Pooled Unit Root Tests in Panels with a Common Factor*

Joakim Westerlund†

August 4, 2005

Abstract

This paper proposes new pooled panel unit root tests that are appropriate when the data exhibit cross-sectional dependence that is generated by a single common factor. Using sequential limit arguments, we show that the tests have a limiting normal distribution that is free of nuisance parameters and that they are consistent and unbiased against heterogenous local alternatives. Our Monte Carlo results indicate that the tests perform well in comparison to several other popular tests that also presumes a common factor structure for the cross-sectional dependence.

JEL Classification: C12; C31; C33.

Keywords: Pooled Unit Root Tests; Panel Data; Common Factor; Cross-Sectional Dependence; Asymptotic Local Power.

1 Introduction

During the last few years there has been an immense proliferation of research concerned with the problem of testing for unit roots in panel data. Some of the most influential contributions within this field include Choi (2001), Im *et al.* (2003), Levin *et al.* (2002) and Harris and Tzavalis (1999). A common feature of these studies is that they all assume that the members of the panel are independent of each other. Under this assumption, various central limit theorems can be applied to obtain test statistics that achieve asymptotic normality. Although it has been widely recognized that cross-sectional independence may be an overly restrictive assumption, is has long been thought that subtracting the cross-sectional average from the data before application of the panel unit root test could be employed to, at least partially, deal with this problem. Recently,

*The author would like to thank David Edgerton, Peter Jochumzen and seminar participants at Lund University for many valuable comments and suggestions. Financial support from the Jan Wallander and Tom Hedelius Foundation, research grant number P2005-0117:1, is gratefully acknowledged.

†Department of Economics, Lund University, P. O. Box 7082, S-220 07 Lund, Sweden. Telephone: +46 46 222 4970; Fax: +46 46 222 4118; E-mail address: joakim.westerlund@nek.lu.se.

however, it has become increasingly clear that cross-sectional demeaning of this sort may not work in general as it cannot be used to accommodate correlations that differ between pairs of individual time series, which seem like a more realistic assumption in many empirical applications such as macroeconomics and international finance.

Recognizing this deficiency, several new panel data unit root tests have been proposed in the literature. These tests are distinct in that they make explicit assumptions regarding the structure of the cross-sectional correlation. Chang (2004), Maddala and Wu (1999) and Smith *et al.* (2003) avoid the restrictive nature of the cross-sectional demeaning by employing bootstrap techniques, which make the unit root test valid under quite general forms of cross-sectional dependence. Other tests that allow for general correlation structures include those of Chang (2002), O'Connell (1998) and Jönsson (2005). To eliminate the dependence, Chang (2002) proposes a nonlinear instrumental variables estimator, while O'Connell (1998) suggests estimation by generalized least squares. The test of Jönsson (2005) relies on a robust estimator of the standard errors to handle the impact of the dependence.

The tests proposed by Phillips and Sul (2003), Pesaran (2005), Bai and Ng (2004) and Moon and Perron (2004) also allow for more general forms of cross-sectional dependence by assuming that the correlation can be modelled using a common factor structure. These studies recognize that the common factors are likely to have differential effects on different cross-sectional units by allowing for the possibility of heterogeneous factor loadings. This avenue is potentially very fruitful as it takes an important step in the direction of reducing the dimension of the cross-sectional correlation structure, and it shall therefore be employed in this paper.

The basic idea here is to first estimate the common factors, and then to test whether the defactored data possesses a unit root or not. The suggestions of Bai and Ng (2004) and Moon and Perron (2004) are very similar in that they both rely on using principal components method to estimate the factors. Unfortunately, this makes the resulting test procedures quite difficult to implement. Another disadvantage with using this method to estimate the factors is that it requires not only the time series but also the cross-sectional dimension to be large, a requirement that is rarely fulfilled in practice. It also involves serious asymptotical difficulties as the error coming from the first stage estimation must be bounded somehow in order to obtain a valid test unit root test in the second.

Phillips and Sul (2003) uses an iterated method of moments approach to estimate the factor, which is simpler than principal components in many respects. In particular, it only requires that the number of time series observations is large, which makes both the practical implementation and the asymptotic analysis more manageable. Unfortunately, many of the tests discussed by Phillips and Sul (2003) are based on meta analysis, which means that the asymptotic distribution of the individual statistics must be estimated somehow before the panel statistic can be constructed. Thus, these tests are not very easy to compute.

In this paper, we follow Phillips and Sul (2003) and Pesaran (2005), and assume a common factor structure for the cross-sectional dependence. In particular, it is assumed that there is single common factor, which may have disparate effects on the different individual time series of the panel. Due to their simplicity, their extensive use in the empirical literature, and their increased availability through various econometric software packages, we propose two new tests based on the pooled unit root tests developed by Levin *et al.* (2002) and Harris and Tzavalis (1999).

In order to defactor the data, we propose using a version of the method of moments procedure developed by Phillips and Sul (2003). Since the consistency of this procedure only requires passing the number of time series observations to infinity, the asymptotic analysis of our new tests can be carried out by using simple sequential limit arguments. The results reveal that the tests have limiting normal distributions, and that they are unbiased against heterogeneous local alternative hypotheses. For a fixed specification of the alternative hypothesis, the tests are consistent. In our Monte Carlo study, we demonstrate that the tests have reasonable size properties and good power. We also find that the proposed tests compares favorably to a number of popular tests that also presumes a common factor structure for the cross-sectional dependence.

The paper proceeds as follows. Section 2 provides a brief presentation of the model that we use. Section 3 introduces the unit root test statistics, whereas Section 4 is concerned with their asymptotic properties. Sections 5 then present our Monte Carlo study. Section 6 concludes. For notational convenience, the Bownian motion $B(r)$ defined on the unit interval $r \in [0,1]$ will be written as only B and integrals such as $\int_0^1 W(r)\mathrm{d}r$ will be written $\int_0^1 W$ and $\int_0^1 W(r)\mathrm{d}W(r)$ as $\int_0^1 W\mathrm{d}W$. The symbols $\Rightarrow$ and $\xrightarrow{p}$ will be used to signify weak convergence and convergence in probability, respectively.

2 Model and assumptions

Let y_{it} be a vector of observed data on individual $i = 1, ..., N$ and time series $t = 1, ..., T$. The data generating process (DGP) of this variable is given by

$$y_t = \gamma d_t + \rho y_{t-1} + e_t, \tag{1}$$

where $y_t = (y_{1t}, ..., y_{Nt})'$, $e_t = (e_{1t}, ..., e_{Nt})'$ and $\rho = \mathrm{diag}(\rho_1, ..., \rho_N)$. For notational simplicity, we use d_t to indicate the vector of deterministic components and γ is used to indicate the corresponding vector of parameters. We shall distinguish between three different deterministic specifications. In Model 1, $d_t = 0$, which correspond to a model with no deterministic components. In Model 2, $d_t = 1$ so the deterministic component comprises an individual specific constant term. In Model 3, $d_t = (1, t)'$, which is the most general specification with both individual specific constants and linear time trends. For the present, y_t is assumed to evolve according to Model 1. Generalizations to Models 2 and 3 are straightforward and will be discussed when appropriate.

