

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

Kaddour Hadri¹
Queen's University Management School
Queen's University

Eiji Kurozumi²
Department of Economics
Hitotsubashi University

February 11, 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. We also develop a Lagrange multiplier (LM) test allowing for cross-sectional dependence and, under restrictive assumptions, compare our augmented KPSS test with the extended LM test under the null of stationarity, under the local alternative and under the fixed alternative, and discuss the differences between these two tests. The latter results are derived using the joint asymptotic limits where T and $N \rightarrow \infty$ simultaneously (under the additional condition that $N/T \rightarrow 0$) and the limiting distributions under the null are shown to be standard normal. 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 augmented KPSS test.

JEL classification: C12, C33

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

¹Corresponding 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

²Kurozumi'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 (2005)). 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. to de Silva, Hadri and Tremayne (2007) 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 (2005) 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 $T \rightarrow \infty$ followed by $N \rightarrow \infty$, and joint limits where $T, N \rightarrow \infty$ 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 $(T(N), N) \rightarrow \infty$ 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.

In this paper, we adapt Pesaran (2007) approach to the panel stationarity test of Hadri (2000) due to its conceptual simplicity. Our test is basically the same as the Kwiatkowski et al. (1992) test (KPSS test hereafter) with the regression augmented by cross-sectional average of the observations. We show that the limiting null distribution is the same as the test suggested by Hadri (2000), which is an Lagrange multiplier (LM) test without cross-sectional dependence. We also extend Hadri's test and develop the LM test allowing for cross-sectional dependence. We compare our augmented KPSS test with the extended LM test under the null of stationarity, under the local alternative and under the fixed alternative, and discuss the difference between these two tests. All these theoretical results 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 test to the case of the serially correlated shocks, and use Monte Carlo simulations to examine the finite sample properties of the augmented KPSS test.

The paper is organized as follows. Section 2 sets up the model and assumptions, and define the augmented test statistic. We also develop the LM test allowing for cross-sectional dependence. Section 3 is devoted to the comparison of our augmented KPSS test under restrictive assumptions with the extended LM test under the null of stationarity, under the local alternative and under the fixed alternative. We show that the limiting null distribution of the augmented KPSS test is the same as that of Hadri's (2000) test. We also examine whether our theoretical result is valid in finite samples via simple Monte Carlo simulations. In Section 4, we relax Assumption 1 in order to allow for serial correlation in the error term and propose a modification of the 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 $\xrightarrow{p(N,T)}$ and $\xrightarrow{(N,T)}$ mean joint convergence in probability and joint

weak convergence respectively when both T and N go to infinity simultaneously, while \xrightarrow{T} or \xrightarrow{N} signifies weak convergence when only T or N go to infinity.

2. Model and Test Statistics

2.1. Model and assumptions

Let us consider the following model:

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

for $i = 1, \dots, N$ and $t = 1, \dots, T$ 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 $z_t = z_t^T = [1, t]'$. 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 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

$$\begin{aligned} \mathbf{y}_i &= \mathbf{Z} \delta_i + \mathbf{r}_i + \mathbf{f} \gamma_i + \boldsymbol{\varepsilon}_i \\ &= \mathbf{Z} \delta_i + \mathbf{L} \mathbf{v}_i + \mathbf{f} \gamma_i + \boldsymbol{\varepsilon}_i, \end{aligned} \quad (2)$$

where $\mathbf{Z} = [\boldsymbol{\tau}, \mathbf{d}]$ with $\boldsymbol{\tau} = [1, 1, \dots, 1]'$ and $\mathbf{d} = [1, 2, \dots, T]'$ being $T \times 1$ vectors. Further,

we have

$$\begin{bmatrix} \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \\ \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} Z & & & \\ & Z & & \\ & & \ddots & \\ & & & Z \end{bmatrix} \begin{bmatrix} \delta_1 \\ \delta_2 \\ \vdots \\ \delta_N \end{bmatrix} + \begin{bmatrix} L & & & \\ & L & & \\ & & \ddots & \\ & & & L \end{bmatrix} \begin{bmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_N \end{bmatrix} + \begin{bmatrix} \mathbf{f}\gamma_1 \\ \mathbf{f}\gamma_2 \\ \vdots \\ \mathbf{f}\gamma_N \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_N \end{bmatrix}$$

or

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

In this section, we assume the following simple assumption:

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

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

and σ_ε^2 , σ_f^2 and σ_v^2 are 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 $\bar{\gamma} = N^{-1} \sum_{i=1}^N \gamma_i$.*

(iii) *$N/T \rightarrow 0$ as $T, N \rightarrow \infty$ simultaneously .*

Assumption 1(i) is restrictive and not practical. The assumption of normality with homoskedasticity is required to derive the LM test and to discuss the optimal property of the tests. The variances of the innovations are assumed to be known in order to make the theoretical investigation as simple as possible. In fact, if the variances are unknown, we need to estimate them and the asymptotic property of the tests depends on those estimators. Our purpose in this section is to examine the theoretical effect of ‘‘augmentation’’, which is explained below, on stationarity tests. Assumption 1(i) will be relaxed in Section 4 to a more practical one. 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). Assumption 1(iii) is often imposed when we consider the joint convergence where T and N go to infinity simultaneously.

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

$$H_0 : \rho \equiv \frac{\sigma_v^2}{\sigma_\varepsilon^2} = 0 \quad \forall i \quad \text{v.s.} \quad H_1 : \rho \geq 0 \quad (4)$$

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 some or all of the cross-sectional units have a unit root under the alternative.

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 and 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 exists 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:

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

where $\bar{y}_t = N^{-1} \sum_{i=1}^N y_{it}$, $\bar{\delta}_t = N^{-1} \sum_{i=1}^N \delta_i$, $\bar{r}_t = N^{-1} \sum_{i=1}^N r_{it}$, $\bar{\gamma}_t = N^{-1} \sum_{i=1}^N \gamma_i$ 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:

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

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

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

where $\tilde{\delta}_i = \delta_i - \tilde{\gamma}_i \bar{\delta}$, $\tilde{\gamma}_i = \gamma_i / \tilde{\gamma}$ 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,

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

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

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

Note that ST_i can also be expressed as

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

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

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 it has the limiting distribution. We call Z_A the augmented KPSS test statistic.

2.3. An LM test for panel stationarity

Although the augmented KPSS test is easy to implement, we do not know whether it has an optimal property. Note that if \bar{y}_t were deterministic, we could see that Z_A is equivalent to the LM test statistic and, because the LM test is a locally best invariant test under the

assumption of normality, the augmented KPSS test would also be locally optimal. One might think that this local optimality holds because \bar{y}_t converges in probability to its mean, $z_t\bar{\delta} + f_t\bar{\gamma}$, as $N \rightarrow \infty$ and f_t is independent of ε_{it} . In order to investigate the local optimality of the augmented KPSS test, we consider the LM test and compare the two tests.

Under the assumption of normality, the log-likelihood function of \mathbf{y} , denoted by ℓ , is expressed as

$$\ell = \text{const} - \frac{1}{2} \log |\Omega| - \frac{1}{2} \{\mathbf{y} - (I_N \otimes Z)\delta\}'^{-1} \{\mathbf{y} - (I_N \otimes Z)\delta\},$$

$$\text{where } \Omega \equiv \text{Var}(\mathbf{y}) = \rho (\sigma_\varepsilon^2 I_N \otimes LL') + A \otimes I_T \quad \text{with } A = \sigma_f^2 \gamma \gamma' + \sigma_\varepsilon^2 I_N.$$

The partial derivative of ℓ with respect to ρ is given by

$$\frac{\partial \ell}{\partial \rho} = \text{const} + \frac{1}{2} \{\mathbf{y} - (I_N \otimes Z)\delta\}'^{-1} \frac{\partial \Omega}{\partial \rho} \Omega^{-1} \{\mathbf{y} - (I_N \otimes Z)\delta\}. \quad (8)$$

Noting that

$$\Omega|_{H_0} = A \otimes I_T, \quad \text{and} \quad \left. \frac{\partial \Omega}{\partial \rho} \right|_{H_0} = \sigma_\varepsilon^2 I_N \otimes LL', \quad (9)$$

the maximum likelihood (ML) estimator of δ under H_0 is given by

$$\begin{aligned} \hat{\delta} &= \left[(I_N \otimes Z') \Omega^{-1} \Big|_{H_0} (I_N \otimes Z) \right]^{-1} (I_N \otimes Z') \Omega^{-1} \Big|_{H_0} \mathbf{y} \\ &= (A^{-1} \otimes Z'Z)^{-1} (A^{-1} \otimes Z') \mathbf{y} \\ &= [I_N \otimes (Z'Z)^{-1} Z'] \mathbf{y}. \end{aligned} \quad (10)$$

Thus, the maximum likelihood estimator of δ under H_0 is the same as the OLS estimator. By evaluating (8) under the null hypothesis using (9) and (10), the LM test statistic is given by

$$\begin{aligned} \overline{LM} &= \frac{1}{NT^2} \{\mathbf{y} - (I_N \otimes Z)\hat{\delta}\}' (A^{-1} \otimes I_T) (\sigma_\varepsilon^2 I_N \otimes LL') (A^{-1} \otimes I_T) \{\mathbf{y} - (I_N \otimes Z)\hat{\delta}\} \\ &= \frac{1}{NT^2} \mathbf{y}' (\sigma_\varepsilon^2 A^{-2} \otimes M_z LL' M_z) \mathbf{y}. \end{aligned}$$

By normalizing \overline{LM} , the extended LM test statistic for panel stationarity is given by

$$Z_{LM} = \frac{\sqrt{N}(\overline{LM} - \xi)}{\zeta} \quad (11)$$

where ξ and ζ are the same as in Z_A . Note that \overline{LM} can also be expressed as

$$\overline{LM} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \sigma_\varepsilon^2 a_2^{ij} LM_{ij} \quad \text{with} \quad LM_{ij} = \frac{1}{T^2} \sum_{t=1}^T S_{it}^z S_{jt}^z \quad \text{and} \quad S_{it}^z = \sum_{s=1}^t \hat{u}_{is},$$

where a_2^{ij} being the (i, j) element of A^{-2} and \hat{u}_{it} for each i is obtained by regressing y_{it} on z_t .

