

## Article

# Testing Cross-Sectional Correlation in Large Panel Data Models with Serial Correlation

Badi H. Baltagi <sup>1</sup>, Chihwa Kao <sup>2</sup> and Bin Peng <sup>3,\*</sup>

<sup>1</sup> Department of Economics & Center for Policy Research, 426 Eggers Hall, Syracuse University, Syracuse, NY 13244-1020, USA; bbaltagi@maxwell.syr.edu

<sup>2</sup> Department of Economics, 365 Fairfield Way, U-1063, University of Connecticut, Storrs, CT 06269-1063, USA; chih-hwa.kao@uconn.edu

<sup>3</sup> Department of Finance, 523 School of Economics, Huazhong University of Science and Technology, Wuhan 430074, China

\* Correspondence: bpeng01@hust.edu.cn; Tel.: +86-188-276-19485

Academic Editors: In Choi and Ryo Okui

Received: 23 July 2016; Accepted: 19 October 2016; Published: 4 November 2016

**Abstract:** This paper considers the problem of testing cross-sectional correlation in large panel data models with serially-correlated errors. It finds that existing tests for cross-sectional correlation encounter size distortions with serial correlation in the errors. To control the size, this paper proposes a modification of Pesaran's Cross-sectional Dependence (CD) test to account for serial correlation of an unknown form in the error term. We derive the limiting distribution of this test as  $(N, T) \rightarrow \infty$ . The test is distribution free and allows for unknown forms of serial correlation in the errors. Monte Carlo simulations show that the test has good size and power for large panels when serial correlation in the errors is present.

**Keywords:** cross-sectional correlation test; serial correlation; large panel data model

**JEL Classification:** C13; C33

## 1. Introduction

This paper studies testing for cross-sectional correlation in panel data when serial correlation is also present in the disturbances. It does that for the case of strictly-exogenous regressors<sup>1</sup>. Cross-sectional correlation could be due to unknown common shocks, spatial effects or interactions within social networks. Ignoring cross-sectional correlation in panels can have serious consequences. In time series with serial correlation, existing cross-sectional correlation leads to efficiency loss for least squares and invalidates inference. In some cases, it results in inconsistent estimation; see Lee [1] and Andrews [2]. Testing the cross-sectional correlation of panel residuals is therefore important.

One could test for a specific form of correlation in the error like spatial correlation; see Anselin and Bera [3] for cross-sectional data and Baltagi et al. [4] for panel data, to mention a few. Alternatively, one could test for correlation without imposing any structure on the form of correlation among the disturbances. The null hypothesis, in that case, is testing the diagonality of the covariance or correlation matrix of the  $N$ -dimensional disturbance vector  $u_t = (u_{1t}, \dots, u_{Nt})'$ , which is usually assumed to be independent over time, for  $t = 1, \dots, T$ . When  $N$  is fixed and  $T$  is large, the traditional multivariate statistics techniques, including log-likelihood ratio and Lagrange multiplier tests, are applicable; see, for example, Breusch and Pagan [5], who propose a Lagrange

<sup>1</sup> The inclusion of predetermined variables, which is the weakly-exogenous case, alters the results.

Multiplier (LM) test, which is based on the average of the squared pair-wise correlation coefficients of the least squares residuals.

However, as  $N$  becomes large because of the growing availability of the comprehensive databases in macro and finance, this so-called “high dimensional” phenomenon brings challenges to classical statistical inference. As shown in the Random Matrix Theory (RMT) literature, the sample covariance and correlation matrices are ill-conditioned since their eigenvectors are not consistent with their population counterparts; see Johnstone [6] and Jiang [7]. New approaches have been considered in the statistics literature for the testing the diagonality of the sample covariance or correlation matrices; see Ledoit and Wolf [8], Schott [9] and Chen et al. [10], to mention a few.

The above tests for raw data cannot be used directly to test cross-sectional correlation in panel data regressions since the disturbances are not observable. Noise caused by substituting residuals for the actual disturbances may accumulate due to large dimensions, and this in turn may lead to biased inference. The bias for cross-sectional correlation tests in large panels depends on the model specification, the estimation method and the sample sizes  $N$  and  $T$ , among other things. For example, Pesaran et al. [11] consider an LM test and correct its bias in a large heterogeneous panel data model; Baltagi et al. [12] extend Schott’s test [9] to a fixed effects panel data model and correct the bias caused by estimating the disturbances with fixed effects residuals in a homogeneous panel data model. Following Ledoit and Wolf [8], Baltagi et al. [13] propose a bias-adjusted test for testing the null of sphericity in the fixed effects homogeneous panel data model. However, this method does not test cross-sectional correlation directly. Rejection of the null could be due to cross-sectional correlation or heteroscedasticity or both. A general test for cross-sectional correlation was proposed by Pesaran [14]. His test statistic is based on the average of pair-wise correlation coefficients, defined as  $CD_P$  ( $CD$ , Cross-sectional Dependence). The test is exactly centered at zero under the null and does not need bias correction. Pesaran [15] extends his test statistic to test the null of weak cross-sectional correlation and derives its asymptotic distribution using joint limits. This test is robust to many model specifications and has many applications. Recent surveys for cross-sectional correlation or dependence tests in large panels are provided by Moscone and Tosetti [16], Sarafidis and Wansbeek [17] and Chudik and Pesaran [18].

The asymptotics and bias-correction of existing tests for cross-sectional correlation in large panels are carried out under some, albeit restrictive, assumptions. For instance, the errors are normally distributed;  $N/T \rightarrow c \in (0, \infty)$  as  $(N, T) \rightarrow \infty$ , and so on. One fundamental restriction is that the errors are independent over time. In fact, the presence of serial correlation in panel data applications is likely to be the rule rather than the exception, especially for macro applications and when  $T$  is large. Ignoring serial correlation does not affect the consistency of estimates, but it leads to incorrect inference. In RMT, when  $u_1, u_2, \dots, u_T$  are independent across  $t = 1, \dots, T$ , and  $N$  is large, the Limiting Spectral Distribution (LSD) of the corresponding sample covariance matrix is the Marchenko-Pastur (M-P) law; see Bai and Silverstein [19]. Existing correlation among these disturbances may cause a deviation of the LSD from the M-P law. Indeed, Bai and Zhou [20] show that the LSD of the sample covariance matrix with correlations in columns is different from the M-P law. Gao et al. [21] show similar results for the sample correlation matrix. Therefore, the cross-sectional correlation tests, which heavily depend on the assumption of independence over time, could lead to misleading inference if there is a serial correlation in the disturbances.

To better understand the effects of potential serial correlation on the existing tests of cross-sectional correlation, let us assume that the  $T \times 1$  independent random vectors  $u_i = (u_{i1}, \dots, u_{iT})'$ , for  $i = 1, \dots, N$  are observable. The correlation coefficients  $\rho_{ij}$  of any  $u_i$  and  $u_j$  ( $i \neq j$ ) are defined by  $u_i' u_j / (\|u_i\| \cdot \|u_j\|)$ . Their means are zero vectors. If all of the elements of each  $u_i$  are independent and identically spherically distributed, Muirhead [22] shows that  $E(\rho_{ij}^2) = 1/T$ . When  $N$  is fixed, the summation of all distinct  $N(N-1)/2$  terms of  $\rho_{ij}^2$  will be small, as  $T \rightarrow \infty$ . In Section 3, we show that if all of the elements of each  $u_i$  follow a multiple Moving Average model of order one (MA(1)) with parameter  $\theta$ , then  $E(\rho_{ij}^2) = [1/T + \theta^2/(T + T\theta^2)]$ . As  $N \rightarrow \infty$ , the extra term  $\theta^2/(T + T\theta^2)$  can

accumulate and lead to extra bias for the existing LM type tests in panels. Although  $CD_P$  is centered at zero, it may still encounter size distortions because serial correlation is ignored.

This paper proposes a modification of Pesaran's CD test of cross-sectional correlation when the error terms are serially correlated in large panel data models. First, using results from RMT, we study the first two moments of the test statistic and propose an unbiased and consistent estimate of the variance with unknown serial correlation under the null. Second, we derive the limiting distribution of the test under the asymptotic framework with  $(N, T) \rightarrow \infty$  simultaneously in any order without any distribution assumption. We also discuss its local power properties under a multi-factor alternative. Monte Carlo simulations are conducted to study the performance of our test statistic in finite samples. The results confirm our theoretical findings.

The plan for the paper is as follows. The next section introduces the model and notation, existing LM type tests and the Cross-sectional Dependence (CD) test. It then presents our assumptions and the proposed modified Pesaran's CD test statistic. Section 3 derives the asymptotics of this test statistic. Section 4 reports the results of the Monte Carlo experiments. Section 5 provides some concluding remarks. All of the mathematical proofs are provided in the Appendix.

Throughout the paper, we adopt the following notation. For a squared matrix  $B$ ,  $\text{tr}(B)$  is the trace of  $B$ ;  $\|B\| = (\text{tr}(B'B))^{1/2}$  denotes the Frobenius norm of a matrix or the Euclidean norm of a vector  $B$ ;  $\xrightarrow{d}$  denotes convergence in distribution; and  $\xrightarrow{p}$  denotes convergence in probability. We use  $(N, T) \rightarrow \infty$  to denote the joint convergence of  $N$  and  $T$  when  $N$  and  $T$  pass to infinity simultaneously.  $K$  is a generic positive number not depending on  $N$  nor  $T$ .

## 2. Model and Tests

Consider the following heterogeneous panel data model

$$y_{it} = \beta_i' x_{it} + u_{it}, \text{ for } i = 1, \dots, N; t = 1, \dots, T, \quad (1)$$

where  $i$  and  $t$  index the cross-section dimension and time dimension, respectively;  $y_{it}$  is the dependent variable, and  $x_{it}$  is a  $k \times 1$  vector of exogenous regressors. The individual coefficients  $\beta_i$  are defined on a compact set and allowed to vary across  $i$ . The null hypothesis of no cross-sectional correlation is

$$H_0 : \text{cov}(u_{it}, u_{jt}) = 0, \text{ for all } t, i \neq j,$$

or equivalently as

$$H_0 : \rho_{ij} = 0, \text{ for } i \neq j, \quad (2)$$

where  $\rho_{ij}$  is the pair-wise correlation coefficients of the disturbances defined by

$$\rho_{ij} = \frac{\sum_{t=1}^T u_{it} u_{jt}}{\left(\sum_{t=1}^T u_{it}^2\right)^{1/2} \left(\sum_{t=1}^T u_{jt}^2\right)^{1/2}}.$$

Under the alternative, there exists at least one  $\rho_{ij} \neq 0$ , for some  $i \neq j$ . For the panel regression model 1, the residuals are unobservable. In this case, the test statistic is based on the residual-based correlation coefficients  $\hat{\rho}_{ij}$ . Specifically,

$$\hat{\rho}_{ij} = \frac{\sum_{t=1}^T e_{it} e_{jt}}{\left(\sum_{t=1}^T e_{it}^2\right)^{1/2} \left(\sum_{t=1}^T e_{jt}^2\right)^{1/2}}, \quad (3)$$

where  $e_{it}$  is the Ordinary Least Squares (OLS) residuals using  $T$  observations for each  $i = 1, \dots, N$ . These OLS residuals are given by

$$e_{it} = y_{it} - x_{it}' \hat{\beta}_i, \quad (4)$$

with  $\hat{\beta}_i$  being the OLS estimates of  $\beta_i$  from (1) for  $i = 1, \dots, N$ . Let  $M_i = I_T - P_{X_i}$ , where  $P_{X_i} = X_i (X_i' X_i)^{-1} X_i'$ , and  $X_i$  is a  $T \times k$  matrix of regressors with the  $t$ -th row being the  $1 \times k$  vector  $x_{it}'$ . We also define  $u_i = (u_{i1}, \dots, u_{iT})'$ ,  $e_i = (e_{i1}, \dots, e_{iT})'$  and  $v_i = e_i / \|e_i\|$ , for  $i = 1, \dots, N$ . The OLS residuals can be rewritten in vector form as  $e_i = M_i u_i$ , and the residual-based pair-wise correlation coefficients can be rewritten as  $\hat{\rho}_{ij} = v_i' v_j$ , for any  $1 \leq i \neq j \leq N$ .

