# Unit root tests for dependent micropanels*

In Choi†

First Draft: January 2014, Revised: November 2016

**Abstract**

This paper proposes a new test for the null hypothesis of panel unit roots for micropanels with short time dimensions ($T$) and large cross sections ($N$). There are several distinctive features of this test. First, the test is based on a panel AR(1) model allowing for cross-sectional dependency, which is introduced by a factor structure of the initial condition. Second, the test employs the panel AR(1) model with AR(1) coefficients that are heterogeneous for finite $N$. Third, the test can be used both for the alternative hypothesis of stationarity and for that of explosive roots. Fourth, the test does not use the AR(1) coefficient estimator. The effectiveness of the test rests on the fact that the initial condition has permanent effects on the trajectory of a time series in the presence of a unit root. To measure the effects of the initial condition, this paper employs cross-sectional regressions using the first time series observations as a regressor and the last as a dependent variable. If there is a unit root in every individual time series, the coefficient of the regressor is equal to one. The $t$-ratios for the coefficient

*This paper was written while I was visiting the Department of Economics, University of Texas, Austin. I thank Jason Abrevaya, Sukjin Han, Haiquing Xu and Stephen Douglas for their hospitality. I also thank participants of the Econometrics Lunch Seminar of Syracuse University and the 2014 Hitotsubashi-Sogang conference on Econometrics for their insightful comments, in particular, Badi Baltagi, Chihwa Kao, Yoonseok Lee, Ryo Okui, Jan Ondrich and Gary Engelhardt. My thanks also go to Mototsugu Shintani and two anonymous referees for insightful comments. Matlab codes for the empirical section of this paper can be downloaded from my personal Web page http:\\inchoi.sogang.ac.kr.

†Sogang University, Seoul, Korea. Personal Webpage: http://inchoi.sogang.ac.kr. E-mail: inchoi@gmail.com, inchoi@sogang.ac.kr.

are this paper's test statistics and have a standard normal distribution in the limit. The $t$-ratios are based on the ordinary least squares estimator and the instrumental variables estimator that uses reshuffled regressors as instruments. The test proposed in this paper makes it possible to test for a unit root even at $T = 2$ as long as $N$ is large. Simulation results show that test statistics have reasonable empirical size and power. The test is applied to college graduates' monthly real wage in South Korea. The number of time series observations for this data is only 2. The null hypothesis of a unit root is rejected against the alternative of stationarity.

Keywords: Unit root, panel data, factor model, internal instrument, earnings dynamics

# 1 Introduction

In recent years, there has been much interest in testing for a unit root using panel data. Nowadays, it is quite common in empirical applications to use panel data rather than individual time series for the purpose of testing for a unit root. Many panel unit root tests are now programmed in commercial software so they are now widely available. Panel unit root tests that have often been used in applications (e.g., Choi, 2001; Im, Pesaran and Shin, 2003; Levin, Lin and Chu, 2002; Maddala and Wu, 1999; etc.) assume a large time dimension ($T$) and a large cross section ($N$), making them more appropriate for applications to macropanels than to micropanels. Extensions of these tests to cross-sectionally correlated panels have also been made (e.g., Bai and Ng, 2004; Breitung and Das, 2005; Choi and Chue, 2007; Demetrescu, Hassler and Tarcolea, 2006; Moon and Perron, 2004; Phillips and Sul, 2003; Pesaran, 2007; Sheng and Yang, 2013; etc.). More discussions and references related to panel unit root tests can be found in Choi (2006, 2015) and Breitung and Pesaran (2008).

There are several panel unit root tests designed for short $T$ and large $N$. Breitung and Meyer (1994), De Blander and Dhaene (2012), De Wachter, Harris and Tzavalis (2007), Harris and Tzavalis (1999, 2004), Karavias and Tzavalis (2014) and Kruiniger (2008, 2009) belong to this category. These tests assume homogeneous AR(1) coefficients and cross-sectional independence of observations, and are more suitable for micropanels. In addition, there are tests designed for cross-sectionally dependent panels with short $T$. Robertson, Sarafidis and Westerlund (2014) employ the panel AR(1) model with errors having a factor structure and use GMM for inference. Their test requires a trend variable in the model and cannot be used for models with only fixed effects. Karavias and Tzavalis (2014) assume a spatial dependence structure for errors to deal with cross-sectional dependency of micro units. But the variance-covariance matrix of the limiting distribution of their estimator cannot be estimated consistently, which requires using bootstrapping.

This paper proposes a panel unit root test for micropanels with short $T$ and large $N$. There are several distinctive features of this test. First, the test is based on a panel AR(1) model allowing for cross-sectional dependency, which is introduced by a factor structure of the initial condition. Our panel data model is new in the literature and should be useful for some applications.

Second, our test uses a panel AR(1) model with AR(1) coefficients that are heterogeneous for finite $N$. All the existing tests based on short $T$ assume homogeneous AR(1) coefficients. Although those tests must have nonnegligible power even when the AR(1) coefficients are heterogeneous, assuming homogeneous AR(1) coefficients seems to be a conceptual drawback.

Third, the test of this paper can be used both for the alternative hypothesis of stationarity and for that of explosive roots. The alternative hypothesis of stationarity is the main concern of many studies. But that of explosive roots is also potentially important because it is related to testing for the presence of asset bubble as discussed in Phillips, Wu and Yu (2011) and Homm and Breitung (2012)

Fourth, this paper's test does not use the AR(1) coefficient estimator.[1] Instead,

[1] There are a few papers in the time series literature that do not use the AR coefficient estimators for unit root testing (see, for example, Breitung, 2002; Breitung and Gouriéroux, 1997; Burridge and Guerre, 1996; Cavaliere, 2001; Hasan and Koenker, 1997; Hallin, Van den Akker and Werker, 2011; Nielsen, 2009). See Chapter 4 of Choi (2015) for detailed discussions on these papers. But the ideas of these papers have not yet been extended to panel data. Moreover, none of these papers use the idea of this paper.

it rests on the fact that the initial condition has permanent effects on the trajectory of a time series in the presence of a unit root. To measure the effects of the initial condition, this paper employs cross-sectional regressions using the first time series observations as a regressor and the last as a dependent variable. If there is a unit root in every individual time series, the coefficient of the regressor is equal to 1. The $t$-ratios for the coefficient are this paper's test statistics and have a standard normal distribution in the limit. The $t$-ratios are based on the OLS estimator and the instrumental-variables (IV) estimator that uses reshuffled regressors as instruments. We will call the IV estimator the "internal IV estimator" since the instruments originate from the given sample. If the AR(1) coefficients are less than 1 in absolute value or greater than 1, the $t$-ratios diverge to minus or plus infinity in probability, making the test consistent. The regressions we will use make it possible to test for a unit root even when $T$ is very small. In fact, this paper's test can be implemented as long as $T \geq 2$. By contrast, all the extant panel unit root tests require at least $T \geq 3$ in the case of the panel AR(1) model with individual level effects, and $T \geq 4$ in the case of the panel AR(1) model with individual level and trend effects. Panel unit root tests for very small $T$ are useful now and in the future as well: New kinds of panel data are constantly being collected throughout the world and these will not have large $T$ at least for some years. This paper's test can be applied to those very short panel data sets as long as $N$ is large enough.

It will be shown that this paper's particular instruments become valid due to the assumption on the initial variable. Thus, the assumption is crucial for the consistency of the test as well as for the cross-sectional dependency of the panel data model. Whether the assumption is appropriate or not is an empirical matter that should be examined for each panel data set. In addition, as in other IV regressions, choosing the number of instruments is an important issue for the internal IV regression. Because the number of available instruments is greater than the sample size for the internal IV regression, the issue is quite complex and existing methods (e.g., Andrews, 1999; and Donald and Newey,2001) cannot be used. This paper does not deal with this issue, leaving it to future work.

As an example, this paper's test is applied to the monthly real wage of college graduates in South Korea hired in 2006 for the first time. The number of time series observations for this data is only 2, whereas that of cross-sectional observations is large. The test rejects the null hypothesis of a unit root at conventional significance levels, rendering support to the similar evidence obtained by Guvenen (2009) and Okubo (2015), who use US and Japanese data, respectively.

This paper is planned as follows. Section 2 introduces the model, basic assumptions and hypotheses. Section 3 introduces this paper's OLS and internal IV regressions and studies asymptotic properties of the test statistics stemming from the OLS and IV regressions. Section 4 reports simulation results. Section 5 contains an empirical application of our test. Section 6 provides summary and further remarks. Appendix I contains technical assumptions and Appendix II proofs.

A few words on notation: (i) All the limits are taken as $N \to \infty$. (ii) The sample mean of $\{a_i : i = 1, ..., N\}$ is denoted as $\bar{a}$. (iii) Definitional relations are denoted by $:=$ .

## 2 The model, basic assumptions and hypotheses

Consider the unobserved components models for the panel data $\{y_{it}\}$

$$y_{it} = \mu_i + x_{it}, \ (i = 1, \ldots, N; \ t = 1, \ldots, T), \tag{1}$$

and

$$y_{it} = \mu_i + \beta_i t + x_{it}, \ (i = 1, \ldots, N; \ t = 1, \ldots, T), \tag{2}$$

where $\{x_{it}\}$ follows the AR(1) model given by[2]

$$x_{it} = \alpha_i x_{i,t-1} + u_{it}.$$

As usual, $i$ and $t$ are indices for individuals and time, respectively, and $\{\mu_i\}$ and $\{\beta_i\}$ denote the unobserved individual level and trend effects, respectively. The individual

[2] Instead, one may wish to assume $x_{it} = \alpha_i x_{i,t-1} + u_{it} + \lambda_i' f_t$ $(t = 2, ..., T)$. If $f_1$ is independent of $\{f_t\}_{t=2,...,T}$, all the theory of this paper works with changes in notation.

trend effects, $\{\beta_i\}$, indicate how fast $\{y_{it}\}$ grows. In these model specifications, $\{x_{it}\}$ is not observed and brings dynamics to the evolvement of $\{y_{it}\}$. Models (1) and (2) have often been used for unit root testing. See, for example, Schmidt and Phillips (1992) and Elliott, Rothenberg and Stock (1996) for earlier references.

Assumption 1 lays out the basic characteristics of the individual effects $\{\mu_i\}$ and $\{\beta_i\}$ and the error terms $\{u_{it}\}$.

**Assumption 1** *(i)* $\mu_i = \mu + m_i$, $m_i \sim \mathrm{i.i.d.}(0, \sigma^2_{\mathrm{m}})$ *and* $\sigma^2_{\mathrm{m}} > 0$*;*

*(ii)* $\beta_i = \beta + b_i$, $b_i \sim \mathrm{i.i.d.}(0, \sigma^2_{\mathrm{b}})$ *and* $\sigma^2_{\mathrm{b}} > 0$*;*

*(iii)* $\{u_{it}\}$ *are independent with* $E(u_{it}) = 0$, $E(u^2_{it}) = \sigma^2_{\mathrm{u}_i} > 0$ *for every* $i$ *and* $t$*;*

*(iv)* $\{m_i\}, \{b_i\}$ *and* $\{u_{it}\}$ *are independent.*

Assumption 1 is of standard nature. Part (iii) implies heteroskedasticity of $\{u_{it}\}$, which will be dealt with as in White (1980).