The focus of interest in this paper is the problem of testing for the presence of a common unit root in y_t against the following local alternative hypothesis

$$\rho = I_N - \frac{\phi}{TN^{1/2}}, \tag{2}$$

where $\phi = \text{diag}(\phi_1, ..., \phi_N)$ is the matrix of local-to-unity parameters. One way to formulate the null and alternative hypotheses is in terms of ϕ_i in which case $H_0 : \phi_i = 0$ for all i is tested versus $H_1 : \phi_i > 0$ for at least some i. Another more convenient way is to assume that ϕ_i is nonnegative and independent and identically distributed (i.i.d.) in which case the null and alternative hypotheses can be formulated using only the expected value of ϕ_i. Specifically, if we denote this expected value by μ, then we have the following equivalent formulation

$$H_0 : \mu = 0 \text{ versus } H_1 : \mu > 0.$$

To model the cross-sectional correlation, we follow Phillips and Sul (2003) and assume that the process e_t is generated by the following single factor model

$$e_t = \tau f_t + v_t, \tag{3}$$

where f_t is a scalar unobservable random factor, $\tau = (\tau_1, ..., \tau_N)'$ is a nonrandom vector of factor loadings and $v_t = (v_{1t}, ..., v_{Nt})'$ is a vector of idiosyncratic disturbances. For convenience in deriving the test statistics and their asymptotic distributions, we shall assume that v_t satisfy the following set of conditions.

Assumption 1. (Error process.) The vector v_t is i.i.d.$(0, \Omega_v)$, where Ω_v is positive definite and block-diagonal.

In addition, the factor model in (3) is introduced to model the cross-sectional dependence. To this end, we make the following assumption.

Assumption 2. (Factor model.) The common factor f_t is i.i.d.$(0, 1)$.

Assumption 1 and 2 establish the basic conditions needed for developing the new unit root tests. Many of these are quire restrictive but are made here to simplify the derivation of the tests and their limit distributions, and will be relaxed later on. Consider first Assumption 1. This assumption ensures that a functional central limit theorem holds individually for each cross-section as T increases. That is, we have $T^{-1/2} \sum_{t=1}^{[Tr]} v_t \Rightarrow B_v$ as $T \longrightarrow \infty$ for a fixed N, where $B_v = (B_{1v}, ..., B_{Nv})'$ is a vector Brownian motion with covariance matrix Ω_v. Note also that Assumption 2 implies that the elements of v_t are cross-sectionally independent, which follows directly from the block-diagonal shape of Ω_v.

Assumption 2 together with equation (3) characterizes the common factor model. As with Assumption 1, Assumption 2 also ensures the applicability of a functional central limit theorem. Thus, we deduce that $T^{-1/2} \sum_{t=1}^{[Tr]} f_t \Rightarrow B_f$ as $T \longrightarrow \infty$ with N held fixed, where B_f is a scalar Brownian motion. Also, the fact that f_t is assumed to be stationary ensures that y_t is the only source of nonstationarity in (1). As required by Assumption 3, the covariance of B_f is normalized to unity, which entails no loss of generality.

Most unit root tests in panel data, such as those of Im *et al.* (2003) and Levin *et al.* (2002), requires that y_t is cross-sectionally independent. This assumption is very strong and it is therefore unlikely to hold in many empirical applications. When the structure of the dependence is completely unknown, it is generally infeasible to deal with the unrestricted correlation structure because of the degree of freedom constraint. Therefore, in order to reduce the dimensionality of the covariance structure of the errors, it is common to make at least some simplifying assumption. The most common way to achieve this is to include a common time effect in the process driving y_t. The justification for doing so is that certain comovements of multidimensional time series may be due to a common factor. For instance, in international macroeconomics it might be argued that the common time effect represents some global shock, such as an oil price shock.

The model we use in this paper allows for a common time effect that may affect the individual time series differently. The extent of the cross-sectional correlation is determined by the loadings τ_i, which can be seen by writing

$$E(e_{it}e_{jt}) = \tau_i \tau_j \text{ for } i \neq j.$$

It follows that there is no cross-sectional correlation when $\tau_i = 0$ for all i and there is identical correlation when $\tau_i = \tau_j$ for all $i \neq j$. The model that we use is the same single factor model studied by Phillips and Sul (2003), which is different from the simple common time effect specification where the factor is assumed to have an identical impact on all the individuals of the panel. Our model also presumes a single common factor but it is less restrictive since it allows the factor to impact the individuals of the panel differently.

3 Pooled unit root tests

In this section, we develop the pooled unit root tests. Towards this end, notice that Assumption 2 and 3 enables us to write the covariance of e_t as

$$\Omega_e = E(e_t e_t') = \tau\tau' + \Omega_v. \tag{4}$$

Making use of this result, under the unit root null, it is possible to show that $T^{-1/2}y_{[Tr]} = T^{-1/2}\sum_{t=2}^{[Tr]} e_t \Rightarrow B$ as $T \longrightarrow \infty$ with N held fixed, where $B = (B_1, ..., B_N)'$ is a vector Brownian motion with covariance matrix equal to Ω_e. This implies that y_t is unsuitable for unit root testing purposes since its limit distribution depend on the nuisance parameters associated with the cross-sectional dependencies of the data as captured by the off-diagonal elements in Ω_e. Note, however, that B may be decomposed as

$$B = \tau B_f + B_v. \tag{5}$$

If we let $\tau_\perp$ be the $N \times (N-1)$ matrix that spans the orthogonal complement of the vector τ, then we may define

$$F_\perp \equiv (\tau_\perp' \Omega_v \tau_\perp)^{-1/2} \tau_\perp'. \tag{6}$$

Making use of (5) and (6), since $\tau_\perp' \tau = 0$ by definition, we may deduce that $T^{-1/2} F_\perp y_{[Tr]} \Rightarrow F_\perp B = F_\perp B_v \equiv W$ as $T \longrightarrow \infty$ for a fixed N, where $W = (W_1, ..., W_{N-1})'$ is an $N-1$ dimensional vector standard Brownian motion with covariance matrix equal to identity. Notice that, since the covariance matrix of B_v is given by Ω_v, it is easy to verify that the covariance of $F_\perp B_v$ is indeed identity. Notice also that the multiplication of B_v by $F_\perp$ reduces the dimension of W from N to $N-1$. This discussion suggests that y_t may be employed to test the unit root hypothesis after multiplying it by $F_\perp$. However, since $\tau_\perp$ and Ω_v are unknown, this means that $F_\perp$ is unobservable and that it needs to be consistently estimated. If we denote this estimator by $\hat{F}_\perp$, then we propose using the following two statistics for testing the hypothesis of H_0 versus H_1.

