Incidental Trends and the Power of Panel Unit Root Tests*

Hyungsik Roger Moon
Department of Economics
University of Southern California

Benoit Perron
Département de sciences économiques, CIREQ, and CIRANO
Université de Montréal

Peter C.B. Phillips
Cowles Foundation, Yale University
University of Auckland & University of York

July 2006

Abstract

The asymptotic local power of various panel unit root tests are investigated. The (Gaussian) power envelope is obtained under homogeneous and heterogeneous alternatives. The envelope is compared with the asymptotic power functions for the pooled t- test, the Ploberger-Phillips (2002) test, and a point optimal test in neighborhoods of unity that are of order $n^{-1/4}T^{-1}$ and $n^{-1/2}T^{-1}$, depending on whether or not incidental trends are extracted from the panel data. In the latter case, when the alternative hypothesis is homogeneous across individuals, it is shown that the point optimal test and the Ploberger-Phillips test both achieve the power envelope and are uniformly most powerful, in contrast to point optimal unit root tests for time series. Some simulations examining the finite sample performance of the tests are reported.

JEL Classification: C22 & C23

Keywords and Phrases: Asymptotic power envelope, common point optimal test, heterogeneous alternatives, incidental trends, local to unity, power function, panel unit root test.

*We thank Peter Robinson, an associate editor, and a referee for their helpful comments and suggestions. Moon thanks the Faculty Development Awards of USC for research support. Email: moonr@usc.edu. Tel: 213-740-2108. Perron thanks SSHRCC, FQRSC, and MITACS for financial support. Phillips thanks the NSF for support under Grant No. SES 04-142254.

1 Introduction

In the past decade, much research has been conducted on panels in which both the cross-sectional and time dimensions are large. Testing for a unit root in such panels has been a major focus of this research. For example, Quah (1994), Levin *et al* (2002), Im *et al* (2003), Maddala and Wu (1999), and Choi (2001) have all proposed various tests. These studies derived the limit theory for the tests under the null hypothesis of a common panel unit root and power properties were investigated by simulation. On the other hand, Bowman (2002) studies the exact power of panel unit root tests against fixed alternative hypotheses. He characterizes the class of admissible tests for unit roots in panels and shows that the averaging-up tests of Im, Pesaran, and Shin (2003) and the test based on Fisher-type statistics in Maddala and Wu (1999) and Choi (2001) are not admissible.

The asymptotic local power properties of some panel unit root tests have become known recently. Breitung (2000) [1] and Moon and Perron (2004) independently find that *without incidental trends* in the panel, their panel unit root test, which is based on a t-ratio type statistic, has significant asymptotic local power in a neighborhood of unity that shrinks to the null at the rate of $n^{-1/2}T^{-1}$ (where n and T denote the size of the cross-section and time dimensions, respectively). However, *in the presence of incidental trends*, Moon and Perron (2004) show that their t-ratio type test statistic constructed from ordinary least squares (OLS) detrended data has no power (beyond size) in a $n^{-\kappa}T^{-1}$- neighborhood of unity with $\kappa > 1/6$. For a panel with incidental trends, Ploberger and Phillips (2002) proposed an optimal invariant panel unit root test that maximizes average local power. They show that the optimal invariant test has asymptotic local power in a neighborhood of unity that shrinks at the rate $n^{-1/4}T^{-1}$, thereby dominating the t-ratio test of Moon and Perron (2004) when there are incidental trends.

The present study makes three contributions. First, the local asymptotic power envelope of the panel unit root testing problem is derived under Gaussian assumptions for four scenarios: (i) with no fixed effects; (ii) with fixed effects that are parameterized by heterogeneous intercept terms (deemed incidental intercepts); (iii) with fixed effects that are parameterized by heterogeneous linear deterministic trends (deemed incidental trends); and (iv) with incidental intercepts but with a common trend. For cases (ii), (iii), and (iv) we restrict the class of tests to be invariant with respect to the incidental intercepts and trends. We show that in cases (i) and (ii), the power envelope is defined within $n^{-1/2}T^{-1}$-neighborhoods of unity and that it depends on the first two moments of the local-to-unity parameters. On the other hand, in case (iii), the power envelope is defined within $n^{-1/4}T^{-1}$- neighborhoods of unity and it depends on the first four moments of the local-to-unity parameters. Finally, in case (iv), we demon-

[1] We thank a referee for bringing this paper to our attention. Breitung (2000) derives his results under a homogeneous local alternative and with cross-sectional independence, while Moon and Perron (2004a) consider a more general model with heterogeneous local alternatives and cross-sectional dependence arising from the presence of common factors.

strate that the power envelope is defined within $n^{-1/2}T^{-1}$- neighborhoods of unity and that it is identical to that of cases (i) and (ii) [2].

Second, we derive the asymptotic local power of some existing panel unit root tests and compare these to the power envelope. For case (i), we investigate the t-ratio statistics studied by Quah (1994), Levin *et al* (2002), and Moon and Perron (2004). For case (ii), we discuss results from Moon and Perron (2005) on a modified t-ratio statistic that is asymptotically equivalent to the test proposed by Levin *et al*. For case (iii), we compare the optimal invariant test proposed by Ploberger and Phillips (2002), the LM test proposed by Moon and Phillips (2004), the unbiased test proposed by Breitung (2000), and a new t-test that is asymptotically equivalent to the Levin *et al*. (2002) test. First, we show that in all three cases the existing tests do not achieve maximal power. Next, when the alternative hypothesis is homogeneous across individuals, it is shown that some tests (the t-test in case (i) and the optimal invariant test of Ploberger and Phillips (2002) in cases (ii) and (iii)) do achieve the power envelope and are uniformly most powerful.

Third, we propose a simple point optimal invariant panel unit root test for each case. These tests are uniformly most powerful (UMP) when the alternative hypothesis is homogeneous, in contrast to point optimal unit root tests for time series (Elliot et al., 1996) where no UMP test exists.

The paper is organized as follows. Section 2 lays out the model, the hypotheses to test, and the assumptions maintained throughout the paper. Section 3 studies the model where there are no fixed effects (or where the fixed effects are known), develops the Gaussian power envelope, gives a point optimal test and performs some power comparisons. Sections 4 and 5 perform similar analyses for panel models with incidental intercepts and trends. Section 6 discusses various extensions and generalizations of our framework. Section 7 reports some simulations comparing the finite sample properties of the main tests studied in Sections 4 and 5. Section 8 concludes, and the Appendix contains the main technical derivations and proofs; the remaining proofs can be found in a companion paper, Moon, Perron, and Phillips (2006b).

2 Model

The observed panel z_{it} is assumed to be generated by the following component model

$$\begin{aligned} z_{it} &= b_i' g_t + y_{it} \\ y_{it} &= \rho_i y_{it-1} + u_{it},\ i = 1, ...;\ t = 0, 1..., \end{aligned} \tag{1}$$

where u_{it} is a mean zero error, $g_t = (1, t)'$, and $b_i = (b_{0i}, b_{1i})'$.

The focus of interest is the problem of testing for the presence of a common unit root in the panel against local alternatives when both n and T are large.

[2] This result can also be found in Breitung (1999), the working paper version of Breitung (2000).

For a local alternative specification, we assume that

$$\rho_i = 1 - \frac{\theta_i}{n^\kappa T} \text{ for some constant } \kappa > 0, \tag{2}$$

where θ_i is a sequence of iid random variables. [3] The main goal of the paper is to find efficient tests for the null hypothesis

$$\mathbb{H}_0 : \ \theta_i = 0 \text{ a.s. } (i.e., \rho_i = 1) \text{ for all } i, \tag{3}$$

against the alternative

$$\mathbb{H}_1 : \theta_i \neq 0 \ \ (i.e., \ \rho_i \neq 1) \ \text{ for some } i's. \tag{4}$$

A common special case of interest for the alternative hypothesis $\mathbb{H}_1$ is

$$\mathbb{H}_2 : \theta_i = \theta > 0 \text{ for all } i, \tag{5}$$

where the local-to-unity coefficients take on a common value $\theta > 0$ for all i. In this case, the series are homogeneously locally stationary, that is $\rho_i = \rho = 1 - \frac{\theta}{n^\kappa T} < 1$ for all i.

In (1) the nonstationary panel z_{it} has two different types of trends. The first component $b_i' g_t$ is a deterministic linear trend that is heterogeneous across individuals i. This component characterizes individual effects in the panel. The second component y_{it} is a stochastic trend or near unit-root process with ρ_i close to unity.

The following sections look at four different cases. In the first case, there are no fixed effects in the panel that have to be estimated, *i.e.* $b_i = (0,0)'$ (or alternatively b_i is known). The second case arises when the panel data z_{it} contain fixed effects that are parameterized by heterogeneous intercept terms b_{0i}, which are incidental parameters to be estimated. The third case arises when the panel contains fixed effects that are parameterized by heterogeneous linear deterministic trends, $b_{0i} + b_{1i}t$ where both sets of parameters b_{0i} and b_{1i} need to be estimated. A final case considers panels with heterogeneous intercepts and a common trend of the form $b_{0i} + b_1 t$.

In each case, under the assumptions that the error terms u_{it} are iid normal with zero mean and known variance σ_i^2 and that the initial conditions are $y_{i.t-1} = 0$ for all i, we construct point optimal test statistics. By deriving the limits of the test statistics, we establish the asymptotic power envelopes of the panel unit root testing problems. Then, we discuss the implementation of these procedures using feasible point optimal test statistics. To develop these, we relax some of the assumptions made in deriving the power envelopes.

We maintain the following assumptions in deriving the limits of the feasible point optimal tests and some other tests available in the literature.

Assumption 1 *For* $i = 1, 2...$ *and over* $t = 0, 1, ...,$ $u_{it} \sim iid\left(0, \sigma_i^2\right)$ *with* $\sup_i E\left[u_{it}^8\right] < \overline{M}$ *and* $\inf_i \sigma_i^2 \geq \underline{M} > 0$ *for some finite constants* $\overline{M}$ *and* $\underline{M}$.

[3] Notice that under the local altenative, ρ_i depends on n and T. Thus, the sequences of panel data z_{it} and y_{it} should be understood as triangular arrays.

Assumption 2 *The initial observations y_{i0} are iid with $E|y_{i0}|^8 < M$ for some constant M and are independent of u_{it}, $t \geq 1$ for all i.*

Assumption 3 $\frac{1}{T} + \frac{1}{n} + \frac{n}{T} \to 0.$

Before proceeding, we introduce the following notation. Define

$$\begin{aligned} z_t &= (z_{1t}, ..., z_{nt})', \ y_t = (y_{1t}, ..., y_{nt})', \ u_t = (u_{1t}, ..., u_{nt})', \\ Z &= (z_1, ..., z_T), \ Y = (y_1, ..., y_T), \ Y_{-1} = (y_0, y_1, ..., y_{T-1}), \ U = (u_1, ..., u_T), \end{aligned}$$

so the $(i,t)^{th}$ elements of Z, Y, Y_{-1}, and U are z_{it}, y_{it}, y_{it-1}, and u_{it}, respectively. Define the $T-$ vectors $G_0 = (1, ..., 1)'$, $G_1 = (1, 2, ..., T)'$, set $G = (G_0, G_1) = (g_1, ..., g_T)'$, and define

$$\begin{aligned} \beta_0 &= (b_{01}, ..., b_{0n})', \ \beta_1 = (b_{11}, ..., b_{1n})', \\ \beta &= (\beta_0, \beta_1) = (b_1, ..., b_n)'. \end{aligned}$$

Let $\underline{Z}_i$, $\underline{Y}_i$, $\underline{Y}_{-1,i}$, and $\underline{U}_i$ denote the transpose of the i^{th} row of Z, Y, Y_{-1}, and U, respectively, and write the model in matrix form as

$$\begin{aligned} Z &= \beta G' + Y, \\ Y &= \rho Y_{-1} + U, \end{aligned}$$

where $\rho = diag(\rho_1, ..., \rho_n)$. Define $\Sigma = diag(\sigma_1^2, ..., \sigma_n^2)$.

3 No Fixed Effects

This section investigates the model in which $b_i' g_t$ is observable or equivalently $g_t = 0$. In this case, the model becomes

$$\begin{aligned} Z &= Y, \\ Y &= \rho Y_{-1} + U. \end{aligned}$$

We consider local neighborhoods of unity that shrink at the rate of $\frac{1}{n^{1/2}T}$ and one-sided alternatives, as indicated in the following assumptions.

Assumption 4 $\kappa = 1/2$ *in (2).*

Assumption 5 *θ_i is a sequence of iid random variables whose support is a subset of a bounded interval $[0, M_\theta]$ for some $M_\theta \geq 0$.*

Let $\mu_{\theta,k} = E\left(\theta_i^k\right)$. The assumption of a bounded support for θ_i is made for convenience, and could be relaxed at the cost of stronger moment conditions. It is also convenient to assume that the θ_i are identically distributed, and this assumption could be relaxed as long as cross sectional averages of the moments $\frac{1}{n}\sum_{i=1}^n E\left(\theta_i^k\right)$ have limits such as $\mu_{\theta,k}$.

According to Assumption 5, $\theta_i \geq 0$ for all i, so that $\rho_i \leq 1$. In this case, the null hypothesis of a unit root in (3) is equivalent to $\mu_{\theta,1} = 0$ or $M_\theta = 0$ (*i.e.* $\theta_i = 0$ *a.s.* and the variance of θ, σ_θ^2, is 0), and the alternative hypothesis in (4) implies $\mu_{\theta,1} > 0$. Hence, in this section we set the hypotheses in terms of the first moment of θ_i as follows:

$$\mathbb{H}_0 : \mu_{\theta,1} = 0, \tag{6}$$

and

$$\mathbb{H}_1 : \mu_{\theta,1} > 0. \tag{7}$$

To test these hypotheses, Moon and Perron (2004) proposed t - ratio tests based on a modified pooled OLS estimator of the autoregressive coefficient and show that they have significant asymptotic local power in neighborhoods of unity shrinking at the rate $\frac{1}{\sqrt{n}T}$. This section first derives the (asymptotic) power envelope and shows that the power function of a feasible point optimal test for $\mathbb{H}_0$ achieves the envelope for the hypotheses above. We then compare the asymptotic local power of this point-optimal test with that of the Moon-Perron test.

3.1 Power Envelope

The power envelope is found by computing the upper bound of power of all point optimal tests for each local alternative. To proceed, we define

$$\rho_{c_i} = 1 - \frac{c_i}{n^{1/2}T},$$

where c_i is an *iid* sequence of random variables on $[0, M_c]$ for some $M_c > 0$. Denote by $\mu_{c,k}$ the k^{th} raw moment of c_i, *i.e.*, $\mu_{c,k} = E\left(c_i^k\right)$.

Define

$$\underset{((T+1)\times(T+1))}{\Delta_{c_i}} = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ -\rho_{c_i} & 1 & \ddots & \vdots & \vdots \\ 0 & \ddots & \ddots & 0 & 0 \\ \vdots & & -\rho_{c_i} & 1 & 0 \\ 0 & \dots & 0 & -\rho_{c_i} & 1 \end{bmatrix},$$

$\mathbb{C} = diag\left(c_1, ..., c_n\right)$, and $\Delta_{\mathbb{C}} = diag\left(\Delta_{c_1}, ..., \Delta_{c_n}\right)$.

When u_{it} are iid $N\left(0, \sigma_i^2\right)$ with σ_i^2 known and the initial conditions $y_{i,-1}$ are all zeros, so that $y_{i0} = u_{i0}$ for all i, the log-likelihood function is

$$L_{nT}\left(\mathbb{C}\right) = -\frac{1}{2}\left(vec\left(Y'\right)\right)' \Delta_{\mathbb{C}}' \left(\Sigma^{-1} \otimes I_{T+1}\right) \Delta_{\mathbb{C}} \left(vec\left(Y'\right)\right).$$

Denote by $L_{nT}(0)$ the log-likelihood function when $c_i = 0$ for all i.

Define

$$V_{nT}\left(\mathbb{C}\right) = -2L_{nT}\left(\mathbb{C}\right) + 2L_{nT}(0) - \frac{1}{2}\mu_{c,2}.$$

The statistic $V_{nT}(\mathbb{C})$ is the (Gaussian) likelihood ratio statistic of the null hypothesis $\rho_i = 1$ against an alternative hypothesis $\rho_i = \rho_{c_i}$ for $i = 1, ..., n$. According to the Neyman-Pearson lemma, rejecting the null hypothesis for small values of $V_{nT}(\mathbb{C})$ is the most powerful test of the null hypothesis $\mathbb{H}_0$ against the alternative hypothesis $\rho_i = \rho_{c_i}$. When the alternative hypothesis is given by $\mathbb{H}_1$, the test is a point optimal test (see, e.g., King (1988)). Let $\Psi_{nT}(\mathbb{C})$ be the test that rejects $\mathbb{H}_0$ for small values of $V_{nT}(\mathbb{C})$.

Theorem 6 *Assume that $b_i = 0$ for all i or $g_t = 0$ in* (1). *Suppose that Assumptions 1 – 5 hold. Then,*

$$V_{nT}(\mathbb{C}) \Rightarrow N\left(-E(c_i\theta_i), 2\mu_{c,2}\right).$$

The asymptotic critical values of the test $\Psi_{nT}(\mathbb{C})$ can be readily computed. In a notation we will use throughout the paper, let $\bar{z}_\alpha$ denote the $(1-\alpha)-$quantile of the standard normal distribution, *i.e.*, $P(Z \leq -\bar{z}_\alpha) = \alpha$, where $Z \sim N(0,1)$. Then, the size α asymptotic critical value $\psi(\mathbb{C}, \alpha)$ of the test $\Psi_{nT}(\mathbb{C})$ is $\psi(\mathbb{C}, \alpha) = -\sqrt{2\mu_{c,2}}\bar{z}_\alpha$, and its asymptotic local power is

$$\Phi\left(\frac{E(c_i\theta_i)}{\sqrt{2\mu_{c,2}}} - \bar{z}_\alpha\right), \tag{8}$$

where $\Phi(x)$ is the cumulative distribution function of Z.

Using (8), it is easy to find the power envelope, i.e., the values of c_i for which power is maximized. By the Cauchy-Schwarz inequality

$$\Phi\left(\frac{E(c_i\theta_i)}{\sqrt{2\mu_{c,2}}} - \bar{z}_\alpha\right) \leq \Phi\left(\sqrt{\frac{\mu_{\theta,2}}{2}} - \bar{z}_\alpha\right),$$

and the upper bound of $\Phi\left(\sqrt{\frac{\mu_{\theta,2}}{2}} - \bar{z}_\alpha\right)$ is achieved with $c_i = \theta_i$. Then, by the Neyman-Pearson lemma, $\Phi\left(\sqrt{\frac{\mu_{\theta,2}}{2}} - \bar{z}_\alpha\right)$ traces out a power envelope and we have the following theorem.

Theorem 7 *Assume that $b_i = 0$ for all i or $g_t = 0$ for all t in* (1). *Suppose that Assumptions 1 – 5 hold. Then, the power envelope for testing $\mathbb{H}_0$ in* (3) *against $\mathbb{H}_1$ in* (4) *is $\Phi\left(\sqrt{\frac{\mu_{\theta,2}}{2}} - \bar{z}_\alpha\right)$, where $\mu_{\theta,2} = E\left(\theta_i^2\right)$ and $\bar{z}_\alpha$ is the $(1-\alpha)-$quantile of the standard normal distribution.*

3.2 Implementation of the test

In order to implement a test that achieves the power envelope, estimates of the variances, σ_i^2, are necessary. The estimator we propose computes the variances under the null hypothesis. To simplify notation, let the first difference matrix

Δ_0 be simply denoted by Δ. Our estimator just takes the sample average of the squared first differences for each cross-section:

$$\hat{\sigma}_{1,iT}^2 = \frac{1}{T}\left(\Delta \underline{Z}_i\right)' \Delta \underline{Z}_i = \frac{1}{T}\left(y_{i0}^2 + \sum_{t=1}^{T}\left(\Delta y_{it}\right)^2\right).$$

Denote by $\hat{\Sigma}_1 = diag\left(\hat{\sigma}_{1,1T}^2, ..., \hat{\sigma}_{1,nT}^2\right)$ the estimated covariance matrix and by $\hat{L}_{nT}(\mathbb{C})$ and $\hat{L}_{nT}(0)$ the log-likelihood functions where the unknown Σ has been replaced by $\hat{\Sigma}_1$.