3. Theoretical Property of the Augmented KPSS and LM Tests

3.1. The limiting distributions of the test statistics

In this section we compare the augmented KPSS test with the extended LM test. Note that the LM test is known to be a locally best invariant test under Assumption 1. Because there is no one-to-one transformation between Z_A and Z_{LM} , we can see that the augmented KPSS test does not have the local optimality in finite samples. Then, our interest moves on to whether it is asymptotically locally optimal or not.

In order to investigate the asymptotic local optimality of the augmented KPSS test, we compare it with the LM test statistic under the null hypothesis, under the local alternative and under the fixed alternative. We first give the limiting distributions of the two test statistics under the null hypothesis.

Theorem 1 *Assume that Assumption 1 holds. Under H_0 , as N and T go to infinity simultaneously, the augmented KPSS and LM test statistics have a limiting standard normal distribution for both cases of $z_t = 1$ and $z_t = [1, t]'$,*

$$Z_A, Z_{LM} \xrightarrow{(N, T)} N(0, 1).$$

Proof. see Appendix.

Note that the rejection region of Z_A and Z_{LM} is the right hand tail as in Hadri's (2000) test.

Theorem 1 shows that Pesaran's (2007) method works well in order to eliminate cross-sectional dependence even for testing the null hypothesis of stationarity and the augmented KPSS test is asymptotically equivalent to the LM test statistic under the null hypothesis.

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

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

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 $\rho = c^2/(\sqrt{NT}^2)$.

Theorem 2 *Assume that Assumption 1 holds. Under H_1^ℓ , as N and T go to infinity simultaneously, the augmented KPSS and LM test statistics have the same limiting distribution given by*

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

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

Proof see Appendix.

This result implies that both the augmented KPSS and extended LM test statistics have the same asymptotic local distribution. Since the LM test is locally best invariant, we can see that the augmented KPSS test has the same asymptotic local optimality.

We can also deduce from Theorem 2 that both tests are more powerful when only a constant is included in the regression than the trending case as is the univariate KPSS test, because $1/90 > 11/12600$.

We finally investigate the asymptotic property of the test statistics under the fixed alternative H_1 . The following theorem gives the difference of powers between the two tests when the alternative is not local but far away from $\rho = 0$.

Theorem 3 *Assume that Assumption 1 holds. Under H_1 , as N and T go to infinity simultaneously,*

$$\begin{aligned} \frac{1}{\sqrt{NT^2}} Z_A &\xrightarrow{(N,T)} \frac{\rho}{\zeta} E_{vi} \left[\int_0^1 G_i^v(r)^2 dr \right] \\ \frac{1}{\sqrt{NT^2}} Z_{LM} &\xrightarrow{p(N,T)} \frac{\rho}{\zeta} E \left[\int_0^1 F_i^{v*}(r)^2 dr \right], \end{aligned}$$

where $G_i^v(r) = \int_0^r B_i^v(s) ds - \int_0^r z_2'(s) ds (\int_0^1 z_2(s) z_2'(s) ds)^{-1} \int_0^1 z_2(s) B_i^v(s) ds$ with $z_2(r) = [z'(r), \underline{B}^v(r)]'$ and $\underline{B}^v(r)$ is a standard Brownian motion independent of $B_i^v(r)$ and E_{vi} denotes the expectation operator with respect to $B_i^v(r)$.

Proof. see Appendix.

Note that since $G_i^v(r)$ depends on $B_i^v(r)$ and $\underline{B}^v(r)$ and they are independent, we can see that $E_{vi}[\int_0^1 G_i^v(r)^2 dr]$ still depends on $\underline{B}^v(\cdot)$ and then it is stochastic, while $E[\int_0^1 F_i^{v*}(r)^2 dr]$ is deterministic. This is an interesting result because, when the asymptotic local powers are the same for two tests, it is often the case that they also have the same limiting distribution under the fixed alternative. For panel stationarity tests, the two tests have the same local asymptotic power from Theorem 2 but the powers are different under the fixed alternative from Theorem 3. This implies that although the two tests are locally optimal, they are not equivalent in a wide range under the alternative.

3.2. Finite sample property under restrictive assumptions

In this subsection we investigate how accurately does the asymptotic theory approximate the finite sample behavior of the augmented KPSS and LM tests. We consider the following data generating process for finite sample simulations:

$$y_{it} = z_i' \delta_i + r_{it} + f_t \gamma_i + \varepsilon_{it}, \quad f_t \sim i.i.d.N(0, 1), \quad \varepsilon_{it} \sim i.i.d.N(0, 1)$$

$$r_{it} = r_{it-1} + v_{it}, \quad v_{it} \sim i.i.d.N(0, \rho), \quad \begin{cases} H_0 : \rho = 0 \\ H_1 : \rho = 0.0001, 0.001, 0.01 \end{cases}$$

where $\delta_i = \alpha_i$ 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. We consider all the pairs of $N = 10, 20, 30, 50, 100$ and $T = 10, 20, 30, 50, 100, 200$. The level of significance is 0.05 and the number of replications is 10,000 in all experiments.

Table 1 shows the sizes of the tests. We can observe that the empirical size of the augmented KPSS test is close to the nominal one when T is equal to or greater than 50 for SCC case while it is slightly undersized for WCC case. On the other hand, the size of the LM test is close to the nominal one irrespective of N and T but it is slightly undersized for SCC case while it is slightly oversized for WCC case. Overall, the null distributions of the two tests seem to be well approximated by a standard normal distribution suggested by Theorem 1 in view of the size of the tests.

Table 2 reports the powers of the tests. For given N and T , the upper, middle and lower entries are the powers of the tests for $\rho = 0.0001, 0.001$ and 0.01 , respectively. From the table the powers of the tests become higher for larger ρ and T , although the tests have low power when T is small. We can also observe that the powers become higher for larger N . For example, the size of the augmented KPSS test for $T = 50$, SCC and constant case is relatively close to 0.05 for all the values of N while the empirical power when $\rho = 0.001$ is 0.145, 0.202, 0.254, 0.342 and 0.539 for $N = 10, 20, 30, 50$ and 100, respectively. Table 2 implies that the tests are consistent as proved by Theorem 3.

In order to see if the augmented KPSS test can be seen as the asymptotically locally best test indicated by Theorem 2, we calculated the size adjusted power of the tests. Figure 1 draws the power curves for selected cases. From the figure we observe that the power of the augmented KPSS test is almost the same as that of the LM test for the constant case. When a linear trend is included, the augmented KPSS test is as powerful as the LM test when ρ is small while the former is slightly less powerful than the latter for the trend case.

As a whole, the finite sample behavior of the augmented KPSS and LM tests is well approximated by the asymptotic theory established in the previous section when N and T are of moderate size.

4. Extension to general case

4.1. Modification of the augmented KPSS test

So far, we have investigated the theoretical property of the augmented KPSS test under restrictive assumptions. In this section we relax Assumptions 1(i) and consider a more practical situation. Because the LM test statistic will depend on nuisance parameters in a complicated way under general assumptions and it would be difficult to correct Z_{LM} so that it becomes free of nuisance parameters, we concentrate on the modification of the augmented KPSS test.

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:

$$y_{it} = z_t' \delta_i + f_t \gamma_i + \varepsilon_{it}, \quad \varepsilon_{it} = \phi_{i1} \varepsilon_{it-1} + \cdots + \phi_{ip} \varepsilon_{it-p} + \nu_{it}. \quad (12)$$

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} are independent of f_t and i.i.d. $(0, \sigma_{\nu_i}^2)$ across i and t with a finite fourth moments.*

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

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

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

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

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 $w_t = [z_t', \bar{y}_t, \bar{y}_{t-1}, \dots, \bar{y}_{t-p}]$ because ε_{it} are AR(p) processes and construct S_{it}^w using this regression residuals. In this case it is not difficult to see that the numerator of each ST_i weakly converges to

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

where $\sigma_i^2 = \sigma_{\varepsilon_i}^2 / (1 - \phi_{i1} - \dots - \phi_{ip})^2$ 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 for parametric model have been proposed in the literature for a 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 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

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

and construct the estimator of the long-run variance by

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

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

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

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:

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

and construct the test statistic using

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

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

The consistency of $\hat{\sigma}_{iSPC}^2$ and $\hat{\sigma}_{iLA}^2$ 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 fixed 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 behavior of these tests.

4.2. Finite sample property under general assumptions

In this section we conduct Monte Carlo simulations to investigate the finite sample properties of the 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 $f_t \sim i.i.d.(0, 1)$, $\nu_{it} \sim i.i.d.N(0, 1)$, f_t and ν_{it} are independent of each other, α_i , β_i and γ_i are set as in Subsection 3.2., 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 $\varepsilon_{1t}, \varepsilon_{2t}, \dots, \varepsilon_{Nt}$. 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 2 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 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 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 4 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 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 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 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 augmented KPSS test is the same as the original Hadri's test that is the LM test without cross-sectional dependence. We also extended Hadri's test by developing an LM test correcting for cross-sectional dependence. Then, we compared our augmented KPSS test with the extended LM test. We found that the augmented KPSS test is asymptotically locally optimal but it is not asymptotically equivalent to the LM test under the fixed alternative. All the theoretical results are derived via the joint asymptotic theory and the limiting distributions under the null are shown to be standard normal. The Monte Carlo simulations indicated that we should carefully use the panel stationarity tests because they are undersized in some cases but suffer

from over rejection in other cases.

References

- [1] Bai, J. and S. Ng (2004). A panic attack on unit roots and cointegration. *Econometrica* 72, 1127-1177.
- [2] 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.
- [3] Baltagi, B. H. (2001). *Econometric Analysis of Panel Data* . Chichester, Wiley.
- [4] Baltagi, B. H. and C. Kao (2000). Nonstationary panels, cointegration in panels and dynamic panels: A survey. *Advances in Econometrics* 15, 7-51.
- [5] Banerjee, A. (1999). Panel data unit roots and cointegration: An overview. *Oxford Bulletin of Economics and Statistics Special issue*, 607-29.
- [6] Banerjee, A., M. Marcellino and C. Osbat (2001). Testing for PPP: Should we use panel methods? *Empirical Economics* 30, 77-91.
- [7] 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.
- [8] 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.
- [9] 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.
- [10] Chang, Y. (2004). Bootstrap unit root tests in panels with cross-sectional dependency. *Journal of Econometrics* 110, 261-292.