Definition 1. (Pooled defactored panel unit root statistics.) Let $y_t^* = \hat{F}_\perp(y_t - \hat{\gamma} d_t)$, where $\hat{\gamma}$ is the least squares estimate of γ, then the pooled defactored unit root statistics are defined as follows

$$PD_\rho \equiv \left(\sum_{t=2}^{T} y_{t-1}^{*\prime} y_{t-1}^* \right)^{-1} \sum_{t=2}^{T} \Delta y_t^{*\prime} y_{t-1}^*, \tag{7}$$

$$PD_t \equiv \left(\sum_{t=2}^{T} y_{t-1}^{*\prime} y_{t-1}^* \right)^{-1/2} \sum_{t=2}^{T} \Delta y_t^{*\prime} y_{t-1}^*. \tag{8}$$

Remark 1. Essentially, the above statistics are nothing but simple modifications of the pooled normalized bias and t-ratio statistics studied in the earlier literature by Harris and Tzavalis (1999) and Levin *et al.* (2002) using the defactored data. Note the particularly simple form of the t-ratio type statistic PD_t as it does not require any estimation of the error variance in the denominator, which is otherwise standard in the literature. This is a direct consequence of the defactoring procedure that we use, which makes the asymptotic covariance matrix of $T^{-1/2} F_\perp y_{[Tr]}$ equal to the identity matrix.

Remark 2. For the estimation of $F_\perp$, we propose using a version of the moment based procedure discussed in Phillips and Sul (2003). The purpose of this procedure is to retrieve $\hat{\tau}$, the vector of estimated loadings, together with $\hat{\Omega}_v$, the estimated covariance of v_t, as the minimizers of the sum of squared errors of τ and Ω_v from $\hat{\Omega}_e = T^{-1} \sum_{t=2}^{T} \hat{e}_t \hat{e}_t'$, where $\hat{e}_t = y_t - \hat{\gamma} d_t - \hat{\rho} y_{t-1}$ is the least squares estimator of e_t. This is different from the procedure employed by Phillips and Sul (2003), which uses $T^{-1} \sum_{t=2}^{T} \Delta e_t \Delta e_t'$ to estimate Ω_e. However, this estimator is only consistent under the null hypothesis suggesting that there should be some merit in using $\hat{\Omega}_e$, which is consistent under the null as well as under the alternative hypothesis. In particular, since $\hat{\Omega}_e$ is consistent in both cases, this suggests that the minimization of the sum of squared errors may be carried out with respect to only τ and Ω_v. Thus, $\hat{\tau}$ and $\hat{\Omega}_v$ may be obtained as

$$(\hat{\tau}, \hat{\Omega}_v) = \arg \min_{\tau, \Omega_v} \text{tr} \left((\hat{\Omega}_e - \Omega_v - \tau\tau')(\hat{\Omega}_e - \Omega_v - \tau\tau')' \right). \tag{9}$$

The solution of this problem satisfies the two equations $\hat{\tau} = (\hat{\tau}'\hat{\tau})^{-1}(\hat{\Omega}_e\hat{\tau} - \hat{\Omega}_v\hat{\tau})$ and $\hat{\Omega}_v = \text{diag}(\hat{\Omega}_e - \hat{\tau}\hat{\tau}')$. Obviously, since these equations have no closed form solutions, the minimization problem in (9) must be solved iteratively. To this end, we engage in an recursive procedure using the following updating scheme

$$\begin{aligned} \tau^r &= (\tau^{r-1\prime}\tau^{r-1})^{-1}(\hat{\Omega}_e\tau^{r-1} - \Omega_v^{r-1}\tau^{r-1}), \\ \Omega_v^r &= \text{diag}(\hat{\Omega}_e - \tau^r\tau^{r\prime}). \end{aligned}$$

The procedure is initiated by choosing τ^0, a vector of starting values for τ, which is used to compute $\Omega_v^0 = \hat{\Omega}_e - \tau^0\tau^{0\prime}$. The updating then continues until convergence. Once $\hat{\tau}$ and $\hat{\Omega}_v$ has been obtained, a consistent estimator of $\tau_\perp$ may be constructed by taking the eigenvectors of the projection matrix $I_N - \hat{\tau}(\hat{\tau}'\hat{\tau})^{-1}\hat{\tau}'$ that correspond to unit eigenvalues. If we denote the resulting estimator by $\hat{\tau}_\perp$, then $\hat{F}_\perp$ may be constructed as $\hat{F}_\perp = (\hat{\tau}_\perp'\hat{\Omega}_v\hat{\tau}_\perp)^{-1/2}\hat{\tau}_\perp'$.

Remark 3. As will be shown in Section 4, the asymptotic distributions of the above statistics are free of nuisance parameters as long as $d_t = 0$. However, if $d_t = 1$ or $d_t = (1, t)'$ as in Models 2 and 3, then these distributions will depend on various nuisance parameters associated with the underlying DGP. Therefore, in order to obtain statistics that are asymptotically similar in Model 2, the data should be demeaned prior to multiplying y_t by $\hat{F}_\perp$, whereas, in Model 3, it should be both demeaned and detrended. Thus, as in the case of a single time series, if a deterministic element is present but not accounted for when constructing the test statistics, the ensuing unit root test will be inconsistent. This is the reason for using $y_t - \hat{\gamma}d_t$ rather than y_t when constructing the statistics.

Remark 4. The relaxation of Assumption 1 implies that PD_t and PD_ρ may be both heteroskedastic and serially correlated. This imply that the statistics are no longer asymptotically similar and that they need to be modified to account for the temporal dependencies of the DGP. To facilitate this, however, Assumption 1 must be replaced with something else. Specifically, it is assumed that the linear process conditions of Phillips and Solo (1992) are satisfied, which allow for a very general class of errors including all stationary autoregressive moving average processes. Under these conditions, similar statistics may be obtained by simply augmenting the right-hand side of (1) with lagged values of Δy_t. This suggests that similar statistics may be obtained by replacing $y_t - \hat{\gamma}d_t$ with $y_t - \hat{\gamma}d_t - \sum_{k=1}^{K}\hat{\delta}_k\Delta y_{t-k}$, where K should be sufficiently large to whiten the remaining error.