### 2.1. LM and CD Tests

For  $N$  fixed and  $T \rightarrow \infty$ , Breusch and Pagan [5] propose an LM test to test the null of no cross-sectional correlation in (2) without imposing any structure on this correlation. It is given by

$$LM_{BP} = T \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij}^2. \quad (5)$$

$LM_{BP}$  is asymptotically distributed as a Chi-squared distribution with  $N(N-1)/2$  degrees of freedom under the null. However, for a typical micro-panel dataset,  $N$  is larger than  $T$ ; and the Breusch-Pagan LM test statistic is not valid under this “large  $N$ , small  $T$ ” setup. In fact, Pesaran [14] proposes a scaled version of this LM test as follows

$$LM_P = \sqrt{\frac{1}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^N (T \hat{\rho}_{ij}^2 - 1). \quad (6)$$

Pesaran [14] shows that  $LM_P$  is distributed as  $N(0, 1)$  with  $T \rightarrow \infty$  first, then  $N \rightarrow \infty$  under the null. However,  $E(T \hat{\rho}_{ij}^2 - 1)$  is not correctly centered at zero with fixed  $T$  and large  $N$ . Hence, Pesaran et al. [11] propose a bias-adjusted version of this LM test, denoted by  $LM_{PUY}$ . They show that the exact mean and variance of  $(T-k) \hat{\rho}_{ij}^2$  are given by

$$\mu_{Tij} = E[(T-k) \hat{\rho}_{ij}^2] = \frac{1}{T-k} \text{tr}[E(M_i M_j)], \quad (7)$$

and

$$v_{Tij}^2 = \text{var}[(T-k) \hat{\rho}_{ij}^2] = \left\{ \text{tr}^2[E(M_i M_j)] \right\} a_{1T} + 2 \text{tr} \left\{ [E(M_i M_j)]^2 \right\} a_{2T}, \quad (8)$$

where  $a_{1T} = a_{2T} - \frac{1}{(T-k)^2}$ , and  $a_{2T} = 3 \left[ \frac{(T-k-8)(T-k+2)}{(T-k+2)(T-k-2)(T-k-4)} \right]^2$ .  $LM_{PUY}$  is given by

$$LM_{PUY} = \sqrt{\frac{2}{N(N-1)}} \frac{(T-k) \hat{\rho}_{ij}^2 - \mu_{Tij}}{v_{Tij}}. \quad (9)$$

Pesaran et al. [11] show that  $LM_{PUY}$  is asymptotically distributed as  $N(0, 1)$  under the null (2) and the normality assumption of the disturbances as  $T \rightarrow \infty$  followed by  $N \rightarrow \infty$ . Alternatively, Pesaran [14] proposes a test based on the average of pair-wise correlation coefficients rather than their squares. The test statistic is given by

$$CD_P = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij}. \quad (10)$$

Pesaran [15] shows that this test is asymptotically distributed as  $N(0, 1)$  with  $(N, T) \rightarrow \infty$ . He also extends this to test the null of weak cross-sectional correlation.

### 2.2. Assumptions and the Modified CD Test Statistic

So far, all of the methods surveyed above for testing cross-sectional correlation in panel data models assume that the disturbances are independent over time. Ignoring serial correlation usually

results in efficiency loss and biased inference. In fact, we show in Section 3 that the existence of serial correlation leads to extra bias in the LM-type tests. For the  $CD_P$  test in (10), it is still centered at zero with serial correlation, but its variance is affected by serial correlation. As a result, we also expect size distortions in  $CD_P$ . To correct for this, we consider a modification of this test statistic that accounts for an unknown form of serial correlation in the disturbances. First, we introduce the assumptions needed:

**Assumption 1.** Define  $\xi_i = (\xi_{i0}, \xi_{i1}, \dots, \xi_{iT})'$  and  $\varepsilon_i = (\varepsilon_{i0}, \varepsilon_{i1}, \dots, \varepsilon_{iT})'$ . We also assume that  $\xi_i = \sigma_i \varepsilon_i$ , for  $i = 1, \dots, N$ , where  $\varepsilon_i$  is a random vector with mean vector zero and covariance matrix  $I_T$ . Let  $\varepsilon_{it}$  denote the  $t$ -th entry of  $\varepsilon_i$ , for any  $i = 1, \dots, N$ .  $\varepsilon_{it}$  has a uniformly bounded fourth moment, and there exists a finite constant  $\Delta$ , such that  $E(\varepsilon_{it}^4) = 3 + \Delta$ . Following Bai and Zhou [20], the disturbances  $u_t = (u_{1t}, u_{2t}, \dots, u_{Nt})'$  are generated by

$$u_t = \sum_{s=0}^{\infty} d_s \xi_{t-s}, \text{ for } t = 1, \dots, T, \quad (11)$$

where  $\xi_t = (\xi_{1t}, \xi_{2t}, \dots, \xi_{Nt})'$ , for  $t = 0, 1, \dots, T$ , are IID random vectors across time, and  $\{d_s\}_{s=0}^{\infty}$  is a sequence of numbers satisfying  $\sum_{s=0}^{\infty} |d_s| < K < \infty$ .

Assumption 1 allows the error term  $u_{it}$  to be correlated over time. The condition  $\sum_{s=0}^{\infty} |d_s| < K < \infty$  excludes long memory-type strong dependence. We need bounded moment conditions to ensure large  $(N, T)$  asymptotics for panel data models with serial correlation. The conditions in Assumption 1 are quite relaxable; they are satisfied by many parametric weak dependence processes, such as stationary and invertible finite-order Auto-Regressive and Moving Average (ARMA) models. Under Assumption 1, the covariance matrix of each  $u_i$  is  $\Sigma_i = \sigma_i^2 \Sigma$ , where  $\Sigma$  is a  $T \times T$  symmetric positive definite matrix. The random vector  $u_i$  can be written as  $u_i = \sigma_i \Gamma \varepsilon_i$ , where  $\Gamma \Gamma' = \Sigma$ . The generic covariance matrix  $\Sigma_i$  of each  $u_i$  captures the serial correlation. Bai and Zhou [20] use this representation and show that  $1/T \text{tr}(\Sigma^\kappa)$  is bounded for any fixed positive integer  $\kappa$ . More specifically, considering a multiple Moving Average model of order one (MA(1))

$$u_t = \xi_t + \theta \xi_{t-1}, t = 1, \dots, T, \quad (12)$$

where  $|\theta| < 1$  and  $u_t, \xi_t, u_i$  and  $\xi_i$  are defined in Assumption 1. For this case,  $\Sigma^{\text{MA}} = (\delta_{lr})_{T \times T}$ , where

$$\delta_{lr} = \begin{cases} (1 + \theta^2), & l = r; \\ \theta, & |l - r| = 1; \\ 0, & |l - r| > 1. \end{cases} \quad (13)$$

One can also verify that for (11), we have the following generic representation,

$$\Sigma = (\omega_{lr})_{T \times T}, \text{ where } \omega_{lr} = \sum_{s=0}^{\infty} d_s d_{(|l-r|+s)}. \quad (14)$$

We use this representation throughout the paper for convenience.

**Assumption 2.** The regressors,  $x_{it}$ , are strictly exogenous, such that

$$E(u_{it} | X_i) = 0, \text{ for all } i = 1, \dots, N \text{ and } t = 1, \dots, T, \quad (15)$$

and  $X_i' X_i$  is a positive definite matrix.

**Assumption 3.**  $T > k$  and the OLS residuals,  $e_{it}$ , defined by (4), are not all zeros with probability approaching one.

Assumptions 2 and 3 are standard for model (1); see Pesaran [14] and Pesaran et al. [11]. We impose the assumption that the regressors are strictly exogenous. We do not impose any restrictions

on the distribution of the errors or the relative convergence speed of  $(N, T)$ . This framework is quite relaxable while LM-type tests usually impose the normality assumption and restrictions on the relative speed of  $N$  and  $T$ , namely  $N/T \rightarrow c \in (0, \infty)$ .

Under these assumptions, the OLS estimates for model (1) are consistent, but inefficient. We focus on the term used in Pesaran's CD test [14]

$$T_n = \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \hat{\rho}_{ij}. \quad (16)$$

In the next section, we derive the first two moments of this test statistic and later derive its limiting distribution under this general unknown form of serial correlation over time.

### 3. Asymptotics

#### 3.1. Asymptotic Distribution under the Null

In this section, we study the asymptotics of the test statistic  $T_n$  defined in (16). To derive its limiting distribution, we first consider its first two moments.

**Theorem 1.** Under Assumptions 1–3 and the null given in (2),

$$E(T_n) = 0 \quad (17)$$

and

$$\gamma^2 = \text{var}(T_n) = \frac{2}{N(N-1)} \sum_{i=2}^N \sum_{j=1}^{i-1} E(\hat{\rho}_{ij}^2) = \frac{2}{N(N-1)} \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\text{tr}(M_j \Sigma M_j M_i \Sigma M_i)}{\text{tr}(M_i \Sigma) \text{tr}(M_j \Sigma)}, \quad (18)$$

where  $M_i = I_T - X_i'(X_i'X_i)^{-1}X_i'$ , and  $\Sigma$  is defined by (14).

Theorem 1 shows that the mean of the test statistic is zero. Its variance depends on  $\Sigma$ , which is a generic form containing serial correlation.

In fact, as shown in the proof of Theorem 1 (see the Appendix B),  $E(\hat{\rho}_{ij}^2) = \text{tr}(M_j \Sigma M_j M_i \Sigma M_i) / [\text{tr}(M_i \Sigma) \text{tr}(M_j \Sigma)]$ . In the special case where the error terms are independent over time,  $\Sigma = I_T$ , and  $E(\hat{\rho}_{ij}^2)$  reduces to  $\text{tr}(M_j M_i) / (T - k)^2$ , which yields the results given in Equation (7) for the  $\text{LM}_{PUY}$  test statistic with no serial correlation. However, with serial correlation in the errors, an extra bias term is introduced in  $\text{LM}_{PUY}$  since

$$\frac{\text{tr}(M_j \Sigma M_j M_i \Sigma M_i)}{\text{tr}(M_i \Sigma) \text{tr}(M_j \Sigma)} - \frac{\text{tr}(M_j M_i)}{(T - k)^2} \neq 0, \text{ if } \Sigma \neq I_T.$$

More specifically, let us assume that  $u_i, i = 1, \dots, N$ , are observable, then  $E(\hat{\rho}_{ij}^2) = \text{tr}(\Sigma^2) / \text{tr}^2(\Sigma)$ . For the MA(1) process defined by (12),  $\text{tr}(\Sigma^2) / \text{tr}^2(\Sigma) = 1/T + \theta^2 / (T + T\theta^2)$  and  $\text{tr}(\Sigma^2) / \text{tr}^2(\Sigma) = 1/T$ , for  $\theta = 0$ . The extra term  $\theta^2 / (T + T\theta^2)$  accumulates in the LM test statistic and leads to extra bias as  $N \rightarrow \infty$ . As discussed above, we expect that  $\text{LM}_{PUY}$  to have serious size distortions when serial correlation is present in the disturbances.

Unlike LM-type tests, the test statistic  $T_n$  is centered at zero; it does not need bias adjustment. Note that if  $u_{it}$  are independent over time, our model reduces to that of Pesaran [14]. Let  $\gamma_0^2$  be the variance of  $T_n$  without serial correlation; it can be written as

$$\gamma_0^2 = \frac{2}{N(N-1)} \sum_{i=2}^N \sum_{j=1}^{i-1} \left[ \frac{T-2k}{(T-k)^2} + \frac{\text{tr}(P_{X_i} P_{X_j})}{(T-k)^2} \right], \quad (19)$$

where  $P_{X_i} = X_i (X_i' X_i)^{-1} X_i'$  and  $P_{X_j} = X_j (X_j' X_j)^{-1} X_j'$ . The above result is the exact variance for  $T_n$  without serial correlation; it is derived by Pesaran [15]. A modified version of  $CD_P$  is also given by Pesaran [15] using this exact variance. From Theorem 1,  $\gamma^2$  is different from  $\gamma_0^2$  if  $\Sigma \neq I_T$ . Hence, we also expect  $CD_P$  to have size distortions when serial correlation is present in the disturbances. Next, we consider the limiting distribution of the proposed test. The result is given in the following theorem.

**Theorem 2.** Under Assumptions 1–3 and the null in (2), as  $(N, T) \rightarrow \infty$ , we have

$$\gamma^{-1} T_n \xrightarrow{d} N(0, 1). \quad (20)$$

Theorem 2 shows that appropriately standardized  $\gamma^{-1} T_n$  is asymptotically distributed as a standard normal. It is valid for  $N$  and  $T$  tending to infinity jointly in any order. However, we do not observe  $\Sigma$  in a panel data regression model; and an estimate of the variance  $\gamma^2$  is needed for practical applications. Following Chen and Qin [23], an unbiased and consistent estimator of  $\gamma^2$  under the null is obtained using the cross-validation approach proposed in the following theorem:

**Theorem 3.** Let  $\hat{\gamma}^2 = \frac{1}{N(N-1)} \sum_{i=2}^N \sum_{j=1}^{i-1} v_i'(v_j - \bar{v}_{(i,j)}) v_j'(v_i - \bar{v}_{(i,j)})$ , where  $\bar{v}_{(i,j)} = \frac{1}{N-2} \sum_{1 \leq \tau \neq i, j \leq N} v_\tau$ . Under Assumptions 1–3 and the null in (2),  $E(\hat{\gamma}^2) = \gamma^2$ . As  $(N, T) \rightarrow \infty$ ,

$$\hat{\gamma}^2 \xrightarrow{p} \gamma^2. \quad (21)$$

Define  $CD_R = \hat{\gamma}^{-1} T_n$ . As  $(N, T) \rightarrow \infty$ ,

$$CD_R \xrightarrow{d} N(0, 1). \quad (22)$$

Theorem 3 shows that  $\hat{\gamma}^2$  is a good approximation for the variance, and we do not need to specify the structure of  $\Sigma$ . In other words, the test statistic allows the error terms of model (1) to be dependent over time. Furthermore,  $CD_R$  is a modified version of  $CD_P$ , so they are likely to perform very similarly with respect to many model specifications (see Pesaran [14]).

