A RMT-based LM test for error cross-sectional independence in large heterogeneous panel data models

Natalia Bailey\textsuperscript{a}, Dandan Jiang\textsuperscript{b}, and Jianfeng Yao\textsuperscript{c}

\textsuperscript{a}Department of Econometrics and Business Statistics, Monash University, Melbourne, Victoria, Australia; \textsuperscript{b}School of Mathematics and Statistics, Xi’an Jiaotong University, Xi’an, Shaanxi, China; \textsuperscript{c}Department of Statistics and Actuarial Science, The University of Hong Kong, Pok Fu Lam, Hong Kong

ABSTRACT
This paper introduces a new test for error cross-sectional independence in large panel data models with exogenous regressors having heterogenous slope coefficients. The proposed statistic, $LM_{RMT}$, is based on the Lagrange Multiplier (LM) principle and the sample correlation matrix $R_N$ of the model’s residuals. Since in large panels $R_N$ poorly estimates its population counterpart, results from Random Matrix Theory (RMT) are used to establish the high-dimensional limiting distribution of $LM_{RMT}$ under heteroskedastic normal errors and assuming that both the panel size $N$ and the sample size $T$ grow to infinity in comparable magnitude. Simulation results show that $LM_{RMT}$ is largely correctly sized (except for some small values of $N$ and $T$). Further, the empirical size and power outcomes show robustness of our statistic to deviations from the assumptions of normality for the error terms and of strict exogeneity for the regressors. The test has comparable small sample properties to related tests in the literature which have been developed under different asymptotic theory.

KEYWORDS
Large panels; Cross-sectional independence; High-dimensional Lagrange-Multiplier test; Random Matrix theory

JEL CLASSIFICATIONS
C12; C13; C33

1. Introduction
Analysis of possible cross-sectional dependence (CD) in panel data models has received renewed attention in the past decade, mainly due to the advent of increasingly large data sets that describe rich interactions among economic agents at local, regional and global level. This paper focuses on the problem of testing for error cross-sectional independence when studying a panel data specification. Two main strands of literature have been developed to address this issue. Early contributions imposed an \textit{a priori} ordering of cross section units when formulating such tests. For example, Moran’s test (Moran, 1948) and further extensions discussed in the spatial econometrics literature such as Cliff and Ord (1973, 1981), Anselin (1988, 2001) and more recently Conley and Molinari (2007), depend on the choice of spatial adjacency matrix which characterizes the interconnections between these ordered units.

The second line of literature abstracts from firm assumptions about the form of the error cross section structure and makes use of the Lagrange Multiplier (LM) principle. The Breusch and Pagan (1980) LM test is an early example and is expressed as the average of squared pair-wise correlation coefficients between the errors from a panel data model with cross section dimension $N$ and time dimension $T$, where $N$ is assumed to be considerably smaller than $T$ ($T \gg N$).

CONTACT Jianfeng Yao \textsuperscript{c} jeffyao@hku.hk Department of Statistics and Actuarial Science, The University of Hong Kong, Pok Fu Lam, Hong Kong.

\textsuperscript{c}The authors would like to acknowledge helpful comments and suggestions from the Editor in charge of our paper and three anonymous reviewers. N. Bailey and D. Jiang are joint first authors with equal contribution.

\textsuperscript{b}Supplemental data for this article is available online at https://doi.org/10.1080/07474938.2021.2009705

© 2021 Taylor & Francis Group, LLC
However, in the current era of 'big data', panel data sets can easily comprise of thousands of cross section units with a considerable number of observations each, thus invalidating the above assumption of $T \gg N$. Examples include large data sets on household income or consumption behavior, house price dynamics across neighborhoods or districts, production patterns of firms within and across industries. In settings where $N$ is larger than $T$, the original $LM$ test is biased and suffers from severe size distortions. Subsequent contributions have attempted to correct for this bias.

First, Frees (1995) provides a distribution-free statistic that does not suffer from this drawback but imposes limitations on the number of regressors that can be included in the panel data models that he considers. Next, Pesaran et al. (2008) compute exact mean and variance expressions associated with the $LM$ statistic for the purpose of bias-adjustment, thus obtaining asymptotic normality for their modified $LM_{adj}$ test. For this, the sequential asymptotic where $T \to \infty$ first, followed by $N \to \infty$ is required. In order to improve on the sequential asymptotics of $LM_{adj}$, Pesaran (2004, 2015a)\(^1\) then develops the CD test which provides favorable performance when $N > T$. While Pesaran (2004) assumes error independence as null hypothesis, the Pesaran (2015a) refinement sets as implicit null of the CD test weak cross-sectional dependence, as defined in Chudik et al. (2011) and further developed in Bailey et al. (2016, 2019). This null depends on the relative expansion rates of $N$ and $T$. More recently, Juodis and Reese (2021) discuss bias correction of the CD statistic in panel data models when they incorporate period-specific parameters.

Extensions of testing for error cross-sectional dependence to dynamic panel data models are more limited. Examples include the GMM approach of Sarafidis et al. (2009) applied to panels with homogeneous slope coefficients, and the LM test of Halunga et al. (2017) which studies heterogeneous panels and is robust to time series heteroskedasticity, but requires $N^2/T \to 0$ when $N, T \to \infty$ in their asymptotic theory. For a comprehensive survey and comparisons of tests for cross-sectional dependence in panel data models see Moscone and Tosetti (2009), Sarafidis and Wansbeek (2012) or Pesaran (2015b).

Implementation of the original $LM$ test and its variants discussed above, make use of the sample correlation matrix of the residuals obtained from regressing the panel data model, $R_N=(\hat{\rho}_{ij}), i,j=1,2,...,N$. As is well known from multivariate analysis, when $N \ll T$, $R_N$ is a consistent estimator of the population correlation matrix $R_N = (\rho_{ij})$. However, as the cross section dimension $N$ increases and eventually surpasses $T$, this property no longer holds. This is reasonable since the error in estimating each individual $\hat{\rho}_{ij}$ is of order $1/T$ which multiplies as the number of units $N$ rises and can no longer be corrected with a larger time dimension once $N \ll T$ does not hold anymore. This result is established in the Random Matrix Theory (RMT) literature, which also suggests a different concept of consistency for $R_N$ to account for change in dimensionality. RMT treats $R_N$ as a random matrix and uses the spectral and/or Frobenius norms as more appropriate metrics to achieve convergence - see for example Paul and Aue (2014), and Yao et al. (2015).

In this paper, we introduce a new $LM$ based statistic for testing for error cross-sectional independence which we denote by $LM_{RMT}$, and rely on asymptotic results from the RMT literature to establish its high-dimensional limiting distribution. It assumes that both $T$ and $N = N(T) \to \infty$ and $\lim N/T \to c > 0$, and is applied to a heterogeneous panel data model with strictly exogenous regressors, as well as normally distributed and heteroskedastic errors. Normality of the error terms is fundamental in the theoretical derivation of the statistic. To the best of our knowledge, the only other paper that utilizes random matrix theory to test for sphericity (John test) under the more restrictive homogeneous panel data setting is Baltagi et al. (2011). This contribution uses different technical derivations to our proposed statistic to obtain its limiting distribution. The follow-up $LM_{bc}$ test of Baltagi et al. (2012) for error cross-sectional dependence under the

\(^1\)For a published version of Pesaran (2004) see Pesaran (2021).
same panel setting is closely related to the test statistic proposed by Schott (2005) and its theory does not rely on RMT.

Our Monte Carlo simulation results suggest that our $LM_{RMT}$ statistic is correctly sized when both $N$ and $T$ are large, though some size distortions are evident for small values of $N$ and $T$. It performs favorably to the original $LM$ and $LM_{bc}$ tests, the latter in the case of slope heterogeneity, and comparably to $LM_{adj}$ in terms of size and power. The CD test is universally correctly sized but lacks power under local alternatives. Finally, small sample results suggest that $LM_{RMT}$ is robust to non-normality in the error terms, and inclusion of temporal dynamics.

