Nonlinear IV Panel Unit Root Tests

Yoosoon Chang[1]
Department of Economics
Rice University

Abstract

This paper presents the nonlinear IV methodology as an effective inferential basis for nonstationary panels. The nonlinear IV method resolves the inferential difficulties in testing for unit roots arising from the intrinsic heterogeneities and cross-dependencies of panel models. Individual units are allowed to be dependent through correlations among innovations, interrelatedness of short-run dynamics and/or cross-sectional cointegrations. If based on the instrumental variables that are nonlinear transformations of the lagged levels, the usual IV estimation of the augmented Dickey-Fuller type regressions yields asymptotically normal unit root tests for panels with general dependencies and heterogeneities. Moreover, the nonlinear IV estimation allows for the use of covariates to further increase power, and order statistics to test for more flexible forms of hypotheses, which are especially important in heterogeneous panels.

This version: March 7, 2003

JEL Classification: C12, C15, C32, C33.

Key words and phrases: Dependent panels, heterogeneity, cross-sectional cointegration, covariates, nonlinear IV t-ratios, order statistics, panel unit root tests.

[1]I am grateful to Joon Y. Park for helpful discussions and comments. This research was supported by the National Science Foundation under Grant SES-0233940. 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

Nonstationary panels have recently drawn much attention in both theoretical and empirical research. This is largely due to the availability of panel data sets covering relatively long time periods and the growing use of cross-country and cross-region data over time to investigate many important economic inter-relationships, especially those involving convergence/divergene of economic variables. Various statistics for testing unit roots in panel models have been proposed, and frequently used for testing growth theories and purchasing power parity hypothesis. They are also used to study the long-run relationships between migration flows and income and unemployment differentials across regions, and among macroeconomic and international financial series including exchange rates and spot and future interest rates. Panel data based tests also appeared attractive to many empirical researchers, since they offer alternatives to the tests based only on individual time series observations that are known to have low discriminatory power. The earlier contributors in theoretical research on the subject include Levin, Lin and Chu (2002), Quah (1994), Im, Pesaran and Shin (1997) and Maddala and Wu (1999). Phillips and Moon (2000) and Baltagi and Kao (2000) provide surveys on the recent developments in testing for unit roots in panels.

In general panels are intrinsically heterogeneous and dependent across cross-sections, and these very properties bring the inferential difficulties in testing for unit roots. The cross-sectional dependency in particular is very hard to deal with in nonstationary panels. For instance, 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 representing correlations across individual units. As a result, most of the existing panel unit root tests have assumed various homogeneities and spatial independence in order to obtain tractable distribution theories. These assumptions are, however, highly restrictive and thus unrealistic for many economic panels of practical interests. Of course the limit distributions of the tests constructed under such assumptions are no longer valid and become unknown when any of the assumptions are violated. Indeed, Maddala and Wu (1999) show through simulations that the tests based on cross-sectional independence, such as those considered in Levin, Lin and Chu (2002) and Im, Pesaran and Shin (1997), yield biased results and have substantial size distortions when applied to cross-sectionally dependent panels.

This paper presents the nonlinear IV methodology which was explored in Chang (2002) and Chang and Song (2002) to overcome the inferential difficulties in panel unit root testing, which arise from the intrinsic heterogeneities and dependencies of general nonstationary panels. Phillips, Park and Chang (1999) and Chang (2002) show that some nonlinear transformations of integrated processes have important statistical properties that can be exploited to construct unit root tests which have Gaussian limit distributions. This is a striking result given that the limit theories for unit root models are generally non-Gaussian, rendering all standard testing procedures based on Gaussian limit theory invalid. Chang (2002) finds an effective inferential basis for the unit root testing in dependent panels in the marriage of this important statistical property and the asymptotic orthogonalities among the nonlinear transfromations of integrated processes. If the transformations of the lagged

levels by an integrable function are used as instruments, the t-ratio based on the usual IV estimator for the autoregressive coefficient in the augmented Dickey-Fuller type regression yields asymptotically normal unit root tests for each cross-section, and more importantly the nonlinear IV t-ratios from different cross-sections are asymptotically independent even when the cross-sections are dependent. The asymptotic orthogonalities follow from the fact that the nonlinear transformations of dependent integrated processes are asymptotically independent so long as they are not cointegrated. This means that the panel unit root tests constructed as a simple standardized sum of the individual nonlinear IV t-ratios have a standard normal limiting distribution.

The asymptotic independence and normality of the nonlinear IV t-ratios essentially provide a set of N asymptotically iid normal random variables, which we may use as a basis for constructing statistics capable of testing various unit root hypotheses in panels. Hence, the hypotheses from economic theories and problems can be tested under more flexible forms of null and alternative hypotheses, which are especially important for heterogeneous panels, using the order statistics based on the nonlinear IV t-ratios whose limit distributions are given by simple functionals of the standard normal distribution functions. Moreover, the nonlinear IV estimation allows for the use of relevant covariates to further increase power, without incurring nuisance parameter problems. The nonlinear IV approach to panel unit root testing is explored further in Chang and Song (2002) to deal with cointegration among cross-sections, which was excluded in the study of Chang (2002). The presence of cointegration invalidates the tests by Chang (2002) constructed from the nonlinear IV t-ratios based the instruments generated by a single integrable function for all cross-sections, because the nonlinear IV t-ratios are no longer asymptotically independent in the presence of cross-sectional cointegrations. Chang and Song (2002) show that we may still obtain asymptotically normal panel unit root tests if the instruments for each of the cross-sections are constructed from orthogonal functions even when the cross-sections are cointegrated.

The rest of the paper is organized as follows. Section 2 introduces the nonlinear IV methodology in a simple univariate setting. The nonlinear IV t-ratio statistic is defined, and the adaptive demeaning and detrending schemes needed to deal with models with deterministic trends are presented. Section 3 discusses various issues in panel unit root testing, and introduces the nonlinear IV panel unit root test for panels with cross-sectional dependency induced by cross-correlations in innovations. Section 4 presents an improved nonlinear IV method to construct valid tests for cointegrated panels, and proposes to use nonlinear IV order statistics to test for more flexible forms of unit root hypotheses in panels. Section 5 concludes.

2. Nonlinear IV Methodology

2.1 Model and Preliminaries

We consider the following unit root regression

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

for $t = 1, \ldots, T$. We are interested in testing the null of a unit root, $\alpha = 1$, against the stationarity alternative, $|\alpha| < 1$. The initial value y_0 does not affect our subsequent asymptotic analysis as long as they are stochastically bounded, and therefore we set it at zero for expositional brevity. The errors (u_t) in the model (1) are serially correlated and specified as an AR(p) process given by

$$\alpha(L)u_t = \varepsilon_t \tag{2}$$

where L is the usual lag operator and $\alpha(z) = 1 - \sum_{k=1}^{p} \alpha_k z^k$. We assume

Assumption 2.1 $\alpha(z) \neq 0$ for all $|z| \leq 1$.

Assumption 2.2 (ε_t) is an iid $(0, \sigma^2)$ sequence of random variables with $\mathbf{E}|\varepsilon_t|^\ell < \infty$ for some $\ell > 4$, and its distribution is absolutely continuous with respect to the Lebesgue measure and has characteristic function φ such that $\lim_{\lambda \to \infty} \lambda^r \varphi(\lambda) = 0$ for some $r > 0$.