The feasible point-optimal statistic is:

$$\begin{aligned}
\hat{V}_{nT}(\mathbb{C}) &= -2\hat{L}_{nT}(\mathbb{C}) + 2\hat{L}_{nT}(0) - \frac{1}{2}\mu_{c,2} \\
&= \sum_{i=1}^{n}\frac{1}{\hat{\sigma}_{1,iT}^2}\left[z_{i0}^2 + \sum_{t=1}^{T}\left(\Delta_{c_i} z_{it}\right)^2\right] - \sum_{i=1}^{n}\frac{1}{\hat{\sigma}_{1,iT}^2}\left[z_{i0}^2 + \sum_{t=1}^{T}\left(\Delta z_{it}\right)^2\right] - \frac{1}{2}\mu_{c,2}.
\end{aligned}$$

The following theorem establishes asymptotic equivalence between the feasible and infeasible versions of the test:

Theorem 8 *Assume that $b_i = 0$ for all i or $g_t = 0$ for all t in* (1). *Suppose that Assumptions 1 – 5 hold. Then,* $\hat{V}_{nT}(\mathbb{C}) = V_{nT}(\mathbb{C}) + o_p(1)$.

3.3 Power Comparison

3.3.1 The $t-$ ratio Test

We start by investigating the $t-$ ratio test of Quah (1994), Levin *et al* (2002), and Moon and Perron (2004), which is based on the pooled OLS estimator[4]. For simplicity we assume that the error variances σ_i^2 are known. Let

$$\hat{\rho} = \frac{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\sum_{t=1}^{T} y_{it}y_{it-1}}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\sum_{t=1}^{T} y_{it-1}^2},$$

be the pooled OLS estimator with corresponding t statistic

$$t = \frac{\hat{\rho} - 1}{\sqrt{\frac{1}{\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\sum_{t=1}^{T} y_{it-1}^2}}}.$$

Under the conditions assumed above, we have $t \Rightarrow N\left(-\frac{\mu_{\theta,1}}{\sqrt{2}}, 1\right)$ (see Moon and Perron (2004)). The power of the t test with size α is then

$$\Phi\left(\frac{\mu_{\theta,1}}{\sqrt{2}} - \bar{z}_\alpha\right). \tag{9}$$

[4]When the error term u_{it} is serially correlated, one can use a modified version of the pooled OLS estimator. Details of this modification can be found in Moon and Perron (2004a). A more detailed discussion of the case where the errors are serially correlated can be found in section 6.4 below.

Remarks

(a) By the Cauchy-Schwarz inequality, it is straightforward to show that

$$\Phi\left(\frac{\mu_{\theta,1}}{\sqrt{2}}-\bar{z}_{\alpha}\right)\leq\Phi\left(\sqrt{\frac{\mu_{\theta,2}}{2}}-\bar{z}_{\alpha}\right). \tag{10}$$

In view of (10), the t ratio test achieves optimal power only when the alternative is homogeneous as in $\mathbb{H}_2$, that is when $\theta_i = \theta$ *a.s.*, so that $E(\theta_i) = \sqrt{E(\theta_i^2)}$. Otherwise, the power of the t ratio test is strictly sub-optimal. This implies that the $t-$ ratio test is the uniformly most powerful test for testing $\mathbb{H}_0$ against $\mathbb{H}_2$ but not against $\mathbb{H}_1$. The result is not surprising since the t ratio test is constructed based on the pooled OLS estimator and pooling is efficient only under the homogeneous alternative.

(b) Notice from (9) that the asymptotic local power of the t-test is determined by $\mu_{\theta,1}$, the mean of the local to unity parameters θ_i. In the given formulation, the local alternative is restricted to be one sided in Assumption 5. If we allow two-sided alternatives, this opens the possibility that $\mu_{\theta,1} = 0$ even under the alternative hypothesis, in which case the power of the pooled $t-$ test is equivalent to size.

(c) The pooled OLS estimator defined above can be interpreted as a GLS estimator since it gives weights that are inversely related to the variance of each observation. Moon and Perron (2004) do not make this adjustment and use a conventional OLS estimator. However, Levin et al. (2002) first correct for heteroskedasticity by dividing through by the estimated standard deviation before using pooled OLS on this transformed data. Their procedure can thus also be interpreted as a GLS estimator although it is commonly called pooled OLS. To avoid confusion with the previous literature, we will keep referring to estimators with weights that are the reciprocal of the standard deviation as pooled OLS estimators.

3.3.2 A Common-Point Optimal Test with $c_i = c$

As shown earlier, to achieve the power envelope, one needs to choose $c_i = \theta_i$ a.s. for $\Psi_{nT}(\mathbb{C})$. Denote this test $\Psi_{nT}(\Theta)$. Of course, the test $\Psi_{nT}(\Theta)$ is infeasible because it is not possible to identify the distribution of θ_i in the panel and generate a sequence from its distribution. Indeed, if the θ_i were known, there would be no need to test the null of a panel unit root.

One way of implementing the test $\Psi_{nT}(\mathbb{C})$ is to use randomly generated c_i's from some domain that is considered relevant. The variates c_i are independent of θ_i and the power of the test $\Psi_{nT}(\mathbb{C})$ is

$$\Phi\left(\frac{\mu_{c,1}\mu_{\theta,1}}{\sqrt{2\mu_{c,2}}}-\bar{z}_{\alpha}\right). \tag{11}$$

Since $\mu_{c,1} \leq \sqrt{\mu_{c,2}}$, the power (11) is bounded by

$$\Phi\left(\frac{\mu_{\theta,1}}{\sqrt{2}} - \bar{z}_{\alpha}\right), \tag{12}$$

which is achieved when we choose $c_i = c$, where c is any positive constant. We denote this test $\Psi_{nT}(c)$.

Remarks

(a) Not surprisingly, the power (12) of the test $\Psi_{nT}(c)$ is identical to that of the t - ratio test in the previous section. Of course, both tests are based on the homogeneous alternative hypothesis.

(b) Note that the power of the test $\Psi_{nT}(c)$ does not depend on c. The test is optimal against the special *homogeneous* alternative hypothesis $\mathbb{H}_2$ for any choice of c. This result is in contrast to the power of the point optimal test for unit root time series in Elliot *et al* (1996), where power does depend on the value of c. The reason is that the local alternative in the panel unit root case, $\rho_{c_i} = 1 - \frac{c_i}{n^{1/2}T}$, is closer to the null hypothesis than the alternative $\rho_{c_i} = 1 - \frac{c}{T}$ that applies in the case where there is only time series data. In effect, when we are this close to the null hypothesis with a homogeneous local alternative, it suffices to use any common local alternative in setting up the panel point optimal test.

4 Fixed Effects I: Incidental Intercepts Case

We extend the analysis in the previous section by allowing for fixed effects, *i.e.* $b_i' g_t = b_{0i}$, so that $g_t = 1$. In this case, the model has the matrix form $Z = \beta_0 G_0' + Y$.

4.1 Power Envelope

This section derives the power envelope of panel unit root tests for $\mathbb{H}_0$ that are invariant to the transformation $Z \to Z + \beta_0^* G_0'$ for arbitrary β_0^*. When u_{it} are iid $N\left(0, \sigma_i^2\right)$ with σ_i^2 known and the initial conditions $y_{i,-1}$ are zeros, *i.e.* $y_{i0} = u_{i0}$, the log-likelihood function is

$$L_{nT}(\mathbb{C}, \beta_0) = -\frac{1}{2}\left[vec\left(Z' - G_0\beta_0'\right)\right]' \Delta_{\mathbb{C}}' \left(\Sigma^{-1} \otimes I_{T+1}\right) \Delta_{\mathbb{C}} \left[vec\left(Z' - G_0\beta_0'\right)\right].$$

We denote by $L_{nT}(0, \beta_0)$ the log-likelihood function when $c_i = 0$ for all i.

A (Gaussian) point optimal invariant test statistic for this case can be constructed as follows (see, for example, Lehmann (1959), Dufour and King (1991), and Elliott *et al* (1996)):

$$V_{fe1,nT}(\mathbb{C}) = -2\left[\min_{\beta_0} L_{nT}(\mathbb{C}, \beta_0) - \min_{\beta_0} L_{nT}(0, \beta_0)\right] - \frac{1}{2}\mu_{c,2}.$$

For given c_i's, the point optimal invariant test, say $\Psi_{fe1,nT}(\mathbb{C})$, rejects the null hypothesis for small values of $V_{fe1,nT}(\mathbb{C})$.

Theorem 9 *Suppose Assumptions 1 – 5 hold and that $b_{1i} = 0$ or are known. Then, as $(n,T) \to \infty$*

(a) $V_{fe1,nT}(\mathbb{C}) \Rightarrow N\left(-E(c_i\theta_i), 2\mu_{c,2}\right).$

(b) *The power envelope for invariant testing of $\mathbb{H}_0$ in (3) against $\mathbb{H}_1$ in (4) is $\Phi\left(\sqrt{\frac{\mu_{\theta,2}}{2}} - \bar{z}_\alpha\right)$, where $\mu_{\theta,2} = E\left(\theta_i^2\right)$ and $\bar{z}_\alpha$ is the $(1-\alpha)-$ quantile of the standard normal distribution.*

Remarks

(a) As in the case of $\Psi_{nT}(c)$, we define the test $\Psi_{fe1,nT}(c)$ with a common constant point $c_i = c$. Then, the power of the test $\Psi_{fe1,nT}(c)$ is

$$\Phi\left(\frac{\mu_{\theta,1}}{\sqrt{2}} - \bar{z}_\alpha\right), \tag{13}$$

which is the same as for the $\Psi_{nT}(c)$ test in the previous section without fixed effects.

(b) Note that the asymptotic power envelope is the same as in the case without incidental intercepts, so estimation of intercepts does not affect maximal achievable power. The result is analogous to the time series case in Elliott et at (1996, p. 816).

(c) With incidental intercepts in the model, Levin et al. (2002) proposed a panel unit root test based on the pooled OLS estimator. Let $\tilde{z}_{it} = z_{it} - \frac{1}{T}\sum_{t=1}^T z_{it}$ and $\tilde{z}_{it-1} = z_{it-1} - \frac{1}{T}\sum_{t=1}^T z_{it-1}$. When the error variances σ_i^2 are known, the t - statistic proposed by Levin et al. is asymptotically equivalent to the following t - statistic

$$t^+ = \sqrt{\frac{30}{51}}\sqrt{\sum_{i=1}^n \frac{1}{\sigma_i^2}\sum_{t=1}^T \tilde{z}_{it-1}^2}\left(\hat{\rho}_{pool}^+ - 1\right),$$

where

$$\hat{\rho}_{pool}^+ = \left[\sum_{i=1}^n \frac{1}{\sigma_i^2}\sum_{t=1}^T \tilde{z}_{it-1}^2\right]^{-1}\left[\sum_{i=1}^n \frac{1}{\sigma_i^2}\sum_{t=1}^T \tilde{z}_{it-1}\tilde{z}_{it}\right] + \frac{3}{T}.$$

As shown by Moon and Perron (2005), the t^+ test also has significant asymptotic local power within $n^{-1/2}T^{-1}$ neighborhoods of unity, and its power is given by

$$\Phi\left(\frac{3}{2}\sqrt{\frac{5}{51}}\mu_{\theta,1} - \bar{z}_\alpha\right),$$

which is below that of the $\Psi_{fe1,nT}(c)$ test.

4.2 Implementation of the test

As in the case without fixed effects, we need to estimate the unknown quantities to make the point-optimal test feasible. In this case, the unknown quantities are the intercepts, b_{i0}, and variances, σ_i^2. The fixed effects will be estimated by generalized least squares (GLS) under the null hypothesis, or

$$\hat{b}_{0i}(0) = \left(\Delta G_0' \Delta G_0\right)^{-1} \Delta G_0' \Delta \underline{Z}_i.$$

where $\Delta G_0 = (1, 0. \ldots, 0)'$, and the resulting estimate is simply the first observation, z_{i0}. The variance estimator for each cross-section is then:

$$\hat{\sigma}^2_{2,iT} = \frac{1}{T}\left[\Delta \underline{Z}_i - \Delta G_0 \hat{b}_{0i}(0)\right]'\left[\Delta \underline{Z}_i - \Delta G_0 \hat{b}_{0i}(0)\right] = \frac{1}{T}\sum_{t=1}^{T}(\Delta z_{it})^2.$$

Define $\hat{\Sigma}_2 = diag\left(\hat{\sigma}^2_{2,1T}, ..., \hat{\sigma}^2_{2,nT}\right)$ as before, and let $\hat{L}_{nT}(\mathbb{C}, \beta_0)$ and $\hat{L}_{nT}(0, \beta_0)$ be the log-likelihood function values with the unknown Σ replaced by $\hat{\Sigma}_2$. The feasible statistic is then

$$\hat{V}_{fe1,nT}(\mathbb{C}) = -2\left[\min_{\beta_0} \hat{L}_{nT}(\mathbb{C}, \beta_0) - \min_{\beta_0} \hat{L}_{nT}(0, \beta_0)\right] - \frac{1}{2}\mu_{c,2},$$

leading to an asymptotically equivalent test.

Theorem 10 *Suppose that Assumptions 1 – 5 hold and that $b_{1i} = 0$ or are known. Then, $\hat{V}_{fe1,nT}(\mathbb{C}) = \hat{V}_{nT}(\mathbb{C}) + o_p(1)$.*

5 Fixed Effects II: Incidental Trends Case

This section considers the important practical case where heterogeneous linear trends need to be estimated. Set $g_t = (1, t)'$ and for this case, we consider local neighborhoods of unity that shrink at the slower rate of $\frac{1}{n^{1/4}T}$.

Assumption 11 $\kappa = 1/4$ *in (2).*

We relax Assumption 5 to allow for two-sided alternatives, so that the time series behavior of y_{it} can be either stationary or explosive under the alternative hypothesis.

Assumption 12 $\theta_i \sim$ *iid with mean μ_θ and variance σ^2_θ with a support that is a subset of a bounded interval $[-M_{l\theta}, M_{u\theta}]$, where $M_{l\theta}, M_{u\theta} \geq 0$.*

Under Assumption 12, we can re-express hypotheses (3) and (4) using the second raw moment of θ_i as follows:

$$\mathbb{H}_0 : \mu_{\theta,2} = 0, \tag{14}$$

and

$$\mathbb{H}_1 : \mu_{\theta,2} > 0. \tag{15}$$

The usual one-sided version where the series has a unit root or is stationary is the special case with $M_{l\theta} = 0$. We proceed as above by first deriving the power envelope, developing a feasible implementation of the resulting statistic, and then investigating the asymptotic local power of different panel unit root tests.

5.1 Power Envelope

This section derives the Gaussian power envelope of panel unit root tests for $\mathbb{H}_0$ that are invariant to the transformation $Z \to Z + \beta^* G'$ for arbitrary β^*. When u_{it} are iid $N\left(0, \sigma_i^2\right)$ with σ_i^2 known and the initial conditions $y_{i,-1}$ are zeros, that is, $y_{i0} = u_{i0}$, the log-likelihood function is

$$L_{nT}\left(\mathbb{C}, \beta\right) = -\frac{1}{2}\left[vec\left(Z' - G\beta'\right)\right]' \Delta_{\mathbb{C}}' \left(\Sigma^{-1} \otimes I_{T+1}\right) \Delta_{\mathbb{C}} \left[vec\left(Z' - G\beta'\right)\right].$$

We denote by $L_{nT}\left(0, \beta\right)$ the log-likelihood function when $c_i = 0$ for all i. As above, a (Gaussian) point optimal invariant test statistic can be constructed as

$$\begin{aligned} V_{fe2,nT}\left(\mathbb{C}\right) &= -2\left[\min_{\beta} L_{nT}\left(\mathbb{C}, \beta\right) - \min_{\beta} L_{nT}\left(0, \beta\right)\right] \\ &\quad + \left(\frac{1}{n^{1/4}}\sum_{i=1}^{n} c_i\right) + \left(\frac{1}{n^{1/2}}\sum_{i=1}^{n} c_i^2\right)\omega_{p2T} + \left(\frac{1}{n}\sum_{i=1}^{n} c_i^4\right)\omega_{p4T}, \end{aligned}$$

where

$$\begin{aligned} \omega_{p2T} &= -\frac{1}{T}\sum_{t=1}^{T}\frac{t-1}{T} + \frac{2}{T}\sum_{t=1}^{T}\frac{t}{T}\left(\frac{t-1}{T}\right) - \frac{1}{3}, \\ \omega_{p4T} &= \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{t-1}{T}\frac{s-1}{T}\min\left(\frac{t-1}{T}, \frac{s-1}{T}\right) - \frac{2}{3}\frac{1}{T}\sum_{t=1}^{T}\left(\frac{t-1}{T}\right)^2 + \frac{1}{9}. \end{aligned}$$

For given c_i's, the point optimal invariant test, say $\Psi_{fe2,nT}\left(\mathbb{C}\right)$, rejects the null hypothesis for small values of $V_{fe2,nT}\left(\mathbb{C}\right)$.

The asymptotic behavior of $V_{fe2,nT}\left(\mathbb{C}\right)$ is given in the following result.

Theorem 13 *Suppose that Assumptions 1 – 3, 11, and 12. Then,* $V_{fe2,nT}\left(\mathbb{C}\right) \Rightarrow N\left(-\frac{1}{90}E\left(c_i^2\theta_i^2\right), \frac{1}{45}E\left(c_i^4\right)\right)$.

>From Theorem 13, the size α asymptotic critical value is

$$\psi_{fe2}\left(\mathbb{C}, \alpha\right) = -\sqrt{\frac{\mu_{c,4}}{45}}\bar{z}_\alpha,$$

and the asymptotic power of the test is given by

$$\Phi\left(\frac{1}{6\sqrt{5}}\frac{E\left(c_i^2\theta_i^2\right)}{\sqrt{E\left(c_i^4\right)}} - \bar{z}_\alpha\right). \tag{16}$$

By the Cauchy-Schwarz inequality, we have

$$\Phi\left(\frac{1}{6\sqrt{5}}\frac{E\left(c_i^2\theta_i^2\right)}{\sqrt{E\left(c_i^4\right)}}-\bar{z}_\alpha\right)\leq\Phi\left(\frac{1}{6\sqrt{5}}\sqrt{\mu_{\theta,4}}-\bar{z}_\alpha\right). \tag{17}$$

Again, the maximal power, $\Phi\left(\frac{1}{6\sqrt{5}}\sqrt{\mu_{\theta,4}}-\bar{z}_\alpha\right)$, is achieved by choosing $c_i=\theta_i$. According to the Neyman-Pearson lemma, $\Phi\left(\frac{1}{6\sqrt{5}}\sqrt{\mu_{\theta,4}}-\bar{z}_\alpha\right)$ traces out the power envelope. Summarizing, we have the following theorem.