The rest of the paper is organized as follows: Section 2 sets up the model and discusses existing Lagrange Multiplier statistics that test for error cross-sectional independence. Section 3 introduces our LM test, $LM_{RMT}$, and presents the main theoretical results for the statistic for a panel data model with large cross section and time dimensions. Section 4 presents a detailed Monte Carlo simulation study. Additional proofs of the lemma and theorem are provided in the Appendices.

2. The Lagrange multiplier test and its variant

Consider a panel data model with $N$ units which are observed during times $t = 1, 2, \ldots, T$. For each panel unit $1 \leq i \leq N$, the response variable $y$ obeys a regression model of type

$$y_{it} = \beta_i'x_{it} + u_{it}, \quad 1 \leq t \leq T. \quad (1)$$

Here $x_{it} = (x_{iti})_{1 \leq i \leq k}$ is a $k \times 1$ vector of time-varying covariates, and $\beta_i$ is a $k \times 1$ vector of regression parameters for unit $i$. An individual-specific intercept can be included by setting the first element of $x_{it}$ to unity. Finally, $u_i = (u_{iti})_{1 \leq t \leq T}$ is a sequence of regression error which is assumed to be independent and identically distributed (IID). Throughout the paper we will use the notation $T_k = T - k$.

Consider the null hypothesis of cross-sectional independence

$$H_0 : \text{the } N \text{ error vectors } u_1, \ldots, u_N \text{ are mutually independent.} \quad (2)$$

This implies in particular that $\text{Cov}(u_{it}, u_{jt}) = 0$, for all $t$ and $i \neq j$. To test this hypothesis, Breusch and Pagan (1980) proposed a Lagrange Multiplier (LM) statistic, as follows. For each unit $i$, first a regression model is fitted by ordinary least squares (OLS) producing residuals

$$e_{it} = y_{it} - \hat{\beta}_i'x_{it}, \quad 1 \leq t \leq T. \quad (3)$$

Here $\hat{\beta}_i$ denotes the vector of OLS estimates of regression parameters. Next, the empirical correlations between these residuals associated with the $N$ units are obtained:

$$\hat{\rho}_{ij} = \frac{\sum_{t=1}^{T} e_{it} e_{jt}}{\left( \sum_{t=1}^{T} e_{it}^2 \right)^{1/2} \left( \sum_{t=1}^{T} e_{jt}^2 \right)^{1/2}}, \quad 1 \leq i, j \leq N. \quad (4)$$

Breusch and Pagan’s LM statistic is given by

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

A large value of LM indicates significance of the test. By assuming that $T \rightarrow \infty$ while the number of units $N$ is held fixed, LM is proved to have an asymptotic Chi-squared distribution with $N(N-1)/2$ degrees of freedom under the null (and mild conditions on the model).

As explained in the introduction, the LM test suffers important bias distortions when $T \gg N$ does not hold: this happens for instance either when $N$ is large or when the sample size $T$ is
small. The situation where the ratio $N/T$ is not close enough to zero is frequent in real data analysis and we refer to it as a high-dimensional setting. In an attempt to correct for such bias, Pesaran et al. (2008) proposed a bias-adjusted LM test. Precisely, their mean-variance-bias-adjusted LM test statistic is

$$LM_{adj} = \sqrt{\frac{2}{N(N-1)} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \frac{T_k \hat{\rho}_{ij}^2 - \mu_{Tij}}{\nu_{Tij}}}$$

where $\mu_{Tij}$ and $\nu_{Tij}$ are the mean and variance of $T_k \hat{\rho}_{ij}^2$ given by

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

and

$$\nu_{Tij} = \text{Var}(T_k \hat{\rho}_{ij}^2) = \left\{ \text{tr}[E(M_i M_j)] \right\}^2 a_{1T} + 2\text{tr}\left\{ E\left((M_i M_j)^2\right) \right\} a_{2T},$$

with $M_i = I_T - X_i(X_i'X_i)^{-1}X_i'$, $X_i = (x_{i1},...,x_{iT})'$, and

$$a_{1T} = a_{2T} = \frac{1}{T_k}, \quad a_{2T} = \frac{3}{(T_k + 2)^2},$$

respectively, as shown in Theorem 1 in Pesaran et al. (2008) (PUY). Under the null hypothesis and considering a sequential “$T$-then-$N$” limiting scheme where first $T \to \infty$ and then $N \to \infty$, they establish that the $LM_{adj}$ statistic is asymptotically distributed as standard normal. The main issue here is that such limiting distribution derived under the sequential “$T$-then-$N$” scheme may not be accurate enough in the high-dimensional setting considered in this paper.

3. High-dimensional asymptotics of the Lagrange multiplier test

It is apparent that the LM statistic (5) is connected to some $N \times N$ sample correlation matrix. More precisely, assume that within each unit $i$, the errors $u_{it}$ are IID$(0,\sigma^2)$ distributed, or equivalently $u_{it} = \sigma_i \epsilon_{it}$ with IID$(0,1)$ random variables $\epsilon_{it}$ ($1 \leq t \leq T$). Collect them into a $T \times 1$ vector $\epsilon_i = (\epsilon_{i1},...,\epsilon_{iT})'$ and define the projection matrix $M_i = I_T - X_i(X_i'X_i)^{-1}X_i'$, where $X_i = (x_{i1},...,x_{iT})'$ is the $T \times k$ design matrix for unit $i$. Then the vector of OLS residuals $e_i = (e_{i1},...,e_{iT})'$ in (3) is

$$e_i = \sigma_i M_i \epsilon_i.$$  

Next, the empirical cross-sectional correlations (4) can be expressed as $\hat{\rho}_{ij} = e_i'e_j/\|e_i\| \|e_j\|$, where $\|e_i\|$ denotes the Euclidean norm of the vector. Define the associated empirical correlation matrix,

$$\hat{R}_N = (\hat{\rho}_{ij}) = \left( \frac{e_i'e_j}{\|e_i\| \|e_j\|} \right), \quad 1 \leq i,j \leq N.$$

Then clearly, $\sum_{i,j=1}^{N} \hat{\rho}_{ij}^2 = \text{tr}\hat{R}_N^2$. Hence the LM statistic in (5) is equal to

$$LM = T \sum_{i<j} \hat{\rho}_{ij}^2 = \frac{T}{2} \left( \sum_{i,j=1}^{N} \hat{\rho}_{ij}^2 - N \right) = \frac{T}{2} \left( \text{tr}\hat{R}_N^2 - N \right).$$

By applying recent RMT results to the correlation matrix $\hat{R}_N$, we find an asymptotic normal distribution for the statistic $\text{tr}\hat{R}_N^2$ under the high-dimensional setting. This asymptotic distribution is easily transferred to the LM test statistic of our interest using the identity above.
3.1. Preliminary reductions

In order to develop a rigorous central limit theorem for statistic $\text{tr}\hat{\mathbf{R}}_{N}^{2}$, the following assumptions are needed.

Assumption (a): The number of covariates $k$ is fixed. Both $T$ and $N = N(T)$ tend to infinity in such a way that $\lim N/T \rightarrow c > 0$ as $T \rightarrow \infty$.

Assumption (b): Under the null hypothesis $H_{0}$, the random errors are of the form $u_{it} = \sigma_{i}e_{it}$ with some positive $\sigma_{i}$ and where the $N \times T$ array $\{e_{it}, 1 \leq i \leq N, 1 \leq t \leq T\}$ are IID standard normal random variables.

Assumption (c): Within each unit $i$, the $T \times k$ design matrix $\mathbf{X}_{i} = (x_{i1}, ..., x_{iT})'$ is of full rank $k$.

