



ELSEVIER

Journal of Econometrics 110 (2002) 261–292

JOURNAL OF
Econometrics

www.elsevier.com/locate/econbase

Nonlinear IV unit root tests in panels with cross-sectional dependency

Yoosoon Chang*

Department of Economics-MS22, Rice University, 6100 Main Street, Houston, TX 77005-1892, USA

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 augmented Dickey–Fuller type regression for each cross-sectional unit, using as instruments nonlinear transformations of the lagged levels. The actual test statistic 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 with 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.

© 2002 Elsevier Science B.V. All rights reserved.

JEL classification: C1; C15; C32; C33

Keywords: Panels with cross-sectional dependency; Unit root tests; Nonlinear instruments; Average IV t -ratio statistics

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

* Corresponding author. Tel.: +713-348-2796; fax: +713-348-5278.

E-mail address: yoosoon@rice.edu (Yoosoon Chang).

divergence, and purchasing power parity (PPP), 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 et al. (1997), Im et al. (1997), Maddala and Wu (1999), Choi (2001) 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 (1999) 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 et al. (1997) and Im et al. (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 eliminate the 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 statistic 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 statistic for testing the unit root has limiting standard normal distribution under the 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 -statistic 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 -statistic based on the ordinary least-squares estimator.

More importantly, we show that the individual IV t -ratio statistics are asymptotically 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 et al. (2001). We are therefore led to consider the average of these independent individual IV t -ratio statistics as a statistic for testing joint unit root null hypothesis for the entire panel. The actual test statistic 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 $T_{\min} \rightarrow \infty$ and $T_{\max}^{1/4} \log T_{\max}/T_{\min}^{3/4} \rightarrow 0$, 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 $i = 1, \dots, N$. 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 or large.¹ Our test is applicable for all panels that have large numbers of individual time series observations and are asymptotically balanced in the sense mentioned above.

Finite sample performance of our average IV t -ratio statistic, which we call S_N statistic, is examined via a set of simulations, and compared with that of the commonly used average statistic t -bar by Im et al. (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 very 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 (1999) for example. The discriminatory powers of S_N are yet noticeably higher than the t -bar test. We also apply our nonlinear IV method to test for the PPP hypothesis using the data sets from Papell (1997) and Oh (1996). Our test S_N supports strongly the PPP relationships, contrary to most of the previous empirical findings which are usually inconclusive.

The rest of the paper is organized as follows. Section 2 introduces the model, assumptions and background 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 components such as constant and linear time trend. In Section 6, we conduct simulations to investigate finite sample performance of the average IV t -ratio statistic. Section 7 provides empirical illustrations for testing the PPP using our nonlinear panel IV unit root test. Section 8 concludes, and mathematical proofs are provided in the Appendix.

2. Model, assumptions and background theory

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

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

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.

¹ The asymptotics developed here are T -asymptotics. Throughout the paper, we assume that N is fixed.

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 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 $\alpha^i(z)$ and p_i (which is assumed to be fixed) 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 (ε_t) is assumed to satisfy the following assumption.

Assumption 2.2. (ε_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 Lebesgue measure and has characteristic function φ such that $\lim_{\lambda \rightarrow \infty} |\lambda|^r \varphi(\lambda) = 0$, 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. Note that here the errors are

assumed to be iid across time periods. However, they are allowed to be cross-sectionally dependent. The technical assumption on the characteristic function is required for our subsequent asymptotics on nonlinear functions of integrated processes as used in Park and Phillips (1999, 2001).

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 Assumption 2.2, an invariance principle holds for U_T , viz.,

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

as $T \rightarrow \infty$, where U is an N -dimensional vector Brownian motion with covariance matrix Σ . It is also convenient to define $B_T(r)$ from $u_t = (u_{1t}, \dots, u_{Nt})'$, similarly as $U_T(r)$. Then we have $B_T \rightarrow_d B$, where $B = (B_1, \dots, B_N)'$ 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 as in Park and Phillips (1999, 2001), Chang and Park (1999) and Chang et al. (2001). The reader is referred to these papers and the references cited there for the concept of local time and its use in the asymptotics for nonlinear models with integrated time series. 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 $U(r) = (U_1(r), \dots, U_N(r))'$. We denote by L_i the (scaled) local time of U_i , for $i = 1, \dots, N$, which is defined by

$$L_i(t, s) = \lim_{\epsilon \rightarrow 0} \frac{1}{2\epsilon} \int_0^t 1\{|U_i(r) - s| < \epsilon\} dr.$$

The local time L_i is therefore the time that the Brownian motion U_i spends in the neighborhood of s , up to time t , measured in chronological units.² Then we may have an important relationship

$$\int_0^t G(U_i(r)) dr = \int_{-\infty}^{\infty} G(s)L_i(t, s) ds \tag{5}$$

which we refer to as the occupation times formula.

In addition to the Brownian motions $U = (U_1, \dots, U_N)'$, we need to introduce another set of standard Brownian motions $W = (W_1, \dots, W_N)'$. 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.

² Usually, the local time is defined in units of quadratic variation time. Therefore, the local time L_{U_i} of U_i defined in the usual manner is given by $L_{U_i} = \sigma_i^2 L_i$ in terms of our local time L_i , where σ_i^2 is the variance of U_i .

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 (ε_t) . Let ℓ be the maximum order of the existing moments. If $\ell > 8$, any piecewise Lipschitz 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 definition of the regularly integrable function in Definition 2.3 is identical to the one introduced in Park and Phillips (1999, 2001).

The asymptotic behaviors of the nonlinear functions of an integrated time series are analyzed by Park and Phillips (1999, 2001). 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 those 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

$$y_{it} = \sum_{k=1}^t u_{ik} = \pi_i(1) \zeta_{it} + (\tilde{u}_{i0} - \tilde{u}_{it}), \tag{6}$$

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

Using the specification of the regression error u_{it} given in (2), we write 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 $\Delta y_{it} = u_{it}$ under the unit root null hypothesis, the above regression may be written as

$$y_{it} = \alpha_i y_{i,t-1} + \sum_{k=1}^{p_i} \alpha_{i,k} \Delta y_{i,t-k} + \varepsilon_{it} \tag{7}$$

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 (7). To deal with the cross-sectional dependency, we use the instrument generated by a nonlinear function F as

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

for the lagged level $y_{i,t-1}$. For the lagged differences $x'_{it} = (\Delta y_{i,t-1}, \dots, \Delta y_{i,t-p_i})$, 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

$$(F(y_{i,t-1}), x'_{it})' = (F(y_{i,t-1}), \Delta y_{i,t-1}, \dots, \Delta y_{i,t-p_i})'. \tag{8}$$

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 $\int_{-\infty}^{\infty} xF(x) dx \neq 0$.

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 et al. (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 IGF is orthogonal to the regression function.

Examples of the regularly integrable IGFs 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 $\text{sgn}(x)1\{|x| < K\}$ and $x1\{|x| < K\}$. Also included are functions of the type $xe^{-|x|}$. The IV estimator, for example, constructed from the indicator function $F(x) = 1\{|x| \leq K\}$ is simply the trimmed OLS estimator,

i.e., the OLS estimator which uses only the observations taking values in the interval $[-K, K]$.