Theorem 14 *Suppose that the trends $b_i'g_t$ in* (1) *are unknown and need to be estimated and Assumptions 1 – 3, 11, and 12 hold. Then, the power envelope for testing the null hypothesis $\mathbb{H}_0$ in* (3) *against the alternative hypothesis $\mathbb{H}_1$ in* (4) *is $\Phi\left(\frac{1}{6\sqrt{5}}\sqrt{\mu_{\theta,4}}-\bar{z}_\alpha\right)$, where $\mu_{\theta,4}=E\left(\theta_i^4\right)$ and $\bar{z}_\alpha$ is the $(1-\alpha)-$ quantile of the standard normal distribution.*

Remarks

(a) An important finding of Theorem 14 is that in the panel unit root model with *incidental trends,* the POI test has significant asymptotic local power in local neighborhoods of unity that shrink at the rate $\frac{1}{n^{1/4}T}$. By contrast, in the panel unit root model either without fixed effects or only with incidental intercepts, the POI test has significant asymptotic power in local neighborhoods of unity that shrink at the faster rate $\frac{1}{n^{1/2}T}$. This difference in the neighborhood radius of non-negligible power is a manifestation of the difficulty in detecting unit roots in panels in the presence of heterogeneous trends, a problem that was originally discovered in Moon and Phillips (1999) and called the 'incidental trend' problem.

(b) The power envelope of invariant tests of $\mathbb{H}_0$ in (3) against $\mathbb{H}_1$ depends on the fourth moment of the local to unity parameters $\theta_i's$. This dependence suggests that panels with more dispersed autoregressive coefficients will tend to more easily reject the null hypothesis.

(c) When the alternative hypothesis is the homogeneous alternative $\mathbb{H}_2$ (i.e., $\theta_i=\theta$), the power envelope is

$$\Phi\left(\frac{1}{6\sqrt{5}}\theta^2-\bar{z}_\alpha\right). \tag{18}$$

and, in this case, the power envelope is attained by using $c_i=c$ for any choice of c.

(d) If the θ_i are symmetrically distributed about $\mu_{\theta,1}$ and κ_4 is the 4^{th} cumulant, then $\sqrt{\mu_{\theta,4}}=\mu_{\theta,1}^2\left\{1+\frac{6\sigma_\theta^2}{\mu_{\theta,1}^2}+\frac{3\sigma_\theta^4+\kappa_4}{\mu_{\theta,1}^4}\right\}^{1/2}$ and this will be close to $\mu_{\theta,1}^2$ when the ratios $\frac{6\sigma_\theta^2}{\mu_{\theta,1}^2}$ and $\frac{3\sigma_\theta^4+\kappa_4}{\mu_{\theta,1}^4}$ are both small. In such cases, it is clear from (17) that the test with $c_i=c$ for any choice of c will be close to the power envelope.

5.2 Implementation of the test

Again, the covariance matrix Σ is generally unknown and needs to be estimated. To do so, we use the GLS estimator of b_i under the null hypothesis,

$$\hat{b}_i(0) = (\Delta G' \Delta G)^{-1} \Delta G' \Delta \underline{Z}_i = \begin{pmatrix} z_{i0} \\ \frac{1}{T}\sum_{t=1}^{T} \Delta z_{it} \end{pmatrix},$$

where $\Delta G = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 1 \end{pmatrix}'$, and define the estimator of the error variance for cross-section i as:

$$\hat{\sigma}^2_{3,iT} = \frac{1}{T}\left[\Delta \underline{Z}_i - \Delta G \hat{b}_i(0)\right]' \left[\Delta \underline{Z}_i - \Delta G \hat{b}_i(0)\right] = \frac{1}{T}\sum_{t=1}^{T}\left(\Delta z_{it} - \frac{1}{T}\sum_{t=1}^{T}\Delta z_{it}\right)^2.$$

Denote $\hat{\Sigma}_3 = diag\left(\hat{\sigma}^2_{3,1T}, ..., \hat{\sigma}^2_{3,nT}\right)$. Let $\hat{L}_{nT}(\mathbb{C})$ and $\hat{L}_{nT}(0)$ be the log-likelihood function with the unknown Σ replaced with $\hat{\Sigma}_3$. The feasible statistic is then:

$$\begin{aligned} \hat{V}_{fe2,nT}(\mathbb{C}) &= -2\left[\min_{\beta} \hat{L}_{nT}(\mathbb{C},\beta) - \min_{\beta} \hat{L}_{nT}(0,\beta)\right] \\ &\quad + \left(\frac{1}{n^{1/4}}\sum_{i=1}^{n} c_i\right) + \left(\frac{1}{n^{1/2}}\sum_{i=1}^{n} c_i^2\right)\omega_{p2T} + \left(\frac{1}{n}\sum_{i=1}^{n} c_i^4\right)\omega_{p4T}. \end{aligned}$$

Again, we have an asymptotically equivalent test.

Theorem 15 *Suppose that Assumptions 1 – 5 hold. Then,* $\hat{V}_{fe2,nT}(\mathbb{C}) = V_{fe2,nT}(\mathbb{C}) + o_p(1)$.

5.3 Power Comparison

We compare the power of five tests, and for simplicity assume that the error variances σ_i^2 are known.

5.3.1 The Optimal Invariant Test of Ploberger and Phillips (2002)

We start with the optimal invariant panel unit root test proposed by Ploberger and Phillips (2002). To construct the test statistic, we first estimate the trend coefficients β by GLS $\bar{\beta} = (\Delta Z \Delta G)(\Delta G' \Delta G)^{-1}$, and detrend the panel data Z giving $E = Z - \bar{\beta}G'$. Define

$$V_{g,nT} = \sqrt{n}\left(\frac{1}{nT^2} tr\left(\Sigma^{-1/2} E E' \Sigma^{-1/2}\right) - \omega_{1T}\right), \tag{19}$$

where $\omega_{1T} = \frac{1}{T}\sum_{t=1}^{T}\frac{t}{T}\left(1 - \frac{t}{T}\right)$. In summation notation, we have

$$V_{g,nT} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{1}{T\sigma_i^2}\sum_{t=1}^{T}\bar{Z}^2_{it,T} - \omega_{1T}\right], \tag{20}$$

where

$$\bar{Z}_{it,T} = \frac{1}{\sqrt{T}} \left[(z_{it} - z_{i0}) - \frac{t}{T} (z_{iT} - z_{i0}) \right],$$

a maximal invariant statistic. In view of (19) and (20), we may interpret $V_{g,nT}$ as the standardized *information* of the GLS detrended panel data. The test $\Psi_{g,nT}$ proposed by Ploberger and Phillips (2002) rejects the null hypothesis $\mathbb{H}_0$ for small values of $V_{g,nT}$.

To investigate the asymptotic power of $\Psi_{g,nT}$, we first derive the asymptotic distribution of $V_{g,nT}$.

Lemma 1 *Suppose Assumptions 1 – 3, 11, and 12 hold. Then,* $V_{g,nT} \Rightarrow N\left(-\frac{1}{90}\mu_{\theta,2}, \frac{1}{45}\right)$.

Using Lemma 1, it is straightforward to find the size α asymptotic critical values $\phi_g(\alpha)$ of the test $\Psi_{g,nT}$. For $\bar{z}_\alpha$, the $(1-\alpha)-$ quantile of Z, the critical value is $\phi_g(\alpha) = -\frac{1}{3\sqrt{5}}\bar{z}_\alpha$, and the asymptotic local power is given by

$$\Phi\left(\frac{\mu_{\theta,2}}{6\sqrt{5}} - \bar{z}_\alpha\right), \tag{21}$$

showing that the test $\Psi_{g,nT}$ has significant asymptotic power against the local alternative $\mathbb{H}_1$.

Remarks

(a) Notice that the asymptotic power of the test $\Psi_{g,nT}$ is determined by the second moment of θ_i, $\mu_{\theta,2}$, so that it relies on the variance of θ_i as well as the mean of θ_i.

(b) According to Ploberger and Phillips (2002), the test $\Psi_{g,nT}$ is an optimal invariant test. Let $Q_{\theta,nT}(\theta)$ be the joint probability measure of the data for the given $\theta_i's$ and let v be the probability measure on the space of θ_i. Ploberger and Phillips (2002) show that the test $\Psi_{g,nT}$ is asymptotically the optimal invariant test that maximizes the average power $\int\left(\int \Psi_{g,nT} dQ_{\theta,nT}(\theta)\right) dv$, a quantity which also represents the power of $\Psi_{g,nT}$ against the Bayesian mixture $\int Q_{\theta,nT}(\theta)\, dv$.

(c) Comparing the power (21) of the test $\Psi_{g,nT}$ to the power envelope is straightforward. By the Cauchy-Schwarz inequality we have

$$\Phi\left(\frac{\mu_{\theta,2}}{6\sqrt{5}} - \bar{z}_\alpha\right) \leq \Phi\left(\frac{\sqrt{\mu_{\theta,4}}}{6\sqrt{5}} - \bar{z}_\alpha\right).$$

The test $\Psi_{g,nT}$ achieves the power envelope if the θ_i are constant *a.s.* That is, the power envelope is achieved against the special alternative hypothesis $\mathbb{H}_2$.

5.3.2 The LM Test in Moon and Phillips (2004)

The second test we investigate is the LM test proposed by Moon and Phillips (2004), which is constructed in a fashion similar to $V_{g,nT}$. The main difference is that Moon and Phillips (2004) use ordinary least squares (OLS) to detrend the data. To fix ideas, define $Q_G = I_T - P_G$ with $P_G = G\left(G'G\right)^{-1} G'$. Let $D_T = diag\left(1, T\right)$. and

$$V_{o,nT} = \sqrt{n}\left(\frac{1}{nT^2} tr\left(\Sigma^{-1/2} Z Q_G Z' \Sigma^{-1/2}\right) - \omega_{2T}\right),$$

where

$$\begin{aligned} \omega_{2T} &= \frac{1}{T}\sum_{t=1}^{T}\frac{t}{T} - \frac{1}{T^2}\sum_{t=1}^{T}\sum_{s=1}^{T}\frac{\min\left(t,s\right)}{T} h_T\left(t,s\right), \\ h_T\left(t,s\right) &= g_t' D_T^{-1}\left(\frac{1}{T}\sum_{p=1}^{T} D_T^{-1} g_p g_p' D_T^{-1}\right)^{-1} D_T^{-1} g_s. \end{aligned}$$

Define

$$\tilde{Z}_{it,T} = \frac{1}{\sqrt{T}}\left[z_{it} - g_t'\left(\sum_{t=1}^{T} g_t g_t'\right)^{-1}\left(\sum_{t=1}^{T} g_t' z_{it}\right)\right],$$

a scaled version of the OLS detrended panel. Then, we can write

$$V_{o,nT} = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{1}{T\sigma_i^2}\sum_{t=1}^{T}\tilde{Z}_{it,T}^2 - \omega_{2T}\right],$$

which can also be interpreted as the standardized *information* of the detrended panel data. The LM test, say $\Psi_{o,nT}$, of Moon and Phillips (2004) is to reject the null hypothesis $\mathbb{H}_0$ for small values of $V_{o,nT}\left(c\right)$.

The following theorem gives the limit distribution of $V_{o,nT}\left(c\right)$.

Lemma 2 *Suppose Assumptions 1 – 3, 11, and 12 hold. Then,* $V_{o,nT} \Rightarrow N\left(-\frac{1}{420}\mu_{\theta,2}, \frac{11}{6300}\right)$.

The size α asymptotic critical value of $\Psi_{o,nT}$, say $\phi_o\left(\alpha\right)$, is given by $\phi_o\left(\alpha\right) = -\sqrt{\frac{11}{6300}}\bar{z}_\alpha$, and the asymptotic power is $\Phi\left(\frac{\mu_{\theta,2}}{2\sqrt{77}} - \bar{z}_\alpha\right)$.

Remarks

(a) Similar to the test $\Psi_{g,nT}$, the test $\Psi_{o,nT}$ has significant asymptotic power against the local alternative $\mathbb{H}_1$, and its power depends on the second moment of θ_i, $\mu_{\theta,2}$.

(b) The asymptotic power of the optimal invariant test $\Psi_{g,nT}$ dominates that of the test $\Psi_{o,nT}$ because $\frac{\mu_\theta^2+\sigma_\theta^2}{2\sqrt{77}} < \frac{\mu_\theta^2+\sigma_\theta^2}{2\sqrt{45}}$. This is not so surprising since the optimal invariant test $\Psi_{g,nT}$ is based on GLS-detrended data, while the test $\Psi_{o,nT}$ is based on OLS-detrended data.

(c) As remarked earlier, the test $V_{fe2,nT}(c)$ will achieve power close to the power envelope when the ratios $\frac{6\sigma_\theta^2}{\mu_{\theta,1}^2}$ and $\frac{3\sigma_\theta^4+\kappa_4}{\mu_{\theta,1}^4}$ are both small.

5.3.3 The unbiased test of Breitung (2000)

Breitung (2000) has proposed an alternative test to the Levin et al. (2002) test that does not require bias adjustment. The idea is to transform the data as

$$\begin{aligned} y_{it}^* &= s_t \left[\Delta z_{it} - \frac{1}{T-t} \left(\Delta z_{it+1} + ... + \Delta z_{iT} \right) \right], \\ x_{it}^* &= z_{it-1} - z_{i0} - \frac{t-1}{T} \left(z_{iT} - z_{i0} \right), \end{aligned}$$

and note that y_{it}^* and x_{it}^* are orthogonal to each other. The pooled estimator proposed by Breitung is then

$$\rho^* = 1 + \frac{\sum_{i=1}^n \sum_{t=2}^{T-1} \sigma_i^{-2} y_{it}^* x_{it}^*}{\sum_{i=1}^n \sum_{t=2}^{T-1} \sigma_i^{-2} x_{it}^{*2}},$$

and is correctly centered and does not require bias adjustment in contrast to the Levin et al. (20002) pooled OLS estimator. Breitung suggests testing the panel unit root null hypothesis by looking at the corresponding t-statistic:

$$UB_{nT} = \frac{\sum_{i=1}^n \sum_{t=2}^{T-1} \sigma_i^{-2} y_{it}^* x_{it}^*}{\sqrt{\sum_{i=1}^n \sum_{t=2}^{T-1} \sigma_i^{-2} x_{it}^{*2}}}.$$

Under a homogeneous local alternative, Breitung claims (theorem 5, p. 172) that this statistic has power in a local neighborhood defined with $\kappa = 1/2$, and that the expectation in the asymptotic normal distribution under the alternative is

$$\theta\sqrt{6} \left[\lim_{T\to\infty} \left. \frac{\partial E\left(T^{-1} \sum x_{it}^* y_{it}^*\right)}{\partial\left(\theta/\sqrt{n}\right)} \right|_{\theta=0} \right].$$

In a separate paper (Moon, Perron, and Phillips (2006a)), we show analytically that the limit above is 0, and therefore that Breitung's test does not have power in a neighborhood that shrinks at the faster rate $\frac{1}{n^{1/2}T}$ towards the null. Instead, we show that the necessary rate is the same slower $\frac{1}{n^{1/4}T}$ rate that applies to the other tests with incidental trends. Indeed, we show that under the assumptions in this section, the UB statistic has the following distribution.

Lemma 3 *Suppose Assumptions 1 – 3, 11, and 12 hold. Then,* $UB_{nT} \Rightarrow N\left(\frac{\mu_{\theta,2}}{6\sqrt{6}}, 1\right)$.

Remark The above lemma shows that the asymptotic power of Breitung's test is $\Phi\left(\frac{\mu_{\theta,2}}{6\sqrt{6}} - \bar{z}_\alpha\right)$, which is obviously below the power envelope.

5.3.4 A Common-Point Optimal Invariant Test

The test $V_{fe2,nT}(\Theta)$ that achieves the power envelope is infeasible. If we use randomly generated $c_i's$ that are independent of θ_i and the panel data z_{it} in constructing the test, according to (16), the power of the test $V_{fe2,nT}(\mathbb{C})$ is

$$\Phi\left(\frac{1}{6\sqrt{5}}\frac{\mu_{c,2}\mu_{\theta,2}}{\sqrt{\mu_{c,4}}} - \bar{z}_\alpha\right). \tag{22}$$

Since $\mu_{c,2} \leq \sqrt{\mu_{c,4}}$, the power (22) is bounded by

$$\Phi\left(\frac{1}{6\sqrt{5}}\mu_{\theta,2} - \bar{z}_\alpha\right), \tag{23}$$

which is achieved when we choose $c_i = c$ for $V_{fe2,nT}(\mathbb{C})$, where c is any positive constant. We denote this test $V_{fe2,nT}(c)$.

Remarks

(a) The power (23) of the test $V_{fe2,nT}(c)$ is identical to that of the Ploberger-Phillips optimal invariant test $V_{g,nT}$.

(b) The power of the test $V_{fe2,nT}(c)$ also does not depend on c. It is optimal against the special homogeneous alternative hypothesis $\mathbb{H}_2$ for any choice of c.

5.3.5 A t-test

In a manner similar to Moon and Perron (2005), we can define statistics that are asymptotically equivalent to the Levin et al. (2002) statistic based on the pooled OLS estimator for this case. When there are incidental trends, the Levin et al. statistic is asymptotically equivalent to the following t - statistic

$$t^+ = \sqrt{\frac{112}{193}}\sqrt{tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}_{-1}'\Sigma^{-1/2}\right)}\left(\hat{\rho}_{pool}^+ - 1\right),$$

where the bias-corrected pooled OLS estimator is

$$\hat{\rho}_{pool}^+ = \left[tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}_{-1}'\Sigma^{-1/2}\right)\right]^{-1}\left[tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}'\Sigma^{-1/2}\right)\right] + \frac{7.5}{T}.$$

On the other hand, Moon and Perron (2004) consider the following t-ratio test based on a different bias-corrected pooled estimator

$$t^\# = \sqrt{tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}_{-1}'\Sigma^{-1/2}\right)}\left(\hat{\rho}_{pool}^\# - 1\right),$$

where

$$\hat{\rho}_{pool}^{\#} = \left[tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}_{-1}'\Sigma^{-1/2}\right)\right]^{-1}\left[tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}'\Sigma^{-1/2}\right) + \frac{nT}{2}\right].$$

By definition,

$$\hat{\rho}_{pool}^{+} - \hat{\rho}_{pool}^{\#} = \frac{15}{2T}\left(\frac{tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}_{-1}'\Sigma^{-1/2}\right) - \frac{nT^2}{15}}{tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}_{-1}'\Sigma^{-1/2}\right)}\right),$$

and

$$t^{+} = \sqrt{\frac{112}{193}}t^{\#} + \frac{15}{2}\sqrt{\frac{112}{193}}\frac{\sqrt{n}\left\{\frac{1}{nT^2}tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}_{-1}'\Sigma^{-1/2}\right) - \frac{1}{15}\right\}}{\frac{1}{nT^2}tr\left(\Sigma^{-1/2}\tilde{Z}_{-1}\tilde{Z}_{-1}'\Sigma^{-1/2}\right)^{1/2}}.$$

Using Theorem 4 of Moon and Perron (2004) and Lemma 2, it is possible to show the following.

Lemma 4 *Suppose Assumptions 1 – 3, 11, and 12 hold. Then,* $t^{+} \Rightarrow N\left(-\frac{15\sqrt{15}}{2}\sqrt{\frac{112}{193}}\frac{\mu_{\theta,2}}{420}, 1\right)$.

6 Discussion

6.1 Case with Incidental Intercepts but a Common Trend

This section investigates the panel model for z_{it} in (1) where there are incidental intercepts but a *common* trend, viz.,

$$\begin{aligned} z_{it} &= b_{0i} + b_1 t + y_{it}, \\ y_{it} &= \rho_i y_{it-1} + u_{it},\ i = 1, ...;\ t = 0, 1.... \end{aligned}$$

This model is relevant because there is a tradition of imposing such a common trend in empirical work in microeconometrics. In addition, the analysis of asymptotic local power for this model provides further evidence that it is the presence of incidental trends, $b_{1i}t$, rather than incidental intercepts b_{0i} that makes the detection of unit roots more challenging.

