Global COE Hi-Stat Discussion Paper Series 016

A Simple Panel Stationarity Test in the Presence of Cross-Sectional Dependence

Kaddour Hadri
Eiji Kurozumi

December 2008
(Revised: June 2010)

Hi-Stat
Institute of Economic Research
Hitotsubashi University
2-1 Naka, Kunitatchi Tokyo, 186-8601 Japan
http://gcoe.ier.hit-u.ac.jp

A Simple Panel Stationarity Test in the Presence of Cross-Sectional Dependence[1]

Kaddour Hadri[2]
Queen's University Management School
Queen's University Belfast

Eiji Kurozumi[3]
Department of Economics
Hitotsubashi University

Revised December 2009

Abstract

This paper develops a simple test for the null hypothesis of stationarity in heterogeneous panel data with cross-sectional dependence in the form of a common factor in the disturbance. We do not estimate the common factor but mop-up its effect by employing the same method as the one proposed in Pesaran (2007) in the unit root testing context. Our test is basically the same as the KPSS test but the regression is augmented by cross-sectional average of the observations. The limiting distribution under the null is shown to be a standard normal. The latter result is derived using the joint asymptotic limits where T and $N \to \infty$ simultaneously (under the additional condition that $N/T \to 0$). We also extend our test to the more realistic case where the shocks are serially correlated. We use Monte Carlo simulations to examine the finite sample property of the panel augmented KPSS test.

JEL classification: C12, C33

Key words: Panel data; stationarity; KPSS test; cross-sectional dependence; joint asymptotic.

[1]We would like to thank the editor, Pierre Perron, and three anonymous referees for their comments, suggestions and constructive criticisms which helped to greatly improve the paper. Evidently, any remaining errors are ours.

[2]Corresponding author: Kaddour Hadri, 25 University Square, Queen's University Management School, Queen's University, Belfast, BT7 1NN, UK. Tel +44 (0)28 9097 3286. Fax +44 (0)28 90975156. Email: k.hadri@qub.ac.uk

[3]Kurozumi's research was partially supported by the Ministry of Education, Culture, Sports, Science and Technology under Grants-in-Aid No.18730142.

1. Introduction

Since the beginning of the 90's, the theoretical and empirical econometrics literature witnessed a formidable output on testing unit root and stationarity in panel data with large T (time dimension) and N (cross-section dimension). The main motive for applying unit root and stationarity tests to panel data is to improve the power of the tests relative to their univariate counterparts. This was supported by the ensuing applications and simulations. The early theoretical contributions are by Breitung and Meyer (1994), Choi (2001), Hadri (2000), Hadri and Larsson (2005), Im, Pesaran and Shin (2003), Levin, Lin and Chu (2002), Maddala and Wu (1999), Phillips and Moon (1999), Quah (1994) and Shin and Snell (2006). On the application side, the early contributions were the work of O'Connell (1998), Oh (1996), Papell (1997, 2002), Wu (1996) and Wu and Wu (2001), who focused on testing the existence of purchasing power parity. Culver and Papell (1997) applied panel unit root tests to the inflation rate for a subset of OECD countries. They have also been employed in testing output convergence and more recently in the analysis of business cycle synchronization, house price convergence, regional migration and household income dynamics (cf. Breitung and Pesaran (2008)). All these "first generation" panel tests are based on the incredible assumption that the cross-sectional units are independent or at least not cross-sectionally correlated. Banerjee (1999), Baltagi and Kao (2000), Baltagi (2001) provide comprehensive surveys on the first generation panel tests. However, in most empirical applications this assumption is erroneous. O'Connell (1998) was the first to show via simulation that the panel tests are considerably distorted when the independence assumption is violated, whether the null hypothesis is a unit root or stationarity. Banerjee, Marcellino and Osbat (2001, 2004) argued against the use of panel unit root tests due to this problem. Therefore, it became imperative that in applications using panel tests to account for the possibility of cross-sectional dependence. This led, recently, to a flurry of papers accounting for cross-sectional dependence of different forms or second generation panel unit root tests. The most noticeable proposals in this area are by Chang (2004), Phillips and Sul (2003), Bai and Ng (2004), Moon and Perron (2004), Choi and Chue (2007) and Pesaran (2007) for unit root panel tests. For panel stationarity tests, the only contributions so far are by Bai and Ng (2005) and Harris, Leybourne and McCabe (2005), both of which corrected for cross-sectional dependence by using the principal component analysis proposed by Bai and Ng (2004).

Choi and Chue (2007) utilize subsampling technique to tackle cross-sectional dependence. Phillips and Sul (2003), Bai and Ng (2004), Moon and Perron (2004) and Pesaran (2007) employ factor models to allow for cross-sectional correlation (cf. de Silva, Hadri and Tremayne (2009) for the comparison of the three last tests). Pesaran (2007) considers only one factor and instead of estimating it, he augments the ADF regressions with the cross-sectional averages of lagged levels and first-differences of the individual series to account for the cross-sectional dependence generated by this one factor. Other contributions are by Maddala and Wu (1999) and Chang (2004) who exploited the flexibility of the bootstrap method to deal with the pervasive problem of cross-sectional dependence of general form. Breitung and Pesaran (2008) give an excellent survey of the first and second generation panel tests.

The transfer of testing for unit root and stationarity from univariate time series to large panel data contributed to a significant increase of the power of those tests. However, this transfer led to a number of difficulties besides the problem of cross-sectional dependence. In particular, the asymptotic theory is by far more intricate due to the presence of two indices: the time dimension and the number of cross-sections. The limit theory for this class of panel data has been developed in a seminal paper by Phillips and Moon (1999). In their paper they study *inter alia* the limit theory that allows for both sequential limits, wherein $T \to \infty$ followed by $N \to \infty$, and joint limits where $T, N \to \infty$ simultaneously. They also mention, in the same paper, the diagonal path limit theory in which the passage to infinity is done along a specific diagonal path. The drawback of sequential limits is that in certain cases, they can give asymptotic results which are misleading. The downside of diagonal path limit theory is that the assumed expansion path $(T(N), N) \to \infty$ may not provide an appropriate approximation for a given (T, N) situation. Finally, the joint limit theory requires, generally, a moment condition as well as a rate condition on the relative speed of T and N going to infinity.

The main contribution of this paper[4] is the derivation under the null and under the local alternative of the limiting distribution of the Kwiatkowski et al. (1992) test (KPSS test hereafter) corrected for cross-sectional dependency generated by a one factor error structure. For each unit, our test is basically the KPSS test with the regression augmented by the cross-sectional average of the observations. In the panel data context, this amounts to adapting Pesaran (2007) approach accounting for cross-sectional dependence to the panel stationarity test of Hadri (2000). The choice of the Pesaran (2007) approach is due essentially to its conceptual simplicity. We show that the limiting null distribution of our panel augmented KPSS test is the same as the panel stationary test proposed by Hadri (2000), which is an Lagrange multiplier (LM) test without cross-sectional dependence. Our theoretical result are obtained via the joint asymptotic theory where T , $N \rightarrow \infty$ concurrently (under the added condition that $N/T \rightarrow 0$) and the limiting distributions under the null are shown to be standard normal .We then extend our panel augmented KPSS test to the more realistic and useful case of the serially correlated shocks. The test is shown to have a standard normal distribution as a limiting null distribution employing the sequential asymptotic theory where $T \rightarrow \infty$ followed by $N \rightarrow \infty$. We conjecture that the sequential limit is the same as the joint limit under the additional condition that $N/T \rightarrow 0$. The test is very easy to implement. We use Monte Carlo simulations to examine the finite sample properties of the panel augmented KPSS test allowing for serial correlation.

The paper is organized as follows. Section 2 sets up the model and assumptions, and define the augmented panel test statistic. In section 3, we show that the limiting null distribution of the panel augmented KPSS test is the same as that of Hadri's (2000) test. We also examine whether our theoretical results are valid in finite samples via simple Monte Carlo simulations. In Section 4, we relax Assumption 1 in order to allow for serial correlation

[4] In the original version of the paper, we derived a Lagrange multiplier (LM) test, which is a locally best invariant test under the assumption of normality, allowing for cross-sectional dependence. We also compared a panel augmented KPSS test with the extended LM test under the null of stationarity, under the local alternative and under the fixed alternative and discussed the differences between the two tests. The latter results were derived using the joint asymptotic limits where T and $N \rightarrow \infty$ jointly.

in the error term and propose a modification of the panel augmented KPSS test statistic to correct for the presence of this serial correlation. Once again, we examine the finite sample properties of the proposed test statistic via Monte Carlo simulations. Section 5 concludes the paper. All the proofs are presented in the Appendix.

A summary word on notation. We define $M_A = I_T - A(A'A)^{-1}A'$ for a full column rank matrix A. The symbols $\overset{p\,(N,T)}{\longrightarrow}$ and $\overset{(N,T)}{\Longrightarrow}$ mean joint convergence in probability and joint weak convergence respectively when both T and N go to infinity simultaneously, while $\overset{T}{\Longrightarrow}$ or $\overset{N}{\Longrightarrow}$ signifies weak convergence when only T or N goes to infinity.

2. Model and Test Statistics

2.1. Model and assumptions

Let us consider the following model:

$$y_{it} = z_t'\delta_i + r_{it} + u_{it}, \quad r_{it} = r_{it-1} + v_{it}, \quad u_{it} = f_t\gamma_i + \varepsilon_{it} \tag{1}$$

for $i = 1, \cdots, N$ and $t = 1, \cdots, T$ where z_t is deterministic and $r_{i0} = 0$ for all i. The commonly used specification of z_t in the literature is either $z_t = z_t^{\mu} = 1$ or $z_t = z_t^{\tau} = [1, t]'$. In this paper, we consider these two cases. Accordingly, we define $\delta_i = \alpha_i$ when $z = 1$ and $\delta_i = [\alpha_i, \beta_i]'$ when $z = [1, t]'$. In model (1), $z_t'\delta_i$ is the individual effect while f_t is one dimensional unobserved common factor, γ_i is the loading factor and ε_{it} is the individual-specific (idiosyncratic) error.

By stacking y_{it} with respect to t, model (1) can be expressed as

$$\begin{bmatrix} y_{i1} \\ y_{i2} \\ \vdots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} z_1' \\ z_2' \\ \vdots \\ z_T' \end{bmatrix} \delta_i + \begin{bmatrix} r_{i1} \\ r_{i2} \\ \vdots \\ r_{iT} \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_T \end{bmatrix} \gamma_i + \begin{bmatrix} \varepsilon_{i1} \\ \varepsilon_{i2} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix},$$

$$\begin{bmatrix} r_{i1} \\ r_{i2} \\ \vdots \\ r_{iT} \end{bmatrix} = \begin{bmatrix} 1 & & & 0 \\ 1 & 1 & & \\ \vdots & \vdots & \ddots & \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} v_{i1} \\ v_{i2} \\ \vdots \\ v_{iT} \end{bmatrix}$$

or

$$\begin{aligned} \mathbf{y}_i &= Z\delta_i + \mathbf{r}_i + \mathbf{f}\gamma_i + \boldsymbol{\varepsilon}_i \\ &= Z\delta_i + B\mathbf{v}_i + \mathbf{f}\gamma_i + \boldsymbol{\varepsilon}_i, \end{aligned} \tag{2}$$

where $Z = [\boldsymbol{\tau}, \mathbf{d}]$ with $\boldsymbol{\tau} = [1, 1, \cdots, 1]'$, $\mathbf{d} = [1, 2, \cdots, T]'$ being $T \times 1$ vectors and B is a $T \times T$ matrix with ones on the main diagonal and everywhere below it. Tanaka (1996) calls matrix B, the *random walk generating matrix*. Further, we have

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} Z & & & \\ & Z & & \\ & & \ddots & \\ & & & Z \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_N \end{bmatrix} + \begin{bmatrix} B & & & \\ & B & & \\ & & \ddots & \\ & & & B \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_N \end{bmatrix} + \begin{bmatrix} \mathbf{f}\gamma_1 \\ \mathbf{f}\gamma_2 \\ \vdots \\ \mathbf{f}\gamma_N \end{bmatrix} + \begin{bmatrix} \boldsymbol{\varepsilon}_1 \\ \boldsymbol{\varepsilon}_2 \\ \vdots \\ \boldsymbol{\varepsilon}_N \end{bmatrix}$$

or

$$\begin{aligned} \mathbf{y} &= (I_N \otimes Z)\delta + \mathbf{r} + (\gamma \otimes \mathbf{f}) + \boldsymbol{\varepsilon} \\ &= (I_N \otimes Z)\delta + (I_N \otimes B)\mathbf{v} + (\gamma \otimes \mathbf{f}) + \boldsymbol{\varepsilon}. \end{aligned} \tag{3}$$

In this section, we assume the following simple assumption:

Assumption 1 *(i) The stochastic processes $\{\varepsilon_{it}\}$, $\{f_t\}$ and $\{v_{it}\}$ are independent,*

$$\varepsilon_{it} \sim i.i.d.N(0, \sigma_\varepsilon^2), \quad f_t \sim i.i.d.N(0, \sigma_f^2), \quad v_{it} \sim i.i.d.N(0, \sigma_v^2),$$

for simplicity, σ_ε^2 is assumed to be known.

