CIRJE-F-85

Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency

Yoosoon Chang
Rice University / CIRJE, University of Tokyo

July 2000

Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency[1]

Yoosoon Chang
Department of Economics
Rice University

Abstract

We propose a unit root test for panels with cross-sectional dependency. We allow general dependency structure among the innovations that generate data for each of the cross-sectional units. Each unit may have different sample size, and therefore unbalanced panels are also permitted in our framework. Yet, the test is asymptotically normal, and does not require any tabulation of the critical values. Our test is based on nonlinear IV estimation of the usual ADF type regression for each cross-sectional unit, using as instruments nonlinear transformations of the lagged levels. The actual test statistics is simply defined as a standardized sum of individual IV t-ratios. We show in the paper that such a standardized sum of individual IV t-ratios has limit normal distribution as long as the panels have large individual time series observations and are asymptotically balanced in a very weak sense. We may have the number of cross-sectional units arbitrarily small or large. In particular, the usual sequential asymptotics, upon which most of the available asymptotic theories for panel unit root models heavily rely, are not required. Finite sample performance of our test is examined via a set of simulations, and compared to those of other commonly used panel unit root tests. Our test generally performs better than the existing tests in terms of both finite sample sizes and powers. We apply our nonlinear IV method to test for the purchasing power parity hypothesis in panels.

This version: July 2000

JEL Classification: C12, C15, C32, C33.

Key words and phrases: Panels with cross-sectional dependency, unit root tests, nonlinear instruments, average IV t-ratio statistics.

---

[1]The paper was prepared for the presentation at the Cardiff Conference on Long Memory and Nonlinear Time Series, and completed while I was visiting the Center for the International Research on the Japanese Economy (CIRJE), University of Tokyo. I am very grateful to Fumio Hayashi, Naoto Kunitomo and Yoshihiro Yajima for their hospitality and support. David Pappel and Keun-Yeob Oh kindly provided the data sets used here for empirical illustrations. I would also like to thank Joon Park and Peter Phillips 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

It is now widely perceived that the panel unit root test is important. The test helps us to answer some of the important economic questions like growth convergence and divergence, and purchasing power parity, among many others. Moreover, it also provides a means to improve the power of the unit root test, which is known to often yield very low discriminatory power if performed on individual time series. A number of unit root tests for panel data are now available in the literature. Examples include the tests proposed by Levin, Lin and Chu (1997), Im, Pesaran and Shin (1997), Maddala and Wu (1996), Choi (1997) and Chang (2000). The reader is referred to Banerjee (1999) for some detailed discussions on the existing panel unit root tests and other related issues.

Rather unsatisfactorily, however, most existing panel unit root tests assume cross-sectional independence, which is quite restrictive given the nature of economic panel data. Such tests are, of course, likely to yield biased results if applied to the panels with cross-sectional dependency. Maddala and Wu (1996) conduct a set of simulations to evaluate the performances of the commonly used panel unit root tests that are developed under the cross-sectional independence when in fact the panel is spatially dependent. They, in particular, show that the panel unit root tests based on independence across cross-sectional units, such as those considered in Levin, Lin and Chu (1997) and Im, Pesaran and Shin (1997), perform poorly for cross-sectionally correlated panels.

The cross-sectional dependency is very hard to deal with in nonstationary panels. In the presence of cross-sectional dependency, the usual Wald type unit root tests based upon the OLS and GLS system estimators have limit distributions that are dependent in a very complicated way upon various nuisance parameters defining correlations across individual units. There does not exist any simple way to get rid of nuisance parameters in such systems. This was shown in Chang (2000). None of the existing tests, except for Chang (2000) which relies on the bootstrap method, successfully overcomes the nuisance parameter problem in panels with cross-sectional dependence.

In this paper, we take the IV approach to solve the nuisance parameter problem for the unit root test in panels with cross-sectional dependency. Our approach here is based upon nonlinear IV estimation of the autoregressive coefficient. We first estimate the AR coefficient from the usual augmented Dickey-Fuller (ADF) regression for each cross-sectional unit using the instruments generated by an integrable transformation of the given time series. We then construct the t-ratio statistics for testing the unit root based on the nonlinear IV estimator for the AR coefficient. We show for each cross-sectional unit that such nonlinear IV t-ratio statistics for testing the unit root has limiting standard normal distribution under the unit root null hypothesis, just as in the stationary alternative cases. The asymptotic normality under the null indeed establishes continuity of the limit theory for the t-statistics over the entire parameter space covering both null and alternative hypotheses. This clearly makes a drastic contrast with the limit theory of the standard t-statistics based on the ordinary least squares estimator.

More importantly, we show that the limit standard normal distributions for each individual IV t-ratio statistics are independent even across dependent cross-sectional units. The cross-sectional independence of the individual IV t-ratio statistics follows readily from the asymptotic orthogonality for the nonlinear transformations of integrated processes by an integrable function, which is established in Chang, Park and Phillips (1999). We are therefore led to consider the average of these independent individual IV t-ratio statistics as a statistics for testing joint unit root null hypothesis for the entire panel. The actual test statistics is simply defined as a standardized sum of the individual IV t-ratios. We show in the paper that such a normalized sum of the individual IV t-ratios has standard normal limit distribution as long as $T_{\min} \to \infty$ and $T_{\max}{}^{1/4} \log T_{\max}/T_{\min}{}^{3/4} \to 0$, where $T_{\min}$ and $T_{\max}$ denote respectively the minimum and maximum numbers of the time series observations T_i's for the cross-sectional units $i = 1, \ldots, N$. The usual sequential asymptotics, upon which most of the available asymptotic theories for panel unit root models heavily rely, are therefore not required. We may thus allow the number of cross-sectional units to be arbitrarily small. Our test is applicable for all panels that have large numbers of individual time series observations and are asymptotically balanced in a very weak sense.

Finite sample performance of our average IV t-ratio statistics, which we call S_N statistics, is examined via a set of simulations, and compared to that of the commonly used average statistics t-bar by Im, Pesaran and Shin (1997). Our test generally performs better than the t-bar test in terms of both finite sample sizes and powers. The simulations conducted indeed corroborate the standard normal limit theory we provide here. The finite sample sizes of S_N are computed using the standard normal critical values, and shown to quite well approximate the nominal sizes. This is quite contrary to the well known finite sample size distortions of the t-bar test, see Maddala and Wu (1996) for example. The discriminatory powers of S_N are yet noticeably higher than the t-bar. We also apply our nonlinear IV method to test for the purchasing power parity hypothesis (PPP) using the data sets from Pappel (1997) and Oh (1996). Our test S_N supports unambiguously the PPP relationships, contrary to most of the previous empirical findings which are usually mixed and inconclusive.

The rest of the paper is organized as follows. Section 2 introduces the model, assumptions and back ground theory. Section 3 presents the nonlinear IV estimation of the augmented autoregression and derives the limit theory for the nonlinear IV t-ratio statistics for each cross-sectional unit. In Section 4, we introduce a nonlinear IV panel unit root test and establish its limit theory. It is in particular shown that the test is asymptotically standard normal. Section 5 extends our nonlinear IV methodology to models with deterministic trends such as constant and linear time trend. In Section 6, we conduct simulations to investigate finite sample performance of the average IV t-ratio statistics. Section 7 provides empirical illustrations for testing the purchasing power parity (PPP) using our nonlinear panel IV unit root test. Section 8 concludes, and mathematical proofs are provided in an Appendix.

2. Assumptions and Background Theory

We consider a panel model generated as the following first order autoregressive regression:

$$y_{it} = \alpha_i y_{i,t-1} + u_{it}, \quad i = 1, \dots, N; \ t = 1, \dots, T_i. \tag{1}$$

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 number of time series observations T_i for each individual i may differ across cross-sectional units. Hence, unbalanced panels are allowed in our model. We are interested in testing the unit root null hypothesis, $\alpha_i = 1$ for all y_{it} given as in (1), against the alternative, $|\alpha_i| < 1$ for some y_{it}, $i = 1, \dots, N$. Thus, the null implies that all y_{it}'s have unit roots, and is rejected if any one of y_{it}'s is stationary with $|\alpha_i| < 1$. The rejection of the null therefore does not imply that the entire panel is stationary. The initial values $(y_{10}, \dots, y_{N0})$ of $(y_{1t}, \dots, y_{Nt})$ do not affect our subsequent asymptotic analysis as long as they are stochastically bounded, and therefore we set them at zero for expositional brevity.

It is assumed that the error term u_{it} in the model (1) is given by an AR(p_i) process specified as

$$\alpha_i(L) u_{it} = \varepsilon_{it} \tag{2}$$

where L is the usual lag operator and

$$\alpha_i(z) = 1 - \sum_{k=1}^{p_i} \alpha_{i,k} z^k$$

for $i = 1, \dots, N$. Note that we let α_i vary across i, thereby allowing heterogeneity in individual serial correlation structures. We assume:

Assumption 2.1 For $i = 1, \dots, N$, $\alpha_i(z) \neq 0$ for all $|z| \leq 1$.

Under Assumption 2.1, the AR(p_i) process u_{it} is invertible, and has a moving-average representation

$$u_{it} = \pi_i(L) \varepsilon_{it}$$

where $\pi_i(z) = \alpha_i(z)^{-1}$ and is given by

$$\pi_i(z) = \sum_{k=0}^{\infty} \pi_{i,k} z^k$$

We allow for the cross-sectional dependency through the cross-correlation of the innovations ε_{it}, $i = 1, \dots, N$, that generate the errors u_{it}'s. To define the cross-sectional dependency more explicitly, we define $(\varepsilon_t)_{t=1}^T$ by

$$\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})' \tag{3}$$

and denote by $|\cdot|$ the Euclidean norm: for a vector $x = (x_i)$, $|x|^2 = \sum_i x_i^2$, and for a matrix $A = (a_{ij})$, $|A| = \sum_{i,j} a_{ij}^2$. The data generating process for the innovations $\{\varepsilon_t\}$ is assumed to satisfy the following assumption.

Assumption 2.2 $\{\varepsilon_t\}$ is an iid $(0,\Sigma)$ sequence of random variables with $\mathbf{E}|\varepsilon_t|^\ell < \infty$ for some $\ell > 4$, and its distribution is absolutely continuous with respect to Lebesque measure and has characteristic function φ such that $\lim_{\lambda\to\infty} |\lambda|^r \varphi(\lambda) = 0$, for some $r > 0$.

Assumption 2.2 is strong, but is still satisfied by a wide class of data generating processes including all invertible Gaussian ARMA models. Define a stochastic processes U_T for ε_t as

$$U_T(r) = T^{-1/2} \sum_{t=1}^{[Tr]} \varepsilon_t$$

on $[0,1]$, where $[s]$ denotes the largest integer not exceeding s. The process $U_T(r)$ takes values in $D[0,1]^N$, where $D[0,1]$ is the space of cadlag functions on $[0,1]$. Under Assumptions 2.1 and 2.2, an invariance principle holds for U_T, viz.

$$U_T \to_d U \tag{4}$$

as $T \to \infty$, where U is an N-dimensional vector Brownian motion with covariance matrix Σ. It is also convenient to define $B_T(r)$ from $u_t = (u_{1t}, \ldots, u_{Nt})'$, similarly as $U_T(r)$. Then we have $B_T \to_d B$, where $B = (B_1, \ldots, B_N)'$ and $B_i = \pi_i(1)U_i$. This is shown in Phillips and Solo (1992).