- [11] Choi, I. (1993). Asymptotic normality of the least-squares estimates for higher order autoregressive integrated processes with some applications. *Econometric Theory* 9, 263-282.
- [12] Choi, I. (2001). Unit root tests for panel data. *Journal of International Money and Banking* 20, 249-272.
- [13] 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.
- [14] Chung, K. L. (1974). *A Course in Probability Theory (2nd ed.)*, Academic Press, San Diego.
- [15] 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.
- [16] de Silva, S., K. Hadri and A. R. Tremayne (2007). Panel unit root tests in the presence of cross-sectional dependence: Finite sample performance and an application. Unpublished manuscript, University of Sydney.
- [17] Hadri, K. (2000). Testing for stationarity in heterogeneous panel data. *Econometrics Journal* 3, 148-161.
- [18] Hadri, K. and Larsson, R. (2005). Testing for stationarity in heterogeneous panel data where the time dimension is finite. *Econometrics Journal* 8, 55-69.
- [19] Hadri K. and Y. Rao, (2008) "Panel stationarity test with breaks", *Oxford Bulletin of Economics and Statistics*, 70(2), 245-269.
- [20] Harris, D., B. McCabe and S. Leybourne (2003). Some limit theory for autocovariances whose order depends on sample size. *Econometric Theory* 19, 829-864.

- [21] 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.
- [22] Im, K., M. H. Pesaran and Y. Shin (2003). Testing for unit roots in heterogeneous panels. *Journal of Econometrics* 115, 53-74.
- [23] Kurozumi, E. (2002). Testing for stationarity with a break. *Journal of Econometrics*, Vol. 108, pp. 63-69.
- [24] 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.
- [25] 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.
- [26] 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.
- [27] Moon, R. and B. Perron (2004). Testing for unit roots in panels with dynamic factors. *Journal of Econometrics* 122, 81-126.
- [28] O'Connell, P. G. J. (1998). The overvaluation of purchasing power parity. *Journal of International Economics* 44, 1-19.
- [29] Oh, K. Y. (1996). Purchasing power parity and unit root tests using panel data. *Journal of International Money and Finance* 15, 405-418.
- [30] Papell, D. H. (1997). Searching for stationarity: Purchasing power parity under the current float. *Journal of International Economics* 43, 313-332.
- [31] Papell, D. H. (2002). The great appreciation, the great depreciation and the purchasing power parity hypothesis. *Journal of International Economics* 57, 51-82.

- [32] Pesaran, M. H. (2007). A simple panel unit root test in the presence of cross-section dependence. *Journal of Applied Econometrics* 22, 265-312.
- [33] Phillips, P. C. B. and H. R. Moon (1999). Linear regression limit theory for nonstationary panel data. *Econometrica* 67, 1057-1111.
- [34] Phillips, P. C. B. and D. Sul (2003). Dynamic panel estimation and homogeneity testing under cross section dependence. *Econometrics Journal* 6, 217-25.
- [35] Quah, D. (1994). Exploiting cross section variation for unit root inference in dynamic data. *Economics Letters* 44, 9-19.
- [36] Shin, Y. and A. Snell (2006). Mean group tests for stationarity in heterogeneous panels. *Econometrics Journal* 9, 123-158.
- [37] 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.
- [38] Toda, H. Y., and T. Yamamoto (1995). Statistical inference in vector autoregressions with possibly integrated processes. *Journal of Econometrics* 66, 225-250.
- [39] Wu, S. and J. L. Wu (2001). Is purchasing power parity overvalued? *Journal of Money, Credit and Banking* 33, 804-812.
- [40] 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 , \dots . 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

$$\bar{\mathbf{y}} = Z\bar{\delta} + \bar{\mathbf{r}} + \mathbf{f}\bar{\gamma} + \bar{\boldsymbol{\varepsilon}}. \quad (13)$$

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

$$\mathbf{y}_i = Z(\delta_i - \tilde{\gamma}_i \bar{\delta}) + \tilde{\gamma}_i \bar{\mathbf{y}} + (\mathbf{r}_i - \tilde{\gamma}_i \bar{\mathbf{r}}) + (\boldsymbol{\varepsilon}_i - \tilde{\gamma}_i \bar{\boldsymbol{\varepsilon}}) \quad (14)$$

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

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

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 $i = 1, \dots, N$ and $t = 1, \dots, T$, $r_{it} = \sum_{s=1}^t v_{is}$ and $\bar{r}_t = N^{-1} \sum_{i=1}^N r_{it}$. Then,

$$E[r_{is}r_{it}] = \sigma_v^2 \min(s, t) \quad (15)$$

$$E \left[\left(\sum_{s=1}^t r_{is} \right)^2 \right] = \frac{\sigma_v^2}{6} t(t+1)(2t+1) \quad (16)$$

$$E \left[\left(\sum_{s=1}^t s r_{is} \right)^2 \right] = \frac{\sigma_v^2}{30} t(t+1)(2t+1)(2t^2 + 2t + 1) \quad (17)$$

$$E[\bar{r}_s \bar{r}_t] = \frac{\sigma_v^2}{N} \min(s, t) \quad (18)$$

$$E \left[\left(\sum_{s=1}^t \bar{r}_s \right)^2 \right] = \frac{\sigma_v^2}{6N} t(t+1)(2t+1) \quad (19)$$

$$E \left[\left(\sum_{s=1}^t s \bar{r}_s \right)^2 \right] = \frac{\sigma_v^2}{30N} t(t+1)(2t+1)(2t^2 + 2t + 1), \quad (20)$$

$$E \left[\left(\sum_{s=1}^t r_{is} \right) \left(\sum_{t=1}^T r_{it} \right) \right] = \frac{\sigma_v^2}{6} t(t+1)(3T - t + 1) \quad (21)$$

$$E \left[\left(\sum_{s=1}^t r_{is} \right) \left(\sum_{t=1}^T t r_{it} \right) \right] = \frac{\sigma_v^2}{24} t(t+1)(6T^2 + 6T - t^2 - t + 2) \quad (22)$$

$$E \left[\left(\sum_{t=1}^T r_{it} \right) \left(\sum_{t=1}^T t r_{it} \right) \right] = \frac{\sigma_v^2}{24} T(T+1)(5T^2 + 5T + 2), \quad (23)$$

$$E[r_{is}r_{it}r_{iu}r_{iv}] = \sigma_v^4(2st + su) \quad \text{for } s \leq t \leq u \leq v. \quad (24)$$

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 $T \rightarrow \infty$, 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 ($i = 1, \dots, N$) 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 $S_{iT} \xrightarrow{T} S_{i\infty}$ as $T \rightarrow \infty$.

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

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

(ii) If (a) $N^{-1/2} \sum_{i=1}^N S_{i\infty} \xrightarrow{N} S$ as $N \rightarrow \infty$, (b) S_{iT} does not depend on N and $\sup_T E[S_{iT}^4] \leq \infty$ or $E[S_{iT}^2] \rightarrow \mu_2 \equiv E[S_{i\infty}^2] < \infty$ as $T \rightarrow \infty$, (c) $\sup_T E[S_{iT}^2] < \infty$ and $E[S_{i\infty}^4] < \infty$, then,

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N S_{iT} \xrightarrow{(N,T)} S.$$

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

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

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

(ii) Since S_{iT} and $S_{i\infty}$ are i.i.d. sequences, we have, for any arbitrary $\varepsilon > 0$,

$$P \left(\left| \frac{1}{\sqrt{N}} \sum_{i=1}^N (S_{iT} - S_{i\infty}) \right| > \varepsilon \right) \leq \frac{1}{\varepsilon^2} (E[S_{iT}^2] + E[S_{i\infty}^2] - 2E[S_{iT}S_{i\infty}]). \quad (25)$$

If $\sup_T E[S_{iT}^4] \leq \infty$, we can replace the limit and the expectation by Theorem 4.5.2 of Chung (1974) and hence $\lim_T E[S_{iT}^2] = E[S_{i\infty}^2]$ under condition (b). On the other hand, by Hölder's inequality, we have, for any arbitrary $0 < \delta < 1$,

$$E \left[|S_{iT}S_{i\infty}|^{1+\delta} \right] \leq (E[S_{iT}^2])^{(1+\delta)/2} \left(E \left[|S_{i\infty}|^{2(1+\delta)/(1-\delta)} \right] \right)^{(1-\delta)/2}.$$

The right hand side of the above inequality with $\delta = 1/3$ is bounded above uniformly over T because of condition (c). This implies that $\sup_T E[|S_{iT}S_{i\infty}|^{4/3}] < \infty$ and gain, we can replace the limit and the expectation, so that $\lim_T E[S_{iT}S_{i\infty}] = E[S_{i\infty}^2]$. As a result, the right hand side of (25) goes to zero as N and T go to infinity simultaneously because both S_{iT} and $S_{i\infty}$ depend only on T . Combining this result with condition (a), we obtain (ii). \square

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

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

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

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

it is sufficient to prove that

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N (ST_{1i} - ST_{1i}^0) \xrightarrow{p(N,T)} 0, \quad (26)$$

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N ST_{2i} \xrightarrow{p(N,T)} 0, \quad (27)$$

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N ST_{3i} \xrightarrow{p(N,T)} 0. \quad (28)$$

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

$$\begin{aligned}
\frac{1}{T} L M_{w^*} \boldsymbol{\varepsilon}_i &= \frac{1}{T} L \boldsymbol{\varepsilon}_i - \frac{1}{T} L W^* D^{-1} (D^{-1} W^* W^* D^{-1})^{-1} D^{-1} W^* \boldsymbol{\varepsilon}_i \\
&= (J_{0i} - J_1 K^{11} L_{1i}) - \{J_2 K^{21} L_{1i} + (J_1 K^{12} + J_2 K^{22}) L_{2i}\}. \quad (29)
\end{aligned}$$

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

$$\frac{1}{T} L M_{w^*} \bar{\boldsymbol{\varepsilon}}_i = \bar{J}_0 - (J_1 K^{11} + J_2 K^{21}) \bar{L}_1 - (J_1 K^{12} + J_2 K^{22}) \bar{L}_2. \quad (30)$$

We first prove (26). Using expression (29), ST_{1i} can be decomposed into

$$\begin{aligned}
ST_{1i} &= (J_{0i} - J_1 K^{11} L_{1i})' (J_{0i} - J_1 K^{11} L_{1i}) \\
&\quad - 2(J_{0i} - J_1 K^{11} L_{1i})' \{J_2 K^{21} L_{1i} + (J_1 K^{12} + J_2 K^{22}) L_{2i}\} \\
&\quad + \{J_2 K^{21} L_{1i} + (J_1 K^{12} + J_2 K^{22}) L_{2i}\}' \{J_2 K^{21} L_{1i} + (J_1 K^{12} + J_2 K^{22}) L_{2i}\} \\
&= ST_{1i}^a + ST_{1i}^b + ST_{1i}^c, \quad \text{say.} \tag{31}
\end{aligned}$$