### 3.2. Local Power Properties

We now consider the power analysis of the test. Naturally, the power properties depend on the specifications of the alternatives. One general alternative specification that allows for global cross-sectional correlation in panels is the unobserved multi-factor model. Under this alternative, the new error terms are defined by

$$u_i' = u_i + \sigma_i F \lambda_i = \sigma_i (\Gamma \varepsilon_i + F \lambda_i), \quad (23)$$

where  $F = (f_1, f_2, \dots, f_T)'$  denotes the  $T \times r$  common factor matrix and  $\lambda_i$  is the  $r$  factor loading vector. Under the null hypothesis,  $\lambda_i = 0$ , for all  $i$ . We now consider the following Pitman-type local alternative<sup>2</sup>

$$H_a : \lambda_i = \frac{1}{T^{1/4} N^{1/2}} \delta_i, \text{ for some } i, \quad (24)$$

where  $\delta_i$  is a non-random and non-zero  $r \times 1$  vector, which does not depend on  $N$  or  $T$ . To simplify the analysis, we add the following assumption:

<sup>2</sup> We only consider the case that the number of non-zero factor loading vectors is  $N$  or of order  $N$ , which means the model has strong error cross-sectional correlation. For the weak error cross-sectional correlation case, we conjecture that it is similar to Pesaran [15].



**Assumption 4.** (1)  $f_t \sim \text{IID}(0, I_r)$ ; (2)  $f_t$  are independent of  $\varepsilon_{it}$ ,  $x_{it}$ , for all  $i$  and  $t$ ; (3) for each  $i$ ,  $T^{-1/2} \sum_{t=1}^T u_{it} f_t = O_p(1)$ ;  $T^{-1/2} \sum_{t=1}^T x_{it} f_t' = O_p(1)$  and  $T^{-1} \sum_{t=1}^T f_t f_t' = I_r + O_p(T^{-1/2})$ ; (4)  $T^{-1/2} X_i' X_j = O_p(1)$  and  $T^{-1/2} X_i' u_i = O_p(1)$ , for all  $i$  and  $j$ .

The following theorem gives the power properties under the local alternative (24).

**Theorem 4.** Under Assumptions 1–4 and local alternative (24), as  $(N, T) \rightarrow \infty$ ,

$$\gamma^{-1} T_n \xrightarrow{d} N(\psi, 1), \quad (25)$$

where  $\psi = \text{plim}_{(N,T) \rightarrow \infty} \gamma^{-1} \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \left( \frac{T^{1/2} N^{-1} \delta_i' \delta_j}{tr^{1/2}(M_i \Sigma) tr^{1/2}(M_j \Sigma)} \right) \neq 0$ .

From Theorem 4, the test has nontrivial power against the local alternative that contracts to the null at the rate of  $T^{-1/4} N^{-1/2}$ . Hence, Theorem 4 establishes the consistency of the proposed test at the rate of  $N\sqrt{T}$  under the alternative, as long as  $\psi \neq 0$ .

#### 4. Monte Carlo Simulations

This section conducts Monte Carlo simulations to examine the empirical size and power of the proposed test ( $CD_R$ ) defined in (22) in heterogeneous panel data regression models. We also look at the performance of  $LM_{PUY}$  and  $CD_P$  defined by (9) and (10), respectively, for comparison purposes. We consider four scenarios: (1) the errors are independent over time, with no serial correlation; (2) the errors follow a moving average model of order one (MA(1)) over time; (3) the errors follow an Auto-Regressive model of order one (AR(1)) over time; (4) the errors follow an Auto-Regressive and Moving Average of order (1,1) (ARMA(1,1)) over time. Finally, we provide small sample evidence on the power performance of the modified  $CD_R$  test against a factor and spatial auto-regressive model of order one alternatives, which are popular in economics for modeling cross-sectional correlation.

##### 4.1. Experimental Design

Following Pesaran et al. [11], our experiments use the following data-generating process

$$y_{it} = \alpha_i + \beta_i x_{it} + u_{it}, \quad i = 1, \dots, N; \quad t = 1, \dots, T, \quad (26)$$

$$x_{it} = \eta x_{it-1} + v_{it}, \quad (27)$$

where  $\alpha_i \sim \text{IIDN}(1, 1)$ ;  $\beta_i \sim \text{IIDN}(1, 0.04)$ .  $x_{it}$  is a strictly exogenous regressor, and we set  $\eta = 0.6$  and  $v_{it} \sim \text{IIDN}(\phi_i^2 / (1 - 0.6^2))$  with  $\phi_i \sim \text{IID}\chi^2(6)/6$ , for  $i = 1, \dots, N$ . The error terms of (26) are generated using the following four data generating processes

$$(1) \text{ IID : } u_{it} = \xi_{it}; \quad (28)$$

$$(2) \text{ MA(1) : } u_{it} = \xi_{it} + \theta \xi_{it-1}; \quad (29)$$

$$(3) \text{ AR(1) : } u_{it} = \rho u_{it-1} + \xi_{it}; \quad (30)$$

$$(4) \text{ ARMA(1,1) : } u_{it} = \rho u_{it-1} + \xi_{it} + \theta \xi_{it-1}, \quad (31)$$

where  $\xi_{it} = \sigma_i \varepsilon_{it}$ ;  $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$  and  $\varepsilon_{it} \sim \text{IID}(0, 1)$ . We further set  $\theta = 0.8$  and  $\rho = 0.6$ . To check the robustness of the tests to non-normal distributions,  $\varepsilon_{it}$  are generated from a Normal(0,1) and a Chi-squared distribution ( $\chi^2(2)/2 - 1$ ).

To examine the empirical power of the tests, we consider two different cross-sectional correlation alternatives: factor and spatial models. The factor model is generated by

$$u_{it}^* = \lambda_i f_t + u_{it}, \quad \text{for } i = 1, \dots, N; \quad t = 1, \dots, T, \quad (32)$$



where  $f_i \sim \text{IIDN}(0,1)$  and  $\lambda_i \sim \text{IIDU}[0.1,0.3]$ ; In this case,  $u_{it}^*$  replaces  $u_{it}$  in (26) for the power studies.  $u_{it}$  is generated by the four scenarios defined by (28)–(31), respectively. For the spatial model, we consider a first-order spatial auto-correlation model (SAR(1)),

$$u_{it}^* = \delta (0.5u_{i-1,t}^* + 0.5u_{i+1,t}^*) + u_{it}, \quad (33)$$

where  $\delta = 0.4$  and  $u_{it}$  are defined by (28)–(31), respectively.

The experiments are conducted for  $N = 10, 20, 30, 50, 100, 200$  and  $T = 10, 20, 30, 50, 100$ . For each pair of  $(N, T)$ , we run 2000 replications. To obtain the empirical size, we conduct the proposed test ( $\text{CD}_R$ ) and  $\text{CD}_P$  at the two-sided 5% nominal significance level and  $\text{LM}_{PUY}$  at the positive one-sided 5% nominal significance level.

#### 4.2. Simulation Results

Table 1 reports the empirical size of  $\text{CD}_P$ ,  $\text{LM}_{PUY}$  and  $\text{CD}_R$  for normal and chi-squared distributed errors. The error terms are assumed to be independent over time. The results show that all of the tests have the correct size with different  $(N, T)$  combinations under both normal and chi-squared scenarios. Those are consistent with the theoretical findings. The only exceptions are for small  $N$  or  $T$  equal to 10, especially for  $\text{LM}_{PUY}$ . Table 2 reports the empirical size of the three tests with MA(1) error terms defined by (29). The results show that  $\text{CD}_R$  has the correct size for all  $(N, T)$ , but  $\text{CD}_P$  has size distortions for different  $(N, T)$  combinations because the disturbances are MA(1) over time. For example, under the normality scenario, the size of  $\text{CD}_P$  is 9.35% for  $N = 10$  and  $T = 20$ ; it becomes 11.1% when  $T$  grows to 100.  $\text{LM}_{PUY}$  suffers serious size distortions, because of the extra bias caused by ignoring serial correlation. From Table 2, the empirical size of  $\text{LM}_{PUY}$  is 100% as  $N$  or  $T$  becomes larger than 30. Tables 3 and 4 report the empirical size of the tests with AR(1) and ARMA(1,1) errors under the two distributions: normal and chi-squared scenarios. Note that  $\text{CD}_R$  is over-sized in Table 4 for the chi-squared case when  $T = 10$ . However, it has the correct size as  $T$  gets larger than 20. In contrast,  $\text{LM}_{PUY}$  has serious size issues, rejecting 100% of the time, and  $\text{CD}_P$  is over sized by as much as 25%. Overall, in comparison with  $\text{CD}_P$  and  $\text{LM}_{PUY}$ , the proposed test  $\text{CD}_R$  controls for size distortions when serial correlation in the disturbances is present and is not much affected when serial correlation is not present.

**Table 1.** Size of tests with IID errors over time.

Tests	(N,T)	Normal					Chi-Squared				
		10	20	30	50	100	10	20	30	50	100
$\text{CD}_R$	10	5.75	5.90	5.50	4.75	6.45	5.90	4.80	5.55	5.15	6.45
	20	3.85	4.55	5.05	4.70	5.15	4.60	4.50	4.50	5.85	5.40
	30	4.45	4.10	4.70	5.10	4.60	4.40	4.80	4.45	4.50	6.25
	50	4.45	4.75	5.40	5.25	4.50	4.10	3.65	4.75	4.05	4.60
	100	4.65	4.85	4.20	5.65	5.30	4.35	4.80	4.70	4.35	4.95
	200	4.05	4.65	3.90	4.60	5.00	5.65	5.05	4.85	4.65	5.40
$\text{CD}_P$	10	5.60	5.50	5.25	4.10	6.00	5.60	4.70	5.05	4.70	5.65
	20	4.05	4.75	5.05	4.90	5.30	4.90	4.70	4.65	5.85	5.30
	30	4.90	4.45	4.85	5.20	5.00	5.20	5.20	4.55	5.00	6.05
	50	4.95	5.20	5.60	5.55	4.45	5.00	4.15	5.00	4.55	4.70
	100	5.65	5.15	4.50	5.95	5.45	5.15	5.65	5.05	4.50	5.05
	200	5.00	5.00	4.45	4.85	5.15	6.35	5.75	5.15	4.70	5.55
$\text{LM}_{PUY}$	10	6.75	6.05	6.10	6.00	5.60	6.60	6.85	7.65	7.95	6.60
	20	6.20	5.45	6.75	7.00	5.50	7.05	6.40	6.40	7.15	5.60
	30	6.20	6.25	5.40	6.35	5.95	7.65	5.95	6.35	5.85	7.00
	50	6.55	4.95	5.25	5.60	5.40	7.00	6.85	7.20	5.40	5.85
	100	8.10	5.45	5.40	4.60	4.55	7.00	5.85	6.10	5.85	5.90
	200	8.60	5.75	6.50	5.90	5.35	8.00	7.20	6.30	6.40	6.70

Notes: This table reports the size of  $\text{CD}_P$ ,  $\text{LM}_{PUY}$  and  $\text{CD}_R$  with  $u_{it} = \zeta_{it}$ , where  $\zeta_{it} = \sigma_i \varepsilon_{it}$ ;  $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$ .  $\varepsilon_{it} \sim \text{IID}(0,1)$  and are generated from normal and Chi-squared distributions. The tests are conducted at the 5% nominal significance level.

Table 2. Size of tests with MA(1) errors.

Tests	(N,T)	Normal					Chi-Squared				
		10	20	30	50	100	10	20	30	50	100
$CD_R$	10	6.10	6.25	4.45	5.35	6.25	6.30	5.40	5.90	5.85	6.50
	20	5.15	4.80	5.05	4.60	5.30	5.20	5.35	4.70	6.15	4.75
	30	4.50	4.35	4.20	5.35	4.95	5.55	4.75	4.90	5.30	6.15
	50	5.25	4.50	5.30	5.70	4.30	5.00	4.65	4.60	4.35	4.85
	100	4.75	5.35	4.50	5.45	5.60	5.80	4.15	5.45	4.35	4.90
	200	4.35	4.95	3.50	4.50	4.90	6.20	6.30	4.30	4.30	5.50
$CD_P$	10	7.60	9.35	8.40	10.05	11.10	7.80	7.75	10.30	10.25	10.95
	20	6.60	8.30	9.95	9.10	10.90	7.00	8.95	9.30	10.70	10.50
	30	6.45	8.35	8.30	10.50	10.60	7.90	9.65	9.50	10.80	10.60
	50	7.45	7.95	10.75	11.30	9.65	7.55	7.90	9.20	9.70	9.15
	100	6.50	9.35	9.00	10.85	11.55	7.85	8.35	10.60	9.30	10.20
	200	6.65	8.45	8.45	9.70	10.95	9.90	9.50	9.35	9.65	11.20
$LM_{PUY}$	10	37.95	54.40	57.10	59.55	60.70	39.15	53.00	56.50	60.75	61.55
	20	81.55	96.00	96.80	98.25	97.90	83.25	95.45	97.05	97.70	98.20
	30	98.30	100.00	100.00	100.00	100.00	98.45	100.00	100.00	100.00	100.00
	50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	200	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: This table reports the size of  $CD_P$ ,  $LM_{PUY}$  and  $CD_R$  with  $u_{it} = \zeta_{it} + \theta \zeta_{it-1}$ , where  $\zeta_{it} = \sigma_i \varepsilon_{it}$ ,  $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$ .  $\varepsilon_{it} \sim \text{IID}(0,1)$  and are generated from normal and Chi-squared distributions. The tests are conducted at the 5% nominal significance level.

