CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Apollo

A Simple Panel Unit Root Test in the Presence of Cross Section Dependence

M. Hashem Pesaran

October 2003

CWPE 0346

Not to be quoted without permission

A Simple Panel Unit Root Test in the Presence of Cross Section Dependence*

M. Hashem Pesaran
University of Southern California & Cambridge University

September 24, 2003

Abstract

A number of panel unit root tests that allow for cross section dependence have been proposed in the literature, notably by Bai and Ng (2002), Moon and Perron (2003), and Phillips and Sul (2002) who use orthogonalization type procedures to asymptotically eliminate the cross dependence of the series before standard panel unit root tests are applied to the transformed series. In this paper we propose a simple alternative test where the standard DF (or ADF) regressions are augmented with the cross section averages of lagged levels and first-differences of the individual series. A truncated version of the CADF statistics is also considered. New asymptotic results are obtained both for the individual CADF statistics, and their simple averages. It is shown that the $CADF_i$ statistics are asymptotically similar and do not depend on the factor loadings under joint asymptotics where N (cross section dimension) and T (time series dimension) $\rightarrow \infty$, such that $N/T \rightarrow k$, where k is a fixed finite non-zero constant. But they are asymptotically correlated due to their dependence on the common factor. Despite this it is shown that the limit distribution of the average CADF statistic exists and its critical values are tabulated. The small sample properties of the proposed tests are investigated by Monte Carlo experiments, for a variety of models. It is shown that the cross sectionally augmented panel unit root tests have satisfactory size and power even for relatively small values of N and T. This is particularly true of cross sectionally augmented and truncated versions of the simple average t-test of Im, Pesaran and Shin, and Choi's inverse normal combination test.

JEL Classification: C12, C15, C22, C23.
Key Words: Panel unit root tests, Cross-section dependence, Heterogeneous dynamic panels, Finite sample properties.

*I would like to thank Soren Johansen and Chris Rogers for helpful discussions with respect to the analysis of exchangeable processes, and Ron Smith for useful general comments. I am also particularly grateful to Yongcheol Shin for carrying out some preliminary computations that were instrumental in helping me form some of the ideas developed formally in the paper. Excellent research assistant by Mutita Akusuwan is also gratefully acknowledged who carried out all the computations reported in this paper.

1 Introduction

Over the past decade the problem of testing for unit roots in heterogeneous panels has attracted a great of deal attention. See, for example, Bowman (1999), Choi (2001), Hadri (2000), Im, Pesaran and Shin (1995, 2003), Levin, Lee, and Lu (2002), Maddala and Wu (1999), and Shin and Snell (2000). Baltagi and Kao (2000) provide an early review. This literature, however, assumed that the individual time series in the panel were cross-sectionally independently distributed. While it was recognized that this was a rather restrictive assumption, particularly in the context of cross country (region) regressions, it was thought that cross-sectionally de-meaning the series before application of the panel unit root test could partly deal with the problem. (see Im, Pesaran and Shin (1995)). However, it was clear that cross-section de-meaning could not work in general where pair-wise cross-section covariances of the error terms differed across the individual series. Recognizing this deficiency new panel unit root tests have been proposed in the literature by Bai and Ng (2002), Chang (2002), Choi (2002), Harvey and Bates (2002), Moon and Perron (2003), and Phillips and Sul (2002).

Chang (2002) proposes a non-linear instrumental variable approach to deal with the cross section dependence of a general form and establishes that individual Dickey-Fuller (DF) or the Augmented DF (ADF) statistics are asymptotically independent when an integrable function of the lagged dependent variables are used as instruments. From this she concludes that her test is valid for both T (the time series dimension) and N (the cross section dimension) are large. However, as shown by Im and Pesaran (2003), her test is valid only if N is fixed as $T \to \infty$. Using Monte Carlo techniques, Im and Pesaran show that Chang's test is grossly over-sized for moderate degrees of cross section dependence, even for relatively small values of N.[1]

The test proposed by Harvey and Bates is also valid for general specifications of error cross-correlations, but is limited as it requires the parameters to be the same across all the series. The Harvey-Bates procedure also seems to work only when N is small and T relatively large.[2]

Choi (2002) models the cross dependence using a two-way error-component model which imposes the same pair-wise error covariances across the different cross section units. This provide a generalization of the cross-section de-meaning procedure proposed in Im, Pesaran and Shin (1995) but it can still be restrictive in the context of heterogeneous panels.

Bai and Ng (2002), Moon and Perron (2003), and Phillips and Sul (2002) avoid the restrictive nature of the cross-section de-meaning procedure by allowing the common factors to have differential effects on different cross section units. In the context of a residual one-factor model Phillips and Sul (2002) show that in the presence of cross section dependence the standard panel unit root tests are no longer asymptotically similar, and propose an orthogonalization procedure which in effect asymptotically eliminates the common factors before preceding to the application of standard panel unit root tests. Sequential asymptotic

[1] Following Maddala and Wu (1999) bootstrap techniques have also been utilized to deal with cross section dependence in panel unit root tests. See, for example, Smith, Leybourne, Kim and Newbold (2003), and Chang (2003). These procedures are also valid if N is fixed as $T \to \infty$.

[2] In their Monte Carlo experiments they report results for $N = 2, 5$ and $T = 100, 200, 500$, in the case of homogeneous panel data models without fixed effects.

results are provided in the case where $T \to \infty$, and then $N \to \infty$.

Independently, similar orthogonalization procedures are used by Bai and Ng (2002) and Moon and Perron (2003) in a more general set up. Moon and Perron (2003) propose a pooled panel unit root test based on "de-factored" observations and suggest estimating the factor loadings that enter their proposed statistic by the principal component method. They derive asymptotic properties of their test under the unit root null and local alternatives, assuming in particular that $N/T \to 0$, as N and $T \to \infty$. They show that their proposed test has good asymptotic power properties if the model does not contain deterministic (incidental) trends. In a related paper, Moon, Perron and Phillips (2003) propose a point optimal invariant panel unit root test which is shown to have local power even in the presence of deterministic trends. Bai and Ng (2002) consider a more general set up and allow for the possibility of unit roots (and cointegration) in the common factors, but continue to assume that $N/T \to 0$, as N and $T \to \infty$. To deal with such a possibility they apply the principle component procedure to the first-differenced version of the model, and estimate the factor loadings and the first differences of the common factors. Standard unit root tests are then applied to the factors and the individual de-factored series, both computed as partial sums of the estimated first differences.

In this paper we adopt a different approach to dealing with the problem of cross section dependence. Instead of basing the unit root tests on deviations from the estimated factors, we augment the standard DF (or ADF) regressions with the cross section averages of lagged levels and first-differences of the individual series. Standard panel unit root tests can now be based on the simple averages of the individual cross sectionally augmented ADF statistics (denoted by CADF), or suitable transformations of the associated rejection probabilities. The individual CADF statistics or the rejection probabilities can then be used to develop modified versions of the t-bar test proposed by Im, Pesaran and Shin (IPS), the inverse chi-squared test (or the P test) proposed by Maddala and Wu (1999), and the inverse normal test (or the Z test) suggested by Choi (2001). A truncated version of the test is also considered where the individual CADF statistics are suitably truncated to avoid undue influences of extreme outcomes that could arise when T is small (in the region of $10-20$). New asymptotic results are obtained both for the individual CADF statistics, and their simple averages, referred to as the cross-sectionally augmented IPS (CIPS) test. The asymptotic null distribution of the individual $CADF_i$ and the associated $CIPS = N^{-1}\sum_{i=1}^{N} CADF_i$ statistics are investigated as $N \to \infty$ followed with $T \to \infty$, as well as jointly when N and T tending to infinity such that $N/T \to k$, where k is a fixed finite non-zero constant. It is shown that the $CADF_i$ statistics are asymptotically similar and do not depend on the factor loadings. But they are asymptotically correlated due to their dependence on the common factor. As a result the standard central limit theorems do not apply to the CIPS statistic (or the other combination or meta type tests proposed by Maddala and Wu, and Choi). However, it is shown that the limit distribution of the truncated version of the $CIPS$ statistic (denoted by $CIPS^*$) exits and is free of nuisance parameters. The critical values of $CIPS$ and $CIPS^*$ statistics are tabulated for the three main specifications of the deterministics, namely in the case of models without intercepts or trends, models with individual-specific intercepts, and models with incidental linear trends.[3]

[3]Critical values for the cross sectionally augmented combination (or meta) tests are available from the

The small sample properties of the proposed tests are investigated by Monte Carlo experiments, for a variety of models with incidental deterministics (intercepts as well as linear trends), cross dependence (low and high) and individual specific residual serial correlation (positive and negative), and sample sizes, N and $T = 10, 20, 30, 50, 100$. The simulations show that the cross sectionally augmented panel unit root tests have satisfactory size and power even for relatively small values of N and T. This is particularly true of the truncated version of the CIPS test and the cross sectionally augmented version of Choi's inverse normal combination test. These tests show satisfactory size properties even for very small sample sizes, namely when $N = T = 10$, and there is a high degree of cross section dependence with a moderate degree of residual serial correlation. Perhaps not surprisingly, the power of the tests critically depends on the sample sizes N and T, and on whether the model contains linear time trends. In the case of models with linear time trends power starts to rise with N only if T is 30 or more. For $T > 30$ the power rises quite rapidly with both N and T. In their respective simulations Bai and Ng (2002) report Monte Carlo results for $T = 100$ and $N = 20, 100$, Moon and Perron (2003) for $T = 100, 300$, $N = 10, 20$, and Phillips and Sul (2002) for $T = 50, 100, 200$ and $N = 10, 20, 30$. All these studies consider experiments where T is much larger than N, and hence are difficult to evaluate in relation to our simulation results where T could be small relative to N and *vice versa*. The simulations conducted by Bai and Ng (2002) and Moon and Perron (2003) are also confined to models without deterministic trends. Further simulations would be needed for comparisons of the small sample properties of the various tests based on the orthogonalization procedure and our cross sectionally augmented tests. The proposed tests have the added advantage of being very simple to compute.

Clearly, it is also possible to construct the CADF test based on recent modifications of the ADF test proposed in the literature, for example, the ADF-GLS test of Elliott et al. (1996), the weighted symmetric ADF (WS-ADF) test of Fuller and Park (1995) and Fuller (1996, Section 10.1.3), or the Max-ADF test of Leybourne (1995). The use of the latter two modifications of ADF statistics in IPS panel unit root test have been recently considered by Smith, Leybourne, Kim and Newbold (2003) who report significant gain in power as compared to the IPS test based on standard ADF statistics.

The plan of the paper is as follows. Section 2 sets out the basic model. Section 3 introduces the cross sectionally augmented regressions for the individual series for models without residual serial correlations, and shows that in this case the CADF statistic does not depend on nuisance parameters as $N \to \infty$ for any fixed $T > 3$. The null distribution of the CADF statistic is derived under sequential and joint asymptotics and it is shown that the CADF statistics for different series are asymptotically correlated and form an exchangeable sequence. Asymptotic critical values for the CADF distribution are provided in Section 3.1, together with simulated values of its moments, as well as the asymptotic correlation coefficient for any pair of CADF statistics. All the three main specifications of the deterministics (no intercept or trend, intercept only, and a linear trend) are covered. The various CADF based panel unit root tests (the cross sectionally augmented versions of the IPS, P and the Z tests) are discussed in Section 4. Section 5 extends the results to the case where the individual specific errors are serially correlated. Three alternative specifications of residual

author on request.

serial correlations are considered and it is shown that the individual CADF statistics have the same asymptotic distribution as in the serially uncorrelated case, so long as the CADF regressions are further augmented with the lagged changes of the individual series as well as the lagged changes of the cross section averages. Small sample performance of the proposed tests are investigated in Section 6 using Monte Carlo experiments. Concluding remarks and direction for future research are provided in Section 7.

Notations: $a_n = O(b_n)$ states the deterministic sequence $\{a_n\}$ is at most of order b_n, $\mathbf{x_n} = \mathbf{O_p}(\mathbf{y_n})$ states the vector of random variables, $\mathbf{x_n}$, is at most of order $\mathbf{y_n}$ in probability,and $\mathbf{x_n} = \mathbf{o_p}(\mathbf{y_n})$ is of smaller order in probability than $\mathbf{y_n}$. $\rightarrow$ denotes convergence in quadratic mean (q.m.) or mean square errors and $\Longrightarrow$ convergence in distribution. All asymptotics are carried out under $N \rightarrow \infty$, either with a fixed T, sequentially, or *jointly* with $T \rightarrow \infty$. In particular, $\overset{N}{\Longrightarrow}$ ($\overset{N}{\rightarrow}$) denotes convergence in distribution (q.m.) with T fixed as $N \rightarrow \infty$, $\overset{T}{\Longrightarrow}$ ($\overset{T}{\rightarrow}$) denotes convergence in distribution (q.m.) for N fixed (or when there is no N-dependence) as $T \rightarrow \infty$, $\overset{N,T}{\Longrightarrow}$ denotes sequential convergence with $N \rightarrow \infty$ first followed by $T \rightarrow \infty$ (similarly $\overset{T,N}{\Longrightarrow}$), $\overset{(N,T)_j}{\Longrightarrow}$ denotes joint convergence with $N, T \rightarrow \infty$ jointly such that $N/T \rightarrow k$, where k is a fixed finite non-zero constant. $\backsim$ denotes asymptotic equivalence in distribution, with $\overset{N}{\backsim}$, $\overset{T}{\backsim}$, $\overset{N,T}{\backsim}$, $\overset{T,N}{\backsim}$, and $\overset{(N,T)_j}{\backsim}$, similarly defined as $\overset{N}{\Longrightarrow}$, $\overset{T}{\Longrightarrow}$, etc.

2 A Simple Dynamic Panel with Cross-Section Dependence

Let y_{it} be the observation on the i^{th} cross-section unit at time t and suppose that it is generated according to the following simple dynamic linear heterogeneous panel data model

$$y_{it} = (1 - \phi_i)\,\mu_i + \phi_i y_{i,t-1} + u_{it},\ i = 1, ..., N;\ t = 1, ..., T, \tag{2.1}$$

where initial value, y_{i0}, is given, and the error term, u_{it}, has the one-factor structure

$$u_{it} = \gamma_i f_t + \varepsilon_{it}, \tag{2.2}$$

in which f_t is the unobserved common effect, and ε_{it} is the individual-specific (idiosyncratic) error.

It is convenient to write (2.1) and (2.2)

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \gamma_i f_t + \varepsilon_{it}, \tag{2.3}$$

where $\alpha_i = (1 - \phi_i)\,\mu_i$, $\beta_i = -(1 - \phi_i)$ and $\Delta y_{it} = y_{it} - y_{i,t-1}$. The unit root hypothesis of interest, $\phi_i = 1$, can now be expressed as

$$H_0 : \beta_i = 0 \text{ for all } i, \tag{2.4}$$

against the possibly heterogeneous alternatives,

$$H_1 : \beta_i < 0,\ i = 1, 2, \ldots, N_1,\ \beta_i = 0,\ i = N_1 + 1, N_1 + 2, \ldots, N. \tag{2.5}$$

We shall assume that N_1/N, the fraction of the individual processes that are stationary, is non-zero and tends to the fixed value δ such that $0 < \delta \leq 1$ as $N \to \infty$. As noted in Im, Pesaran and Shin (2003) this condition is necessary for the consistency of the panel unit root tests.

We shall make the following assumptions:

Assumption 1: The idiosyncratic shocks, ε_{it}, $i = 1, 2, ..., N$, $t = 1, 2, ..., T$, are independently distributed both across i and t, have mean zero, variance σ_i^2, and finite fourth-order moment.

Assumption 2: The common factor, f_t, is serially uncorrelated with mean zero and a constant variance, σ_f^2, and finite fourth-order moment. Without loss of generality σ_f^2 will be set equal to unity.

Assumption 3: ε_{it}, f_t, and γ_i are independently distributed for all i.

The cross-section independence of ε_{it} (across i) is standard in one factor models, although its validity in more general settings may require specification of more than one common factor in (2.2).[4] Assumptions 1 and 2 together imply that the composite error, u_{it}, is serially uncorrelated. This restriction can be relaxed by considering stationary error processes of the type

$$u_{it} = \sum_{j=1}^{p} \rho_{ij} u_{i,t-j} + \gamma_i f_t + \varepsilon_{it}.$$

A simple example of this generalization will be considered in Section 5. Other approaches for dealing with residual serial correlation in DF regressions, such as the Phillips and Perron (1988) non-parametric procedure, could also be considered. Initially, however, we shall develop the proposed test in the simple case where u_{it} is serially uncorrelated.

3 Unit Root Tests For One-Factor Residual Models with Serially Uncorrelated Errors

Let $\bar{\gamma} = N^{-1} \sum_{j=1}^{N} \gamma_j$ and suppose that $\bar{\gamma} \neq 0$ for a fixed N and as $N \to \infty$. Then following the line of reasoning in Pesaran (2002), the common factor f_t can be proxied by the cross section mean of y_{it}, namely $\bar{y}_t = N^{-1} \sum_{j=1}^{N} y_{jt}$, and its lagged value(s), $\bar{y}_{t-1}$, $\bar{y}_{t-2}, ...$ for N sufficiently large. In the simple case where u_{it} is serially uncorrelated, it turns out that $\bar{y}_t$ and $\bar{y}_{t-1}$ are sufficient for asymptotically filtering out the effects of the unobserved common factor, f_t. We shall therefore base our test of the unit root hypothesis, (2.4), on the t-ratio of the OLS estimate of b_i ($\hat{b}_i$) in the following cross-sectionally augmented DF (CADF) regression

$$\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_i \Delta \bar{y}_t + e_{it}. \tag{3.6}$$

Denoting this t-ratio by $t_i(N, T)$ we have

$$t_i(N, T) = \frac{\Delta \mathbf{y}_i' \bar{\mathbf{M}}_w \mathbf{y}_{i,-1}}{\hat{\sigma}_i \left(\mathbf{y}_{i,-1}' \bar{\mathbf{M}}_w \mathbf{y}_{i,-1} \right)^{1/2}}, \tag{3.7}$$

[4] Assumption 2 is also employed by Phillips and Sul (2002) who also require f_t to be normally distributed.

where

$$\Delta \mathbf{y}_i = (\Delta y_{i1}, \Delta y_{i2}, \ldots, \Delta y_{iT})', \mathbf{y}_{i,-1} = (y_{i0}, y_{i1}, \ldots, y_{i,T-1})' \tag{3.8}$$

$$\bar{\mathbf{M}}_w = \mathbf{I}_T - \bar{\mathbf{W}}\left(\bar{\mathbf{W}}'\bar{\mathbf{W}}\right)^{-1}\bar{\mathbf{W}}',\ \ \bar{\mathbf{W}} = \left(\boldsymbol{\tau}, \Delta\bar{\mathbf{y}}, \bar{\mathbf{y}}_{-1}\right), \tag{3.9}$$

$$\boldsymbol{\tau} = (1,1,\ldots,1)', \Delta\bar{\mathbf{y}} = (\Delta\bar{y}_1, \Delta\bar{y}_2, \ldots, \Delta\bar{y}_T)', \bar{\mathbf{y}}_{-1} = (\bar{y}_0, \bar{y}_1, \ldots, \bar{y}_{T-1})', \tag{3.10}$$

$$\hat{\sigma}_i^2 = \frac{\Delta\mathbf{y}_i'\mathbf{M}_{i,w}\Delta\mathbf{y}_i}{T-4}, \tag{3.11}$$

$$\mathbf{M}_{i,w} = \mathbf{I}_T - \mathbf{G}_i(\mathbf{G}_i'\mathbf{G}_i)^{-1}\mathbf{G}_i',\ \text{and}\ \mathbf{G}_i = \left(\mathbf{y}_{i,-1}, \bar{\mathbf{W}}\right). \tag{3.12}$$

In computing the t-ratio of $\hat{b}_i$ it is analytically more convenient to use the following alternative estimator of σ_i^2,

$$\tilde{\sigma}_i^2 = \frac{\Delta\mathbf{y}_i'\bar{\mathbf{M}}_w\Delta\mathbf{y}_i}{T-3}, \tag{3.13}$$

As we shall, see under the null hypothesis $\hat{\sigma}_i^2$ and $\tilde{\sigma}_i^2$ are both consistent for σ_i^2, as N and T tend to infinity. But for investigating the limiting properties of the proposed test as $N \to \infty$ with T fixed, the use of $\tilde{\sigma}_i^2$ simplifies the analysis considerably. The t-ratio associated with $\tilde{\sigma}_i^2$ is given by

$$\tilde{t}_i(N,T) = \frac{\sqrt{T-3}\,\Delta\mathbf{y}_i'\bar{\mathbf{M}}_w\mathbf{y}_{i,-1}}{\left(\Delta\mathbf{y}_i'\bar{\mathbf{M}}_w\Delta\mathbf{y}_i\right)^{1/2}\left(\mathbf{y}_{i,-1}'\bar{\mathbf{M}}_w\mathbf{y}_{i,-1}\right)^{1/2}}. \tag{3.14}$$

Under $\beta_i = 0$ we have $\Delta y_{it} = \gamma_i f_t + \varepsilon_{it}$ and hence

$$\Delta\mathbf{y}_i = \gamma_i\mathbf{f} + \boldsymbol{\varepsilon}_i, \tag{3.15}$$

$$\mathbf{y}_{i,-1} = y_{i0}\boldsymbol{\tau} + \gamma_i\mathbf{s}_{f,-1} + \mathbf{s}_{i,-1}, \tag{3.16}$$

$$\Delta\bar{\mathbf{y}} = \bar{\gamma}\mathbf{f} + \bar{\boldsymbol{\varepsilon}}, \tag{3.17}$$

$$\bar{\mathbf{y}}_{-1} = \bar{y}_0\boldsymbol{\tau} + \bar{\gamma}\mathbf{s}_{f,-1} + \bar{\mathbf{s}}_{-1}, \tag{3.18}$$

where $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{iT})'$, $\mathbf{f} = (f_1, f_2, ..., f_T)'$, $\bar{\boldsymbol{\varepsilon}} = (\bar{\varepsilon}_1, \bar{\varepsilon}_2, ..., \bar{\varepsilon}_T)'$, $\bar{\varepsilon}_t = N^{-1}\sum_{j=1}^N \varepsilon_{jt}$, y_{i0} is a given initial value (fixed or random), $\bar{y}_0 = N^{-1}\sum_{j=1}^N y_{j0}$, $\mathbf{s}_{i,-1} = (0, s_{i1}, \ldots, s_{i,T-1})'$, $\mathbf{s}_{f,-1} = (s_{f0}, s_{f1}, ..., s_{f,T-1})'$, $\bar{\mathbf{s}}_{-1} = (0, \bar{s}_1, ..., \bar{s}_{T-1})$ with $s_{it} = \sum_{j=1}^t \varepsilon_{ij}$, and $\bar{s}_t = N^{-1}\sum_{j=1}^N s_{jt}$, for

$t=1,2,..,$ and $s_{ft}=\sum_{j=1}^{t}f_j$. Using (3.17) to eliminate $\mathbf{f}$ from (3.15), and noting that by assumption $\bar{\gamma}\neq 0$, we have

$$\Delta \mathbf{y}_i=\delta_i\Delta\bar{\mathbf{y}}+\xi_{it}, \tag{3.19}$$

where

$$\delta_i=\gamma_i/\bar{\gamma},\text{ and }\xi_{it}=\varepsilon_{it}-\delta_i\bar{\varepsilon}_t. \tag{3.20}$$

Therefore,

$$\bar{\mathbf{M}}_w\Delta\mathbf{y}_i=\bar{\mathbf{M}}_w\boldsymbol{\xi}_i,\text{ and }\mathbf{M}_{i,w}\Delta\mathbf{y}_i=\mathbf{M}_{i,w}\boldsymbol{\xi}_i, \tag{3.21}$$

where $\boldsymbol{\xi}_i=(\xi_{i1},\xi_{i2},...,\xi_{iT})'$ is distributed with mean zero and the covariance matrix $\omega_i^2\mathbf{I}_T$, with the ω_i^2 given by

$$\omega_i^2=\sigma_i^2\left(1-\frac{2\delta_i}{N}\right)+\frac{\delta_i^2}{N}\bar{\sigma}^2=\sigma_i^2+O\left(\frac{1}{N}\right), \tag{3.22}$$

where $\bar{\sigma}^2=N^{-1}\sum_{j=1}^{N}\sigma_j^2<\infty$. Therefore,

$$\bar{\mathbf{M}}_w\Delta\mathbf{y}_i=\omega_i\bar{\mathbf{M}}_w\boldsymbol{\upsilon}_i,\text{ and }\mathbf{M}_{i,w}\Delta\mathbf{y}_i=\omega_i\mathbf{M}_{i,w}\boldsymbol{\upsilon}_i, \tag{3.23}$$

where $\boldsymbol{\upsilon}_i=\boldsymbol{\xi}_i/\omega_i\sim(\mathbf{0},\mathbf{I}_T)$.

Similarly, using (3.18) to eliminate $\mathbf{s}_{f,-1}$ from (3.16) we obtain

$$\mathbf{y}_{i,-1}=(y_{i0}-\delta_i\bar{y}_0)\boldsymbol{\tau}+\delta_i\bar{\mathbf{y}}_{-1}+\mathbf{s}_{i,-1}-\delta_i\bar{\mathbf{s}}_{-1}, \tag{3.24}$$

and hence

$$\bar{\mathbf{M}}_w\mathbf{y}_{i,-1}=\omega_i\bar{\mathbf{M}}_w\mathfrak{s}_{\mathrm{i},-1}, \tag{3.25}$$

where $\mathfrak{s}_{\mathrm{i},-1}=(\mathbf{s}_{i,-1}-\delta_i\bar{\mathbf{s}}_{-1})/\omega_i$. It is easily seen that $\mathfrak{s}_{\mathrm{i}}$ is the random walk associated to $\boldsymbol{\upsilon}_i$.