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 $x'_{it} = (\Delta y_{i,t-1}, \dots, \Delta y_{i,t-p_i})$. Then the augmented autoregression (7) can be written in matrix form as

$$y_i = y_{\ell i} \alpha_i + X_i \beta_i + \varepsilon_i = Y_i \gamma_i + \varepsilon_i, \tag{9}$$

where $\beta_i = (\alpha_{i,1}, \dots, \alpha_{i,p_i})'$, $Y_i = (y_{\ell i}, X_i)$, and $\gamma_i = (\alpha_i, \beta_i)'$. For the augmented autoregression (9), we consider the estimator $\hat{\gamma}_i$ of γ_i given by

$$\hat{\gamma}_i = \begin{pmatrix} \hat{\alpha}_i \\ \hat{\beta}_i \end{pmatrix} = (W_i' Y_i)^{-1} W_i' y_i = \begin{pmatrix} F(y_{\ell i})' y_{\ell i} & F(y_{\ell i})' X_i \\ X_i' y_{\ell i} & X_i' X_i \end{pmatrix}^{-1} \begin{pmatrix} F(y_{\ell i})' y_i \\ X_i' y_i \end{pmatrix}, \tag{10}$$

where $W_i = (F(y_{\ell i}), X_i)$ with $F(y_{\ell i}) = (F(y_{i,p_i}), \dots, F(y_{i,T_i-1}))'$. 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 (10). Under the null, we have

$$\hat{\alpha}_i - 1 = B_{T_i}^{-1} A_{T_i}, \tag{11}$$

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 A_{T_i} is given by

$$\sigma_i^2 \mathbf{E} C_{T_i}$$

under Assumption 2.2, where

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

For testing the unit root hypothesis $H_0 : \alpha_i = 1$ for each $i = 1, \dots, N$, we construct the t -ratio statistic from the nonlinear IV estimator $\hat{\alpha}_i$ defined in (11). More specifically, we construct such an IV t -ratio statistic for testing for a unit root in (1) or (7) as

$$Z_i = \frac{\hat{\alpha}_i - 1}{s(\hat{\alpha}_i)}, \tag{12}$$

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

$$s(\hat{\alpha}_i)^2 = \hat{\sigma}_i^2 B_{T_i}^{-2} C_{T_i}. \tag{13}$$

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 (7), viz.,

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

It is natural in our context to use the IV estimate $(\hat{\alpha}_i, \hat{\beta}_i)$ given in (10) 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 statistic Z_i introduced in (12), we need to obtain the asymptotics for various sample product moments appearing in A_{T_i} , B_{T_i} and C_{T_i} . For this, we need to introduce a set of independent standard Brownian motions W_1, \dots, W_N, x which are independent of the Brownian motions U_1, \dots, U_N . The limit theories are presented in the following lemma.

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

- (a) $T_i^{-1/4} \sum_{t=1}^{T_i} F(y_{i,t-1}) \varepsilon_{it} \rightarrow_d \sigma_i(\alpha^i(1) L_i(1, 0) \int_{-\infty}^{\infty} F(s)^2 ds)^{1/2} W_i(1)$,
- (b) $T_i^{-1/2} \sum_{t=1}^{T_i} F(y_{i,t-1})^2 \rightarrow_d \alpha^i(1) L_i(1, 0) \int_{-\infty}^{\infty} F(s)^2 ds$,
- (c) $T_i^{-3/4} \sum_{t=1}^{T_i} F(y_{i,t-1}) \Delta y_{i,t-k} \rightarrow_p 0$, for $k = 1, \dots, p_i$

jointly as $T_i \rightarrow \infty$, 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 5 in Chang and Park (1999). For the detailed discussion on the asymptotics here, the reader is referred to Park and Phillips (1999, 2001) and Chang et al. (2001). 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, 2001) and Chang et al. (2001).

The limit null distribution of the IV t -ratio statistic Z_i defined in (12) now follows readily from the results in Lemma 3.2.

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

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

as $T_i \rightarrow \infty$ for all $i = 1, \dots, N$.

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 $i = 1, \dots, N$, as we show in the next section. Our limit theory here is thus fundamentally different from the usual unit root asymptotics. This is due to the local time asymptotics and mixed normality of the results in Lemma 3.2. The nonlinearity of the IV is essential for our Gaussian limit theory. The result in Theorem 3.3, and thereby any of the subsequent result, is not applicable for the usual linear unit root tests such as those by Phillips (1987) and Phillips and Perron (1988).

We now consider the limit behavior of our IV t -ratio statistic under the alternative of stationarity to discuss the consistency of the test. Note that under the alternative, i.e., $\alpha_i = \alpha_{i0} < 1$, our IV t -ratio Z_i given in (12) can be expressed as

$$Z_i = Z_i(\alpha_{i0}) + \frac{\sqrt{T_i}(\alpha_{i0} - 1)}{\sqrt{T_i}s(\hat{\alpha}_i)}, \tag{14}$$

where $s(\hat{\alpha}_i)$ is defined in (13) and

$$Z_i(\alpha_{i0}) = \frac{\hat{\alpha}_i - \alpha_{i0}}{s(\hat{\alpha}_i)} \tag{15}$$

which is the IV t -ratio statistic for testing $\alpha_i = \alpha_{i0} < 1$. Under the alternative, we may expect that $Z_i(\alpha_{i0}) \rightarrow_d \mathbf{N}(0, 1)$ if the usual mixing conditions for (y_{it}) are assumed to hold. Moreover, if we let $B_{i0} = \text{plim}_{T_i \rightarrow \infty} T_i^{-1}B_{T_i}$ and $C_{i0} = \text{plim}_{T_i \rightarrow \infty} T_i^{-1}C_{T_i}$ exist under suitable mixing conditions for (y_{it}) , then the second term in the right-hand side of Eq. (14) diverges to $-\infty$ at the rate of $\sqrt{T_i}$. This is because

$$\sqrt{T_i}(\alpha_{i0} - 1) \rightarrow -\infty \quad \text{and} \quad \sqrt{T_i}s(\hat{\alpha}_i) \rightarrow_p v_i,$$

where $v_i^2 = \sigma_i^2 B_{i0}^{-2} C_{i0} > 0$. Hence, the IV t -ratio Z_i diverges at the $\sqrt{T_i}$ -rate under the alternative of stationarity, just as in the case of the usual OLS-based t -type unit root tests such as the augmented Dickey–Fuller test.

We also note that the IV t -ratios constructed with regularly integrable IGF are normally distributed asymptotically, for all $|\alpha_i| \leq 1$. The continuity of the distribution across the values of α_i of the t -ratio $Z_i(\alpha_i)$ defined in (15) also allows us to construct the confidence intervals for α_i from the IV estimator. As we have noticed above, $Z_i(\alpha_i) \rightarrow_d \mathbf{N}(0, 1)$ 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

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

using the IV estimator 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 statistic we propose here to test for the unit root hypothesis in a panel is basically an average of t -ratio statistics for testing the unity of the AR coefficient computed individually from each cross-sectional unit. More specifically, we test for the joint unit root null hypothesis $H_0: \alpha_i = 1$ for all $i = 1, \dots, N$ using an average statistic based on the individual t -statistics for testing $\alpha_i = 1$ in (7) constructed from the nonlinear IV estimator $\hat{\alpha}_i$ defined in (11). The *average IV t-ratio statistic* is thus defined as

$$S_N = \frac{1}{\sqrt{N}} \sum_{i=1}^N Z_i, \tag{16}$$

where Z_i is the individual nonlinear IV t -ratio statistic, defined in (12), for testing $\alpha_i = 1$ for the i th cross-sectional unit. For the average statistic S_N , we allow each of the cross-sectional units $i = 1, \dots, N$ 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 statistic S_N , we first investigate how the individual IV t -ratio statistics Z_i 's interact in the limit. We have

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

$$T_j^{-1/4} \sum_{t=1}^{T_j} F(y_{j,t-1}) \varepsilon_{jt} \approx_d \sqrt[4]{T_j} \int_0^1 F(\sqrt{T_j} B_{jT_j}) dU_j,$$

which become asymptotically independent if their quadratic covariation

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

converges a.s. to zero, where σ_{ij} denotes the covariance between U_i and U_j . This was shown in Chang et al. (2001). 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 $i = 1, \dots, N$.