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) $E\|J_{0i}\|^2 \leq C$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|J_{0i}\|^2 = O_p(\sqrt{N})$ and $\|\bar{J}_0\| = O_p(\frac{1}{\sqrt{N}})$, (ii) $\|J_1\| = O(1)$, (iii) $E\|J_2\|^2 \leq \frac{C}{T}$ 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) $E\|L_{1i}\|^2 \leq C$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|L_{1i}\|^2 = O_p(\sqrt{N})$ and $\|\bar{L}_1\| = O_p(\frac{1}{\sqrt{N}})$, (vi) $E\|L_{2i}\|^2 \leq C$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|L_{2i}\|^2 = O_p(\sqrt{N})$ 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

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

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),

$$\begin{aligned}
E\|J_2\|^2 &= \frac{1}{T^3} E \left[\sum_{t=1}^T \left\{ \sum_{s=1}^t (\bar{\gamma}f_s + \bar{\varepsilon}_s) \right\}^2 \right] \\
&= \frac{1}{T^3} \sum_{t=1}^T \left\{ \bar{\gamma}^2 E \left[\left(\sum_{s=1}^t f_s \right)^2 \right] + E \left[\left(\sum_{s=1}^t \bar{\varepsilon}_s \right)^2 \right] \right\} \\
&= \frac{1}{T^3} \sum_{t=1}^T \left(\bar{\gamma}^2 \sigma_f^2 t + \frac{t}{N} \right) \leq \frac{C}{T}
\end{aligned} \tag{32}$$

and thus we obtain (iii).

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

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

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 $K^{12} = K^{21'}$ 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 Assumption 1 (iii), we have $\|\bar{L}_2\| = O_p(\sqrt{T}/N)$.

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

Since $ST_{1i}^0 = (J_{0i} - J_1 K_{11}^{-1} L_{1i})'(J_{0i} - J_1 K_{11}^{-1} L_{1i})$, we have

$$\begin{aligned} ST_{1i}^a - ST_{1i}^0 &= \{J_1(K^{11} - K_{11}^{-1})L_{1i}\}'\{J_1(K^{11} - K_{11}^{-1})L_{1i}\} \\ &\quad - 2(J_{0i} - J_1 K_{11}^{-1} L_{1i})'\{J_1(K^{11} - K_{11}^{-1})L_{1i}\}. \end{aligned} \quad (33)$$

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 (33) becomes $O_p(\sqrt{N}/T)$. Hence, the right hand side of (33) converges to 0 in probability when both N and T go to infinity because $N/T \rightarrow 0$ by Assumption 1(iii).

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 (26).

In order to prove (27) and (28), notice that

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

Then, it is sufficient to show that

$$\frac{\sqrt{N}}{T^2} \bar{\varepsilon}' M_{w^*} L' L M_{w^*} \bar{\varepsilon} \xrightarrow{p(N,T)} 0,$$

which can be proved by noting that

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

from expression (30) and Lemma 3. We thus obtain the result for the augmented KPSS test statistic.

In order to derive the limiting distribution of the LM test statistic, we first note that

under H_0 ,

$$\begin{aligned}
(I_N \otimes M_z)\mathbf{y} &= (\gamma \otimes M_z \mathbf{f}) + (I_N \otimes M_z)\boldsymbol{\varepsilon} \\
&\sim N(0, (\sigma_f^2 \gamma \gamma' \otimes M_z) + (\sigma_\varepsilon^2 I_N \otimes M_z)) \\
&= N(0, A \otimes M_z) \\
&= (A^{1/2} \otimes M_z)\boldsymbol{\eta}
\end{aligned}$$

where $\boldsymbol{\eta} = [\boldsymbol{\eta}'_1, \dots, \boldsymbol{\eta}'_N]'$ $\sim N(0, I_N \otimes I_T)$. Then, \overline{LM} can be expressed as

$$\begin{aligned}
\overline{LM} &= \frac{1}{NT^2} \boldsymbol{\eta}' (A^{1/2} \otimes M_z) (A^{-2} \otimes LL') (A^{1/2} \otimes M_z) \boldsymbol{\eta} \\
&= \frac{1}{NT^2} \boldsymbol{\eta}' (I_N \otimes M_z L) (A^{-1} \otimes I_T) (I_N \otimes L' M_z) \boldsymbol{\eta}.
\end{aligned} \tag{34}$$

We first investigate the matrix A . Note that A^{-1} can be expressed as

$$A^{-1} = (\sigma_f^2 \gamma \gamma' + I_N)^{-1} = \left(I_N - \frac{1}{1 + \sigma_f^2 \gamma' \gamma} \sigma_f^2 \gamma \gamma' \right).$$

Since $\text{rk}(\gamma \gamma') = 1$ and $(\gamma \gamma') \gamma = (\gamma' \gamma) \gamma$, the $(N-1)$ eigenvalues of $\gamma \gamma'$ are 0 and the non-zero eigenvalue is $\gamma' \gamma$, for which the corresponding eigenvector is γ . Then, there exists an $N \times N$ orthonormal matrix P such that

$$P' P = P P' = I_N \quad \text{and} \quad P' \gamma \gamma' P = \text{diag}\{\gamma' \gamma, 0, \dots, 0\} \equiv \Lambda_\gamma.$$

This implies that

$$P' A^{-1} P = I_N - \frac{1}{1 + \sigma_f^2 \gamma' \gamma} \sigma_f^2 \Lambda_\gamma = \begin{bmatrix} \frac{1}{1 + \sigma_f^2 \gamma' \gamma} & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \equiv \Lambda_A^{-1}. \tag{35}$$

By inserting (35) into (34), we obtain

$$\begin{aligned}
\overline{LM} &= \frac{1}{NT^2} \boldsymbol{\eta}' (I_N \otimes M_z L) (P P' A^{-1} P P' \otimes I_T) (I_N \otimes L' M_z) \boldsymbol{\eta} \\
&= \frac{1}{NT^2} \boldsymbol{\eta}' (I_N \otimes M_z L) (\Lambda_A^{-1} \otimes I_T) (I_N \otimes L' M_z) \boldsymbol{\eta}^* \\
&= \frac{1}{1 + \sigma_f^2 \gamma' \gamma} \frac{1}{NT^2} \boldsymbol{\eta}'_1 M_z L L' M_z \boldsymbol{\eta}_1^* + \frac{1}{NT^2} \sum_{i=2}^N \boldsymbol{\eta}'_i M_z L L' M_z \boldsymbol{\eta}_i^*,
\end{aligned}$$

where $\boldsymbol{\eta}^* = [\boldsymbol{\eta}_1^{*'}, \dots, \boldsymbol{\eta}_N^{*'}]' = (P \otimes I_T)\boldsymbol{\eta} \sim N(0, I_N \otimes I_T)$. Note that the first term converges to zero in probability as both N and T go to infinity, while the second term has the same structure as ST_{1i}^0 . We then obtain the result for the LM 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 $\sigma_v^2 = \rho\sigma_\varepsilon^2 = c^2/(\sqrt{NT}^2)$, (iii) is proved by noting that

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

To show (iv), note that

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

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 $E[(T^{-1} \sum_{t=1}^T \bar{r}_t)^2] \leq C/(N\sqrt{NT})$ and $E[(T^{-2} \sum_{t=1}^T t\bar{r}_t)^2] \leq C/(N\sqrt{NT})$ from (19) and (20). 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 $E[T^{-1} \sum_{t=1}^T \bar{r}_t^2] \leq C/(N\sqrt{NT})$ from (18).

(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 $E[(T^{-1/2} \sum_{t=1}^T \bar{r}_t \varepsilon_{it})^2] \leq C/(N\sqrt{NT})$, which is proved using (18) and the independence between \bar{r}_t and ε_{it} .

Next, we give a lemma similar to Lemma 3. Let $J_{0i}^r = T^{-1}L\mathbf{r}_i$, $[L_{1i}^r, L_{2i}^r]' = D^{-1}W^{*'}\mathbf{r}_i = [(D_\tau^{-1}Z'\mathbf{r}_i)', (T^{-1/2}\bar{\mathbf{y}}^{*'}\mathbf{r}_i)']'$, $\bar{J}_0^r = N^{-1}\sum_{i=1}^N J_{0i}^r$, $\bar{L}_1^r = N^{-1}\sum_{i=1}^N L_{1i}^r$ 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 $\|\bar{L}_1^r\| = O_p(\frac{1}{N^{3/4}})$, (iii) $E\|L_{2i}^r\|^2 \leq \frac{C}{\sqrt{NT}}$, $\frac{1}{\sqrt{N}}\sum_{i=1}^N \|L_{2i}^r\|^2 = O_p(\frac{1}{T})$ and $\|\bar{L}_2^r\| = O_p(\frac{1}{N^{3/4}\sqrt{T}})$, (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 (16). 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 (19) 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 (16), (17), (19) and (20) by noting that

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

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

$$\begin{aligned} \|L_{2i}^r\|^2 &\leq C \left\{ \left(\frac{1}{N\sqrt{T}} \sum_{t=1}^T r_{it}^2 \right)^2 + \left(\frac{1}{N\sqrt{T}} \sum_{t=1}^T \sum_{j \neq i} r_{it} r_{jt} \right)^2 \right. \\ &\quad \left. + \left(\frac{\tilde{\gamma}}{\sqrt{T}} \sum_{t=1}^T f_t r_{it} \right)^2 + \left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \bar{\varepsilon}_t r_{it} \right)^2 \right\}. \end{aligned} \quad (36)$$

Since $E[r_{is}^2 r_{it}^2] \leq 3\sigma_v^4 T^2$ from (24) and $\sigma_v^4 = c^2/(NT^4)$, it is observed that

$$\begin{aligned} E \left[\left(\frac{1}{N\sqrt{T}} \sum_{t=1}^T r_{it}^2 \right)^2 \right] &\leq \frac{1}{N^2 T} \sum_{s=1}^T \sum_{t=1}^T E[r_{is}^2 r_{it}^2] \\ &\leq \frac{\sigma_v^4}{N^2 T} \sum_{s=1}^T \sum_{t=1}^T 3T^2 \leq \frac{C}{N^3 T}. \end{aligned} \quad (37)$$

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 (15) we have

$$\begin{aligned} E \left[\left(\frac{1}{N\sqrt{T}} \sum_{t=1}^T \sum_{j \neq i} r_{it} r_{jt} \right)^2 \right] &= \frac{1}{N^2 T} \sum_{s=1}^T \sum_{t=1}^T \sum_{j \neq i} E[r_{is} r_{it}] E[r_{js} r_{jt}] \\ &\leq \frac{\sigma_v^4 C}{NT} \sum_{s=1}^T \sum_{t=1}^T s^2 \leq \frac{C}{N^2 T}. \end{aligned} \quad (38)$$

$$E \left[\left(\frac{\bar{\gamma}}{\sqrt{T}} \sum_{t=1}^T f_t r_{it} \right)^2 \right] = \frac{\bar{\gamma}^2}{T} \sum_{t=1}^T E[f_t^2] E[r_{it}^2] \leq \frac{C}{\sqrt{NT}}. \quad (39)$$

$$E \left[\left(\frac{1}{\sqrt{T}} \sum_{t=1}^T \bar{\varepsilon}_t r_{it} \right)^2 \right] = \frac{1}{T} \sum_{t=1}^T E[\bar{\varepsilon}_t^2] E[r_{it}^2] \leq \frac{C}{N\sqrt{NT}}. \quad (40)$$

From (36)–(40) the first and second assertions of (iii) hold.