Using (3.11), (3.23) and (3.25) in (3.7) we have

$$t_i(N,T)=\frac{\sqrt{T-4}\,\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_w\mathfrak{s}_{\mathrm{i},-1}}{(\boldsymbol{\upsilon}_i'\mathbf{M}_{i,w}\boldsymbol{\upsilon}_i)^{1/2}\left(\mathfrak{s}_{\mathrm{i},-1}'\bar{\mathbf{M}}_w\mathfrak{s}_{\mathrm{i},-1}\right)^{1/2}}. \tag{3.26}$$

Similarly using (3.23) and (3.25) in (3.14) we have

$$\tilde{t}_i(N,T)=\frac{\sqrt{T-3}\,\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_w\mathfrak{s}_{\mathrm{i},-1}}{\left(\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_w\boldsymbol{\upsilon}_i\right)^{1/2}\left(\mathfrak{s}_{\mathrm{i},-1}'\bar{\mathbf{M}}_w\mathfrak{s}_{\mathrm{i},-1}\right)^{1/2}}. \tag{3.27}$$

Hence, the exact null distribution of $t_i(N,T)$ or $\tilde{t}_i(N,T)$ will depend on the nuisance parameters only through their effects on $\bar{\mathbf{M}}_w$ and $\mathbf{M}_{i,w}$. But, as shown in the Appendix this dependence vanishes as $N\to\infty$, irrespective of whether T is fixed or is allowed to tend to infinity jointly with N.[5] In the case where T is fixed to ensure that the CADF statistics, $t_i(N,T)$ or $\tilde{t}_i(N,T)$, do not depend on the nuisance parameters the effect of the initial cross-section mean, $\bar{y}_0$, must also be eliminated. This can be achieved by applying the test to the deviations $y_{it}-\bar{y}_0$. The following theorems provide a formal statement of these results.

[5]In the case where N is small seemingly unrelated regression techniques can be applied and will not be considered here.

Theorem 3.1 *Suppose the series y_{it}, for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$, are generated under (2.4) according to (2.3) and by construction $\bar{y}_0$ (the cross-section mean of the initial observations) is set to zero. Then under Assumptions 1 and 2 the null distribution of $\tilde{t}_i(N,T)$ given by (3.14), will be free of nuisance parameters as $N \to \infty$ for any fixed $T > 3$. In particular, we have (in quadratic mean)*

$$\tilde{t}_i(N,T) \xrightarrow{N} \frac{\frac{\boldsymbol{\varepsilon}_i'\mathbf{s}_{i,-1}}{\sigma_i^2 T} - \mathbf{q}_{iT}'\Psi_{fT}^{-1}\mathbf{h}_{iT}}{\left(\frac{\boldsymbol{\varepsilon}_i'\boldsymbol{\varepsilon}_i}{\sigma_i^2(T-3)} - \frac{\mathbf{q}_{iT}'\Psi_{fT}^{-1}\mathbf{q}_{iT}}{T-3}\right)^{1/2}\left(\frac{\mathbf{s}_{i,-1}'\mathbf{s}_{i,-1}}{\sigma_i^2 T^2} - \mathbf{h}_{iT}'\Psi_{fT}^{-1}\mathbf{h}_{iT}\right)^{1/2}}, \tag{3.28}$$

where

$$\Psi_{fT} = \begin{pmatrix} \frac{\mathbf{f}'\mathbf{f}}{T} & \frac{\mathbf{f}'\boldsymbol{\tau}}{T} & \frac{\mathbf{f}'\mathbf{s}_{f,-1}}{T^{3/2}} \\ \frac{\boldsymbol{\tau}'\mathbf{f}}{T} & 1 & \frac{\boldsymbol{\tau}'\mathbf{s}_{f,-1}}{T^{3/2}} \\ \frac{\mathbf{s}_{f,-1}'\mathbf{f}}{T^{3/2}} & \frac{\mathbf{s}_{f,-1}'\boldsymbol{\tau}}{T^{3/2}} & \frac{\mathbf{s}_{f,-1}'\mathbf{s}_{f,-1}}{T^2} \end{pmatrix}, \mathbf{q}_{iT} = \begin{pmatrix} \frac{\mathbf{f}'\boldsymbol{\varepsilon}_i}{\sigma_i\sqrt{T}} \\ \frac{\boldsymbol{\tau}'\boldsymbol{\varepsilon}_i}{\sigma_i\sqrt{T}} \\ \frac{\mathbf{s}_{f,-1}'\boldsymbol{\varepsilon}_i}{\sigma_i T} \end{pmatrix}, \mathbf{h}_{iT} = \begin{pmatrix} \frac{\mathbf{f}'\mathbf{s}_{i,-1}}{\sigma_i T^{3/2}} \\ \frac{\boldsymbol{\tau}'\mathbf{s}_{i,-1}}{\sigma_i T^{3/2}} \\ \frac{\mathbf{s}_{f,-1}'\mathbf{s}_{i,-1}}{\sigma_i T^2} \end{pmatrix},$$

$\boldsymbol{\varepsilon}_i/\sigma_i$ and $\mathbf{f}$ are independently distributed as $(\mathbf{0}, \mathbf{I}_T)$, $\mathbf{s}_i = \mathbf{s}_{i,-1} + \boldsymbol{\varepsilon}_i$, and $\mathbf{s}_f = \mathbf{s}_{f,-1} + \mathbf{f}$.

The critical values of the CADF test can be computed by stochastic simulation for any fixed $T > 3$, and given distributional assumptions for the random variables ($\boldsymbol{\varepsilon}_i$,$\mathbf{f}$).

Theorem 3.2 *Let y_{it} be defined by (2.3) and consider the statistics $t_i(N,T)$ and $\tilde{t}_i(N,T)$ defined by (3.7) and (3.14), respectively. Suppose that assumptions 1-3 hold and $\bar{\gamma}$ tends to a finite non-zero limit as $N \to \infty$, then under (2.4) and as N and $T \to \infty$, $t_i(N,T)$ and $\tilde{t}_i(N,T)$ have the same sequential $(N \to \infty, T \to \infty)$ and joint $\left[(N,T)_j \to \infty\right]$ limit distributions, referred to as Cross-sectionally Augmented Dickey-Fuller (CADF) distribution given by*

$$CADF_{if} = \frac{\int_0^1 W_i(r)\,dW_i(r) - \boldsymbol{\psi}_{if}'\Lambda_f^{-1}\boldsymbol{\kappa}_{if}}{\left(\int_0^1 W_i^2(r)\,dr - \boldsymbol{\kappa}_{if}'\Lambda_f^{-1}\boldsymbol{\kappa}_{if}\right)^{1/2}}, \tag{3.29}$$

where

$$\Lambda_f = \begin{pmatrix} 1 & \int_0^1 W_f(r)\,dr \\ \int_0^1 W_f(r)\,dr & \int_0^1 W_f^2(r)\,dr \end{pmatrix}, \tag{3.30}$$

and

$$\boldsymbol{\psi}_{if} = \begin{pmatrix} W_i(1) \\ \int_0^1 W_f(r)\,dW_i(r) \end{pmatrix}, \boldsymbol{\kappa}_{if} = \begin{pmatrix} \int_0^1 W_i(r)\,dr \\ \int_0^1 W_f(r)\,W_i(r)\,dr \end{pmatrix}, \tag{3.31}$$

with $W_i(r)$ and $W_f(r)$ being independent standard Brownian motions. For the joint limit distribution to hold it is also required that as $(N,T)_j \to \infty$, $N/T \to k$, where k is a non-zero, finite constant.

Remark 3.1 *The critical values of $CADF_{if}$ can be computed by stochastic simulation assuming that*

$$\begin{pmatrix} \varepsilon_t \\ f_t \end{pmatrix} \backsim N(\mathbf{0}, \mathbf{I}_{N+1}), \text{ for } t = 1, 2, ..., T.$$

Remark 3.2 *From (3.29) it is clear that*

$$CADF_{if} = G\left(W_i, W_f\right), \tag{3.32}$$

where $G(.)$ is a general non-linear function common for all i. Therefore, $CADF_{if}$ and $CADF_{jf}$ are dependently distributed, with the same degree of dependence for all $i \neq j$.

Remark 3.3 *The random variables $CADF_{1f}, CADF_{2f}, ..., CADF_{Nf}$ form an exchangeable sequence. This follows from the fact that under Assumption 3, conditional on W_f the random variables $\{CADF_{if}\}$ are identically and independently distributed.*[6]

Remark 3.4 *The distribution of $CADF_{if}$ reduces to the standard DF distribution under $f_t = 0$. The singularlity of Λ_f in this case can be dealt with by use of generalized inverse. It is easily seen that*

$$\lim_{f\to 0} CADF_{if} = DF_i = \frac{\int_0^1 W_i(r)\, dW_i(r) - W_i(1) \int_0^1 W_i(r)\, dr}{\left[\int_0^1 W_i^2(r)\, dr - \left(\int_0^1 W_i(r)\, dr\right)^2\right]^{1/2}},$$

as required. Therefore, erroneous use of CADF in cases where a simple DF statistic would have sufficed, although inefficient, will not be invalid.

Remark 3.5 *The CADF test can be applied to test the unit root hypothesis in the case of a single time series when information on the cross-section average, $\bar{y}_t$, is available. For example, when testing for a unit root in the UK output one could use OECD output as a proxy for a possible common technological effect in output series across countries, and apply the CADF test instead of the standard DF test. Of course, different critical values would now apply, and must be computed using the CADF distribution given by (3.29).*

Remark 3.6 *In the above analysis we have opted for simple averages (i.e. $\bar{y}$) in dealing with the cross dependence problem. But it is clear that weighted averages could also be used instead. For example, $\bar{y}_t$ in the i^{th} CADF regression can be replaced with $y_{it}^* = \sum_{j=1}^N w_{ij} y_{jt}$ where w_{ij}, $j = 1, ..., N$ are weights specific to the series i. However, in the case where the factor loadings (γ_i) are independent random draws from a common distribution the choice of $\mathbf{w}_i = (w_{i1}, w_{i2}, ..., w_{iN})'$ has no effect on the asymptotic properties of the proposed tests so long as for each i, $\sum_{j=1}^N w_{ij}^2 \to 0$ as $N \to \infty$. There would be some small sample differences across the tests using different weighting schemes, but this is unlikely to be important.*

Remark 3.7 *Our assumption that $\bar{\gamma} \neq 0$, with a non-zero limit as $N \to \infty$ might be viewed by some as restrictive. But it is instructive to recall from (3.17) that in the case where $\bar{\gamma} = 0$, we have $\Delta\bar{y}_t = \bar{\varepsilon}_t$, and therefore $\Delta\bar{y}_t$ would tend to zero as $N \to \infty$, and $\bar{y}_t$ to a fixed constant for all t. In most economic and financial panels of interest this does not seem to be a very likely outcome. However, the case where $\bar{\gamma} \to 0$, as $N \to \infty$ could be of theoretical interest and is worth further consideration.*

[6] See, for example, Theorem 1.2.2 in Taylor, Daffer and Patterson (1985, p.13).

3.1 Critical Values of the Individual CADF Test

Figure 1 displays the simulated cumulative distribution function of the CADF statistic under the null hypothesis using 50,000 replications for $N = 100$ and $T = 500$. For comparison the simulated cumulative distribution function of the standard DF statistic is also provided.[7] We expect the simulated distribution to be very close to theoretical distribution given by (3.29). Perhaps not surprisingly the CADF distribution is more skewed to the left as compared to the standard DF distribution.

This is clearly reflected in the critical values of the two distributions summarized in Table A for the three standard cases considered in the literature: no intercept, intercept only, intercept and a linear trend.

Table A: **Critical Values of the CADF and DF Distributions**
(N=100,T=500, 50,000 replications)

	1%	2.5%	5%	10%
No intercept				
DF	-2.60	-2.23	-1.94	-1.61
CADF	-3.23	-2.88	-2.58	-2.24
Intercept				
DF	-3.46	-3.14	-2.86	-2.57
CADF	-3.84	-3.50	-3.23	-2.91
Linear Trend				
DF	-3.98	-3.68	-3.43	-3.14
CADF	-4.29	-3.97	-3.70	-3.39

Critical values of the individual CADF distribution for values of T and N in the range of 10 to 200 for the three standard cases (of no intercept and no trend, intercept only, and intercept and trend) are given in Tables 1a to 1c, respectively.

Another interesting aspect of the CADF distribution, which becomes important when the test is used in a panel data context, is the pair-wise dependence of the $CADF_{if}$ statistics across i, mentioned above. The simulated values of the simple pair-wise correlation coefficient, $Corr\,(CADF_{if},\ CADF_{if})$, together with simulated mean and standard deviation of the $CADF$ distribution for different values of N and T are given in Tables 2a to 2c for the three standard cases. These simulated moments are remarkably stable for different values of N and T in excess of 20. The simulated estimate of the correlation coefficient is around 0.03 for the intercept case and in the range 0.01-0.02 for the linear trend case, both quite small but non-zero.

[7] The series $y_{it} = y_{i,t-1} + f_t + \varepsilon_{it}$, for $i = 1, 2, ..., 100$, and $t = -50, -49, ..., 1, 2, ..., 500$ were first generated from $y_{i,-50} = 0$, with f_t and ε_{it} as iid $N(0, 1)$. Then 50,000 CADF regressions of Δy_{1t} on an intercept, $y_{1,t-1}$, $\bar{y}_{t-1}$ and $\Delta\bar{y}_t$ were computed over the sample $t = 1, 2, ..., 500$. Figure 1 plots the ordered values of the OLS t-ratios of $y_{1,t-1}$ in these regressions.

We also double-checked our computations by deriving the value of $Corr\left(CADF_{if},\ CADF_{if}\right)$ from the variances of the individual $CADF_{if}$ and the variance of the average value of these statistics, which we denote by $\overline{CADF}_f = N^{-1}\sum_{j=1}^{N} CADF_{jf}$. It is easily seen that

$$Corr\left(CADF_{if},\ CADF_{if}\right) = \left(\frac{1}{N-1}\right)\left(\frac{N \times Var(\overline{CADF}_f)}{Var(CADF_{if})} - 1\right).$$

Recall that $CADF_{if}$'s are identically distributed over i. These implied simulated values are also given in Tables 2a to 2c and confirm the estimates computed directly.

3.2 Normal Approximation to the Distribution of CADF

The CADF distribution, like the standard DF distribution, departs from normality in two important respects: It has a substantially negative mean and its standard deviation is less than unity, although not by a large amount. However, the distribution of a standardized version of the CADF statistic, defined by $\left[t_i\left(N,T\right) - E\left(CADF_{if}\right)\right]/\sqrt{Var\left(CADF_{if}\right)}$, looks remarkably like a standard normal distribution, where $E\left(CADF_{if}\right)$ and $\sqrt{Var\left(CADF_{if}\right)}$ are given in Tables 2a to 2c. The simulated density functions of the the standardized $CADF$ for the intercept and the linear trend cases, computed with $N = 100$, $T = 500$, and 50,000 replications are displayed in Figures 2 and 3, respectively. The skewness and Kurtosis -3 coefficients of the standardized $CADF$ distributions are 0.17 and 0.27 for the intercept case, and 0.06 and 0.29 for the linear trend case. They are quite small, although statistically highly significant. Nevertheless, the closeness of the approximation particularly for the left tail of the distribution suggests a relatively simple Normal test, once the mean and the standard deviation of CADF distribution is computed.

4 CADF Panel Unit Root Tests

Given that the null distribution of the individual CADF statistics are asymptotically independent of the nuisance parameters, the various panel unit root tests developed in the literature for the case of the cross-sectionally independent errors can also be applied to the present more general case. Here we focus on a generalization of the t-bar test proposed by IPS and consider a cross-sectionally augmented version of the IPS test based on

$$CIPS(N,T) = N^{-1}\sum_{i=1}^{N} t_i\left(N,T\right), \tag{4.33}$$

where $t_i\left(N,T\right)$ is the cross-sectionally augmented Dickey-Fuller statistic for the i^{th} cross section unit given by the t-ratio of the coefficient of $y_{i,t-1}$ in the CADF regression defined by (3.6).[8]

[8]The DF regressions can be augmented both for cross section dependence as well as for residual serial correlation. The serial correlation case will be discussed in the following section.

One could also consider combining the p-values of the individual tests as proposed by Maddala and Wu (1999) and Choi (2001). Examples are the inverse chi-squared (or Fisher) test defined by

$$P(N,T) = -2\sum_{i=1}^{N} \ln\left(p_{iT}\right), \tag{4.34}$$

where p_{iT} is the p-value corresponding to the unit root test of the i^{th} individual cross section unit. Another possibility would be to use the inverse normal test defined by

$$Z(N,T) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \Phi^{-1}\left(p_{iT}\right). \tag{4.35}$$

Here we focus on the t-bar version of the panel unit root test, (4.33), and consider the mean deviations

$$D(N,T) = N^{-1}\sum_{i=1}^{N}\left[t_i\left(N,T\right) - CADF_{if}\right],$$

where $CADF_{if}$ is the stochastic limit of $t_i\left(N,T\right)$ as N and T tend to infinity such that $N/T \rightarrow k$ $(0 < k < \infty)$. See (3.29). It seems reasonable to expect that $D(N,T) = o_p(1)$ for N and T sufficiently large. This conjecture would clearly hold in the case where $\bar{t}_i\left(N,T\right)$ have finite moments for all N and T above some given threshold values, say N_0, and T_0. However, such moment conditions are difficult to establish even under cross-section independence. (see IPS)

One possible method of dealing with these technical difficulties would be to base the t-bar test on a suitably truncated version of the CADF statistics. The simulations reported in Section 3.2 suggest that the standardized version of these statistic are very close to being standard Normal with finite first and second order moments. Therefore, for the purpose of the panel unit root test it would be equally valid to base the test on an average of the truncated versions of $t_i(N,T)$, say $t_i^*(N,T)$, where

$$\begin{cases} t_i^*(N,T) = t_i(N,T), \text{ if } & -K_1 < t_i(N,T) < K_2, \\ t_i^*(N,T) = -K_1, \text{ if } & t_i(N,T) \leq -K_1, \\ t_i^*(N,T) = K_2, \text{ if } & t_i(N,T) \geq K_2. \end{cases} \tag{4.36}$$

where K_1 and K_2 are positive constants that are sufficiently large so that $Pr\left[-K_1 < t_i(N,T) < K_2\right]$ is sufficiently large, say in excess of 0.9999. Using the normal approximation of $t_i(N,T)$ as a crude benchmark we would have

$$K_1 = -E\left(CADF_{if}\right) - \Phi^{-1}(\varepsilon/2)\sqrt{Var\left(CADF_{if}\right)},$$

and

$$K_2 = E\left(CADF_{if}\right) + \Phi^{-1}(1-\varepsilon/2)\sqrt{Var\left(CADF_{if}\right)},$$

where ε is a sufficiently small positive constant. For example, setting $\varepsilon = 1\times 10^{-6}$, in the case of models without an intercept or trend (using the mean and the standard deviations in Table

2a) we would have $K_1 = 0.98 + 4.8917(1.05) = 6.12$, and $K_2 = -0.98 + 4.8917(1.05) = 4.16$. Similarly, for models with an intercept we have $K_1 = 6.19$, and $K_2 = 2.61$, and finally for models with a linear trend we obtain $K_1 = 6.42$, and $K_2 = 1.70$

The associated truncated panel unit root test is now given by

$$CIPS^*(N,T) = N^{-1}\sum_{i=1}^{N} t_i^*(N,T). \tag{4.37}$$

Since, by construction all moments of $t_i^*(N,T)$ exist it then follows that

$$CIPS^*(N,T) = N^{-1}\sum_{i=1}^{N} CADF_{if}^* + o_p(1), \tag{4.38}$$

where $CADF_{if}^*$ is given by

$$\begin{cases} CADF_{if}^* = CADF_{if}, \text{ if } \; -K_1 < CADF_{if} < K_2, \\ CADF_{if}^* = -K_1, \text{ if } \; CADF_{if} \leq -K_1, \\ CADF_{if}^* = K_2, \text{ if } \; CADF_{if} \geq K_2, \end{cases} \tag{4.39}$$

and $CADF_{if}$ is defined by (3.29) in Theorem 3.2. The distributions of the average CADF statistic or its truncated counterpart, $\overline{CADF}^* = N^{-1}\sum_{i=1}^{N} CADF_{if}^*$, are non-standard even for sufficiently large N. This is due to the dependence of the individual $CADF_{if}$ variates on the common process W_f which invalidates the application of the standard central limit theorems to $\overline{CADF}$ or $\overline{CADF}^*$, and is in contrast to the results obtained by IPS under cross-section independence where a standardized version of $\overline{CADF} = N^{-1}\sum_{i=1}^{N} CADF_{if}$, was shown to be normally distributed for N sufficiently large. Nevertheless, it is possible to show that $\overline{CADF}^*$ converges in distribution as $N \to \infty$, without any need for further normalization. Recall that $CADF_{if} = G(W_i, W_f)$, $i = 1, 2, ..., N$, where $W_1, W_2, ..., W_N$ and W_f are *i.i.d.* Brownian motions. Similarly, $CADF_{if}^*$ defined by (4.39) will be a non-linear function of W_i and W_f and hence conditional on W_f , $CADF_{if}^*$ will be independently distributed across i. Therefore, since by construction

$$E\left|CADF_{if}^*\right| < \infty,$$

it follows that

$$N^{-1}\sum_{i=1}^{N} CADF_{if}^* \overset{a.s.}{\to} \pi_2 K_2 - \pi_1 K_1 + E\left(CADF_{1f} \mid W_f, \; -K_1 < CADF_{1f} < K_2\right), \tag{4.40}$$

where $\pi_1 = \Pr\left(CADF_{if} \leq -K_1 \mid W_f\right)$ and $\pi_2 = \Pr\left(CADF_{if} \geq K_2 \mid W_f\right)$. This result simplifies further if we could also establish that $E\left|CADF_{if}\right| < \infty$, a property that we conjecture to be true. By letting K_1 and $K_2 \to \infty$, and noting that in this case $\pi_2 K_2 - \pi_1 K_1 \to 0$, we have

$$N^{-1}\sum_{i=1}^{N} CADF_{if} \overset{a.s.}{\to} E\left(CADF_{1f} \mid W_f\right).$$

The above results establish that the $\overline{CADF}^*$ converges almost surely to a distribution which depends on K_1, K_2 and W_f. This distribution does not seem analytically tractable, but can be readily simulated using (4.37). We simulated the distribution of $CIPS^*$ setting $N = 100$, $T = 500$, and using 50,000 replications under the following cases:

1. Models without intercepts or trends (I), with $K_1 = 6.12$, and $K_2 = 4.16$,

2. Models with intercept only (II), with $K_1 = 6.19$, and $K_2 = 2.61$,

3. Models with a linear trend (III), with $K_1 = 6.42$, and $K_2 = 1.70$.

The simulated density functions for these three cases are displayed in Figures 4-6. All the three densities show marked departures from normality, although the extent of the departure depends on the nature of the deterministic included in the model. The density in the case of the model without any deterministics show the greatest degree of departure from normality and is in fact bimodal. The density for the other two models are uni-modal but are highly skewed. The density for the model with a linear trend is closest to being normal. This pattern of departures from normality is in accordance with the estimates of pair-wise correlation coefficients of the individual CADF statistics reported in Tables 2a-2c. The larger the value of this correlation coefficient, the greater one would expect the density to depart from normality. Recall that the asymptotic correlation coefficients of the individual CADF statistics are 0.10, 0.03 and 0.01 for the models I to III, respectively.

We carried out the same analysis for the non-truncated version, using $CIPS$ defined by (4.33), and obtained identical results. The finite sample distributions of $CIPS^*(N,T)$ and $CIPS(N,T)$ differ only for very small values of T and are indistinguishable for $T > 20$. The comparative small sample performances of the $CIPS$ and the $CIPS^*$ tests will be considered in Section 6.

The 1%, 5% and 10% critical values of $\overline{CADF}^*$ and $\overline{CADF}$ are given in Tables 3a-3c for models I-III, respectively. In most cases the critical values for the two versions of the CIPS test are identical and only one value is reported. In cases where the two critical values differ the truncated version is included in brackets.

Similar arguments also apply to the other forms of the panel unit root tests given by (4.34) and (4.35). The cross sectionally augmented versions of these statistics, where the rejection probabilities, p_{iT}, are computed using the CADF regressions, (3.6), will be denoted by $CP(N,T)$ and $CZ(N,T)$. Note that in the presence of cross section dependence these statistics are no longer asymptotically normally distributed and their critical values must be obtained by stochastic simulations. The 1%, 5% and 10% critical values of $CP(N,T)$ and $CZ(N,T)$ are computed by Mutita Akusuwan for all pairs of $N, T = 10, 15, 20, 30, 50, 70, 100, 200$, and are available from the author on request.

5 Case of Serially Correlated Errors

The CIPS* testing procedure can be readily extended to the case where in addition to the cross dependence, the individual-specific error terms are also serially correlated.

5.1 Alternative Residual Serial Correlation Models with Cross Dependence

The residual serial correlation can be modelled in a number of different ways, directly via the idiosyncractic components, the common effects, or a mixture of the two. To simplify the exposition we shall confine our analysis to stationary first-order autoregressive processes and consider three general types of specifications.[9]. In the case where only the idiosyncratic components are serially correlated we have

$$u_{it} = \gamma_i f_t + v_{it}, \tag{5.41}$$

where

$$v_{it} = \rho_i v_{i,t-1} + \varepsilon_{it}, \quad |\rho_i| < 1 \tag{5.42}$$

and $\varepsilon_{it} \sim i.i.d.(0, \sigma_i^2)$. In conjunction with (2.3) this would yield the following (time series) augmented Dickey-Fuller regression

$$\Delta y_{it} = -\mu_i \beta_i (1 - \rho_i) + \beta_i y_{i,t-1} + \rho_i (1 + \beta_i) \Delta y_{i,t-1} + \gamma_i \left(f_t - \rho_i f_{t-1} \right) + \varepsilon_{it}, \tag{5.43}$$

with cross-sectionally dependent errors.

In the case where the residual serial correlation is confined to the common effects we have

$$u_{it} = \gamma_i f_t + \varepsilon_{it}, \tag{5.44}$$

where

$$f_t = \lambda f_{t-1} + \xi_t, \quad |\lambda| < 1, \tag{5.45}$$

and ξ_t are serially uncorrelated with mean zero and a constant variance.[10] The serial correlation in the common effects induces moving average errors in the individual ADF regressions, and we have

$$\Delta y_{it} = -\mu_i \beta_i (1 - \lambda) + \beta_i (1 - \lambda) y_{i,t-1} + \lambda (1 + \beta_i) \Delta y_{i,t-1} + \gamma_i \xi_t + \varepsilon_{it} - \lambda \varepsilon_{i,t-1}. \tag{5.46}$$

In this case the cross-section dependence is characterized through the residual common effects, ξ_t.