To proceed, we make the same assumptions as in Sections 2, 3, and 4, so that

$$\rho_i = 1 - \frac{\theta_i}{n^{1/2}T}.$$

Let $l_n = (1, ..., 1)'$, $n-$ vector of ones. Using notation defined in Section 2, we write the model as

$$\begin{aligned} Z &= \beta_0 G_0' + b_1 l_n G_1' + Y, \\ Y &= \rho Y_{-1} + U. \end{aligned}$$

In the following theorem we show that the power envelope of panel unit root tests for $\mathbb{H}_0$ that are invariant to the transformation $Z \to Z + \beta_0^* G_0' + b_1^* l_n G_1'$ for arbitrary β_0^* and b_1^* is the same as the one we found in Sections 3 and 4.

When u_{it} are iid $N\left(0, \sigma_i^2\right)$ with σ_i^2 known and the initial conditions $y_{i,-1}$ are zeros, that is, $y_{i0} = u_{i0}$, the log-likelihood function is

$$L_{nT}\left(\mathbb{C}, \beta_0, b_1\right) = -\frac{1}{2}\left[vec\left(Z' - G_0\beta_0' - G_1 l_n' b_1\right)\right]' \Delta_{\mathbb{C}}'\left(\Sigma^{-1} \otimes I_{T+1}\right)\Delta_{\mathbb{C}}\left[vec\left(Z' - G_0\beta_0' - G_1 l_n' b_1\right)\right].$$

As before, a (Gaussian) point optimal invariant test statistic for this case can be constructed as follows:

$$V_{fe3,nT}\left(\mathbb{C}\right) = -2\left[\min_{\beta_0, b_1} L_{nT}\left(\mathbb{C}, \beta_0, b_1\right) - \min_{\beta_0, b_1} L_{nT}\left(0, \beta_0, b_1\right)\right] - \frac{1}{2}\mu_{c,2}.$$

For given c_i's, the point optimal invariant test, say $\Psi_{fe3,nT}\left(\mathbb{C}\right)$, rejects the null hypothesis for small values of $V_{fe3,nT}\left(\mathbb{C}\right)$.

Theorem 16 *Suppose Assumptions 1 – 5 hold. Then,*

$$V_{fe3,nT}\left(\mathbb{C}\right) = V_{fe1,nT}\left(\mathbb{C}\right) + o_p\left(1\right).$$

6.2 Initial conditions

In the derivations above, we have assumed that all series in the panel were initialized at the origin ($y_{i,-1} = 0$). It is well-known in the time series case that the initial condition can play an important role in the performance of unit root tests (Evans and Savin (1984), Phillips (1987), Elliott (1999) and Müller and Elliott (2004)). A common assumption made in the time series context is that the initial condition is drawn from the unconditional distribution under the stationary alternative, *i.e.* $y_0 \sim N\left(0, \frac{1}{1-\rho^2}\right)$. In the local to unity case, $\rho = 1 - \frac{\theta}{T}$, this formulation of the initial condition gives $y_0 = O_p\left(\sqrt{T}\right)$, which has some appeal because the order of magnitude of the initial condition is the same as that of the sample data y_t.

This commonly used set up for the time series case does not extend naturally to the panel model. Indeed, under the assumption $y_{i,-1} \sim N\left(0, \frac{1}{1-\rho_i^2}\right)$, and with local alternatives $\rho_i = 1 - \frac{\theta_i}{n^{1/2}T}$ or $\rho_i = 1 - \frac{\theta_i}{n^{1/4}T}$ (depending on whether trends are present or not), we have $y_{i,-1} = O_p\left(n^{1/4}\sqrt{T}\right)$ or $y_{i,-1} = O_p\left(n^{1/8}\sqrt{T}\right)$, respectively, in which case $y_{i,-1}$ diverges with n. The sample data $y_{i,t}$ for this series is then dominated by the initial condition $y_{i,-1}$. There is, of course, no reason in empirical panels why the order of magnitude of the initial condition for an individual series should depend on the total number of individuals (n) observed in the panel and such a formulation would be hard to justify. In this sense, the situation is quite different from the time series case, where there are good reasons for expecting initial observations for nonstationary or nearly nonstationary time series to have stochastic orders comparable to those of the sample. Moreover, under the initialization $y_{i,-1} \sim N\left(0, \frac{1}{1-\rho_i^2}\right)$, the likelihood ratio statistic diverges to negative infinity under the local alternative, as we show below.

To illustrate, consider the case with no fixed effect and with $u_{it} \sim iid\, N(0,1)$ across i over t. Here we assume that $\rho_i = 1 - \frac{\theta_i}{n^{1/2}T}$, as in Sections 3 and 4. Assume that if $\theta_i \neq 0$, $y_{i,-1}$ are $iid\ N\left(0, \frac{1}{1-\rho_i^2}\right)$ and independent of u_{jt}, and if $\theta_i = 0$, $y_{i,-1}$ are iid $N(0,1)$ and independent of u_{jt}. Denote deviations from the initial condition as

$$\begin{aligned}
\tilde{y}_{it} &= y_{it} - y_{i,-1} \\
&= u_{it} + \rho_i u_{it-1} + \rho_i^2 u_{it-2} + ... + \rho_i^t u_{i0} + \left(\rho_i^{t+1} - 1\right) y_{i,-1}.
\end{aligned}$$

All quantities based on $\tilde{y}_{it}$ will behave as in the case of a fixed initial condition. Define the notation

$$\begin{aligned}
\Delta_{\theta_i} \underline{\mathrm{Y}}_i &= \left(\left(1-\rho_i^2\right)^{1/2} y_{i,-1}, \Delta_{\theta_i} y_{i0}, ..., \Delta_{\theta_i} y_{iT}\right)' \\
&= \left(\left(1-\rho_i^2\right)^{1/2} y_{i,-1}, \Delta y_{i0} - \frac{\theta_i}{n^{1/2}T} y_{i,-1}, ..., \Delta y_{iT} - \frac{\theta_i}{n^{1/2}T} y_{iT-1}\right)', \\
\Delta_0 \underline{\mathrm{Y}}_i &= \left(y_{i,-1}, \Delta y_{i0}, ..., \Delta y_{iT}\right)'.
\end{aligned}$$

Then, the likelihood ratio is

$$\begin{aligned}
& -\frac{1}{2} L_{nT,A} + \frac{1}{2} L_{nT,0} \\
= & \sum_{i=1}^n \left(\Delta_{\theta_i} \underline{\mathrm{Y}}_i\right)' \Delta_{\theta_i} \underline{\mathrm{Y}}_i - \sum_{i=1}^n \left(\Delta_0 \underline{\mathrm{Y}}_i\right)' \Delta_0 \underline{\mathrm{Y}}_i \\
= & \sum_{i=1}^n \left[\left(1-\rho_i^2\right) y_{i0}^2 + \sum_{t=0}^T \left(\Delta \tilde{y}_{it} - \frac{\theta_i}{n^{1/2}T} \tilde{y}_{it-1} - \frac{\theta_i}{n^{1/2}T} y_{i,-1}\right)^2 - y_{i,-1}^2 - \sum_{t=0}^T \Delta \tilde{y}_{it}^2\right] \\
= & \sum_{i=1}^n \left(\frac{\theta_i^2}{nT} - \rho_i^2\right) y_{i,-1}^2 - 2\frac{1}{n^{1/2}T} \sum_{i=1}^n \theta_i y_{i,-1} \sum_{t=0}^T \left(\Delta \tilde{y}_{it}\right) + 2\frac{1}{nT^2} \sum_{i=1}^n \theta_i^2 y_{i,-1} \sum_{t=0}^T \tilde{y}_{it-1} \\
& -2\frac{1}{n^{1/2}T} \sum_{i=1}^n \theta_i \sum_{t=0}^T \Delta \tilde{y}_{it} \left(\tilde{y}_{it-1}\right) + \frac{1}{nT^2} \sum_{i=1}^n \theta_i^2 \sum_{t=0}^T \tilde{y}_{it-1}^2.
\end{aligned}$$

The last two terms behave as in the case of fixed initial conditions in the limit since they are deviations from the initial condition. As for the other three terms, we concentrate on the homogeneous case, $\theta_i = \theta$ for simplicity. We can show that the first term is

$$\sum_{i=1}^n \left(\frac{\theta^2}{nT} - \rho_i^2\right) y_{i,-1}^2 = O_p\left(n^{3/2}T\right),$$

while the second term is

$$\frac{1}{n^{1/2}T} \sum_{i=1}^n \theta y_{i,-1} \left(y_{iT} - y_{i,-1}\right) = O_p\left(n^{1/4}\right),$$

and the third term is

$$\frac{1}{nT^2}\sum_{i=1}^{n}\theta^2 y_{i,-1}\sum_{t=0}^{T}\tilde{y}_{it-1} = O_p\left(\frac{1}{n^{1/4}}\right).$$

Thus, the behavior of the likelihood ratio statistic is dominated by the first term. This first term has a negative mean and thus the likelihood ratio statistic diverges to negative infinity under the local alternative.

This example makes it clear that mechanical extensions of time series formulations that are commonly used for initial conditions can lead to quite unrealistic and unjustifiable features in a panel context. It is therefore necessary to consider initializations that are sensible for panel models, while at the same time having realistic time series properties. Given the more limited focus of the present study, we will not pursue this discussion of initial conditions further here but retain the (simplistic) assumption of zero initial conditions. Clearly, it is an important matter for future research to extend the theory and relax this condition.

6.3 Cross-sectional dependence

As with most of the early panel unit root tests that have been proposed in the literature, the above analysis supposes that the observational units that make up the panel are independent of each other. This assumption is not realistic in many applications, such as the analysis of cross-country macroeconomic series, where individual series are likely to be affected by common, worldwide shocks. Accordingly, more recent panel tests such as those in Bai and Ng (2004) , Moon and Perron (2004), Phillips and Sul (2003) , Chang (2002), and Pesaran (2005) allow for the presence of cross-sectional dependence among the units, typically through the presence of dynamic factors.

In order to handle such cross-sectional dependence, we can combine the defactoring method of Bai and Ng (2004) , Moon and Perron (2004) or Phillips and Sul (2003) to the analysis of this paper. The idea is to apply the optimal tests developed here to the data after the common factors have been extracted. Once the extraction process has been completed, there is, of course, no claim of optimality in the resulting tests, and we do not prove here that this approach has any optimality property. However, intuition suggests that this approach should perform well in practice, and simulation evidence provided in Moon and Perron (2006) confirms this.

For illustration, we will use the model of Moon and Perron (2004) . Thus, the assumption is that the disturbance in (1) has a factor structure

$$u_{it} = \gamma_i^{'} f_t + e_{it}. \tag{24}$$

The proposed procedure is as follows:

1. Estimate the deterministic components (b_i) by GLS to obtain $\hat{y}_{it}$;

2. Use the pooled OLS estimate to compute residuals $\hat{u}_{it}$;

3. Use principal components on the covariance matrix of these estimated residuals to estimate the common factor(s), $\hat{f}_t$ and factor loadings, $\hat{\gamma}_i$. Post-multiply the data matrix Z by $Q_{\hat{\gamma}} = I - \hat{\gamma}\left(\hat{\gamma}'\hat{\gamma}\right)^{-1}\hat{\gamma}'$ so that $ZQ_{\hat{\gamma}}$ is no longer affected by the common factors;

4. Use the common point optimal test proposed earlier in the paper on $ZQ_{\hat{\gamma}}$.

6.4 Serial correlation

Serial correlation can be accounted for in the construction of the test statistics by replacing variances with long-run variances, $\omega_i^2 = \sum_{j=-\infty}^{\infty} \gamma_{ij}$, where $\gamma_{ij} = E\left(u_{i,t}u_{i,t-j}\right)$. Since serial correlation is not accommodated in the above derivation of the power envelope, this procedure will not in general be optimal, but should result in tests with correct asymptotic size under quite general short memory autocorrelation (as in Elliott *et al.* (1996)). Standard kernel-based estimators of the long-run variance as in Andrews (1991) and Newey and West (1994) can be used to estimate the long-run variances. The development of optimal procedures that accommodate serial correlation is of interest but beyond the scope of the present contribution.

7 Simulations

This section reports the results of a small Monte Carlo experiment designed to assess and compare the finite-sample properties of the tests presented earlier in the paper. For this purpose, we use the following data generating process:

$$\begin{aligned}
z_{it} &= b_{0i} + b_{1i}t + y_{it}, \\
y_{it} &= \rho_i y_{it-1} + u_{it}, \\
y_{i,-1} &= 0, \quad u_{it} \sim iid\ N\left(0, \sigma_i^2\right) \\
\sigma_i^2 &\sim U\left[0.5, 1.5\right].
\end{aligned}$$

We consider both the incidental intercepts case ($b_{1i} = 0$) of section 4 and the incidental trends case ($b_{1i} \neq 0$) of section 5. In each case, the heterogeneous intercepts and/or trends are $iidN\left(0,1\right)$. We assume that the error term is independent in both the time and cross-sectional dimensions with a Gaussian distribution and heteroskedastic variances. Initial conditions are set to zero and, as discussed earlier, this is a limitation of the experiments and may lead to more favorable results for many of the tests than under random initializations where there is some dependence on the localization parameters.

We focus the study on three main questions. The first is the sensitivity of the point-optimal invariant test to the choice of c_i. The second is how far the feasible and infeasible point-optimal tests are from the theoretical power envelope in finite samples. Finally, we look at the impact of the distribution of the local-to-unity parameters under the alternative hypothesis.

We consider the following nine distributions for the local-to-unity parameters: $\theta_i = 0 \quad \forall i$ for size, and for local power, (1) $\theta_i \sim iidU[0,2]$, (2) $\theta_i \sim iidU[0,4]$, (3) $\theta_i \sim iidU[0,8]$, (4) $\theta_i \sim iid\chi^2(1)$, (5) $\theta_i \sim iid\chi^2(2)$, (6) $\theta_i \sim iid\chi^2(4)$, (7) $\theta_i = 1 \; \forall i$, and $\theta_i = 2 \; \forall i$. These distributions enable us to examine performance of the tests as the mass of the distribution of the localizing parameters moves away from the null hypothesis. We can also look at the effect of homogeneous versus heterogeneous alternatives (cases (1) and (4) versus (7), and cases (2) and (5) versus (8)) together with the role of the higher-order moments of the distribution. For instance, case (1) has the same mean as case (4) but smaller higher-order moments. The same situation arises for cases (2) and (5), and cases (3) and (6). Note that the alternatives with χ^2 distributions do not fit our asymptotic framework since they have unbounded support.

We take three values for each of n (10, 25, and 100) and T (50, 100, and 250). All tests are conducted at the 5% significance level, and the number of replications is set at 10,000.

Table 1 presents the results for the incidental intercepts case. The tests we consider are the infeasible point-optimal test with $c_i = \theta_i$ (the finite-sample analog of the power envelope which uses the local-to-unity parameters generated in the simulation), our common point-optimal (CPO) invariant test for three values of c (1, 2, and 0.5), the t-ratio type test as in Moon and Perron (2005), and the $t - bar$ statistic of Im, Pesaran, and Shin for which no analytical power result is available [5]. The first panel of the table provides the size and power predicted by the asymptotic theory in section 4 using the moments of θ_i and c_i. The other panels in the table report the size and size-adjusted power of the tests for the various combinations of n and T. Thus, if asymptotic theory were a reliable guide to finite-sample behavior, subsequent panels in the table would mirror the first panel.

The main outcomes from the first panel of the table can be summarized as follows:

- The power envelope is higher for the χ^2 alternatives than for the uniform alternatives with the same mean. This is because the power envelope depends on the second uncentered moment of θ_i;
- The power of the feasible CPO test is the same for the uniform and χ^2 alternatives since power in this case depends only on the mean of θ_i;
- The test based on the t^+ statistic is less powerful than the CPO test;
- The power envelope is higher for the heterogeneous alternatives than the homogeneous alternatives with the same mean.

[5] We have also considered tests with randomly generated values for the $c_i's$. Since the results were inferior to those with fixed choices of c, we do not report them here, but they are available from the authors upon request.

For the other panels of the table, the second column gives the expected value of the autoregressive parameter implied by the distribution of the local-to-untiy parameter and the values of n and T. As can be seen, the alternatives considered are very close to 1, and at a qualitative level, the results match closely the asymptotic predictions. The main conclusions are:

- The size properties of the common point-optimal test appear to be mildly sensitive to the choice of c. The size of the test tends to increase with c;
- In terms of power, the choice of c is much less important, as predicted by asymptotic theory. In fact, most of the variation is within 2 simulation standard deviations, and much of the difference is probably due to experimental randomness;
- In all cases, power is far below what is predicted by theory and below the power envelope defined by $c_i = \theta_i$. The differences are reduced as both n and T are increased;
- In all cases, the t^+ test is less powerful than the CPO tests, but it does dominate the t-bar statistic;
- In the homogeneous cases, there is less power difference between the CPO tests and the optimal test. This is expected since the CPO test is most powerful against these alternatives;
- Finally, despite the theoretical predictions that they should be equal, the actual power for the χ^2 alternatives is slightly below that for the corresponding uniform alternatives.

Table 2 reports the same information as Table 1 for the incidental trends case. In addition to the above tests, in this case we also consider the optimal test of Ploberger and Phillips (2002) the LM test of Moon and Phillips (2004), and the unbiased test of Breitung (2000). Once again, the first panel of the table gives the predictions for size and power based on our asymptotic theory.

Just as in unit root testing with time series models, power is much lower when trends are present or fitted. In fact, power is much lower than it may first appear in the table since the actual local alternative approaches the null hypothesis at the slower rate $O\left(n^{-1/4}T^{-1}\right)$ than in the incidental intercepts case. Thus, for the same distribution of the local-to-unity parameters, the alternative hypothesis is actually further from unity than in Table 1.

The main predictions contained in the first panel of the table for the incidental trends case are as follows:

- In contrast to the incidental intercepts case, power of the CPO test is higher for χ^2 alternatives than for uniform alternatives since it depends on higher-order moments in this case;
- The Moon and Phillips test, although dominated, is expected to perform well;

- The t^+ test has lowest power as is expected;
- Breitung's unbiased test has power that lies between the common point-optimal test and the Moon and Phillips test;
- The power envelope is lower for homogeneous alternatives.

The simulation findings reported in the remaining panels of table 2 conform well to these predictions. We have not reported the finite-sample analog of the power envelope because of numerical problems encountered in the computation. In a finite sample, the terms involving high powers of c_i dominate for distant alternatives, and this pushes the distribution of the statistic to the right, leading to negligible rejection probabilities.

Our other findings for this case are:

- The size properties of the point-optimal test are much more sensitive to the choice of c and values of n and T than for the incidental intercepts case. It is therefore difficult to come up with a good choice of c based on these results, although values between 1 and 2 seem to provide a good balance for all values of n and T;
- Both the Ploberger-Phillips and Moon-Phillips tests tend to underreject, sometimes quite severely;
- The t-type test tends to overreject, and its power is close to that of Moon and Phillips;
- As in the incidental intercepts case, the power properties of the CPO test do not appear sensitive to the choice of c. There is a slight tendency for $c = 2$ to achieve highest power;
- The fatter-tailed distributions have higher power than the corresponding uniform distributions for the two closest alternatives. For the alternatives that are furthest away (cases (3) and (6)), the reverse is true;
- The Ploberger-Phillips test behaves in a similar way to the CPO test, as predicted by the asymptotics;
- The LM test of Moon and Phillips has good power but appears to be slightly dominated by the other two tests, as again predicted by our theory;
- Power of the unbiased test of Breitung is generally between that of the Ploberger-Phillips and Moon-Phillips test, again as predicted;
- When the alternative hypothesis is homogeneous (cases (7) and (8)), the tests based on a common value of c_i have higher power than for the corresponding heterogeneous alternative case. This phenomenon is more pronounced for the χ^2 alternative hypothesis.

