

He, Ming; Lin, Kuan-Pin

## Article

# Testing in a random effects panel data model with spatially correlated error components and spatially lagged dependent variables

Econometrics

### Provided in Cooperation with:

MDPI – Multidisciplinary Digital Publishing Institute, Basel

*Suggested Citation:* He, Ming; Lin, Kuan-Pin (2015) : Testing in a random effects panel data model with spatially correlated error components and spatially lagged dependent variables, *Econometrics*, ISSN 2225-1146, MDPI, Basel, Vol. 3, Iss. 4, pp. 761-796, <https://doi.org/10.3390/econometrics3040761>

This Version is available at:

<https://hdl.handle.net/10419/171856>

#### Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

#### Terms of use:

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



<http://creativecommons.org/licenses/by/4.0/>

Article

## Testing in a Random Effects Panel Data Model with Spatially Correlated Error Components and Spatially Lagged Dependent Variables

Ming He <sup>1,\*</sup> and Kuan-Pin Lin <sup>2</sup>

<sup>1</sup> Department of Economics, University of Washington, Seattle, WA 98195, USA

<sup>2</sup> Department of Economics, Portland State University, Portland, OR 97201, USA;  
E-Mail: link@pdx.edu

\* Author to whom correspondence should be addressed; E-Mail: hem611@uw.edu;  
Tel.: +1-503-863-9798.

Academic Editor: Kerry Patterson

Received: 6 August 2015 / Accepted: 26 October 2015 / Published: 9 November 2015

---

**Abstract:** We propose a random effects panel data model with both spatially correlated error components and spatially lagged dependent variables. We focus on diagnostic testing procedures and derive Lagrange multiplier (LM) test statistics for a variety of hypotheses within this model. We first construct the joint LM test for both the individual random effects and the two spatial effects (spatial error correlation and spatial lag dependence). We then provide LM tests for the individual random effects and for the two spatial effects separately. In addition, in order to guard against local model misspecification, we derive locally adjusted (robust) LM tests based on the Bera and Yoon principle (Bera and Yoon, 1993). We conduct a small Monte Carlo simulation to show the good finite sample performances of these LM test statistics and revisit the cigarette demand example in Baltagi and Levin (1992) to illustrate our testing procedures.

**Keywords:** individual random effects; spatial error correlation; spatial lag dependence; lagrange multiplier (LM) test; (robust) LM test

**JEL classifications:** C12; C21; C23

---

## 1. Introduction

Spatial econometric models have been extensively used to study regional effects and interdependence between different spatial units. Most of the widely used spatial models are variants of the benchmark models developed in Cliff and Ord (1973, 1981) [1,2] and Anselin (1988a) [3]. Based on the form of spatially correlated error components and/or spatially lagged dependent variables, these models better fit the real world data generating process by explicitly considering the spatial interdependence. Hypothesis testing for spatial dependence has been developed rapidly in the recent literature. For standard LM tests of spatial dependence in cross section models, see Anselin (1988a, b) [3,4], Anselin and Bera (1998) [5], and Anselin (2001) [6]. For standard LM tests of spatial dependence in panel data models, Baltagi *et al.*, (2003) [7] provide tests for random effects and/or spatial error correlation. Baltagi and Liu (2008) [8] provide tests for random effects and/or spatial lag dependence. Debarsy and Ertur (2010) [9] derive tests in the spatial panel data model with individual fixed effects based on Lee and Yu (2010) [10]. Qu and Lee (2012) [11] consider tests in spatial models with limited dependent variables. Baltagi *et al.*, (2013) [12] extend the model in Kapoor *et al.*, (2007) [13] by allowing for different spatial correlation parameters in the individual random effects and in the disturbances, and they derive the corresponding LM tests. Further, standardized versions of the LM tests are discussed in Yang (2010) [14], Baltagi and Yang (2013a) [15] to remedy distributional misspecifications in finite sample and sensitivity to spatial layout. Born and Breitung (2011) [16], Baltagi and Yang (2013b) [17] discuss versions of LM tests that are robust against unknown heteroskedasticity. Recently, Yang (2015) [18] provides residual-based bootstrap procedure to obtain improved approximations to the finite sample critical values of the LM test statistics in spatial econometric models.

However, to the best of our knowledge, there are no test statistics treating the individual random effects, the spatial error correlation, and the spatial lag dependence simultaneously. We contribute to the literature by constructing various LM test statistics in such a general framework, or the so-called spatial autoregressive model with autoregressive disturbances (SARAR). Our results are useful for applied researchers to implement and perform model diagnostic testing in the SARAR framework. In particular, we first derive the joint LM test for the individual random effects and the two spatial effects. We next derive LM tests for the individual random effects. Finally, we derive LM tests for the two spatial effects. In addition, we provide robust LM tests in some cases as needed in order to guard against local misspecification. We emphasize some key features of the robust LM test in the following.

Bera and Yoon (1993) [19] argue that the LM test with specific values of the nuisance parameters (marginal LM test) might suffer from local misspecification in the nuisance parameters. They propose robust LM test to guard against such local misspecification, see also Anselin *et al.*, (1996) [20], Bera *et al.*, (2001, 2009, 2010) [21–23], and He and Lin (2013) [24]. Here, we emphasize two advantages of the robust LM test. First, the asymptotic size of marginal LM test will be distorted under local misspecification in the nuisance parameters since it follows a non-central  $\chi^2$  distribution. On the other hand, robust LM test follows a  $\chi^2$  distribution under such misspecification, thus it can provide valid asymptotic size as long as the misspecification is local. Second, while LM test without specifying values of the nuisance parameters (conditional LM test) does not suffer from such size distortion, it generally needs the maximum likelihood estimator (MLE), which could be costly in computation. In contrast, the

robust LM test only requires restricted estimator under the relevant joint null hypothesis, which is simply the ordinary least square (OLS) estimator in most cases. Therefore, the robust LM test can provide result as good as the conditional LM test at a lower computational cost, provided that the deviation of nuisance parameters is local. However, there is one potential loss in using the robust LM test. If the values of nuisance parameters are correctly specified, the robust LM test is in general less powerful than the marginal LM test. Also, when the nuisance parameters deviate far away from the pre-specified values, the robust LM test is generally invalid. In sum, the standard LM tests (marginal and conditional LM tests) and the robust LM tests complement each other, and they should be used together for inference purposes.

In this paper, we maintain the assumption of random effects model, while an alternative specification is the fixed effects model with spatial dependence as in Lee and Yu (2010) [10], Debarsy and Ertur (2010) [9], and He and Lin (2013) [24]. On the one hand, the random effects specification is a parsimonious way to allow for individual effects in different spatial units and it will be particularly useful for testing and selection in microeconomic applications when the number of units is very large. On the other hand, the fixed effects specification, which allows for correlation between the individual effects and the covariates, is more suited for many macro studies when the number of units is not very large (see Elhorst (2014) [25] for more discussion on comparison of the random effects model and the fixed effects model).

The rest of the paper is organized as follows. Model specification is discussed in Section 2. The LM test statistics are presented In Section 3. In Section 4, we report the Monte Carlo simulation results to show their satisfactory finite sample size and power performances. In Section 5, we provide an empirical example to illustrate our testing procedures. Section 6 concludes with suggestions for future research. All mathematical derivations are relegated to the Appendices.

## 2. The Model

Suppose that the data is generated according to the following spatial panel data model, for  $t = 1, 2, \dots, T$ ,

$$\begin{cases} y_t = \lambda W y_t + X_t \beta + \epsilon_t, \\ \epsilon_t = \rho M \epsilon_t + \mu + v_t. \end{cases} \quad (2.1)$$

In the above specification,  $y_t = (y_{t1}, y_{t2}, \dots, y_{tN})'$  is an  $N \times 1$  vector of dependent variable for period  $t$ . It is spatially interdependent, as reflected by the spatial lag dependence coefficient  $\lambda$ .  $X_t$  is an  $N \times (K + 1)$  matrix of non-stochastic regressors for period  $t$ , with the first column to be ones, and  $\beta = (\beta_0, \beta_1, \dots, \beta_K)'$  is the corresponding  $(K + 1) \times 1$  slope parameter vector.  $\epsilon_t = (\epsilon_{t1}, \epsilon_{t2}, \dots, \epsilon_{tN})'$  is an  $N \times 1$  vector of the regression error term for period  $t$ . The error vector is also spatially correlated, as reflected by the spatial error correlation coefficient  $\rho$ .  $\mu = (\mu_1, \mu_2, \dots, \mu_N)'$  is an  $N \times 1$  vector representing the individual random effects. The random effects terms  $\{\mu_i\}, i = 1, \dots, N$  are i.i.d. across  $i$ , with zero mean, variance  $\sigma_\mu^2$  and  $E|\mu_i|^{4+c_1} < \infty$  for some  $c_1 > 0$ .  $v_t = (v_{t1}, v_{t2}, \dots, v_{tN})'$  is an  $N \times 1$  vector of innovation terms. The innovation terms  $\{v_{ti}\}, t = 1, \dots, T, i = 1, \dots, N$  are i.i.d. across  $i$  and  $t$ , with zero mean, variance  $\sigma_v^2$  and  $E|v_{ti}|^{4+c_2} < \infty$  for some  $c_2 > 0$  (see Lee and Yu (2012) [26] for other regularity and identification conditions as needed for asymptotic theory).  $W$  and  $M$  are nonstochastic spatial weights

matrices of size  $N \times N$ . Typically, they are specified by the first-order rook contiguity criterion and are row-standardized so that each row sums up to one.

By stacking across  $t$ , the model can be written as

$$\begin{cases} y = \lambda(I_T \otimes W)y + X\beta + \epsilon, \\ \epsilon = \rho(I_T \otimes M)\epsilon + \iota_T \otimes \mu + v, \end{cases} \tag{2.2}$$

where  $y = (y'_1, y'_2, \dots, y'_T)'$ ,  $X = (X'_1, X'_2, \dots, X'_T)'$ ,  $\epsilon = (\epsilon'_1, \epsilon'_2, \dots, \epsilon'_T)'$ , and  $v = (v'_1, v'_2, \dots, v'_T)'$ .  $\iota_T$  is a  $T \times 1$  vector of ones,  $I$  denotes the identity matrix with its dimension in the subscript, and  $\otimes$  denotes the kronecker product. Let  $A = [I_T \otimes (I_N - \rho M)]$ , and  $B = [I_T \otimes (I_N - \lambda W)]$ . The error component  $\epsilon$  is expressed as  $\epsilon = A^{-1}(\iota_T \otimes \mu + v)$ , with  $E[\epsilon] = 0$ ,  $\text{Var}[\epsilon] = \Omega_\epsilon = A^{-1}\Omega(A^{-1})'$ , where  $\Omega = [J_T \otimes (\sigma_\mu^2 I_N) + I_T \otimes (\sigma_v^2 I_N)]$ . Using the results in Magnus (1982) [27], we get  $\Omega^{-1} = (T\sigma_\mu^2 + \sigma_v^2)^{-1} \bar{J}_T \otimes I_N + (\sigma_v^2)^{-1} E_T \otimes I_N$ , and  $|\Omega| = (T\sigma_\mu^2 + \sigma_v^2)^N (\sigma_v^2)^{N(T-1)}$ , where  $\bar{J}_T = J_T/T$ ,  $J_T = \iota_T \iota'_T$ ,  $E_T = I_T - \bar{J}_T$ . Notice that  $y = B^{-1}X\beta + B^{-1}\epsilon$ , with  $E[y] = B^{-1}X\beta$  and  $\text{Var}[y] = \Omega_y = B^{-1}A^{-1}\Omega(A^{-1})'(B^{-1})'$ . We have  $\Omega_y^{-1} = B'A'\Omega^{-1}AB$ , and  $|\Omega_y| = |B|^{-2}|A|^{-2}(T\sigma_\mu^2 + \sigma_v^2)^N (\sigma_v^2)^{N(T-1)}$ .

Let  $\delta = (\beta', \rho, \lambda, \sigma_\mu^2, \sigma_v^2)'$  and  $\epsilon = By - X\beta$ , the log-likelihood function of the random vector  $y$  as it is normally distributed is

$$L(\delta) = -\frac{NT}{2} \ln(2\pi) - \frac{N}{2} \ln(T\sigma_\mu^2 + \sigma_v^2) - \frac{N(T-1)}{2} \ln(\sigma_v^2) + \ln|A| + \ln|B| - \frac{1}{2} \epsilon' A' \Omega^{-1} A \epsilon \tag{2.3}$$

### 3. LM and Robust LM Test Statistics

In this section, we provide explicit formulae for the LM test statistics. We first present the joint LM test for both the individual random effects and the spatial effects. We then provide LM test statistics for the individual random effects. Lastly, we present LM test statistics for the spatial effects, namely, the spatial error correlation and/or the spatial lag dependence.<sup>1</sup> In addition, we provide formulae for robust LM tests when necessary.