A third possibility would be to model the residual serial correlation first as

$$u_{it} = \rho_i \ u_{i,t-1} + \eta_{it}, \quad |\rho_i| < 1, \text{ for } i = 1, 2, ..., N, \tag{5.47}$$

and then allow for the cross section dependence by assuming a one-factor model for the residuals

$$\eta_{it} = \gamma_i f_t + \varepsilon_{it}. \tag{5.48}$$

Under this specification we have

$$\Delta y_{it} = -\mu_i \beta_i (1 - \rho_i) + \beta_i (1 - \rho_i) y_{i,t-1} + \rho_i (1 + \beta_i) \Delta y_{i,t-1} + \gamma_i f_t + \varepsilon_{it}. \tag{5.49}$$

---

[9] The analysis can be readily extended to higher order processes.

[10] The case of non-stationary common effects will not be considered here.

5.2 Individual-Specific CADF Statistics for the Serially Correlated Case

All the three models yield the same ADF regressions, but with different error specifications and parameter heterogeneity. The asymptotic theory to be developed in this section can be adapted to deal with *all* the three specifications, but to save space here we focus on (5.49). Also this specification has the advantage that it can be readily generalized to higher order processes, to deal with the moving average error processes in (5.43) and (5.46). We shall also confine our attention to the case where the autoregressive coefficients, ρ_i are homogeneous across i, but shall consider the implications of relaxing this assumption using Monte Carlo simulations. The mathematical details become much more complicated if ρ_i is allowed to differ across i.

To deal with the unobserved common effects, f_t, we first note that in this case under the unit root hypothesis, $\beta_i = 0$, we have (using (5.49))

$$\Delta y_{it} = \rho \Delta y_{i,t-1} + \gamma_i f_t + \varepsilon_{it},$$

and

$$f_t = \bar{\gamma}^{-1} \left(\Delta \bar{y}_t - \rho \Delta \bar{y}_{t-1} \right) - \bar{\gamma}^{-1} \bar{\varepsilon}_t.$$

Hence, for sufficiently large N, and under our assumption that $\bar{\gamma}$ tends to a non-zero limit as $N \to \infty$, the common effects can be proxied by a linear combinations of $\Delta \bar{y}_t$ and $\Delta \bar{y}_{t-1}$. In addition the DF regressions must be augmented for residual serial correlation and the lagged levels of the cross section means of the processes, namely $\Delta y_{i,t-1}$ and $\bar{y}_{t-1}$. Accordingly, we propose running the following CADF regressions which are augmented to asymptotically filter out the effects of both the cross section and the time dependence patterns in the residuals:

$$\Delta \mathbf{y}_i = b_i \mathbf{y}_{i,-1} + \bar{\mathbf{W}}_i \mathbf{c}_i + \mathbf{e}_i,$$

where $\bar{\mathbf{W}}_i = \left(\Delta \mathbf{y}_{i,-1}, \Delta \bar{\mathbf{y}}, \Delta \bar{\mathbf{y}}_{-1}, \boldsymbol{\tau}_T, \bar{\mathbf{y}}_{-1} \right)$ is a $T \times 5$ matrix of observations defined in Section 3. The individual CADF statistics are given by

$$t_i(N,T) = \frac{\Delta \mathbf{y}_i' \bar{\mathbf{M}}_i \mathbf{y}_{i,-1}}{\hat{\sigma}_i \left(\mathbf{y}_{i,-1}' \bar{\mathbf{M}}_i \mathbf{y}_{i,-1} \right)^{1/2}}, \tag{5.50}$$

where

$$\hat{\sigma}_i^2 = \frac{\Delta \mathbf{y}_i' \mathbf{M}_{i,w} \Delta \mathbf{y}_i}{T-6},$$

$\bar{\mathbf{M}}_i = \mathbf{I}_T - \bar{\mathbf{W}}_i \left(\bar{\mathbf{W}}_i' \bar{\mathbf{W}}_i \right)^{-1} \bar{\mathbf{W}}_i'$, $\mathbf{M}_{i,w} = \mathbf{I}_T - \mathbf{G}_i (\mathbf{G}_i' \mathbf{G}_i)^{-1} \mathbf{G}_i'$, and $\mathbf{G}_i = \left(\mathbf{y}_{i,-1}, \bar{\mathbf{W}}_i \right)$. Similarly, we have the Lagrange multiplier version defined

$$\tilde{t}_i(N,T) = \frac{\sqrt{T-5}\, \Delta \mathbf{y}_i' \bar{\mathbf{M}}_i \mathbf{y}_{i,-1}}{\left(\Delta \mathbf{y}_i' \bar{\mathbf{M}}_i \Delta \mathbf{y}_i \right)^{1/2} \left(\mathbf{y}_{i,-1}' \bar{\mathbf{M}}_i \mathbf{y}_{i,-1} \right)^{1/2}}. \tag{5.51}$$

As with the serially uncorrelated case both versions of the CADF tests are asymptotically equivalent, although $\tilde{t}_i(N,T)$ is simpler to study analytically.

To establish the asymptotic invariance of the above CADF statistics to the coefficients of the common effects, γ_i, we first note that under $\beta_i = 0$

$$\Delta y_{it} = \rho \Delta y_{i,t-1} + \delta_i \left(\Delta \bar{y}_t - \rho \Delta \bar{y}_{t-1} \right) + \left(\varepsilon_{it} - \delta_i \bar{\varepsilon}_t \right), \tag{5.52}$$

or

$$\Delta y_{it} = \delta_i \Delta \bar{y}_t + z_{it}(\rho) - \delta_i \bar{z}_t(\rho), \tag{5.53}$$

and

$$y_{it} = (y_{i0} - \delta_i \bar{y}_0) + \delta_i \bar{y}_t + s_{z,it} - \delta_i \bar{s}_{zt}, \tag{5.54}$$

where

$$z_{it}(\rho) = (1 - \rho L)^{-1} \varepsilon_{it}, \tag{5.55}$$

$$s_{z,it} = \sum_{j=1}^{t} z_{ij}(\rho), \;\; \bar{s}_{zt} = N^{-1} \sum_{i=1}^{N} s_{z,it}, \tag{5.56}$$

and L is a one-period lag operator. Using these results we now have the following generalizations of (3.19) and (3.24):

$$\Delta \mathbf{y}_i = \rho \Delta \mathbf{y}_{i,-1} + \delta_i \left(\Delta \bar{\mathbf{y}} - \boldsymbol{\rho} \Delta \bar{\mathbf{y}}_{-1} \right) + \left(\boldsymbol{\varepsilon}_i - \delta_i \bar{\boldsymbol{\varepsilon}} \right), \tag{5.57}$$

and

$$\mathbf{y}_{i,-1} = (y_{i0} - \delta_i \bar{y}_0) \boldsymbol{\tau} + \delta_i \bar{\mathbf{y}}_{-1} + \left(\mathbf{s}_{zi,-1} - \delta_i \bar{\mathbf{s}}_{z,-1} \right), \tag{5.58}$$

which if used in (5.51) yields (under $\beta_i = 0$)

$$\tilde{t}_i(N,T) = \frac{\frac{\boldsymbol{\upsilon}_i' \bar{\mathbf{M}}_i \mathfrak{s}_{zi,-1}}{T}}{\left(\frac{\boldsymbol{\upsilon}_i' \bar{\mathbf{M}}_i \boldsymbol{\upsilon}_i}{T-5} \right)^{1/2} \left(\frac{\mathfrak{s}_{zi,-1}' \bar{\mathbf{M}}_i \mathfrak{s}_{zi,-1}}{T^2} \right)^{1/2}}, \tag{5.59}$$

where as before $\boldsymbol{\upsilon}_i = \left(\boldsymbol{\varepsilon}_i - \delta_i \bar{\boldsymbol{\varepsilon}} \right) / \omega_i \sim (\mathbf{0}, \mathbf{I}_T)$, ω_i is defined by (3.22), and

$$\mathfrak{s}_{zi,-1} = \left(\mathbf{s}_{zi,-1} - \delta_i \bar{\mathbf{s}}_{z,-1} \right) / \omega_i.$$

The elements of $\mathbf{s}_{zi,-1}$ and $\bar{\mathbf{s}}_{z,-1}$ are defined in (5.56).

The exact sample distribution of $\tilde{t}_i(N,T)$ depends on δ_i, $\bar{\gamma}$ and ρ, but as stated in the following theorem this dependence vanishes for N and $T \to \infty$, such that $N/T \to k$, where k is a finite, non-zero constant.

Theorem 5.1 *Let y_{it} be defined by (5.49) with $|\rho_i| = |\rho| < 1$, and consider the statistics $t_i(N,T)$ and $\tilde{t}_i(N,T)$ defined by (5.50) and (5.51), respectively. Suppose that Assumptions 1-3 hold and $\bar{\gamma}$ tends to a finite non-zero limit as $N \to \infty$, then under (2.4) and as N and $T \to \infty$, $t_i(N,T)$ and $\tilde{t}_i(N,T)$ have the same sequential $(N \to \infty, T \to \infty)$ and joint $\left[(N,T)_j \to \infty \right]$ limit distributions given by (3.29), obtained under $\rho = 0$.*

For a proof see Section A.3 in the Appendix. It is worth noting that Theorem 5.1 also holds under (5.43) and (5.46), so long as $|\lambda| < 1$.[11]

This theorem establishes that the familiar Augmented Dickey-Fuller regression results in pure time series contexts also applies to cross sectionally augmented regressions. Although, our proof assumes a first-order error process, the approach readily extends to higher order processes. For example, for an $AR(p)$ error specification the relevant individual CADF statistics will be given by the OLS t-ratio of b_i in the following p^{th} order cross-section/time-series augmented regression:

$$\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + \sum_{j=0}^{p} d_{ij} \Delta \bar{y}_{t-j} + \sum_{j=1}^{p} \delta_{ij} \Delta y_{i,t-j} + e_{it}. \quad (5.60)$$

This testing procedure also readily extend to models containing linear trends. Clearly, the same critical values reported in Tables 3a-3c will also be applicable here.

5.3 Panel Unit Root Tests for Panels with Serially Correlated Errors

It is now relatively easy to construct panel unit root tests that simultaneously take account of cross-section dependence and residual serial correlation. Once again we focus on the truncated version of the CIPS statistic given by (4.37), with $t_i^*(N,T)$ computed using the cross-section/time–series augmented regression, (5.60), subject to the truncation scheme defined by (4.36).[12] Using theorem 5.1 and noting that the result of the theorem applies equally to the truncated version of the CADF statistics we have

$$t_i^*(N,T) = CADF_{if}^* + o_p(1).$$

Hence

$$CIPS^* \overset{(N,T)_j}{\backsim} N^{-1} \sum_{i=1}^{N} CADF_{if}^*,$$

and $CIPS^*$ in the case of serially correlated errors has the same limit distribution as (4.40) obtained under $\rho = 0$ and the critical values reported in Tables 3a-3c also applies equally to the serially correlated case.

6 Small Sample Performance: Monte Carlo Evidence

In this section we consider the small sample performance of the cross sectionally augmented unit root tests proposed in the paper using Monte Carlo techniques. Initially we shall consider

[11] Details of the proof are very similar and can be obtained from the author on request.

[12] The order of augmentation, p, can be estimated using model selection criteria such as Akaike or Schwartz applied to the underlying time series specification, namely (5.60) without the cross-section variables, $\bar{y}_{t-1}$, $\Delta \bar{y}_{t-j}$, $j = 1, 2, .., p$.

dynamic panels with fixed effects and cross section dependence, but without residual serial correlation or linear trends. The data generating process (DGP) in this case is given by

$$y_{it} = (1 - \phi_i)\,\mu_i + \phi_i y_{i,t-1} + u_{it},\ i = 1, 2, ..., N, t = -51, -50, ..., 1, 2, ..., T; \tag{6.61}$$

where

$$u_{it} = \gamma_i f_t + \varepsilon_{it}, \quad f_t \sim iidN(0, 1), \tag{6.62}$$

and

$$\varepsilon_{it} \sim i.i.d.N(0, \sigma_i^2), \text{ with } \sigma_i^2 \sim iidU\,[0.5, 1.5]\,. \tag{6.63}$$

We shall consider two levels of cross section dependence where we generate $\gamma_i \sim iidU\,[0, 0.20]$ as an example of "low cross section dependence", and $\gamma_i \sim iidU\,[-1, 3]$ to represent the case of "high cross section dependence". The average pair-wise cross correlation coefficient of u_{it} and u_{jt} under these two scenarios are 1% and 50%, respectively, and cover a wide range of values applicable in practice.

To examine the impact of the residual serial correlation on the proposed tests we considered a number of experiments where the errors ε_{it} were generated as

$$\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it},\, e_{it} \sim i.i.d.N(0, \sigma_i^2),\, \sigma_i^2 \sim iidU\,[0.5, 1.5]\,, \tag{6.64}$$

with $\rho_i \sim iidU\,[0.2, 0.4]$, as an example of positive residual serial correlations, and $\rho_i \sim iidU\,[-0.4, -0.2]$, as an example of negative residual serial correlations. This yields the augmented ADF model given by (5.43). These experiments were also carried out under low and high cross section dependence scenarios. This DGP differs from model (5.49) that underlies the theoretical derivations in Section 5, and is intended to check the robustness of our analysis to alternative residual serial correlation models. It also allows the residual serial correlation coefficients, ρ_i, to differ across i; thus providing an opportunity to check the robustness of our results to such heterogeneities.

In a third set of experiments we allow for deterministic trends in the DGP and the CADF regressions. For this case y_{it} were generated as follows:

$$y_{it} = \mu_i + (1 - \phi_i)\,\delta_i t + \phi_i y_{i,t-1} + u_{it},$$

with $\mu_i \sim iidU\,[0.0, 0.02]$ and $\delta_i \sim iidU\,[0.0, 0.02]$. This ensures that y_{it} has the same average trend properties under the null and the alternative hypotheses. The errors, u_{it}, were generated according to (6.62), (6.63) and (6.64) for different values of ρ_i as set out above.

Size and power of the tests were computed under the null $\phi_i = 1$ for all i, and the heterogeneous alternatives $\phi_i \sim iidU\,[0.85, 0.95]$, using 1,000 replications per experiment.[13] The tests were one-sided with the nominal size set at 5%, and were conducted for all combinations of N and $T = 10, 20, 30, 50, 100$. All the parameters, μ_i , δ_i, ϕ_i, ρ_i, σ_i^2, and γ_i were generated independently of the errors, e_{it} (ε_{it}) and f_t; with f_t also generated independently of e_{it} (ε_{it}).

[13] Under the alternative hypothesis μ_i are drawn as $\mu_i \sim iidU[0, 0.02]$.

6.1 Size Distortion of the Standard Panel Unit Root Tests

Before reporting the results for the proposed cross sectionally augmented tests, it would be helpful first to examine the extent to which the size of the standard panel unit root tests (that assume cross section independence) are distorted in the presence of cross section dependence. Table 4 reports the empirical sizes of the IPS, truncated IPS, the inverse chi-squared (P), and the inverse normal (Z) tests when the DGP is subject to cross section dependence with serially uncorrelated errors as defined by (6.61) and (6.62).[14] All these tests are based on simple DF regressions and utilize the individual DF statistics, or the associated rejection probabilities.[15] The IPS statistic is the familiar standardized $t - bar$ statistic define by (see IPS (2003)) Standard Panel

$$IPS(N,T) = \frac{\sqrt{N}\left\{t\text{-}bar_{NT} - E\left[t_{iT} \mid \beta_i = 0\right]\right\}}{\sqrt{Var\left[t_{iT} \mid \beta_i = 0\right]}} \overset{T,N}{\Longrightarrow} N(0,1).$$

where $t\text{-}bar_{NT} = N^{-1}\sum_{i=1}^{N} t_{iT}$, and t_{iT} is the t-ratio of the estimated coefficient of $y_{i,t-1}$ in the OLS regression of Δy_{it} on an intercept and $y_{i,t-1}$.[16] The truncated version of the IPS test uses the same formula as above but replaces t_{iT} with the individually truncated statistic, t_{iT}^{*}, defined by (4.36) with $K_1 = 6.19$, and $K_2 = 2.61$. In the case of P and Z tests we report two sets of results: one set based on normal approximations as originally proposed by Maddala and Wu (1999), and Choi (2001), and another set based on empirical critical values obtained from the simulated distribution of these statistics under the null hypothesis. We refer to the latter versions of these tests as $\tilde{P}$ and the $\tilde{Z}$ tests.[17]

As to be expected the extent of over-rejection of the tests very much depends on the degree of cross section dependence. Under the low cross section dependence the different version of the IPS and the Z tests perform reasonably well. The standard P test tends to over-reject for small values of T, but the normal approximation begins to work as T is allowed to increase. Overall, when the cross section dependence is low all tests (possibly except for the P test) have the correct size, which is in line with the results in the literature, reported, for example, by Choi (2001). However, the situation is very different if we consider the results under the high cross section scenario. In this case all the tests exhibit substantial size distortions, with the extent of size distortion tending to increase with N and T. This is true of all the tests except for the P test which behaves rather erratically. It is clear that the standard panel unit root tests that do not allow for cross section dependence can be seriously biased if the the degree of cross section dependence is sufficiently large. It would now be interesting to see if the cross sectionally augmented versions of these tests can resolve their size distortions under the high cross section dependence scenario.

[14]In calculation of P and Z statistics the rejection probabilities, p_{iT}, are truncated to lie in the range $[0.000001, 0.999999]$, in order to avoid very extreme values affecting these test statistics. This is in effect a kind of truncation, similar to the truncated version of the IPS statistics.

[15]See (4.34), (4.35), and the notes to Table 4 for further details.

[16]The IPS and other panel unit root tests can be readily adapted for use with unbalanced panels where the available time periods differ across i. In the case of standard IPS test this generalization is considered in Im, Pesaran and Shin (2003).

[17]The critical values of the $\tilde{P}$ and $\tilde{Z}$ tests are available from the author on request.

6.2 Size and Power in the Case of Models with Serially Uncorrelated Errors

The tests to be considered are the cross sectionally augmented IPS test, $CIPS(N,T)$, and its truncated version, $CIPS^*(N,T)$, and the cross sectionally augmented versions of the inverse chi-squared and the inverse normal tests, denoted by $CP(N,T)$ and $CZ(N,T)$, respectively. The $CIPS$ and $CIPS^*$ statistics are defined by (4.33) and (4.37), respectively, and are very simple to compute. In contrast, the computation of the CP and CZ statistics require the estimation of individual-specific rejection probabilities by stochastic simulations. In particular, the cross-sectionally augmented inverse chi-squared test statistic is given by

$$CP\left(N,T\right) = -2\sum_{i=1}^{N} \ln\left[\hat{p}_i\left(N,T\right)\right], \tag{6.65}$$

where the rejection probabilities are computed as

$$\widehat{p}_i(N,T) = \frac{1}{S}\sum_{s=1}^{S} I\left[t_i(N,T) - \mathfrak{CDF}^{(s)}\right] \tag{6.66}$$

$\mathfrak{CDF}^{(s)}$ is the s^{th} random draw from the distribution of $\mathfrak{CDF}$, and S is the number of replications used to compute $\widehat{p}_i(N,T)$, which we also set equal to 50,000. $I\left[A\right]$ is the indicator function that takes the value of 1 when $A > 0$, and 0 otherwise. Similarly,

$$CZ(N,T) = \frac{1}{\sqrt{N}}\sum_{i=1}^{N} \Phi^{-1}\left[\hat{p}_i(N,T)\right]. \tag{6.67}$$

To avoid very extreme values the rejection probabilities were truncated to lie in the range $[0.000001, 0.999999]$.

The size and power characteristics of these tests are summarized in Tables 5a and 5b, respectively. There are no evidence of size distortions in the case of the $CIPS$, $CIPS^*$, and the versions of the CP, and CZ tests (denoted by $\widetilde{CP}$, and $\widetilde{CZ}$) that use the correct critical values. Not surprisingly, the use of normal approximations for the CP, and CZ tests does not work here since due to the cross section dependence these test statistics are not normally distributed even for sufficiently large N and T. Therefore, it is only valid to consider a power comparison of $CIPS$, $CIPS^*$, $\widetilde{CP}$, and $\widetilde{CZ}$ tests, as summarized in Table 5b. It is clear that $\widetilde{CP}$ test is generally dominated by the other three tests which are very similar indeed. None of the tests exhibit much power when $T = 10$, irrespective of the size of N. Only when T is increased to 20 and beyond one can begin to see the benefit of increasing N on the power of the tests. Finally, in the present simple case of serially uncorrelated residuals little seems to be gained by the truncation procedure. But as we shall see the truncated version of CIPS test could avoid size distortions in the case of models with residual serial correlations and linear trends.

6.3 Size and Power in the Case of Models with Serially Correlated Errors

In this case the cross sectionally augmented tests ($CIPS$, $CIPS^*$, $\widetilde{CP}$, and $\widetilde{CZ}$) are computed both for the basic CADF regressions without time series augmentation (which we denote by CADF(0)), and the CADF regressions are augmented with lagged changes of y_{it} and $\bar{y}_t$, as in (5.60), which we refer to by CADF(p), where p is the order of the time series augmentation. We computed the tests for $p = 0, 1$, and focussed on the high cross section scenario. The size and power results for the experiments with positive residual serial correlation are summarized in Tables 6a and 6b, and the ones for negative residual serial correlation are given in Tables 7a and 7b. As to be expected significant size distortions will be present if CADF regressions are not augmented for the time series dependence. There are substantial under rejections for positive residual serial correlation, and substantial over-rejections in the case of negative residual serial correlations. But the test sizes stabilize at around 5% when the CADF regressions are augmented with $\Delta y_{i,t-1}$. Recall that $\Delta \bar{y}_{t-1}$ is already included in the base CADF regressions. Irrespective of whether the residual serial correlations are positive or negative, the tests based on CADF(1) regressions tend to have the correct size. There is, however, some evidence that for small T (less than 20 in these experiments) the $\widetilde{CP}$ test, and to a lesser extent, the $CIPS$ test are over-sized. But the truncated version of the $CIPS$ does not seem to suffer from this problem even for T as small as 10. The truncation of the extreme individual CADF statistics seem to have paid out in the present application where T is very small relative to the number of parameters of the underlying CADF(1) regressions.

Focussing on $CIPS^*$ and $\widetilde{CZ}$ we note from Tables 6b and 7b that both tests have very similar power properties. Neither of the tests seem to have any power for $T = 10$ or less, and as in the serially uncorrelated case, the power does not rise with N if T is too small. However, with $T = 20$ or higher the power of both tests begin to rise quite rapidly with N. The tests tend to show higher power for negative as compared to positive residual serial correlations. Finally, there is very little to choose between the two tests, although as noted earlier the $CIPS^*$ statistic is much simpler to compute.

6.4 Size and Power in the Case of Models with Linear Trends and Serially Correlated Errors

Size and power of $CIPS$, $CIPS^*$, $\widetilde{CP}$, and $\widetilde{CZ}$ tests in the case of models with linear deterministic trends are summarized in Tables 8-10. Tables 8a and 8b give the results for models without residual serial correlation and show that all the various test continue to have sizes very close to the nominal value of 5%. However, as to be expected the inclusion of linear trends in the CADF regressions come at the cost of a lower power. We now need T to be 30 or more before power begins to increase with N. For example, when $T = 20$ the power of the tests stays around 7% irrespective of the value of N. But when $T = 50$ the power of the CIPS test rises from 18% to 62% as N is increased from 10 to 100. Once again the $\widetilde{CP}$ test is dominated by the other three tests which have very similar power characteristics.

The results for the linear trend case combined with residual serial correlation are presented in Tables 9 and 10. Table 9a and 9b give the results for positive residual correlations, and Tables 10a and 10b summarize the results for the negative residual serial correlations. The sizes of the $CIPS^*$ and $\widetilde{CZ}$ tests continue to be satisfactory, even for $T = 10$ once the augmentation for the residual serial correlation is implemented (see the results under ACDF(1)). In contrast the $CIPS$ and $\widetilde{CP}$ tests are grossly over-sized when $T = 10$, and gets worse as N is increased. Note that in the case of CADF(1) regressions with linear trends the number of parameters being estimated is 7, and with only 3 degrees of freedom remaining the non-truncated individual $CADF_i(1)$ statistics might not have moments, which could be the reason why the $CIPS$ test breaks down. The application of the truncation procedure fixes the lack of the moment problem and renders the truncated $CIPS$ test valid even when the degrees of freedom of the underlying CADF regressions is as low as 3. Similarly, the $\widetilde{CZ}$ statistic overcome the problem of extreme values by using the inverse probability transformation and by the fact that the rejection probabilities used in the construction of $\widetilde{CZ}$ are truncated to avoid very extreme values. See (6.67).

7 Concluding Remarks

This paper presents a new and simple procedure for testing unit roots in dynamic panels subject to (possibly) cross sectionally dependent as well as serially correlated errors. The procedure involves augmenting the standard ADF regressions for the individual series with current and lagged cross section averages of all the series in the panel. This is a natural extension of the DF approach to dealing with residual serial correlation where lagged changes of the series are used to filter out the time series dependence when T is sufficiently large. Here we propose to use cross section averages to perform a similar task in dealing with the cross dependence problem. Our approach should be seen as providing a simple alternative to the orthogonalization type procedures advanced in the literature by Bai and Ng (2002), Moon and Perron (2003), and Phillips and Sul (2002). Although we have provided extensive simulation results in support of our proposed tests, further simulation experiments are needed to shed light on the relative merits of the various panel unit roots that are now available in the literature.

Our analysis and testing approach can also be extended in a number of directions. One obvious generalization is to allow for a richer pattern of cross dependence by including additional common factors in the model. This is likely to pose additional technical difficulties, but can be dealt with by augmenting the individual ADF regressions with additional cross section averages formed over sub-groups, such as regions, sectors or industries. Another worthwhile extension would be to consider cross section augmented versions of unit root tests due to Elliott et al. (1996), Fuller and Park (1995), and Leybourne (1995). Such tests are likely to have better small sample power properties.