Assumption (d): Within each unit $i$, the covariates, $\{x_{it}, 1 \leq t \leq T\}$ and the errors $\{u_{it}, 1 \leq t \leq T\}$ are independent.

Assumption (a) defines the high-dimensional framework considered in this paper. It will soon become clear that although in theory an asymptotic limit of the ratios $N/T$ is assumed, the limiting results can be used with the ratio at hand $N/T$ from a given data set, without relying on the asymptotic limit $c$. This important fact makes the whole procedure feasible in practice. In contrast to our Assumption (a), Pesaran et al. (2008) assumes an almost equivalent condition, which is $T = O(N^{r})$ for some $0 < \epsilon \leq 1$. In this case, the implicit null of the cross-sectional dependence test is given by $0 \leq \alpha < (2 - \epsilon)/4$, or $0 \leq \alpha < 1/4$ when $N$ and $T$ tend to infinity at the same rate such that $T/N \rightarrow \kappa$ with $\kappa$ being a finite positive constant.

Our Assumption (b) requires that for each unit $i$, $u_{it}$’s are IID for all $i$ and $t$, which is in line with both Pesaran et al. (2008) and Pesaran (2015a). While our Assumption (b) also requires normality of $u_{it}$ which is not needed in Pesaran et al. (2008), Pesaran (2015a) further assumes symmetry of the error distribution. Our normality assumption for the errors is fundamental for the derivation of our high-dimensional results. However, simulations in Section 4 confirm robustness of the test with respect to non normal errors.

Finally, our Assumptions (c) and (d) require that within each unit $i$, the regression vectors $x_{it}$’s can exhibit time dependence but are required to be independent of the errors $u_{it}$. By comparison, both Pesaran et al. (2008) and Pesaran (2015a) consider $x_{it}$ to be a vector of strictly exogenous regressors such that $E(u_{it}|x_{it}) = 0$, for all $i$ and $t$.

A key step in our technical proofs is the following reduction for the residual vectors $\{e_{i}\}$ in (9): there is a $T_{k} \times 1$ standard normal vector $z_{i}$ and a $T \times T_{k}$ random matrix $\mathbf{U}_{i}$ with orthonormal columns, $z_{i}$ and $T \times T_{k}$ matrix $\mathbf{U}_{i}$ being independent, such that

$$e_{i} = \sigma_{i}\mathbf{M}_{i}e_{i} = \sigma_{i}\mathbf{U}_{i}z_{i}.$$  \hspace{1cm} (12)

To see why this claim is true, note that $\mathbf{X}_{i}(\mathbf{X}'_{i}\mathbf{X}_{i})^{-1}\mathbf{X}_{i}$ and $\mathbf{M}_{i} = I_{T} - \mathbf{X}_{i}(\mathbf{X}'_{i}\mathbf{X}_{i})^{-1}\mathbf{X}'_{i}$ are the projection matrices onto the sub-space generated by the $k$ columns of $\mathbf{X}_{i}$, and onto its orthogonal complement, respectively. Then we can find a $T \times T_{k}$ matrix $\mathbf{U}_{i}$ whose $T_{k}$ columns are orthonormal such that $\mathbf{M}_{i} = \mathbf{U}_{i}\mathbf{U}'_{i}$. We have

$$e_{i} = \sigma_{i}\mathbf{M}_{i}e_{i} = \sigma_{i}\mathbf{U}_{i}\mathbf{U}'_{i}e_{i} = \sigma_{i}\mathbf{U}_{i}z_{i},$$

where $z_{i} = \mathbf{U}'_{i}e_{i} \equiv (z_{i1}, ..., z_{iT})'$ has zero mean and covariance matrix $\sigma_{i}^{2}\mathbf{I}_{T_{k}}$. Note that despite the multiplication by $\mathbf{U}'_{i}$, $z_{i}$ is independent of $\mathbf{U}_{i}$ (since its conditional distribution given $\mathbf{U}_{i}$ is independent of $\mathbf{U}_{i}$). The claim (12) is thus established.

Recall that our target is the empirical correlation matrix $\hat{\mathbf{R}}_{N}$ of the residuals $\{e_{i}\}$ in (10). We now rescale the residuals in order to get a normalized population with unit covariance matrix. Define
with the corresponding population \( v = \sqrt{T/T_k} U z \). Clearly,
\[
\mathbf{R}_N = \begin{pmatrix} \mathbf{e}' \mathbf{e}_j & \mathbf{v}_i \mathbf{v}_j \\ \|\mathbf{e}_i\| \|\mathbf{e}_j\| & \|\mathbf{v}_i\| \|\mathbf{v}_j\| \end{pmatrix}, \quad 1 \leq i, j \leq N,
\]
so it is also the empirical correlation matrix of the \( v_i \)’s. The advantage of this scaling is that the population for the \( v_i \)’s is standardized as shown in the next lemma.

**Lemma 3.1.** For the population \( v = \sqrt{T/T_k} U z = (v_1, \ldots, v_T) \), we have \( \mathbb{E}[v] = 0 \), \( \mathbb{E}[vv'] = \mathbf{I}_T \). Moreover, its distribution is invariant under permutation of its coordinates \( v_j \). In particular, they have the same marginal distribution and
\[
\kappa = \mathbb{E}[v_i^4] = \frac{3T(T_k + 2)}{(T + 2)T_k}. \tag{13}
\]
The proof of this lemma is given in Appendix A.

### 3.2. High-dimensional limit of the LM statistic

The main theoretical result of the paper establishes the asymptotic normality of the LM statistic.

**Theorem 3.1.** Assume that Assumptions (a)-(b)-(c) hold. Then under the null hypothesis \( H_0 \) shown in (2),
\[
LM_{RMT} = v_T^{-1} \left( \frac{2LM}{T} - Nc_T - c_T^2 + c_T \right) \Rightarrow N(0, 1), \tag{14}
\]
where
\[
c_T = \frac{N}{T}, \\
v_T^2 = 4c_T(1 + 2c_T)(c_T + 2) - 4(k - 1)c_T(1 + c_T)^2 \\
+ (k - 3)c_T(c_T - 4)^2(c_T + 1)^2
\]
and \( \kappa \) is given in (13).

The proof of this theorem is given in Appendix B.

It is worth comparing the asymptotic normal distribution found in Theorem 3.1 and the one derived in PUY for the statistic \( LM_{adj} \). As mentioned earlier, from a theoretical perspective, Theorem 3.1 assumes a full double limit scheme \( N, T \to \infty \) where \( N, T \) are of comparable magnitude, while for \( LM_{adj} \) a \( T \)-first-then-\( N \) sequential limiting scheme is used. Thus, the two asymptotic schemes are very distinct. Empirically through the simulation experiments in Section 4, we find that the two tests have comparable performance in small samples, with a slightly higher power in favor of the proposed test in some cases.