Additionally, we assume a factor structure for the initial variables $\{x_{i1}\}$ as follows.

**Assumption 2** *(i) For every* $i$, $x_{i1} = \lambda_i' f_1$, *where* $\lambda_i$ *and* $f_1$ *are the vectors of unobserved factor loadings and factors, respectively, and* $\{\lambda_i\}$ *is a sequence of constant vectors.*

*(ii)* $E(f_1) = 0$ *and* $E(f_1 f_1') = \Sigma_f > 0$*.*

*(iii)* $f_1$ *is independent of* $\{m_i\}, \{b_i\}$ *and* $\{u_{it}\}$*.*

According to Assumption 2 (i), the initial variable $y_{i1}$ is affected by the individual effects and the random variable $\lambda_i' f_1$, which differs across individuals. In Assumption 2 (i), $\{\lambda_i\}$ is a sequence of constant vectors. Because assuming a random sequence for $\{\lambda_i\}$ entails no analytical difficulties and because it brings simplicity, we will keep that assumption. Assumption 2 (i) also introduces nonzero cross-sectional correlations of $\{y_{it}\}$. The exact nature of this depends on the locations of $\{\alpha_i\}$ and will be discussed in the next section. Parts (ii) and (iii) of Assumption 2 are of standard nature. Because the AR(1) coefficients, error terms and initial conditions are heterogeneous across individuals, so are the dynamics of $\{y_{it}\}$ even though $\{x_{i1}\}$ depends on $f_1$

across individuals under Assumption 2. To put it differently, $\{y_{it}\}$ evolves following a different path for each individual under Assumption 2.

We now use an example to motivate Model (2). Suppose $y_{i1}$ is the initial wage of individual $i$, a new job entrant. The initial wage is usually a function of individual $i$'s characteristics such as educational level, type of industry she works for, intelligence quotient (IQ), emotional quotient (EQ), unobserved ability in workplaces that are represented by $\mu_i$. In addition, the wage must be affected by economy-wide variables in that particular year. These variables are, for example, the level and growth rate of GDP in that year, the size of new job seeker pool, the level of per capita capital stock, the degree of technological progress and the state of the world economy. These variables are represented by $f_1$. Since each individual responds to these economy-wide variables differently, the impact of the economy-wide variables can be written as $\lambda_i' f_1$, where $\lambda_i$ denotes individual $i$'s response to the economy-wide variables. As the new job entrant accumulates her work experience, her wage tends to grow and the rate of growth depends once again on her individual characteristics represented by $\beta_i$. In particular, if log-wages are used, $\beta_i$ is the growth rate of individual $i$'s wage. In light of these discussions, modelling individual wages as Model (2) seems reasonable. Model (1) is more specialized than Model (2) and is suitable for panel data without trends.

The null hypothesis we consider is

$$H_0 : \alpha_i = 1 \text{ for all } i, \tag{3}$$

and the alternative hypothesis is

$$H_A : \alpha_i = \alpha + \frac{\eta_i}{N^{1/2+\varepsilon}}, \tag{4}$$

where $|\alpha| < 1$ or $\alpha > 1$, $\varepsilon > 0$ and $\{\eta_i\}$ is a sequence of finite constants. Under the null hypothesis, every individual has a unit root. Under the alternative hypothesis, every individual has the AR coefficient less than 1 in absolute value or greater than 1 for large $N$. In finite samples, the AR(1) coefficients are heterogenous. Implications of the null and alternative hypotheses on the properties of $\{y_{it}\}$ will further be discussed in the next section. The alternative hypothesis (4) does not represent the usual local alternatives because $\alpha_i$ converges to $\alpha$ faster than in the usual local alternatives. But it serves the purpose of this paper well.

# 3 Test statistics and asymptotic properties

This section introduces tests statistics for the null and alternative hypotheses (3) and (4), and reports their asymptotic properties.

Model (1) can be written as

$$y_{it} = \mu_i + \alpha_i^{t-1} x_{i1} + w_{it}, \ (i = 1, \ldots, N; \ t = 2, \ldots, T), \tag{5}$$

where $w_{it} := \sum_{j=0}^{t-2} \alpha_i^j u_{i,T-j}$. Since $y_{i1} = \mu_i + x_{i1}$, individual $i$'s last data under the null hypothesis (3) can be represented as

$$\begin{aligned} y_{iT} &= \mu_i + x_{i1} + s_{iT} \\ &= y_{i1} + s_{iT}, \end{aligned} \tag{6}$$

where $s_{iT} := \sum_{j=2}^{T} u_{ij}$. The initial variable $y_{i1}$ and the shocks $\{u_{it}\}_{t=2,\ldots,T}$ have permanent effects on $y_{iT}$ under the null hypothesis. Moreover, $\mathrm{Cov}(y_{iT}, y_{jT}) = \lambda_i' \Sigma_f \lambda_j$, implying cross-sectional covariances of $\{y_{iT}\}$ that do not change with $T$. Since $\mathrm{Var}(y_{iT}) = \sigma_{\mathrm{m}}^2 + \lambda_i' \Sigma_f \lambda_i + (T-1)\sigma_{\mathrm{u}_i}^2$, the variance of $y_{it}$ grows with $T$.

Now consider the alternative hypothesis (4) with $|\alpha| < 1$. Relation (5) gives

$$\begin{aligned} y_{iT} &= \mu_i + \alpha_i^{T-1} x_{i1} + w_{iT} \\ &= \mu_i (1 - \alpha_i^{T-1}) + \alpha_i^{T-1} y_{i1} + w_{iT}, \end{aligned} \tag{7}$$

where $w_{iT} := \sum_{j=0}^{T-2} \alpha_i^j u_{i,T-j}$. In contrast to the behavior of $\{y_{iT}\}$ under the null hypothesis, effects of the initial variable, $y_{i1}$, and the shocks, $\{u_{it}\}_{t=2,\ldots,T}$, on $y_{iT}$ are weakened as $y_{iT}$ progresses into the future. Likewise, cross-sectional covariances diminish for large $N$ as $T$ grows since $\mathrm{Cov}(y_{iT}, y_{jT}) = \alpha_i^{T-1} \alpha_j^{T-1} \lambda_i' \Sigma_f \lambda_j$. The variance of $y_{iT}$ is represented by $\sigma_{\mathrm{m}}^2 + \alpha_i^{2(T-1)} \lambda_i' \Sigma_f \lambda_i + \sum_{j=0}^{T-2} \alpha^{2j} \sigma_{\mathrm{u}_j}^2$, which does not grow in proportion to $T$ and is smaller than that under the null hypothesis for large $N$.

Combining data-generating processes (6) and (7), the regression model that can be used to test the null hypothesis (3) is written as

$$y_{iT} = \phi_\mu + \phi_1 y_{i1} + v_{iT},\ (i = 1, \ldots, N), \tag{8}$$

where $\{v_{iT}\}$ denotes the regression errors. The null hypothesis (3) is equivalent to $H_0 : \phi_1 = 1$.

Unlike the conventional dynamic panel data models, Model (8) uses only the cross-sectional variations of the data at periods 1 and $T$ to estimate the coefficients $\phi_\mu$ and $\phi_1$. Let $\tilde{\phi}_\mu$ and $\tilde{\phi}_1$ be the OLS estimators of $\phi_\mu$ and $\phi_1$, respectively. The heteroskedasticity-robust $t$-ratio for the null hypothesis (3) using $\tilde{\phi}_1$ is defined as

$$t_{\tilde{\phi}_1} = \frac{\tilde{\phi}_1 - 1}{\sqrt{\left(\sum_{i=1}^N (y_{i1} - \bar{y}_1)^2\right)^{-2} \sum_{i=1}^N (y_{i1} - \bar{y}_1)^2 \tilde{v}_{iT}^2}},$$

where $\{\tilde{v}_{iT}\}$ denotes the regression residuals using the OLS estimators. Properties of the OLS estimator and the t-ratio are reported in the following theorem.

**Theorem 1** *Assume that Assumptions 1–5 hold.*

*(i) Under the null hypothesis (3),*

*(a)* $\tilde{\phi}_\mu = O_p\left(\frac{1}{\sqrt{N}}\right)$ *and* $\tilde{\phi}_1 = 1 + O_p\left(\frac{1}{\sqrt{N}}\right)$;

*(b)* $t_{\tilde{\phi}_1} \xrightarrow{d} N(0, 1)$

*(ii) Under the alternative hypothesis (4),*

*(a)* $\tilde{\phi}_1 \xrightarrow{p} \frac{\sigma_{\mathrm{m}}^2 + \alpha^{T-1}\left(f_1' M_\lambda f_1 - f_1' q_\lambda q_\lambda' f_1\right)}{\sigma_{\mathrm{m}}^2 + f_1' M_\lambda f_1 - f_1' q_\lambda q_\lambda' f_1}$, *where* $M_\lambda$ *and* $q_\lambda$ *are defined in Assumption 3;*

*(b)* $t_{\tilde{\phi}_1} \xrightarrow{p} -\infty$ *if* $|\alpha| < 1$ *and* $t_{\tilde{\phi}_1} \xrightarrow{p} \infty$ *if* $\alpha > 1$.

Under the null hypothesis, the regressor is uncorrelated with the error term and the OLS-based $t$-ratio has a standard normal distribution in the limit. However, under the alternative hypothesis, the regressor is correlated with the error term since $m_i$ is present in both $y_{i1}$ and $v_{iT}$. Nonetheless, the OLS-based t-ratio is consistent because

$\tilde{\phi}_1$ converges to a random variable that is less (greater) than 1 almost surely if $|\alpha| < 1$ ($\alpha > 1$).

But because

$$\frac{\sigma_{\mathrm{m}}^2 + \alpha^{T-1}\left(f_1' M_\lambda f_1 - f_1' q_\lambda q_\lambda' f_1\right)}{\sigma_{\mathrm{m}}^2 + f_1' M_\lambda f_1 - f_1' q_\lambda q_\lambda' f_1} = \alpha^{T-1} + \frac{\sigma_{\mathrm{m}}^2 (1 - \alpha^{T-1})}{\sigma_{\mathrm{m}}^2 + f_1' M_\lambda f_1 - f_1' q_\lambda q_\lambda' f_1},$$

the OLS estimator $\tilde{\phi}_1$ has a positive asymptotic bias that has negative impact on the finite sample power of the test. The asymptotic bias is caused by the regressor-error dependence under the alternative hypothesis. The farther away the value of $\alpha$ is from 1, the large the bias becomes. Moreover, the bias is an increasing (decreasing) function of $\sigma_{\mathrm{m}}^2$ when $|\alpha| < 1$ ($\alpha > 1$). Thus, the finite-sample power of the OLS-based t-test decreases as $\sigma_{\mathrm{m}}^2$ increases at fixed $\alpha$ as one of the referees indicated. In spite of the asymptotic bias, the OLS-based t-ratio works well in finite samples as shown in Section 4.

