THE MAXWELL SCHOOL

CENTER FOR POLICY RESEARCH

WORKING PAPER SERIES

On Testing for Sphericity with Non-normality in a Fixed Effects Panel Data Model

Badi H. Baltagi, Chihwa Kao, and Bin Peng

ISSN: 1525-3066

426 Eggers Hall
Syracuse University
Syracuse, NY 13244-1020
(315) 443-3114 / email: ctrpol@syr.edu

http://www.maxwell.syr.edu/CPR_Working_Papers.aspx

Paper No. 176
December 2014

CENTER FOR POLICY RESEARCH – Fall 2014

Leonard M. Lopoo, Director
Associate Professor of Public Administration and International Affairs (PAIA)

Associate Directors

Margaret Austin
Associate Director

Budget and Administration
John Yinger
Trustee Professor of Economics and PAIA
Associate Director, Metropolitan Studies Program

SENIOR RESEARCH ASSOCIATES

Badi H. Baltagi... Economics
Robert Bifulco..PAIA
Thomas Dennison ....PAIA
Alfonso Flores-Lagunes .... Economics
Sarah Hamersma ....PAIA
William C. Horrace .... Economics
Yilin Hou ....PAIA
Duke Kao...Economics
Sharon Kioko .... PAIA
Jeffrey Kubik...Economics
Yoonseok Lee .... Economics
Amy Lutz... Sociology
Yingyi Ma... Sociology
Jerry Miner.. Economics
Cynthia Morrow ....PAIA
Jan Ondrich...Economics
John Palmer.. PAIA
Eleonora Patacchini ....Economics
David Popp .... PAIA
Stuart Rosenthal ....Economics
Ross Rubenstein... PAIA
Rebecca Schewe ....Sociology
Amy Ellen Schwartz .... PAIA/Economics
Perry Singleton…………………………......Economics
Abbey Steele... PAIA
Michael Wasylenko... ……………………….Economics
Peter Wilcoxen..…PAIA

GRADUATE ASSOCIATES

Christian Buerger .... PAIA
Emily Cardon.. PAIA
Hannah Dalager .... PAIA
Maidel De La Cruz.. PAIA
Carlos Diaz ....Economics
Vantiel Elizabeth Duncan .... PAIA
Alex Falevich ....Economics
Lincoln Groves .... PAIA
Ruby Jennings.. PAIA
Yusun Kim .... PAIA
Bridget Lenkiewicz .... PAIA
Michelle Lofton .... PAIA
Roberto Martinez ....PAIA
Qing Miao ....PAIA
Nuno Abreu Faro E Mota............................ Economics
Judson Murchie ....PAIA
Sun Jung Oh .... Social Science
Brian Ohl ....PAIA
Laura Rodriquez-Ortiz ....PAIA
Timothy Smilnak ....PAIA
Kelly Stevens ....PAIA
Rebecca Wang .... Sociology
Pengju Zhang .... Economics
Xirui Zhang .... Economics

STAFF

Kelly Bogart.....……………..….Administrative Specialist
Karen Cimilluca....…………….…..Office Coordinator
Kathleen Nasto........................Administrative Assistant
Mary Santy.…………………….....Administrative Assistant
Katrina Wingle.......……………….Administrative Assistant

Candi Patterson……………….……Computer Consultant

Abstract

Building upon the work of Chen et al. (2010), this paper proposes a test for sphericity of the variance-covariance matrix in a fixed effects panel data regression model without the normality assumption on the disturbances.

JEL No. C13; C33

Keywords: Panel Data; Cross-sectional Dependence; John Test

We would like to thank the Associate editor and two anonymous referees for their helpful comments and suggestions.

Badi H. Baltagi: Corresponding author, Department of Economics and Center for Policy Research, 426 Eggers Hall, Syracuse, NY 13244-1020
Email Address: bbaltagi@maxwell.syr.edu

Chihwa Kao: Department of Economics and Center for Policy Research, 426 Eggers Hall, Syracuse, NY 13244-1020
Email Address: dkao@maxwell.syr.edu

Bin Peng: Department of Economics, 110 Eggers Hall, Syracuse, NY, 13244-1020
Email Address: bpeng01@syr.edu

On Testing for Sphericity with Non-normality in a Fixed Effects Panel Data Model

Badi H. Baltagi*, Chihwa Kao**, Bin Peng

Abstract

Building upon the work of Chen et al. (2010), this paper proposes a test for sphericity of the variance-covariance matrix in a fixed effects panel data regression model without the normality assumption on the disturbances.

Keywords: Sphericity; Panel Data; Cross-sectional Dependence; John Test.
JEL Classification: C13; C33.

1. Introduction

This paper proposes testing the null of sphericity of the variance-covariance matrix in a fixed effects panel data model which does not require the normality assumption on the disturbances. This builds on the paper by Chen et al. (2010) who use U-statistics to test for sphericity of the variance-covariance matrix in statistics. The null of sphericity means that the variance-covariance matrix is proportional to the identity matrix. Rejecting the null means having cross-sectional dependence among the individual units of observation or heteroskedasticity or both. In empirical economic studies, individuals are affected by common shocks. For example, investors' decisions may be influenced by the way they interact with each other and also by common macro-economic shocks or public policies. These potentially cause cross-sectional dependence among the units.

We would like to thank the Associate editor and two anonymous referees for their helpful comments and suggestions.
*Badi H. Baltagi: Corresponding author, Department of Economics and Center for Policy Research, 426 Eggers Hall, Syracuse, NY 13244-1020, USA; Email Address: bbaltagi@maxwell.syr.edu.
**Chihwa Kao: Department of Economics and Center for Policy Research, 426 Eggers Hall, Syracuse, NY 13244-1020, USA; Email Address: dkao@maxwell.syr.edu.
Bin Peng: Department of Economics, 110 Eggers Hall, Syracuse, NY, 13244-1020, USA; Email Address: bpeng01@syr.edu.