In their analysis Bai and Ng (2002) also consider the possibility of unit root in the common factors. However, under their set up the unit properties of the common factor(s) and the idiosyncractic component of the individual series are unrelated. As a result they are able to carry out separate unit root tests in the common and the idiosyncractic components. The specification used by Bai and Ng is given by the static factor model (assuming one factor for ease of comparison)

$$y_{it} = \alpha_{i0} + \alpha_{i1}t + \gamma_i f_t + v_{it},$$

where f_t is the common factor, γ_i the associated factor loadings, and v_{it} the idiosyncractic component assumed independently distributed of f_t. The unit root properties of y_{it} is determined by the maximum order of integration of the two series f_t and v_{it}. Hence, y_{it} will be $I(1)$ if either v_{it} and/or f_t contain a unit root. Averaging across i and letting $N \to \infty$, for each t, $\bar{v}_t \overset{q.m.}{\to} 0$, if v_{it} is stationary, and $\bar{v}_t \overset{q.m.}{\to} c$, where c is a fixed constant if v_{it} is $I(1)$. Therefore, a unit root in f_t may be tested by testing the presence of a unit root in $\bar{y}_t$ independently of whether the idiosyncractic components are $I(0)$ or $I(1)$. By contrast, in our specifications (see (5.43), (5.46), and (5.49)), the common factor is introduced to model cross section dependence of the stationary components. As a result when testing $\phi_i = 1$, the order of integration of y_{it} changes from being $I(1)$ if f_t is stationary, to $I(2)$ if f_t is $I(1)$. Therefore, in our set up it makes sense not to allow f_t to have a unit root. The models advanced here and the static factor model used by Bai and Ng serve different purposes.

Appendix A: Mathematical Proofs

A.1 Some Preliminary Order Results

Recall from (3.20) and (3.22) that $\boldsymbol{\upsilon}_i = \boldsymbol{\xi}_i/\omega_i = (\boldsymbol{\varepsilon}_i - \delta_i\bar{\boldsymbol{\varepsilon}})/\omega_i$, where $\omega_i^2 = \sigma_i^2 + O\left(\frac{1}{N}\right)$. Also $\mathfrak{s}_{i,-1} = (\mathbf{s}_{i,-1} - \delta_i\bar{\mathbf{s}}_{-1})/\omega_i$ where

$$\mathfrak{s}_i = \mathfrak{s}_{i,-1} + \boldsymbol{\upsilon}_i. \tag{A.1}$$

Hence

$$\frac{\boldsymbol{\tau}'\boldsymbol{\upsilon}_i}{\sqrt{T}} = \frac{\boldsymbol{\tau}'\boldsymbol{\varepsilon}_i - \delta_i\left(\boldsymbol{\tau}'\bar{\boldsymbol{\varepsilon}}\right)}{\omega_i\sqrt{T}},\ \frac{\boldsymbol{\tau}'\mathfrak{s}_{i,-1}}{T^{3/2}} = \frac{\boldsymbol{\tau}'\mathbf{s}_{i,-1} - \delta_i\left(\boldsymbol{\tau}'\bar{\mathbf{s}}_{-1}\right)}{\omega_i T^{3/2}}, \tag{A.2}$$

$$\frac{\bar{\boldsymbol{\varepsilon}}'\boldsymbol{\upsilon}_i}{\sqrt{T}} = \frac{\bar{\boldsymbol{\varepsilon}}'\boldsymbol{\varepsilon}_i - \delta_i\left(\bar{\boldsymbol{\varepsilon}}'\bar{\boldsymbol{\varepsilon}}\right)}{\omega_i\sqrt{T}},\ \frac{\bar{\boldsymbol{\varepsilon}}'\mathfrak{s}_{i,-1}}{T^{3/2}} = \frac{\bar{\boldsymbol{\varepsilon}}'\mathbf{s}_{i,-1} - \delta_i\left(\bar{\boldsymbol{\varepsilon}}'\bar{\mathbf{s}}_{-1}\right)}{\omega_i T^{3/2}}, \tag{A.3}$$

$$\frac{\mathbf{f}'\boldsymbol{\upsilon}_i}{\sqrt{T}} = \frac{\mathbf{f}'\boldsymbol{\varepsilon}_i - \delta_i\left(\mathbf{f}'\bar{\boldsymbol{\varepsilon}}\right)}{\omega_i\sqrt{T}},\ \frac{\mathbf{f}'\mathfrak{s}_{i,-1}}{T^{3/2}} = \frac{\mathbf{f}'\mathbf{s}_{i,-1} - \delta_i\left(\mathbf{f}'\bar{\mathbf{s}}_{-1}\right)}{\omega_i T^{3/2}}, \tag{A.4}$$

$$\frac{\bar{\mathbf{s}}_{-1}'\mathfrak{s}_{i,-1}}{T^2} = \frac{\bar{\mathbf{s}}_{-1}'\mathbf{s}_{i,-1} - \delta_i\left(\bar{\mathbf{s}}_{-1}'\bar{\mathbf{s}}_{-1}\right)}{\omega_i T^2},\ \frac{\mathbf{s}_{f,-1}'\mathfrak{s}_{i,-1}}{T^2} = \frac{\mathbf{s}_{f,-1}'\mathbf{s}_{i,-1} - \delta_i\left(\mathbf{s}_{f,-1}'\bar{\mathbf{s}}_{-1}\right)}{\omega_i T^2}, \tag{A.5}$$

$$\frac{\mathbf{s}_{f,-1}'\boldsymbol{\upsilon}_i}{T} = \frac{\mathbf{s}_{f,-1}'\boldsymbol{\varepsilon}_i - \delta_i\left(\mathbf{s}_{f,-1}'\bar{\boldsymbol{\varepsilon}}\right)}{\omega_i T},\ \frac{\bar{\mathbf{s}}_{-1}'\boldsymbol{\upsilon}_i}{T} = \frac{\bar{\mathbf{s}}_{-1}'\boldsymbol{\varepsilon}_i - \delta_i\left(\bar{\mathbf{s}}_{-1}'\bar{\boldsymbol{\varepsilon}}\right)}{\omega_i T}, \tag{A.6}$$

$$\frac{\boldsymbol{\upsilon}_i'\boldsymbol{\upsilon}_i}{T} = \frac{\boldsymbol{\varepsilon}_i'\boldsymbol{\varepsilon}_i + \delta_i^2\left(\bar{\boldsymbol{\varepsilon}}'\bar{\boldsymbol{\varepsilon}}\right) - 2\delta_i\left(\boldsymbol{\varepsilon}_i'\bar{\boldsymbol{\varepsilon}}\right)}{\omega_i^2 T}, \tag{A.7}$$

$$\frac{\mathfrak{s}_{i,-1}'\mathfrak{s}_{i,-1}}{T^2} = \frac{\mathbf{s}_{i,-1}'\mathbf{s}_{i,-1} + \delta_i^2\ \left(\bar{\mathbf{s}}_{-1}'\bar{\mathbf{s}}_{-1}\right) - 2\delta_i\left(\mathbf{s}_{i,-1}'\bar{\mathbf{s}}_{-1}\right)}{\omega_i^2 T^2}, \tag{A.8}$$

Now using results in Pesaran (2002, Appendix) it is easily seen that[18]

$$E\left(\frac{\boldsymbol{\tau}'\bar{\boldsymbol{\varepsilon}}}{\sqrt{T}}\right) = 0,\ Var\left(\frac{\boldsymbol{\tau}'\bar{\boldsymbol{\varepsilon}}}{\sqrt{T}}\right) = O\left(\frac{1}{N}\right), \tag{A.9}$$

$$E\left(\frac{\mathbf{f}'\bar{\boldsymbol{\varepsilon}}}{\sqrt{T}}\right) = 0, Var\left(\frac{\mathbf{f}'\bar{\boldsymbol{\varepsilon}}}{\sqrt{T}}\right) = O\left(\frac{1}{N}\right), \tag{A.10}$$

$$E\left(\frac{\bar{\boldsymbol{\varepsilon}}'\bar{\boldsymbol{\varepsilon}}}{\sqrt{T}}\right) = O\left(\frac{\sqrt{T}}{N}\right),\ Var\left(\frac{\bar{\boldsymbol{\varepsilon}}'\bar{\boldsymbol{\varepsilon}}}{\sqrt{T}}\right) = O\left(\frac{T}{N^2}\right), \tag{A.11}$$

$$E\left(\frac{\bar{\boldsymbol{\varepsilon}}'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right) = \frac{\sqrt{T}\sigma_i^2}{N},\ Var\left(\frac{\bar{\boldsymbol{\varepsilon}}'\boldsymbol{\varepsilon}_i}{\sqrt{T}}\right) = O\left(\frac{1}{N}\right). \tag{A.12}$$

[18] Note that in the present application f_t and ε_{it} are serially uncorrelated and independently distributed, while the results in Pesaran (2002) are more generally applicable.

Also using results in Fuller (1996, p. 547) and carrying out similar derivations we also have

$$E\left(\frac{\bar{\mathbf{s}}'_{-1}\bar{\boldsymbol{\varepsilon}}}{T^{3/2}}\right)=0,\ Var\left(\frac{\bar{\mathbf{s}}'_{-1}\bar{\boldsymbol{\varepsilon}}}{T^{3/2}}\right)=O\left(\frac{1}{TN^2}\right), \tag{A.13}$$

$$E\left(\frac{\bar{\mathbf{s}}'_{-1}\boldsymbol{\varepsilon}_i}{T}\right)=0,\ Var\left(\frac{\bar{\mathbf{s}}'_{-1}\boldsymbol{\varepsilon}_i}{T}\right)=O\left(\frac{1}{N}\right), \tag{A.14}$$

$$E\left(\frac{\bar{\mathbf{s}}'_{-1}\boldsymbol{\tau}}{T^{3/2}}\right)=0,\ Var\left(\frac{\bar{\mathbf{s}}'_{-1}\boldsymbol{\tau}}{T^{3/2}}\right)=O\left(\frac{1}{N}\right), \tag{A.15}$$

$$E\left(\frac{\bar{\mathbf{s}}'_{-1}\bar{\mathbf{s}}_{-1}}{T^2}\right)=O\left(\frac{1}{N}\right),\ Var\left(\frac{\bar{\mathbf{s}}'_{-1}\bar{\mathbf{s}}_{-1}}{T^2}\right)=O\left(\frac{1}{N^2}\right), \tag{A.16}$$

$$E\left(\frac{\bar{\mathbf{s}}'_{-1}\mathbf{s}_{i,-1}}{T^2}\right)=O\left(\frac{1}{N}\right),\ Var\left(\frac{\bar{\mathbf{s}}'_{-1}\mathbf{s}_{i,-1}}{T^2}\right)=O\left(\frac{1}{N}\right), \tag{A.17}$$

$$E\left(\frac{\bar{\mathbf{s}}'_{-1}\mathbf{s}_{f,-1}}{T^2}\right)=0,\ Var\left(\frac{\bar{\mathbf{s}}'_{-1}\mathbf{s}_{f,-1}}{T^2}\right)=O\left(\frac{1}{N}\right), \tag{A.18}$$

$$E\left(\frac{\mathbf{s}'_{f,-1}\bar{\boldsymbol{\varepsilon}}}{T}\right)=0,\ Var\left(\frac{\mathbf{s}'_{f,-1}\bar{\boldsymbol{\varepsilon}}}{T}\right)=O\left(\frac{1}{N}\right), \tag{A.19}$$

$$E\left(\frac{\mathbf{f}'\bar{\boldsymbol{\varepsilon}}}{\sqrt{T}}\right)=0,\ Var\left(\frac{\mathbf{f}'\bar{\boldsymbol{\varepsilon}}}{\sqrt{T}}\right)=O\left(\frac{1}{N}\right), \tag{A.20}$$

$$E\left(\frac{\mathbf{f}'\bar{\mathbf{s}}_{-1}}{T^{3/2}}\right)=0,\ Var\left(\frac{\mathbf{f}'\bar{\mathbf{s}}_{-1}}{T^{3/2}}\right)=O\left(\frac{1}{TN}\right), \tag{A.21}$$

A.2 Asymptotic Distribution of $\tilde{t}_i(N,T)$ - Serially Uncorrelated Case

The $\tilde{t}_i(N,T)$ statistic defined by (3.27) may be written as

$$\tilde{t}_i(N,T)=\frac{\left(\sqrt{\frac{T-3}{T}}\right)\left(\frac{\boldsymbol{\upsilon}'_i\bar{\mathbf{M}}_w\mathfrak{s}_{i,-1}}{T}\right)}{\left(\frac{\boldsymbol{\upsilon}'_i\bar{\mathbf{M}}_w\boldsymbol{\upsilon}_i}{T}\right)^{1/2}\left(\frac{\mathfrak{s}'_{i,-1}\bar{\mathbf{M}}_w\mathfrak{s}_{i,-1}}{T^2}\right)^{1/2}}. \tag{A.22}$$

where $\boldsymbol{\upsilon}_i\sim(\mathbf{0},\mathbf{I}_T)$, and $\mathfrak{s}_{i,-1}$ is defined by (A.1) which is the standardized random walk associated to $\boldsymbol{\upsilon}_i$.

First consider the numerator of $\tilde{t}_i(N,T)$, and note that

$$\frac{\boldsymbol{\upsilon}'_i\bar{\mathbf{M}}_w\mathfrak{s}_{i,-1}}{T}=\frac{\boldsymbol{\upsilon}'_i\mathfrak{s}_{i,-1}}{T}-\left(\boldsymbol{\upsilon}'_i\bar{\mathbf{W}}\mathbf{D}\right)\left(\mathbf{D}\bar{\mathbf{W}}'\bar{\mathbf{W}}\mathbf{D}\right)^{-1}\left(\frac{\mathbf{D}\bar{\mathbf{W}}'\mathfrak{s}_{i,-1}}{T}\right), \tag{A.23}$$

where

$$\mathbf{D}=\begin{pmatrix}T^{-1/2}&0&0\\0&T^{-1/2}&0\\0&0&T^{-1}\end{pmatrix}. \tag{A.24}$$

Also

$$\mathbf{D}\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i = \begin{pmatrix} \frac{\Delta\bar{\mathbf{y}}'\boldsymbol{\upsilon}_i}{\sqrt{T}} \\ \frac{\boldsymbol{\tau}'\boldsymbol{\upsilon}_i}{\sqrt{T}} \\ \frac{\bar{\mathbf{y}}'_{-1}\boldsymbol{\upsilon}_i}{T} \end{pmatrix}, \quad \frac{\mathbf{D}\bar{\mathbf{W}}'\mathfrak{s}_{i,-1}}{T} = \begin{pmatrix} \frac{\Delta\bar{\mathbf{y}}'\mathfrak{s}_{i,-1}}{T^{3/2}} \\ \frac{\boldsymbol{\tau}'\mathfrak{s}_{i,-1}}{T^{3/2}} \\ \frac{\bar{\mathbf{y}}'_{-1}\mathfrak{s}_{i,-1}}{T^2} \end{pmatrix}, \tag{A.25}$$

$$\mathbf{D}\bar{\mathbf{W}}'\bar{\mathbf{W}}\mathbf{D} = \begin{pmatrix} \frac{\Delta\bar{\mathbf{y}}'\Delta\bar{\mathbf{y}}}{T} & \frac{\Delta\bar{\mathbf{y}}'\boldsymbol{\tau}}{T} & \frac{\Delta\bar{\mathbf{y}}'\bar{\mathbf{y}}_{-1}}{T^{3/2}} \\ \frac{\boldsymbol{\tau}'\Delta\bar{\mathbf{y}}}{T} & 1 & \frac{\boldsymbol{\tau}'\bar{\mathbf{y}}_{-1}}{T^{3/2}} \\ \frac{\bar{\mathbf{y}}'_{-1}\Delta\bar{\mathbf{y}}}{T^{3/2}} & \frac{\bar{\mathbf{y}}'_{-1}\boldsymbol{\tau}}{T^{3/2}} & \frac{\bar{\mathbf{y}}'_{-1}\bar{\mathbf{y}}_{-1}}{T^2} \end{pmatrix}. \tag{A.26}$$

Using (3.17) and (3.18) we have

$$\frac{\Delta\bar{\mathbf{y}}'\boldsymbol{\upsilon}_i}{\sqrt{T}} = \bar{\gamma}\left(\frac{\mathbf{f}'\boldsymbol{\upsilon}_i}{\sqrt{T}}\right) + \left(\frac{\bar{\boldsymbol{\varepsilon}}'\boldsymbol{\upsilon}_i}{\sqrt{T}}\right), \tag{A.27}$$

$$\frac{\bar{\mathbf{y}}'_{-1}\boldsymbol{\upsilon}_i}{T} = \bar{y}_0\left(\frac{\boldsymbol{\tau}'\boldsymbol{\upsilon}_i}{T}\right) + \bar{\gamma}\left(\frac{\mathbf{s}'_{f,-1}\boldsymbol{\upsilon}_i}{T}\right) + \left(\frac{\bar{\mathbf{s}}'_{-1}\boldsymbol{\upsilon}_i}{T}\right), \tag{A.28}$$

$$\frac{\Delta\bar{\mathbf{y}}'\mathfrak{s}_{i,-1}}{T^{3/2}} = \bar{\gamma}\left(\frac{\mathbf{f}'\mathfrak{s}_{i,-1}}{T^{3/2}}\right) + \left(\frac{\bar{\boldsymbol{\varepsilon}}'\mathfrak{s}_{i,-1}}{T^{3/2}}\right), \tag{A.29}$$

$$\frac{\bar{\mathbf{y}}'_{-1}\mathfrak{s}_{i,-1}}{T^2} = \bar{y}_0\left(\frac{\boldsymbol{\tau}'\mathfrak{s}_{i,-1}}{T^2}\right) + \bar{\gamma}\left(\frac{\mathbf{s}'_{f,-1}\mathfrak{s}_{i,-1}}{T^2}\right) + \left(\frac{\bar{\mathbf{s}}'_{-1}\mathfrak{s}_{i,-1}}{T^2}\right), \tag{A.30}$$

$$\frac{\boldsymbol{\tau}'\Delta\bar{\mathbf{y}}}{T} = \bar{\gamma}\left(\frac{\boldsymbol{\tau}'\mathbf{f}}{T}\right) + \left(\frac{\boldsymbol{\tau}'\bar{\boldsymbol{\varepsilon}}}{T}\right), \tag{A.31}$$

$$\frac{\boldsymbol{\tau}'\bar{\mathbf{y}}_{-1}}{T^{3/2}} = \bar{y}_0\left(\frac{1}{\sqrt{T}}\right) + \bar{\gamma}\left(\frac{\boldsymbol{\tau}'\mathbf{s}_{f,-1}}{T^{3/2}}\right) + \left(\frac{\boldsymbol{\tau}'\bar{\mathbf{s}}_{-1}}{T^{3/2}}\right), \tag{A.32}$$

$$\frac{\Delta\bar{\mathbf{y}}'\Delta\bar{\mathbf{y}}}{T} = \bar{\gamma}^2\left(\frac{\mathbf{f}'\mathbf{f}}{T}\right) + 2\bar{\gamma}\left(\frac{\mathbf{f}'\bar{\boldsymbol{\varepsilon}}}{T}\right) + \left(\frac{\bar{\boldsymbol{\varepsilon}}'\bar{\boldsymbol{\varepsilon}}}{T}\right), \tag{A.33}$$

$$\begin{aligned} \frac{\Delta\bar{\mathbf{y}}'\bar{\mathbf{y}}_{-1}}{T^{3/2}} &= \bar{\gamma}\bar{y}_0\left(\frac{\mathbf{f}'\boldsymbol{\tau}}{T^{3/2}}\right) + \bar{y}_0\left(\frac{\bar{\boldsymbol{\varepsilon}}'\boldsymbol{\tau}}{T^{3/2}}\right) + \bar{\gamma}^2\left(\frac{\mathbf{f}'\mathbf{s}_{f,-1}}{T^{3/2}}\right) \\ &+ \bar{\gamma}\left(\frac{\mathbf{f}'\bar{\mathbf{s}}_{-1}}{T^{3/2}}\right) + \bar{\gamma}\left(\frac{\bar{\boldsymbol{\varepsilon}}'\mathbf{s}_{f,-1}}{T^{3/2}}\right) + \left(\frac{\bar{\boldsymbol{\varepsilon}}'\bar{\mathbf{s}}_{-1}}{T^{3/2}}\right) \end{aligned} \tag{A.34}$$

$$\begin{aligned} \frac{\bar{\mathbf{y}}'_{-1}\bar{\mathbf{y}}_{-1}}{T^2} &= \frac{\bar{y}_0^2}{T} + \bar{\gamma}^2\left(\frac{\mathbf{s}'_{f,-1}\mathbf{s}_{f,-1}}{T^2}\right) + \left(\frac{\bar{\mathbf{s}}'_{-1}\bar{\mathbf{s}}_{-1}}{T^2}\right) \\ &+ 2\bar{y}_0\bar{\gamma}\left(\frac{\boldsymbol{\tau}'\mathbf{s}_{f,-1}}{T^2}\right) + 2\bar{y}_0\left(\frac{\boldsymbol{\tau}'\bar{\mathbf{s}}_{-1}}{T^2}\right) + 2\bar{\gamma}\left(\frac{\mathbf{s}'_{f,-1}\bar{\mathbf{s}}_{-1}}{T^2}\right). \end{aligned} \tag{A.35}$$

Similarly, apart from $T^{-2}\left(\mathfrak{s}'_{i,-1}\mathfrak{s}_{i,-1}\right)$ and $T^{-1}\left(\boldsymbol{\upsilon}'_i\boldsymbol{\upsilon}_i\right)$, the remaining terms in the denominator of $\tilde{t}_i\left(N,T\right)$ may also be written in terms of the above expressions.

A.2.1 T Fixed and $N \to \infty$

In this case

$$Var(\bar{\varepsilon}_t) = \frac{\bar{\sigma}^2}{N},\; Var(\bar{s}_t) = \frac{t\bar{\sigma}^2}{N},$$

where $\bar{\sigma}^2 = N^{-1}\sum_{j=1}^{N}\sigma_j^2 < \infty$. Hence, for a fixed T those elements in $\tilde{t}_i(N,T)$ that involve $\bar{\boldsymbol{\varepsilon}}$ and $\bar{\mathbf{s}}_{-1}$ will converge to zero in mean square errors as $N \to \infty$. Assuming also that the series, y_{it}, are in the form of deviations from the cross-sectional mean of the initial observations so that $\bar{y}_0 = 0$, using the above results for a fixed T and as $N \to \infty$ we have (in mean square errors)

$$\mathbf{D}\bar{\mathbf{W}}'\bar{\mathbf{W}}\mathbf{D} \overset{N}{\to} \Gamma_*\Psi_T\Gamma_*,$$

$$\mathbf{D}\bar{\mathbf{W}}'\boldsymbol{\upsilon}_i \overset{N}{\to} \Gamma_*\,\mathbf{q}_{iT}, \quad \frac{\mathbf{D}\bar{\mathbf{W}}'\mathfrak{s}_{i,-1}}{T} \overset{N}{\to} \Gamma_*\,\mathbf{h}_{iT}$$

$$T^{-2}\left(\mathfrak{s}_{i,-1}'\mathfrak{s}_{i,-1}\right) \overset{N}{\to} \frac{\mathbf{s}_{i,-1}'\mathbf{s}_{i,-1}}{\sigma_i^2T^2},\; T^{-1}\left(\boldsymbol{\upsilon}_i'\boldsymbol{\upsilon}_i\right) \overset{N}{\to} \frac{\boldsymbol{\varepsilon}_i'\boldsymbol{\varepsilon}_i}{\sigma_i^2T},$$

$$\frac{\boldsymbol{\upsilon}_i'\mathfrak{s}_{i,-1}}{T} \overset{N}{\to} \frac{\boldsymbol{\varepsilon}_i'\mathbf{s}_{i,-1}}{\sigma_i^2T},$$

where

$$\Gamma_* = \begin{pmatrix} \gamma_* & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \gamma_* \end{pmatrix},\; \Psi_{fT} = \begin{pmatrix} \frac{\mathbf{f}'\mathbf{f}}{T} & \frac{\mathbf{f}'\boldsymbol{\tau}}{T} & \frac{\mathbf{f}'\mathbf{s}_{f,-1}}{T^{3/2}} \\ \frac{\boldsymbol{\tau}'\mathbf{f}}{T} & 1 & \frac{\boldsymbol{\tau}'\mathbf{s}_{f,-1}}{T^{3/2}} \\ \frac{\mathbf{s}_{f,-1}'\mathbf{f}}{T^{3/2}} & \frac{\mathbf{s}_{f,-1}'\boldsymbol{\tau}}{T^{3/2}} & \frac{\mathbf{s}_{f,-1}'\mathbf{s}_{f,-1}}{T^2} \end{pmatrix},$$

γ_* is the limit of $\bar{\gamma}$ as $N \to \infty$, and

$$\mathbf{q}_{iT} = \begin{pmatrix} \frac{\mathbf{f}'\boldsymbol{\varepsilon}_i}{\sigma_i\sqrt{T}} \\ \frac{\boldsymbol{\tau}'\boldsymbol{\varepsilon}_i}{\sigma_i\sqrt{T}} \\ \frac{\mathbf{s}_{f,-1}'\boldsymbol{\varepsilon}_i}{\sigma_iT} \end{pmatrix},\; \mathbf{h}_{iT} = \begin{pmatrix} \frac{\mathbf{f}'\mathbf{s}_{i,-1}}{\sigma_iT^{3/2}} \\ \frac{\boldsymbol{\tau}'\mathbf{s}_{i,-1}}{\sigma_iT^{3/2}} \\ \frac{\mathbf{s}_{f,-1}'\mathbf{s}_{i,-1}}{\sigma_iT^2} \end{pmatrix}.$$

Using these results in (A.22) we finally obtain:

$$\tilde{t}_i(N,T) \overset{N}{\to} \frac{\frac{\boldsymbol{\varepsilon}_i'\mathbf{s}_{i,-1}}{\sigma_i^2T} - \mathbf{q}_{iT}'\Psi_{fT}^{-1}\mathbf{h}_{iT}}{\left(\frac{\boldsymbol{\varepsilon}_i'\boldsymbol{\varepsilon}_i}{\sigma_i^2(T-3)} - \frac{\mathbf{q}_{iT}'\Psi_{fT}^{-1}\mathbf{q}_{iT}}{T-3}\right)^{1/2}\left(\frac{\mathbf{s}_{i,-1}'\mathbf{s}_{i,-1}}{\sigma_i^2T^2} - \mathbf{h}_{iT}'\Psi_{fT}^{-1}\mathbf{h}_{iT}\right)^{1/2}}, \tag{A.36}$$

which is free of nuisance parameters and its probability distribution can be simulated for any given value of $T > 3$. Recall that f_t and $\varepsilon_{it}/\sigma_i$ are independently distributed as $iid(0,1)$.