Assumption 4.1. Assume

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

Then we have

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

$$\sqrt[4]{T_i T_j} \int_0^1 F(\sqrt{T_i} B_{iT_i}(r)) F(\sqrt{T_j} B_{jT_j}(r)) dr \rightarrow_p 0 \tag{17}$$

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

The result in Lemma 4.2 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, even when the variables $y_{i,t-1}$ and $y_{j,t-1}$ generating the instruments are correlated. 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 below.

The limit theory for S_N follows immediately from Theorem 3.3 and Lemma 4.2, and is provided in

Theorem 4.3. We have

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

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 Assumption 4.1. It should be noted that the usual sequential asymptotics is not used here.³ The factor $N^{-1/2}$ in the definition of the test statistic S_N in (16) 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 implies that we can do simple inference based on the standard normal distribution even for unbalanced panels with general cross-sectional dependencies.

³ 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.

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 et al. (1997) and the group mean t -bar statistic by Im et al. (1997) have normal asymptotics. However, they all presume cross-sectional independence and their normal limit theories are obtained through sequential asymptotics.⁴ 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 IV 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 using properly demeaned or detrended data. A proper demeaning or detrending scheme required here must be able to successfully remove the nonzero mean and time trend, while maintaining the martingale property of the errors and ultimately the Gaussian limit theory for our nonlinear IV unit root tests. We now introduce our demeaning and detrending schemes.

If the time series (z_{it}) with a nonzero mean is given by

$$z_{it} = \mu_i + y_{it}, \tag{18}$$

where the stochastic component (y_{it}) is generated as in (1), then we may test for the presence of a unit root in (y_{it}) from the augmented regression (7) defined with the demeaned series y_{it}^μ and $y_{i,t-1}^\mu$ of z_{it} and $z_{i,t-1}$, viz.,

$$y_{it}^\mu = \alpha_i y_{i,t-1}^\mu + \sum_{k=1}^{p_i} \alpha_{i,k} \Delta y_{i,t-k}^\mu + e_{it}, \tag{19}$$

where

$$y_{it}^\mu = z_{it} - \frac{1}{t-1} \sum_{k=1}^{t-1} z_{ik}, \tag{20}$$

$$y_{i,t-1}^\mu = z_{i,t-1} - \frac{1}{t-1} \sum_{k=1}^{t-1} z_{ik}, \tag{21}$$

$$\Delta y_{i,t-k}^\mu = \Delta z_{i,t-k}, \quad k = 1, \dots, p_i \tag{22}$$

and (e_{it}) are regression errors.

⁴ They also consider the case $N/T \rightarrow k$, where k is a fixed constant, but the relevant asymptotics is not rigorously developed. Moreover, they introduce common time effects to their model, and thereby allow for limited cross-sectional dependency.

The term $(1/(t - 1)) \sum_{k=1}^{t-1} z_{ik}$ is the least-squares estimator of μ_i obtained from the preliminary regression

$$z_{ik} = \mu_i + y_{ik} \quad \text{for } k = 1, \dots, t - 1.$$

We note that the parameter μ_i is estimated from model (18) using the observations up to time $t - 1$. That we use the data up to the period $t - 1$ 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 up to $t - 1$ as given in (20) and (21), which we call *adaptive demeaning*.⁵ Note from (20) that even for the t th observation z_{it} , we use $(t - 1)$ -adaptive demeaning to maintain the martingale property. No further demeaning is needed for the lagged differences $\Delta z_{i,t-k}$, $i = 1, \dots, p_i$, since the differencing has already removed the mean.

We may then construct the nonlinear IV t -ratio statistic Z_i^μ based on the nonlinear IV estimator for α_i from regression (19), just as in (12). 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 statistic.

We may also test for a 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

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

where (y_{it}) is generated as in (1). Similarly, we may test for the unit root in (y_{it}) from regression (7) defined with the properly detrended series y_{it}^τ , $y_{i,t-1}^\tau$ and $\Delta y_{i,t-k}^\tau$ of z_{it} , $z_{i,t-1}$ and $\Delta z_{i,t-k}$, $k = 1, \dots, p_i$ as

$$y_{it}^\tau = \alpha_i y_{i,t-1}^\tau + \sum_{k=1}^{p_i} \alpha_{i,k} \Delta y_{i,t-k}^\tau + e_{it}, \tag{24}$$

where

$$y_{it}^\tau = z_{it} + \frac{2}{t-1} \sum_{k=1}^{t-1} z_{ik} - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k z_{ik} - \frac{1}{T_i} z_{iT_i}, \tag{25}$$

$$y_{i,t-1}^\tau = z_{i,t-1} + \frac{2}{t-1} \sum_{k=1}^{t-1} z_{ik} - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k z_{ik}, \tag{26}$$

$$\Delta y_{i,t-k}^\tau = \Delta z_{i,t-k} - \frac{1}{T_i} z_{iT_i}, \quad k = 1, \dots, p_i \tag{27}$$

and (e_{it}) are regression errors.

The variables z_{it} and $z_{i,t-1}$ are detrended using the least-squares estimators of the drift and trend coefficients, μ_i and δ_i , from model (23) using again the observations

⁵ This method was formerly used by So and Shin (1999) to demean positively correlated stationary AR processes. They found that the method reduces the biases of the parameter estimators.

up to time $t - 1$ only, viz.,

$$z_{ik} = \mu_i + \delta_i k + y_{ik} \quad \text{for } k = 1, \dots, t - 1.$$

The term z_{iT_i}/T_i appearing in the definitions of y_{it}^μ and $\Delta y_{i,t-k}^\tau$ given in (25) and (27) is the grand sample mean of $\Delta z_{it}, \frac{1}{T_i} \sum_{k=1}^{T_i} \Delta z_{ik}$. The term is used in (25) to eliminate the remaining drift term of $z_{it} + (2/(t - 1)) \sum_{k=1}^{t-1} z_{ik} - (6/t(t - 1)) \sum_{k=1}^{t-1} k z_{ik}$, and in (27) to remove the nonzero mean of $\Delta z_{i,t-k}$, for $k = 1, \dots, p_i$.

The *adaptive detrending* of the data as given in (25) and (26) above preserves the predictability of our instrument $F(y_{i,t-1}^\tau)$. The nonlinear IV t -ratio statistic Z_i^τ is then defined as in (12) from the nonlinear IV estimator for α_i from the regression (24).⁶

We may now derive the limit theories for the statistics Z_i^μ and Z_i^τ in the similar manner as we did to establish the limit theory given in Theorem 3.3. In order to define the limit distributions properly, we first introduce some notations. Define *adaptively demeaned* Brownian motion by

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

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

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

and analogously denote its local time by L_i^τ for $i = 1, \dots, N$. If we let $U_i^\mu(0) = 0$ and $U_i^\tau(0) = 0$, then both processes have well-defined continuous versions on $[0, \infty)$, as shown in the proof of Corollary 5.1.

The limit theories given in Lemma 3.2 extend easily to the models with nonzero means and deterministic trends if we replace the lagged level $y_{i,t-1}$ with the lagged detrended series $y_{i,t-1}^\mu$ and $y_{i,t-1}^\tau$ defined, respectively, in (21) and (26). They are indeed given similarly with the local times L_i^μ and L_i^τ of the adaptively demeaned and detrended Brownian motions U_i^μ and U_i^τ in the place of the local time L_i of the Brownian motion U_i . Then the limit theories for the nonlinear IV t -ratio statistics Z_i^μ and Z_i^τ for the models with nonzero means and deterministic trends follow immediately, and are given in

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

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

as $T_i \rightarrow \infty$ for all $i = 1, \dots, N$.