Our theory relies heavily on the local time of Brownian motion, which we discuss only briefly here for convenience. The reader is referred to e.g., Chung and Williams (1990) and Revuz and Yor (1994) for the concept of local time and a more detailed discussion. The local time of a Brownian motion V is a two parameter process, written as $L_V(t,s)$, with t and s respectively being the time and spatial parameters, satisfying the important (so-called *occupation time*) formula

$$\int_0^t G(V(r))\, d[V]_r = \int_{-\infty}^{\infty} G(s) L_V(t,s)\, ds, \tag{5}$$

for locally integrable $G : \mathbf{R} \to \mathbf{R}^k$, where $[V]_r$ is the quadratic variation process of the Brownian motion V. If we apply (5) to the function $G(x) = 1\{a \le x \le b\}$ for $a, b \in \mathbf{R}$, then

$$\int_0^t 1\{a \le V(r) \le b\}\, d[V]_r = \int_a^b L_V(t,s)\, ds,$$

and, correspondingly, when the local time $L_V(t,s)$ is treated as a function of its spatial parameter s, it can be viewed as an occupation time (or sojourn) density. The time that V stays in the interval $[a,b]$ is measured by $d[V]_r$, which can be thought of as a natural time scale for V. Also, due to the continuity of $L_V(t,\cdot)$, we have

$$L_V(t,s) = \lim_{\epsilon\to 0} \frac{1}{2\epsilon} \int_0^t 1\{|V(r) - s| < \epsilon\}\, d\,[V]_r\,.$$

for $\epsilon > 0$. Therefore, $L_V(t,s)$ measures the time (in units of quadratic variation) that V spends in the neighborhood of s, up to time t.

To define local times that appear in our limit theory more precisely, we first write the limit vector Brownian motion given in (4) explicitly as $U(r)=(U_1(r),\ldots,U_N(r))'$. We denote by L_{U_i} the local time of U_i, for $i=1,\ldots,N$, and define

$$L_i(t,s)=(1/\sigma_i^2)L_{U_i}(t,s)=\lim_{\epsilon\to 0}\frac{1}{2\varepsilon}\int_0^t 1\{|U_i(r)-s|<\epsilon\}\,dr,$$

where σ_i^2 is the variance of U_i, for $i=1,\ldots,N$. Clearly, L_i is just a scaled local time of U_i that measures time in chronological units. Our asymptotic results will be presented using L_i, instead of L_{U_i}. Using L_i, the occupation time formula (5) is rewritten as

$$\int_0^t G(U_i(r))\,dr=\int_{-\infty}^{\infty}G(s)L_i(t,s)\,ds, \tag{6}$$

since $d[U_i]_r=\sigma_i^2\,dr$. In the rest of the paper, we refer to (6) as the occupation time formula.

In addition to the Brownian motions $U=(U_1,\ldots,U_N)'$, we need to introduce another set of the standard Brownian motions $W=(W_1,\ldots,W_N)'$. Throughout the paper, the Brownian motion W will be assumed to be standard vector Brownian motion that is independent of U.

We now introduce the class of *regularly integrable* transformations in $\mathbf{R}$, which plays an important role in the subsequent development of our theory.

Definition 2.3 A transformation G on $\mathbf{R}$ is said to be *regularly integrable* if G is a bounded integrable function such that for some constants $c>0$ and $k>6/(\ell-2)$ with $\ell>4$ given in Assumption 2.2, $|G(x)-G(y)|\leq c\,|x-y|^k$ on each piece A_i of its support $A=\bigcup_{i=1}^m A_i\subset\mathbf{R}$.

The regularly integrable transformations are roughly integrable functions that are reasonably smooth on each piece of their supports. The required smoothness depends on the moment condition of the innovation sequence $\{\varepsilon_t\}$. Let ℓ be the maximum order of the existing moments. If $\ell>8$, any piecewise Lipschiz continuous function is allowed. For the indicator function on a compact interval to be regularly integrable, on the other hand, it is sufficient to have $\ell>4$.

The asymptotic behaviors of the nonlinear functions of an integrated time series are analyzed by Park and Phillips (1999,2000). For $\{y_{it}\}$ generated as in (1), they provide, in particular, the asymptotic theories for the sample moments given by $\sum_{t=1}^{T_i}G(y_{it})$ and $\sum_{t=1}^{T_i}G(y_{i,t-1})\,\varepsilon_{it}$, which are referred to in their paper as the *mean* and *covariance asymptotics* respectively, for various types of function G. Our subsequent theory is based upon the mean and covariance asymptotics for G regularly integrable. The conditions in Assumption 2.2 are required to obtain the relevant asymptotics. They are stronger than are required for the usual unit root asymptotics, because we need the convergence and invariance of the sample local time, as well as those of the sample Brownian motion, for the asymptotics of integrable transformations of integrated time series.

We now obtain the Beveridge-Nelson representations for u_{it} and y_{it}. Let $\alpha_i(1) = 1 - \sum_{k=1}^{p_i} \alpha_{i,k}$. Then it is indeed easy to get

$$\begin{aligned} u_{it} &= \frac{1}{\alpha_i(1)} \varepsilon_{it} + \sum_{k=1}^{p_i} \frac{\sum_{j=k}^{p_i} \alpha_{i,j}}{\alpha_i(1)} (u_{i,t-k} - u_{i,t-k+1}) \\ &= \pi_i(1)\varepsilon_{it} + (\tilde{u}_{i,t-1} - \tilde{u}_{it}) \end{aligned}$$

where $\pi_i(1) = 1/\alpha_i(1)$ and $\tilde{u}_{it} = \sum_{k=1}^{p_i} \tilde{\alpha}_{i,k} u_{i,t-k+1}$, with $\tilde{\alpha}_{i,k} = \pi_i(1) \sum_{j=k}^{p_i} \alpha_{i,j}$. Under our condition in Assumption 2.1, $\{\tilde{u}_{it}\}$ is well defined both in a.s. and L^r sense [see Brockwell and Davis (1991, Proposition 3.1.1)]. Under the unit root hypothesis $\alpha_i = 1$, we may now write

$$y_{it} = \sum_{k=1}^{t} u_{ik} = \pi_i(1)\xi_{it} + (\tilde{u}_{i0} - \tilde{u}_{it}) \tag{7}$$

where $\xi_{it} = \sum_{k=1}^{t} \varepsilon_{ik}$, for all $i = 1, \ldots, N$. Consequently, y_{it} behaves asymptotically as the constant $\pi_i(1)$ multiple of ξ_{it}. Note that $\tilde{u}_{it}$ is stochastically of smaller order of magnitude than ξ_{it}, and therefore will not contribute to our limit theory.

Using the specification of the regression error u_{it} given in (2), we write the model (1) as

$$y_{it} = \alpha_i y_{i,t-1} + \sum_{k=1}^{p_i} \alpha_{i,k} u_{i,t-k} + \varepsilon_{it}$$

Since $\triangle y_{it} = u_{it}$ under the unit root null hypothesis, the above regression may be written as

$$y_{it} = \alpha_i y_{i,t-1} + \sum_{k=1}^{p_i} \alpha_{i,k} \triangle y_{i,t-k} + \varepsilon_{it} \tag{8}$$

on which our unit root test will be based.

3. IV Estimation and Limit Theory

In this section, we consider the IV estimation of the augmented autoregression (8). To deal with the asymptotic endogeneity in the lagged level variable $y_{i,t-1}$, we use the instrument generated by a nonlinear instrument generating function F as

$$F(y_{i,t-1})$$

For the lagged differences $x'_{it} = (\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-p_i})$, we use the variables themselves as the instruments. Hence for the entire regressors $(y_{i,t-1}, x'_{it})'$, we use the instruments given by

$$(F(y_{i,t-1}), x'_{it})' = (F(y_{i,t-1}), \triangle y_{i,t-1}, \ldots, \triangle y_{i,t-p_i})' \tag{9}$$

The transformation F will be called the *instrument generating function* (IGF) throughout the paper. We assume that

Assumption 3.1 Let F be regularly integrable and satisfy $\int_{-\infty}^{\infty} xF(x) \neq 0$.

Roughly speaking, the condition given in Assumption 3.1 requires that the instrument $F(y_{i,t-1})$ is correlated with the regressor $y_{i,t-1}$. It is shown in Phillips, Park and Chang (1999, Theorem 3.2(a)) that IV estimators become inconsistent when the instrument is generated by a regularly integrable function F such that $\int_{-\infty}^{\infty} xF(x)\, dx = 0$. In this case, the IGF F is orthogonal to the regression function, which is the identity in this case, in the Hilbert space $L^2(\mathbf{R})$ of square integrable functions. In the standard stationary regression, an instrument is invalid and the resulting IV estimator becomes inconsistent if in particular it is uncorrelated with the regressor. Such an instrument failure also arises in our nonstationary regression with an integrated regressor when the instrument generating function is orthogonal to the regression function.

Examples of the regularly integrable IGF's satisfying Assumption 3.1 include $1\{|x| \leq K\}$, any indicator function on a compact interval defined by a truncation parameter K, and its variates such as $\mathrm{sgn}(x)1\{|x| < K\}$ and $x1\{|x| < K\}$. Also included are functions of the type $xe^{-|x|}$. For example, the IV estimator constructed from the indicator function on the interval $[0,1]$

$$F(x) = 1\{|x| \leq 1\}$$

is simply the *trimmed* OLS *estimator*, i.e., the OLS estimator which uses only the observations taking values in some compact interval.

Define

$$y_i = \begin{pmatrix} y_{i,p_i+1} \\ \vdots \\ y_{i,T_i} \end{pmatrix}, \quad y_{\ell i} = \begin{pmatrix} y_{i,p_i} \\ \vdots \\ y_{i,T_i-1} \end{pmatrix}, \quad X_i = \begin{pmatrix} x'_{i,p_i+1} \\ \vdots \\ x'_{i,T_i} \end{pmatrix}, \quad \varepsilon_i = \begin{pmatrix} \varepsilon_{i,p_i+1} \\ \vdots \\ \varepsilon_{i,T_i} \end{pmatrix}$$

where $x'_{it} = (\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-p_i})$. Then the augmented autoregression (8) can be written in matrix form as

$$y_i = y_{\ell i}\alpha_i + X_i\beta_i + \varepsilon_i = Y_i\gamma_i + \varepsilon_i \tag{10}$$

where $\beta_i = (\alpha_{i,1}, \ldots, \alpha_{i,p_i})'$, $Y_i = (y_{\ell i}, X_i)$, and $\gamma_i = (\alpha_i, \beta'_i)'$. For the augmented autoregression (10), we consider the estimator $\hat{\gamma}_i$ of γ_i given by

$$\hat{\gamma}_i = \begin{pmatrix} \hat{\alpha}_i \\ \hat{\beta}_i \end{pmatrix} = (W'_iY_i)^{-1}W'_iy_i = \begin{pmatrix} F(y_{\ell i})'y_{\ell i} & F(y_{\ell i})'X_i \\ X'_iy_{\ell i} & X'_iX_i \end{pmatrix}^{-1} \begin{pmatrix} F(y_{\ell i})'y_i \\ X'_iy_i \end{pmatrix} \tag{11}$$

where $W_i = (F(y_{\ell i}), X_i)$ with $F(y_{\ell i}) = (F(y_{i,p_i}), \ldots, F(y_{i,T_i-1}))'$. The estimator $\hat{\gamma}_i$ is thus defined to be the IV estimator using the instruments W_i.