Table 3. Size of tests with AR(1) errors.

Tests	(N,T)	Normal					Chi-Squared				
		10	20	30	50	100	10	20	30	50	100
$CD_R$	10	6.10	6.25	4.90	6.15	6.75	6.05	4.80	6.10	6.00	5.65
	20	4.75	5.65	4.65	4.70	5.00	4.85	5.60	4.50	5.55	4.80
	30	4.15	4.85	4.00	4.55	4.65	5.50	4.25	5.75	5.10	6.65
	50	4.15	4.50	5.20	5.45	4.40	5.25	5.35	4.60	4.40	4.35
	100	4.35	4.80	4.80	5.45	4.80	5.75	4.15	5.30	4.05	5.10
	200	4.85	4.60	4.05	4.55	5.05	7.80	5.35	4.95	4.20	4.55
$CD_P$	10	6.80	9.65	10.20	14.55	16.80	6.55	8.25	12.25	13.90	16.30
	20	5.75	9.50	11.35	13.25	16.85	5.90	9.60	11.50	15.05	15.45
	30	5.65	9.80	10.00	13.30	14.05	7.35	9.65	12.00	15.20	17.15
	50	5.90	8.45	11.95	14.80	14.10	7.10	9.55	9.70	12.40	15.80
	100	6.05	10.00	10.40	14.70	16.55	7.25	8.70	12.25	13.85	15.00
	200	6.65	9.00	10.25	13.30	16.70	9.40	10.3	10.85	13.70	16.10
$LM_{PUY}$	10	37.95	54.40	57.10	59.55	60.70	27.60	66.30	82.45	90.80	95.35
	20	55.50	97.90	99.85	100.00	100.00	59.95	98.40	99.85	100.00	100.00
	30	98.30	99.95	100.00	100.00	100.00	82.75	100.00	100.00	100.00	100.00
	50	97.80	100.00	100.00	100.00	100.00	98.60	100.00	100.00	100.00	100.00
	100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	200	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: This table reports the size of  $CD_P$ ,  $LM_{PUY}$  and  $CD_R$  with  $u_{it} = \rho u_{it-1} + \zeta_{it}$ , where  $\zeta_{it} = \sigma_i \varepsilon_{it}$ ,  $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$ .  $\varepsilon_{it} \sim \text{IID}(0,1)$  and are generated from normal and Chi-squared distributions. The tests are conducted at the 5% nominal significance level.

**Table 4.** Size of tests with ARMA(1,1) errors.

Tests	(N,T)	Normal					Chi-Squared				
		10	20	30	50	100	10	20	30	50	100
$CD_R$	10	6.95	6.45	4.90	6.20	5.85	7.20	5.25	6.40	5.40	5.45
	20	5.40	5.55	4.95	4.75	4.95	6.40	5.70	4.95	5.55	4.70
	30	4.65	4.75	4.05	4.80	4.65	7.45	4.60	5.95	5.10	6.50
	50	4.95	4.95	5.25	5.30	4.50	7.50	5.70	4.80	4.35	4.80
	100	5.05	5.15	4.60	5.10	4.90	10.25	5.10	4.65	4.00	4.80
	200	5.75	4.65	4.45	4.85	5.20	17.45	6.60	5.75	4.50	4.25
$CD_P$	10	9.10	15.95	16.35	22.50	24.30	10.95	13.80	19.20	21.70	25.15
	20	8.30	14.40	17.80	20.15	25.05	10.10	14.80	18.90	22.85	23.15
	30	8.30	15.40	17.70	21.55	22.55	10.95	15.25	19.25	23.55	24.25
	50	8.70	14.85	18.80	22.70	23.40	11.75	15.40	17.30	19.15	23.95
	100	9.35	15.90	17.50	22.15	24.20	17.20	14.45	17.95	22.05	22.70
	200	9.50	14.05	18.35	20.00	24.95	25.45	17.00	18.55	21.35	24.65
$LM_{PUY}$	10	83.65	98.45	99.45	99.75	99.80	83.65	98.40	99.70	99.90	100.00
	20	99.85	100.00	100.00	100.00	100.00	99.85	100.00	100.00	100.00	100.00
	30	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	50	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	100	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
	200	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00

Notes: This table reports the size of  $CD_P$ ,  $LM_{PUY}$  and  $CD_R$  with  $u_{it} = \rho u_{it-1} + \xi_{it} + \theta \xi_{it-1}$ , where  $\xi_{it} = \sigma_i \varepsilon_{it}$ ;  $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$ .  $\varepsilon_{it} \sim \text{IID}(0,1)$  and are generated from normal and Chi-squared distributions. The tests are conducted at the 5% nominal significance level.

Table 5 summarizes the size-adjusted power of  $CD_R$  with MA(1), AR(1) and ARMA(1,1) errors under the factor model alternative. Results show that  $CD_R$  performs reasonably well under the two distribution scenarios especially for  $N$  and  $T > 10$ . Table 6 confirms the power properties of  $CD_R$  for MA(1), AR(1) and ARMA(1,1) errors under the SAR(1) alternative, especially for large  $N$  and  $T$ .

**Table 5.** Size adjusted power of  $CD_R$ : factor model.

DGP	(N,T)	Normal					Chi-Squared				
		10	20	30	50	100	10	20	30	50	100
MA(1)	10	14.55	23.95	30.30	45.40	63.05	21.95	30.75	33.65	46.00	66.10
	20	35.70	56.65	68.95	84.05	95.95	47.30	63.25	75.80	86.00	97.40
	30	59.65	81.70	91.75	97.65	99.95	69.75	87.50	92.60	98.00	99.95
	50	83.65	96.60	99.30	100.00	100.00	88.75	98.00	99.55	100.00	100.00
	100	96.75	99.95	100.00	100.00	100.00	98.90	99.90	100.00	100.00	100.00
	200	99.70	100.00	100.00	100.00	100.00	99.70	100.00	100.00	100.00	100.00
AR(1)	10	18.95	23.95	32.40	38.10	56.75	26.95	35.00	28.90	37.15	61.25
	20	45.60	62.10	69.95	81.45	94.20	55.10	67.45	74.85	85.65	96.60
	30	68.80	83.50	92.30	97.60	99.75	78.15	90.85	92.70	97.40	99.85
	50	88.55	97.45	99.40	100.00	100.00	92.90	98.50	99.65	100.00	100.00
	100	98.80	100.00	100.00	100.00	100.00	99.60	99.95	100.00	100.00	100.00
	200	99.90	100.00	100.00	100.00	100.00	99.85	100.00	100.00	100.00	100.00
ARMA(1,1)	10	7.70	7.70	10.00	10.80	14.80	9.65	10.35	8.80	9.60	19.60
	20	22.05	18.85	24.25	27.80	39.50	24.85	22.35	23.40	30.60	46.20
	30	37.75	37.45	46.15	48.90	75.00	41.75	47.35	44.15	53.15	71.25
	50	66.50	66.75	71.60	83.10	96.20	66.25	72.35	82.45	88.20	98.00
	100	91.15	96.60	98.75	99.90	100.00	90.45	98.55	99.40	99.95	100.00
	200	98.95	100.00	100.00	100.00	100.00	98.45	99.95	100.00	100.00	100.00

Notes: This table computes the size adjusted power for  $CD_R$  with a factor model that allows for cross-sectional correlation in the errors:  $u_{it}^* = \lambda_i f_t + u_{it}$ .  $u_{it}$  are generated by MA(1), AR(1) and ARMA (1,1) defined by (29)–(31).  $\xi_{it} = \sigma_i \varepsilon_{it}$ ;  $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$ .  $\varepsilon_{it} \sim \text{IID}(0,1)$  and are generated from normal and Chi-squared distributions.

**Table 6.** Size adjusted power of  $CD_R$ : SAR(1) model.

DGP	(N,T)	Normal					Chi-Squared				
		10	20	30	50	100	10	20	30	50	100
MA(1)	10	38.85	60.55	72.20	88.25	97.30	43.05	67.15	72.55	88.45	97.70
	20	37.45	61.70	76.00	92.15	99.05	39.25	61.25	76.80	89.55	99.10
	30	39.60	64.55	78.60	92.00	99.60	40.30	65.65	78.80	91.90	99.35
	50	40.05	66.45	79.15	92.70	99.75	39.95	66.55	78.65	94.65	99.70
	100	33.60	62.70	80.55	92.55	99.65	37.85	64.65	79.20	94.40	99.90
	200	40.65	64.50	80.65	94.70	99.8	37.75	62.50	81.25	95.65	99.80
AR(1)	10	37.20	53.95	68.20	79.20	92.10	42.85	63.20	61.15	78.00	94.80
	20	38.25	56.50	69.30	82.90	95.85	38.55	55.50	68.65	83.70	97.20
	30	37.90	56.90	71.80	84.65	98.10	38.70	62.00	66.25	85.70	96.90
	50	38.80	59.80	71.40	86.60	98.60	39.70	59.15	71.25	89.00	99.00
	100	38.85	57.85	70.90	86.60	98.75	35.25	59.85	72.55	88.95	98.60
	200	40.75	55.95	74.40	87.75	98.80	33.80	56.00	70.85	90.40	99.10
ARMA(1,1)	10	29.00	43.40	58.05	70.20	85.90	32.75	49.75	51.30	67.40	88.20
	20	31.05	43.55	56.65	72.10	89.10	28.35	43.45	54.80	71.35	91.35
	30	30.00	45.70	59.35	71.35	94.20	28.10	48.10	54.00	73.05	91.90
	50	33.05	45.30	54.40	71.70	93.30	27.30	43.90	58.00	75.75	94.45
	100	30.60	45.15	55.50	75.40	94.95	21.80	45.45	57.85	77.35	94.75
	200	30.30	42.05	58.15	75.75	95.15	21.05	38.80	55.70	77.50	95.80

Notes: This table computes the size adjusted power for  $CD_R$  with a SAR(1) model that allows for cross-sectional correlation in the error:  $u_{it}^* = \delta(0.5u_{i-1,t}^* + 0.5u_{i+1,t}^*) + u_{it}$  with  $\delta = 0.4$ .  $u_{it}$  are generated by MA(1), AR(1) and ARMA (1,1) defined by (29)–(31).  $\xi_{it} = \sigma_i \varepsilon_{it}$ ;  $\sigma_i^2 \sim \text{IID}\chi^2(2)/2$ .  $\varepsilon_{it} \sim \text{IID}(0,1)$  and are generated from normal and Chi-squared distributions.

## 5. Conclusions

In this paper, we find that in the large heterogeneous panel data model,  $LM_{PUY}$  exhibits serious size bias when there is serial correlation in the disturbances. While  $CD_P$  is centered at zero, it still encounters size distortions caused by ignoring serial correlation. We modify Pesaran's  $CD_P$  test to account for serial correlation of an unknown form in the error term and call it  $CD_R$ . This paper has several novel aspects: first, an unbiased and consistent estimate of the variance under the assumptions and the null of no cross-section correlation is proposed without knowing the form of serial correlation over time. Second, the limiting distribution of the test is derived as  $(N, T) \rightarrow \infty$  in any order. Third, it is distribution free. Simulations show that the proposed test  $CD_R$  successfully controls for size distortions with serial correlation in the error term. It also has reasonable power under the alternatives of a factor model and a spatial auto-correlation SAR(1) model for different serial correlation specifications.

**Author Contributions:** All authors contributed equally to the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix

This Appendix includes proofs of the main results in the text. The Appendix includes two parts: Part A includes some useful lemmas, which are frequently used in the proofs of the theorems; Part B gives the proofs of all of the theorems included in the paper.

Let us introduce some notation before proceeding: For two matrices  $B = (b_{ij})$  and  $C = (c_{ij})$ , we define  $B \circ C = (b_{ij}c_{ij})$ .  $\sum$  denotes summation over mutually-different indices, e.g.,  $\sum_{(i_1, i_2, j_1, j_2)}$  means summation over  $\{(i_1, i_2, j_1, j_2) : i_1, i_2, j_1, j_2 \text{ are mutually different}\}$ .

## Appendix A. Some Useful Lemmas

**Lemma A1.** Let  $F$  and  $G$  be non-stochastic  $N \times N$  symmetric and positive definite matrices. Define  $r = \frac{u_i' F u_i}{u_i' G u_i}$ . Under Assumptions 1, we have

$$\begin{aligned} (a) \quad E(r^k) &= \frac{E[(\varepsilon_i' F \varepsilon_i)^k]}{[E(\varepsilon_i' G \varepsilon_i)]^k}; \\ (b) \quad E(\varepsilon_i' F \varepsilon_i) &= \text{tr}(F); \\ (c) \quad E(\varepsilon_i' F \varepsilon_i)^2 &= \text{tr}(F^2) + 2\text{tr}^2(F) + \Delta \text{tr}(F \circ F); \\ (d) \quad \text{tr}(F \circ F) &\leq \text{tr}(F^2). \end{aligned}$$

The Proof of part (a) is given by Lieberman [24], and the proofs of (b)–(d) are from Proposition 1 of Chen et al. [10]; hence, we omit the proof here.