In statistics, the $n \times n$ sample covariance matrix S_n is widely used for tests of sphericity since it is a consistent estimator for the variance-covariance matrix Σ_n. One could either use the likelihood ratio test, see Anderson (2003), or test the Frobenius norm of the difference between S_n and Σ_n, see John (1971, 1972). However, with panel data sets where n the number of individuals is larger than the time series dimension of the data T, the sample covariance matrix becomes singular. This causes problems for the likelihood ratio test which is based on the inverse of S_n. Even when n is smaller than T, the sample covariance matrix S_n is ill-conditioned as shown in the Random Matrix Theory (RMT) literature. In fact, the eigenvalues of the sample covariance matrix S_n are no longer consistent for their population counterpart, see Johnstone (2001). Ledoit and Wolf (2004) show that the scaled Frobenius norm of S_n does not converge to that of Σ_n with $n/T \to c \in (0, \infty)$. As a result, John's test, see John (1971, 1972), is no longer applicable. Hence, Ledoit and Wolf (2002) propose a new test for the null of sphericity which could be applied even when n is relatively as large as T. However, these statistical tests for raw data are not directly applicable to testing sphericity in panel data regressions since the disturbances are unobservable. Baltagi et al. (2011) extend the Ledoit and Wolf (2002)'s John test to the fixed effects panel data model and correct for the bias due to substituting within residuals for the actual disturbances. However, their test relies on the normality assumption and their simulation results show that the test has size distortion under non-normality of the disturbances.

To account for the possible "non-normality" of the disturbances as well as the "large n, small T" issues in testing the null of sphericity, Chen et al. (2010) propose a modified John test by constructing U-statistics of observable samples for estimating $\mathrm{tr}\Sigma_n$ and $\mathrm{tr}\Sigma_n^2$. Based on their work, this paper proposes a new test for the null of sphericity of the disturbances in a fixed effects regression panel data model. This test does not require the assumption of normality of the disturbances, and can be applied to the case where n is larger than T. The limiting distribution of this test statistic under the null is derived. Also, its finite sample properties are studied using Monte Carlo simulations.

The paper is organized as follows. Section 2 specifies the fixed effects panel data regression model and the assumptions required. Section 3 introduces the test statistic. Section 4 derives the limiting distribution of this test statistic under the null and discusses its power properties. Section 5 reports the results of Monte Carlo simulations, while Section 6 concludes. All the proofs and technical details can be found in an Appendix available upon request from the authors.

Notation: $||B|| = (\text{tr}(B'B))^{1/2}$ is the Frobenius norm of a matrix B or the Euclidean norm of a vector B, and $\text{tr}(B)$ is the trace of B. $\xrightarrow{d}$ denotes convergence in distribution and $\xrightarrow{p}$ denotes convergence in probability. For two matrices $B = (b_{ij})$ and $C = (c_{ij})$, we define $B \circ C = (b_{ij}c_{ij})$.

2. The Model and Assumptions

Consider the following fixed effects panel data regression model

$$y_{it} = \alpha + x_{it}'\beta + \mu_i + v_{it}, \text{ for } i = 1, 2, \ldots, n; \; t = 1, 2, \ldots, T, \tag{2.1}$$

where i indexes the cross-sectional dimension and t indexes the time series dimension. y_{it} is the dependent variable, x_{it} denotes the $k \times 1$ vector of exogenous regressors, and β is the corresponding $k \times 1$ vector of parameters. μ_i denotes the time-invariant individual effects which can be fixed or random and could be correlated with the regressors. Define the vector of disturbances $v_t = (v_{1t}, \ldots, v_{nt})'$ and its corresponding variance-covariance matrix Σ_n. The null hypothesis of interest is sphericity:

$$H_0 : \Sigma_n = \sigma_v^2 I_n \quad \text{vs} \quad H_1 : \Sigma_n \neq \sigma_v^2 I_n. \tag{2.2}$$

The alternative hypothesis allows cross-sectional dependence or heteroskedasticity or both.

For the panel data regression model, v_{it} is unobserved, and the test statistic is based upon consistent estimates of variance-covariance matrix, denoted by S_n or its correlation coefficients matrix counterpart, see Breusch and Pagan (1980). Baltagi et al. (2011) extend the Ledoit and Wolf (2002) test to a fixed effects panel data model with large n and large T. They show that the noise resulting from using within residuals rather than the actual disturbances accumulates and causes bias for the proposed test statistic. However, their simulations show that their test is oversized under non-normality of the disturbances. This paper extends Chen et al. (2010) to test the null of sphericity of the variance-covariance matrix of the disturbances in a fixed effects panel data regression model without assuming normality of the disturbances. We use the within residuals which are given by

$$\hat{v}_{it} = \tilde{y}_{it} - \tilde{x}_{it}'\tilde{\beta} = v_{it} - \bar{v}_{i.} - \tilde{x}_{it}'\left(\tilde{\beta} - \beta\right), \tag{2.3}$$

where $\tilde{x}_{it} = x_{it} - \bar{x}_{i.}$ and $\bar{x}_{i.} = \frac{1}{T}\sum_{t=1}^{T} x_{it}$. Similarly, $\tilde{y}_{it} = y_{it} - \bar{y}_{i.}$, $\bar{y}_{i.} = \frac{1}{T}\sum_{t=1}^{T} y_{it}$, and $\bar{v}_{i.} = \frac{1}{T}\sum_{t=1}^{T} v_{it}$. The within estimator of β is given by $\tilde{\beta} = \left(\sum_{t=1}^{T}\sum_{i=1}^{n} \tilde{x}_{it}\tilde{x}_{it}'\right)^{-1}\left(\sum_{t=1}^{T}\sum_{i=1}^{n} \tilde{x}_{it}\tilde{y}_{it}\right)$.