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 the space of cadlag functions on $[0, 1]$. Under Assumption 2.2, an invariance principle holds for U_T, viz.,

$$U_T \to_d U \tag{3}$$

as $T \to \infty$, where U is Brownian motion withvariance σ^2. It is also convenient to define $B_T(r)$ from (u_t), similarly as $U_T(r)$. Let $\alpha(1) = 1 - \sum_{k=1}^{p} \alpha_k$ and $\pi(1) = 1/\alpha(1)$. Then we have $B_T(r) = T^{-1/2} \sum_{t=1}^{[Tr]} u_t \to_d B(r)$ where $B = \pi(1)U$, as shown in Phillips and Solo (1992). This implies that the process (y_t) when scaled by $T^{-1/2}$ behaves asymptotically as the Brownian motion B under the null $\alpha = 1$ of unit root.

Our limit theory involves the two-parameter stochastic process called *local time* of Brownian motion.[2] We denote by L the (scaled) local time of U, which is defined by

$$L(t, s) = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t 1\{|U(r) - s| < \epsilon\} \, dr. \tag{4}$$

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

$$\int_0^t G(U(r)) \, dr = \int_{-\infty}^{\infty} G(s) L(t, s) \, ds$$

which we refer to as the occupation time formula. The conditions in Assumption 2.2 are required to obtain 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.

[2]The reader is referred to Chung and Williams (1990) and the references cited in Park and Phillips (1999, 2001) for the concept of local time and its use in the asymptotics for nonlinear models with integrated time series.

2.2 IV Estimation and Its Limit Theories

Taking the serial correlations in the error (u_t) specified in (2) into the model (1) gives

$$y_t = \alpha y_{t-1} + \sum_{k=1}^{p} \alpha_k \triangle y_{t-k} + \varepsilon_t \tag{5}$$

where $\triangle$ is the difference operator, since $\triangle y_t = u_t$ under the unit root null hypothesis. We consider the IV estimation of the augmented autoregression (5), using the following instruments

$$(F(y_{t-1}), \triangle y_{t-1}, \ldots, \triangle y_{t-p})' \tag{6}$$

where F is an integrable function. Notice that the instrument we use for the lagged level y_{t-1} is an integrable transformation of itself, i.e., $F(y_{t-1})$. For the lagged differences $(\triangle y_{t-1}, \ldots, \triangle y_{t-p})$, we use the variables themselves as the instruments. The transformation F is called *instrument generating function* (IGF) and assumed to satisfy

Assumption 2.3 Let F be integrable and satisfy $\displaystyle\int_{-\infty}^{\infty} xF(x)dx \neq 0$.

The condition given in Assumption 2.3 requires that the instrument $F(y_{t-1})$ is correlated with the regressor y_{t-1} that it is instrumenting for. It is shown for simple random walk models in Phillips, Park and Chang (1999) that IV estimators become inconsistent when the instrument is generated by a regularly integrable function F such that $\int_{-\infty}^{\infty} xF(x)\,dx = 0$. It is analogous to the non-orthogonality (between the instruments and regressors) requirement for the validity of IV estimation in standard stationary regressions. In our nonstationary IV estimation, where an integrable function of an integrated regressor is used as the instrument, such instrument failure occurs when the instrument generating function is orthogonal to the regression function.

Define $x_t = (\triangle y_{t-1}, \ldots, \triangle y_{t-p})'$, and let y, y_ℓ and X be the observations on y_t, y_{t-1} and x_t, respectively. Using this notation, we write the regression (5) in matrix form as

$$y = y_\ell \alpha + X\beta + \varepsilon = Y\gamma + \varepsilon$$

where $\beta = (\alpha_1, \ldots, \alpha_p)'$, $Y = (y_\ell, X)$, and $\gamma = (\alpha, \beta')'$. Denote by w_t the instrumental variables given in (6), and let $W = (F(y_\ell), X)$ be the matrix of observations on w_t, where $F(y_\ell) = (F(y_p), \ldots, F(y_{T-1}))'$. We consider the following estimator $\hat{\gamma}$ for γ:

$$\hat{\gamma} = \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = (W'Y)^{-1}W'y = \begin{pmatrix} F(y_\ell)'y_\ell & F(y_\ell)'X \\ X'y_\ell & X'X \end{pmatrix}^{-1} \begin{pmatrix} F(y_\ell)'y \\ X'y \end{pmatrix} \tag{7}$$

which is the usual IV estimator of γ using the instruments w_t. The IV estimator $\hat{\alpha}$ for the AR coefficient α corresponds to the first element of $\hat{\gamma}$, and is given by

$$\hat{\alpha} - 1 = B_T^{-1} A_T \tag{8}$$

under the null, where

$$\begin{aligned} A_T &= \sum_{t=1}^{T} F(y_{t-1})\varepsilon_t - \sum_{t=1}^{T} F(y_{t-1})x_t' \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \sum_{t=1}^{T} x_t \varepsilon_t \\ B_T &= \sum_{t=1}^{T} F(y_{t-1})y_{t-1} - \sum_{t=1}^{T} F(y_{t-1})x_t' \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \sum_{t=1}^{T} x_t y_{t-1} \end{aligned}$$

The variance of A_T is given by

$$\sigma^2 \mathbf{E} C_T$$

under Assumption 2.2, where

$$C_T = \sum_{t=1}^{T} F^2(y_{t-1}) - \sum_{t=1}^{T} F(y_{t-1})x_t' \left(\sum_{t=1}^{T} x_t x_t'\right)^{-1} \sum_{t=1}^{T} x_t F(y_{t-1}).$$

For testing the unit root hypothesis $\alpha = 1$, we construct the t-ratio statistic from the nonlinear IV estimator $\hat{\alpha}$ defined in (8) as

$$\tau = \frac{\hat{\alpha} - 1}{s(\hat{\alpha})} \tag{9}$$

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

$$s^2(\hat{\alpha}) = \hat{\sigma}^2 B_T^{-2} C_T \tag{10}$$

The $\hat{\sigma}^2$ is the usual variance estimator given by $T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_t^2$, where $\hat{\varepsilon}_t$ is the fitted residual from the augmented regression (5), i.e., $\hat{\varepsilon}_t = y_t - \hat{\alpha} y_{t-1} - x_t'\hat{\beta}$. It is natural in our context to use the IV estimate $(\hat{\alpha}, \hat{\beta})$ given in (7) to get the fitted residual $\hat{\varepsilon}_t$. However, we may obviously use any other estimator of (α, β) as long as it yields a consistent estimate for the residual error variance. The nonlinear IV t-ratio τ reduces to $\tau = \hat{\sigma}^{-1} A_T C_T^{-1/2}$, due to (8), (9) and (10).

The following lemma presents the asymptotics for the sample product moments appearing in the definition of τ.

Lemma 2.4 Under Assumptions 2.1, 2.2 and 2.3, we have

(a) $T^{-1/4} \sum_{t=1}^{T} F(y_{t-1})\, \varepsilon_t \;\to_d\; \mathbf{MN}\left(0,\; \sigma^2 \alpha(1)\, L(1,0) \int_{-\infty}^{\infty} F^2(s) ds\right)$

(b) $T^{-1/2} \sum_{t=1}^{T} F^2(y_{t-1}) \;\to_d\; \alpha(1)\, L(1,0) \int_{-\infty}^{\infty} F^2(s) ds$

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

as $T \to \infty$.

Parts (a) and (b) can easily be obtained as in Park and Phillips (1999, 2001), using the recent extension made by Park (2003a).[3] Part (c) is shown in Chang (2002).

[3] Park and Phillips (1999, 2001) impose piecewise higher-order Lipschitz conditions on F to derive the asymptotics here. Such conditions, however, are shown to be unnecessary in Park (2003a).