4 Asymptotic distribution

In this section, we characterize the asymptotic distributions of the proposed test statistics. In so doing, we shall exploit the fact that the orthogonalization procedure applied here is consistent passing $T \longrightarrow \infty$ for a fixed N. This means that the limiting distributions of the tests may be derived in a relatively straightforward fashion using the simple sequential limit theory. In particular, it will be shown that both statistics require standardization based on the first two moments of the following vector Brownian motion functional

$$K_i \equiv \left(\int_0^1 W_i^{*2}, \ \int_0^1 W_i^* \mathrm{d}W_i \right)', \tag{10}$$

where

$$W_i^* = W_i - \left(\int_0^1 W_i d' \right) \left(\int_0^1 dd' \right)^{-1} d.$$

The vector functional K_i takes the scalar Browninan motion W_i^* as its only argument. This scalar is the Hilbert projection of W_i onto the space orthogonal to the vector d, which is the limiting trend function. Specifically, let $D_T = \text{diag}(1, T)$ denote the matrix of normalizing orders that is conformable with $d_t = (1, t)'$, then $D_T^{-1} d_{[Tr]} \Rightarrow d = (1, r)'$ as $T \longrightarrow \infty$. In Model 1, $d_t = 0$ so W_i^* reduces to W_i. In Model 2, we have $d_t = 1$ so W_i^* is the demeaned standard Brownian motion $W_i - \int_0^1 W_i$. Similarly, in Model 3, $d_t = (1, t)'$ in which case W_i^* represents the demeaned and detrended standard Brownian motion $W_i + (6r - 4) \int_0^1 W_i + (6 - 12r) \int_0^1 r W_i$. Now, Theorem 1 provides the asymptotic distributions of our test statistics. To this end, it is useful to define Θ and Σ as the mean and the variance of K_i, respectively. In addition, it is useful to let $\psi = (-\Theta_2 \Theta_1^{-2}, \Theta_1^{-1})'$ and $\varphi = (-(1/2)\Theta_2 \Theta_1^{-3/2}, \Theta_1^{-1/2})'$. The theorem is given next.

Theorem 1. (Asymptotic distribution.) Under Assumption 1 and 2, and the local alternative hypothesis in (2), as $T \longrightarrow \infty$ followed by $N \longrightarrow \infty$

$$TN^{1/2} PD_\rho - N^{1/2} \Theta_1 \Theta_2^{-1} \Rightarrow N(-\mu, \psi' \Sigma \psi), \tag{11}$$

$$PD_t - N^{1/2} \Theta_1 \Theta_2^{-1/2} \Rightarrow N(-\mu \Theta_1^{1/2}, \varphi' \Sigma \varphi). \tag{12}$$

Remark 5. The proof of Theorem 1 is outlined in the appendix. It proceeds by showing that the intermediate limiting distributions passing $T \longrightarrow \infty$ with N held fixed can be written entirely in terms of the elements of the vector Brownian motion functional K_i. Therefore, by virtue of cross-sectional independence of the defactored data, the limiting distribution of the test statistics can be described in terms of differentiable functions of i.i.d. random variables to which the Delta method is applicable. Hence, by subsequently passing $N \longrightarrow \infty$, we obtain a limiting normal distribution for the test statistics, which depend only on the first two moments of K_i.

Remark 6. Notice the difference between Theorem 1 and the asymptotic theory for univariate time series, which typically involves moments of diffusion processes rather than standard Brownian motions. In particular, Theorem 1 shows that the statistics are asymptotically normal under both the null and alternative hypotheses. Under the null, we have $\mu = 0$ in which case Theorem 1 shows that the distributions are mean zero and that they only depend on the moments of K_i. On the other hand, when $\mu > 0$, although still asymptotically normal and similar with respect to the variance of the error process, the distributions depend on the nuisance parameter μ. Theorem 1 shows that this dependency causes a miscentering of the limiting distributions of the test statistics. The expected value of these distributions are negative, which imply that they will tend to shift leftwards as we move away from the null hypothesis. This means that the tests are unbiased and that their asymptotic local powers therefore are greater than their size. The drift of the distributions depend on the average of the deviations ϕ_i. Thus, the statistics will tend to diverge towards negative infinity as ϕ_i grows arbitrarily large for at least some i. For a given value $\phi > 0$, optimal power is obtained when $\phi_i = \phi_j$ for all i and j, which is not unexpected given our pooling approach. Note also that the local power only depends on μ and therefore not on the heterogeneity of the local alternative.

Remark 7. Under a fixed specification of the alternative hypothesis, the value taken by ρ_i do not depend on N or T. In this case, the probability that the statistics take on a more negative value than the critical value provided by the standard normal distribution approaches one asymptotically as $T \longrightarrow \infty$ followed by $N \longrightarrow \infty$. Thus, in spite of the fact that the statistics are pooled, they may be used to construct consistent tests against the heterogenous type of alternative hypothesis considered here.

Remark 8. Theorem 1 indicate that each of the standardized statistics converges to a normal distribution whose moments depend on the underlying vector Brownian motion functional K_i. As we have seen, since it is a relatively straightforward matter to adjust the formulaes in (7) and (8) to account for the impact of temporal dependency, the results of Theorem 1 are quite general and they apply regardless of the deterministic specification of (1). If $d_t = 0$, then d is the empty set and the results of Theorem 1 apply directly to the standard Brownian motion W_i. If $d_t = 1$, the asymptotic distributions still have the same form as in (11) and (12) but now with moments based on the demeaned standard Brownian motion $W_i - \int_0^1 W_i$. Analogously, if $d_t = (1, t)'$, the limiting distributions in (11) and (12) retain their stated forms but involve moments of the demeaned and detrended standard Brownian motion $W_i + (6r - 4) \int_0^1 W_i + (6 - 12r) \int_0^1 rW_i$. Numerical values of these moments are presented in Table 1. They have been obtained by direct calculation using the properties of Brownian motion.

5 Monte Carlo simulations

In this section, we compare and evaluate the small-sample properties of our proposal relative to those of several competing tests. For this purpose, a small set of Monte Carlo experiment were conducted using the following DGP

$$y_{it} = \rho_i y_{it-1} + e_{it}, \qquad (13)$$

$$e_{it} = \lambda \tau_i f_t + v_{it}, \qquad (14)$$

where $y_{i0} = 0$ and $(\tau_i, f_t, v_{it})' \sim N(0, I_3)$. For each experiment, we generate $1,000$ panels with N individual and $T + 50$ time series observations. The first

50 observations for each cross-section is then disregarded to attenuate the effect of the initial condition.

For each model, we consider a small set of experiments, which correspond to different parameterizations of ρ_i and λ. The autoregressive parameter ρ_i determines whether the null hypothesis is true or not. Under the null, $\rho_i = 1$ for all i whereas $\rho_i < 1$ under alternative. Also, to be able to infer the increased power that derives from increasing the size of the panel, the alternative hypothesis is assumed to be fixed and therefore independent of N and T.

The parameter λ controls the relative importance of the common and idiosyncratic disturbances. A larger value of λ represents a greater weight being attached to the common disturbances relative to the idiosyncratic ones. There are two cases. In the first, $\lambda = 0$ so there is no common factor and the individuals of the panel are therefore independent. In the second, we set $\lambda = 2$, which represents a situation when the common disturbances are twice as important as the idiosyncratic ones. Thus, in this case, there is a common factor present, which induces correlation among the cross-sectional units.