### 4. Monte Carlo simulations

We investigate the small sample properties of our proposed \( LM_{RMT} \) statistic by use of a Monte Carlo simulation study and make comparisons with related tests in the literature. For this purpose, we consider a similar data generating process to Pesaran et al. (2008):
Finally, we set interesting comparisons with other tests, but additional values of applicable, for each innovations in (17), (i.e. (17) produces weakly cross-sectionally dependent errors, mined by the value of the exponent, of units that are significantly affected by factor additional dependence in (2016, 2019). Under the null hypothesis, These are generated as: The bias-corrected LM test of Baltagi et al. (2012) which assumes a static panel data model in all settings, we run OLS regressions of The original one-sided Lagrange Multiplier (LM) statistic of Breusch and Pagan (1980), given in (4) and (5). This is asymptotically distributed as \( LM \sim \chi^2(N(N-1)/2) \). The bias-adjusted version of the LM test, developed by Pesaran et al. (2008), given in (6). The \( LM_{\text{adj}} \) statistic is a one-sided test and is distributed as \( LM_{\text{adj}} \sim N(0,1) \). The bias-corrected LM test of Baltagi et al. (2012) which assumes a static panel data model (i.e. \( \hat{\lambda}_i = 0 \), for all \( i \)), homogeneous slope coefficients (i.e. \( \beta_{\ell i} = \beta_{\ell} \), for all \( i \)) and allows for heteroskedastic errors in (15). This is given by:

\[
y_{it} = a_i(1 - \hat{\lambda}_i) + \sum_{\ell=2}^{k} \beta_{\ell i}x_{\ell it} + \hat{\lambda}_iy_{i,t-1} + u_{it}, \quad 1 \leq i \leq N; \quad 1 \leq t \leq T, \tag{15}
\]

where \( a_i \sim \text{IIDN}(1,1) \), \( \hat{\lambda}_i \sim \text{IIDU}(0,0.04) \), \( \beta_{\ell i} \sim \text{IIDN}(1,0.04) \) and \( k \) is the number of rhs exogenous regressors in (15) including the individual-specific intercepts. The exogenous covariates are generated as:

\[
x_{\ell it} = 0.6x_{\ell i,t-1} + \sigma_{\ell i}\nu_{\ell it}; \quad \ell = 2,3,\ldots,k, \quad i = 1,2,\ldots,N; \quad t = -50,49,\ldots,0,\ldots,T, \tag{16}
\]

with \( x_{\ell i,-51} = 0 \), \( \sigma_{\ell i}^2 = \tau_{\ell i}^2/(1-0.6^2) \) and \( \tau_{\ell i}^2 \sim \text{IID}\chi^2(6)/6 \). The first 50 observations in (16) are disregarded. The innovations in (16), \( \nu_{\ell it} \), are generated under two different schemes: (i) normal, IIDN(0,1), and (ii) non-normal, IID(\( \chi^2(1)-1 \))/\( \sqrt{2} \). Of importance are the disturbances in (15). These are generated as:

\[
u_{\ell it} = c_{(i;k)}(\gamma_{fi} + \sigma_{fi}e_{fi}), \quad 1 \leq i \leq N; \quad 1 \leq t \leq T, \tag{17}
\]

where \( f_i \sim \text{IIDN}(0,1) \), and \( \sigma_{fi}^2 \sim \text{IID}\chi^2(2)/2 \). The factor loadings are constructed as:

\[
\left\{
\begin{array}{ll}
\gamma_i = v_{i}, & \text{for } i = 1,2,\ldots,[N^z], \\
\gamma_i = 0, & \text{for } i = [N^z] + 1, [N^z] + 2,\ldots,N,
\end{array}
\right. \tag{18}
\]

where \( v_i \sim \text{IIDU}(0.5,1.5) \), \( [N^z] \) is the integer part of \( N^z \), and parameter \( z \) determines the number of units that are significantly affected by factor \( f_i \), and ranges between zero and unity - for further details on the definition of the exponent of cross-sectional dependence, \( z \), see Bailey et al. (2016, 2019). Under the null hypothesis, \( H_0 : \text{Cov}(u_{it}, u_{jt}) = 0 \), for all \( t \) and \( i \neq j \), which implies that \( z = 0 \) and \( \gamma_i = 0 \) for all \( i \). A number of alternatives (\( H_1 \)) can be considered which are determined by the value of the exponent, \( z \). For values of \( z \) less than 0.5 the alternatives are local and (17) produces weakly cross-sectionally dependent errors, \( u_{it} \). For \( z > 0.5 \) then \( u_{it} \) become semi-strongly cross-sectionally dependent, until \( z = 1 \) which is the extreme case of strong cross-sectional dependence in \( u_{it} \). We focus on local alternatives \( H_{1,z} \) with \( z \in \{0.2,0.4\} \) which make interesting comparisons with other tests, but additional values of \( z \) also can be considered. Finally, we set \( c_{(i;k)} = 1 \) under both \( H_0 \) and all versions of \( H_{1,z} \) as well as for all \( k \). Again, for the innovations in (17), \( e_{fi} \), these are generated under two separate schemes: (i) normal, IIDN(0,1), and (ii) non-normal, \( \text{IID}(\chi^2(1)-1)/\sqrt{2} \).

In all settings, we run OLS regressions of \( y_{it} \) on a constant, \( x_{\ell it}, \ell = 2,3,\ldots,k \), and \( y_{i,t-1} \) where applicable, for each \( i \) and obtain residuals:

\[
e_{it} = y_{it} - \hat{\alpha}_i(1 - \hat{\lambda}_i) - \sum_{\ell=2}^{k} \hat{\beta}_{\ell i}x_{\ell it} - \hat{\lambda}_iy_{i,t-1}, \quad \text{with } \hat{\alpha}_i, \hat{\lambda}_i \text{ and } \hat{\beta}_{\ell i} \text{ being estimates of } a_i, \lambda_i \text{ and } \beta_{\ell i} \text{ for } i \text{ and } \ell, \text{ respectively.}
\]

We investigate the properties of our proposed one-sided test statistic, \( LM_{RMT} \), with its asymptotic normal distribution given in Theorem 3.1. We also consider the following test statistics in the literature:

The original one-sided Lagrange Multiplier (LM) statistic of Breusch and Pagan (1980), given in (4) and (5). This is asymptotically distributed as \( LM \sim \chi^2(N(N-1)/2) \).

The bias-adjusted version of the LM test, developed by Pesaran et al. (2008), given in (6). The \( LM_{\text{adj}} \) statistic is a one-sided test and is distributed as \( LM_{\text{adj}} \sim N(0,1) \).
where \( \hat{\rho}_{ij} \) are the correlation coefficients corresponding to within residuals of units \( i \) and \( j \), for \( i,j = 1,2,\ldots,N \). These are denoted by \( \hat{v}_{it} = \tilde{y}_{it} - \tilde{x}_{it}' \hat{\beta} \), where \( \tilde{y}_{it} = y_{it} - \frac{1}{T} \sum_{t=1}^{T} y_{it} \), \( \tilde{x}_{it} = (\tilde{x}_{2it}, \tilde{x}_{3it}, \ldots, \tilde{x}_{kit})' \), \( \tilde{x}_{sit} = x_{sit} - \frac{1}{T} \sum_{t=1}^{T} x_{sit} \) and \( \hat{\beta} = \left( \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{x}_{it}' \tilde{x}_{it} \right)^{-1} \left( \sum_{i=1}^{N} \sum_{t=1}^{T} \tilde{x}_{it} \tilde{y}_{it} \right) \). \( LM_{bc} \) is a one-sided test and distributed as \( N(0, 1) \) under the null.

Finally, the CD statistic of Pesaran (2015a):

\[
CD = \sqrt{\frac{2T}{N(N-1)} \left( \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij} \right)},
\]

where \( \hat{\rho}_{ij} \) is given by (4). CD is a two-sided test and is distributed as \( CD \sim N(0, 1) \).

In all experiments the values of \( a_i, \beta_{i\ell}, \lambda_i, x_{i\ell} \) and \( \sigma_i^2 \) are drawn for each \( i = 1,2,\ldots,N \) and then fixed across replications. We commence with the static version of (15) where \( \lambda_i = 0 \) for all \( i \), and examine the effects of changing the number of exogenous regressors (which include individual-specific intercepts) by setting \( k = 2,4 \). Additionally, we study the case where one of the exogenous regressors in (15) assumes a quadratic form, i.e. \( k = 3 \) and \( x_{i3t} = x_{i2t}^2 \) where \( \tau_i^2 = 0.5 \), for \( \ell = 2,3, i = 1,2,\ldots,N \). Further, we consider a homogeneous slope coefficient setting where \( \beta_{i\ell} = \beta_{i} \), for all \( i \) in (15) and \( \beta_{i} \sim \text{IIDN}(1,0.04) \). Finally, we examine dynamic versions of (15): (i) a pure AR(1) process with individual-specific intercepts, where \( \lambda_i \sim \text{IIDU}(0.0,0.4) \) and \( \beta_{i\ell} = 0 \), for all \( i \) and \( \ell \), and (ii) a process with one time lag of \( y_{i\ell t} \), individual-specific intercepts and one exogenous regressor (i.e. \( k = 2 \) plus the time lag). The combinations of sample sizes \( N = 50,100,200,500,800,1000 \) and \( T/N = 0.5,1.0,2.0 \) are considered. The nominal size of the tests is set at the 5% significance level. All experiments are based on 2000 replications. The simulated results are listed in Table 1 and Figures 1a–3.