The result in part (a) shows that the covariance asymptotics of the integrable F yields a mixed normal limiting distribution with a mixing variate depending upon the local time L of the limit Brownian motion U, as well as the integral of the square of the instrument generating function F. It is indeed an intriguing result, which will lead to the normal limit theory for the nonlinear IV t-ratio. It is very useful to note

$$\begin{aligned} T^{-1/4}\sum_{t=1}^{T} F(y_{t-1})\varepsilon_t &\approx_d \sqrt[4]{T}\int_0^1 F(\sqrt{T}B(r))dU(r) \\ T^{-1/2}\sum_{t=1}^{T} F^2(y_{t-1}) &\approx_d \sqrt{T}\int_0^1 F^2(\sqrt{T}B(r))dr \end{aligned}$$

from which we may easily deduce the results in parts (a) and (b) using the theory of continuous martingales and the relationship $B = \pi(1)U$ with $\pi(1) = 1/\alpha(1)$. The reader is referred to Park (2003b) for a heuristic of the results in parts (a) and (b) along this line.

The result in part (c) implies that the lagged differences $\triangle y_{t-k}$, $k = 1, \ldots, p$, and the nonlinear instrument $F(y_{t-1})$ are asymptotically independent. This in turn implies that the presence of the stationary lagged differences in our base regression (5) does not affect the limit theory of the nonlinear IV estimator for the coefficient on the lagged level, which is an integrated process under the null. The rate $T^{3/4}$ at which the sample moment vanishes is obtained in Chang, Park and Phillips (2001). The asymptotic orthogonalities we have here are analogous to the familiar asymptotic orthogonalities between the stationary and integrated regressors we have seen in the usual cointegrating regressions, shown earlier in Park and Phillips (1988).

The limit null distribution of the IV t-ratio statistic τ now follows readily from the results in Lemma 2.4, as in Phillips, Park and Chang (1999) and Chang (2002).

Theorem 2.5 Under Assumption 2.1, 2.2 and 2.3, we have

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

as $T \rightarrow \infty$.

The limiting distribution of τ for testing $\alpha = 1$ is standard normal if a regularly integrable function is used to generate the instrument. The limit theory here is fundamentally different from the asymptotics for the usual linear unit root tests such as those by Phillips (1987) and Phillips and Perron (1988). The use of nonlinear IV is essential for our Gaussian limit theory which is obtained from the local time asymptotics and mixed normality result given in Lemma 2.4.

The nonlinear IV t-ratio statistic is a consistent test for testing the null of unit root. To see this, consider the limit behavior of the IV t-ratio τ given in (9) under the alternative of stationarity, i.e., $\alpha = \alpha_0 < 1$. We may express τ as

$$\tau = \tau(\alpha_0) + \frac{\sqrt{T}(\alpha_0 - 1)}{\sqrt{T}\, s(\hat{\alpha})} \tag{11}$$

where

$$\tau(\alpha_0) = \frac{\hat{\alpha} - \alpha_0}{s(\hat{\alpha})} \tag{12}$$

which is the IV t-ratio for testing $\alpha = \alpha_0 < 1$. Under the alternative, we may expect that $\tau(\alpha_0) \to_d \mathbf{N}(0,1)$ if the usual mixing conditions for (y_t) are assumed to hold. Moreover, if we let $B_0 = \text{plim}_{T\to\infty} T^{-1} B_T$ and $C_0 = \text{plim}_{T\to\infty} T^{-1} C_T$ exist under suitable mixing conditions for (y_t), then the second term in the right hand side of equation (11) diverges to $-\infty$ at the rate of $\sqrt{T}$ under the alternative of stationarity. This is because $\sqrt{T}(\alpha_0 - 1) \to -\infty$ and $\sqrt{T}\, s(\hat{\alpha}) \to_p \nu$, as $T \to \infty$, where $\nu^2 = \sigma^2 B_0^{-2} C_0 > 0$. Hence, under the alternative

$$\tau \to_p -\infty$$

at $\sqrt{T}$-rate, implying that the IV t-ratio τ is $\sqrt{T}$-consistent, just as in the case of the usual OLS-based t-type unit root tests such as the augmented Dickey-Fuller test.

We note that the IV t-ratios constructed from integrable IGFs are asymptotically normal, for all $|\alpha| \leq 1$, which makes a drastic contrast with the limit theory of the standard t-statistic based on the ordinary least squares estimator. The continuity of the limit distribution of the t-ratio in (12) across all values of α including the unity, in particular, allows us to construct the confidence intervals for α from the nonlinear IV estimator $\hat{\alpha}$ given in (8). We may construct $100\,(1-\lambda)\,\%$ asymptotic confidence interval for α as

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

where $z_{\lambda/2}$ is the $(1-\lambda/2)$-percentile from the standard normal distribution. This is another important advantage of using the nonlinear IV method. The OLS-based standard t-ratio, such as ADF test, has non-Gaussian limiting null distribution called Dickey-Fuller distribution. The DF distribution is tabulated in Fuller (1996) and known to be asymmetric and skewed to the left, rendering it impossible to construct an confidence interval which is valid for all $|\alpha| \leq 1$ from the standard OLS based t-ratios.

2.3 IV Estimation for Models with Deterministic Trends

The models with deterministic components can be analyzed similarly as in the previous section, if properly demeaned or detrended data are used. A proper demeaning or detrending scheme required here must be able to successfully remove the nonzero mean or time trend, while maintaining the orthogonality of the demeaned or detrended series with the error term, which is essential in retaining the martingale property of the error and ultimately the Gaussian limit theory of the covariance asymptotics for the nonlinear instrument. To meet the needs described above, Chang (2002) devises the *adaptive demeaning* and *detrending*, which we introduce below.

For a time series (z_t) with a nonzero mean given by

$$z_t = \mu + y_t \tag{13}$$

where the stochastic component (y_t) is generated as in (1), we may test for a unit root in (y_t) from the regression

$$y_t^{\mu} = \alpha y_{t-1}^{\mu} + \sum_{k=1}^{p} \alpha_k \triangle y_{t-k}^{\mu} + \varepsilon_t \tag{14}$$

where y_t^μ and y_{t-1}^μ are, respectively, the *adaptively demeaned* series of z_t and z_{t-1} that are defined as

$$\begin{aligned} y_t^\mu &= z_t - \frac{1}{t-1}\sum_{k=1}^{t-1} z_k \\ y_{t-1}^\mu &= z_{t-1} - \frac{1}{t-1}\sum_{k=1}^{t-1} z_k \\ \triangle y_{t-k}^\mu &= \triangle z_{t-k}, \qquad k = 1, \ldots, p. \end{aligned}$$

The term $(t-1)^{-1}\sum_{k=1}^{t-1} z_k$ appearing in the definitions above is the least squares estimator of μ obtained from the preliminary regression

$$z_k = \mu + y_k, \quad \text{for} \quad k = 1, \ldots, t-1.$$

We note that the mean μ is estimated from the model (13) using the observations up to time $(t-1)$ only, rather than using the full sample. This leads to the demeaning based on the partial sum of the data up to $(t–1)$, and for this reason the above demeaning scheme is called adaptive. Notice that even for the t-th observation z_t, we use $(t-1)$-adaptive demeaning to maintain the martingale property. The lagged differences $\triangle z_{t-k}$, $k = 1, \ldots, p$, need no further demeaning, since the differencing has already removed the mean.