To be able to evaluate the comparative merit of our tests, we compute nine alternative tests. The first four are constructed under the assumption of cross-sectional independence. They are the Z_{tbar} and $Z_{\tilde{t}bar}$ statistics by Im *et al.* (2003), the $\hat{\varphi}$ statistic by Harris and Tzavalis (1999), and the t_{δ}^{*} statistic by Levin *et al.* (2002). The remaining five statistics, which do not require the independence assumption, are the IPS^{*} statistic by Pesaran (2005), $P_{\hat{e}}$ statistic by Bai and Ng (2004), and the G_{ols}^{++}, Z and P_m statistics by Phillips and Sul (2003).[1]

The Pesaran (2005) test is based on a long autoregression, which has been augmented with cross-sectional averages of lagged levels and differences of the dependent variable, whose presence eliminates the cross-sectional correlation in the case of a single common factor. As mentioned earlier, Phillips and Sul (2003) also presume that there is a single common factor, which can be eliminated using the moment based procedure outlined in Section 3. The authors discuss several statistics that can be used together with this procedure. Among the three statistics considered here, G_{ols}^{++} is based on averaging, while Z and P_m are based on meta analysis, wherein the p-values of each cross-section are combined into a panel test.

The Bai and Ng (2004) test is more general than the other tests considered here, and allows for multiple factors that may be nonstationary. The idea is to eliminate the cross-sectional dependence under the null by projecting the data onto the nonstationary factors. Bai and Ng (2004) show that this can be accomplished by first estimating the factors by using principal components on the differentiated data, which can then be recumulated to retrieve estimates of the nonstationary factors. Due to the consistency of the factor estimates, the resulting projection errors will be cross-sectionally independent and therefore suitable for testing the unit root hypothesis.

[1] The Moon and Perron (2004) test is only valid in the specification with a fitted intercept term and it is therefore excluded.

The tests are computed as follows. For the implementation of the iterated method of moments procedure laid out in Section 3, we set the maximum number of iterations to 50 and the convergence criterion for the loading parameters is set to 0.0001. For sake of comparison, the Bai and Ng (2004) test is constructed using only one factor. Also, all statistics are based on the same choice of lag length and bandwidth. To this end, since there is no obvious choice, we set these parameters equal to the largest integer less than $4(T/100)^{2/9}$. The t_{δ}^{*} test performed very poorly for small bandwidths, and we therefore followed the recommendation of Levin *et al.* (2002) and used the formula $3.21T^{1/3}$. Moreover, to be able to construct the statistics that are based on combining p-values, we need to obtain their asymptotic distributions. To this end, we employ the method of MacKinnon (1994), in which the distributions are approximated by means of response surface regressions. All tests are carried out on the 5% level and all powers are adjusted for size so that each test has the same rejection frequency of 5% under the null hypothesis. All computational work was performed in GAUSS.

Consider first the empirical size of the tests presented in Table 2. All tests should have a rejection rate of 5% when $\tau = 0$. In agreement with this, Table 2 illustrates that the size of most tests generally lies close to the nominal leveleven though there is an overall tendency to over-reject the null hypothesis. When $\tau > 0$, we expect the tests based on the assumption of cross-sectional independence to suffer from size distortions. This is confirmed by Table 2, which indicates that the size can be very unreliable for these tests when the errors admit a common factor structure. The t_{δ}^{*} and $\hat{\varphi}$ tests appear to suffer the greatest with massive distortions that tend to get larger and to become very serious as N increases. As expected, the distortions increase as the degree of cross-sectional correlation increases. The results for the tests that do not maintain the assumption of cross-sectional independence are much more encouraging with only small size distortions in most cases. The results for the Phillips and Sul (2003) tests are particularly good with the proposed tests being only slightly less so.

Next, we continue to the results on the power of the tests, which are presented in Tables 3 and 4. In this case, the results suggest that the $\hat{\varphi}$ test generally is most powerful when $\lambda = 0$ and there is no cross-sectional correlation. As expected, the power is higher the larger is the deviation from the null hypothesis. In addition, it is seen that the power of the tests in Model 1 with no deterministic components is higher than that for Model 2 with a fitted intercept, which in turn is higher than that for Model 3 with fitted intercept and trend terms. This effect is particularly striking for the t_{δ}^{*} test, which has almost no power in excess of its size when ρ_i is lies in the vicinity of one.

When $\lambda > 0$, and the data is cross-sectionally correlated, the power of the proposed tests rises relative to the others. This effect is particularly pronounced when considering those tests that allow for cross-sectional correlation. In addition, in agreement with the simulation results presented by Moon and Perron (4004), we see that the power of all tests decline significantly as the value of λ increases. In particular, the results suggest that the power may be very poor, and practically nonexisting in some cases when the tests are fitted with both intercept and time trend terms. The power usually improves, however, as the size of the panel increases, especially along the T dimension.

6 Conclusions

During the last few years there has been an immense proliferation of research concerned with the problem of testing for unit roots in panel data. Most of these studies assume that the members of the panel are independent. For many empirical applications, however, cross-sectional independence seem like a very restrictive assumption. This apparent shortcoming has recently inspired several authors to develop tests that relax this assumption. However, many of these tests are not very easy to analyze or to implement. Recognizing this shortcoming, this paper proposes two simple panel unit root tests that may be used when the cross-sectional dependence can be described by a single common factor that may have disparate effects on the individual members of the panel. In order to defactor the data, we propose using a version of the procedure recently developed by Phillips and Sul (2003), which is based on estimating the factor loadings by iterated method of moments. Sequential limit arguments reveal that the tests have limiting normal distributions, and that they are unbiased when compared to local heterogeneous alternatives. In our Monte Carlo study, we demonstrate that the tests have reasonable size properties and good power.

Appendix: Mathematical proofs

In this appendix, we prove Theorem 1. For illustrative purposes, we focus on the simple case without any deterministic components. Throughout, we assume that the necessary moment conditions for the relevant central limit theory to apply are satisfied.

Proof of Theorem 1.