---

We have also considered the case of \( k = 6 \). These results are qualitatively similar to those for \( k = 4 \) and are omitted in order to save space.
4.1. Results

We commence with the baseline scenario of a static heterogeneous panel data model and consider both the exogenous regressors and error terms to be Gaussian in (15). Table 1 shows size results under this scenario for the original $LM$ and $LM_{bc}$ statistics. The $LM$ test is clearly oversized for all $(N, T)$ combinations, which is expected especially when $N \gg T$. Similarly, $LM_{bc}$ records very high empirical size universally which again makes intuitive sense since the test has been designed for homogeneous panel models. Given the uniformly poor performance of both tests, we will omit $LM$ and $LM_{bc}$ from all subsequent size and power comparisons.\(^3\) In what follows, we concentrate on the remaining adjusted versions of $LM$,

\(^3\)When considering homogeneous slope coefficients, $\beta_j$, in (15), $LM_{bc}$ naturally displays correct size as $N \to \infty$, for all $T$, as do the remaining $LM_{RMT}$, $LM_{adj}$ and CD tests which are based on a slope heterogeneity assumption. Results for this experiment are shown in Figure A4 of the Online Supplement.
namely $LM_{RMT}$, $LM_{adj}$, and the $CD$ statistic. Figure 1a and b show size and local power results such that $\alpha = 0.0$, 0.2 and 0.4 in (18) for these tests under the same scenario for all $(N, T)$ combinations, when the number of regressors (inclusive of individual-specific intercepts) is set to $k = 2$ and 4, respectively.

Focusing on size first, it is evident from the left column of Figure 1a and b that with a nominal size of $p = 0.05$, our proposed RMT based $LM$ test produces empirical sizes which hover around the nominal values of $p = 0.05$ as $N \rightarrow \infty$, for all $T$ and irrespective of the number of regressors included in the panel data model. Some incremental size distortions can be detected when both $N$ and $T$ are small and as $k$ rises. For example, the empirical size of $LM_{RMT}$ rises from 6.5% to 8.8% when considering combination $T/N = 0.5$ with $N = 50$ and when $k$ rises from 2 to 4. This anomaly is expected when $N$ and $T$ take small values since random matrix theory results apply to high dimensional problems, but it disappears as $T$ increases, for all values of $N$.

Compared to the other tests considered, $LM_{adj}$ follows a near identical pattern to $LM_{RMT}$ for all $(N, T)$ combinations with the exception of small sample sizes and when the number of regressors increases where it exhibits superior performance. The $CD$ test fluctuates around the nominal size
irrespective of the values of \( N \) and \( T \) or the number of regressors, and hence outperforms LMRMT in small samples with larger \( k \). However, performance of all tests is considered comparable under increasingly large sample settings with their empirical sizes converging to \( p \) as \( N \) rises, for all \( T \).

Next, we turn to the middle column of Figure 1a and b which present empirical local power results when \( \alpha = 0.2 \) for the LM adjusted statistics and the CD test. Given that the departure from the null of \( \alpha = 0.0 \) is very small, the LMRMT test does not produce very high empirical power. However, as \( T \) increases the empirical power rises as \( N \to \infty \) and irrespective of the number of regressors in the model. The elevated numbers shown for \( T/N = 0.5 \) for all \( k \) should be treated with care, since they correspond to the cases when size of the LMRMT tests are distorted. Compared to the LM_{adj} test, the LMRMT statistic slightly outperforms the former but not to a great extent. However, the RMT based statistic clearly outperforms the CD test for all \( N, T \) and \( k \). The empirical power of the CD statistic still hovers around the 5% mark throughout, which is expected since the null hypothesis for this test coincides with the assumption that \( 0 \leq \alpha < (2 - \epsilon)/4 \), where \( \epsilon \) is in the range \((0,1]\) and \( T = O(N^\alpha) \).
Finally, the right column of Figure 1a and b analyses the more distant alternative hypothesis of $\alpha = 0.4$. As $N$ increases the empirical power for the $LM_{RMT}$ test tends to unity for all values of $T$ and $k$. It appears that setting as alternative hypothesis a factor structure with $N^{0.4}$ non-zero loadings is sufficient to achieve very high power. The performance of the $LM_{RMT}$ statistic is comparable to the $LM_{adj}$ test, which also produces empirical power tending to unity for all $N$, $T$ and $k$ combinations. Further, both $LM_{RMT}$ and $LM_{adj}$ statistics continue to outperform the CD test, which lags behind considerably.

Empirical size and power results corresponding to the scenarios when: (i) the exogenous regressors are Gaussian but the error terms are non-normally distributed, (ii) the exogenous regressors are non-normally distributed but the error terms are Gaussian, and (iii) both the exogenous regressors and error terms are non-Gaussian, are presented in Figures A1a-A1b, A2a-A2b and A3a-A3b of the Online Supplement, respectively. Results are qualitatively similar to those shown in Figure 1a and b, which implies that $LM_{RMT}$ is robust to departures from the Gaussianity assumption for the errors shown in Section 3. Similarly, assigning a non-linear setting by letting one of the exogenous variables in (15) take a quadratic form, produces size and power results that comply to those of the baseline scenario. These

Figure 3. Empirical size of $LM_{RMT}$ and alternative tests applied to residuals of a dynamic panel data model with one lag of the dependent variable and $k = 1, 2$ Gaussian exogenous regressors (including individual-specific intercepts), all with heterogeneous slope coefficients, and Gaussian or non-Gaussian errors, when $N = 50, 100, 200, 500, 800, 1000$ and $T/N = 0.5, 1.0, 2.0$. See notes in Figure 1a. The dynamic panel data model is generated as in (15). $G$ and $NG$ stand for Gaussian and non-Gaussian errors, respectively. Computation of $LM_{adj}$ for $N = 800, 1000$ has been omitted for computational efficiency.
are shown in Figure 2. All the above findings are not discussed in length for brevity of exposition.

It is interesting to examine the dynamic version of (15) also considered in the simulations section of Pesaran et al (2008). Here, we focus on the size of the $LM_{RMT}$ and $LM_{adj}$ CD tests. Assuming an AR(1) dynamic process with Gaussian errors, we obtain similar size results for the three tests to the baseline scenario under all $(N, T)$ combinations, as shown in the first column of Figure 3. Adding an extra exogenous regressor does not alter the general pattern of size for the $LM_{RMT}$ and CD statistics. However, in this case $LM_{adj}$ becomes considerably oversized, especially when $T/N = 0.5$ - see middle and left columns of Figure 3 where data are generated under (15) with Gaussian and non-Gaussian errors, respectively.

Overall, for practitioners interested in testing for error independence rather than weak dependence in large dimensional static panel data models where $T \gg N$ no longer holds, $LM_{RMT}$ and $LM_{adj}$ statistics are preferred, given their similar size and power performance based on the sample sizes considered in this simulation study. From a practical viewpoint, $LM_{RMT}$ is also comparatively computationally more efficient as $N$ takes on very large values since it does not require computation of $\mu_{Tij}$ and $\nu_{Tij}$ as in (6). Outperformance of $LM_{RMT}$ over $LM_{adj}$ in terms of empirical size when considering dynamic panel data models is an interesting outcome that warrants further investigation in future research.

