PANEL UNIT ROOT TESTS BASED ON SAMPLE VARIANCE

XIJIA LIU[†]

ABSTRACT. In this paper, we propose a novel way to test for unit root in a panel setting. The new tests are based on the observation that the trajectory of the cross sectional sample variance behaves differently for stationary than for non-stationary processes. Three different test statistics are considered and their limiting distributions are derived. Interestingly, one of the statistics has a non-standard limiting distribution which can be described in terms of functionals of a Gaussian process. A small scale simulation study indicates that our proposed tests have good power properties, quite close to the test of Levin, Lin and Chu (2002)(LLC). However, the empirical size of one of our tests is better than LLC when T is small and N is large, and this suggest a good property for unit root tests in micro panels. In addition, the study also suggests that our tests are robust to cross section dependence for a particular covariance structure.

1. Introduction

In time series econometrics, the construction of statistical tests, with good size and power properties, able to distinguish between stationary and non-stationary process is of special interest. However, it is well known that existing test statistics, for example the augmented Dickey-Fuller (ADF) test by Dickey and Fuller (1979) and semi-parametric test procedures of Phillips and Perron (1988), have low power in small sample sizes. Furthermore, in general, such tests have non-standard limiting distributions which can be described in terms of functionals of Brownian motion.

To overcome these drawbacks, the so called 'first generation' panel unit root tests, for example, Levin, Lin and Chu (2002) and Im, Pesaran and Shin (2003) proposed to add a cross-sectional dimension to univariate unit root tests. There are two primary benefits: Firstly, the power of unit root tests can be increased by borrowing the strength from cross sectional units. Secondly, asymptotic normality can be obtained for the test statistics, even though the asymptotic theory is considerably more complicated due to the introduction of a cross sectional dimension. However, the mythology of power improving was broken by cross section dependence very soon. Some studies, for example, O'Connell (1998), show that the nice size and power properties will be destroyed if the convenient but unrealistic assumption of cross sectional independence is included. To overcome this problem, the second generation panel unit root tests emerged. The most popular idea is to apply the factor structure approach, e.g. Moon and Perron (2004), Bai and Ng (2004) and Pesaran (2005). Intuitively, the dependence of cross section units is removed by the factor model and then methods based on cross section independence can be applied. However, no matter if the first or second generation panel unit root tests are used, most of them barely do the test beyond the Dickey-Fuller test. So a natural question is coming up

Key words and phrases. Panel unit root test, Micro panels, Robustness, Cross section dependence, Asymptotic results.

The author wish to gratefully thank Rolf Larsson and Søren Johansen for the valuable comments and suggestions. [†]Uppsala University. Address correspondence to Xijia Liu, Department of Statistics, Uppsala University, Box 513, 751 20 Uppsala, Sweden; e-mail: xijia.liu@statistik.uu.se.

immediately: since after adding cross-sectional dimension the data structure is changed, can we try to do test for unit roots in an entirely different way?

Before announcing the answer, let us recall a very special and tricky situation in time series econometrics. One uppermost difficulty is that we can not repeatedly observe the realization of one process several times. For example, it is unfeasible to observe the GDP of China from 1949 to 2012 twice. That means we can not study the properties of mean and variance of economic variables at any given time point directly. However this impossible issue becomes "feasible" in a panel setting, since one economic variable can be repeatedly observed from different regions or countries. Assuming the economic variable has some similar properties among all cross section units, then the properties of mean and variance of the process at each time point can be applied. In this paper, based on that thinking, we propose a new way to take advantage of the information from cross section units to improve the power properties of the unit root test. Thereby the basic idea is the observation that the trajectory of the cross sectional sample variance behaves differently for stationary than for non-stationary processes. More specifically, the trajectory rises along a straight line for non-stationary, but almost remains at the same level for stationary processes.

Given this idea, three test statistics are considered. We should emphasize that our study belongs to the framework of the first generation panel unit root tests, since this is just the beginning of this research. And we will consider how to extend it to more general setting in the next step. A small scale simulation study indicates that our proposed tests statistics have good power properties, similar to the test of Levin, Lin and Chu (2002). And it also shows that the empirical size of one of our test statistics has better performance than LLC when N is very much larger than T and without considering cross sectional dependence. This suggest a good property when doing unit root tests in micro panels. (For the example of research on unit root test in micro panels, see Harris and Tzavalis (1999) and Bond, Nauges and Windmeijer (2002)). Furthermore, the study suggests that our test statistics are robust to the cross section dependence for a particular covariance structure. About the limiting distribution of our test, asymptotic normality can still be obtained, but a new non-standard limiting distribution has to be introduced for one of our test statistics in order to keep the robustness to cross section dependence of our test statistics. Additionally this non-standard limiting distribution is invariant with the variance of error. At last, it is worth to mention the variance ratio test proposed by Lo and MacKinlay (1988), since the inherent idea is similar to ours. However the way of our methods applying this idea is dramatically different from the variance ratio. Comparing to the variance ratio test, although our tests only can be applied in the panel setting, the way of applying this idea is more direct.

A word on notation. $\doteq$ denotes asymptotic equivalence. The symbols $\xrightarrow{P}$ and $\xrightarrow{L}$ will be used to signify convergence in probability and in law respectively. And $T, N \to \infty$ signifies the sequential convergence, i.e. let $T \to \infty$ first, then $N \to \infty$. And $(N, T) \to \infty$ denotes the joint convergence, i.e. let N and T simultaneously converges to infinity. The rest of the paper is organized as follows. In Section 2, we describe the motivation for our tests, the data generating process (DGP) and test statistics. The robustness to cross section dependence is also discussed. The asymptotic results for our test statistics are presented in Section 3. Section 4 reports the results of a small scale simulation study to illustrate the properties and the robustness of our test statistics. Section 5 provides the conclusion and discussion of the proposed panel unit root test. All proofs and derivations are included in the Appendix.

2. Panel Unit root test

2.1. **Motivation.** It is well known that the first-order autoregressive process is weakly stationary for a suitable choice of distribution of its initial value, provided that the autoregressive coefficient is less than one in absolute value. Consider the simple $AR(1)$ process:

$$y_t = \rho y_{t-1} + \varepsilon_t$$

where $t = 1, ..., T$ and $\varepsilon_t \overset{iid}{\sim} N\left(0, \sigma^2\right)$. By recursion in t

$$y_t = \rho^t y_0 + \sum_{i=1}^{t} \rho^{t-i} \varepsilon_i$$

hence

$$E\left(y_t\right) = \rho^t E\left(y_0\right)$$

and

$$Var\left(y_t\right) = \rho^{2t} Var\left(y_0\right) + \frac{1-\rho^{2t}}{1-\rho^2}\sigma^2$$

If $|\rho| < 1$, we can find a distribution for the initial value such that $E\left(y_t\right)$ and $Var\left(y_t\right)$ do not change over time. For example, let

$$E\left(y_0\right) = 0, \tag{1}$$

and

$$Var\left(y_0\right) = \frac{\sigma^2}{1-\rho^2}. \tag{2}$$

Then $y_0 \sim N(0, \frac{\sigma^2}{1-\rho^2})$. This can be illustrated by a simple simulation study. Consider the DGP

$$y_t = \begin{cases} y_0 + \sum_{i=1}^{t} \varepsilon_i & \text{if } \rho = 1 \\ \rho^t y_0 + \sum_{i=1}^{t} \rho^{t-i} \varepsilon_i & \text{if } |\rho| < 1 \end{cases} \tag{3}$$

where the distribution of the initial value y_0 satisfies equation (1) and (2), if $|\rho| < 1$, otherwise set $y_0 = 0$. Set $\rho = 1, 0.9, 0.5, 0.1$ respectively. And for each ρ repeat generating a sequence $y_t, T = 10$, by the DGP $N = 100$ times. Then we get a data matrix $\{y_{it}\}$, for $i = 1, 2, ..., N$ and $t = 1, 2, ..., T$. For each t, we calculate the sample variances which should be close to each other except for the case when $\rho = 1$. In other words, the sample variance will increase in t in the unit root case, and remain at the same level in the stationary case, cf. the left panel of Figure 1. However, it is unfeasible to give a distribution to the initial value in an empirical study. Thus, now consider the DGP when setting the initial value to 0. We still can get a very "nice" plot by which the unit root process still can be identified, see the left panel of Figure 1. We can see that for the non stationary process the sample variance increases rather quickly and goes through on a straight line, but for the stationary process this is not so. Then one may ask if we can apply this phenomenon to panel unit root testing? The answer is positive and three natural test statistics will be introduced in the next subsection.

2.2. **Model and test statistics.** For simplicity, consider the DGP:

$$y_{it} = \rho_i y_{it-1} + \varepsilon_{it} \tag{4}$$

where $i = 1, ..., N$ and $t = 1, ..., T$.

Assumption 1. *Let y_{it} ($i = 1, ..., N$ and $t = 1, ..., T$) be given by Equation (4), suppose that*

1. *The initial values are fixed as 0, i.e. $y_{i0} = 0$ for all i,*
2. *ε_{it} are independent standard Gaussian noise, i.e. $\varepsilon_{it} \sim \mathcal{N}(0, \sigma^2)$*

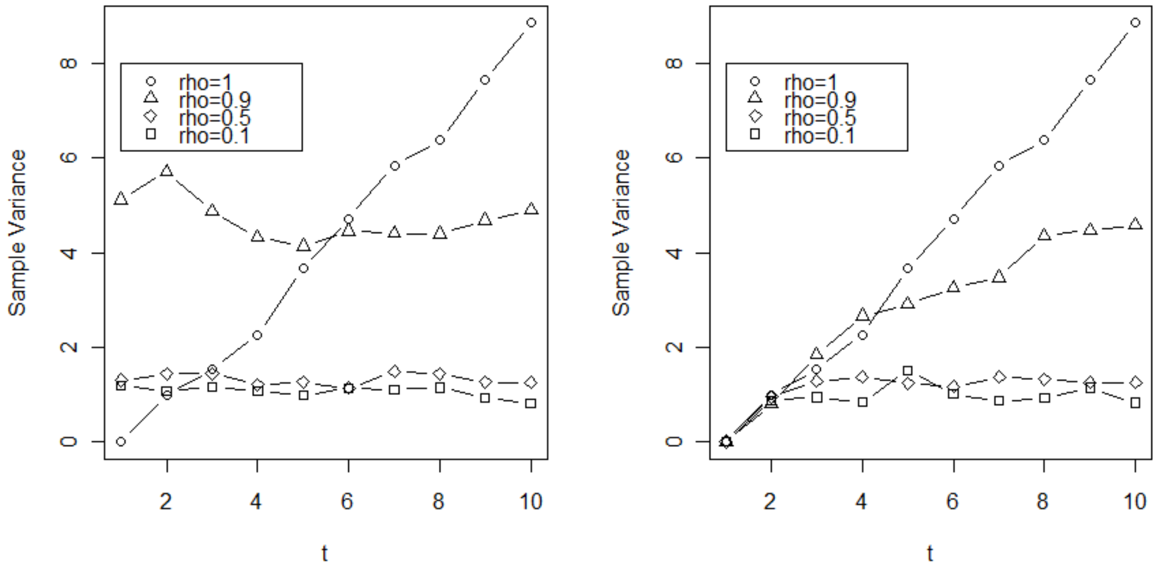

FIGURE 1. The plots of sample variances at each time points. Left panel: The distribution of the initial value is given according to Equation (1) and (2). Right panel: The initial values are fixed as 0