(ii) There exist real numbers M_1, $\underline{M}$ and $\overline{M}$ such that $|\gamma_i| < M_1 < \infty$ for all i and $0 < \underline{M} < |\bar{\gamma}| < \overline{M} < \infty$ for all N, where $\bar{\gamma} = N^{-1} \sum_{i=1}^{N} \gamma_i$.

Assumption 1(i) is too restrictive and not practical. It will be relaxed in Section 4 to a more realistic one. The assumption of normality with homoskedasticity is adopted here to simplify the derivation of the theoretical results. The cross-sectional independence of ε_{it} across i is standard in one-factor models. The independence over t will be relaxed in Section 4 to allow for serial correlation. In section 4 the common factor f_t will be allowed to follow a general stationary process. The independence of v_{it} across i and t is standard in KPSS models. Instead of considering σ_ε^2 known, we can also estimate it employing the estimator $\widehat{\sigma}_\varepsilon^2 = \sum_{i=1}^{N}\sum_{t=1}^{T}\widehat{\varepsilon}_{it}^2/NT$, where $\widehat{\varepsilon}_{it}$ is the residual from the augmented regression. It can be shown that this estimator jointly converges in probability to σ_ε^2 under both the null hypothesis and the local alternative. Then, we could replace σ_ε^2 in the definition of ST_i by its estimator $\widehat{\sigma}_\varepsilon^2$. Assumption 1(ii) is concerned with the weights of the common factor f_t. This assumption implies that each individual is possibly affected by the common factor with the finite weight γ_i and that the absolute value of the average of γ_i is bounded away from 0 and above both in finite samples and in asymptotics. The latter property is important in order to eliminate the common factor effect from the regression. A similar assumption is also entertained in Pesaran (2007). The main purpose of this section is to examine the theoretical effect of "augmentation", which is explained below, on stationarity tests.

We consider a test for the null hypothesis of (trend) stationarity against the alternative of a unit root for model (1). Since all the innovations are homoskedastic, the testing problem is given by

$$H_0: \ \rho \equiv \frac{\sigma_v^2}{\sigma_\varepsilon^2} = 0 \quad \text{v.s.} \quad H_1 : \rho > \ 0 \tag{4}$$

where $\rho = \sigma_v^2/\sigma_\varepsilon^2$ is a signal-to-noise ratio. Under H_0, r_{it} becomes equal to zero for all i so that y_{it} is stationary whereas all of the cross-sectional units have a unit root under the alternative.

Remark 1 *Allowing for only one factor could be considered as too restrictive. Bai and Ng (2005) examined a more general case than here in which they allowed for more than one factor. However, they were disappointed by their results. In the case of more than one factor, it is not sufficient to augment the regression by $\overline{y}_t$ in order to eliminate the effect of the cross-sectional dependence from the test statistic. In such a case, the limiting distribution of our test statistic would depend on nuisance parameters. To correct for multifactor error structure, we may be able to adapt the method proposed by Pesaran, Smith and Yamagata (2009) in the unit root context to our case. Pesaran et al. (2009) method can be considered as a generalisation of Pesaran (2007) employed here. But, the application of the new proposal* *to our test is not straightforward and is therefore left for future research. The downside of allowing for multifactor error structure is that the number of factors has to be estimated employing information criteria. Simulations carried out by de Silva et al. (2009) to compare panel unit root tests allowing for multifactor error structure, show that the number of factors estimated is often greater than the true number of factors in small samples and this may be responsible for the upward size distorsions observed in their simulations.*

2.2. A simple stationarity test

A panel stationarity test has already been proposed by Hadri (2000) and Shin and Snell (2006) for cross-sectionally independent data. Here, we extend Hadri's test to the cross-sectionally dependent case. Hadri (2000) showed that if there is no cross-sectional dependence in a model, we can construct the LM test using the regression residuals of y_{it} on z_t in the same way as KPSS (1992) and that the limiting distribution of the standardized LM test statistic is standard normal under the null hypothesis. However, it can be shown that Hadri's (2000) test depends on nuisance parameters even asymptotically if there exits cross-sectional dependence; we then need to develop a stationarity test that takes into account cross-sectional dependence.

In order to eliminate the effect of the common factor from the test statistic, we make use of the simple method proposed by Pesaran (2007), which develops panel unit root tests with cross-sectional dependence. As in Pesaran (2007), we first take a cross-sectional average of the model:

$$\bar{y}_t = z_t'\bar{\delta} + \bar{r}_t + f_t\bar{\gamma} + \bar{\varepsilon}_t, \tag{5}$$

where $\bar{y}_t = N^{-1}\sum_{i=1}^{N} y_{it}$, $\bar{\delta} = N^{-1}\sum_{i=1}^{N} \delta_i$, $\bar{r}_t = N^{-1}\sum_{i=1}^{N} r_{it}$, $\bar{\gamma} = N^{-1}\sum_{i=1}^{N} \gamma_i$ and $\bar{\varepsilon}_t = N^{-1}\sum_{i=1}^{N} \varepsilon_{it}$. Since $\bar{\gamma} \neq 0$ by assumption, we can solve equation (5) with respect to f_t as follows:

$$f_t = \frac{1}{\bar{\gamma}}\left(\bar{y}_t - z_t'\bar{\delta} - \bar{r}_t - \bar{\varepsilon}_t\right).$$

By inserting this solution of f_t into model (1) we obtain the following augmented regression model:

$$y_{it} = z_t' \tilde{\delta}_i + \tilde{\gamma}_i \bar{y}_t + \epsilon_{it}, \tag{6}$$

where $\tilde{\delta}_i = \delta_i - \tilde{\gamma}_i \bar{\delta}$, $\tilde{\gamma}_i = \gamma_i / \bar{\gamma}$ and $\epsilon_{it} = r_{it} - \tilde{\gamma}_i \bar{r}_t + \varepsilon_{it} - \tilde{\gamma}_i \bar{\varepsilon}_t$. Based on (6) we propose to regress y_{it} on z_t and $\bar{y}_t$ for each i and construct the test statistic in the same way as Hadri (2000). That is,

$$Z_A = \frac{\sqrt{N}(\overline{ST} - \xi)}{\zeta} \tag{7}$$

$$\text{where} \quad \overline{ST} = \frac{1}{N} \sum_{i=1}^{N} ST_i \quad \text{with} \quad ST_i = \frac{1}{\sigma_\varepsilon^2 T^2} \mathbf{y}_i' M_w B' B M_w \mathbf{y}_i$$

$$\text{and} \quad \begin{cases} \xi = \xi_\mu = \frac{1}{6}, \quad \zeta^2 = \zeta_\mu^2 = \frac{1}{45} & \text{when } z_t = z_t^\mu = 1, \\ \xi = \xi_\tau = \frac{1}{15}, \quad \zeta^2 = \zeta_\tau^2 = \frac{11}{6300} & \text{when } z_t = z_t^\tau = [1, t]'. \end{cases}$$

Note that ST_i can also be expressed as

$$ST_i = \frac{1}{\sigma_\varepsilon^2 T^2} \sum_{t=1}^{T} (S_{it}^w)^2 \quad \text{where} \quad S_{it}^w = \sum_{s=1}^{t} \hat{\epsilon}_{is}$$

with $\hat{\epsilon}_{it}$ obtained for each i by regressing y_{it} on $w_t = [z_t', \bar{y}_t]'$ for $t = 1, \cdots, T$.

From (7) we can see that $\overline{ST}$ is the average of the KPSS test statistic across i and Z_A is normalized so that we obtain a limiting distribution. We call Z_A the panel augmented KPSS test statistic.

3. Theoretical Property of the panel Augmented KPSS test

In this section we derive the joint limit distribution, where $T, N \to \infty$ simultaneously, of the panel augmented KPSS test under the null hypothesis and under the local alternative.

Theorem 1 *Assume that Assumption 1 holds. Under H_0, as N and T go to infinity simultaneously with $N/T \to 0$, the panel augmented KPSS test statistic has a limiting standard normal distribution for both cases of $z_t = 1$ and $z_t = [1, t]'$,*

$$Z_A \overset{(N,T)}{\Longrightarrow} N(0, 1).$$

Proof. see Appendix.

This shows that the sequential asymptotic is equivalent to the joint asymptotic only under the condition that $N/T \to 0$ as T and N go to infinity jointly. The use of the sequential asymptotic on its own will not uncover this condition. Note that the rejection region of Z_A is the right hand tail as in Hadri's (2000) test. The condition that $N/T \to 0$ as T and $N \to \infty$ jointly, means that the test is suitable for panels with T larger than N.

Theorem 1 shows that Pesaran's (2007) method works well in order to eliminate cross-sectional dependence for testing the null hypothesis of stationarity.

We next investigate the asymptotic property of the test statistics under the local alternative, which is expressed as

$$H_1^\ell : \ \rho = \frac{c^2}{\sqrt{N}T^2} \quad \text{where } c \text{ is some constant.}$$

Note that for a single time series analysis, the local alternative is given by $\rho = c^2/T^2$. Since the sum of ST_i is normalized by $\sqrt{N}$ as in Z_A, the local alternative for panel stationarity tests becomes $\rho = c^2/(\sqrt{N}T^2)$.

Theorem 2 *Assume that Assumption 1 holds. Under H_1^ℓ, as N and T go to infinity simultaneously with $N/T \to 0$, the panel augmented KPSS test statistics has a limiting distribution given by*

$$Z_A \overset{(N,T)}{\Longrightarrow} N(0,1) + \frac{c^2}{\zeta} E\left[\int_0^1 F_i^v(r)^2 dr\right]$$

where $F_i^v(r) = \int_0^r B_i^v(s)ds - \int_0^r z(s)'ds \left(\int_0^1 z(s)z(s)'ds\right)^{-1} \int_0^1 z(s)B_i^v(s)ds$ *and* $B_i^v(r)$ *are independent Brownian motions across* i*,* $z(r) = 1$ *and* $E[\int_0^1 F_i^v(r)^2 dr]/\zeta = \sqrt{45}/90$ *when* $z_t = 1$ *and* $z(r) = [1, r]'$ *and* $E[\int_0^1 F_i^v(r)^2 dr]/\zeta = \sqrt{6300/11}(11/12600)$ *when* $z_t = [1, t]'$.

Proof see Appendix.

Theorem 2 shows that the test is more powerful when only a constant is included in the regression than the trending case as is the univariate KPSS test, because $\sqrt{45}/90 > \sqrt{6300/11}(11/12600)$.

We investigated via Monte Carlo simulations how accurately does the asymptotic theory approximate the finite sample behavior of the panel augmented KPSS. As a whole, we found that the finite sample behaviour of the panel augmented KPSS test is well approximated by the asymptotic theory established in this section when N and T are of moderate size. Due to space constraint the results are not reported here but can be requested from the authors.