The IV estimator $\hat{\alpha}_i$ for the AR coefficient α_i corresponds to the first element of $\hat{\gamma}_i$ given in (11). Under the null, we have

$$\hat{\alpha}_i - 1 = B_{T_i}^{-1}A_{T_i} \tag{12}$$

where

$$\begin{aligned}
A_{T_i} &= F(y_{\ell i})'\varepsilon_i - F(y_{\ell i})'X_i(X_i'X_i)^{-1}X_i'\varepsilon_i \\
&= \sum_{t=1}^{T_i} F(y_{i,t-1})\varepsilon_{it} - \sum_{t=1}^{T_i} F(y_{i,t-1})x_{it}' \left(\sum_{t=1}^{T_i} x_{it}x_{it}' \right)^{-1} \sum_{t=1}^{T_i} x_{it}\varepsilon_{it} \\
B_{T_i} &= F(y_{\ell i})'y_{\ell i} - F(y_{\ell i})'X_i(X_i'X_i)^{-1}X_i'y_{\ell i} \\
&= \sum_{t=1}^{T_i} F(y_{i,t-1})y_{i,t-1} - \sum_{t=1}^{T_i} F(y_{i,t-1})x_{it}' \left(\sum_{t=1}^{T_i} x_{it}x_{it}' \right)^{-1} \sum_{t=1}^{T_i} x_{it}y_{i,t-1}
\end{aligned}$$

and the variance of $\hat{\alpha}_i$ is given by

$$\sigma_i^2 \mathbf{E} B_{T_i}^{-2} C_{T_i} \tag{13}$$

under Assumption 2.2, where

$$\begin{aligned}
C_{T_i} &= F(y_{\ell i})'F(y_{\ell i}) - F(y_{\ell i})'X_i(X_i'X_i)^{-1}X_i'F(y_{\ell i}) && (14) \\
&= \sum_{t=1}^{T_i} F(y_{i,t-1})^2 - \sum_{t=1}^{T_i} F(y_{i,t-1})x_{it}' \left(\sum_{t=1}^{T_i} x_{it}x_{it}' \right)^{-1} \sum_{t=1}^{T_i} x_{it}F(y_{i,t-1}) && (15)
\end{aligned}$$

For testing the unit root hypothesis $H_0 : \alpha_i = 1$ for each $i = 1, \ldots, N$, we construct the t-ratio statistics from the nonlinear IV estimator $\hat{\alpha}_i$ defined in (12). More specifically, we construct such IV t-ratio statistics for testing for a unit root in (1) or (8) as

$$Z_i = \frac{\hat{\alpha}_i - 1}{s(\hat{\alpha}_i)} \tag{16}$$

where $s(\hat{\alpha}_i)$ is the standard error of the IV estimator $\hat{\alpha}_i$ given by

$$s(\hat{\alpha}_i)^2 = \hat{\sigma}_i^2 B_{T_i}^{-2} C_{T_i} \tag{17}$$

due to (13). The $\hat{\sigma}_i^2$ is the usual variance estimator given by $T_i^{-1} \sum_{t=1}^{T_i} \hat{\varepsilon}_{it}^2$, where $\hat{\varepsilon}_{it}$ is the fitted residual from the augmented regression (8), viz.

$$\hat{\varepsilon}_{it} = y_{it} - \hat{\alpha}_i y_{i,t-1} - \sum_{k=1}^{p_i} \hat{\alpha}_{i,k} \triangle y_{i,t-k} = y_{it} - \hat{\alpha}_i y_{i,t-1} - x_{it}'\hat{\beta}_i$$

It is natural in our context to use the IV estimate $(\hat{\alpha}_i, \hat{\beta}_i)$ given in (11) to get the fitted residual $\hat{\varepsilon}_{it}$. However, we may obviously use any other estimator of (α_i, β_i) as long as it yields a consistent estimate for the residual error variance.

To derive the limit null distribution of the IV t-ratio statistics Z_i introduced in (16), we need to obtain the asymptotics for various sample product moments appearing in A_{T_i}, B_{T_i} and C_{T_i}. They are presented in the following lemma.

Lemma 3.2 Under Assumptions 2.1, 2.2 and 3.1, we have

(a) $T_i^{-1/4}\sum_{t=1}^{T_i} F(y_{i,t-1})\,\varepsilon_{it} \;\to_d\; \sigma_i\left(\alpha_i(1)\,L_i(1,0)\int_{-\infty}^{\infty}F(s)^2ds\right)^{1/2} W_i(1)$

(b) $T_i^{-1/2}\sum_{t=1}^{T_i} F(y_{i,t-1})^2 \;\to_d\; \alpha_i(1)\,L_i(1,0)\int_{-\infty}^{\infty}F(s)^2ds$

(c) $T_i^{-3/4}\sum_{t=1}^{T_i} F(y_{i,t-1})\triangle y_{i,t-k} \;\to_p\; 0,$ for $k = 1,\ldots,p_i$

as $T_i \to \infty$, where $\alpha_i(1) = 1 - \sum_{k=1}^{p_i}\alpha_{i,k}$.

The results in Lemma 3.2 are simple extensions of the results in parts (c), (i) and (e) of Lemma 3.1 in Chang, Park and Phillips (1999). For the detailed discussion on the asymptotics here, the reader is referred to Park and Phillips (1999,2000) and Chang, Park and Phillips (1999). For the regularly integrable IGF F, the covariance asymptotics yields a mixed normal limiting distribution with a mixing variate depending upon the local time L_i of the limit Brownian motion U_i, as well as the integral of the square of the transformation function F.

It is very useful to note that

$$T_i^{-1/4}\sum_{t=1}^{T_i} F(y_{i,t-1})\varepsilon_{it} \;\approx_d\; \sqrt[4]{T_i}\int_0^1 F(\sqrt{T_i}B_{iT_i})dU_i$$

$$T_i^{-1/2}\sum_{t=1}^{T_i} F(y_{i,t-1})^2 \;\approx_d\; \sqrt{T_i}\int_0^1 F(\sqrt{T_i}B_{iT_i})^2$$

from which we may easily deduce the results in parts (a) and (b) of Lemma 3.2 using elementary martingale theory as in Park and Phillips (1999,2000) and Chang, Park and Phillips (1999).

The limit null distribution of the IV t-ratio statistics Z_i defined in (16) now follows readily from the results in Lemma 3.2.

Theorem 3.3 Under Assumption 2.1, 2.2 and 3.1, we have

$$Z_i \;\to_d\; W_i(1) \equiv \mathbf{N}(0,1)$$

as $T_i \to \infty$ for all $i = 1,\ldots,N$.

The limiting distribution of the IV t-ratio Z_i for testing $\alpha_i = 1$ is standard normal if a regularly integrable function is used as an IGF. Moreover, the limit standard normal distributions, $W_i(1)$'s, are independent across cross-sectional units $i = 1,\ldots,N$.

Note that we have under the alternative of stationarity, i.e., $|\alpha_i| < 1$,

$$T_i^{1/2}\,(\hat{\alpha}_i - \alpha_i) \to_d \mathbf{N}(0, v^2)$$

where

$$v^2 = \mathrm{plim}_{T_i\to\infty}\frac{\sum_{t=1}^{T_i} F(y_{i,t-1})^2}{\left(\sum_{t=1}^{T_i} F(y_{i,t-1})y_{i,t-1}\right)^2}$$

Consequently, if we let

$$Z_i(\alpha_i) = \frac{\hat{\alpha}_i - \alpha_i}{s(\hat{\alpha}_i)} \tag{18}$$

where $s(\hat{\alpha}_i)$ is defined in (17), then

$$Z_i(\alpha_i) \to_d \mathbf{N}(0,1) \tag{19}$$

Therefore, the IV t-ratio constructed with regularly integrable IGF are normally distributed asymptotically, for all $|\alpha_i| \leq 1$.

Continuity of the distribution across the values α_i of the t-ratio $Z_i(\alpha_i)$ defined in (18) also allows us to construct the confidence intervals for α_i from the IV estimators. As we have noticed above,

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

for all values of α_i including unity, when the IGF F is a regularly integrable function. We may therefore construct $100\,(1-\lambda)\,\%$ asymptotic confidence interval for α_i as

$$\Big[\hat{\alpha}_i - z_{\lambda/2}\, s(\hat{\alpha}_i), \quad \hat{\alpha}_i + z_{\lambda/2}\, s(\hat{\alpha}_i)\Big] \tag{20}$$

using the IV estimators generated by any integrable function F, where $z_{\lambda/2}$ is the $(1-\lambda/2)$-percentile from the standard normal distribution.

This is one important advantage of using the nonlinear IV method. The OLS-based standard t-ratio has non-Gaussian asymptotic null distribution, called the Dickey-Fuller distribution. It is asymmetric and skewed to the left, as tabulated in Fuller (1996). Therefore, the confidence interval which is valid for all $|\alpha_i| \leq 1$ cannot be constructed from the OLS based t-ratio.

4. Panel Nonlinear IV Unit Root Test

The test statistics that we propose here to test the unit root null hypothesis H_0 : $\alpha_i = 1$ for all $i = 1, \ldots, N$ is basically an average of the individual t-ratio statistics for testing $\alpha_i = 1$ in (8) for each cross-sectional unit $i = 1, \ldots, N$. For testing the joint unit root hypothesis $H_0: \ \alpha_1 = \ldots = \alpha_N = 1$, we propose to use an average statistics based on the individual t-ratios constructed from the nonlinear IV estimators $\hat{\alpha}_i$ defined in (12). More specifically, the *average IV t-ratio statistics* is defined as

$$S_N = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Z_i \tag{21}$$

where Z_i is the nonlinear IV t-ratio statistics, defined (16), for testing $\alpha_i = 1$ for the i-th cross-sectional unit.

For the average statistics S_N, we allow each of the cross-sectional units $i = 1, \ldots, N$ to have a different sample size T_i, and therefore unbalanced panels are permitted in our framework. Our test is based on nonlinear IV estimation of the usual

ADF type regression for each cross-sectional unit, using as instruments nonlinear transformations of the lagged levels $y_{i,t-1}$'s.