**Lemma A2.** Define  $B_j = M_j \Sigma M_j$ , for any  $j$ , respectively. Under Assumptions 1–3 and the null in (2), we have

$$\begin{aligned} (a) \quad E(\hat{\rho}_{ij}^2) &= \frac{\text{tr}(B_i B_j)}{\text{tr}(B_i) \text{tr}(B_j)}; \\ (b) \quad E(\hat{\rho}_{ij}^4) &\leq (3 + \Delta) \frac{(2 + \Delta) \text{tr}(B_i B_j)^2 + \text{tr}^2(B_i B_j)}{\text{tr}^2(B_i) \text{tr}^2(B_j)}; \\ (c) \quad \text{For any } j_1 \neq j_2, E(\hat{\rho}_{ij_1}^2 \hat{\rho}_{ij_2}^2) &\leq \frac{\left( (2 + \Delta) \text{tr}(B_i B_{j_1})^2 + \text{tr}^2(B_i B_{j_1}) \right)^{1/2} \left( (2 + \Delta) \text{tr}(B_i B_{j_2})^2 + \text{tr}^2(B_i B_{j_2}) \right)^{1/2}}{\text{tr}(B_{j_1}) \text{tr}(B_{j_2}) \text{tr}^2(B_i)}. \end{aligned}$$

**Proof.** Recall that the pair-wise correlation coefficients is defined as

$$\hat{\rho}_{ij} = v_i' v_j = \sum_{t=1}^T v_{it} v_{jt},$$

where  $v_i$  are the scaled residual vectors defined by  $v_i = \frac{e_i}{(e_i' e_i)^{1/2}}$ .  $e_i$  is the OLS residual vector from the individual-specific least squares regression, and it is given by

$$e_i = M_i u_i = M_i \sigma_i \varepsilon_i, \text{ with } M_i = I_T - P_{X_i} = I_T - X_i (X_i' X_i)^{-1} X_i',$$

where  $M_i$  is idempotent. Consider part (a),

$$E(\hat{\rho}_{ij}^2) = E(v_i' v_j)^2 = E\left(\frac{e_i' e_j}{(e_i' e_i)^{1/2} (e_j' e_j)^{1/2}}\right)^2 = E\left(\frac{e_i' A_j e_i}{e_i' e_i}\right),$$

where  $A_j = \frac{e_j e_j'}{e_j' e_j}$ . Then

$$E(\hat{\rho}_{ij}^2) = E[E(\hat{\rho}_{ij}^2 | \varepsilon_j)] = E\left[E\left(\frac{e_i' A_j e_i}{e_i' e_i} | \varepsilon_j\right)\right].$$

Since  $e_i = M_i \sigma_i \Gamma \varepsilon_i$ , and using parts (a) and (b) of Lemma A1, we have  $E\left(\frac{e_i' A_j e_i}{e_i' e_i} | \varepsilon_j\right) = \frac{\text{tr}(\Gamma' M_i A_j M_i \Gamma)}{\text{tr}(\Gamma' M_i \Gamma)}$ . Moreover,

$$\begin{aligned} E[\text{tr}(\Gamma' M_i A_j M_i \Gamma)] &= E\left(\frac{\varepsilon_j' \Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma \varepsilon_j}{\varepsilon_j' \Gamma' M_j \Gamma \varepsilon_j}\right) \\ &= \frac{\text{tr}(\Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma)}{\text{tr}(\Gamma' M_j \Gamma)} \\ &= \frac{\text{tr}(M_j \Sigma M_j M_i \Sigma M_i)}{\text{tr}(M_j \Sigma)}. \end{aligned}$$

Together with the above results, we have:

$$E(\hat{\rho}_{ij}^2) = \frac{\text{tr}(M_j \Sigma M_j M_i \Sigma M_i)}{\text{tr}(M_i \Sigma) \text{tr}(M_j \Sigma)} = \frac{\text{tr}(B_i B_j)}{\text{tr}(B_i) \text{tr}(B_j)}.$$

Consider part (b),

$$\begin{aligned} E(\rho_{ij}^4) &= E[E(\rho_{ij}^4 | v_j)] = E\left(E\left[\left(\frac{e_i' A_j e_i}{e_i' e_i}\right)^2 | v_j\right]\right) = E\left[\frac{E(\varepsilon_j' \Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma \varepsilon_j)^2}{\text{tr}^2(\Gamma' M_j \Gamma)} | v_j\right] \\ &= E\left[\frac{2\text{tr}(\Gamma' M_i A_j M_i \Gamma)^2 + \text{tr}^2(\Gamma' M_i A_j M_i \Gamma) + \Delta \text{tr}(\Gamma' M_i A_j M_i \Gamma \circ \Gamma' M_i A_j M_i \Gamma)}{\text{tr}^2(B_i)}\right]. \end{aligned}$$

Using part (a) of Lemma A1, we have

$$E[\text{tr}^2(\Gamma' M_i A_j M_i \Gamma)] = E\left(\frac{\varepsilon_j' \Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma \varepsilon_j}{\varepsilon_j' \Gamma' M_j \Gamma \varepsilon_j}\right)^2 = \frac{E(\varepsilon_j' \Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma \varepsilon_j)^2}{[E(\varepsilon_j' \Gamma' M_j \Gamma \varepsilon_j)]^2}.$$

Using part (c) of Lemma A1, we also have

$$\begin{aligned} E(\varepsilon_j' \Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma \varepsilon_j)^2 &= 2\text{tr}(\Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma)^2 + \text{tr}^2(\Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma) \\ &\quad + \Delta \text{tr}(\Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma \circ \Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma) \\ &= 2\text{tr}(B_i B_j)^2 + \text{tr}^2(B_i B_j) + \Delta \text{tr}(\Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma \circ \Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma) \\ &\leq (2 + \Delta) \text{tr}(B_i B_j)^2 + \text{tr}^2(B_i B_j). \end{aligned}$$

With the fact that  $E(\varepsilon_j' \Gamma' M_j \Gamma \varepsilon_j) = \text{tr}(B_j)$ , we obtain

$$E[\text{tr}^2(\Gamma' M_i A_j M_i \Gamma)] \leq \frac{(2 + \Delta) \text{tr}(B_i B_j)^2 + \text{tr}^2(B_i B_j)}{\text{tr}^2(B_j)}.$$

Next, we consider  $E[\text{tr}(\Gamma' M_i A_j M_i \Gamma)^2]$ .

$$\begin{aligned} E[\text{tr}(\Gamma' M_i A_j M_i \Gamma)^2] &= E\left[\left(\frac{\varepsilon_j' \Gamma' M_j M_i \Gamma \Gamma' M_i M_j \Gamma \varepsilon_j}{\varepsilon_j' \Gamma' M_j \Gamma \varepsilon_j}\right)^2\right] \\ &\leq \frac{(2 + \Delta) \text{tr}(B_i B_j)^2 + \text{tr}^2(B_i B_j)}{\text{tr}^2(B_j)}. \end{aligned}$$

Hence,

$$E\left(\hat{\rho}_{ij}^4\right) \leq (3+\Delta) \frac{(2+\Delta) \operatorname{tr}\left(B_i B_j\right)^2+\operatorname{tr}^2\left(B_i B_j\right)}{\operatorname{tr}^2\left(B_i\right) \operatorname{tr}^2\left(B_j\right)}.$$

Consider part (c); since

$$\begin{aligned} E\left(\hat{\rho}_{i j_1}^2 \hat{\rho}_{i j_2}^2\right) &= E E\left(\hat{\rho}_{i j_1}^2 \hat{\rho}_{i j_2}^2 \mid v_i\right)=E\left(E\left(\hat{\rho}_{i j_1}^2 \mid v_i\right) E\left(\hat{\rho}_{i j_2}^2 \mid v_i\right)\right) \\ &= \frac{E\left(v_i' B_{j_1} v_i v_i' B_{j_2} v_i\right)}{\operatorname{tr}\left(B_{j_1}\right) \operatorname{tr}\left(B_{j_2}\right)}. \end{aligned}$$

Note that  $\left|E\left(v_i' B_{j_1} v_i v_i' B_{j_2} v_i\right)\right| \leq\left[E\left(v_i' B_{j_1} v_i\right)^2\right]^{1 / 2}\left[E\left(v_i' B_{j_2} v_i\right)^2\right]^{1 / 2}$  by using the Cauchy–Schwarz inequality and

$$E\left(v_i' B_{j_1} v_i\right)^2=E\left(\frac{\varepsilon_i' \Gamma' M_i M_{j_1} \Gamma \Gamma' M_{j_1} M_i \Gamma \varepsilon_i}{\varepsilon_i' \Gamma' M_i \Gamma \varepsilon_i}\right)^2 \leq \frac{(2+\Delta) \operatorname{tr}\left(B_i B_{j_1}\right)^2+\operatorname{tr}^2\left(B_i B_{j_1}\right)}{\operatorname{tr}^2\left(B_i\right)}.$$

Hence:

$$E\left(\hat{\rho}_{i j_1}^2 \hat{\rho}_{i j_2}^2\right) \leq \frac{\left((2+\Delta) \operatorname{tr}\left(B_i B_{j_1}\right)^2+\operatorname{tr}^2\left(B_i B_{j_1}\right)\right)^{1 / 2}\left((2+\Delta) \operatorname{tr}\left(B_i B_{j_2}\right)^2+\operatorname{tr}^2\left(B_i B_{j_2}\right)\right)^{1 / 2}}{\operatorname{tr}\left(B_{j_1}\right) \operatorname{tr}\left(B_{j_2}\right) \operatorname{tr}^2\left(B_i\right)}.$$

□

**Lemma A3.** Under Assumptions 1–3 and the null in (2), for any fixed positive number  $k$ , we have

- (a)  $\frac{1}{T} \operatorname{tr}\left(\Sigma^k\right)=O(1)$ ;
- (b)  $\frac{1}{T} \operatorname{tr}\left(B_i^k\right)=O(1)$ ;
- (c)  $\frac{1}{T} \operatorname{tr}\left(B_{i_1} B_{i_2} \cdots B_{i_k}\right)=O(1)$ , for  $i_1 \neq i_2 \neq \cdots \neq i_k$ .

**Proof.** Part (a) is directly from Bai and Zhou [20]; hence, we omit it here. Next, we consider part (b). Since  $I_T - P_{X_i}$  is idempotent, for any  $i = 1, \dots, N$ ; hence,  $\operatorname{tr}\left(B_i^k\right)=\operatorname{tr}\left[\left(I_T - P_{X_i}\right) \Sigma\left(I_T - P_{X_i}\right)\right]^k=\operatorname{tr}\left(\left[\left(I_T - P_{X_i}\right) \Sigma\right]^k\right)$ . By using the inequality that for any positive definite matrices  $A$  and  $B$  (see Bushell and Trustring [25])

$$\operatorname{tr}(A B)^k \leq \operatorname{tr}\left(A^k B^k\right),$$

we have

$$\operatorname{tr}\left(B_i^k\right) \leq \operatorname{tr}\left(\left(I_T - P_{X_i}\right) \Sigma^k\right) \leq \operatorname{tr}\left(\Sigma^k\right).$$

Using part (a), then

$$\frac{1}{T} \operatorname{tr}\left(B_i^k\right) \leq \frac{1}{T} \operatorname{tr}\left(\Sigma^k\right)=O(1).$$

For part (c), since for each  $B_{i_l}$ ,  $l = 1, \dots, k$ , it is positive semi-definite. We also have  $B_{i_l} \leq \Sigma$ ,  $l = 1, \dots, k$ . By using the facts that for any matrices  $A, B$ , with  $A \leq B$  and  $C$  positive definite,  $\operatorname{tr}(A C) \leq \operatorname{tr}(B C)$ , we conclude that

$$\frac{1}{T} \operatorname{tr}\left(B_{i_1} B_{i_2} \cdots B_{i_k}\right) \leq \frac{1}{T} \operatorname{tr}\left(\Sigma^k\right)=O(1).$$

Part (c) holds. □



## Appendix B. Proof of the Theorems

### Appendix B.1. Proof of Theorem 1

**Proof.** Since  $E(e_i|X_i) = 0$  and  $\varepsilon_i$ ,  $i = 1, \dots, N$ , are independent, it is easy to show that

$$E(\hat{\rho}_{ij}) = 0,$$

which further implies  $E(T_n) = 0$ . Next, we consider the variance of  $T_n$ .

$$\text{var} \left( \sum_{i=1}^N \sum_{j=1}^{i-1} \hat{\rho}_{ij} \right) = E \left( \sum_{i=1}^N \sum_{j=1}^{i-1} \hat{\rho}_{ij} \right)^2 = E \left( \sum_{i_1=1}^N \sum_{j_1=1}^{i_1-1} \sum_{i_2=1}^N \sum_{j_2=1}^{i_2-1} \hat{\rho}_{i_1 j_1} \hat{\rho}_{i_2 j_2} \right).$$

To calculate the above term, we have three cases to discuss:

- (1)  $i_1, i_2, j_1, j_2$  are mutually different.  $E(\hat{\rho}_{i_1 j_1} \hat{\rho}_{i_2 j_2}) = 0$ .
- (2)  $i_1 = i_2, j_1 = j_2$ . By using Lemma A2, we have  $E(\hat{\rho}_{ij}^2) = \frac{\text{tr}(B_i B_j)}{\text{tr}(B_i) \text{tr}(B_j)}$ .
- (3)  $i_1 = i_2, i_1 \neq j_1 \neq j_2$ . Since  $v_{i_1}, v_{j_1}, v_{i_1}$  and  $v_{j_2}$  are independent, we have  $E(\hat{\rho}_{i_1 j_1} \hat{\rho}_{i_1 j_2}) = E(v_{i_1}' v_{j_1} v_{i_1}' v_{j_2}) = 0$ .