Let $\tilde{y}_t = (\tilde{y}_{1t}, \ldots, \tilde{y}_{nt})'$, $\hat{v}_t = (\hat{v}_{1t}, \ldots, \hat{v}_{nt})'$, $\bar{v}_{.} = (\bar{v}_{1.}, \ldots, \bar{v}_{n.})'$, and $\tilde{x}_t = (\tilde{x}_{1t}, \ldots, \tilde{x}_{nt})$. The within residuals can be rewritten in matrix form as $\hat{v}_t = v_t - \bar{v}_{.} - \tilde{x}_t' \left(\tilde{\beta} - \beta\right)$. To facilitate our analysis, we require the following assumptions:

Assumption 1. *The $n \times 1$ vectors $v_1, v_2, \ldots, v_T$ are independent and identically distributed (i.i.d.) with mean vector 0 and covariance matrix $\Sigma_n = \Gamma\Gamma'$, where Γ is an $n \times m$ ($m \leq \infty$) matrix, v_t can be written as $v_t = \Gamma Z_t$, where $Z_t = (z_{t1}, \ldots, z_{tm})$ are i.i.d.random vectors with mean vector 0 and covariance matrix I_m. We also assume that each v_{it}, for $i = 1, \ldots, n$ has uniformly bounded 8th moment and there exists a finite constant Δ such that $E(z_{1l}^4) = 3 + \Delta$, for $l = 1, \ldots, m$.*

Assumption 2. *The regressors x_{it}, $i = 1, ..., n$, $t = 1, ..., T$ are independent of the idiosyncratic disturbances v_{it}, $i = 1, \ldots, n$, $t = 1, \ldots, T$. The regressors x_{it} have finite fourth moments: $E[||x_{it}||^4] \leq K < \infty$, where K is a positive constant.*

Assumption 3. *As $(n, T) \to \infty$, $tr(\Sigma_n^2) \to \infty$, $tr(\Sigma_n^4)/tr^2(\Sigma_n^2) \to 0$.*

The asymptotics follow the framework employed by Chen et al. (2010). Assumption 3 requires $\mathrm{tr}(\Sigma_n^4)$ to grow at a slower rate than $\mathrm{tr}^2(\Sigma_n^2)$. This assumption is flexible. In fact, if all the eigenvalues of Σ_n are bounded away from zero and infinity, $\mathrm{tr}(\Sigma_n^4)/\mathrm{tr}^2(\Sigma_n^2) \to 0$, is always true for any n as $n \to \infty$. Moreover, this assumption allows n to be much larger than T, which is more suitable for micro-panel data.

3. J_u Test

For testing the null hypothesis (2.2), the test statistic is based on the scaled distance measure between $\sigma_v^{-2}\Sigma_n$ and I_n:

$$U_0 = \frac{1}{n}\mathrm{tr}\left[S_n\left(\frac{1}{n}\mathrm{tr}S_n\right)^{-1} - I_n\right]^2 = \left(\frac{1}{n}\mathrm{tr}S_n\right)^{-2}\left(\frac{1}{n}\mathrm{tr}(S_n^2)\right) - 1, \tag{3.1}$$

where S_n is the $n \times n$ sample covariance matrix and I_n is an $n \times n$ identity matrix. With the normality assumption, John (1972) shows that for fixed n, and as $T \to \infty$:

$$\frac{nT}{2}U_0 \xrightarrow{d} \chi^2_{n(n+1)/2-1}. \tag{3.2}$$

But when n goes to infinity, the test statistic diverges. Ledoit and Wolf (2002) propose a modified test statistic under the null, as $(n,T) \to \infty$ and $n/T \to c \in (0,\infty)$:

$$TU_0 - n \xrightarrow{d} N(1,4). \tag{3.3}$$

Define $J_0 = \frac{TU_0 - n}{2} - \frac{1}{2}$, then under the null $J_0 \xrightarrow{d} N(0,1)$. However, this test cannot be used directly in a fixed effects panel data regression model. The raw data sample covariance matrix S_n is replaced by its counterpart $\hat{S}_n = \frac{1}{T}\sum_{t=1}^{T} \hat{v}_t \hat{v}_t'$, where $\hat{v}_t$ is the within residual given by (2.3). The residual-based $\hat{U}_0$ is defined as $\hat{U}_0 = \left(\frac{1}{n}\text{tr}\hat{S}_n\right)^{-2} \frac{1}{n}\text{tr}\left(\hat{S}_n^2\right) - 1$ and the corresponding residual-based J_0 test is $\hat{J}_0 = \frac{T\hat{U}_0 - n}{2} - \frac{1}{2}$. Baltagi et al. (2011) propose a bias correction:

$$J_{BFK} = \hat{J}_0 - \frac{n}{2(T-1)}. \tag{3.4}$$

They show that in a fixed effects panel data regression, as $(n,T) \to \infty$ and $n/T \to c$, $J_{BFK} \xrightarrow{d} N(0,1)$ under the null. However, their result relies on the normality assumption of v_t. Without the normality assumption, the bias-corrected John test is not robust, see the simulations in Baltagi et al. (2011).

Chen et al. (2010) propose a new test statistic for the sphericity of the variance-covariance matrix of the disturbances without the normality assumption and under much relaxed conditions where n could be much larger than T. They construct the U-statistics for estimating $\text{tr}\Sigma_n$ and $\text{tr}\Sigma_n^2$. Following their framework, we propose a residual-based test statistic for testing the null of sphericity described in (2.2) in a fixed effects panel data model. Define

$$\hat{M}_{1,T} = \frac{1}{T}\sum_{t=1}^{T} \hat{v}_t'\hat{v}_t; \quad \hat{M}_{2,T} = \frac{1}{C_T^2}\sum_{t\neq s}^{T}\sum_{s=1}^{T} \hat{v}_t'\hat{v}_s; \quad \hat{M}_{3,T} = \frac{1}{C_T^2}\sum_{t\neq s}^{T}\sum_{s=1}^{T} (\hat{v}_t'\hat{v}_s)^2;$$

