Available online at www.sciencedirect.com
SCIENCE DIRECT
JOURNAL OF Econometrics
Journal of Econometrics 115 (2003) 53 – 74
www.elsevier.com/locate/econbase

Testing for unit roots in heterogeneous panels☆

Kyung So Im[a,*], M. Hashem Pesaran[b], Yongcheol Shin[c]

[a] *Department of Economics, University of Central Florida, P.O. Box 161400, Orlando, FL 32816-1400, USA*
[b] *Trinity College, Cambridge CB2 1TQ, UK*
[c] *School of Economics and Management, University of Edinburgh, 50, George Square, Edinburgh EH8 9JY, UK*

Accepted 12 January 2003

Abstract

This paper proposes unit root tests for dynamic heterogeneous panels based on the mean of individual unit root statistics. In particular it proposes a standardized t-bar test statistic based on the (augmented) Dickey–Fuller statistics averaged across the groups. Under a general setting this statistic is shown to converge in probability to a standard normal variate sequentially with T (the time series dimension) $\to \infty$, followed by N (the cross sectional dimension) $\to \infty$. A diagonal convergence result with T and $N \to \infty$ while $N/T \to k$, k being a finite non-negative constant, is also conjectured. In the special case where errors in individual Dickey–Fuller (DF) regressions are serially uncorrelated a modified version of the standardized t-bar statistic is shown to be distributed as standard normal as $N \to \infty$ for a fixed T, so long as $T > 5$ in the case of DF regressions with intercepts and $T > 6$ in the case of DF regressions with intercepts and linear time trends. An exact fixed N and T test is also developed using the simple average of the DF statistics. Monte Carlo results show that if a large enough lag order is selected for the underlying ADF regressions, then the small sample performances of the t-bar test is reasonably satisfactory and generally better than the test proposed by Levin and Lin (Unpublished manuscript, University of California, San Diego, 1993).

JEL classification: C12; C15; C22; C23

Keywords: Heterogeneous dynamic panels; Tests of unit roots; t-bar statistics; Finite sample properties

☆ This is a substantially revised version of the DAE Working Papers Amalgamated Series No. 9526, University of Cambridge.

* Corresponding author.
E-mail addresses: kim@bus.ucf.edu (K.S. Im), mhp1@econ.cam.ac.uk (M.H. Pesaran), y.shin@ed.ac.uk (Y. Shin).

0304-4076/03/$ - see front matter
doi:10.1016/S0304-4076(03)00092-7

1. Introduction

Since the seminal work of Balestra and Nerlove (1966), dynamic models have played an increasingly important role in empirical analysis of panel data in economics. Given the small time dimension of most panels, the emphasis has been put on models with homogeneous dynamics, and until recently little attention has been paid to the analysis of dynamic heterogeneous panels. However, over the past decade a number of important panel data set covering different industries, regions, or countries over relatively long time spans have become available, the most prominent example of which is the Summers and Heston (1991) data. The availability of this type of "pseudo" panels raises the issue of the plausibility of the dynamic homogeneity assumption that underlies the traditional analysis of panel data models, and poses the problem of how best to analyze them. The inconsistency of pooled estimators in dynamic heterogeneous panel models has been demonstrated by Pesaran and Smith (1995), and Pesaran et al. (1996). The present paper builds on these works and considers the problem of testing for unit roots in such pseudo panels.

Panel based unit root tests have been advanced by Quah (1992, 1994) and Levin and Lin (1993, *LL* hereafter). The tests proposed by Quah do not accommodate heterogeneity across groups such as individual specific effects and different patterns of residual serial correlations. LL's test is more generally applicable, allows for individual specific effects as well as dynamic heterogeneity across groups, and requires $N/T \to 0$ as both N (the cross section dimension) and T (the time series dimension) tend to infinity.[1]

Using the likelihood framework, this paper proposes an alternative testing procedure based on averaging individual unit root test statistics for panels. In particular, we propose a test based on the average of (augmented) Dickey–Fuller (Dickey and Fuller, 1979) statistics computed for each group in the panel, which we refer to as the t-bar test. Like the *LL* procedure, our proposed test allows for residual serial correlation and heterogeneity of the dynamics and error variances across groups. Under very general settings this statistic is shown to converge in probability to a standard normal variate sequentially with $T \to \infty$, followed by $N \to \infty$. A diagonal convergence result with T and $N \to \infty$ while $N/T \to k$, k being a finite non-negative constant, is also conjectured.

In the special case where errors in individual Dickey–Fuller (DF) regressions are serially uncorrelated a modified version of the (standardized) t-bar statistic, denoted by $Z_{\tilde{t}bar}$, is shown to be distributed as standard normal as $N \to \infty$ for a fixed T, so long as $T > 5$ in the case of DF regressions with intercepts, and $T > 6$ in the case of DF regressions with intercepts and linear time trends. An exact fixed N and T test is also developed using the simple average of the DF statistics. Based on stochastic simulations it is shown that the standardized t-bar statistic provides an excellent approximation to the exact test even for relatively small values of N.

[1] Since the release of the working paper version of this paper in 1995 a number of other approaches to unit root testing in heterogeneous panels have also been proposed in the literature. See, for example, Bowman (1999), Choi (2001), Hadri (2000), Maddala and Wu (1999), and Shin and Snell (2000). A review of this literature is provided in Baltagi and Kao (2000).

Mean and variance of the individual t statistics needed for the implementation of the test for N and T large are also provided. The finite sample performances of the proposed t-bar test and the LL test are examined using Monte Carlo methods. The simulation results clearly show that if a large enough lag order is selected for the underlying ADF regressions, then the finite sample performance of the t-bar test is reasonably satisfactory and generally better than that of the LL test.

The plan of the paper is as follows. Section 2 sets out the model and provides a brief review of the previous studies. Section 3 sets out the likelihood framework for heterogeneous panels, and derives test statistics in the case where errors in individual Dickey–Fuller regressions are serially uncorrelated. Section 4 considers a more general case with serially correlated errors. Section 5 presents the Monte Carlo evidence. Section 6 provides some concluding remarks.

2. The basic framework

Consider a sample of N cross sections (industries, regions or countries) observed over T time periods. We suppose that the stochastic process, y_{it}, is generated by the first-order autoregressive process:

$$y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + \varepsilon_{it}, \quad i = 1,\ldots,N, \ t = 1,\ldots,T, \tag{2.1}$$

where initial values, y_{i0}, are given. We are interested in testing the null hypothesis of unit roots $\phi_i = 1$ for all i. (2.1) can be expressed as

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \varepsilon_{it}, \tag{2.2}$$

where $\alpha_i = (1-\phi_i)\mu_i$, $\beta_i = -(1-\phi_i)$ and $\Delta y_{it} = y_{it} - y_{i,t-1}$. The null hypothesis of unit roots then becomes

$$H_0 : \beta_i = 0 \quad \text{for all } i, \tag{2.3}$$

against the alternatives,

$$H_1 : \beta_i < 0, \quad i = 1,2,\ldots,N_1, \quad \beta_i = 0, \quad i = N_1+1, N_1+2,\ldots,N. \tag{2.4}$$

This formulation of the alternative hypothesis allows for β_i to differ across groups, and is more general than the homogeneous alternative hypothesis, namely $\beta_i = \beta < 0$ for all i, which is implicit in the testing approach of Quah and Levin and Lin (LL) to be discussed below. It also allows for some (but not all) of the individual series to have unit roots under the alternative hypothesis. Formally we assume that under the alternative hypothesis the fraction of the individual processes that are stationary is non-zero, namely if $\lim_{N\to\infty}(N_1/N) = \delta$, $0 < \delta \leqslant 1$. This condition is necessary for the consistency of the panel unit root tests.

Quah (1994) considers the following simple dynamic panel:

$$y_{it} = \phi y_{i,t-1} + \varepsilon_{it}, \quad i = 1,\ldots,N, \ t = 1,\ldots,T, \tag{2.5}$$

where ε_{it} are independently and identically distributed across i and t with finite variance, σ^2. Under additional conditions that $T = \kappa N$ with $\kappa > 0$, and $c = \lim_{T\to\infty} \mathrm{E}(y_{i0}T^{-1/2}\sum_{t=1}^{T}\varepsilon_{it})$ is the same for all i, Quah shows that under the unit root hypothesis, $\phi = 1$, as $N \to \infty$ and $T \to \infty$,

$$Q_{NT}(c,\sigma^2) = \sqrt{\frac{N}{2}}\, T\left(\hat{\phi}_{NT} - 1 - 2\,\frac{c}{\sigma^2}T^{-3/2}\right) \Rightarrow \mathrm{N}(0,1), \tag{2.6}$$

where $\hat{\phi}_{NT}$ is the pooled OLS estimator of ϕ in (2.5) and "$\Rightarrow$" represents the weak convergence in distribution. $Q_{NT}(c,\sigma^2)$ is of limited practical use as it does not allow for the group specific effects and serially correlated and heterogeneous errors. In another paper Quah (1992) considers the more general case where ε_{it} and y_{i0} follow mixing processes, thus allowing for errors in (2.5) to be serially correlated and to have a certain degree of variance heterogeneity. [2]

Levin and Lin provide a more general testing framework and consider the following three models:

$$\Delta y_{it} = \beta y_{i,t-1} + \alpha_{mi} d_{mt} + u_{it}, \quad i = 1,\ldots,N; \quad t = 1,\ldots,T; \quad m = 1,2,3, \tag{2.7}$$

where d_{mt} contains deterministic variables; $d_{1t} = \{\emptyset\}$, $d_{2t} = \{1\}$, and $d_{3t} = \{1,t\}$. In specifying u_{it}, *LL* also allow for different dynamics across groups (see (4.2) below), and argue that under $\beta = 0$, their proposed statistic weakly converges to a standard normal variate, as $T \to \infty$ and $N \to \infty$ with $N/T \to 0$. [3]

A number of other tests have also been proposed more recently. Baltagi and Kao (2000) provide a review.

3. Fixed T unit root tests for heterogeneous panels with serially uncorrelated errors

In this section we develop our proposed test in the context of the panel data model, (2.1), where the errors are serially uncorrelated but T is fixed. For this purpose we make the following assumption:

Assumption 3.1. ε_{it}, $i = 1,\ldots,N$, $t = 1,\ldots,T$, *in* (2.1) *are independently and normally distributed random variables for all i and t with zero means and finite heterogeneous variances*, σ_i^2.