⁶ The adaptive demeaning and detrending, in particular, make our statistics independent of the starting values. Note that the nonlinear instruments $F(y_{i,t-1}^\mu)$ and $F(y_{i,t-1}^\tau)$ are now generated, respectively, from the adaptively demeaned and detrended series, $(y_{i,t-1}^\mu)$ and $(y_{i,t-1}^\tau)$, for $t = 2, \dots, T_i$, and that they are forced to start at the origin, i.e., $F(y_{i1}^\mu) = F(0)$ and $F(y_{i1}^\tau) = F(0)$ a.s.

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

6. Simulations

We conduct a set of simulations to investigate the finite sample performance of the average IV t -statistic S_N based on integrable IGFs for testing the unit root null hypothesis $H_0: \alpha_i = 1$ for all $i = 1, \dots, N$ against the stationarity alternative $H_1: |\alpha_i| < 1$ 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 $N(0, 1)$ distribution, and compare its sizes and powers with those of the commonly used average statistic t -bar proposed by Im et al. (1997).

For the simulations, we consider the time series (z_{it}) with a drift given by model (18) with (y_{it}) generated as in (1) and (u_{it}) as AR(1) processes, viz.,

$$u_{it} = \rho_i u_{i,t-1} + \varepsilon_{it}. \tag{28}$$

The innovations $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})'$ that generate $u_t = (u_{1t}, \dots, u_{Nt})'$ 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., $\rho_i \sim \text{Uniform}[0.2, 0.4]$. 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 $\text{Uniform}[0, 1]$.
- (2) Construct from M an orthogonal matrix $H = M(M'M)^{-1/2}$.
- (3) Generate a set of N eigenvalues, $\lambda_1, \dots, \lambda_N$. Let $\lambda_1 = r > 0$ and $\lambda_N = 1$ and draw $\lambda_2, \dots, \lambda_{N-1}$ from $\text{Uniform}[r, 1]$.
- (4) Form a diagonal matrix A with $(\lambda_1, \dots, \lambda_N)$ on the diagonal.
- (5) Construct the covariance matrix Σ as a spectral representation $\Sigma = HAH'$.

The covariance matrix constructed in this way will surely be symmetric and nonsingular with eigenvalues taking values from r to 1. We set the maximum eigenvalue 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 eigenvalue and the ratio of the minimum to the maximum eigenvalues 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 0.1.

For the estimation of the model (7) for $i = 1, \dots, N$, we consider the IV estimator $\hat{\gamma}_i$ defined in (10) which uses the instrument $(F(y_{i,t-1}), \Delta y_{i,t-1}, \dots, \Delta y_{i,t-p_i})'$. The

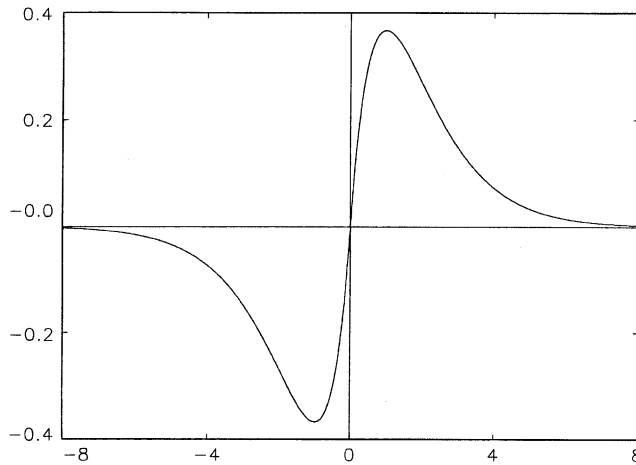


Fig. 1. IGF $F(x) = xe^{-|x|}$.

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 inversely proportional to the sample standard error of $\Delta y_{it} = u_{it}$ and $T_i^{1/2}$. That is,

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

where K is a constant. The value of K is fixed at 3 for all $i = 1, \dots, N$, and for all combinations of N and T considered here.⁷ 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 $i = 1, \dots, N$. The shape of the integrable IV generating function F is given in Fig. 1.

Our asymptotics requires the factor c_i to be constant. For practical applications, however, we found it desirable to make c_i dependent upon T_i as suggested in the previous equation. With the choice of larger (smaller) value of c_i , we may have better size (power) at the cost of power (size). This is well expected. Notice that the larger the value of c_i , the more integrable the IGF F becomes. Our asymptotics thus better predicts finite sample behavior of the test. On the other hand, the test loses finite sample powers as the factor c_i gets larger. As is well known from the standard regression

⁷The test S_N constructed from the IGF F with a larger value of K tends to have smaller rejection probabilities uniformly over all the choices of N and T . The IGF defined with $K = 3$ seems to work best overall, and thus chosen for our simulations. For the cases where the time dimension is small $T = 25$, the average nonlinear 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.

theory, the optimal IGF is given by the identity $F(x) = x$, which reduces our nonlinear IV estimator to the OLS estimator. As the IGF F tends to be more integrable, the resulting nonlinear IV estimator becomes less efficient, which may lead to the power loss in our test.

To test the unit root hypothesis, we set $\alpha_i = 1$ for all $i = 1, \dots, N$, and investigate the finite sample sizes in relation to the corresponding nominal test sizes. To examine the rejection probabilities under the alternative, 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 nonlinear IV t -ratio statistic S_N is compared with that of the t -bar statistic by Im et al. (1997), which is based on the average of the individual t -tests computed from the sample ADF regressions (7) with mean and variance modifications. More explicitly, the t -bar statistic is defined as

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

where t_i is the t -statistic for testing $\alpha_i = 1$ for the i th sample ADF regression (7), and $\bar{t}_N = N^{-1} \sum_{i=1}^N t_i$. The values of the expectation and variance, $\mathbf{E}(t_i)$ and $\text{var}(t_i)$, 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 $i = 1, \dots, N$ is required to be the same.⁸

The panels with the cross-sectional dimensions $N = 5, 15, 25, 50, 100$ and the time series dimensions $T = 25, 50, 100$ are considered for the 1%, 5% and 10% size tests.⁹ Since we are using randomly drawn parameter values, we simulate 20 times and report the ranges of the finite sample performances of the average nonlinear IV t -ratio statistic 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, the finite sample rejection probabilities and the size adjusted finite sample powers of the two tests. For each statistic, we report the minimum, mean, median and maximum of the rejection probabilities under the null and the alternative hypotheses.

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 provide reasonably good approximations 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 statistic exhibits noticeable size distortions, as reported, for

⁸ Table 2 in Im et al. (1997) tabulates the values of $\mathbf{E}(t_i)$ and $\text{var}(t_i)$ for $T = 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 100$ and for $p_i = 0, \dots, 8$.

⁹ For simplicity we use the same T for all cross-sectional units in our simulations. However, our theory does permit heterogeneity in the number T_i of time series observations. It is also true that the t -bar test can be practically implemented for unbalanced panels, though Im et al. (1997) do not explicitly allow for heterogeneous T_i 's in their theoretical developments.