4. Extension to the case of serially correlated errors

4.1. Modification of the panel augmented KPSS test

So far, we have investigated the theoretical property of the panel augmented KPSS test under restrictive assumptions. In this section we relax Assumptions 1(i) and consider a more practical and useful situation in which we allow for serial correlation.

Since it is often the case that the observed process can be approximated by an autoregressive (AR) model, we do not consider the error component model (1) but an AR(p) model instead in this section[5]:

$$y_{it} = z_t'\delta_i + f_t\gamma_i + \varepsilon_{it}, \quad \varepsilon_{it} = \phi_{i1}\varepsilon_{it-1} + \cdots + \phi_{ip}\varepsilon_{it-p} + \nu_{it}. \tag{8}$$

The lag length p may change depending on the cross-sectional units but we suppress the dependence of p on i for notational convenience.

Assumption 2 *(i) The stochastic process f_t is stationary with a finite fourth moment and the functional central limit theorem (FCLT) holds for the partial sum process of f_t. (ii) The stochastic process ν_{it} is independent of f_t and i.i.d.$(0, \sigma^2_{\nu i})$ across i and t with a finite fourth moments.*

This assumption allows the common factor to be stationary but still presumes that it is independent of the idiosyncratic errors, which are finite order AR processes with i.i.d. innovations. We assume Assumptions 1 (ii) and 2 in the rest of this section.

[5] We do not consider a general linear process instead of an AR(p) model because in the case of a general linear process the long-run variance estimator based on Toda and Yamamoto (1995), used here, will diverge to infinity at a rate T^2 under the alternative when the process is AR(∞). As a result, our test based on the lag-augmented method becomes inconsistent.

Since our interest is whether y_{it} are (trend) stationary or unit root processes, the testing problem is given by

$$H_0' : \ \phi_i(1) \neq 0 \ \ \forall i \quad \text{v.s.} \quad H_1' : \ \phi_i(1) = 0 \quad \text{for some } i,$$

where $\phi_i(L) = 1 - \phi_{i1}L - \cdots - \phi_{ip}L^p$.

In this case we need to modify the original KPSS test statistic for serial correlation as well as cross-sectional dependence. For the correction of cross-sectional dependence, we regress y_{it} on $w_t = [z_t', \bar{y}_t, \bar{y}_{t-1}, \cdots, \bar{y}_{t-p}]$ because ε_{it} are AR(p) processes and construct S_{it}^w using these regression residuals. In this case it is not difficult to see that the numerator of each ST_i weakly converges to

$$\frac{1}{T^2}\sum_{t=1}^{T}(S_{it}^w)^2 \overset{T}{\Longrightarrow} \sigma_i^2 \int_0^1 \left(V_i^{\varepsilon}(r) + O_p\left(\frac{1}{\sqrt{N}}\right)\right)^2 dr$$

where $\sigma_i^2 = \sigma_{\nu i}^2/(1-\phi_{i1}-\cdots-\phi_{ip})^2$ and $V_i^{\varepsilon}(r) = B_i^{\varepsilon}(r) - \int_0^r z(t)'dt \left(\int_0^1 z(t)z(t)'dt\right)^{-1} \int_0^1 z(t)dB_i^{\varepsilon}(t)$ with $B_i^{\varepsilon}(t)$ are independent standard Brownian motions. This suggests that we should divide the numerator of each ST_i by a consistent estimator of the long-run variance σ_i^2 in order to correct for serial correlation.

Several consistent estimators of the long-run variance[6] for parametric model have been proposed in the literature for univariate time series. For example, Leybourne and McCabe (1994) propose to correct the stationarity test for serial correlation by estimating the AR coefficients based on the ML method for the ARIMA model. Their method is also applied to panel data with no cross-sectional dependence by Shin and Snell (2006). However, our preliminary simulation shows that this method does not work well in finite samples and we do not use this method in this paper.

We next consider to make use of the new truncation rule proposed by Sul, Phillips and Choi (2005). Their method is originally developed for the prewhitening method, but it is

[6] We cannot use here the estimator of the long-run variance proposed in Perron and Ng (1998), despite its good properties in finite samples, because it is consistent under the null of a unit root but not under the null of stationarity which we are considering in this paper.

also applicable to parametric model. We first estimate the AR(p) model augmented by the lags of $\bar{y}_t$ for each i by the least squares method

$$y_{it} = z_t'\hat{\delta}_i + \hat{\phi}_{i1} y_{it-1} + \cdots + \hat{\phi}_{ip} y_{it-p} + \hat{\psi}_{i0}\bar{y}_t + \cdots + \hat{\psi}_{ip}\bar{y}_{t-p} + \hat{\nu}_{it},$$

and construct the estimator of the long-run variance by

$$\hat{\sigma}^2_{iSPC} = \frac{\hat{\sigma}^2_{\nu i}}{(1-\hat{\phi}_i)^2} \quad \text{where} \quad \hat{\phi}_i = \min\left\{1 - \frac{1}{\sqrt{T}}, \sum_{j=1}^{p} \hat{\phi}_{ij}.\right\} \quad \text{and} \quad \hat{\sigma}^2_{\nu i} = \frac{1}{T}\sum_{t=1}^{T} \hat{\nu}^2_{it}.$$

We then propose to construct the test statistic (7) using

$$ST_i^{SPC} = \frac{1}{\hat{\sigma}^2_{iSPC} T^2} \sum_{t=1}^{T} (S^w_{it})^2.$$

We denote this test statistic as Z_A^{SPC}.

The other method we consider is the lag-augmented method proposed by Choi (1993) and Toda and Yamamoto (1995). According to these papers, we intentionally add an additional lag of y_t and estimate an AR($p+1$) model instead of an AR(p) model:

$$y_{it} = z_t'\tilde{\delta}_i + \tilde{\phi}_{i1} y_{it-1} + \cdots + \tilde{\phi}_{ip} y_{it-p} + \tilde{\phi}_{ip+1} y_{it-p-1} + \tilde{\psi}_{i0}\bar{y}_t + \cdots + \tilde{\psi}_{ip}\bar{y}_{t-p} + \tilde{\nu}_{it},$$

and construct the test statistic using

$$ST_i^{LA} = \frac{1}{\hat{\sigma}^2_{iLA} T^2} \sum_{t=1}^{T} (S^w_{it})^2 \quad \text{where} \quad \hat{\sigma}^2_{iLA} = \frac{\hat{\sigma}^2_{\nu i}}{(1 - \tilde{\phi}_{i1} - \cdots - \tilde{\phi}_{ip})^2}.$$

We denote this test statistic as Z_A^{LA}.

The consistency of $\hat{\sigma}^2_{iSPC}$ and $\hat{\sigma}^2_{iLA}$ under the null hypothesis is established in the standard way and we omit here the details. On the other hand, they are shown to diverge to infinity at a rate of T under the alternative, so that ST_i can be seen as a consistent stationarity test for univariate time series. It is also shown by using the sequential limit that the null distributions of Z_A^{SPC} and Z_A^{LA} are asymptotically standard normal in the same way as Theorem 1 while they diverge to infinity under the alternative. Unfortunately, it is tedious to derive the joint limit of Z_A^{SPC} or Z_A^{LA} under general assumptions and we do not pursue it. Instead, we shall conduct Monte Carlo simulations in the next section in order to see whether or not the sequential limit theory can approximate the finite sample behaviour of these tests. However, we conjecture that the sequential limit is the same as the joint limit under the additional condition that $N/T \to 0$.

4.2. Finite sample property under general assumptions

In this section we conduct Monte Carlo simulations to investigate the finite sample properties of the panel augmented KPSS test using the long-run variance estimated by the SPC or the LA methods in order to correct for serial correlation in the innovations. The data generating process in this subsection is given as follows:

$$y_{it} = z_t'\delta_i + f_t\gamma_i + \varepsilon_{it}, \quad \varepsilon_{it} = \phi_i\varepsilon_{it-1} + \nu_{it},$$

where $f_t \sim i.i.d.N(0,1)$, $\nu_{it} \sim i.i.d.N(0,1)$, f_t and ν_{it} are independent of each other, $\delta_i = \alpha_i$ for the constant case while $\delta_i = [\alpha_i, \beta_i]'$ for the trend case with α_i and β_i being drawn from independent $U(0, 0.02)$, γ_i are drawn from $-1+U(0,4)$ for strong cross-sectional correlation case (SCC) and from $U(0, 0.02)$ for weak cross-sectional correlation case (WCC), and α_i, β_i and γ_i are fixed throughout the iterations. The ϕ_i are drawn from $0.1 + U(0, 0.8)$ under the null hypothesis and they remain fixed throughout the iterations. On the other hand, the ϕ_i are set to be equal to 1 for all i under the alternative. For the purpose of comparison, we also calculate the test statistic proposed by Harris, Leybourne and McCabe (2005) (HLM hereafter). According to HLM, we first estimate the idiosyncratic errors ε_{it} by the principal component method proposed by Bai and Ng (2004) and next apply the stationarity test proposed by Harris, McCabe and Leybourne (2003) to the estimated series of $\varepsilon_{1t}, \varepsilon_{2t}, \cdots, \varepsilon_{Nt}$. HLM method requires to predetermine the order of the autocovariance and the bandwidth parameter for the kernel estimate of the long-run variance; we set these parameters as recommended in HLM (2005).

Table 1 reports the sizes of the tests. There are no entries for HLM test when $T = 10$ because the time dimension is too short to calculate their test statistic. When only a constant is included in the model, the panel augmented KPSS test corrected by the SPC method tends to be undersized for moderate size of T for SCC (strong cross-correlation) case while it is oversized for small or large size of T, although the over-rejection is not so severe when $N = 100$ and $T = 200$. For WCC (weak cross-correlation) case Z_A^{SPC} is undersized[7] except for the case of $T = 10$. The augmented KPSS test corrected by the LA method has a similar property as Z_A^{SPC} for SCC case while the size of the test is relatively well controlled for WCC case. On the other hand, the size of HLM test seems to be better controlled for moderate or large size of T, although the test becomes undersized for large size of N and small or moderate size of T.

When both a constant and a linear trend are included in the model, the overall property of Z_A^{SPC} and Z_A^{LA} is preserved while HLM test tends to be undersized for N larger than 20.

Table 2 shows the nominal powers of the tests. Because of the size distortion of the tests it is not easy to compare the powers of these tests but we observe that all the tests are less powerful for the moderate size of T due to the undersize property of the tests. In some cases the panel augmented KPSS test apparently dominates HLM test but the reversed relation is observed in other cases. For example, the empirical sizes of Z_A^{SPC}, Z_A^{LA} and HLM test are 0.009, 0.022 and 0.078 when $N = 10$ and $T = 30$ for the constant case with SCC, while the powers of these tests are 0.437, 0.262 and 0.218. On the other hand, the sizes of these tests are 0.058, 0.076 and 0.054 when $N = 10$ and $T = 100$ for the constant case with WCC while the powers are 0.878, 0.812 and 1.00.

Although our simulations are limited, it is difficult to recommend one of these tests because none of them dominates the others. It seems that HLM test tends to work relatively

[7] It seems that the long-run variance is well estimated by the method proposed by Sul *et al.* (2005). But it is well known that the numerator of the KPSS statistic has a downward bias (cf *inter alia* Shin and Snell (2006) and Kurozumi and Tanaka (2009)). As a result, each test statistic ST_i is downward biased. These downward biases accumulate as N increases leading to the undersize of the tests based on Z_A^{SPC}. Another problem raised by one of the referees is that the centering and scaling constants are derived asymptotically, $T \to \infty$. When T is finite these constants may be inappropriate. We did some additional simulations employing the bias corrections proposed by Kurozumi and Tanaka (2009) and the centering and scaling constants for fixed T suggested by Hadri and Larsson (2005). We found that the results after corrections are very similar to those before corrections except for the case where $T = 10$. In the latter case, the finite sample corrections are effective in reducing the severe over-size distorsions.

well in the constant case because the size of the test is more or less controlled in many cases and it has moderate power, whereas the panel augmented KPSS test with SPC correction seems to perform best in many cases corresponding to the trend case (all the other tests tend to be undersized in this case) and is most powerful in many cases.