In this case the relevant Dickey–Fuller (1979) regressions are given by (2.2) with the following pooled log-likelihood function:

$$\ell_{NT}(\beta,\varphi) = \sum_{i=1}^{N}\left\{-\frac{T}{2}\log 2\pi\sigma_i^2 - \frac{1}{2\sigma_i^2}\sum_{t=1}^{T}(\Delta y_{it} - \alpha_i - \beta_i y_{i,t-1})^2\right\}, \tag{3.1}$$

[2] The application of the mixing condition to y_{i0} with respect to i pre-assumes an immutable ordering of the groups, which is unlikely to exist in practice.

[3] However, Bowman (1999, Section 6) argues that the conditions stated by *LL* are too weak to establish their claim.

where $\beta = (\beta_1, \ldots, \beta_N)'$, $\varphi_i = (\alpha_i, \sigma_i^2)'$ and $\varphi = (\varphi_1', \ldots, \varphi_N')'$. Using this likelihood framework one could in principle develop alternative panel unit root tests based on the average of the log-likelihood ratio, Wald or the Lagrange multiplier statistics. Here we follow the time series unit root literature and primarily focus on panel unit root tests based on the average of individual Dickey–Fuller (DF) statistics. The standard DF statistic for the ith group is given by the t-ratio of β_i in the regression of $\Delta \mathbf{y}_i = (\Delta y_{i1}, \Delta y_{i2}, \ldots, \Delta y_{iT})'$ on $\tau_T = (1, 1, \ldots, 1)'$ and $\mathbf{y}_{i,-1} = (y_{i0}, y_{i1}, \ldots, y_{i,T-1})'$. Denoting it by t_{iT} we have

$$t_{iT} = \frac{\hat{\beta}_{iT}(\mathbf{y}_{i,-1}'\mathbf{M}_\tau \mathbf{y}_{i,-1})^{1/2}}{\hat{\sigma}_{iT}} = \frac{\Delta \mathbf{y}_i' \mathbf{M}_\tau \mathbf{y}_{i,-1}}{\hat{\sigma}_{iT}(\mathbf{y}_{i,-1}'\mathbf{M}_\tau \mathbf{y}_{i,-1})^{1/2}}, \tag{3.2}$$

where $\hat{\beta}_{iT}$ is the OLS estimator of β_i, $\mathbf{M}_\tau = \mathbf{I}_T - \tau_T(\tau_T'\tau_T)^{-1}\tau_T'$,

$$\hat{\sigma}_{iT}^2 = \frac{\Delta \mathbf{y}_i' \mathbf{M}_{X_i} \Delta \mathbf{y}_i}{T-2}, \tag{3.3}$$

$\mathbf{M}_{X_i} = \mathbf{I}_T - \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'$, and $\mathbf{X}_i = (\tau_T, \mathbf{y}_{i,-1})$.

For analytical tractability we shall also consider the following simplified version of this statistic:

$$\tilde{t}_{iT} = \frac{\Delta \mathbf{y}_i' \mathbf{M}_\tau \mathbf{y}_{i,-1}}{\tilde{\sigma}_{iT}(\mathbf{y}_{i,-1}'\mathbf{M}_\tau \mathbf{y}_{i,-1})^{1/2}}, \tag{3.4}$$

where

$$\tilde{\sigma}_{iT}^2 = \frac{\Delta \mathbf{y}_i' \mathbf{M}_\tau \Delta \mathbf{y}_i}{T-1}. \tag{3.5}$$

The two statistics, t_{iT} and $\tilde{t}_{iT}$, differ only as far as the choice of the estimator of σ_i^2 is concerned. It is easily seen that both of the estimators, $\hat{\sigma}_{iT}^2$ and $\tilde{\sigma}_{iT}^2$ are consistent under the null hypothesis.

As is well known, under $\beta_i = 0$ both test statistics have the same asymptotic distribution (for a given i and as $T \to \infty$), although they will have different distributional properties for a fixed T. It turns out that for a fixed T and a sufficiently large N, the analysis of panel unit root tests is analytically much more manageable if we base the test on the average of $\tilde{t}_{iT}$, although both procedures yield equivalent panel unit root tests if $T \to \infty$, followed by $N \to \infty$. Therefore, for a fixed T (the focus of this section) we consider the following average statistic:

$$\tilde{t}\text{-}bar_{NT} = \frac{1}{N}\sum_{i=1}^{N} \tilde{t}_{iT}, \tag{3.6}$$

which we shall refer to as the $\tilde{t}$-bar statistic. To establish the asymptotic distribution of this statistic we first note that under $\beta_i = 0$, $\mathbf{y}_{i,-1}$ can be written as

$$\mathbf{y}_{i,-1} = y_{i0}\tau_T + \mathbf{s}_{i,-1}, \tag{3.7}$$

where y_{i0} is a given initial value (fixed or random), $\mathbf{s}_{i,-1} = (s_{i0}, s_{i1}, \ldots, s_{i,T-1})'$, with $s_{it} = \sum_{j=1}^{t} \varepsilon_{ij}$. Also $\mathbf{s}_{i,-1} = \mathbf{H}\varepsilon_i$ where $\mathbf{H}$ is the $T \times T$ matrix given by

$$\mathbf{H} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 0 & 0 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{bmatrix}. \tag{3.8}$$

Using these results in (3.4), under the null hypothesis we have

$$\tilde{t}_{iT} = \frac{\sqrt{T-1}\,\boldsymbol{\upsilon}_i'\mathbf{A}\boldsymbol{\upsilon}_i}{(\boldsymbol{\upsilon}_i'\mathbf{C}\boldsymbol{\upsilon}_i)^{1/2}(\boldsymbol{\upsilon}_i'\mathbf{B}\boldsymbol{\upsilon}_i)^{1/2}}, \qquad \boldsymbol{\upsilon}_i = \frac{\varepsilon_i}{\sigma_i} \sim \mathrm{N}(\mathbf{0}, \mathbf{I}_T), \tag{3.9}$$

where $\varepsilon_i = (\varepsilon_{i1}, \varepsilon_{i2}, \ldots, \varepsilon_{i,T})'$, $\mathbf{A}, \mathbf{B}$ and $\mathbf{C}$ defined by

$$\mathbf{A} = \frac{1}{2}(\mathbf{M}_\tau\mathbf{H} + \mathbf{H}'\mathbf{M}_\tau), \quad \mathbf{B} = \mathbf{H}'\mathbf{M}_\tau\mathbf{H}, \quad \mathbf{C} = \mathbf{M}_\tau$$

are fixed symmetric matrices, $\mathbf{B}$ and $\mathbf{C}$ being semi-positive definite with rank $T-1$. Therefore, by Theorem 1(iii) of Magnus (1990), $\mathrm{E}(\boldsymbol{\upsilon}_i'\mathbf{A}\boldsymbol{\upsilon}_i/\boldsymbol{\upsilon}_i'\mathbf{B}\boldsymbol{\upsilon}_i)^s$ exists for $0 \leqslant s < (T-1)/2$. In particular, the second moments of $\boldsymbol{\upsilon}_i'\mathbf{A}\boldsymbol{\upsilon}_i/\boldsymbol{\upsilon}_i'\mathbf{B}\boldsymbol{\upsilon}_i$ and $\boldsymbol{\upsilon}_i'\mathbf{A}\boldsymbol{\upsilon}_i/\boldsymbol{\upsilon}_i'\mathbf{C}\boldsymbol{\upsilon}_i$ exist if $T > 5$. Hence by Cauchy–Schwarz inequality

$$\mathrm{E}(\tilde{t}_{iT}^2) \leqslant (T-1)\left[\mathrm{E}\left(\frac{\boldsymbol{\upsilon}_i'\mathbf{A}\boldsymbol{\upsilon}_i}{\boldsymbol{\upsilon}_i'\mathbf{B}\boldsymbol{\upsilon}_i}\right)^2\right]^{1/2}\left[\mathrm{E}\left(\frac{\boldsymbol{\upsilon}_i'\mathbf{A}\boldsymbol{\upsilon}_i}{\boldsymbol{\upsilon}_i'\mathbf{C}\boldsymbol{\upsilon}_i}\right)^2\right]^{1/2},$$

and for a fixed T, the second-order moment of $\tilde{t}_{iT}$ also exists if $T > 5$.[4]

The case of $T \to \infty$ requires a different treatment. As is well known, for each i both statistics, t_{iT} and $\tilde{t}_{iT}$, converge to the Dickey–Fuller distribution, η_i, defined by[5]

$$\eta_i = \frac{\frac{1}{2}\{[W_i(1)]^2 - 1\} - W_i(1)\int_0^1 W_i(u)\,du}{\{\int_0^1 [W_i(u)]^2\,du - [\int_0^1 W_i(u)\,\mathrm{d}u]^2\}^{1/2}}, \tag{3.10}$$

where $W_i(u)$ is the standard Brownian motion associated with $\boldsymbol{\upsilon}_i$ defined on $u \in [0,1]$.[6] The problem of the existence of the moments of η_i is addressed by Nabeya (1999)

[4] Note that this condition is sufficient and not necessary for the existence of the moments of the modified t-ratio. In fact since $\tilde{t}_{iT}^2 = (T-1)R_{iT}^2$, where R_{iT}^2 is the squared multiple correlation coefficient of the DF regression (2.1), then all moments of $\tilde{t}_{iT}^2$ exists for a fixed T, assuming that $T > 3$ and the DF regression can be computed.

[5] See, for example, Hamilton (1994, Section 17.4).

[6] This result holds more generally and does not require the normality assumption which is invoked here to deal with the case where T is fixed.

who also provides numerical values for the first six moments of the DF distribution for the three standard specifications of the DF regressions, namely with and without intercepts and linear trends. In particular, for the specification (2.1) he reports $\mathrm{E}(\eta_i) = -1.53296244$ and $Var(\eta_i) = 0.706022$. See Table 4, Model 2 in Nabeya (1999).