Remark 1. *Here we only consider the simplest case which is random walk without drift term under the null hypothesis. Despite the lack of realism in an empirical study, it is still worth to mention that the properties of cross section sample variances at each time point will not be affected by the drift terms if we assume the drift terms are identical through all cross section units. More generally, we consider the model as*

$$y_{it} = D_t + \rho y_{i,t-1} + \epsilon_{it}$$

where D_t denotes the deterministic terms. Thus the cross section sample variance at each time is invariant with D_t, no matter how complicated it is.

Consider the null hypothesis

$$H_0 : \rho_1 = ... = \rho_N = 1$$

and alternative hypothesis

$$H_{1a} : \rho_1 = ... = \rho_N \equiv \rho \text{ and } |\rho| < 1$$

or

$$H_{1b} : |\rho_i| < 1$$

for all i. The cross sectional sample variance at time t is

$$S_t^2 = \frac{1}{N}\sum_{i=1}^{N}(y_{it} - \overline{y}_t)^2 = \frac{1}{N}\sum_{i=1}^{N} y_{it}^2 - \overline{y}_t^2$$

Next we propose three test statistics motivated by the different behavior of the unit root process and a covariance stationary process. Under the null, we have

$$y_{it} = \sum_{j=1}^{t} \varepsilon_{ij} \sim N\left(0, t\sigma^2\right). \tag{5}$$

While under the alternative hypothesis,

$$y_{it} = \sum_{j=1}^{t} \rho^{t-j} \epsilon_{ij} \sim N\left(0, \frac{1-\rho^{2t}}{1-\rho^2}\sigma^2\right). \tag{6}$$

It is obvious that the sample variance will diverge for the unit root case, and converge to $\frac{1}{1-\rho^2}$ for a stationary process. Since the variance of y_{it} increases faster in a unit root process than in a stationary process, we can consider the sample variance of $\left\{S_t^2\right\}_{t=1}^{T}$ as a test statistic:

$$\psi = \frac{\sqrt{\sum_{t=1}^{T}\left(S_t^2 - \overline{S^2}\right)^2 / T}}{\overline{S^2}}$$

where $\overline{S^2} = \sum_{t=1}^{T} S_t^2 / T$. In ψ, the variance of the sample variance at each time point is normalized by the mean. There are two benefits of this normalization. First, it makes sure that the test statistic ψ will converge as $t \to \infty$, since the variance of the sample variance at each time point increases with the increase of t. Another benefit is related to the robustness of cross section dependence, and it will be discussed in the next subsection. From the left panel of Figure 1, we also find that the circle points can be fitted by a straight line very well under the null hypothesis, however, it is not so for the stationary case. This is not very difficult to explain by Equations 5 and 6. Thus, we can consider to apply the OLS method to fit a line to go through the variance at each time point and use R^2 statistics and F statistics to measure the goodness of fit. Then we have another two test statistics. That is we consider the auxiliary model,

$$S_t^2 = \beta_0 + \beta_1 t + u_t \tag{7}$$

and the statistics

$$\psi_{R^2} = \frac{\widehat{\beta}_1^2 \sum_{t=1}^{T}\left(t - \bar{t}\right)^2}{\sum_{t=1}^{T}\left(S_t^2 - \overline{S^2}\right)^2}$$

and

$$\psi_F = \frac{\widehat{\beta}_1^2 \sum_{t=1}^{T}\left(t - \bar{t}\right)^2}{\sum_{t=1}^{T}\left(S_t^2 - \widehat{S_t^2}\right)^2 / (T-2)}$$

where $\widehat{S_t^2} = \widehat{\beta}_0 + \widehat{\beta}_1 t$. Intuitively, the statistic ψ_{R^2} should be close to 1 under the null hypothesis, but close to 0 for the stationary case. Similarly, under the null hypothesis, the statistic ψ_F should be larger than in the case when the processes are stationary. So we can expect a good power performance when we apply those test statistics to detect unit roots.

Remark 2. *We emphasize that there is no exact meaning of u_t in Equation 7, and we can not give any proper assumptions on it. The test statistic ψ_F is not used to do any hypothesis test in the scenario of regression analysis, for example, $H_0 : \beta_1 = 0$. We just want to use this "F statistic" to measure the distance between a unit root process and stationary process. In other words, we alert the reader that the distribution (asymptotic distribution) of test statistic ψ_F is derived under $H_0 : \rho = 1$ rather than $H_0 : \beta_1 = 0$.*

Remark 3. *In fact, by the indication of the simulation study we can see that the test statistics* ψ_F *and* ψ_{R^2} *have the same size adjusted power. However we study them separatively since different asymptotic distributions can be derived for each of them.*

2.3. **Robustness on cross section dependence.** As we have discussed in the introduction, the main drawback of the first generation panel unit root tests is that those tests are valid only when the cross sections are uncorrelated. However, for a particular situation this drawback can be solved by our methods. Again, let us do a simple simulation study. Using the same DGP, Equation (3), but there are some changes in the assumptions about the error term. Assume $\epsilon_{i\cdot} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$ where

$$\mathbf{\Sigma} = \begin{pmatrix} 1 & \tau & \cdots & \tau \\ \tau & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \tau \\ \tau & \cdots & \tau & 1 \end{pmatrix} \tag{8}$$

Setting the initial value $y_0 = 0$, $\rho = 1, 0.9$, and $\tau = 0, 0.8$ respectively. And set the sample size $T = 30$ and the number of individuals $N = 100$. In the same way, we calculate the sample variances and plot them. From Figure 2 (L), we can see that the variances at each time point are varied from 0 to 30 for the unit root process, however the variances are only changed from 0 to 5 under the cross section dependence settings. That is to say, if we only consider the variance of those sample variances at each time point, then the size properties of our test statistics ψ will be destroyed by the cross section dependence. As we mentioned before, that is the second benefit of the normalization of the test statistic ψ, i.e. the distribution of our test statistic ψ is invariant from adding this type of cross section dependence. Comparing these two interesting graphs again, we can see that there is no difference between the alteration pattern of the sample variances in the two different settings. This "nice" picture motivates that we could obtain the same robustness to cross section dependence when we use another two statistics which are based on the measure of goodness of fit. Intuitively, the robustness of our tests can be explained by the cross section demeaning procedure which is suggested by Pedroni (1999). More specifically, all of our test statistics are based on sample variance at each time point, and sample variance is calculated by the sum of squares of $y_{it} - \bar{y}_t$ which is the cross section demeaned data. And we can see that the LLC test based on cross section demeaned data also has this robustness by a simulation study, and we call it demeaned LLC (DLLC). More studies about robustness will be shown in Section 4.

2.4. **About the nuisance parameters.** As we have discussed in Remark 1, under the strong homogeneous assumption, no matter how complicated the model is, the sample variance at each time point will not be affected by any nuisance parameters. More specifically, for example we consider the true DGP as

$$y_{it} = D_t + \rho y_{i,t-1} + u_{it} \tag{9}$$

where $u_{it} = D(L)\epsilon_{it}$, $D(L)$ is a lag polynomial and $\epsilon \sim N(0, \sigma^2)$. Namely the deterministic term D_t, the lag polynomial $D(L)$ and the variance of error terms σ^2 are all homogenous among cross section units. Then at least, the limiting distribution of ψ_F does not depend on the nuisance parameters. However, this is quite strong and unrealistic assumption in the empirical study. If we relax the homogeneity assumption, we need to consider the estimation of nuisance parameters. For example, the variance of error terms can be estimated consistently from each cross section unit and be used to normalize y_{it} before applying the test statistics. More troubles come from the deterministic terms and it will be discussed in the last section.

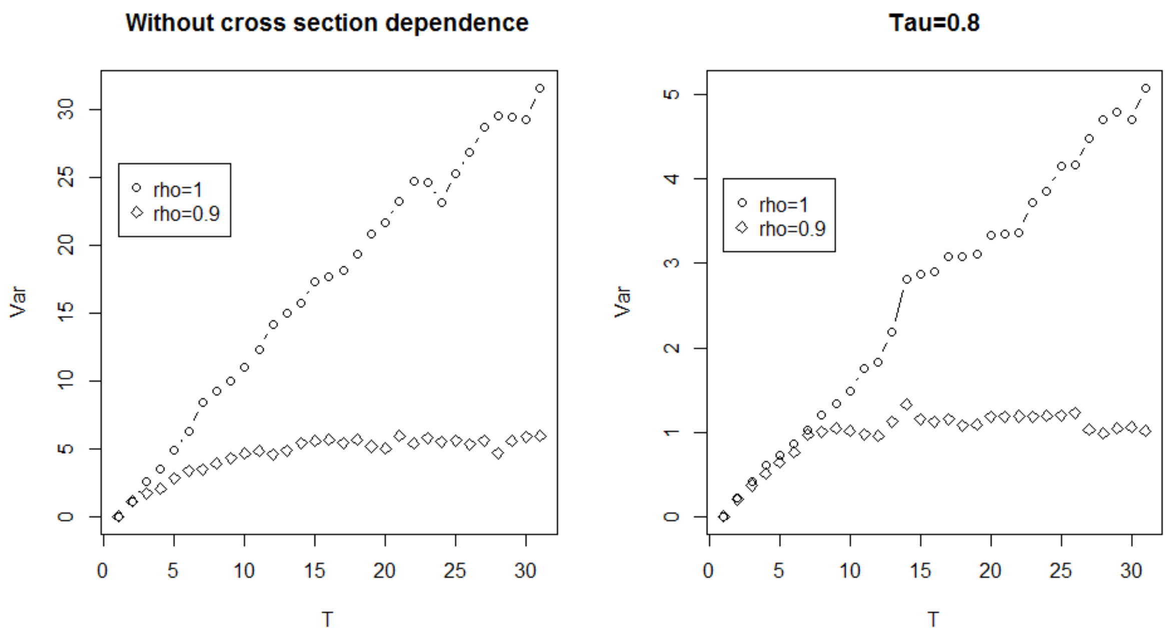

FIGURE 2. The plot of sample variances at each time points. Left panel: The data are generated without considering the cross section dependence. Right panel: The covariance structure is fixed as Equation (8) and set to $\tau = 0.8$

.