A.2.2 Sequential Asymptotic : $N \to \infty$ then $T \to \infty$

First, using familiar results from the unit root literature as $T \to \infty$, we have (See, for example, Hamilton (1994, p.486))

$$\frac{\boldsymbol{\tau}'\mathbf{s}_{i,-1}}{\sigma_iT^{3/2}} \overset{T}{\Longrightarrow} \int_0^1 W_i(r)\,dr,\; \frac{\mathbf{s}_{i,-1}'\mathbf{s}_{i,-1}}{\sigma_i^2T^2} \overset{T}{\Longrightarrow} \int_0^1 W_i^2(r)\,dr,\; \frac{\boldsymbol{\tau}'\boldsymbol{\varepsilon}_i}{\sigma_i\sqrt{T}} \overset{T}{\Longrightarrow} W_i(1),$$

$$\frac{\boldsymbol{\varepsilon}_i'\mathbf{s}_{i,-1}}{\sigma_i^2T} \overset{T}{\Longrightarrow} \int_0^1 W_i(r)\,dW_i(r) = (1/2)\left(W^2(1) - 1\right),$$

$$\frac{\boldsymbol{\tau}'\mathbf{s}_{f,-1}}{T^{3/2}} \overset{T}{\Longrightarrow} \int_0^1 W_f(r)\,dr,\; \frac{\mathbf{s}_{f,-1}'\mathbf{s}_{f,-1}}{T^2} \overset{T}{\Longrightarrow} \int_0^1 W_f^2(r)\,dr$$

where $W_i(r)$ and $W_f(r)$ are independently distributed standard Brownian motions defined on $[0,1]$. Similarly,

$$\frac{\mathbf{s}'_{f,-1}\mathbf{s}_{i,-1}}{\sigma_i T^2} \overset{T}{\Longrightarrow} \int_0^1 W_f(r) W_i(r)\, dr,\ \frac{\mathbf{f}'\boldsymbol{\varepsilon}_i}{\sigma_i\sqrt{T}} \overset{T}{\Longrightarrow} W_{fi}(1),\ \frac{\mathbf{s}'_{f,-1}\boldsymbol{\varepsilon}_i}{\sigma_i T} \overset{T}{\Longrightarrow} \int_0^1 W_f(r)\, dW_i(r),$$

where $W_f(r)$ and $W_{fi}(r)$ are also distributed as standard Brownian motions. Finally, it is easily seen that

$$\frac{\mathbf{f}'\mathbf{f}}{T} \overset{T}{\to} 1,\ \frac{\boldsymbol{\tau}'\mathbf{f}}{T} \overset{T}{\to} 0,\ \frac{\mathbf{f}'\mathbf{s}_{i,-1}}{\sigma_i T^{3/2}} \overset{T}{\to} 0,$$

$$\frac{\boldsymbol{\varepsilon}'_i\boldsymbol{\varepsilon}_i}{\sigma_i^2(T-3)} \overset{T}{\to} 1,\ \frac{\mathbf{q}'_{iT}\Psi_{fT}^{-1}\mathbf{q}_{iT}}{T-3} \overset{T}{\to} 0.$$

Using these results in (A.36) as $T \to \infty$, we obtain the following sequential limit distribution

$$\tilde{t}_i(N,T) \overset{N,T}{\Longrightarrow} \frac{\int_0^1 W_i(r)\, dW_i(r) - \boldsymbol{\psi}'_{if}\Lambda_f^{-1}\boldsymbol{\kappa}_{if}}{\left(\int_0^1 W_i^2(r)\, dr - \boldsymbol{\kappa}'_{if}\Lambda_f^{-1}\boldsymbol{\kappa}_{if}\right)^{1/2}}, \tag{A.37}$$

where

$$\Lambda_f = \begin{pmatrix} 1 & \int_0^1 W_f(r)\, dr \\ \int_0^1 W_f(r)\, dr & \int_0^1 W_f^2(r)\, dr \end{pmatrix}, \tag{A.38}$$

$$\boldsymbol{\psi}_{if} = \begin{pmatrix} W_i(1) \\ \int_0^1 W_f(r)\, dW_i(r) \end{pmatrix},\ \boldsymbol{\kappa}_{if} = \begin{pmatrix} \int_0^1 W_i(r)\, dr \\ \int_0^1 W_f(r) W_i(r)\, dr \end{pmatrix}. \tag{A.39}$$

A.2.3 Joint Asymptotics

Using the order results (A.9) to)A.21) it is easily seen that all the terms in (A.2) to (A.8) that contain the cross-section means, $\bar{\boldsymbol{\varepsilon}}$ and $\bar{\mathbf{s}}_{-1}$, converge in quadratic mean to zero as $N,T \to \infty$, jointly so long as $\sqrt{T}/N \to 0$. This latter condition is satisfied if $N/T \to k$, where where k is a fixed finite non-zero constant. It therefore follows that the asymptotic result, (A.37) also holds under joint asymptotic and so long as $\sqrt{T}/N \to 0$ we have

$$\tilde{t}_i(N,T) \overset{(N,T)_j}{\Longrightarrow} \frac{\int_0^1 W_i(r)\, dW_i(r) - \boldsymbol{\psi}'_{if}\Lambda_f^{-1}\boldsymbol{\kappa}_{if}}{\left(\int_0^1 W_i^2(r)\, dr - \boldsymbol{\kappa}'_{if}\Lambda_f^{-1}\boldsymbol{\kappa}_{if}\right)^{1/2}}. \tag{A.40}$$

It is also easily seen that with $\sqrt{T}/N \to 0$

$$\hat{\sigma}_i^2 = \frac{\Delta\mathbf{y}'_i\mathbf{M}_{i,w}\Delta\mathbf{y}_i}{T-4} \overset{(N,T)_j}{\to} \sigma_i^2,$$

and hence we also have

$$t_i(N,T) \overset{(N,T)_j}{\Longrightarrow} \frac{\int_0^1 W_i(r)\, dW_i(r) - \boldsymbol{\psi}'_{if}\Lambda_f^{-1}\boldsymbol{\kappa}_{if}}{\left(\int_0^1 W_i^2(r)\, dr - \boldsymbol{\kappa}'_{if}\Lambda_f^{-1}\boldsymbol{\kappa}_{if}\right)^{1/2}}, \tag{A.41}$$

where $t_i(N,T)$ is defined by (3.7).

A.3 Asymptotic Distribution of $\tilde{t}_i(N,T)$ - Serially Correlated Case

Consider first the numerator of (5.59) and note that

$$\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_i\mathfrak{s}_{zi,-1}}{T} = \frac{\boldsymbol{\upsilon}_i'\mathfrak{s}_{zi,-1}}{T} - \boldsymbol{\upsilon}_i'\bar{\mathbf{W}}_i\mathbf{D}\left(\mathbf{D}\bar{\mathbf{W}}_i'\bar{\mathbf{W}}_i\mathbf{D}\right)^{-1}\left(\frac{\mathbf{D}\bar{\mathbf{W}}_i'\mathfrak{s}_{zi,-1}}{T}\right),$$

where

$$\begin{aligned}
\mathbf{D} &= \begin{pmatrix} T^{-1/2}\mathbf{I}_4 & \mathbf{0} \\ \mathbf{0} & T^{-1} \end{pmatrix}, \\
\bar{\mathbf{W}}_i &= \left(\Delta\mathbf{y}_{i,-1}, \Delta\bar{\mathbf{y}}, \Delta\bar{\mathbf{y}}_{-1}, \boldsymbol{\tau}_T, \bar{\mathbf{y}}_{-1}\right), \\
\mathfrak{s}_{zi,-1} &= \left(\mathbf{s}_{zi,-1} - \delta_i\bar{\mathbf{s}}_{z,-1}\right)/\omega_i.
\end{aligned}$$

The elements of $\mathbf{s}_{zi,-1}$ and $\bar{\mathbf{s}}_{z,-1}$ are defined by (5.55) and (5.56), and can be written in terms of general first-difference stationary processes. Recall also that $\omega_i^2 = \sigma_i^2 + O\left(N^{-1}\right)$.

Using the results set out above, together with the familiar results on stationary first-difference processes summarized, for example, in Proposition 17.3 of Hamilton (1994), the following limits can now be established under joint asymptotics (with N and $T \to \infty$, such that $N/T \to k$, $\infty > k > 0$)

$$\mathbf{D}\bar{\mathbf{W}}_i'\bar{\mathbf{W}}_i\mathbf{D} \overset{(N,T)_j}{\rightarrow} \begin{pmatrix} \mathbf{V}_i & \mathbf{0}_{3\times 2} \\ \mathbf{0}_{2\times 3} & \Gamma_\rho\Lambda_f\Gamma_\rho \end{pmatrix},$$

$$\mathbf{D}\bar{\mathbf{W}}_i'\boldsymbol{\upsilon}_i \overset{(N,T)_j}{\Longrightarrow} \begin{pmatrix} \sqrt{\frac{\gamma_i^2+\sigma_i^2}{1-\rho^2}}W_i(1) \\ \frac{\gamma_*}{\sqrt{1-\rho^2}}W_i(1) \\ \frac{\gamma_*}{\sqrt{1-\rho^2}}W_i(1) \\ W_i(1) \\ \frac{\gamma_*}{1-\rho}\int_0^1 W_i(r)\,dW_i(r) \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{\gamma_i^2+\sigma_i^2}{1-\rho^2}}W_i(1) \\ \frac{\gamma_*}{\sqrt{1-\rho^2}}W_i(1) \\ \frac{\gamma_*}{\sqrt{1-\rho^2}}W_i(1) \\ \Gamma_\rho\boldsymbol{\psi}_{if} \end{pmatrix},$$

$$\frac{\mathbf{D}\bar{\mathbf{W}}_i'\mathfrak{s}_{zi,-1}}{T} \overset{(N,T)_j}{\Longrightarrow} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{1-\rho}\int_0^1 W_i(r)\,dr \\ \frac{\gamma_*}{(1-\rho)^2}\int_0^1 W_f(r)\,W_i(r)\,dr \end{pmatrix} = \begin{pmatrix} \mathbf{0}_{3\times 1} \\ \frac{1}{1-\rho}\Gamma_\rho\boldsymbol{\kappa}_{if} \end{pmatrix}$$

$$\frac{\boldsymbol{\upsilon}_i'\mathfrak{s}_{zi,-1}}{T} \overset{(N,T)_j}{\Longrightarrow} \frac{1}{1-\rho}\int_0^1 W_i(r)\,dW_i(r),$$

where Λ_f , $\boldsymbol{\psi}_{if}$,and $\boldsymbol{\kappa}_{if}$, are defined by (A.38) and (A.39), $\gamma_* \neq 0$ is the limit of $\bar{\gamma}$ as $N \to \infty$, $W_i(r)$ and $W_f(r)$ are independent standard Brownian motions, and

$$\mathbf{V}_i = \frac{1}{1-\rho^2}\begin{pmatrix} \gamma_i^2+\sigma_i^2 & \rho\gamma_i\gamma_* & \gamma_i\gamma_* \\ \rho\gamma_i\gamma_* & \gamma_*^2 & \rho\gamma_*^2 \\ \gamma_i\gamma_* & \rho\gamma_*^2 & \gamma_*^2 \end{pmatrix}, \; \Gamma_\rho = \begin{pmatrix} 1 & 0 \\ 0 & \frac{\gamma_*}{1-\rho} \end{pmatrix}.$$

Similarly,

$$\frac{\mathfrak{s}_{zi,-1}'\bar{\mathbf{M}}_i\mathfrak{s}_{zi,-1}}{T^2} = \frac{\mathfrak{s}_{zi,-1}'\mathfrak{s}_{zi,-1}}{T^2} - \left(\frac{\mathfrak{s}_{zi,-1}'\bar{\mathbf{W}}_i\mathbf{D}}{T}\right)\left(\mathbf{D}\bar{\mathbf{W}}_i'\bar{\mathbf{W}}_i\mathbf{D}\right)^{-1}\left(\frac{\mathbf{D}\bar{\mathbf{W}}_i'\mathfrak{s}_{zi,-1}}{T}\right),$$

$$\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_i\boldsymbol{\upsilon}_i}{T-5} = \frac{\boldsymbol{\upsilon}_i'\boldsymbol{\upsilon}_i}{T-5} - \frac{\mathbf{1}}{T-5}\boldsymbol{\upsilon}_i'\bar{\mathbf{W}}_i\mathbf{D}\left(\mathbf{D}\bar{\mathbf{W}}_i'\bar{\mathbf{W}}_i\mathbf{D}\right)^{-1}\mathbf{D}\bar{\mathbf{W}}_i'\boldsymbol{\upsilon}_i,$$

where it is also easily seen that

$$\frac{\mathfrak{s}'_{zi,-1}\mathfrak{s}_{zi,-1}}{T^2} \overset{(N,T)_j}{\Longrightarrow} \frac{1}{(1-\rho)^2}\int_0^1 W_f^2(r)\,dr,$$

and

$$\frac{\boldsymbol{\upsilon}_i'\bar{\mathbf{M}}_i\boldsymbol{\upsilon}_i}{T-5} \overset{(N,T)_j}{\backsim} \frac{\boldsymbol{\upsilon}_i'\mathbf{M}_{i,w}\boldsymbol{\upsilon}_i}{T-6} \overset{(N,T)_j}{\to} 1.$$

Using the above results in (5.59) we now have ($|\rho|<1$)

$$\tilde{t}_i(N,T) \overset{(N,T)_j}{\backsim} t_i(N,T) \overset{(N,T)_j}{\Longrightarrow} \frac{\frac{1}{1-\rho}\int_0^1 W_i(r)\,dW_i(r) - \boldsymbol{\psi}_{if}'\Gamma_\rho\left(\Gamma_\rho\Lambda_f\Gamma_\rho\right)^{-1}\frac{1}{1-\rho}\Gamma_\rho\boldsymbol{\kappa}_{if}}{\left\{\frac{1}{(1-\rho)^2}\int_0^1 W_f^2(r)\,dr - \frac{1}{1-\rho}\boldsymbol{\kappa}_{if}'\Gamma_\rho\left(\Gamma_\rho\Lambda_f\Gamma_\rho\right)^{-1}\frac{1}{1-\rho}\Gamma_\rho\boldsymbol{\kappa}_{if}\right\}^{1/2}},$$

which reduces to the desired result:

$$\frac{\int_0^1 W_i(r)\,dW_i(r) - \boldsymbol{\psi}_{if}'\Lambda_f^{-1}\boldsymbol{\kappa}_{if}}{\left(\int_0^1 W_i^2(r)\,dr - \boldsymbol{\kappa}_{if}'\Lambda_f^{-1}\boldsymbol{\kappa}_{if}\right)^{1/2}},$$

the joint asymptotic limit distribution of the CADF obtained in the case of serially uncorrelated errors given by (A.40) or (A.41).

References

[1] Bai, J. and S. Ng, (2002), "A Panic Attack on Unit Roots and Cointegration", Department of Economics, Boston College, Unpublished Manuscript.

[2] Baltagi, B. H., Kao, C., (2000), Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey. *Advances in Econometrics* 15, 7-51.

[3] 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.

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

[5] Chang, Y., (2003), "Bootstrap Unit Root Test in Panels with Cross-Sectional Dependency" *Journal of Econometrics*, (forthcoming).

[6] Choi, I., (2001), "Unit Root Tests for Panel Data", *Journal of International Money and Banking* 20, 249-272.

[7] Choi, I. (2002), "Combination Unit Root Tests for Cross-Sectionally Correlated Panels", *Mimeo*, Hong Kong University of Science and Technology.

[8] Elliott, G., T.J. Rothenberg, and J.H. Stock, (1996), "Efficient Tests for an Autoregressive Unit Root", *Econometrica*, 64, 4, 813-836.

[9] Fuller, W.A, (1996), *Introduction to Statistical Time Series*, Second edition, John Wiley & Sons, New York.

[10] Fuller, W.A. and H.J. Park (1995), "Alternative estimators and unit root tests for the autoregressive process", *Journal of Time Series Analysis*, 16, 415-429.

[11] Hadri, K., (2000), Testing for Stationarity in Heterogeneous Panel Data. *Econometrics Journal* 3, 148-161.

[12] Hamilton, J.D., (1994), *Time Series Analysis.* Princeton University Press, Princeton.

[13] Harvey, A. and D. Bates, (2002), "Multivariate Unit Root Tests and Testing for Convergence", *University of Cambridge DAE Working Paper No. 0301* (December 2002).

[14] Im, K.S. and M.H. Pesaran, (2003), "On the Panel Unit Root Tests Using Nonlinear Instrumental Variables", Unpublished menuscript.

[15] Im, K., Pesaran, H., and Y. Shin, (1995), "Testing for Unit Roots in Heterogeneous Panels,"DAE Working Papers Amalgamated Series No. 9526, University of Cambridge.

[16] Im, K., Pesaran, H., and Y. Shin, (2003), "Testing for Unit Roots in Heterogeneous Panels," *Journal of Econometrics* 115, 53-74.

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

[18] Leybourne, S.J., (1995), "Testing for Unit Roots using Forward and Reverse Dickey-Fuller Regressions", *Oxford Bulletin of Economics and Statistics*, Vol. 57, No.4, 559-571.

[19] 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.

[20] Moon, H.R., and B. Perron, (2003), "Testing for a Unit Root in Panels with Dynamic Factors", *Journal of Econometrics* (forthcoming).

[21] Moon, H.R., and B. Perron, P.C.B. Phillips (2003), "Power Comparisons of Panel Unit Root Tests under Incidental Trends", Department of Economics, *University of Southern Califonia*, unpublished manuscript.

[22] Pesaran, M.H., (2002) "Estimation and Inference in Large Heterogeneous Panels with Cross Section Dependence", *University of Cambridge DAE Working Paper No.0305* and CESifo Working Paper Series No 869, October 2002.

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

[24] Phillips, P.C.B. and D. Sul, (2002), "Dynamic Panel Estimation and Homogeneity Testing Under Cross Section Dependence", *Cowles Foundation Discussion Paper No. 1362*, Yale University.

[25] Shin, Y. and A. Snell, (2000), "Testing for Stationarity in Heterogeneous Panels with Serially Correlated Errors", *Department of Economics Discussion Paper* No. 2000-01, University of Edinburgh.

[26] Smith, V., S. Leybourne, T-H Kim, and P. Newbold, (2003), "More Powerful Panel Data Unit Root Tests with an Application to Mean Reversion in Real Exchange Rates, *Journal of Applied Econometrics*, (forthcoming).

[27] Taylor, R.L., Daffer, P.Z., and Patterson, R.F.,(1985), *Limit Theorems for Sums of Exchangeable Random Variables*, Rowman and Allanheld, N.J.

Figure 1: Cumulative Distribution Function of DF and Cross-Sectionally Augmented DF Statistics (The Intercept Case)

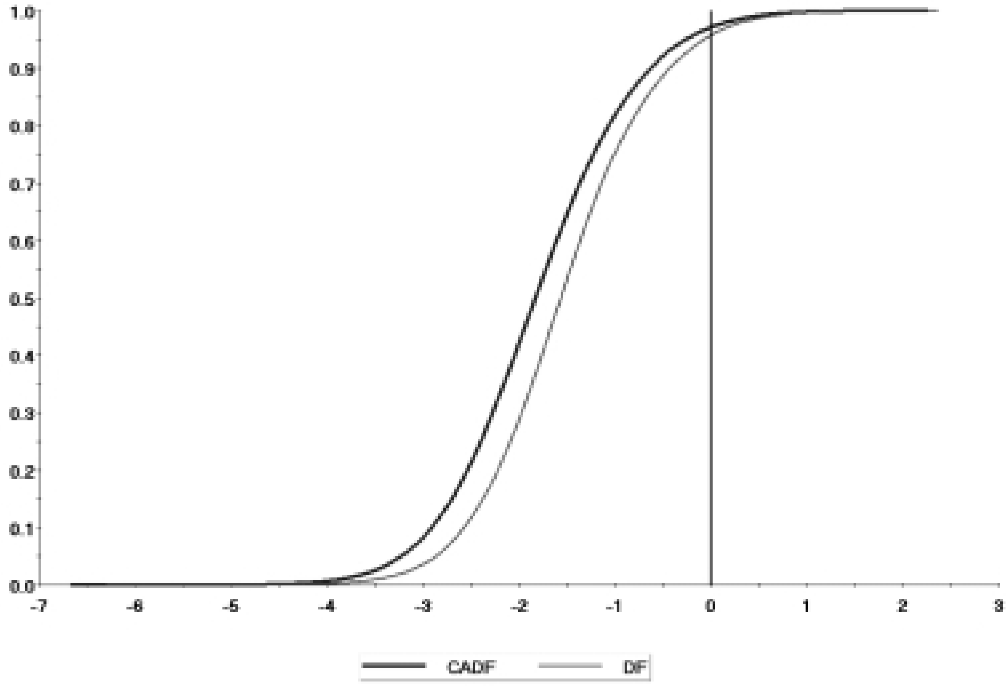

Figure 2: Simulated Density Function of the Standardized $CADF_i$ Distribution as Compared to the Normal Density (The Intercept Case)

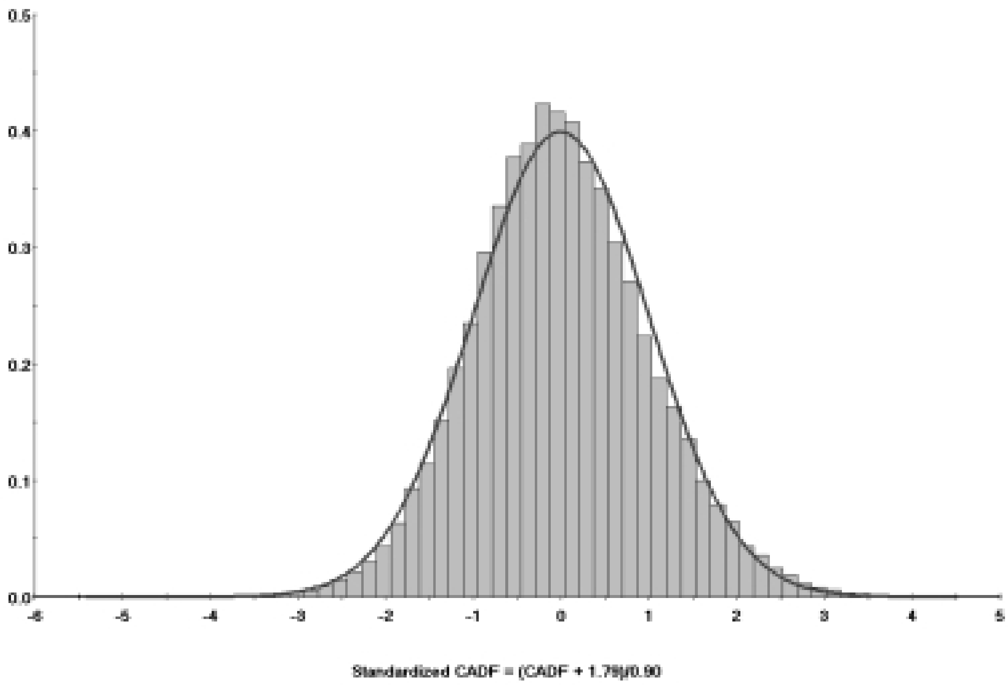

Figure 3: Simulated Density Function of the Standardized $CADF_i$ Distribution as Compared to the Normal Density (The Linear Trend Case)

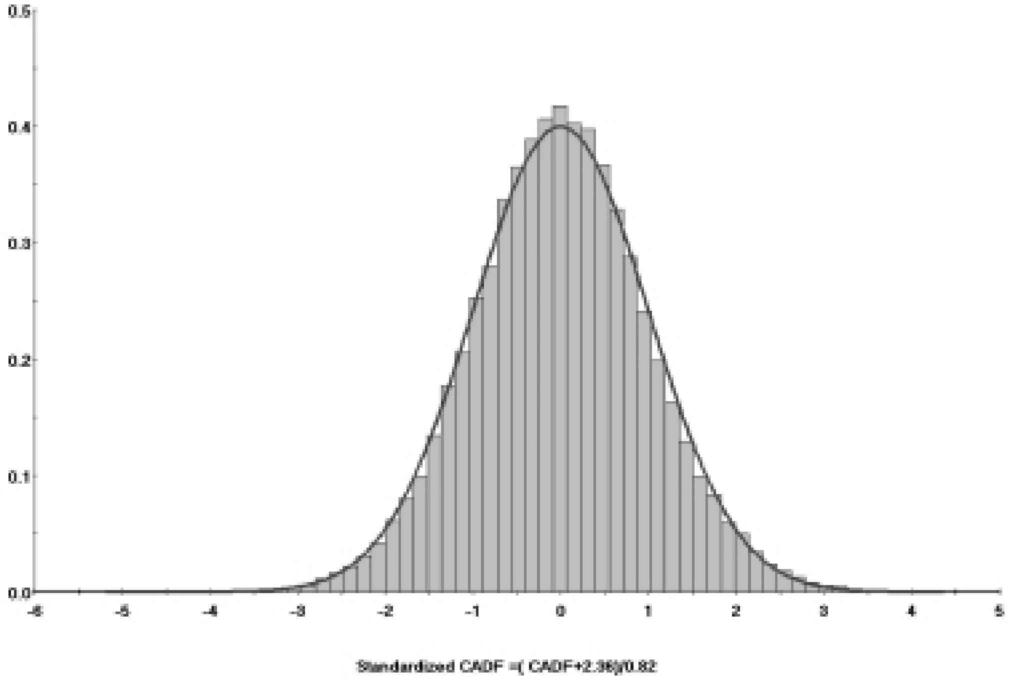

Figure 4: Simulated Density of the CIPS* Statistic - Case of no Intercept or Trend