The above results now establish the following theorem:

Theorem 3.1. *Under Assumption* 3.1, *and for* $T > 5$, *the individual statistics*, $\tilde{t}_{iT}$, $i = 1, \ldots, N$, *given by* (3.4) *are identically and independently distributed with finite second order moments, and therefore by Lindberg–Levy central limit theorem under the null hypothesis* (2.3) *and as* $N \to \infty$ *the standardized* $\tilde{t}$*-bar statistic*

$$Z_{\tilde{t}bar} = \frac{\sqrt{N}\{\tilde{t}\text{-}bar_{NT} - \mathrm{E}(\tilde{t}_T)\}}{\sqrt{Var(\tilde{t}_T)}}, \tag{3.11}$$

converges to a standard normal variate [*written more compactly as* $Z_{\tilde{t}bar} \overset{N}{\Rightarrow} \mathrm{N}(0,1)$], *where* $\tilde{t}\text{-}bar_{NT}$ *is defined by* (3.6), *and* $\mathrm{E}(\tilde{t}_T)$ *and* $Var(\tilde{t}_T)$ *are respectively the mean and the variance of* $\tilde{t}_{iT}$, *defined by* (3.9).

Although, we have not been able to establish the conditions under which the moments of t_{iT} exist, our simulations using t_{iT} instead of $\tilde{t}_{iT}$ suggest that the same conditions are likely to apply to t_{iT}. In particular, in the case of Dickey–Fuller regressions containing intercepts it appears that the condition $T > 5$ is in fact sufficient for the second moment of t_{iT} to exist. For DF regressions containing both intercepts and linear trends, using Magnus' Theorem it is easily seen that the second moment of $\tilde{t}_{iT}$ exist if $T > 6$. Again the same condition seems to be applicable to t_{iT}.[7] These moments computed for different values of T using stochastic simulations with 50,000 replications are presented in Table 1. As to be expected the estimated first two moments of t_{iT} and $\tilde{t}_{iT}$ converge as T is increased, although they differ for small values of T. The last row of Table 1 also gives the asymptotic moment estimates reported by Nabeya (1999), which are very close to those given for $T = 1000$ in Table 1. However, for small values of T, the use of Nabeya's asymptotic moment values could lead to poor test results.

Remark 3.1. In the more general case where T_i differs across groups, $\tilde{t}_{iT_i}$ are independently but not identically distributed across i. So long as $T_i > 9$ so that the third moment of $\tilde{t}_{iT_i}$ exists, by the Lyapunov central limit theorem, we have

$$Z_{\tilde{t}bar} = \frac{\sqrt{N}\{\tilde{t}\text{-}bar_{NT} - N^{-1}\sum_{i=1}^{N}\mathrm{E}(\tilde{t}_{T_i})\}}{\sqrt{N^{-1}\sum_{i=1}^{N} Var(\tilde{t}_{T_i})}} \overset{N}{\Rightarrow} \mathrm{N}(0,1). \tag{3.12}$$

For future reference the corresponding standardized t-bar statistic is given by

$$Z_{tbar} = \frac{\sqrt{N}\{t\text{-}bar_{NT} - N^{-1}\sum_{i=1}^{N}\mathrm{E}(t_{T_i})\}}{\sqrt{N^{-1}\sum_{i=1}^{N} Var(t_{T_i})}}, \tag{3.13}$$

[7] Detailed stochastic simulation results are available upon request.

Table 1
Moments of the Individual $\tilde{t}_{iT}$ and t_{iT} Statistics[a]

T	Moments of $\tilde{t}_{iT}$		Moments of t_{iT}	
	$\mathrm{E}(\tilde{t}_T)$	$Var(\tilde{t}_T)$	$\mathrm{E}(t_T)$	$Var(t_T)$
6	−1.125	0.497	−1.520	1.745
7	−1.178	0.506	−1.514	1.414
8	−1.214	0.506	−1.501	1.228
9	−1.244	0.527	−1.501	1.132
10	−1.274	0.521	−1.504	1.069
15	−1.349	0.565	−1.514	0.923
20	−1.395	0.592	−1.522	0.851
25	−1.423	0.609	−1.520	0.809
30	−1.439	0.623	−1.526	0.789
40	−1.463	0.639	−1.523	0.770
50	−1.477	0.656	−1.527	0.760
100	−1.504	0.683	−1.532	0.735
500	−1.526	0.704	−1.531	0.715
1000	−1.526	0.702	−1.529	0.707
∞[b]	−1.533	0.706	−1.533	0.706

[a]Based on 50,000 replications.
[b]From Table 4, Model 2 in Nabeya (1999).

where

$$t\text{-}bar_{NT} = \frac{1}{N}\sum_{i=1}^{N} t_{iT_i}. \tag{3.14}$$

3.1. A fixed N and T test

When N and T are both fixed, the sample distribution of t-bar statistic, $t\text{-}bar_{NT}$, under the null hypothesis (2.3) is non-standard, but does not depend on any nuisance parameters. Exact sample critical values for the t-bar statistic in this case can be computed via stochastic simulation. The results, using 50,000 replications, are summarized in Table 2 for different values of N and T, for the case of DF regressions with and without linear time trends. The test based on the standardized t-bar statistic, Z_{tbar}, is equivalent to that based on $t\text{-}bar_{NT}$ when N is sufficiently large, but in principle could differ from it for a fixed N. However, our computations suggest that the two testing procedures have very similar performance even for relatively small values of N. To see this note that

$$\Pr(Z_{tbar} < z_\alpha) = \Pr(t\text{-}bar_{NT} < c_T(\alpha)),$$

where z_α is the α percent critical value of the standard normal distribution, and

$$c_T(\alpha) = z_\alpha\sqrt{N^{-1}Var(t_T)} + \mathrm{E}(t_T).$$

Table 2
Exact critical values of the $tbar_{NT}$ statistic

$N\backslash T$	5	10	15	20	25	30	40	50	60	70	100
Panel A: DF regressions containing only intercepts											
1 percent											
5	−3.79	−2.66	−2.54	−2.50	−2.46	−2.44	−2.43	−2.42	−2.42	−2.40	−2.40
7	−3.45	−2.47	−2.38	−2.33	−2.32	−2.31	−2.29	−2.28	−2.28	−2.28	−2.27
10	−3.06	−2.32	−2.24	−2.21	−2.19	−2.18	−2.16	−2.16	−2.16	−2.16	−2.15
15	−2.79	−2.14	−2.10	−2.08	−2.07	−2.05	−2.04	−2.05	−2.04	−2.04	−2.04
20	−2.61	−2.06	−2.02	−2.00	−1.99	−1.99	−1.98	−1.98	−1.98	−1.97	−1.97
25	−2.51	−2.01	−1.97	−1.95	−1.94	−1.94	−1.93	−1.93	−1.93	−1.93	−1.92
50	−2.20	−1.85	−1.83	−1.82	−1.82	−1.82	−1.81	−1.81	−1.81	−1.81	−1.81
100	−2.00	−1.75	−1.74	−1.73	−1.73	−1.73	−1.73	−1.73	−1.73	−1.73	−1.73
5 percent											
5	−2.76	−2.28	−2.21	−2.19	−2.18	−2.16	−2.16	−2.15	−2.16	−2.15	−2.15
7	−2.57	−2.17	−2.11	−2.09	−2.08	−2.07	−2.07	−2.06	−2.06	−2.06	−2.05
10	−2.42	−2.06	−2.02	−1.99	−1.99	−1.99	−1.98	−1.98	−1.97	−1.98	−1.97
15	−2.28	−1.95	−1.92	−1.91	−1.90	−1.90	−1.90	−1.89	−1.89	−1.89	−1.89
20	−2.18	−1.89	−1.87	−1.86	−1.85	−1.85	−1.85	−1.85	−1.84	−1.84	−1.84
25	−2.11	−1.85	−1.83	−1.82	−1.82	−1.82	−1.81	−1.81	−1.81	−1.81	−1.81
50	−1.95	−1.75	−1.74	−1.73	−1.73	−1.73	−1.73	−1.73	−1.73	−1.73	−1.73
100	−1.84	−1.68	−1.67	−1.67	−1.67	−1.67	−1.67	−1.67	−1.67	−1.67	−1.67
10 percent											
5	−2.38	−2.10	−2.06	−2.04	−2.04	−2.02	−2.02	−2.02	−2.02	−2.02	−2.01
7	−2.27	−2.01	−1.98	−1.96	−1.95	−1.95	−1.95	−1.95	−1.94	−1.95	−1.94
10	−2.17	−1.93	−1.90	−1.89	−1.88	−1.88	−1.88	−1.88	−1.88	−1.88	−1.88
15	−2.06	−1.85	−1.83	−1.82	−1.82	−1.82	−1.81	−1.81	−1.81	−1.81	−1.81
20	−2.00	−1.80	−1.79	−1.78	−1.78	−1.78	−1.78	−1.78	−1.78	−1.77	−1.77
25	−1.96	−1.77	−1.76	−1.75	−1.75	−1.75	−1.75	−1.75	−1.75	−1.75	−1.75
50	−1.85	−1.70	−1.69	−1.69	−1.69	−1.69	−1.68	−1.68	−1.68	−1.68	−1.69
100	−1.77	−1.64	−1.64	−1.64	−1.64	−1.64	−1.64	−1.64	−1.64	−1.64	−1.64
Panel B: DF regressions containing intercepts and linear time trends											
5	−8.12	−3.42	−3.21	−3.13	−3.09	−3.05	−3.03	−3.02	−3.00	−3.00	−2.99
7	−7.36	−3.20	−3.03	−2.97	−2.94	−2.93	−2.90	−2.88	−2.88	−2.87	−2.86
10	−6.44	−3.03	−2.88	−2.84	−2.82	−2.79	−2.78	−2.77	−2.76	−2.75	−2.75
15	−5.72	−2.86	−2.74	−2.71	−2.69	−2.68	−2.67	−2.65	−2.66	−2.65	−2.64
20	−5.54	−2.75	2.67	−2.63	−2.62	−2.61	−2.59	−2.60	−2.59	−2.58	−2.58
25	−5.16	−2.69	−2.61	−2.58	−2.58	−2.56	−2.55	−2.55	−2.55	−2.54	−2.54
50	−4.50	−2.53	−2.48	−2.46	−2.45	−2.45	−2.44	−2.44	−2.44	−2.44	−2.43
100	−4.00	−2.42	−2.39	−2.38	−2.37	−2.37	−2.36	−2.36	−2.36	−2.36	−2.36
5 percent											
5	−4.66	−2.98	−2.87	−2.82	−2.80	−2.79	−2.77	−2.76	−2.75	−2.75	−2.75
7	−4.38	−2.85	−2.76	−2.72	−2.70	−2.69	−2.68	−2.67	−2.67	−2.66	−2.66
10	−4.11	−2.74	−2.66	−2.63	−2.62	−2.60	−2.60	−2.59	−2.59	−2.58	−2.58
15	−3.88	−2.63	−2.57	−2.55	−2.53	−2.53	−2.52	−2.52	−2.52	−2.51	−2.51
20	−3.73	−2.56	−2.52	−2.49	−2.48	−2.48	−2.48	−2.47	−2.47	−2.46	−2.46
25	−3.62	−2.52	−2.48	−2.46	−2.45	−2.45	−2.44	−2.44	2.44	−2.44	−2.43