$$\hat{M}_{4,T} = \frac{1}{C_T^3}\sum_{t\neq s\neq \tau}^{T}\sum_{s\neq\tau}^{T}\sum_{\tau=1}^{T} \hat{v}_t'\hat{v}_s\hat{v}_s'\hat{v}_\tau; \quad \hat{M}_{5,T} = \frac{1}{C_T^4}\sum_{t\neq s\neq\tau\neq\eta}^{T}\sum_{s\neq\tau\neq\eta}^{T}\sum_{\tau\neq\eta}^{T}\sum_{\eta=1}^{T} \hat{v}_t'\hat{v}_s\hat{v}_\tau'\hat{v}_\eta,$$

where $C_T^i = T!/(T-i)!$. Also, let $\hat{R}_1 = \hat{M}_{1,T} - \hat{M}_{2,T}$ and $\hat{R}_2 = \hat{M}_{3,T} - 2\hat{M}_{4,T} + \hat{M}_{5,T}$. If we observe the true v_t, then R_1, R_2 and $M_{j,T}$, for $j = 1,2,3,4,5$ are obtained similarly by replacing $(\hat{v}_t, \hat{v}_s, \hat{v}_\tau, \hat{v}_\eta)$ with $(v_t, v_s, v_\tau, v_\eta)$. R_1 and R_2 are unbiased estimators for $\text{tr}\Sigma_n$ and $\text{tr}\Sigma_n^2$, respectively. The scaled distance measure between $\sigma_v^{-2}\Sigma_n$ and I_n is given by $U_T = \frac{nR_2}{R_1^2} - 1$. Define

$A = \Gamma'\Gamma$ and $\psi^2 = \frac{4}{T^2} + \frac{8}{T}\text{tr}\left[\left(\frac{\Sigma_n^2}{\text{tr}(\Sigma_n^2)} - \frac{\Sigma_n}{\text{tr}(\Sigma_n)}\right)^2\right] + \frac{4\Delta}{T}\text{tr}\left[\left(\frac{A^2}{\text{tr}(\Sigma_n^2)} - \frac{A}{\text{tr}(\Sigma_n)}\right) \circ \left(\frac{A^2}{\text{tr}(\Sigma_n^2)} - \frac{A}{\text{tr}(\Sigma_n)}\right)\right]$. Chen et al. (2010) show that as $(n, T) \to \infty$:

$$\psi^{-1}\left[\left(\frac{U_T + 1}{n}\right)\left(\frac{\text{tr}^2(\Sigma_n)}{\text{tr}\left(\Sigma_n^2\right)}\right) - 1\right] \xrightarrow{d} N(0,1). \tag{3.5}$$

Let $J_{CZZ} = \frac{TU_T}{2}$, then under the null $J_{CZZ} \xrightarrow{d} N(0,1)$. Following this framework, we propose the following test statistic:

$$J_u = \frac{T}{2}\hat{U}_T = \frac{T}{2}\left(n\frac{\hat{R}_2}{\hat{R}_1^2} - 1\right). \tag{3.6}$$

J_u is the residual-based statistic corresponding to J_{CZZ}. There are two important issues to be considered. First, whether the residual-based $\hat{R}_1$ and $\hat{R}_2$ are consistent estimates for $\text{tr}\Sigma_n$ and $\text{tr}\Sigma_n^2$ under the null, respectively. Second, the asymptotics of the proposed test need to be derived. Both concerns are tackled in the next Section.

4. Asymptotics of the J_u Test

In this Section, we prove that, under the null, $\frac{1}{n}\hat{R}_1$ and $\frac{1}{n}\hat{R}_2$ are consistent estimators for $\frac{1}{n}\text{tr}\Sigma_n = \sigma_v^2$ and $\frac{1}{n}\text{tr}\Sigma_n^2 = \sigma_v^4$, respectively. Next, we show J_u converges to $N(0,1)$ under the null and we discuss its power properties. To examine the asymptotics of J_u, we rewrite it as

$$J_u = J_{CZZ} + (J_u - J_{CZZ}) = J_{CZZ} + \frac{T\left(\hat{U}_T - U_T\right)}{2}. \tag{4.1}$$

The first term J_{CZZ} is asymptotically standard normal under the null. The second term $J_u - J_{CZZ}$ is the scaled difference between the residual-based $\hat{U}_T$ and the true U_T. From Section 3, this difference can be rewritten as follows:

$$J_u - J_{CZZ} = \frac{T}{2}\left[\left(\frac{1}{n}\hat{R}_2\right)\left(\frac{1}{n}R_1\right)^2 - \left(\frac{1}{n}R_2\right)\left(\frac{1}{n}\hat{R}_1\right)^2\right]\left(\frac{1}{n}\hat{R}_1\right)^{-2}\left(\frac{1}{n}R_1\right)^{-2}. \tag{4.2}$$

From equation (4.2), it is clear that this term depends upon the two differences: $\frac{1}{n}\hat{R}_1 - \frac{1}{n}R_1$ and $\frac{1}{n}\hat{R}_2 - \frac{1}{n}R_2$. Their asymptotic behavior is given in the following propositions:

Proposition 1. *Under Assumptions 1-2 and the null, (1)* $\frac{1}{n}\hat{M}_{1,T} = \frac{1}{n}M_{1,T} - \frac{\sigma_v^2}{T} + O_p\left(\frac{1}{T\sqrt{n}}\right)$; *(2)* $\frac{1}{n}\hat{M}_{2,T} = \frac{1}{n}M_{2,T} - \frac{\sigma_v^2}{T} + O_p\left(\frac{1}{T\sqrt{n}}\right)$; *(3)* $\frac{1}{n}\hat{R}_1 - \frac{1}{n}R_1 = O_p\left(\frac{1}{nT}\right)$.

