行政院國科會專題研究計劃成果報告

一致性多變量單根檢定與比較

A Consistent Test for Unit Root in Panel Data

計劃編號：NSC 89-2415-H-004-002

執行期限：88/08/01 ~ 89/07/31

主持人　：郭炳伸　　國立政治大學國貿系

E-mail　：bsku@nccu.edu.tw

A Consistent Test for Unit Root in Panel Data *

Biing-Shen Kuo † and Cheng-Feng Lee ‡

National Chengchi University

Feb., 2000

Abstract

The paper develops a new test for panel unit root. The test suggested is a panel version of the Dicky-Fuller-type test. By taking full advantage of trending properties in data, the test is consistent at a rate faster than that considered in Levin, Lin and Chu (1997). The use of a pooled hyper-consistent estimator of unit root in panel regressions renders this feasible. The limit distribution of the test under the null, established by letting time series (T) and cross-sectional units (N) go to infinity is shown to be a standard normal. Our bootstrap tests are found to have correct rejection probability even for narrow and short panels, and to exhibit better power than the Im-Pesaran-Shin test statistics in large panels.

JEL classification: C12, C15, C22
Keywords: unit root, hyper-consistency, panel data

---

*This is a revised version of the paper presented at 1999 North American Summer Meeting of Econometric Society, and was originally written while the first author visited the Department of Economics of University of Wisconsin - Madison in 1999. Kuo thanks the department for the hospitality and the National Science Council, Taiwan, for research support. Both authors thank Bruce Hansen, Kaddour Hadri, Chung-Ming Kuan, Yuichi Kitamura, seminar participants of the Meeting, Institute of Economics as well as Institute of Statistics, Academia Sinica for numerous useful comments and discussions on earlier drafts.

†Corresponding author, Graduate Institute of International Trade, National Chengchi University, Taipei 116, Taiwan; Tel: (886 2) 2939 3091 ext. 81046, Tel: (886 2) 2939 3091 ext. 81029, Fax: (886 2) 2938 7699, e-mail: bsku@nccu.edu.tw

‡Tel: (886 2) 2939 3091 ext. 81046, Fax: (886 2) 2938 7699, e-mail: g5351501@grad.cc.nccu.edu.tw

1 Introduction

There has been a surge of interest in using panel data to test for nonstationarity in recent empirical work. Notable examples may include testing the income convergence hypothesis based on the growth theory (e.g., Bernard and Jones, 1996), and the long-run purchasing power parity in international finance (e.g., Frankel *et al.*, 1996). These studies were made possible by more available panel data sets, covering different countries over a relatively long time period. Quah (1994) pioneers the research by proposing the tests that exploit information from cross-sectional dimensions in inferring nonstationarity from panel data. Levin, Lin and Chu (1997, henceforth LLC[1]) and Im, Pesaran and Shin (1997, henceforth IPS) constitute further important contributions along the line. Extending the work of Quah (1994), Levin *et al.* consider an panel version of ADF-t test statistics by pooling estimates of unit root. Im *et al.*, on the other hand, mount a panel unit root test by averaging individual LM-statistics. As an important feature, the limit distribution of these available panel unit root tests are all characterized by a standard normal, remarkably different from those with univariate unit root tests. Closely related to the development in the panel unit root testing, Kao (1999) and Phillips and Moon (1998), on the other hand, study the asymptotic theory of cointegration in the panel data.[2]

One of the original motivations to develop tests for unit root in panel data is due to the lack power of conventional univariate unit root tests against persistent alternatives, typically for sample sizes that occur in practice. Recognizing data of longer span may lead to more reliable inference, researchers then employ the amount of available information as much as possible in applied time series work. This has been proven quite satisfactory in improving the power of unit root tests. Alternatively, applied researchers appeals to panel data where additional information from cross-sectional units help identify the parameters of concern, when long time series is not available. Panel data sets used in applied work, such as those aforementioned, consist of time series and cross-sections of comparable dimension, or often, short time series but very wide cross-sectional units. For example, the real exchange rate data in the study of Frankel *et al.* on the PPP hypothesis is a panel of more than 100 countries over less than 25 years. As far as the applicability and usefulness in applied work are concerned, it appears to be more needed developing tests for panel unit root valid in the context with short time series but large cross-sections than with long time series and large cross-sections. A test statistic is thus preferred, if it can more efficiently extract information from time series while summarizing the same or more information from cross-sections than the existing ones, or technically speaking, if it is powerful when the panel under study is

[1]Originally Levin and Lin, 1993. The revision is concerned with proofs of the asymptotic results. It does not affect conduct of the tests or critical values derived in the earlier version.

[2]A related contribution to the literature is by Hadri (1998) who considers tests for the null of stationarity in panel data.

wide but short.

This paper proposes tests for unit root in panel data that makes the notion come into effect. The use of a hyper-consistent estimator of a unit root by Phillips (1995) makes possible the construction of the test statistics. Faster than the OLS estimator of a unit root, the estimator, by taking advantage of the trending property in the data, is consistent at a rate of $O_p(3/2)$, thus suggesting using information from time series more efficiently. Constructing a test of unit root based on the estimator, in the univariate context, however, is appealing in theory but seems implausible in practice. This is because the asymptotics of the hyper-consistent estimator depends crucially on the nature of regression errors that are generally unknown. Our paper, in view of this, contributes to bringing the construction of an efficient test of unit root a revival by marrying the estimator and panel data. The new test for panel unit root we propose is essentially a pooled-t test statistic. By pooling cross-sectional information, random variations from the regression errors thus can be smoothed out, paving the way for building the new test that can efficiently utilize time series information. Indeed, the new test is consistent at a rate faster than the LLC test, as a direct result of the use of the hyper-consistent estimator. An implication of this faster rate is simply that the new test could be more powerful than alternative tests, in particular with panels of short time series but wide cross-sections.