Table 2 (*continued*)

$N\backslash T$	5	10	15	20	25	30	40	50	60	70	100
50	−3.35	−2.42	−2.38	−2.38	−2.37	−2.37	−2.36	−2.36	−2.36	−2.36	−2.36
100	−3.13	−2.34	−2.32	−2.32	−2.31	−2.31	−2.31	−2.31	−2.31	−2.31	−2.31
10 percent											
5	−3.73	−2.77	−2.70	−2.67	−2.65	−2.64	−2.63	−2.62	−2.63	−2.62	−2.62
7	−3.60	−2.68	−2.62	−2.59	−2.58	−2.57	−2.57	−2.56	−2.56	−2.55	−2.55
10	−3.45	−2.59	−2.54	−2.52	−2.51	−2.51	−2.50	−2.50	−2.50	−2.49	−2.49
15	−3.33	−2.52	−2.47	−2.46	−2.45	−2.45	−2.44	−2.44	−2.44	−2.44	−2.44
20	−3.26	−2.47	−2.44	−2.42	−2.41	−2.41	−2.41	−2.40	−2.40	−2.40	−2.40
25	−3.18	−2.44	−2.40	−2.39	−2.39	−2.38	−2.38	−2.38	−2.38	−2.38	−2.38
50	−3.02	−2.36	−2.33	−2.33	−2.33	−2.32	−2.32	−2.32	−2.32	−2.32	−2.32
100	−2.90	−2.30	−2.29	−2.28	−2.28	−2.28	−2.28	−2.28	−2.28	−2.28	−2.28

Notes: The critical values reported in this table are computed via stochastic simulations with 50,000 replications. The $tbar_{NT}$ statistic defined in (3.14) is the sample average of the t-statistics obtained from individual DF regressions with and without time trend. The underlying data are generated by $y_{it} = y_{i,t-1} + e_{it}$, $e_{it} \sim \mathrm{N}(0,1)$, $t = 1,2,\ldots,T$, $i = 1,2,\ldots,N$, with $y_{i0} = 0$.

Hence, the tests based on Z_{tbar} and $t\text{-}bar_{NT}$ differ only to the extent that $c_T(\alpha)$ is different from the corresponding exact critical value of the t-bar test reported in Table 2. Using the mean and the variance of t_T given in Table 1 for $T = 10$ we have

	Approximate 5% critical values		Exact 5% critical values	
N	Without trend	With trend	Without trend	With trend
2	−2.71	−3.40	−2.73	−3.47
3	−2.49	−3.18	−2.50	−3.22
4	−2.35	−3.04	−2.38	−3.06
5	−2.26	−2.95	−2.28	−2.98
6	−2.20	−2.88	−2.22	−2.91
7	−2.15	−2.83	−2.17	−2.85
8	−2.11	−2.78	−2.12	−2.81
9	−2.07	−2.75	−2.08	−2.77
10	−2.04	−2.72	−2.06	−2.74

It is clear from these results that the standardized t-bar test provides an excellent approximation to the exact test even for relatively small values of N. Similar results are also obtained if the tests are based on $\tilde{t}_{iT}$ instead of t_{iT}.

4. Unit root tests for heterogeneous panels with serially correlated errors

In this section we consider the more general case where the errors in (2.1) may be serially correlated, possibly with different serial correlation patterns across groups, but with T and N sufficiently large.

Suppose that y_{it} are generated according to the following finite-order AR($p_i + 1$) processes:

$$y_{it} = \mu_i \phi_i(1) + \sum_{j=1}^{p_i+1} \phi_{ij} y_{i,t-j} + \varepsilon_{it}, \quad i = 1,\dots,N, \ t = 1,\dots,T, \tag{4.1}$$

which can be written equivalently as the ADF(p_i) regressions:

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \sum_{j=1}^{p_i} \rho_{ij} \Delta y_{i,t-j} + \varepsilon_{it}, \quad i = 1,\dots,N, \ t = 1,\dots,T, \tag{4.2}$$

where $\phi_i(1) = 1 - \sum_{j=1}^{p_i+1} \phi_{ij}$, $\alpha_i = \mu_i \phi_i(1)$, $\beta_i = -\phi_i(1)$, and $\rho_{ij} = -\sum_{h=j+1}^{p_i+1} \phi_{ih}$. Writing the ADF regressions for each i in matrix notations we have

$$\Delta \mathbf{y}_i = \beta_i \mathbf{y}_{i,-1} + \mathbf{Q}_i \gamma_i + \varepsilon_i, \tag{4.3}$$

where $\mathbf{Q}_i = (\tau_T, \Delta\mathbf{y}_{i,-1}, \Delta\mathbf{y}_{i,-2}, \dots, \Delta\mathbf{y}_{i,-p_i})$ and $\gamma_i = (\alpha_i, \rho_{i1}, \rho_{i2,\dots,}\rho_{ip_i})'$.

We also make the following assumptions:

Assumption 4.1. *All the roots of* $\phi_i(z) = 1 - \sum_{j=1}^{p_i+1} \phi_{ij} z^j = 0$, $i = 1,2,\dots,N$, *fall on or outside the unit circle, while all the roots of* $\rho_i(z) = 1 - \sum_{j=1}^{p_i} \rho_{ij} z^j = 0$, $i = 1,2,\dots,N$, *fall strictly outside the unit circle.*

Assumption 4.2. ε_{it}, $i = 1,2,\dots,N$, $t = 1,2,\dots,T$, *in* (4.2) *are independently distributed as normal variates with zero means and finite (possibly) heterogeneous variances,* σ_i^2, *and the initial values,* $y_{i0}, y_{i,-1},\dots,y_{i,-p_i}$, *are given (either fixed or stochastic).*

As in the previous section the t-bar statistic is formed as a simple average of the individual t statistic for testing $\beta_i = 0$ in (4.2), namely

$$t\text{-}bar_{NT} = \frac{1}{N} \sum_{i=1}^{N} t_{iT}(p_i, \boldsymbol{\rho}_i), \tag{4.4}$$

where $t_{iT}(p_i, \boldsymbol{\rho}_i)$ is given by

$$t_{iT}(p_i, \boldsymbol{\rho}_i) = \frac{\sqrt{T - p_i - 2}(\mathbf{y}_{i,-1}' \mathbf{M}_{Q_i} \Delta \mathbf{y}_i)}{(\mathbf{y}_{i,-1}' \mathbf{M}_{Q_i} \mathbf{y}_{i,-1})^{1/2} (\Delta \mathbf{y}_i' \mathbf{M}_{X_i} \Delta \mathbf{y}_i)^{1/2}}, \tag{4.5}$$

$\boldsymbol{\rho}_i = (\rho_{i1}, \rho_{i2}, \dots, \rho_{ip_i})'$, $\mathbf{M}_{Q_i} = \mathbf{I}_T - \mathbf{Q}_i(\mathbf{Q}_i'\mathbf{Q}_i)^{-1}\mathbf{Q}_i'$, $\mathbf{M}_{X_i} = \mathbf{I}_T - \mathbf{X}_i(\mathbf{X}_i'\mathbf{X}_i)^{-1}\mathbf{X}_i'$, and $\mathbf{X}_i = (\mathbf{y}_{i,-1}, \mathbf{Q}_i)$.

When T is fixed, the individual ADF statistics, $t_{iT}(p_i, \boldsymbol{\rho}_i)$, will depend on the nuisance parameters, $\boldsymbol{\rho}_i$, $i = 1,\dots,p_i$, even under $\beta_i = 0$. Therefore, the standardization using $\mathrm{E}[t_{iT}(p_i, \boldsymbol{\rho}_i)]$ and $Var[t_{iT}(p_i, \boldsymbol{\rho}_i)]$ will not be practical. But when T and N are sufficiently large it is possible to develop asymptotically valid t-bar type panel unit root tests that are free from the nuisance parameters.

To this end first note that under Assumptions 4.1 and 4.2, and as $T \to \infty$ the individual ADF statistics, (4.5), converge to η_i, defined by (3.10).[8] Let

$$x_{iT}(p_i, \boldsymbol{\rho}_i) = \frac{t_{iT}(p_i, \boldsymbol{\rho}_i) - \mathrm{E}(\eta_i)}{\sqrt{Var(\eta_i)}},$$

and consider the following standardized t-bar statistic

$$Z_{tbar}(\mathbf{p}, \boldsymbol{\rho}) = \frac{\sqrt{N}\{t\text{-}bar_{NT} - \mathrm{E}(\eta)\}}{\sqrt{Var(\eta)}} = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} x_{iT}(p_i, \boldsymbol{\rho}_i), \tag{4.6}$$

where $\boldsymbol{\rho} = (\boldsymbol{\rho}_1', \boldsymbol{\rho}_2', \ldots, \boldsymbol{\rho}_N')'$ and $\mathbf{p} = (p_1, p_2, \ldots, p_N)'$. For a fixed N, first let $T \to \infty$ to obtain

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} x_{iT}(p_i, \boldsymbol{\rho}_i) \overset{T}{\Rightarrow} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} x_i,$$

where $x_i = (\eta_i - \mathrm{E}(\eta_i))/\sqrt{Var(\eta_i)}$ is the limiting distribution of x_{iT}, as $T \to \infty$. Then let $N \to \infty$. Since $x_i's$ are identically and independently distributed with zero means and unit variances, by Lindberg–Levy's central limit theorem we have

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} x_i \overset{N}{\Rightarrow} \mathrm{N}(0, 1).$$

This establishes that the standardized t-bar statistic, $Z_{tbar}(\mathbf{p}, \boldsymbol{\rho})$, defined by (4.6) converges in distribution to a standard normal variate *sequentially*, as $T \to \infty$ first and then $N \to \infty$, denoted by $\overset{T,N}{\Rightarrow} \mathrm{N}(0, 1)$.