5. Conclusion

This paper proposes a new statistic for testing for error cross-sectional independence in a panel data model with strictly exogenous regressors of large cross section $(N)$ and time $(T)$ dimensions. The test is based on the Lagrange Multiplier principle and makes use of the pairwise sample correlation coefficients of residuals obtained by regressing the panel data model. Recognizing that in a high dimensional setting where $T \gg N$ no longer holds, the sample correlation matrix derived from the linear associations between all pairs of these $N$ residuals is not consistent, we use results from the Random Matrix Theory literature to prove the asymptotic limit of our $LM_{RMT}$ statistic. This assumes that both $T$ and $N(T) \to \infty$ and $N/T \to \zeta > 0$, and that the error terms are normally distributed. The small sample properties of the proposed $LM_{RMT}$ test are investigated by use of Monte Carlo experiments. These show that $LM_{RMT}$ is correctly sized for large $N$ and for all $T$ considered, with some size distortions for small values of $N$ and $T$. The statistic shows satisfactory power under local alternatives of weak cross-sectional dependence expressed in terms of the exponent of cross-sectional dependence, $\zeta$, and tends toward unity even for values of $\zeta$ as small as 0.4. Compared to existing related tests in the literature which have been developed under different asymptotics, $LM_{RMT}$ universally outperforms the original $LM$ test which is severely oversized in high dimensional settings, and has similar size properties to the bias adjusted $LM$ statistic and marginally better power as $N$ rises. The CD test has universally correct size but $LM_{RMT}$ has superior power performance under local alternatives. The $LM_{bc}$ test which has been developed for homogeneous panels is severely oversized under a heterogeneous panel setting. Finally, the test is robust to inclusion of temporal dynamics. A limitation of our test statistic is the assumption that the time-varying regressors and errors are fully independent, as prescribed by the random matrix theory framework. Theoretical extensions that deal with non-normality, serial correlation or endogeneity are interesting avenues for future research.

576 N. BAILEY ET AL.

4From the expressions of $\mu_{Tij}$ and $\nu_{Tij}$ shown in (7) and (8) respectively, the added complexity of operations amounts to an order of $O(N^2T)$, which augments the computational time for the $LM_{adj}$ statistic.
Appendix A: Proof of Lemma 3.1

Let \( y = Uz \) so that \( v = \sqrt{T/T_k}y \). For the first claim of the lemma, it is sufficient to prove that

\[
E[y] = 0, \quad \text{and} \quad E(yy') = \frac{T_k}{T} I_T.
\]

(A.1)

Obviously \( E[y] = E[U]E[z] = 0 \). Next, since \( U \) is a \( T_k \) -frame, that is,

\[
U = (u_1, \ldots, u_{T_k}) = \begin{pmatrix}
u_{11} & \cdots & u_{1T_k} \\
\vdots & \ddots & \vdots \\
u_{T1} & \cdots & u_{TT_k}
\end{pmatrix}
\]

by the properties of a Haar matrix, it follows that

\[
E(u^2_{i1}) = \frac{1}{T^2}; \quad E(u_{i1}u_{j1}) = 0, \quad \text{for} \quad i \neq j.
\]

Therefore, \( E(u_1u_1') = \frac{1}{T}I_T \). It follows that

\[
E(yy') = E(UU') = E\left(\sum_{j=1}^{T_k} u_j u_j'\right) = T_k E(u_1u_1') = \frac{T_k}{T} I_T.
\]

It remains to calculate the value of \( E[u_i^4] \). Write \( U = (u_j) \) and \( z = (z_1, \ldots, z_{T_k})' \), and

\[
v = \sqrt{T/T_k}y = \sqrt{T/T_k}Uz \triangleq (v_1, \ldots, v_T)'.
\]

Clearly, the components \( v_i = \sqrt{T/T_k} \sum_{j=1}^{T_k} u_{ij}z_j \) are identically distributed for \( 1 \leq i \leq T \). We know already that

\[
E(v) = 0; \quad E(v^2) = I_T.
\]

Next, we have

\[
E(v_i^4) = \frac{T^2}{T_k^2} E\left(\sum_{j=1}^{T_k} u_{ij}z_j\right)^4
\]

\[
= \frac{T^2}{T_k^2} E\left(\sum_{j_1\neq j_2\neq j_3\neq j_4} u_{i1}u_{i2}u_{i3}u_{i4}z_{j1}z_{j2}z_{j3}z_{j4}\right)
\]

\[
= \frac{T^2}{T_k^2} \sum_{j_1\neq j_2\neq j_3\neq j_4} E(u_{i1}u_{i2}u_{i3}u_{i4})E(z_{j1}z_{j2}z_{j3}z_{j4})
\]

where \( E(u_{i1}u_{i2}u_{i3}u_{i4}) = 0 \) except for four cases:

(i) \( j_1 = j_2 = j_3 = j_4 \);

(ii) \( j_1 = j_2 \neq j_3 = j_4 \);

(iii) \( j_1 = j_3 \neq j_2 = j_4 \);

(iv) \( j_1 = j_4 \neq j_2 = j_3 \).

So

\[
E(v_i^4) = \frac{T^2}{T_k^2} \left(T_k E_{u_{i1}}^4 E_{z_1}^4 + 3T_k(T_k - 1) E(u_{i1}^2u_{i2}^2 E(z_{i1}^2z_{i2}^2))\right).
\]

\( U \) is a Haar matrix and \( (u_{i1}, \ldots, u_{iT_k}, u_{iT_k+1}, \ldots, u_{iT})' \) follows a uniform distribution on a \( T \)-dimensional sphere. The elements \( \{u_{ij}\}_{1 \leq i \leq T_k} \) have the same marginal distribution by symmetry and \( u_{ij}^2 \) has a beta distribution with parameter \( \left(\frac{1}{2}, \frac{1}{2}\right) \). Then, it can be easily obtained that

\[
E(u_{i1}^2) = \frac{1}{T}; \quad E(u_{i1}^4) = \frac{3}{T(T+2)}.
\]

Further, by

\[
1 = (u_{i1}^2 + \cdots + u_{iT_k}^2)^2 = T E_{u_{i1}}^4 + T(T - 1) E(u_{i1}^2u_{i2}^2)
\]

we have
\[ \mathbb{E}(u^2_{i1} u^2_{i2}) = \frac{1}{T(T+2)}. \]

Under the standard normal assumption of \( z \), we have

\[ \mathbb{E}(z_i) = 0; \quad \mathbb{E}(z^2_i) = 1; \quad \mathbb{E}(z_i^2 z_j^2) = 1; \quad \mathbb{E}(z_i^4) = 3. \]

Therefore, we have

\[ \mathbb{E}(v_i^4) = \frac{3T(T+2)}{T(T+2)}. \]

**Appendix B: Proof of Theorem 3.1**

The proof is based on a central limit theorem for linear spectral statistics of a high-dimensional correlation matrix established in Gao et al. (2017). First recall that the Marcenko-Pastur law \( F_c \) with index \( c > 0 \) is the distribution with density

\[ p_c(x) = \frac{1}{2\pi c x} \sqrt{(b(c) - x)(x - a(c))}, \quad a(c) \leq x \leq b(c), \]

where the support interval has endpoints \( a(c) = (1 - \sqrt{c})^2 \) and \( b(c) = (1 + \sqrt{c})^2 \). In addition, the distribution has a point mass at the origin with weight \( 1 - 1/c \) if \( c > 1 \). The first two moments of the distribution are

\[ \int xdF_c(x) = c \quad \text{and} \quad \int x^2dF_c(x) = 1 + c^2. \]

More information on this distribution can be found in Yao et al. (2015), Chapter 2. In particular, the Stieltjes transform \( m(z) \) of the distribution has an explicit form that solves the second-degree equation \( czm^2 + (z + c - 1)m + 1 = 0 \). In the following, the companion Stieltjes transform \( \mathcal{m}(z) \), defined as \( \mathcal{m}(z) = (1-c)/z + cm(z) \), is also used.