In addition to the OLS estimator, we consider an IV estimator using reshuffles of the regressor as instruments. We call the IV estimator the "internal IV estimator" because the instruments originate from given sample. Under Assumption 1, $\mathrm{Cov}(v_{iT}, y_{j1}) = 0$ for every $i$ ($\neq j$). Moreover, if $\mathrm{Cov}(y_{i1}, y_{j1}) = \lambda_i' \Sigma_f \lambda_j \neq 0$ for enough number of individuals, the internal IV becomes valid. More precise conditions are given in Assumption 4. In a simple case where there is only 1 instrument and $\{\lambda_i\}$ is a sequence of scalars, the essential condition for the validity of the internal instrument is

$$\frac{1}{N}\sum_{i=1}^{N} y_{i1} z_i = f_1^2 \frac{1}{N}\sum_{i=1}^{N} \lambda_i^* \lambda_i + o_p(1) \neq 0 \text{ almost surely,} \tag{9}$$

where $\{z_1, ..., z_N\}$ and $\{\lambda_1^*, ..., \lambda_N^*\}$ are reshuffles of $\{y_{11}, ..., y_{N1}\}$ and $\{\lambda_1, ..., \lambda_N\}$, respectively. Using the example in Section 2, condition (9) is tantamount to requiring that individual reactions to the economy-wide variables are somehow related, which seems reasonable.

The internal IV is defined formally as follows:

**Definition 1** *Let $\{z_{k1}, ..., z_{kN}\}$ be the k-th reshuffle (or permutation ) of the initial observations $\{y_{11}, ..., y_{N1}\}$ ($k = 1, ..., K$) that is not equal to $\{y_{11}, ..., y_{N1}\}$. Assume* *that the reshuffles are all distinct. The internal IV for* $y_{i1}$ *is* $z_i := [z_{1i}, ..., z_{Ki}]'$.

This definition means that there are $K$ internal instruments. Each instrument is different from the regressor, and each value of the regressor is used only once to construct an instrument. The latter is not strictly required, but it is assumed for analytical simplicity. In addition, instruments are all different. Under Definition 1, there are $N!-1$ instrument variables. In practice, a set of instruments that minimizes the standard error of the IV estimator of $\phi_1$ can be used. This will require a massive amount of calculation when $N$ is large.

The IV estimators of the coefficients $\phi_\mu$ and $\phi_1$ are defined as

$$\begin{pmatrix} \hat{\phi}_\mu \\ \hat{\phi}_1 \end{pmatrix} = (Y_1' P_Z Y_1)^{-1} Y_1' P_Z y_T,$$

where $P_Z := Z(Z'Z)^{-1}Z'$, $Z := \begin{bmatrix} 1 & z_1' \\ \vdots & \vdots \\ 1 & z_N' \end{bmatrix}$ and $Y_1 := \begin{bmatrix} 1 & y_{11} \\ \vdots & \vdots \\ 1 & y_{11} \end{bmatrix}$. The heteroskedasticity-robust $t$-ratio for the null hypothesis (3) is defined as

$$t_{\hat{\phi}_1} = \frac{\hat{\phi}_1 - 1}{\sqrt{(y_1^{*\prime} P_{Z^*} y_1^*)^{-2} y_1^{*\prime} Z^* (Z^{*\prime} Z^*)^{-1} \left( Z^{*\prime} diag(\hat{v}_{1T}^2, ..., \hat{v}_{NT}^2) Z^* \right) (Z^{*\prime} Z^*)^{-1} Z^{*\prime} y_1^*}},$$

where $y_1^* := [y_{11} - \bar{y}_1, ..., y_{N1} - \bar{y}_1]'$, $Z^*$ is similarly defined, $\{\hat{v}_{iT}\}$ denotes the regression residuals using the IV estimators $\hat{\phi}_\mu$ and $\hat{\phi}_1$, and $diag(\hat{v}_{1T}^2, ..., \hat{v}_{NT}^2)$ is a diagonal matrix having $\hat{v}_{1T}^2, ..., \hat{v}_{NT}^2$ as its diagonal elements.

Theorem 2 reports limiting properties of $\hat{\phi}_1$ and $t_{\hat{\phi}_1}$ under both the null and alternative hypotheses (3) and (4).

**Theorem 2** *Assume that Assumptions 1–5 hold.*

*(i) Under the null hypothesis (3),*

*(a)* $\hat{\phi}_\mu = O_p\left(\frac{1}{\sqrt{N}}\right)$ *and* $\hat{\phi}_1 = 1 + O_p\left(\frac{1}{\sqrt{N}}\right)$;

*(b)* $t_{\hat{\phi}_1} \xrightarrow{d} N(0,1)$.

*(ii) Under the alternative hypothesis (4),*

(a) $\hat{\phi}_\mu = \mu\left(1-\alpha^{T-1}\right) + O_p\left(\frac{1}{\sqrt{N}}\right)$ and $\hat{\phi}_1 = \alpha^{T-1} + O_p\left(\frac{1}{\sqrt{N}}\right)$;

(b) $t_{\hat{\phi}_1} \xrightarrow{p} -\infty$ if $|\alpha| < 1$ and $t_{\hat{\phi}_1} \xrightarrow{p} \infty$ if $\alpha > 1$.

Part (i) of Theorem 2 shows that the internal IV estimators are $\sqrt{N}$-consistent under the null hypothesis and that the $t$-ratio has a standard normal distribution in the limit. The latter result does not require using the functional central limit theorem unlike extant panel unit root tests like Levin, Lin and Chu (2002), Im, Pesaran and Shin (2003), and Choi (2001). This is because only the cross-sectional variations of the data at periods 1 and $T$ are used for the IV estimators. Part (ii) of Theorem 2 shows that the IV estimators $\hat{\phi}_\mu$ and $\hat{\phi}_1$ are $\sqrt{N}$-consistent for $\mu(1-\alpha^{T-1})$ and $\alpha^{T-1}$, respectively, and that the $t$-ratio diverges to $-\infty$ or $\infty$ in probability under the alternative, making our unit root test consistent. The former result is of independent interest, since it shows that the common asymptotic value of the AR(1) coefficients $\{\alpha_i\}$ can be estimated using the internal IV estimator even in the presence of individual effects and cross-sectional dependency. Confidence intervals of $\alpha$ can also be formulated regardless of its value. In the univariate case, constructions of the confidence intervals are not straightforward and various methods have been proposed (see Choi, 2015).

Now, we extend the discussions so far to Model (2). The model is rewritten as

$$y_{it} = \mu_i + \beta_i t + \alpha_i^{t-1} x_{i1} + w_{it},\ (i = 1, \ldots, N;\ t = 2, \ldots, T), \tag{10}$$

which gives the null representation of $\{y_{iT}\}$ as

$$\begin{aligned} y_{iT} &= \mu_i + \beta_i T + x_{i1} + s_{iT} \\ &= \beta_i(T-1) + y_{i1} + s_{iT}, \end{aligned} \tag{11}$$

since $y_{i1} = \mu_i + \beta_i + x_{i1}$. If the observations $\{y_{it}\}$ are logarithmic, the parameter $\beta_i$ is the average growth rate of $y_{it}$ over the $T-1$ periods since $E\left((y_{iT} - y_{i1})/(T-1)\right) = \beta_i$. The initial variable, $y_{i1}$, and the shocks, $\{u_{it}\}_{t=2,\ldots,T}$, have permanent effects on $y_{iT}$ as for the data-generating process (6). Cross-sectional covariances of $\{y_{iT}\}$ also behave in the same manner as for the data-generating process (6). The variance of $y_{iT}$ grows with $T$ since $\mathrm{Var}(y_{iT}) = \sigma^2_{\mathrm{m}} + T^2\sigma^2_{\mathrm{b}} + \lambda_i'\Sigma_f\lambda_i + (T-1)\sigma^2_{\mathrm{u}_i}$.

Under the alternative hypothesis (4) under $|\alpha| < 1$, relation (10) gives

$$\begin{aligned} y_{iT} &= \mu_i + \beta_i T + \alpha_i^{T-1} x_{i1} + w_{iT} \\ &= \mu_i \left(1 - \alpha_i^{T-1}\right) + \beta_i \left(T - \alpha_i^{T-1}\right) + \alpha_i^{T-1} y_{i1} + w_{iT}. \end{aligned} \tag{12}$$

As for the data-generating process (7), the initial variable, $y_{i1}$, and the shocks, $\{u_{it}\}_{t=2,\ldots,T}$, have weakening effects on $y_{iT}$ as $T$ grows, while cross-sectional covariances of $\{y_{iT}\}$ diminish for large $N$ as $T$ grows. The variance of $y_{iT}$ is $\sigma_{\mathrm{m}}^2 + T^2\sigma_{\mathrm{b}}^2 + \alpha_i^{2(T-1)}\lambda_i'\Sigma_f\lambda_i + \sum_{j=0}^{T-2}\alpha_i^{2j}\sigma_{\mathrm{u}_j}^2$. This is smaller than the variance of $y_{iT}$ under the null hypothesis for large $N$.

In light of the data-generating processes (11) and (12), we are allowed to use Model (8) for hypothesis testing even in the presence of the trend variable $t$ because it does not introduce any additional regressor. This is in contrast to the usual time series and panel unit root tests where the linear time trend introduces an additional regressor that results in lower power for the tests.

However, the OLS estimator of $\phi$ is inconsistent under the null and alternative hypotheses. In fact, we have under Assumptions 1, 2 and 3

$$\hat{\phi}_1 \xrightarrow{p} \left(\sigma_{\mathrm{m}}^2 + \sigma_{\mathrm{b}}^2 + M_\lambda - f_1' q_\lambda q_\lambda' f_1\right)^{-1} \left(\sigma_{\mathrm{m}}^2 + T\sigma_{\mathrm{b}}^2 + \alpha^{T-1}\left(M_\lambda - f_1' q_\lambda q_\lambda' f_1\right)\right),$$

which implies that $\hat{\phi}_1$ converges in probability to a random variable greater than 1 almost surely even when $\alpha = 1$ as long as $\sigma_{\mathrm{b}}^2 > 0$. The OLS-based t-ratio should not be used for unit root testing for this reason.

By contrast, $\hat{\phi}_1$ and $t_{\hat{\phi}_1}$ have sound limiting properties as reported below.

**Theorem 3** *Assume that Assumptions 1–5 hold.*

*(i) Under the null hypothesis (3),*

*(a)* $\hat{\phi}_\mu = \beta(T-1) + O_p\left(\frac{1}{\sqrt{N}}\right)$ *and* $\hat{\phi}_1 = 1 + O_p\left(\frac{1}{\sqrt{N}}\right)$;

*(b)* $t_{\hat{\phi}_1} \xrightarrow{d} N(0,1)$.

*(ii) Under the alternative hypothesis (4),*

*(a)* $\hat{\phi}_\mu = \mu\left(1 - \alpha^{T-1}\right) + \beta\left(T - \alpha^{T-1}\right) + O_p\left(\frac{1}{\sqrt{N}}\right)$ *and* $\hat{\phi}_1 = \alpha^{T-1} + O_p\left(\frac{1}{\sqrt{N}}\right)$;

(b) $t_{\hat{\phi}_1} \xrightarrow{p} -\infty$ *if* $|\alpha| < 1$ *and* $t_{\hat{\phi}_1} \xrightarrow{p} \infty$ *if* $\alpha > 1$.