A number of other sequentially asymptotically equivalent tests can also be obtained. For example, the asymptotic version of the result in Theorem 3.1 can be written as

$$Z_{\tilde{t}bar}(\mathbf{p}, \boldsymbol{\rho}) = \frac{\sqrt{N}\{\tilde{t}\text{-}bar_{NT} - \mathrm{E}(\tilde{t}_T)\}}{\sqrt{Var(\tilde{t}_T)}} \overset{T,N}{\Rightarrow} \mathrm{N}(0, 1), \tag{4.7}$$

where

$$\tilde{t}\text{-}bar_{NT} = \frac{1}{N} \sum_{i=1}^{N} \tilde{t}_{iT}(p_i, \boldsymbol{\rho}_i), \tag{4.8}$$

$$\tilde{t}_{iT}(p_i, \boldsymbol{\rho}_i) = \frac{\sqrt{T - p_i - 1}(\mathbf{y}_{i,-1}' \mathbf{M}_{Q_i} \Delta \mathbf{y}_i)}{(\mathbf{y}_{i,-1}' \mathbf{M}_{Q_i} \mathbf{y}_{i,-1})^{1/2} (\Delta \mathbf{y}_i' \mathbf{M}_{Q_i} \Delta \mathbf{y}_i)^{1/2}}. \tag{4.9}$$

Another alternative of practical relevance would be to carry out the standardization of the t-bar statistic using the means and variances of $t_{iT}(p_i, \mathbf{0})$ evaluated under $\beta_i = 0$. This is likely to yield better approximations, since $\mathrm{E}[t_{iT}(p_i, \mathbf{0}) \mid \beta_i = 0]$, for example, makes use of the information contained in p_i while $\mathrm{E}[t_{iT}(0, \mathbf{0}) \mid \beta_i = 0]$ does not. In

[8] See, for example, Hamilton (1994, Section 17.7).

view of this we propose the alternative standardized t-bar statistic

$$W_{tbar}(\mathbf{p},\boldsymbol{\rho}) = \frac{\sqrt{N}\{t\text{-}bar_{NT} - \frac{1}{N}\sum_{i=1}^{N}\mathrm{E}[t_{iT}(p_i,\mathbf{0})|\beta_i=0]\}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}Var[t_{iT}(p_i,\mathbf{0})|\beta_i=0]}} \overset{T,N}{\Rightarrow} \mathrm{N}(0,1). \tag{4.10}$$

As in Remark 3.1, the preceding analysis can be readily extended to unbalanced panels and/or to dynamic panels with intercepts and linear time trends. Table 3 gives the values of $\mathrm{E}[t_{iT}(p,\mathbf{0})\,|\,\beta_i=0]$ and $Var[t_{iT}(p,\mathbf{0})\,|\,\beta_i=0]$ for different values of T and p, computed via stochastic simulations with 50,000 replications, when the underlying ADF(p) regression is estimated with and without a linear time trend.

The above sequential asymptotic results while useful may not provide an adequate approximation in cases where both N and T tend to infinity simultaneously.[9] We believe it is possible to generalize the asymptotic results of this paper to the case where N and $T \to \infty$ such that $N/T \to k$, for a finite non-negative constant k. The basis of this conjecture is as follows: Consider first the following decomposition of $Z_{\tilde{t}bar}(\mathbf{p},\boldsymbol{\rho})$, defined by (4.7)

$$Z_{\tilde{t}bar}(\mathbf{p},\boldsymbol{\rho}) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\tilde{t}_{iT}(p_i,\boldsymbol{\rho}_i)-\mathrm{E}(\tilde{t}_T)}{\sqrt{Var(\tilde{t}_T)}} = Z_{\tilde{t}bar} + D_{NT}, \tag{4.11}$$

where[10]

$$Z_{\tilde{t}bar} = \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\frac{\tilde{t}_{iT}-\mathrm{E}(\tilde{t}_T)}{\sqrt{Var(\tilde{t}_T)}},$$

$$D_{NT} = \frac{1}{\sqrt{Var(\tilde{t}_T)}}\left(\sqrt{\frac{N}{T}}\right)\left(\frac{\sum_{i=1}^{N}\sqrt{T}d_{iT}(p_i,\boldsymbol{\rho}_i)}{N}\right), \tag{4.12}$$

and

$$d_{iT}(p_i,\boldsymbol{\rho}_i) = \tilde{t}_{iT}(p_i,\boldsymbol{\rho}_i) - \tilde{t}_{iT}.$$

By Theorem 3.1, the first part of (4.11) converges to the standard normal variate as $N \to \infty$ for $T > 5$. Since $Z_{\tilde{t}bar}(\mathbf{p},\boldsymbol{\rho}) \overset{T,N}{\Rightarrow} \mathrm{N}(0,1)$, it also follows that $Z_{\tilde{t}bar} \overset{T,N}{\Rightarrow} \mathrm{N}(0,1)$. Therefore, $Z_{\tilde{t}bar} \Rightarrow \mathrm{N}(0,1)$ irrespective of whether $T \to \infty$ followed by $N \to \infty$, or *vice versa*. Hence, for $Z_{\tilde{t}bar}(\mathbf{p},\boldsymbol{\rho}) \Rightarrow \mathrm{N}(0,1)$ along the diagonal path $N/T \to k$, it is sufficient to establish that under the null hypothesis

$$\operatorname*{plim}_{N,T\to\infty,\,N/T\to k} D_{NT} = 0. \tag{4.13}$$

[9] See, for example, Phillips and Moon (1999) for a discussion of different types of asymptotics in panel data models.

[10] Note that $Z_{\tilde{t}bar} = Z_{\tilde{t}bar}(\mathbf{0},\mathbf{0})$, and $\tilde{t}_{iT} = \tilde{t}_{iT}(0,\mathbf{0})$.

Table 3
Mean and variance of $t_T(p,0)$ in ADF(p) regression

p	T	10	15	20	25	30	40	50	60	70	100
Without time trend											
0	Mean	−1.504	−1.514	−1.522	−1.520	−1.526	−1.523	−1.527	−1.519	−1.524	−1.532
	Variance	1.069	0.923	0.851	0.809	0.789	0.770	0.760	0.749	0.736	0.735
1	Mean	−1.488	−1.503	−1.516	−1.514	−1.519	−1.520	−1.524	−1.519	−1.522	−1.530
	Variance	1.255	1.011	0.915	0.861	0.831	0.803	0.781	0.770	0.753	0.745
2	Mean	−1.319	−1.387	−1.428	−1.443	−1.460	−1.476	−1.493	−1.490	−1.498	−1.514
	Variance	1.421	1.078	0.969	0.905	0.865	0.830	0.798	0.789	0.766	0.754
3	Mean	−1.306	−1.366	−1.413	−1.433	−1.453	−1.471	−1.489	−1.486	−1.495	−1.512
	Variance	1.759	1.181	1.037	0.952	0.907	0.858	0.819	0.802	0.782	0.761
4	Mean	−1.171	−1.260	−1.329	−1.363	−1.394	−1.428	−1.454	−1.458	−1.470	−1.495
	Variance	2.080	1.279	1.097	1.005	0.946	0.886	0.842	0.819	0.801	0.771
5	Mean			−1.313	−1.351	−1.384	−1.421	−1.451	−1.454	−1.467	−1.494
	Variance			1.171	1.055	0.980	0.912	0.863	0.839	0.814	0.781
6	Mean				−1.289	−1.331	−1.380	−1.418	−1.427	−1.444	−1.476
	Variance				1.114	1.023	0.942	0.886	0.858	0.834	0.795
7	Mean				−1.273	−1.319	−1.371	−1.411	−1.423	−1.441	−1.474
	Variance				1.164	1.062	0.968	0.910	0.875	0.851	0.806
8	Mean				−1.212	−1.266	−1.329	−1.377	−1.393	−1.415	−1.456
	Variance				1.217	1.105	0.996	0.929	0.896	0.871	0.818
With time trend											
0	Mean	−2.166	−2.167	−2.168	−2.167	−2.172	−2.173	−2.176	−2.174	−2.174	−2.177
	Variance	1.132	0.869	0.763	0.713	0.690	0.655	0.633	0.621	0.610	0.597
1	Mean	−2.173	−2.169	−2.172	−2.172	−2.173	−2.177	−2.180	−2.178	−2.176	−2.179
	Variance	1.453	0.975	0.845	0.769	0.734	0.687	0.654	0.641	0.627	0.605
2	Mean	−1.914	−1.999	−2.047	−2.074	−2.095	−2.120	−2.137	−2.143	−2.146	−2.158
	Variance	1.627	1.036	0.882	0.796	0.756	0.702	0.661	0.653	0.634	0.613
3	Mean	−1.922	−1.977	−2.032	−2.065	−2.091	−2.117	−2.137	−2.142	−2.146	−2.158
	Variance	2.482	1.214	0.983	0.861	0.808	0.735	0.688	0.674	0.650	0.625
4	Mean	−1.750	−1.823	−1.911	−1.968	−2.009	−2.057	−2.091	−2.103	−2.114	−2.135
	Variance	3.947	1.332	1.052	0.913	0.845	0.759	0.705	0.685	0.662	0.629
5	Mean			−1.888	−1.955	−1.998	−2.051	−2.087	−2.101	−2.111	−2.135
	Variance			1.165	0.991	0.899	0.792	0.730	0.705	0.673	0.638
6	Mean				−1.868	−1.923	−1.995	−2.042	−2.065	−2.081	−2.113
	Variance				1.055	0.945	0.828	0.753	0.725	0.689	0.650
7	Mean				−1.851	−1.912	−1.986	−2.036	−2.063	−2.079	−2.112
	Variance				1.145	1.009	0.872	0.786	0.747	0.713	0.661
8	Mean				−1.761	−1.835	−1.925	−1.987	−2.024	−2.046	−2.088
	Variance				1.208	1.063	0.902	0.808	0.766	0.728	0.670

Notes: The means and variances reported in this table are computed via stochastic simulations with 50,000 replications. The data are generated by $\Delta y_t = e_t,\ e_t \sim \mathrm{N}(0,1)$ for $t=-p+1,-p+2,\ldots,T$, with $y_{-p}=0$. $t_T(p,0)$ is the t-statistic for testing $\beta=0$ in the ADF(p) regression without time trend: $\Delta y_t = \alpha + \beta y_{t-1} + \sum_{j=1}^{p} \rho_j \Delta y_{t-j} + e_t$, and with time trend: $\Delta y_t = \alpha + \delta t + \beta y_{t-1} + \sum_{j=1}^{p} \rho_j \Delta y_{t-j} + e_t,\ t=1,\ldots,T$.