In order to derive the limit theory for the statistics S_N, we first investigate how the individual IV t-ratio statistics Z_i's interact in the limit. We have

$$
\begin{aligned}
T_i^{-1/4}\sum_{t=1}^{T_i} F(y_{i,t-1})\varepsilon_{it} &\approx_d \sqrt[4]{T_i}\int_0^1 F(\sqrt{T_i}B_{iT_i})dU_i \\
T_j^{-1/4}\sum_{t=1}^{T_j} F(y_{j,t-1})\varepsilon_{jt} &\approx_d \sqrt[4]{T_j}\int_0^1 F(\sqrt{T_j}B_{jT_j})dU_j
\end{aligned}
$$

which become asymptotically independent if their quadratic covariation

$$
\sigma_{ij}\sqrt[4]{T_iT_j}\int_0^1 F(\sqrt{T_i}B_{iT_i}(r))F(\sqrt{T_j}B_{jT_j}(r))dr
$$

converges a.s. to zero, where σ_{ij} denotes the covariance between U_i and U_j. This was shown in Chang, Park and Phillips (1999). Below we introduce a sufficient condition and establish their asymptotic independence subsequently.

Let $T_{\min}$ and $T_{\max}$ respectively be the minimum and the maximum of T_i's for $i = 1, \ldots, N$.

Assumption 4.1 Assume

$$
T_{\min} \to \infty \quad \text{and} \quad \frac{{T_{\max}}^{1/4}\log T_{\max}}{{T_{\min}}^{3/4}} \to 0
$$

Then we have

Lemma 4.2 Under Assumptions 2.1, 2.2, 3.1 and 4.1, the following holds:

$$
\sqrt[4]{T_iT_j}\int_0^1 F(\sqrt{T_i}B_{iT_i}(r))F(\sqrt{T_j}B_{jT_j}(r))dr \to_p 0 \tag{22}
$$

and the results in Lemma 3.2 hold jointly for all $i = 1, \ldots, N$ with independent W_i's across $i = 1, \ldots, N$.

The result Lemma 4.1 is new, and shows that the product of the nonlinear instruments $F(y_{i,t-1})$ and $F(y_{j,t-1})$ from different cross-sectional units i and j are asymptotically uncorrelated. This implies that the individual IV t-ratio statistics Z_i and Z_j constructed from the nonlinear IV's $F(y_{i,t-1})$ and $F(y_{j,t-1})$ are asymptotically independent. This asymptotic orthogonality plays a crucial role in developing limit theory for our panel unit root test S_N defined above, as can be seen in below.

The limit theory for S_N follows immediately from Theorem 3.3 and Lemma 4.1, and is provided in

Theorem 4.3 We have

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

under Assumptions 2.1, 2.2, 3.1 and 4.1.

The limit theory is derived using T-asymptotics only. It holds as long as all T_i's go to infinity and T_i's are asymptotically balanced in a very weak sense, as we specify in Lemma 4.1. It should be noted that the usual sequential asymptotics is not used here.[2] The factor $N^{-1/2}$ in the definition of the test statistics S_N in (21) is used just as a normalization factor, since S_N is based on the sum of N independent random variables. Therefore, the dimension of the cross-sectional units N may take any value, small as well as large. The above result therefore implies that we can do simple inference based on the standard normal distribution even for unbalanced panels with general cross-sectional dependencies.

The normal limit theory is also obtained for the existing panel unit root tests, but the theory holds only under cross-sectional independence, and obtained only through sequential asymptotics. For example, the pooled OLS test by Levin, Lin and Chu (1997) and the group mean t-bar statistics by Im, Pesaran and Shin (1997) have normal asymptotics. However, they all presume cross-sectional independence and their normal limit theories are obtained through sequential asymptotics. The independence assumption was crucial for their tests to have normal limiting distributions, since the individual t -statistics contributing to the average become independent only when the innovations ε_{it} generating the individual units are independent. Moreover, the sequential asymptotics is an essential tool to derive their results, and they do not provide joint asymptotics. Here we achieve the asymptotic independence of individual t-statistics by establishing asymptotic orthogonalities of the nonlinear instruments used in the construction of the individual t-ratio statistics without having to impose independence across cross-sectional units, or relying on sequential asymptotics.

5. Nonlinear IV Estimation for Models with Deterministic Trends

The models with deterministic components can be analyzed similarly. If the time series (z_{it}) with a nonzero mean is given by

$$z_{it} = \mu_i + y_{it} \tag{23}$$

where (y_{it}) is generated as in (1). We can test for the presence of the unit root in the process y_{it} from the augmented regression (8) defined with the fitted values of $\{y_{it}\}$ obtained from the preliminary regression (23), viz.

$$y_{it}^{\mu} = \alpha_i y_{i,t-1}^{\mu} + \sum_{k=1}^{p_i} \alpha_{i,k} \triangle y_{i,t-k}^{\mu} + \varepsilon_{it} \tag{24}$$

[2]The usual sequential asymptotics is carried out by first passing T to infinity with N fixed, and subsequently let N go to infinity, usually under cross-sectional independence.

where

$$y_{it}^{\mu} = z_{it} - \hat{\mu}_i^t = z_{it} - \frac{1}{t}\sum_{k=1}^{t} z_{ik} \tag{25}$$

We note that the parameter μ_i is estimated from the model (23) using the observations up to time t. That is, the $\hat{\mu}_i^t$ is the least squares estimator from the regression

$$z_{ik} = \mu_i + y_{ik}, \quad \text{for} \quad k = 1, \ldots, t \tag{26}$$

That we use the data upto the current period t only, instead of using the full sample, for the estimation of the constant μ_i leads to the demeaning based on the partial sum of the data upto t as given in (25), which we call *adaptive demeaning*. We may then construct the nonlinear IV t-ratio statistics Z_i^{μ} based on the nonlinear IV estimator for α_i from the regression (24), just as in (16). With the adaptive demeaning the predictability of our nonlinear instrument $F(y_{i,t-1}^{\mu})$ is retained, and consequently our previous results continue to apply, including the normal distribution theory for the IV t-ratio statistics.

We may also test for the unit root in the models with more general deterministic time trends. As in the cases with the models with nonzero means, we may derive nonlinear IV unit root test Z_i^{τ} in the same manner. More explicitly, consider the time series with a linear time trend

$$z_{it} = \mu_i + \delta_i t + y_{it} \tag{27}$$

where (y_{it}) is generated as in (1). Similarly, we may test for the unit root in y_{it} from the regression (8) defined with the fitted values of (y_{it})

$$y_{it}^{\tau} = \alpha_i y_{i,t-1}^{\tau} + \sum_{k=1}^{p_i} \alpha_{i,k} \triangle y_{i,t-k}^{\tau} + \varepsilon_{it} \tag{28}$$

from running the regression (27), where

$$y_{it}^{\tau} = z_{it} - \hat{\mu}_i^t - \hat{\delta}_i^t\, t \tag{29}$$

The parameter estimates $\hat{\mu}_i^t$ and $\hat{\delta}_i^t$ are estimated using again the observations upto time t only, from the model (27). That is, $\hat{\mu}_i^t$ and $\hat{\delta}_i^t$ are the least squares estimator from

$$z_{ik} = \mu_i + \delta_i k + y_{ik}, \quad \text{for} \quad k = 1, \ldots, t \tag{30}$$

This leads to the *adaptive detrending* of the data y_{it} as given in (29) above, and this in turn preserves the predictability of our instrument $F(y_{i,t-1}^{\tau})$. The nonlinear IV t-ratio statistics Z_i^{τ} is then defined as in (16) from the nonlinear IV estimator for α_i from the regression (28).

We may now derive the limit theory for the statistics Z_i^{μ} and Z_i^{τ}. We may do so in the similar manner as we did to establish the limit theory given in Theorem 3.3.

In order to define the limit distribution properly, we first introduce some notation. Define the *adaptively demeaned* Brownian motion

$$U_i^{\mu}(r) = U_i(r) - \frac{1}{r}\int_0^r U_i(s)ds \tag{31}$$

for $i = 1, \ldots, N$, and denote its local time by L_i^{μ} scaled as for L_i. Similarly we also define *adaptively detrended* Brownian motion

$$U_i^{\tau}(r) = U_i(r) - \frac{4}{r}\int_0^r U_i(s)ds + \frac{6}{r^2}\int_0^r sU_i(s)ds \tag{32}$$

and analogously denote the local time of the adaptively detrended Brownian motion U_i^{τ} by L_i^{τ} for $i = 1, \ldots, N$.

The processes U_i^{μ} and U_i^{τ} introduced in (31) and (32) are not defined at the origin. However, due to the well known Brownian law of iterated logarithm [see, for instance, Revuz and Yor (1994, p.53)], we have

$$U_i(r) = O(r^{1/2}(\log\log(1/r))^{1/2}) \;\; \text{a.s.}$$

and therefore,

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

and

$$\int_0^r sU_i(s)ds = O(r^{5/2}(\log\log(1/r))^{1/2}) \;\; \text{a.s.}$$

It then follows that

$$\frac{1}{r}\int_0^r U_i(s)ds, \;\; \frac{1}{r^2}\int_0^r sU_i(s)ds \;\to\; 0 \quad \text{as} \;\; r \to 0$$

Therefore, if we let $U_i^{\mu}(0) = U_i^{\tau}(0) = 0$, then both $U_i^{\mu}(r)$ and $U_i^{\tau}(r)$ become continuous stochastic processes defined on $[0, \infty)$. We make this convention through the rest of the paper.

The limit theories given in Lemma 3.2 extend easily to the models with nonzero means and deterministic trends if we replace y_{it} with the fitted residuals y_{it}^{μ} and y_{it}^{τ} defined respectively in (25) and (29), and are given similarly with the local times L_i^{μ} and L_i^{τ} in the places of the local time L_i of the Brownian motion U_i. Then the limit theories for Z_i^{μ} and Z_i^{τ} follows immediately, and are given in

Corollary 5.1 Under Assumption 2.1, 2.2 and 3.1, we have

$$Z_i^{\mu}, \; Z_i^{\tau} \;\to_d\; \mathbf{N}(0,1)$$

as $T_i \to \infty$ for all $i = 1, \ldots, N$.

The standard normal limit theory of the nonlinear IV t-ratio statistics continues to hold for the models with determininstic components.

6. Simulations

We conduct a set of simulations to investigate the finite sample performance of the average IV t-statistics S_N based on integrable IGF's for testing the unit root null hypothesis $H_0 : \alpha_i = 1$ for all $i = 1, \ldots, N$ against the stationarity alternatives $H_1 : |\alpha_i| < 1$ for some i. In particular, we explore how close are the finite sample sizes of the test S_N in relation to the corresponding nominal test sizes, using the critical values from its limit $\mathbf{N}(0,1)$ distribution, and compare its sizes and powers to those of the commonly used average statistics t-bar proposed by Im, Pesaran and Shin (1997).

For the simulation, we consider the (y_{it}) given by the model (1) with (u_{it}) generated as AR(1) processes, viz.,

$$u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it} \tag{33}$$

The innovations $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$ that generate $u_t = (u_{1t}, \ldots, u_{Nt})'$ are drawn from an N-dimensional multivariate normal distribution with mean zero and covariance matrix Σ. The AR coefficients, ρ_i's, used in the generation of the errors (u_{it}) are drawn randomly from the uniform distribution, i.e., $\rho_i \sim$ Uniform[0.2,0.4]. The parameter values for the $(N \times N)$ covariance matrix $\Sigma = (\sigma_{ij})$ are also randomly drawn, but with particular attention. To ensure that Σ is a symmetric positive definite matrix and to avoid the near singularity problem, we generate Σ following the steps outlined in Chang (2000). The steps are presented here for convenience:

(1) Generate an $(N \times N)$ matrix M from Uniform[0,1].
(2) Construct from M an orthogonal matrix $H = M(M'M)^{-1/2}$.
(3) Generate a set of N eigen values, $\lambda_1, \ldots, \lambda_N$. Let $\lambda_1 = r > 0$ and $\lambda_N = 1$ and draw $\lambda_2, \ldots, \lambda_{N-1}$ from Uniform[r,1].
(4) Form a diagonal matrix Λ with $(\lambda_1, \ldots, \lambda_N)$ on the diagonal.
(5) Construct the covariance matrix Σ as a spectral representation $\Sigma = H \Lambda H'$.