Regarding to \bar{L}_2^r , we can see using (18) 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). \square

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,

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

Using Lemmas 3 and 4, it is shown that

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

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

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

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 (29) and Lemmas 3 and 4, it is observed that

$$\begin{aligned} &\left\| \frac{1}{T} LM_{w^*} \boldsymbol{\varepsilon}_i - (J_{0i} - J_1 K_{11}^{-1} L_{1i}) \right\| \\ &\leq \|J_1\| \|K^{11} - K_{11}^{-1}\| \|L_{1i}\| + \|J_2 K^{21} L_{1i} + (J_1 K^{12} + J_2 K^{22}) L_{2i}\| = O_p \left(\frac{1}{\sqrt{T}} \right), \\ &\left\| \frac{1}{T} LM_{w^*} \mathbf{r}_i - (J_{0i}^r - J_1 K_{11}^{-1} L_{1i}^r) \right\| \\ &\leq \|J_1\| \|K^{11} - K_{11}^{-1}\| \|L_{1i}^r\| + \|J_2 K^{21} L_{1i}^r + (J_1 K^{12} + J_2 K^{22}) L_{2i}^r\| = O_p \left(\frac{1}{N^{1/4} T} \right), \end{aligned}$$

which imply we have only to consider $(J_{0i} - J_1 K_{11}^{-1} L_{1i})$ and $(J_{0i}^r - J_1 K_{11}^{-1} L_{1i}^r)$ in the limit. Moreover, the cross product between these two terms is shown to be negligible by noting that

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

where we used the fact that the deterministic term $\|J_1 K_{11}^{-1}\|$ is bounded above by a constant and ε_{it} is independent of r_{it} .

Using these results, the augmented KPSS test statistic becomes

$$\begin{aligned} Z_A &= \frac{1}{\zeta\sqrt{N}} \sum_{i=1}^N \left\{ (J_{0i} - J_1 K_{11}^{-1} L_{1i})' (J_{0i} - J_1 K_{11}^{-1} L_{1i}) - \xi \right\} \\ &\quad + \frac{1}{\zeta\sqrt{N}} \sum_{i=1}^N (J_{0i}^r - J_1 K_{11}^{-1} L_{1i}^r)' (J_{0i} - J_1 K_{11}^{-1} L_{1i}) + o_p(1). \end{aligned}$$

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

$$\begin{aligned} &\frac{1}{\zeta N} \sum_{i=1}^N \sqrt{N} (J_{0i}^r - J_1 K_{11}^{-1} L_{1i}^r)' (J_{0i} - J_1 K_{11}^{-1} L_{1i}) \tag{41} \\ &= \frac{1}{\zeta N} \sum_{i=1}^N \sqrt{N} \sum_{t=1}^T \left\{ \frac{1}{T} \sum_{s=1}^t r_{is} - \frac{t(4T-3t-3)}{T^3} \sum_{t=1}^T r_{it} + \frac{6t(T-t-1)}{T^4} \sum_{t=1}^T t r_{it} \right\}^2, \end{aligned}$$

where we used the fact that

$$\text{the } t\text{-th row of } J_1 K_{11}^{-1} = \left[\frac{t(4T-3t-3)}{T^2\sqrt{T}}, -\frac{6t(T-t-1)}{T^2\sqrt{T}} \right].$$

The joint probability limit of (41) is obtained by Lemma 2 (i). Using Lemma 1 it can be shown that

$$\begin{aligned} E \left[\sqrt{N} (J_{0i}^r - J_1 K_{11}^{-1} L_{1i}^r)' (J_{0i} - J_1 K_{11}^{-1} L_{1i}) \right] &= \sigma_v^2 \sqrt{N} \left(\frac{11}{12600} T^2 + O(1) \right) \\ &= \frac{11}{12600} c^2 + O\left(\frac{1}{T^2}\right), \tag{42} \end{aligned}$$

while the second moment is bounded above uniformly over N and T using (24). For example, since $t(4T-3t-3)/T^2 \leq 4$ for all t and T ,

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

On the other hand, since $N^{1/4}\sqrt{T}r_{[Tr]} \xrightarrow{T} cB_i^v(r)$, we can see using expression (41) that

$$\begin{aligned} & \sqrt{N} (J_{0i}^r - J_1 K_{11}^{-1} L_{1i}^r)' (J_{0i} - J_1 K_{11}^{-1} L_{1i}^r) \\ & \xrightarrow{T} c^2 \int_0^1 \left\{ \int_0^r B_i^v(s) ds - r(4-3r) \int_0^1 B_i^v(s) ds + 6r(1-r) \int_0^1 s B_i^v(s) ds \right\}^2 dr \\ & = c^2 \int_0^1 F_i^v(r)^2 dr, \end{aligned}$$

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

$$\frac{1}{\zeta\sqrt{N}} \sum_{i=1}^N (J_{0i}^r - J_1 K_{11}^{-1} L_{1i}^r)' (J_{0i} - J_1 K_{11}^{-1} L_{1i}^r) \xrightarrow{p(N,T)} \frac{c^2}{\zeta} E \left[\int_0^1 F_i^v(r)^2 dr \right] = \frac{11}{12600} \frac{c^2}{\zeta}.$$

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

In order to derive the limiting distribution of the LM test statistic, note that

$$(I_N \otimes M_z)\mathbf{y} = (I_N \otimes M_z)\mathbf{r} + (A^{1/2} \otimes M_z)\boldsymbol{\eta}.$$

Then, the denominator of the LM test statistic can be expressed as

$$\begin{aligned} \sqrt{N}(\overline{LM} - \xi) &= \frac{1}{\sqrt{NT^2}} \{ \boldsymbol{\eta}'(A^{-2} \otimes M_z LL' M_z)\boldsymbol{\eta} - \xi \} \\ &+ \frac{1}{\sqrt{NT^2}} \mathbf{r}'(A^{-2} \otimes M_z LL' M_z)\mathbf{r} + \frac{2}{\sqrt{NT^2}} \boldsymbol{\eta}'(A^{-2} \otimes M_z LL' M_z)\boldsymbol{\varepsilon}. \end{aligned} \quad (43)$$

The first term on the right hand side of (43) converges in distribution to a standard normal distribution as proved in Theorem 1.

Since $A^{-2} = P\Lambda_A^{-2}P'$, the second term on the right hand side of (43) is expressed as

$$\frac{1}{\sqrt{NT^2}} \mathbf{r}'(A^{-2} \otimes M_z LL' M_z)\mathbf{r} = \frac{1}{\sqrt{NT^2}} \mathbf{r}'(P \otimes I_T)(\Lambda_A^{-2} \otimes M_z LL' M_z)(P \otimes I_T)\mathbf{r}.$$

Note that

$$(P \otimes I_T)\mathbf{r} = (P \otimes I_T)(I_N \otimes L)\mathbf{v} = (I_N \otimes L)\mathbf{v}^* = \mathbf{r}^*$$

where $\mathbf{v}^* \equiv (P \otimes I_T)\mathbf{v} \sim N(0, \sigma_v^2(I_N \otimes I_T))$ and $\mathbf{r}^* \equiv (I_N \otimes L)\mathbf{v}^*$. Using this expression, we

have

$$\begin{aligned}
& \frac{1}{\sqrt{NT^2}} \mathbf{r}'(A^{-2} \otimes M_z LL' M_z) \mathbf{r} \\
&= \frac{1}{\sqrt{NT^2}} \mathbf{r}^{*'} (\Lambda_A^{-2} \otimes M_z LL' M_z) \mathbf{r}^* \\
&= \frac{1}{(1 + \sigma_f^2 \gamma' \gamma)} \frac{1}{\sqrt{NT^2}} \mathbf{r}_1^{*'} M_z LL' M_z \mathbf{r}_1^* + \frac{1}{\sqrt{NT^2}} \sum_{i=2}^N \mathbf{r}_i^{*'} M_z LL' M_z \mathbf{r}_i^*.
\end{aligned}$$

It is not difficult to see that the first term on the right hand side converges to zero as both N and T go to infinity, while

$$\frac{1}{\sqrt{NT^2}} \sum_{i=2}^N \mathbf{r}_i^{*'} M_z LL' M_z \mathbf{r}_i^* = \frac{1}{\sqrt{N}} \sum_{i=2}^N (J_{0i}^{r^*} - J_1 K_{11}^{-1} L_{1i}^{r^*})' (J_{0i}^{r^*} - J_1 K_{11}^{-1} L_{1i}^{r^*}), \quad (44)$$

where $J_{0i}^{r^*}$ and $L_{1i}^{r^*}$ are defined in the same way as J_{0i}^r and L_{1i}^r with r_t replaced by r_t^* . Since r_{it}^* has the same distribution as r_{it} , (44) converges in probability to $11c^2/12600$ as proved for the case of Z_A .