Another contribution of the paper lies in our simplifying the hyper-consistent estimator by making further use of trending property in time series. As a common practice in the unit root literature, the estimator is corrected so that asymptotically it is free of nuisance parameters controlling error temporal dependence. However, we find that the correction, in the unit root context, is invariant to the dependence structure of error process. This suggests using the probability limit in place of semi-parametric estimates of the correction originally by Phillips and Hansen (1990) and Phillips (1995). The importance of the simplification thus has in fact much to do with eliminating sampling errors between the estimates and the true parameters. Practically, this could give a better asymptotic approximation to the finite-sample distribution of the suggested test.

Like those univariate unit root tests, the existing tests for panel unit root, including ours, subject to small-sample problems. The situation is in particular severe when time series is short, and heterogeneity and nuisance parameters, such as time trend and individual mean, are involved in estimation. The LLC and IPS tests, to cope with the problem, take the approach of simulation-based bias correction. Here, however, we deviate from their routes by taking the bootstrap method, another simulation-based approach. The preliminary simulation results we have reveal that the bootstrap test seems promising. Our bootstrap test not only has the right empirical rejection probability even for short and narrow panels, but also presents significantly better power than the test with level-adjusted critical values. Of more practical relevance is that instances of the latter take place more often for panels of large cross-sections but short time series that usually are associated with very low size-adjusted power in our simulations. This evidence shows the potential merits of using our test in combination with the bootstrap method in the practical analysis.

The rest of the paper is organized as follows. Section 2 discusses the hyper-consistent estimator of unit root and its simplification. Section 3 looks at models of concern and assumptions. The new test statistic for panel unit root, and its asymptotic properties are studied in Section 4. Section 5 reports finite-sample performance of the new tests. Section 6 concludes. Appendix contains mathematical proofs.

2 An Efficient Estimator of Unit Root

The idea of our test for panel unit root is built on the hyper-consistency of the fully-modified OLS (FM-OLS) estimator for unit root in Phillips (1995). We star by considering a simple univariate autoregressive time series:

$$y_t = \phi y_{t-1} + u_t, \text{ where } \phi = 1,$$

and u_t is a linear process with $Eu_t = 0$, and $E(u_t^2) = \sigma_u^2$. The FM-OLS estimator for ϕ is then defined as

(1)
$$\widehat{\phi}_{FM} = \frac{\sum_{t=2}^{T} y_t^+ y_{t-1} - T\widehat{\Delta}_{01}^+}{\sum_{t=2}^{T-1} y_{t-1}^2},$$

where

(2)
$$y_t^+ = y_t - \frac{\widehat{\Omega}_{01}}{\widehat{\Omega}_{11}} \triangle y_{t-1}$$

(3)
$$\widehat{\Delta}_{01}^+ = \widehat{\Delta}_{01} - \frac{\widehat{\Omega}_{01}}{\widehat{\Omega}_{11}} \widehat{\Delta}_{11},$$

where $\triangle y_t = y_t - y_{t-1}$. Here $\widehat{\Omega}_{01}$, $\widehat{\Omega}_{11}$, $\widehat{\Delta}_{01}$, and $\widehat{\Delta}_{11}$ are the elements in any consistent estimates of the long-run covariance and one-sided long-run covariance matrix, respectively,

$$\Omega = \begin{bmatrix} \Omega_{00} & \Omega_{01} \\ \Omega_{10} & \Omega_{11} \end{bmatrix} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=j}^{T} \sum_{j=-T}^{T} \begin{bmatrix} u_{t-j}u_t & u_{t-j}\triangle y_{t-1} \\ \triangle y_{t-1-j}u_t & \triangle y_{t-1-j}\triangle y_{t-1} \end{bmatrix}$$

$$\Delta = \begin{bmatrix} \Delta_{00} & \Delta_{01} \\ \Delta_{10} & \Delta_{11} \end{bmatrix} = \lim_{T\to\infty} \frac{1}{T} \sum_{t=j}^{T} \sum_{j=0}^{T} \begin{bmatrix} u_{t-j}u_t & u_{t-j}\triangle y_{t-1} \\ \triangle y_{t-1-j}u_t & \triangle y_{t-1-j}\triangle y_{t-1} \end{bmatrix}.$$

Phillips (1995) shows that

$$T(\widehat{\phi}_{FM} - 1) \to_p 0,$$

indicating that $\widehat{\phi}_{FM}$ converges in probability to the unit root at a faster rate than the OLS estimator. The fully-modified estimator by Phillips and Hansen (1990) is designed to yield a theory of inference invariant to parameters in the context with cointegrating variables. It works for the unit root process because y_t and y_{t-1} can be regarded as two variables cointegrating with a vector $[1, -1]$. The common trending component in y_t and y_{t-1} is actually the same. As a result,

$$\widehat{\Omega} \rightarrow_p \Omega = \begin{bmatrix} \omega^2 & \omega^2 \\ \omega^2 & \omega^2 \end{bmatrix}, \text{ and } \widehat{\Delta} \rightarrow_p \Delta = \frac{1}{2} \begin{bmatrix} \omega^2 + \sigma_u^2 & \omega^2 - \sigma_u^2 \\ \omega^2 + \sigma_u^2 & \omega^2 + \sigma_u^2 \end{bmatrix}$$