These results suggest that the asymptotic theory generally provides a useful guide to the finite sample performance of the tests statistics in the vicinity of the panel unit root null. However, the presence of more complex deterministic components and increasing distance from the null hypothesis reduces the accuracy of the analytic results from asymptotic theory. Overall, the simulation findings strongly suggest that use of the CPO test (and the Ploberger-Phillips test in the trends case) improves power over the commonly-used t-ratio type statistics.

8 Conclusion

In terms of their asymptotic power functions, the Ploberger-Phillips (2002) test and the common point optimal test have good discriminatory power against a unit root null in shrinking neighborhoods of unity. When the alternative is homogeneous it is possible to attain the Gaussian asymptotic power envelope and both the Ploberger- Phillips test and the common point optimal test are uniformly most powerful in this case. Interestingly, the common point optimal test has this property irrespective of the point chosen to set up the test. This is in contrast to point optimal tests of a unit root that are based solely on time series data (Elliott et. al. 1996), where no test is uniformly most powerful, and an arbitrary selection of a common point is needed in the construction of the test.

An important empirical consequence of the present investigation is that increasing the complexity of the fixed effects in a panel model inevitably reduces the potential power of unit root tests. This reduction in power has a quantitative manifestation in the radial order of the shrinking neighborhoods around unity for which asymptotic power is non negligible. When there are no fixed effects or constant fixed effects, tests have power in a neighborhood of unity of order $n^{-1/2}T^{-1}$. When incidental trends are fitted, the tests only have power in a larger neighborhood of order $n^{-1/4}T^{-1}$. A continuing reduction in power is to be expected as higher order incidental trends are fitted in a panel model. The situation is analogous to what happens in time series models where unit root nonstationary data is fitted by a lagged variable and deterministic trends. In such cases, both the lagged variable and the deterministic trends compete to model the nonstationarity in the data with the upshot that the rate of convergence is affected. In particular, Phillips (2001) showed that rate of convergence to a unit root is slowed by the presence of increasing numbers of deterministic regressors. In the panel model context, the present paper shows that discriminatory power against a unit root is generally weakened as more complex deterministic regressors are included in the panel model.

9 Appendix: Technical Results and Proofs

Let $z_{it}(0)$ and $y_{it}(0)$, respectively, denote the panel observations z_{it} and y_{it} that are generated by model (1) with $\rho_i = 1$, that is, $\theta_i = 0$. Also define $Z(0)$, $Y(0)$, $Y_{-1}(0)$, respectively, in a similar fashion to Z, Y, and Y_{-1}. For notational simplicity, set $u_{i0} = y_{i0}$. Throughout the proofs, we will use the notation

$$\tilde{\sigma}_{iT}^2 = \frac{1}{T}\sum_{t=1}^{T} u_{it}^2,$$

and

$$h(r,s) = (1,r)\begin{pmatrix} 1 & \int_0^1 r dr \\ \int_0^1 r dr & \int_0^1 r^2 dr \end{pmatrix}^{-1}\begin{pmatrix} 1 \\ s \end{pmatrix} = 4 - 6r - 6s + 12rs.$$

9.1 Preliminary Results

Lemma 5 *Suppose that Assumption 1 is satisfied. Then, as $n, T \to \infty$ with $\frac{n}{T} \to 0$, the following hold.*

(a) $\sum_{i=1}^{n}\left(\tilde{\sigma}_{iT}^2 - \sigma_i^2\right)^2 = o_p(1)$.

(b) $\sup_{1\le i\le n}\left|\tilde{\sigma}_{iT}^2 - \sigma_i^2\right| = o_p(1)$.

(c) *With probability approaching one, there exists a constant $M > 0$ such that $\inf_i \tilde{\sigma}_{iT}^2 \ge M$.*

Proof: see Moon, Perron, and Phillips $(2006b)$.

Suppose that c_i is a sequence of *iid* random variables, independent of u_{it} for all i and t, with a bounded support.

Lemma 6 *Suppose that Assumptions 1 – 3, 11, and 12 hold. Then, the following hold as $(n, T \to \infty)$ with $\frac{n}{T} \to 0$.*

(a) $\frac{1}{\sqrt{n}}\sum_{i=1}^{n} c_i^2\left[\frac{1}{T^2\sigma_i^2}\sum_{t=1}^{T}\left\{(y_{it} - y_{i0}) - \frac{t}{T}(y_{iT} - y_{i0})\right\}^2 - \omega_{1T}\right] \Rightarrow N\left(-\frac{E\left(c_i^2\theta_i^2\right)}{90}, \frac{E\left(c_i^4\right)}{45}\right)$

(b) $\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\left[\frac{1}{T^2\sigma_i^2}\sum_{t=1}^{T} y_{it}^2 - \frac{1}{T^3\sigma_i^2}\sum_{t=1}^{T}\sum_{s=1}^{T} y_{it}y_{is}h_T(t,s) - \omega_{2T}\right] \Rightarrow N\left(-\frac{E\left(\theta_i^2\right)}{420}, \frac{11}{6300}\right)$.

Proof: see Moon, Perron, and Phillips $(2006b)$.

9.2 Proofs and Derivations for Section 3

Proof of Theorem 6

Since $\Delta y_{it} = -\frac{\theta_i}{n^{1/2}T} y_{it-1} + u_{it}$ under Assumption 4, we can write

$$
\begin{aligned}
& V_{nT}(\mathbb{C}) \\
= & \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left[y_{i0}^2 + \sum_{t=1}^{T} (\Delta_{c_i} y_{it})^2 \right] - \frac{1}{\sigma_i^2} \sum_{i=1}^{n} \left[y_{i0}^2 + \sum_{t=1}^{T} (\Delta y_{it})^2 \right] - \frac{1}{2}\mu_{c,2} \\
= & \frac{2}{n^{1/2}T} \sum_{i=1}^{n} \frac{c_i}{\sigma_i^2} \sum_{t=1}^{T} \Delta y_{it} y_{it-1} + \frac{1}{nT^2} \sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2} \sum_{t=1}^{T} y_{it-1}^2 - \frac{1}{2}\mu_{c,2} \\
= & -\frac{2}{nT^2} \sum_{i=1}^{n} \frac{c_i\theta_i}{\sigma_i^2} \sum_{t=1}^{T} y_{it-1}^2 + \frac{2}{n^{1/2}T} \sum_{i=1}^{n} \frac{c_i}{\sigma_i^2} \sum_{t=1}^{T} u_{it} y_{it-1} \\
& + \frac{1}{nT^2} \sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2} \sum_{t=1}^{T} y_{it-1}^2 - \frac{1}{2}\mu_{c,2}.
\end{aligned}
$$

Direct calculation shows that under the assumptions of the theorem, we have

$$
\begin{aligned}
-\frac{2}{nT^2} \sum_{i=1}^{n} \sum_{t=1}^{T} \frac{c_i\theta_i}{\sigma_i^2} y_{it-1}^2 & \to_p \quad -E(c_i\theta_i), \\
\frac{1}{nT^2} \sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2} \sum_{t=1}^{T} y_{it-1}^2 & \to_p \quad \frac{1}{2}\mu_{c,2},
\end{aligned}
$$

and

$$
\frac{2}{n^{1/2}T} \sum_{i=1}^{n} \frac{c_i}{\sigma_i^2} \sum_{t=1}^{T} u_{it} y_{it-1} \Rightarrow N\left(0, 2\mu_{c,2}\right),
$$

thereby giving the required result. ■

Lemma 7 *Let M be a finite constant. Under Assumptions 1 and 2, the following hold.*

(a) $\sup_i E\left[\left(\frac{1}{T}\sum_{t=1}^{T} u_{it} y_{it-1}\right)^2\right] < M.$

(b) $\sup_i E\left[\left(\frac{1}{T^2}\sum_{t=1}^{T} y_{it-1}^2\right)^2\right] < M.$

(c) $\sup_i E\left[y_{i0}^2\right] < M.$

Proof. The lemma follows by direct calculation and we omit the proof. ■

Lemma 8 *Suppose that Assumptions 1 – 3, and 4 hold. Then, the following hold.*

(a) $\sum_{i=1}^{n} \left(\hat{\sigma}_{1,iT}^{2} - \sigma_i^2\right)^2 = o_p(1)$.

(b) $\sup_{1 \le i \le n} \left|\hat{\sigma}_{1,iT}^{2} - \sigma_i^2\right| = o_p(1)$.

(c) *With probability approaching one, there exists a constant* $M > 0$ *such that* $\inf_i \hat{\sigma}_{1,iT}^{2} \ge M$.

Proof: see Moon, Perron, and Phillips (2006*b*).

Proof of Theorem 8.
By definition,

$$\hat{V}_{nT}(\mathbb{C}) = -\frac{2}{nT^2}\sum_{i=1}^{n}\frac{c_i\theta_i}{\hat{\sigma}_{1,iT}^{2}}\sum_{t=1}^{T} y_{it-1}^2 + \frac{2}{n^{1/2}T}\sum_{i=1}^{n}\frac{c_i}{\hat{\sigma}_{1,iT}^{2}}\sum_{t=1}^{T} u_{it}y_{it-1}$$
$$+\frac{1}{nT^2}\sum_{i=1}^{n}\frac{c_i^2}{\hat{\sigma}_{1,iT}^{2}}\sum_{t=1}^{T} y_{it-1}^2 - \frac{1}{2}\mu_{c,2}.$$

First, by the Cauchy-Schwarz inequality,

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{1,iT}^{2}} - \frac{1}{\sigma_i^2}\right)\frac{c_i\theta_i}{T^2}\sum_{t=1}^{T} y_{it-1}^2\right| \\
\le\ & \left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\hat{\sigma}_{1,iT}^{2} - \sigma_i^2}{\hat{\sigma}_{1,iT}^{2}}\right)^2\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{c_i\theta_i}{\sigma_i^2}\frac{1}{T^2}\sum_{t=1}^{T} y_{it-1}^2\right)^2\right)^{1/2} \\
\le\ & \frac{\sup_i \left|\hat{\sigma}_{1,iT}^{2} - \sigma_i^2\right|}{\inf_i \hat{\sigma}_{1,iT}^{2}}\frac{M}{\inf_i \sigma_i^2}\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T^2}\sum_{t=1}^{T} y_{it-1}^2\right)^2\right)^{1/2} \\
=\ & o_p(1)\,O_p(1) = o_p(1),
\end{aligned}$$

where the last line holds by Lemmas 7 and 8, the assumption that c_i and θ_i have uniformly bounded supports, and $\inf_i \sigma_i^2 > 0$. Similarly, by Lemmas 7 and 8, the assumption that c_i has a bounded support, and $\inf_i \sigma_i^2 > 0$, we have

$$\begin{aligned}
&\left|\frac{1}{n^{1/2}}\sum_{i=1}^{n}\left(\frac{\hat{\sigma}_{1,iT}^{2} - \sigma_i^2}{\hat{\sigma}_{1,iT}^{2}}\right)\frac{c_i}{T\sigma_i^2}\sum_{t=1}^{T} u_{it}y_{it-1}\right| \\
\le\ & \left(\sum_{i=1}^{n}\left(\frac{\hat{\sigma}_{1,iT}^{2} - \sigma_i^2}{\hat{\sigma}_{1,iT}^{2}}\right)^2\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{c_i}{T\sigma_i^2}\sum_{t=1}^{T} u_{it}y_{it-1}\right)^2\right)^{1/2} \\
\le\ & \frac{\left(\sum_{i=1}^{n}\left(\hat{\sigma}_{1,iT}^{2} - \sigma_i^2\right)^2\right)^{1/2}}{\inf_i \hat{\sigma}_{1,iT}^{2}}\frac{M}{\inf_i \sigma_i^2}\left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{T}\sum_{t=1}^{T} u_{it}y_{it-1}\right)^2\right)^{1/2} \\
=\ & o_p(1)\,O_p(1),
\end{aligned}$$

and

$$\frac{1}{nT^2}\sum_{i=1}^{n}\frac{c_i^2}{\hat{\sigma}_{1,iT}^2}\sum_{t=1}^{T}y_{it-1}^2=\frac{1}{nT^2}\sum_{i=1}^{n}\frac{c_i^2}{\sigma_i^2}\sum_{t=1}^{T}y_{it-1}^2+o_p\left(1\right).$$

Combining these, we complete the proof that $\hat{V}_{nT}\left(\mathbb{C}\right)=V_{nT}\left(\mathbb{C}\right)+o_p\left(1\right).$ ■

9.3 Proofs and Derivations for Section 4

Proof of Theorem 9.

For the theorem, it is enough to show that

$$V_{fe1,nT}\left(\mathbb{C}\right)=V_{nT}\left(\mathbb{C}\right)+o_p\left(1\right).$$

Let $\hat{b}_{0i}\left(c_i\right)=\left(\Delta_{c_i}G_0'\Delta_{c_i}G_0\right)^{-1}\left(\Delta_{c_i}G_0'\Delta_{c_i}\underline{Z}_i\right).$ Then $\underline{Z}_i-G_0\hat{b}_{0i}\left(c_i\right)=\underline{Y}_i-G_0\left(\hat{b}_{0i}\left(c_i\right)-b_{0i}\right)$, and we can rewrite $V_{fe1,nT}\left(\mathbb{C}\right)$ as

$$\begin{aligned}
&V_{fe1,nT}\left(\mathbb{C}\right)\\
=&\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\left[\begin{array}{c}\left(\Delta_{c_i}\underline{Y}_i-\Delta_{c_i}G_0\left(\hat{b}_{0i}\left(c_i\right)-b_{0i}\right)\right)'\left(\Delta_{c_i}\underline{Y}_i-\Delta_{c_i}G_0\left(\hat{b}_{0i}\left(c_i\right)-b_{0i}\right)\right)\\-\left(\Delta\underline{Y}_i-\Delta G_0\left(\hat{b}_{0i}\left(c_i\right)-b_{0i}\right)\right)'\left(\Delta\underline{Y}_i-\Delta G_0\left(\hat{b}_{0i}\left(c_i\right)-b_{0i}\right)\right)\end{array}\right]\\
&-\frac{1}{2}\mu_{c,2}\\
=&V_{nT}\left(\mathbb{C}\right)+V_{fe11,nT}\left(\mathbb{C}\right),
\end{aligned}$$

where

$$V_{fe11,nT}\left(\mathbb{C}\right)=\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\left[\begin{array}{c}\left(\Delta\underline{Y}_i'\Delta G_0\right)\left(\Delta G_0'\Delta G_0\right)^{-1}\left(\Delta G_0'\Delta\underline{Y}_i\right)\\-\left(\Delta_{c_i}\underline{Y}_i'\Delta_{c_i}G_0\right)\left(\Delta_{c_i}G_0'\Delta_{c_i}G_0\right)^{-1}\left(\Delta_{c_i}G_0'\Delta_{c_i}\underline{Y}_i\right)\end{array}\right].$$

For the required result, it is enough to show that

$$V_{fe11,nT}\left(\mathbb{C}\right)=o_p\left(1\right)$$

as $n,T\to\infty$ with $\frac{n}{T}\to0$, which follows by Lemmas 7(c) and 9 and the assumption that $\inf_i\sigma_i^2>0$, since

$$\begin{aligned}
&V_{fe11,nT}\left(\mathbb{C}\right)\\
=&\sum_{i=1}^{n}\frac{1}{\sigma_i^2}\left[y_{i0}^2-\frac{1}{1+\frac{c_i^2}{n}\frac{1}{T}}\left(y_{i0}+\frac{c_i}{n^{1/2}}\frac{1}{T}\left(y_{iT}-y_{i0}\right)+\frac{c_i^2}{n}\frac{1}{T^2}\sum_{t=1}^{T}y_{it-1}\right)^2\right]\\
=&I_1-I_2-I_3-2I_4-2I_5-2I_6,
\end{aligned}$$

and

$$I_1=\frac{1}{nT}\sum_{i=1}^{n}\frac{y_{i0}^2}{\sigma_i^2}\left(\frac{c_i^2}{1+\frac{c_i^2}{nT}}\right)=O_p\left(\frac{1}{T}\right)=o_p\left(1\right),$$

$$I_2 = \frac{1}{nT}\sum_{i=1}^{n}\frac{c_i^2}{\sigma_i^2\left(1+\frac{c_i^2}{nT}\right)}\left(\frac{y_{iT}-y_{i0}}{\sqrt{T}}\right)^2 = O_p\left(\frac{1}{T}\right) = o_p\left(1\right),$$

$$I_3 = \frac{1}{n^2T}\sum_{i=1}^{n}\frac{c_i^4}{\sigma_i^2\left(1+\frac{c_i^2}{nT}\right)}\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}y_{it-1}\right)^2 = O_p\left(\frac{1}{nT}\right) = o_p\left(1\right),$$

$$\begin{aligned}
|I_4| &= \left|\frac{1}{\sqrt{nT}}\sum_{i=1}^{n}\frac{c_i}{\sigma_i^2\left(1+\frac{c_i^2}{nT}\right)}y_{i0}\left(\frac{y_{iT}-y_{i0}}{\sqrt{T}}\right)\right| \\
&\leq \sqrt{\frac{n}{T}}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{c_i}{\sigma_i^2\left(1+\frac{c_i^2}{nT}\right)}y_{i0}^2\right)^{1/2}\left(\frac{1}{n}\sum_{i=1}^{n}\frac{c_i}{\sigma_i^2\left(1+\frac{c_i^2}{nT}\right)}\left(\frac{y_{iT}-y_{i0}}{\sqrt{T}}\right)^2\right)^{1/2} \\
&= \sqrt{\frac{n}{T}}O_p\left(1\right)O_p\left(1\right) = o_p\left(1\right),
\end{aligned}$$

and, similarly,

$$I_5 = \frac{1}{n^{3/2}T}\sum_{i=1}^{n}\frac{c_i^3}{\sigma_i^2\left(1+\frac{c_i^2}{nT}\right)}\left(\frac{y_{iT}-y_{i0}}{\sqrt{T}}\right)\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}y_{it-1}\right) = o_p\left(1\right),$$

$$I_6 = \frac{1}{n\sqrt{T}}\sum_{i=1}^{n}\frac{c_i^2}{\sigma_i^2\left(1+\frac{c_i^2}{nT}\right)}y_{i0}\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}y_{it-1}\right) = o_p\left(1\right),$$

as required. ■

Lemma 9 *Let M be a finite constant. Under Assumptions 1 and 2, the following hold.*

(a) $\sup_i E\left[\left(\frac{y_{iT}-y_{i0}}{\sqrt{T}}\right)^2\right] < M.$

(b) $\sup_i E\left[\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}y_{it-1}\right)^2\right] < M.$

Proof. The lemma follows by direct calculation, and its proof is omitted. ■

Lemma 10 *Suppose that Assumptions 1 – 3, and 4 hold. Then, the following hold.*

(a) $\sup_{1\leq i\leq n}\left(\hat{\sigma}_{2,iT}^2-\sigma_i^2\right) = o_p\left(1\right).$

(b) *With probability approaching one, there exists a constant $M > 0$ such that* $\inf_i \hat{\sigma}_{2,iT}^2 \geq M$.

Proof: see Moon, Perron, and Phillips (2006*b*).