Part (i) of Theorem 3 implies that $\hat{\phi}_\mu$ and $\hat{\phi}_1$ are $\sqrt{N}$-consistent for $\beta(T-1)$ and 1, respectively, and that the test statistic, $t_{\hat{\phi}_1}$, has a standard normal distribution in the limit. Part (ii) of Theorem 3 shows that the IV estimators $\hat{\phi}_\mu$ and $\hat{\phi}_1$ are $\sqrt{N}$-consistent for $\mu(1-\alpha^{T-1}) + \beta(T - \alpha^{T-1})$ and $\alpha^{T-1}$, respectively, and that the $t$-ratio diverges to $-\infty$ or $\infty$ in probability under the alternative hypothesis. Thus, the unit root test continues to be consistent for Model (2).

What effects would $T$ have on the size and power of our unit root tests? As $T$ becomes larger, so does the difference in the values of $\phi_1$ under the null and alternative hypotheses, which will induce higher power of our unit root test. This will be studied more closely in the next section via simulation.

# 4 Simulation

This section reports simulation results for the test statistics of this paper. Such well-known tests as Im, Pesaran and Shin's (2003), and Levin, Lin and Chu's (2002), Maddala and Wu's (2001), and Choi's (2001) are not appropriate for the sample sizes considered in this paper and are not simulated here. Simulation results for Harris and Tzavalis' (1999) test are also reported in this section. Harris and Tzavalis' test, essentially the Dickey-Fuller coefficient test with bias corrections and variance adjustments, is designed for short $T$ and independent panels.

Data for our simulation were generated by Models (1) and (2). We set $\mu_i \sim$ i. i. d. $N(0,1)$, $\beta_i \sim$ i. i. d. $N(0,1)$ and $u_{it} \sim$ i. i. d. $N(0, \sigma^2_{\mathrm{u}_i})$ with $\sigma^2_{\mathrm{u}_i} \sim U[0.5, 1.5]$. The initial variables, $\{x_{i1}\}$, are generated by $x_{i1} = \lambda_i f_1$ with $f_1 \sim$ i. i. d. $N(0,1)$, and $\lambda := [\lambda_1, ..., \lambda_N]' \sim N(0, \Omega)$, $\Omega := \begin{bmatrix} 10 & \delta & \cdots & \delta \\ \delta & 10 & & \vdots \\ \vdots & \ddots & \ddots & \delta \\ \delta & \cdots & \delta & 10 \end{bmatrix}$. Under this data-generating scheme, any random permutation of $\{y_{i,1}\}$ can be used as an instrument. We use $\delta = 1, 3$ in our simulation. Since $\{\sigma^2_{\mathrm{u}_i}\}$ and $\{\lambda_i\}$ are sequences of constants, they are generated once before iterations and kept to have the same values throughout the iterations.[3] Under the null hypothesis, $\alpha_i = 1$ for every $i$; and under the alternative hypothesis, $\alpha_i = \alpha + \frac{\eta_i}{N^{0.8}}$ for every $i$ with $\alpha = 1, 1.01, 0.99, 0.98, 0.95$ and $\eta_i \sim$ i. i. d. $U[-0.25, 0.25]$. The numbers of time series observations considered are 2, 3 and 7. The numbers of cross-sectional observations are 50, 100, 200 and 400. The reported results are based on 5,000 iterations, and the nominal size is set at 0.05. We report both size-unadjusted and size-adjusted empirical power. The size-adjusted empirical power means that of the test which uses an empirical critical value that always makes the empirical size 0.05.

Table 1 reports simulation results using Model (1). The IV-based t-ratio employs 25 instruments[4] that are selected by a random permutation of the regressor. Empirical sizes of both left-sided and right-sided tests are reported. The sample size $T = 2$ deserves an attention because only this paper's test can be used at this sample size. All extant panel unit root tests require at least 3 time series observations, 1 reserved for the initial variable and the remaining 2 used for the parameter estimation. The OLS-based t-ratio, $t_{\hat{\phi}_1}$, keeps the nominal size well at all cross-sectional and time series sample sizes and its power quickly converges to 1 as $N$ increases. The IV-based t-ratio, $t_{\hat{\phi}_1}$, also keeps the nominal size well under the null hypothesis and has lower empirical power than $t_{\hat{\phi}_1}$ at all $T$ and $N$. As $N$ increases, its power increases as $N$ does, confirming consistency of the test. In addition, as well expected, the empirical powers of our test statistics improve as $\alpha$ takes smaller values and $T$ becomes larger. The value of $\delta$ does not influence empirical sizes and powers much in our data-generating scheme.

Table 2 reports simulation results using Model (2) and $t_{\hat{\phi}_1}$ using 25 instruments. The OLS-based t-ratio is not experimented with, because it is not consistent. The sample sizes $T = 2, 3$ are worthy of our attention, because all panel unit root tests

[3] If $\{\lambda_i\}$ is treated as a random sequence, much more computation time is needed while the results are not qualitatively different.

[4] Using more than 25 instruments did not bring any notable changes in empirical sizes of the test. See the next section for discussions on the choice of instruments.

other than ours require at least 4 time series observations, 1 for the initial variable and 3 for the coefficient estimation. According to Table 3, $t_{\hat{\phi}_1}$ tends to be well sized. The behavior of the empirical power in Table 2 is similar to that in Table 1: It increases with $N$ and $T$ and decreases with the value of $\alpha$. Compared to Table 1, empirical power decreases in Table 2. This can be explained by the lower efficiency of $\hat{\phi}_1$ for Model (2) than for Model (1).

Table 3 reports simulation results for Harris and Tzavalis' (1999) test using Model (1) at $T = 3$. Dependent panels were generated as for Table 1 with $\delta = 1$. Independent panels were generated by setting $\lambda_i = 0$ for all $i$. Harris and Tzavalis' test keeps the nominal size reasonably well. But it shows quite low power for dependent panels. In many cases, its empirical power is zero. For independent panels, the test behaves reasonably: Its power increases as $N$ does and as the alternative points become farther away from the null point. Results of Table 3 confirm that Harris and Tzavalis' test is suitable only for independent panels as it is designed to be.

# 5 An empirical example

This section applies our panel unit root test to the monthly real wages of new college graduates in South Korea. The data have been collected regularly by the Korea Employment Information Service since 2006. We will use the panel data collected in 2006 and 2008 regarding 15,226 two-year and four-year college graduates, all of whom were employed in 2006 for the first time and remained to be hired in 2008.[5] With small $T$ and large $N$, our panel unit root test is suitable for this data. Moreover, since the individuals have the same tenure in the job market, there will be less heterogeneity in individual profiles than for other samples.

There has been much research about modelling individual earnings dynamics in labor economics. Some studies report empirical evidence that the shocks to earnings (i.e., $x_{it}$ in Models (1) and (2), if $y_{it}$ denotes the log earnings) seem stationary (see Lillard and Weiss, 1979; Baker, 1997; Haider, 2001; Guvenen, 2009; Okubo, 2015).

[5] Some data points obviously subject to measurement errors are excluded.

Others assume the presence of a unit root in individual earnings (see, for example, MaCurdy, 1982; Abowd and Card, 1989; Topel and Ward, 1992; Dickens, 2000; Moffitt and Gottschalk, 2002; Ramos, 2003) and study individual earnings dynamics, reporting mainly that the autocorrelations of the log difference of income are small and negative. This practice was motivated by MaCurdy(1982) who reports empirical evidence favoring the presence of a unit root.[6] If there is a unit root in the shocks to earnings, a large proportion of the individual variability of earnings can be explained by the shocks and the variability increases as individuals accumulate job-market experience. If not, the individual variability of earnings can better be explained by the heterogeneous job-market profiles of the individuals and there is no guarantee that the variability increases as time goes by. Beyond this, the nature of idiosyncratic shocks crucially affects individuals' consumption–savings decisions, which are related to a host of economic issues. See Guvenen (2009) for further discussion. Notwithstanding the perceived importance of a unit root for earnings data, panel unit root tests have not been used much for earnings data. The only result from using panel data is Ng (2008) who reports that one-fifth to as many as one third of earnings of male household heads in the US ($N = 104$ and $T = 25$) have a unit root. The lack of interest in using panel unit root tests for earnings data is indeed surprising given their widespread use in other research areas.

Table 4 reports estimation and test results using the monthly real wages of new college graduates in South Korea. We report the results both from the full sample and from the restricted sample of four-year college graduates. The IV regressions were run using 100, 150,..., 2950, 3000 instruments. At each number of instruments, 11 sets of instruments were randomly chosen[7] and the results corresponding to the median of the AR(1) coefficient estimates were recorded. Then, letting the AR(1) coefficient using $k$ instruments be $\hat{\phi}_1(k),$ the sample variances of $\{\hat{\phi}_1(l)\}_{l=k-100,\ k-50,\ k,\ k+50,\ k+100}$ were calculated. Table 4 reports results corresponding to the number of instruments

[6] Topel and Ward (1992) also report similar results.

[7] Using more sets of instruments does not bring qualitatively different results. But we find that it is computationally very costly. So we used only 11 sets of instruments.

showing the minimum sample variance.

Part (i) of Table 4 show that the null hypothesis of a unit root is rejected at the 1% significance level for both samples when the OLS-based t-ratio is used. The AR(1) coefficient estimates are 0.6961 and 0.6567 for the full and restricted samples, respectively. Part (ii) of Table 4 reports results of IV-based estimation and test. For both samples, the null hypothesis of a unit root is rejected at the 1% level. The AR(1) coefficient estimates are 0.6994 and 0.6603 for the full and restricted samples, respectively, which are similar to the corresponding results using OLS. We report here only the results chosen by the criterion of minimum variance of the AR(1) coefficient estimates, but all the values of the IV-based t-ratio using 100, 150,..., 2950, 3000 instruments reject the null hypothesis of a unit root at the 1% level.

The test results of Table 4 strongly indicate that the individual earnings follow a stationary autoregressive process rather than a unit root process. This provides support to the similar evidence obtained by Guvenen (2009) and Okubo (2015) using US and Japanese data, respectively.

One of the important issues in IV estimation is selection of the number of instruments. Andrews (1999) and Donald and Newey (2001) are representative works studying this issue. However, the number of available instruments in this paper is much larger than the sample size, making existing methods inapplicable. In this section, we used the minimum variance criterion, but other methods can also be considered.

## 6 Summary and further remark

We have proposed a panel unit root test suitable for micropanels with short time dimensions and large cross sections. This test is based on a heterogeneous panel AR(1) model that allows for cross-sectional dependency introduced by the initial condition, which assumes a factor structure. Most importantly, the test does not use the AR(1) coefficient estimator. Instead, it rests on the fact that the initial condition has permanent effects on the trajectory of a time series in the presence of a unit root.