But as shown by Said and Dickey (1984), $d_{iT}(p_i, \boldsymbol{\rho}_i) = O_p(T^{-1/2})$,[11] and it seems reasonable to conjecture that under $\beta_i = 0$ and Assumptions 4.1 and 4.2,

$$E(Td_{iT}(p_i, \boldsymbol{\rho}_i)) = O(1);\ Var(\sqrt{T}d_{iT}(p_i, \boldsymbol{\rho}_i)) = O(1). \tag{4.14}$$

This conjecture is supported by stochastic simulations we have carried out using 100,000 replications in the simple case where $p_i = 1$. The results for the values of $\rho_i = 0.5$, 0.9 and 0.98 for all i are summarized below:

T	$E\{T[\tilde{t}_{iT}(1,\rho) - \tilde{t}_{iT}]\}$			$\sqrt{Var\{\sqrt{T}[\tilde{t}_{iT}(1,\rho) - \tilde{t}_{iT}]\}}$		
	$\rho = 0.5$	$\rho = 0.9$	$\rho = 0.98$	$\rho = 0.5$	$\rho = 0.9$	$\rho = 0.98$
10	1.15	4.65	8.28	2.03	3.53	4.43
25	0.90	6.41	18.10	2.48	5.02	7.22
50	0.77	6.26	27.70	2.70	5.89	9.58
100	0.65	5.51	36.60	2.84	6.51	11.80
500	0.64	4.11	29.00	2.99	7.29	15.20
1000	0.67	3.45	23.50	3.02	7.50	16.10
5000	0.73	3.69	17.40	3.01	7.56	17.10
10000	0.82	1.81	16.00	3.04	7.66	17.40

Assuming that our conjecture is valid, using the independence of $d_{iT}(p_i, \boldsymbol{\rho}_i)$ across i for any T, we have

$$Var(D_{NT}) = \frac{1}{Var(\tilde{t}_T)}\left(\frac{1}{T}\right)\left(\frac{\sum_{i=1}^{N} Var[\sqrt{T}d_{iT}(p_i, \boldsymbol{\rho}_i)]}{N}\right),$$

and

$$E(D_{NT}) = \frac{1}{\sqrt{Var(\tilde{t}_T)}}\left(\frac{\sqrt{N}}{T}\right)\left(\frac{\sum_{i=1}^{N} E[Td_{iT}(p_i, \boldsymbol{\rho}_i)]}{N}\right).$$

Using (4.14), and recalling that $Var(\tilde{t}_T)$ exists and is finite for $T > 5$ we now have

$$\lim_{N,T\to\infty,\frac{N}{T}\to k} Var(D_{NT}) = 0, \quad \text{and} \quad \lim_{N,T\to\infty,\frac{N}{T}\to k} E(D_{NT}) = 0,$$

and D_{NT} weakly converges to zero as required by (4.13).

5. Monte Carlo simulation results

In this section we use Monte Carlo experiments to examine finite sample properties of the alternative panel-based unit root tests. We consider four sets of Monte Carlo experiments. The first set focuses on the benchmark model,

$$y_{it} = (1 - \phi_i)\mu_i + \phi_i y_{i,t-1} + \varepsilon_{it}, \quad t = 1, \dots, T,\ i = 1, \dots, N, \tag{5.1}$$

[11] The result in Said and Dickey (1984) is proved for $t_{iT}(p_i, \boldsymbol{\rho}_i) - t_{iT}$ but applies equally to $\tilde{t}_{iT}(p_i, \boldsymbol{\rho}_i) - \tilde{t}_{iT}$.

where $\varepsilon_{it} \sim \mathrm{N}(0, \sigma_i^2)$. The second set of experiments allows for the presence of positive (heterogeneous) AR(1) serial correlations in ε_{it},

$$\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}, \quad t = 1, \ldots, T, \; i = 1, \ldots, N, \tag{5.2}$$

where $e_{it} \sim \mathrm{N}(0, \sigma_i^2)$, $\rho_i \sim \mathrm{U}[0.2, 0.4]$, U stands for a uniform distribution and ρ_i's are generated independently of e_{it}. The third set of experiments considers the MA(1) error processes,

$$\varepsilon_{it} = e_{it} + \psi_i e_{i,t-1}, \quad t = 1, \ldots, T, \; i = 1, \ldots, N, \tag{5.3}$$

where $\psi_i \sim \mathrm{U}[-0.4, -0.2]$, and ψ_i's are generated independently of e_{it}. The fourth set of experiments allows for a linear trend in estimation of the ADF regressions using the same data generating process employed in the second set of experiments.

In all of the experiments e_{it} (or ε_{it} in (5.1)) are generated as *iid* normal variates with zero means and heterogeneous variances, σ_i^2. The parameters μ_i and σ_i^2 are generated according to

$$\mu_i \sim \mathrm{N}(0, 1), \quad \sigma_i^2 \sim \mathrm{U}[0.5, 1.5], \quad i = 1, 2, \ldots, N. \tag{5.4}$$

Under the null $\phi_i = 1$ for all i, while $\phi_i = 0.9$ for all i under the alternative hypothesis.[12] All of the parameter values such as μ_i, σ_i^2, ρ_i or ψ_i are generated independently of ε_{it} once, and then fixed throughout replications. The first set of experiments were carried out for $N = 1, 5, 10, 25, 50, 100$, $T = 10, 25, 50, 100$. The other experiments were carried out for $N = 10, 25, 50, 100$, $T = 10, 25, 50, 100$, and $p = 0, 1, 2, 3, 4$. We used 2,000 replications to compute empirical size and power of the tests at the 5% nominal level.[13]

All experiments are carried out using the following two statistics: the t-bar and the LL tests. The results for the first set of experiments are summarized in Table 4.[14] As a benchmark, in the first row of this table we give empirical size and power of the standard DF test. When $T = 10$, we report only the results for the t-bar tests, because the adjustment factors needed to compute the LL statistic are not supplied by Levin and Lin (1993) for $T = 10$. The t-bar test (i.e. Z_{tbar} defined by (3.13) for simple DF regressions) perform well even for small values of T. It has the correct size, and its power rises monotonically with N and T. In the case of $T = 10$, the power of the t-bar test increases steadily from 0.071 for $N = 5$ to 0.384 for $N = 100$. Overall, both tests

[12] We also carried out a number of other experiments where we considered heterogeneous alternatives with $\phi_i \sim \mathrm{U}[0.85, 0.95]$. The results are similar to the ones reported in this section. A complete set of the Monte Carlo results for these experiments are available from the authors on request.

[13] We have also carried out a number of other experiments including the cases with negatively correlated AR(1) disturbances or positively correlated MA(1) disturbances in conjunction with the ADF regression models with or without time trend being estimated. Most results are qualitatively similar to what follows. We have further conducted simulation exercises where the ADF order is chosen by information criteria such as the Akaike Information or the Schwarz Criterion. The tests show a significant degree of size distortions, largely reflecting that in finite samples these criteria tend to select a low order for the ADF regressions. See also Ng and Perron (1995) for the effect of the ADF order selection on the finite sample performance of unit root tests in the case of single time series.

[14] We also carried out a number of experiments allowing for linear trends in the DF regressions and obtained similar results, although as to be expected the tests tended to have little power for T and N less than 25.

Table 4
Size and power of unit root tests in heterogeneous panels. Experiment 1: No serial correlation, no time trend

N	Test	$T=10$		$T=25$		$T=50$		$T=100$	
		Size	Power	Size	Power	Size	Power	Size	Power
1	DF	0.089	0.095	0.069	0.091	0.058	0.151	0.053	0.351
5	Z_{tbar}	0.052	0.071	0.050	0.153	0.050	0.441	0.042	0.972
10	Z_{tbar}	0.050	0.090	0.049	0.261	0.054	0.752	0.050	1.00
	LL			0.061	0.260	0.057	0.555	0.046	0.964
25	Z_{tbar}	0.052	0.141	0.048	0.549	0.050	0.992	0.054	1.00
	LL			0.064	0.532	0.054	0.925	0.053	1.00
50	Z_{tbar}	0.050	0.229	0.044	0.838	0.051	1.00	0.050	1.00
	LL			0.070	0.809	0.065	0.998	0.062	1.00
100	Z_{tbar}	0.046	0.384	0.053	0.990	0.051	1.00	0.046	1.00
	LL			0.084	0.983	0.068	1.00	0.059	1.00

Notes: This table reports the size and power of the t-bar test statistic Z_{tbar} defined by (3.13) and the Levin and Lin (LL) test, respectively. The underlying data is generated by $y_{it}=(1-\phi)\mu_i+\phi y_{i,t-1}+\varepsilon_{it}$, $i=1,\dots,N$, $t=-51,-50,\dots,T$, where we generate $\mu_i\sim \mathrm{N}(0,1)$ and $\varepsilon_{it}\sim \mathrm{N}(0,\sigma_i^2)$ with $\sigma_i^2\sim \mathrm{U}[0.5,1.5]$. μ_i and σ_i^2 are generated once and then fixed in all replications. The test statistics are obtained using the DF regressions: $\Delta y_{it}=\alpha_i+\beta_i y_{i,t-1}+\varepsilon_{it}$, $t=1,2,\dots,T$, $i=1,2,\dots,N$. The size ($\phi=1$) and power ($\phi=0.9$) of the tests are computed at the five percent nominal level. Number of replications is set to 2,000. The result for $N=1$ is reported for comparison, and DF refers to the Dicky–Fuller test. The LL test for $T=10$ is not included, since the adjustment factors necessary for computing the LL statistic are not reported in Levin and Lin (1993).

have similar features, although for $T=25$, the LL test seems to have a tendency to over-reject as N is allowed to increase. In addition, for small T the t-bar test has a slightly higher power even though the LL test has a larger size.