Table 1
Finite sample sizes

N	T	Tests	1% Test				5% Test				10% Test			
			Min	Mean	Med	Max	Min	Mean	Med	Max	Min	Mean	Med	Max
5	25	<i>t</i> -bar	0.024	0.027	0.026	0.030	0.090	0.095	0.094	0.100	0.156	0.166	0.166	0.173
		S_N	0.018	0.021	0.021	0.024	0.065	0.071	0.071	0.075	0.117	0.124	0.124	0.130
15	25	<i>t</i> -bar	0.037	0.041	0.041	0.045	0.129	0.136	0.135	0.144	0.217	0.225	0.226	0.230
		S_N	0.017	0.020	0.020	0.022	0.063	0.071	0.071	0.076	0.114	0.124	0.124	0.131
25	25	<i>t</i> -bar	0.044	0.049	0.049	0.053	0.153	0.162	0.163	0.169	0.258	0.265	0.266	0.271
		S_N	0.015	0.017	0.018	0.020	0.057	0.066	0.067	0.071	0.108	0.118	0.118	0.125
50	25	<i>t</i> -bar	0.077	0.081	0.082	0.086	0.225	0.232	0.232	0.238	0.346	0.353	0.353	0.360
		S_N	0.015	0.017	0.017	0.019	0.058	0.063	0.063	0.068	0.106	0.115	0.115	0.121
100	25	<i>t</i> -bar	0.146	0.154	0.154	0.159	0.349	0.358	0.358	0.364	0.486	0.494	0.494	0.505
		S_N	0.015	0.017	0.018	0.019	0.058	0.063	0.064	0.066	0.109	0.114	0.115	0.118
5	50	<i>t</i> -bar	0.016	0.019	0.018	0.021	0.071	0.075	0.076	0.077	0.132	0.138	0.138	0.142
		S_N	0.016	0.018	0.018	0.021	0.059	0.065	0.066	0.072	0.109	0.116	0.116	0.123
15	50	<i>t</i> -bar	0.020	0.024	0.024	0.029	0.086	0.092	0.091	0.100	0.156	0.163	0.163	0.172
		S_N	0.014	0.016	0.017	0.017	0.059	0.062	0.062	0.065	0.104	0.111	0.111	0.120
25	50	<i>t</i> -bar	0.024	0.026	0.026	0.028	0.095	0.101	0.101	0.103	0.168	0.177	0.177	0.181
		S_N	0.010	0.014	0.014	0.016	0.050	0.055	0.056	0.060	0.096	0.102	0.101	0.109
50	50	<i>t</i> -bar	0.030	0.034	0.034	0.039	0.115	0.123	0.122	0.132	0.197	0.209	0.209	0.218
		S_N	0.010	0.012	0.012	0.013	0.045	0.050	0.050	0.053	0.087	0.093	0.094	0.098
100	50	<i>t</i> -bar	0.049	0.053	0.053	0.059	0.162	0.167	0.167	0.174	0.260	0.267	0.267	0.277
		S_N	0.009	0.011	0.011	0.014	0.040	0.045	0.045	0.048	0.079	0.085	0.085	0.089
5	100	<i>t</i> -bar	0.014	0.016	0.016	0.018	0.062	0.066	0.067	0.071	0.120	0.124	0.125	0.129
		S_N	0.014	0.017	0.017	0.019	0.060	0.063	0.064	0.068	0.109	0.115	0.115	0.120
15	100	<i>t</i> -bar	0.016	0.018	0.019	0.020	0.072	0.076	0.076	0.080	0.134	0.139	0.140	0.144
		S_N	0.013	0.015	0.015	0.017	0.052	0.059	0.059	0.063	0.100	0.107	0.107	0.112
25	100	<i>t</i> -bar	0.017	0.018	0.018	0.022	0.070	0.076	0.076	0.080	0.134	0.141	0.142	0.145
		S_N	0.011	0.013	0.013	0.015	0.049	0.052	0.052	0.056	0.091	0.098	0.098	0.102
50	100	<i>t</i> -bar	0.018	0.021	0.021	0.023	0.077	0.084	0.084	0.092	0.146	0.153	0.154	0.159
		S_N	0.010	0.011	0.011	0.013	0.042	0.046	0.047	0.050	0.083	0.087	0.087	0.092
100	100	<i>t</i> -bar	0.025	0.029	0.029	0.035	0.099	0.103	0.103	0.107	0.173	0.178	0.178	0.183
		S_N	0.007	0.010	0.009	0.012	0.038	0.041	0.040	0.045	0.073	0.078	0.078	0.084

instance, in the previous simulation work by Maddala and Wu (1999). The direction of the size distortions are upward in all cases for all 1%, 5% and 10% tests. The *t*-bar statistic suffers from severe size distortions especially when the number of cross-sectional units is large relative to the number of time series observations. For example, when $N = 100$ and $T = 25$, the average finite sample sizes of the *t*-bar statistics for 1%, 5% and 10% tests are, respectively, 15%, 36% and 49%. The size distortions become less serious as the time dimension gets large; however, they are still quite noticeable.

The test S_N is more powerful than the *t*-bar statistic 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,

Table 2
Finite sample rejection probabilities

N	T	Tests	1% Test				5% Test				10% Test			
			Min	Mean	Med	Max	Min	Mean	Med	Max	Min	Mean	Med	Max
5	25	<i>t</i> -bar	0.059	0.094	0.088	0.139	0.191	0.258	0.248	0.346	0.306	0.390	0.377	0.498
		S_N	0.073	0.132	0.119	0.225	0.207	0.316	0.297	0.465	0.319	0.445	0.428	0.607
15	25	<i>t</i> -bar	0.184	0.277	0.278	0.368	0.417	0.540	0.546	0.652	0.568	0.684	0.691	0.778
		S_N	0.223	0.364	0.364	0.526	0.449	0.617	0.625	0.767	0.590	0.741	0.750	0.862
25	25	<i>t</i> -bar	0.346	0.472	0.464	0.628	0.627	0.736	0.732	0.854	0.761	0.844	0.841	0.926
		S_N	0.397	0.577	0.575	0.780	0.655	0.804	0.808	0.928	0.778	0.887	0.892	0.968
50	25	<i>t</i> -bar	0.725	0.831	0.847	0.908	0.901	0.951	0.962	0.980	0.951	0.979	0.983	0.993
		S_N	0.780	0.901	0.917	0.969	0.930	0.975	0.980	0.995	0.965	0.990	0.992	0.998
100	25	<i>t</i> -bar	0.972	0.988	0.989	0.994	0.996	0.998	0.999	0.999	0.999	1.000	1.000	1.000
		S_N	0.987	0.996	0.997	0.999	0.998	1.000	1.000	1.000	0.999	1.000	1.000	1.000
5	50	<i>t</i> -bar	0.081	0.194	0.163	0.372	0.246	0.438	0.404	0.672	0.385	0.591	0.563	0.809
		S_N	0.189	0.382	0.337	0.651	0.417	0.642	0.614	0.877	0.566	0.765	0.752	0.943
15	50	<i>t</i> -bar	0.341	0.601	0.611	0.817	0.626	0.834	0.849	0.953	0.760	0.911	0.925	0.980
		S_N	0.648	0.861	0.879	0.979	0.867	0.959	0.971	0.998	0.929	0.982	0.990	1.000
25	50	<i>t</i> -bar	0.675	0.854	0.855	0.979	0.887	0.962	0.967	0.997	0.945	0.985	0.988	0.999
		S_N	0.912	0.981	0.988	1.000	0.984	0.997	0.999	1.000	0.994	0.999	1.000	1.000
50	50	<i>t</i> -bar	0.979	0.996	0.998	1.000	0.998	1.000	1.000	1.000	0.999	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	<i>t</i> -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	<i>t</i> -bar	0.215	0.609	0.558	0.955	0.498	0.828	0.833	0.996	0.668	0.906	0.921	0.999
		S_N	0.530	0.830	0.843	0.998	0.791	0.946	0.964	1.000	0.883	0.975	0.987	1.000
15	100	<i>t</i> -bar	0.870	0.981	0.995	1.000	0.975	0.997	1.000	1.000	0.991	0.999	1.000	1.000
		S_N	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25	100	<i>t</i> -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
50	100	<i>t</i> -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	<i>t</i> -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

reported, respectively, in Tables 2 and 3. The discriminatory power of S_N is noticeably much higher than that of the *t*-bar statistic 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 statistic. 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 statistic 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* increase.