Proposition 2. *Under Assumptions 1-2 and the null, (1)* $\frac{1}{n}\hat{M}_{3,T} = \frac{1}{n}M_{3,T} + \frac{n-2T-6}{T^2}\sigma_v^4 + O_p\left(\frac{\sqrt{n}}{T^2}\right) + O_p\left(\frac{1}{T\sqrt{T}}\right) + O_p\left(\frac{1}{T\sqrt{n}}\right)$; *(2)* $\frac{1}{n}\hat{M}_{4,T} = \frac{1}{n}M_{4,T} + \frac{n-T-5}{T^2}\sigma_v^4 + O_p\left(\frac{1}{T\sqrt{n}}\right) + O_p\left(\frac{\sqrt{n}}{T^2}\right) + O_p\left(\frac{1}{T\sqrt{T}}\right)$; *(3)* $\frac{1}{n}\hat{M}_{5,T} = \frac{1}{n}M_{5,T} + \frac{n-4}{T^2}\sigma_v^4 + O_p\left(\frac{\sqrt{n}}{T^2}\right)$; *(4)* $\frac{1}{n}\hat{R}_2 - \frac{1}{n}R_2 = O_p\left(\frac{1}{T^2}\right) + O_p\left(\frac{1}{nT}\right) + O_p\left(\frac{1}{T\sqrt{nT}}\right)$.

Propositions 1 and 2 show that the differences $\frac{1}{n}\hat{R}_1 - \frac{1}{n}R_1$ and $\frac{1}{n}\hat{R}_2 - \frac{1}{n}R_2$ vanish as $(n,T) \to \infty$. Therefore, since $\frac{1}{n}R_1 \xrightarrow{p} \sigma_v^2$ and $\frac{1}{n}R_2 \xrightarrow{p} \sigma_v^4$, we conclude that $\frac{1}{n}\hat{R}_1$ and $\frac{1}{n}\hat{R}_2$ are consistent estimates for σ_v^2 and σ_v^4 respectively. The following corollary gives these conclusions:

Corollary 1. *Under Assumptions 1-2 and the null, as* $(n,T) \to \infty$*, (1)* $\frac{1}{n}\hat{R}_1 \xrightarrow{p} \sigma_v^2$*; (2)* $\frac{1}{n}\hat{R}_2 \xrightarrow{p} \sigma_v^4$.

Note that $\frac{1}{n}\hat{R}_2$ is a consistent estimator of σ_v^4 under the null with large n and large T. However, $\frac{1}{n}\mathrm{tr}\hat{S}_n^2$ is not consistent, see Baltagi et al. (2011).

Proposition 3. *Under Assumptions 1-2 and the null,* $\frac{T(\hat{U}_T - U_T)}{2} = O_p\left(\frac{1}{n}\right) + O_p\left(\frac{1}{T}\right) + O_p\left(\frac{1}{\sqrt{nT}}\right)$.

Propositions 1, 2 and 3 give the asymptotics of the bias term $J_u - J_{CZZ}$. Compared with the statistic based on raw data, the test statistic based on the within residuals, defined below equation (2.3), can be expressed by $\hat{v}_t = v_t - \bar{v}_{.} - \tilde{x}_t'\left(\tilde{\beta} - \beta\right)$. This has the additional terms $\bar{v}_{.}$ and $\tilde{x}_t'\left(\tilde{\beta} - \beta\right)$. These two terms can be regarded as extra noise resulting from the regression. Based on equations (4.1) and (4.2), the extra noise $\frac{T(\hat{U}_T - U_T)}{2}$ depends upon $\frac{1}{n}\hat{R}_1 - \frac{1}{n}R_1 = \frac{1}{n}(\hat{M}_{1,T} - M_{1,T}) - \frac{1}{n}(\hat{M}_{2,T} - M_{2,T})$ and $\frac{1}{n}\hat{R}_2 - \frac{1}{n}R_2 = \frac{1}{n}(\hat{M}_{3,T} - M_{3,T}) - \frac{2}{n}(\hat{M}_{4,T} - M_{4,T}) + \frac{1}{n}(\hat{M}_{5,T} - M_{5,T})$. Hence the magnitude of $\frac{T(\hat{U}_T - U_T)}{2}$ depends upon how $\bar{v}_{.}$ and $\tilde{x}_t'(\tilde{\beta} - \beta)$ accumulate in $\left(\frac{1}{n}\hat{M}_{j,T} - \frac{1}{n}M_{j,T}\right)$, for $j = 1,2,3,4,5$. Note that $\bar{v}_{.}$ is an n dimensional vector, although each element of $\bar{v}_{.}$ is $O_p(\frac{1}{\sqrt{T}})$, $\bar{v}_{.}$ may still accumulate in the above five terms as $(n,T) \to \infty$, depending upon the relative speed of n and T. $\tilde{x}_t'(\tilde{\beta} - \beta)$ is $O_p\left(\frac{1}{\sqrt{nT}}\right)$ which is related to both n and T. We may expect its convergence speed $\frac{1}{\sqrt{nT}}$ to be fast enough so that $\tilde{x}_t'(\tilde{\beta} - \beta)$ vanishes as $(n,T) \to \infty$. More specifically, Proposition 2 shows the leading terms of $\frac{1}{n}\hat{M}_{j,T} - \frac{1}{n}M_{j,T}$, for $j = 3,4,5$ will not vanish if $\frac{n}{T^2}$ does not converge to zero. These terms are caused by the accumulation of $\bar{v}_{.}$. However, due to the subtraction formulation of the test statistic, the leading terms cancel each other in both $\frac{1}{n}\hat{R}_1 - \frac{1}{n}R_1$ and $\frac{1}{n}\hat{R}_2 - \frac{1}{n}R_2$. Similar cancellations occur for other terms which are $O_p\left(\frac{\sqrt{n}}{T^2}\right)$, $O_p\left(\frac{1}{T\sqrt{n}}\right)$ and $O_p\left(\frac{1}{T\sqrt{T}}\right)$ since their expressions are exactly the same. These cancellations lead us to $\frac{1}{n}\hat{R}_1 - \frac{1}{n}R_1 = O_p\left(\frac{1}{nT}\right)$ and $\frac{1}{n}\hat{R}_2 - \frac{1}{n}R_2 = O_p\left(\frac{1}{T^2}\right) + O_p\left(\frac{1}{nT}\right) + O_p\left(\frac{1}{T\sqrt{nT}}\right)$, and consequently $\frac{T(\hat{U}_T - U_T)}{2} = O_p(\frac{1}{n}) + O_p(\frac{1}{T}) + O_p\left(\frac{1}{\sqrt{nT}}\right)$. Therefore, $J_u - J_{CZZ} \xrightarrow{p} 0$ as $(n,T) \to \infty$ and we do not need to correct the bias in the fixed effects panel data regression model. This result is based on our detailed calculation of how $\bar{v}_{.}$ and $\tilde{x}_t'(\tilde{\beta}-\beta)$ are accumulating in $\left(\frac{1}{n}\hat{M}_{j,T}-\frac{1}{n}M_{j,T}\right)$, for $j=1,2,3,4,5$ and the special formulation of J_{CZZ}. As discussed above, the convergence of J_u is given by the following theorem:

Theorem 1. *Under Assumptions 1-3 and the null, in the fixed effects panel data regression model (2.1), as* $(n,T)\to\infty$

$$J_u \xrightarrow{d} N(0,1). \tag{4.3}$$

Under the alternative, the limiting distribution of J_u is the same as (3.5) if $\frac{T\left(\hat{U}_T-U_T\right)}{2}$ vanishes as $(n,T)\to 0$. Similar to Chen et al. (2010), we consider an alternative: $H_1:\Sigma_n=\left(\sigma_i\sigma_j\rho^{|j-i|}\right)_{n\times n}$, where $\rho\in(-1,1)$ and $\rho\neq 0$. $\sigma_l^2=\mathrm{var}(v_{lt})$, which is uniformly bounded away from infinity and zero, for $l=1,\ldots,n$. Under this alternative, we can show that $J_u-J_{CZZ}=o_p(1)$, which in turn implies that J_u and J_{CZZ} have the same power properties. Define $\delta_{1,T}=1-\frac{\mathrm{tr}^2(\Sigma_n)}{n\mathrm{tr}(\Sigma_n^2)}$ and $\delta_{2,T}=\mathrm{tr}\left[\left(\frac{\Sigma_n^2}{\mathrm{tr}(\Sigma_n^2)}-\frac{\Sigma_n}{\mathrm{tr}(\Sigma_n)}\right)^2\right]$. One can show that $T\delta_{1,T}\to\infty$ and $\delta_{2,T}/(T\delta_{1,T}^2)\to 0$ as $(n,T)\to\infty$. This satisfies the conditions of Theorem 4 in Chen et al. (2010). By using this Theorem, the corresponding power function $P(J_u\geq z_\alpha|\Sigma_n=(\sigma_i\sigma_j\rho^{|j-i|})_{n\times n})\to 1$, as $(n,T)\to\infty$, where z_α is the upper quantile of $N(0,1)$. Let us consider a special case under this alternative. More specifically, assume that $\Delta=0$, $\sigma_i=\sigma_j=\sigma_v$ for any (i,j) and $T/n\to 0$ as $(n,T)\to\infty$. It follows that $\psi^{-1}\to\frac{T}{2}$ and $\left(1-\rho^2\right)J_u-T\rho^2/2\xrightarrow{d}N(0,1)$.

5. Monte Carlo Simulations

We conduct Monte Carlo experiments to assess the empirical size and power of the J_u test proposed in this paper. We follow the design of Baltagi et al. (2011) and assume homoskedasticity on the remainder error term. In this case, the J_u test becomes a test for cross-sectional dependence. We also report the performance of J_{BFK} proposed by Baltagi et al. (2011) for comparison purposes.

5.1. Experiment Design

Consider the following data-generating process:

$$y_{it} = \alpha+\beta x_{it}+\mu_i+v_{it},\quad i=1,\ldots,n;\ t=1,\ldots,T, \tag{5.1}$$

$$x_{it} = \lambda x_{i,t-1}+\mu_i+\eta_{it}, \tag{5.2}$$

where μ_i is the fixed effects and v_{it} is the idiosyncratic error, $\eta_{it} \sim i.i.d.\ N(\phi_\eta, \sigma_\eta^2)$. The regressor x_{it} is allowed to be correlated with the μ_i's. This follows the design by Im et al. (1999).

To study the power of the tests, we consider two different types of cross-sectional dependence models: a factor model and a spatial model. For the factor model, see Pesaran (2004), Pesaran and Tosetti (2011), Baltagi et al. (2011), we assume:

$$v_{it} = \gamma_i f_t + \epsilon_{it}, \tag{5.3}$$

where f_t $(t = 1, \ldots, T)$ are the factors and γ_i $(i = 1, \ldots, n)$ are the loadings. For the spatial model, we consider a first-order spatial autocorrelation model SAR(1), see Anselin and Bera (1998) and Baltagi et al. (2003), given by:

$$v_{it} = \delta(0.5v_{i-1,t} + 0.5v_{i+1,t}) + \epsilon_{it}. \tag{5.4}$$

The ϵ_{it} in (5.3) and (5.4) are assumed to be $i.i.d.(0, \sigma_\epsilon^2)$ across individuals and over time. Under the null, we have $\gamma_i = 0$ and $\delta = 0$.