The covariance matrix constructed in this way will surely be symmetric and nonsingular with eigen values taking values from r to 1. We set the maximum eigen value at 1 since the scale does not matter. The ratio of the minimum eigenvalue to the maximum is therefore determined by the same parameter r. We now have some control over the size of the minimum eigen value and the ratio of the minimum to the maximum eigen values through the choice of r. The covariance matrix becomes singular as r tends to zero, and becomes spherical as r approaches to 1. For the simulations, we set r at $r = 0.1$.

For the estimation of the model (8) for $i = 1, \ldots, N$, we consider the IV estimator $\hat{\gamma}_i$ defined in (11) which uses the instrument $(F(y_{i,t-1}), \triangle y_{i,t-1}, \ldots, \triangle y_{i,t-p_i})'$. The instrument used for the lagged level $y_{i,t-1}$ is generated by the integrable IGF

$$F(y_{i,t-1}) = y_{i,t-1} e^{-c_i |y_{i,t-1}|}$$

where the factor c_i is proportional to the sample standard error of $\triangle y_{it} = u_{it}$ and

$T_i^{-1/2}$. That is,

$$c_i = KT_i^{-1/2} s(\triangle y_{it}) \quad \text{with} \quad s^2(\triangle y_{it}) = T_i^{-1} \sum_{t=1}^{T_i} \triangle y_{it}^2$$

where K is a constant. Notice that the larger the value of the factor c_i, the more integrable the IGF F becomes. The value of K is fixed at 5 for all $i = 1, \ldots, N$, and for all combinations of N and T considered here.[3] We note that the factor c_i in the definition of the instrument generating function F is data-dependent through the sample standard error of the difference of the data y_{it}. Hence, the value of c_i will be determined for each cross-sectional unit $i = 1, \ldots, N$. The shape of the integrable IV generating function F is given in Figure 1.

To test the unit root hypothesis, we set $\alpha_i = 1$ for all $i = 1, \ldots, N$, and investigate the finite sample sizes in relation to the corresponding nominal test sizes. To examine the rejection probabilities under the alternative of stationarity, we generate α_i's randomly from Uniform[0.8,1]. The model is thus heterogeneous under the alternative. The finite sample performance of the average IV t-ratio statistics S_N is compared with that of the t-bar statistics by Im, Pesaran and Shin (1997), which is based on the average of the individual t-tests computed from the sample ADF regressions (8) with mean and variance modifications. More explicitly, the t-bar statistics is defined as

$$t\text{-bar} = \frac{\sqrt{N}(\bar{t}_N - N^{-1} \sum_{i=1}^N \mathbf{E}(t_i))}{\sqrt{N^{-1} \sum_{i=1}^N \mathrm{var}(t_i)}}$$

where t_i is the t-statistics for testing $\alpha_i = 1$ for the i-th sample ADF regression (8), and $\bar{t}_N = N^{-1} \sum_{i=1}^N t_i$. The values of the expectation and variance, $\mathbf{E}(t_i)$ and $\mathrm{var}(t_i)$, for each individual t_i depend on T_i and the lag order p_i, and computed via simulations from independent normal samples. The number of time series observation T_i for each $i = 1, \ldots, N$ is required to be the same.[4]

The panels with the cross-sectional dimensions $N = 5, 15, 25, 50, 100$ and the time series dimensions $T = 25, 50, 100$[5] are considered for the 1%, 5% and 10% size tests. Since we are using randomly drawn parameter values, we simulate 20 times and report the ranges of the finite sample performances of the average IV nonlinear t-ratio statistics S_N and the t-bar test. Each simulation run is carried out with 10,000 simulation iterations. Tables 1, 2 and 3 report, respectively, the finite sample sizes,

[3]The test S_N constructed from the IGF F with a lager value of K tends to have smaller rejection probabilities uniformly over all the choices of N and T. The IGF defined with $K = 5$ seems to work best overally, and thus chosen for our simulations. For the cases where the time dimension is small $T = 25$, the average IV test S_N slightly over-rejects the null. In such cases, one might use a little larger value of K to correct the upward size distortion.

[4]Table 2 in Im, Pesaran and Shin (1997) tabulates the values of $\mathbf{E}(t_i)$ and $\mathrm{var}(t_i)$ for $T = 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 100$ and for $p_i = 1, \ldots, 8$.

[5]Though we allow heterogeneity in the number T_i of time series observations in our theoretical framework, we use the same T for all the cross-sectional units to be able to make our average IV t-statistics S_N comparable to the t-bar test by Im, Pesaran and Shin (1997), which requires T_i's to be the same across i.

the finite sample rejection probabilities and the size adjusted finite sample powers of the two tests. For each statistics, we report the minimum, mean, median and maximum of the rejection probabilities under the null and under the alternative hypothesis.

As can be seen from Table 1, the finite sample sizes of the test S_N are quite close to the corresponding nominal sizes. The sizes are calculated using the critical values from the standard normal distribution, and therefore the simulation results corroborate the asymptotic normal theory for S_N. The limit theory seems to work reasonably well even when the number of time series observation is relatively small, i.e., when $T = 25$, for all of the cross-sectional dimensions considered. On the other hand, the t-bar statistics exhibits noticeable size distortions, as reported, for instance, in the previous simulation work by Maddala and Wu (1996). The size distortions seem to be mostly upward for the 1% tests, and downward for 5% and 10% tests. For the cases where the number of cross-sectional units is larger relative to the number of time series observations, however, the t-bar suffers from serious upward size distortions for all tests. For example, when $N=100$ and $T=25$, the finite sample sizes of t-bar for 1%, 5% and 10% tests are, respectively, 32.9%, 35.0% and 36.1%.

The test S_N is more powerful than the t-bar statistics for all 1%, 5% and 10% tests and for all N and T combinations considered, as can be seen clearly from the results on the finite sample rejection probabilities and the size adjusted powers, reported repetitively in Tables 2 and 3. The discriminatory power of S_N is noticeably much higher than that of the t-bar statistics for the cases with smaller T and N. For the 1% tests with the combinations $(N,T) = \{(15,25),(25,25),(5,50)\}$, the power of the test S_N is more than twice as large as that of the t-bar statistics. The S_N still performs much better than the t-bar even when T is large, if the cross-sectional dimension is small. The performance of the t-bar statistics improves as both N and T increase, though the improvement is more noticeable with the growth in T. The differences in the finite sample powers of S_N and t-bar vanish as both N and T increases.

7. Empirical Illustrations

In this section, we apply the newly developed panel unit root test S_N to test whether the purchasing power parity (PPP) hypothesis holds. The PPP hypothesis has been tested numerously many times by many researchers using various unit root tests, both in panel as well as univariate models. Examples include MacDonald (1996), Frankel and Rose (1996), Oh (1996), Papell (1997), O'Connell (1998), just to name a few. There have been, however, conflicting evidences, and the issue does not seem to be completely settled.

We consider the data used in Papell (1997), which consists of the real exchange rates for twenty countries computed from the IMF's International Financial Statistics (IFS) tape, covering the period 1973:1 – 1998:4.[6] We also consider the data from Penn

[6]The data used in Pappel (1997) covers the period 1973:1 – 1994:3, but the data used here is extended to 1998:4. The data is quarterly, and is taken from the International Monetary Fund's

World Table (PWT) analyzed in Oh (1996).[7] The empirical results are summarized in Tables 4 and 5, respectively for the results from using data from Pappel (1997) and Oh (1996). We allow the models to have heterogeneous dynamic structure, i.e., the models may have different AR orders for individual cross-sectional units. For each cross-sectional unit the AR order is selected using BIC criterion with the maximum number of lags 4 and 6. To see how sensitive are the test results with respect to the specifications of individual dynamics, we also look at the panels with homogeneous dynamics, where we do not allow the AR order to vary across the individual units and fix the AR order at 2 and 4 for all cross-sectional units.

For the analysis of the PWT data, we looked at four different groups of countries. For each group of countries, the numbers of the time series observations are different, varying from 30 to 41.[8] The IFS data has total 104 time series observations. To examine the dependency on the sample size also for the test results from the IFS data set, we considered three sub-samples of the sizes 25, 50 and 100. The sub-samples are obtained by retaining the most recent observations.

For both data from Papell (1997) and Oh (1996), our test strongly and unambiguously support the PPP relationship. As seen from Tables 4 and 5, our test rejects the presence of the unit root in all cases.[9] The values of the test statistics S_N of course vary for different choices of the sample size T and the specifications of the dynamic structures, but they consistently reject the null hypothesis of the unit root. Our test appears to be fairly robust with respect to the specifications of model dynamics and the sizes of the samples.

In sharp contrast, the t-bar test by Im, Pesaran and Shin (1997) produces the results that are inconclusive. The test results are, in particular, sensitive to the specifications of the individual dynamic structures, and to the dimensions of the cross-sectional and time series observations. For the IFS data, we get contradictory results for each choice of the number of time series observations and maximum order in the BIC criterion. It appears that the test has the tendency to support the PPP when the sample size is large. However, this tendency is not observed when we do not allow for heterogeneous dynamics across individual units. The results from the

International Financial Statistics data tape. The countries considered include Austria, Belgium, Denmark, Finland, France, Germany, Italy, Japan, Netherlands, Norway, Spain, Sweden, Switzerland, United Kingdom, Ireland, Australia, Greece, New Zealand, Portugal, and Canada. The real exchange rate r_{it} for the i-th country is computed using the US dollar as the numeraire currency, and calculated as $r_{it} = \log(e_{it}p_{*t}/p_{it})$, where e_{it}, p_{*t} and p_{it} denote respectively the nominal spot exchange rate for the i-th country, the US CPI, and the CPI for the i-th country.

[7] The data used in Oh (1996) are yearly observations from the Penn World Table, Mark 5.5. The data is collected for 111 countries for the period 1960 – 1989, and extended to a longer period 1950 – 1990 for a group of 51 countries. For the longer sample, the data is analyzed for two sub-samples, the twenty-two OECD countries and G6 countries (Canada, France, Germany, Italy, Japan and United Kingdom).

[8] For the group of 111 countries, there are 30 annual observations. But for the the group of 51 countries (including its sub-samples of 22 OECD countries and G6 countries), there are 41 time series observations.

[9] There is only one exception. Our test is not able to reject the absence of PPP for G6 countries based on PWT data, when the dynamics is restricted to be AR(4) for all cross-sectional units.

PWT data are also mixed. The t-bar test supports or rejects the PPP depending upon how we select the countries and the time series observations. The test fails to reject the presence of the unit root except for one case where we have largest number of total observations.