Table 3
Finite sample powers

N	T	Tests	1% Test				5% Test				10% Test			
			Min	Mean	Med	Max	Min	Mean	Med	Max	Min	Mean	Med	Max
5	25	<i>t</i> -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	0.077	0.069	0.143	0.157	0.250	0.236	0.393	0.271	0.392	0.379	0.554
15	25	<i>t</i> -bar	0.062	0.110	0.108	0.165	0.208	0.310	0.305	0.398	0.340	0.462	0.456	0.567
		S_N	0.150	0.257	0.245	0.390	0.367	0.540	0.543	0.692	0.529	0.693	0.695	0.827
25	25	<i>t</i> -bar	0.138	0.208	0.196	0.336	0.342	0.473	0.462	0.627	0.486	0.625	0.619	0.759
		S_N	0.344	0.484	0.464	0.714	0.607	0.761	0.760	0.901	0.748	0.866	0.867	0.957
50	25	<i>t</i> -bar	0.354	0.482	0.501	0.584	0.627	0.753	0.772	0.854	0.761	0.860	0.869	0.927
		S_N	0.707	0.858	0.879	0.944	0.911	0.966	0.973	0.993	0.958	0.987	0.990	0.998
100	25	<i>t</i> -bar	0.713	0.806	0.806	0.868	0.895	0.944	0.946	0.966	0.950	0.976	0.978	0.987
		S_N	0.980	0.993	0.994	0.998	0.997	1.000	1.000	1.000	0.999	1.000	1.000	1.000
5	50	<i>t</i> -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	0.128	0.293	0.251	0.537	0.353	0.586	0.548	0.845	0.526	0.733	0.713	0.926
15	50	<i>t</i> -bar	0.192	0.451	0.483	0.687	0.480	0.734	0.752	0.898	0.640	0.847	0.859	0.959
		S_N	0.569	0.813	0.826	0.963	0.842	0.949	0.958	0.996	0.922	0.979	0.986	1.000
25	50	<i>t</i> -bar	0.496	0.735	0.732	0.943	0.783	0.915	0.921	0.991	0.884	0.962	0.967	0.997
		S_N	0.891	0.974	0.984	1.000	0.978	0.997	0.999	1.000	0.994	0.999	1.000	1.000
50	50	<i>t</i> -bar	0.919	0.980	0.990	0.999	0.987	0.998	0.999	1.000	0.997	0.999	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	<i>t</i> -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	<i>t</i> -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	0.432	0.774	0.771	0.993	0.753	0.929	0.950	1.000	0.864	0.970	0.984	1.000
15	100	<i>t</i> -bar	0.777	0.966	0.988	1.000	0.950	0.994	1.000	1.000	0.982	0.998	1.000	1.000
		S_N	0.986	0.998	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
25	100	<i>t</i> -bar	0.992	0.999	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	<i>t</i> -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	<i>t</i> -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

7. Empirical illustrations

In this section, we apply the newly developed panel unit root test S_N to test whether the PPP hypothesis holds. The PPP hypothesis has been tested by many researchers using various unit root tests, both in panel as well as in 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 evidence, 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 20 countries computed from the IMF’s International Financial Statistics (IFS)

Table 4
PPP tests for IFS data

		$t\text{-bar}$		S_N	
		$T = 50$	$T = 100$	$T = 50$	$T = 100$
AR	2	-1.589 ^b	-1.183	-2.872 ^a	-2.554 ^a
Order	4	-7.108 ^a	-4.525 ^a	-5.969 ^a	-5.119 ^a
BIC	4	-2.740 ^a	-3.490 ^a	-3.719 ^a	-4.695 ^a
Max order	8	-1.127	-2.646 ^a	-0.249	-3.937 ^a

Note: The superscripts a, b and c denote, respectively, the statistical significance at 1%, 5% and 10% levels.

tape, covering the period 1973:1–1998:4.¹⁰ We also consider the data from the Penn World Table (PWT) analyzed in Oh (1996).¹¹ The empirical results are summarized in Tables 4 and 5, respectively, for the results obtained from the data from Papell (1997) and Oh (1996). We allow the models to have heterogeneous dynamic structures, 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 the BIC criterion with the maximum number of lags 4 or 8 for the quarterly IFS data, and with 2 or 4 for the annual PWT data. 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 for all cross-sectional units at 2 or 4 for the IFS data and at 1 or 2 for the PWT data.

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 41.¹² The IFS data have 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 two sub-samples of sizes 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 rejects the unit root hypothesis, which is used in the empirical studies as an indirect evidence for the PPP relationship. As seen from Tables 4 and 5, our test rejects the presence of the

¹⁰ The quarterly data used in Papell (1997) covers the period 1973:1–1994:3, but the data used here is extended to 1998:4. 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.

¹¹ The data used in Oh (1996) are yearly observations from the Penn World Table, Mark 5.5. The data are 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 are analyzed for two sub-samples, the 22 OECD countries and G6 countries (Canada, France, Germany, Italy, Japan and United Kingdom).

¹² For the group of 111 countries, there are 30 annual observations. But for the group of 51 countries (including its sub-samples of 22 OECD countries and G6 countries), there are 41 time series observations.

Table 5
PPP tests for PWT data

		$t\text{-bar}$				S_N			
		G6	OECD	51Con	111Con	G6	OECD	51Con	111Con
AR	1	-1.014	-1.997 ^b	-2.694 ^a	-6.111 ^a	-1.471 ^c	-3.217 ^a	-4.116 ^a	-8.066 ^a
Order	2	-0.400	-0.669	-0.912	-3.420 ^a	-0.629	-1.528 ^c	-1.975 ^b	-5.033 ^a
BIC	2	-1.014	-1.997 ^b	-2.066 ^b	-4.279 ^a	-1.471 ^c	-3.217 ^a	-3.841 ^a	-6.708 ^a
Max order	4	-1.014	-1.723 ^b	-2.062 ^b	-3.899 ^a	-1.471 ^c	-3.103 ^a	-3.740 ^a	-6.002 ^a

Note: The superscripts a, b and c denote, respectively, the statistical significance at 1%, 5% and 10% levels.

unit root in most of the cases considered here.¹³ The values of the test statistic S_N of course vary for different choices of the sample size T and the specifications of the dynamic structures, but overall they provide strong evidence against 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 $t\text{-bar}$ test by Im et al. (1997) produces the results that are inconclusive. The test results are, in particular, quite 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 PWT data are also inconclusive. The $t\text{-bar}$ test supports or rejects the PPP depending upon how we select the countries and the time series observations. The test fails to consistently 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 statistic for the test of the unit root constructed from such nonlinear

¹³ Our test is not able to reject the absence of PPP for 20 OECD countries from the IFS data when short-run dynamics is selected by the BIC with the larger maximum number of lags 8 and the smaller time dimension $T = 50$. This might be due to the fact that the 50 quarterly time series observations (possibly with upto 8 losses in data points for constructing lagged differences) amount to only about 12 years of time span, which is too short for uncovering long-run properties of the underlying stochastic processes. Our test also fails to reject the null for G6 countries based on the PWT data, when the dynamics is restricted at AR(2). This indeed is the case with the smallest number of total observations, which may well have led to the low power.

IV estimator is shown to have standard normal limit distribution, for each individual cross-sectional unit $i = 1, \dots, N$. The nonlinear IV t -ratio statistic 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 statistic, 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 et al. (1997). The panel nonlinear IV unit root test seems to improve significantly upon the t -bar test under cross-sectional dependency, especially for the panels with smaller time and spatial dimensions. The new statistic S_N is applied to test whether the PPP 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 supports the PPP relationship, while the t -bar test by Im et al. (1997) produces inconclusive results.