Proof of Theorem 10

Using Lemmas 7(c), 9, and 10 and the assumptions that the supports of θ_i and c_i are bounded and $\inf_i \sigma_i^2 > 0$, we can show using arguments similar to those used in the proof of Theorem 8 that

$$\begin{aligned}
& \hat{V}_{fe11,nT}(\mathbb{C}) \\
= & \sum_{i=1}^{n} \frac{1}{\hat{\sigma}_i^2} \left[y_{i0}^2 - \frac{1}{1+\frac{c_i^2}{n}\frac{1}{T}} \left(y_{i0} + \frac{c_i}{n^{1/2}} \frac{1}{T} (y_{iT} - y_{i0}) + \frac{c_i^2}{n} \frac{1}{T^2} \sum_{t=1}^{T} y_{it-1} \right)^2 \right] \\
= & V_{fe11,nT}(\mathbb{C}) + o_p(1).
\end{aligned}$$

The required result now follows. ■

9.4 Proofs and Derivations for Section 5

Lemma 11 *Under Assumptions 1 – 3, 11, and 12,*

$$\begin{aligned}
& V_{fe2,nT}(\mathbb{C}) \\
= & \frac{1}{n^{1/4}} \sum_{i=1}^{n} \frac{c_i}{\sigma_i^2} \left[\frac{2}{T} \sum_{t=1}^{T} \Delta y_{it} y_{it-1} - \left(\frac{y_{iT}}{\sqrt{T}} \right)^2 + \left(\frac{y_{i0}}{\sqrt{T}} \right)^2 + \sigma_i^2 \right] \\
& + \frac{1}{n^{1/2}} \sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2} \left[\begin{array}{c} \frac{1}{T^2} \sum_{t=1}^{T} y_{it-1}^2 - 2 \left(\frac{y_{iT}}{\sqrt{T}} \right) \left(\frac{1}{T\sqrt{T}} \sum_{t=1}^{T} \frac{t}{T} y_{it-1} \right) \\ + \frac{1}{3} \left(\frac{y_{iT}}{\sqrt{T}} \right)^2 + \sigma_i^2 \omega_{p2T} \end{array} \right] \\
& + \frac{1}{n} \sum_{i=1}^{n} \frac{c_i^4}{\sigma_i^2} \left[\begin{array}{c} -\left(\frac{1}{T\sqrt{T}} \sum_{t=1}^{T} \frac{t}{T} y_{it-1} \right)^2 + \frac{2}{3} \left(\frac{y_{iT}}{\sqrt{T}} \right) \left(\frac{1}{T\sqrt{T}} \sum_{t=1}^{T} \frac{t}{T} y_{it-1} \right) \\ -\frac{1}{9} \left(\frac{y_{iT}}{\sqrt{T}} \right)^2 + \sigma_i^2 \omega_{p4T} \end{array} \right] \\
& + \frac{1}{n^{1/4} T} \sum_{i=1}^{n} \frac{\mathcal{S}_{1iT}}{\sigma_i^2} + \frac{1}{n^{1/2} T^{1/2}} \sum_{i=1}^{n} \frac{\mathcal{S}_{2iT}}{\sigma_i^2} + \frac{1}{n^{5/4}} \sum_{i=1}^{n} \frac{\mathcal{S}_{3iT}}{\sigma_i^2},
\end{aligned}$$

with $\frac{1}{n} \sum_{i=1}^{n} E\left[\mathcal{S}_{kiT}^2 \right] = O(1)$, *for* $k = 1, 2, 3$ *when* $(n, T \to \infty)$ *with* $\frac{n}{T} \to 0$.

Proof: see Moon, Perron, and Phillips (2006*b*).

Lemma 12 *Under Assumptions 1 – 3, 11, and 12, the following hold:*

(a) $\frac{1}{n^{1/4}} \sum_{i=1}^{n} \frac{c_i}{\sigma_i^2} \left[\frac{2}{T} \sum_{t=1}^{T} \Delta y_{it} y_{it-1} - \left(\frac{y_{iT}}{\sqrt{T}} \right)^2 + \left(\frac{y_{i0}}{\sqrt{T}} \right)^2 + \sigma_i^2 \right] = o_p(1);$

(b) $\frac{1}{n^{1/2}} \sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2} \left[\begin{array}{c} \left(\frac{1}{T^2} \sum_{t=1}^{T} y_{it-1}^2 - \sigma_i^2 \frac{1}{T} \sum_{t=1}^{T} \frac{t-1}{T} \right) + \frac{1}{3} \left\{ \left(\frac{y_{iT}}{\sqrt{T}} \right)^2 - \sigma_i^2 \right\} \\ -\left\{ 2 \left(\frac{y_{iT}}{\sqrt{T}} \right) \left(\frac{1}{T\sqrt{T}} \sum_{t=1}^{T} \frac{t}{T} y_{it-1} \right) - \sigma_i^2 \frac{2}{T} \sum_{t=1}^{T} \left(\frac{t}{T} \right) \left(\frac{t-1}{T} \right) \right\} \end{array} \right] \Rightarrow$
$N\left(-\frac{1}{90} E\left(c_i^2 \theta_i^2 \right), \frac{1}{45} E\left(c_i^4 \right) \right);$

(c) $\frac{1}{n}\sum_{i=1}^{n}\frac{c_i^4}{\sigma_i^2}\left[\begin{array}{c} -\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}\frac{t}{T}y_{it-1}\right)^2+\frac{2}{3}\left(\frac{y_{iT}}{\sqrt{T}}\right)\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}\frac{t}{T}y_{it-1}\right) \\ -\frac{1}{9}\left(\frac{y_{iT}}{\sqrt{T}}\right)^2+\sigma_i^2\omega_{p4T} \end{array}\right]=$ $o_p(1)$.

Proof: see Moon, Perron, and Phillips (2006*b*).

Lemma 13 *Let M be a finite constant. Under Assumptions 1 – 3, 11, and 12, the following hold.*

(a) $\sup_i E\left[y_{i0}^4\right]<M.$

(b) $\sup_i E\left[\left(\frac{y_{iT}}{\sqrt{T}}\right)^4\right]<M.$

(c) $\sup_i E\left[\left(\frac{1}{T}\sum_{t=1}^{T}y_{it-1}u_{it}\right)^2\right]<M.$

(d) $\sup_i E\left[\left(\frac{1}{T^2}\sum_{t=1}^{T}y_{it-1}^2\right)^2\right]<M.$

(e) $\sup_i E\left[\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}y_{it-1}\right)^4\right]<M.$

(f) $\sup_i E\left[\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}\frac{t-1}{T}y_{it-1}\right)^4\right]<M.$

Proof. The lemma follows by direct calculations and we omit the proof. ■

Lemma 14 *Suppose that Assumptions 1 – 3, and 11 hold. Then, the following hold.*

(a) $\sum_{i=1}^{n}\left(\hat{\sigma}_{1,iT}^2-\sigma_i^2\right)^2=o_p(1).$

(b) $\sup_{1\leq i\leq n}\left(\hat{\sigma}_{1,iT}^2-\sigma_i^2\right)=o_p(1).$

(c) $\sum_{i=1}^{n}\left(\hat{\sigma}_{3,iT}^2-\sigma_i^2\right)^2=o_p(1).$

(d) $\sup_{1\leq i\leq n}\left(\hat{\sigma}_{3,iT}^2-\sigma_i^2\right)=o_p(1).$

(e) *With probability approaching one, there exists a constant $M>0$ such that* $\inf_i\hat{\sigma}_{3,iT}^2\geq M$.

Proof: see Moon, Perron, and Phillips (2006*b*).

Proof of Theorem 15

For the required result, it is enough to show that

$$(a)\quad : \quad \frac{1}{n^{1/4}}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{3,iT}^{2}}-\frac{1}{\sigma_i^2}\right)c_i\left[\frac{2}{T}\sum_{t=1}^{T}\Delta y_{it}y_{it-1}-\left(\frac{y_{iT}}{\sqrt{T}}\right)^2+\left(\frac{y_{i0}}{\sqrt{T}}\right)^2+\sigma_i^2\right]=o_p\left(1\right),$$

$$(b)\quad : \quad \frac{1}{n^{1/4}}\sum_{i=1}^{n}\left(\frac{\hat{\sigma}_{3,iT}^{2}-\sigma_i^2}{\hat{\sigma}_{3,iT}^{2}}\right)=o_p\left(1\right),$$

$$(c)\quad : \quad \frac{1}{n^{1/2}}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{3,iT}^{2}}-\frac{1}{\sigma_i^2}\right)c_i^2\left[\begin{array}{c}\frac{1}{T^2}\sum_{t=1}^{T}y_{it-1}^2-2\left(\frac{y_{iT}}{\sqrt{T}}\right)\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}\frac{t}{T}y_{it-1}\right)\\ +\frac{1}{3}\left(\frac{y_{iT}}{\sqrt{T}}\right)^2+\sigma_i^2\omega_{p2T}\end{array}\right]=o_p\left(1\right),$$

$$(d)\quad : \quad \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{3,iT}^{2}}-\frac{1}{\sigma_i^2}\right)c_i^4\left[\begin{array}{c}-\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}\frac{t}{T}y_{it-1}\right)^2+\frac{2}{3}\left(\frac{y_{iT}}{\sqrt{T}}\right)\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}\frac{t}{T}y_{it-1}\right)\\ -\frac{1}{9}\left(\frac{y_{iT}}{\sqrt{T}}\right)^2+\sigma_i^2\omega_{p4T}\end{array}\right]=o_p\left(1\right),$$

$$(e)\quad : \quad \frac{1}{n^{1/4}T}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{3,iT}^{2}}-\frac{1}{\sigma_i^2}\right)\mathcal{S}_{1iT}=o_p\left(1\right),$$

$$(f)\quad : \quad \frac{1}{n^{1/2}T^{1/2}}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{3,iT}^{2}}-\frac{1}{\sigma_i^2}\right)\mathcal{S}_{2iT}=o_p\left(1\right),$$

$$(g)\quad : \quad \frac{1}{n^{5/4}}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{3,iT}^{2}}-\frac{1}{\sigma_i^2}\right)\mathcal{S}_{3iT}=o_p\left(1\right).$$

Parts (c) – (g) hold by arguments similar to those used in the proof of Theorem 8, that is, use the Cauchy-Schwarz inequality, Lemmas 13, 14 and the assumptions that the supports of θ_i and c_i are uniformly bounded and $\inf_i \sigma_i^2 > 0$. For Part (a), notice by definition that

$$\begin{aligned}
&\frac{1}{n^{1/4}}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{3,iT}^{2}}-\frac{1}{\sigma_i^2}\right)c_i\left[\frac{2}{T}\sum_{t=1}^{T}\Delta y_{it}y_{it-1}-\left(\frac{y_{iT}}{\sqrt{T}}\right)^2+\left(\frac{y_{i0}}{\sqrt{T}}\right)^2+\sigma_i^2\right]\\
=\;&\frac{1}{n^{1/4}}\sum_{i=1}^{n}\left(\frac{1}{\hat{\sigma}_{3,iT}^{2}}-\frac{1}{\sigma_i^2}\right)c_i\left[-\left(\rho_i-1\right)^2\frac{1}{T}\sum_{t=1}^{T}y_{it-1}^2+2\left(1-\rho_i\right)\frac{1}{T}\sum_{t=1}^{T}y_{it-1}u_{it}-\left(\frac{1}{T}\sum_{t=1}^{T}u_{it}^2-\sigma_i^2\right)\right]\\
=\;&\frac{1}{n^{3/4}T}\sum_{i=1}^{n}\left(\frac{\hat{\sigma}_{3,iT}^{2}-\sigma_i^2}{\hat{\sigma}_{3,iT}^{2}\sigma_i^2}\right)\frac{c_i\theta_i^2}{T^2}\sum_{t=1}^{T}y_{it-1}^2-\frac{2}{n^{1/2}T}\sum_{i=1}^{n}\left(\frac{\hat{\sigma}_{3,iT}^{2}-\sigma_i^2}{\hat{\sigma}_{3,iT}^{2}\sigma_i^2}\right)\frac{c_i\theta_i}{T}\sum_{t=1}^{T}y_{it-1}u_{it}\\
&+\frac{1}{n^{1/4}T^{1/2}}\sum_{i=1}^{n}\left(\frac{\hat{\sigma}_{3,iT}^{2}-\sigma_i^2}{\hat{\sigma}_{3,iT}^{2}\sigma_i^2}\right)\left(\frac{1}{T^{1/2}}\sum_{t=1}^{T}\left(u_{it}^2-\sigma_i^2\right)\right).
\end{aligned}$$

Using similar arguments to those in the proofs of Parts (c) – (g), we can show that the first and the second terms are of $O_p\left(\frac{1}{n^{1/4}T}\right)$ and $O_p\left(\frac{1}{T}\right)$, respectively. Also the third term is $o_p\left(1\right)$ since by the Cauchy-Schwarz inequality, Lemma

14, and the assumption $\inf_i \sigma_i^2 > 0$, it follows that

$$
\begin{aligned}
& \left| \frac{1}{n^{1/4} T^{1/2}} \sum_{i=1}^{n} \left(\frac{\hat{\sigma}_{3,iT}^2 - \sigma_i^2}{\hat{\sigma}_{3,iT}^2 \sigma_i^2} \right) \left(\frac{1}{T^{1/2}} \sum_{t=1}^{T} \left(u_{it}^2 - \sigma_i^2 \right) \right) \right| \\
\leq\ & \frac{n^{1/4}}{T^{1/2}} \left(\sum_{i=1}^{n} \left(\frac{\hat{\sigma}_{3,iT}^2 - \sigma_i^2}{\hat{\sigma}_{3,iT}^2 \sigma_i^2} \right)^2 \right)^{1/2} \left(\frac{1}{n} \sum_{i=1}^{n} \left(T^{1/2} \sum_{t=1}^{T} \left(u_{it}^2 - \sigma_i^2 \right) \right)^2 \right)^{1/2} \\
\leq\ & \frac{n^{1/4}}{T^{1/2}} \frac{1}{\inf_i \hat{\sigma}_{3,iT}^2} \frac{1}{\inf_i \sigma_i^2} \left(\sum_{i=1}^{n} \left(\hat{\sigma}_{3,iT}^2 - \sigma_i^2 \right)^2 \right)^{1/2} \left(\frac{1}{n} \sum_{i=1}^{n} \left(T^{1/2} \sum_{t=1}^{T} \left(u_{it}^2 - \sigma_i^2 \right) \right)^2 \right)^{1/2} \\
=\ & \frac{n^{1/4}}{T^{1/2}} O_p(1)\, o_p(1)\, O_p(1) = o_p(1),
\end{aligned}
$$

which yields Part (a).

For Part (b), notice that

$$
\frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\frac{\hat{\sigma}_{3,iT}^2 - \sigma_i^2}{\hat{\sigma}_{3,iT}^2} \right) = \frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\frac{\hat{\sigma}_{3,iT}^2 - \sigma_i^2}{\sigma_i^2} \right) + \frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\hat{\sigma}_{3,iT}^2 - \sigma_i^2 \right) \left(\frac{1}{\hat{\sigma}_{3,iT}^2} - \frac{1}{\sigma_i^2} \right).
$$

The second term is $o_p(1)$ by Lemma 14 and the assumption $\inf_i \sigma_i^2 > 0$, since

$$
\begin{aligned}
\left| \frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\hat{\sigma}_{3,iT}^2 - \sigma_i^2 \right) \left(\frac{1}{\hat{\sigma}_{3,iT}^2} - \frac{1}{\sigma_i^2} \right) \right| &= \left| \frac{1}{n^{1/4}} \sum_{i=1}^{n} \frac{\left(\hat{\sigma}_{3,iT}^2 - \sigma_i^2 \right)^2}{\hat{\sigma}_{3,iT}^2 \sigma_i^2} \right| \\
&\leq \frac{1}{n^{1/4}} \frac{1}{\inf_i \hat{\sigma}_{3,iT}^2} \frac{1}{\inf_i \sigma_i^2} \left(\sum_{i=1}^{n} \left(\hat{\sigma}_{3,iT}^2 - \sigma_i^2 \right)^2 \right) = o_p(1).
\end{aligned}
$$

To complete the proof of Part (b), it is enough to show that the first term is $o_p(1)$. Write the first term as

$$
\begin{aligned}
& \frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\frac{\hat{\sigma}_{3,iT}^2 - \sigma_i^2}{\sigma_i^2} \right) \\
=\ & \frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\frac{\hat{\sigma}_{3,iT}^2 - \hat{\sigma}_{1,iT}^2}{\sigma_i^2} \right) + \frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\frac{\hat{\sigma}_{1,iT}^2 - \tilde{\sigma}_{iT}^2}{\sigma_i^2} \right) + \frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\frac{\tilde{\sigma}_{iT}^2 - \sigma_i^2}{\sigma_i^2} \right).
\end{aligned}
$$

By definition and by Lemma 13, we have

$$
\frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\frac{\hat{\sigma}_{3,iT}^2 - \hat{\sigma}_{1,iT}^2}{\sigma_i^2} \right) = \frac{1}{n^{1/4} T} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(y_{i0}^2 + \left(\frac{y_{iT} - y_{i0}}{\sqrt{T}} \right)^2 \right) = O_p\left(\frac{n^{3/4}}{T} \right) = o_p(1),
$$

$$
\begin{aligned}
\frac{1}{n^{1/4}} \sum_{i=1}^{n} \left(\frac{\hat{\sigma}_{1,iT}^2 - \tilde{\sigma}_{iT}^2}{\sigma_i^2} \right) &= \frac{1}{n^{1/4}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{y_{i0}^2}{T} + \frac{\theta_i^2}{n^{1/2} T} \left(\frac{1}{T^2} \sum_{t=1}^{T} y_{it-1}^2 \right) - 2 \frac{\theta_i}{n^{1/4} T} \left(\frac{1}{T} \sum_{t=1}^{T} u_{it} y_{it-1} \right) \right) \\
&= O_p\left(\frac{n^{3/4}}{T} \right) + O_p\left(\frac{n^{1/4}}{T} \right) + O_p\left(\frac{1}{T} \right) = o_p(1),
\end{aligned}
$$

and

$$\begin{aligned}
\frac{1}{n^{1/4}}\sum_{i=1}^{n}\left(\frac{\tilde{\sigma}_{iT}^{2}-\sigma_{i}^{2}}{\sigma_{i}^{2}}\right) &= \frac{1}{n^{1/4}T^{1/2}}\sum_{i=1}^{n}\frac{1}{\sigma_{i}^{2}}\left(\frac{1}{T^{1/2}}\sum_{t=1}^{T}\left(u_{it}^{2}-\sigma_{i}^{2}\right)\right)\\
&= O_{p}\left(\frac{n^{1/4}}{T^{1/2}}\right)=o_{p}\left(1\right),
\end{aligned}$$

the last line holding because $E\left[\frac{1}{n^{1/2}}\sum_{i=1}^{n}\frac{1}{\sigma_{i}^{2}}\left(\frac{1}{T^{1/2}}\sum_{t=1}^{T}\left(u_{it}^{2}-\sigma_{i}^{2}\right)\right)\right]^{2}=O\left(1\right).$ Combining these, we have

$$\frac{1}{n^{1/4}}\sum_{i=1}^{n}\left(\frac{\hat{\sigma}_{3,iT}^{2}-\sigma_{i}^{2}}{\sigma_{i}^{2}}\right)=o_{p}\left(1\right),$$

as required. ■

Proof of Lemma 1

The lemma holds by Lemma 6(a) with $c_i = 1$. ■

Proof of Lemma 2

The lemma holds by Lemma 6(b). ■

9.5 Proofs and Derivations for Section 6

Proof of Theorem 16:

Denote $Z_{\mathbb{C}}^{*}=\left(\Sigma^{-1/2}\otimes I_{T+1}\right)\Delta_{\mathbb{C}}vec\left(Z'\right), G_{0,\mathbb{C}}^{*}=\left(\Sigma^{-1/2}\otimes I_{T+1}\right)\Delta_{\mathbb{C}}\left(I_{n}\otimes G_{0}\right),$ $G_{1,\mathbb{C}}^{*}=\left(\Sigma^{-1/2}\otimes I_{T+1}\right)\Delta_{\mathbb{C}}\left(I_{n}\otimes G_{0}\right)l_{n}$, $Y_{\mathbb{C}}^{*}=\left(\Sigma^{-1/2}\otimes I_{T+1}\right)\Delta_{\mathbb{C}}vec\left(Y'\right)$, and $M_{0,\mathbb{C}}^{*}=I_{n(T+1)}-G_{0,\mathbb{C}}^{*}\left(G_{0,\mathbb{C}}^{*\prime}G_{0,\mathbb{C}}^{*}\right)^{-1}G_{0,\mathbb{C}}^{*\prime}$. Under the null, when $\mathbb{C}$ =0, we denote these quantities by Z_{0}^{*}, $G_{0,0}^{*}$, $G_{1,0}^{*}$ Y_{0}^{*}, and M_{0}^{*} respectively. Then, by definition

$$Z_{\mathbb{C}}^{*}=G_{0,\mathbb{C}}^{*}\beta_{0}+G_{1,\mathbb{C}}^{*}b_{1}+Y_{\mathbb{C}}^{*}.$$

Using this notation, we may express

$$\begin{aligned}
V_{fe3,nT}\left(\mathbb{C}\right) &= -2\left[\min_{\beta_{0},b_{1}}L_{nT}\left(\mathbb{C},\beta_{0},b_{1}\right)-\min_{\beta_{0},b_{1}}L_{nT}\left(0,\beta_{0},b_{1}\right)\right]-\frac{1}{2}\mu_{c,2}\\
&= Y_{\mathbb{C}}^{*\prime}M_{0,\mathbb{C}}^{*}Y_{\mathbb{C}}^{*}-Y_{\mathbb{C}}^{*\prime}M_{0,\mathbb{C}}^{*}G_{1,\mathbb{C}}^{*}\left(G_{1,\mathbb{C}}^{*\prime}M_{0,\mathbb{C}}^{*}G_{1,\mathbb{C}}^{*}\right)^{-1}G_{1,\mathbb{C}}^{*\prime}M_{0,\mathbb{C}}^{*}Y_{\mathbb{C}}^{*}\\
&\quad -Y_{0}^{*\prime}M_{1,0}^{*}Y_{0}^{*}+Y_{0}^{*\prime}M_{1,0}^{*}G_{1,0}^{*}\left(G_{1,0}^{*\prime}M_{1,0}^{*}G_{1,0}^{*}\right)^{-1}G_{1,0}^{*\prime}M_{1,0}^{*}Y_{0}^{*}-\frac{1}{2}\mu_{c,2}.
\end{aligned}$$

In what follows we show that

$$\begin{aligned}
& Y_{\mathbb{C}}^{*\prime}M_{0,\mathbb{C}}^{*}G_{1,\mathbb{C}}^{*}\left(G_{1,\mathbb{C}}^{*\prime}M_{0,\mathbb{C}}^{*}G_{1,\mathbb{C}}^{*}\right)^{-1}G_{1,\mathbb{C}}^{*\prime}M_{0,\mathbb{C}}^{*}Y_{\mathbb{C}}^{*}\\
& -Y_{0}^{*\prime}M_{1,0}^{*}G_{1,0}^{*}\left(G_{1,0}^{*\prime}M_{1,0}^{*}G_{1,0}^{*}\right)^{-1}G_{1,0}^{*\prime}M_{1,0}^{*}Y_{0}^{*}\\
= & \; o_{p}\left(1\right). \qquad (25)
\end{aligned}$$

Then, by definition, it follows that

$$\begin{aligned}
V_{fe3,nT}(\mathbb{C}) &= Y_{\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} Y_{\mathbb{C}}^{*} - Y_{0}^{*\prime} M_{1,0}^{*} Y_{0}^{*} - \frac{1}{2}\mu_{c,2} + o_p(1) \\
&= -2\left[\min_{\beta_0} L_{nT}(\mathbb{C}, \beta_0) - \min_{\beta_0} L_{nT}(0, \beta_0)\right] - \frac{1}{2}\mu_{c,2} + o_p(1) \\
&= V_{fe1,nT}(\mathbb{C}) + o_p(1),
\end{aligned}$$

as required for the theorem. ■

Proof of (25)

By definition

$$\begin{aligned}
& Y_{\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*} \left(G_{1,\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*}\right)^{-1} G_{1,\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} Y_{\mathbb{C}}^{*} - Y_{0}^{*\prime} M_{1,0}^{*} G_{1,0}^{*} \left(G_{1,0}^{*\prime} M_{1,0}^{*} G_{1,0}^{*}\right)^{-1} G_{1,0}^{*\prime} M_{1,0}^{*} Y_{0}^{*} \\
= & \frac{\left(\frac{1}{\sqrt{nT}} Y_{\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*}\right)^2 - \left(\frac{1}{\sqrt{nT}} Y_{0}^{*\prime} M_{1,0}^{*} G_{1,0}^{*}\right)^2}{\frac{1}{nT} G_{1,\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*}} \\
& + \left(\frac{1}{\sqrt{nT}} Y_{0}^{*\prime} M_{1,0}^{*} G_{1,0}^{*}\right)^2 \left(\frac{1}{\frac{1}{nT} G_{1,\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*}} - \frac{1}{\frac{1}{nT} G_{1,0}^{*\prime} M_{1,0}^{*} G_{1,0}^{*}}\right) \\
= & I + II, \text{ say.}
\end{aligned}$$

For term I, with probability approaching one,

$$\frac{G_{1,\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*}}{nT} > 0,$$

since

$$\begin{aligned}
\frac{G_{1,\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*}}{nT} &= \frac{1}{nT}\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\Delta_{c_i} G_1\right)' \left(\Delta_{c_i} G_1\right) - \frac{1}{nT}\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \frac{\left[\left(\Delta_{c_i} G_1\right)' \left(\Delta_{c_i} G_0\right)\right]^2}{\left(\Delta_{c_i} G_0\right)' \left(\Delta_{c_i} G_0\right)} \\
&= \frac{1}{nT}\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \sum_{t=1}^{T} \left(1 + \frac{c_i}{n^{1/2}}\frac{t}{T}\right)^2 - \frac{1}{n^2 T}\sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2} \frac{\left[\frac{1}{T}\sum_{t=1}^{T}\left(1 + \frac{c_i}{n^{1/2}}\frac{t}{T}\right)\right]^2}{1 + \frac{c_i^2}{nT}} \\
&\geq \frac{1 + o(1)}{\inf_i \sigma_i^2}.
\end{aligned}$$

Next,

$$\begin{aligned}
\frac{1}{\sqrt{nT}} Y_{\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*} &= \frac{1}{\sqrt{nT}} Y_{\mathbb{C}}^{*\prime} G_{1,\mathbb{C}}^{*} - \frac{1}{\sqrt{nT}} Y_{\mathbb{C}}^{*\prime} G_{0,\mathbb{C}}^{*} \left(G_{0,\mathbb{C}}^{*\prime} G_{0,\mathbb{C}}^{*}\right)^{-1} G_{0,\mathbb{C}}^{*\prime} G_{1,\mathbb{C}}^{*} \\
&= \frac{1}{\sqrt{n}}\sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{y_{iT} - y_{i0}}{\sqrt{T}} + \frac{c_i}{\sqrt{n}}\frac{y_{iT}}{\sqrt{T}} + \frac{c_i^2}{n}\frac{1}{T\sqrt{T}}\sum_{t=1}^{T}\frac{t}{T} y_{it-1}\right) + o_p(1)
\end{aligned}$$

because

$$
\begin{aligned}
& \frac{1}{\sqrt{nT}} Y_{\mathbb{C}}^{*\prime} G_{0,\mathbb{C}}^{*} \left(G_{0,\mathbb{C}}^{*\prime} G_{0,\mathbb{C}}^{*} \right)^{-1} G_{0,\mathbb{C}}^{*\prime} G_{1,\mathbb{C}}^{*} \\
= \quad & \frac{1}{n\sqrt{T}} \sum_{i=1}^{n} \frac{c_i}{\sigma_i^2} \frac{\left(y_{i0} + \frac{c_i}{n^{1/2}} \frac{(y_{iT} - y_{i0})}{T} + \frac{c_i^2}{n} \frac{1}{T^2} \sum_{t=1}^{T} y_{it-1} \right) \left(1 + \frac{c_i}{n^{1/2}} \frac{1}{T} \sum_{t=1}^{T} \frac{t}{T} \right)}{1 + \frac{c_i^2}{nT}} = O_p\left(\frac{1}{\sqrt{T}} \right).
\end{aligned}
$$

Similarly, we have

$$
\frac{1}{\sqrt{nT}} Y_0^{*\prime} M_{1,0}^{*} G_{1,0}^{*} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{y_{iT} - y_{i0}}{\sqrt{T}} \right) + o_p(1).
$$

Then, since $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{y_{iT} - y_{i0}}{\sqrt{T}} \right) = O_p(1)$ and $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{y_{iT} - y_{i0}}{\sqrt{T}} + \frac{c_i}{\sqrt{n}} \frac{y_{iT}}{\sqrt{T}} + \frac{c_i^2}{n} \frac{1}{T\sqrt{T}} \sum_{t=1}^{T} \frac{t}{T} y_{it-1} \right) = O_p(1)$, the numerator of term I is

$$
\begin{aligned}
& \left(\frac{1}{\sqrt{nT}} Y_{\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*} \right)^2 - \left(\frac{1}{\sqrt{nT}} Y_0^{*\prime} M_{1,0}^{*} G_{1,0}^{*} \right)^2 \\
= \quad & \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{y_{iT} - y_{i0}}{\sqrt{T}} + \frac{c_i}{\sqrt{n}} \frac{y_{iT}}{\sqrt{T}} + \frac{c_i^2}{n} \frac{1}{T\sqrt{T}} \sum_{t=1}^{T} \frac{t}{T} y_{it-1} \right) \right\}^2 \\
& - \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{y_{iT} - y_{i0}}{\sqrt{T}} \right) \right\}^2 + o_p(1) \\
= \quad & 2 \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{y_{iT} - y_{i0}}{\sqrt{T}} \right) \right\} \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{c_i}{\sqrt{n}} \frac{y_{iT}}{\sqrt{T}} + \frac{c_i^2}{n} \frac{1}{T\sqrt{T}} \sum_{t=1}^{T} \frac{t}{T} y_{it-1} \right) \right\} \\
& + \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{c_i}{\sqrt{n}} \frac{y_{iT}}{\sqrt{T}} + \frac{c_i^2}{n} \frac{1}{T\sqrt{T}} \sum_{t=1}^{T} \frac{t}{T} y_{it-1} \right) \right\}^2 + o_p(1) \\
= \quad & o_p(1),
\end{aligned}
$$

where the last line holds since

$$
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{c_i}{\sqrt{n}} \frac{y_{iT}}{\sqrt{T}} + \frac{c_i^2}{n} \frac{1}{T\sqrt{T}} \sum_{t=1}^{T} \frac{t}{T} y_{it-1} \right) = O_p\left(\frac{1}{n^{1/2}} \right) = o_p(1).
$$

Therefore, we have

$$
I = o_p(1).
$$

Next, we show that $II = o_p(1)$. Since $\frac{1}{\sqrt{nT}} Y_0^{*\prime} M_{1,0}^{*} G_{1,0}^{*} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left(\frac{y_{iT} - y_{i0}}{\sqrt{T}} \right) + o_p(1) = O_p(1)$, the required result $II = o_p(1)$ follows if we show that

$$
\left(\frac{1}{\frac{1}{nT} G_{1,\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*}} - \frac{1}{\frac{1}{nT} G_{1,0}^{*\prime} M_{1,0}^{*} G_{1,0}^{*}} \right) = o_p(1),
$$

which follows because with probability approaching one,

$$\frac{G_{1,0}^{*\prime} M_{1,0}^{*} G_{1,0}^{*}}{nT} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \geq \frac{1}{\inf_i \sigma_i^2},$$

and

$$\begin{aligned}
& \frac{G_{1,\mathbb{C}}^{*\prime} M_{0,\mathbb{C}}^{*} G_{1,\mathbb{C}}^{*}}{nT} - \frac{G_{1,0}^{*\prime} M_{1,0}^{*} G_{1,0}^{*}}{nT} \\
= & \frac{1}{n} \sum_{i=1}^{n} \frac{1}{\sigma_i^2} \left\{ \frac{1}{T} \sum_{t=1}^{T} \left(1 + \frac{c_i}{n^{1/2}} \frac{t}{T} \right)^2 - 1 \right\} - \frac{1}{n^2 T} \sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2} \frac{\left[\frac{1}{T} \sum_{t=1}^{T} \left(1 + \frac{c_i}{n^{1/2}} \frac{t}{T} \right) \right]^2}{1 + \frac{c_i^2}{nT}} \\
= & \frac{1}{n^{3/2}} \sum_{i=1}^{n} \frac{c_i}{\sigma_i^2} \frac{1}{T} \sum_{t=1}^{T} \left(2 \frac{t}{T} + \frac{c_i}{n^{1/2}} \left(\frac{t}{T} \right)^2 \right) - \frac{1}{n^2 T} \sum_{i=1}^{n} \frac{c_i^2}{\sigma_i^2} \frac{\left[\frac{1}{T} \sum_{t=1}^{T} \left(1 + \frac{c_i}{n^{1/2}} \frac{t}{T} \right) \right]^2}{1 + \frac{c_i^2}{nT}} \\
= & O\left(\frac{1}{n^{1/2}} \right) + O\left(\frac{1}{n} \right) = o(1).
\end{aligned}$$

Therefore,

$$II = o_p(1). \blacksquare$$

References

[1] Andrews, D. (1991): Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica*, 59, 817–858.

[2] Bai, J. and S. Ng (2004): A PANIC Attack on Unit Roots and Cointegration, *Econometrica*, 72, 1127-1177.

[3] Bowman, D. (2002): Efficient Tests for Autoregressive Unit Roots in Panel Data, *Mimeo.*

[4] Breitung, J. (1999): The Local Power of Some Unit Root Tests for Panel Data, SFB 373 *Discussion paper*, No 69–1999, Humboldt University, Berlin.

[5] Breitung, J. (2000): The Local Power of Some Unit Root Tests for Panel Data, in: B. Baltagi (ed.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels, Advances in Econometrics, Vol. 15,* JAI: Amsterdam, 161-178.

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

[7] Dufour, J. and M. King (1991): Optimal Invariant Tests for the Autocorrelation Coefficient in Linear Regressions with Stationary or Nonstationary AR(1) Errors, *Journal of Econometrics*, 47, 115–143.

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

[9] Evans, G. B. A. and N. E. Savin (1984), "Testing for unit roots: 2," *Econometrica* 52, 1241–1269.

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

[11] King, M. (1988): Towards a Theory of Point Optimal Testing, *Econometric Reviews*, 6, 169–218.

[12] Lehmann, E. (1959): *Testing Statistical Hypotheses*, New York, Wiley.

[13] Levin, A., F. Lin, and C. Chu (2002): Unit Root Tests in Panel Data: Asymptotic and Finite-Sample Properties, *Journal of Econometrics*, 108, 1–24.

[14] Maddala, G.S. and S. Wu (1999): A Comparative Study of Unit Root Tests with Panel Data and a New Simple Test, *Oxford Bulletin of Economics and Statistics*, 61, 631–651.

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

[16] Moon, H.R. and B. Perron (2005): Asymptotic Local Power of Pooled t-ratio Tests for Unit Roots in Panels with Fixed Effects, *Mimeo.*

[17] Moon, H.R. and B. Perron (2006): An Empirical Analysis of Nonstationarity in Panels of Exchange Rates and Interest Rates with Factors, *Journal of Applied Econometrics*, forthcoming.

[18] Moon, H.R., B. Perron, and P.C.B Phillips (2006a): A Note on "The Local Power of Some Unit Root Tests For Panel Data" by J. Breitung, *Econometric Theory,* forthcoming.

[19] Moon, H.R., B. Perron, and P.C.B Phillips (2006b): Unpublished Appendix for "Incidental Trends and the Power of Panel Unit Root Tests", *Mimeo.*

[20] H.R. Moon and Phillips, P.C.B. and (1999): Maximum Likelihood Estimation in Panels with Incidental Trends, *Oxford Bulletin of Economics and Statistics*, 61, 771-748.

[21] Moon, H.R. and P.C.B. Phillips (2004): GMM Estimation of Autoregressive Roots Near Unity with Panel Data, *Econometrica*, 72, 467-522.

[22] Newey, W. K. and K. D. West (1994): Automatic Lag Selection in Covariance Matrix Estimation, *Review of Economic Studies*, 61, 631-653.

[23] Phillips, P. C. B. (1987). "Time Series Regression with a Unit Root," *Econometrica,* 55, 277–302.

[24] Phillips, P. C. B. (2001): New Unit Root Asymptotics in the Presence of Deterministic Trends, *Journal of Econometrics*, 11, 323-353.

[25] Phillips, P.C.B. and H.R. Moon (1999): Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica*, 67, 1057–1111.

[26] Phillips, P. C. B. and D. Sul (2003): Dynamic Panel Estimation and Homogeneity Testing under Cross Section Dependence. *Econometrics Journal*, 6, 217-239.

[27] Ploberger, W. and P.C.B. Phillips (2002): Optimal Testing for Unit Roots in Panel Data, *Mimeo.*

[28] Quah, D. (1994): Exploiting Cross-Section Variations for Unit Root Inference in Dynamic Panels, *Economics Letters*, 44, 9–19.