Under the null, the v_{it} comes from some $i.i.d.$ distribution across individuals and over time with mean zero and variance σ_v^2. These are not necessarily normally distributed. For models (5.1) and (5.2), we set $\alpha = 1$ and $\beta = 2$; μ_i is drawn from $i.i.d.\ N(\phi_\mu, \sigma_\mu^2)$ with $\phi_\mu = 0$ and $\sigma_\mu^2 = 0.25$. We also set $\lambda = 0.7$, $\phi_\eta = 0$ and $\sigma_\eta^2 = 1$. For models (5.3) and (5.4), $\gamma_i \sim i.i.d.\ U(-0.5, 0.55)$; f_t is set to be $i.i.d.\ N(0, 1)$ and $\delta = 0.4$. Various distributions are considered in generating the model errors, v_{it} in (5.1) and ϵ_{it} in (5.3) and (5.4) are assumed be normal, lognormal, gamma, chi-squared with mean zero and variance 0.5.

The Monte Carlo experiments are conducted for $n = 20, 40, 60, 80, 100, 200, 400$ and $T = 20, 40, 60, 80$. We perform 1,000 replications to compute the J_u and J_{BFK} test statistics. We conduct the tests at the positive one-sided 5% nominal significance level to obtain the empirical size.

5.2. Results

Table 1 gives the empirical size of the J_u and J_{BFK} tests allowing v_{it} to be generated from different distributions. When the disturbances are normally distributed, the size of J_u and J_{BFK} are both close to 5%, which is consistent with the theoretical results. The rest of Table 1 shows the results with v_{it} coming from alternative non-normal distributions. The size of J_u is close to 5% when n and T are large; for small n or small T, it is slightly oversized. However, J_{BFK} is no longer robust to non-normality and suffers from size distortions.

Table 2 presents the size adjusted power of the tests under the alternative specification of a factor model. Both tests have size adjusted power that is almost 1 when n and T are large with v_{it} normally distributed. For small n and small T, the size adjusted power of J_u works as well as J_{BFK}. Note that the size adjusted power of J_{BFK} is quite good even when n is a lot larger than T for the normal distribution scenario. However, for non-normal distributions, the size adjusted power of J_u is 1 as n and T become large; and it is larger than the size adjusted power of J_{BFK} for all (n, T) combinations.

Table 3 reports the size adjusted power of both tests under the alternative specification of SAR(1). The results are similar to the factor model. J_u works as well as J_{BFK} for the normal distribution scenario, but better for all combinations of n and T for non-normal distribution scenarios.

6. Conclusion

Though the John test proposed by Baltagi et al. (2011) has been shown to perform well for a large panel data regression model with fixed effects, it relies heavily on the normality assumption. This paper proposes a new test, J_u, for the null of sphericity of the disturbances which does not rely on the normality assumption. Instead of $n/T \to c$, we allow n to be a larger order of T which is consistent with micro-panel data sets with "large n and small T" .

References

Anderson, T. W., 2003. An Introduction to Multivariate Statistical Analysis, Wiley.

Anselin, L., and Bera, A. K., 1998. Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. Handbook of Applied Economic Statistics, Amman Ullah and David E. A. Giles, eds., Marcel Dekker, New York.

Bai, Z.D., and Saranadasa, H., 1996. Effect of High Dimension: By an Example of a Two Sample Problem. Statistical Sinica. 6, 311-329.

Baltagi, B. H., Song, S. H., and Koh, W., 2003. Testing Panel Data Regression Models with Spatial Error Correlation. Journal of Econometrics. 117, 123-150.

Baltagi, B. H., Feng, Q., and Kao, C., 2011. Testing for Sphericity in a Fixed Effects Panel Data Model. The Econometrics Journal. 14, 25-47.

Breusch, T. S., and Pagan, A. R., 1980. The Lagrange Multiplier Test and Its Application to Model Specification in Econometrics. Review of Economic Studies. 47, 239-254.

Chen, S., Zhang, L., and Zhong, P., 2010. Tests for High-dimensional Covariance Matrices. Journal of the American Statistical Association. 105, 810-819.

Im, K. S., Ahn, S. C., Schmidt, P., and Wooldridge, J. M., 1999. Efficient Estimation of Panel Data Models with Strictly Exogenous Explanatory Variables. Journal of Econometrics. 93, 177-201.

John, S., 1971. Some Optimal Multivariate Test. Biometrika. 58, 123-127.

John, S., 1972. The Distribution of a Statistic Used for Testing Sphericity of Normal Distributions. Biometrika. 59, 169–173.

Johnstone, I., 2001. On the Distribution of the Largest Principal Component. Annals of Statistics. 29, 295-327.

Ledoit, O., and Wolf, M., 2002. Some Hypothesis Tests for the Covariance Matrix When the Dimension is Large Compared to the Sample Size. Annals of Statistics. 30, 1081-1102.

Ledoit, O., and Wolf, M., 2004. A Well-conditioned Estimator for Large-Dimensional Covariance Matrices. Journal of Multivariate Analysis. 88, 365-411.

Pesaran, M. H., 2004. General Diagnostic Test for Cross Section Dependence in Panels. CESifo Working Paper, No.1229, University of Cambridge.

Pesaran, M. H., and Tosetti, E., 2011. Large Panels with Common Factors and Spatial Correlations. Journal of Econometrics. 161, 182-202.