We measure the effects of the initial condition by cross-sectional regressions using the first time series observations as a regressor and the last as the dependent variable. Our test statistic is the $t$-ratio for the coefficient of the regressor, and it has a standard normal distribution in the limit. The $t$-ratio is based on the OLS and IV estimators. The IV estimator uses reshuffled regressors as instruments. The test can be used for very small $T$ including $T = 2$ as long as $N$ is large. Simulation results show that this paper's test has reasonable empirical size and power. As an example, we apply the test to the monthly real wage of new college graduates in South Korea who were hired in the same year for the first time. Even with just 2 time-series observations, the test rejects the null hypothesis of a unit root at conventional significance levels.

Our panel AR(1) model is new in the literature and should be useful for some applications. The IV estimator consistently estimates the model's AR(1) coefficients regardless of their locations. Choosing the number of instruments is an important problem for the internal IV regression. This paper does not deal with this problem, because it is beyond the scope of this paper. More work for the problem is needed to make the internal IV regression more useful.

# Appendix I Technical assumptions

This section presents some assumptions of purely technical nature that are required to prove the main results of the paper. Assumption 3 involves the factor loadings.

**Assumption 3** *(i)* $\frac{1}{N}\sum_{i=1}^{N}\lambda_i \to q_\lambda$;

*(ii)* $\frac{1}{N}\sum_{i=1}^{N}\lambda_i\lambda_i' \to M_\lambda$ $(>0)$.

Assumption 4 is required for proper limiting distributions of the OLS and internal IV estimators of Section 3.

**Assumption 4** *(i)* $\text{plim}\,\frac{1}{N}\begin{bmatrix} 1 & \sum_{i=1}^{N} y_{i1} \\ \sum_{i=1}^{N} y_{i1} & \sum_{i=1}^{N} y_{i1}^2 \end{bmatrix} := A$ *is invertible almost surely.*

*(ii)* $\text{plim}\,\frac{1}{N}\sum_{i=1}^{N}\sigma_{\mathrm{u}_i}^2 \begin{bmatrix} 1 & y_{1i} \\ y_{1i} & y_{1i}^2 \end{bmatrix} := B$ *is positive definite almost surely.*

*(iii)* $\text{plim}\,\frac{1}{N}Z'Z := M_{ZZ}$ *is invertible almost surely.*

*(iv)* $\text{plim}\,\frac{1}{N}Y_1'Z := M_{Y_1Z}$ *exists.*

*(v)* $\text{plim}\,\frac{1}{N}Y_1'P_ZY_1 := C$ *is invertible almost surely.*

*(vi)* $\text{plim}\,\frac{1}{N}\sum_{i=1}^{N}\sigma_{\mathrm{u}_i}^2 \begin{bmatrix} 1 & z_i' \\ z_i & z_iz_i' \end{bmatrix} := D$ *is positive definite almost surely.*

*(vii)* $\text{plim}\,\frac{1}{N}\sum_{i=1}^{N}\left(\left(\sum_{j=0}^{T-2}\alpha^{2j}\right)\sigma_{\mathrm{u}_i}^2 + \left(1-\alpha^{T-1}\right)^2\sigma_{\mathrm{m}}^2\right) \begin{bmatrix} 1 & z_i' \\ z_i & z_iz_i' \end{bmatrix} := E$ *is positive definite almost surely.*

*(viii)* $\text{plim}\,\frac{1}{N}\sum_{i=1}^{N}\left(\sigma_{\mathrm{u}_i}^2 + (T-1)\sigma_{\mathrm{b}}^2\right) \begin{bmatrix} 1 & z_i' \\ z_i & z_iz_i' \end{bmatrix} := G$ *is positive definite almost surely.*

*(ix)* $\text{plim}\,\frac{1}{N}\sum_{i=1}^{N}\left(\left(\sum_{j=0}^{T-2}\alpha^{2j}\right)\sigma_{\mathrm{u}_i}^2 + \left(1-\alpha^{T-1}\right)^2\sigma_{\mathrm{m}}^2 + \left(T-\alpha^{T-1}\right)^2\sigma_{\mathrm{b}}^2\right) \begin{bmatrix} 1 & z_i' \\ z_i & z_iz_i' \end{bmatrix} :=$ $K$ *is positive definite almost surely.*

Assumption 5 is required to apply the central limit theorem for independent and heterogeneous sequences (Davidson, 1994, chapter 23).

**Assumption 5** *(i)* $\frac{1}{N}\sum_{i=1}^{N}\sigma_{\mathrm{u}_i}^2 > 0$ *uniformly in* $N$*;*

*(ii) For* $\delta > 0$, $E\left|u_{iT}\right|^{2+\delta} < \infty$ *uniformly in* $i$*.*

*(iii) For* $\delta > 0$, $E\left|m_i\right|^{2+\delta} < \infty$ *uniformly in* $i$*.*

*(iv) For* $\delta > 0$, $E\left|b_i\right|^{2+\delta} < \infty$ *uniformly in* $i$*.*

# Appendix II Proofs

**Proof of Theorem 1**:

(i) (a) Under the null hypothesis, $\phi_\mu = 0$ and $\phi_1 = 1$. Thus, we have

$$\sqrt{N}\begin{pmatrix} \tilde{\phi}_\mu \\ \tilde{\phi}_1 - 1 \end{pmatrix} = \begin{bmatrix} 1 & \sum_{i=1}^{N} y_{i1}/N \\ \sum_{i=1}^{N} y_{i1}/N & \sum_{i=1}^{N} y_{i1}^2/N \end{bmatrix}^{-1} \begin{bmatrix} \sum_{i=1}^{N} s_{iT}/\sqrt{N} \\ \sum_{i=1}^{N} y_{i1}s_{iT}/\sqrt{N} \end{bmatrix}$$

$$:= A_N^{-1}B_N.$$

Under Assumption 4, we have

$$A_N \xrightarrow{p} A. \tag{A.II.1}$$

The sequence $\{s_{iT}\}$ is independent with $E(s_{iT}) = 0$ and $\mathrm{Var}(s_{iT}) = (T-1)\sigma_{\mathrm{u}_i}^2$ for every $i$. Moreover, $E\left|s_{iT}\right|^{2+\delta} < \infty$ and $E\left|m_i s_{iT}\right|^{2+\delta} < \infty$ uniformly in $i$ for $\delta > 0$ due to Assumption 5 and the Minkowski inequality. Thus, for given $f_1$, the central limit theorem for independent and heterogeneous sequences gives

$$B_N \xrightarrow{d} N\left(0, (T-1)\,\mathrm{plim}\,\frac{1}{N}\sum_{i=1}^{N}\sigma_{\mathrm{u}_i}^2\begin{bmatrix} 1 & y_{i1} \\ y_{i1} & y_{i1}^2 \end{bmatrix}\right) = N(0, (T-1)B). \tag{A.II.2}$$

Asymptotic results (A.II.1) and (A.II.2) under Assumption 4 indicate that the limiting distribution of the OLS estimator is well defined and that it is mixture normal. The consistency rates of the OLS estimators $\tilde{\phi}_\mu$ and $\tilde{\phi}_1$ stated in part (i) follow from this.

(b) Because $\tilde{\phi}_1 - 1 = \sum_{i=1}^{N}(y_{i1} - \bar{y}_1)\,v_{iT}/\sum_{i=1}^{N}(y_{i1} - \bar{y}_1)^2$ and, because $\tilde{v}_{iT}^2 = v_{iT}^2 + O_p\left(\frac{1}{\sqrt{N}}\right)$, $\sqrt{N}$ times the denominator of $t_{\tilde{\phi}_1}$ converges in probability to the square root of the asymptotic variance of $\tilde{\phi}_1 - 1$. This is represented by the southeast element of $(T-1)A^{-1}BA^{-1\prime}$. The stated result follows from this.

(ii) (a) The binomial expansion of $\alpha_i^{T-1}$ gives

$$\alpha_i^{T-1} = \left(\alpha + \eta_i/N^{1/2+\varepsilon}\right)^{T-1} = \alpha^{T-1} + \sum_{j=1}^{T-1} c_j \left(\eta_i/N^{1/2+\varepsilon}\right)^j, \tag{A.II.3}$$

where $\{c_j\}$ is a sequence of finite constants. Using this representation, write the data-generating process (7) as

$$y_{iT} = \mu + m_i \ + \alpha^{T-1}x_{i1} + w_{iT} + \zeta_{iN}x_{i1},$$

where $\zeta_{iN} := \sum_{j=1}^{T-1} c_j \left(\eta_i/N^{1/2+\varepsilon}\right)^j$. Because

$$\begin{aligned} w_{iT} &= \sum_{j=0}^{T-2} \alpha_i^j u_{i,T-j} \\ &= u_{i,T} + \sum_{j=1}^{T-2} \left(\alpha + \eta_i/N^{1/2+\varepsilon}\right)^j u_{i,T-j} \\ &= u_{i,T} + \sum_{j=1}^{T-2} \left(\alpha^j + \sum_{k=1}^{j} d_k \left(\eta_i/N^{1/2+\varepsilon}\right)^k\right) u_{i,T-j} \\ &= \sum_{j=0}^{T-2} \alpha^j u_{i,T-j} + \sum_{j=1}^{T-2} \left(\sum_{k=1}^{j} d_k \left(\eta_i/N^{1/2+\varepsilon}\right)^k\right) u_{i,T-j} \\ &= \sum_{j=0}^{T-2} \alpha^j u_{i,T-j} + O_p\left(\frac{1}{N^{1/2+\varepsilon}}\right), \end{aligned} \tag{A.II.4}$$

where $\{d_k\}$ is a sequence of finite constants, $\{w_{iT}\}$ behaves as if it were $\left\{\sum_{j=0}^{T-2} \alpha^j u_{i,T-j}\right\}$ in the limit. Thus, we obtain

$$\begin{aligned} \tilde{\phi}_1 &= \frac{\sum_{i=1}^N (y_{i1} - \bar{y}_1)(y_{iT} - \bar{y}_T)}{\sum_{i=1}^N (y_{i1} - \bar{y}_1)^2} \\ &= \left(\sum_{i=1}^N (m_i - \bar{m} + x_{i1} - \bar{x}_1)^2/N\right)^{-1} \\ &\quad \times \left(\sum_{i=1}^N (m_i - \bar{m} + x_{i1} - \bar{x}_1)(m_i - \bar{m} + \alpha^{T-1}(x_{i1} - \bar{x}_1) + w_{iT} - \bar{w}_T)/N + o_p(1)\right) \\ &\xrightarrow{p} \frac{\sigma_{\mathrm{m}}^2 + \alpha^{T-1}\left(f_1' M_\lambda f_1 - f_1' q_\lambda q_\lambda' f_1\right)}{\sigma_{\mathrm{m}}^2 + f_1' M_\lambda f_1 - f_1' q_\lambda q_\lambda' f_1}, \end{aligned}$$

as desired.

(b) Because the denominator of the t-ratio is $O_p(1)$, the stated result follows from part (a).