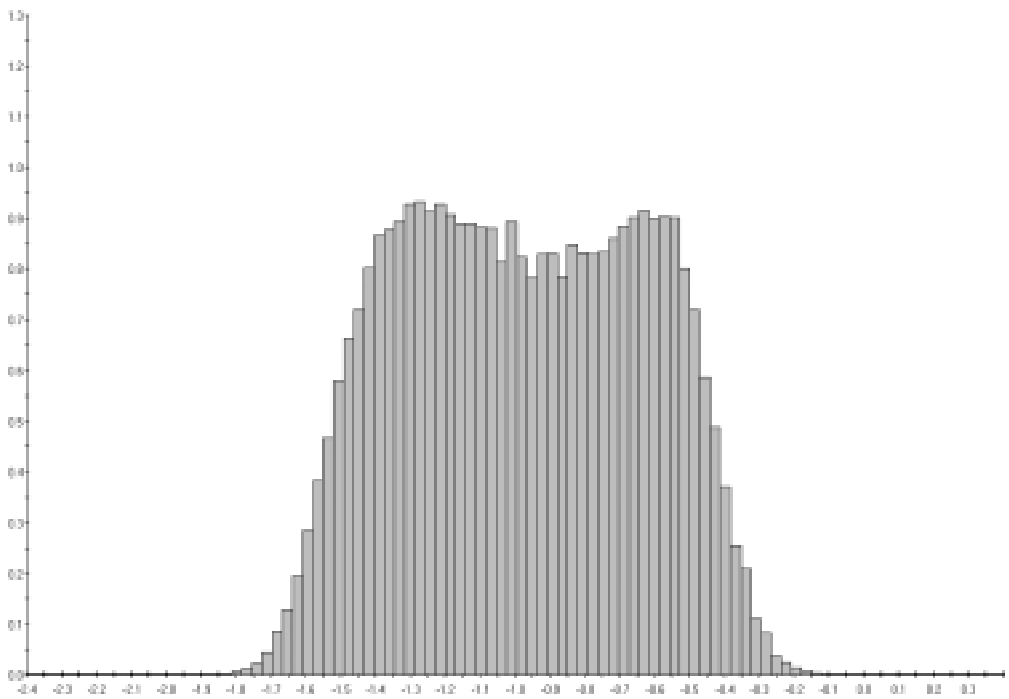

Figure 5: Simulated Density of the CIPS* Statistic - The Intercept Case

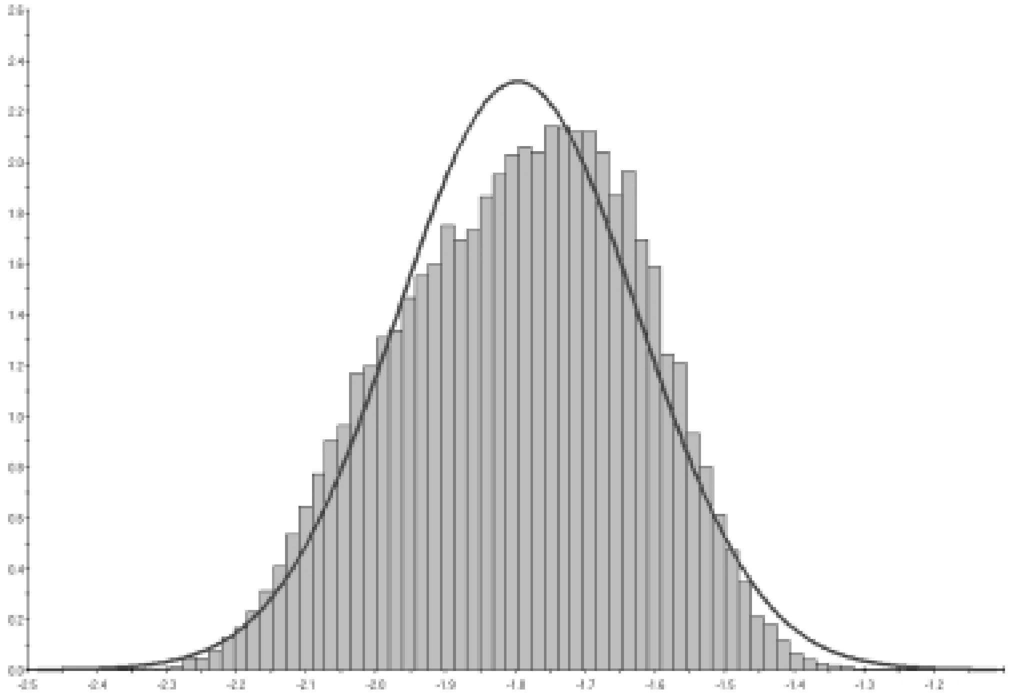

Figure 6: Simulated Density of the CIPS* Statistic - The Linear Trend Case

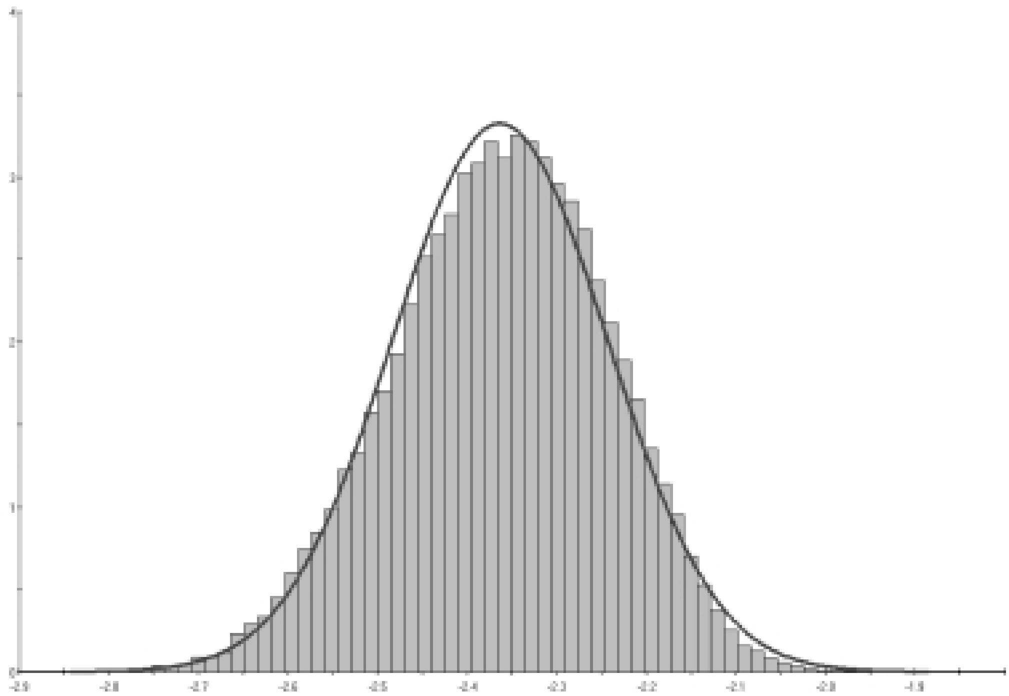

Table 1a
Critical Values of Individual Cross-Sectionally Augmented Dickey-Fuller Distribution (Case I: No Intercept and No Trend)[1,2]

1% ($CADF_i$)

T/N	**10**	**15**	**20**	**30**	**50**	**70**	**100**	**200**
10	-4.23	-4.26	-4.21	-4.25	-4.22	-4.25	-4.31	-4.28
15	-3.69	-3.72	-3.67	-3.71	-3.75	-3.75	-3.70	-3.65
20	-3.53	-3.52	-3.48	-3.51	-3.55	-3.51	-3.52	-3.53
30	-3.40	-3.43	-3.38	-3.39	-3.41	-3.41	-3.40	-3.40
50	-3.33	-3.32	-3.33	-3.31	-3.31	-3.35	-3.33	-3.30
70	-3.28	-3.24	-3.29	-3.25	-3.28	-3.30	-3.28	-3.31
100	-3.26	-3.27	-3.25	-3.28	-3.30	-3.26	-3.25	-3.25
200	-3.21	-3.22	-3.25	-3.24	-3.24	-3.24	-3.25	-3.24

5% ($CADF_i$)

T/N	**10**	**15**	**20**	**30**	**50**	**70**	**100**	**200**
10	-2.95	-2.92	-2.91	-2.91	-2.93	-2.94	-2.96	-2.94
15	-2.75	-2.75	-2.75	-2.76	-2.77	-2.77	-2.75	-2.74
20	-2.69	-2.71	-2.70	-2.69	-2.71	-2.69	-2.68	-2.69
30	-2.66	-2.65	-2.66	-2.66	-2.66	-2.66	-2.65	-2.66
50	-2.63	-2.64	-2.63	-2.63	-2.61	-2.63	-2.62	-2.62
70	-2.60	-2.61	-2.61	-2.60	-2.61	-2.62	-2.62	-2.62
100	-2.60	-2.61	-2.61	-2.61	-2.60	-2.62	-2.60	-2.61
200	-2.60	-2.61	-2.60	-2.60	-2.59	-2.60	-2.60	-2.60

10% ($CADF_i$)

T/N	**10**	**15**	**20**	**30**	**50**	**70**	**100**	**200**
10	-2.39	-2.39	-2.37	-2.38	-2.39	-2.38	-2.39	-2.39
15	-2.31	-2.32	-2.31	-2.32	-2.32	-2.32	-2.32	-2.32
20	-2.29	-2.31	-2.30	-2.30	-2.30	-2.30	-2.30	-2.29
30	-2.29	-2.28	-2.28	-2.28	-2.28	-2.28	-2.27	-2.29
50	-2.28	-2.28	-2.26	-2.28	-2.27	-2.28	-2.26	-2.27
70	-2.26	-2.26	-2.28	-2.26	-2.27	-2.26	-2.27	-2.27
100	-2.26	-2.26	-2.27	-2.27	-2.26	-2.27	-2.26	-2.26
200	-2.26	-2.27	-2.26	-2.27	-2.26	-2.27	-2.27	-2.26

Notes: [1] The calculations are carried out for 50,000 replications based on the OLS regression of Δy_{it} on $y_{i,t-1}$, $\bar{y}_{t-1}$ and $\Delta\overline{y}_t$, where $\bar{y}_t = N^{-1}\,\Sigma_{j=1}^{N} y_{jt}$. The $CADF_i$ refers to the OLS t-ratio of the coefficient of $y_{i,t-1}$.

[2] The critical values for the truncated version of the test statistics are indicated in brackets if they differ from the non-truncated ones.

Table 1b
Critical Values of Individual Cross-Sectionally Augmented Dickey-Fuller Distribution (Case II: Intercept only)[1,2]

1% ($CADF_i$)

T/N	10	15	20	30	50	70	100	200
10	-5.75	-5.73	-5.78	-5.73	-5.71	-5.72	-5.89	-5.72
15	-4.65	-4.65	-4.62	-4.68	-4.66	-4.64	-4.69	-4.61
20	-4.35	-4.34	-4.32	-4.35	-4.35	-4.33	-4.36	-4.34
30	-4.11	-4.12	-4.11	-4.12	-4.11	-4.12	-4.11	-4.09
50	-3.94	-4.00	-3.99	-3.97	-3.95	-3.99	-3.96	-3.96
70	-3.92	-3.90	-3.91	-3.92	-3.94	-3.93	-3.91	-3.94
100	-3.88	-3.86	-3.87	-3.90	-3.86	-3.85	-3.85	-3.89
200	-3.81	-3.83	-3.84	-3.84	-3.83	-3.85	-3.83	-3.84

5% ($CADF_i$)

T/N	10	15	20	30	50	70	100	200
10	-3.93	-3.96	-3.94	-3.97	-3.94	-3.93	-3.96	-3.99
15	-3.53	-3.57	-3.54	-3.55	-3.55	-3.55	-3.57	-3.55
20	-3.43	-3.43	-3.42	-3.43	-3.43	-3.42	-3.44	-3.43
30	-3.36	-3.36	-3.34	-3.34	-3.34	-3.34	-3.33	-3.34
50	-3.29	-3.30	-3.28	-3.27	-3.27	-3.28	-3.28	-3.28
70	-3.26	-3.26	-3.27	-3.27	-3.27	-3.28	-3.26	-3.29
100	-3.24	-3.25	-3.24	-3.27	-3.26	-3.24	-3.24	-3.24
200	-3.22	-3.23	-3.23	-3.24	-3.24	-3.23	-3.24	-3.22

10% ($CADF_i$)

T/N	10	15	20	30	50	70	100	200
10	-3.26	-3.27	-3.24	-3.26	-3.25	-3.25	-3.28	-3.27
15	-3.06	-3.08	-3.06	-3.07	-3.07	-3.07	-3.07	-3.06
20	-3.00	-3.02	-3.01	-3.01	-3.01	-3.00	-3.02	-3.01
30	-2.97	-2.98	-2.96	-2.97	-2.97	-2.97	-2.95	-2.97
50	-2.94	-2.95	-2.94	-2.93	-2.94	-2.94	-2.93	-2.94
70	-2.93	-2.94	-2.94	-2.94	-2.93	-2.94	-2.93	-2.94
100	-2.92	-2.92	-2.92	-2.93	-2.93	-2.92	-2.91	-2.92
200	-2.91	-2.92	-2.91	-2.92	-2.92	-2.91	-2.92	-2.91

Notes: [1] The calculations are carried out for 50,000 replications based on the OLS regression of Δy_{it} on an intercept, $y_{i,t-1}$, $\bar{y}_{t-1}$ and $\Delta \bar{y}_t$, where $\bar{y}_t = N^{-1} \Sigma_{j=1}^{N} y_{jt}$. The $CADF_i$ refers to the OLS t-ratio of the coefficient of $y_{i,t-1}$.

[2] The critical values for the truncated version of the test statistics are indicated in brackets if they differ from the non-truncated ones.

Table 1c
Critical Values of Individual Cross-Sectionally Augmented Dickey-Fuller Distribution (Case III: Intercept and Trend)[1,2]

1% ($CADF_i$)

T/N	10	15	20	30	50	70	100	200
10	-7.49	-7.67	-7.50	-7.64	-7.69	-7.44	-7.40	-7.51
	(-6.40)	(-6.40)	(-6.40)	(-6.40)	(-6.40)	(-6.40)	(-6.40)	(-6.40)
15	-5.44	-5.46	-5.40	-5.50	-5.48	-5.42	-5.49	-5.41
20	-4.97	-4.98	-4.96	-4.97	-5.01	-5.00	-5.02	-4.95
30	-4.67	-4.67	-4.68	-4.69	-4.69	-4.64	-4.68	-4.68
50	-4.49	-4.51	-4.52	-4.51	-4.47	-4.46	-4.48	-4.47
70	-4.41	-4.41	-4.39	-4.41	-4.41	-4.41	-4.40	-4.42
100	-4.35	-4.35	-4.35	-4.34	-4.37	-4.35	-4.35	-4.35
200	-4.28	-4.32	-4.32	-4.30	-4.32	-4.28	-4.30	-4.31

5% ($CADF_i$)

T/N	10	15	20	30	50	70	100	200
10	-4.89	-4.93	-4.89	-4.87	-4.91	-4.90	-4.88	-4.88
15	-4.17	-4.17	-4.14	-4.18	-4.17	-4.19	-4.19	-4.17
20	-3.99	-3.99	-4.00	-4.01	-4.01	-4.00	-4.01	-4.01
30	-3.87	-3.88	-3.87	-3.88	-3.87	-3.86	-3.87	-3.87
50	-3.78	-3.79	-3.79	-3.80	-3.78	-3.78	-3.79	-3.79
70	-3.76	-3.75	-3.76	-3.75	-3.76	-3.76	-3.77	-3.78
100	-3.72	-3.74	-3.74	-3.74	-3.74	-3.73	-3.73	-3.74
200	-3.69	-3.71	-3.71	-3.71	-3.72	-3.72	-3.72	-3.71

10% ($CADF_i$)

T/N	10	15	20	30	50	70	100	200
10	-4.00	-4.00	-3.99	-4.00	-4.02	-3.99	-4.01	-4.02
15	-3.64	-3.63	-3.62	-3.65	-3.63	-3.63	-3.65	-3.64
20	-3.55	-3.54	-3.55	-3.56	-3.56	-3.55	-3.56	-3.56
30	-3.49	-3.49	-3.49	-3.49	-3.49	-3.49	-3.48	-3.49
50	-3.44	-3.44	-3.44	-3.45	-3.44	-3.43	-3.45	-3.45
70	-3.43	-3.43	-3.43	-3.43	-3.43	-3.44	-3.42	-3.44
100	-3.41	-3.42	-3.42	-3.43	-3.42	-3.42	-3.41	-3.42
200	-3.39	-3.39	-3.41	-3.40	-3.41	-3.41	-3.41	-3.41

Notes: [1] The calculations are carried out for 50,000 replications based on the OLS regression of Δy_{it} on an intercept, trend, $y_{i,t-1}$, $\bar{y}_{t-1}$ and $\Delta \bar{y}_t$, where $\bar{y}_t = N^{-1} \Sigma_{j=1}^{N} y_{jt}$. The $CADF_i$ refers to the OLS t-ratio of the coefficient of $y_{i,t-1}$.

[2] The critical values for the truncated version of the test statistics are indicated in brackets if they differ from the non-truncated ones.

Table 2a
Summary Statistics of Individual Cross-Sectionally Augmented Dickey-Fuller Distribution (Case I: No Intercept and No Trend)[1]

Mean: $E(CADF_i)$

T/N	10	15	20	30	50	70	100	200
10	-0.91	-0.92	-0.91	-0.91	-0.91	-0.92	-0.91	-0.91
15	-0.93	-0.93	-0.92	-0.93	-0.93	-0.94	-0.93	-0.93
20	-0.94	-0.95	-0.94	-0.94	-0.94	-0.94	-0.94	-0.94
30	-0.95	-0.96	-0.95	-0.95	-0.96	-0.95	-0.95	-0.96
50	-0.97	-0.96	-0.96	-0.96	-0.97	-0.96	-0.96	-0.97
70	-0.97	-0.96	-0.98	-0.97	-0.97	-0.97	-0.97	-0.97
100	-0.97	-0.97	-0.97	-0.97	-0.97	-0.97	-0.97	-0.98
200	-0.98	-0.98	-0.97	-0.98	-0.97	-0.98	-0.98	-0.98

Standard deviation: $[Var(CADF_i)]^{1/2}$

T/N	10	15	20	30	50	70	100	200
10	1.26	1.25	1.25	1.25	1.25	1.25	1.27	1.26
15	1.14	1.14	1.14	1.14	1.15	1.15	1.14	1.14
20	1.11	1.11	1.11	1.11	1.11	1.11	1.11	1.11
30	1.09	1.08	1.08	1.09	1.09	1.09	1.09	1.09
50	1.07	1.07	1.06	1.07	1.07	1.07	1.07	1.07
70	1.05	1.06	1.06	1.06	1.06	1.06	1.06	1.06
100	1.05	1.05	1.05	1.06	1.05	1.06	1.06	1.05
200	1.05	1.05	1.05	1.05	1.05	1.05	1.05	1.05

Correlation ($CADF_i$, $CADF_j$) - Direct Calculations

T/N	10	15	20	30	50	70	100	200
10	0.11	0.10	0.11	0.10	0.10	0.10	0.10	0.10
15	0.11	0.11	0.11	0.11	0.11	0.10	0.10	0.12
20	0.12	0.10	0.11	0.11	0.11	0.10	0.11	0.11
30	0.11	0.11	0.11	0.11	0.11	0.10	0.10	0.10
50	0.11	0.10	0.10	0.10	0.11	0.10	0.10	0.10
70	0.11	0.10	0.11	0.10	0.10	0.10	0.10	0.10
100	0.11	0.10	0.11	0.10	0.10	0.10	0.09	0.10
200	0.11	0.11	0.10	0.10	0.10	0.10	0.10	0.10

Derived correlation ($CADF_i$, $CADF_j$)

T/N	10	15	20	30	50	70	100	200
10	0.11	0.11	0.10	0.10	0.10	0.10	0.10	0.10
15	0.11	0.11	0.11	0.10	0.11	0.10	0.11	0.11
20	0.11	0.11	0.11	0.11	0.11	0.11	0.11	0.11
30	0.11	0.11	0.11	0.10	0.11	0.10	0.10	0.10
50	0.11	0.10	0.11	0.10	0.10	0.10	0.10	0.10
70	0.11	0.11	0.10	0.10	0.10	0.10	0.10	0.10
100	0.11	0.10	0.10	0.10	0.10	0.10	0.10	0.10
200	0.11	0.11	0.10	0.10	0.10	0.10	0.10	0.10

Notes: [1] See note 1 to Table 1a.

Table 2b
Summary Statistics of Individual Cross-Sectionally Augmented Dickey-Fuller Distribution (Case II: Intercept only)[1]

Mean: $E(CADF_i)$

T/N	10	15	20	30	50	70	100	200
10	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69	-1.69
15	-1.71	-1.73	-1.71	-1.72	-1.72	-1.72	-1.71	-1.71
20	-1.73	-1.75	-1.73	-1.74	-1.73	-1.73	-1.73	-1.74
30	-1.76	-1.77	-1.75	-1.75	-1.76	-1.75	-1.74	-1.76
50	-1.78	-1.77	-1.77	-1.77	-1.77	-1.77	-1.77	-1.78
70	-1.78	-1.78	-1.78	-1.78	-1.78	-1.78	-1.78	-1.78
100	-1.78	-1.78	-1.78	-1.79	-1.79	-1.78	-1.78	-1.78
200	-1.79	-1.79	-1.79	-1.79	-1.79	-1.79	-1.79	-1.79

Standard deviation: $[Var(CADF_i)]^{1/2}$

T/N	10	15	20	30	50	70	100	200
10	1.35	1.34	1.34	1.34	1.33	1.33	1.35	1.34
15	1.09	1.09	1.09	1.10	1.09	1.10	1.10	1.09
20	1.03	1.02	1.02	1.02	1.02	1.02	1.03	1.02
30	0.97	0.97	0.97	0.98	0.97	0.97	0.97	0.97
50	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
70	0.92	0.92	0.92	0.92	0.92	0.93	0.92	0.93
100	0.91	0.91	0.91	0.92	0.92	0.92	0.91	0.92
200	0.90	0.90	0.90	0.91	0.91	0.90	0.91	0.90

Correlation $(CADF_i, CADF_j)$ - Direct Calculations

T/N	10	15	20	30	50	70	100	200
10	0.03	0.03	0.03	0.03	0.02	0.04	0.03	0.03
15	0.04	0.03	0.02	0.03	0.02	0.03	0.02	0.03
20	0.03	0.03	0.03	0.03	0.02	0.04	0.03	0.03
30	0.04	0.02	0.03	0.03	0.03	0.03	0.03	0.03
50	0.03	0.02	0.03	0.03	0.03	0.02	0.03	0.03
70	0.03	0.03	0.03	0.03	0.03	0.02	0.03	0.02
100	0.03	0.03	0.03	0.03	0.03	0.02	0.03	0.03
200	0.04	0.03	0.02	0.03	0.03	0.02	0.02	0.03

Derived correlation $(CADF_i, CADF_j)$

T/N	10	15	20	30	50	70	100	200
10	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
15	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
20	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
30	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
50	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
70	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
100	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
200	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03

Notes: [1] See note 1 to Table 1b.

Table 2c
Summary Statistics of Individual Cross-Sectionally Augmented Dickey-Fuller Distribution (Case III: Intercept and Trend)[1]

Mean: $E(CADF_i)$

T/N	10	15	20	30	50	70	100	200
10	-2.24	-2.24	-2.24	-2.24	-2.26	-2.23	-2.24	-2.25
15	-2.25	-2.26	-2.25	-2.26	-2.26	-2.26	-2.26	-2.25
20	-2.28	-2.28	-2.28	-2.29	-2.29	-2.28	-2.29	-2.29
30	-2.31	-2.32	-2.31	-2.31	-2.31	-2.31	-2.31	-2.32
50	-2.34	-2.34	-2.34	-2.34	-2.34	-2.33	-2.33	-2.34
70	-2.35	-2.34	-2.34	-2.34	-2.35	-2.35	-2.34	-2.35
100	-2.35	-2.35	-2.35	-2.35	-2.36	-2.36	-2.35	-2.35
200	-2.36	-2.36	-2.36	-2.36	-2.36	-2.36	-2.35	-2.36

Standard deviation: $[Var(CADF_i)]^{1/2}$

T/N	10	15	20	30	50	70	100	200
10	1.57	1.60	1.58	1.58	1.63	1.56	1.55	1.59
15	1.11	1.11	1.11	1.12	1.11	1.11	1.13	1.11
20	1.01	1.00	1.01	1.01	1.01	1.01	1.01	1.01
30	0.93	0.93	0.93	0.93	0.93	0.93	0.93	0.93
50	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
70	0.86	0.86	0.86	0.86	0.86	0.86	0.86	0.86
100	0.84	0.84	0.84	0.84	0.85	0.84	0.84	0.85
200	0.83	0.83	0.83	0.83	0.83	0.83	0.84	0.83

Correlation $(CADF_i, CADF_j)$ - Direct Calculations

T/N	10	15	20	30	50	70	100	200
10	0.04	0.03	0.03	0.02	0.03	0.04	0.03	0.04
15	0.03	0.03	0.02	0.03	0.02	0.02	0.01	0.02
20	0.04	0.02	0.02	0.02	0.03	0.02	0.01	0.02
30	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02
50	0.02	0.02	0.01	0.01	0.02	0.01	0.02	0.01
70	0.02	0.01	0.02	0.02	0.02	0.01	0.02	0.01
100	0.02	0.01	0.02	0.01	0.02	0.00	0.01	0.01
200	0.02	0.02	0.02	0.01	0.02	0.02	0.01	0.01

Derived correlation $(CADF_i, CADF_j)$

T/N	10	15	20	30	50	70	100	200
10	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
15	0.03	0.02	0.03	0.02	0.02	0.02	0.02	0.02
20	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02
30	0.03	0.02	0.02	0.02	0.02	0.02	0.02	0.02
50	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01
70	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01
100	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01
200	0.02	0.02	0.01	0.01	0.01	0.01	0.01	0.01

Notes: [1] See note 1 to Table 1c.