Before presenting the LM test statistics, we introduce the following notations for easy reference. Let  $R_1 = M(I_N - \rho M)^{-1}$ ,  $R_2 = W(I_N - \rho M)^{-1}$ ,  $R_3 = W(I_N - \lambda W)^{-1}$ , and  $R_4 = (I_N - \rho M')(I_N - \rho M)$ . Let  $\hat{z}_\rho = \hat{\epsilon}' \hat{A}' \hat{\Omega}^{-1} (I_T \otimes M) \hat{\epsilon}$ ,  $\hat{z}_\lambda = \hat{\epsilon}' \hat{A}' \hat{\Omega}^{-1} \hat{A} (I_T \otimes W) y$ , and  $\hat{z}_{\sigma_\mu^2} = [\hat{\epsilon}' \hat{A}' (\bar{J}_T \otimes I_N) \hat{A} \hat{\epsilon}] / \hat{\sigma}_v^2 - N$ , where  $\hat{\epsilon} = \hat{B}y - X\hat{\beta}$ ,  $\hat{A}$ ,  $\hat{B}$ ,  $\hat{\Omega}$ ,  $\hat{\beta}$ , and  $\hat{\sigma}_v^2$  are restricted MLEs of  $A$ ,  $B$ ,  $\Omega$ ,  $\beta$ , and  $\sigma_v^2$  under the null hypotheses, respectively. Next, define  $\hat{v} = \hat{y}' (I_T \otimes W') \hat{A}' \hat{\Omega}^{-1} \hat{A} (I_T \otimes W) \hat{y}$ ,  $\hat{\tau} = T^2(b_1 b_3 - b_2^2) + T b_1 \hat{\omega}$ , where  $\hat{\omega} = \hat{y}' (I_T \otimes W') \hat{A}' [\hat{\Omega}^{-1} - \hat{\Omega}^{-1} \hat{A} X (X' \hat{A}' \hat{\Omega}^{-1} \hat{A} X)^{-1} X' \hat{A}' \hat{\Omega}^{-1}] \hat{A} (I_T \otimes W) \hat{y}$ ,  $\hat{y} = \hat{B}^{-1} X \hat{\beta}$ ,  $b_1 = \text{tr}(M'M + MM)$ ,  $b_2 = \text{tr}(M'W + MW)$ ,  $b_3 = \text{tr}(W'W + WW)$ , and  $\text{tr}(\cdot)$  is the trace operator. Finally, let

$$\hat{\xi} = \frac{NT\hat{\vartheta}_2 + N\hat{\omega} - 2T\hat{\vartheta}_3^2}{Tb_1(NT\hat{\vartheta}_2 + N\hat{\omega} - 2T\hat{\vartheta}_3^2) - N(T\hat{\vartheta}_1)^2}, \quad \hat{\zeta} = \frac{N\hat{\theta}_1 - 2\hat{\theta}_3^2}{(N\hat{\theta}_1 - 2\hat{\theta}_3^2)(T\hat{\theta}_4 + \hat{\omega}) - NT\hat{\theta}_2^2}$$

<sup>1</sup> Notice that there are four cases for which we do not present the LM tests formulae, since these four cases are not particularly interesting. The null hypotheses for the four cases are:  $\sigma_\mu^2 = \rho = 0$  ( $\lambda = 0$ ),  $\sigma_\mu^2 = \rho = 0$  ( $\lambda \neq 0$ ),  $\sigma_\mu^2 = \lambda = 0$  ( $\rho = 0$ ) and  $\sigma_\mu^2 = \lambda = 0$  ( $\rho \neq 0$ ). The LM tests formulae for these four cases are not presented in the paper, but they are available upon request from the authors.

where  $\hat{\vartheta}_1 = \text{tr}[(M + M')\hat{R}_3]$ ,  $\hat{\vartheta}_2 = \text{tr}(\hat{R}_3\hat{R}_3 + \hat{R}_3\hat{R}_3')$ ,  $\hat{\vartheta}_3 = \text{tr}(\hat{R}_3)$ ,  $\hat{\theta}_1 = \text{tr}(\hat{R}_1\hat{R}_1 + \hat{R}_1\hat{R}_1')$ ,  $\hat{\theta}_2 = \text{tr}[W\hat{R}_1 + \hat{R}_2\hat{R}_1'(I_N - \hat{\rho}M)]$ ,  $\hat{\theta}_3 = \text{tr}(\hat{R}_1)$ ,  $\hat{\theta}_4 = \text{tr}(WW) + \text{tr}(\hat{R}_2\hat{R}_2'\hat{R}_4)$ , and  $\hat{R}_1, \hat{R}_2, \hat{R}_3, \hat{R}_4$  are restricted MLEs of  $R_1, R_2, R_3, R_4$ , respectively.

### 3.1. Jointly Testing for Random Effects and Spatial Effects

Now we are ready to present the test statistics. We first construct the joint LM test statistic for the individual random effects and the spatial effects, namely, the spatial error correlation and the spatial lag dependence. The joint hypothesis is  $H_0^a: \sigma_\mu^2 = \rho = \lambda = 0$  vs.  $H_1^a$ : at least one of  $\sigma_\mu^2, \rho$  and  $\lambda$  is not zero. Thus we are testing the classical pooled panel data model against the full specification in (2.1), the LM test in this case is given by

$$LM_a = \frac{Tb_3 + \hat{\omega}_a}{\hat{\tau}_a} \hat{z}_{\rho,a}^2 + \frac{Tb_1}{\hat{\tau}_a} \hat{z}_{\lambda,a}^2 - \frac{2Tb_2}{\hat{\tau}_a} \hat{z}_{\rho,a} \hat{z}_{\lambda,a} + \frac{T}{2N(T-1)} \hat{z}_{\sigma_\mu^2,a}^2, \tag{3.1}$$

where  $\hat{\omega}_a = [\hat{y}'_a(I_T \otimes W')(I_{NT} - X(X'X)^{-1}X')(I_T \otimes W)\hat{y}_a]/\hat{\sigma}_{v,a}^2$ ,  $\hat{\sigma}_{v,a}^2 = (\hat{\epsilon}'_a\hat{\epsilon}_a)/(NT)$ ,  $\hat{\epsilon}_a = y - \hat{y}_a$ ,  $\hat{y}_a = X\hat{\beta}_a$ ,  $\hat{\tau}_a = T^2(b_1b_3 - b_2^2) + Tb_1\hat{\omega}_a$ ,  $\hat{z}_{\rho,a} = [\hat{\epsilon}'_a(I_T \otimes M)\hat{\epsilon}_a]/\hat{\sigma}_{v,a}^2$ ,  $\hat{z}_{\lambda,a} = [\hat{\epsilon}'_a(I_T \otimes W)y]/\hat{\sigma}_{v,a}^2$ ,  $\hat{z}_{\mu,a} = [\hat{\epsilon}'_a(\bar{J}_T \otimes I_N)\hat{\epsilon}_a]/\hat{\sigma}_{v,a}^2 - N$ , and  $\hat{\beta}_a$  is the OLS estimator of  $\beta$ . Under  $H_0^a$ ,  $LM_a$  is asymptotically distributed as  $\chi_d^2$ , where  $\chi_d^2$  denotes the  $\chi^2$  distribution with degree of freedom  $d$ .

The test statistic  $LM_a$  is useful in practice, and it is simple to compute since only the OLS estimator is required. Researchers should first use this joint LM test to determine if there is individual random effects and/or spatial effects in the general specification (2.1). If the joint null hypothesis cannot be rejected, then it is reasonable to just adopt the classical pooled panel data model. Otherwise either the individual random effects, or the spatial error correlation, or the spatial lag dependence need to be considered. As in Baltagi *et al.*, (2003) [7], we do not provide formal proofs for the asymptotic null distributions of the LM test statistics in this paper, but these distributions are likely to hold by using the Central Limit Theorems (CLTs) in Kelejian and Prucha (2001, 2010) [28,29] under similar sets of low level assumptions in their papers.

### 3.2. Testing for Random Effects

In this section, we focus on testing for the individual random effects in various spatial panel data models. We provide formulae for the standard LM tests as well as formulae for the robust LM tests when necessary. The first hypothesis we consider is  $H_0^b: \sigma_\mu^2 = 0$  ( $\rho = \lambda = 0$ ) vs.  $H_1^b: \sigma_\mu^2 > 0$  ( $\rho = \lambda = 0$ ). The null model is the classical pooled panel data model, and the alternative model is the random effects panel data model without spatial effects. The LM test denoted as  $LM_b$ , is available in Baltagi *et al.*, (2003) [7]. Moreover, it can be shown that the robust LM test is the same as  $LM_b$  in this case. We thus omit these formulae for the sake of compactness.

The second hypothesis is  $H_0^c: \sigma_\mu^2 = 0$  ( $\rho \neq 0, \lambda = 0$ ) vs.  $H_1^c: \sigma_\mu^2 > 0$  ( $\rho \neq 0, \lambda = 0$ ).<sup>2</sup> Under the null hypothesis, it is the pooled panel data model with spatial error correlation. Under the alternative

---

<sup>2</sup> By  $\rho \neq 0$ , we mean that  $\rho$  is allowed to be nonzero, and the notation “ $\neq$ ” has similar meaning in the rest of this paper.

hypothesis, it is the random effects panel data model with spatial error correlation. The LM test statistic in this case is given by

$$LM_c = \frac{T}{2N(T-1)} \widehat{z}_{\sigma_\mu^2, c}^2, \quad (3.2)$$

where  $\widehat{z}_{\sigma_\mu^2, c}$  is  $\widehat{z}_{\sigma_\mu^2}$  evaluated at the restricted MLE under  $H_0^c$ .<sup>3</sup> Under  $H_0^c$ ,  $LM_c$  is asymptotically distributed as  $\chi_1^2$ . Further, it can be easily shown that the robust LM test in this case is the same as  $LM_c$ . Thus  $LM_c$  itself is robust against local deviation of  $\lambda$  from 0, and this will be confirmed by the simulation results.

The third hypothesis is  $H_0^d: \sigma_\mu^2 = 0$  ( $\rho = 0, \lambda \neq 0$ ) vs.  $H_1^d: \sigma_\mu^2 > 0$  ( $\rho = 0, \lambda \neq 0$ ). Under the null hypothesis, it is the pooled panel data model with spatial lag dependence. Under the alternative hypothesis, it is the random effects panel data model with spatial lag dependence. The LM test statistic, denoted as  $LM_d$ , is available in Baltagi and Liu (2008) [8]. Moreover, it can be shown that the robust LM test in this case is the same as  $LM_d$ . We thus omit these formulae here.

The last hypothesis concerning testing for the individual random effects is  $H_0^e: \sigma_\mu^2 = 0$  ( $\rho \neq 0, \lambda \neq 0$ ) vs.  $H_1^e: \sigma_\mu^2 > 0$  ( $\rho \neq 0, \lambda \neq 0$ ). Under the null hypothesis, it is the pooled panel data model with both spatial error correlation and spatial lag dependence. Under the alternative hypothesis, it is the full model in (2.1). The LM test statistic is given by

$$LM_e = \frac{T}{2N(T-1)} \widehat{z}_{\sigma_\mu^2, e}^2. \quad (3.3)$$

Under  $H_0^e$ ,  $LM_e$  is asymptotically distributed as  $\chi_1^2$ .  $LM_e$  tests for the individual random effects in the most general spatial model, and it is particularly useful when the researcher does not have any prior knowledge about whether the spatial error correlation and/or spatial lag dependence exist or not. In practice, the above four LM tests for the individual random effects correspond to different prior information on the nuisance parameters, and they need to be analyzed together to lead to the most appropriate model.

Notice that the LM tests in Section 3.2 are all designed for two-sided alternative hypothesis, while the parameter involved in the hypotheses,  $\sigma_\mu^2$ , is by definition nonnegative. While our LM tests have good power against the one-sided alternative (see the simulation results), we point out that power of these tests can be further improved by following the ideas in Honda (1985, 1991) [30,31].

### 3.3. Testing for Spatial Effects

In this section, we focus on testing for the spatial effects. We provide formulae for the standard LM tests as well as formulae for the robust LM tests when necessary.

---

<sup>3</sup> In order to save space and avoid notational complication, we do not present the formulae of quantities involved in the LM test statistics case by case. The subscripts of these quantities indicate these quantities evaluated at the restricted MLE for each case. If the restricted MLE is just the OLS estimator, we use subscript "a" to indicate this, which is in line with the fact that only OLS estimation is needed in  $LM_a$ .

### 3.3.1. Joint Tests for Spatial Effects

In practice, researcher may be interested in jointly testing for the spatial error correlation and the spatial lag dependence. The first joint hypothesis is  $H_0^f: \rho = \lambda = 0$  ( $\sigma_\mu^2 = 0$ ) vs.  $H_1^f$ : at least one of  $\rho$  and  $\lambda$  is not zero ( $\sigma_\mu^2 \neq 0$ ). Under the null hypothesis, it is the classical pooled panel data model. Under the alternative hypothesis, it is the pooled panel data model with at least one type of the spatial effects. The LM test statistic in this case is given by

$$LM_f = \frac{Tb_3 + \widehat{\omega}_a}{\widehat{\tau}_a} \widehat{z}_{\rho,a}^2 + \frac{Tb_1}{\widehat{\tau}_a} \widehat{z}_{\lambda,a}^2 - \frac{2Tb_2}{\widehat{\tau}_a} \widehat{z}_{\rho,a} \widehat{z}_{\lambda,a}, \tag{3.4}$$

where all quantities involved are defined Section 3.1 since only the OLS estimator is needed in this case. The test  $LM_f$  is a useful extension of the result in Anselin *et al.*, (1996) [20] to the panel data case. Interestingly,  $LM_f$  is the sum of the first three terms of  $LM_a$ . Under  $H_0^f$ ,  $LM_f$  is asymptotically distributed as  $\chi_2^2$ . It can be easily shown that the robust LM test in this case is the same as  $LM_f$ .

The second joint hypothesis is  $H_0^g: \rho = \lambda = 0$  ( $\sigma_\mu^2 \geq 0$ ) vs.  $H_1^g$ : at least one of  $\rho$  and  $\lambda$  is not zero ( $\sigma_\mu^2 > 0$ ). Under the null hypothesis, it is the random effects panel data model without any spatial effects. Under the alternative hypothesis, it is the random effects panel data model with at least one type of spatial effects. The LM test statistic in this case is given by

$$LM_g = \frac{Tb_3 + \widehat{\omega}_g}{\widehat{\tau}_g} \widehat{z}_{\rho,g}^2 + \frac{Tb_1}{\widehat{\tau}_g} \widehat{z}_{\lambda,g}^2 - \frac{2Tb_2}{\widehat{\tau}_g} \widehat{z}_{\rho,g} \widehat{z}_{\lambda,g}. \tag{3.5}$$

Under  $H_0^g$ ,  $LM_g$  is asymptotically distributed as  $\chi_2^2$ . Notice that both  $LM_f$  and  $LM_g$  are useful for jointly testing the spatial error correlation and spatial lag dependence. However,  $LM_f$  assumes that it is pooled panel data model, while  $LM_g$  allows for the individual random effects. Since  $LM_f$  is the same as its robust version, then it can guard against local deviation of  $\sigma_\mu^2$  from zero. However,  $LM_g$  will work well even when  $\sigma_\mu^2$  deviates far away from zero.