**Proof of Theorem 2**:

(i) (a) Under the null hypothesis, we may write

$$\sqrt{N}\begin{pmatrix} \hat{\phi}_\mu \\ \hat{\phi}_1 - 1 \end{pmatrix} = (Y_1' P_Z Y_1)^{-1} Y_1' Z (Z'Z)^{-1} \begin{bmatrix} \sum_{i=1}^N s_{iT}/\sqrt{N} \\ \sum_{i=1}^N z_i s_{iT}/\sqrt{N} \end{bmatrix}$$

$$:= C_N^{-1} Y_1' Z (Z'Z)^{-1} D_N.$$

Under Assumption 4, we have

$$\frac{1}{N} C_N \xrightarrow{p} C;\ \frac{1}{N} Y_1' Z \xrightarrow{p} M_{Y_1 Z};\ \frac{1}{N} Z'Z \xrightarrow{p} M_{ZZ} \qquad \text{(A.II.5)}$$

and, for given $f_1$, the central limit theorem for independent and heterogeneous sequences gives as for part (a) of Theorem 1

$$D_N \xrightarrow{d} N\left(0, (T-1)D\right). \qquad \text{(A.II.6)}$$

The consistency rates of the IV estimators $\hat{\phi}_\mu$ and $\hat{\phi}_1$ stated in part (i) follow from relations (A.II.5) and (A.II.6).

(b) This is obtained as for part (i) (b) of Theorem 1.

(ii) (a) Using the expansion (A.II.3), write the data-generating process (7) as

$$y_{iT} = \mu\left(1 - \alpha^{T-1}\right) + \alpha^{T-1} y_{i1} + w_{iT} + m_i\left(1 - \alpha^{T-1}\right) + \zeta_{iN}(y_{i1} - \mu - m_i).$$

Using this, we have

$$\sqrt{N}\begin{pmatrix} \hat{\phi}_\mu - \mu(1-\alpha^{T-1}) \\ \hat{\phi}_1 - \alpha^{T-1} \end{pmatrix} = (\frac{1}{N} Y_1' P_Z Y_1)^{-1} Y_1' Z (Z'Z)^{-1} \begin{bmatrix} \sum\limits_{i=1}^N \tau_{iNT}/\sqrt{N} \\ \sum\limits_{i=1}^N z_i \tau_{iNT}/\sqrt{N} \end{bmatrix}$$

$$:= C_N^{-1} Y_1' Z (Z'Z)^{-1} E_N,$$

where $\tau_{iNT} := w_{iT} + m_i\left(1 - \alpha^{T-1}\right) + \zeta_{iN}(y_{i1} - \mu - m_i)$. Since $\zeta_{iN} = O\left(\frac{1}{N^{1/2+\varepsilon}}\right)$ for every $i$, $\frac{1}{\sqrt{N}} \sum\limits_{i=1}^N \zeta_{iN}(y_{i1} - \mu - m_i) = o_p(1)$ and $\frac{1}{\sqrt{N}} \sum_{i=1}^N z_i \zeta_{iN}(y_{i1} - \mu - m_i) = o_p(1)$.

Therefore, we have

$$E_N = \begin{bmatrix} \sum_{i=1}^{N} \left( w_{iT} + m_i(1-\alpha^{T-1}) \right) / \sqrt{N} \\ \sum_{i=1}^{N} z_i \left( w_{iT} + m_i(1-\alpha^{T-1}) \right) / \sqrt{N} \end{bmatrix} + o_p(1).$$

Moreover, $\{w_{iT}\}$ behaves as if it were $\left\{ \sum_{j=0}^{T-2} \alpha^j u_{i,T-j} \right\}$ in the limit as shown in (A.II.4). Thus, the standard theory of linear regression yields, for given $f_1$,

$$\sqrt{N} \begin{pmatrix} \hat{\phi}_\mu - \mu(1-\alpha^{T-1}) \\ \hat{\phi}_1 - \alpha^{T-1} \end{pmatrix} \xrightarrow{d} N(0, C^{-1} M_{Y_1 Z} M_{ZZ} E M_{ZZ} M_{Y_1 Z}' C^{-1}),$$

where $E$ is defined in Assumption 4.

(b) Let the usual heteroskedasticity-robust $t$-ratio for the null hypothesis $H_0 : \phi_1 = \alpha^{T-1}$ be denoted as $\upsilon_{\phi_1}$. Then we have $t_{\hat{\phi}_1} = \upsilon_{\phi_1} + \sqrt{N}(\alpha^{T-1} - 1)/(\sqrt{N}F)$, where $F$ is the denominator of $t_{\hat{\phi}_1}$. Since $\upsilon_{\phi_1}$ has a standard normal distribution and $F = O_p\left(1/\sqrt{N}\right)$, the stated result follows.

**Proof of Theorem 3:**

(i) (a) Write $y_{iT} = \beta(T-1) + y_{i1} + s_{iT} + b_i(T-1)$. Then, as in the proof of Theorem 2 (i) (a), for given $f_1$,

$$\sqrt{N} \begin{pmatrix} \hat{\phi}_\mu - \beta(T-1) \\ \hat{\phi}_1 - 1 \end{pmatrix} = (\frac{1}{N} Y_1' P_Z Y_1)^{-1} Y_1' Z (Z'Z)^{-1} \begin{bmatrix} \frac{\sum_{i=1}^{N} (s_{iT} + b_i(T-1))}{\sqrt{N}} \\ \frac{\sum_{i=1}^{N} z_i (s_{iT} + b_i(T-1))}{\sqrt{N}} \end{bmatrix}$$

$$\xrightarrow{d} N(0, (T-1) C^{-1} M_{Y_1 Z} M_{ZZ} G M_{ZZ} M_{Y_1 Z}' C^{-1}),$$

where $G$ is defined in Assumption 4. The stated results follow from this.

(b) Use the same arguments as for Theorem 2 (i) (b).

(ii) (a) Using (A.II.3), the data-generating process (12) can be written as

$$\begin{aligned} y_{iT} &= \mu_i \left(1 - \alpha_i^{T-1}\right) + \beta_i (T - \alpha_i^{T-1}) + \alpha_i^{T-1} y_{i1} + w_{iT} \\ &= \mu \left(1 - \alpha^{T-1}\right) + \beta \left(T - \alpha^{T-1}\right) + \alpha^{T-1} y_{i1} + w_{iT} \\ &\quad + m_i \left(1 - \alpha^{T-1}\right) + b_i \left(T - \alpha^{T-1}\right) + \zeta_{iN}(y_{i1} - \mu - \beta - m_i - b_i). \end{aligned}$$

Now, $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\zeta_{iN}(y_{i1}-\mu-\beta-m_i-b_i)$ and $\frac{1}{\sqrt{N}}\sum_{i=1}^{N}z_i\zeta_{iN}(y_{i1}-\mu-\beta-m_i-b_i)$ are $O_p(1)$ under the assumptions, implying that the terms involving $\zeta_{iN}$ do not affect the limiting distribution of the IV estimator. Thus, according to the standard theory of linear regression, we have, for given $f_1$,

$$\sqrt{N}\begin{pmatrix}\hat{\phi}_\mu-\mu(1-\alpha^{T-1})-\beta(T-\alpha^{T-1})\\ \hat{\phi}_1-\alpha^{T-1}\end{pmatrix}\xrightarrow{d}N(0,(T-1)C^{-1}M_{Y_1Z}M_{ZZ}KM_{ZZ}M'_{Y_1Z}C^{-1}),$$

where $K$ is as defined in Assumption 4. Thus, the stated results are obtained.

(b) This follows as in the proof of part (ii) (b) of Theorem 2.

## References


Abowd, J.M., and D. Card. (1989). "On the Covariance Structure of Earnings and Hours Changes," Econometrica, 57, 411–445.

Andrews, D. W. (1999). Consistent moment selection procedures for generalized method of moments estimation. Econometrica, 67(3), 543-563.

Bai, J., and S. Ng. (2004). "A Panic Attack on Unit Roots and Cointegration," Econometrica, 72, 1127–1177.

Baker, M. (1997). "Growth-Rate Heterogeneity and the Covariance Structure of Life-Cycle Earnings," Journal of Labor Economics, 21, 338–375.

Breitung, J. (2002). "Nonparametric Tests for Unit Root and Cointegration," Journal of Econometrics, 108, 343–363.

Breitung, J., and S. Das. (2005). "Panel Unit Root Tests Under Cross-Sectional Dependence," Statistica Neerlandica, 59, 414–433.

Breitung, J., and C. Gouriéroux. (1997). "Rank Tests for Unit Roots," Journal of Econometrics, 81, 7–27.

Breitung, J., and W. Meyer. (1994). "Testing for Unit Roots in Panel Data: Are Wages on Different Bargaining Levels Cointegrated?" Applied Economics, 26, 353–361.

Breitung, J., and M.H. Pesaran. (2008). "Unit Roots and Cointegration in Panels," in The Econometrics of Panel Data, edited by L. Mátyás and P. Sevestre, pp. 279–322. Berlin: Springer.

Burridge, P., and E. Guerre. (1996). "The Limit Distribution of Level Crossings of a Random Walk, and a Simple Unit Root Test," Econometric Theory, 12, 705–723.

Cavaliere, G. (2001). "Testing the Unit Root Hypothesis Using Generalized Range Statistics," Econometrics Journal, 4, 70–88.

Choi, I. (2001). "Unit Root Tests for Panel Data," Journal of International Money and Finance, 20, 249–272.

Choi, I. (2006). "Nonstationary Panels," in Palgrave Handbook of Econometrics, vol. 1, edited by T.C. Mills and K. Patterson, pp. 511–539. New York: Palgrave Macmillan.

Choi, I. (2015). Almost All About Unit Roots: Foundations, Developments and Applications. Cambridge, UK: Cambridge University Press.

Choi, I., and T.K. Chue. (2007). "Subsampling Hypothesis Tests for Nonstationary Panels with Applications to Exchange Rates and Stock Prices," Journal of Applied Econometrics, 22, 233–264.

Davidson, J. (1994). Stochastic Limit Theory. New York: Oxford University Press.

De Blander, R., and G. Dhaene. (2012). Unit Root Tests for Panel Data with AR(1) Errors and Small $T$," The Econometrics Journal, 15, 101–124.

Demetrescu, M., U. Hassler, and A. Tarcolea. (2006). "Combining Significance of Correlated Statistics with Application to Panel Data," Oxford Bulletin of Economics and Statistics, 68, 647–663.

De Wachter, S., R.D.F. Harris, and E. Tzavalis. (2007). "Panel Data Unit Roots Tests: The Role of Serial Correlation and the Time Dimension," Journal of Statistical Planning and Inference, 137, 230–244.

Dickens, R. (2000). "The Evolution of Individual Male Earnings in Great Britain: 1975–95," The Economic Journal, 110, 27–49.

Donald, S. G., & Newey, W. K. (2001). "Choosing the Number of Instruments," Econometrica, 69, 1161–1191.

Elliott, G., T.J. Rothenberg, and J.H. Stock. (1996). "Efficient Tests for an Autoregressive Unit Root," Econometrica, 64, 813–836.

Guvenen, F. (2009). "An Empirical Investigation of Labor Income Processes," Review of Economic Dynamics, 12, 58–79.

Haider, S.J. (2001). "Earnings Instability and Earnings Inequality of Males in the United States: 1967–1991," Journal of Labor Economics, 19, 799–836.