Acknowledgements

I thank the Co-Editors and two anonymous referees for many constructive comments and suggestions. 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. My thanks also go to David Papell and Keun-Yeob Oh for kindly providing the data sets used here for empirical illustrations, and to Paul Evans, Joon Park and Peter Phillips for helpful discussions and comments.

Appendix: Mathematical proofs

Proof of Lemma 3.2. We have from the Beveridge–Nelson representation of y_{it} given in (6) 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 $T_i \rightarrow \infty$

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

since $U_{iT_i} \rightarrow_d U_i$ as $T_i \rightarrow \infty$, due to the invariance principle in (4). Then it follows from Lemma 5 (c) of Chang et al. (2001) that

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

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 5 (i) of Chang et al. (2001) as follows:

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

again by a simple change of variables.

For part (c), just note that $\Delta y_{i,t-1}, \dots, \Delta y_{i,t-p_i}$ 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 5 in Chang et al. (2001). \square

Proof of Theorem 3.3. We begin by investigating the limit behavior of A_{T_i} and C_{T_i} defined below (11). Recall $x'_{it} = (\Delta y_{i,t-1}, \dots, \Delta y_{i,t-p_i})$. 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

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

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 (12) 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 (11) and (13). Then it follows immediately from Lemma 3.2 (a) and (b) that

$$\begin{aligned} Z_i &= \frac{T_i^{-1/4}A_{T_i}}{\hat{\sigma}_i(T_i^{-1/2}C_{T_i})^{1/2}} \\ &= \frac{T_i^{-1/4} \sum_{t=1}^{T_i} F(y_{i,t-1})\varepsilon_{it}}{\hat{\sigma}_i(T_i^{-1/2} \sum_{t=1}^{T_i} F(y_{i,t-1})^2)^{1/2}} + o_p(1) \end{aligned}$$

$$\begin{aligned} &\xrightarrow{d} \frac{\sigma_i(\alpha^i(1)L_i(1,0) \int_{-\infty}^{\infty} F(s)^2 ds)^{1/2} W_i(1)}{\sigma_i(\alpha^i(1)L_i(1,0) \int_{-\infty}^{\infty} F(s)^2 ds)^{1/2}} \\ &\equiv W_i(1) \end{aligned}$$

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

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

$$\begin{aligned} &\sqrt[4]{T_i T_j} \int_0^1 F(\sqrt{T_i} B_{iT_i}(r)) F(\sqrt{T_j} B_{jT_j}(r)) dr \\ &\leq \sqrt[4]{T_i T_j} \int_0^1 F(\sqrt{T_i} B_{iT_i}(r)) F(\sqrt{T_i} B_{jT_j}(r)) dr. \end{aligned} \tag{A.2}$$

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

$$\int_0^1 F(\sqrt{T_i} B_i(r)) F(\sqrt{T_i} B_j(r)) dr = O_p\left(\frac{\log T_i}{T_i}\right) \text{ a.s.} \tag{A.3}$$

and as shown in Chang et al. (2001)

$$\begin{aligned} &\frac{T_i}{\log T_i} \int_0^1 F(\sqrt{T_i} B_{iT_i}(r)) F(\sqrt{T_i} B_{jT_j}(r)) dr \\ &= \frac{T_i}{\log T_i} \int_0^1 F(\sqrt{T_i} B_i(r)) F(\sqrt{T_i} B_j(r)) dr + o_p(1). \end{aligned} \tag{A.4}$$

We now have from (A.2)–(A.4) that

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

which is of order $o_p(1)$ if

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

as is assumed in Assumption 4.1. This establishes the stated result in (17).

As shown in Chang et al. (2001), the result in (17) implies that

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

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

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. \square

Proof of Corollary 5.1. From $z_{it} = \mu_i + y_{it}$ and (21), it follows that

$$\begin{aligned} y_{i,t-1}^\mu &= z_{i,t-1} - \frac{1}{t-1} \sum_{k=1}^{t-1} z_{ik} \\ &= (\mu_i + y_{i,t-1}) - \frac{1}{t-1} \sum_{k=1}^{t-1} (\mu_i + y_{ik}) \\ &= y_{i,t-1} - \frac{1}{t-1} \sum_{k=1}^{t-1} y_{ik}. \end{aligned}$$

We have $y_{it} = y_{i,t-1} + u_{it}$ under the unit root null, and this implies $z_{it} = z_{i,t-1} + u_{it}$. Then we have from (20)

$$\begin{aligned} y_{it}^\mu &= z_{it} - \frac{1}{t-1} \sum_{k=1}^{t-1} z_{ik} \\ &= \left(z_{i,t-1} - \frac{1}{t-1} \sum_{k=1}^{t-1} z_{ik} \right) + u_{it} \\ &= y_{i,t-1}^\mu + \sum_{k=1}^{p_i} \alpha_{i,k} u_{i,t-k} + \varepsilon_{it} \end{aligned}$$

since $u_{it} = \sum_{k=1}^{p_i} \alpha_{i,k} u_{i,t-k} + \varepsilon_{it}$ as specified in (2). Now regression (19), which is used for testing the unit root in the stochastic component y_{it} of the series z_{it} with nonzero mean, follows directly from the previous equation, since $u_{i,t-k} = \Delta y_{i,t-k}$ under the null

and

$$\Delta y_{i,t-k} = \Delta z_{i,t-k} = \Delta y_{i,t-k}^\mu$$

for $k = 1, \dots, p_i$, using the notation defined in (22). Note that $e_{it} = \varepsilon_{it}$ under the null.

To examine the limit behavior of $y_{i,t-1}^\mu$, we may write it as

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

for $r \in [0, 1]$ such that $[T_i r] = t - 1$. Then it follows from (A.1) and the invariance principle given in (4) that

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

as $T_i \rightarrow \infty$, where

$$U_i^\mu = U_i(r) - \frac{1}{r} \int_0^r U_i(s) ds. \tag{A.5}$$

We now turn to detrending the time series $z_{it} = \mu_i + \delta_i t + y_{it}$ with linear time trend. We begin by writing the detrended series $y_{i,t-1}^\tau$ of the lagged level $z_{i,t-1}$ defined in (26) more explicitly as

$$\begin{aligned} y_{i,t-1}^\tau &= z_{i,t-1} + \frac{2}{t-1} \sum_{k=1}^{t-1} z_{ik} - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k z_{ik} \\ &= y_{i,t-1} + \frac{2}{t-1} \sum_{k=1}^{t-1} y_{ik} - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k y_{ik}. \end{aligned} \tag{A.6}$$

Note that $z_{it} = \delta_i + z_{i,t-1} + u_{it}$ since $y_{it} = y_{i,t-1} + u_{it}$ under the null, and that $z_{iT_i} = \sum_{k=1}^{T_i} \Delta z_{ik}$. Then we use these to rewrite the detrended series y_{it}^τ of z_{it} defined in (25) as

$$\begin{aligned} y_{it}^\tau &= z_{it} - \frac{1}{T_i} z_{iT_i} + \frac{2}{t-1} \sum_{k=1}^{t-1} z_{ik} - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k z_{ik} \\ &= z_{i,t-1} + \frac{2}{t-1} \sum_{k=1}^{t-1} z_{ik} - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k z_{ik} + u_{it} - \left(\frac{1}{T_i} \sum_{t=1}^{T_i} \Delta z_{it} - \delta_i \right) \\ &= y_{i,t-1} + \frac{2}{t-1} \sum_{k=1}^{t-1} y_{ik} - \frac{6}{t(t-1)} \sum_{k=1}^{t-1} k y_{ik} + u_{it} - \frac{1}{T_i} \sum_{t=1}^{T_i} u_{it} \end{aligned} \tag{A.7}$$

as $\Delta y_{it} = u_{it}$ under the null. Therefore, the suggested detrending indeed effectively remove all the deterministic components in both z_{it} and $z_{i,t-1}$.