5. Conclusion

In this paper we extended Hadri's (2000) test to correct for cross-sectional dependence à la Pesaran (2007). We showed that the limiting null distribution of this panel augmented KPSS test is the same as the original Hadri's test that is the LM test without cross-sectional dependence. All the theoretical results are derived via the joint asymptotic theory and the limiting distributions under the null are shown to be standard normal. We also propose a more practical panel augmented KPSS under more realistic assumptions and allowing for serial correlation in the error disturbances. The Monte Carlo simulations indicated that we should use the panel stationarity tests with care because they are undersized in some cases but suffer from over rejection in other cases.

References

[1] Abadir, K. M. (1993). OLS Bias in nonstationary autoregression. *Econometric Theory 9,* 81-93.

[2] Abadir, K. M. and K. Hadri (2000). Is more information a good thing? Bias non-monotonicity in stochastic difference equations. *Bulletin of Economic Research 52*, 91-100.

[3] Bai, J. and S. Ng (2004). A panic attack on unit roots and cointegration. *Econometrica 72,* 1127-1177.

[4] Bai, J. and S. Ng (2005). A new look at panel testing of stationarity and the PPP hypothesis, D. W. K. Andrews and J. H. Stock, ed., *Identification and Inference for Econometric Models. Essays in Honor of Thomas Rothenberg.* Cambridge University Press, Cambridge.

[5] Baltagi, B. H. (2001). *Econometric Analysis of Panel Data* . Chichester, Wiley.

[6] Baltagi, B. H. and C. Kao (2000). Nonstationary panels, cointegration in panels and dynamic panels: A survey. *Advances in Econometrics 15*, 7-51.

[7] Banerjee, A. (1999). Panel data unit roots and cointegration: An overview. *Oxford Bulletin of Economics and Statistics Special issue*, 607-29.

[8] Banerjee, A., M. Marcellino and C. Osbat (2001). Testing for PPP: Should we use panel methods? *Empirical Economics 30,* 77-91.

[9] Banerjee, A., M. Marcellino and C. Osbat (2004). Some cautions on the use of panel methods for integrated series of macroeconomic data. *Econometrics Journal 7,* 322-340.

[10] Breitung, J. and W. Meyer (1994). Testing for unit roots in panel data: Are wages on different bargaining levels cointegrated? *Applied Economics 26,* 353-61.

[11] Breitung, J. and Pesaran, M. H. (2008). Unit roots and cointegration in panels. In: L. Matyas and P. Sevestre (eds): *The Econometrics of Panel Data: Fundamentals and Recent Developments in Theory and Practice. Dordrecht:* Kluwer Academic Publishers, 3rd edition, Chapt. 9, pp. 279-322.

[12] Chang, Y. (2004). Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics 110,* 261-292.

[13] Choi, I. (1993). Asymptotic normality of the least-squares estimates for higher order autoregressive integrated processes with some applications. *Econometric Theory 9,* 263-282.

[14] Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Banking 20,* 249-272.

[15] 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.

[16] Chung, K. L. (1974). *A Course in Probability Theory (2nd ed.)*, Academic Press, San Diego.

[17] Culver, S. E. and D. H. Papell (1997). Is there a unit root in the inflation rate? Evidence from sequential break and panel data models, *Journal of Applied Econometrics 12, 4,* 435-444.

[18] de Silva, S., K. Hadri and A. R. Tremayne (2009). Panel unit root tests in the presence of cross-sectional dependence: Finite sample performance and an application, forthcoming in *Econometrics Journal.*

[19] Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *Econometrics Journal 3,* 148-161.

[20] Hadri, K. and Larsson, R. (2005). Testing for stationarity in heterogeneous panel data where the time dimension is finite. *Econometrics Journal 8,* 55-69.

[21] Hadri, K. and G. D. A. Phillips (1999). The accuracy of the higher order bias approximation for the 2SLS estimator. *Economic Letters 62*, 167-74.

[22] Hadri K. and Y. Rao, (2008) Panel stationarity test with breaks. *Oxford Bulletin of Economics and Statistics, 7*0(2), 245-269.

[23] Harris, D., B. McCabe and S. Leybourne (2003). Some limit theory for autocovariances whose order depends on sample size. *Econometric Theory 19,* 829-864.

[24] Harris, D., S. Leybourne and B. McCabe (2005). Panel stationarity tests for purchasing power parity with cross-sectional dependence. *Journal of Business and Economic Statistics 23,* 395-409.

[25] Im, K., M. H. Pesaran and Y. Shin (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics 115,* 53-74.

[26] Kurozumi, E. (2002). Testing for stationarity with a break. *Journal of Econometrics,* Vol. 108, pp. 63-69.

[27] Kurozumi, E. and S. Tanaka (2009). Reducing the size distortion of the KPSS test, Global COE Hi-Stat Discussion Paper Series 085, Hitotsubashi University, Japan.

[28] Kwiatkowski, D., P. C. B. Phillips, P. Schmidt and Y. Shin (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics 54,* 159-178.

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

[30] 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,* 631-52.

[31] Moon, R. and B. Perron (2004). Testing for unit roots in panels with dynamic factors. *Journal of Econometrics 122,* 81-126.

[32] O'Connell, P. G. J. (1998). The overvaluation of purchasing power parity. *Journal of International Economics 44,* 1-19.

[33] Oh, K. Y. (1996). Purchasing power parity and unit root tests using panel data. *Journal of International Money and Finance 15,* 405-418.

[34] Papell, D. H. (1997). Searching for stationarity: Purchasing power parity under the current float. *Journal of International Economics 43,* 313-332.

[35] Papell, D. H. (2002). The great appreciation, the great depreciation and the purchasing power parity hypothesis. *Journal of International Economics 57,* 51-82.

[36] Perron, P. and S. Ng (1998). An autoregressive spectral density estimator at frequency zero for non-stationary tests. *Econometric Theory*, Vol. 4(05), 560-603.

[37] Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics 22,* 265-312.

[38] Pesaran, M. H., L. V. Smith and T. Yamagata (2009). A panel unit root test in the presence of a multifactor error structure. working paper, University of Cambridge.

[39] Phillips, P. C. B. and H. R. Moon (1999). Linear regression limit theory for nonstationary panel data. *Econometrica 67,* 1057-1111.

[40] Phillips, P. C, B. and D. Sul (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal 6,* 217-25.

[41] Quah, D. (1994). Exploiting cross section variation for unit root inference in dynamic data. *Economics Letters 44,* 9-19.

[42] Shin, Y. and A. Snell (2006). Mean group tests for stationarity in heterogeneous panels. *Econometrics Journal 9,* 123-158.

[43] Sul, D., P. C. B. Phillips and C. Y. Choi (2005). Prewhitening bias in HAC estimation. *Oxford Bulletin of Economics and Statistics 67,* 517-546.

[44] Tanaka, K. (1996). Time Series Analysis, Nonstatinary and Noninvertible Distribution Theory. John Wiley.

[45] Toda, H. Y., and T. Yamamoto (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics 66,* 225-250.

[46] Wu, S. and J. L. Wu (2001). Is purchasing power parity overvalued? *Journal of Money, Credit and Banking 33,* 804-812.

[47] Wu, Y. (1996). Are real exchange rates nonstationary? Evidence from a panel data test. *Journal of Money, Credit and Banking 28,* 54-63.

Appendix

In this appendix, we denote some constant that is independent of N, T and the subscripts i and t as C, C_1, C_2, $\cdots$. We prove the theorems only for the case where $z_t = [1, t]$. The proof for the level case with $z_t = 1$ proceeds in exactly the same way and thus we omit it. We also assume $\sigma_\varepsilon^2 = 1$ in this appendix without loss of generality because we know σ_ε^2 under Assumption 1(i).

We first express $\bar{y}_t$ in matrix form. Since

$$\bar{y}_t = z_t'\bar{\delta} + \bar{r}_t + f_t\bar{\gamma} + \bar{\varepsilon}_t,$$

we have

$$\begin{bmatrix} \bar{y}_1 \\ \bar{y}_2 \\ \vdots \\ \bar{y}_T \end{bmatrix} = \begin{bmatrix} z_1' \\ z_2' \\ \vdots \\ z_T' \end{bmatrix} \bar{\delta} + \begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \\ \vdots \\ \bar{r}_T \end{bmatrix} + \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_T \end{bmatrix} \bar{\gamma} + \begin{bmatrix} \bar{\varepsilon}_1 \\ \bar{\varepsilon}_2 \\ \vdots \\ \bar{\varepsilon}_T \end{bmatrix}$$

or

$$\bar{\mathbf{y}} = Z\bar{\delta} + \bar{\mathbf{r}} + \mathbf{f}\bar{\gamma} + \bar{\boldsymbol{\varepsilon}}. \tag{9}$$

Since $\bar{\gamma} \neq 0$, we have $\mathbf{f} = (\bar{\mathbf{y}} - Z\bar{\delta} - \bar{\mathbf{r}} - \bar{\boldsymbol{\varepsilon}})/\bar{\gamma}$. By inserting this into (2), the model becomes

$$\mathbf{y}_i = Z(\delta_i - \tilde{\gamma}_i\bar{\delta}) + \tilde{\gamma}_i\bar{\mathbf{y}} + (\mathbf{r}_i - \tilde{\gamma}_i\bar{\mathbf{r}}) + (\boldsymbol{\varepsilon}_i - \tilde{\gamma}_i\bar{\boldsymbol{\varepsilon}}) \tag{10}$$

where $\tilde{\gamma}_i = \gamma_i/\bar{\gamma}$.

Let $W = [\boldsymbol{\tau}, \mathbf{d}, \bar{\mathbf{y}}] = [Z, \bar{\mathbf{y}}]$ and $W^* = WQ = [Z, \bar{\mathbf{y}}^*]$ where $\bar{\mathbf{y}}^* = \bar{\mathbf{y}} - Z\bar{\delta} = \bar{\mathbf{r}} + \mathbf{f}\bar{\gamma} + \bar{\boldsymbol{\varepsilon}}$,

$$Q = \begin{bmatrix} I_2 & -\bar{\delta} \\ 0 & 1 \end{bmatrix}, \quad D = \begin{bmatrix} D_\tau & 0 \\ 0 & \sqrt{T} \end{bmatrix} \quad \text{and} \quad D_\tau = \begin{bmatrix} \sqrt{T} & 0 \\ 0 & T\sqrt{T} \end{bmatrix}.$$

Because $M_w = M_{w^*}$, ST_i in the augmented KPSS test statistic can be expressed in matrix form as

$$ST_i = \frac{1}{T^2}\mathbf{y}_i' M_{w^*} B' B M_* \mathbf{y}_i.$$

Before proceeding with the proof of theorems, we state two lemmas, which will be used in the proof repeatedly. The first lemma states various moments related to r_{it}. Since these can be obtained by direct calculation, we omit the proof.