Hallin, M., R. Van den Akker, and B.J. Werker. (2011). "A Class of Simple Distribution-Free Rank-Based Unit Root Tests," Journal of Econometrics, 163, 200–214.

Harris, R.D.F., and E. Tzavalis. (1999). "Inference for Unit Roots in Dynamic Panels Where the Time Dimension Is Fixed," Journal of Econometrics, 91, 201–226.

Harris, R.D.F., and E. Tzavalis. (2004). "Testing for Unit Roots in Dynamic Panels in the Presence of a Deterministic Trend: Re-examining the Unit Root Hypothesis for Real Stock Prices and Dividends," Econometric Reviews, 23, 149–166.

Hasan, M.N., and R.W. Koenker. (1997). "Robust Rank Tests of the Unit Root Hypothesis," Econometrica, 65, 133–161.

Homm, U., and J. Breitung. (2012). "Testing for Speculative Bubbles in Stock Markets: A Comparison of Alternative Methods," Journal of Financial Econometrics, 10, 198–231.

Im, K.S., M.H. Pesaran, and Y. Shin. (2003). "Testing for Unit Roots in Heterogeneous Panels," Journal of Econometrics, 115, 53–74.

Karavias, Y., & Tzavalis, E. (2014). "Testing for Unit Roots in Short Panels Allowing for a Structural Break," Computational Statistics & Data Analysis, 76, 391-407.

Karavias, Y., & Tzavalis, E. (2014). "Testing for Unit Roots in Panels with Structural Changes, Spatial and Temporal Dependence When the Time Dimension is Finite," Working Paper, Athens University of Economiocs and finance.

Kruiniger, H. (2008). "Maximum Likelihood Estimation and Inference Methods for the Covariance Stationary Panel AR(1)/Unit Root Model," Journal of Econometrics, 144, 447–464.

Kruiniger, H. (2009). "GMM Estimation and Inference in Dynamic Panel Data Models with Persistent Data," Econometric Theory, 25, 1348–1391.

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

Lillard, L.A., & Y. Weiss. (1979). "Components of Variation in Panel Earnings Data: American Scientists 1960–70," Econometrica, 47, 437–454.

MaCurdy, T.E. (1982). "The Use of Time Series Processes to Model the Error Structure of Earnings in a Longitudinal Data Analysis," Journal of Econometrics, 18, 83–114.

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, 631–652.

Moffitt, R.A., and P. Gottschalk, P. (2002). "Trends in the Transitory Variance of Earnings in the United States," Economic Journal, 112, C68–C73.

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

Nielsen, M.O. (2009). "A Powerful Test of the Autoregressive Unit Root Hypothesis Based on a Tuning Parameter Free Statistic," Econometric Theory, 25, 1515–1544.

Okubo, M. (2015). "Earnings Dynamics and Profile Heterogeneity: Estimates from Japanese Panel Data," Japanese Economic Review, 66, 112-146.

Pesaran, M.H. (2007). "A Simple Panel Unit Root Test in the Presence of Cross-Section Dependence," Journal of Applied Econometrics, 22, 265–312.

Phillips, P.C.B., and D. Sul. (2003). "Dynamic Panel Estimation and Homogeneity Testing under Cross Section Dependence," Econometrics Journal, 6, 217–259.

Phillips, P.C.B., Y. Wu and J. Yu (2011). "Explosive Behavior in the 1990s NASDAQ:

When did Exuberance Escalate Asset Values?" International Economic Review 71, 421–436.

Ramos, X. (2003). "The Covariance Structure of Earnings in Great Britain, 1991–1999," Economica, 70, 353–374.

Robertson, D., Sarafidis, V., & Westerlund, J. (2014). "GMM Unit Root Inference in Generally Trending and Cross-Correlated Dynamic Panels," Working Paper, Monash University.

Schmidt, P., and P.C.B. Phillips. (1992). "LM Tests for a Unit Root in the Presence of Deterministic Trends," Oxford Bulletin of Economics and Statistics, 54, 257–287.

Sheng, X., and J. Yang. (2013). "Truncated Product Methods for Panel Unit Root Tests," Oxford Bulletin of Economics and Statistics, 75, 624–636.

Topel, R.H., and M.P. Ward (1992). "Job Mobility and the Careers of Young Men," Quarterly Journal of Economics, 107, 439–479.

White, H. (1980). "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," Econometrica, 48, 817–838.

Table 1: Empirical Size and Power of $t_{\tilde{\phi}_1}$ and $t_{\hat{\phi}_1}$ for the Case of an Intercept

Note: 1. The number of iterations is 5,000.
2. NSA: not size-adjusted; SA: size-adjusted.
3. Two figures for each $N$ at $\alpha = 1$ are empirical sizes of left-sided and right-sided tests, respectively. The nominal size is kept at 0.05.

(i) $T = 2$

| | | $\delta = 1$ | | | | $\delta = 3$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| $\alpha$ | $N$ | OLS | | IV | | OLS | | IV | |
| | | NSA | SA | NSA | SA | NSA | SA | NSA | SA |
| 1.00 | 50 (L) | 0.072 | - | 0.058 | - | 0.067 | - | 0.060 | - |
| | 50 (R) | 0.063 | - | 0.059 | - | 0.063 | - | 0.055 | - |
| | 100 (L) | 0.059 | - | 0.048 | - | 0.056 | - | 0.051 | - |
| | 100 (R) | 0.061 | - | 0.048 | - | 0.062 | - | 0.051 | - |
| | 200 (L) | 0.055 | - | 0.041 | - | 0.057 | - | 0.042 | - |
| | 200 (R) | 0.045 | - | 0.045 | - | 0.045 | - | 0.046 | - |
| | 400 (L) | 0.049 | - | 0.043 | - | 0.051 | - | 0.043 | - |
| | 400 (R) | 0.050 | - | 0.046 | - | 0.049 | - | 0.046 | - |
| 1.01 | 50 | 0.451 | 0.410 | 0.308 | 0.282 | 0.432 | 0.373 | 0.289 | 0.274 |
| | 100 | 0.871 | 0.848 | 0.418 | 0.429 | 0.842 | 0.821 | 0.401 | 0.401 |
| | 200 | 1.000 | 1.000 | 0.703 | 0.715 | 1.000 | 1.000 | 0.675 | 0.686 |
| | 400 | 1.000 | 1.000 | 0.882 | 0.890 | 1.000 | 1.000 | 0.857 | 0.865 |
| 0.99 | 50 | 0.510 | 0.437 | 0.342 | 0.320 | 0.506 | 0.440 | 0.335 | 0.318 |
| | 100 | 0.881 | 0.866 | 0.454 | 0.463 | 0.859 | 0.841 | 0.439 | 0.436 |
| | 200 | 1.000 | 1.000 | 0.654 | 0.686 | 1.000 | 0.999 | 0.645 | 0.661 |
| | 400 | 1.000 | 1.000 | 0.868 | 0.878 | 1.000 | 1.000 | 0.848 | 0.858 |
| 0.98 | 50 | 0.907 | 0.871 | 0.713 | 0.694 | 0.887 | 0.853 | 0.692 | 0.673 |
| | 100 | 1.000 | 1.000 | 0.868 | 0.870 | 0.999 | 0.999 | 0.842 | 0.841 |
| | 200 | 1.000 | 1.000 | 0.981 | 0.983 | 1.000 | 1.000 | 0.970 | 0.973 |
| | 400 | 1.000 | 1.000 | 0.999 | 0.999 | 1.000 | 1.000 | 0.998 | 0.998 |
| 0.95 | 50 | 1.000 | 1.000 | 0.995 | 0.995 | 1.000 | 1.000 | 0.994 | 0.993 |
| | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 400 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

(ii) $T = 3$

| $\alpha$ | $N$ | $\delta = 1$ | | | | $\delta = 3$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | OLS | | IV | | OLS | | IV | |
| | | NSA | SA | NSA | SA | NSA | SA | NSA | SA |
| 1.00 | 50 (L) | 0.063 | - | 0.051 | - | 0.064 | - | 0.058 | - |
| | 50 (R) | 0.067 | - | 0.062 | - | 0.065 | - | 0.059 | - |
| | 100 (L) | 0.056 | - | 0.046 | - | 0.055 | - | 0.049 | - |
| | 100 (R) | 0.052 | - | 0.054 | - | 0.051 | - | 0.054 | - |
| | 200 (L) | 0.055 | - | 0.044 | - | 0.056 | - | 0.041 | - |
| | 200 (R) | 0.056 | - | 0.047 | - | 0.052 | - | 0.046 | - |
| | 400 (L) | 0.051 | - | 0.039 | - | 0.050 | - | 0.039 | - |
| | 400 (R) | 0.053 | - | 0.046 | - | 0.052 | - | 0.047 | - |
| 1.01 | 50 | 0.641 | 0.594 | 0.453 | 0.415 | 0.612 | 0.560 | 0.428 | 0.392 |
| | 100 | 0.982 | 0.982 | 0.643 | 0.633 | 0.971 | 0.970 | 0.614 | 0.603 |
| | 200 | 1.000 | 1.000 | 0.900 | 0.902 | 1.000 | 1.000 | 0.880 | 0.888 |
| | 400 | 1.000 | 1.000 | 0.980 | 0.981 | 1.000 | 1.000 | 0.970 | 0.972 |
| 0.99 | 50 | 0.715 | 0.664 | 0.500 | 0.494 | 0.699 | 0.655 | 0.487 | 0.455 |
| | 100 | 0.986 | 0.985 | 0.643 | 0.651 | 0.975 | 0.972 | 0.621 | 0.624 |
| | 200 | 1.000 | 1.000 | 0.874 | 0.885 | 1.000 | 1.000 | 0.856 | 0.869 |
| | 400 | 1.000 | 1.000 | 0.976 | 0.980 | 1.000 | 1.000 | 0.969 | 0.973 |
| 0.98 | 50 | 0.987 | 0.981 | 0.899 | 0.898 | 0.978 | 0.974 | 0.877 | 0.860 |
| | 100 | 1.000 | 1.000 | 0.970 | 0.972 | 1.000 | 1.000 | 0.958 | 0.959 |
| | 200 | 1.000 | 1.000 | 0.999 | 0.999 | 1.000 | 1.000 | 0.997 | 0.997 |
| | 400 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.95 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 0.999 | 0.999 |
| | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 400 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