8. Conclusions

This paper introduces an asymptotically normal unit root test for panels with cross-sectional dependency. The test is based on nonlinear IV estimation of the autoregressive coefficient using the instruments generated by the class of regularly integrable functions. The t-ratio statistics for the test of the unit root constructed from such nonlinear IV estimators is shown to have standard normal limit distribution, for each individual cross-sectional unit $i = 1, \ldots, N$, The nonlinear IV t-ratio statistics has simple symmetric confidence intervals both under the unit root null as well as under the stationarity alternatives. Therefore, there are no more discontinuity problems in the confidence intervals in the transition from stationary to nonstationary cases. The same results extend to the models with deterministic trends. More importantly, we show that the limit distributions of the nonlinear IV t-ratio statistics for testing for the unit root in individual cross-sectional units are cross-sectionally independent.

The asymptotic orthogonalities among the individual nonlinear IV t-ratio statistics naturally lead us to propose a standardized sum of the individual IV t-ratios for the test of the unit root for panels with cross-sectional dependency. We show that the limit theory of such standardized sum of individual nonlinear IV t-ratios, which we call the S_N statistics, is also standard normal. The limit theory is derived via T-asymptotics, which is not followed by N-asymptotics. The spatial dimension consequently is not required to be large, and therefore it may take any value, large or small. Moreover, the number of time series observations is allowed to be different across cross-sectional units, and thus our panel nonlinear IV method permits unbalanced panels. This implies that we can do simple inference based on the standard normal distribution even for unbalanced panels with general cross-sectional dependency.

The simulation results seem to well support our theoretical findings. The finite sample sizes of S_N calculated from using the standard normal critical values quite closely approximate the nominal test sizes. Moreover, the test S_N has noticeably higher discriminatory power than the commonly used average panel unit root test t-bar by Im, Pesaran and Shin (1997). The panel nonlinear IV unit root test seems to improve significantly upon the t-bar test under cross-sectional dependency, especially for smaller time and spatial dimensions. The new statistics S_N is applied to test whether the purchasing power parity hypothesis holds, using the data sets from the International Financial Statistics and the Penn World Table. Our test appears to be fairly robust to the specifications of the model dynamics and the sizes of the samples, and strongly and unambiguously support the PPP relationships, while the t-bar test by Im, Pesaran and Shin (1997) produces inconclusive results.

Appendix: Mathematical Proofs

Proof of Lemma 3.2 We have from the Beveridge-Nelson representation of y_{it} given in (7) that

$$T_i^{-1/2} y_{i[T_i r]} = \pi_i(1) T_i^{-1/2} \sum_{t=1}^{[T_i r]} \varepsilon_{it} + o_p(1)$$

where $\pi_i(1) = \alpha_i(1)^{-1}$. Then we have as $T_i \to \infty$

$$T_i^{-1/2} y_{i[T_i r]} = \pi_i(1) U_{iT_i}(r) + o_p(1) \to_d \pi_i(1) U_i(r) \tag{34}$$

since $U_{iT_i} \to_d U_i$ as $T_i \to \infty$, due to the invariance principle in (4). Then it follows from Lemma 3.1 (c) of Chang, Park and Phillips (1999) that

$$\begin{aligned} T_i^{-1/4} \sum_{t=1}^{T_i} F(y_{i,t-1})\, \varepsilon_{it} &\to_d \sigma_i \left(L_i(1,0) \int_{-\infty}^{\infty} F(\pi_i(1)\, s)^2 ds \right)^{1/2} W_i(1) \\ &= \sigma_i \left(\alpha_i(1)\, L_i(1,0) \int_{-\infty}^{\infty} F(s)^2 ds \right)^{1/2} W_i(1) \end{aligned}$$

by a simple change of variables. This establishes the result in part (a).

The stated result in part (b) is obtained similarly using the result in Lemma 3.1 (i) of Chang, Park and Phillips as follows (1999)

$$\begin{aligned} T_i^{-1/2} \sum_{t=1}^{T_i} F(y_{i,t-1})^2 &\to_d L_i(1,0) \int_{-\infty}^{\infty} F(\pi_i(1)\, s)^2 ds \\ &= \alpha_i(1)\, L_i(1,0) \int_{-\infty}^{\infty} F(s)^2 ds \end{aligned}$$

again by a simple change of variables.

For part (c), just note that $\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-k}$ are stationary regressors, and then the proof follows directly from the asymptotic orthogonality between the integrable transformations of integrated processes and stationary regressors established in part (e) of Lemma 3.1 in Chang, Park and Phillips (1999).

Proof of Theorem 3.3 We begin by investigating the limit behavior of A_{T_i} and C_{T_i} defined below (12) and (13), respectively. Recall $x'_{it} = (\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-p_i})$. Then it follows from Lemma 3.2 (c) that

$$\sum_{t=1}^{T_i} F(y_{i,t-1}) x'_{it} = o_p(T_i^{3/4})$$

which gives

$$\left| \sum_{t=1}^{T_i} F(y_{i,t-1}) x'_{it} \left(\sum_{t=1}^{T_i} x_{it} x'_{it} \right)^{-1} \sum_{t=1}^{T_i} x_{it} \varepsilon_{it} \right|$$

$$
\begin{aligned}
&\leq \left|\sum_{t=1}^{T_i} F(y_{i,t-1})x_{it}'\right| \left|\left(\sum_{t=1}^{T_i} x_{it}x_{it}'\right)^{-1}\right| \left|\sum_{t=1}^{T_i} x_{it}\varepsilon_{it}\right| \\
&= o_p(T_i^{3/4})O_p(T_i^{-1})O_p(T_i^{1/2}) \\
&= o_p(T_i^{1/4})
\end{aligned}
$$

and

$$
\begin{aligned}
&\left|\sum_{t=1}^{T_i} F(y_{i,t-1})x_{it}' \left(\sum_{t=1}^{T_i} x_{it}x_{it}'\right)^{-1} \sum_{t=1}^{T_i} x_{it}F(y_{i,t-1})\right| \\
&\leq \left|\sum_{t=1}^{T_i} F(y_{i,t-1})x_{it}'\right| \left|\left(\sum_{t=1}^{T_i} x_{it}x_{it}'\right)^{-1}\right| \left|\sum_{t=1}^{T_i} x_{it}F(y_{i,t-1})\right| \\
&= o_p(T_i^{3/4})O_p(T_i^{-1})o_p(T_i^{3/4}) \\
&= o_p(T_i^{1/2})
\end{aligned}
$$

Then we have

$$
T_i^{-1/4}A_{T_i} = T_i^{-1/4}\sum_{t=1}^{T_i} F(y_{i,t-1})\varepsilon_{it} + o_p(1)
$$

and

$$
T_i^{-1/2}C_{T_i} = T_i^{-1/2}\sum_{t=1}^{T_i} F(y_{i,t-1})^2 + o_p(1)
$$

Next we write Z_i defined in (16) as

$$
Z_i = \frac{\hat{\alpha}_i - 1}{s(\hat{\alpha}_i)} = \frac{B_{T_i}^{-1}A_{T_i}}{(\hat{\sigma}_i^2 B_{T_i}^{-2}C_{T_i})^{1/2}} = \frac{A_{T_i}}{\hat{\sigma}_i C_{T_i}^{1/2}}
$$

using the results in (12) and (17). Then it follows immediately from Lemma 3.2 (a) and (b) that

$$
\begin{aligned}
Z_i &= \frac{T_i^{-1/4}A_{T_i}}{\hat{\sigma}_i\left(T_i^{-1/2}C_{T_i}\right)^{1/2}} \\
&= \frac{T_i^{-1/4}\sum_{t=1}^{T_i} F(y_{i,t-1})\varepsilon_{it}}{\hat{\sigma}_i\left(T_i^{-1/2}\sum_{t=1}^{T_i} F(y_{i,t-1})^2\right)^{1/2}} + o_p(1) \\
&\to_d \frac{\sigma_i\left(\alpha_i(1)\, L_i(1,0)\int_{-\infty}^{\infty} F(s)^2 ds\right)^{1/2} W_i(1)}{\sigma_i\left(\alpha_i(1)\, L_i(1,0)\int_{-\infty}^{\infty} F(s)^2 ds\right)^{1/2}} \\
&\equiv W_i(1)
\end{aligned}
$$

as $T_i \to \infty$, and this establishes the stated result.

Proof of Lemma 4.2 We may assume w.l.o.g. that F is monotonic decreasing, since otherwise we may redefine F as such a function dominating the given F. Let $T_i \leq T_j$ for any $1 \leq i, j \leq N$. Then we have

$$\begin{aligned}
&\sqrt[4]{T_iT_j}\int_0^1 F(\sqrt{T_i}B_{iT_i}(r))F(\sqrt{T_j}B_{jT_j}(r))dr \\
&\leq \sqrt[4]{T_iT_j}\int_0^1 F(\sqrt{T_i}B_{iT_i}(r))F(\sqrt{T_i}B_{jT_j}(r))dr
\end{aligned} \tag{35}$$

However, due to the result in Kasahara and Kotani (1979), we have

$$\int_0^1 F(\sqrt{T_i}B_i(r))F(\sqrt{T_i}B_j(r))dr = O_p\left(\frac{\log T_i}{T_i}\right) \quad \text{a.s.} \tag{36}$$

and as shown in Chang, Park and Phillips (1999)

$$\begin{aligned}
&\frac{T_i}{\log T_i}\int_0^1 F(\sqrt{T_i}B_{iT_i}(r))F(\sqrt{T_j}B_{jT_j}(r))dr \\
&= \frac{T_i}{\log T_i}\int_0^1 F(\sqrt{T_i}B_i(r))F(\sqrt{T_j}B_j(r))dr + o_p(1)
\end{aligned} \tag{37}$$

We now have from (35), (36) and (37) that

$$\begin{aligned}
&\sqrt[4]{T_iT_j}\int_0^1 F(\sqrt{T_i}B_{iT_i}(r))F(\sqrt{T_j}B_{jT_j}(r))dr \\
&= \sqrt[4]{T_iT_j}\int_0^1 F(\sqrt{T_i}B_{iT_i}(r))F(\sqrt{T_i}B_{jT_j}(r))dr + o_p\left(\frac{T_j^{1/4}\log T_i}{T_i^{3/4}}\right) \\
&= O_p\left(\frac{T_j^{1/4}\log T_i}{T_i^{3/4}}\right) \\
&= O_p\left(\frac{{T_{\max}}^{1/4}\log T_{\max}}{{T_{\min}}^{3/4}}\right)
\end{aligned}$$

which is of order $o_p(1)$ if

$$\frac{{T_{\max}}^{1/4}\log T_{\max}}{{T_{\min}}^{3/4}} \to 0$$

as assumed. This establishes the stated result in (22).

As shown in Chang, Park and Phillips (1999), the result in (22) implies that

$$T_i^{-1/4}\sum_{t=1}^{T_i} F(y_{i,t-1})\varepsilon_{it}$$

are asymptotically independent for all $i = 1, \ldots, N$. It is clear that the results in Lemma 3.2 hold jointly, if $T_{\min} \to \infty$, and thus the proof is complete.