3. Asymptotic properties

Introduce some notation which is useful to simplify our test statistics.

$$A = \sum_{t=1}^{T} \left(S_t^2\right)^2 = \frac{1}{N^2} \sum_{t=1}^{T} \left(\sum_{i=1}^{N} y_{it}^2 \right)^2 - \frac{2}{N} \sum_{t=1}^{T} \overline{y}_t^2 \left(\sum_{i=1}^{N} y_{it}^2 \right) + \sum_{t=1}^{T} \overline{y}_t^4$$

$$B = T\left(\overline{S^2}\right)^2 = \left(\frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{N} y_{it}^2 - \sum_{t=1}^{T} \overline{y}_t^2 \right)^2 /T$$

$$C = \frac{1}{N} \sum_{i=1}^{N} \sum_{t=1}^{T} t y_{it}^2 - \sum_{t=1}^{T} t\overline{y}_t^2 - \frac{1}{N} \frac{T+1}{2} \sum_{i=1}^{N} \sum_{t=1}^{T} y_{it}^2 + \frac{T+1}{2} \sum_{t=1}^{T} \overline{y}_t^2$$

Then our test statistics can be rewritten as

$$\psi = \sqrt{\frac{A-B}{B}}$$

$$\psi_{R^2} \doteq \frac{12C^2}{T^3 (A-B)}$$

$$\psi_F \doteq \frac{12C^2 T}{T^3 (A-B) - 12C^2}$$

see the Appendix. Before constructing the limiting distribution for those three test statistics, we need to do some preparations. Some limiting results which are used to describe the asymptotic behavior of each part are contained in the following propositions. It should be noted that the results from iv to vii are sequential asymptotic results, i.e. let $T \to \infty$ first, then $N \to \infty$.

Lemma 1. *Given Assumption 1, we have the following asymptotic results*

- **i:** $\frac{A}{\sigma^4 T^3} \underset{T\to\infty}{\overset{L}{\to}} \int_0^1 \left(S_W^2\right)^2 dr$ *where* $S_W^2 = \frac{1}{N}\sum_{i=1}^N W_i(r)^2 - \left(\frac{1}{N}\sum_{i=1}^N W_i(r)\right)^2$
- **ii:** $\frac{B}{T^3\sigma^4} \underset{T\to\infty}{\overset{L}{\to}} \left(\int_0^1 S_W^2 dr\right)^2$
- **iii:** $\frac{12C^2}{T^6\sigma^4} \underset{T\to\infty}{\overset{L}{\to}} 3\left(\int_0^1 (2r-1) S_W^2 dr\right)^2$
- **iv:** $\frac{A}{\sigma^4 T^3} \underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{3}$
- **v:** $\frac{1}{T^3\sigma^4} B \underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{4}$
- **vi:** $\frac{C}{T^3\sigma^2} \underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{12}$
- **vii:** $\sqrt{N}\left(\frac{C}{T^3\sigma^2} - \frac{1}{12}\right) \underset{T,N\to\infty}{\overset{L}{\to}} \mathcal{N}\left(0, \frac{1}{60}\right)$

Firstly, a quick limit result of test statistic ψ_{R^2} is followed by the results in Lemma 1.

Theorem 1. *Given Assumption 1,*

i: *In view of Lemma 1, the sequential limiting distribution of* ψ_{R^2} *is*

$$\sqrt{N}\left(\psi_{R^2} - \frac{1}{12}/\frac{A-B}{T^3\sigma^4}\right) \underset{T,N\to\infty}{\overset{L}{\to}} \mathcal{N}\left(0, \frac{48}{5}\right)$$

where $\frac{1}{12}/\frac{A-B}{T^3\sigma^4} \underset{T,N\to\infty}{\overset{P}{\to}} 1.$

ii: *The joint limiting distribution of* ψ_R^2 *is equivalent to the sequential limiting distribution in i.*

However for the test statistic ψ and ψ_F, the limit distribution is not very easy to derive. And we only create the sequential limit distribution for the other two. For this, we need to prove the following two lemmas first.

Lemma 2. *Suppose* $W_i(t)$ *are independent Brownian motions,* $i = 1, 2, ..., n$, *then the process*

$$P_n(t) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^n W_i(t)^2 - \left(\frac{1}{n}\sum_{i=1}^n W_i(t)\right)^2 - t\right)$$

converges weakly to the Gaussian process $P(t)$ *on* $C[0,1]$. *For finite points* $0 \le t_1 < t_2 < ... < t_m \le 1$, *the finite dimensional distribution of* $P_n(t)$ *is* $N(\mathbf{0}, \mathbf{\Sigma_m})$ *where*

$$\mathbf{\Sigma}_m = \begin{pmatrix} 2t_1^2 & 2t_1^2 & \cdots & 2t_1^2 \\ & 2t_2^2 & \cdots & 2t_2^2 \\ & & \ddots & \vdots \\ & & & 2t_m^2 \end{pmatrix}$$

Lemma 2 shows that the stochastic process $P_n(t)$ will converge to some kind of Gaussian process $P(t)$ with certain covariance matrix as $n \to \infty$. In addition, for any real-valued measurable functional $f(\cdot)$ defined on the space $C[0,1]$, we have $f(P_n(\cdot)) \to f(P(\cdot))$ as $n \to \infty$ weakly. And we can see from next theorem that the scaled statistic ψ_F can be written as a function of the stochastic process $P_n(t)$ as $T \to \infty$.

Remark 4. *Note that the covariance matrix of the finite dimensional distribution of the Gaussian process $P(r)$ is very close to that of a Brownian motion except that t is replaced by t^2. Thus $P(r)$ is a time transformed Brownian motion, and $P(r)/\sqrt{2}$ is just Brownian motion evaluated at t^2.*

Lemma 3. *Given Assumption 1, we have*

$$\sqrt{N}\left(\frac{A}{\sigma^4 T^3}-\frac{1}{3}\right)\underset{T,N\to\infty}{\overset{L}{\to}}\int_0^1 2rP\left(r\right)dr$$

where $P(\cdot)$ is the Gaussian process in Lemma 2.

Remark 5. *By Lemma 3 we know that by proper normalization, A converges sequentially in distribution to random variable $\int_0^1 2rP\left(r\right)dr$. In fact, it describes a random variable that has a $\mathcal{N}\left(0,2/3\right)$ distribution, since it can be rewritten as $\sqrt{2}\int_0^1 P\left(r\right)/\sqrt{2}dr^2$ and this can be represented as $\sqrt{2}\int_0^1 W(t)dt$ (see Remark 4), where $\int_0^1 W(t)dt$ describes a normal random variable with zero mean and variance $1/3$.*

Theorem 2. *In view of Lemmas 1 and 3 and the statements in Remark 5, the sequential limiting distribution of ψ is*

$$\sqrt{N}\left(\psi^2+1-\frac{1/3}{B/T^3\sigma^4}\right)\underset{T,N\to\infty}{\overset{L}{\to}}\mathcal{N}\left(0,\frac{32}{3}\right)$$

where $\frac{1/3}{B/T^3\sigma^4}\underset{T,N\to\infty}{\overset{P}{\to}}\frac{4}{3}$.

From Theorem 1 and 2, we can see that the test statistic ψ_{R^2} is asymptotic normal. However to obtain the asymptotic normality we must sacrifice the robustness, for example since $\frac{1}{12}/\frac{A-B}{T^3\sigma^4}$ is not robust to the covariance structure which is equal correlation. And this will be demonstrated by simulation. Furthermore if we want to apply these results, we must estimate the nuisance parameter σ.

Remark 6. *The reader might be misled and argue that since we know that $\frac{1}{12}/\frac{A-B}{T^3\sigma^4}\underset{T,N\to\infty}{\overset{P}{\to}}1$, then we can use*

$$\sqrt{N}\left(\psi_{R^2}-1\right)\underset{T,N\to\infty}{\overset{L}{\to}}\mathcal{N}\left(0,\frac{48}{5}\right)$$

instead of the limiting distribution in Theorem 1, such that the robustness can be retained. However this argument is not true, for this see Claim 1 in Appendix. For the test statistic ψ, a similar statement can be given.

Comparing with this drawback, we will see that a more useful limit result can be obtained for ψ_F. The merit is that to obtain the limiting distribution, we do not need to modify our test statistic except by multiplying with some scalars.

Theorem 3. *By Lemma 1 and 2, the sequential limiting distribution of ψ_F is*

$$\frac{\psi_F}{NT}\underset{T,N\to\infty}{\overset{L}{\to}}\frac{1}{12\left\{\int_0^1 P\left(r\right)^2dr-\left[\int_0^1 P\left(r\right)dr\right]^2-3\left[\int_0^1\left(2r-1\right)P\left(r\right)dr\right]^2\right\}}$$

where $P(\cdot)$ is Gaussian process in Lemma 2.