Table 3a
Critical Values of Average of Individual Cross-Sectionally Augmented Dickey-Fuller Distribution (Case I: No Intercept and No Trend)[1]

1% ($\overline{CADF}$)

T/N	10	15	20	30	50	70	100	200
10	-2.16	-2.02	-1.93	-1.85	-1.78	-1.74	-1.71	-1.70
	(-2.14)	(-2.00)	(-1.91)	(-1.84)	(-1.77)	(-1.73)		(-1.69)
15	-2.03	-1.91	-1.84	-1.77	-1.71	-1.68	-1.66	-1.63
20	-2.00	-1.89	-1.83	-1.76	-1.70	-1.67	-1.65	-1.62
30	-1.98	-1.87	-1.80	-1.74	-1.69	-1.67	-1.64	-1.61
50	-1.97	-1.86	-1.80	-1.74	-1.69	-1.66	-1.63	-1.61
70	-1.95	-1.86	-1.80	-1.74	-1.68	-1.66	-1.63	-1.61
100	-1.94	-1.85	-1.79	-1.74	-1.68	-1.65	-1.63	-1.61
200	-1.95	-1.85	-1.79	-1.73	-1.68	-1.65	-1.63	-1.61

5% ($\overline{CADF}$)

T/N	10	15	20	30	50	70	100	200
10	-1.80	-1.71	-1.67	-1.61	-1.58	-1.56	-1.54	-1.53
	(-1.79)		(-1.66)		(-1.57)	(-1.55)	(-1.53)	(-1.52)
15	-1.74	-1.67	-1.63	-1.58	-1.55	-1.53	-1.52	-1.51
20	-1.72	-1.65	-1.62	-1.58	-1.54	-1.53	-1.52	-1.50
30	-1.72	-1.65	-1.61	-1.57	-1.55	-1.54	-1.52	-1.50
50	-1.72	-1.64	-1.61	-1.57	-1.54	-1.53	-1.52	-1.51
70	-1.71	-1.65	-1.61	-1.57	-1.54	-1.53	-1.52	-1.51
100	-1.71	-1.64	-1.61	-1.57	-1.54	-1.53	-1.52	-1.51
200	-1.71	-1.65	-1.61	-1.57	-1.54	-1.53	-1.52	-1.51

10% ($\overline{CADF}$)

T/N	10	15	20	30	50	70	100	200
10	-1.61	-1.56	-1.52	-1.49	-1.46	-1.45	-1.44	-1.43
		(-1.55)		(-1.48)			(-1.43)	
15	-1.58	-1.53	-1.50	-1.48	-1.45	-1.44	-1.44	-1.43
20	-1.58	-1.52	-1.50	-1.47	-1.45	-1.45	-1.44	-1.43
30	-1.57	-1.53	-1.50	-1.47	-1.46	-1.45	-1.44	-1.43
50	-1.58	-1.52	-1.50	-1.47	-1.45	-1.45	-1.44	-1.43
70	-1.57	-1.52	-1.50	-1.47	-1.46	-1.45	-1.44	-1.43
100	-1.56	-1.52	-1.50	-1.48	-1.46	-1.45	-1.44	-1.43
200	-1.57	-1.53	-1.50	-1.47	-1.45	-1.45	-1.44	-1.43

Notes: [1] $\overline{CADF}$ statistic is computed as the simple average of the individual-specific $CADF_i$ statistics. See notes to Table 1a.

Table 3b
Critical Values of Average of Individual Cross-Sectionally Augmented Dickey-Fuller Distribution (Case II: Intercept only)[1]

1% ($\overline{CADF}$)

T/N	10	15	20	30	50	70	100	200
10	-2.97	-2.76	-2.64	-2.51	-2.41	-2.37	-2.33	-2.28
	(-2.85)	(-2.66)	(-2.56)	(-2.44)	(-2.36)	(-2.32)	(-2.29)	(-2.25)
15	-2.66	-2.52	-2.45	-2.34	-2.26	-2.23	-2.19	-2.16
20	-2.60	-2.47	-2.40	-2.32	-2.25	-2.20	-2.18	-2.14
30	-2.57	-2.45	-2.38	-2.30	-2.23	-2.19	-2.17	-2.14
50	-2.55	-2.44	-2.36	-2.30	-2.23	-2.20	-2.17	-2.14
70	-2.54	-2.43	-2.36	-2.30	-2.23	-2.20	-2.17	-2.14
100	-2.53	-2.42	-2.36	-2.30	-2.23	-2.20	-2.18	-2.15
200	-2.53	-2.43	-2.36	-2.30	-2.23	-2.21	-2.18	-2.15

5% ($\overline{CADF}$)

T/N	10	15	20	30	50	70	100	200
10	-2.52	-2.40	-2.33	-2.25	-2.19	-2.16	-2.14	-2.10
	(-2.47)	(-2.35)	(-2.29)	(-2.22)	(-2.16)	(-2.13)	(-2.11)	(-2.08)
15	-2.37	-2.28	-2.22	-2.17	-2.11	-2.09	-2.07	-2.04
20	-2.34	-2.26	-2.21	-2.15	-2.11	-2.08	-2.07	-2.04
30	-2.33	-2.25	-2.20	-2.15	-2.11	-2.08	-2.07	-2.05
50	-2.33	-2.25	-2.20	-2.16	-2.11	-2.10	-2.08	-2.06
70	-2.33	-2.25	-2.20	-2.15	-2.12	-2.10	-2.08	-2.06
100	-2.32	-2.25	-2.20	-2.16	-2.12	-2.10	-2.08	-2.07
200	-2.32	-2.25	-2.20	-2.16	-2.12	-2.10	-2.08	-2.07

10% ($\overline{CADF}$)

T/N	10	15	20	30	50	70	100	200
10	-2.31	-2.22	-2.18	-2.12	-2.07	-2.05	-2.03	-2.01
	(-2.28)	(-2.20)	(-2.15)	(-2.10)	(-2.05)	(-2.03)	(-2.01)	(-1.99)
15	-2.22	-2.16	-2.11	-2.07	-2.03	-2.01	-2.00	-1.98
20	-2.21	-2.14	-2.10	-2.07	-2.03	-2.01	-2.00	-1.99
30	-2.21	-2.14	-2.11	-2.07	-2.04	-2.02	-2.01	-2.00
50	-2.21	-2.14	-2.11	-2.08	-2.05	-2.03	-2.02	-2.01
70	-2.21	-2.15	-2.11	-2.08	-2.05	-2.03	-2.02	-2.01
100	-2.21	-2.15	-2.11	-2.08	-2.05	-2.03	-2.03	-2.02
200	-2.21	-2.15	-2.11	-2.08	-2.05	-2.04	-2.03	-2.02

Notes: [1] $\overline{CADF}$ statistic is computed as the simple average of the individual-specific $CADF_i$ statistics. See notes to Table 1b.

Table 3c
Critical Values of Average of Individual Cross-Sectionally Augmented Dickey-Fuller Distribution (Case III: Intercept and Trend)[1]

1% ($\overline{CADF}$)

T/N	**10**	**15**	**20**	**30**	**50**	**70**	**100**	**200**
10	-3.88	-3.61	-3.46	-3.30	-3.15	-3.10	-3.05	-2.98
	(-3.51)	(-3.31)	(-3.20)	(-3.10)	(-3.00)	(-2.96)	(-2.93)	(-2.88)
15	-3.24	-3.09	-3.00	-2.89	-2.81	-2.77	-2.74	-2.71
	(-3.21)	(-3.07)	(-2.98)	(-2.88)	(-2.80)	(-2.76)		(-2.70)
20	-3.15	-3.01	-2.92	-2.83	-2.76	-2.72	-2.70	-2.65
30	-3.10	-2.96	-2.88	-2.81	-2.73	-2.69	-2.66	-2.63
50	-3.06	-2.93	-2.85	-2.78	-2.72	-2.68	-2.65	-2.62
70	-3.04	-2.93	-2.85	-2.78	-2.71	-2.68	-2.65	-2.62
100	-3.03	-2.92	-2.85	-2.77	-2.71	-2.68	-2.65	-2.62
200	-3.03	-2.91	-2.85	-2.77	-2.71	-2.67	-2.65	-2.62

5% ($\overline{CADF}$)

T/N	**10**	**15**	**20**	**30**	**50**	**70**	**100**	**200**
10	-3.27	-3.11	-3.02	-2.94	-2.86	-2.82	-2.79	-2.75
	(-3.10)	(-2.97)	(-2.89)	(-2.82)	(-2.75)	(-2.73)	(-2.70)	(-2.67)
15	-2.93	-2.83	-2.77	-2.70	-2.64	-2.62	-2.60	-2.57
	(-2.92)	(-2.82)	(-2.76)	(-2.69)			(-2.59)	
20	-2.88	-2.78	-2.73	-2.67	-2.62	-2.59	-2.57	-2.55
30	-2.86	-2.76	-2.72	-2.66	-2.61	-2.58	-2.56	-2.54
50	-2.84	-2.76	-2.71	-2.65	-2.60	-2.58	-2.56	-2.54
70	-2.83	-2.76	-2.70	-2.65	-2.61	-2.58	-2.57	-2.54
100	-2.83	-2.75	-2.70	-2.65	-2.61	-2.59	-2.56	-2.55
200	-2.83	-2.75	-2.70	-2.65	-2.61	-2.59	-2.57	-2.55

10% ($\overline{CADF}$)

T/N	**10**	**15**	**20**	**30**	**50**	**70**	**100**	**200**
10	-2.98	-2.89	-2.82	-2.76	-2.71	-2.68	-2.66	-2.63
	(-2.87)	(-2.78)	(-2.73)	(-2.67)	(-2.63)	(-2.60)	(-2.58)	(-2.56)
15	-2.76	-2.69	-2.65	-2.60	-2.56	-2.54	-2.52	-2.50
		(-2.68)	(-2.64)	(-2.59)	(-2.55)	(-2.53)	(-2.51)	
20	-2.74	-2.67	-2.63	-2.58	-2.54	-2.53	-2.51	-2.49
30	-2.73	-2.66	-2.63	-2.58	-2.54	-2.52	-2.51	-2.49
50	-2.73	-2.66	-2.63	-2.58	-2.55	-2.53	-2.51	-2.50
70	-2.72	-2.66	-2.62	-2.58	-2.55	-2.53	-2.52	-2.50
100	-2.72	-2.66	-2.63	-2.59	-2.55	-2.53	-2.52	-2.50
200	-2.73	-2.66	-2.63	-2.59	-2.55	-2.54	-2.52	-2.51

Notes: [1] $\overline{CADF}$ statistic is computed as the simple average of the individual-specific $CADF_i$ statistics. See notes to Table 1c.

Table 4: Size of Panel Unit Root Tests that Do Not Allow for Cross-Sectional Dependence No Serial Correlation, Low and High Cross Section Dependence, Intercept Case[1]

		Low Cross Section Dependence					High Cross Section Dependence				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	IPS	.041	.046	.056	.051	.046	.093	.117	.122	.111	.114
	IPS*	.040	.046	.056	.051	.046	.095	.118	.122	.111	.114
	P-test (DF)										
	- Normal approximation	.186	.105	.092	.074	.063	.207	.159	.155	.126	.124
	- Empirical Distribution	.030	.042	.048	.048	.041	.054	.082	.108	.092	.095
	Z-test (DF)										
	- Normal approximation	.078	.055	.066	.057	.048	.139	.135	.135	.122	.122
	- Empirical Distribution	.039	.043	.056	.052	.043	.094	.118	.123	.114	.119
20	IPS	.043	.054	.045	.048	.051	.179	.218	.213	.215	.246
	IPS*	.045	.054	.045	.048	.051	.182	.219	.213	.215	.246
	P-test (DF)										
	- Normal approximation	.227	.123	.084	.083	.071	.242	.216	.178	.188	.207
	- Empirical Distribution	.035	.053	.041	.049	.051	.104	.148	.134	.157	.191
	Z-test (DF)										
	- Normal approximation	.067	.058	.047	.056	.052	.214	.227	.217	.221	.243
	- Empirical Distribution	.043	.051	.039	.052	.052	.178	.213	.204	.216	.243
30	IPS	.037	.035	.052	.047	.060	.208	.234	.236	.256	.283
	IPS*	.038	.034	.052	.047	.060	.210	.234	.236	.256	.283
	P-test (DF)										
	- Normal approximation	.242	.119	.103	.071	.067	.299	.237	.222	.221	.231
	- Empirical Distribution	.027	.039	.056	.047	.051	.125	.157	.163	.180	.208
	Z-test (DF)										
	- Normal approximation	.066	.042	.056	.045	.052	.251	.245	.239	.253	.277
	- Empirical Distribution	.042	.035	.050	.042	.052	.214	.234	.237	.249	.277
50	IPS	.046	.032	.050	.047	.039	.285	.314	.318	.334	.370
	IPS*	.047	.032	.050	.047	.039	.291	.314	.318	.334	.370
	P-test (DF)										
	- Normal approximation	.314	.145	.128	.089	.059	.327	.274	.271	.277	.291
	- Empirical Distribution	.040	.032	.053	.054	.044	.168	.214	.223	.246	.271
	Z-test (DF)										
	- Normal approximation	.075	.035	.052	.047	.041	.307	.314	.322	.330	.354
	- Empirical Distribution	.043	.032	.050	.048	.042	.285	.305	.316	.330	.362

Table 4: Size of Panel Unit Root Tests that Do Not Allow for Cross-Sectional Dependence No Serial Correlation, Low and High Cross Section Dependence, Intercept Case (continued)

		Low Cross Section Dependence					High Cross Section Dependence				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
100	IPS	.035	.067	.059	.057	.059	.330	.382	.372	.383	.406
	IPS*	.036	.067	.059	.057	.059	.336	.383	.372	.383	.406
	P-test (DF)										
	- Normal approximation	.409	.181	.120	.090	.091	.379	.340	.306	.310	.325
	- Empirical Distribution	.026	.049	.040	.048	.054	.218	.259	.263	.289	.314
	Z-test (DF)										
	- Normal approximation	.052	.062	.041	.048	.049	.345	.367	.354	.376	.394
	- Empirical Distribution	.038	.062	.044	.049	.052	.330	.367	.355	.378	.394

Notes: [1] This table reports the size of various test statistics defined in the paper. The underlying data is generated by $y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + u_{it}, i = 1,2,.., N, t = -51,-50,..., T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[0,0.2]$ for low cross section dependence, and $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and f_t ~ iidN(0,1), and $\varepsilon_{it} \sim iidN(0, \sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$. The test statistics are computed using various regression specifications as specified in the text for the case with intercept only and no lag augmentation. The size (under the null $\phi_i = 1$) of the tests are computed at the 5% nominal level. 'DF' refers to the Dickey-Fuller regression with an intercept defined as $\Delta y_{it} = a_i + b_i y_{i,t-1} + u_{it}$, $i = 1,2,.., N, t = 1,2,..., T$. The IPS test statistic is defined as $IPS = \frac{\sqrt{N}\{\bar{t} - E[t_i \mid \phi_i = 0]\}}{\sqrt{Var[t_i \mid \phi_i = 0]}}$, where t_i is the OLS t-ratio of b_i in the 'DF' regression defined above, $\bar{t}$ is the simple average of these t-ratios. The IPS* is the truncated IPS statistic computed by applying the truncation procedure simlar to that used in (4.36) to the individual DF (or ADF) statistics. For P-test and Z-test, the normal and empirical approximations distinguish the use of standard normal critical values and empirical critical values obtained by stochastic simulations.

[2] Simulation results reported in this and the subsequent tables are based on 1,000 replications.

Table 5a: Size of Cross-Sectionally Augmented Panel Unit Root Tests
No Serial Correlation, Low and High Cross Section Dependence, Intercept Case[1]

		Low Cross Section Dependence					High Cross Section Dependence				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.032	.035	.060	.051	.059	.043	.048	.057	.052	.063
	CIPS*	.041	.035	.060	.051	.059	.049	.048	.057	.052	.063
	CP-tests										
	- Normal approximation	.251	.114	.124	.096	.094	.248	.125	.117	.091	.087
	- Empirical Distribution	.027	.040	.053	.057	.056	.036	.038	.060	.052	.056
	CZ-tests										
	- Normal approximation	.102	.062	.090	.077	.087	.124	.078	.084	.073	.082
	- Empirical Distribution	.039	.034	.060	.051	.060	.050	.049	.055	.050	.060
20	CIPS	.034	.051	.053	.051	.047	.034	.062	.055	.064	.057
	CIPS*	.039	.049	.054	.051	.047	.041	.062	.055	.064	.057
	CP-tests										
	- Normal approximation	.268	.162	.136	.104	.085	.259	.171	.129	.116	.107
	- Empirical Distribution	.027	.044	.047	.046	.039	.028	.057	.056	.045	.057
	CZ-tests										
	- Normal approximation	.091	.094	.099	.100	.084	.088	.097	.093	.109	.105
	- Empirical Distribution	.039	.048	.052	.052	.047	.037	.063	.057	.063	.057
30	CIPS	.044	.045	.050	.053	.044	.046	.054	.054	.043	.045
	CIPS*	.047	.045	.050	.053	.044	.045	.054	.054	.043	.045
	CP-tests										
	- Normal approximation	.337	.164	.146	.119	.102	.342	.187	.133	.111	.103
	- Empirical Distribution	.031	.038	.056	.042	.037	.029	.053	.049	.044	.041
	CZ-tests										
	- Normal approximation	.114	.097	.105	.103	.109	.132	.123	.099	.099	.115
	- Empirical Distribution	.047	.043	.051	.051	.044	.047	.056	.052	.042	.044
50	CIPS	.038	.048	.050	.047	.053	.039	.045	.048	.046	.051
	CIPS*	.043	.048	.050	.047	.053	.040	.045	.048	.046	.051
	CP-tests										
	- Normal approximation	.405	.210	.175	.140	.140	.425	.193	.168	.155	.125
	- Empirical Distribution	.029	.042	.059	.047	.058	.026	.031	.052	.053	.052
	CZ-tests										
	- Normal approximation	.127	.120	.133	.129	.140	.123	.114	.135	.133	.136
	- Empirical Distribution	.041	.048	.051	.050	.054	.041	.046	.049	.046	.052

Table 5a: Size of Cross-Sectionally Augmented Panel Unit Root Tests
No Serial Correlation, Low and High Cross Section Dependence, Intercept Case (continued)

		Low Cross Section Dependence					High Cross Section Dependence				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
100	CIPS	.041	.046	.047	.052	.040	.037	.049	.046	.061	.044
	CIPS*	.046	.046	.047	.052	.040	.041	.049	.046	.061	.044
	CP-tests										
	- Normal approximation	.456	.279	.233	.194	.185	.467	.269	.226	.208	.192
	- Empirical Distribution	.030	.054	.035	.047	.039	.026	.049	.044	.064	.040
	CZ-tests										
	- Normal approximation	.158	.170	.183	.190	.218	.163	.154	.183	.217	.220
	- Empirical Distribution	.046	.045	.047	.050	.041	.040	.049	.048	.060	.044

Notes: [1] The underlying data is generated by $y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[0,0.2]$ for low cross section dependence, and $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and $f_t \sim$ iidN(0,1), and $\varepsilon_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$. Size (under the null $\phi_i = 1)$ is computed at the 5% nominal level, based on cross section augmented Dickey-Fuller (CADF) regressions: $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$. The CIPS test statistic is defined as $CIPS(N,T) = N^{-1}\sum_{i=1}^{N} t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regression. Similarly, the truncated CIPS statistic (CIPS*) is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$, defined by (4.36). The statitics underlying the two versions of the CP and CZ tests (normal and empirical approximations) are computed using (6.64) and (6.66), respectively.

Table 5b: Power of Cross-Sectionally Augmented Panel Unit Root Tests
No Serial Correlation, Low and High Cross Section Dependence, Intercept Case[1]

		Low Cross Section Dependence					High Cross Section Dependence				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.049	.101	.172	.357	.959	.062	.115	.192	.382	.958
	CIPS*	.054	.101	.172	.357	.959	.068	.115	.192	.382	.958
	$\tilde{C}P$	.041	.077	.118	.272	.907	.048	.095	.141	.300	.890
	$\tilde{C}Z$	.054	.100	.171	.352	.959	.066	.114	.188	.381	.957
20	CIPS	.068	.135	.240	.683	1.00	.071	.114	.243	.688	1.00
	CIPS*	.071	.134	.240	.683	1.00	.071	.114	.243	.688	1.00
	$\tilde{C}P$	.048	.099	.168	.518	1.00	.048	.087	.175	.520	1.00
	$\tilde{C}Z$	.071	.135	.238	.677	1.00	.072	.113	.237	.687	1.00
30	CIPS	.056	.135	.253	.676	1.00	.058	.122	.231	.674	1.00
	CIPS*	.067	.134	.252	.676	1.00	.067	.122	.231	.674	1.00
	$\tilde{C}P$	.035	.104	.192	.517	.999	.043	.085	.181	.536	1.00
	$\tilde{C}Z$	.066	.133	.249	.670	1.00	.068	.121	.229	.667	1.00
50	CIPS	.052	.157	.335	.854	1.00	.061	.146	.298	.849	1.00
	CIPS*	.057	.157	.335	.854	1.00	.069	.145	.298	.849	1.00
	$\tilde{C}P$	.041	.117	.237	.684	1.00	.045	.121	.228	.708	1.00
	$\tilde{C}Z$	.060	.157	.332	.844	1.00	.065	.146	.298	.843	1.00
100	CIPS	.060	.192	.387	.949	1.00	.055	.189	.379	.963	1.00
	CIPS*	.072	.193	.386	.949	1.00	.060	.188	.379	.963	1.00
	$\tilde{C}P$	.038	.152	.303	.864	1.00	.035	.141	.292	.887	1.00
	$\tilde{C}Z$	.068	.195	.382	.945	1.00	.060	.189	.377	.960	1.00

Notes: [1] The underlying data is generated by $y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + u_{it}$, , $i = 1,2,..,N$, $t = -51,-50,...,T$ and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[0,0.2]$ for low cross section dependence, and $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and f_t ~ iidN(0,1), and $\varepsilon_{it} \sim iidN(0,\sigma_i^{2})$ with $\sigma_i^{2} \sim iidU[0.5,1.5]$. Power, under the heterogeneous alternatives $\phi_i \sim iidU[0.85,0.95]$, is computed at the 5% nominal level based on the cross section augmented Dickey-Fuller regression (CADF): $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$. The CIPS test statistic is defined as $CIPS(N,T) = N^{-1}\sum_{i=1}^{N} t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regression. Similarly, truncated CIPS statistic is computed with $t_i(N,T)$ replaced by $t_i^{*}(N,T)$, defined by (4.36). The $\tilde{C}P$ and $\tilde{C}Z$ tests are based on the statitics defined by (6.64) and (6.66), respectively, using their corresponding empirical critical values.

Table 6a: Size of Cross-Sectionally Augmented Panel Unit Root Tests
Positive Serial Correlation, High Cross Section Dependence, Intercept Case[1]

		CADF(0)					CADF(1)				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.005	.012	.005	.005	.000	.090	.062	.051	.048	.051
	CIPS*	.005	.011	.005	.005	.000	.055	.062	.051	.048	.051
	$\tilde{CP}$	.015	.015	.006	.013	.002	.122	.077	.062	.054	.048
	$\tilde{CZ}$	.005	.011	.004	.005	.000	.056	.065	.051	.050	.050
20	CIPS	.007	.000	.000	.000	.000	.098	.056	.047	.042	.053
	CIPS*	.005	.000	.000	.000	.000	.062	.054	.047	.042	.053
	$\tilde{CP}$	.008	.003	.004	.001	.001	.148	.077	.063	.065	.066
	$\tilde{CZ}$	.005	.000	.000	.000	.000	.060	.054	.048	.043	.053
30	CIPS	.004	.000	.000	.000	.000	.083	.033	.049	.037	.047
	CIPS*	.001	.000	.000	.000	.000	.052	.033	.049	.037	.047
	$\tilde{CP}$	.010	.003	.003	.000	.001	.177	.053	.068	.051	.048
	$\tilde{CZ}$	.001	.000	.000	.000	.000	.052	.034	.049	.038	.048
50	CIPS	.000	.000	.000	.000	.000	.072	.039	.035	.028	.047
	CIPS*	.000	.000	.000	.000	.000	.040	.039	.035	.028	.047
	$\tilde{CP}$	.005	.001	.000	.000	.000	.163	.069	.063	.044	.051
	$\tilde{CZ}$	.000	.000	.000	.000	.000	.039	.039	.037	.029	.046
100	CIPS	.000	.000	.000	.000	.000	.069	.037	.029	.037	.040
	CIPS*	.000	.000	.000	.000	.000	.026	.037	.029	.037	.040
	$\tilde{CP}$	.000	.000	.000	.000	.000	.194	.080	.056	.053	.055
	$\tilde{CZ}$	.000	.000	.000	.000	.000	.027	.039	.031	.039	.041

Notes: [1] The underlying data is generated by $y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and $f_t \sim$ iidN(0,1), and $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}$, where $e_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$ and $\rho_i \sim iidU[0.2,0.4]$. Size (under the null $\phi_i = 1$) is computed at the 5% nominal level based on cross section augmented Dickey-Fuller regressions: $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$, and the CADF(1) regressions: $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + d_{i1}\Delta\bar{y}_{t-1} + \delta_{i1}\Delta y_{i,t-1} + u_{it}$. The CIPS test statistic is defined by $CIPS(N,T) = N^{-1}\sum_{i=1}^{N} t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regressions. Similarly, truncated CIPS statistic is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$ as defined by (4.36). The $\tilde{CP}$ and $\tilde{CZ}$ tests are based on the statitics defined by (6.64) and (6.66), respectively, using their corresponding empirical critical values.

Table 6b: Power of Cross-Sectionally Augmented Panel Unit Root Tests Positive Serial Correlation, High Cross Section Dependence, Intercept Case[1]

		CADF(0)					CADF(1)				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.015	.010	.013	.011	.101	.131	.094	.109	.269	.873
	CIPS*	.018	.010	.013	.011	.101	.100	.093	.109	.269	.873
	$\tilde{CP}$	.020	.019	.017	.011	.052	.161	.105	.111	.222	.778
	$\tilde{CZ}$	.017	.010	.013	.011	.094	.100	.094	.111	.268	.868
20	CIPS	.006	.002	.003	.004	.143	.125	.075	.129	.412	.992
	CIPS*	.007	.002	.003	.004	.143	.080	.074	.129	.412	.992
	$\tilde{CP}$	.010	.006	.000	.002	.040	.183	.093	.124	.330	.973
	$\tilde{CZ}$	.007	.001	.003	.004	.133	.081	.075	.129	.403	.992
30	CIPS	.002	.000	.002	.000	.213	.126	.072	.171	.420	.999
	CIPS*	.003	.000	.002	.000	.213	.065	.071	.171	.420	.999
	$\tilde{CP}$	.005	.001	.003	.002	.112	.211	.122	.173	.358	.990
	$\tilde{CZ}$	.003	.000	.002	.000	.204	.064	.075	.172	.416	.999
50	CIPS	.000	.000	.000	.000	.348	.113	.089	.160	.595	1.00
	CIPS*	.000	.000	.000	.000	.348	.065	.089	.160	.595	1.00
	$\tilde{CP}$	.003	.001	.001	.000	.129	.222	.119	.175	.477	1.00
	$\tilde{CZ}$	.000	.000	.000	.000	.332	.065	.094	.162	.586	1.00
100	CIPS	.001	.000	.000	.000	.398	.090	.088	.187	.713	1.00
	CIPS*	.001	.000	.000	.000	.398	.044	.087	.188	.713	1.00
	$\tilde{CP}$	.001	.000	.000	.000	.149	.235	.167	.213	.641	1.00
	$\tilde{CZ}$	.001	.000	.000	.000	.371	.047	.091	.193	.713	1.00

Notes: [1] The underlying data is generated by $y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + u_{it}, i = 1,2,.., N, t = -51,-50,..., T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and $f_t \sim$ iidN(0,1), and $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}$, where $e_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$ and $\rho_i \sim iidU[0.2,0.4]$. Power, under the heterogeneous alternatives $\phi_i \sim iidU[0.85,0.95]$, is computed at the 5% nominal level based on the cross section augmented Dickey-Fuller regressions: $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,.., N, t = 1,2,..., T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$, and the CADF(1) regressions: $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + d_{i1}\Delta\bar{y}_{t-1} + \delta_{i1}\Delta y_{i,t-1} + u_{it}$. The CIPS test statistic is given by $CIPS(N,T) = N^{-1}\sum_{i=1}^{N} t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regressions. Similarly, truncated CIPS statistic is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$, defined by (4.36). The $\tilde{CP}$ and $\tilde{CZ}$ tests are based on the statitics defined by (6.64) and (6.66), respectively, using their corresponding empirical critical values.