Under Assumptions (a)-(b)-(c), we can apply the main central limit theorem of Gao et al. (2017) to the quadratic function \( f(x) = x^2 \) and the sample correlation matrix \( R_N \) in (10). That is when \( T \to \infty \) and \( c_T = N/T \to c > 0 \),

\[ \text{tr}\mathcal{R}_N^2 - N(1 + c_T) \Rightarrow N(\mu(f), \nu(f)), \quad \text{(B.1)} \]

where

\[
\mu(f) = \frac{\kappa - 1}{2\pi i} \int_c \left\{ f(z) \frac{cm')(z)}{[1 + m(z)]^2 - cm^2(z)[1 + m(z)]} \right. \\
- \frac{\kappa - 3}{2\pi i} \int_c \left\{ f(z) \frac{cm^2(z)(1 + m(z))^2}{[1 + m(z)]^2 - cm^2(z)[1 + m(z)]} \right. \\
+ \frac{1}{2\pi i} \int_c \left\{ f(z) \frac{cm(z)m'(z)}{z[1 + cm(z)]^2 - cm^2(z)]} \right\} \, dz, \\
\]

and

\[
\nu(f) = -\frac{\kappa - 1}{2\pi i} \int_{c_1} \int_{c_2} f(z_1)f(z_2) \frac{cm'(z_1)m'(z_2)}{[1 + cm(z_1) + m(z_2)] + c(c - 1)m(z_1)m(z_2)} \, dz_1 \, dz_2 \\
+ \frac{\kappa - 3}{4\pi i} \int_{c_1} \int_{c_2} f(z_1)f(z_2) \frac{cm'(z_1)m'(z_2)}{(1 + m(z_1))^2[1 + m(z_2)]^2} \, dz_1 \, dz_2 \\
- \frac{\kappa - 3}{4\pi i} \int_{c_1} \int_{c_2} f(z_1)f(z_2) V(c, m(z_1), m(z_2)) \, dz_1 \, dz_2 \\
\]

where
\[ V\{c,m(z_1),m(z_2)\} = c\{m(z_1)m(z_1) + z_1m(z_1)m'(z_1) + z_1m'(z_1)m(z_1)\} \]
\[ \{m(z_2)m(z_2) + z_2m(z_2)m'(z_2) + z_2m'(z_2)m(z_2)\}. \]

Here the contours \( C, C_1 \) and \( C_2 \) are non-overlapping, closed, taken in the positive direction in the complex plane and enclosing the support of the Marčenko-Pastur distribution \( F_c \).

The main task is to evaluate these complex contour integrals to get the limiting parameters \( \mu(f) \) and \( \nu(f) \). These calculations are detailed in the two subsections below, leading to the results:

\[ \mu(f) = c^2 - c, \quad (B.2) \]

and

\[ \nu(f) = 4c(1+2c)(c+2) - 4(\kappa-1)c(1+c)^2 + (\kappa-3)c(c-4)(c+1)^2, \quad (B.3) \]

respectively. Thus, the main theorem follows by plugging these values in \( (B.1) \) and by noticing that the replacement of \( c \) with \( c_T = N/T \) does not modify the asymptotic distribution.

**Calculation of \( \mu(f) \) in \( (B.2) \)**

By equation \( (B.71) \) given in Gao et al. \( (2017) \), we have

\[ z = -\frac{1}{m(z)} + \frac{c}{1 + m(z)} \quad (B.4) \]

Denote \( m(z) \) as \( m \) for simplicity. It is easily obtained that

\[ z^2 = \frac{(c-1)m - 1}{m^2(1+m)^2} \quad dz = \frac{(1+m)^2 - cm^2}{m^2(1+m)^2} dm \]

According to the Lemma 2.1 and \( f(x) = x^2 \), the first part of \( \mu(f) \) can also be expressed as

\[ \mu_1(f) = \frac{(\kappa-1)}{2\pi i} \int z^2 \frac{cm^3}{[(1+m)^2 - cm^2](1+m)} \quad dz \]
\[ = \frac{(\kappa-1)}{2\pi i} \int \frac{c(cm - m - 1)^2}{m(1+m)^5} \quad dm \]

For \( z \in C^+ \), solving \( m \) from \( (B.4) \), the contour for the above integral should enclose the interval

\[ \left[ \min\left( -\frac{1}{1-\sqrt{c}}, -\frac{1}{1+\sqrt{c}} \right), \max\left( -\frac{1}{1-\sqrt{c}}, -\frac{1}{1+\sqrt{c}} \right) \right]. \]

So it is obtained that \(-1\) is the residue with the same integral direction if \( c \leq 1 \), and \( 0 \) is the residue with the opposite integral direction if \( c > 1 \). Therefore, the integral is calculated as

\[ \mu_1(f) = -(\kappa-1)c, \]

which is the same result for both cases of \( c \leq 1 \) and \( c > 1 \).

Similarly, the above substitution \( (B.4) \) is still used, by Lemma 2.1, then it follows that the second part of \( \mu(f) \) is

\[ \mu_2(f) = \frac{(\kappa-3)}{2\pi i} \int z^2 \frac{cm^2(1+m)m^2}{[(1+m)^2 - cm^2](1+cm)} \quad dz \]
\[ = \frac{(\kappa-3)}{2\pi i} \int \frac{-c((1-c)m + 1)^2}{m(1+m)^4} \quad dm \]
\[ = (\kappa-3)c \]

where
\[ m = -\frac{1 - \epsilon}{\bar{z}} + cm. \]  

(B.5)

This result is also valid for both cases of \( \bar{z} \leq 1 \) and \( \bar{z} > 1 \).

The third part of \( \mu(f) \) is

\[
\mu_3(f) = -\frac{1}{2\pi i} \int \frac{cm m'}{(1 + m)^2 - cm^2} \frac{dm}{1 + m} \\
= -\frac{1}{2\pi i} \int \frac{c(1 - \epsilon)m + 1)^2}{m(1 + m)^4 \left[ (1 + m)^2 - cm^2 \right]} \frac{dm}{m} \\
= c
\]

The fourth part of \( \mu(f) \) is

\[
\mu_4(f) = -\frac{1}{2\pi i} \int \frac{c^2(1 - \epsilon)m + 1)^2(1 + 2m)}{m(1 + m)^4} \frac{dm}{m} \\
= -c^2
\]

The fifth part of \( \mu(f) \) is

\[
\mu_5(f) = \frac{1}{2\pi i} \int \frac{c^2(\epsilon - \epsilon)m - 2]}{m(1 + m)} \frac{dm}{m} = 2c^2
\]

Finally, we obtain

\[
\mu(f) = \mu_1(f) + \mu_2(f) + \mu_3(f) + \mu_4(f) + \mu_5(f) \\
= c^2 - c.
\]

**Calculation of \( u(f) \) in (B.3)**

Let \( m_i \) denote \( \bar{m}(z_i), i = 1, 2 \), for simplicity. By equation (B.5), the first part of \( u(f) \) should be

\[
u_1(f) = -\frac{1}{2\pi i} \int \int \frac{cm_1 m'_2}{[1 + c(m_1 + m_2)](1 + m_1)\left(1 + m_2\right)^3} \frac{m_1}{m_2} \frac{m_2}{m_1} \frac{dm_1}{dm_2} \\
= \frac{1}{2\pi i} \int \frac{c}{m_1(1 + m_2)^3} \frac{dm_1}{m_1} \left[ \frac{(c - 1)m_1 + 1)^2}{m_1(1 + m_2)^3} \frac{dm_2}{m_2} \\
- \frac{(c - 1)m_2 - 1)^2}{m_2(1 + m_2)^3} \frac{dm_1}{m_1} \right] \\
= \frac{1}{2\pi i} \left[ \frac{c^2(\epsilon - \epsilon)m_1 - 2]}{m_1(1 + m_2)^3} + \frac{c^2(\epsilon - \epsilon)m_2 - 1)^2}{m_2(1 + m_2)^3} \right] \frac{dm_1}{m_1} \frac{dm_2}{m_2} \\
= \frac{4c(\epsilon + 1)}{c_2 + 1}(c_2 + 1)
\]

Clearly, similar calculations are conducted for two cases \( 0 \leq \epsilon \leq \bar{z} \leq 1 \) and \( \bar{z} > 1 \) with different residues. It is found that the results are the same for both cases of \( \bar{z} \).