In order to derive the asymptotic distributions, we multiply the test statistics by T in which case (7) and (8) can be rewritten as

$$TZ_\rho = \left(T^{-2}\sum_{t=2}^{T} y_{it-1}^{*\prime} y_{it-1}^{*}\right)^{-1} T^{-1}\sum_{t=2}^{T} y_{it-1}^{*\prime}\Delta y_{it}^{*}, \tag{A1}$$

$$Z_t = \left(T^{-2}\sum_{t=2}^{T} y_{it-1}^{*\prime} y_{it-1}^{*}\right)^{-1/2} T^{-1}\sum_{t=2}^{T} y_{it-1}^{*\prime}\Delta y_{it}^{*}. \tag{A2}$$

Under the local alternative hypothesis, we have $\rho_i = 1 - T^{-1}N^{-1/2}\phi_i$, which imply that Δy_{it}^* in (A1) and (A2) can be written as

$$\begin{aligned}\Delta y_{it}^* &= (\rho_i - 1)y_{it-1}^* + e_{it}^* \\ &= N^{-1/2}T^{-1}\phi_i y_{it-1}^* + e_{it}^*,\end{aligned} \tag{A3}$$

where $e_t^* = \hat{F}_\perp e_t$ and $e_t^* = (e_{1t}^*, ..., e_{Nt}^*)'$. Define the scaled quantities $Q_1 \equiv T^{-2}\sum_{t=2}^{T} y_{t-1}^{*\prime} y_{t-1}^*$, $Q_2 \equiv T^{-1}\sum_{t=2}^{T} y_{t-1}^{*\prime} e_t^*$ and $Q_3 \equiv T^{-2}\sum_{t=2}^{T} y_{t-1}^{*\prime}\phi y_{t-1}^*$. Now, by using Theorem 4.4 of Hansen (1992), it is possible to show that $T^{-1/2}\hat{F}_\perp y_{[Tr]} \Rightarrow F_\perp J \equiv J^*$ as $T \longrightarrow \infty$ with N held fixed, where $J = (J_1, ..., J_{N-1})'$ is a $N-1$ dimensional diffusion process that satisfy

$$J_i \equiv \int_0^r e^{\phi_i(r-s)N^{-1/2}} \mathrm{d}B_i = B_i + N^{-1/2}\phi_i \int_0^r e^{\phi_i(r-s)N^{-1/2}} B_i. \tag{A4}$$

But $F_\perp B = F_\perp B_v = W$ as explained in the text. Therefore, the expression in (A4) implies that J^* can be written as

$$J_i^* = W_i + N^{-1/2}\phi_i \int_0^r e^{\phi_i(r-s)N^{-1/2}} W_i. \tag{A5}$$

By using this result, it is straightforward to obtain the following weak limits as $T \longrightarrow \infty$ for a fixed N

$$Q_1 \Rightarrow \int_0^1 J^{*\prime} J^* = \sum_{i=1}^{N-1}\int_0^1 J_i^{*2}, \tag{A6}$$

$$Q_2 \Rightarrow \int_0^1 J^{*\prime}\mathrm{d}W = \sum_{i=1}^{N-1}\int_0^1 J_i^{*}\mathrm{d}W_i, \tag{A7}$$

$$Q_3 \quad \Rightarrow \quad \int_0^1 J^{*\prime}\phi J^* = \sum_{i=1}^{N-1}\phi_i \int_0^1 J_i^{*2}. \tag{A8}$$

Now, since W_i is i.i.d. over the cross-section, (A5) implies that J_i^* must be so too. Thus, the limiting distributions of Q_1, Q_2 and Q_3 passing $T \longrightarrow \infty$ is i.i.d. over the cross-section. Another important implication of (A5) is that $J_i^* \Rightarrow W_i$ as $N \longrightarrow \infty$ and that $J_i^* = W_i$ if $\phi_i = 0$ for a fixed N. Hence, the limit of J_i^* passing $N \longrightarrow \infty$ is the same under both the null and the local alternative hypothesis. This implies that the expected value of each element of the vector $(Q_1, Q_2)'$ is equal to Θ. Furthermore, by using the results of Levin *et al.* (2002), we infer that Σ must be block-diagonal.

To derive the limiting distributions of the test statistics, we shall make use of the Delta method, which provides the limiting distribution for continuously differentiable transformations of i.i.d. vector sequences. In so doing, we substitute Δy_{it}^* from (A3), which means that (A1) and (A2) can be rewritten as

$$TZ_\rho = Q_1^{-1}\left(Q_2 - N^{-1/2}Q_3\right), \tag{A9}$$

$$Z_t = Q_1^{-1/2}\left(Q_2 - N^{-1/2}Q_3\right). \tag{A10}$$

These expressions may be expanded in the following fashion

$$\begin{aligned} TN^{1/2}Z_\rho + \mu - N^{1/2}\Theta_2\Theta_1^{-1} &= \mu + N^{1/2}\left(N^{-1}Q_2 - \Theta_2\right)\left(N^{-1}Q_1\right)^{-1} \\ &+ N^{1/2}\Theta_2\left(\left(N^{-1}Q_1\right)^{-1} - \Theta_1^{-1}\right) \\ &- \left(N^{-1}Q_1\right)^{-1}N^{-1}Q_3, \end{aligned} \tag{A11}$$

$$\begin{aligned} Z_t + \mu\Theta_1^{1/2} - N^{1/2}\Theta_2\Theta_1^{-1/2} &= \mu\Theta_1^{1/2} + N^{1/2}\left(N^{-1}Q_2 - \Theta_2\right)\left(N^{-1}Q_1\right)^{-1/2} \\ &+ N^{1/2}\Theta_2\left(\left(N^{-1}Q_1\right)^{-1/2} - \Theta_1^{-1/2}\right) \\ &- \left(N^{-1}Q_1\right)^{-1/2}N^{-1}Q_3. \end{aligned} \tag{A12}$$

The terms appearing in (A11) and (A12) with normalizing order N^{-1} converge in probability to the means of the corresponding random variables by virtue of a law of large numbers as $T \longrightarrow \infty$ and then $N \longrightarrow \infty$. Hence, $N^{-1}Q_1 \xrightarrow{p} \Theta_1$ and $N^{-1}Q_2 \xrightarrow{p} \Theta_2$. In addition, by Corollary 1 of Phillips and Moon (1999), we have $N^{-1}Q_3 \xrightarrow{p} \mu\Theta_1$ as $T \longrightarrow \infty$ prior to N. Moreover, by the Lindberg-Lévy central limit theorem, $N^{1/2}\left(N^{-1}Q_2 - \Theta_2\right) \Rightarrow N(0, \Sigma_{22})$. The remaining expressions involve continuously differentiable transformations of i.i.d. random variables. Thus, by the Delta method, as $T \longrightarrow \infty$ prior to N

$$N^{1/2}\left(\left(N^{-1}Q_1\right)^{-1} - \Theta_1^{-1}\right) \Rightarrow N(0, \Theta_1^{-4}\Sigma_{11}), \tag{A13}$$

$$N^{1/2}\left(\left(N^{-1}Q_1\right)^{-1/2} - \Theta_1^{-1/2}\right) \Rightarrow N(0, (1/4)\Theta_1^{-3}\Sigma_{11}). \tag{A14}$$

Together, these results imply that the expressions appearing on the right hand side of equations (A11) and (A12) are mean zero with variance $\Theta_1^{-4}\Theta_2^2\Sigma_{11} +$

$\Theta_1^{-2}\Sigma_{22}$ and $(1/4)\Theta_1^{-3}\Theta_2^2\Sigma_{11} + \Theta_1^{-1}\Sigma_{22}$, respectively. This establishes the required results. ∎

Table 1: Asymptotic moments

Model	$\Theta_2\Theta_1^{-1}$	$\phi'\Sigma\phi$	$\Theta_2\Theta_1^{-1/2}$	$\varphi'\Sigma\varphi$
1	0	2	0	1
2	-3	51/5	$-\sqrt{3/2}$	8/10
3	$-15/2$	2895/112	$-\sqrt{15/4}$	277/448

Notes: Model 1 refers to the regression without any deterministic intercepts or trends, Model 2 refers to the regression with a fitted intercept and Model 3 refers to the regression with fitted intercept and trend terms.