Table 1. Empirical critical values for ψ_F/NT

N	T	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
25	25	0.068	0.104	0.147	0.214	1.490	1.860	2.257	2.805
	50	0.072	0.111	0.156	0.226	1.532	1.898	2.264	2.734
	100	0.075	0.117	0.165	0.237	1.566	1.930	2.306	2.808
	250	0.081	0.121	0.169	0.241	1.586	1.945	2.320	2.779
50	25	0.105	0.145	0.190	0.256	1.495	1.881	2.228	2.742
	50	0.117	0.159	0.206	0.275	1.506	1.841	2.174	2.633
	100	0.119	0.163	0.212	0.281	1.543	1.886	2.212	2.642
	250	0.119	0.165	0.214	0.285	1.562	1.902	2.233	2.679
100	25	0.134	0.173	0.215	0.277	1.438	1.783	2.132	2.624
	50	0.139	0.181	0.228	0.294	1.486	1.793	2.106	2.528
	100	0.144	0.189	0.237	0.305	1.526	1.845	2.150	2.564
	250	0.147	0.191	0.241	0.312	1.551	1.876	2.189	2.597
250	25	0.150	0.188	0.231	0.292	1.436	1.765	2.100	2.582
	50	0.159	0.201	0.246	0.312	1.478	1.794	2.098	2.497
	100	0.165	0.209	0.254	0.321	1.528	1.840	2.141	2.526
	250	0.168	0.212	0.260	0.328	1.539	1.861	2.170	2.578

4. Simulation study

In this section, we discuss the results of several Monte Carlo (MC) experiments. Firstly, the table of critical values of the test statistic ψ_F are produced by a MC simulation which is based on 100,000 replicates is presented in Table 1.

The second set of simulation experiments investigates the size adjusted power of our tests. We also compare the size adjusted power of our tests with Levin, Lin and Chu (2002)(LLC). The true DGP is $y_{i,t} = \rho y_{i,t-1} + \epsilon_{i,t}$, where $\epsilon_{i,\cdot} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma})$. The covariance matrix is defined by Equation (8), and we set $\tau = 0$ and 0.8 respectively. For the alternative hypothesis, we set $\rho_i = 0.9$ and 0.95 for all i. The size adjusted power for each test is presented in Table 2. From Table 2 we can see that the powers of our tests all are very close to LLC. More interestingly, the power of LLC will be drawn down a little bit by adding the cross section dependence, but the power properties of our tests escape unscathed from releasing this improper assumption by equal correlation assumption. As we have discussed, this property is due to the demeaning procedure. Thus we also exam the power of DLCC (demeaning the data first, then do the LLC test), and from Table 2 we can see that DLLC has the same good performance.

Finally, the size property and the robustness to the cross section dependence is examined. We generate the data from the null hypothesis by the same DGP. Again, we use Equation (8) to describe the cross section dependence. For $\tau = 0$ and 0.8, the empirical size of our test statistics, LLC and DLLC are presented in Table 3. From this table we can see that the empirical size of ψ is the best among those five statistics when $T = 25$, $N = 100$ and $\tau = 0$. However it will increase again as T increases and for fixed $N = 100$, although it is still the best. Intuitively, the main reason is that the asymptotic result of the test ψ is based on the sequential limiting distribution and some condition of the ratio of N and T is missed. Then for fixed $T = 50$ and N increased to 200, the empirical size of ψ decreases to 0.05 again and it is still better than LLC and DLLC. Thus this also provides us a good property for unit root tests in micro panels that the asymptotic approximation of ψ works very well when N is very much larger than T. And this

TABLE 2. Size adjusted power: Each table entry (based on 10 000 Monte Carlo replications) reports the frequency of the simulation in which the false null is rejected using the empirical critical value for $\alpha = 0.05$ which presented in Table 1

		$\tau = 0$									
		LLC		ψ		ψ_{R^2}		ψ_F		DLLC	
T	N	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.95$
25	25	1.000	0.960	0.909	0.637	0.927	0.613	0.927	0.613	1.000	0.960
	50	1.000	1.000	0.998	0.910	0.990	0.824	0.990	0.824	1.000	1.000
	100	1.000	1.000	1.000	0.996	1.000	0.946	1.000	0.946	1.000	1.000
50	25	1.000	1.000	0.985	0.906	0.998	0.931	0.998	0.931	1.000	1.000
	50	1.000	1.000	1.000	0.997	1.000	0.992	1.000	0.992	1.000	1.000
	100	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
100	25	1.000	1.000	0.998	0.982	1.000	0.997	1.000	0.997	1.000	1.000
	50	1.000	1.000	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	1.000	1.000
		$\tau = 0.8$									
		LLC		ψ		ψ_{R^2}		ψ_F		DLLC	
T	N	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.95$	$\rho = 0.9$	$\rho = 0.95$
25	25	0.945	0.812	0.894	0.631	0.926	0.623	0.926	0.623	1.000	0.970
	50	0.960	0.845	0.998	0.901	0.991	0.809	0.991	0.809	1.000	0.999
	100	0.968	0.865	1.000	0.996	0.999	0.941	0.999	0.941	1.000	1.000
50	25	0.945	0.997	0.984	0.901	0.998	0.925	0.998	0.925	1.000	1.000
	50	0.998	0.961	1.000	0.997	1.000	0.991	1.000	0.991	1.000	1.000
	100	0.917	0.635	1.000	1.000	1.000	0.999	1.000	0.999	1.000	1.000
100	25	0.998	0.900	0.998	0.984	1.000	0.998	1.000	0.998	1.000	1.000
	50	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	100	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

TABLE 3. Empirical size: Each table entry (based on 10 000 Monte Carlo replications) reports the frequency of the simulation in which the false null is rejected using the critical value based on the asymptotic distribution for $\alpha = 0.05$.

T	N	ψ	ψ_{R^2}	ψ_F	LLC	DLLC	ψ	ψ_{R^2}	ψ_F	LLC	DLLC
		$\tau = 0$					$\tau = 0.8$				
25	25	0.107	0.162	0.139	0.093	0.099	1.000	1.000	0.141	0.527	0.105
	50	0.079	0.133	0.099	0.080	0.086	1.000	1.000	0.098	0.567	0.083
	100	0.053	0.115	0.084	0.072	0.074	1.000	1.000	0.081	0.604	0.075
	200	0.037	0.098	0.078	0.069	0.069	1.000	1.000	0.069	0.618	0.067
50	25	0.112	0.159	0.127	0.098	0.104	1.000	1.000	0.124	0.541	0.095
	50	0.090	0.131	0.083	0.078	0.084	1.000	1.000	0.084	0.578	0.085
	100	0.067	0.111	0.072	0.075	0.071	1.000	1.000	0.068	0.606	0.073
	200	0.052	0.098	0.066	0.069	0.065	1.000	1.000	0.059	0.636	0.068
100	25	0.127	0.157	0.113	0.096	0.096	1.000	1.000	0.113	0.547	0.103
	50	0.100	0.129	0.082	0.079	0.088	1.000	1.000	0.082	0.590	0.084
	100	0.078	0.113	0.063	0.064	0.069	1.000	1.000	0.064	0.612	0.070
	200	0.062	0.087	0.054	0.062	0.061	1.000	1.000	0.056	0.630	0.067

also suggest a further research which we will discuss in the last section. Furthermore from Table 3 we can see that only the asymptotic distribution of ψ_F and DLLC have robustness to the cross section dependence with equal correlation structure. However, for the rest of test statistics, the size properties are destroyed by cross section dependence. Just as we have discussed before, for the test statistics ψ and ψ_{R^2}, in order to achieve asymptotic normality, we must subtract a part which is not robust to the equal correlation. (We want to emphasize that under Assumption 1, the distribution of ψ and ψ_{R^2} are also invariant with equal correlation, and this can be shown by the empirical critical value of them. See tables 4 and 5 in appendix.)

5. Conclusion and further discussion

In this paper, we have introduced a new idea to borrow the strength from cross section units to improve the power properties of unit root tests. Three test statistics are proposed to test the null hypothesis that each time series contains a unit root against the alternative hypothesis that each unit is stationary. So far the sequential limiting results are provided (for test statistic ψ_{R^2}, the joint limiting distribution is equivalent to the sequential limiting distribution). For all the three test statistics, asymptotic normality can be obtained. However, to obtain the asymptotic normal limiting distribution, we must sacrifice the robustness. For this reason, we construct a non standard limiting distribution which can be expressed in terms of functionals on some kind of Gaussian process for the test statistic ψ_F. And another benefit of this limiting distribution is that it is invariant with the variance of the error term. The Monte Carlo simulation studies show that the power of our tests is quite close to LLC's. However, the size property of test statistic ψ is better than LLC, especially when N is very much larger than T, and this provide a good property for unit root test in micro panels. In addition, for the particular covariance structure (equal correlation), simulation studies also indicate that the size and power properties of our tests will not be destroyed by cross section dependence.

Since this paper is just the beginning of this research, all results are discussed in a very simple setting or under some quite strong assumptions. Next, there could be several substantial directions to generalize this idea. Firstly, we could consider a more general stochastic process, for example with non-normal innovations and serial correlations which is non-identical among all cross section units. Another important issue is about the heterogenous deterministic terms.

As we have discussed, if the deterministic terms are not homogenous, then the distribution of test statistics will depend on the variation of the coefficients of deterministic terms, and we will study this problem in the next paper. Except continuing to consider this framework, one could consider the DGP as

$$z_{it} = D_{it} + y_{it} \tag{10}$$

where D_{it} is the deterministic terms and y_{it} is the same as Equation (4). After removing D_{it} for each cross section unit by proper estimations, then our methods can be applied. Thirdly, even though our methods have robustness to a particular covariance structure, they still should be called first generation panel unit root tests. So how to deal with a more general covariance structure of cross section units in this framework could be an interesting problem. Fourthly, as we have mentioned in the section of simulation studies of empirical size, the test statistic ψ works very well when N is very much larger than T. And intuitively, the sample variance at each time point will consistently converge to the population variance. That means we can study the asymptotic property for fixed T and infinitive N, and this is very interesting for the unit root test in a micro panels. Finally, we only create the complete asymptotic results for ψ_R^2. It is still worth and interesting to find the joint limiting distribution for another two.