We may then construct the nonlinear IV t-ratio τ^μ to test for unit root in (y_t) based on the nonlinear IV estimator for α from the regression (14), just as in (9). The adaptively demeaned lagged level y_{t-1}^μ is orthogonal to the error ε_t, which ensures the predictability of the nonlinear instrument $F(y_{t-1}^\mu)$, thereby retaining the martingale property of the error. Consequently, the sample moment $T^{-1/4}\sum_{t=1}^{T} F(y_{t-1}^\mu)\varepsilon_t$ has mixed normal limit theory, as in the models with no deterministic trends given in Lemma 2.4 (a), leading to the normal distribution theory for the IV t-ratio τ^μ for the models with nonzero mean.

We may also test for a unit root in the stochastic component of the time series with linear time trend using the nonlinear IV t-ratio τ^τ constructed in the similar manner using *adaptively detrended* data. For a time series with a linear time trend given by

$$z_t = \mu + \delta t + y_t \tag{15}$$

where (y_t) is generated as in (1), we may use the following regression as the basis for testing unit root in the stochastic component (y_t):

$$y_t^\tau = \alpha y_{t-1}^\tau + \sum_{k=1}^{p} \alpha_k \triangle y_{t-k}^\tau + e_t \tag{16}$$

where y_t^τ, y_{t-1}^τ and $\triangle y_{t-k}^\tau$ are adaptively detrended series of z_t, z_{t-1} and $\triangle z_{t-k}$, $k = 1, \ldots, p$, given as

$$y_t^\tau = z_t + \frac{2}{t-1}\sum_{k=1}^{t-1} z_k - \frac{6}{t(t-1)}\sum_{k=1}^{t-1} k z_k - \frac{1}{T} z_T$$

$$
\begin{aligned}
y_{t-1}^{\tau} &= z_{t-1} + \frac{2}{t-1}\sum_{k=1}^{t-1} z_k - \frac{6}{t(t-1)}\sum_{k=1}^{t-1} k z_k \\
\triangle y_{t-k}^{\tau} &= \triangle z_{t-k} - \frac{1}{T_i} z_T, \qquad k = 1, \ldots, p,
\end{aligned}
$$

and (e_t) are the regression errors. The variables z_t and z_{t-1} are detrended using the least squares estimators of the drift and trend coefficients, μ and δ, from the model (15) using again the observations up to time $(t-1)$ only, viz.,

$$
z_k = \mu + \delta k + y_{ik}, \quad \text{for} \;\; k = 1, \ldots, t-1.
$$

The term $T^{-1} z_T$ apprearing in the definitions of y_t^{τ} and $\triangle y_{t-k}^{\tau}$, $k = 1, \ldots, p$, is the grand sample mean of $\triangle z_t$, i.e., $T^{-1}\sum_{k=1}^{T} \triangle z_k$. The grand sample mean is needed for y_t^{τ} to eliminate the remaining the drift term of $z_t + 2(t-1)^{-1}\sum_{k=1}^{t-1} z_k - 6(t(t-1))^{-1}\sum_{k=1}^{t-1} k z_k$, and for $\triangle y_{t-k}^{\tau}$ to remove the nonzero mean of $\triangle z_{t-k}$, for $k = 1, \ldots, p$.

The nonlinear IV t-ratio τ^{τ} is then defined as in (9) from the nonlinear IV estimator for α from the regression (16) with the detrended data. The adaptive detrending of the data given above also preserves the predictability of the instrument $F(y_{t-1}^{\tau})$, and the martingale property of the error e_t is also retained in this case, as shown in Chang (2002). These ensure the mixed normal limit theory of the the sample moment $T^{-1/4}\sum_{t=1}^{T} F(y_{t-1}^{\tau}) e_t$ and the normal limit theory for the IV t-ratio τ^{τ}. For the actual derivation of the limit theories for the statistics τ^{μ} and τ^{τ}, we need to characterize the limit processes of the adaptively demeaned and detrended series, (y_t^{μ}) and (y_t^{τ}). When scaled by $T^{-1/2}$, (y_t^{μ}) and (y_t^{τ}) converge in distribution to the constant $\pi(1)$ multiple of the adaptively demeanded Brownian motion U^{μ} and the adaptively detrended Brownian motion U^{τ}, resectively. More explicitly, the *adaptively demeaned* Brownian is defined as

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

and similarly the *adaptively detrended* Brownian motion as

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

If we let $U^{\mu}(0) = 0$ and $U^{\tau}(0) = 0$, then both processes have well defined continuous versions on $[0, \infty)$, as shown in Chang (2002).

The asymptotic results given in Lemma 2.4 extend easily to the models with nonzero means and deterministic trends if we replace the lagged level y_{t-1} with the lagged detrended series y_{t-1}^{μ} and y_{t-1}^{τ}. They are indeed given similarly with the local times L^{μ} and L^{τ} of the adaptively demeaned and detrended Brownian motions U^{μ} and U^{τ} in the place of the local time L of the Brownian motion U. Then the limit theories for the nonlinear IV t-ratio statistics τ^{μ} and τ^{τ} for the models with nonzero means and deterministic trends follow immediately, as shown in Chang (2002), and are given in

Corollary 2.6 Under Assumption 2.1, 2.2 and 2.3, we have

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

as $T \to \infty$.

The standard normal limit theory of the nonlinear IV t-ratio for the models with no deterministic trends given in Theorem 2.5, therefore, continues to hold for the models with nonzero mean and linear time trend.

3. Panel Unit Root Tests

3.1 Issues in Panel Unit Root Testing

A number of unit root tests for panel data have been developed in the recent literature, including most notably those by Levin, Lin and Chu (2002) and Im, Pesaran and Shin (1997). However, they all have some important drawbacks and limitations. The test proposed by Levin, Lin and Chu (2002) is applicable only for homogeneous panels, where the AR coefficients for unit roots are in particular assumed to be the same across cross-sections. Their tests are based on the pooled OLS estimator for the unit root AR coefficient, and therefore, cannot be used for heterogeneous panels with different individual unit root AR coefficients. In addition, they assume cross-sectional independence. Im, Pesaran and Shin (1997) allow for the heterogeneous panels and propose the unit root tests which are based on the average of the individual unit root tests, t-statistics and LM statistics, computed from each individual unit. The validity of their tests, however, also require cross-sectional independence. Needless to say, the cross-sectional independence and homogeneity are quite restrictive assumptions for most of the economic panel data we encounter.

Chang (2002) explores the nonlinear IV methodology introduced in the previous section to solve the inferential difficulties in the panel unit root testing which arise from the instrinsic heterogeneities and dependencies of panel models. For each cross-section, the t-statistic for testing the unit root is constructed from the IV estimator of the autoregressive coefficient obtained from using an integrable transformation of the lagged level as instrument. As expected from our earlier results, each individual nonlinear IV t-ratio statistic constructed as such has standard normal limit null distribution. What is more important, however, is that the individual IV t-ratio statistics are asymptotically independent even across dependent cross-sectional units. This is indeed an intriguing result and follows from the asymptotic orthogonality results established in Chang, Park and Phillips (2001) for the nonlinear transformations of integrated processes by an integrable function.

The most important implication of the asymptotic normality and orthogonalities of the individual nonlinear IV t-ratios is that we now have a set of N asymptotically independent standard normal random variables to construct the unit root test for panels with N cross-sections. Of course, a simple average, defined as a standardized sum, of the individual IV t-ratios is a valid statistic for testing joint unit root null hypothesis for the entire panel. Chang (2002) shows that such a normalized sum of the individual IV t-ratios also has standard normal limit distribution, as long as the number of observations in each individual unit is large and the panel is asymptotically balanced in a weak sense. The standard limit theory is thus obtained without having to require the sequential asymptotics,[4] upon which

[4]The usual sequential asymptotics is carried out by first passing T to infinity with N fixed, and subse-

most of the available asymptotic theories for panel unit root models heavily rely. As a result, we may allow the number of cross-sectional units N to be arbitrarily small as well as large. The asymptotics developed here are T-asymptotics, and we assume that N is fixed throughout the paper.