Lemma 1 *Let $v_{it} \sim i.i.d.N(0, \sigma_v^2)$ for $i = 1, \cdots, N$ and $t = 1, \cdots, T$, $r_{it} = \sum_{s=1}^{t} v_{is}$ and $\bar{r}_t = N^{-1} \sum_{i=1}^{N} r_{it}$. Then,*

$$
\begin{aligned}
E\left[r_{is} r_{it}\right] &= \sigma_v^2 \min(s,t) && (11)\\
E\left[\left(\sum_{s=1}^{t} r_{is}\right)^2\right] &= \frac{\sigma_v^2}{6} t(t+1)(2t+1) && (12)\\
E\left[\left(\sum_{s=1}^{t} s r_{is}\right)^2\right] &= \frac{\sigma_v^2}{30} t(t+1)(2t+1)(2t^2+2t+1) && (13)\\
E\left[\bar{r}_s \bar{r}_t\right] &= \frac{\sigma_v^2}{N} \min(s,t) && (14)\\
E\left[\left(\sum_{s=1}^{t} \bar{r}_s\right)^2\right] &= \frac{\sigma_v^2}{6N} t(t+1)(2t+1) && (15)\\
E\left[\left(\sum_{s=1}^{t} s\bar{r}_s\right)^2\right] &= \frac{\sigma_v^2}{30N} t(t+1)(2t+1)(2t^2+2t+1), && (16)
\end{aligned}
$$

$$
\begin{aligned}
E\left[\left(\sum_{s=1}^{t} r_{is}\right)\left(\sum_{t=1}^{T} r_{it}\right)\right] &= \frac{\sigma_v^2}{6} t(t+1)(3T-t+1) && (17)\\
E\left[\left(\sum_{s=1}^{t} r_{is}\right)\left(\sum_{t=1}^{T} t r_{it}\right)\right] &= \frac{\sigma_v^2}{24} t(t+1)(6T^2+6T-t^2-t+2) && (18)\\
E\left[\left(\sum_{t=1}^{T} r_{it}\right)\left(\sum_{t=1}^{T} t r_{it}\right)\right] &= \frac{\sigma_v^2}{24} T(T+1)(5T^2+5T+2), && (19)\\
E\left[r_{is} r_{it} r_{iu} r_{iv}\right] &= \sigma_v^4 (2st + su) \quad \text{for} \quad s \le t \le u \le v. && (20)
\end{aligned}
$$

The next lemma gives the sufficient condition on the equivalence of the sequential limit to the joint limit. Notice that when the statistic S_{iT} weakly converges to $S_{i\infty}$ as $T \to \infty$, we can construct the probability space on which both S_{iT} and $S_{i\infty}$ exist, as discussed in Phillips and Moon (1999).

Lemma 2 *Let S_{iT} and $S_{i\infty}$ are i.i.d. sequences across i ($i = 1, \cdots, N$) on the same probability space. Assume that $S_{i\infty}$ does not depend on N, S_{iT} is independent of $S_{j\infty}$ for $i \neq j$*

and $S_{iT} \overset{T}{\Longrightarrow} S_{i\infty}$ *as* $T \to \infty$. *If (a)* $E[S_{iT}] \to \mu_1 \equiv E[S_{i\infty}] < \infty$ *as both* N *and* T *go to infinity, and (b)* $\sup_{N,T} E[S_{iT}^2] < \infty$, *then,*

$$\frac{1}{N}\sum_{i=1}^{N} S_{iT} \overset{p\ (N,T)}{\longrightarrow} \mu_1.$$

Proof of Lemma 2: Since S_{iT} is an i.i.d. sequence, we have, for any arbitrary $\varepsilon > 0$,

$$P\left(\left|\frac{1}{N}\sum_{i=1}^{N} S_{iT} - E[S_{iT}]\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^2 N} E[(S_{iT} - E[S_{iT}])^2] \leq \frac{1}{\varepsilon^2 N}\sup_{N,T} E[S_{iT}^2] \to 0$$

by condition (b) as both N and T go to infinity. Because $E[S_{iT}] \to \mu_1$ by condition (a), we can see that $N^{-1}\sum_{i=1}^{N} S_{iT} \overset{p\ (N,T)}{\longrightarrow} \mu_1$.□

Proof of Theorem 1

Because $\mathbf{r}_i$ and $\bar{\mathbf{r}}$ disappear under the null hypothesis, ST_i can be expressed in matrix form under H_0 as

$$\begin{aligned}
ST_i &= \frac{1}{T^2}\mathbf{y}_i' M_{w^*} B' B M_{w^*} \mathbf{y}_i \\
&= \frac{1}{T^2}(\boldsymbol{\varepsilon}_i - \tilde{\gamma}_i \bar{\boldsymbol{\varepsilon}})' M_{w^*} B' B M_{w^*} (\boldsymbol{\varepsilon}_i - \tilde{\gamma}_i \bar{\boldsymbol{\varepsilon}}) \\
&= \frac{1}{T^2}\boldsymbol{\varepsilon}_i' M_{w^*} B' B M_{w^*} \boldsymbol{\varepsilon}_i - \frac{2\tilde{\gamma}_i}{T^2}\bar{\boldsymbol{\varepsilon}}' M_{w^*} B' B M_{w^*} \boldsymbol{\varepsilon}_i + \frac{\tilde{\gamma}_i^2}{T^2}\bar{\boldsymbol{\varepsilon}}' M_{w^*} B' B M_{w^*} \bar{\boldsymbol{\varepsilon}} \\
&= ST_{1i} - 2\tilde{\gamma}_i ST_{2i} + \tilde{\gamma}_i^2 ST_{3i}, \quad \text{say.}
\end{aligned}$$

Let $ST_{1i}^0 = T^{-2}\boldsymbol{\varepsilon}_i' M_z B' B M_z \boldsymbol{\varepsilon}_i$. Since Shin and Snell (2006) showed that

$$\frac{\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(ST_{1i}^0 - \xi)}{\zeta} \overset{(N,T)}{\Longrightarrow} N(0,1),$$

it is sufficient to prove that

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}(ST_{1i} - ST_{1i}^0) \overset{p\ (N,T)}{\longrightarrow} 0, \tag{21}$$

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N} ST_{2i} \overset{p\ (N,T)}{\longrightarrow} 0, \tag{22}$$

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N} ST_{3i} \overset{p\ (N,T)}{\longrightarrow} 0. \tag{23}$$

Let $J_{0i} = T^{-1}B\boldsymbol{\varepsilon}_i$, $[J_1, J_2] = T^{-1}BW^*D^{-1} = [T^{-1}BZD_\tau^{-1}, T^{-3/2}B\bar{\mathbf{y}}^*]$, $[L_{1i}', L_{2i}'] = D^{-1}W^{*\prime}\boldsymbol{\varepsilon}_i = [(D_\tau^{-1}Z'\boldsymbol{\varepsilon}_i)', (T^{-1/2}\bar{\mathbf{y}}^{*\prime}\boldsymbol{\varepsilon}_i)']'$, $K = [[K_{ij}]] = D^{-1}W^{*\prime}W^*D^{-1}$ and $K^{-1} = [[K^{ij}]]$ for $i, j = 1, 2$. Then, we have

$$\begin{aligned} \frac{1}{T}BM_{w^*}\boldsymbol{\varepsilon}_i &= \frac{1}{T}B\boldsymbol{\varepsilon}_i - \frac{1}{T}BW^*D^{-1}\left(D^{-1}W^{*\prime}W^*D^{-1}\right)^{-1}D^{-1}W^{*\prime}\boldsymbol{\varepsilon}_i \\ &= (J_{0i} - J_1K^{11}L_{1i}) - \{J_2K^{21}L_{1i} + (J_1K^{12} + J_2K^{22})L_{2i}\}. \qquad (24) \end{aligned}$$

Similarly, by letting $\bar{J}_0 = N^{-1}\sum_{i=1}^N J_{0i}$, $\bar{L}_1 = N^{-1}\sum_{i=1}^N L_{1i}$ and $\bar{L}_2 = N^{-1}\sum_{i=1}^N L_{2i}$, we can see that

$$\frac{1}{T}BM_{w^*}\bar{\boldsymbol{\varepsilon}}_i = \bar{J}_0 - (J_1K^{11} + J_2K^{21})\bar{L}_1 - (J_1K^{12} + J_2K^{22})\bar{L}_2. \qquad (25)$$

We first prove (21). Using expression (24), ST_{1i} can be decomposed into

$$\begin{aligned} ST_{1i} &= (J_{0i} - J_1K^{11}L_{1i})'(J_{0i} - J_1K^{11}L_{1i}) \\ &\quad -2(J_{0i} - J_1K^{11}L_{1i})'\{J_2K^{21}L_{1i} + (J_1K^{12} + J_2K^{22})L_{2i}\} \\ &\quad +\{J_2K^{21}L_{1i} + (J_1K^{12} + J_2K^{22})L_{2i}\}'\{J_2K^{21}L_{1i} + (J_1K^{12} + J_2K^{22})L_{2i}\} \\ &= ST_{1i}^a + ST_{1i}^b + ST_{1i}^c, \quad \text{say}. \qquad (26) \end{aligned}$$

In order to evaluate each term, we use the following lemma.

Lemma 3 *Under the null hypothesis, as both N and T go to infinity simultaneously, (i)* $E\|J_{0i}\|^2 \le C$, $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|J_{0i}\|^2 = O_p(\sqrt{N})$ *and* $\|\bar{J}_0\| = O_p(\frac{1}{\sqrt{N}})$, *(ii)* $\|J_1\| = O(1)$, *(iii)* $E\|J_2\|^2 \le \frac{C}{T}$ *and* $\|J_2\| = O_p(\frac{1}{\sqrt{T}})$, *(iv)* $K_{11} = O(1)$, $K_{12} = K_{21}' = O_p(\frac{1}{\sqrt{T}})$, $K_{22} = O_p(1)$, $K^{11} = K_{11}^{-1} + O_p(\frac{1}{T})$, $K_{21} = K_{21}' = O_p(\frac{1}{\sqrt{T}})$ *and* $K^{22} = O_p(1)$, *(v)* $E\|L_{1i}\|^2 \le C$, $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|L_{1i}\|^2 = O_p(\sqrt{N})$ *and* $\|\bar{L}_1\| = O_p(\frac{1}{\sqrt{N}})$, *(vi)* $E\|L_{2i}\|^2 \le C$, $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|L_{2i}\|^2 = O_p(\sqrt{N})$ *and* $\bar{L}_2 = O_p(\frac{\sqrt{T}}{N})$, *(vii)* $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|J_{0i}\|\|L_{\ell i}\| = O_p(\sqrt{N})$ *for* $\ell, m = 1, 2$ *and* $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|L_{1i}\|\|L_{2i}\| = O_p(\sqrt{N})$.

Proof of Lemma 3: (i) Since ε_{it} is i.i.d.$N(0,1)$, we have

$$E\|J_{0i}\|^2 = E\left[\frac{1}{T^2}\sum_{t=1}^T\left(\sum_{s=1}^t \varepsilon_{is}\right)^2\right] = \frac{1}{T^2}\sum_{t=1}^T t \le C.$$

Since C does not depend on i and J_{0i} is independent of J_{0j} for $i \neq j$, we can see that $E[\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\|J_{0i}\|^2] \leq C\sqrt{N}$ and $E[\|\bar{J}_0\|^2] \leq C/N$, which imply (i).

(ii) It is easy to see that

$$J_1'J_1 = \begin{bmatrix} \frac{1}{T^3}\sum_{t=1}^{T} t^2 & \frac{1}{T^4}\sum_{t=1}^{T} t\sum_{s=1}^{t} s \\ \frac{1}{T^4}\sum_{t=1}^{T} t\sum_{s=1}^{t} s & \frac{1}{T^5}\sum_{t=1}^{T}(\sum_{s=1}^{t} s)^2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{1}{8} \\ \frac{1}{8} & \frac{1}{20} \end{bmatrix} + O\left(\frac{1}{T}\right).$$

(iii) Since $\bar{y}_t^* = \bar{\gamma} f_t + \bar{\varepsilon}_t$ under H_0, f_t and $\bar{\varepsilon}_t$ are independent and $\bar{\varepsilon}_t$ is $i.i.d.N(0, 1/N)$,

$$\begin{aligned} E\|J_2\|^2 &= \frac{1}{T^3}E\left[\sum_{t=1}^{T}\left\{\sum_{s=1}^{t}(\bar{\gamma} f_s + \bar{\varepsilon}_s)\right\}^2\right] \\ &= \frac{1}{T^3}\sum_{t=1}^{T}\left\{\bar{\gamma}^2 E\left[\left(\sum_{s=1}^{t} f_s\right)^2\right] + E\left[\left(\sum_{s=1}^{t}\bar{\varepsilon}_s\right)^2\right]\right\} \\ &= \frac{1}{T^3}\sum_{t=1}^{T}\left(\bar{\gamma}^2\sigma_f^2 t + \frac{t}{N}\right) \leq \frac{C}{T} \end{aligned} \tag{27}$$

and thus we obtain (iii).