### 3.3.2. Testing for Spatial Error Correlation

In this section, we focus on testing for spatial error correlation. The first hypothesis we consider is  $H_0^h: \rho = 0$  ( $\sigma_\mu^2 = \lambda = 0$ ) vs.  $H_1^h: \rho \neq 0$  ( $\sigma_\mu^2 = \lambda = 0$ ). Under the null hypothesis, it is the classical pooled panel data model. Under the alternative hypothesis, it is the pooled panel data model with spatial error correlation. The LM and robust LM (denoted as  $LM^*$ ) test statistics are given by

$$LM_h = \frac{1}{Tb_1} \widehat{z}_{\rho,a}^2, LM_h^* = \frac{Tb_3 + \widehat{\omega}_a}{\widehat{\tau}_a} \left( \widehat{z}_{\rho,a} - \frac{Tb_2}{Tb_3 + \widehat{\omega}_a} \widehat{z}_{\lambda,a} \right)^2, \tag{3.6}$$

where all quantities involved are defined in Section 3.1 since only the OLS estimator is needed in this case. Notice that although the formula of  $LM_h$  is available in Baltagi *et al.*, (2003) [7], we provide it here for comparison purposes as it is different from the robust test  $LM_h^*$ , which has not been considered previously in the literature. Under  $H_0^h$ , if the nuisance parameters  $\lambda$  and  $\sigma_\mu^2$  do not deviate from 0, both  $LM_h$  and  $LM_h^*$  are asymptotically distributed as  $\chi_1^2$ . However, when  $\rho = 0$ , but either  $\lambda$  or  $\sigma_\mu^2$  deviates locally from 0, the distribution of  $LM_h$  becomes non-centralized, tending to over reject the null hypothesis. On the other hand,  $LM_h^*$  is still asymptotically distributed as  $\chi_1^2$ , thus it does not suffer from size distortion as  $LM_h$ .



The second hypothesis is  $H_0^i: \rho = 0 (\sigma_\mu^2 = 0, \lambda \neq 0)$  vs.  $H_1^i: \rho \neq 0 (\sigma_\mu^2 = 0, \lambda \neq 0)$ . Under the null hypothesis, it is the pooled panel data model with spatial lag dependence. Under the alternative hypothesis, it is the pooled panel data model with both spatial error correlation and spatial lag dependence. The LM test statistic is given by

$$LM_i = \widehat{\xi}_i \widehat{z}_{\rho,i}^2 \tag{3.7}$$

$LM_i$  is a useful extension of the results in Anselin *et al.*, (1996) [20] to the panel data case. Under  $H_0^i$ ,  $LM_i$  is asymptotically distributed as  $\chi_1^2$ . Moreover, it can be easily shown that the robust LM test in this case is the same as  $LM_i$ . Thus  $LM_i$  itself is robust against local deviation of  $\sigma_\mu^2$  from 0, and this will be confirmed by the simulation results.

The third hypothesis is  $H_0^j: \rho = 0 (\sigma_\mu^2 \geq 0, \lambda = 0)$  vs.  $H_1^j: \rho \neq 0 (\sigma_\mu^2 \geq 0, \lambda = 0)$ . Under the null hypothesis, it is the random effects panel data model without spatial effects. Under the alternative hypothesis, it is the random effects panel data model with spatial error correlation. The LM and robust LM test statistics in this case are given by

$$LM_j = \frac{1}{Tb_1} \widehat{z}_{\rho,j}^2, \quad LM_j^* = \frac{Tb_3 + \widehat{\omega}_j}{\widehat{\tau}_j} \left( \widehat{z}_{\rho,j} - \frac{Tb_2}{Tb_3 + \widehat{\omega}_j} \widehat{z}_{\lambda,j} \right)^2 \tag{3.8}$$

When  $\rho = 0$ , and if the nuisance parameter  $\lambda$  does not deviate from 0, both  $LM_j$  and  $LM_j^*$  are asymptotically distributed as  $\chi_1^2$ . However, when  $\rho = 0$ , but  $\lambda$  deviates locally from 0, the distribution of  $LM_j$  becomes non-centralized, tending to over reject the null hypothesis. On the other hand,  $LM_j^*$  is still asymptotically distributed as  $\chi_1^2$  in this case, thus it does not suffer from size distortion as  $LM_j$ .

The last hypothesis for spatial error correlation is  $H_0^k: \rho = 0 (\sigma_\mu^2 \geq 0, \lambda \neq 0)$  vs.  $H_1^k: \rho \neq 0 (\sigma_\mu^2 \geq 0, \lambda \neq 0)$ . Under the null hypothesis, it is the random effects panel data model with spatial lag dependence. Under the alternative hypothesis, it is the random effects panel data model with both spatial error correlation and spatial lag dependence. The LM test statistic in this case is given by

$$LM_k = \widehat{\xi}_k \widehat{z}_{\rho,k}^2 \tag{3.9}$$

Under  $H_0^k$ ,  $LM_k$  is asymptotically distributed as  $\chi_1^2$ . In practice, to test for the spatial error correlation, the above test statistics correspond to different prior information on the nuisance parameters, and they need to be analyzed together to lead to the most appropriate model.

### 3.3.3. Testing for Spatial Lag Dependence

In this section, we focus on testing for the spatial lag dependence. The first hypothesis we consider is  $H_0^l: \lambda = 0 (\sigma_\mu^2 = \rho = 0)$  vs.  $H_1^l: \lambda \neq 0 (\sigma_\mu^2 = \rho = 0)$ . Under the null hypothesis, it is the classical pooled panel data model. Under the alternative hypothesis, it is the pooled panel data model with spatial lag dependence. The LM and robust LM test statistic in this case are given by

$$LM_l = \frac{1}{Tb_3 + \widehat{\omega}_a} \widehat{z}_{\lambda,a}^2, \quad LM_l^* = \frac{Tb_1}{\widehat{\tau}_a} \left( \widehat{z}_{\lambda,a} - \frac{b_2}{b_1} \widehat{z}_{\rho,a} \right)^2 \tag{3.10}$$

where all the related quantities are defined in Section 3.1 since only the OLS estimator is needed in this case. The formulae for  $LM_l$  and  $LM_l^*$  are useful extensions of the results in Anselin *et al.*, (1996) [20]

to the panel data case. When  $\lambda = 0$ , and if the nuisance parameters  $\sigma_\mu^2$  and  $\rho$  do not deviate from 0, both  $LM_l$  and  $LM_l^*$  are asymptotically distributed as  $\chi_1^2$ . However, when  $\lambda = 0$ , but either  $\sigma_\mu^2$  or  $\rho$  deviates locally from 0, the distribution of  $LM_l$  becomes non-centralized, tending to over reject the null hypothesis. On the other hand,  $LM_l^*$  is still asymptotically distributed as  $\chi_1^2$  in this case, thus it does not suffer from size distortion as  $LM_l$ .

The second hypothesis is  $H_0^m: \lambda = 0 (\sigma_\mu^2 = 0, \rho \neq 0)$  vs.  $H_1^m: \lambda \neq 0 (\sigma_\mu^2 = 0, \rho \neq 0)$ . Under the null hypothesis, it is the pooled panel data model with spatial error correlation. Under the alternative hypothesis, it is the pooled panel data model with both spatial error correlation and spatial lag dependence. The LM test statistic in this case is given by

$$LM_m = \hat{\zeta}_m \hat{z}_{\lambda,m}^2 \tag{3.11}$$

$LM_m$  is a useful extension of the result in Anselin *et al.*, (1996) [20] to the panel data case, and it is asymptotically distributed as  $\chi_1^2$  under  $H_0^m$ . It can be easily shown that the robust LM test in this case is the same as  $LM_m$ . Thus  $LM_m$  itself is robust against local deviation of  $\sigma_\mu^2$  from 0, and this will be confirmed by the simulation results.

The third hypothesis is  $H_0^n: \lambda = 0 (\sigma_\mu^2 \geq 0, \rho = 0)$  vs.  $H_1^n: \lambda \neq 0 (\sigma_\mu^2 \geq 0, \rho = 0)$ . Under the null hypothesis, it is the random effects panel data model without spatial effects. Under the alternative hypothesis, it is the random effects panel data model with spatial lag dependence. The LM and robust LM test statistics are given by

$$LM_n = \frac{1}{Tb_3 + \hat{\omega}_n} \hat{z}_{\lambda,n}^2, \quad LM_n^* = \frac{Tb_1}{\hat{\tau}_n} \left( \hat{z}_{\lambda,n} - \frac{b_2}{b_1} \hat{z}_{\rho,n} \right)^2 \tag{3.12}$$

Notice that although the formula of  $LM_n$  is available in Baltagi and Liu (2008) [8], we provide it here for comparison purposes since it is different from the robust test  $LM_n^*$ , which has not been considered previously in the literature. When  $\lambda = 0$ , and if the nuisance parameters  $\rho$  does not deviate from 0, both  $LM_n$  and  $LM_n^*$  are asymptotically distributed as  $\chi_1^2$ . However, when  $\lambda = 0$ , but  $\rho$  deviates locally from 0, the distribution of  $LM_n$  becomes non-centralized, tending to over reject the null hypothesis. On the other hand,  $LM_n^*$  is still asymptotically distributed as  $\chi_1^2$  in this case, thus it does not suffer from size distortion as  $LM_n$ .

The last hypothesis is  $H_0^o: \lambda = 0 (\sigma_\mu^2 \geq 0, \rho \neq 0)$  vs.  $H_1^o: \lambda \neq 0 (\sigma_\mu^2 \geq 0, \rho \neq 0)$ . Under the null hypothesis, it is the random effects panel data model with spatial error correlation. Under the alternative hypothesis, it is the random effects panel data model with both spatial error correlation and spatial lag dependence. The LM test statistic in this case is given by

$$LM_o = \hat{\zeta}_o \hat{z}_{\lambda,o}^2 \tag{3.13}$$

Under  $H_0^o$ ,  $LM_o$  is asymptotically distributed as  $\chi_1^2$ . In practice, to test for the spatial lag dependence, the above test statistics correspond to different prior information on the nuisance parameters, and they need to be analyzed together to lead to the most appropriate model.

#### 4. Monte Carlo Experiment

In this section, we conduct and present a small Monte Carlo experiment to show satisfactory performances of the above LM test statistics. In our Monte Carlo experiment, the data generating process is, for  $t = 1, 2, \dots, T$ ,

$$(I_N - \lambda W)y_t = \alpha \iota_N + X_t \beta + (I_N - \rho M)^{-1}(\mu + v_t),$$

where  $\alpha = 5, \beta = 0.5$ .  $x_{it}$  is a single variable and is generated as  $x_{it} = 0.1t + 0.5x_{i,t-1} + z_{it}$ , where  $z_{it}$  is generated according to a uniform distribution on  $[-0.5, 0.5]$ . The initial value  $x_{i0}$  is set to be  $5 + 10z_{i0}$ . The random effects term  $\mu_i$  is generated according to  $\mu_i \sim \text{i.i.n.}(0, \sigma_\mu^2)$ , and the innovation term  $v_{it}$  is generated according to  $v_{it} \sim \text{i.i.n.}(0, \sigma_v^2)$ .  $\sigma_\mu^2$  takes values 0, 0.2, 0.5, and 0.8, while  $\sigma_v^2$  is fixed to be 1. The spatial weights matrices  $M$  and  $W$  are set to be first-order rook and queen contiguity matrices with row-standardization, respectively. The spatial error correlation parameter  $\rho$  and spatial lag parameter  $\lambda$  vary in  $[-0.8, 0.8]$ , with increment 0.2. Two combinations of sample size  $(N, T)$  are considered, namely  $(49, 7)$  and  $(100, 10)$ . Each experiment is replicated 1000 times, and the nominal size is set to be 0.05.

Frequency of rejection (FoR) of the joint test  $LM_a$  is summarized in Table 1, where the upper part and lower part correspond to different sample sizes. We only report the case when  $\sigma_\mu^2 = 0$ . For other cases, that is,  $\sigma_\mu^2 = 0.2, 0.5, 0.8$ , FoRs are uniformly higher than that when  $\sigma_\mu^2 = 0$ . This is because we are jointly testing  $\sigma_\mu^2 = \rho = \lambda = 0$ . The empirical sizes of  $LM_a$  are 0.049 and 0.050 for the  $(49, 7)$  and the  $(100, 10)$  sample, respectively. They are almost the same as the nominal size, reflecting that the limiting  $\chi_3^2$  distribution approximates the finite sample null distribution very well. As  $\rho$  or  $\lambda$  deviates from 0, FoR increases very fast. For example, when the sample size is  $(49, 7)$ , FoR is 0.995 when  $\rho = \lambda = 0.2$ , and it reaches 1 when  $\rho = \lambda = 0.4$ . The power performance when sample size is  $(100, 10)$  is better than that when the sample size is  $(49, 7)$ . The good size and power performance demonstrates that the joint test  $LM_a$  should be very useful for applied researcher to determine whether there are individual random effects and spatial effects in a preliminary diagnostic testing process.

Experiment results for  $LM_c$  and  $LM_e$  are summarized in Table 2. For  $LM_c$ , we only report results when  $\rho = -0.4$  and  $0.4$  here to save space. The results when  $\rho$  takes other values are very similar, and they are available upon request. First, for the  $(49, 7)$  sample, the empirical sizes of  $LM_c$  are 0.043 and 0.036 when  $\rho = -0.4$  and  $0.4$ , respectively. As  $\sigma_\mu^2$  increases from 0, the FoR of  $LM_c$  increases very fast. Actually, when  $\rho = -0.4$ , FoR of  $LM_c$  is 0.976 even when  $\sigma_\mu^2$  is only 0.2. Second, we discussed in Section 3.2 that the robust LM test corresponding to  $LM_c$  is the same as  $LM_c$ , which implies that  $LM_c$  itself is robust against local deviation of  $\lambda$  from 0. This is confirmed by the column for  $\sigma_\mu^2 = 0$  in Table 2. Given  $\sigma_\mu^2 = 0$ , for the  $(49, 7)$  sample, when  $\lambda$  varies in  $[-0.4, 0.4]$ , the FoRs of  $LM_c$  are in the range of  $[0.036, 0.057]$ . Thus  $LM_c$  itself does not suffer from size distortion under local misspecification of the nuisance parameter  $\lambda$ . As expected, the performance of  $LM_c$  for the  $(100, 10)$  sample is even better than that for the  $(49, 7)$  sample. Next, for  $LM_e$ , we choose to report the simulation results for a few combinations of  $\rho$  and  $\lambda$  as shown in the table. Results for other cases are very similar. For the  $(49, 7)$  sample, the empirical sizes of  $LM_e$  vary in  $[0.035, 0.044]$ , and the FoR increases rapidly as  $\sigma_\mu^2$  increases from 0. As expected, the performance of  $LM_e$  for the  $(100, 10)$  sample is even better than that for the  $(49, 7)$  sample. Both  $LM_c$  and  $LM_e$  are useful in testing for the random effects.  $LM_e$  is useful in testing

for the random effects when the researcher does not have any knowledge about the spatial effects, while  $LM_c$  is particularly useful when the researcher has information that there is only spatial error correlation.