Similarly, we can see that

$$\begin{aligned}
& \frac{1}{\sqrt{NT^2}} \boldsymbol{\eta}'(A^{-2} \otimes M_z LL' M_z) \mathbf{r} \\
&= \frac{1}{(1 + \sigma_f^2 \gamma' \gamma)} \frac{1}{\sqrt{NT^2}} \boldsymbol{\eta}_1^{*'} M_z LL' M_z \mathbf{r}_1^* + \frac{1}{\sqrt{NT^2}} \sum_{i=2}^N \boldsymbol{\eta}_i^{*'} M_z LL' M_z \mathbf{r}_i^*,
\end{aligned}$$

which converges to zero in probability as proved for the case of Z_A . ■

Proof of Theorem 3

We first summarize the orders of the various terms in Lemmas 3 and 4 under the fixed alternative. Since these orders are obtained in exactly the same manner as before, we omit the proof.

Lemma 5 *Under the fixed alternative H_1 , as both N and T go to infinity simultaneously, (i) $E\|J_{0i}\|^2 \leq C$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|J_{0i}\|^2 = O_p(\sqrt{N})$ and $\|\bar{J}_0\| = O_p(\frac{1}{\sqrt{N}})$, (ii) $\|J_1\| = O(1)$, (iii) $E\|J_2\|^2 \leq C\frac{T}{N}$ and $\|J_2\| = O_p(\frac{\sqrt{T}}{\sqrt{N}})$, (iv) $K_{11} = O(1)$, $K_{12} = K'_{21} =$*

$O_p(\frac{\sqrt{T}}{\sqrt{N}})$ and $K_{22} = O_p(\frac{T}{N})$, (v) $E\|L_{1i}\|^2 \leq C$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|L_{1i}\|^2 = O_p(\sqrt{N})$ and $\|\bar{L}_1\| = O_p(\frac{1}{\sqrt{N}})$, (vi) $E\|L_{2i}\|^2 \leq C\frac{T}{N}$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|L_{2i}\|^2 = O_p(\frac{T}{\sqrt{N}})$ and $\bar{L}_2 = O_p(\frac{\sqrt{T}}{\sqrt{N}})$, (vii) $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|J_{0i}\| \|L_{1i}\| = O_p(\sqrt{N})$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|J_{0i}\| \|L_{2i}\| = O_p(\sqrt{T})$ and $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|L_{1i}\| \|L_{2i}\| = O_p(\sqrt{T})$.

Lemma 6 Under the fixed alternative H_1 , as both N and T go to infinity simultaneously, (i) $E\|J_{0i}^r\|^2 \leq CT^2$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|J_{0i}^r\|^2 = O_p(\sqrt{N}T^2)$ and $\|\bar{J}_0^r\| = O_p(\frac{T}{\sqrt{N}})$, (ii) $E\|L_{1i}^r\|^2 \leq CT^2$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|L_{1i}^r\|^2 = O_p(\sqrt{N}T^2)$ and $\|\bar{L}_1\| = O_p(\frac{T}{\sqrt{N}})$, (iii) $E\|L_{2i}^r\|^2 \leq C\frac{T^3}{N}$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|L_{2i}^r\|^2 = O_p(\frac{T^3}{\sqrt{N}})$ and $\|\bar{L}_2\| = O_p(\frac{T\sqrt{T}}{\sqrt{N}})$, (iv) $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|J_{0i}^r\| \|L_{1i}^r\| = O_p(\sqrt{N}T^2)$, $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|J_{0i}^r\| \|L_{2i}^r\| = O_p(T^2\sqrt{T})$ and $\frac{1}{\sqrt{N}} \sum_{i=1}^N \|L_{1i}^r\| \|L_{2i}^r\| = O_p(T^2\sqrt{T})$.

Using Lemmas 5 and 6 it is shown that

$$\left\| \frac{1}{T^2} LM_{w^*} \boldsymbol{\varepsilon}_i \right\| = O_p\left(\frac{1}{T}\right), \quad \left\| \frac{1}{T^2} LM_{w^*} \bar{\boldsymbol{\varepsilon}} \right\| = O_p\left(\frac{1}{\sqrt{NT}}\right), \quad \left\| \frac{1}{T^2} LM_{w^*} \bar{\mathbf{r}} \right\| = O_p\left(\frac{1}{\sqrt{N}}\right),$$

which implies that

$$\frac{1}{\sqrt{NT^2}} Z_A = \frac{1}{\zeta} \frac{1}{N} \sum_{i=1}^N \frac{1}{T^4} \mathbf{r}'_i M_{w^*} L' LM_{w^*} \mathbf{r}_i + o_p(1).$$

We decompose the first term on the right hand side such that

$$\frac{1}{T^2} LM_{w^*} \mathbf{r}_i = \frac{1}{T} J_{0i}^r - \begin{bmatrix} J_1 & \frac{\sqrt{N}}{\sqrt{T}} J_2 \end{bmatrix} \begin{bmatrix} K_{11} & \frac{\sqrt{N}}{\sqrt{T}} K_{12} \\ \frac{\sqrt{N}}{\sqrt{T}} K_{21} & \frac{N}{T} K_{22} \end{bmatrix}^{-1} \begin{bmatrix} \frac{1}{T} L_{1i}^r \\ \frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^r \end{bmatrix}.$$

Using this expression and letting $\underline{K} = D_2^{-1} K D_2^{-1}$ where $D_2 = \text{diag}\{1, \sqrt{T}/\sqrt{N}\}$, we have

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \frac{1}{T^4} \mathbf{r}'_i M_{w^*} L' LM_{w^*} \mathbf{r}_i \\ &= \frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} J_{0i}^r J_{0i}^r - 2tr \left(\underline{K}^{-1} \frac{1}{N} \sum_{i=1}^N \begin{bmatrix} \frac{1}{T} L_{1i}^r \\ \frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^r \end{bmatrix} \begin{bmatrix} \frac{1}{T} J_{0i}^r J_1 & \frac{\sqrt{N}}{T\sqrt{T}} J_{0i}^r J_2 \end{bmatrix} \right) \\ &+ tr \left(\underline{K}^{-1} \begin{bmatrix} J_1' J_1 & \frac{\sqrt{N}}{\sqrt{T}} J_1' J_2 \\ \frac{\sqrt{N}}{\sqrt{T}} J_2' J_1 & \frac{N}{T} J_2' J_2 \end{bmatrix} \underline{K}^{-1} \frac{1}{N} \sum_{i=1}^N \begin{bmatrix} \frac{1}{T} L_{1i}^r \\ \frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^r \end{bmatrix} \begin{bmatrix} \frac{1}{T} L_{1i}^r & \frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^r \end{bmatrix} \right). \end{aligned} \quad (45)$$

Using (16) the expectation of the summand of the first term of (45) becomes

$$E \left[\frac{1}{T^2} J_{0i}^{r'} J_{0i}^r \right] = E \left[\frac{1}{T^4} \sum_{t=1}^T \left(\sum_{s=1}^t r_{is} \right)^2 \right] = \frac{\sigma_v^2}{12} + O \left(\frac{1}{T} \right),$$

while its second moment is uniformly bounded because

$$\begin{aligned} E \left[\left(\frac{1}{T^2} J_{0i}^{r'} J_{0i}^r \right)^2 \right] &= \frac{1}{T^8} \sum_{t=1}^T \sum_{s=1}^t \sum_{s'=1}^t \sum_{u=1}^T \sum_{v=1}^u \sum_{v'=1}^u E[r_{is} r_{is'} r_{iv} r_{iv'}] \\ &\leq \frac{C}{T^8} \sum_{t=1}^T \sum_{s=1}^t \sum_{s'=1}^t \sum_{u=1}^T \sum_{v=1}^u \sum_{v'=1}^u 3T^2 \leq C. \end{aligned}$$

Because $T^{-2} J_{0i}^{r'} J_{0i}^r \xrightarrow{T} \sigma_v^2 \int_0^1 \left(\int_0^r B_i^v(s) ds \right)^2 dr$, whose expectation is $\sigma_v^2/12$, we can see by applying Lemma 2 (i) that

$$\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} J_{0i}^{r'} J_{0i}^r \xrightarrow{p(N,T)} \sigma_v^2 E_{vi} \left[\int_0^1 \left(\int_0^r B_i^v(s) ds \right)^2 dr \right]. \quad (46)$$

For the second term on the right hand side of (45), we first derive the weak limit of \underline{K} . It is easy to show that $J_1' J_1 \rightarrow \int_0^1 z(t) z(t)' dt$. For the (2,1) block of \underline{K} , note that

$$\frac{\sqrt{N}}{\sqrt{T}} K_{21} = \left[\frac{\sqrt{N}}{T\sqrt{T}} \sum_{t=1}^T \bar{y}_t^*, \frac{\sqrt{N}}{T^2\sqrt{T}} \sum_{t=1}^T t \bar{y}_t^* \right] = \frac{1}{\sqrt{N}} \sum_{i=1}^N \left[\frac{1}{T\sqrt{T}} \sum_{t=1}^T r_{it}, \frac{1}{T^2\sqrt{T}} \sum_{t=1}^T t r_{it} \right] + o_p(1).$$

The sequential limit of the first term on the right hand side becomes

$$\frac{1}{\sqrt{N}} \sum_{i=1}^N \left[\frac{1}{T\sqrt{T}} \sum_{t=1}^T r_{it}, \frac{1}{T^2\sqrt{T}} \sum_{t=1}^T t r_{it} \right] \xrightarrow{T} \frac{\sigma_v}{\sqrt{N}} \sum_{i=1}^N \left[\int_0^1 B_i^v(r) dr, \int_0^1 r B_i^v(r) dr \right] \quad (47)$$

$$\xrightarrow{N} \sigma_v \left[\int_0^1 \underline{B}^v(r) dr, \int_0^1 r \underline{B}^v(r) dr \right]. \quad (48)$$

It is not difficult to see that the second moments of both $T^{-3/2} \sum_t r_{it}$ and $T^{-5/2} \sum_t t r_{it}$ are bounded above uniformly over T and they have the finite fourth moments using Lemma 1, while their weak limits in (47) when $T \rightarrow \infty$ have the finite fourth moments. Therefore, we can apply Lemma 2 (ii) and the sequential limit (48) is the same as the joint limit.