APPENDIX

Derivation of the representation of test statistics. We have

$$\psi = \frac{\sqrt{\sum_{t=1}^{T}\left(S_t^2 - \overline{S^2}\right)^2 / T}}{\overline{S^2}} = \frac{\sqrt{(A-B)/T}}{\sqrt{B/T}} = \sqrt{\frac{A-B}{B}} \tag{11}$$

$$\begin{aligned}
\psi_{R^2} &= \frac{\widehat{\beta}_1^2 \sum_{t=1}^{T}\left(t-\bar{t}\right)^2}{\sum_{t=1}^{T}\left(S_t^2 - \overline{S^2}\right)^2} = \frac{\left(\sum_{t=1}^{T}\left(t-\bar{t}\right)S_t^2\right)^2 / C_T}{\sum_{t=1}^{T}\left(S_t^2 - \overline{S^2}\right)^2} = \frac{\left(\sum_{t=1}^{T} tS_t^2 - \frac{T+1}{2}\sum_{t=1}^{T} S_t^2\right)^2 / C_T}{\sum_{t=1}^{T}\left(S_t^2\right)^2 - T\overline{S^2}^2} \\
&= \frac{\left(\sum_{t=1}^{T} t\left(\frac{1}{N}\sum_{i=1}^{N} y_{it}^2 - \overline{y}_t^2\right) - \frac{T+1}{2}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N} y_{it}^2 - \overline{y}_t^2\right)\right)^2 / C_T}{\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N} y_{it}^2 - \overline{y}_t^2\right)^2 - \left(\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N} y_{it}^2 - \overline{y}_t^2\right)\right)^2 / T} \\
&= \frac{\left(\underbrace{\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T} t y_{it}^2 - \sum_{t=1}^{T} t\overline{y}_t^2 - \frac{1}{N}\frac{T+1}{2}\sum_{i=1}^{N}\sum_{t=1}^{T} y_{it}^2 + \frac{T+1}{2}\sum_{t=1}^{T}\overline{y}_t^2}_{C}\right)^2 / C_T}{\left(\underbrace{\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N} y_{it}^2\right)^2 - \frac{2}{N}\sum_{t=1}^{T}\overline{y}_t^2\sum_{i=1}^{N} y_{it}^2 + \sum_{t=1}^{T}\overline{y}_t^4}_{A}\right) - \underbrace{\left(\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{N} y_{it}^2 - \sum_{t=1}^{T}\overline{y}_t^2\right)^2 / T}_{B}} \\
&\doteq \frac{12C^2}{T^3\left(A-B\right)}
\end{aligned}$$

where

$$\bar{t} = \frac{T+1}{2}$$

and

$$C_T=\sum_{t=1}^{T}\left(t-\overline{t}\right)^2=\frac{T\left(T+1\right)\left(2T+1\right)}{6}-\frac{T\left(T+1\right)^2}{4}=\frac{\left(T^3-T\right)}{12}$$

At last,

$$\psi_F=\frac{\left(T-2\right)\widehat{\beta}_1^2\sum_{t=1}^{T}\left(t-\overline{t}\right)^2}{\sum_{t=1}^{T}\left(S_t^2-\overline{S^2}\right)^2-\widehat{\beta}_1^2\sum_{t=1}^{T}\left(t-\overline{t}\right)^2}\doteq T\frac{12C^2}{T^3\left(A-B\right)-12C^2}$$

Proof of Lemma 1. Firstly, about part $A:$

$$A=\underbrace{\frac{1}{N^2}\sum_{t=1}^{T}\left(\sum_{i=1}^{N}y_{it}^2\right)^2}_{A_1}-\underbrace{\frac{2}{N}\sum_{t=1}^{T}\overline{y}_t^2\left(\sum_{i=1}^{N}y_{it}^2\right)}_{A_2}+\underbrace{\sum_{t=1}^{T}\overline{y}_t^4}_{A_3}$$

where

$$A_1=\frac{1}{N^2}T^3\sum_{i,j=1}^{N}\left(T^{-1}\sum_{t=1}^{T}\left(\frac{y_{it}}{\sqrt{T}}\right)^2\left(\frac{y_{it}}{\sqrt{T}}\right)^2\right)$$

then by functional CLT and LLN we have

$$\frac{A_1}{\sigma^4T^3}\underset{T\to\infty}{\overset{L}{\to}}\frac{1}{N^2}\sum_{i,j=1}^{N}\left(\int_0^1\left[W_i\left(r\right)\right]^2\left[W_j\left(r\right)\right]^2dr\right) \tag{12}$$

$$=\int_0^1\left(\frac{1}{N}\sum_{i=1}^{N}\left[W_i\left(r\right)\right]^2\right)^2dr\underset{N\to\infty}{\overset{P}{\to}}\frac{1}{3}$$

Next

$$A_2=\frac{2T^3}{N^3}\sum_{i,j,k=1}^{N}\left[T^{-1}\left(\sum_{t=1}^{T}\frac{y_{it}}{\sqrt{T}}\frac{y_{jt}}{\sqrt{T}}\frac{y_{kt}^2}{T}\right)\right]$$

$$\frac{A_2}{\sigma^4T^3}\underset{T\to\infty}{\overset{L}{\to}}\frac{2}{N^3}\sum_{i,j,k=1}^{N}\left(\int_0^1W_i\left(r\right)W_j\left(r\right)\left[W_k\left(r\right)\right]^2dr\right) \tag{13}$$

$$=2\int_0^1\left(\frac{1}{N}\sum_{k=1}^{N}\left[W_k\left(r\right)\right]^2\right)\left(\frac{1}{N}\sum_{i=1}^{N}W_i\left(r\right)\right)^2dr$$

thus

$$\frac{A_2}{\sigma^4T^3}\underset{T,N\to\infty}{\overset{p}{\to}}2\int_0^1\left(EW_i\left(r\right)\right)^2E\left[W_k\left(r\right)\right]^2dr=0$$

Next

$$A_3=\frac{T^3}{N^4}\sum_{i,j,k,l=1}^{N}\left(T^{-1}\sum_{t=1}^{T}\left(\frac{y_{it}}{\sqrt{T}}\frac{y_{jt}}{\sqrt{T}}\frac{y_{kt}}{\sqrt{T}}\frac{y_{lt}}{\sqrt{T}}\right)\right)$$

$$\frac{A_3}{\sigma^4T^3}\underset{T\to\infty}{\overset{L}{\to}}\frac{1}{N^4}\sum_{i,j,k,l=1}^{N}\left(\int_0^1W_i\left(r\right)W_j\left(r\right)W_k\left(r\right)W_l\left(r\right)dr\right)=\int_0^1\left(\frac{1}{N}\sum_{i=1}^{N}W_i\left(r\right)\right)^4dr \tag{14}$$

thus

$$\frac{A_3}{\sigma^4T^3}\underset{T,N\to\infty}{\overset{P}{\to}}\int_0^1\left(EW_i\left(r\right)\right)^4dr=0$$

Then we have

$$\frac{A}{\sigma^4 T^3} \underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{3}$$

Furthermore, by Equation (12) to (14)

$$\frac{A}{\sigma^4 T^3} \underset{T\to\infty}{\overset{L}{\to}} \int_0^1 \left(\frac{1}{N}\sum_{i=1}^{N} \left[W_i\left(r\right)\right]^2 - \left(\frac{1}{N}\sum_{i=1}^{N} W_i\left(r\right)\right)^2\right)^2 dr = \int_0^1 \left(S_W^2\right)^2 dr$$

Next, about part B

$$B = \left(\underbrace{\frac{1}{\sqrt{T}N}\sum_{t=1}^{T}\sum_{i=1}^{N} y_{it}^2}_{B_1} - \underbrace{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\overline{y}_t^2}_{B_2} \right)^2$$

where

$$B_1 = \frac{T^{\frac{3}{2}}}{N}\sum_{i=1}^{N}\left(T^{-2}\sum_{t=1}^{T} y_{it}^2\right)$$

$$\frac{B_1}{T^{\frac{3}{2}}\sigma^2} \underset{T\to\infty}{\overset{L}{\to}} \frac{1}{N}\sum_{i=1}^{N}\left(\int_0^1 \left[W_i\left(r\right)\right]^2 dr\right) = \int_0^1 \frac{1}{N}\sum_{i=1}^{N}\left[W_i\left(r\right)\right]^2 dr \underset{N\to\infty}{\overset{P}{\to}} \frac{1}{2} \tag{15}$$

$$\frac{B_1}{T^{\frac{3}{2}}\sigma^2} \underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{2}$$

and

$$B_2 = \frac{T^{\frac{3}{2}}}{N^2}\sum_{i,j=1}^{N}\left(T^{-1}\sum_{t=1}^{T}\frac{y_{it}}{\sqrt{T}}\frac{y_{jt}}{\sqrt{T}}\right)$$

$$\frac{B_2}{T^{\frac{3}{2}}\sigma^2} \underset{T\to\infty}{\overset{L}{\to}} \frac{1}{N^2}\sum_{i,j=1}^{N}\left(\int_0^1 W_i\left(r\right)W_j\left(r\right)dr\right) = \int_0^1 \left(\frac{1}{N}\sum_{i=1}^{N} W_i\left(r\right)\right)^2 dr \underset{N\to\infty}{\overset{P}{\to}} 0 \tag{16}$$

Then we have

$$\frac{B_1 - B_2}{T^{\frac{3}{2}}\sigma^2} \underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{2}$$

and

$$\frac{B}{T^3\sigma^4} \underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{4}$$

Furthermore by Equations (15) and (16)

$$\frac{B}{T^3\sigma^4} \underset{T,N\to\infty}{\overset{L}{\to}} \left(\int_0^1 \frac{1}{N}\sum_{i=1}^{N}\left[W_i\left(r\right)\right]^2 - \left(\frac{1}{N}\sum_{i=1}^{N} W_i\left(r\right)\right)^2 dr\right)^2 = \left(\int_0^1 S_W^2 dr\right)^2$$

Next, about C

$$C = \underbrace{\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{T}\left(t y_{it}^2 - \frac{T+1}{2} y_{it}^2\right)}_{C_1} - \underbrace{\left(\sum_{t=1}^{T} t\overline{y}_t^2 - \frac{T+1}{2}\sum_{t=1}^{T}\overline{y}_t^2\right)}_{C_2}$$

where

$$C_1 \doteq \frac{1}{2N}\sum_{i=1}^{N}\sum_{t=1}^{T} T\left(2\frac{t}{T}-1\right) y_{it}^2 = \frac{T^3}{2N}\sum_{i=1}^{N} T^{-2}\sum_{t=1}^{T}\left(2\frac{t}{T}-1\right) y_{it}^2$$

$$\frac{C_1}{T^3\sigma^2} \underset{T\to\infty}{\overset{L}{\to}} \frac{1}{2N}\sum_{i=1}^{N}\underbrace{\int_0^1 (2r-1)\left[W_i(r)\right]^2 dr}_{X_i} \tag{17}$$

$$2\sqrt{N}\left(\frac{C_1}{T^3\sigma^2}-\frac{EX_i}{2}\right) \underset{N\to\infty}{\overset{L}{\to}} \mathcal{N}\left(0, VarX_i\right)$$

i.e.