**Table 1.** Frequency of rejection (FoR) of  $LM_a$ ,  $\sigma_\mu^2 = 0$ , Sample Sizes: Upper Part: (49, 7); Lower Part: (100, 10).

$\rho$	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8
$\lambda$									
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.996	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	0.980	0.942	0.998	1.000
-0.4	1.000	1.000	1.000	1.000	0.946	0.692	0.861	1.000	1.000
-0.2	1.000	1.000	1.000	0.984	0.389	0.290	0.954	1.000	1.000
0.0	1.000	1.000	1.000	0.681	0.049	0.623	1.000	1.000	1.000
0.2	1.000	1.000	0.994	0.515	0.552	0.995	1.000	1.000	1.000
0.4	1.000	1.000	0.998	0.957	0.997	1.000	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
$\rho$	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8
$\lambda$									
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	1.000	1.000	0.988	1.000	1.000	1.000
-0.2	1.000	1.000	1.000	1.000	0.864	0.758	1.000	1.000	1.000
0.0	1.000	1.000	1.000	0.989	0.050	0.986	1.000	1.000	1.000
0.2	1.000	1.000	1.000	0.911	0.926	1.000	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

The simulation results for  $LM_f$  and  $LM_g$  are presented in Tables 3 and 4, respectively. From Table 3, the empirical sizes of  $LM_f$  are 0.053 and 0.050 for the (49, 7) and (100, 10) sample, respectively. As either  $\rho$  or  $\lambda$  deviates from 0, the FoR of increases very fast, indicating good power performance of  $LM_f$ . For example, when  $\rho = \lambda = 0.2$ , FoR of  $LM_f$  is 0.994 for the (49, 7) sample, while it is 1 for the (100, 10) sample. In Table 4, we only report the results for the case when  $\sigma_\mu^2 = 0.5$  to save space, results for the other two cases are similar and available upon request. The empirical sizes of  $LM_g$  are 0.050 and 0.043 for the (49, 7) and the (100, 10) sample, respectively. Similar to  $LM_f$ , as either  $\rho$  or  $\lambda$  deviates from 0, the FoR of increases rapidly. Both  $LM_f$  and  $LM_g$  are useful in jointly detecting spatial error correlation and spatial lag dependence.  $LM_f$  is useful when we assume pooled panel data model, while  $LM_g$  is useful when we assume random effects panel data model.

**Table 2.** FoR of LM Tests for Random Effects, Sample Sizes: Upper Part: (49, 7); Lower Part: (100, 10).

			$\sigma_\mu^2 = 0$	0.2	0.5	0.8
<i>LM<sub>c</sub></i>	$\rho = -0.4$	$\lambda = -0.4$	0.057	0.974	1.000	1.000
		-0.2	0.043	0.978	1.000	1.000
		0.0	0.043	0.976	1.000	1.000
		0.2	0.048	0.973	1.000	1.000
		0.4	0.051	0.970	1.000	1.000
	$\rho = 0.4$	$\lambda = -0.4$	0.047	0.968	1.000	1.000
		-0.2	0.046	0.969	1.000	1.000
		0.0	0.036	0.976	1.000	1.000
		0.2	0.052	0.982	1.000	1.000
		0.4	0.053	0.969	1.000	1.000
<i>LM<sub>e</sub></i>	$\rho = -0.4$	$\lambda = -0.4$	0.035	0.976	1.000	1.000
		0.4	0.043	0.982	1.000	1.000
	0.0	0.0	0.035	0.970	1.000	1.000
	0.4	-0.4	0.039	0.976	1.000	1.000
	0.4	0.4	0.044	0.983	1.000	1.000
			$\sigma_\mu^2 = 0$	0.2	0.5	0.8
<i>LM<sub>c</sub></i>	$\rho = -0.4$	$\lambda = -0.4$	0.052	1.000	1.000	1.000
		-0.2	0.047	1.000	1.000	1.000
		0.0	0.049	1.000	1.000	1.000
		0.2	0.045	1.000	1.000	1.000
		0.4	0.057	1.000	1.000	1.000
	$\rho = 0.4$	$\lambda = -0.4$	0.053	1.000	1.000	1.000
		-0.2	0.054	1.000	1.000	1.000
		0.0	0.038	1.000	1.000	1.000
		0.2	0.047	1.000	1.000	1.000
		0.4	0.049	1.000	1.000	1.000
<i>LM<sub>e</sub></i>	$\rho = -0.4$	$\lambda = -0.4$	0.041	1.000	1.000	1.000
		0.4	0.045	1.000	1.000	1.000
	0.0	0.0	0.041	1.000	1.000	1.000
	0.4	-0.4	0.047	1.000	1.000	1.000
	0.4	0.4	0.042	1.000	1.000	1.000



**Table 4.** *Cont.*

$\rho$	-0.8	-0.6	-0.4	-0.2	0.0	0.2	0.4	0.6	0.8
$\lambda$									
-0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
-0.4	1.000	1.000	1.000	1.000	1.000	0.993	1.000	1.000	1.000
-0.2	1.000	1.000	1.000	1.000	0.898	0.815	1.000	1.000	1.000
0.0	1.000	1.000	1.000	0.993	0.043	0.991	1.000	1.000	1.000
0.2	1.000	1.000	1.000	0.928	0.944	1.000	1.000	1.000	1.000
0.4	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Now we discuss simulation results of the test statistics for spatial error correlation, their FoRs are summarized in Tables 5 and 6 for different sample sizes. We focus on the (49, 7) sample for discussion. For  $LM_h$ , the empirical size is 0.048. Moreover,  $LM_h$  is very powerful in detecting spatial error correlation given that  $\sigma_\mu^2 = \lambda = 0$ . For example, FoR of  $LM_h$  is 0.856 when  $\rho = -0.2$ . However, as discussed in Section 3.3.2,  $LM_h$  is not robust against local misspecification, and this is confirmed by the simulation results. Given  $\rho = 0$ , when  $\sigma_\mu^2$  and  $\lambda$  deviate locally from 0, the large FoRs of  $LM_h$  are undesirable. For example, when  $\rho = 0$  but  $\sigma_\mu^2 = \lambda = 0.2$ , FoR of  $LM_h$  is 0.416. However, this size distortion is avoided by using  $LM_h^*$ . When  $\rho = 0$ ,  $\sigma_\mu^2$  takes values 0, 0.2 and  $\lambda \in [-0.4, 0.4]$ , FoRs of  $LM_h^*$  are in the range of [0.039, 0.081]. Although it is a little oversized in some cases, it is much better than that of  $LM_h$ . On the other hand, the robust LM test is supposed to be less powerful than the corresponding LM test when the nuisance parameters are correctly specified. This is also confirmed by the simulation results, that is, when  $\sigma_\mu^2 = \lambda = 0$ ,  $LM_h^*$  is less powerful than  $LM_h$ . For example, when  $\rho = -0.2$ , FoR of  $LM_h$  is 0.856, while that of  $LM_h^*$  is 0.626. Next,  $LM_i$  tests for spatial error correlation in a pooled panel data model with spatial lag dependence. When  $\sigma_\mu^2 = 0$ , the empirical sizes of  $LM_i$  vary in [0.040, 0.060]. As  $\rho$  deviates from 0, FoR increases as expected. Furthermore, as discussed in Section 3.3.2, the robust LM test in this case is the same as  $LM_i$ , implying that  $LM_i$  is robust against local deviation of  $\sigma_\mu^2$  from 0. This is confirmed by the simulation results. When  $\sigma_\mu^2 = 0.2, \rho = 0$ , the empirical sizes of  $LM_i$  vary in [0.058, 0.080]. Next, we discuss  $LM_j$  and  $LM_j^*$ .  $LM_j$  tests for spatial error correlation in a random effects panel data model without spatial lag dependence. We only present the representative result when  $\sigma_\mu^2 = 0.5$  to save space, results in other cases are very similar. The empirical size of  $LM_j$  is 0.045, and the FoR increases as  $\rho$  deviates away from 0. For example, FoR is 0.835 when  $\rho = -0.2$ . However,  $LM_j$  suffers from size distortion when the nuisance parameter  $\lambda$  deviates away from 0. For example, when  $\lambda = -0.4, \rho = 0$ , the FoR of  $LM_j$  is 0.747, which is undesirably high. For the same case, FoR of  $LM_j^*$  is 0.064. On the other hand, the robust LM test is less powerful when the nuisance parameter is correctly specified. This is also confirmed by the simulation results. For example, when  $\lambda = 0, \rho = -0.2$ , FoR of  $LM_j^*$  is 0.602, while that of  $LM_j$  is 0.835. Lastly, we discuss the performance of  $LM_k$ .  $LM_k$  tests for spatial error correlation in a random effects panel data model with spatial lag dependence. The empirical sizes of  $LM_k$  vary in [0.038, 0.058]. As expected, when  $\rho$  moves away from 0, FoR increases rapidly, suggesting its good power performance. For all of the above tests, their performances are even better for the (100, 10) sample than for the (49, 7) sample

and follow a similar discussion. In sum, the test statistics  $LM_h$ ,  $LM_h^*$ ,  $LM_i$ ,  $LM_j$ ,  $LM_j^*$ , and  $LM_k$  are all useful for detecting the spatial error correlation, but they are suited for different assumptions about the nuisance parameters. In practice, researchers are suggested to analyze them together to draw correct inference on  $\rho$ .

**Table 5.** FoR of LM Tests for the Spatial Error Correlation, Sample Size: (49, 7).

			$\rho = -0.8$	$-0.6$	$-0.4$	$-0.2$	$0$	$0.2$	$0.4$	$0.6$	$0.8$
$LM_h$	$\sigma_\mu^2 = 0$	$\lambda = -0.4$	1.000	1.000	1.000	1.000	0.738	0.023	0.671	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.986	0.305	0.193	0.975	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.856	0.048	0.754	1.000	1.000	1.000
		$0.2$	1.000	1.000	0.996	0.340	0.452	0.993	1.000	1.000	1.000
		$0.4$	1.000	1.000	0.753	0.210	0.965	1.000	1.000	1.000	1.000
$LM_h$	$\sigma_\mu^2 = 0.2$	$\lambda = -0.4$	1.000	1.000	1.000	0.998	0.747	0.037	0.620	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.984	0.324	0.189	0.971	1.000	1.000
		$0.0$	1.000	1.000	0.999	0.822	0.065	0.743	1.000	1.000	1.000
		$0.2$	1.000	1.000	0.992	0.372	0.416	0.989	1.000	1.000	1.000
		$0.4$	1.000	1.000	0.756	0.206	0.954	1.000	1.000	1.000	1.000
$LM_h^*$	$\sigma_\mu^2 = 0$	$\lambda = -0.4$	1.000	1.000	0.976	0.485	0.071	0.539	0.961	1.000	0.999
		$-0.2$	1.000	1.000	0.990	0.562	0.056	0.482	0.941	1.000	0.999
		$0.0$	1.000	1.000	0.995	0.626	0.039	0.431	0.952	0.999	0.989
		$0.2$	1.000	1.000	0.997	0.639	0.052	0.460	0.969	1.000	0.961
		$0.4$	1.000	1.000	0.996	0.524	0.046	0.597	0.980	1.000	0.880
	$\sigma_\mu^2 = 0.2$	$\lambda = -0.4$	1.000	1.000	0.975	0.495	0.081	0.535	0.937	1.000	0.999
		$-0.2$	1.000	1.000	0.987	0.553	0.080	0.465	0.923	1.000	0.999
		$0.0$	1.000	1.000	0.988	0.607	0.075	0.432	0.926	0.998	0.991
		$0.2$	1.000	1.000	0.990	0.648	0.056	0.461	0.943	0.999	0.957
		$0.4$	1.000	1.000	0.994	0.533	0.054	0.510	0.967	0.999	0.869
$LM_i$	$\sigma_\mu^2 = 0$	$\lambda = -0.4$	1.000	1.000	0.999	0.665	0.050	0.479	0.957	0.998	1.000
		$-0.2$	1.000	1.000	0.999	0.641	0.046	0.469	0.940	0.999	1.000
		$0.0$	1.000	1.000	0.997	0.646	0.040	0.448	0.951	1.000	1.000
		$0.2$	1.000	1.000	0.997	0.646	0.060	0.457	0.961	0.999	1.000
		$0.4$	1.000	1.000	0.996	0.639	0.046	0.497	0.966	1.000	1.000
	$\sigma_\mu^2 = 0.2$	$\lambda = -0.4$	1.000	1.000	0.995	0.651	0.058	0.458	0.939	0.999	1.000
		$-0.2$	1.000	1.000	0.995	0.619	0.080	0.455	0.927	0.998	1.000
		$0.0$	1.000	1.000	0.991	0.619	0.072	0.454	0.920	0.997	1.000
		$0.2$	1.000	1.000	0.992	0.662	0.069	0.463	0.939	1.000	1.000
		$0.4$	1.000	1.000	0.995	0.610	0.063	0.487	0.951	1.000	1.000
$LM_j$	$\sigma_\mu^2 = 0.5$	$\lambda = -0.4$	1.000	1.000	1.000	1.000	0.747	0.036	0.657	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.987	0.301	0.192	0.975	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.835	0.045	0.771	1.000	1.000	1.000
		$0.2$	1.000	1.000	0.991	0.353	0.427	0.993	1.000	1.000	1.000
		$0.4$	1.000	1.000	0.758	0.194	0.957	1.000	1.000	1.000	1.000
$LM_j^*$	$\sigma_\mu^2 = 0.5$	$\lambda = -0.4$	1.000	1.000	0.979	0.480	0.064	0.522	0.945	1.000	0.998
		$-0.2$	1.000	1.000	0.990	0.551	0.060	0.460	0.934	1.000	0.998
		$0.0$	1.000	1.000	0.993	0.602	0.040	0.412	0.943	0.999	0.985
		$0.2$	1.000	1.000	0.993	0.617	0.035	0.429	0.954	1.000	0.944
		$0.4$	1.000	1.000	0.991	0.508	0.042	0.550	0.972	0.999	0.852
$LM_k$	$\sigma_\mu^2 = 0.5$	$\lambda = -0.4$	1.000	1.000	0.997	0.660	0.049	0.417	0.933	0.999	1.000
		$-0.2$	1.000	1.000	0.998	0.607	0.049	0.422	0.936	0.996	1.000
		$0.0$	1.000	1.000	0.992	0.627	0.058	0.433	0.952	0.999	1.000
		$0.2$	1.000	1.000	0.996	0.602	0.054	0.427	0.952	1.000	1.000
		$0.4$	1.000	1.000	0.995	0.630	0.038	0.480	0.948	1.000	1.000