Hence, the above results give us the variance of  $T_n$ , which is

$$\begin{aligned} \gamma^2 = \text{var}(T_n) &= \frac{2}{N(N-1)} \sum_{i=1}^N \sum_{j=1, j \neq i}^N \frac{\text{tr}(M_j \Sigma M_j M_i \Sigma M_i)}{\text{tr}(M_i \Sigma) \text{tr}(M_j \Sigma)} \\ &= \frac{2}{N(N-1)} \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\text{tr}(B_i B_j)}{\text{tr}(B_i) \text{tr}(B_j)}, \end{aligned}$$

and Theorem 1 is proven.  $\square$

### Appendix B.2. Proof of Theorem 2

**Proof.** To prove this theorem, we need to employ the Martingale central limit theorem (Billingsley [26]). For that purpose, we define  $\mathcal{F}_0 = \{\phi, \Omega\}$ ,  $\mathcal{F}_{Ni}$  as the  $\sigma$ -field generated by  $\{\varepsilon_1, \varepsilon_2, \dots, \varepsilon_i\}$  for  $1 \leq i \leq N$ . Let  $E_{N_r}(\cdot)$  denote the conditional expectation given filtration  $\mathcal{F}_{N_r}$  [ $E_0(\cdot) = E(\cdot)$ ]. Write  $L_n = \sum_{i=1}^N D_{Ni}$  with  $D_{N,1} = 0$ . More specifically,

$$D_{Ni} = \left( \frac{1}{N(N-1)} \right)^{1/2} \sum_{j=1}^{i-1} v_i' v_j.$$

For every  $N$ , we can further show that

$$E(D_{Ni} | \mathcal{F}_{N,i-1}) = 0.$$

Hence,  $D_{Ni}$  ( $1 \leq i \leq N$ ) is a martingale difference sequence with respect to  $\mathcal{F}_{Ni}$  ( $1 \leq i \leq N$ ). Let  $\delta_{Ni}^2 = E[(D_{Ni})^2 | \mathcal{F}_{N,i-1}]$ . By applying the Martingale central limit theorem, it is sufficient to show that, as  $(N, T) \rightarrow \infty$ ,

$$\frac{\sum_{i=1}^N \delta_{Ni}^2}{\text{var}(T_n)} \xrightarrow{p} 1 \quad \text{and} \quad \frac{\sum_{i=1}^N E(D_{Ni}^4)}{\text{var}^2(T_n)} \rightarrow 0.$$

Lemmas B1 and B2 prove the above conditions. Hence, we can apply the Martingale central limit theorem, and as  $(N, T) \rightarrow \infty$ , we have

$$\gamma^{-1}T_n \xrightarrow{d} N(0, 1).$$

□

**Lemma B1.** Under Assumptions 1–3 and the null (2), as  $(N, T) \rightarrow \infty$ ,

$$\frac{\sum_{i=1}^N \delta_{Ni}^2}{\text{var}(T_n)} \xrightarrow{p} 1,$$

where  $\delta_{Ni}^2 = E[(D_{Ni})^2 | \mathcal{F}_{N,i-1}]$ .

**Proof.** To prove Lemma B1, we first show that  $E(\sum_{i=1}^N \delta_{Ni}^2) = \text{var}(T_n)$ . Then, we will show that as  $(N, T) \rightarrow \infty$ ,  $\text{var}(\sum_{i=1}^N \delta_{Ni}^2) / \text{var}^2(T_n) \rightarrow 0$ . It is easy to show that

$$E\left(\sum_{i=1}^N \delta_{Ni}^2\right) = \sum_{i=1}^N E\left\{E[(D_{Ni})^2 | \mathcal{F}_{N,i-1}]\right\} = \text{var}(T_n).$$

Next, we only need to show that the second condition is satisfied. We first consider the magnitude of  $\text{var}(T_n)$ . From Lemma A3, we know that

$$\frac{\text{tr}(B_j B_i)}{\text{tr}(B_i) \text{tr}(B_j)} = O(T^{-1}),$$

which implies  $\text{var}^2(T_n) = O(T^{-2})$ . Now, consider  $\text{var}(\sum_{i=1}^N \delta_{Ni}^2)$ . Let  $Q_j = \sum_{j=1}^{i-1} v_j$ , then:

$$\begin{aligned} \delta_{Ni}^2 &= E[(D_{Ni})^2 | \mathcal{F}_{N,i-1}] \\ &= \frac{2}{N(N-1)} E(v_i' Q_j Q_j' v_i | \mathcal{F}_{N,i-1}) \\ &= \frac{2}{N(N-1)} E\left(\frac{\epsilon_i' \Gamma' M_i Q_j Q_j' M_i \Gamma \epsilon_i}{(\epsilon_i' M_i \Gamma' \Gamma M_i \epsilon_i)} | \mathcal{F}_{N,i-1}\right) \\ &= \frac{2}{N(N-1)} \frac{(Q_j' M_i \Gamma' \Gamma M_i Q_j)}{\text{tr}(B_i)}. \end{aligned}$$

Therefore, we need to show the magnitude of  $\text{var}(\sum_{i=1}^N Q_j' M_i \Gamma' \Gamma M_i Q_j)$ . Rewrite  $Q_j' M_i \Gamma' \Gamma M_i Q_j = \sum_{j_1=1}^{i-1} \sum_{j_2=1}^{i-1} v_{j_1}' B_i v_{j_2}$  and:

$$E\left(\sum_{j_1=1}^{i-1} \sum_{j_2=1}^{i-1} v_{j_1}' B_i v_{j_2}\right) = E\left(\sum_{j=1}^{i-1} v_j' B_i v_j\right) = \sum_{j=1}^{i-1} E\left[\frac{\epsilon_j' \Gamma' M_j B_i M_j \Gamma \epsilon_j}{(\epsilon_j' \Gamma' M_j \Gamma \epsilon_j)}\right] = \sum_{j=1}^{i-1} \frac{\text{tr}(B_j B_i)}{\text{tr}(B_j)}.$$

Next, we consider  $E(\sum_{j_1=1}^{i-1} \sum_{j_2=1}^{i-1} v_{j_1}' B_i v_{j_2})^2$ .

$$E\left(\sum_{j_1=1}^{i-1} \sum_{j_2=1}^{i-1} v_{j_1}' B_i v_{j_2}\right)^2 = E\sum_{j_1=1}^{i-1} \sum_{j_2=1}^{i-1} \sum_{j_3=1}^{i-1} \sum_{j_4=1}^{i-1} (v_{j_1}' B_i v_{j_2} v_{j_3}' B_i v_{j_4}).$$

To calculate the magnitude order of the above term, we have three cases to discuss:

$$(1) \ j_1 = j_2 = j_3 = j_4 = j.$$

$$\begin{aligned} E(v'_{j_1} B_i v_j)^2 &= E \frac{(\epsilon'_j \Gamma' M_j B_i M_j \Gamma \epsilon_j)^2}{(\epsilon'_j \Gamma' M_j \Gamma \epsilon_j)^2} = \frac{E(\epsilon'_j \Gamma' M_j B_i M_j \Gamma \epsilon_j)^2}{[E(\epsilon'_j \Gamma' M_j \Gamma \epsilon_j)]^2} \\ &= \frac{\text{tr}^2(B_j B_i) + 2\text{tr}(B_j B_i)^2 + \Delta \text{tr}(B_j B_i \circ B_j B_i)}{\text{tr}^2(B_j)} \leq \frac{(3 + \Delta) \text{tr}^2(B_j B_i)}{\text{tr}^2(B_j)}. \end{aligned}$$

$$(2) \ j_1 = j_2 \neq j_3 = j_4.$$

$$E(v'_{j_1} B_i v_{j_1})(v'_{j_3} B_i v_{j_3}) = E(v'_{j_1} B_i v_{j_1}) E(v'_{j_3} B_i v_{j_3}) = \frac{\text{tr}(B_{j_1} B_i)}{\text{tr}(B_{j_1})} \frac{\text{tr}(B_{j_3} B_i)}{\text{tr}(B_{j_3})}.$$

$$(3) \ j_1 = j_3 \neq j_2 = j_4.$$

$$\begin{aligned} E(v'_{j_1} B_i v_{j_2})(v'_{j_1} B_i v_{j_2}) &= E[E(v'_{j_1} B_i v_{j_2} v'_{j_2} B_i v_{j_1} | v_{j_2})] \\ &= E\left[\frac{\text{tr}(\Gamma' M_{j_1} B_i M_{j_2} \Gamma \epsilon_{j_2} \epsilon'_{j_2} \Gamma' M_{j_2} B_i M_{j_1} \Gamma)}{\text{tr}(M_{j_1} \Sigma) \epsilon'_{j_2} \Gamma' M_{j_2} \Gamma \epsilon_{j_2}}\right] \\ &= \frac{\text{tr}(B_{j_2} B_i B_{j_1} B_i)}{\text{tr}(B_{j_1}) \text{tr}(B_{j_2})}. \end{aligned}$$

Hence,

$$\begin{aligned} \text{var}(Q'_j \Gamma M_i \Gamma' Q_j) &= E(Q'_j \Gamma M_i \Gamma' Q_j)^2 - [E(Q'_j \Gamma M_i \Gamma' Q_j)]^2 \\ &\leq \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2 \neq j_1}^{i-1} \frac{\text{tr}(B_{j_1} B_i) \text{tr}(B_{j_2} B_i)}{\text{tr}(B_{j_1}) \text{tr}(B_{j_2})} + (3 + \Delta) \sum_{j=1}^{i-1} \frac{\text{tr}^2(B_j B_i)}{\text{tr}^2(B_j)} \\ &\quad + 2 \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2 \neq j_1}^{i-1} \frac{\text{tr}(B_{j_2} B_i B_{j_1} B_i)}{\text{tr}(B_{j_1}) \text{tr}(B_{j_2})} - \left(\sum_{j=1}^{i-1} \frac{\text{tr}(B_j B_i)}{\text{tr}(B_j)}\right)^2 \\ &= 2 \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2 \neq j_1}^{i-1} \frac{\text{tr}(B_{j_2} B_i B_{j_1} B_i)}{\text{tr}(B_{j_1}) \text{tr}(B_{j_2})} + (2 + \Delta) \sum_{j=1}^{i-1} \frac{\text{tr}^2(B_j B_i)}{\text{tr}^2(B_j)}. \end{aligned}$$

It further leads to

$$\begin{aligned} \text{var}\left(\sum_{i=1}^N \delta_{Ni}^2\right) &\leq \frac{4}{N^2(N-1)^2} N \sum_{i=1}^N \text{var}(\delta_{Ni}^2) \\ &\leq \frac{8}{N(N-1)^2} \sum_{i=1}^N \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2 \neq j_1}^{i-1} \frac{\text{tr}(B_{j_2} B_i B_{j_1} B_i)}{\text{tr}^2(B_i) \text{tr}(B_{j_1}) \text{tr}(B_{j_2})} \\ &\quad + \frac{4(2 + \Delta)}{N(N-1)^2} \sum_{i=1}^N \sum_{j=1}^{i-1} \frac{\text{tr}^2(B_j B_i)}{\text{tr}^2(B_i) \text{tr}^2(B_j)}. \end{aligned}$$

By using Lemma A3, we have

$$\text{var}\left(\sum_{i=1}^N \delta_{Ni}^2\right) \leq K \left[O\left(\frac{1}{T^3}\right) + O\left(\frac{1}{NT^2}\right)\right].$$

As  $(N, T) \rightarrow \infty$ ,  $\text{var}\left(\sum_{i=1}^N \delta_{Ni}^2\right) / \text{var}^2(T_n) \rightarrow 0$ . Lemma B1 is proven.  $\square$

**Lemma B2.** Under Assumptions 1–3 and the null (2), as  $(N, T) \rightarrow \infty$ ,

$$\frac{\sum_{i=1}^N E\left(D_{Ni}^4\right)}{\text{var}^2(T_n)} \rightarrow 0.$$

**Proof.** Rewrite

$$\begin{aligned} E\left(D_{Ni}^4\right) &= E\left[E\left(D_{Ni}^4 | \mathcal{F}_{N,i-1}\right)\right] = E\left\{E\left[\left(v_i' Q_j Q_j' v_i\right)^2 | \mathcal{F}_{N,i-1}\right]\right\} \\ &= E\left[\frac{\text{tr}^2\left(\Gamma' M_i Q_j Q_j' M_i \Gamma\right) + 2\text{tr}\left(\Gamma' M_i Q_j Q_j' M_i \Gamma\right)^2 + \Delta \text{tr}\left(\Gamma' M_i Q_j Q_j' M_i \Gamma \circ \Gamma' M_i Q_j Q_j' M_i \Gamma\right)}{\text{tr}^2\left(B_i\right)}\right]. \end{aligned}$$