Note here that this implies that regardless of the dependence structure in u_t,

$$\widehat{\Omega}_{11}^{-1} \widehat{\Omega}_{01} \rightarrow_p 1.$$

Replacing $\widehat{\Omega}_{11}^{-1} \widehat{\Omega}_{01}$ by its probability limit, 1, thus would not alter the limit of $\widehat{\phi}_{FM}$. When the sample is sufficiently large,

$$\widehat{\phi}_{FM} = \frac{\sum_{t=2}^{T} (y_t y_{t-1} - y_{t-1} \Delta y_{t-1}) + (T-1)\sigma_u^2}{\sum_{t=2}^{T} y_{t-1}^2} + o_p(1),$$

where we also use the fact that in (3), $\widehat{\Delta}_{01}^{+} = \Delta_{01} - \Delta_{11} + o_p(1) = -\sigma_u^2 + o_p(1)$. In other words,

$$(\widehat{\phi}_{FM} - 1) = \frac{y_{T-1} u_T}{\sum_{t=2}^{T} y_{t-1}^2} + o_p(1),$$

Because $y_{T-1} u_T$ is $O_p(\sqrt{T})$ and $\sum_{t=2}^{T} y_{t-1}^2$ is $O_p(T^2)$, we immediately see why the hyper-consistency holds.

It can be further shown that

$$T^{3/2}(\widehat{\phi}_{FM} - 1) \Longrightarrow \frac{W(1) u_\infty}{\omega \int_0^1 W^2(r) dr}, \tag{4}$$

where $W(r) \equiv BM(1)$ and is independent of the random variable u_∞ which has the same distribution as the error u_t. Phillips (1992) first establishes the asymptotic distribution of $\widehat{\phi}_{FM}$, using the kernel-based estimates of the long-run variance, under the assumption that the bandwidth parameter has to grow at a rate faster than the optimal rate based on minimizing the mean square error. Here, without relying on the requirement on the growth rate of the bandwidth number, we show that it is yet able to obtain the hyper-consistency.

Test statistics based on (1) in the univariate context could be more powerful than conventional unit root tests, but is not readily applicable in practice. The reason for this is that the asymptotic distribution of the estimator in (4) depends crucially on the nature of u_∞ that is generally unknown. However, when panel data is available, it is possible to develop such a useful test by pooling data. The additional information coming from cross-sectional units helps to smooth out random deviations of u_∞, and renders the construction of the test feasible. In what follows, we shall explore this possibility.

3 Models and Assumptions

In this section, we lay out our models that generate the data, and specify the assumptions under which the limit distribution of our proposed test can be obtained.

Consider a sample of N cross-sectional units over T time periods, denoted by $\{y_{it}\}$ for $i = 1, \cdots, N$ and $t = 1, \cdots, T$. Following the literature, we assume that $\{y_{it}\}$ are generated from one of the following first-order autoregressive processes:

$$\begin{array}{ll} \text{Model 1:} & y_{i,t} = \phi_i y_{i,t-1} + u_{i,t} \\ \text{Model 2:} & y_{i,t} = \alpha_i(1-\phi) + \phi y_{i,t-1} + u_{i,t} \\ \text{Model 3:} & y_{i,t} = \alpha_i + \beta_i(1-\phi)t + \phi y_{i,t-1} + u_{i,t} \end{array}$$

We are interested in testing for unit root in $\{y_{it}\}$. The null hypothesis we are looking at is

$$H_0 : \phi_i = 1 \text{ for all } i,$$

while the alternative hypothesis is

$$H_a : |\phi_i| < 1 \text{ for some } i.$$

The model allows for heterogeneity to some degrees. The time trend effect and individual effect could differ across units. Under the alternative, part of the series, not necessarily all of them, in the panel could be stationary. To complete model specification, we also make the following assumptions about the error processes $\{u_{it}\}$ allowing for heterogeneity in their moments.

Assumption 1: $\{u_{i,t}\}$ is a linear process such that
(a) $u_{i,t} = C_i(L)\varepsilon_{i,t} = \sum_{t=0}^{\infty} c_i \varepsilon_{i,t}$, with $\sum_{j=0}^{\infty} j^2 c_j^2 < \infty$ for all i;
(b) $\varepsilon_{i,t}$ is i.i.d. across i and over t with $E(\varepsilon_{i,t}) = 0$ $E(\varepsilon_{i,t}^2) = 1$ and $E(\varepsilon_{i,t}^4) = \sigma_\varepsilon < M$, a finite number.

Assumption 1 allows for individual time series to exhibit varying serial correlations, and to have a heterogeneous variance. Under Assumption 1, the variance of $u_{i,t}$, $\sigma_{u_i}^2$, is equal to $\sum_{j=0}^{\infty} c_j^2$, and the long run variance of $u_{i,t}$ is given by $\omega_i^2 = C_i^2(1) < \infty$. The average of individual variance, that of individual long-run variance, and that of individual long-run variance weighted by its own variance, however, has to be bounded to carry out the derivation of the asymptotics, as we now state:

Assumption 2:
(a) $\sigma_u^2 = \lim_{N\to\infty} \frac{1}{N} \sum_i \sigma_{u_i}^2 < \infty$;
(b) $\omega^2 = \lim_{N\to\infty} \frac{1}{N} \sum_i \omega_i^2 < \infty$;
(c) $\omega_u^2 = \lim_{N\to\infty} \frac{1}{N} \sum_i \sigma_{u_i}^2 \omega_i^2 < \infty$.