**Table 6.** FoR of LM Tests for the Spatial Error Correlation, Sample Size: (100, 10).

			$\rho = -0.8$	$-0.6$	$-0.4$	$-0.2$	$0$	$0.2$	$0.4$	$0.6$	$0.8$
$LM_h$	$\sigma_\mu^2 = 0$	$\lambda = -0.4$	1.000	1.000	1.000	1.000	0.989	0.016	0.996	1.000	1.000
		$-0.2$	1.000	1.000	1.000	1.000	0.654	0.573	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.999	0.048	0.994	1.000	1.000	1.000
		$0.2$	1.000	1.000	1.000	0.628	0.819	1.000	1.000	1.000	1.000
		$0.4$	1.000	1.000	0.984	0.369	1.000	1.000	1.000	1.000	1.000
$LM_h$	$\sigma_\mu^2 = 0.2$	$\lambda = -0.4$	1.000	1.000	1.000	1.000	0.983	0.030	0.993	1.000	1.000
		$-0.2$	1.000	1.000	1.000	1.000	0.654	0.548	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.999	0.086	0.993	1.000	1.000	1.000
		$0.2$	1.000	1.000	1.000	0.633	0.781	1.000	1.000	1.000	1.000
		$0.4$	1.000	1.000	0.976	0.344	1.000	1.000	1.000	1.000	1.000
$LM_h^*$	$\sigma_\mu^2 = 0$	$\lambda = -0.4$	1.000	1.000	1.000	0.852	0.076	0.938	1.000	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.934	0.061	0.899	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.952	0.051	0.899	1.000	1.000	1.000
		$0.2$	1.000	1.000	1.000	0.950	0.036	0.930	1.000	1.000	1.000
		$0.4$	1.000	1.000	1.000	0.901	0.067	0.977	1.000	1.000	1.000
	$0.2$	$\lambda = -0.4$	1.000	1.000	1.000	0.826	0.103	0.914	1.000	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.917	0.086	0.870	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.948	0.079	0.864	1.000	1.000	1.000
		$0.2$	1.000	1.000	1.000	0.921	0.052	0.890	1.000	1.000	1.000
		$0.4$	1.000	1.000	1.000	0.896	0.070	0.961	1.000	1.000	1.000
$LM_i$	$\sigma_\mu^2 = 0$	$\lambda = -0.4$	1.000	1.000	1.000	0.961	0.046	0.910	1.000	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.958	0.047	0.896	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.954	0.048	0.892	1.000	1.000	1.000
		$0.2$	1.000	1.000	1.000	0.948	0.047	0.902	1.000	1.000	1.000
		$0.4$	1.000	1.000	1.000	0.950	0.048	0.913	1.000	1.000	1.000
	$0.2$	$\lambda = -0.4$	1.000	1.000	1.000	0.955	0.064	0.876	1.000	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.948	0.064	0.865	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.933	0.066	0.864	1.000	1.000	1.000
		$0.2$	1.000	1.000	1.000	0.934	0.061	0.870	1.000	1.000	1.000
		$0.4$	1.000	1.000	1.000	0.942	0.056	0.885	1.000	1.000	1.000
$LM_j$	$\sigma_\mu^2 = 0.5$	$\lambda = -0.4$	1.000	1.000	1.000	1.000	0.990	0.018	0.998	1.000	1.000
		$-0.2$	1.000	1.000	1.000	1.000	0.658	0.577	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.999	0.049	0.996	1.000	1.000	1.000
		$0.2$	1.000	1.000	1.000	0.642	0.810	1.000	1.000	1.000	1.000
		$0.4$	1.000	1.000	0.983	0.351	1.000	1.000	1.000	1.000	1.000
$LM_j^*$	$\sigma_\mu^2 = 0.5$	$\lambda = -0.4$	1.000	1.000	1.000	0.853	0.086	0.932	1.000	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.933	0.060	0.893	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.949	0.049	0.884	1.000	1.000	1.000
		$0.2$	1.000	1.000	1.000	0.940	0.035	0.921	1.000	1.000	1.000
		$0.4$	1.000	1.000	1.000	0.903	0.058	0.974	1.000	1.000	1.000
$LM_k$	$\sigma_\mu^2 = 0.5$	$\lambda = -0.4$	1.000	1.000	1.000	0.958	0.055	0.899	1.000	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.949	0.048	0.888	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.947	0.050	0.884	1.000	1.000	1.000
		$0.2$	1.000	1.000	1.000	0.949	0.046	0.897	1.000	1.000	1.000
		$0.4$	1.000	1.000	1.000	0.953	0.046	0.911	1.000	1.000	1.000

Finally, the performances of the test statistics for spatial lag dependence are summarized in Tables 7 and 8. We focus on the (49, 7) sample for discussion. For  $LM_l$ , the empirical size is 0.044, and it is very powerful in detecting the spatial lag dependence given that  $\sigma_\mu^2 = \rho = 0$ . For example, FoR of  $LM_l$  is 0.987 when  $\lambda = -0.4$ . However, as discussed in Section 3.3.3,  $LM_l$  is not robust against local misspecification, and this is confirmed by the simulation results. Given  $\lambda = 0$ , when  $\sigma_\mu^2$  and  $\rho$  deviate locally from 0, the large FoRs of  $LM_l$  are undesirable. For example, when  $\lambda = 0$  but  $\sigma_\mu^2 = \rho = 0.2$ ,

FoR of  $LM_l$  is 0.526. This size distortion is to some extent avoided by  $LM_l^*$ . When  $\lambda = 0$  and  $\sigma_\mu^2$  takes values 0, 0.2 and  $\rho \in [-0.2, 0.2]$ , FoRs of  $LM_l^*$  are in the range of [0.035, 0.088]. However,  $LM_l^*$  should be used with caution in this case, since the range of  $\rho$  in which  $LM_l^*$  provides valid size is narrow, and this becomes even worse for the (100, 10) sample. On the other hand, the robust LM test is supposed to be less powerful than the corresponding LM test when the nuisance parameters are correctly specified. When  $\sigma_\mu^2 = \rho = 0$ ,  $LM_l^*$  is less powerful than  $LM_l$ . For example, when  $\lambda = -0.4$ , FoR of  $LM_l$  is 0.987, while that of  $LM_l^*$  is 0.809. Next,  $LM_m$  tests for the spatial lag dependence in a pooled panel data model with spatial error correlation. The empirical sizes of  $LM_m$  vary in [0.037, 0.055]. As  $\lambda$  deviates from 0, FoR increases. Moreover, as discussed in Section 3.3.3, the robust LM test in this case is the same as  $LM_m$ , implying that  $LM_m$  itself is robust against local deviation of  $\sigma_\mu^2$  from 0. This is confirmed in the simulation results. When  $\sigma_\mu^2 = 0.2, \lambda = 0$ , the FoRs of  $LM_m$  vary in [0.060, 0.74]. Next, we discuss  $LM_n$  and  $LM_n^*$ .  $LM_n$  tests for the spatial lag dependence in a random effects panel data model without spatial error correlation. We only present the representative result when  $\sigma_\mu^2 = 0.5$ , results in other cases are very similar. The empirical size of  $LM_n$  is 0.035, and the FoR increases as  $\lambda$  deviates away from 0. For example, FoR is 0.981 when  $\lambda = -0.4$ . However,  $LM_n$  suffers from size distortion when the nuisance parameter  $\rho$  deviates away from 0. For example, when  $\lambda = 0, \rho = -0.2$ , the FoR of  $LM_n$  is 0.476, which is undesirably high. For the same case, FoR of  $LM_n^*$  is 0.063. However, as  $LM_l^*$ ,  $LM_n^*$  should also be used with caution, since the range of  $\rho$  in which  $LM_n^*$  provides valid size is narrow, and it becomes even worse for the (100, 10) sample. On the other hand,  $LM_n^*$  is less powerful than  $LM_n$  when  $\rho = 0$ . For example, when  $\rho = 0, \lambda = -0.4$ , FoR of  $LM_n^*$  is 0.778, while that of  $LM_n$  is 0.981. Lastly, we discuss the performance of  $LM_o$ .  $LM_o$  tests for the spatial lag dependence in a random effects panel data model with spatial error correlation. The empirical sizes of  $LM_o$  vary in [0.063, 0.079]. As expected, when  $\lambda$  moves away from 0, FoR increases, suggesting its good power performance. For all of the above tests, their performances for the (100, 10) sample follows a similar discussion. In sum, the test statistics  $LM_l, LM_l^*, LM_m, LM_n, LM_n^*$ , and  $LM_o$  are all useful for detecting the spatial lag dependence, but they are suited for different assumptions about the nuisance parameters. In practice, researchers are suggested to analyze them together to draw correct inference on  $\lambda$ .

**Table 7.** FoR of LM Tests for the Spatial Lag Dependence, Sample Size: (49, 7).

			$\lambda = -0.8$	$-0.6$	$-0.4$	$-0.2$	$0$	$0.2$	$0.4$	$0.6$	$0.8$
$LM_l$	$\sigma_\mu^2 = 0$	$\rho = -0.4$	1.000	1.000	1.000	1.000	0.965	0.221	0.405	1.000	1.000
		$-0.2$	1.000	1.000	1.000	0.982	0.446	0.063	0.938	1.000	1.000
		$0.0$	1.000	1.000	0.987	0.591	0.044	0.693	0.998	1.000	1.000
		$0.2$	1.000	0.984	0.631	0.067	0.576	0.988	1.000	1.000	1.000
		$0.4$	0.979	0.510	0.061	0.593	0.992	1.000	1.000	1.000	1.000
	$\sigma_\mu^2 = 0.2$	$\rho = -0.4$	1.000	1.000	1.000	1.000	0.964	0.285	0.305	0.999	1.000
		$-0.2$	1.000	1.000	1.000	0.968	0.505	0.071	0.892	1.000	1.000
		$0.0$	1.000	1.000	0.985	0.620	0.051	0.637	0.998	1.000	1.000
		$0.2$	1.000	0.987	0.637	0.094	0.526	0.981	1.000	1.000	1.000
		$0.4$	0.963	0.502	0.081	0.599	0.981	1.000	1.000	1.000	1.000

Table 7. Cont.

			$\lambda = -0.8$	$-0.6$	$-0.4$	$-0.2$	$0$	$0.2$	$0.4$	$0.6$	$0.8$		
$LM_l^*$	$\sigma_\mu^2 = 0$	$\rho = -0.4$	0.626	0.318	0.128	0.121	0.337	0.801	0.995	1.000	1.000		
		$-0.2$	0.968	0.855	0.547	0.163	0.054	0.497	0.983	1.000	1.000		
		$0.0$	0.999	0.974	0.809	0.338	0.035	0.459	0.978	1.000	1.000		
		$0.2$	0.999	0.991	0.858	0.305	0.066	0.649	0.993	1.000	1.000		
		$0.4$	1.000	0.976	0.651	0.121	0.323	0.922	1.000	1.000	1.000		
	$\sigma_\mu^2 = 0.2$	$\rho = -0.4$	0.582	0.318	0.157	0.153	0.353	0.753	0.992	1.000	1.000		
		$-0.2$	0.961	0.812	0.510	0.181	0.088	0.450	0.964	1.000	1.000		
		$0.0$	0.998	0.963	0.773	0.366	0.063	0.421	0.963	1.000	1.000		
		$0.2$	0.999	0.985	0.824	0.324	0.082	0.614	0.991	1.000	1.000		
		$0.4$	1.000	0.952	0.641	0.145	0.309	0.904	1.000	1.000	1.000		
		$LM_m$	$\sigma_\mu^2 = 0$	$\rho = -0.4$	1.000	1.000	0.958	0.490	0.050	0.484	0.981	1.000	1.000
				$-0.2$	1.000	0.998	0.941	0.444	0.037	0.422	0.966	1.000	0.973
$0.0$	1.000			0.998	0.900	0.406	0.055	0.371	0.909	0.997	0.757		
$0.2$	1.000			0.988	0.862	0.360	0.049	0.285	0.769	0.923	0.556		
$\sigma_\mu^2 = 0.2$	$\rho = -0.4$		0.998	0.971	0.756	0.263	0.046	0.233	0.575	0.766	0.366		
	$-0.4$		1.000	1.000	0.978	0.543	0.067	0.493	0.990	1.000	1.000		
	$-0.2$		1.000	0.999	0.929	0.478	0.074	0.446	0.964	1.000	0.962		
$LM_n$	$\sigma_\mu^2 = 0.5$	$\rho = -0.4$	1.000	1.000	1.000	1.000	0.970	0.269	0.291	0.999	1.000		
		$-0.2$	1.000	1.000	1.000	0.979	0.476	0.063	0.917	1.000	1.000		
		$0.0$	1.000	1.000	0.981	0.591	0.035	0.676	1.000	1.000	1.000		
		$0.2$	1.000	0.984	0.612	0.067	0.557	0.989	1.000	1.000	1.000		
		$0.4$	0.965	0.461	0.080	0.629	0.986	1.000	1.000	1.000	1.000		
	$LM_n^*$	$\sigma_\mu^2 = 0.5$	$\rho = -0.4$	0.577	0.285	0.130	0.137	0.353	0.779	0.991	1.000	1.000	
			$-0.2$	0.966	0.823	0.515	0.153	0.063	0.453	0.972	1.000	1.000	
$LM_o$	$\sigma_\mu^2 = 0.5$	$\rho = -0.4$	0.999	0.971	0.778	0.328	0.051	0.421	0.973	1.000	1.000		
		$0.0$	0.999	0.983	0.830	0.287	0.072	0.642	0.996	1.000	1.000		
		$0.2$	1.000	0.952	0.623	0.103	0.326	0.928	1.000	1.000	1.000		
		$0.2$	1.000	1.000	0.961	0.504	0.064	0.532	0.992	1.000	0.998		
		$0.4$	1.000	0.999	0.933	0.477	0.070	0.484	0.982	1.000	0.964		
		$0.0$	1.000	0.997	0.899	0.442	0.078	0.438	0.935	0.996	0.644		
		$0.2$	1.000	0.999	0.918	0.475	0.079	0.347	0.818	0.892	0.371		
		$0.4$	1.000	0.995	0.855	0.381	0.063	0.205	0.460	0.522	0.275		

Table 8. FoR of LM Tests for the Spatial Lag Dependence, Sample Size: (100, 10).