(iii) $T = 7$

| $\alpha$ | $N$ | $\delta = 1$ | | | | $\delta = 3$ | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | OLS | | IV | | OLS | | IV | |
| | | NSA | SA | NSA | SA | NSA | SA | NSA | SA |
| 1.00 | 50 (L) | 0.070 | - | 0.054 | - | 0.066 | - | 0.057 | - |
| | 50 (R) | 0.059 | - | 0.059 | - | 0.062 | - | 0.058 | - |
| | 100 (L) | 0.054 | - | 0.050 | - | 0.054 | - | 0.051 | - |
| | 100 (R) | 0.054 | - | 0.051 | - | 0.057 | - | 0.051 | - |
| | 200 (L) | 0.053 | - | 0.039 | - | 0.053 | - | 0.043 | - |
| | 200 (R) | 0.055 | - | 0.045 | - | 0.054 | - | 0.039 | - |
| | 400 (L) | 0.050 | - | 0.036 | - | 0.050 | - | 0.041 | - |
| | 400 (R) | 0.055 | - | 0.041 | - | 0.052 | - | 0.041 | - |
| 1.01 | 50 | 0.923 | 0.910 | 0.758 | 0.737 | 0.898 | 0.881 | 0.724 | 0.702 |
| | 100 | 1.000 | 1.000 | 0.923 | 0.922 | 1.000 | 1.000 | 0.903 | 0.902 |
| | 200 | 1.000 | 1.000 | 0.995 | 0.996 | 1.000 | 1.000 | 0.991 | 0.993 |
| | 400 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.99 | 50 | 0.953 | 0.935 | 0.803 | 0.793 | 0.940 | 0.925 | 0.783 | 0.770 |
| | 100 | 1.000 | 1.000 | 0.932 | 0.932 | 1.000 | 1.000 | 0.916 | 0.914 |
| | 200 | 1.000 | 1.000 | 0.995 | 0.996 | 1.000 | 1.000 | 0.992 | 0.994 |
| | 400 | 1.000 | 1.000 | 0.999 | 0.999 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.98 | 50 | 1.000 | 1.000 | 0.994 | 0.994 | 1.000 | 0.999 | 0.991 | 0.990 |
| | 100 | 1.000 | 1.000 | 0.999 | 0.999 | 1.000 | 1.000 | 0.999 | 0.999 |
| | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 400 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| 0.95 | 50 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 100 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 200 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 400 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 | 1.000 |

Table 2: Empirical Size and Power of $t_{\hat{\phi}_1}$ for the Case of an Intercept and a Linear Time Trend

Note: 1. The number of iterations is 5,000.

2. NSA: not size-adjusted; SA: size-adjusted.

3. Two figures for each $N$ at $\alpha = 1$ are empirical sizes of left-sided and right-sided tests, respectively. The nominal size is kept at 0.05.

(i) $T = 2$

| $\alpha$ | $N$ | $\delta = 1$ | | $\delta = 3$ | |
|---|---|---|---|---|---|
| | | NSA | SA | NSA | SA |
| 1.00 | 50 (L) | 0.041 | - | 0.039 | - |
| | 50 (R) | 0.087 | - | 0.086 | - |
| | 100 (L) | 0.044 | - | 0.042 | - |
| | 100 (R) | 0.064 | - | 0.068 | - |
| | 200 (L) | 0.031 | - | 0.032 | - |
| | 200 (R) | 0.057 | - | 0.055 | - |
| | 400 (L) | 0.033 | - | 0.036 | - |
| | 400 (R) | 0.051 | - | 0.047 | - |
| 1.01 | 50 | 0.265 | 0.188 | 0.255 | 0.173 |
| | 100 | 0.326 | 0.285 | 0.317 | 0.271 |
| | 200 | 0.526 | 0.499 | 0.516 | 0.500 |
| | 400 | 0.703 | 0.699 | 0.685 | 0.690 |
| 0.99 | 50 | 0.179 | 0.200 | 0.176 | 0.200 |
| | 100 | 0.265 | 0.281 | 0.257 | 0.280 |
| | 200 | 0.404 | 0.477 | 0.396 | 0.469 |
| | 400 | 0.653 | 0.700 | 0.637 | 0.682 |
| 0.98 | 50 | 0.430 | 0.461 | 0.413 | 0.447 |
| | 100 | 0.642 | 0.659 | 0.614 | 0.636 |
| | 200 | 0.870 | 0.898 | 0.836 | 0.876 |
| | 400 | 0.978 | 0.985 | 0.968 | 0.974 |
| 0.95 | 50 | 0.948 | 0.957 | 0.931 | 0.940 |
| | 100 | 0.992 | 0.992 | 0.988 | 0.989 |
| | 200 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 400 | 1.000 | 1.000 | 1.000 | 1.000 |

(ii) $T = 3$

| $\alpha$ | $N$ | $\delta = 1$ | | $\delta = 3$ | |
|---|---|---|---|---|---|
| | | NSA | SA | NSA | SA |
| 1.00 | 50 (L) | 0.039 | - | 0.040 | - |
| | 50 (R) | 0.081 | - | 0.083 | - |
| | 100 (L) | 0.038 | - | 0.038 | - |
| | 100 (R) | 0.068 | - | 0.071 | - |
| | 200 (L) | 0.036 | - | 0.037 | - |
| | 200 (R) | 0.054 | - | 0.050 | - |
| | 400 (L) | 0.037 | - | 0.038 | - |
| | 400 (R) | 0.055 | - | 0.052 | - |
| 1.01 | 50 | 0.309 | 0.225 | 0.299 | 0.215 |
| | 100 | 0.403 | 0.351 | 0.388 | 0.331 |
| | 200 | 0.611 | 0.599 | 0.589 | 0.589 |
| | 400 | 0.805 | 0.799 | 0.781 | 0.777 |
| 0.99 | 50 | 0.206 | 0.240 | 0.205 | 0.237 |
| | 100 | 0.304 | 0.351 | 0.296 | 0.342 |
| | 200 | 0.511 | 0.556 | 0.496 | 0.542 |
| | 400 | 0.749 | 0.780 | 0.729 | 0.759 |
| 0.98 | 50 | 0.513 | 0.554 | 0.488 | 0.528 |
| | 100 | 0.729 | 0.776 | 0.704 | 0.745 |
| | 200 | 0.932 | 0.946 | 0.909 | 0.926 |
| | 400 | 0.992 | 0.994 | 0.986 | 0.991 |
| 0.95 | 50 | 0.972 | 0.979 | 0.958 | 0.967 |
| | 100 | 0.998 | 0.998 | 0.997 | 0.997 |
| | 200 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 400 | 1.000 | 1.000 | 1.000 | 1.000 |

(iii) $T = 7$

| $\alpha$ | $N$ | $\delta = 1$ | | $\delta = 3$ | |
|---|---|---|---|---|---|
| | | NSA | SA | NSA | SA |
| 1.00 | 50 (L) | 0.041 | - | 0.041 | - |
| | 50 (R) | 0.091 | - | 0.089 | - |
| | 100 (L) | 0.040 | - | 0.039 | - |
| | 100 (R) | 0.072 | - | 0.072 | - |
| | 200 (L) | 0.040 | - | 0.039 | - |
| | 200 (R) | 0.055 | - | 0.053 | - |
| | 400 (L) | 0.035 | - | 0.032 | - |
| | 400 (R) | 0.046 | - | 0.046 | - |
| 1.01 | 50 | 0.365 | 0.252 | 0.353 | 0.254 |
| | 100 | 0.495 | 0.426 | 0.474 | 0.412 |
| | 200 | 0.704 | 0.689 | 0.686 | 0.671 |
| | 400 | 0.888 | 0.895 | 0.864 | 0.871 |
| 0.99 | 50 | 0.227 | 0.256 | 0.223 | 0.247 |
| | 100 | 0.360 | 0.405 | 0.349 | 0.386 |
| | 200 | 0.587 | 0.622 | 0.564 | 0.603 |
| | 400 | 0.820 | 0.857 | 0.793 | 0.839 |
| 0.98 | 50 | 0.560 | 0.594 | 0.526 | 0.556 |
| | 100 | 0.802 | 0.832 | 0.767 | 0.792 |
| | 200 | 0.959 | 0.965 | 0.938 | 0.948 |
| | 400 | 0.997 | 0.998 | 0.994 | 0.996 |
| 0.95 | 50 | 0.977 | 0.981 | 0.964 | 0.968 |
| | 100 | 0.998 | 0.998 | 0.996 | 0.997 |
| | 200 | 1.000 | 1.000 | 1.000 | 1.000 |
| | 400 | 1.000 | 1.000 | 1.000 | 1.000 |

Table 3: Empirical Size and Power of Harris and Tzavalis' (1999) Test at $T = 3$ for the Case of an Intercept

Note: 1. The number of iterations is 5,000.

3. NSA: not size-adjusted; SA: size-adjusted.

4. Two figures for each $N$ at $\alpha = 1$ are empirical sizes of left-sided and right-sided tests, respectively. The nominal size is kept at 0.05.

| $\alpha$ | $N$ | Dependent panels | | Independent panels | |
|---|---|---|---|---|---|
| | | NSA | SA | NSA | SA |
| 1.00 | 50 (L) | 0.071 | - | 0.072 | - |
| | 50 (R) | 0.078 | - | 0.076 | - |
| | 100 (L) | 0.065 | - | 0.068 | - |
| | 100 (R) | 0.065 | - | 0.074 | - |
| | 200 (L) | 0.072 | - | 0.084 | - |
| | 200 (R) | 0.072 | - | 0.075 | - |
| | 400 (L) | 0.071 | - | 0.067 | - |
| | 400 (R) | 0.068 | - | 0.078 | - |
| 1.01 | 50 | 0.177 | 0.119 | 0.086 | 0.059 |
| | 100 | 0.300 | 0.252 | 0.086 | 0.059 |
| | 200 | 0.813 | 0.758 | 0.090 | 0.068 |
| | 400 | 1.000 | 1.000 | 0.106 | 0.075 |
| 0.99 | 50 | 0.033 | 0.022 | 0.081 | 0.058 |
| | 100 | 0.011 | 0.007 | 0.082 | 0.061 |
| | 200 | 0.000 | 0.000 | 0.105 | 0.063 |
| | 400 | 0.000 | 0.000 | 0.095 | 0.070 |
| 0.98 | 50 | 0.007 | 0.004 | 0.091 | 0.065 |
| | 100 | 0.000 | 0.000 | 0.098 | 0.072 |
| | 200 | 0.000 | 0.000 | 0.127 | 0.081 |
| | 400 | 0.000 | 0.000 | 0.127 | 0.100 |
| 0.95 | 50 | 0.000 | 0.000 | 0.120 | 0.091 |
| | 100 | 0.000 | 0.000 | 0.153 | 0.121 |
| | 200 | 0.000 | 0.000 | 0.217 | 0.152 |
| | 400 | 0.000 | 0.000 | 0.273 | 0.225 |

Table 4: Results of Estimation and Panel Unit Root Test

Note: 1. For the full-sample, $(T, N)$ =(2, 15226); and for the sample from four-year college graudates, $(T, N)$ =(2, 9568).

2. (**): significant at the 1% level; SE.: standard error

(i) OLS-based inference

| | $\tilde{\phi}_1$ | SE | $t_{\tilde{\phi}_1}$ |
|---|---|---|---|
| Full-sample | 0.6961 | 0.0071 | -42.68** |
| Four-year college | 0.6567 | 0.0099 | -34.82** |

(ii) IV-based inference

| | # of instruments | $\hat{\phi}_1$ | SE | $t_{\hat{\phi}_1}$ |
|---|---|---|---|---|
| Full-sample | 2,650 | 0.6994 | 0.0136 | -22.18** |
| Four-year college | 1,950 | 0.6603 | 0.0175 | -19.39** |