4 Test Statistics and Limit Results

Our test statistics is easy to implement. It is a multivariate version of t ratio for unit root as the LLC test. But our test statistics differ from the LLC test by the use of the simplified FM-OLS estimator of unit root, as discussed in (??).[3] We now formally describe our testing procedure.

1. The first step is to de-mean (and de-trend) $\{y_{it}\}$, when Model 2 (Model 3) is of concern. Specifically, regress y_{it} against constant (and time trend),[4] and the OLS residuals obtained then form the de-meaned (and de-trended) panel, denoted by $\{\widehat{y}_{it}\}$. That is, $\widehat{y}_{it} = y_{it}$ for Model 1, $\widehat{y}_{i,t} = y_{i,t} - \widehat{\alpha}_i$ for Model 2, and $\widehat{y}_{i,t} = y_{i,t} - \widehat{\alpha}_i - \widehat{\beta}_i t$ for Model 3.

2. Stack series all together and obtain the pooled estimate of ϕ_i, denoted as $\widetilde{\phi}_{PFM}$, by applying the simplified FM-OLS to the whole panel:

$$\widetilde{\phi}_{PFM} = \frac{\left(\sum_{i=1}^{N}\sum_{t=2}^{T} \widehat{y}_{i,t}\widehat{y}_{i,t-1} - \Delta\widehat{y}_{i,t-1}\widehat{y}_{i,t-1}\right) + N(T-1)\widehat{\sigma}_u^2}{\sum_{i=1}^{N}\sum_{t=2}^{T} \widehat{y}_{i,t-1}^2}, \quad (5)$$

   where

$$\widehat{\sigma}_u^2 = \frac{1}{N}\sum_{i=1}^{n} \widehat{\sigma}_{u_i}^2$$

   with $\widehat{\sigma}_{u_i}^2$ being estimates of $\sigma_{u_i}^2$.

3. The last step comes to form the pooled t-statistics:

$$t_{FM} = \sqrt{T}\frac{\widetilde{\phi}_{PFM} - 1}{\widehat{\sigma}\left[\sum_{i=1}^{N}\sum_{t=2}^{T} \widehat{y}_{i,t-1}^2\right]^{-1/2}}, \quad (6)$$

   where

$$\hat{\sigma} = C\frac{\widehat{\omega}_u}{\widehat{\omega}},$$

   with $\widehat{\omega}_u$ and $\widehat{\omega}$ are estimates of ω_u and ω, as defined in Assumption 2, respectively. Here, C is a model-specific adjustment factor,

$$C = \begin{cases} \sqrt{2} & \text{for Model 1 and 2} \\ \sqrt{1.95} & \text{for Model 3.} \end{cases}$$

[3] In principle, it is possible to construct an average-t test using the simplified FM-OLS estimator. We do not pursue this approach here.

[4] When heterogeneity in intercept (time trend) is allowed for, the de-meaning (de-trending) procedure has to be done series by series.

The proposed t-statistics is the one that adjusts for nuisance parameters. The adjustment implicit in $\hat{\sigma}$ is made such that the limit distribution of the suggested test is free of parameters, a quite common practice in the theory of unit root and cointegration. It should be noted that unlike the LLC test, our pooled-t test is with a norming factor, $\sqrt{T}$, as a direct consequence of using the hyper-consistent estimator.

Having defined the test statistics, we focus our attention on its large-sample behavior. The asymptotic derivation generally would be involved a treatment of two indices, N and T. The limit theory dealing with double indexed processes has been made transparent in Phillips and Moon (1998). For a clear exposition, we mainly take the sequential limit argument in the text when deriving the limiting results. This approach gives an immediate limit by first letting T and subsequently N go to ∞. However, to be more rigorous, in the appendix, we also establish, under stronger conditions, the same limit results using the joint limit theory where both T and N grow simultaneously.

Indeed, the limit distribution of the test statistics is a standard normal with appropriate normalization. This is the main limit result of the paper as given below.

Theorem 1: Suppose $\widehat{\sigma}_u^2 \to_p \sigma_u^2$, $\widehat{\omega}_u^2 \to_p \omega_u^2$, $\widehat{\omega}^2 \to_p \omega^2$, and Assumptions 1 and 2 hold. Under $H_0 : \phi_i = 1$ for all i,

$$t_{FM} \Longrightarrow N(0,1)$$

as T and N pass to ∞ in a sequential order.

While the normality of the test statistics obtained is the same as those with the LLC and IPS tests, our test to converge to the normal is distinct from others by multiplying an extra $\sqrt{T}$, as a result of using the simplified FM-OLS estimator. Table 1 reports the appropriateness of the asymptotic normal approximation for Model 1 for different sample sizes. The data generating processes considered are those in IPS which allow for heterogeneous variances and serial correlations over panels. The normality approximation indeed appears quite accurate even for small and narrow panels, a remarkable feature revealing in the table.