Table 1: Size of Tests

		n						
Normal Errors	T	20	40	60	80	100	200	400
J_u	20	6.4	7.1	6.5	8.0	7.8	7.2	6.3
	40	5.6	7.0	4.9	4.7	5.8	5.9	6.0
	60	6.7	7.1	6.5	5.7	5.4	5.5	5.9
	80	5.2	5.6	4.9	7.1	5.8	5.1	4.4
J_{BFK}	20	6.4	6.7	5.6	6.6	6.9	6.1	5.8
	40	5.8	6.7	5.0	4.9	5.6	6.5	5.1
	60	6.5	6.6	6.7	5.9	4.8	5.0	5.9
	80	5.0	5.3	4.6	6.7	6.1	4.7	4.7
Gamma Errors								
J_u	20	7.2	6.8	8.0	7.3	5.5	7.7	6.9
	40	7.1	5.6	7.0	6.8	5.1	5.3	5.1
	60	7.4	7.0	5.2	6.0	6.0	5.4	5.1
	80	6.2	5.9	5.6	5.2	6.5	5.3	5.6
J_{BFK}	20	16.0	17.8	19.4	20.1	18.6	21.3	18.3
	40	17.4	17.5	21.0	19.3	17.0	19.7	18.2
	60	19.5	19.9	16.4	19.3	18.0	18.6	18.5
	80	18.5	18.8	17.7	18.5	19.3	18.2	18.8
Lognormal Errors								
J_u	20	9.3	7.9	6.8	7.6	7.2	6.2	7.1
	40	8.0	8.0	5.7	6.3	6.9	6.7	6.4
	60	8.3	6.4	6.6	6.5	5.9	5.3	5.4
	80	7.0	6.3	7.1	6.0	5.8	5.0	6.0
J_{BFK}	20	27.1	26.9	27.9	27.8	28.5	28.0	29.0
	40	26.5	30.2	27.0	29.0	28.3	29.7	28.7
	60	25.4	27.1	29.9	29.7	30	30.3	30.9
	80	26.2	26.7	29.0	28.1	28.4	32.0	30.1
Chi-squared Errors								
J_u	20	8.4	8.2	7.6	7.9	7.7	6.9	7.6
	40	8.6	6.8	6.6	6.3	5.0	7.3	6.2
	60	8.4	8.1	7.6	5.6	4.6	5.5	5.4
	80	8.0	6.4	7.1	7.3	4.9	6.5	6.0
J_{BFK}	20	26.6	26.2	28.7	29.8	30.7	29.6	31.9
	40	27.9	29.3	31.5	31.9	31.0	32.4	33.2
	60	30.6	33.5	33.6	31.6	32.3	32.5	28.9
	80	30.7	30.1	35.0	34.1	32.9	32.4	31.8

Note: This table reports the size of J_u and J_{BFK} with different error distribution specification in a fixed effects panel data model without cross-sectional dependence among the errors. The tests are one-sided and are conducted at the 5% nominal significance level. We conduct the simulation with four distributions: normal, gamma, lognormal and chi-squared with mean 0, and variance 0.5.

Table 2: Size adjusted power of tests: factor model

		n						
Normal Errors	T	20	40	60	80	100	200	400
Ju	20	73.1	94.0	98.3	99.5	99.7	99.8	100
	40	95.6	99.8	99.9	100	100	100	100
	60	99.3	100	100	100	100	100	100
	80	99.8	100	100	100	100	100	100
J_{BFK}	20	73.4	94.7	98.3	99.5	99.9	99.9	100
	40	95.8	99.8	99.9	100	100	100	100
	60	99.4	100	100	100	100	100	100
	80	99.8	100	100	100	100	100	100
Gamma Errors								
J_u	20	68.2	93.3	97.1	99.1	99.6	100	100
	40	94.6	99.7	100	100	100	100	100
	60	99.1	100	100	100	100	100	100
	80	99.6	100	100	100	100	100	100
J_{BFK}	20	60.1	89.0	96.1	98.5	99.2	99.9	100
	40	91.5	99.5	99.9	100	100	100	100
	60	98.1	100	100	100	100	100	100
	80	99.2	100	100	100	100	100	100
Lognormal Errors								
Ju	20	68.1	91.6	98.2	99.2	99.4	99.9	100
	40	95.3	99.7	100	100	100	100	100
	60	99.6	100	100	100	100	100	100
	80	99.9	100	100	100	100	100	100
J_{BFK}	20	48.7	85.2	95.5	97.6	98.3	99.9	100
	40	88.3	99	100	100	100	100	100
	60	97.9	100	100	100	100	100	100
	80	99.5	100	100	100	100	100	100
Chi-squared Errors								
J_u	20	70.1	90.3	98	98.9	99.4	100	100
	40	94.1	100	100	100	100	100	100
	60	98.6	100	100	100	100	100	100
	80	99.6	100	100	100	100	100	100
J_{BFK}	20	53.8	80.4	93.5	97.8	98.3	100	100
	40	84.8	99.3	100	100	100	100	100
	60	96.1	100	100	100	100	100	100
	80	98.6	100	100	100	100	100	100

Note: This table computes the size adjusted power for a factor structure model that allows for cross-sectional dependence in the error. We conduct the simulation with four distributions: normal, gamma, lognormal and chi-squared with mean 0, and variance 0.5.

Table 3: Size adjusted power of tests: SAR(1) model

		n						
Normal Errors	T	20	40	60	80	100	200	400
J_u	20	81.4	83.7	86.0	82.1	83.3	84.8	88.0
	40	99.9	99.9	100	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
J_{BFK}	20	82	87.1	89.7	87.5	86.4	88.3	90.6
	40	100	99.9	100	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
Gamma Errors								
J_u	20	75.1	81.6	84.3	84.1	87.3	85.0	86.2
	40	99.9	100	100	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
J_{BFK}	20	74	78.3	82.2	83.3	83.8	84.1	84.9
	40	99.9	100	99.9	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
Lognormal Errors								
J_u	20	71.4	80.2	83.5	83.5	85.2	87.8	88.6
	40	99.9	100	100	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
J_{BFK}	20	61.4	72.3	79.8	80.4	80.1	86.1	84.8
	40	99.4	100	99.9	99.9	100	99.9	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
Chi-squared Errors								
J_u	20	74.2	79.5	83.6	84.2	84.8	86.3	84.5
	40	99.7	100	99.7	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100
J_{BFK}	20	65.6	70.3	79.5	77.9	76.7	82.6	78.4
	40	99.4	99.7	99.7	100	100	100	100
	60	100	100	100	100	100	100	100
	80	100	100	100	100	100	100	100

Note: This table computes the size adjusted power for a SAR(1) structure model that allows for cross-sectional dependence in the error. We conduct the simulation with four distributions: normal, gamma, lognormal and chi-squared with mean 0, and variance 0.5.