Table 7a: Size of Cross-Sectionally Augmented Panel Unit Root Tests
Negative Serial Correlation, High Cross Section Dependence, Intercept Case[1]

		CADF(0)					CADF(1)				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.237	.436	.498	.531	.579	.101	.065	.057	.044	.049
	CIPS*	.258	.436	.498	.531	.579	.074	.065	.057	.044	.049
	$\tilde{C}P$	.185	.403	.501	.542	.602	.117	.068	.052	.054	.051
	$\tilde{C}Z$	.255	.437	.501	.532	.582	.072	.065	.056	.045	.050
20	CIPS	.365	.595	.646	.724	.746	.106	.054	.051	.046	.056
	CIPS*	.372	.596	.646	.724	.746	.076	.053	.051	.046	.056
	$\tilde{C}P$	.298	.582	.674	.751	.798	.155	.061	.059	.048	.059
	$\tilde{C}Z$	.375	.593	.646	.724	.747	.078	.053	.050	.045	.055
30	CIPS	.423	.694	.748	.808	.822	.103	.046	.056	.034	.046
	CIPS*	.445	.694	.748	.807	.822	.065	.046	.056	.034	.046
	$\tilde{C}P$	.346	.708	.779	.847	.881	.133	.048	.060	.045	.052
	$\tilde{C}Z$	.444	.695	.751	.806	.825	.064	.046	.056	.034	.047
50	CIPS	.499	.795	.854	.886	.919	.103	.042	.041	.030	.047
	CIPS*	.513	.795	.854	.886	.919	.068	.042	.041	.030	.047
	$\tilde{C}P$	.387	.816	.887	.928	.955	.165	.051	.046	.032	.044
	$\tilde{C}Z$	.508	.795	.857	.890	.920	.067	.045	.041	.028	.047
100	CIPS	.587	.835	.894	.941	.967	.094	.044	.040	.045	.045
	CIPS*	.602	.834	.894	.941	.966	.051	.045	.040	.045	.045
	$\tilde{C}P$	.520	.898	.945	.969	.986	.173	.060	.046	.047	.053
	$\tilde{C}Z$	.601	.837	.898	.941	.967	.052	.045	.041	.044	.044

Notes: [1] The underlying data is generated by $y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and $f_t \sim$ iidN(0,1), and $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}$, where $e_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$ and $\rho_i \sim iidU[-0.4,-0.2]$. Size (under the null $\phi_i = 1$) is computed at the 5% nominal level based on the cross section augmented Dickey-Fuller regressions: $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^N y_{it}$, and the CADF(1) regressions: $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + d_{i1}\Delta\bar{y}_{t-1} + \delta_{i1}\Delta y_{i,t-1} + u_{it}$. The CIPS test statistic is defined by $CIPS(N,T) = N^{-1}\sum_{i=1}^N t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regressions. Similarly, truncated CIPS statistic is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$ as defined by (4.36). The $\tilde{C}P$ and $\tilde{C}Z$ tests are based on the statitics defined by (6.64) and (6.66), respectively, using their corresponding empirical critical values.

Table 7b: Power of Cross-Sectionally Augmented Panel Unit Root Tests
Negative Serial Correlation, High Cross Section Dependence, Intercept Case[1]

		CADF(0)					CADF(1)				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.272	.617	.820	.980	1.00	.130	.096	.134	.344	.938
	CIPS*	.292	.618	.819	.980	1.00	.094	.096	.134	.344	.938
	$\widetilde{CP}$	.183	.540	.750	.960	1.00	.146	.081	.107	.266	.857
	$\widetilde{CZ}$	.289	.617	.817	.980	1.00	.092	.096	.134	.337	.937
20	CIPS	.418	.864	.973	.999	1.00	.144	.088	.160	.536	.999
	CIPS*	.442	.862	.973	.999	1.00	.098	.088	.160	.536	.999
	$\widetilde{CP}$	.319	.807	.953	.999	1.00	.162	.088	.121	.403	.993
	$\widetilde{CZ}$	.443	.861	.974	.999	1.00	.093	.090	.158	.528	.999
30	CIPS	.453	.882	.990	1.00	1.00	.133	.093	.187	.579	1.00
	CIPS*	.475	.883	.990	1.00	1.00	.088	.092	.187	.579	1.00
	$\widetilde{CP}$	.347	.831	.975	1.00	1.00	.184	.097	.163	.448	.998
	$\widetilde{CZ}$	.476	.882	.990	1.00	1.00	.089	.091	.186	.572	1.00
50	CIPS	.617	.984	1.000	1.00	1.00	.129	.107	.208	.751	1.00
	CIPS*	.634	.984	1.000	1.00	1.00	.090	.107	.208	.751	1.00
	$\widetilde{CP}$	.512	.957	.999	1.00	1.00	.186	.104	.178	.607	1.00
	$\widetilde{CZ}$	.635	.984	1.000	1.00	1.00	.091	.108	.209	.745	1.00
100	CIPS	.665	.998	1.00	1.00	1.00	.121	.108	.254	.880	1.00
	CIPS*	.685	.998	1.00	1.00	1.00	.072	.107	.254	.880	1.00
	$\widetilde{CP}$	.565	.995	1.00	1.00	1.00	.218	.129	.225	.776	1.00
	$\widetilde{CZ}$	.680	.998	1.00	1.00	1.00	.074	.110	.255	.875	1.00

Notes: [1] The underlying data is generated by $y_{it} = (1-\phi_i)\mu_i + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and $f_t \sim$ iidN(0,1), and $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}$, where $e_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$ and $\rho_i \sim iidU[-0.4,-0.2]$. Power, under the heterogeneous alternatives $\phi_i \sim iidU[0.85,0.95]$, is computed at the 5% nominal level based on the cross section augmented Dickey-Fuller regressions: $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^N y_{it}$, and the CADF(1) regressions: $\Delta y_{it} = a_i + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + d_{i1}\Delta\bar{y}_{t-1} + \delta_{i1}\Delta y_{i,t-1} + u_{it}$. The CIPS test statistic is given by $CIPS(N,T) = N^{-1}\sum_{i=1}^N t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regressions. Similarly, truncated CIPS statistic is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$, defined by (4.36). The $\widetilde{CP}$ and $\widetilde{CZ}$ tests are based on the statitics defined by (6.64) and (6.66), respectively, using their corresponding empirical critical values.

Table 8a : Size of Cross-Sectionally Augmented Panel Unit Root Tests
No Serial Correlation, Low and High Cross Section Dependence, Linear Trend Case[1]

		Low Cross Section Dependence					High Cross Section Dependence				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.022	.054	.054	.046	.040	.038	.059	.062	.059	.049
	CIPS*	.034	.053	.054	.046	.040	.045	.058	.062	.059	.049
	$\tilde{CP}$	.018	.051	.049	.050	.045	.032	.047	.059	.058	.046
	$\tilde{CZ}$	.035	.055	.052	.045	.041	.047	.058	.063	.060	.046
20	CIPS	.027	.048	.059	.047	.045	.032	.048	.061	.052	.041
	CIPS*	.034	.047	.059	.047	.045	.050	.048	.061	.051	.041
	$\tilde{CP}$	.028	.045	.061	.055	.047	.031	.044	.060	.051	.038
	$\tilde{CZ}$	.035	.047	.058	.046	.043	.050	.049	.060	.053	.041
30	CIPS	.030	.061	.053	.044	.047	.028	.060	.052	.049	.047
	CIPS*	.042	.062	.053	.043	.047	.039	.060	.053	.049	.047
	$\tilde{CP}$	.029	.054	.057	.044	.049	.025	.054	.047	.048	.050
	$\tilde{CZ}$	.042	.062	.052	.044	.046	.040	.063	.053	.051	.047
50	CIPS	.027	.045	.035	.051	.063	.020	.058	.043	.048	.058
	CIPS*	.034	.044	.035	.051	.063	.028	.058	.044	.048	.058
	$\tilde{CP}$	.021	.041	.037	.046	.050	.008	.048	.049	.046	.055
	$\tilde{CZ}$	.035	.043	.035	.051	.059	.028	.057	.044	.047	.057
100	CIPS	.019	.057	.050	.062	.043	.016	.047	.058	.058	.045
	CIPS*	.026	.057	.050	.062	.043	.029	.047	.058	.058	.045
	$\tilde{CP}$	.013	.046	.041	.066	.041	.009	.036	.052	.048	.041
	$\tilde{CZ}$	.026	.057	.050	.062	.043	.029	.045	.056	.060	.044

Notes: [1] The underlying data is generated by $y_{it} = \alpha_i + d_i(1-\phi_i)t + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, with $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[0,0.2]$ for low cross section dependence, and $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. α_i and $d_i \sim iidU[0,0.02]$, f_t ~ iidN(0,1), and $\varepsilon_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$. Size (under the null $\phi_i = 1$) is computed at the 5% nominal level based on a cross section augmented Dickey-Fuller regressions with linear trends: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$. The CIPS test statistic is defined as $CIPS(N,T) = N^{-1}\sum_{i=1}^{N} t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regression. Similarly, the truncated CIPS statistic (CIPS*) is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$, defined by (4.36). The $\tilde{CP}$ and $\tilde{CZ}$ tests are defined by (6.64) and (6.66) respectively using their corresponding empirical critical values.

Table 8b: Power of Cross-Sectionally Augmented Panel Unit Root Tests
No Serial Correlation, Low and High Cross Section Dependence, The Linear Trend Case[1]

		Low Cross Section Dependence					High Cross Section Dependence				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.029	.078	.085	.159	.696	.040	.070	.078	.183	.692
	CIPS*	.039	.078	.085	.159	.696	.054	.069	.079	.183	.692
	$\widetilde{CP}$	.025	.057	.076	.126	.537	.038	.067	.075	.143	.577
	$\widetilde{CZ}$	.040	.079	.083	.160	.697	.057	.070	.076	.187	.697
20	CIPS	.034	.071	.089	.274	.928	.042	.070	.093	.270	.935
	CIPS*	.044	.072	.090	.274	.928	.052	.070	.095	.270	.935
	$\widetilde{CP}$	.027	.051	.072	.218	.832	.038	.056	.086	.209	.866
	$\widetilde{CZ}$	.044	.073	.091	.274	.927	.053	.070	.092	.269	.937
30	CIPS	.030	.064	.105	.332	.991	.032	.081	.097	.332	.987
	CIPS*	.040	.062	.105	.332	.991	.038	.082	.097	.332	.987
	$\widetilde{CP}$	.028	.058	.081	.240	.940	.022	.059	.081	.224	.954
	$\widetilde{CZ}$	.041	.064	.103	.330	.991	.040	.082	.100	.332	.987
50	CIPS	.029	.088	.157	.496	.999	.025	.075	.133	.467	1.00
	CIPS*	.040	.090	.158	.496	.999	.034	.075	.133	.467	1.00
	$\widetilde{CP}$	.019	.062	.114	.359	.999	.017	.059	.099	.325	.996
	$\widetilde{CZ}$	.040	.087	.157	.495	.999	.034	.075	.132	.460	1.00
100	CIPS	.040	.086	.146	.619	1.00	.034	.074	.137	.621	1.00
	CIPS*	.050	.088	.146	.619	1.00	.043	.076	.137	.621	1.00
	$\widetilde{CP}$	.021	.061	.112	.467	1.00	.020	.058	.106	.478	1.00
	$\widetilde{CZ}$	.051	.087	.147	.618	1.00	.042	.076	.135	.617	1.00

Notes: [1] The underlying data is generated by $y_{it} = \alpha_i + d_i(1-\phi_i)t + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, with $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[0,0.2]$ for low cross section dependence, and $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. α_i and $d_i \sim iidU[0,0.02]$, $f_t \sim$ iidN(0,1), and $\varepsilon_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$. Power, under the heterogeneous alternative $\phi_i \sim iidU[0.85,0.95])$, is computed at the 5% nominal level based cross section augmented Dickey-Fuller regressions with linear trends: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^N y_{it}$. The CIPS test statistic is defined as $CIPS(N,T) = N^{-1}\sum_{i=1}^N t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regression. Similarly, the truncated CIPS statistic (CIPS*) is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$, defined by (4.36). The $\widetilde{CP}$ and $\widetilde{CZ}$ tests are defined by (6.64) and (6.66) respectively using their corresponding empirical critical values.

Table 9a : Size of Cross-Sectionally Augmented Panel Unit Root Tests
Positive Serial Correlation, High Cross Section Dependence, The Linear Trend Case

		ACDF(0)					ACDF(1)				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.010	.002	.001	.000	.000	.181	.054	.039	.058	.044
	CIPS*	.012	.002	.001	.000	.000	.080	.052	.038	.058	.044
	$\widetilde{CP}$	.012	.006	.001	.002	.000	.199	.075	.054	.061	.040
	$\widetilde{CZ}$	.012	.002	.001	.000	.000	.083	.052	.039	.062	.046
20	CIPS	.004	.001	.000	.000	.000	.198	.048	.038	.055	.050
	CIPS*	.005	.001	.000	.000	.000	.069	.047	.039	.055	.050
	$\widetilde{CP}$	.007	.003	.000	.000	.000	.207	.095	.056	.065	.053
	$\widetilde{CZ}$	.005	.001	.000	.000	.000	.068	.048	.038	.055	.049
30	CIPS	.003	.000	.000	.000	.000	.226	.049	.048	.049	.054
	CIPS*	.001	.000	.000	.000	.000	.053	.048	.048	.049	.054
	$\widetilde{CP}$	.004	.000	.000	.000	.000	.268	.084	.069	.059	.064
	$\widetilde{CZ}$	.001	.000	.000	.000	.000	.054	.050	.048	.050	.053
50	CIPS	.000	.000	.000	.000	.000	.261	.049	.052	.041	.039
	CIPS*	.000	.000	.000	.000	.000	.052	.049	.052	.041	.039
	$\widetilde{CP}$	.000	.000	.000	.000	.000	.313	.106	.080	.057	.050
	$\widetilde{CZ}$	.000	.000	.000	.000	.000	.056	.050	.053	.043	.041
100	CIPS	.000	.000	.000	.000	.000	.283	.048	.034	.047	.033
	CIPS*	.002	.000	.000	.000	.000	.038	.042	.034	.047	.033
	$\widetilde{CP}$	.000	.000	.000	.000	.000	.369	.105	.076	.060	.046
	$\widetilde{CZ}$	.002	.000	.000	.000	.000	.039	.046	.034	.049	.035

Notes: [1] The underlying data is generated by $y_{it} = \alpha_i + d_i(1-\phi_i)t + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and $f_t \sim$ iidN(0,1), and $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}$, where $e_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$ and $\rho_i \sim iidU[0.2,0.4]$. Size (under the null $\phi_i = 1$) is computed at the 5% nominal level based on the cross section augmneted Dickey-Fuller regressions: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^N y_{it}$, and the CADF(1) regressions: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + d_{i0}\Delta\bar{y}_t + c_i\bar{y}_{t-1} + \delta_{i1}\Delta y_{i,t-1} + d_{i1}\Delta\bar{y}_{t-1} + u_{it}$. The CIPS test statistic is defined by $CIPS(N,T) = N^{-1}\sum_{i=1}^N t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regressions. Similarly, truncated CIPS statistic is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$ as defined by (4.36). The $\widetilde{CP}$ and $\widetilde{CZ}$ tests are based on the statitics defined by (6.64) and (6.66), respectively, using their corresponding empirical critical values.

Table 9b : Power of Cross-Sectionally Augmented Panel Unit Root Tests
Positive Serial Correlation, High Cross Section Dependence, The Linear Trend Case

		CADF(0)					CADF(1)				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.005	.003	.002	.004	.010	.177	.064	.074	.131	.521
	CIPS*	.006	.003	.002	.004	.010	.078	.062	.074	.131	.521
	$\tilde{C}P$	.009	.006	.001	.006	.002	.197	.084	.080	.118	.417
	$\tilde{C}Z$	.006	.003	.002	.004	.010	.079	.061	.073	.132	.523
20	CIPS	.001	.001	.001	.001	.007	.205	.066	.069	.174	.816
	CIPS*	.002	.001	.001	.001	.007	.070	.066	.069	.174	.816
	$\tilde{C}P$	.002	.003	.001	.001	.004	.238	.105	.076	.157	.709
	$\tilde{C}Z$	.002	.001	.001	.001	.007	.071	.066	.069	.173	.814
30	CIPS	.004	.000	.000	.000	.003	.206	.056	.079	.201	.921
	CIPS*	.004	.000	.000	.000	.003	.050	.056	.078	.201	.921
	$\tilde{C}P$	.002	.000	.000	.000	.002	.245	.092	.096	.182	.817
	$\tilde{C}Z$	.005	.000	.000	.000	.003	.049	.056	.080	.201	.922
50	CIPS	.001	.000	.000	.000	.000	.255	.041	.073	.223	.988
	CIPS*	.001	.000	.000	.000	.000	.037	.039	.072	.223	.988
	$\tilde{C}P$	.001	.000	.000	.000	.000	.319	.103	.112	.198	.957
	$\tilde{C}Z$	.001	.000	.000	.000	.000	.040	.039	.072	.223	.990
100	CIPS	.000	.000	.000	.000	.001	.274	.044	.094	.290	1.00
	CIPS*	.000	.000	.000	.000	.001	.046	.041	.094	.290	1.00
	$\tilde{C}P$	.001	.000	.000	.000	.001	.361	.095	.108	.287	.998
	$\tilde{C}Z$	.000	.000	.000	.000	.001	.047	.045	.096	.289	1.00

Notes: [1] The underlying data is generated by $y_{it} = \alpha_i + d_i(1-\phi_i)t + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and f_t ~ iidN(0,1), and $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}$, where $e_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$ and $\rho_i \sim iidU[0.2,0.4]$. Power, under the heterogeneous alternative $\phi_i \sim iidU[0.85,0.95]$, is computed at the 5% nominal level based cross section augmented Dickey-Fuller regressions with linear trends: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + +d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$, and the CADF(1) regressions: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + d_{i0}\Delta\bar{y}_t + c_i\bar{y}_{t-1} + \delta_{i1}\Delta y_{i,t-1} + d_{i1}\Delta\bar{y}_{t-1} + u_{it}$. The CIPS test statistic is defined by $CIPS(N,T) = N^{-1}\sum_{i=1}^{N} t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regressions. Similarly, truncated CIPS statistic is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$ as defined by (4.36). The $\tilde{C}P$ and $\tilde{C}Z$ tests are based on the statitics defined by (6.64) and (6.66), respectively, using their corresponding empirical critical values.

Table 10a : Size of Cross-Sectionally Augmented Panel Unit Root Tests
Negative Serial Correlation, High Cross Section Dependence, The Linear Trend Case

		CADF(0)					CADF(1)				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.202	.616	.709	.805	.828	.192	.060	.039	.059	.040
	CIPS*	.250	.614	.709	.805	.828	.092	.059	.039	.059	.040
	$\tilde{C}P$	.166	.557	.690	.789	.822	.190	.066	.045	.062	.037
	$\tilde{C}Z$	.253	.610	.707	.804	.826	.091	.060	.039	.059	.042
20	CIPS	.337	.797	.904	.948	.964	.227	.053	.035	.057	.051
	CIPS*	.386	.795	.903	.948	.965	.104	.056	.035	.057	.051
	$\tilde{C}P$	.249	.753	.890	.955	.970	.232	.081	.056	.059	.055
	$\tilde{C}Z$	.388	.794	.905	.948	.963	.105	.051	.036	.059	.050
30	CIPS	.359	.895	.967	.983	.994	.266	.052	.059	.062	.059
	CIPS*	.433	.892	.967	.983	.994	.068	.051	.059	.062	.059
	$\tilde{C}P$	.284	.866	.971	.979	.995	.254	.066	.061	.052	.067
	$\tilde{C}Z$	.430	.894	.967	.981	.994	.067	.052	.061	.062	.061
50	CIPS	.420	.965	.987	.998	.999	.293	.056	.057	.049	.048
	CIPS*	.494	.965	.988	.998	.999	.084	.055	.057	.049	.048
	$\tilde{C}P$	.334	.952	.988	.998	.999	.309	.064	.056	.052	.044
	$\tilde{C}Z$	.491	.965	.986	.997	.999	.084	.056	.057	.050	.049
100	CIPS	.552	.995	1.000	1.000	1.000	.316	.058	.048	.053	.040
	CIPS*	.610	.995	1.000	1.000	1.000	.088	.059	.048	.053	.040
	$\tilde{C}P$	.471	.993	1.000	1.000	1.000	.368	.079	.063	.052	.037
	$\tilde{C}Z$	.607	.994	1.000	1.000	1.000	.087	.060	.047	.054	.040

Notes: [1] The underlying data is generated by $y_{it} = \alpha_i + d_i(1-\phi_i)t + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and f_t ~ iidN(0,1), and $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}$, where $e_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$ and $\rho_i \sim iidU[-0.4,-0.2]$. Size (under the null $\phi_i = 1$) is computed at the 5% nominal level based on the cross section augmneted Dickey-Fuller regressions: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^N y_{it}$, and the CADF(1) regressions: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + d_{i0}\Delta\bar{y}_t + c_i \bar{y}_{t-1} + \delta_{i1}\Delta y_{i,t-1} + d_{i1}\Delta\bar{y}_{t-1} + u_{it}$. The CIPS test statistic is defined by $CIPS(N,T) = N^{-1}\sum_{i=1}^N t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regressions. Similarly, truncated CIPS statistic is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$ as defined by (4.36). The $\tilde{C}P$ and $\tilde{C}Z$ tests are based on the statitics defined by (6.64) and (6.66), respectively, using their corresponding empirical critical values.

Table 10b : Power of Cross-Sectionally Augmented Panel Unit Root Tests
Negative Serial Correlation, High Cross Section Dependence, The Linear Trend Case

		CADF(0)					CADF(1)				
		T					T				
N	Test	10	20	30	50	100	10	20	30	50	100
10	CIPS	.217	.686	.836	.969	1.00	.198	.072	.083	.151	.640
	CIPS*	.277	.684	.836	.969	1.00	.096	.071	.083	.151	.640
	$\tilde{C}P$	.185	.631	.806	.949	1.00	.175	.064	.073	.127	.512
	$\tilde{C}Z$	.281	.686	.836	.969	1.00	.096	.070	.084	.152	.639
20	CIPS	.338	.882	.977	.998	1.00	.243	.071	.071	.208	.919
	CIPS*	.419	.884	.977	.998	1.00	.083	.070	.071	.208	.919
	$\tilde{C}P$	.274	.827	.966	.999	1.00	.237	.086	.079	.172	.832
	$\tilde{C}Z$	.417	.884	.978	.998	1.00	.085	.069	.071	.207	.916
30	CIPS	.391	.961	.997	1.00	1.00	.246	.064	.089	.256	.977
	CIPS*	.463	.963	.997	1.00	1.00	.084	.064	.088	.256	.977
	$\tilde{C}P$	.322	.939	.993	1.00	1.00	.247	.079	.080	.197	.908
	$\tilde{C}Z$	.459	.963	.997	1.00	1.00	.082	.065	.089	.255	.977
50	CIPS	.449	.981	.999	1.00	1.00	.290	.066	.087	.317	1.00
	CIPS*	.523	.981	.999	1.00	1.00	.078	.066	.087	.317	1.00
	$\tilde{C}P$	.344	.963	.999	1.00	1.00	.303	.089	.085	.255	.995
	$\tilde{C}Z$	.523	.980	.999	1.00	1.00	.082	.065	.088	.318	1.00
100	CIPS	.600	.999	1.00	1.00	1.00	.311	.061	.103	.407	1.00
	CIPS*	.654	.999	1.00	1.00	1.00	.088	.063	.103	.407	1.00
	$\tilde{C}P$	.528	.999	1.00	1.00	1.00	.356	.080	.101	.352	1.00
	$\tilde{C}Z$	.650	.999	1.00	1.00	1.00	.087	.062	.103	.407	1.00

Notes: [1] The underlying data is generated by $y_{it} = \alpha_i + d_i(1-\phi_i)t + \phi_i y_{i,t-1} + u_{it}, i = 1,2,..,N, t = -51,-50,...,T$, and $u_{it} = \gamma_i f_t + \varepsilon_{it}$ where we generate $\gamma_i \sim iidU[-1,3]$ for high cross section dependence. μ_i and $f_t \sim$ iidN(0,1), and $\varepsilon_{it} = \rho_i \varepsilon_{i,t-1} + e_{it}$, where $e_{it} \sim iidN(0,\sigma_i^2)$ with $\sigma_i^2 \sim iidU[0.5,1.5]$ and $\rho_i \sim iidU[-0.4,-0.2]$. Power, under the heterogeneous alternative $\phi_i \sim iidU[0.85,0.95]$, is computed at the 5% nominal level based cross section augmented Dickey-Fuller regressions with linear trends: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + c_i \bar{y}_{t-1} + +d_{i0}\Delta\bar{y}_t + u_{it}$, $i = 1,2,..,N, t = 1,2,...,T$ where $\bar{y}_t = N^{-1}\sum_{i=1}^N y_{it}$, and the CADF(1) regressions: $\Delta y_{it} = a_{i0} + a_{i1}t + b_i y_{i,t-1} + d_{i0}\Delta\bar{y}_t + c_i\bar{y}_{t-1} + \delta_{i1}\Delta y_{i,t-1} + d_{i1}\Delta\bar{y}_{t-1} + u_{it}$. The CIPS test statistic is defined by $CIPS(N,T) = N^{-1}\sum_{i=1}^N t_i(N,T)$, where $t_i(N,T)$ is the OLS t-ratio of b_i in the above CADF regressions. Similarly, truncated CIPS statistic is computed with $t_i(N,T)$ replaced by $t_i^*(N,T)$ as defined by (4.36). The $\tilde{C}P$ and $\tilde{C}Z$ tests are based on the statitics defined by (6.64) and (6.66), respectively, using their corresponding empirical critical values.