It is easy to derive the normality by the sequential limit argument. To see this, we first derive the limit of t_{FM}, aside from the adjustment $\widehat{\sigma}$. So for fixed N, as $T \to \infty$, due to (4) and the cross-sectional independence,

$$\sqrt{T}\left(\tilde{\phi}_{PFM} - 1\right)\left(\sum_{i=1}^{N}\sum_{t=2}^{T}\widehat{y}_{i,t-1}^2\right)^{1/2} \Longrightarrow \left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\omega_i \bar{W}_i(1)u_{i,\infty}\right)\left(\frac{1}{N}\sum_{i=1}^{N}\int_0^1 \omega_i^2 \bar{W}_i^2(r)dr\right)^{-1/2}$$

where

$$\bar{W}_i(r) = \begin{cases} W_i(r) & \text{for Model 1} \\ W_i(r) - \int_0^1 W_i(s)ds & \text{for Model 2} \\ W_i(r) + 2\int_0^1 W_i(s)ds - 6\int_0^1 sW_i(s)ds & \text{for Model 3.} \end{cases}$$

Note that

$$E\left(\omega_i \bar{W}_i(1) u_{i,\infty}\right) = 0, \text{ and } E\left(\omega_i \bar{W}_i(1) u_{i,\infty}\right)^2 = \left(A\sigma_{u_i}\omega_i\right)^2,$$

with $A = 1$, $\sqrt{.333}$ and $\sqrt{.13}$ for Model 1, 2 and 3, respectively. Because $\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \left(A\sigma_{u_i}\omega_i\right)^2 < \infty$, by the Lindeberg Levy central limit theorem, as $N \to \infty$,

$$(7) \qquad \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \omega_i \bar{W}_i(1) u_{i,\infty} \Longrightarrow N(0, A^2\omega_u^2).$$

Also, because

$$E\left(\int_0^1 \omega_i^2 \bar{W}_i^2(r)dr\right) = \left(\frac{A}{C}\omega_i\right)^2,$$

it is straightforward to see that as $N \to \infty$

$$(8) \qquad \frac{1}{N} \sum_{i=1}^{N} \int_0^1 \omega_i^2 \bar{W}_i^2(r)dr \to_p \left(\frac{A}{C}\omega\right)^2.$$

Combining (7) and (8), by the continuous mapping theorem, we have

$$\left(\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \omega_i \bar{W}_i(1) u_{i,\infty}\right) \left(\frac{1}{N} \sum_{i=1}^{N} \int_0^1 \omega_i^2 \bar{W}_i^2(r)dr\right)^{-1/2} \Longrightarrow N(0, \sigma^2)$$

where $\sigma = C\omega_u/\omega$. We now have established the normality of the test in the limit.

In the preceding limit derivation, it is important to have consistent estimates of $\sigma_{u_i}^2$, ω_u, ω. The choice of the estimator is also of much practical relevance in terms of finite-sample performance of the suggested test. We will discuss how to obtain them in the Monte-Carlo simulation.

We come to examine the large-sample power implication of the suggested test, comparing to other existing tests. We investigate the test consistency under the alternative that $|\phi_i| < 1$ and $\phi_i = \phi$ for all i. Letting T and $N \to \infty$ sequentially,

$$t_{FM} = \frac{\sqrt{T}}{\widehat{\sigma}} \left[(\widetilde{\phi}_{PFM} - \phi) + (\phi - 1)\right] \left[\sum_{i=1}^{N} \sum_{t=2}^{T} \widehat{y}_{i,t-1}^2\right]^{1/2} \equiv O_p(\sqrt{T}) + O_p(T\sqrt{N}).$$

The first term in the LHS is due to that the FM-OLS estimation of stationary coefficient is consistent at a rate of $O_p(\sqrt{T})$ by Phillips (1995), and the second term comes from the fact $\sum_{i=1}^{N} \sum_{t=2}^{T} \widehat{y}_{i,t-1}^2/TN = O_p(1)$ as $\widehat{y}_{i,t-1}$ is stationary. The result is summarized in the following theorem.

Theorem 2: Suppose $\widehat{\sigma}_u^2 \to_p \sigma_u^2$ $\widehat{\omega}_u^2 \to_p \omega_u^2$, $\widehat{\omega}^2 \to_p \omega^2$, and Assumptions 1 and 2 hold. Under $H_a : \phi_i = |\phi| < 1$ for all i,

$$t_{FM} \equiv O_p(T\sqrt{N})$$

as T and N pass to ∞ in a sequential order.

Table 1 also gives the simulation evidence for the test consistency under Model 1. As seen from the table, the power of the test increases with N and T.

IPS prove that the average-LM test diverges at the same rate of $O_p(T\sqrt{N})$. This suggests that as a reflection of the rate, our test may in the small sample exhibit similar power performance against fixed stationarity alternatives to the IPS test. [5] LLC do not prove the test consistency, however.

5 Simulation

This section is devoted to investigating small-sample performance of the proposed test for unit root in panel data. Our simulation is featured by the use of bootstrap method to give a good control over the rejection probability. It has been well documented that conventional unit root tests subject to size distortion problem, based on the asymptotic critical values for sample size that occurs in practice. The existing tests for panel unit root, including ours, as a multivariate extension of the univariate unit root test, have the same sort of the problem. The downward bias with estimates of autoregressive coefficients when time series dimension is not sufficiently large compounds further the problem. Like other panel unit root tests, our test can be seen as an average of the sum of a sequence of random variables. The summation across the random variables, however, exacerbates the bias as cross-sectional units grows for a fixed time length.