			$\lambda = -0.8$	$-0.6$	$-0.4$	$-0.2$	$0$	$0.2$	$0.4$	$0.6$	$0.8$
$LM_l$	$\sigma_\mu^2 = 0$	$\rho = -0.4$	1.000	1.000	1.000	1.000	1.000	0.608	0.699	1.000	1.000
		$-0.2$	1.000	1.000	1.000	1.000	0.853	0.108	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.953	0.056	0.975	1.000	1.000	1.000
		$0.2$	1.000	1.000	0.939	0.068	0.922	1.000	1.000	1.000	1.000
		$0.4$	1.000	0.825	0.102	0.978	1.000	1.000	1.000	1.000	1.000
	$\sigma_\mu^2 = 0.2$	$\rho = -0.4$	1.000	1.000	1.000	1.000	1.000	0.660	0.581	1.000	1.000
		$-0.2$	1.000	1.000	1.000	1.000	0.861	0.105	1.000	1.000	1.000
		$0.0$	1.000	1.000	1.000	0.929	0.082	0.959	1.000	1.000	1.000
		$0.2$	1.000	1.000	0.930	0.087	0.889	1.000	1.000	1.000	1.000
		$0.4$	1.000	0.787	0.148	0.961	1.000	1.000	1.000	1.000	1.000

Table 8. Cont.

			$\lambda = -0.8$	$-0.6$	$-0.4$	$-0.2$	$0$	$0.2$	$0.4$	$0.6$	$0.8$		
$LM_t^*$	$\sigma_\mu^2 = 0$	$\rho = -0.4$	0.955	0.579	0.152	0.198	0.731	0.995	1.000	1.000	1.000		
		$-0.2$	1.000	0.997	0.908	0.316	0.101	0.893	1.000	1.000	1.000		
		$0.0$	1.000	1.000	0.993	0.723	0.058	0.844	1.000	1.000	1.000		
		$0.2$	1.000	1.000	0.997	0.642	0.111	0.959	1.000	1.000	1.000		
		$0.4$	1.000	1.000	0.946	0.128	0.704	1.000	1.000	1.000	1.000		
	$\sigma_\mu^2 = 0.2$	$\rho = -0.4$	0.927	0.542	0.175	0.238	0.715	0.988	1.000	1.000	1.000		
		$-0.2$	1.000	0.989	0.860	0.327	0.137	0.849	1.000	1.000	1.000		
		$0.0$	1.000	1.000	0.994	0.684	0.093	0.798	1.000	1.000	1.000		
		$0.2$	1.000	1.000	0.992	0.612	0.131	0.948	1.000	1.000	1.000		
		$0.4$	1.000	1.000	0.921	0.153	0.679	1.000	1.000	1.000	1.000		
		$LM_m$	$\sigma_\mu^2 = 0$	$\rho = -0.4$	1.000	1.000	1.000	0.897	0.052	0.904	1.000	1.000	1.000
				$-0.2$	1.000	1.000	1.000	0.864	0.054	0.860	1.000	1.000	1.000
$0.0$	1.000			1.000	1.000	0.794	0.053	0.794	1.000	1.000	0.995		
$0.2$	1.000			1.000	0.997	0.712	0.055	0.679	0.994	1.000	0.883		
$\sigma_\mu^2 = 0.2$	$\rho = -0.4$		1.000	1.000	1.000	0.870	0.068	0.858	1.000	1.000	1.000		
	$-0.2$		1.000	1.000	1.000	0.810	0.075	0.800	1.000	1.000	1.000		
	$0.0$		1.000	1.000	1.000	0.795	0.075	0.732	0.999	1.000	0.989		
$LM_n$	$\sigma_\mu^2 = 0.5$	$\rho = -0.4$	1.000	1.000	1.000	1.000	1.000	0.648	0.634	1.000	1.000		
		$-0.2$	1.000	1.000	1.000	1.000	0.868	0.086	1.000	1.000	1.000		
		$0.0$	1.000	1.000	1.000	0.941	0.053	0.975	1.000	1.000	1.000		
		$0.2$	1.000	1.000	0.930	0.056	0.930	1.000	1.000	1.000	1.000		
		$0.4$	1.000	0.795	0.124	0.980	1.000	1.000	1.000	1.000	1.000		
$LM_n^*$	$\sigma_\mu^2 = 0.5$	$\rho = -0.4$	0.941	0.531	0.132	0.227	0.726	0.995	1.000	1.000	1.000		
		$-0.2$	1.000	0.995	0.884	0.305	0.105	0.874	1.000	1.000	1.000		
		$0.0$	1.000	1.000	0.995	0.702	0.055	0.813	1.000	1.000	1.000		
		$0.2$	1.000	1.000	0.999	0.624	0.100	0.961	1.000	1.000	1.000		
		$0.4$	1.000	1.000	0.940	0.137	0.731	1.000	1.000	1.000	1.000		
$LM_o$	$\sigma_\mu^2 = 0.5$	$\rho = -0.4$	1.000	1.000	1.000	0.884	0.062	0.992	1.000	1.000	1.000		
		$-0.2$	1.000	1.000	1.000	0.858	0.076	0.880	1.000	1.000	1.000		
		$0.0$	1.000	1.000	1.000	0.830	0.084	0.812	1.000	1.000	0.979		
		$0.2$	1.000	1.000	1.000	0.759	0.076	0.696	0.994	1.000	0.761		
		$0.4$	1.000	1.000	0.994	0.647	0.070	0.497	0.925	0.966	0.746		

### 5. Empirical Illustration

In this section, we revisit the empirical example in Baltagi and Levin (1992) [32]. For the purpose of illustrating usefulness of the test statistics in our framework, we estimate a static demand model for cigarettes as in Elhorst (2014) [25]. The data is obtained from the Wiley website, it is a panel data of 45 U.S. states and Washington D.C. over the period of 1963–1992.<sup>4</sup> We estimate the regression equation

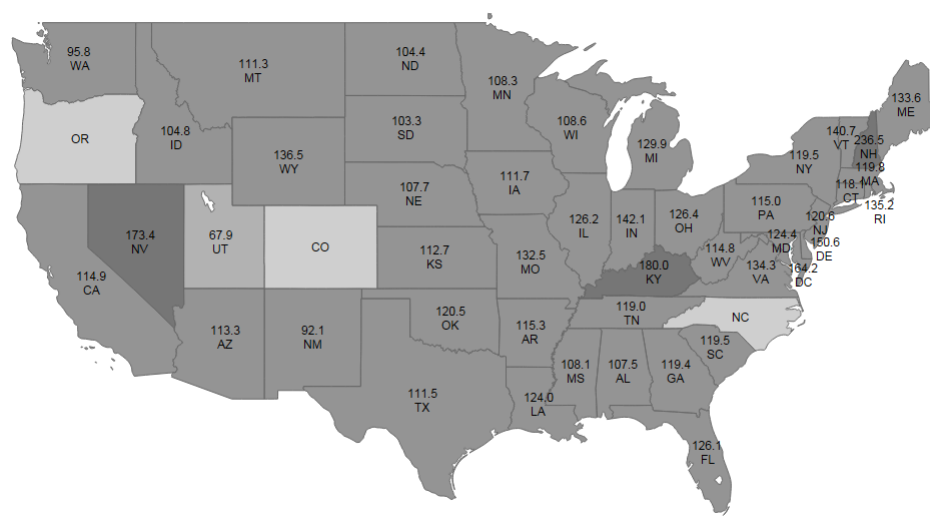
$$\begin{cases} \ln C_t = \lambda W \ln C_t + \beta_0 + \beta_1 \ln P_t + \beta_2 \ln Y_t + \epsilon_t, \\ \epsilon_t = \rho M \epsilon_t + \mu + v_t, \end{cases}$$

for  $t = 1, \dots, T$ . In the above specification,  $C_t$  is the vector of average cigarettes consumption (in packs) per capita (14 years and older) for all the states in a given year  $t$ .  $P_t$  is the corresponding price (per pack)

<sup>4</sup> The data for Colorado, Oregon and Pennsylvania are not available.

vector in year  $t$ , with  $\beta_1$  capturing the price effect on the demand for cigarettes.  $Y_t$  is the vector of per capita disposable income in year  $t$ , with  $\beta_2$  capturing the income effect on the demand for cigarettes. In the regression analysis, in addition to the standard price and income effects, we are particularly interested in whether and how the consumption of cigarettes in one state is related to those of its neighboring states. This is reflected by the spatial lag dependence structure (or the corresponding parameter  $\lambda$ ). Further, we use the error component structure to allow for individual random effects and spatial correlation in the error term. For the spatial weights matrices, we choose two specifications and try different combinations of them. One is the standard row-normalized first-order rook contiguity matrix. The other one is the row-standardized border length weights matrix as suggested in Debarsy *et al.*, (2012) [33].<sup>5</sup>

Since we are particularly interested in whether the consumptions of cigarettes are spatially correlated and how they are correlated, we plot the average consumption of cigarettes (in packs) per capita over the 30 years for all the states in Figure 1. On the one hand, we see that the cigarette consumptions in some states are negatively correlated with those of its neighboring states. For example, Utah, which has a high percentage of Mormon population,<sup>6</sup> has the lowest consumption of cigarettes, with the average consumption to be 67.9 packs per capita per year. In sharp contrast, the consumption of cigarettes in Nevada is nearly 2.6 times of that of Utah, and this could be attributed to the fact that Nevada is a highly tourist state with many casinos. The cigarette consumption in Utah is negatively correlated with those of its neighboring states. This similar pattern also holds for Nevada, New Mexico, Kentucky, and New Hampshire. On the other hand, we see that the consumptions of cigarettes among many other states tend to be similar and thus positively correlated, examples include Iowa, Wisconsin, Pennsylvania, Alabama, Georgia, and so on. The overall spatial dependence is not clear and we will use the formal regression analysis to assess it below.



**Figure 1.** Average Consumption of Cigarettes (in packs) per Capita per Year by States, 1963–1992.

<sup>5</sup> The data for border length is obtained from Holmes [34].

<sup>6</sup> According to the 2010 United States Census, the Mormons represent 62.1% of Utah's total population.

Before the regression analysis, we first perform the diagnostic procedures and the results are summarized in Table 9. The first row in Table 9 shows that the joint null hypothesis  $H_0^a$  is strongly rejected, thus the classical pooled panel data model not appropriate for this data set, and the OLS estimator is biased. The next four rows show that the random effects model is strongly favored against the pooled panel data model, regardless of the assumptions imposed on the spatial parameters. Next, when we test for the spatial error effect and the spatial lag effect jointly, the values of  $LM_f$  and  $LM_g$  show that the null hypothesis  $\rho = \lambda = 0$  is rejected no matter we assume  $\sigma_\mu^2 = 0$  or allow for the case  $\sigma_\mu^2 > 0$ . This points us to a random effects model with at least one type of spatial effects. In the next six rows, we focus on testing for the spatial error correlation. Although these test statistics unanimously reject the null hypothesis  $\rho = 0$  in favor of a model with spatial error correlation, we point out that the appropriate statistics should be  $LM_j$ ,  $LM_j^*$ , and  $LM_k$  given the presence of random effects but not spatial lag dependence yet at this stage. Finally, we test for the spatial lag dependence, the results in the last six rows suggest rejecting the null hypothesis of no spatial lag dependence, although there is a small value of  $LM_l^*$ . Given that the random effects and the spatial error correlation are present by results above, we point out that the appropriate statistics should be  $LM_l^*$ ,  $LM_n^*$ , and  $LM_o$ . From the estimation results in Table 10, the estimated  $\rho$  is between 0.353 and 0.622 and thus it is not proper to consider such value of  $\rho$  as local deviation from 0. As a result, the robust LM tests do not apply. We are left with  $LM_o$ , and it points to the presence of spatial lag dependence. In the following, we first estimate the random effects model with spatial error correlation using two types of spatial weights matrices, then we add in the spatially lagged dependent variable as suggested by our test statistics.

The estimation results are summarized in Table 10. The first column provides the OLS benchmark estimation result for comparison purpose. Although it is biased, all the parameter estimates have expected signs. Price has a negative effect on the consumption of cigarettes per capita, and the disposable income per capita has a positive effect on the consumption of cigarettes per capita, which are consistent with the standard consumption theory. In Columns 2 and 4, we include the spatial error correlation for the two different spatial weights matrices. Compared to the OLS estimates, the price and income elasticity all become smaller in magnitude but have the same signs. The estimated spatial error correlation are positive as 0.353 and 0.364, which do not differ much for the two different spatial weights matrices. In Columns 3, 5, 6, and 7, we estimate the full model with different combinations of the weights matrices. The estimates for the spatial lag dependence is negative and statistically significant, which is consistent with our diagnostic testing results. We thus conclude that the full model is the more appropriate specification, and overall we find that the negative correlation dominates in the spatial dependence among cigarette consumptions.