(iv) The result of K_{11} is obvious, while K_{12} and K_{22} are expressed as

$$K_{12} = \frac{1}{\sqrt{T}}D_\tau^{-1}Z'\bar{\mathbf{y}}^* = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T}(\bar{\gamma} f_t + \bar{\varepsilon}_t) \\ \frac{1}{T^2}\sum_{t=1}^{T} t(\bar{\gamma} f_t + \bar{\varepsilon}_t) \end{bmatrix} \quad \text{and} \quad K_{22} = \frac{1}{T}\sum_{t=1}^{T}(\bar{\gamma} f_t + \bar{\varepsilon}_t)^2.$$

We can see that each component of K_{12} has mean zero and variance bounded above by C/T while the expectation of K_{22} is bounded above by C. On the other hand, the orders of K^{ij} for $i, j = 1, 2$ are obtained by using the inversion formula of a partitioned matrix. For example,

$$K^{11} = K_{11}^{-1} + K_{11}^{-1}K_{12}(K_{22} - K_{21}K_{11}^{-1}K_{12})^{-1}K_{21}K_{11}^{-1} = K_{11}^{-1} + O_p\left(\frac{1}{T}\right).$$

The orders of $K^{12} = K^{21\prime}$ and K^{22} are obtained similarly.

The first and second assertions of (v) and (vi) are obtained by noting that

$$E\|L_{1i}\|^2 = E\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\varepsilon_{it}\right)^2\right] + E\left[\left(\frac{1}{T\sqrt{T}}\sum_{t=1}^{T} t\varepsilon_{it}\right)^2\right] \leq C,$$

and

$$
\begin{aligned}
E\|L_{2i}\|^2 &= E\left[\left\{\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\bar{\gamma}f_t+\bar{\varepsilon}_t)\,\varepsilon_{it}\right\}^2\right] \\
&\leq CE\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}f_t\varepsilon_{it}\right)^2+\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\bar{\varepsilon}_t\varepsilon_{it}\right)^2\right]\leq C.
\end{aligned}
$$

The third assertion of (v) is established from

$$
E\|\bar{L}_1\|^2=\frac{1}{T}E\left[\left(\sum_{t=1}^{T}\bar{\varepsilon}_t\right)^2\right]+\frac{1}{T^3}E\left[\left(\sum_{t=1}^{T}t\bar{\varepsilon}_t\right)^2\right]\leq\frac{C}{N}.
$$

For $\bar{L}_2$, note that

$$
\bar{L}_2=\frac{\bar{\gamma}}{\sqrt{T}}\sum_{t=1}^{T}f_t\bar{\varepsilon}_t+\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\bar{\varepsilon}_t^2.
$$

The first term is easily shown to be $O_p(1/\sqrt{N})$ because f_t is independent of $\bar{\varepsilon}_t$, whereas the expectation of the second term equals $\sqrt{T}/N$. Since $\sqrt{T}/N$ dominates $1/\sqrt{N}$, which follows from condition $N/T\to 0$, we have $\|\bar{L}_2\|=O_p(\sqrt{T}/N)$.

(vii) Since $E\|J_{0i}\|\|L_{\ell i}\|\leq(E\|J_{0i}\|^2E\|L_{\ell i}\|^2)^{1/2}$ and $E\|L_{1i}\|\|L_{2i}\|\leq(E\|L_{1i}\|^2E\|L_{2i}\|^2)^{1/2}$ by Cauchy-Schwarz inequality, we obtain the result by using (i), (v) and (vi).□

Since $ST_{1i}^0=(J_{0i}-J_1K_{11}^{-1}L_{1i})'(J_{0i}-J_1K_{11}^{-1}L_{1i})$, we have

$$
\begin{aligned}
ST_{1i}^a-ST_{1i}^0 &= \{J_1(K^{11}-K_{11}^{-1})L_{1i}\}'\{J_1(K^{11}-K_{11}^{-1})L_{1i}\} \\
&\quad -2(J_{0i}-J_1K_{11}^{-1}L_{1i})'\{J_1(K^{11}-K_{11}^{-1})L_{1i}\}. \qquad (28)
\end{aligned}
$$

Using Lemma 3 (ii), (iv) and (v), we can see that

$$
\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\|\{J_1(K^{11}-K_{11}^{-1})L_{1i}\}'\{J_1(K^{11}-K_{11}^{-1})L_{1i}\}\| \\
&\leq \|J_1\|^2\|K^{11}-K_{11}^{-1}\|^2\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\|L_{1i}\|^2=O_p\left(\frac{\sqrt{N}}{T^2}\right).
\end{aligned}
$$

Similarly, the second term on the right hand side of (28) becomes $O_p(\sqrt{N}/T)$. Hence, the right hand side of (28) converges to 0 in probability when both N and T go to infinity and $N/T\to 0$.

In exactly the same manner, we have

$$
\begin{aligned}
\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\|ST_{1i}^{b}\| &\leq \frac{C}{\sqrt{N}}\sum_{i=1}^{N}(\|J_{0i}\|+\|J_1\|\|K^{11}\|\|L_{1i}\|) \\
&\quad \times(\|J_2\|\|K^{21}\|\|L_{1i}\|+\|J_1\|\|K^{12}\|\|L_{2i}\|+\|J_2\|\|K^{22}\|\|L_{2i}\|) \\
&= O_p\left(\frac{\sqrt{N}}{\sqrt{T}}\right), \\
\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\|ST_{1i}^{c}\| &\leq \frac{1}{\sqrt{N}}\sum_{i=1}^{N}\{\|J_2\|\|K^{21}\|\|L_{1i}\|+(\|J_1\|\|K^{12}\|+\|J_2\|\|K^{22}\|)\|L_{2i}\|\}^2 \\
&= O_p\left(\frac{\sqrt{N}}{T}\right).
\end{aligned}
$$

Therefore, we obtained (21).

In order to prove (22) and (23), notice that

$$
\begin{aligned}
\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}ST_{2i}\right| &\leq C\left|\frac{1}{T^2}\bar{\boldsymbol{\varepsilon}}'M_{w^*}B'BM_{w^*}\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\boldsymbol{\varepsilon}_i\right| \\
&\leq C\frac{\sqrt{N}}{T^2}\bar{\boldsymbol{\varepsilon}}'M_{w^*}B'BM_{w^*}\bar{\boldsymbol{\varepsilon}}, \\
\left|\frac{1}{\sqrt{N}}\sum_{i=1}^{N}ST_{3i}\right| &\leq C\frac{\sqrt{N}}{T^2}\bar{\boldsymbol{\varepsilon}}'M_{w^*}B'BM_{w^*}\bar{\boldsymbol{\varepsilon}}.
\end{aligned}
$$

Then, it is sufficient to show that

$$
\frac{\sqrt{N}}{T^2}\bar{\boldsymbol{\varepsilon}}'M_{w^*}B'BM_{w^*}\bar{\boldsymbol{\varepsilon}} \overset{p\ (N,T)}{\longrightarrow} 0,
$$

which can be proved by noting that

$$
\left\|\frac{1}{T}BM_{w^*}\bar{\varepsilon}\right\| = O_p\left(\frac{1}{\sqrt{N}}\right)
$$

from expression (25) and Lemma 3. We thus obtain the result for the augmented KPSS test statistic.

Proof of Theorem 2

We first show that Lemma 3 still holds under the local alternative H_1^ℓ. Since $\bar{y}_t^* = \bar{r}_t + \bar{\gamma} f_t + \bar{\varepsilon}_t$ under H_1^ℓ, it is sufficient to show that Lemma 3 (iii), (iv), (vi) and (vii) hold because only J_2, K_{12}, K_{22} and L_{2i} are related to $\bar{y}_t^*$. Using Lemma 1 with $\sigma_v^2 = \rho\sigma_\varepsilon^2 = c^2/(\sqrt{N}T^2)$, (iii) is prove by noting that

$$\begin{aligned} E\|J_2\|^2 &= \sum_{t=1}^T E\left[\left(\frac{1}{T\sqrt{T}}\sum_{s=1}^t \bar{y}_s^*\right)^2\right] \\ &\leq 2\sum_{t=1}^T E\left[\left(\frac{1}{T\sqrt{T}}\sum_{s=1}^t \bar{r}_s\right)^2 + \left\{\frac{1}{T\sqrt{T}}\sum_{s=1}^t (\bar{\gamma} f_s + \bar{\varepsilon}_s)\right\}^2\right] \\ &\leq C_1 \sum_{t=1}^T \left(\frac{t^3}{N\sqrt{N}T^5}\right) + \frac{C_2}{T} \leq \frac{C}{T}. \end{aligned}$$

To show (iv), note that

$$K_{12} = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^T (\bar{r}_t + \bar{\gamma} f_t + \bar{\varepsilon}_t) \\ \frac{1}{T^2}\sum_{t=1}^T t(\bar{r}_t + \bar{\gamma} f_t + \bar{\varepsilon}_t) \end{bmatrix} \quad \text{and} \quad K_{22} = \frac{1}{T}\sum_{t=1}^T (\bar{r}_t + \bar{\gamma} f_t + \bar{\varepsilon}_t)^2.$$

As shown in the proof of Lemma 3 (iv) the variances of the second and third terms in each row of K_{12} are bounded by C/T, whereas $E[(T^{-1}\sum_{t=1}^T \bar{r}_t)^2] \leq C/(N\sqrt{N}T)$ and $E[(T^{-2}\sum_{t=1}^T t\bar{r}_t)^2] \leq C/(N\sqrt{N}T)$ from (15) and (16). This shows that $K_{12} = O_p(1/\sqrt{T})$. Similarly, we can see that $K_{22} = O_p(1)$ using Lemma 3 and noting that $E[T^{-1}\sum_{t=1}^T \bar{r}_t^2] \leq C/(N\sqrt{N}T)$ from (14).

(vi) is obtained by noting that

$$L_{2i} = \frac{1}{\sqrt{T}}\sum_{t=1}^T \bar{r}_t \varepsilon_{it} + \frac{1}{\sqrt{T}}\sum_{t=1}^T (\bar{\gamma} f_t + \bar{\varepsilon}_t)\varepsilon_{it},$$

and $E[(T^{-1/2}\sum_{t=1}^T \bar{r}_t \varepsilon_{it})^2] \leq C/(N\sqrt{N}T)$,which is proved using (14) and the independence between $\bar{r}_t$ and ε_{it}.

Next, we give a lemma similar to Lemma 3. Let $J_{0i}^r = T^{-1}B\mathbf{r}_i$, $[L_{1i}^{r\prime}, L_{2i}^{r\prime}]' = D^{-1}W^{*\prime}\mathbf{r}_i = [(D_\tau^{-1}Z'\mathbf{r}_i)', (T^{-1/2}\bar{\mathbf{y}}^{*\prime}\mathbf{r}_i)']'$, $\bar{J}_0^r = N^{-1}\sum_{i=1}^N J_{0i}^r$, $\bar{L}_1^r = N^{-1}\sum_{i=1}^N L_{1i}^r$ and $\bar{L}_2^r = N^{-1}\sum_{i=1}^N L_{2i}^r$.

Lemma 4 *Under the local alternative H_1^ℓ, as both N and T go to infinity simultaneously, (i) $E\|J_{0i}^r\|^2 \leq \frac{C}{\sqrt{N}}$, $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|J_{0i}^r\|^2 = O_p(1)$ and $\|\bar{J}_0^r\| = O_p(\frac{1}{N^{3/4}})$, (ii) $E\|L_{1i}^r\|^2 \leq \frac{C}{\sqrt{N}}$,*

$\frac{1}{\sqrt{N}}\sum_{i=1}^N \|L_{1i}^r\|^2 = O_p(1)$ *and* $\|\bar{L}_1\| = O_p(\frac{1}{N^{3/4}})$, *(iii)* $E\|L_{2i}^r\|^2 \leq \frac{C}{\sqrt{N}T}$, $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|L_{2i}^r\|^2 = O_p(\frac{1}{T})$ *and* $\|\bar{L}_2\| = O_p(\frac{1}{N^{3/4}\sqrt{T}})$, *(iv)* $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|J_{0i}^r\|\|L_{1i}^r\| = O_p(1)$, $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|J_{0i}^r\|\|L_{2i}^r\| = O_p(\frac{1}{\sqrt{T}})$ *and* $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|L_{1i}^r\|\|L_{2i}^r\| = O_p(\frac{1}{\sqrt{T}})$.