To deal with the problem, while LLC and IPS take the route of correcting for small-sample bias that depends on sample sizes, we adopt the bootstrap method as another simulation-based approach. The difficulty to do small-sample corrections here is because the nuisance parameters on which the test statistics (6) depend is not invariant to the data generating process of the errors, unlike in the LLC and IPS tests.

To allow for a comparison, we consider the average LM test by IPS, denoted by $\Psi_{\overline{LM}}$. The pooled t test by LLC is not taken into account. The test is less applicable in panels with heterogeneity as considered here. Simulation evidence presented Im *et al.* (1998) indeed demonstrates that the LLC-test does not perform well, comparing to the IPS-test, in many instances with heterogeneity.

Our Monte-Carlo designs follow those in IPS as in Table 1. Since our panel unit root test is found to perform well based on the asymptotic critical value for Model 1, the bootstrap simulation is conducted only for Model 2 and 3. The data is thus generated according to

[5] Here we only establish the test consistency under the alternative that $|\phi_i| < 1$ for all i. When some of the processes are stationary and some are not under the alternative, the proof is a little more involved, but the conclusion is the same.

the following process,

Model 2: $y_{i,t} = \alpha_i(1-\phi) + \phi y_{i,t-1} + u_{i,t}$, or Model 3: $y_{i,t} = \alpha_i + \alpha_i(1-\phi)t + \phi y_{i,t-1} + u_{i,t}$,

where

$$u_{i,t} = \rho_i u_{i,t-1} + \varepsilon_{i,t},$$

$\alpha_i \sim N(0,1)$, $\varepsilon_{i,t} \sim N(0,\sigma_i^2)$ and $\sigma_i^2 \sim U[0.5, 1.5]$. We consider both cases that errors are with and without autocorrelation, controlled by ρ_i. $\rho_i \sim U[.2,.4]$ in the presence of autocorrelation. $\rho_i \sim U[.2,.4]$. α_i, and σ_i^2 are generated once and fixed in all replications. The initial values, y_{i0} are set to be 0. To evaluate the finite sample performance, the parameter of major interest, ϕ_i, are set to 1 under H_0 and to .9 under H_a for all units. Monte-Carlo replications are 1,000 and bootstrap replications take 100. Larger bootstrap seems desirable yet much time-consuming. Bootstrap replications of 100 have been typically used in the literature. Rejection frequencies based on 5% asymptotic and bootstrap critical values are, respectively, used to evaluate the test performance. The dimensions of the panels are chosen to be: $N = (10, 25, 50, 100)$ and $T = (25, 50)$. The simulation without autocorrelation basically serves as a baseline. On the other hand, to account for autocorrelation, the long-run variance for each cross-sectional unit is estimated non-parametrically using the QS kernel with a bandwidth set to $T^{1/3}$ and no pre-whitening. The choice of the kernel and the bandwidth are based on the criterion of the mean square error minimum in Andrews (1990).

5.1 Asymptotic Approximation

Entries labelled under 'Asym.' in Table 2 and 3 illustrate the finite-sample performance of our test and the IPS-test for Model 2 without and with autocorrelation, respectively, based on the 5 % asymptotic critical value. The performance of the tests is evaluated using the empirical rejection frequencies in 2,000 replications. It should be noted that the power reported for t_{FM} is size-adjusted.

Very different from Table 1, unfortunately, the test is now subject to a substantial size distortion that grows as N increases for a fixed T. Our asymptotics offers a partial explanation of the size distortion where a comparable growing rate between N and T has to be maintained in order to achieve the normality weak convergence. When N is of relatively larger in dimension than T, the small-sample distribution of the proposed test shifts leftwards as the imprecise estimate of the long-run variance from each unit accumulates when N grows. The slow convergence of the FM-OLS estimator documented in Phillips (1992) may also explain the size distortion.

While with the size distortion problem, the test for Model 2 displays an excellent size-corrected power. Indeed, for panels of medium size $(T, N = 25)$ which are of practical interest, the rejection frequencies of the proposed test in general exceeds that of the IPS-test by more than 20% under the alternative, whether or not autocorrelation exists. This suggests the merit of using the efficient FM-OLS estimator to gain the power in the context of testing for panel unit root. Table 3 and 4 show for Model 3 the same size distortion problem and power excellence of the suggested test.

5.2 Bootstrap Approximation

The size distortion problem is clearly indicative of the poor large-sample approximation under the null. We take a bootstrap procedure that enforces the restriction of a unit root to overcome the problem. The bootstrap distribution in the standard statistical contexts can yield consistent estimates for the sampling distribution, and thus the bootstrap test gives a better control over the size. The bootstrap estimate is implemented as follows.

1. Given a draw $\{y_{i,t}\}$, calculate the constrained least square (CLS) residuals under the null for each individual i. The residual for each i is given by $\bar{u}_{i,t} = \triangle y_{i,t}$ for Model 2, and $\bar{u}_{i,t} = \triangle y_{i,t} - \bar{\alpha}_i$ for Model 3, where $\bar{\alpha}_i = \frac{1}{T}\sum_{t=2}^{T} \triangle y_{i,t}$.

2. Fit $\bar{u}_{i,t}$ with an $AR(p)$. i.e. For each i,

$$\bar{u}_{i,t} = \sum_{k=1}^{p} \bar{\rho}_{i,k}\bar{u}_{i,t-k} + \bar{e}_{i,t}$$

3. Center the residuals $\{\bar{e}_{i,t}\}$ by

$$\bar{e}_{i,t} = \sqrt{\frac{T-1}{T-1-p}}\left(\bar{e}_{i,t} - \frac{\sum_{t=2+p}^{T}\bar{e}_{i,t}}{T-1}\right)$$

4. Draw a random sample of T for each i with replacement from the empirical distribution function of $\bar{e}_{i,t}$, denoted by $\bar{e}^*_{i,t}$.