**Table 9.** LM Testing Statistics.

	<b>W = Rook M = Rook</b>	<b>W = Border M = Border</b>	<b>W = Rook M = Border</b>	<b>W = Border M=Rook</b>
Joint Test $LM_a$	12559 ***	12542 ***	12532 ***	12555 ***
Testing for Random Effects				
$LM_b$	12471 ***	12471 ***	12471 ***	12471 ***
$LM_c$	12207 ***	12260 ***	12260 ***	12207 ***
$LM_d$	12471 ***	12530 ***	12471 ***	12530 ***
$LM_e$	1354.7 ***	1662.7 ***	1573.1 ***	1457.8 ***
Testing for Spatial Effects				
Spatial Error and Lag				
$LM_f$	88.13 ***	70.78 ***	60.82 ***	83.76 ***
$LM_g$	172.81 ***	188.36 ***	174.62 ***	143.68 ***
Spatial Error				
$LM_h$	76.35 ***	60.72 ***	60.72 ***	76.35 ***
$LM_h^*$	51.78 ***	42.07 ***	24.47 ***	55.05 ***
$LM_i$	32.39 ***	23.68 ***	10.33 ***	38.02 ***
$LM_j$	138.96 ***	156.08 ***	156.08 ***	138.96 ***
$LM_j^*$	126.82 ***	126.41 ***	128.63 ***	81.72 ***
$LM_k$	94.01 ***	112.90 ***	99.03 ***	58.83 ***
Spatial Lag				
$LM_l$	36.35 ***	28.71 ***	36.35 ***	28.71 ***
$LM_l^*$	11.77 ***	10.06 ***	0.10	7.41 ***
$LM_m$	1147.00 ***	1385.40 ***	1028.20 ***	580.53 ***
$LM_n$	45.99 ***	61.95 ***	45.99 ***	61.95 ***
$LM_n^*$	33.85 ***	32.28 ***	18.53 ***	4.72 ***
$LM_o$	133.96 ***	283.36 ***	108.62 ***	80.05 ***

\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

**Table 10.** Estimation of the Cigarette Demand Function.

<b>Model</b>	<b>OLS</b>	<b>MLE</b>	<b>MLE</b>	<b>MLE</b>	<b>MLE</b>	<b>MLE</b>	<b>MLE</b>
		<b>(W = M = Rook)</b>		<b>(W = M = Border)</b>		<b>(W = Rook) (M = Border)</b>	<b>(W = Border) (M = Rook)</b>
$\hat{\beta}_0$	2.825 *** (0.098)	2.918 *** (0.086)	4.267 *** (0.241)	2.949 *** (0.087)	3.737 *** (0.196)	4.634 *** (0.261)	3.227 *** (0.169)
$\hat{\beta}_1$	-0.773 *** (0.026)	-0.739 *** (0.021)	-0.867 *** (0.026)	-0.729 *** (0.021)	-0.821 *** (0.027)	-0.851 *** (0.025)	-0.788 *** (0.030)
$\hat{\beta}_2$	0.586 *** (0.022)	0.559 *** (0.018)	0.645 *** (0.022)	0.551 *** (0.018)	0.616 *** (0.022)	0.628 *** (0.022)	0.595 *** (0.024)
$\hat{\lambda}$			-0.329 *** (0.044)		-0.204 *** (0.039)	-0.388 *** (0.044)	-0.088 ** (0.038)
$\hat{\rho}$		0.353 *** (0.030)	0.586 *** (0.034)	0.364 *** (0.028)	0.510 *** (0.034)	0.622 *** (0.032)	0.419 *** (0.039)
$\hat{\sigma}_\mu^2$		0.152 *** (0.015)	0.140 *** (0.014)	0.154 *** (0.016)	0.145 *** (0.015)	0.140 *** (0.014)	0.149 *** (0.015)
$\hat{\sigma}_v^2$		0.075 *** (0.001)	0.069 *** (0.002)	0.073 *** (0.001)	0.071 *** (0.002)	0.067 *** (0.002)	0.074 *** (0.001)
log-likelihood	450.94	1489.2	1514.7	1502.4	1514.6	1537.9	1491.7

\*  $p < 0.1$ ; \*\*  $p < 0.05$ ; \*\*\*  $p < 0.01$ .

## 6. Conclusions

In this paper, we propose a panel data random effects models with both spatially correlated error components and spatially lagged dependent variables. We consider diagnostic testing within such a framework. We first derive the joint LM test for the individual random effects, the spatial error correlation and the spatial lag dependence. In practice, applied researchers should first consider this joint test. If the joint null hypothesis cannot be rejected, it is reasonable to adopt the classical pooled panel data model. Otherwise, either the individual random effects, or the spatial error correlation, or the spatial lag dependence must be taken into consideration. Next, we derive a wide range of LM tests for the individual random effects and for the two spatial effects separately. In addition, in order to guard against possible local model misspecification, we apply the Bera and Yoon (1993) [19] principle and construct robust LM tests in some cases. These test statistics complement each other and should be used together in performing diagnostic test to search for the most appropriate model. A small Monte Carlo experiment is carried out and the size and power performances of these test statistics are satisfactory. We further use the cigarette demand data set in Baltagi and Levin (1992) [32] to illustrate our testing procedures.

Some future research directions can be considered. First, for the model specification of spatially correlated error components used in this paper, although it allows for both spatial spillovers of permanent and temporary shocks, it does not permit different intensities of these two shocks. It would be of interest to relax this assumption and further generalize our model (see Baltagi *et al.*, 2013 [12]). Second, one can borrow the ideas in Baltagi and Yang (2013a, 2013b) [15,17] to modify our test statistics in order to remedy distributional misspecifications in finite sample, sensitivity to spatial layout, or to be robust against unknown heteroskedasticity.

## Acknowledgments

We thank the Editor-in-Chief, Kerry Patterson, and two anonymous referees for their constructive comments that have greatly improved the article. We also thank Anil Bera and participants at the IVth World Conference of Spatial Econometric Association for helpful discussions. All remaining errors are our own.

## Author Contributions

Kuan-Pin Lin suggested the problem and Ming He solved the problem. Both authors contributed to the final paper.

## Conflicts of Interest

The authors declare no conflict of interest.



**Appendix**

In the appendices, we provide detailed derivations of the score vector, information matrix and the LM test statistics in each case.

**Appendix A**

In this appendix, we provide formulae of the score vector and the information matrix for our general model specification. Let  $R_1 = M(I_N - \rho M)^{-1}$ ,  $R_2 = W(I_N - \rho M)^{-1}$ ,  $R_3 = W(I_N - \lambda W)^{-1}$  and  $R_4 = (I_N - \rho M')(I_N - \rho M)$ . The score vector is

$$\frac{\partial L(\delta)}{\partial \delta} = \begin{pmatrix} s_\beta \\ s_\rho \\ s_\lambda \\ s_{\sigma_\mu^2} \\ s_{\sigma_v^2} \end{pmatrix} = \begin{pmatrix} X' A' \Omega^{-1} A \epsilon \\ -T \text{tr}(R_1) + \epsilon' A' \Omega^{-1} (I_T \otimes M) \epsilon \\ -T \text{tr}(R_3) + \epsilon' A' \Omega^{-1} A (I_T \otimes W) y \\ -\frac{NT}{2(T\sigma_\mu^2 + \sigma_v^2)} + \frac{T}{2(T\sigma_\mu^2 + \sigma_v^2)^2} \epsilon' A' (\bar{J}_T \otimes I_N) A \epsilon \\ -\frac{N}{2(T\sigma_\mu^2 + \sigma_v^2)} - \frac{N(T-1)}{2\sigma_v^2} - \frac{1}{2} \epsilon' A' \frac{\partial \Omega^{-1}}{\partial \sigma_v^2} A \epsilon \end{pmatrix},$$

where  $\epsilon = By - X\beta$ . Let  $\mathcal{I}$  be the information matrix, that is,  $\mathcal{I} = -E[\partial^2 L(\delta)/\partial \delta \partial \delta']$ . Then after some routine calculation, the elements of  $\mathcal{I}$  are given by

$$\begin{aligned} \mathcal{I}_{\beta\beta'} &= X' A' \Omega^{-1} A X, \quad \mathcal{I}_{\beta\rho} = 0, \quad \mathcal{I}_{\beta\lambda} = X' A' \Omega^{-1} A (I_T \otimes W) B^{-1} X \beta, \quad \mathcal{I}_{\beta\sigma_\mu^2} = 0 \quad \mathcal{I}_{\beta\sigma_v^2} = 0, \\ \mathcal{I}_{\rho\rho} &= T \text{tr}(R_1 R_1 + R_1 R_1'), \quad \mathcal{I}_{\rho\lambda} = T \text{tr}(R_3 R_1) + T \text{tr}[R_3 (I_N - \rho M)^{-1} R_1' (I_N - \rho M)], \\ \mathcal{I}_{\rho\sigma_\mu^2} &= \frac{T}{T\sigma_\mu^2 + \sigma_v^2} \text{tr}(R_1), \quad \mathcal{I}_{\rho\sigma_v^2} = \left( \frac{1}{T\sigma_\mu^2 + \sigma_v^2} + \frac{T-1}{\sigma_v^2} \right) \text{tr}(R_1), \\ \mathcal{I}_{\lambda\lambda} &= T \text{tr}(R_3 R_3) + T \text{tr}(R_4 R_3 R_4^{-1} R_3') + (B^{-1} X \beta)' (I_T \otimes W') A' \Omega^{-1} A (I_T \otimes W) B^{-1} X \beta, \\ \mathcal{I}_{\lambda\sigma_\mu^2} &= \frac{T}{T\sigma_\mu^2 + \sigma_v^2} \text{tr}(R_3), \quad \mathcal{I}_{\lambda\sigma_v^2} = \left( \frac{1}{T\sigma_\mu^2 + \sigma_v^2} + \frac{T-1}{\sigma_v^2} \right) \text{tr}(R_3), \\ \mathcal{I}_{\sigma_\mu^2 \sigma_\mu^2} &= \frac{NT^2}{2(T\sigma_\mu^2 + \sigma_v^2)^2}, \quad \mathcal{I}_{\sigma_\mu^2 \sigma_v^2} = \frac{NT}{2(T\sigma_\mu^2 + \sigma_v^2)^2}, \quad \mathcal{I}_{\sigma_v^2 \sigma_v^2} = \frac{N}{2(T\sigma_\mu^2 + \sigma_v^2)^2} + \frac{N(T-1)}{2(\sigma_v^2)^2}. \end{aligned}$$

Due to the large number of test statistics in this paper, it turns out to be convenient to introduce some general notations for reference and easy exposition. Let

$$\hat{z}_\rho = \hat{\epsilon}' \hat{A}' \hat{\Omega}^{-1} (I_T \otimes M) \hat{\epsilon}, \quad \hat{z}_\lambda = \hat{\epsilon}' \hat{A}' \hat{\Omega}^{-1} \hat{A} (I_T \otimes W) y, \quad \hat{z}_{\sigma_\mu^2} = \frac{\hat{\epsilon}' \hat{A}' (\bar{J}_T \otimes I_N) \hat{A} \hat{\epsilon}}{\hat{\sigma}_v^2} - N,$$

where  $\hat{\epsilon} = \hat{B}y - X\hat{\beta}$ , and  $\hat{A}, \hat{B}, \hat{\Omega}, \hat{\beta}, \hat{\sigma}_v^2$  are restricted MLEs of  $A, B, \Omega, \beta, \sigma_v^2$ , respectively. Next, define

$$\begin{aligned} \hat{\nu} &= \hat{y}' (I_T \otimes W') \hat{A}' \hat{\Omega}^{-1} \hat{A} (I_T \otimes W) \hat{y}, \quad \hat{\tau} = T^2 (b_1 b_3 - b_2^2) + T b_1 \hat{\omega}, \\ \hat{\omega} &= \hat{y}' (I_T \otimes W') \hat{A}' \left[ \hat{\Omega}^{-1} - \hat{\Omega}^{-1} \hat{A} X (X' \hat{A}' \hat{\Omega}^{-1} \hat{A} X)^{-1} X' \hat{A}' \hat{\Omega}^{-1} \right] \hat{A} (I_T \otimes W) \hat{y}, \end{aligned}$$

where  $\hat{y} = \hat{B}^{-1} X \hat{\beta}$ ,  $b_1 = \text{tr}(M' M + M M)$ ,  $b_2 = \text{tr}(M' W + M W)$ ,  $b_3 = \text{tr}(W' W + W W)$ . Finally, let

$$\hat{\xi} = \frac{NT\hat{\nu}_2 + N\hat{\omega} - 2T\hat{\nu}_3^2}{Tb_1(NT\hat{\nu}_2 + N\hat{\omega} - 2T\hat{\nu}_3^2) - N(T\hat{\nu}_1)^2}, \quad \hat{\zeta} = \frac{N\hat{\theta}_1 - 2\hat{\theta}_3^2}{(N\hat{\theta}_1 - 2\hat{\theta}_3^2)(T\hat{\theta}_4 + \hat{\omega}) - NT\hat{\theta}_2^2},$$

where  $\hat{\vartheta}_1 = \text{tr}[(M + M')\hat{R}_3]$ ,  $\hat{\vartheta}_2 = \text{tr}(\hat{R}_3\hat{R}_3 + \hat{R}_3\hat{R}_3')$ ,  $\hat{\vartheta}_3 = \text{tr}(\hat{R}_3)$ ,  $\hat{\theta}_1 = \text{tr}(\hat{R}_1\hat{R}_1 + \hat{R}_1\hat{R}_1')$ ,  $\hat{\theta}_2 = \text{tr}(W\hat{R}_1 + \hat{R}_2\hat{R}_1'(I_N - \hat{\rho}M))$ ,  $\hat{\theta}_3 = \text{tr}(\hat{R}_1)$ ,  $\hat{\theta}_4 = \text{tr}(WW) + \text{tr}(\hat{R}_2\hat{R}_2\hat{R}_4)$ , and  $\hat{R}_1, \hat{R}_2, \hat{R}_3, \hat{R}_4$  are restricted MLEs of  $R_1, R_2, R_3, R_4$ , respectively.