In the case of models with serially correlated errors, we experimented with both versions of the standardized t-bar tests; that is, $Z_{tbar}(\mathbf{p},\boldsymbol{\rho})$ defined in (4.6), and $W_{tbar}(\mathbf{p},\boldsymbol{\rho})$ defined in (4.10), respectively. Although the tests are asymptotically equivalent, we have found that $W_{tbar}(\mathbf{p},\boldsymbol{\rho})$ statistic, which explicitly takes account of the underlying ADF orders in computing the mean and the variance adjustment factors, perform much better in small samples. Therefore, in what follows, we report only the results based on this statistic.

Table 5 summarizes the results for the second set of experiments. These results clearly show the importance of correctly choosing the order of the underlying ADF regressions. The problem is particularly serious when the order is underestimated. For example, incorrectly using the ADF(0) regressions to compute the test statistics causes the size of all the tests to go to zero.[15] The situation is different, however, when the

[15] The direction of the size distortion of the tests crucially depends on whether ε_{it} are positively or negatively autocorrelated. Failure to adequately allow for positive residual serial correlation results in under-rejection of the null hypothesis, while gross over-rejection results in the presence of neglected negative residual serial correlation. This is due to the fact that the distribution of the ADF(0) t-statistic gets shifted to the right with larger dispersions when there is a (neglected) positive residual serial correlation, and to the left with larger dispersions when there is a (neglected) negative residual serial correlation. Using 50,000 replications $\mathrm{E}(t_T|\phi=1)$ and $Var(t_T|\phi=1)$ are simulated to be -1.5320 and 0.7346, respectively, when $\rho=0$, while they are -0.8979 and 1.2789 when $\rho=0.5$; and -2.6291 and 1.2061 when $\rho=-0.5$.

Table 5
Size and power of unit root tests in heterogeneous panels. Experiment 2: AR(1) Errors with $\rho_i \sim$ U[0.2, 0.4], No Time Trend

T	N	Test	ADF(0)		ADF(1)		ADF(2)		ADF(3)		ADF(4)	
			Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
10	10	W_{tbar}	0.010	0.010	0.060	0.110	0.067	0.108	0.073	0.110	0.065	0.095
	25	W_{tbar}	0.002	0.003	0.046	0.133	0.066	0.136	0.067	0.163	0.075	0.151
	50	W_{tbar}	0.000	0.001	0.046	0.213	0.061	0.211	0.066	0.238	0.074	0.217
	100	W_{tbar}	0.000	0.000	0.044	0.343	0.055	0.333	0.062	0.372	0.069	0.337
25	10	W_{tbar}	0.005	0.004	0.046	0.241	0.043	0.207	0.042	0.185	0.053	0.175
		LL	0.015	0.004	0.067	0.224	0.043	0.088	0.038	0.061	0.029	0.021
	25	W_{tbar}	0.000	0.000	0.056	0.483	0.058	0.431	0.051	0.384	0.054	0.339
		LL	0.006	0.001	0.081	0.464	0.043	0.155	0.035	0.081	0.017	0.008
	50	W_{tbar}	0.000	0.000	0.050	0.758	0.049	0.691	0.045	0.633	0.042	0.559
		LL	0.002	0.001	0.093	0.722	0.030	0.258	0.021	0.101	0.008	0.006
	100	W_{tbar}	0.000	0.000	0.052	0.959	0.054	0.922	0.046	0.892	0.057	0.845
		LL	0.000	0.000	0.113	0.930	0.019	0.398	0.008	0.129	0.002	0.002
50	10	W_{tbar}	0.003	0.018	0.051	0.672	0.045	0.579	0.044	0.545	0.046	0.486
		LL	0.011	0.002	0.062	0.379	0.046	0.174	0.037	0.087	0.029	0.017
	25	W_{tbar}	0.000	0.015	0.064	0.972	0.055	0.941	0.046	0.917	0.048	0.878
		LL	0.002	0.001	0.084	0.778	0.052	0.451	0.043	0.203	0.026	0.026
	50	W_{tbar}	0.000	0.021	0.060	1.00	0.049	0.999	0.054	0.998	0.057	0.993
		LL	0.000	0.000	0.085	0.977	0.047	0.740	0.039	0.397	0.020	0.035
	100	W_{tbar}	0.000	0.015	0.063	1.00	0.042	1.00	0.048	1.00	0.051	1.00
		LL	0.000	0.000	0.095	1.00	0.042	0.963	0.033	0.685	0.012	0.058

Notes: The results in this table are computed using the same data generating process as in Table 4 except that ε_{it}'s now follow the AR(1) processes: $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}$, $t = 1, \ldots, T$, $i = 1, \ldots, N$, where $e_{it} \sim \mathrm{N}(0, \sigma_i^2)$ and $\rho_i \sim$ U[0.2, 0.4]. ρ_i's are generated once and then fixed in all replications. The reported size and power of the tests are based on the individual ADF regressions: $\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \sum_{j=1}^{p} \rho_{ij} \Delta y_{i,t-j} + e_{it}$, $t = 1, \ldots, T$, $i = 1, \ldots, N$. The test statistic W_{tbar} is defined in (4.10). See also the notes to Table 4.

orders of the underlying ADF regressions are correctly chosen or over-estimated. In the case where the order is correctly chosen, the t-bar test, based on $W_{tbar}(\mathbf{p}, \boldsymbol{\rho})$, has empirical sizes very close to the nominal size, for all sample sizes, even when $T = 10$. In contrast, the LL test tends to over-reject the null hypothesis, and the extent of this over-rejection worsens as N increases. For example, when $T = 50$, the size of the LL test for $p = 1$ rises from 0.062 for $N = 10$ to 0.095 for $N = 100$. For $p > 1$, the size of the LL test tends towards zero as N increases. Turning to the power performance of the tests, we find that the t-bar test is more powerful than the LL test. For example, when $T = 50$ and $N = 25$, although the LL test for $p = 1$ is slightly over-sized (0.084) and the t-bar test has correct size (0.064), the powers of the t-bar test (0.972) substantially exceeds that of the LL test (0.778).

In general, over-fitting is less harmful for inference on unit roots than under-fitting, but there is also a trade-off between power and size of the tests, which is a well-established empirical finding when testing for a unit root in the case of single time series. This trade-off is, however, much less favorable to the LL test than it is to the t-bar test. For example, when $T = 50$ and $N = 25$, the power of the t-bar test based

Table 6
Size and power of unit root tests in heterogeneous panels. Experiment 3: MA(1) errors with $\psi_i \sim$ U$[-0.4, -0.2]$, no time trend

T	N	Test	ADF(0)		ADF(1)		ADF(2)		ADF(3)		ADF(4)	
			Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
10	10	W_{tbar}	0.522	0.714	0.099	0.195	0.051	0.086	0.043	0.074	0.043	0.066
	25	W_{tbar}	0.696	0.904	0.120	0.281	0.065	0.145	0.043	0.110	0.042	0.090
	50	W_{tbar}	0.945	0.998	0.166	0.483	0.070	0.237	0.047	0.151	0.041	0.124
	100	W_{tbar}	0.999	1.00	0.265	0.776	0.088	0.415	0.044	0.291	0.040	0.186
25	10	W_{tbar}	0.810	0.998	0.186	0.618	0.076	0.310	0.050	0.213	0.047	0.188
		LL	0.787	0.995	0.103	0.391	0.020	0.044	0.013	0.015	0.007	0.002
	25	W_{tbar}	0.926	1.00	0.179	0.853	0.078	0.574	0.048	0.451	0.045	0.379
		LL	0.899	1.00	0.100	0.601	0.018	0.067	0.009	0.019	0.004	0.001
	50	W_{tbar}	0.999	1.00	0.335	0.995	0.096	0.876	0.048	0.742	0.042	0.645
		LL	0.999	1.00	0.141	0.894	0.009	0.090	0.004	0.009	0.000	0.000
	100	W_{tbar}	1.00	1.00	0.531	1.00	0.124	0.997	0.058	0.963	0.048	0.908
		LL	1.00	1.00	0.203	0.994	0.002	0.117	0.001	0.004	0.000	0.000
50	10	W_{tbar}	0.892	1.00	0.213	0.978	0.072	0.829	0.046	0.682	0.038	0.579
		LL	0.867	1.00	0.114	0.774	0.029	0.133	0.018	0.031	0.013	0.003
	25	W_{tbar}	0.964	1.00	0.215	1.00	0.073	0.995	0.053	0.973	0.046	0.941
		LL	0.940	1.00	0.094	0.971	0.018	0.353	0.011	0.086	0.005	0.004
	50	W_{tbar}	1.00	1.00	0.426	1.00	0.120	1.00	0.070	1.00	0.055	0.999
		LL	0.999	1.00	0.172	1.00	0.021	0.639	0.006	0.129	0.001	0.005
	100	W_{tbar}	1.00	1.00	0.663	1.00	0.150	1.00	0.070	1.00	0.060	1.00
		LL	1.00	1.00	0.281	1.00	0.008	0.908	0.002	0.199	0.000	0.001

Notes: The results in this table are computed using the same data generation process as in Table 4 except that ε_{it}'s now follow the MA(1) processes, $\varepsilon_{it} = e_{it} + \psi_i e_{i,t-1}$, $t = 1, \ldots, T$, $i = 1, \ldots, N$, where $e_{it} \sim \mathrm{N}(0, \sigma_i^2)$ and $\psi_i \sim$ U$[-0.4, -0.2]$. ψ_i's are generated once and then fixed in all replications. See also the notes to Tables 4 and 5.

on ADF(1) regressions decreases from 0.972 to 0.878 when ADF(4) regressions are used, but the power of the *LL* test drops from 0.778 to 0.026!