Proof of Lemma 4: (i) Since $\|J_{0i}^r\|^2 = \sum_{t=1}^T (T^{-1}\sum_{s=1}^t r_{is})^2$, we have $E\|J_{0i}^r\|^2 \leq C/\sqrt{N}$ using (12). We can also see that

$$E\|\bar{J}_0^r\|^2 = \frac{1}{T^2}\sum_{t=1}^T E\left[\left(\sum_{s=1}^t \bar{r}_s\right)^2\right] \leq \frac{C}{N\sqrt{N}}$$

using (15) and $\sigma_v^2 = c^2/(\sqrt{N}T^2)$. This implies $\|\bar{J}_0^r\| = O_p(1/N^{3/4})$.

(ii) is obtained using (12), (13), (15) and (16) by noting that

$$\begin{aligned} E\|L_{1i}^r\|^2 &= E\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T r_{it}\right)^2 + \left(\frac{1}{T\sqrt{T}}\sum_{t=1}^T t r_{it}\right)^2\right] \\ E\|\bar{L}_1^r\|^2 &= E\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \bar{r}_t\right)^2 + \left(\frac{1}{T\sqrt{T}}\sum_{t=1}^T t\bar{r}_t\right)^2\right]. \end{aligned}$$

(iii) From the definition of L_{2i}^r, we can see that

$$\begin{aligned} \|L_{2i}^r\|^2 \leq\ & C\left\{\left(\frac{1}{N\sqrt{T}}\sum_{t=1}^T r_{it}^2\right)^2 + \left(\frac{1}{N\sqrt{T}}\sum_{t=1}^T\sum_{j\neq i} r_{it}r_{jt}\right)^2\right. \\ & \left. + \left(\frac{\bar{\gamma}}{\sqrt{T}}\sum_{t=1}^T f_t r_{it}\right)^2 + \left(\frac{1}{\sqrt{T}}\sum_{t=1}^T \bar{\varepsilon}_t r_{it}\right)^2\right\}. \end{aligned} \tag{29}$$

Since $E[r_{is}^2 r_{it}^2] \leq 3\sigma_v^4 T^2$ from (20) and $\sigma_v^4 = c^2/(NT^4)$, it is observed that

$$\begin{aligned} E\left[\left(\frac{1}{N\sqrt{T}}\sum_{t=1}^T r_{it}^2\right)^2\right] &\leq \frac{1}{N^2T}\sum_{s=1}^T\sum_{t=1}^T E[r_{is}^2 r_{it}^2] \\ &\leq \frac{\sigma_v^4}{N^2T}\sum_{s=1}^T\sum_{t=1}^T 3T^2 \leq \frac{C}{N^3T}. \end{aligned} \tag{30}$$

Similarly, since $E[r_{is}r_{it}r_{js}r_{kt}] = 0$ for $j, k \neq i$ and $j \neq k$ and $E[r_{is}r_{it}] \leq \sigma_v^2 s$ from (11) we have

$$
\begin{aligned}
E\left[\left(\frac{1}{N\sqrt{T}}\sum_{t=1}^{T}\sum_{j\neq i} r_{it}r_{jt}\right)^2\right] &= \frac{1}{N^2T}\sum_{s=1}^{T}\sum_{t=1}^{T}\sum_{j\neq i} E[r_{is}r_{it}]E[r_{js}r_{jt}] \\
&\leq \frac{\sigma_v^4 C}{NT}\sum_{s=1}^{T}\sum_{t=1}^{T} s^2 \leq \frac{C}{N^2T}. \qquad (31)
\end{aligned}
$$

$$
E\left[\left(\frac{\bar{\gamma}}{\sqrt{T}}\sum_{t=1}^{T} f_t r_{it}\right)^2\right] = \frac{\bar{\gamma}^2}{T}\sum_{t=1}^{T} E\left[f_t^2\right]E\left[r_{it}^2\right] \leq \frac{C}{\sqrt{N}T}. \qquad (32)
$$

$$
E\left[\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T} \bar{\varepsilon}_t r_{it}\right)^2\right] = \frac{1}{T}\sum_{t=1}^{T} E\left[\bar{\varepsilon}_t^2\right]E\left[r_{it}^2\right] \leq \frac{C}{N\sqrt{N}T}. \qquad (33)
$$

From (29)–(33) the first and second assertions of (iii) hold.

Regarding to $\bar{L}_2^r$, we can see using (14) that

$$
\begin{aligned}
\bar{L}_2^r &= \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\bar{r}_t^2 + \frac{\bar{\gamma}}{\sqrt{T}}\sum_{t=1}^{T} f_t\bar{r}_t + \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\bar{\varepsilon}_t\bar{r}_t \\
&= O_p\left(\frac{1}{N\sqrt{NT}}\right) + O_p\left(\frac{1}{N^{3/4}\sqrt{T}}\right) + O_p\left(\frac{1}{N^{5/4}\sqrt{T}}\right) = O_p\left(\frac{1}{N^{3/4}\sqrt{T}}\right).
\end{aligned}
$$

(iv) is proved in exactly the same manner as Lemma 3 (vii).□

Using Lemmas 3 and 4, we can distinguish between the main term that weakly converges and the negligible term in the test statistic. Under the local alternative,

$$
\frac{1}{T}BM_{w^*}\mathbf{y}_i = \frac{1}{T}BM_{w^*}\boldsymbol{\varepsilon}_i - \frac{\tilde{\gamma}_i}{T}BM_{w^*}\bar{\boldsymbol{\varepsilon}} + \frac{1}{T}BM_{w^*}\mathbf{r}_i - \frac{\tilde{\gamma}_i}{T}BM_{w^*}\bar{\mathbf{r}}.
$$

Using Lemmas 3 and 4, it is shown that

$$
\left\|\frac{1}{T}BM_{w^*}\bar{\boldsymbol{\varepsilon}}\right\| = O_p\left(\frac{1}{\sqrt{N}}\right), \quad \left\|\frac{1}{T}BM_{w^*}\bar{\mathbf{r}}\right\| = O_p\left(\frac{1}{N^{3/4}}\right),
$$

and thus, since $|\tilde{\gamma}_i| \leq C$, we can see that

$$
\begin{aligned}
\left\|\frac{1}{\sqrt{N}T^2}\sum_{i=1}^{N}\tilde{\gamma}_i\boldsymbol{\varepsilon}_i' M_{w^*}B'BLM_{w^*}\bar{\mathbf{r}}\right\| &\leq C\sqrt{N}\left\|\frac{1}{T}BM_{w^*}\bar{\boldsymbol{\varepsilon}}\right\|\left\|\frac{1}{T}BM_{w^*}\bar{\mathbf{r}}\right\| = O_p\left(\frac{1}{N^{3/4}}\right), \\
\left\|\frac{1}{\sqrt{N}T^2}\sum_{i=1}^{N}\tilde{\gamma}_i\mathbf{r}_i' M_{w^*}B'BM_{w^*}\bar{\boldsymbol{\varepsilon}}\right\| &\leq C\sqrt{N}\left\|\frac{1}{T}BM_{w^*}\bar{\boldsymbol{\varepsilon}}\right\|\left\|\frac{1}{T}BM_{w^*}\bar{\mathbf{r}}\right\| = O_p\left(\frac{1}{N^{3/4}}\right).
\end{aligned}
$$

Therefore, the cross products between the terms related with $\boldsymbol{\varepsilon}_i$ and $\bar{\mathbf{r}}$, $\mathbf{r}_i$ and $\bar{\boldsymbol{\varepsilon}}$, and $\bar{\boldsymbol{\varepsilon}}$ and $\bar{\mathbf{r}}$ converge to zero in probability as both N and T go to infinity.

In addition, using expression (24) and Lemmas 3 and 4, it is observed that

$$
\begin{aligned}
&\frac{1}{T}BM_{w^*}\boldsymbol{\varepsilon}_i - (J_{0i} - J_1K_{11}^{-1}L_{1i}) \\
&\quad \leq \|J_1\|\|K^{11} - K_{11}^{-1}\|\|L_{1i}\| + \|J_2K^{21}L_{1i} + (J_1K^{12} + J_2K^{22})L_{2i}\| = O_p\left(\frac{1}{\sqrt{T}}\right), \\
&\left\|\frac{1}{T}BM_{w^*}\mathbf{r}_i - (J_{0i}^r - J_1K_{11}^{-1}L_{1i}^r)\right\| \\
&\quad \leq \|J_1\|\|K^{11} - K_{11}^{-1}\|\|L_{1i}^r\| + \|J_2K^{21}L_{1i}^r + (J_1K^{12} + J_2K^{22})L_{2i}^r\| = O_p\left(\frac{1}{N^{1/4}T}\right),
\end{aligned}
$$

which imply we have only to consider $(J_{0i} - J_1K_{11}^{-1}L_{1i})$ and $(J_{0i}^r - J_1K_{11}^{-1}L_{1i}^r)$ in the limit. Moreover, the cross product between these two terms is shown to be negligible by noting that

$$
\begin{aligned}
E\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} J_{0i}'J_{0i}^r\right)^2\right] &\leq \frac{C}{N}\sum_{i=1}^{N} E\|J_{0i}\|^2E\|J_{0i}^r\|^2 \leq \frac{C}{\sqrt{N}}, \\
E\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} J_{0i}'J_1K_{11}^{-1}L_{1i}^r\right)^2\right] &\leq \frac{C}{N}\sum_{i=1}^{N} E\|J_{0i}\|^2E\|L_{1i}^r\|^2 \leq \frac{C}{\sqrt{N}}, \\
E\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} J_{0i}^{r\prime}J_1K_{11}^{-1}L_{1i}\right)^2\right] &\leq \frac{C}{N}\sum_{i=1}^{N} E\|J_{0i}^r\|^2E\|L_{1i}\|^2 \leq \frac{C}{\sqrt{N}}, \\
E\left[\left(\frac{1}{\sqrt{N}}\sum_{i=1}^{N} L_{1i}'K_{11}^{-1}J_1'J_1K_{11}^{-1}L_{1i}^r\right)^2\right] &\leq \frac{C}{N}\sum_{i=1}^{N} E\|L_{1i}\|^2E\|L_{1i}^r\|^2 \leq \frac{C}{\sqrt{N}},
\end{aligned}
$$

where we used the fact that the deterministic term $\|J_1K_{11}^{-1}\|$ is bounded above by a constant and ε_{it} is independent of r_{it}.