However, there remains three important issues yet to be addressed. First, the presence of cointegration across cross-sectional units has never been allowed. It appears that there is a high potential for such possibilities in many panels of practical interests. Yet none of the existing tests, including the nonlinear IV panel unit root test by Chang (2002) discussed above, is applicable for such panels. Second, there is an issue of using covariates. As demonstrated by Hansen (1995) and Chang, Sickles and Song (2001), the inclusion of covariates can dramatically increase the power of the tests. Nevertheless, the pontential has never been investigated in the context of panels. Third, the issue of formulating the panel unit root hypothesis remains largely unresolved. We often need to consider the null and alternative hypotheses that some, not all, of the cross-sectional units have unit roots. Though none has seriously considered such hypotheses, they seem to be more relevant for many practical applications. The nonlinear IV methodology can be extended to resolve all these issues, as shown in Chang and Song (2002).

The improved IV methodology introduced in Chang and Song (2002) is based on the ADF regression further augmented with stationary covariates, and uses the instruments generated by a set of orthogonal functions. The use of orthogonal instrument generating functions for each cross-section is crucial for retaining the Gaussian limit theory and asymptotic orthogonalities of the nonlinear IV t-ratios in cointegrated panels. With these properties, the nonlinear IV t-ratios computed from N cross-sections continue to serve as a basis of N asymptotically iid standard normal random variables, which can be used to construct various unit root tests. In particular, we may use the order statistics, minimum or maximum, of the nonlinear IV t-ratios to test for or againt the existence of the unit roots in not all, but only in a fraction of the cross-sectional units. The limit distributions of the order statistics computed from the nonlinear IV t-ratios are nuisance parameter free and given by simple functionals of the standard normal distribution functions, as shown in Chang and Song (2002). This implies in particular that the critical values of the tests can be obtained from the standard normal table.

The nonlinear IV based panel unit root testing is indeed very general and applicable for a wide class of panel models. The nonlinear IV methodology, more specifically the asymptotic independence and normality of the nonlinear IV t-ratios, allows us to derive simple Gaussian limit theories for the panel unit root tests in panels with cross-sectional dependency at all levels. We may allow for cross-correlations of the innovations and/or cross-sectional dynamics in the short-run, and also for the comovements of the individual stochastic trends in the long-run. The nonlinear IV approach also allows us to formulate hypotheses in a much sharper form, which in turn makes it possible to test for and against the existence of unit roots in a sub-group of the cross-sections. Such general results are not entertained in other existing testing procedures, since their results require either cross-sectional independence or rely on a specific form of cross-sectional correlation structure,

quently let N go to infinity, usually under cross-sectional independence.

which may have a limited applicability in practice.

In what follows we first present the nonlinear IV unit root test for the dependent panels with no cointegration among cross-sections in Section 3.2, and subsequently in Section 4.1 introduce the improved IV methodology for cointegrated panels. In Section 4.2 the issue of properly formulating unit root hypotheses in panels is then discussed, and nonlinear IV order statistics are presented as vehicles to conduct testing for more flexibly formulated hypotheses.

3.2 Nonlinear IV Panel Unit Root Tests

We consider a panel unit root model given by

$$y_{it} = \alpha_i y_{i,t-1} + u_{it}, \quad i = 1, \ldots, N;\ t = 1, \ldots, T. \tag{17}$$

As usual, the index i denotes individual cross-sectional units, such as individuals, households, industries or countries, and the index t denotes time periods. The cross-sectional dimension N is not restricted and allowed to take large or small values. For expositional simplicity, the cross-sections are assumed to have the same T number of observations; however, the results here easily extend to unbalanced panels, as we discuss later in the section. The error terms u_{it} are allowed to have serial correlations, which we specify by an AR process of order p_i, i.e., $\alpha^i(L)u_{it} = \varepsilon_{it}$, where $\alpha^i(z) = 1 - \sum_{k=1}^{p_i} \alpha_{ik} z^k$, for $i = 1, \ldots, N$. Note that we let $\alpha^i(z)$ and p_i vary across i, thereby allowing heterogeneity in individual short-run dynamic structures. The cross-sections are allowed to be dependent via cross-sectional dependency of the innovations $\varepsilon_{it},\ i = 1, \ldots, N$, that generate the errors u_{it}. To be more explicit about the cross-sectional dependency we allow here, define $(\varepsilon_t)_{t=1}^T$ by

$$\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})' \tag{18}$$

We allow the N-dimensional innovation process (ε_t) to have covariance matrix Σ, which is unrestricted except for being positive definite.

We are interested in testing whether the series (y_{it}) generated as in (17) has a unit root in all cross-sections $i = 1, \ldots, N$, against the alternative that (y_{it}) are stationary in some cross-section i. The null hypothesis is therefore formulated as $H_0 : \alpha_i = 1$ for all i, and tested against the stationarity alternative $H_1 : |\alpha_i| < 1$ for some i. The test statistic we first consider for testing the panel unit root hypothesis is a simple average of the individual nonlinear IV t-ratio statistics for testing the unity of the AR coefficient computed from each cross-sectional unit. Under the AR(p_i) specification of the error u_{it}, the model (17) is written as

$$y_{it} = \alpha_i y_{i,t-1} + \sum_{k=1}^{p_i} \alpha_{ik} \triangle y_{i,t-k} + \varepsilon_{it}, \quad i = 1, \ldots, N;\ t = 1, \ldots, T \tag{19}$$

which is identical, at each cross-section level, to the regression (5) used in the previous section to derive the univariate nonlinear IV t-ratio given in (9). For each cross-section i,

we therefore instrument the lagged level $y_{i,t-1}$ using an integrable transformation of itself $F(y_{i,t-1})$, and accordingly define the nonlinear IV t-ratio τ_i for testing $\alpha_i = 1$ as

$$\tau_i = \frac{\hat{\alpha}_i - 1}{s(\hat{\alpha}_i)}$$

where $\hat{\alpha}_i$ and $s(\hat{\alpha}_i)$ are, respectively, the nonlinear IV estimator of α_i and its standard error, defined as in (8) and (10). The limit theories for each individual IV t-ratio τ_i follow exactly in the same manner as for the univariate nonlinear IV t-ratio given in Lemma 2.4 and Theorem 2.5, under the same set of conditions modified for our panel setting here. Let $|\cdot|$ denote the Euclidean norm: for a vector $x = (x_i)$, $|x|^2 = \sum_i x_i^2$, and for a matrix $A = (a_{ij})$, $|A| = \sum_{i,j} a_{ij}^2$. We assume

Assumption 3.1 For $i = 1, \ldots, N$, $\alpha^i(z) \neq 0$ for all $|z| \leq 1$.

Assumption 3.2 (ε_t) is an iid $(0, \Sigma)$ sequence of random vectors 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 $\lim_{\lambda \to \infty} |\lambda|^r \varphi(\lambda) = 0$, for some $r > 0$.

Under Assumption 3.2, an invariance principle holds for the partial sum process U_T of the N-vector innovations (ε_t), i.e., $U_T \to_d U$ as $T \to \infty$, where $U = (U_1, \ldots, U_N)'$ 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 $U_T(r)$. Let $\alpha^i(1) = 1 - \sum_{k=1}^{p_i} \alpha_{ik}$ and $\pi_i(1) = 1/\alpha^i(1)$. Then under Assumptions 3.1 and 3.2 we have $B_T \to_d B$, where $B = (B_1, \ldots, B_N)'$ and $B_i = \pi_i(1) U_i$ for $i = 1, \ldots, N$. The local times that appear in our limit theory are L_i that are (scaled) local time of U_i, for $i = 1, \ldots, N$. As shown in Chang (2002), we have

Theorem 3.3 Under Assumptions 2.3, 3.1 and 3.2, we have

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

as $T \to \infty$ for all $i = 1, \ldots, N$, and

$$\tau_i, \tau_j \ \sim \ \text{asymptotically independent}$$

for all $i \neq j = 1, \ldots, N$.