Proof of Theorem 4.3 Under the given assumptions, the individual IV t-ratio statistics Z_i's have the standard normal limit distributions as established in Theorem 3.3, and Z_i and Z_j become asymptotically independent for all $i \neq j$, as implied by Lemma 4.2. This suffices to establish the stated result.

Proof of Corollary 5.1 The least squares estimator of μ_i from the regression (26) with sub-samples upto time t is $t^{-1} \sum_{k=1}^{t} z_{ik}$. Then we have from (25) that

$$y_{it}^{\mu} = z_{it} - \hat{\mu}_i^t = z_{it} - \frac{1}{t}\sum_{k=1}^{t} z_{ik} = (\mu_i + y_{it}) - \frac{1}{t}\sum_{k=1}^{t}(\mu_i + y_{ik}) = y_{it} - \frac{1}{t}\sum_{k=1}^{t} y_{ik}$$

To examine the limit behavior of y_{it}^{μ}, we may write it as

$$y_{i[T_i r]}^{\mu} = y_{i[T_i r]} - \frac{1}{[T_i r]}\sum_{k=1}^{[T_i r]} y_{ik}$$

Then it follows from (34) and the invariance principle given in (4) that

$$\begin{aligned}
T_i^{-1/2} y_{i[T_i r]}^{\mu} &= T_i^{-1/2} y_{i[T_i r]} - \frac{1}{[T_i r]}\frac{T_i}{T_i}\sum_{k=1}^{[T_i r]} T_i^{-1/2} y_{ik} \\
&= \pi_i(1) U_{iT_i}(r) - \frac{1}{r}\int_0^r \pi_i(1) U_{iT_i}(s)ds + o_p(1) \\
&\to_d \pi_i(1) U_i^{\mu}(r)
\end{aligned}$$

as $T_i \to \infty$, where

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

To calculate the LS estimators of μ_i and δ_i from the sub-sample regression (30), we may rewrite the model as

$$z_{ik} = \left(\mu_i + \frac{t+1}{2}\delta_i\right) + \delta_i\left(k - \frac{t+1}{2}\right) + y_{ik}, \quad \text{for } k = 1, \ldots, t$$

Note that the regressors are orthogonal in this reparameterized regression. Hence, the LS estimators for the parameters in the reparameterized regression are easily calculated and are given by

$$\begin{aligned}
\hat{\mu}_i^t + \frac{t+1}{2}\hat{\delta}_i^t &= \mu_i + \frac{t+1}{2}\delta_i + \frac{1}{t}\sum_{k=1}^{t} y_{ik} \\
\hat{\delta}_i^t &= \delta_i + \left(\sum_{k=1}^{t}\left(k - \frac{t+1}{2}\right)^2\right)^{-1}\sum_{k=1}^{t}\left(k - \frac{t+1}{2}\right) y_{ik}
\end{aligned}$$

Now, we may write the fitted residual y_{it}^{τ} given in (29) more explicitly as

$$\begin{aligned}
y_{it}^{\tau} &= z_{it} - \hat{\mu}_i^t - \hat{\delta}_i^t t \\
&= z_{it} - \left(\hat{\mu}_i^t + \frac{t+1}{2}\hat{\delta}_i^t\right) - \frac{t-1}{2}\hat{\delta}_i^t \\
&= (\mu_i + \delta_i t + y_{it}) - \left(\mu_i + \frac{t+1}{2}\delta_i + \frac{1}{t}\sum_{k=1}^{t} y_{ik}\right) \\
&\quad -\frac{t-1}{2}\left(\delta_i + \left(\sum_{k=1}^{t}\left(k - \frac{t+1}{2}\right)^2\right)^{-1}\sum_{k=1}^{t}\left(k - \frac{t+1}{2}\right)y_{ik}\right) \\
&= y_{it} - \frac{1}{t}\sum_{k=1}^{t} y_{ik} - \frac{t-1}{2}\left(\sum_{k=1}^{t}\left(k - \frac{t+1}{2}\right)^2\right)^{-1}\sum_{k=1}^{t}\left(k - \frac{t+1}{2}\right)y_{ik} \\
&= y_{it} - \frac{1}{t}\sum_{k=1}^{t} y_{ik} + \frac{6}{t(t+1)}\left(\sum_{k=1}^{t} k\, y_{ik} - \frac{t+1}{2}\sum_{k=1}^{t} y_{ik}\right) \\
&= y_{it} - \frac{4}{t}\sum_{k=1}^{t} y_{ik} + \frac{6}{t(t+1)}\sum_{k=1}^{t} k\, y_{ik}
\end{aligned}$$

Then we have

$$y_{i[T_i r]}^{\tau} = y_{i[T_i r]} - \frac{4}{[T_i r]}\sum_{k=1}^{[T_i r]} y_{ik} + \frac{6}{[T_i r]([T_i r]+1)}\sum_{k=1}^{t} k\, y_{ik}$$

giving

$$\begin{aligned}
T_i^{-1/2} y_{i[T_i r]}^{\tau} &= T_i^{-1/2} y_{i[T_i r]} - \frac{4T_i}{[T_i r]}\frac{1}{T_i}\sum_{k=1}^{[T_i r]} T_i^{-1/2} y_{ik} + \frac{6T_i^2}{[T_i r]([T_i r]+1)}\frac{1}{T_i}\sum_{k=1}^{t}\frac{k}{T_i}T_i^{-1/2} y_{ik} \\
&= \pi_i(1)U_{iT_i}(r) - \frac{4}{r}\int_0^r \pi_i(1)U_{iT_i}(s)ds + \frac{6}{r^2}\int_0^r s\pi_i(1)U_{iT_i}(s)ds + o_p(1) \\
&\to_d \pi_i(1)U_i^{\tau}(r)
\end{aligned}$$

as $T_i \to \infty$ due to (34) and (4), where

$$U_i^{\tau}(r) = U_i(r) - \frac{4}{r}\int_0^r U_i(s)ds + \frac{6}{r^2}\int_0^r sU_i(s)ds$$

Now the limit theories given in Lemma 3.2 for the sample moments for the models with no deterministic trend easily extends to the sample moments from the models with deterministic trends. The limit theories in fact have the identical expressions with the local times of adaptively demeaned and detrended Brownian motions U_i^{μ} and U_i^{τ} in the place of the original U_i. Using these modified limit theories, we may then easily derive the limit distributions of the nonlinear IV t-ratio statistics Z_i^{μ} and Z_i^{τ} for the models with nonzero mean and deterministic trends.

References

Chang, Y. (2000). "Bootstrap unit root tests in panels with cross-sectional dependency," Rice University, mimeographed.

Chang, Y. and J.Y. Park (1999). "Nonstationary index models," mimeographed, Department of Economics, Rice University.

Chang, Y., J.Y. Park and P.C.B. Phillips (1999). "Nonlinear econometric models with cointegrated and deterministically trending regressors," mimeographed, Department of Economics, Rice University.

Choi, I. (1997). "Unit root tests for panel data," mimeographed, Department of Economics, Kookmin University.

Chung, K.L. and Williams, R.J. (1990). *Introduction to Stochastic Integration*, 2nd ed. Birkhäuser, Boston.

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.

Im, K.S., M.H. Pesaran and Y. Shin (1997). "Testing for unit roots in heterogeneous panels," mimeographed.

Levin, A., C.F. Lin and C.S. Chu (1997). "Unit root tests in panel data: Asymptotic and finite sample properties," University of California, San Diego, mimeographed.

Kasahara, Y. and S. Kotani (1979). "On limit processes for a class of additive functionals of recurrent diffusion processes," *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **49**: 133-153.

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

Maddala, G.S. and S. Wu (1996). "A comparative study of unit root tests with panel data and a new simple test: Evidence from simulations and bootstrap". Ohio State University, mimeographed.

O'Connell, P.G.J. (1998). "The overevaluation of purchasing power parity," *Journal of International Economics*, 44, 1-20.

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. (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.

Park, J.Y. and P.C.B. Phillips (2000). "Nonlinear regressions with integrated time series," forthcoming in *Econometrica.*

Phillips, P.C.B., J.Y. Park and Y. Chang (1999). "Nonlinear instrumental variable estimation of an autoregression," mimeographed, Department of Economics, Rice University.

Phillips, P.C.B. and V. Solo (1992). "Asymptotics for linear processes," *Annals of Statistics*, 20, 971-1001.

Revuz, D. And M. Yor (1994). *Continuous Martingale and Brownian Motion.* New York: Springer–Verlag.