5. Generate a bootstrap sample by the autoregression,

$$\bar{u}^*_{i,t} = \sum_{k=1}^{p} \bar{\rho}_{i,k}\bar{u}^*_{i,t-k} + \bar{e}^*_{i,t},$$

   and construct the pseudo-series $\left\{y^*_{i,t}\right\}$ by

$$y^*_{i,t} = y^*_{i,t-1} + \bar{u}^*_{i,t} \text{ (Model 2)}; \ = \bar{\alpha}_i + y^*_{i,t-1} + \bar{u}^*_{i,t} \text{ (Model 3)}.$$

6. Calculate the t_{FM}-statistics using the bootstrap sample $\left\{y^*_{i,t}\right\}$, denoted by t^*_{FM}.

7. Repeat the preceding steps 100 times, and obtain the bootstrap distribution of t^*_{FM}.

8. Calculate the 5% quantile of the bootstrap distribution of t^*_{FM}, denoted by $t^*_{FM,5\%}$.

9. Compute the rejection frequencies under both the null and the alternative, based on the bootstrap critical value, $t^*_{FM,5\%}$. The rejection frequencies obtained are then the bootstrap size and power, respectively.

We now discuss the bootstrap procedure. Step 1 imposes the unit root null hypothesis in obtaining estimates for the errors. This will prove quite important in examining the performance of the bootstrap tests. The constrained bootstrap is first-order correct under the null of unit root, but is not under the alternative. In view of this, the bootstrap critical value calculated, $t^*_{FM,5\%}$, under the alternative may not be appropriate. It might well anticipate that the power performance based on it would not perform as well as that based on the size-corrected critical value.[6]

Step 2 is a parametric recursive bootstrap in order to estimate the autocorrelation in errors. In simulation reported subsequently, the lag order p is chosen to be 1 or 2, while the true order in DGP is 1. By letting $p = 2$, we want to check if the performance of our bootstrap tests is sensitive to the overfitting. This is an importnat information for applications because the true lag order is generally unknown.

The standardization in Step 3 is to correct for downward bias of the autoregression estimates as suggested by Bickel and Freedman (1983). The bias is particularly large when T is small. The normalization factor, $\sqrt{\frac{T-1}{T-1-p}}$, indeed converges to 1 as T increases. The rest of the procedure is typical in the bootstrap approximation, and is self-evident as stated.

We now report the simulation results with the constrained bootstrap test, given by entries labelled under 'Asym.' in Table 2 to 5. Two observations emerge from the tables. First, the bootstrap test exhibits an excellent control over the rejection frequencies under the null for panels that can be as small as $N = 10$ and $T = 10$ for the cases without autocorrelation, and $N = 10$ and $T = 25$ for the cases with autocorrelation. Of more practical relevance is that the overfitting by choosing $p = 2$ in the autoregression assoicated with $\bar{u}_{i,t}$ harm little the size performance of our bootstrap test, as seen from those entries labelled under $t_{FM}(2)$. This bootstrap levels reported are indeed close to the nominal 5%, regardless of which models or cases are considered.

Second, the power performance of our bootstrap test depends on which model is under invesitgation. For Model 2, our bootstrap test again outperforms, even a slight power loss occurs, comparing to its size-adjusted counterpart. The power dominance is not sensitive to the overfitting as well. When Model 3 is considered, in contrast to those reported using size-corrected critical values, the bootstrap tests have a similar power performance to the IPS-test when the panels are large. In some instances, however, particularly cases with an overfitting in narrow and short panels, the bootstrap test is inferior. As expected, this is a consequence of using incorrect bootstrap critical values under the alternative as discussed.

[6]It would be fruitful to evaluate the performance of the 'unconstrained' bootstrap test.

6 Concluding Remarks

This paper proposed a panel unit root test by taking most advantage of the trending properties in time series data. The suggested test takes a form of averaging a series of random variables, and is not different from the existing tests in broad sense. Distinct from others, however, our test employs an efficient estimator of a unit root in construction. The use of such an estimator convert the gain in efficiency into a power gain, evident from the simulation where the tests could perform better than the IPS test in some cases, based on either size-adjusted or bootstrap critical values. The proposed tests therefore serves complementary to the existing test for panel unit root.

While the size distortion could be well controlled for the bootstrap test, the constrained bootstrap adopted here potentially yields inconsistent critical values under the alternatives. Our bootstrap tests in cases with time trend displays a sizable power loss, in contrast to its size-corrected counterpart. It remains a topic of research to investigate whether the unconstrained bootstrap could improve the power, at the same time maintain the correct size.

References

[1] Andrews, D.W.K. (1990), Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica*, 59, 817-858.

[2] Bernard, A. and C. Jones (1996), "Productivity Across Industries and Countries:Time Series Theory and Evidence," Review of Economics and Statistics 78, 135-146.

[3] Bickel, P.J. and Freedman, D.A. (1983), "Bootstrapping Regression Models with Many Parameters," In: A Festschrift for Erich Lehmann, ed. by Bickel, P., Doksum, K. and Hodges, J.L. Belmont, California: Wadsworth, pp. 28 -48.

