059	0.115	0.081	0.245	0.381	0.066	0.216	0.350
	50	IPS	0.019	0.086	0.152	0.104	0.277	0.415	0.064	0.188	0.309
		CP	0.014	0.058	0.113	0.218	0.462	0.607	0.174	0.428	0.579
		Cauchy	0.016	0.069	0.124	0.323	0.601	0.737	0.259	0.538	0.691
	100	IPS	0.030	0.106	0.186	0.202	0.448	0.597	0.096	0.284	0.437
		CP	0.014	0.059	0.113	0.453	0.723	0.831	0.396	0.687	0.814
		Cauchy	0.018	0.073	0.134	0.655	0.872	0.936	0.579	0.822	0.911
	200	IPS	0.040	0.135	0.224	0.438	0.707	0.823	0.232	0.483	0.646
		CP	0.013	0.059	0.111	0.791	0.936	0.970	0.748	0.922	0.966
		Cauchy	0.017	0.075	0.140	0.939	0.989	0.996	0.909	0.979	0.993
50	10	IPS	0.012	0.054	0.107	0.032	0.128	0.225	0.025	0.120	0.214
		CP	0.011	0.054	0.104	0.095	0.281	0.423	0.091	0.267	0.413
		Cauchy	0.012	0.054	0.107	0.111	0.324	0.480	0.098	0.307	0.463
	50	IPS	0.013	0.059	0.116	0.142	0.366	0.526	0.113	0.339	0.482
		CP	0.014	0.055	0.102	0.512	0.774	0.875	0.439	0.753	0.870
		Cauchy	0.014	0.058	0.111	0.671	0.890	0.951	0.617	0.871	0.942
	100	IPS	0.014	0.070	0.135	0.380	0.671	0.800	0.312	0.605	0.739
		CP	0.011	0.056	0.106	0.923	0.983	0.996	0.917	0.980	0.995
		Cauchy	0.014	0.059	0.112	0.983	0.998	1.000	0.976	0.998	0.999
	200	IPS	0.018	0.079	0.145	0.751	0.922	0.967	0.657	0.875	0.942
		CP	0.012	0.056	0.110	1.000	1.000	1.000	0.999	1.000	1.000
		Cauchy	0.015	0.065	0.122	1.000	1.000	1.000	1.000	1.000	1.000
100	10	IPS	0.011	0.054	0.102	0.117	0.345	0.507	0.107	0.328	0.498
		CP	0.011	0.057	0.105	0.473	0.754	0.854	0.456	0.735	0.848
		Cauchy	0.010	0.052	0.102	0.516	0.802	0.905	0.509	0.797	0.903
	50	IPS	0.011	0.053	0.110	0.578	0.841	0.921	0.549	0.829	0.912
		CP	0.011	0.056	0.102	0.985	0.998	1.000	0.984	0.998	1.000
		Cauchy	0.012	0.053	0.106	1.000	1.000	1.000	1.000	1.000	1.000
	100	IPS	0.013	0.060	0.118	0.947	0.992	0.998	0.931	0.989	0.997
		CP	0.010	0.050	0.099	1.000	1.000	1.000	1.000	1.000	1.000
		Cauchy	0.011	0.054	0.106	1.000	1.000	1.000	1.000	1.000	1.000
	200	IPS	0.014	0.060	0.122	1.000	1.000	1.000	1.000	1.000	1.000
		CP	0.010	0.050	0.100	1.000	1.000	1.000	1.000	1.000	1.000
		Cauchy	0.013	0.059	0.112	1.000	1.000	1.000	1.000	1.000	1.000

Table 4: Panel Unit Root Tests: Dependent Case

			Sizes			Rejection Probabilities			Size-Adjusted Powers		
T	N	Tests	1%	5%	10%	1%	5%	10%	1%	5%	10%
25	10	PS	0.013	0.063	0.122	0.006	0.039	0.082	0.005	0.031	0.066
		CP	0.011	0.050	0.099	0.032	0.129	0.233	0.030	0.129	0.235
		Cauchy	0.009	0.047	0.099	0.038	0.153	0.266	0.041	0.163	0.268
	50	PS	0.023	0.103	0.183	0.031	0.098	0.169	0.015	0.054	0.094
		CP	0.015	0.060	0.112	0.094	0.254	0.370	0.076	0.226	0.343
		Cauchy	0.017	0.069	0.125	0.129	0.320	0.446	0.089	0.270	0.392
	100	PS	0.038	0.132	0.225	0.058	0.146	0.225	0.028	0.070	0.116
		CP	0.018	0.065	0.119	0.193	0.395	0.507	0.141	0.343	0.470
		Cauchy	0.020	0.079	0.138	0.296	0.507	0.612	0.223	0.425	0.552
	200	PS	0.064	0.197	0.306	0.066	0.149	0.222	0.021	0.057	0.092
		CP	0.025	0.081	0.136	0.425	0.637	0.726	0.285	0.554	0.674
		Cauchy	0.032	0.095	0.158	0.574	0.744	0.809	0.397	0.641	0.752
50	10	PS	0.012	0.058	0.116	0.026	0.089	0.157	0.022	0.079	0.139
		CP	0.009	0.046	0.093	0.023	0.100	0.179	0.025	0.108	0.188
		Cauchy	0.008	0.044	0.092	0.034	0.115	0.205	0.038	0.130	0.218
	50	PS	0.018	0.080	0.144	0.035	0.098	0.160	0.026	0.070	0.121
		CP	0.012	0.055	0.107	0.184	0.381	0.497	0.163	0.363	0.487
		Cauchy	0.012	0.055	0.111	0.268	0.483	0.599	0.254	0.470	0.585
	100	PS	0.022	0.086	0.159	0.114	0.200	0.263	0.083	0.160	0.212
		CP	0.011	0.051	0.105	0.358	0.523	0.600	0.349	0.521	0.595
		Cauchy	0.012	0.059	0.113	0.493	0.628	0.685	0.470	0.615	0.673
	200	PS	0.027	0.114	0.207	0.122	0.195	0.245	0.092	0.144	0.184
		CP	0.012	0.058	0.109	0.638	0.730	0.767	0.624	0.718	0.761
		Cauchy	0.015	0.066	0.122	0.723	0.781	0.806	0.700	0.770	0.798
100	10	PS	0.012	0.056	0.109	0.079	0.164	0.234	0.071	0.156	0.224
		CP	0.009	0.049	0.100	0.043	0.126	0.198	0.045	0.128	0.198
		Cauchy	0.007	0.042	0.089	0.057	0.172	0.259	0.065	0.194	0.275
	50	PS	0.014	0.065	0.126	0.035	0.080	0.118	0.029	0.069	0.104
		CP	0.010	0.053	0.101	0.506	0.677	0.749	0.505	0.669	0.748
		Cauchy	0.011	0.053	0.103	0.675	0.782	0.825	0.665	0.775	0.823
	100	PS	0.017	0.070	0.130	0.131	0.201	0.249	0.116	0.181	0.228
		CP	0.011	0.053	0.098	0.571	0.668	0.714	0.565	0.663	0.715
		Cauchy	0.010	0.053	0.105	0.677	0.738	0.767	0.678	0.736	0.765
	200	PS	0.019	0.084	0.152	0.099	0.148	0.178	0.088	0.126	0.156
		CP	0.011	0.048	0.099	0.808	0.849	0.869	0.806	0.850	0.869
		Cauchy	0.012	0.060	0.113	0.854	0.881	0.896	0.851	0.878	0.893