Using these results, the augmented KPSS test statistic becomes

$$
\begin{aligned}
Z_A &= \frac{1}{\zeta\sqrt{N}}\sum_{i=1}^{N}\left\{\left(J_{0i} - J_1K_{11}^{-1}L_{1i}\right)'\left(J_{0i} - J_1K_{11}^{-1}L_{1i}\right) - \xi\right\} \\
&\quad + \frac{1}{\zeta\sqrt{N}}\sum_{i=1}^{N}\left(J_{0i}^r - J_1K_{11}^{-1}L_{1i}^r\right)'\left(J_{0i} - J_1K_{11}^{-1}L_{1i}^r\right) + o_p(1).
\end{aligned}
$$

The first term weakly converges to a standard normal distribution as proved in Theorem 1, whereas the second term can be expressed as

$$\frac{1}{\zeta N}\sum_{i=1}^{N}\sqrt{N}\left(J_{0i}^{r}-J_1K_{11}^{-1}L_{1i}^{r}\right)\prime\left(J_{0i}^{r}-J_1K_{11}^{-1}L_{1i}^{r}\right) \tag{34}$$

$$= \frac{1}{\zeta N}\sum_{i=1}^{N}\sqrt{N}\sum_{t=1}^{T}\left\{\frac{1}{T}\sum_{s=1}^{t}r_{is}-\frac{t(4T-3t-3)}{T^3}\sum_{t=1}^{T}r_{it}+\frac{6t(T-t-1)}{T^4}\sum_{t=1}^{T}tr_{it}\right\}^2,$$

where we used the fact that

$$\text{the } t\text{-th row of } J_1K_{11}^{-1}=\left[\frac{t(4T-3t-3)}{T^2\sqrt{T}},-\frac{6t(T-t-1)}{T^2\sqrt{T}}\right].$$

The joint probability limit of (34) is obtained by Lemma 2. Using Lemma 1 it can be shown that

$$\begin{aligned} E\left[\sqrt{N}\left(J_{0i}^{r}-J_1K_{11}^{-1}L_{1i}^{r}\right)'\left(J_{0i}^{r}-J_1K_{11}^{-1}L_{1i}^{r}\right)\right] &= \sigma_v^2\sqrt{N}\left(\frac{11}{12600}T^2+O(1)\right)\\ &= \frac{11}{12600}c^2+O\left(\frac{1}{T^2}\right), \end{aligned} \tag{35}$$

while the second moment is bounded above uniformly over N and T using (20). For example, since $t(4T-3t-3)/T^2\le 4$ for all t and T,

$$\begin{aligned} & E\left[\left\{\sqrt{N}\sum_{t=1}^{T}\left(\frac{1}{T}\sum_{s=1}^{t}r_{is}\right)\left(\frac{t(4T-3t-3)}{T^3}\sum_{t\prime=1}^{T}r_{it\prime}\right)\right\}^2\right]\\ \le\ & C\frac{N}{T^4}\sum_{t=1}^{T}\sum_{s=1}^{t}\sum_{t'=1}^{T}\sum_{u=1}^{T}\sum_{v=1}^{u}\sum_{u'=1}^{T}E[r_{is}r_{it'}r_{iv}r_{iu'}]\\ \le\ & C\frac{\sigma_v^4N}{T^4}\sum_{t=1}^{T}\sum_{s=1}^{t}\sum_{t'=1}^{T}\sum_{u=1}^{T}\sum_{v=1}^{u}\sum_{u'=1}^{T}3T^2\le C. \end{aligned}$$

On the other hand, since $N^{1/4}\sqrt{T}r_{[Tr]}\overset{T}{\Longrightarrow}cB_i^v(r)$, we can see using expression (34) that

$$\begin{aligned} & \sqrt{N}\left(J_{0i}^{r}-J_1K_{11}^{-1}L_{1i}^{r}\right)\prime\left(J_{0i}-J_1K_{11}^{-1}L_{1i}^{r}\right)\\ \overset{T}{\Longrightarrow}\ & c^2\int_0^1\left\{\int_0^r B_i^v(s)ds-r(4-3r)\int_0^1 B_i^v(s)ds+6r(1-r)\int_0^1 sB_i^v(s)ds\right\}^2dr\\ =\ & c^2\int_0^1 F_i^v(r)^2dr, \end{aligned}$$

whose moment is $11c^2/12600$ by direct calculation. By applying Lemma 2, we have

$$\frac{1}{\zeta\sqrt{N}}\sum_{i=1}^{N}\left(J_{0i}^{r}-J_{1}K_{11}^{-1}L_{1i}^{r}\right)'\left(J_{0i}^{r}-J_{1}K_{11}^{-1}L_{1i}^{r}\right)\overset{p\ (N,T)}{\longrightarrow}\frac{c^2}{\zeta}E\left[\int_0^1 F_i^v(r)^2dr\right]=\frac{11}{12600}\frac{c^2}{\zeta}.$$

When $z_t = 1$, the above probability limit can be shown to be $c^2/(90\zeta)$ in exactly the same manner.

Table 1. Size of the tests: serially correlated case

		constant case						trend case					
		SCC			WCC			SCC			WCC		
N	T	Z_A^{SPC}	Z_A^{LA}	HLM	Z_A^{SPC}	Z_A^{LA}	HLM	Z_A^{SPC}	Z_A^{LA}	HLM	Z_A^{SPC}	Z_A^{LA}	HLM
10	10	0.075	0.338	-	0.062	0.262	-	0.289	0.650	-	0.535	0.800	-
	20	0.004	0.068	0.033	0.004	0.069	0.038	0.001	0.029	0.027	0.003	0.039	0.034
	30	0.009	0.022	0.078	0.009	0.036	0.086	0.006	0.021	0.059	0.011	0.029	0.068
	50	0.040	0.062	0.086	0.018	0.046	0.079	0.030	0.050	0.056	0.014	0.034	0.056
	100	0.061	0.101	0.064	0.024	0.070	0.064	0.045	0.085	0.033	0.014	0.060	0.033
	200	0.109	0.124	0.058	0.058	0.076	0.054	0.120	0.135	0.051	0.053	0.073	0.053
20	10	0.081	0.425	-	0.080	0.338	-	0.437	0.859	-	0.759	0.922	-
	20	0.002	0.059	0.011	0.002	0.067	0.015	0.001	0.021	0.010	0.001	0.036	0.013
	30	0.002	0.008	0.043	0.004	0.033	0.058	0.001	0.008	0.023	0.006	0.022	0.031
	50	0.025	0.048	0.088	0.009	0.042	0.075	0.013	0.026	0.037	0.006	0.025	0.031
	100	0.041	0.087	0.074	0.013	0.072	0.072	0.026	0.059	0.024	0.006	0.044	0.022
	200	0.122	0.150	0.055	0.040	0.071	0.055	0.121	0.154	0.036	0.029	0.063	0.038
30	10	0.088	0.499	-	0.085	0.386	-	0.488	0.930	-	0.832	0.948	-
	20	0.001	0.068	0.003	0.001	0.063	0.006	0.001	0.020	0.003	0.002	0.039	0.007
	30	0.001	0.005	0.026	0.003	0.029	0.040	0.001	0.006	0.009	0.004	0.016	0.016
	50	0.020	0.045	0.076	0.007	0.041	0.064	0.010	0.027	0.019	0.006	0.023	0.014
	100	0.034	0.078	0.062	0.013	0.063	0.063	0.017	0.050	0.014	0.004	0.036	0.014
	200	0.131	0.176	0.058	0.027	0.071	0.059	0.145	0.187	0.028	0.020	0.059	0.031
50	10	0.089	0.635	-	0.103	0.444	-	0.603	0.984	-	0.917	0.980	-
	20	0.001	0.045	0.000	0.002	0.057	0.002	0.000	0.017	0.001	0.004	0.042	0.003
	30	0.000	0.001	0.013	0.001	0.029	0.020	0.000	0.002	0.003	0.002	0.012	0.005
	50	0.009	0.032	0.050	0.002	0.041	0.042	0.002	0.013	0.005	0.002	0.016	0.004
	100	0.030	0.076	0.059	0.006	0.049	0.060	0.017	0.040	0.010	0.004	0.029	0.009
	200	0.089	0.122	0.056	0.023	0.061	0.058	0.082	0.118	0.025	0.016	0.045	0.026
100	10	0.097	0.752	-	0.163	0.529	-	0.877	1.000	-	0.986	0.996	-
	20	0.000	0.040	0.000	0.001	0.054	0.000	0.000	0.014	0.000	0.007	0.059	0.000
	30	0.000	0.001	0.001	0.001	0.035	0.002	0.000	0.001	0.000	0.001	0.008	0.000
	50	0.003	0.018	0.018	0.001	0.032	0.013	0.001	0.003	0.000	0.001	0.007	0.000
	100	0.028	0.067	0.045	0.005	0.050	0.049	0.015	0.030	0.002	0.001	0.019	0.002
	200	0.084	0.124	0.049	0.016	0.049	0.056	0.078	0.114	0.013	0.009	0.032	0.014

Table 2. Power of the tests: serially correlated case

		constant case						trend case					
		SCC			WCC			SCC			WCC		
N	T	Z_A^{SPC}	Z_A^{LA}	HLM	Z_A^{SPC}	Z_A^{LA}	HLM	Z_A^{SPC}	Z_A^{LA}	HLM	Z_A^{SPC}	Z_A^{LA}	HLM
10	10	0.229	0.564	-	0.092	0.315	-	0.186	0.537	-	0.373	0.726	-
	20	0.323	0.282	0.033	0.162	0.269	0.040	0.000	0.044	0.004	0.000	0.070	0.006
	30	0.437	0.262	0.218	0.267	0.324	0.231	0.003	0.041	0.001	0.002	0.079	0.001
	50	0.695	0.373	0.740	0.454	0.461	0.739	0.039	0.086	0.000	0.023	0.155	0.001
	100	0.843	0.521	0.985	0.669	0.631	0.984	0.374	0.207	0.113	0.260	0.357	0.113
	200	0.944	0.672	1.000	0.878	0.812	1.000	0.831	0.413	0.890	0.700	0.636	0.894
20	10	0.312	0.748	-	0.123	0.392	-	0.222	0.705	-	0.508	0.857	-
	20	0.511	0.445	0.004	0.205	0.336	0.008	0.000	0.035	0.000	0.000	0.068	0.001
	30	0.609	0.407	0.194	0.336	0.413	0.200	0.000	0.036	0.000	0.001	0.085	0.000
	50	0.862	0.587	0.894	0.503	0.541	0.890	0.042	0.100	0.000	0.021	0.187	0.000
	100	0.944	0.748	1.000	0.714	0.722	1.000	0.606	0.297	0.083	0.344	0.464	0.084
	200	0.993	0.861	1.000	0.930	0.903	1.000	0.965	0.606	0.987	0.790	0.777	0.986
30	10	0.367	0.814	-	0.146	0.420	-	0.251	0.817	-	0.586	0.899	-
	20	0.608	0.574	0.001	0.231	0.376	0.002	0.000	0.030	0.000	0.000	0.064	0.000
	30	0.659	0.512	0.151	0.369	0.450	0.160	0.000	0.029	0.000	0.000	0.080	0.000
	50	0.898	0.716	0.949	0.525	0.578	0.950	0.046	0.114	0.000	0.019	0.195	0.000
	100	0.962	0.844	1.000	0.728	0.762	1.000	0.710	0.390	0.063	0.394	0.518	0.060
	200	0.996	0.921	1.000	0.947	0.930	1.000	0.981	0.765	0.998	0.831	0.844	0.999
50	10	0.449	0.928	-	0.167	0.452	-	0.288	0.919	-	0.703	0.945	-
	20	0.807	0.739	0.000	0.267	0.410	0.000	0.000	0.020	0.000	0.000	0.072	0.000
	30	0.762	0.612	0.114	0.399	0.490	0.126	0.000	0.018	0.000	0.000	0.088	0.000
	50	0.977	0.872	0.989	0.546	0.610	0.989	0.051	0.107	0.000	0.019	0.222	0.000
	100	0.995	0.943	1.000	0.752	0.788	1.000	0.895	0.469	0.028	0.456	0.582	0.028
	200	1.000	0.980	1.000	0.967	0.962	1.000	0.999	0.874	1.000	0.858	0.893	1.000
100	10	0.556	0.980	-	0.197	0.490	-	0.393	0.987	-	0.837	0.978	-
	20	0.872	0.870	0.000	0.294	0.440	0.000	0.000	0.010	0.000	0.000	0.074	0.000
	30	0.816	0.739	0.055	0.423	0.526	0.065	0.000	0.012	0.000	0.000	0.092	0.000
	50	0.986	0.947	1.000	0.575	0.646	1.000	0.050	0.118	0.000	0.019	0.254	0.000
	100	0.999	0.977	1.000	0.775	0.825	1.000	0.953	0.659	0.008	0.514	0.650	0.007
	200	1.000	0.993	1.000	0.981	0.979	1.000	1.000	0.973	1.000	0.887	0.938	1.000