$$\sqrt{N}\left(\frac{C_1}{T^3\sigma^2}-\frac{1}{12}\right) \underset{N\to\infty}{\overset{L}{\to}} \mathcal{N}\left(0, \frac{1}{60}\right)$$

where

$$\begin{aligned} EX_i &= E\int_0^1 (2r-1)\left[W_i(r)\right]^2 dr = \int_0^1 (2r-1) E\left[W_i(r)\right]^2 dr \\ &= \int_0^1 2r^2 - r dr = \frac{1}{6} \end{aligned}$$

and

$$\begin{aligned} EX_i^2 &= \int_0^1\int_0^1 (2r-1)(2s-1) E\left[W_i(r)\right]^2\left[W_i(s)\right]^2 dsdr \\ &= \int_0^1\int_0^1 (2r-1)(2s-1)\left(2\left(\min(r,s)\right)^2 + rs\right) dsdr \\ &= \frac{17}{180} \end{aligned}$$

$$VarX_i = \frac{17}{180} - \left(\frac{1}{6}\right)^2 = \frac{1}{15}$$

where about $E\left[W(r)\right]^2\left[W(s)\right]^2$, let $r > s$, then $W(r) = W(s) + \left[W(r) - W(s)\right]$

$$\begin{aligned} E\left[W(r)\right]^2\left[W(s)\right]^2 &= E\left[W(r)W(s)\right]^2 \\ &= E\left[W(s)^2 + \left[W(r)-W(s)\right]W(s)\right]^2 \\ &= E\left[W(s)^4 + 2\left[W(r)-W(s)\right]W(s)^3 + \left[W(r)-W(s)\right]^2 W(s)^2\right] \\ &= 3s^2 + E\left[W(r)^2 + W(s)^2 - 2W(s)W(r)\right]s \\ &= 2s^2 + rs \end{aligned}$$

thus

$$E\left[W(r)\right]^2\left[W(s)\right]^2 = 2\min(r,s)^2 + rs$$

and

$$\begin{aligned} C_2 &\doteq \sum_{t=1}^{T} t\overline{y}_t^2 - \frac{T}{2}\sum_{t=1}^{T}\overline{y}_t^2 = \sum_{t=1}^{T} t\left(\frac{1}{N}\sum_{i=1}^{N} y_{it}\right)^2 - \frac{T}{2}\sum_{t=1}^{T}\left(\frac{1}{N}\sum_{i=1}^{N} y_{it}\right)^2 \\ &= T^3\frac{1}{N^2}\sum_{i,j=1}^{N}\left(T^{-1}\sum_{t=1}^{T}\frac{t}{T}\frac{y_{it}}{\sqrt{T}}\frac{y_{jt}}{\sqrt{T}}\right) - \frac{T^3}{2}\frac{1}{N^2}\sum_{i,j=1}^{N}\left(T^{-1}\sum_{t=1}^{T}\frac{y_{it}}{\sqrt{T}}\frac{y_{jt}}{\sqrt{T}}\right) \end{aligned}$$

$$\frac{1}{T^3\sigma^2}C_2 \underset{T\to\infty}{\overset{L}{\to}} \frac{1}{2N^2}\sum_{i,j=1}^{N}\int_0^1 (2r-1)\,W_i(r)\,W_j(r)\,dr \tag{18}$$

$$= \frac{1}{2}\int_0^1 (2r-1)\left(\frac{1}{N}\sum_{i=1}^{N}W_i(r)\right)^2 dr \tag{19}$$

then

$$\frac{\sqrt{N}}{T^3\sigma^2}C_2 \underset{N,T\to\infty}{\to} 0$$

Thus

$$\frac{C}{T^3\sigma^2} \underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{12}$$

$$\sqrt{N}\left(\frac{C}{T^3\sigma^2}-\frac{1}{12}\right) \underset{T,N\to\infty}{\overset{L}{\to}} \mathcal{N}\left(0,\frac{1}{60}\right)$$

Furthermore

$$\frac{12C^2}{T^6\sigma^4} \underset{T\to\infty}{\overset{L}{\to}} 3\left(\int_0^1 (2r-1)\left(\frac{1}{N}\sum_{i=1}^{N}[W_i(r)]^2-\left(\frac{1}{N}\sum_{i=1}^{N}W_i(r)\right)^2\right)dr\right)^2 = 3\left(\int_0^1 (2r-1)\,S_W^2 dr\right)^2$$

□

Proof of Theorem 1. By rewriting we have

$$\begin{aligned}
\psi_{R^2} &= \frac{12C^2}{T^2(T-1)(A-B)} = 12\left(\frac{C}{T^3\sigma^2}\right)^2 / \frac{A-B}{T^3\sigma^4} \\
&= 12\left(\frac{C}{T^3\sigma^2}-\frac{1}{12}+\frac{1}{12}\right)^2 / \frac{A-B}{T^3\sigma^4} \\
&= 12\left[\left(\frac{C}{T^3\sigma^2}-\frac{1}{12}\right)^2+\left(\frac{1}{12}\right)^2+\frac{1}{6}\left(\frac{C}{T^3\sigma^2}-\frac{1}{12}\right)\right] / \frac{A-B}{T^3\sigma^4} \\
&= 12\left[\left(\frac{C}{T^3\sigma^2}-\frac{1}{12}\right)^2+\frac{1}{6}\left(\frac{C}{T^3\sigma^2}-\frac{1}{12}\right)\right] / \frac{A-B}{T^3\sigma^4}+\frac{1}{12}/\frac{A-B}{T^3\sigma^4}
\end{aligned}$$

Then by the results from Lemma 1, we have

$$\sqrt{N}\left(\psi_{R^2}-\frac{1}{12}/\frac{A-B}{T^3\sigma^4}\right) \underset{T,N\to\infty}{\overset{L}{\to}} \mathcal{N}\left(0,\frac{48}{5}\right)$$

where

$$\frac{1}{12}/\frac{A-B}{T^3\sigma^4} \underset{T,N\to\infty}{\overset{P}{\to}} 1$$

Then the first part of this theorem is proved. Next, to prove the equivalence of sequential and joint limiting distribution. First, prove $A_1/T^3\sigma^4 \underset{(N,T)\to\infty}{\overset{P}{\to}} \frac{1}{3}$. Since

$$\begin{aligned} A_1/T^3\sigma^4 &= \frac{1}{N^2}\sum_{i,j=1}^{N}\left(T^{-1}\sum_{t=1}^{T}\left(\frac{y_{it}}{\sqrt{T}}\right)^2\left(\frac{y_{jt}}{\sqrt{T}}\right)^2\right) \\ &= \underbrace{\frac{1}{N^2}\sum_{i=1}^{N}\left(T^{-1}\sum_{t=1}^{T}\left(\frac{y_{it}}{\sqrt{T}}\right)^4\right)}_{A_{11}} + \underbrace{\frac{1}{N^2}\sum_{i\neq j}^{N}\left(T^{-1}\sum_{t=1}^{T}\left(\frac{y_{it}}{\sqrt{T}}\right)^2\left(\frac{y_{jt}}{\sqrt{T}}\right)^2\right)}_{A_{12}} \end{aligned}$$

where $A_{11} = O_p\left(N^{-1}\right)$. For A_{12}, as $T\to\infty$

$$T^{-1}\sum_{t=1}^{T}\left(\frac{y_{it}}{\sqrt{T}}\right)^2\left(\frac{y_{jt}}{\sqrt{T}}\right)^2 \underset{T\to\infty}{\overset{L}{\to}} \int_0^1 W_i\left(r\right)^2 W_j\left(r\right)^2 dr,$$

Since

$$E\left(T^{-1}\sum_{t=1}^{T}\left(\frac{y_{it}}{\sqrt{T}}\right)^2\left(\frac{y_{jt}}{\sqrt{T}}\right)^2\right) = T^{-3}\sum_{t=1}^{T}Ey_{it}^2Ey_{jt}^2 \to E\int_0^1 W_i\left(r\right)^2 W_j\left(r\right)^2 dr$$

then it follows from Theorem 5.4 in Billingsley (1968) that $T^{-1}\sum_{t=1}^{T}\left(\frac{y_{it}}{\sqrt{T}}\right)^2\left(\frac{y_{jt}}{\sqrt{T}}\right)^2$ is uniformly integrable in T. Thus, by Corollary 1 of Phillips and Moon (1999), $A_1/T^3\sigma^4$ jointly converges to 1/3, as $N, T\to\infty$. A similar result can be proved for A_2, A_3 and B. Thus we have $\left(A-B\right)/T^3\sigma^4 \underset{(N,T)\to\infty}{\overset{P}{\to}} \frac{1}{12}$. Second, prove $\sqrt{N}\left(C_1/T^3\sigma^2 - 1/12\right) \underset{(N,T)\to\infty}{\overset{L}{\to}} \mathcal{N}\left(0, 1/60\right)$. Recall

$$\begin{aligned} \frac{C_1}{T^3\sigma^2} - \frac{1}{12} &\doteq \frac{1}{2N}\sum_{i=1}^{N}T^{-2}\sum_{t=1}^{T}\left(2\frac{t}{T}-1\right)y_{it}^2 - \frac{1}{12} \\ &= \frac{1}{N}\sum_{i=1}^{N}\left(\frac{1}{2}T^{-2}\sum_{t=1}^{T}\left(2\frac{t}{T}-1\right)y_{it}^2 - \frac{1}{12}\right) \end{aligned}$$

Let $\xi_i = N^{-\frac{1}{2}}\left(\frac{1}{2}T^{-2}\sum_{t=1}^{T}\left(2\frac{t}{T}-1\right)y_{it}^2 - \frac{1}{12}\right)$, then by Theorem 2 of Phillips and Moon (1999), in order to show the joint convergence in law, we need to show that $\forall\epsilon > 0$,

$$\lim_{N,T\to\infty}\sum_{i=1}^{N}E\left(\xi_i^2 1_{\{|\xi_i|>\epsilon\}}\right) = 0$$

By Cauchy-Schwarz inequality

$$E\left(\xi_i^2 1_{\{|\xi_i|>\epsilon\}}\right) < \sqrt{E\left(\xi_i^4\right)E\left(1_{\{|\xi_i|>\epsilon\}}\right)}$$

where by Markov inequality we have

$$E\left(1_{\{|\xi_i|>\epsilon\}}\right) \leq \frac{E\left(\xi_i^2\right)}{\epsilon^2}$$