[4] Davidson, A.C. and D.V. Hinkley (1997), Bootstrap Methods and Their Application, Cambridge: Cambridge University Press.

[5] 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.

[6] Hadri, K. (1998), "Tesitng for Stationarity in Heterogeneous Panel Data," Exter University, unpublished manuscript.

[7] Im, K.S. , M.H. Pesaran and Y. Shin (1997), "Testing for Unit Roots in Heterogeneous Panels," Econometrica (forthcoming).

[8] Levin, A., C.-F. Lin and C.S.J. Chu (1997), "Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties ", Journal of Econometrics (forthcoming).

[9] Kao, C. (1999), "Spurious Regression and Residual-Based Tests for Cointegration in Panel Data," Journal of Econometrics, forthcoming.

[10] Nankervis, J.C. and N.E. Savin (1996), "The Level and Power of the Bootstrap t Test in the AR(1) Model with Trend," Journal of Economic and Business Statistics 14, 161-168.

[11] Phillips, P.C.B. (1992), "Hyper-Consistent Estimation of a Unit Root in Time Series Regression," Cowles Foundation Discussion Paper No. 1040.

[12] Phillips, P.C.B. and B. Hansen (1990), "Statistical Inference in Instrumental Variables Regression with I(1) Processes," Review of Economics Studies 57,99-125.

[13] Phillips, P.C.B. (1995), "Fully Modified Least Squares and Vector Autoregression," Econometrica 63, 1023-79

[14] Phillips, P C.B. and H.R. Moon (1998), "Linear Regression Limit Theory for Nonstationary Panel Data," Econometrica (forthcoming).

[15] Quah, D. (1994), "Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Data, Economic Letters 44, 9-19.

Table 1: Finite-sample performance of t_{FM} test
(Model 1: $y_{i,t} = \phi_i y_{i,t-1} + u_{i,t}$)

	T=10		25		50	
	size	power	size	power	size	power
N=10	0.068	0.451	0.059	0.641	0.052	0.962
25	0.079	0.685	0.043	0.949	0.058	1.000
50	0.081	0.864	0.061	0.994	0.060	1.000
100	0.091	0.972	0.052	0.999	0.063	1.000

Note:
1. The data generation process (DGP) is

$$y_{i,t} = \phi y_{i,t-1} + u_{i,t}, u_{i,t} = \rho_i u_{i,t-1} + \varepsilon_{i,t}, i = 1, ..., N; t = 1, ..., T,$$

where $\phi = 1, .9$ under H_0 and H_a respectively, $\varepsilon_{i,t} \sim N(0, \sigma_i^2)$, $\sigma_i^2 \sim U[0.5, 1.5]$ and $\rho_i \sim U[0.2, 0.4]$. σ_i^2 and ρ_i are generated once and fixed in all replications.
2. The reported size ($\phi = 1$) and power ($\phi = .9$) are empirical rejection frequencies using asymptotic 5%-level critical value (-1.645) in 3,000 replications.
3. The long-run variance is estimated using QS kernel with a fixed bandwidth $T^{1/3}$ on pre-whittened residuals.

Table 2: Size and Power of t_{FM} and $\Psi_{\overline{LM}}$ in Heterogenous Panels
(Model 2: $y_{i,t} = \alpha_i(1-\phi_i) + \phi_i y_{i,t-1} + u_{i,t}$, No autocorrelation)

		T=10		25		50	
		t_{FM}	$\Psi_{\overline{LM}}$	t_{FM}	$\Psi_{\overline{LM}}$	t_{FM}	$\Psi_{\overline{LM}}$
N=10	Asym.						
	size	0.823	0.053	0.850	0.064	0.839	0.057
	power	0.223	0.105	0.530	0.252	0.942	0.678
	Boot.						
	size	0.050		0.061		0.056	
	power	0.203		0.571		0.935	
25	Asym.						
	size	0.982	0.062	0.982	0.058	0.978	0.059
	power	0.456	0.165	0.928	0.463	1.000	0.964
	Boot.						
	size	0.051		0.043		0.034	
	power	0.408		0.895		1.000	
50	Asym.						
	size	1.000	0.059	0.998	0.057	0.999	0.052
	power	0.682	0.242	0.996	0.693	1.000	0.999
	Boot.						
	size	0.047		0.043		0.042	
	power	0.670		0.995		1.000	
100	Asym.						
	size	1.000	0.060	1.000	0.068	1.000	0.054
	power	0.904	0.386	1.000	0.931	1.000	1.000
	Boot.						
	size	0.037		0.034		0.034	
	power	0.862		1.000		1.000	

Note:
1. The DGP is

$$y_{i,t} = \alpha_i(1-\phi) + \phi y_{i,t-1} + u_{i,t}$$

where $\phi = 1, .9$ under H_0 and H_a respectively, $\alpha_i \sim N(0,1)$, $u_{i,t} \sim N(0, \sigma_i^2)$, $\sigma_i^2 \sim U[0.5, 1.5]$. α_i and σ_i^2 are generated once and fixed in all replications.
2. Entries labelled 'Asym.' are empirical rejection frequencies in 2000 replications, based on 5%-level asymptotic critical value (-1.645), except that power of t_{FM} is size-adjusted.
3. Entries for t_{FM} labelled 'Boot.' are empirical rejection frequencies in 1,000 Monte Carlo replications of 100 bootstrap replications, based on 5%-level bootstrap critical value under the null.