For the (3,3) block of \underline{K} , we can see that

$$\frac{N}{T} K_{22} = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \frac{1}{T^2} \sum_{t=1}^T r_{it} r_{jt} + o_p(1).$$

Using Lemma 1 it is shown that $E[T^{-2} \sum_t r_{it} r_{jt}] = \sigma_v^2/2 + O(T^{-1})$ when $i = j$ and 0 otherwise while its second moment is bounded above. On the other hand,

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \frac{1}{T^2} \sum_{t=1}^T r_{it} r_{jt} &\xrightarrow{T} \frac{\sigma_v^2}{N} \sum_{i=1}^N \sum_{j=1}^N \int_0^1 B_i^v(r) B_j^v(r) dr \\ &\xrightarrow{N} \sigma_v^2 \int_0^1 \underline{B}^v(r)^2 dr. \end{aligned} \quad (49)$$

Since $E[\int_0^1 B_i^v(r) B_j^v(r) dr] = 1/2$ when $i = j$ and 0 otherwise, the sequential limit is the same as the joint limit if it is shown that

$$E \left[\left\{ \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \left(\frac{1}{T^2} \sum_{t=1}^T r_{it} r_{jt} - \sigma_v^2 \int_0^1 B_i^v(r) B_j^v(r) dr \right) \right\}^2 \right] \rightarrow 0. \quad (50)$$

We prove (50) in a similar manner as the proof of Lemma 2 (ii). Noting that $E[r_{it} r_{jt} r_{ks} r_{\ell s}]$ does not equal zero only when $i = j = k = \ell$ or the two pairs of i, j, k, ℓ are the same, we have, using (15) and (24),

$$\begin{aligned} &E \left[\left(\frac{1}{N} \sum_{i=1}^N \sum_{j=1}^N \frac{1}{T^2} \sum_{t=1}^T r_{it} r_{jt} \right)^2 \right] \\ &= \frac{1}{N^2 T^4} \sum_{t=1}^T \sum_{s=1}^T \left(\sum_{i=1}^N E[r_{it}^2 r_{is}^2] + 2 \sum_{i=2}^N \sum_{j=1}^{i-1} E[r_{it}^2] E[r_{js}^2] + 2 \times 2 \sum_{i=2}^N \sum_{j=1}^{i-1} E[r_{is} r_{it}] E[r_{js} r_{jt}] \right) \\ &= \frac{\sigma_v^4}{N^2 T^4} \sum_{t=1}^T \sum_{s=1}^T \left(O(NT^2) + 2 \sum_{i=2}^N \sum_{j=1}^{i-1} st + 4 \sum_{i=2}^N \sum_{j=1}^{i-1} \min(s, t)^2 \right) \\ &= \frac{7}{12} \sigma_v^4 + o(1). \end{aligned}$$

Similarly, we can show that $E \left[\left(N^{-1} \sum_{i=1}^N \sum_{j=1}^N \int_0^1 B_i^v(r) B_j^v(r) dr \right)^2 \right] = 7/12 + o(1)$. In addition, by noting that the eighth moment of $B_i(r)$ with $i = i_1, \dots, i_8$ does not equal zero only when at least four pairs among i_1, \dots, i_8 are the same, we can see that the fourth moment of the right hand side of (49) is bounded above uniformly over N . Thus, we can replace the limit and the expectation as in the proof of Lemma 2 (ii), so that (50) is established.

Combining the above results, we have

$$\underline{K} \xrightarrow{(N,T)} D_v \int_0^1 z_2(r) z_2(r)' dr D_v \quad (51)$$

where $z_2(r) = [z(r)', \underline{B}^v(r)]'$ and $D_v = \text{diag}\{I_2, \sigma_v\}$.

The derivation of the weak limit of the product of $[\frac{1}{T} L_{1i}^r, \frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^r]'$ and $[\frac{1}{T} J_{0i}^{r'} J_1, \frac{\sqrt{N}}{T\sqrt{T}} J_{0i}^{r'} J_2]$ proceeds similarly. For example, it can be shown by direct calculation that

$$\begin{aligned} & E \left[\left(\frac{1}{T} L_{1i}^r \right) \left(\frac{1}{T} J_{0i}^{r'} J_1 \right) \right] \\ &= E \left[\begin{array}{cc} \frac{1}{T^5} (\sum_{t=1}^T r_{it}) (\sum_{t'=1}^T t' \sum_{s=1}^{t'} r_{is}) & \frac{1}{T^6} (\sum_{t=1}^T r_{it}) \{ \sum_{t'=1}^T (\sum_{s=1}^{t'} s) (\sum_{s'=1}^{t'} r_{is'}) \} \\ \frac{1}{T^6} (\sum_{t=1}^T t r_{it}) (\sum_{t'=1}^T t' \sum_{s=1}^{t'} r_{is}) & \frac{1}{T^7} (\sum_{t=1}^T t r_{it}) \{ \sum_{t'=1}^T (\sum_{s=1}^{t'} s) (\sum_{s'=1}^{t'} r_{is'}) \} \end{array} \right] \\ &= \sigma_v^2 \left[\begin{array}{cc} \frac{11}{120} & \frac{13}{360} \\ \frac{1}{18} & \frac{1}{1680} \end{array} \right] + O\left(\frac{1}{T}\right), \end{aligned} \quad (52)$$

while the second moment of each element can be shown to be uniformly bounded. On the other hand, we can see that

$$\left(\frac{1}{T} L_{1i}^r \right) \left(\frac{1}{T} J_{0i}^{r'} J_1 \right) \xrightarrow{T} \sigma_v^2 \left(\int_0^1 z(r) B_i^v(r) dr \right) \left\{ \int_0^1 \left(\int_0^r z(s)' ds \right) \left(\int_0^r B_i^v(s) ds \right) dr \right\},$$

whose moment can be shown to equal (52) except for $O(1/T)$ term. Hence, by applying Lemma 2 (i), we obtain

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{T} L_{1i}^r \right) \left(\frac{1}{T} J_{0i}^{r'} J_1 \right) \\ & \xrightarrow{p(N,T)} \sigma_v^2 E_{v_i} \left[\left(\int_0^1 z(r) B_i^v(r) dr \right) \left\{ \int_0^1 \left(\int_0^r z(s)' ds \right) \left(\int_0^r B_i^v(s) ds \right) dr \right\} \right]. \end{aligned} \quad (53)$$

For $(\frac{1}{T} L_{1i}^r)(\frac{\sqrt{N}}{T\sqrt{T}} J_{0i}^{r'} J_2)$, the first element becomes

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{T\sqrt{T}} \sum_{t=1}^T r_{it} \right) \left(\frac{\sqrt{N}}{T^4} \sum_{t'=1}^T \sum_{s=1}^{t'} r_{is} \sum_{s'=1}^{t'} \bar{y}_{s'}^* \right) \\ &= \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{T\sqrt{T}} \sum_{t=1}^T r_{it} \right) \left(\frac{\sqrt{N}}{T^4} \sum_{t'=1}^T \sum_{s=1}^{t'} r_{is} \sum_{s'=1}^{t'} \bar{r}_{s'} \right) + o_p(1) \\ &= \frac{1}{\sqrt{N}} \sum_{j=1}^N \left(\frac{1}{T^2\sqrt{T}} \sum_{t=1}^T \sum_{s'=1}^{t'} r_{js'} \right) \left(\frac{1}{N} \sum_{i=1}^N \frac{1}{T^3} \sum_{s=1}^{t'} \sum_{t=1}^T r_{is} r_{it} \right) + o_p(1). \end{aligned} \quad (54)$$

Using Lemma 1 the expectation of the second parentheses on the right hand side of (54) becomes

$$E \left[\frac{1}{N} \sum_{i=1}^N \frac{1}{T^3} \sum_{s=1}^{t'} \sum_{t=1}^T r_{is} r_{it} \right] = \frac{\sigma_v^2 t'^2}{6 T^2} \left(3 - \frac{t'}{T} \right) + O \left(\frac{1}{T} \right)$$

and its variance is bounded above by C/T , while the second moment of the first parentheses on the right hand side of (54) is bounded above. This implies that we can replace the second parentheses by its expectation, so that

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{T\sqrt{T}} \sum_{t=1}^T r_{it} \right) \left(\frac{\sqrt{N}}{T^4} \sum_{t'=1}^T \sum_{s=1}^{t'} r_{is} \sum_{s'=1}^{t'} \bar{y}_{s'}^* \right) \\ &= \frac{1}{\sqrt{N}} \sum_{j=1}^N \left(\frac{1}{T^2\sqrt{T}} \sum_{t=1}^T \sum_{s'=1}^{t'} r_{js'} \right) \frac{\sigma_v^2 t'^2}{6 T^2} \left(3 - \frac{t'}{T} \right) + o_p(1) \\ &\stackrel{(N,T)}{\implies} \sigma_v^3 \int_0^1 \left(\int_0^r \underline{B}^v(s) ds \right) \frac{r^2}{6} (3-r) dr \\ &= \sigma_v^3 E_{vi} \left[\left(\int_0^1 B_i^v(r) dr \right) \left(\int_0^1 \int_0^r B_i^v(s) ds \int_0^r \underline{B}^v(s) ds dr \right) \right], \end{aligned} \quad (55)$$

where the joint convergence is obtained by applying Lemma 2 (ii) while the last equality can be seen to hold by direct calculation.

Similarly, the second element of $(\frac{1}{T} L_{1i}^r)(\frac{\sqrt{N}}{T\sqrt{T}} J_{0i}' J_2)$, $(\frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^{r'}) (\frac{1}{T} J_{0i}' J_1)$ and $(\frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^{r'}) (\frac{\sqrt{N}}{T\sqrt{T}} J_{0i}' J_2)$ can be shown to weakly converge such that

$$\begin{aligned} & \frac{1}{N} \sum_{i=1}^N \left(\frac{1}{T^2\sqrt{T}} \sum_{t=1}^T t r_{it} \right) \left(\frac{\sqrt{N}}{T^4} \sum_{t'=1}^T \sum_{s=1}^{t'} r_{is} \sum_{s'=1}^{t'} \bar{y}_{s'}^* \right) \\ &= \frac{1}{\sqrt{N}} \sum_{j=1}^N \frac{1}{T^2\sqrt{T}} \sum_{t'=1}^T \sum_{s'=1}^{t'} r_{js'} \left(\frac{1}{N} \sum_{i=1}^N \frac{1}{T^4} \sum_{s=1}^{t'} r_{is} \sum_{t=1}^T t r_{it} \right) + o_p(1) \\ &= \frac{1}{\sqrt{N}} \sum_{j=1}^N \frac{1}{T^2\sqrt{T}} \sum_{t'=1}^T \sum_{s'=1}^{t'} r_{js'} \frac{\sigma_v^2 t^2}{24 T^2} \left(6 - \frac{t^2}{T^2} \right) + o_p(1) \\ &\stackrel{(N,T)}{\implies} \frac{\sigma_v^3}{24} \int_0^1 r^2 (6-r^2) \int_0^r \underline{B}^v(s) ds dr \\ &= \sigma_v^3 E_{vi} \left[\left(\int_0^1 r B_i^v(r) dr \right) \left\{ \int_0^1 \left(\int_0^r \underline{B}^v(s) ds \right) \left(\int_0^r B_i^v(s) ds \right) dr \right\} \right], \end{aligned} \quad (56)$$