Then we have

$$\sum_{i=1}^{N}E\left(\xi_i^2 1_{\{|\xi_i|>\epsilon\}}\right) \leq \frac{1}{\epsilon}\sum_{i=1}^{N}\sqrt{E\left(\xi_i^4\right)E\left(\xi_i^2\right)} \leq \frac{1}{\epsilon}\sqrt{\sum_{i=1}^{N}E\left(\xi_i^4\right)}\sqrt{\sum_{i=1}^{N}E\left(\xi_i^2\right)}$$

Given Assumption 1, we have

$$E\left(\xi_i^4\right) = \frac{E\left(\frac{1}{2}T^{-2}\sum_{t=1}^{T}\left(2\frac{t}{T}-1\right)y_{it}^2 - \frac{1}{12}\right)^4}{N^2} = O\left(N^{-2}\right)$$

and

$$E\left(\xi_i^2\right) = \frac{E\left(\frac{1}{2}T^{-2}\sum_{t=1}^{T}\left(2\frac{t}{T}-1\right)y_{it}^2 - \frac{1}{12}\right)^2}{N} = O\left(N^{-1}\right)$$

And since $C_2/T^3\sigma^2 = O_p\left(N^{-1}\right)$, then by the continuous mapping theorem, the joint limiting distribution of ψ_{R^2} is constructed. □

Proof of Lemma 2. Firstly, we decompose this process into two parts:

$$P_n(t) = \underbrace{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}W_i(t)^2 - t\right)}_{G_n(t)} - \underbrace{\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}W_i(t)\right)^2}_{\xrightarrow{P} 0}$$

For all $t \in [0,1]$, by standard result we have

$$P_n(t) \xrightarrow{L} N\left(0, 2t^2\right)$$

Next, let $0 \le t_1 < t_2 \le 1$, we consider the random vector $\left(G_n(t_1), G_n(t_2)\right)'$. It can be represented as

$$\left(G_n(t_1), G_n(t_2)\right) = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{W}_i(\mathbf{t})^2 - \mathbf{t}\right)$$

where $\mathbf{W}_i(\mathbf{t})^2 = \left(W_i(t_1)^2, W_i(t_2)^2\right)'$ and $\mathbf{t} = (t_1, t_2)'$. Again, by CLT we have

$$\left(G_n(t_1), G_n(t_2)\right) \xrightarrow{L} N(\mathbf{0}, \mathbf{\Sigma})$$

where

$$\mathbf{\Sigma} = \begin{pmatrix} Var\left(W_i(t_1)^2\right) & Cov\left(W_i(t_1)^2, W_i(t_2)^2\right) \\ Cov\left(W_i(t_1)^2, W_i(t_2)^2\right) & Var\left(W_i(t_2)^2\right) \end{pmatrix}$$

where $Var\left(W_i(t_1)^2\right) = 2t_1^2$, $Var\left(W_i(t_2)^2\right) = 2t_2^2$ and $Cov\left(W_i(t_1)^2, W_i(t_2)^2\right) = 2t_1^2$.

More generally, for finite points $0 \le t_1 < t_2 < ... < t_m \le 1$, we have the finite dimensional distributions of $P_n(t)$ will converge to Gaussian process, i.e.

$$\left(G_n(t_1), ..., G_n(t_m)\right) \xrightarrow{L} N(\mathbf{0}, \mathbf{\Sigma}_m)$$

where

$$\mathbf{\Sigma}_m = \begin{pmatrix} 2t_1^2 & 2t_1^2 & \cdots & 2t_1^2 \\ & 2t_2^2 & \cdots & 2t_2^2 \\ & & \ddots & \vdots \\ & & & 2t_m^2 \end{pmatrix}$$

And since $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n}W_i(t)\right)^2 \xrightarrow{P} 0$, then by Slutsky Theorems we have the finite dimensional distribution of $P_n(t)$ will converge to Gaussian process $P(t)$. Next, we check the tightness of process $P_n(t)$. This can by done by checking $\exists\ K,\alpha,\beta>0$ s.t. $\forall\ n=0,1,...$ the following compactness inequality holds, for $0<\delta\leq 1$,

$$\sup_{0\leq t<t+\delta\leq 1} E\left|P_n(t+\delta)-P_n(t)\right|^{\alpha}\leq K\delta^{\beta+1}$$

Since by Invariance principle$\frac{1}{n}\sum_{i=1}^{n}W_i(t)$ will converge to wiener process $W(t)$, we only need to check the compactness for $G_n(t)$ part. Check the following moment

$$\begin{aligned}
& E\left[G_n(t+\delta)-G_n(t)\right]^4 \\
= & \frac{1}{n^2}E\left(\sum_{i=1}^{n}\left(\underbrace{W_i(t+\delta)^2-W_i(t)^2-\delta}_{X_i}\right)\right)^4=\frac{1}{n^2}\sum_{i,j,k,l=1}^{n}E\left(X_iX_jX_kX_l\right) \\
= & \frac{1}{n^2}\left(\sum_{i=1}^{n}EX_i^4+3\sum_{i\neq j}^{n}EX_i^2EX_j^2\right)=\frac{1}{n^2}\left(nEX_i^4+3\left(n^2-n\right)\left(EX_i^2\right)^2\right) \\
= & \frac{EX_i^4}{n}+\frac{3(n-1)}{n}\left(EX_i^2\right)^2\leq EX_i^4+3\left(EX_i^2\right)^2 \\
= & 210t^2\delta^2+108t\delta^3+72\delta^4
\end{aligned}$$

Where

$$\begin{aligned}
EX_i^2 &= Var\left[W_i(t+\delta)^2-W_i(t)^2\right] \\
&= Var\left(W_i(t+\delta)^2\right)+Var\left(W_i(t)^2\right)-2Cov\left(W_i(t+\delta)^2,W_i(t)^2\right) \\
&= 2(t+\delta)^2+2t^2-4t^2=4t\delta+2\delta^2
\end{aligned}$$

and

$$\begin{aligned}
& EX_i^4 \\
= & E\left[\left(W_i(t)+\Delta W_i\right)^2-W_i(t)^2-\delta\right]^4 \\
= & E\left[\Delta W_i^2+2W_i(t)\Delta W_i-\delta\right]^4 \\
= & E\Delta W_i^8+16EW_i(t)^4\Delta W_i^4+\delta^4+12E\Delta W_i^6W_i(t)^2+6\delta^2E\Delta W_i^4+24\delta^2EW_i(t)^2\Delta W_i^2 \\
& +6\delta^2EW_i(t)^4-4\delta E\Delta W_i^6-4\delta^3E\Delta W_i^2-48\delta E\Delta W_i^4W_i(t)^2 \\
= & 105\delta^4+144t^2\delta^2+\delta^4+180t\delta^3+18\delta^4+24t\delta^3+18\delta^2t^2-60\delta^4-4\delta^4-144t\delta^3 \\
= & 162t^2\delta^2+60t\delta^3+60\delta^4
\end{aligned}$$

Then

$$\begin{aligned}
\sup_{0\leq t<t+\delta\leq 1}E\left[G_n(t+\delta)-G(t)\right]^4 &\leq \sup_{0\leq t<t+\delta\leq 1}\left(EX_i^4+3\left(EX_i^2\right)^2\right) \\
&= \sup_{0\leq t<t+\delta\leq 1}\left(210t^2\delta^2+108t\delta^3+72\delta^4\right) \\
&= 210\delta^2(1-\delta)^2+108\delta^3(1-\delta)+72\delta^4 \\
&= 174\delta^4+210\delta^2-312\delta^3\leq 210\delta^2
\end{aligned}$$

By taking $K=210$ and $\beta=1$, we have

$$\sup_{0\leq t<t+\delta\leq 1}E\left[G_n(t+\delta)-G(t)\right]^4<K\delta^{\beta+1}$$

□

Proof of Lemma 3. By (i) of Lemma 1, we have

$$\frac{A}{\sigma^4 T^3} \underset{T\to\infty}{\overset{L}{\to}} \int_0^1 \left(S_W^2\right)^2 dr = \int_0^1 \left(S_W^2 - r + r\right)^2 dr = \int_0^1 \left(S_W^2 - r\right)^2 dr + 2\int_0^1 r\left(S_W^2 - r\right) dr + \frac{1}{3}$$

Then by Lemma 2, we have

$$\sqrt{N}\left(\frac{A}{\sigma^4 T^3} - \frac{1}{3}\right) \underset{T\to\infty}{\overset{L}{\to}} \underbrace{\int_0^1 \sqrt{N}\left(S_W^2 - r\right)^2 dr}_{\underset{N\to\infty}{\overset{P}{\to}} 0} + 2\underbrace{\int_0^1 r\sqrt{N}\left(S_W^2 - r\right) dr}_{\underset{N\to\infty}{\overset{P}{\to}} \int_0^1 rP(r)dr}$$

Then we have

$$\sqrt{N}\left(\frac{A}{\sigma^4 T^3} - \frac{1}{3}\right) \underset{T,N\to\infty}{\overset{L}{\to}} \int_0^1 2rP\left(r\right) dr$$

Proof of Theorem 2. By rewriting we have

$$\psi^2 + 1 = \frac{A}{B}$$

and

$$\sqrt{N}\left(\psi^2 + 1 - \frac{1/3}{B/\sigma^4 T^3}\right) = \frac{\sqrt{N}\left(A/\sigma^4 T^3 - 1/3\right)}{B/\sigma^4 T^3}$$

Then by (v) of Lemma 1 and Lemma 3 we have

$$\sqrt{N}\left(\psi^2 + 1 - \frac{1/3}{B/\sigma^4 T^3}\right) = \frac{\underbrace{\sqrt{N}\left(A/\sigma^4 T^3 - 1/3\right)}_{\underset{T,N\to\infty}{\overset{L}{\to}} \int_0^1 2rP(r)dr}}{\underbrace{B/\sigma^4 T^3}_{\underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{4}}}$$

i.e.

$$\sqrt{N}\left(\psi^2 + 1 - \frac{1/3}{B/\sigma^4 T^3}\right) \underset{T,N\to\infty}{\overset{L}{\to}} 4\int_0^1 2rP\left(r\right) dr$$

where $\frac{1/3}{B/\sigma^4 T^3} \underset{T,N\to\infty}{\overset{P}{\to}} \frac{4}{3}$. Then by the statement Remark 5, Theorem 2 is proved □