For the second part of \( u(f) \), we still use the same substitution in (B.4), and consider the calculation in two cases: first, the contour containing \(-1\) as a residue if \( 0 \leq \bar{z} \leq 1 \); second, enclosing \( 0 \) as a residue for \( \bar{z} > 1 \). Then, we get the same results for both cases with different residues, which is

\[
u_2(f) = \frac{(\epsilon - 1)c}{4\pi^2} \int \int \frac{cm_1 m_2}{(1 + m_1)^3 (1 + m_2)^3} \frac{dm_1}{m_1} \frac{dm_2}{m_2} \\
= \frac{(\epsilon - 1)c}{4\pi^2} \int \int \frac{(c - 1)m_1 - 1)^2}{m_1(1 + m_2)^3} \frac{dm_1}{m_1} \frac{(c - 1)m_2 - 1)^2}{m_2(1 + m_2)^3} \frac{dm_2}{m_2} \\
= -4(\epsilon - 1)c(1 + c)^2.
\]
Furthermore, the third part of $v(f)$ is

$$v_3(f) = -\frac{\kappa - 3}{4\pi^2} \left\{ \sum_i \sum_j t_i t_j \left\{ m_1 m_1 + z_i m_1 m_1' + z_1 m_1' m_1 \right\} \right. \left\{ m_2 m_2 + z_i m_2 m_2' + z_2 m_2' m_2 \right\} dz_i dz_j \right.$$ 

$$= -\frac{(\kappa - 3)c^2}{4\pi^2} \left\{ \left\{ (c - 1) m_1 - 1 \right\} \left\{ c m_1^2 - (1 + m_1)^2 \right\} \right.$$ 

$$\frac{m_1^2(1 + m_1)^3}{m_1(1 + m_1)} \left\{ (c - 1) m_1 - 1 \right\} \left\{ c m_1^2 - (1 + m_1)^2 \right\} \frac{d m_1}{m_1^2} \right\}^2 \right.$$ 

$$= (\kappa - 3)c(c - 4)(c + 1)^2.$$

Collecting the calculations, we finally obtain

$$v(f) = 4c(1 + 2c)(c + 2) - 4(\kappa - 1)c(1 + c)^2 + (\kappa - 3)c(c - 4)^2(c + 1)^2.$$

**Funding**

D. Jiang’s research is partly supported by Project 11971371 from NSFC.

**References**

Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

Anselin, L. (2001). Spatial econometrics. In: Baltagi, B., ed., *A Companion to Theoretical Econometrics*. Oxford: Blackwell.

Bailey, N., Kapetanios, G., Pesaran, M. H. (2016). Exponent of cross-sectional dependence: Estimation and inference. *Journal of Applied Econometrics* 31(6):929–960. doi:10.1002/jae.2476

Bailey, N., Kapetanios, G., Pesaran, M. H. (2019). Exponent of cross-sectional dependence for residuals. *Sankhya B* Statistics 81(S1):46–102. doi:10.1007/s13571-019-00196-9

Baltagi, B., Feng, Q., Kao, C. (2011). Testing for sphericity in a fixed effects panel data model. *The Econometrics Journal* 14(1):25–47. doi:10.1111/j.1368-423X.2010.00331.x

Baltagi, B., Feng, Q., Kao, C. (2012). A Lagrange multiplier test for cross-sectional dependence in a fixed effects panel data model. *Journal of Econometrics* 170(1):164–177. doi:10.1016/j.jeconom.2012.04.004

Breusch, T. S., Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification tests in econometrics. *The Review of Economic Studies* 47(1):239–253. doi:10.2307/2297111

Chudik, A., Pesaran, M. H., Tosetti, E. (2011). Weak and strong cross-section dependence and estimation of large panels. *The Econometrics Journal* 14(1):C45–C90. doi:10.1111/j.1368-423X.2010.00330.x

Cliff, A., Ord, J. K. (1973). *Spatial Autocorrelation*. London: Pion.

Cliff, A., Ord, J. K. (1981). *Spatial Processes: Models and Applications*. London: Pion.

Conley, T. G., Molinari, F. (2007). Spatial correlation robust inference with errors in location or distance. *Journal of Econometrics* 140(1):76–96. doi:10.1016/j.jeconom.2006.09.003

Frees, E. W. (1995). Assessing cross-sectional correlation in panel data. *Journal of Econometrics* 69(2):393–414. doi:10.1016/0304-4076(94)01658-M

Gao, J., Han, X., Pan, G., Yang, Y. (2017). High dimensional correlation matrices: the Central limit theorem and its applications. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79(3):677–693. doi:10.1111/rssb.12189

Halunga, A. G., Orme, C. D., Yamagata, T. (2017). A heteroskedasticity robust Breusch-Pagan test for contemporaneous correlation in dynamic panel data models. *Journal of Econometrics* 198(2):209–230. doi:10.1016/j.jeconom.2016.12.005

Juodis, A., Reese, S. (2021). The incidental parameter problem in testing for remaining cross-section correlation. *Journal of Business & Economic Statistics*. doi:10.1080/07350015.2021.1906687

Moran, P. A. P. (1948). The interpretation of statistical maps. *Biometrika* 35(3/4):255–260. doi:10.2307/2332344

Moscone, F., Tosetti, E. (2009). A review and comparisons of tests of cross-section independence in panels. *Journal of Economic Surveys* 23(3):528–561. doi:10.1111/j.1467-6419.2008.00571.x

Paul, D., Aue, A. (2014). Random matrix theory in statistics: a review. *Journal of Statistical Planning and Inference* 150:1–29. doi:10.1016/j.jspi.2013.09.005

Pesaran, M. H. (2015a). Testing weak cross-sectional dependence in large panels. *Econometric Reviews* 34(6–10):1089–1117. doi:10.1080/07474938.2014.956623
Pesaran, M. H. (2015b). *Time Series and Panel Data Econometrics*. Oxford: Oxford University Press.

Pesaran, M. H. (2021). General diagnostic tests for cross section dependence in panels. *Empirical Economics* 60(1): 13–50. doi:10.1007/s00181-020-01875-7

Pesaran, M. H., Ullah, A., Yamagata, T. (2008). A bias adjusted LM test of error cross-section independence. *The Econometrics Journal* 11(1):105–127. doi:10.1111/j.1368-423X.2007.00227.x

Sarafidis, V., Wansbeek, T. (2012). Cross-sectional dependence in panel data analysis. *Econometric Reviews* 31(5): 483–531. doi:10.1080/07474938.2011.611458

Sarafidis, V., Yamagata, T., Robertson, D. (2009). A test of cross section dependence for linear dynamic panel model with regressors. *Journal of Econometrics* 148(2):149–161. doi:10.1016/j.jeconom.2008.10.006

Schott, J. (2005). Testing for complete independence in high dimensions. *Biometrika* 92(4):951–956. doi:10.1093/biomet/92.4.951

Yao, J., Zheng, S., Bai, Z. (2015). *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. New York: Cambridge University Press.