Table 1. .**Size and size-adjusted power of tests - Incidental intercepts case**

DGP: $z_{it} = b_{0i} + z_{it}^0$

$$z_{it}^0 = \left(1 - \frac{\theta_i}{n^{\frac{1}{2}} T}\right) z_{it-1}^0 + \sigma_i e_{it}$$

$$b_{0i}, e_{it} \sim iidN\left(0,1\right)$$

$$\sigma_i \sim iidU\left[0.5, 1.5\right]$$

Theoretical values

	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	5.0	5.0	5.0	5.0	5.0	5.0
$\theta_i \sim U[0,2]$	20.4	17.4	17.4	17.4	12.0	-
$\theta_i \sim U[0,4]$	49.5	40.9	40.9	40.9	24.0	-
$\theta_i \sim U[0,8]$	94.7	88.2	88.2	88.2	59.2	-
$\theta_i \sim \chi^2(1)$	33.7	17.4	17.4	17.4	12.0	-
$\theta_i \sim \chi^2(2)$	63.9	40.9	40.9	40.9	24.0	-
$\theta_i \sim \chi^2(4)$	96.6	88.2	88.2	88.2	59.2	-
$\theta_i = 1$	17.4	17.4	17.4	17.4	12.0	-
$\theta_i = 2$	40.9	40.9	40.9	40.9	24.0	-

$n = 10$, $T = 50$

	$E(\rho_i)$	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	1	-	2.8	5.2	1.9	7.1	5.4
$\theta_i \sim U[0,2]$	.9684	14.0	11.9	11.9	12.0	9.1	8.0
$\theta_i \sim U[0,4]$	.9368	41.0	23.1	23.5	22.9	14.4	9.8
$\theta_i \sim U[0,8]$	.8735	88.9	46.4	48.2	45.6	25.9	14.7
$\theta_i \sim \chi^2(1)$	.9684	15.9	11.2	11.2	11.2	9.1	7.4
$\theta_i \sim \chi^2(2)$	.9368	46.5	20.6	20.9	20.7	13.2	9.5
$\theta_i \sim \chi^2(4)$	.8735	87.8	43.9	45.5	43.1	24.6	15.1
$\theta_i = 1$	.9684	7.9	12.9	12.9	13.1	9.2	7.1
$\theta_i = 2$	.9368	28.5	27.5	27.6	27.5	15.5	10.5

$n = 25$, $T = 50$

	$E(\rho_i)$	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	1	-	3.8	5.5	3.2	8.4	6.4
$\theta_i \sim U[0,2]$	.9817	20.1	13.5	13.7	13.4	10.2	7.9
$\theta_i \sim U[0,4]$	.9635	47.0	27.2	27.8	26.9	16.8	10.7
$\theta_i \sim U[0,8]$	.9270	90.8	58.8	59.7	57.9	32.3	16.9
$\theta_i \sim \chi^2(1)$	.9817	23.4	12.6	12.7	12.4	9.4	7.5
$\theta_i \sim \chi^2(2)$	.9635	55.1	24.6	25.2	24.3	15.1	9.9
$\theta_i \sim \chi^2(4)$	.9270	91.7	56.8	57.8	55.9	32.0	16.8
$\theta_i = 1$	.9817	12.2	15.4	15.2	15.3	10.9	7.8
$\theta_i = 2$	.9635	34.2	32.6	32.6	32.4	18.9	11.3

$n = 100$, $T = 50$

	$E(\rho_i)$	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	1	-	4.4	5.3	4.1	12.9	8.3
$\theta_i \sim U[0,2]$	.99	23.7	14.1	14.2	14.1	10.5	8.0
$\theta_i \sim U[0,4]$	.98	49.4	29.4	29.6	29.4	19.1	11.6
$\theta_i \sim U[0,8]$	.96	91.6	67.7	68.2	67.6	40.1	21.1
$\theta_i \sim \chi^2(1)$	.99	31.7	13.2	13.4	13.2	9.6	7.9
$\theta_i \sim \chi^2(2)$	.98	60.0	27.8	28.1	27.8	17.8	12.0
$\theta_i \sim \chi^2(4)$	.96	93.9	66.8	67.2	66.8	40.9	21.1
$\theta_i = 1$	.99	14.4	15.9	15.8	15.8	10.7	7.8
$\theta_i = 2$	.98	38.6	37.3	37.2	37.4	21.2	12.0

$n = 10$, $T = 100$

	$E(\rho_i)$	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	1	-	2.6	5.1	1.7	6.7	5.1
$\theta_i \sim U[0,2]$	.9968	13.8	13.5	13.8	13.4	9.1	7.6
$\theta_i \sim U[0,4]$	.9937	39.3	23.2	23.9	23.2	13.7	9.3
$\theta_i \sim U[0,8]$	.9874	89.3	48.1	50.4	46.9	24.3	14.1
$\theta_i \sim \chi^2(1)$	.9968	14.4	11.0	11.2	11.0	8.2	7.2
$\theta_i \sim \chi^2(2)$	.9937	44.3	21.1	21.6	20.7	11.8	8.6
$\theta_i \sim \chi^2(4)$	.9874	88.0	47.7	49.7	46.7	24.1	14.9
$\theta_i = 1$	.9968	8.6	14.2	14.3	14.0	9.9	7.9
$\theta_i = 2$	.9937	28.4	27.7	28.0	27.1	16.0	10.4

$n = 25$, $T = 100$

	$E(\rho_i)$	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	1	-	4.0	5.6	3.4	6.7	5.4
$\theta_i \sim U[0,2]$	.9982	19.5	13.7	13.6	13.6	9.9	7.7
$\theta_i \sim U[0,4]$	.9963	45.8	28.5	28.5	28.3	16.6	11.3
$\theta_i \sim U[0,8]$	.9927	91.0	58.9	59.3	58.4	31.0	16.7
$\theta_i \sim \chi^2(1)$	.9982	21.9	12.8	12.9	12.7	9.2	7.7
$\theta_i \sim \chi^2(2)$	.9963	53.5	25.8	25.9	25.5	15.1	10.6
$\theta_i \sim \chi^2(4)$	.9927	91.6	57.9	58.8	57.2	29.8	16.4
$\theta_i = 1$	.9982	12.1	14.7	14.7	14.6	9.6	7.6
$\theta_i = 2$	.9963	33.7	31.9	32.1	31.7	17.4	11.5

$n = 100$, $T = 100$

	$E(\rho_i)$	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	1	-	4.6	5.3	4.3	8.4	6.3
$\theta_i \sim U[0,2]$	.999	22.9	14.7	14.6	14.8	10.8	8.2
$\theta_i \sim U[0,4]$	.998	48.9	31.5	31.4	31.6	19.7	11.4
$\theta_i \sim U[0,8]$	.996	92.8	71.4	71.7	71.5	42.1	20.8
$\theta_i \sim \chi^2(1)$	.999	30.3	13.3	13.2	13.3	10.2	8.2
$\theta_i \sim \chi^2(2)$	.998	59.9	28.7	28.8	28.8	18.6	11.9
$\theta_i \sim \chi^2(4)$	.996	94.1	68.6	68.6	68.5	40.4	20.8
$\theta_i = 1$	.999	14.5	15.5	15.8	15.4	10.2	7.6
$\theta_i = 2$	.998	37.8	36.4	36.7	36.4	19.3	12.0

$n = 10$, $T = 250$

	$E(\rho_i)$	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	1	-	3.2	5.5	2.1	6.2	4.3
$\theta_i \sim U[0,2]$	.9989	12.6	11.7	12.0	11.6	9.4	7.4
$\theta_i \sim U[0,4]$	.9979	37.5	21.8	22.6	21.9	13.3	9.3
$\theta_i \sim U[0,8]$	.9958	88.8	46.4	48.7	45.2	23.5	13.9
$\theta_i \sim \chi^2(1)$	.9989	13.7	10.3	10.6	10.2	8.0	6.8
$\theta_i \sim \chi^2(2)$	.9979	43.9	19.2	20.0	18.8	12.7	8.9
$\theta_i \sim \chi^2(4)$	.9958	87.4	44.3	46.7	43.3	23.7	14.3
$\theta_i = 1$	.9989	8.6	12.9	12.9	12.8	9.5	7.9
$\theta_i = 2$	.9979	28.1	26.0	26.3	25.7	14.4	9.9

$n = 25$, $T = 250$

	$E(\rho_i)$	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	1	-	3.8	5.2	3.2	6.1	4.8
$\theta_i \sim U[0,2]$	.9994	18.7	13.9	14.1	13.9	9.8	7.9
$\theta_i \sim U[0,4]$	.9988	45.1	28.1	28.4	28.0	16.0	10.2
$\theta_i \sim U[0,8]$	.9976	91.1	60.0	60.9	59.6	30.9	16.7
$\theta_i \sim \chi^2(1)$	.9994	21.1	12.2	12.4	12.2	8.5	7.0
$\theta_i \sim \chi^2(2)$	.9988	52.4	25.0	25.6	24.9	15.0	10.7
$\theta_i \sim \chi^2(4)$	.9976	91.5	58.3	59.1	57.9	30.6	17.1
$\theta_i = 1$	.9994	12.0	14.8	14.6	14.9	10.6	8.2
$\theta_i = 2$	.9988	33.6	32.7	32.8	32.9	18.0	11.1

$n = 100$, $T = 250$

	$E(\rho_i)$	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	t^+	IPS
$\theta_i = 0$ (size)	1	-	4.8	5.6	4.5	6.4	5.3
$\theta_i \sim U[0,2]$	.9997	21.6	14.9	14.9	14.8	10.9	7.9
$\theta_i \sim U[0,4]$	.9993	49.7	33.3	33.2	32.9	20.4	12.2
$\theta_i \sim U[0,8]$	.9987	92.3	73.3	73.4	73.1	41.9	20.7
$\theta_i \sim \chi^2(1)$	.9997	30.4	14.5	14.3	14.4	10.5	7.6
$\theta_i \sim \chi^2(2)$	.9993	59.9	30.4	30.5	30.1	18.4	11.3
$\theta_i \sim \chi^2(4)$	.9987	94.4	71.3	71.3	71.0	40.9	20.6
$\theta_i = 1$	.9997	15.2	15.7	15.8	15.7	12.4	8.7
$\theta_i = 2$	.9993	36.8	34.9	35.2	34.9	21.6	12.0

Table 2... **Size and size-adjusted power of tests - Incidental trends case**

DGP: $z_{it} = b_{0i} + b_{1i}t + z_{it}^0$

$$z_{it}^0 = \left(1 - \frac{\theta_i}{n^{\frac{1}{4}}T}\right) z_{it-1}^0 + \sigma_i e_{it}$$

$$b_{0i}, b_{1i}, e_{it} \sim iidN(0,1)$$

$$\sigma_i \sim iidU[0.5, 1.5]$$

Theoretical values

	$c_i = \theta_i$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0	5.0
$\theta_i \sim U[0,2]$	6.5	6.1	6.1	6.1	6.1	5.8	5.8	-	6.0
$\theta_i \sim U[0,4]$	13.3	10.6	10.6	10.6	10.6	9.0	8.6	-	10.0
$\theta_i \sim U[0,8]$	68.7	47.8	47.8	47.8	47.8	33.4	30.1	-	42.3
$\theta_i \sim \chi^2(1)$	18.9	7.8	7.8	7.8	7.8	7.0	6.9	-	7.5
$\theta_i \sim \chi^2(2)$	42.7	14.7	14.7	14.7	14.7	11.7	11.1	-	13.6
$\theta_i \sim \chi^2(4)$	94.7	55.7	55.7	55.7	55.7	39.1	35.2	-	49.5
$\theta_i = 1$	5.8	5.8	5.8	5.8	5.8	5.6	5.6	-	5.7
$\theta_i = 2$	8.9	8.9	8.9	8.9	8.9	7.8	7.6	-	8.5

$n = 10, T = 50$

	$E(\rho_i)$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	1	2.2	0.1	3.2	1.3	1.0	6.1	7.1	6.0
$\theta_i \sim U[0,2]$	.944	5.9	6.0	5.8	5.8	5.8	5.4	5.2	5.9
$\theta_i \sim U[0,4]$	.888	8.3	8.4	8.3	8.3	8.1	7.3	6.2	8.3
$\theta_i \sim U[0,8]$	.775	18.3	18.4	18.2	18.1	15.3	13.3	10.6	16.0
$\theta_i \sim \chi^2(1)$	.944	6.4	6.6	6.4	6.4	6.1	6.2	5.8	6.3
$\theta_i \sim \chi^2(2)$	.888	9.4	9.5	9.3	9.3	8.7	7.7	7.0	8.1
$\theta_i \sim \chi^2(4)$	.775	18.1	18.3	18.0	18.1	15.5	13.5	10.8	15.2
$\theta_i = 1$	.944	5.7	5.6	5.7	5.7	6.0	5.8	5.9	5.9
$\theta_i = 2$	.888	8.3	8.2	8.2	8.2	7.8	7.4	6.9	7.4

$n = 25, T = 50$

	$E(\rho_i)$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	1	5.6	1.8	6.7	2.5	1.3	7.8	9.0	5.0
$\theta_i \sim U[0,2]$	.957	5.3	5.3	5.3	5.3	4.8	4.5	4.8	5.6
$\theta_i \sim U[0,4]$	.915	8.7	8.6	8.7	8.7	7.3	6.0	6.2	7.9
$\theta_i \sim U[0,8]$	.829	22.6	22.6	22.5	22.5	17.7	14.2	11.7	18.8
$\theta_i \sim \chi^2(1)$	.957	6.2	6.1	6.2	6.3	5.7	4.8	5.2	6.7
$\theta_i \sim \chi^2(2)$	.915	9.1	9.0	9.1	9.1	7.9	6.6	6.4	9.2
$\theta_i \sim \chi^2(4)$	.829	22.2	22.3	22.1	22.2	17.4	13.9	11.5	18.5
$\theta_i = 1$	.957	5.6	5.5	5.6	5.6	5.1	5.5	5.0	6.0
$\theta_i = 2$	.915	8.1	8.1	8.1	8.1	6.9	6.9	6.1	7.5

$n = 100,\ T = 50$

	$E(\rho_i)$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	1	12.9	7.9	14.0	3.2	0.1	10.6	12.8	4.2
$\theta_i \sim U[0,2]$	.968	5.4	5.4	5.4	5.4	6.0	6.1	5.3	5.1
$\theta_i \sim U[0,4]$	.937	9.2	9.4	9.3	9.3	8.9	8.7	7.0	7.9
$\theta_i \sim U[0,8]$	.874	29.0	29.3	29.0	29.0	23.6	20.4	13.5	21.7
$\theta_i \sim \chi^2(1)$	.968	7.0	7.2	7.0	7.0	7.2	7.2	5.8	5.6
$\theta_i \sim \chi^2(2)$	.937	10.5	10.6	10.5	10.5	10.1	10.0	8.0	8.8
$\theta_i \sim \chi^2(4)$	.874	27.9	28.3	27.9	27.9	22.6	20.4	13.9	21.4
$\theta_i = 1$	.968	6.4	6.3	6.4	6.4	5.6	5.7	5.4	4.8
$\theta_i = 2$	.937	8.4	8.4	8.4	8.4	7.0	6.8	6.4	7.4

$n = 10,\ T = 100$

	$E(\rho_i)$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	1	1.2	0.1	1.8	1.3	1.5	5.5	5.7	6.2
$\theta_i \sim U[0,2]$	0.994	5.8	5.6	5.7	5.8	5.7	5.4	5.7	6.0
$\theta_i \sim U[0,4]$	0.989	8.6	8.5	8.6	8.6	8.7	7.4	6.6	7.6
$\theta_i \sim U[0,8]$	0.978	19.3	19.3	19.3	19.4	16.6	14.1	11.2	16.1
$\theta_i \sim \chi^2(1)$	0.994	6.7	6.7	6.8	6.8	6.7	6.5	6.4	6.5
$\theta_i \sim \chi^2(2)$	0.989	9.6	9.6	9.6	9.6	8.4	8.1	7.4	8.5
$\theta_i \sim \chi^2(4)$	0.978	18.2	18.1	18.1	18.2	15.7	13.6	11.2	16.0
$\theta_i = 1$	0.994	5.3	5.4	5.3	5.4	5.1	5.3	5.3	5.6
$\theta_i = 2$	0.989	6.9	6.8	6.9	6.9	6.5	6.4	6.9	7.0

$n = 25,\ T = 100$

	$E(\rho_i)$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	1	3.6	1.0	4.6	2.7	2.1	6.0	6.2	5.7
$\theta_i \sim U[0,2]$	.996	5.7	5.7	5.7	5.7	5.7	5.6	5.6	5.8
$\theta_i \sim U[0,4]$	.992	8.8	8.8	8.7	8.7	8.5	7.7	7.1	7.9
$\theta_i \sim U[0,8]$	.983	22.7	22.7	22.7	22.6	18.4	16.4	12.9	18.4
$\theta_i \sim \chi^2(1)$	.996	6.8	6.8	6.7	6.7	6.7	6.7	6.4	6.0
$\theta_i \sim \chi^2(2)$	.992	9.5	9.4	9.4	9.3	8.3	8.4	7.6	8.1
$\theta_i \sim \chi^2(4)$	.983	21.2	21.2	21.1	21.1	17.5	15.7	12.6	18.0
$\theta_i = 1$	.996	5.9	6.0	5.8	5.8	5.9	5.5	5.7	5.4
$\theta_i = 2$	.992	7.4	7.5	7.3	7.3	7.2	6.8	6.2	7.8

$n = 100,\ T = 100$

	$E(\rho_i)$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	1	7.1	3.5	7.9	3.4	1.6	8.0	8.6	4.7
$\theta_i \sim U[0,2]$	.997	6.2	6.3	6.2	6.2	5.9	6.0	5.2	5.6
$\theta_i \sim U[0,4]$	.994	10.2	10.4	10.3	10.3	8.7	8.7	7.4	8.4
$\theta_i \sim U[0,8]$	.988	28.8	29.1	28.8	28.8	21.6	19.6	13.7	23.2
$\theta_i \sim \chi^2(1)$	.997	7.1	7.1	7.1	7.1	6.3	6.4	6.1	6.5
$\theta_i \sim \chi^2(2)$	.994	10.7	10.8	10.7	10.7	9.3	9.0	7.4	9.6
$\theta_i \sim \chi^2(4)$	.987	30.0	30.4	30.0	30.1	22.2	20.1	14.3	23.1
$\theta_i = 1$	.997	5.9	6.0	5.9	5.9	6.2	6.0	5.2	5.6
$\theta_i = 2$	.994	9.2	9.3	9.2	9.3	8.0	7.2	6.2	7.7

$n = 10$, $T = 250$

	$E(\rho_i)$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	1	1.2	0.0	2.0	1.8	2.5	6.0	5.2	6.2
$\theta_i \sim U[0,2]$	.998	5.1	5.1	5.0	5.0	5.1	5.2	5.4	6.1
$\theta_i \sim U[0,4]$	.996	7.8	7.8	7.9	7.9	6.6	6.2	6.0	7.5
$\theta_i \sim U[0,8]$	.993	18.1	18.4	18.2	18.2	14.4	12.6	9.8	16.6
$\theta_i \sim \chi^2(1)$	.998	6.3	6.4	6.3	6.3	6.0	5.8	5.8	6.3
$\theta_i \sim \chi^2(2)$	.996	9.1	9.1	9.2	9.2	7.4	7.0	6.7	8.3
$\theta_i \sim \chi^2(4)$	.993	17.2	17.2	17.2	17.2	13.9	12.1	10.2	15.9
$\theta_i = 1$	.998	5.8	5.7	5.8	5.8	5.7	5.7	5.2	5.2
$\theta_i = 2$	.996	7.2	7.2	7.2	7.3	7.0	6.6	5.9	7.8

$n = 25$, $T = 250$

	$E(\rho_i)$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	1	2.6	0.6	3.2	2.8	2.7	5.4	5.2	5.8
$\theta_i \sim U[0,2]$	.999	6.6	6.5	6.5	6.5	6.2	6.1	5.8	5.4
$\theta_i \sim U[0,4]$	.997	8.9	8.8	8.8	8.8	8.2	7.8	7.1	7.4
$\theta_i \sim U[0,8]$	.994	23.1	23.2	22.9	22.9	19.1	16.1	12.5	19.0
$\theta_i \sim \chi^2(1)$	.999	6.6	6.5	6.5	6.5	6.3	6.0	5.9	6.2
$\theta_i \sim \chi^2(2)$	.997	9.4	9.5	9.3	9.3	9.0	8.5	7.3	8.6
$\theta_i \sim \chi^2(4)$	.994	21.5	21.4	21.4	21.3	17.4	15.0	12.4	18.9
$\theta_i = 1$	.999	5.4	5.4	5.3	5.3	5.4	5.3	5.7	5.6
$\theta_i = 2$	.997	7.4	7.3	7.4	7.4	6.8	6.5	6.5	7.4

$n = 100$, $T = 250$

	$E(\rho_i)$	$c_i = 1$	$c_i = 2$	$c_i = 0.5$	Ploberger-Phillips	Moon-Phillips	t^+	IPS	UB
$\theta_i = 0$ (size)	1	4.7	2.2	5.4	3.9	3.3	6.6	6.2	5.2
$\theta_i \sim U[0,2]$	.999	5.9	5.9	5.8	5.8	5.2	5.2	5.5	5.7
$\theta_i \sim U[0,4]$	.998	9.2	9.2	9.1	9.2	8.3	7.5	7.3	8.5
$\theta_i \sim U[0,8]$	.996	29.6	29.8	29.6	29.6	21.9	18.6	14.3	23.5
$\theta_i \sim \chi^2(1)$	.999	6.6	6.6	6.5	6.5	6.2	5.8	5.4	6.6
$\theta_i \sim \chi^2(2)$	.998	10.4	10.5	10.4	10.4	9.2	8.3	7.7	9.5
$\theta_i \sim \chi^2(4)$	.996	27.4	27.5	27.4	27.4	20.8	17.8	14.5	24.1
$\theta_i = 1$	.999	6.0	5.9	5.9	5.9	5.8	5.9	5.5	5.4
$\theta_i = 2$	.998	9.1	9.0	9.1	9.0	8.1	8.3	7.2	7.7