By using the results from Lemma B1, we have

$$\begin{aligned} E\left[\text{tr}^2\left(\Gamma' M_i Q_j Q_j' M_i \Gamma\right)\right] &= E\left(Q_j' B_i Q_j\right)^2 \\ &\leq \sum_{j_1=1}^{i-1} \sum_{j_3=1, j_3 \neq j_1}^{i-1} \frac{\text{tr}\left(B_{j_1} B_i\right) \text{tr}\left(B_{j_3} B_i\right)}{\text{tr}\left(B_{j_1}\right) \text{tr}\left(B_{j_3}\right)} + (3 + \Delta) \sum_{j=1}^{i-1} \frac{\text{tr}^2\left(B_j B_i\right)}{\text{tr}^2\left(B_j\right)} \\ &\quad + 2 \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2 \neq j_1}^{i-1} \frac{\text{tr}\left(B_{j_2} B_i B_{j_1} B_i\right)}{\text{tr}\left(B_{j_1}\right) \text{tr}\left(B_{j_2}\right)}. \end{aligned}$$

Since

$$\text{tr}\left(\Gamma' M_i Q_j Q_j' M_i \Gamma\right)^2 \leq \text{tr}^2\left(\Gamma' M_i Q_j Q_j' M_i \Gamma\right)$$

and

$$\text{tr}\left(\Gamma' M_i Q_j Q_j' M_i \Gamma \circ \Gamma' M_i Q_j Q_j' M_i \Gamma\right) \leq \text{tr}^2\left(\Gamma' M_i Q_j Q_j' M_i \Gamma\right),$$

thus

$$\begin{aligned} \sum_{i=1}^N E\left(D_{Ni}^4\right) &\leq \frac{K}{N^2(N-1)^2} \sum_{i=1}^N \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2 \neq j_1}^{i-1} \frac{\text{tr}\left(B_{j_1} B_i\right) \text{tr}\left(B_{j_2} B_i\right)}{\text{tr}^2\left(B_i\right) \text{tr}\left(B_{j_1}\right) \text{tr}\left(B_{j_2}\right)} \\ &\quad + \frac{K}{N^2(N-1)^2} \sum_{i=1}^N \sum_{j=1}^{i-1} \frac{\text{tr}^2\left(B_j B_i\right)}{\text{tr}^2\left(B_i\right) \text{tr}^2\left(B_j\right)} \\ &\quad + \frac{K}{N^2(N-1)^2} \sum_{i=1}^N \sum_{j_1=1}^{i-1} \sum_{j_2=1, j_2 \neq j_1}^{i-1} \frac{\text{tr}\left(B_{j_2} B_i B_{j_1} B_i\right)}{\text{tr}^2\left(B_i\right) \text{tr}\left(B_{j_1}\right) \text{tr}\left(B_{j_2}\right)} \\ &\leq \frac{K^2}{NT^2} = O\left(\frac{1}{NT^2}\right). \end{aligned}$$

Hence,  $\frac{\sum_{i=1}^N E\left(D_{Ni}^4\right)}{\text{var}^2(T_n)} \rightarrow 0$ , as  $(N, T) \rightarrow \infty$ . Lemma B2 is proven.  $\square$

### Appendix B.3. Proof of Theorem 3

**Proof.** We want to show

$$E(\hat{\gamma}^2) = \gamma^2 \text{ and } \hat{\gamma}^2 - \gamma^2 = o_p(1).$$

Note that

$$\begin{aligned}
\hat{\gamma}^2 &= \frac{1}{2N(N-1)} \sum_{(i,j)} v_i' (v_j - \bar{v}_{(i,j)}) v_j' (v_i - \bar{v}_{(i,j)}) \\
&= \frac{1}{2N(N-1)} \left[ \sum_{(i,j)} (v_i' v_j)^2 - v_i' v_j v_j' \bar{v}_{(i,j)} - v_i' \bar{v}_{(i,j)} v_j' v_i + v_i' \bar{v}_{(i,j)} v_j' \bar{v}_{(i,j)} \right] \\
&= a_1 + a_2 + a_3 + a_4, \text{ say.}
\end{aligned}$$

It is easy to show that the first term  $E(a_1) = \gamma^2$ , and  $E(a_i) = 0, i = 2, 3, 4$ . Therefore, we prove the first part. By using Lemma A3 and Theorem 1, we have  $\gamma^2 = O(T^{-1})$ . Hence, to prove  $\hat{\gamma}^2 - \gamma^2 = o_p(1)$ , we only need to show that  $\text{var}(a_1) = o_p(T^{-2})$  and  $a_i = o_p(\gamma^2)$ , for  $i = 2, 3, 4$ . Let us consider  $\text{var}(a_1)$ .

$$\begin{aligned}
\text{var}(a_1) &= E(a_1^2) - \gamma^4 \\
&= \frac{4}{N^2(N-1)^2} E \left( \sum_{i=1}^N \sum_{j=1}^{i-1} \hat{\rho}_{ij}^2 \right)^2 - \frac{4}{N^2(N-1)^2} \left( \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\text{tr}(B_i B_j)}{\text{tr}(B_i) \text{tr}(B_j)} \right)^2 \\
&= \frac{4}{N^2(N-1)^2} E \left( \sum_{i_1=2}^N \sum_{j_1=1}^{i_1-1} \sum_{i_2=2}^N \sum_{j_2=1}^{i_2-1} \rho_{i_1 j_1}^2 \rho_{i_2 j_2}^2 \right) - \frac{4}{N^2(N-1)^2} \left( \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\text{tr}(B_i B_j)}{\text{tr}(B_i) \text{tr}(B_j)} \right)^2.
\end{aligned}$$

Now, we only consider the term  $E \left( \sum_{i_1=2}^N \sum_{j_1=1}^{i_1-1} \sum_{i_2=2}^N \sum_{j_2=1}^{i_2-1} \rho_{i_1 j_1}^2 \rho_{i_2 j_2}^2 \right)$ . There are three cases for this term, and Lemma A2 is used frequently:

(1)  $i_1, i_2, j_1$  and  $j_2$  are mutually different.

$$E \left( \rho_{i_1 j_1}^2 \rho_{i_2 j_2}^2 \right) = \frac{\text{tr}(B_{i_1} B_{j_1}) \text{tr}(B_{i_2} B_{j_2})}{\text{tr}(B_{i_1}) \text{tr}(B_{j_1}) \text{tr}(B_{i_2}) \text{tr}(B_{j_2})} = O_p \left( \frac{1}{T^2} \right).$$

(2)  $i_1 = i_2, j_1 = j_2$  and  $i_1 \neq j_1$ .

$$E \left( \rho_{ij}^4 \right) \leq (3 + \Delta) \frac{(2 + \Delta) \text{tr}(B_i B_j)^2 + \text{tr}^2(B_i B_j)}{\text{tr}^2(B_i) \text{tr}^2(B_j)} = O_p \left( \frac{1}{T^2} \right).$$

(3)  $i_1 = i_2, i_1 \neq j_1 \neq j_2$ .

$$\begin{aligned}
E \left( \rho_{i j_1}^2 \rho_{i j_2}^2 \right) &\leq \frac{\left( (2 + \Delta) \text{tr}(B_i B_{j_1})^2 + \text{tr}^2(B_i B_{j_1}) \right)^{1/2} \left( (2 + \Delta) \text{tr}(B_i B_{j_2})^2 + \text{tr}^2(B_i B_{j_2}) \right)^{1/2}}{\text{tr}(B_{j_1}) \text{tr}(B_{j_2}) \text{tr}^2(B_i)} \\
&= O_p \left( \frac{1}{T^2} \right).
\end{aligned}$$

From the above results, we have

$$\text{var}(a_1) = O_p \left( \frac{1}{N^2 T^2} \right).$$

Hence  $a_1 \xrightarrow{p} \gamma^2$ . Consider the second term  $a_2$ , which is equal to  $\frac{1}{2N(N-1)(N-2)} \sum_{(i,j,\tau)}^N v'_i v_j v'_j v_\tau$ . The first term of  $E \left( \sum_{(i,j,\tau)}^N v'_i v_j v'_j v_\tau \right)^2$  is

$$\begin{aligned} & \sum_{(i,j_1,j_2,\tau)}^N E \left( v'_i v_{j_1} v'_{j_1} v_\tau v'_i v_{j_2} v'_{j_2} v_\tau \right) \\ &= \sum_{(i,j_1,j_2,\tau)}^N \frac{\text{tr}(M_{j_2} M_\tau \Sigma M_\tau M_i \Sigma M_i M_{j_1} \Sigma M_{j_1} M_{j_2} \Sigma)}{\text{tr}(B_\tau) \text{tr}(B_{j_2}) \text{tr}(B_{j_1}) \text{tr}(B_i)} \\ &= O(N^4 T^{-3}), \end{aligned}$$

by using Lemmas A2 and A3. By using part (c) of Lemma A3, the second term of  $E \left( \sum_{(i,j,\tau)}^N v'_i v_j v'_j v_\tau \right)^2$  is

$$E \left[ \sum_{(i,j,\tau)}^N \left( v'_i v_j v'_j v_\tau \right)^2 \right] = O_p(N^3 T^{-2}).$$

Hence,  $a_2 = O_p(N^{-1} T^{-3/2}) + O_p(N^{-3/2} T^{-1})$ , which further implies  $a_2 = o_p(\gamma^2)$ . Since  $a_2 = a_3$ ,  $a_3 = o_p(\gamma^2)$ . Consider  $a_4$ ; it can be divided into two terms

$$\frac{1}{2N(N-1)(N-2)^2} \sum_{(i,j,\tau)}^N \left( v'_i v_\tau v'_j v_\tau \right) \text{ and } \frac{1}{2N(N-1)(N-2)^2} \sum_{(i,j,\tau_1,\tau_2)}^N \left( v'_i v_{\tau_1} v'_j v_{\tau_2} \right).$$

It is easy to show that the former term is  $O_p(N^{-1} a_2)$ , then it is  $o_p(\gamma^2)$ . We only need to consider the latter term  $E \left( \sum_{(i,j,\tau_1,\tau_2)}^N \left( v'_i v_{\tau_1} v'_j v_{\tau_2} \right) \right)^2$ .

$$E \left[ \sum_{(i,j,\tau_1,\tau_2)}^N \left( v'_i v_{\tau_1} v'_j v_{\tau_2} \right)^2 \right] = \sum_{(i,j,\tau_1,\tau_2)}^N E \left[ \left( v'_i v_{\tau_1} \right)^2 \left( v'_j v_{\tau_2} \right)^2 \right] = O(N^4 T^{-2}),$$

by using Lemma A2–A3. Hence, the latter term is  $O_p(N^{-2} T^{-1})$ . The above results together lead to  $a_4 = o_p(\gamma^2)$ . The first part of Theorem 3 holds; the second part of Theorem 3 is directly derived by using Theorem 2 and the first part of Theorem 3.  $\square$

#### Appendix B.4. Proof of Theorem 4

**Proof.** The OLS residuals under the local alternative are defined by  $M_i u'_i = \sigma_i (M_i \Gamma \varepsilon_i + M_i F \lambda_i)$ , thus

$$\begin{aligned} T_n &= \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{(M_i \Gamma \varepsilon_i + M_i F \lambda_i)' (M_j \Gamma \varepsilon_j + M_j F \lambda_j)}{\|M_i \Gamma \varepsilon_i + M_i F \lambda_i\| \|M_j \Gamma \varepsilon_j + M_j F \lambda_j\|} \\ &= \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{(M_i \Gamma \varepsilon_i + M_i F \lambda_i)' (M_j \Gamma \varepsilon_j + M_j F \lambda_j)}{(\varepsilon'_i \Gamma' M_i \Gamma \varepsilon_i + 2\varepsilon'_i \Gamma' M_i F \lambda_i + \lambda'_i F' M_i F \lambda_i)^{1/2} (\varepsilon'_j \Gamma' M_j \Gamma \varepsilon_j + 2\varepsilon'_j \Gamma' M_j F \lambda_j + \lambda'_j F' M_j F \lambda_j)^{1/2}}. \end{aligned}$$

Consider the denominator. Note that  $E(\varepsilon'_i \Gamma' M_i \Gamma \varepsilon_i)^2 = \text{tr}(\Sigma M_i)^2 + 2[\text{tr}(\Sigma M_i)]^2 + \Delta \text{tr}(\Sigma M_i \circ \Sigma M_i) = O_p(T^2)$ , which lead to  $\varepsilon'_i \Gamma' M_i \Gamma \varepsilon_i = O(T)$ . Consider the term  $\varepsilon'_i \Gamma' M_i F \lambda_i$ . Since

$$\varepsilon'_i \Gamma' M_i F \lambda_i = \varepsilon'_i \Gamma' F \lambda_i - \varepsilon'_i \Gamma' X_i (X'_i X_i)^{-1} X'_i F \lambda_i.$$