The standard normal limit theory of the univariate nonlinear IV t-ratio τ given in Theorem 2.5 continues to apply for each individual IV t-ratio τ_i for $i = 1, \ldots, N$. Moreover, the nonlinear IV t-ratios from different cross-sections are asymptotically independent even under cross-sectional dependency. The asymptotic orthogonalities of the nonlinear IV t-ratios play a crucial role in developing our panel unit root test, since these, along with the Gaussianity of the individual IV t-ratios, provide a basis of N asymptotically iid normal random variates for us to work with.

Before we proceed to define our panel unit root test, we provide some heuristics to understand the important asymptotic orthogonalities. We have the following distributional

equivalences in the limit for the sample moments determining the limit theories of the IV t-ratios τ_i and τ_j from the cross-sections i and j

$$T^{-1/4}\sum_{t=1}^{T} F(y_{i,t-1})\varepsilon_{it} \approx_d \sqrt[4]{T}\int_0^1 F(\sqrt{T}B_i(r))dU_i(r)$$
$$T^{-1/4}\sum_{t=1}^{T} F(y_{j,t-1})\varepsilon_{jt} \approx_d \sqrt[4]{T}\int_0^1 F(\sqrt{T}B_j(r))dU_j(r)$$

It is well known that the two right hand side stochastic processes become asymptotically independent if their quadratic covariation

$$\sigma_{ij}\sqrt{T}\int_0^1 F(\sqrt{T}B_i(r))F(\sqrt{T}B_j(r))dr \tag{20}$$

converges a.s. to zero, where σ_{ij} denotes the covariance between U_i and U_j. The order of the integral $\int_0^1 F(\sqrt{T}B_i(r))F(\sqrt{T}B_j(r))dr$ in the above equation is known to be $O_p(T^{-1}\log(T))$ a.s., due to Kasahara and Kotani (1979), which in turn implies that the quadratic covariation (20) vanishes almost surely, regardless of the value of σ_{ij}. The reasons why the integral is of such a small order are two-fold. The first is that the instrument generating function F is an integrable function which assigns a non-trivial value only when the argument takes a small value; and the second is that each of the arguments $\sqrt{T}B_i$ and $\sqrt{T}B_j$ takes a large value with increasing probability as T grows, due to the stochastic trends in the Brownian motions B_i and B_j. As a result, the product $F(\sqrt{T}B_i)F(\sqrt{T}B_j)$ takes a non-trivial value only with a small probability, making the above integral of such a small order.

Notice that U_i and U_j are the limit Brownian motions of the innovations (ε_{it}) and (ε_{jt}) generating the (y_{it}) and (y_{jt}). The above result therefore shows that 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_{it}) and (y_{jt}) generating the instruments are correlated. This then implies that the individual IV t-ratio statistics τ_i and τ_j constructed from the nonlinear instruments, $F(y_{i,t-1})$ and $F(y_{j,t-1})$, are asymptotically independent. Hence, we may naturally consider an average of such asymptotically normal and independent nonlinear IV t-ratios to test for the joint unit root null hypothesis $H_0:\ \alpha_i = 1$ for all $i = 1, \ldots, N$. The test is defined as

$$S = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\tau_i \tag{21}$$

and the limit theory for S follows immediately from Theorem 3.3 as

Theorem 3.4 We have

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

as $T \rightarrow \infty$ under Assumptions 2.3, 3.1 and 3.2.

Our limit theories derived here for the balanced panels continue to hold for the unbalanced panels, where each cross-section i may have a different number T_i of observations,

as shown in Chang (2002). To deal with the unbalanced nature of the data, we only need to ensure that the panel is asymptotically balanced in a weak sense. Specifically, it is required that $T_i \to \infty$ for all i, and $T_{\min}^{-3/4} T_{\max}^{1/4} \log(T_{\max}) \to 0$, where $T_{\min}$ and $T_{\max}$ denote, repectively, the minimum and the maximum of T_i, $i = 1, \ldots, N$. Note that our limit theory is derived using T-asymptotics only, and the factor $N^{-1/2}$ in the definition of the test statistic S in (21) is used just as a normalization factor, since S is based on the sum of N independent random variables. This implies the dimension of the cross-sectional units N may take any value, small as well as large. Our results also extend to the panels with heterogeneous deterministic components, such as individual fixed effects and linear time trends. The asymptotic results established here for the models without deterministic components continue to hold for those with nonzero means and linear trends, if the nonlinear IV estimation in each cross-section is based on the properly detrended data, using the adaptive demeaning or detrending scheme introduced in Section 2.3. The standard normal theory for our nonlinear IV panel unit root test constructed from the adaptively detrended data, therefore, continues to hold for the panels with heterogeneous fixed effects and time trends.

The normal limit theory is also obtained for the existing panel unit root tests, such as the pooled OLS test by Levin, Lin and Chu (2002) and the group mean t-bar statistic by Im, Pesaran and Shin (1997); however, their theory holds only under cross-sectional independence, and is obtained only through sequential asymptotics. More recently, several authors have made serious attempt to allow for cross-sectional dependencies. Chang (1999) allows for dependencies of unrestricted form, but her bootstrap procedure requires the dimension of time series T to be substantially larger than that of the cross-section N, which is restrictive for many practical applications. On the other hand, the procedures by Choi (2001), Phillips and Sul (2001) and Moon and Perron (2001) allow for a specific form of cross-sectional correlation structures, which may find only little justification in practical applications. In constrast, here we achieve the standard limit theory from the Gaussianity and independence of the limit distributions of the individual IV t-ratios, without having to assume independence across cross-sectional units, or relying on specific correlation structures. We can now do simple inference for panel unit root testing based on the critical values from the standard normal table in dependent panels driven by cross-correlated innovations and with various heterogeneities.

4. Extensions and Generalizations

4.1 Cointegrated Panels with Covariates

Our nonlinear IV approach can be extended to allow for the presence of cross-sectional cointegration and the use of relevant stationary covariates. We now let $y_t = (y_{1t}, \ldots, y_{Nt})'$ and assume that there are $N - M$ cointegrating relationships in the unit root process (y_t), which are represented by the cointegrating vectors (c_j), $j = 1, \ldots, N - M$. The usual vector autoregression and error correction representation allows us to specify the short-run

dynamics of (y_t) as

$$\triangle y_{it} = \sum_{j=1}^{N} \sum_{k=1}^{P_{ij}} a_{ij} \triangle y_{j,t-k} + \sum_{j=1}^{N-M} b_{ij} c_j' y_{t-1} + \varepsilon_{it} \tag{22}$$

for each cross-sectional unit, $i = 1, \ldots, N$.