Proof of Theorem 3. By the results in Lemma 3, we have

$$\frac{A}{\sigma^4 T^3} \underset{T\to\infty}{\overset{L}{\to}} \int_0^1 \left(S_W^2\right)^2 dr = \int_0^1 \left(S_W^2 - r + r\right)^2 dr = \int_0^1 \left(S_W^2 - r\right)^2 dr + 2\int_0^1 r\left(S_W^2 - r\right) dr + \frac{1}{3} \tag{20}$$

By Equation (15) and (16)

$$\frac{B}{T^3\sigma^4} \underset{T\to\infty}{\overset{L}{\to}} \left(\int_0^1 \frac{1}{N}\sum_{i=1}^N [W_i(r)]^2 - \left(\frac{1}{N}\sum_{i=1}^N W_i(r)\right)^2 dr\right)^2 \tag{21}$$

$$= \left(\int_0^1 S_W^2 dr\right)^2 = \left(\int_0^1 \left(S_W^2 - r + r\right) dr\right)^2 = \left(\int_0^1 \left(S_W^2 - r\right) dr + \frac{1}{2}\right)^2$$

$$= \left(\int_0^1 \left(S_W^2 - r\right) dr\right)^2 + \int_0^1 S_W^2 dr - \frac{1}{4}$$

And by Equation (17) and (18)

$$\frac{12C^2}{T^6\sigma^4} \underset{T\to\infty}{\overset{L}{\to}} 3\left(\int_0^1 (2r-1)\left(\frac{1}{N}\sum_{i=1}^N [W_i(r)]^2 - \left(\frac{1}{N}\sum_{i=1}^N W_i(r)\right)^2\right) dr\right)^2 \tag{22}$$

$$= 3\left(\int_0^1 (2r-1) S_W^2 dr\right)^2 = 3\left(\int_0^1 (2r-1)\left(S_W^2 - r + r\right) dr\right)^2$$

$$= 3\left(\int_0^1 (2r-1)\left(S_W^2 - r\right) dr + \int_0^1 (2r-1) r dr\right)^2$$

$$= 3\left(\int_0^1 (2r-1)\left(S_W^2 - r\right) dr\right)^2 + \int_0^1 (2r-1)\left(S_W^2 - r\right) dr + \frac{1}{12}$$

$$= 3\left(\int_0^1 (2r-1)\left(S_W^2 - r\right) dr\right)^2 + \int_0^1 (2r-1) S_W^2 dr - \frac{1}{12}$$

Then

$$\frac{\psi_F}{NT} = \frac{12C^2}{N\left(T^3(A-B) - 12C^2\right)} = \frac{12C^2/T^6\sigma^4}{N\left(\frac{A-B}{T^3\sigma^4} - \frac{12C^2}{T^6\sigma^4}\right)}$$

where by the result in Lemma 1, we have $12C^2/T^6\sigma^4 \underset{T,N\to\infty}{\overset{P}{\to}} \frac{1}{12}$. And combine from the Equation (20) to (22), we have

$$N\left(\frac{A-B}{T^3\sigma^4} - \frac{12C^2}{T^6\sigma^4}\right) \underset{T\to\infty}{\overset{L}{\to}}$$

$$\int_0^1 N\left(S_W^2 - r\right)^2 dr - \left(\int_0^1 \sqrt{N}\left(S_W^2 - r\right) dr\right)^2 - 3\left(\int_0^1 (2r-1)\sqrt{N}\left(S_W^2 - r\right) dr\right)^2$$

And by Lemma 2 we have

$$N\left(\frac{A-B}{T^3\sigma^4} - \frac{12C^2}{T^6\sigma^4}\right) \underset{T,N\to\infty}{\overset{L}{\to}}$$

$$\int_0^1 P(r)^2 dr - \left(\int_0^1 P(r) dr\right)^2 - 3\left(\int_0^1 (2r-1) P(r) dr\right)^2$$

Then Theorem 3 is proved □

Claim 1. *Given the assumptions in Theorem 1, we do not have*

$$\sqrt{N}\left(\psi_{R^2} - 1\right) \underset{T,N\to\infty}{\overset{L}{\to}} \mathcal{N}\left(0, \frac{48}{5}\right)$$

Proof. We have

$$\sqrt{N}\left(\psi_{R^2}-1\right)=\underbrace{\sqrt{N}\left(\psi_{R^2}-\frac{1}{12}/\frac{A-B}{T^3\sigma^4}\right)}_{\underset{T,N\to\infty}{\overset{L}{\to}}\mathcal{N}\left(0,\frac{48}{5}\right)}-\underbrace{\sqrt{N}\left(1-\frac{1}{12}/\frac{A-B}{T^3\sigma^4}\right)}_{\Phi_1}$$

Now we check Φ_1 part

$$\Phi_1=\sqrt{N}\left(1-\frac{1}{12}/\frac{A-B}{T^3\sigma^4}\right)=\frac{\sqrt{N}\left(\frac{A-B}{T^3\sigma^4}-\frac{1}{12}\right)}{\underbrace{\frac{A-B}{T^3\sigma^4}}_{\underset{T,N\to\infty}{\overset{P}{\to}}\frac{1}{12}}}$$

And by equation (20) and (21) we have

$$\begin{aligned}
&\sqrt{N}\left(\frac{A-B}{T^3\sigma^4}-\frac{1}{12}\right)\\
&\underset{T\to\infty}{\overset{L}{\to}}\sqrt{N}\left(\underbrace{\int_0^1\left(S_W^2-r\right)^2dr-\left(\int_0^1\left(S_W^2-r\right)dr\right)^2}_{\Phi_2}+\int_0^1(2r-1)\,S_W^2dr-\frac{1}{6}\right)\\
=\;&\sqrt{N}\left(\Phi_2+\int_0^1(2r-1)\left(S_W^2-r\right)dr+\int_0^1\left(2r^2-r\right)dr-\frac{1}{6}\right)\\
=\;&\sqrt{N}\left(\Phi_2+\int_0^1(2r-1)\left(S_W^2-r\right)dr\right)
\end{aligned}$$

And by Lemma 2 we have, $\sqrt{N}\Phi_2\underset{T,N\to\infty}{\overset{P}{\to}}0$ and

$$\int_0^1(2r-1)\sqrt{N}\left(S_W^2-r\right)dr\underset{T,N\to\infty}{\overset{L}{\to}}\int_0^1(2r-1)\,P(r)\,dr. \tag{23}$$

Then the statement is proved. □

TABLE 4. Empirical critical value of ψ

T	N	τ	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
25	25	0	0.390	0.416	0.443	0.471	0.691	0.723	0.753	0.787
		0.5	0.391	0.420	0.443	0.471	0.693	0.727	0.756	0.793
		0.8	0.397	0.420	0.443	0.471	0.692	0.724	0.750	0.785
25	50	0	0.434	0.455	0.474	0.494	0.653	0.674	0.695	0.718
		0.5	0.432	0.454	0.472	0.492	0.653	0.676	0.695	0.718
		0.8	0.435	0.452	0.471	0.493	0.652	0.676	0.695	0.719
250	100	0	0.483	0.497	0.510	0.525	0.634	0.649	0.662	0.676
		0.5	0.481	0.497	0.510	0.525	0.632	0.647	0.660	0.675
		0.8	0.483	0.498	0.511	0.526	0.635	0.651	0.664	0.681
50	250	0	0.511	0.520	0.529	0.539	0.608	0.617	0.627	0.637
		0.5	0.508	0.518	0.527	0.538	0.608	0.618	0.627	0.636
		0.8	0.509	0.519	0.527	0.537	0.608	0.618	0.627	0.637

TABLE 5. Empirical critical value of ψ_{R^2}

T	N	τ	0.01	0.025	0.05	0.1	0.9	0.95	0.975	0.99
25	25	0	0.655	0.739	0.803	0.853	0.976	0.981	0.984	0.987
		0.5	0.659	0.745	0.806	0.857	0.977	0.981	0.985	0.987
		0.8	0.662	0.746	0.802	0.857	0.976	0.981	0.984	0.987
25	50	0	0.854	0.888	0.913	0.934	0.988	0.990	0.992	0.993
		0.5	0.851	0.888	0.912	0.933	0.988	0.990	0.992	0.993
		0.8	0.855	0.889	0.914	0.934	0.988	0.990	0.991	0.993
250	100	0	0.938	0.951	0.960	0.969	0.994	0.995	0.996	0.996
		0.5	0.938	0.951	0.960	0.969	0.994	0.995	0.995	0.996
		0.8	0.937	0.951	0.961	0.969	0.994	0.995	0.996	0.996
50	250	0	0.977	0.981	0.985	0.988	0.997	0.998	0.998	0.999
		0.5	0.977	0.982	0.985	0.988	0.997	0.998	0.998	0.998
		0.8	0.977	0.982	0.985	0.988	0.997	0.998	0.998	0.999

REFERENCES

Bai, J. and S. Ng (2004). A Panic Attack on Unit Roots and Cointegration, *Econometrica* **72**: 1127–177.

Billingsley, P. (1968). Convergence of probability measures, *New York: Wiley*

Bond S., Nauges C. and Windmeijer F. (2002). Unit Roots and Identification in Autoregressive Panel Data Models: A Comparison of Alternative Tests, 10th International Conference on Panel Data, Berlin, July 5-6, 2002.

Dickey, D. A. and W. A. Fuller. (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association* **74**: 427–31.

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**: 201–26.

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

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.

Lo, A. W. and A. C. MacKinlay (1988). Stock prices do not follow random walks: Evidence from a simple specification test, *Review of financial studies* **1**: 41–66.

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

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

Pedroni, P., (1999). Critical values for cointegration tests in heterogeneous panels with multiple regressors, *Oxford Bulletin of Economics and statistics,* **61**:653–670.

Pesaran, M.H. (2005). A Simple Panel Unit Root Test in the Presence of Cross Section Dependence, *Cambridge Working Papers in Economics 0346, revised Version, University of Cambridge.*

Phillips, P.C.B. and Pierre Perron. (1988). Testing for a unit root in time series regression, *Biometrika* **75**: 335–46.

Phillips, P.C.B. and H.R. Moon. (1999). Linear Regression Limit Theory for Nonstationary Panel Data, *Econometrica* **67**: 1057–111.