Table 2: Size on the 5% level

Model	N	T	PD_t	PD_ρ	IPS^*	Z_{tbar}	$Z_{\tilde{t}bar}$	t^*_δ	$\hat{\varphi}$	$P_{\hat{e}}$	G^{++}_{ols}	Z	P_m
$\lambda = 0$													
1	5	50	0.088	0.155	0.113	0.026	0.021	0.074	0.147	0.078	0.016	0.056	0.070
1	10	50	0.108	0.155	0.076	0.037	0.032	0.080	0.120	0.109	0.038	0.067	0.091
1	5	100	0.066	0.148	0.123	0.034	0.031	0.060	0.150	0.080	0.015	0.045	0.068
1	10	100	0.094	0.140	0.084	0.056	0.052	0.076	0.122	0.090	0.033	0.065	0.087
2	5	50	0.110	0.157	0.253	0.045	0.024	0.170	0.156	0.099	0.009	0.093	0.115
2	10	50	0.152	0.152	0.170	0.037	0.013	0.256	0.123	0.112	0.032	0.125	0.162
2	5	100	0.097	0.161	0.205	0.037	0.023	0.120	0.133	0.107	0.005	0.083	0.100
2	10	100	0.127	0.150	0.139	0.035	0.026	0.141	0.105	0.072	0.020	0.088	0.114
3	5	50	0.153	0.147	0.223	0.062	0.011	0.644	0.101	0.121	0.006	0.114	0.138
3	10	50	0.244	0.230	0.135	0.050	0.007	0.798	0.095	0.140	0.046	0.187	0.213
3	5	100	0.114	0.145	0.175	0.049	0.024	0.384	0.096	0.099	0.005	0.089	0.109
3	10	100	0.160	0.187	0.087	0.042	0.020	0.578	0.092	0.123	0.024	0.125	0.151
$\lambda = 2$													
1	5	50	0.061	0.148	0.158	0.079	0.070	0.156	0.326	0.073	0.017	0.052	0.073
1	10	50	0.064	0.115	0.132	0.135	0.128	0.225	0.451	0.085	0.027	0.057	0.077
1	5	100	0.086	0.165	0.141	0.101	0.094	0.170	0.357	0.099	0.018	0.059	0.073
1	10	100	0.073	0.117	0.104	0.145	0.138	0.258	0.450	0.088	0.025	0.054	0.077
2	5	50	0.092	0.144	0.271	0.094	0.058	0.230	0.310	0.110	0.006	0.070	0.098
2	10	50	0.096	0.127	0.231	0.140	0.099	0.386	0.426	0.103	0.030	0.091	0.120
2	5	100	0.076	0.126	0.248	0.090	0.079	0.162	0.293	0.115	0.005	0.069	0.094
2	10	100	0.077	0.112	0.208	0.141	0.115	0.260	0.418	0.085	0.010	0.065	0.097
3	5	50	0.121	0.146	0.296	0.140	0.061	0.626	0.289	0.137	0.003	0.108	0.121
3	10	50	0.122	0.133	0.222	0.181	0.081	0.697	0.367	0.147	0.020	0.139	0.166
3	5	100	0.094	0.144	0.182	0.122	0.080	0.445	0.289	0.118	0.002	0.091	0.114
3	10	100	0.086	0.131	0.182	0.166	0.122	0.538	0.355	0.109	0.018	0.110	0.135

Notes: The value λ refers to the importance of the common relative to the idiosyncratic disturbances. See Table 1 for an explanation of the various models.

Table 3: Size-adjusted power on the 5% level when $\rho_i \sim U(0.9, 1)$

Model	N	T	PD_t	PD_ρ	IPS^*	Z_{tbar}	$Z_{\tilde{t}bar}$	t_δ^*	$\hat{\varphi}$	$P_{\hat{e}}$	G_{ols}^{++}	Z	P_m
$\lambda = 0$													
1	5	50	0.477	0.503	0.168	0.599	0.605	0.619	0.674	0.278	0.476	0.425	0.350
1	10	50	0.664	0.720	0.165	0.788	0.798	0.777	0.823	0.420	0.690	0.634	0.487
1	5	100	0.777	0.792	0.318	0.894	0.893	0.844	0.859	0.650	0.834	0.781	0.704
1	10	100	0.915	0.935	0.497	0.988	0.988	0.941	0.967	0.882	0.957	0.910	0.877
2	5	50	0.144	0.230	0.079	0.141	0.141	0.068	0.312	0.096	0.133	0.133	0.098
2	10	50	0.205	0.403	0.138	0.238	0.250	0.066	0.557	0.161	0.219	0.164	0.135
2	5	100	0.267	0.543	0.195	0.342	0.366	0.053	0.670	0.204	0.338	0.258	0.231
2	10	100	0.484	0.782	0.262	0.666	0.670	0.064	0.870	0.403	0.561	0.464	0.336
3	5	50	0.066	0.096	0.056	0.076	0.073	0.076	0.108	0.056	0.066	0.054	0.061
3	10	50	0.080	0.103	0.096	0.118	0.130	0.100	0.141	0.078	0.114	0.086	0.079
3	5	100	0.140	0.218	0.123	0.166	0.173	0.093	0.318	0.120	0.160	0.156	0.133
3	10	100	0.183	0.307	0.169	0.300	0.311	0.079	0.470	0.163	0.255	0.202	0.171
$\lambda = 2$													
1	5	50	0.424	0.472	0.137	0.408	0.409	0.464	0.387	0.215	0.403	0.354	0.255
1	10	50	0.608	0.663	0.132	0.462	0.468	0.532	0.393	0.365	0.591	0.520	0.433
1	5	100	0.539	0.602	0.289	0.678	0.685	0.656	0.640	0.428	0.626	0.558	0.509
1	10	100	0.770	0.791	0.400	0.834	0.834	0.787	0.656	0.708	0.802	0.776	0.652
2	5	50	0.106	0.247	0.076	0.097	0.105	0.068	0.185	0.074	0.098	0.098	0.067
2	10	50	0.136	0.303	0.064	0.117	0.130	0.068	0.186	0.135	0.136	0.131	0.109
2	5	100	0.197	0.387	0.144	0.215	0.212	0.063	0.385	0.132	0.212	0.187	0.139
2	10	100	0.303	0.540	0.176	0.305	0.316	0.065	0.395	0.248	0.378	0.362	0.247
3	5	50	0.075	0.097	0.039	0.056	0.058	0.052	0.074	0.069	0.054	0.076	0.052
3	10	50	0.070	0.105	0.053	0.078	0.082	0.054	0.095	0.068	0.056	0.074	0.054
3	5	100	0.096	0.138	0.106	0.093	0.092	0.064	0.157	0.080	0.114	0.097	0.077
3	10	100	0.153	0.248	0.104	0.139	0.138	0.085	0.189	0.126	0.180	0.133	0.095

Notes: See Tables 1 and 2 for an explanation of the various features of the table.