Our unit root tests at individual levels will be based on the regression

$$y_{it} = \alpha_i y_{i,t-1} + \sum_{k=1}^{P_i} \alpha_{ik} \triangle y_{i,t-k} + \sum_{k=1}^{Q_i} \beta_{ik}' w_{i,t-k} + \varepsilon_{it} \tag{23}$$

for $i = 1, \ldots, N$, where we interpret (w_{it}) as the *covariates* added to the augmented DF regression for the i-th cross-sectional unit. It is important to note that the vector autoregression and error correction formulation of the cointegrated unit root panels in (22) suggests that we use such covariates. Under the null, we may obviously rewrite (1) and (22) as (23) with the covariates which may include several lagged differences of other cross-sections and linear combinations of lagged levels of all cross-sections. Indeed, in many panels of interest, we naturally expect to have inter-related short-run dynamics, which would make it necessary to include the dynamics of other cross-sections to properly model the dynamics in an individual unit. In the presence of cointegration, we also need to incorporate the long-run trends of other cross-sectional units, since the stochastic trends of other cross-sectional units would interfere with the short-run dynamics in an individual unit through the error correction mechanism. Hence, the potential covariates for unit root testing in a cointegrated panel include the lagged differences of other cross-sections to account for interactions in short-run dynamics and the linear combinations of cross-sectional levels in the presence of cointegration. Other stationary covariates may also be included to account for idiosyncratic characteristics of cross-sectional units.

The idea of using stationary covariates in the unit root regression was first entertained by Hansen (1995) in a univariate context, and pursued further by Chang, Sickles and Song (2001) using a bootstrap method. Both papers clearly demonstrate that the use of meaningful covariates offers a great potential in power gains for the test of a unit root. In light of the fact that one of the main motivations to use panels to test for unit roots is to increase the power, we have indeed overlooked this important opportunity in panel unit root testing. Choosing proper covariates can be a difficult issue in univariate applications, and the limit distribution of the standard covariates augmented ADF test by Hansen (1995) involves a nuisance parameter. In a panel context, however, there are many natural candidates for covariates, such as those listed above. Moreover, the nonlinear instruments are asymptotically orthogonal to the stationary covariates, as shown in Chang and Song (2002), and therefore the inclusion of the stationary covariates does not incur the nuisance parameter problem for our nonlinear IV tests.

Denote as before by $U = (U_1, \ldots, U_N)'$ the limit of the partial sum process constructed from the N-vector innovations $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$, which is assumed to satisfy the conditions in Assumption 2.2, and define $B = (B_1, \ldots, B_N)'$ to be the corresponding limit Brownian motion for $u_t = (u_{1t}, \ldots, u_N)'$. As is well known, the presence of cointegration in (y_t) implies

that the vector Brownian motion B is degenerate, i.e., some of the individual limit Brownian motions B_i, $i = 1, \ldots, N$, are linearly dependent. This has important consequences for our asymptotic analyses. Most importantly, the asymptotic orthogonalities of the nonlinear IV t-ratios no longer hold. This is because the quadratic covariation given in (20) no longer vanishes in the limit. To see this, suppose $B_i = B_j$ for some i and j. Then the quadratic covariation becomes

$$\sigma_{ij}\sqrt{T}\int_0^1 F(\sqrt{T}B_i(r))F(\sqrt{T}B_i(r))dr = \sigma_{ij}\sqrt{T}\int_0^1 F^2(\sqrt{T}B_i(r))dr = O_p(1), \text{ a.s.}$$

since the integral $\int_0^1 F^2(\sqrt{T}B_i(r))dr$ is now of order $O_p(T^{-1/2})$, which is higher than the order for the case with no cointegration, given below Theorem 3.3. This is because the integral now involves an integrable function of only one Brownian motion, viz., $F^2(\sqrt{T}B_i)$, which takes a nontrivial value with much larger probability compared to the product $F(\sqrt{T}B_i)F(\sqrt{T}B_j)$.

The presence of cointegration can be dealt with if we use an orthogonal set of functions as IGFs. This idea is exploited in Chang and Song (2002). It is indeed easy to see that the quadratic covariation given above may vanish in the limit if we use orthogonal functions F_i and F_j to generate the instruments for cross-sections i and j. As shown in, e.g., Park (2003b),

$$\sqrt{T}\int_0^1 F_i(\sqrt{T}B_i(r))F_j(\sqrt{T}B_i(r))dr \to_{a.s.} 0$$

if F_i and F_j are orthogonal, and $\int_{-\infty}^{\infty} F_i(x)F_j(x)\,dx = 0$.

The Hermite functions of odd orders $k = 2i-1$, $i = 1, \ldots, N$, for instance, can be used as a valid set of IGFs for the cointegrated panels. The Hermite function G_k of order k, $k = 0, 1, 2, \ldots$, is defined as

$$G_k(x) = (2^k k!\sqrt{\pi})^{-1/2} H_k(x) e^{-x^2/2}$$

where H_k is the Hermite polynomial of order k given by

$$H_k(x) = (-1)^k e^{x^2} \frac{d^k}{dx^k} e^{-x^2}$$

It is well known that the class of Hermite functions introduced above forms an orthonormal basis for $L^2(\mathbf{R})$, i.e., the Hilbert space of square integrable functions on $\mathbf{R}$. We thus have

$$\int_{-\infty}^{\infty} G_j(x)G_k(x)dx = \delta_{jk}$$

for all j and k, where δ_{jk} is the kronecker delta. Moreover, the odd order Hermite functions satisfy the IGF validity conditions in Assumption 2.3.

The IV estimation for the regression (23) is straightforward, given our discussions here. To deal with the cross-sectional cointegration, we use the instrument $F_i(y_{i,t-1})$ for the lagged level $y_{i,t-1}$ for each cross-sectional unit $i = 1, \ldots, N$, where (F_i) is a set of orthogonal IGFs. For the augmented regressors $x_{it} = (\triangle y_{i,t-1}, \ldots, \triangle y_{i,t-P_i}; w_{i,t-1}, \ldots, w_{i,t-Q_i})'$, we use the

variables themselves as the instruments. Hence the instruments $(F(y_{i,t-1}), x'_{it})'$ are used for the entire regressors $(y_{i,t-1}, x'_{it})'$. As is well expected, our previous asymptotic results apply also in this more general context. In particular, the limit theories in Theorems 3.3 and 3.4 continue to hold for the IV t-ratios τ_i's and also for the panel unit root statistic S in the presence of cointegration and covariates, if constructed in the way suggested above. This is shown in Chang and Song (2002).

4.2 Formulations of Hypotheses and Order Statistics

We now turn to the problems of formulating unit root hypotheses in panels, which was considered in Chang and Song (2002). For many practical applications, we are often interested in testing for unit roots collectively for a group of cross-sectional units included in the given panel. In this case, we need to more precisely formulate both the null and the alternative hypotheses. In particular, we may want to test for and against the existence of unit roots in not all, but only a fraction of cross-sectional units. Such formulations, however, seem more relevant and appropriate for many interesting empirical applications, including testing for purchasing power parity and growth convergence, among others. Here we lay out three sets of possible hypotheses in panel unit root testing and propose to use the order statistics constructed from the nonlinear IV t-ratios to effectively test for such hypotheses.

We consider

Hypotheses (A) $\mathrm{H}_0 : \alpha_i = 1$ for all i *versus* $\mathrm{H}_1 : \alpha_i < 1$ for all i

Hypotheses (B) $\mathrm{H}_0 : \alpha_i = 1$ for all i *versus* $\mathrm{H}_1 : \alpha_i < 1$ for some i

Hypotheses (C) $\mathrm{H}_0 : \alpha_i = 1$ for some i *versus* $\mathrm{H}_1 : \alpha_i < 1$ for all i

Hypotheses (A) and (B) share the same null hypothesis that a unit root is present in all individual units. However, the null hypothesis competes with different alternative hypotheses. In Hypotheses (A), the null is tested against the alternative hypothesis that all individual units are stationary, while in Hypothesis (B) it is tested against the alternative that there are some stationary individual units. On the contrary, the null hypothesis in Hypotheses (C) holds as long as unit root exists in at least one individual unit, and is tested against the alternative hypothesis that all individual units are stationary. The alternative hypotheses in both Hypotheses (B) and (C) are the negations of their respective null hypotheses. This is, however, not the case for Hypotheses (A). Hypotheses (C) has never been considered in the literature. Note that the rejection of H_0 in favor of H_1 in Hypotheses (C) directly implies that all (y_{it})'s are stationary, and therefore, purchasing power parities or growth convergence hold if we let (y_{it})'s be real exchange rates or differences in growth rates respectively. No test, however, is available to deal with Hypotheses (C) appropriately.