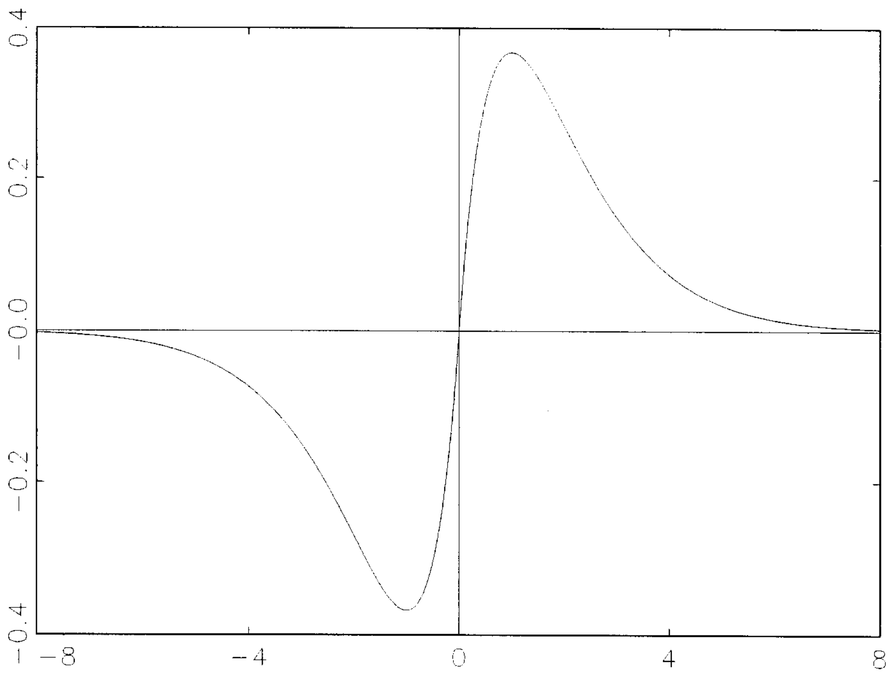

Figure 1: IGF $F(x) = xe^{-|x|}$

Table 1: Finite Sample Sizes

				1%	test			5%	test			10%	test	
N	T	tests	min	mean	med	max	min	mean	med	max	min	mean	med	max
5	25	t-bar	0.017	0.020	0.020	0.022	0.033	0.036	0.036	0.039	0.044	0.047	0.047	0.051
		S_N	0.012	0.016	0.016	0.019	0.053	0.059	0.060	0.065	0.100	0.109	0.109	0.117
15	25	t-bar	0.063	0.067	0.067	0.071	0.085	0.090	0.089	0.094	0.098	0.103	0.102	0.108
		S_N	0.018	0.021	0.021	0.025	0.070	0.077	0.075	0.086	0.125	0.135	0.134	0.147
25	25	t-bar	0.098	0.102	0.102	0.106	0.119	0.124	0.123	0.129	0.134	0.138	0.138	0.142
		S_N	0.017	0.020	0.020	0.022	0.064	0.073	0.073	0.078	0.121	0.131	0.131	0.140
50	25	t-bar	0.179	0.185	0.185	0.189	0.202	0.208	0.208	0.214	0.211	0.219	0.220	0.226
		S_N	0.017	0.019	0.018	0.021	0.067	0.072	0.072	0.079	0.122	0.129	0.128	0.136
100	25	t-bar	0.321	0.329	0.329	0.342	0.342	0.350	0.350	0.362	0.353	0.361	0.361	0.374
		S_N	0.019	0.023	0.022	0.025	0.079	0.084	0.085	0.089	0.139	0.146	0.146	0.152
5	50	t-bar	0.012	0.015	0.015	0.017	0.026	0.028	0.028	0.030	0.034	0.036	0.036	0.039
		S_N	0.011	0.014	0.014	0.016	0.048	0.055	0.055	0.061	0.096	0.103	0.102	0.113
15	50	t-bar	0.041	0.044	0.043	0.047	0.055	0.060	0.060	0.065	0.065	0.070	0.069	0.076
		S_N	0.014	0.017	0.017	0.019	0.059	0.066	0.066	0.071	0.108	0.119	0.119	0.127
25	50	t-bar	0.054	0.059	0.059	0.062	0.068	0.074	0.074	0.077	0.075	0.083	0.084	0.086
		S_N	0.011	0.014	0.014	0.017	0.053	0.058	0.058	0.064	0.100	0.107	0.107	0.116
50	50	t-bar	0.088	0.094	0.094	0.100	0.099	0.108	0.108	0.114	0.108	0.117	0.117	0.122
		S_N	0.011	0.012	0.012	0.015	0.048	0.052	0.052	0.059	0.091	0.097	0.097	0.105
100	50	t-bar	0.140	0.147	0.146	0.158	0.152	0.160	0.160	0.172	0.160	0.167	0.167	0.180
		S_N	0.011	0.013	0.013	0.015	0.050	0.054	0.053	0.057	0.094	0.100	0.100	0.108
5	100	t-bar	0.011	0.013	0.013	0.015	0.021	0.023	0.024	0.026	0.028	0.032	0.032	0.035
		S_N	0.010	0.013	0.013	0.017	0.051	0.056	0.056	0.060	0.096	0.103	0.103	0.109
15	100	t-bar	0.032	0.035	0.035	0.038	0.044	0.047	0.047	0.050	0.052	0.055	0.055	0.059
		S_N	0.013	0.016	0.016	0.018	0.054	0.062	0.063	0.070	0.103	0.112	0.114	0.121
25	100	t-bar	0.042	0.045	0.045	0.049	0.053	0.057	0.057	0.061	0.060	0.064	0.064	0.068
		S_N	0.011	0.013	0.013	0.016	0.049	0.054	0.054	0.060	0.094	0.101	0.101	0.108
50	100	t-bar	0.058	0.062	0.061	0.067	0.068	0.072	0.072	0.078	0.073	0.077	0.078	0.083
		S_N	0.010	0.011	0.010	0.013	0.041	0.046	0.045	0.050	0.078	0.088	0.089	0.095
100	100	t-bar	0.083	0.089	0.090	0.093	0.092	0.099	0.099	0.103	0.098	0.104	0.105	0.108
		S_N	0.010	0.011	0.011	0.013	0.041	0.046	0.046	0.053	0.081	0.087	0.086	0.098

Table 2: Finite Sample Rejection Probabilities

			1% test				5% test				10% test			
N	T	tests	min	mean	med	max	min	mean	med	max	min	mean	med	max
5	25	t-bar	0.047	0.074	0.069	0.115	0.079	0.119	0.113	0.175	0.102	0.149	0.140	0.211
		S_N	0.053	0.097	0.087	0.168	0.166	0.257	0.240	0.392	0.271	0.381	0.364	0.537
15	25	t-bar	0.312	0.375	0.381	0.450	0.377	0.443	0.448	0.519	0.408	0.475	0.481	0.551
		S_N	0.272	0.380	0.388	0.488	0.525	0.638	0.649	0.742	0.663	0.762	0.773	0.847
25	25	t-bar	0.515	0.630	0.632	0.720	0.566	0.677	0.678	0.763	0.595	0.702	0.705	0.786
		S_N	0.414	0.583	0.581	0.726	0.677	0.811	0.817	0.904	0.793	0.893	0.897	0.953
50	25	t-bar	0.870	0.923	0.924	0.969	0.890	0.936	0.937	0.975	0.897	0.941	0.942	0.977
		S_N	0.764	0.882	0.888	0.969	0.920	0.969	0.970	0.995	0.961	0.987	0.989	0.998
100	25	t-bar	0.994	0.998	0.998	1.000	0.995	0.999	0.999	1.000	0.996	0.999	0.999	1.000
		S_N	0.987	0.997	0.998	1.000	0.998	1.000	1.000	1.000	0.999	1.000	1.000	1.000
5	50	t-bar	0.067	0.168	0.138	0.331	0.112	0.248	0.214	0.449	0.141	0.292	0.255	0.508
		S_N	0.146	0.306	0.262	0.552	0.345	0.561	0.525	0.811	0.488	0.695	0.673	0.899
15	50	t-bar	0.589	0.727	0.750	0.853	0.657	0.781	0.806	0.887	0.689	0.805	0.829	0.905
		S_N	0.761	0.885	0.905	0.965	0.925	0.971	0.980	0.994	0.967	0.988	0.993	0.998
25	50	t-bar	0.827	0.934	0.941	0.984	0.858	0.949	0.956	0.989	0.875	0.956	0.962	0.991
		S_N	0.928	0.983	0.988	0.999	0.986	0.998	0.999	1.000	0.995	1.000	1.000	1.000
50	50	t-bar	0.994	0.999	1.000	1.000	0.995	0.999	1.000	1.000	0.996	1.000	1.000	1.000
		S_N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	50	t-bar	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		S_N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	100	t-bar	0.191	0.581	0.524	0.946	0.275	0.673	0.639	0.975	0.333	0.721	0.701	0.985
		S_N	0.435	0.765	0.759	0.991	0.708	0.913	0.931	1.000	0.824	0.956	0.971	1.000
15	100	t-bar	0.984	0.996	0.999	1.000	0.990	0.998	0.999	1.000	0.992	0.998	1.000	1.000
		S_N	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25	100	t-bar	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		S_N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	100	t-bar	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		S_N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	100	t-bar	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		S_N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 3: Finite Sample Powers

				1%	test			5%	test			10%	test	
N	T	tests	min	mean	med	max	min	mean	med	max	min	mean	med	max
5	25	t-bar	0.022	0.041	0.040	0.068	0.109	0.156	0.149	0.217	0.201	0.269	0.260	0.349
		S_N	0.034	0.070	0.065	0.129	0.143	0.228	0.209	0.359	0.249	0.361	0.345	0.515
15	25	t-bar	0.074	0.112	0.112	0.151	0.257	0.317	0.314	0.398	0.399	0.469	0.464	0.549
		S_N	0.187	0.263	0.265	0.359	0.437	0.546	0.552	0.663	0.606	0.697	0.706	0.798
25	25	t-bar	0.136	0.208	0.202	0.298	0.366	0.474	0.472	0.568	0.513	0.626	0.626	0.715
		S_N	0.309	0.466	0.457	0.625	0.596	0.749	0.751	0.852	0.733	0.857	0.862	0.932
50	25	t-bar	0.318	0.455	0.454	0.607	0.614	0.736	0.737	0.853	0.750	0.845	0.850	0.928
		S_N	0.654	0.818	0.823	0.942	0.882	0.951	0.956	0.990	0.941	0.980	0.982	0.997
100	25	t-bar	0.654	0.810	0.817	0.900	0.878	0.947	0.947	0.977	0.942	0.977	0.979	0.992
		S_N	0.969	0.992	0.993	0.999	0.995	0.999	0.999	1.000	0.998	1.000	1.000	1.000
5	50	t-bar	0.047	0.129	0.111	0.280	0.176	0.352	0.320	0.584	0.306	0.507	0.479	0.730
		S_N	0.116	0.258	0.222	0.483	0.310	0.538	0.500	0.803	0.477	0.687	0.662	0.898
15	50	t-bar	0.294	0.455	0.462	0.596	0.612	0.752	0.769	0.866	0.749	0.859	0.877	0.940
		S_N	0.686	0.832	0.847	0.942	0.903	0.959	0.968	0.992	0.961	0.985	0.990	0.998
25	50	t-bar	0.521	0.744	0.748	0.901	0.813	0.922	0.929	0.981	0.898	0.967	0.973	0.994
		S_N	0.882	0.975	0.982	0.998	0.980	0.997	0.998	1.000	0.994	0.999	1.000	1.000
50	50	t-bar	0.905	0.974	0.981	0.998	0.985	0.996	0.998	1.000	0.994	0.999	1.000	1.000
		S_N	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	50	t-bar	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		S_N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
5	100	t-bar	0.182	0.540	0.480	0.925	0.441	0.786	0.785	0.993	0.615	0.882	0.899	0.999
		S_N	0.375	0.731	0.715	0.982	0.678	0.903	0.926	0.999	0.818	0.955	0.971	1.000
15	100	t-bar	0.940	0.982	0.992	0.999	0.990	0.998	1.000	1.000	0.997	0.999	1.000	1.000
		S_N	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25	100	t-bar	0.993	1.000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		S_N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
50	100	t-bar	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		S_N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	100	t-bar	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
		S_N	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Table 4: PPP Tests for IFS Data

		t-bar			S_N		
		T=25	T=50	T=100	T=25	T=50	T=100
AR order	2	-1.200	-1.589	-1.183	-3.425	-4.362	-4.087
	4	-5.362	-7.108	-4.525	-4.099	-8.539	-7.362
BIC max order	4	0.555	-2.623	-2.420	-2.534	-5.438	-5.464
	6	-0.064	-1.127	-2.646	-2.544	-2.962	-5.613
critical values (%)	1	-1.990	-1.980	-1.970		N(0,1)	
	5	-1.850	-1.850	-1.840			
	10	-1.780	-1.780	-1.770			

Table 5: PPP Tests for PWT Data

		t-bar				S_N			
		G6	OECD	51Con	111Con	G6	OECD	51Con	111Con
AR order	2	-0.400	-0.669	-0.912	-3.420	-2.219	-3.865	-4.702	-9.896
	4	0.247	-0.101	-1.213	-3.031	-1.187	-3.157	-4.133	-7.537
BIC max order	4	-0.575	-1.026	-1.636	-2.823	-2.605	-5.012	-6.389	-9.709
	6	-0.575	-0.888	-0.891	-1.789	-2.605	-4.678	-6.151	-9.077
critical values (%)	1	-2.430	-1.980	-1.810	-1.730		N(0,1)		
	5	-2.160	-1.850	-1.730	-1.670				
	10	-2.020	-1.780	-1.680	-1.640				