The results for the third set of experiments are summarized in Table 6 for the case with negative MA(1) errors. Now choosing too small p in the ADF regressions could have more detrimental effects on the finite sample performance of both tests. But, it is notable that the empirical size of t-bar test is reasonably close to the nominal size, once the sufficiently large order of p is selected, e.g. $p = 3$ or 4. At the same time the power of the t-bar test remains comparable to that obtained in benchmark case. For example, when $p = 3$ and $T = 25$, the empirical powers of the t-bar test are $(0.213, 0.451, 0.742, 0.963)$ for $N = (10, 25, 50, 100)$, respectively, which are more or less comparable to the powers of the t-bar test for benchmark case $(0.261, 0.549, 0.838, 0.990)$ (see Table 4 for the corresponding T and N). On the other hand, the *LL* test tends to heavily under-reject both under the null and under the alternative when the higher order of p is selected. For example, when $p = 4$ is selected, both empirical size and power of the *LL* test are close to zero throughout all experiments.

Table 7
Size and power of unit root tests in heterogeneous panels. Experiment 4: time trend, AR(1) errors with $\rho_i \sim$ U[0.2, 0.4]

T	N	Test	ADF(0)		ADF(1)		ADF(2)		ADF(3)		ADF(4)	
			Size	Power	Size	Power	Size	Power	Size	Power	Size	Power
10	10	W_{tbar}	0.000	0.000	0.055	0.069	0.058	0.057	0.065	0.065	0.054	0.058
	25	W_{tbar}	0.000	0.000	0.044	0.065	0.053	0.065	0.056	0.076	0.066	0.077
	50	W_{tbar}	0.000	0.000	0.054	0.067	0.064	0.083	0.086	0.096	0.075	0.096
	100	W_{tbar}	0.000	0.000	0.058	0.077	0.050	0.077	0.080	0.103	0.092	0.103
25	10	W_{tbar}	0.000	0.000	0.051	0.093	0.052	0.085	0.056	0.084	0.057	0.084
		LL	0.095	0.111	0.045	0.088	0.011	0.015	0.005	0.007	0.000	0.000
	25	W_{tbar}	0.000	0.000	0.051	0.123	0.053	0.106	0.057	0.098	0.055	0.093
		LL	0.095	0.205	0.030	0.104	0.003	0.007	0.000	0.000	0.000	0.000
	50	W_{tbar}	0.000	0.000	0.057	0.203	0.056	0.165	0.053	0.136	0.058	0.139
		LL	0.071	0.229	0.024	0.158	0.000	0.004	0.000	0.000	0.000	0.000
	100	W_{tbar}	0.000	0.000	0.053	0.315	0.049	0.246	0.038	0.209	0.058	0.225
		LL	0.046	0.200	0.017	0.212	0.000	0.000	0.000	0.000	0.000	0.000
50	10	W_{tbar}	0.000	0.000	0.052	0.257	0.054	0.234	0.054	0.205	0.052	0.181
		LL	0.129	0.129	0.043	0.239	0.011	0.075	0.006	0.030	0.002	0.004
	25	W_{tbar}	0.000	0.000	0.052	0.488	0.064	0.442	0.058	0.379	0.049	0.312
		LL	0.119	0.316	0.040	0.420	0.007	0.102	0.002	0.018	0.000	0.000
	50	W_{tbar}	0.000	0.000	0.055	0.786	0.055	0.716	0.054	0.648	0.049	0.550
		LL	0.114	0.430	0.036	0.685	0.003	0.138	0.000	0.008	0.000	0.000
	100	W_{tbar}	0.000	0.000	0.048	0.964	0.063	0.937	0.054	0.888	0.043	0.806
		LL	0.087	0.435	0.015	0.925	0.000	0.204	0.000	0.007	0.000	0.000

Notes: The results in this table are computed using the same data generating process as in Table 5. But, the reported size and power of the tests are now based on the individual ADF regressions with time trend included: $\Delta y_{it} = \alpha_i + \delta_i t + \beta_i y_{i,t-1} + \sum_{j=1}^{p} \rho_{ij} \Delta y_{i,t-j} + e_{it},\ t = 1, \ldots, T,\ i = 1, \ldots, N$. See also the notes to Tables 4 and 5.

The results for the fourth set of experiments are summarized in Table 7. Once again an incorrect choice of the order of the ADF regression has detrimental effects for the performance of the test when the order is under-estimated. Correctly choosing the order of the ADF regression seems to be much more important in the present experiments where the panel contains deterministic time trends. In the case of the ADF(1) specification, the size of the t-bar test is around the nominal level even when $T = 10$. On the other hand, the LL test tends to under-reject, as N increases relative to T. This size distortion problem becomes more serious as p increases. In general, the t-bar test has better power performance than the LL test. However, even the t-bar test has a rather poor power performance when $T = 10$. This is in line with the results obtained for the case of the ADF test applied to a single time series. Another important feature of the results lies in the fact that the power of the t-bar test is much more favorably affected by a rise in T than by an equivalent rise in N.

6. Concluding remarks

In this paper we have developed a computationally simple procedure for testing the unit root hypothesis in heterogeneous panels. The small sample properties of the proposed tests are investigated via Monte Carlo methods. It is found that when there are no serial correlations, the t-bar test performs very well even when $T = 10$. In this case it is possible to substantially augment the power of the unit root tests applied to single time series.

The situation is more complicated when the disturbances in the dynamic panel are serially correlated. The t-bar test procedures now require that both N and T should be sufficiently large. Furthermore, in the case of serially correlated errors, it is critically important not to under-estimate the order of the underlying ADF regressions. When we allow for serial correlation and heterogeneity in the underlying data generating process, the simulation results clearly show that if a large enough lag order is selected for the underlying ADF regressions, then the finite sample performances of the t-bar test is reasonably satisfactory and generally better than that of the LL test.

Since the working paper version of this paper first appeared in 1995, the t-bar test has been used extensively in the empirical panel literature and its computations are coded in econometric software packages such as TSP and Stata.[16] Special care needs to be exercised when interpreting the results of the panel unit root tests. Due to the heterogeneous nature of the alternative hypothesis, H_1, as specified by (2.4), rejection of the null hypothesis does not necessarily imply that the unit root null is rejected for all i, but only that the null hypothesis is rejected for $N_1 < N$, members of the group such that as $N \to \infty$, $N_1/N \to \delta > 0$. The test result does not provide any guidance as to the magnitude of δ, or the identity of the particular panel members for which the null hypothesis is rejected.

Acknowledgements

We are grateful to Ron Smith, Stephen Satchell, Michael Binder, James Chu, Cheng Hsiao, Karim Abadir, Rolf Larsson, Clive Granger, Kevin Lee, Bent Nielson, Jan Magnus, Jeffrey Woodridge, Robert de Jong, Peter Robinson and a number of anonymous referees for helpful comments. Partial financial support from the ESRC (Grant No. H519255003) and the Isaac Newton Trust of Trinity College, Cambridge is gratefully acknowledged.

References

Balestra, P., Nerlove, M., 1966. Pooling cross-section and time-series data in the estimation of a dynamic economic model: the demand for natural gas. Econometrica 34, 585–612.

Baltagi, B.H., Kao, C., 2000. Nonstationary panels, cointegration in panels and dynamic panels: a survey. Advances in Econometrics 15, 7–51.

Bornhorst, F., Baum, C.F., 2001. IPSHIN: Stata module to perform Im-Pesaran-Shin panel unit root test. Boston College, Department of Economics, Statistical Software Components, No. S419704.

Bowman, D., 1999. Efficient tests for autoregressive unit roots in panel data. Unpublished manuscript, Board of Governors of the Federal Reserve System, Washington, D.C.

[16] The procedure for performing the t-bar test in TSP is provided by Hall and Mairesse (2001), and for Stata by Bornhorst and Baum (2001).

Choi, I., 2001. Unit root tests for panel data. Journal of International Money and Banking 20, 249–272.
Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427–431.
Hadri, K., 2000. Testing for stationarity in heterogeneous panel data. Econometrics Journal 3, 148–161.
Hall, B.H., Mairesse, J., 2001. Testing for unit roots in panel data: an exploration using real and simulated data. Economics Department, UC Berkeley.
Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton.
Levin, A., Lin, C.F., 1993. Unit root tests in panel data: asymptotic and finite-sample properties. Unpublished manuscript, University of California, San Diego.
Maddala, G.S., Wu, S., 1999. A comparative study of unit root tests with panel data and a new simple test. Oxford Bulletin of Economics and Statistics 61, 631–652.
Magnus, J., 1990. On certain moments relating to ratios of quadratic forms in normal variables: further results. Sankhya, 52, Series B, Pt. 1, 1–13.
Nabeya, S., 1999. Asymptotic moments of some unit root test statistics in the null case. Econometric Theory 15, 139–149.
Ng, S., Perron, P., 1995. Unit root test in ARMA models with data-dependent methods for the selection of the truncation lag. Journal of American Statistical Association 90, 268–281.
Pesaran, M.H., Smith, R., 1995. Estimating long-run relationships from dynamic heterogeneous panels. Journal of Econometrics 68, 79–113.
Pesaran, M.H., Smith, R., Im, K.S., 1996. Dynamic linear models for heterogeneous panels. In: Mátyás L, Sevestre P (Ed.), The Econometrics of Panel Data: A Handbook of Theory with Applications, (2nd Revised Edition). Kluwer Academic Publishers, Dordrecht.
Phillips, P.C.B., Moon, R.H., 1999. Linear regression limit theory and non-stationary panel data. Econometrica 67, 1112–1157.
Quah, D., 1992. International patterns of growth: I. Persistence in cross-country disparities. Unpublished manuscript, London School of Economics.
Quah, D., 1994. Exploiting cross-section variations for unit root inference in dynamic data. Economics Letters 44, 9–19.
Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive moving average models of unknown order. Biometrika 71, 599–607.
Shin, Y., Snell, A., 2000. Joint asymptotic results for mean group tests in heterogenous panels: application to the KPSS stationarity test. Unpublished manuscript, University of Edinburgh.
Summers, R., Heston, A., 1991. The Penn world table (Mark 5): An expanded set of international comparisons 1950–1988. Quarterly Journal of Economics 106, 327–368.