For the tests for Hypotheses (A)–(C), we define

$$S = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \tau_i$$

$$
\begin{aligned}
S_{\min} &= \min_{1 \le i \le N} \tau_i \\
S_{\max} &= \max_{1 \le i \le N} \tau_i
\end{aligned}
$$

where τ_i is the nonlinear IV t-ratio for the i-th cross-sectional unit. The averaged statistic S is comparable to other existing tests and proposed for the test of Hypotheses (A). Virtually all of the existing panel unit root tests effectively test for Hypotheses (A). Some recent work, including Im, Pesaran and Shin (1997) and Chang (2002), formulate their null and alternative hypotheses as in Hypotheses (B). However, their use of average t-ratios can only be justified for the test of Hypotheses (A). The minimum statistic $S_{\min}$ is more appropriate for the test of Hypotheses (B) than the tests based on the averages. The average statistic S may also be used to test for Hypotheses (B), but the test based on $S_{\min}$ would have more power, especially when only a small fraction of cross-sectional units are stationary under the alternative hypothesis. The maximum statistic $S_{\max}$ can be used to test Hypotheses (C). Obviously, the average statistic S cannot be used to test for Hypotheses (C), since it would have incorrect size.

Let M be $0 \le M \le N$, and assume $\alpha_i = 1$ for $1 \le i \le M$, and set $M = 0$ if $\alpha_i < 1$ for all $1 \le i \le N$. It is easy to derive the asymptotic theories for the test statistics defined above. Recall that the nonlinear IV t-ratios for all individual cross-sections have standard normal limiting distributions, and are asymptotically orthogonal to each other. This leaves us a set of M random variates τ_i's that are asymptotically independent and identically distributed as standard normal. To obtain the asymptotic critical values for the statistics S, $S_{\min}$ and $S_{\max}$, we let Φ be the distribution function for the standard normal distribution, and let λ be the size of the tests. For a given size λ, we define $x_M(\lambda)$ (with $x_1(\lambda) = x(\lambda)$) and $y_N(\lambda)$ by

$$
\Phi(x_M(\lambda))^M = \lambda, \quad (1 - \Phi(y_N(\lambda)))^N = 1 - \lambda
$$

These provide the critical values of the statistics S, $S_{\min}$ and $S_{\max}$ for the tests of Hypotheses (A) – (C). The following table shows the tests and critical values that can be used to test each of Hypotheses (A) – (C).

Hypotheses	Test Statistics	Critical Values
Hypotheses (A)	S	$x(\lambda)$
Hypotheses (B)	S $S_{\min}$	$x(\lambda)$ $y_N(\lambda)$
Hypotheses (C)	$S_{\max}$ $S_{\max}$	$x(\lambda)$ $x_M(\lambda)$

It is clear that

$$
\lim_{T \to \infty} \mathbf{P}\{S \le x(\lambda)\} = \lambda
$$

$$\lim_{T\to\infty} \mathbf{P}\{S_{\min} \le y_N(\lambda)\} \quad = \quad \lambda$$

if $M = N$. The tests using statistics S and $S_{\min}$ with critical values $x(\lambda)$ and $y_N(\lambda)$, respectively, have the exact size λ asymptotically under the null hypotheses in Hypotheses (A) and (B).

Moreover, if $1 \le M \le N$, then

$$\lim_{T\to\infty} \mathbf{P}\{S_{\max} \le x_M(\lambda)\} = \lambda, \quad \lim_{T\to\infty} \mathbf{P}\{S_{\max} \le x(\lambda)\} \le \lambda$$

The null hypothesis in Hypotheses (C) is composite, and the rejection probabilities of the test $S_{\max}$ based on the critical value $x(\lambda)$ may not be exactly λ even asymptotically. The size λ in this case is the maximum rejection probabilities that can be obtained under the null hypothesis.

The above results imply that the nonlinear IV order statistics suggested here have limit distributions which are nuisance parameter free and given by simple functions of the standard normal distribution function. The critical values are thus easily derived from those of the standard normal distribution.

5. Conclusion

The paper takes the nonlinear IV approach to testing for unit roots in general panels with dependency and heterogeneity. The nonlinear IV based tests have many desirable properties. First, the tests permit both short and long-run cross-sectional dependencies of the most general form. The presence of cointegration, the inter-relatedness of short-run dynamics and the cross-correlations in the errors are all allowed. This is in sharp constrast with other existing tests, which assume complete cross-sectional independence of the errors. Second, various heterogeneities are allowed. Heterogeneities in dynamics, individual fixed effects and trends, and the number of individual time series observations are all permitted in our framework. Third, the tests are very flexible to accommodate the extensions in various directions. In particular, they enable us to use relevant stationary covariates without having to run into nuisance parameter problems. For the standard approach, the use of covariates yields limit distributions including nuisance parameters that are difficult to deal with. They can also be extended to test for cointegration, as will be reported in a later work. Finally, our nonlinear IV method makes it possible to formulate the panel unit root hypotheses in more flexible and more precise forms, which can be tested by various order statistics. With the tests provided here, the empirical researchers will be able to investigate much more precisely and rigorously a wide variety of important economic issues regarding international and interstate comparisons and interactions including purchasing power parities and growth convergences/divergences.

6. References

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

University.

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

Chang, Y. (2002). "Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency," *Journal of Econometrics* 110, 261-292.

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

Chang, Y., J.Y. Park and P.C.B. Phillips (2001). "Nonlinear econometric models with cointegrated and deterministically trending regressors," *Econometrics Journal* 4, 1-36.

Chang, Y., R.C. Sickles and W. Song (2001). "Bootstrapping unit root tests with covariates," mimeographed, Department of Economics, Rice University.

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

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

Hansen, B.E. (1995). "Rethinking the univariate approach to unit root testing: Using covariates to increase power," *Econometric Theory* 11, 1148-1171.

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

Levin, A., C.F. Lin and C.S. Chu (2002). "Unit root tests in panel data: Asymptotic and finite sample properties," *Journal of Econometrics* 108, 1-24.

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

Moon, H.R. and B. Perron, (2001). "Testing for a unit root in panels with dynamic factors," mimeographed.

Park, J.Y. (2003a). "Strong approximations for nonlinear transformations of integrated time series," mimeographed, Department of Economics, Rice University.

Park, J.Y. (2003b). "Nonstationary nonlinearity: An outlook for new opportunities," mimeographed, Deparment of Economics, Rice University.

Park, J.Y. and P.C.B. Phillips (1988). "Statistical Inference in Regressions with Integrated Processes: Part 1," *Econometric Theory* 5, 468-498.

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 (2001). "Nonlinear regressions with integrated time series," *Econometrica* 69, 1452-1498.

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

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

Phillips, P.C.B., J.Y. Park and Y. Chang (1999). "Nonlinear instrumental variable estimation of an autoregression," forthcoming in *Journal of Econometrics*.

Phillips, P.C.B and P. Perron (1988). "Testing for a unit root in time series regression," *Biometrika* 75, 335-346.

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

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