From Assumption 4, we have  $X_i'F = O_p(T^{1/2})$ ,  $\varepsilon_i'\Gamma'F = O_p(T^{1/2})$  and  $\varepsilon_i'\Gamma'X_i = O_p(T^{1/2})$ , which lead to  $\|\varepsilon_i'\Gamma'F\lambda_i\| = O_p(T^{1/4}N^{-1/2})$  and  $\|\varepsilon_i'\Gamma'X_i(X_i'X_i)^{-1}X_i'F\lambda_i\| = O_p(T^{-1/4}N^{-1/2})$ . Hence,  $\varepsilon_i'\Gamma'M_iF\lambda_i = o_p(\varepsilon_i'\Gamma'M_i\Gamma\varepsilon_i)$ . Similarly, by using Assumption 4, we also have  $\lambda_i'F'M_iF\lambda_i = o_p(\lambda_i'F'M_i\Gamma\varepsilon_i)$ . From the above results, we further have

$$\varepsilon_i'\Gamma'M_i\Gamma\varepsilon_i + 2\varepsilon_i'\Gamma'M_iF\lambda_i + \lambda_i'F'M_iF\lambda_i = (1 + o_p(1))\varepsilon_i'\Gamma'M_i\Gamma\varepsilon_i.$$

It results in

$$T_n = \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\varepsilon_i'\Gamma'M_iM_j\Gamma\varepsilon_j + \varepsilon_i'\Gamma'M_iM_jF\lambda_j + \lambda_i'F'M_iM_j\Gamma\varepsilon_j + \lambda_i'F'M_iM_jF\lambda_j}{((1 + o_p(1))\varepsilon_i'\Gamma'M_i\Gamma\varepsilon_i)^{1/2} ((1 + o_p(1))\varepsilon_j'\Gamma'M_j\Gamma\varepsilon_j)^{1/2}} \\ = T_{n1} + T_{n2} + T_{n3} + T_{n4},$$

where  $T_{n1} = \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\varepsilon_i'\Gamma'M_iM_j\Gamma\varepsilon_j}{D_{ij}}$ ;  $T_{n2} = \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\varepsilon_i'\Gamma'M_iM_jF\lambda_j}{D_{ij}}$ ;  $T_{n3} = \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\lambda_i'F'M_iM_j\Gamma\varepsilon_j}{D_{ij}}$  and  $T_{n4} = \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\lambda_i'F'M_iM_jF\lambda_j}{D_{ij}}$  with  $D_{ij} = ((1 + o(1))\varepsilon_i'\Gamma'M_i\Gamma\varepsilon_i)^{1/2} ((1 + o(1))\varepsilon_j'\Gamma'M_j\Gamma\varepsilon_j)^{1/2}$ . From Theorem 2,  $\gamma^{-1}T_{n1} \xrightarrow{d} N(0, 1)$ .

From Theorem 1,  $T_{n1} = O_p(T^{-1/2})$ . Consider  $T_{n2}$ . We observe that  $E(T_{n2}) = 0$  and

$$E(T_{n2})^2 = \frac{2}{N(N-1)} E \left( \sum_{i=2}^N \sum_{j=1}^{i-1} \frac{\varepsilon_i'\Gamma'M_iM_jF\lambda_j}{D_{ij}} \right)^2 \\ = \frac{2}{N(N-1)} \left[ \sum_{i=2}^N \sum_{j=1}^{i-1} E \left( \frac{\varepsilon_i'\Gamma'M_iM_jF\lambda_j}{D_{ij}} \right)^2 + \sum_{i=2}^N \sum_{j_1=1}^{i-1} \sum_{j_2 \neq j_1}^{i-1} E \frac{\varepsilon_i'\Gamma'M_iM_{j_1}F\lambda_{j_1} \lambda_{j_2}'F'M_{j_2}M_i\Gamma\varepsilon_i}{D_{ij_1}D_{ij_2}} \right].$$

Consider the term  $\varepsilon_i'\Gamma'M_iM_jF\lambda_j$ .

$$\varepsilon_i'\Gamma'M_iM_jF\lambda_j \\ = \varepsilon_i'\Gamma'F\lambda_j - \varepsilon_i'\Gamma'X_i(X_i'X_i)^{-1}X_i'F\lambda_j - \varepsilon_i'\Gamma'X_j(X_j'X_j)^{-1}X_j'F\lambda_j + \varepsilon_i'\Gamma'X_i(X_i'X_i)^{-1}X_i'X_j(X_j'X_j)^{-1}X_j'F\lambda_j.$$

Using Assumption 4 and under the local alternative, we first have  $\|\varepsilon_i'\Gamma'F\lambda_j\| = O_p(T^{1/4}N^{-1/2})$ ; we then have  $\|\varepsilon_i'\Gamma'X_i(X_i'X_i)^{-1}X_i'F\lambda_j\| = O_p(T^{-1/4}N^{-1/2})$  since

$$\varepsilon_i'\Gamma'X_i(X_i'X_i)^{-1}X_i'F = \left( \frac{\varepsilon_i'\Gamma'X_i}{\sqrt{T}} \right) \left( \frac{X_i'X_i}{T} \right)^{-1} \left( \frac{X_i'F}{\sqrt{T}} \right) = O_p(1);$$

we last have  $\|\varepsilon_i'\Gamma'X_i(X_i'X_i)^{-1}X_i'X_j(X_j'X_j)^{-1}X_j'F\lambda_j\| = O_p(T^{-1/4}N^{-1/2})$ . Hence,  $\|\varepsilon_i'\Gamma'M_iM_jF\lambda_j\| = O_p(T^{1/4}N^{-1/2})$ . Together with the fact that  $\|D_{ij}\| = O_p(T)$ , the first term of  $E(T_{n2})^2$  is of order  $O_p(T^{-3/2}N^{-1})$ . Similar to the proof of above,  $\|\varepsilon_i'\Gamma'M_iM_{j_1}F\lambda_{j_1} \lambda_{j_2}'F'M_{j_2}M_i\Gamma\varepsilon_i\| = O_p(T^{1/2}N^{-1})$ ; with the facts that  $\|D_{ij_1}\| = O_p(T)$  and  $\|D_{ij_2}\| = O_p(T)$ ; the second term of  $E(T_{n2})^2$  is of order  $O_p(T^{-3/2})$ . Thus,  $T_{n2} = O_p(T^{-3/4}) = o_p(T_{n1})$ . Similarly,  $T_{n3} = o_p(T_{n1})$ .

Consider  $T_{n4}$ . Note that

$$\lambda_i'F'M_iM_jF\lambda_j \\ = \lambda_i'F'F\lambda_j - \lambda_i'F'X_i(X_i'X_i)^{-1}X_i'F\lambda_j - \lambda_i'F'X_j(X_j'X_j)^{-1}X_j'F\lambda_j + \lambda_i'F'X_i(X_i'X_i)^{-1}X_i'X_j(X_j'X_j)^{-1}X_j'F\lambda_j.$$



From Assumption 4, we know that  $\lambda_i F' F \lambda_j = \lambda_i' T (I_r + O_p(T^{-1/2})) \lambda_j \xrightarrow{p} T \lambda_i' \lambda_j$ . Since

$$F' X_i (X_i' X_i)^{-1} X_i' F = \left( \frac{F' X_i}{\sqrt{T}} \right) \left( \frac{X_i' X_i}{T} \right)^{-1} \left( \frac{X_i' F}{\sqrt{T}} \right) = O_p(1),$$

$\lambda_i' F' X_i (X_i' X_i)^{-1} X_i' F \lambda_j = o_p(\lambda_i' F' F \lambda_j)$ . Similarly, we can also show that the third and the fourth terms are of smaller order of the first term. Hence,  $\lambda_i' F' M_i M_j F \lambda_j = (1 + o_p(1)) \lambda_i' F' F \lambda_j$ .

Note that  $E(\lambda_i' F' F \lambda_j) = T \lambda_i' \lambda_j \neq 0$ , and  $\left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^N \frac{\lambda_i' F' F \lambda_j}{D_{ij}} = O_p(T^{-1/2})$ ; hence,  $\gamma^{-1} T_{n4} = O_p(1)$ . One can also show that  $D_{ij} \xrightarrow{p} \text{tr}^{1/2}(M_i \Sigma) \text{tr}^{1/2}(M_j \Sigma)$ . Let  $\psi = \text{plim}_{(N,T) \rightarrow \infty} \gamma^{-1} \left( \frac{2}{N(N-1)} \right)^{1/2} \sum_{i=2}^N \sum_{j=1}^{i-1} \left( \frac{T^{1/2} N^{-1} \delta_i' \delta_j}{\text{tr}^{1/2}(M_i \Sigma) \text{tr}^{1/2}(M_j \Sigma)} \right)$ ; from all of the above results, as  $(N, T) \rightarrow \infty$ ,

$$\gamma^{-1} T_{n1} - \psi \xrightarrow{d} N(0, 1).$$

□

## References

1. Lee, L. Consistency and Efficiency of Least Squares Estimation for Mixed Regressive, Spatial Autoregressive Models. *Econom. Theory* **2002**, *18*, 252–277.
2. Andrews, D.W.K. Cross-Section Regression with Common Shocks. *Econometrica* **2005**, *73*, 1551–1585.
3. Anselin, L.; Bera, A.K. Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In *Handbook of Applied Economic Statistics*; Ullah, A., Giles, D.E., Eds.; Marcel Dekker: New York, NY, USA, 1998; pp. 237–289.
4. Baltagi, B.H.; Song, S.H.; Koh, W. Testing Panel Data Regression Models with Spatial Error Correlation. *J. Econom.* **2003**, *117*, 123–150.
5. Breusch, T.S.; Pagan, A.R. The Lagrange Multiplier Test and Its Application to Model Specifications in Econometrics. *Rev. Econ. Stud.* **1980**, *47*, 239–253.
6. Johnstone, I.M. On the Distribution of the Largest Eigenvalue in Principal Components Analysis. *Ann. Stat.* **2001**, *29*, 295–327.
7. Jiang, T.F. The Limiting Distributions of Eigenvalues of Sample Correlation Matrices. *Sankhyā* **2004**, *66*, 35–48.
8. Ledoit, O.; Wolf, M. Some Hypothesis Tests for the Covariance Matrix When the Dimension is Large Compared to the Sample Size. *Ann. Stat.* **2002**, *41*, 1055–1692.
9. Schott, J.R. Testing for Complete Independence in High Dimensions. *Biometrika* **2005**, *92*, 951–956.
10. Chen, S.X.; Zhang, L.X.; Zhong, P.S. Tests for High Dimensional Covariance Matrices. *J. Am. Stat. Assoc.* **2010**, *105*, 810–819.
11. Pesaran, M.H.; Ullah, A.; Yamagata, T. A Bias-Adjusted LM Test of Error Cross-Section Independence. *Econom. J.* **2008**, *11*, 105–127.
12. Baltagi, B.H.; Feng, Q.; Kao, C. A Lagrange Multiplier Test for Cross-Sectional Dependence in a Fixed Effects Panel Data Model. *J. Econom.* **2012**, *170*, 164–177.
13. Baltagi, B.H.; Feng, Q.; Kao, C. Testing for Sphericity in a Fixed Effects Panel Data Model. *Econom. J.* **2011**, *14*, 25–47.
14. Pesaran, M.H. General Diagnostic Test for Cross Section Dependence in Panels. CESifo Working Paper Series No. 1229, IZA Discussion Paper No. 1240. Available online: <http://ssrn.com/abstract=572504> (accessed on 2 May 2015).
15. Pesaran, M.H. Testing Weak Cross-Sectional Dependence in Large Panels. *Econom. Rev.* **2015**, *34*, 1089–1117.
16. Moscone, F.; Tosetti, E. A Review and Comparisons of Tests of Cross-Section Dependence in Panels. *J. Econ. Surv.* **2009**, *23*, 528–561.
17. Sarafidis, V.; Wansbeek, T. Cross-Sectional Dependence in Panel Data Analysis. *Econom. Rev.* **2012**, *31*, 483–531.

18. Chudik, A.; Pesaran, M.H. Large Panel Data Models with Cross-Sectional Dependence: A Survey. In *The Oxford Handbook on Panel Data*; Baltagi, B.H., Ed.; Oxford University Press: Oxford, UK, 2015; Chapter 1; pp. 3–45.
19. Bai, Z.D.; Silverstein, J.W. CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices. *Ann. Probab.* **2004**, *32*, 553–605.
20. Bai, Z.D.; Zhou, W. Large Sample Covariance Matrices without Independence Structures in Columns. *Stat. Sin.* **2008**, *18*, 425–442.
21. Gao, J.T.; Han, X.; Pan, G.M.; Yang, Y.R. High Dimensional Correlation Matrices: CLT and Its Applications. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **2016**, doi: 10.1111/rssb.12189.
22. Muirhead, R.J. *Aspects of Multivariate Statistical Theory*; John Wiley & Sons: Hoboken, NJ, USA, 1982.
23. Chen, S.X.; Qin, Y.L. A Two-Sample Test for High Dimensional Data with Application to Gene-Set Testing. *Ann. Stat.* **2010**, *38*, 808–835.
24. Lieberman, O. A Laplace Approximation to the Moments of a Ratio of Quadratic Forms. *Biometrika* **1994**, *81*, 681–690.
25. Bushell, P.J.; Trustrum, G.B. Trace Inequality for Positive Definite Matrix Power Products. *Linear Algebra Appl.* **1990**, *132*, 173–178.
26. Billingsley, P. *Probability and Measure*, 3rd ed.; Wiley: New York, NY, USA, 1995.



© 2016 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).