Next we note from the definitions of the detrended lagged differences $\Delta y_{i,t-k}^\tau$, $k = 1, \dots, p_i$, given in (27) that

$$\Delta y_{i,t-k}^\tau = \Delta z_{i,t-k} - \frac{1}{T_i} z_{iT_i} = \Delta z_{i,t-k} - \frac{1}{T_i} \sum_{s=1}^{T_i} \Delta z_{is} = u_{i,t-k} - \frac{1}{T_i} \sum_{s=1}^{T_i} u_{is}$$

since $\Delta z_{ik} = \delta_i + u_{ik}$. Then it follows that

$$\begin{aligned} u_{it} &= \sum_{k=1}^{p_i} \alpha_{i,k} u_{i,t-k} + \varepsilon_{it} \\ &= \sum_{k=1}^{p_i} \alpha_{i,k} \left(\Delta y_{i,t-k}^\tau + \frac{1}{T_i} \sum_{t=1}^{T_i} u_{it} \right) + \varepsilon_{it} \\ &= \sum_{k=1}^{p_i} \alpha_{i,k} \Delta y_{i,t-k}^\tau + \varepsilon_{it} + \sum_{k=1}^{p_i} \alpha_{i,k} \left(\frac{1}{T_i} \sum_{t=1}^{T_i} u_{it} \right). \end{aligned} \tag{A.8}$$

Then we may deduce from (A.6)–(A.8) that

$$y_{it}^\tau = y_{i,t-1}^\tau + \sum_{k=1}^{p_i} \alpha_{i,k} \Delta y_{i,t-k}^\tau + \varepsilon_{it} + \left(\sum_{k=1}^{p_i} \alpha_{i,k} - 1 \right) \left(\frac{1}{T_i} \sum_{t=1}^{T_i} u_{it} \right)$$

under the null, and this justifies regression (24) used in testing for the unit root in the stochastic component y_{it} in the time series z_{it} with linear time trend defined in (23), if we let

$$e_{it} = \varepsilon_{it} + \left(\sum_{k=1}^{p_i} \alpha_{i,k} - 1 \right) \left(\frac{1}{T_i} \sum_{t=1}^{T_i} u_{it} \right).$$

To develop asymptotics for the adaptively detrended series, we consider

$$y_{i[T_i r]}^\tau = y_{i[T_i r]} + \frac{2}{[T_i r]} \sum_{k=1}^{[T_i r]} y_{ik} - \frac{6}{[T_i r]([T_i r] + 1)} \sum_{k=1}^{[T_i r]} k y_{ik}$$

for $r \in [0, 1]$ such that $[T_i r] = t - 1$, and this gives

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

as $T_i \rightarrow \infty$ due to (A.1) and (4), where

$$U_i^\tau(r) = U_i(r) + \frac{2}{r} \int_0^r U_i(s) ds - \frac{6}{r^2} \int_0^r sU_i(s) ds. \tag{A.9}$$

The processes U_i^μ and U_i^τ introduced in (A.5) and (A.9) 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

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

Hence, 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 $[0, \infty)$.

Now the limit theories given in Lemma 3.2 for the sample moments from the models with no deterministic trend easily extend to the sample moments from the models with deterministic trends. The limit theories in fact have the identical expressions with the local times of the adaptively demeaned and detrended Brownian motions U_i^μ and U_i^τ in the place of the original Brownian motion 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 means and deterministic trends. The proof for the models with nonzero mean is trivial. For the models with linear trend, define

$$v_{iT_i} = \frac{1}{T_i} \sum_{t=1}^{T_i} u_{it}.$$

For the stated results in this case, it suffices to show that

$$R_{iT_i} = \frac{1}{\sqrt[4]{T_i}} \sum_{t=1}^{T_i} F(y_{i,t-1}^\tau) v_{iT_i}$$

is negligible in the limit, and dominated by the leading term

$$\frac{1}{\sqrt[4]{T_i}} \sum_{t=1}^{T_i} F(y_{i,t-1}^\tau) \varepsilon_{it}$$

which follows immediately from

$$R_{iT_i} = \frac{1}{\sqrt[4]{T_i}} (\sqrt{T_i} v_{iT_i}) \left(\frac{1}{\sqrt{T_i}} \sum_{t=1}^{T_i} F(y_{i,t-1}^\tau) \right) = O_p(T_i^{-1/4}). \quad \square$$

References

- Banerjee, A., 1999. Panel data unit roots and cointegration: an overview *Oxford Bulletin of Economics & Statistics* 61, 607–629.
- Brockwell, P.J., Davis, R.A., 1991. *Time series: Theory and Methods*. Springer-Verlag, New York.
- Chang, Y., 2000. Bootstrap unit root tests in panels with cross-sectional dependency. Mimeographed. Rice University.
- Chang, Y., Park, J.Y., 1999. Nonstationary index models. Mimeographed. Department of Economics, Rice University.
- Chang, Y., Park, J.Y., Phillips, P.C.B., 2001. Nonlinear econometric models with cointegrated and deterministically trending regressors. *Econometrics Journal* 4, 1–36.
- Choi, I., 2001. Unit root tests for panel data. *Journal of International Money and Finance* 20, 219–247.
- Frankel, J., Rose, A., 1996. A panel project on purchasing power parity: mean reversion within and between countries *Journal of International Economics* 40, 209–224.
- Fuller, W.A., 1996. *Introduction to Statistical Time Series*, 2nd ed. Wiley, New York.
- Im, K.S., Pesaran, M.H., Shin, Y., 1997. Testing for unit roots in heterogeneous panels. Mimeographed.
- Kasahara, Y., Kotani, S., 1979. On limit processes for a class of additive functionals of recurrent diffusion processes. *Zeitschrift fuer Wahrscheinlichkeitstheorie verw Gebiete* 49, 133–153.
- Levin, A., Lin, C.F., Chu, C.S., 1997. Unit root tests in panel data: asymptotic and finite sample properties. Mimeographed. University of California, San Diego.
- MacDonald, R., 1996. Panel unit root tests and real exchange rates. *Economics Letters* 50, 7–11.
- Maddala, G.S., Wu, S., 1999. A comparative study of unit root tests with panel data and a new simple test: evidence from simulations and bootstrap *Oxford Bulletin of Economics and Statistics* 61, 631.
- O’Connell, P.G.J., 1998. The overvaluation 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.H., 1997. Searching for stationarity: purchasing power parity under the current float *Journal of International Economics* 43, 313–323.
- Park, J.Y., Phillips, P.C.B., 1999. Asymptotics for nonlinear transformations of integrated time series. *Econometric Theory* 15, 269–298.
- Park, J.Y., Phillips, P.C.B., 2001. Nonlinear regressions with integrated time series. *Econometrica* 69, 117–162.
- Phillips, P.C.B., 1987. Time series regression with a unit root. *Econometrica* 55, 277–301.
- Phillips, P.C.B., Perron, P., 1988. Testing for a unit root in time series regression. *Biometrika* 75, 335–346.
- Phillips, P.C.B., Solo, V., 1992. Asymptotics for linear processes. *Annals of Statistics* 20, 971–1001.
- Phillips, P.C.B., Park, J.Y., Chang, Y., 1999. Nonlinear instrumental variable estimation of an autoregression. Mimeographed. Department of Economics, Rice University.
- Revuz, D., Yor, M., 1994. *Continuous Martingale and Brownian Motion*. Springer, New York.
- So, B.S., Shin, D.W., 1999. Recursive mean adjustment in time-series inferences. *Statistics & Probability Letters* 43, 65–73.