$$\begin{aligned}
& \left\{ \left(\frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^r \right) \left(\frac{1}{T} J_{0i}' J_1 \right) \right\}' \\
&= \left[\frac{\frac{1}{\sqrt{NT^5}\sqrt{T}} \sum_{i=1}^N \sum_{t=1}^T \bar{r}_t r_{it} \sum_{t'=1}^T t' \sum_{s=1}^{t'} r_{is}}{\frac{1}{\sqrt{NT^6}\sqrt{T}} \sum_{i=1}^N \sum_{t=1}^T \bar{r}_t r_{it} \sum_{t'=1}^T \sum_{s'=1}^{t'} s' \sum_{s=1}^{t'} r_{is}} \right] + o_p(1) \\
&= \left[\frac{\frac{1}{\sqrt{N}} \sum_{j=1}^N \frac{1}{T\sqrt{T}} \sum_{t=1}^T r_{jt} \left\{ \frac{1}{N} \sum_{i=1}^N \frac{\sigma_v^2}{T} \sum_{t'=1}^T \frac{t'}{T} \frac{1}{T} \sum_{s=1}^{t'} \min\left(\frac{t}{T}, \frac{s}{T}\right) \right\}}{\frac{1}{\sqrt{N}} \sum_{j=1}^N \frac{1}{T\sqrt{T}} \sum_{t=1}^T r_{jt} \left\{ \frac{1}{N} \sum_{i=1}^N \frac{\sigma_v^2}{T} \sum_{t'=1}^T \frac{t'^2}{2T^2} \frac{1}{T} \sum_{s=1}^{t'} \min\left(\frac{t}{T}, \frac{s}{T}\right) \right\}} \right] + o_p(1) \\
&\stackrel{(N,T)}{\implies} \sigma_v^3 \int_0^1 \underline{B}^v(r) \int_0^1 \left[\frac{t}{t^2} \right] \int_0^t \min(r, s) ds dt dr \\
&= \sigma_v^3 E_{vi} \left[\left(\int_0^1 \underline{B}^v(r) B_i^v(r) dr \right) \left(\int_0^1 \int_0^r z(s) ds \int_0^r B_i^v(s) ds dr \right) \right], \tag{57}
\end{aligned}$$

$$\begin{aligned}
& \left(\frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^r \right) \left(\frac{\sqrt{N}}{T\sqrt{T}} J_{0i}' J_2 \right) \\
&= \frac{1}{T^6} \sum_{i=1}^N \sum_{t=1}^T \bar{r}_t r_{it} \sum_{t'=1}^T \sum_{s'=1}^{t'} r_{is'} \sum_{s=1}^{t'} \bar{r}_s + o_p(1) \\
&= \frac{1}{N} \sum_{j=1}^N \sum_{j'=1}^N \frac{1}{T^4} \sum_{t=1}^T r_{jt} \sum_{t'=1}^T \sum_{s=1}^{t'} r_{j's} \left(\frac{1}{N} \sum_{i=1}^N \frac{1}{T^2} r_{it} \sum_{s'=1}^{t'} r_{is'} \right) + o_p(1) \\
&\stackrel{(N,T)}{\implies} \sigma_v^4 \int_0^1 \int_0^1 \underline{B}^v(r) \int_0^t \underline{B}^v(s) ds \int_0^t \min(r, s) ds dt dr \tag{58} \\
&= \sigma_v^4 E_{vi} \left[\left(\int_0^1 \underline{B}^v(r) B_i^v(r) dr \right) \left(\int_0^1 \int_0^r B_i^v(s) ds \int_0^r \underline{B}^v(s) ds dr \right) \right].
\end{aligned}$$

Combining (53) and (55)–(58) we can see that the second term on the right hand side of (45) weakly converges to

$$\begin{aligned}
& 2tr \left(\underline{K}^{-1} \frac{1}{N} \sum_{i=1}^N \left[\frac{1}{T} L_{1i}^r \frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^r \right] \left[\frac{1}{T} J_{0i}' J_1, \frac{\sqrt{N}}{T\sqrt{T}} J_{0i}' J_2 \right] \right) \tag{59} \\
&\stackrel{(N,T)}{\implies} 2\sigma_v^2 E_{vi} \left[\left(\int_0^1 z_2(r)' B_i^v(r) dr \right) \left(\int_0^1 z_2(r) z_2(r)' dr \right)^{-1} \int_0^1 \int_0^r z_s(s) ds \int_0^r B_i^v(s) ds dr \right].
\end{aligned}$$

Completely in the same way, the joint weak limit of the third term on the right hand

side of (45) is shown to be equivalent to the sequential limit:

$$\begin{aligned}
& tr \left(\underline{K}^{-1} \begin{bmatrix} J_1' J_1 & \frac{\sqrt{N}}{\sqrt{T}} J_1' J_2 \\ \frac{\sqrt{N}}{\sqrt{T}} J_2' J_1 & \frac{N}{T} J_2' J_2 \end{bmatrix} \underline{K}^{-1} \frac{1}{N} \sum_{i=1}^N \begin{bmatrix} \frac{1}{T} L_{1i}^r \\ \frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^r \end{bmatrix} \begin{bmatrix} \frac{1}{T} L_{1i}^{r'} & \frac{\sqrt{N}}{T\sqrt{T}} L_{2i}^{r'} \end{bmatrix} \right) \\
& \xrightarrow{(N,T)} \sigma_v^2 E_{vi} \left[\int_0^1 z_2(r)' B_i^v(r) dr \left(\int_0^1 z_2(r) z_2(r)' dr \right)^{-1} \right. \\
& \quad \left. \times \int_0^1 \int_0^r z_2(s) ds \int_0^r z_2(s)' ds dr \left(\int_0^1 z_2(r) z_2(r)' dr \right)^{-1} \int_0^1 z_2(r)' B_i^v(r) dr \right]. \quad (60)
\end{aligned}$$

From the expression (45) and the convergence results of (46), (59) and (60), we obtain the joint weak limit of the augmented KPSS test statistic because $\sigma_v^2 = \sigma_\varepsilon^2 \rho = \rho$.

For the LM test statistic, we can see using expression (43) and Lemmas 5 and 6 that

$$\begin{aligned}
\frac{1}{\sqrt{NT^2}} Z_{LM} &= \frac{1}{\zeta NT^4} \mathbf{r}' (A^{-2} \otimes M_z L L' M_z) \mathbf{r} + o_p(1) \\
&= \frac{1}{\zeta NT^2} \sum_{i=2}^N (J_{0i}^{r*} - J_1 K_{11}^{-1} L_{1i}^{r*})' (J_{0i}^{r*} - J_1 K_{11}^{-1} L_{1i}^{r*}) + o_p(1) \\
&\xrightarrow{p(N,T)} \frac{\sigma_v^2}{\zeta} E \left[\int_0^1 F_i^{v*}(r)^2 dr \right]. \blacksquare
\end{aligned}$$

Table 1. Size of the tests: base case

N	T	constant case				trend case			
		SCC		WCC		SCC		WCC	
		Z_A	Z_{LM}	Z_A	Z_{LM}	Z_A	Z_{LM}	Z_A	Z_{LM}
10	10	0.023	0.038	0.019	0.066	0.002	0.022	0.007	0.055
	20	0.046	0.035	0.026	0.064	0.026	0.026	0.014	0.061
	30	0.035	0.034	0.028	0.065	0.032	0.025	0.017	0.063
	50	0.049	0.033	0.026	0.061	0.040	0.022	0.018	0.062
	100	0.053	0.034	0.033	0.067	0.045	0.026	0.024	0.063
	200	0.057	0.036	0.036	0.066	0.048	0.026	0.026	0.063
20	10	0.012	0.038	0.014	0.057	0.000	0.026	0.005	0.045
	20	0.040	0.037	0.022	0.060	0.016	0.031	0.013	0.058
	30	0.027	0.042	0.031	0.065	0.026	0.030	0.017	0.059
	50	0.052	0.041	0.032	0.065	0.040	0.029	0.023	0.060
	100	0.057	0.042	0.040	0.067	0.043	0.030	0.023	0.058
	200	0.058	0.045	0.041	0.067	0.047	0.031	0.029	0.060
30	10	0.007	0.037	0.010	0.054	0.000	0.024	0.002	0.040
	20	0.042	0.042	0.024	0.063	0.015	0.032	0.012	0.054
	30	0.023	0.040	0.027	0.059	0.027	0.033	0.017	0.059
	50	0.053	0.041	0.034	0.060	0.041	0.031	0.021	0.056
	100	0.056	0.043	0.037	0.060	0.047	0.032	0.027	0.057
	200	0.054	0.040	0.037	0.059	0.046	0.032	0.029	0.057
50	10	0.004	0.042	0.011	0.057	0.000	0.020	0.001	0.032
	20	0.037	0.042	0.021	0.055	0.009	0.030	0.010	0.046
	30	0.015	0.042	0.024	0.057	0.018	0.033	0.015	0.052
	50	0.051	0.046	0.034	0.060	0.041	0.036	0.022	0.058
	100	0.055	0.044	0.036	0.058	0.049	0.039	0.032	0.061
	200	0.056	0.046	0.042	0.063	0.048	0.036	0.033	0.055
100	10	0.001	0.040	0.006	0.051	0.000	0.014	0.000	0.019
	20	0.037	0.045	0.015	0.055	0.006	0.031	0.006	0.043
	30	0.013	0.046	0.026	0.058	0.020	0.037	0.012	0.050
	50	0.061	0.047	0.030	0.058	0.046	0.038	0.019	0.052
	100	0.064	0.046	0.036	0.056	0.060	0.040	0.029	0.056
	200	0.060	0.040	0.038	0.052	0.064	0.040	0.033	0.055

Table 3. Size of the tests: serially correlated case

N	T	constant case						trend case					
		SCC			WCC			SCC			WCC		
		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 4. Power of the tests: serially correlated case

N	T	constant case						trend case					
		SCC			WCC			SCC			WCC		
		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