Table 4: Size-adjusted power on the 5% level when $\rho_i \sim U(0.8, 1)$

Model	N	T	PD_t	PD_ρ	IPS^*	Z_{tbar}	$Z_{\tilde{t}bar}$	t^*_δ	$\hat{\varphi}$	$P_{\hat{e}}$	G^{++}_{ols}	Z	P_m
$\lambda = 0$													
1	5	50	0.741	0.781	0.366	0.887	0.885	0.826	0.869	0.533	0.822	0.742	0.671
1	10	50	0.899	0.933	0.482	0.984	0.985	0.943	0.959	0.818	0.948	0.905	0.839
1	5	100	0.890	0.909	0.716	0.996	0.996	0.935	0.954	0.890	0.953	0.921	0.900
1	10	100	0.954	0.970	0.915	1.000	1.000	0.973	0.988	0.991	0.994	0.978	0.974
2	5	50	0.265	0.480	0.128	0.331	0.334	0.070	0.638	0.176	0.298	0.244	0.173
2	10	50	0.419	0.705	0.256	0.555	0.568	0.091	0.873	0.356	0.534	0.350	0.250
2	5	100	0.550	0.820	0.438	0.756	0.768	0.120	0.899	0.464	0.711	0.567	0.504
2	10	100	0.792	0.948	0.736	0.965	0.964	0.201	0.981	0.872	0.919	0.827	0.736
3	5	50	0.177	0.285	0.103	0.189	0.197	0.139	0.347	0.121	0.193	0.142	0.128
3	10	50	0.204	0.313	0.129	0.259	0.281	0.181	0.498	0.146	0.309	0.195	0.158
3	5	100	0.365	0.531	0.278	0.514	0.510	0.209	0.744	0.302	0.498	0.377	0.323
3	10	100	0.532	0.768	0.516	0.782	0.788	0.300	0.918	0.529	0.771	0.579	0.499
$\lambda = 2$													
1	5	50	0.635	0.688	0.310	0.713	0.714	0.701	0.639	0.479	0.666	0.615	0.523
1	10	50	0.778	0.807	0.291	0.794	0.793	0.768	0.671	0.639	0.798	0.755	0.702
1	5	100	0.751	0.792	0.640	0.917	0.918	0.854	0.858	0.740	0.836	0.822	0.795
1	10	100	0.851	0.868	0.730	0.982	0.981	0.898	0.861	0.886	0.905	0.891	0.860
2	5	50	0.202	0.424	0.126	0.216	0.221	0.067	0.361	0.153	0.263	0.203	0.153
2	10	50	0.309	0.572	0.134	0.243	0.250	0.045	0.403	0.290	0.385	0.313	0.236
2	5	100	0.386	0.623	0.385	0.520	0.517	0.089	0.695	0.377	0.548	0.445	0.375
2	10	100	0.519	0.768	0.512	0.696	0.704	0.085	0.693	0.673	0.722	0.676	0.583
3	5	50	0.085	0.176	0.056	0.086	0.085	0.069	0.161	0.089	0.094	0.088	0.059
3	10	50	0.125	0.250	0.097	0.118	0.131	0.076	0.191	0.104	0.164	0.148	0.091
3	5	100	0.190	0.319	0.258	0.269	0.270	0.134	0.437	0.190	0.333	0.242	0.211
3	10	100	0.300	0.500	0.275	0.336	0.331	0.151	0.423	0.352	0.476	0.362	0.266

Notes: See Tables 1 and 2 for an explanation of the various features of the table.

References

Bai, J., and S. Ng (2004). A Panic Attack on Unit Roots and Cointegration, *Econometrica*, Vol. 72, pp. 1127-1177.

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

Chang, Y. (2004). Bootstrap Unit Root Test in Panels with Cross-Sectional Dependency, *Journal of Econometrics*, Vol. 120, pp. 263-293.

Choi, I. (2002). Unit Root Tests for Panel Data, *Journal of International Money and Finance*, Vol. 20, pp. 249-272.

Hansen, B. (1992). Convergence to Stochastic Integrals for Dependent Heterogeneous Processes, *Econometric Theory*, Vol. 8, pp. 489-500.

Harris, R. D. F., E. Tzavalis (1999). Inference for Unit Roots in Dynamic Panels where the Time Dimension is Fixed, *Journal of Econometrics*, Vol. 91, pp. 201-226.

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

Jönsson, K. (2004). Cross-Sectional Dependency and Size Distortion in a Small-Sample Homogenous Panel Data Unit Root Test, *Oxford Bulletin of Economics and Statistics*, Vol.67, pp. 369-392.

Levin, A., C. F. Lin and J. Chu (2002). Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties, *Journal of Econometrics*, Vol. 98, pp. 1-24.

Maddala, G. S., and S. Wu (1999). A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics*, Vol. 61, pp. 631-652.

Moon, H. R., and B. Perron (2004). Testing for a Unit Root in Panels with Dynamic Factors, *Journal of Econometrics*, Vol. 122, pp. 81-126.

O'Connell, P. (1998). The Overvaluation of Purchasing Power Parity, *Journal of International Economics*, Vol. 44, pp. 1-19.

Pesaran, M. H. (2003). A Simple Panel Data Unit Root Test in the Presence of Cross Section Dependence, Cambridge Working Papers in Economics 0346, University of Cambridge.

Phillips, P. C. B., and D. Sul (2003). Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence, *Econometrics Journal*, Vol. 6, pp. 217-259.

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

Phillips, P. C. B., and V. Solo (1992). Asymptotics for Linear Processes, *Annals of Statistics*, Vol. 20, pp. 971-1001.

Smith, V. L., S. Leybourne, T-H. Kim and P. Newbold (2004). More Powerful Panel Data Unit Root Tests with an Application to Mean Reversion in Real Exchange Rates, *Journal of Applied Econometrics*, Vol. 19, pp. 147-170.