For notational convenience in deriving many LM test statistics in Appendix B, we write the general information matrix as

$$\mathcal{I} = \begin{pmatrix} \mathcal{I}_{\beta\beta'} & \mathcal{I}_{\beta\rho} & \mathcal{I}_{\beta\lambda} & \mathcal{I}_{\beta\sigma_\mu^2} & \mathcal{I}_{\beta\sigma_v^2} \\ \cdot & \mathcal{I}_{\rho\rho} & \mathcal{I}_{\rho\lambda} & \mathcal{I}_{\rho\sigma_\mu^2} & \mathcal{I}_{\rho\sigma_v^2} \\ \cdot & \cdot & \mathcal{I}_{\lambda\lambda} & \mathcal{I}_{\lambda\sigma_\mu^2} & \mathcal{I}_{\lambda\sigma_v^2} \\ \cdot & \cdot & \cdot & \mathcal{I}_{\sigma_\mu^2\sigma_\mu^2} & \mathcal{I}_{\sigma_\mu^2\sigma_v^2} \\ \cdot & \cdot & \cdot & \cdot & \mathcal{I}_{\sigma_v^2\sigma_v^2} \end{pmatrix},$$

and partition it into

$$\mathcal{I} = \begin{pmatrix} \mathcal{I}_{11} & \mathcal{I}_{12} \\ \mathcal{I}_{21} & \mathcal{I}_{22} \end{pmatrix}, \text{ where } \mathcal{I}_{11} = \mathcal{I}_{\beta\beta'}.$$

Now

$$\begin{aligned} & (\mathcal{I}_{22} - \mathcal{I}_{21}\mathcal{I}_{11}^{-1}\mathcal{I}_{12}) \\ &= \begin{pmatrix} \mathcal{I}_{\rho\rho} & \mathcal{I}_{\rho\lambda} & \mathcal{I}_{\rho\sigma_\mu^2} & \mathcal{I}_{\rho\sigma_v^2} \\ \cdot & \mathcal{I}_{\lambda\lambda} & \mathcal{I}_{\lambda\sigma_\mu^2} & \mathcal{I}_{\lambda\sigma_v^2} \\ \cdot & \cdot & \mathcal{I}_{\sigma_\mu^2\sigma_\mu^2} & \mathcal{I}_{\sigma_\mu^2\sigma_v^2} \\ \cdot & \cdot & \cdot & \mathcal{I}_{\sigma_v^2\sigma_v^2} \end{pmatrix} - \begin{pmatrix} \mathcal{I}_{\rho\beta'}\mathcal{I}_{\beta\beta'}^{-1} \\ \mathcal{I}_{\lambda\beta'}\mathcal{I}_{\beta\beta'}^{-1} \\ \mathcal{I}_{\sigma_\mu^2\beta'}\mathcal{I}_{\beta\beta'}^{-1} \\ \mathcal{I}_{\sigma_v^2\beta'}\mathcal{I}_{\beta\beta'}^{-1} \end{pmatrix} \begin{pmatrix} \mathcal{I}_{\beta\rho} & \mathcal{I}_{\beta\lambda} & \mathcal{I}_{\beta\sigma_\mu^2} & \mathcal{I}_{\beta\sigma_v^2} \end{pmatrix} \\ &= \begin{pmatrix} \mathcal{J}_{\rho\rho} & \mathcal{J}_{\rho\lambda} & \mathcal{J}_{\rho\sigma_\mu^2} & \mathcal{J}_{\rho\sigma_v^2} \\ \cdot & \mathcal{J}_{\lambda\lambda} & \mathcal{J}_{\lambda\sigma_\mu^2} & \mathcal{J}_{\lambda\sigma_v^2} \\ \cdot & \cdot & \mathcal{J}_{\sigma_\mu^2\sigma_\mu^2} & \mathcal{J}_{\sigma_\mu^2\sigma_v^2} \\ \cdot & \cdot & \cdot & \mathcal{J}_{\sigma_v^2\sigma_v^2} \end{pmatrix} \equiv \mathcal{J}, \end{aligned}$$

where  $\mathcal{J}_{s_1s_2} = \mathcal{I}_{s_1s_2} - \mathcal{I}_{s_1\beta'}\mathcal{I}_{\beta\beta'}^{-1}\mathcal{I}_{\beta s_2}$  for  $s_1, s_2 = \rho, \lambda, \sigma_\mu^2$  or  $\sigma_v^2$ . Further partition  $\mathcal{J}$  into

$$\mathcal{J} = \begin{pmatrix} \mathcal{J}_{11} & \mathcal{J}_{12} \\ \mathcal{J}_{21} & \mathcal{J}_{22} \end{pmatrix}, \text{ where } \mathcal{J}_{11} = \begin{pmatrix} \mathcal{J}_{\rho\rho} & \mathcal{J}_{\rho\lambda} \\ \cdot & \mathcal{J}_{\lambda\lambda} \end{pmatrix}.$$

Let  $\mathcal{K} = \mathcal{J}_{11} - \mathcal{J}_{12}\mathcal{J}_{22}^{-1}\mathcal{J}_{21}$  and  $\mathcal{L} = \mathcal{J}_{22} - \mathcal{J}_{21}\mathcal{J}_{11}^{-1}\mathcal{J}_{12}$ , thus the upper left 2 by 2 submatrix and lower right 2 by 2 submatrix of  $\mathcal{J}^{-1}$  are  $\mathcal{K}^{-1}, \mathcal{L}^{-1}$ , respectively.<sup>7</sup> Moreover, we use  $\mathcal{I}_{s_1s_2}^\Delta$  to denote the term in  $\mathcal{I}^{-1}$  with the same location as  $\mathcal{I}_{s_1s_2}$  in  $\mathcal{I}$ .

### Appendix B

In this appendix, we provide detailed derivations for all of the LM test statistics.

<sup>7</sup> For computing the inverse of a partitioned symmetric matrix, if  $\Lambda = \begin{pmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{21} & \Lambda_{22} \end{pmatrix}$ , then

$$\Lambda^{-1} = \begin{pmatrix} (\Lambda_{11} - \Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21})^{-1} & -\Lambda_{11}^{-1}\Lambda_{12}(\Lambda_{22} - \Lambda_{21}\Lambda_{11}^{-1}\Lambda_{12})^{-1} \\ -(\Lambda_{22} - \Lambda_{21}\Lambda_{11}^{-1}\Lambda_{12})^{-1}\Lambda_{21}\Lambda_{11}^{-1} & (\Lambda_{22} - \Lambda_{21}\Lambda_{11}^{-1}\Lambda_{12})^{-1} \end{pmatrix}.$$

**B.1. Derivation of LM<sub>a</sub>**

The restricted MLE under  $H_0^a$  is essentially the OLS estimator, *i.e.*,  $\hat{\beta}_a = (X'X)^{-1}X'y$ ,  $\hat{\sigma}_{v,a}^2 = (\hat{\epsilon}'_a \hat{\epsilon}_a)/(NT)$ , where  $\hat{\epsilon}_a = y - \hat{y}_a$ ,  $\hat{y}_a = X\hat{\beta}_a$ . The score vector under  $H_0^a$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_a = \begin{pmatrix} 0 & \hat{z}_{\rho,a} & \hat{z}_{\lambda,a} & \frac{T}{2\hat{\sigma}_{v,a}^2} \hat{z}_{\sigma_{\mu,a}^2} & 0 \end{pmatrix},$$

where  $\hat{z}_{\rho,a} = [\hat{\epsilon}'_a(I_T \otimes M)\hat{\epsilon}_a]/\hat{\sigma}_{v,a}^2$ ,  $\hat{z}_{\lambda,a} = [\hat{\epsilon}'_a(I_T \otimes W)y]/\hat{\sigma}_{v,a}^2$ ,  $\hat{z}_{\sigma_{\mu,a}^2} = [\hat{\epsilon}'_a(\bar{J}_T \otimes I_N)\hat{\epsilon}_a]/\hat{\sigma}_{v,a}^2 - N$ . After straightforward calculation, the information matrix under  $H_0^a$  and evaluated at the restricted MLE is

$$\mathcal{I}_a = \begin{pmatrix} \frac{X'X}{\hat{\sigma}_{v,a}^2} & 0 & \frac{X'(I_T \otimes W)\hat{y}_a}{\hat{\sigma}_{v,a}^2} & 0 & 0 \\ \cdot & Tb_1 & Tb_2 & 0 & 0 \\ \cdot & \cdot & Tb_3 + \hat{\nu}_a & 0 & 0 \\ \cdot & \cdot & \cdot & \frac{NT^2}{2(\hat{\sigma}_{v,a}^2)^2} & \frac{NT}{2(\hat{\sigma}_{v,a}^2)^2} \\ \cdot & \cdot & \cdot & \cdot & \frac{NT}{2(\hat{\sigma}_{v,a}^2)^2} \end{pmatrix},$$

where  $\hat{\nu}_a = [\hat{y}'_a(I_T \otimes (W'W))\hat{y}_a]/\hat{\sigma}_{v,a}^2$ . In this case,

$$\mathcal{J}_a = \begin{pmatrix} Tb_1 & Tb_2 & 0 & 0 \\ \cdot & Tb_3 + \hat{\omega}_a & 0 & 0 \\ \cdot & \cdot & \frac{NT^2}{2(\hat{\sigma}_{v,a}^2)^2} & \frac{NT}{2(\hat{\sigma}_{v,a}^2)^2} \\ \cdot & \cdot & \cdot & \frac{NT}{2(\hat{\sigma}_{v,a}^2)^2} \end{pmatrix}, \mathcal{K}_a = \begin{pmatrix} Tb_1 & Tb_2 \\ \cdot & Tb_3 + \hat{\omega}_a \end{pmatrix}, \mathcal{L}_a = \begin{pmatrix} \frac{NT^2}{2(\hat{\sigma}_{v,a}^2)^2} & \frac{NT}{2(\hat{\sigma}_{v,a}^2)^2} \\ \cdot & \frac{NT}{2(\hat{\sigma}_{v,a}^2)^2} \end{pmatrix},$$

where  $\hat{\omega}_a = [\hat{y}'_a(I_T \otimes W')(I_{NT} - X(X'X)^{-1}X')(I_T \otimes W)\hat{y}_a]/\hat{\sigma}_{v,a}^2$ . The first term of  $(\mathcal{L}_a)^{-1}$  is easily calculated to be  $2(\hat{\sigma}_{v,a}^2)^2/[NT(T - 1)]$ . Also, we have

$$(\mathcal{K}_a)^{-1} = \frac{1}{\hat{\tau}_a} \begin{pmatrix} Tb_3 + \hat{\omega}_a & -Tb_2 \\ -Tb_2 & Tb_1 \end{pmatrix}, \text{ where } \hat{\tau}_a = T^2(b_1b_3 - b_2^2) + Tb_1\hat{\omega}_a.$$

Finally, the joint LM test statistic is given by

$$LM_a = \frac{Tb_3 + \hat{\omega}_a}{\hat{\tau}_a} \hat{z}_{\rho,a}^2 + \frac{Tb_1}{\hat{\tau}_a} \hat{z}_{\lambda,a}^2 - \frac{2Tb_2}{\hat{\tau}_a} \hat{z}_{\rho,a} \hat{z}_{\lambda,a} + \frac{T}{2N(T - 1)} \hat{z}_{\sigma_{\mu,a}^2}^2.$$

**B.2. Derivation of LM<sub>c</sub>**

The score vector under  $H_0^c$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_c = \begin{pmatrix} 0 & 0 & \frac{T}{2\hat{\sigma}_{v,c}^2} \hat{z}_{\sigma_{\mu,c}^2} & 0 \end{pmatrix}, \text{ where } \hat{z}_{\sigma_{\mu,c}^2} = [\hat{\epsilon}'_c \hat{A}'(\bar{J}_T \otimes I_N) \hat{A} \hat{\epsilon}_c]/\hat{\sigma}_{v,c}^2 - N,$$

$\hat{\sigma}_{v,c}^2 = (\hat{\epsilon}'_c \hat{A}' \hat{A} \hat{\epsilon}_c)/(NT)$ ,  $\hat{\epsilon}_c = y - \hat{y}_c$ ,  $\hat{y}_c = X\hat{\beta}_c$ , and  $\hat{\beta}_c, \hat{A}_c$  are the restricted MLEs of  $\beta, A$  under  $H_0^c$ , respectively. After a little calculation, the information matrix under  $H_0^c$  and evaluated at the restricted MLE is

$$\mathcal{I}_c = \begin{pmatrix} \frac{X' \hat{A}'_c \hat{A}_c X}{\hat{\sigma}_{v,c}^2} & 0 & 0 & 0 \\ \cdot & T\hat{\theta}_{1,c} & \frac{T\hat{\theta}_{3,c}}{\hat{\sigma}_{v,c}^2} & \frac{T\hat{\theta}_{3,c}}{\hat{\sigma}_{v,c}^2} \\ \cdot & \cdot & \frac{NT^2}{2(\hat{\sigma}_{v,c}^2)^2} & \frac{NT}{2(\hat{\sigma}_{v,c}^2)^2} \\ \cdot & \cdot & \cdot & \frac{NT}{2(\hat{\sigma}_{v,c}^2)^2} \end{pmatrix},$$

where  $\widehat{\theta}_{1,c}, \widehat{\theta}_{3,c}$  are  $\widehat{\theta}_1, \widehat{\theta}_3$  evaluated under  $H_0^c$ , respectively. Making use of the block diagonal structure of  $\mathcal{I}_c$  and straightforward but tedious calculation gives that  $\mathcal{I}_{\sigma_\mu^2 \sigma_{\mu,c}^2}^\Delta = 2(\widehat{\sigma}_{v,c}^2)^2/(NT(T-1))$ . Finally, the LM test statistic in this case is given by

$$LM_c = \frac{T}{2N(T-1)} \widehat{z}_{\sigma_\mu^2,c}^2.$$

**B.3. Derivation of  $LM_e$**

The score vector under  $H_0^e$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_e = \left( 0 \quad 0 \quad 0 \quad \frac{T}{2\widehat{\sigma}_{v,e}^2} \widehat{z}_{\sigma_\mu^2,e} \quad 0 \right), \text{ where } \widehat{z}_{\sigma_\mu^2,e} = [\widehat{\epsilon}'_e \widehat{A}'_e (\bar{J}_T \otimes I_N) \widehat{A}_e \widehat{\epsilon}_e] / \widehat{\sigma}_{v,e}^2 - N,$$

$\widehat{\sigma}_{v,e}^2 = (\widehat{\epsilon}'_e \widehat{A}'_e \widehat{A}_e \widehat{\epsilon}_e) / (NT)$ ,  $\widehat{\epsilon}_e = \widehat{B}_e y - X \widehat{\beta}_e$ , and  $\widehat{\beta}_e, \widehat{A}_e, \widehat{B}_e$  are restricted MLEs of  $\beta, A, B$  under  $H_0^e$ , respectively. After a little calculation, the information matrix under  $H_0^e$  and evaluated at the restricted MLE is

$$\mathcal{I}_e = \begin{pmatrix} \frac{X' \widehat{A}'_e \widehat{A}_e X}{\widehat{\sigma}_{v,e}^2} & 0 & \frac{X' \widehat{A}'_e \widehat{A}_e (I_T \otimes W) \widehat{y}_e}{\widehat{\sigma}_{v,e}^2} & 0 & 0 \\ \cdot & \widehat{\mathcal{I}}_{\rho\rho,e} & \widehat{\mathcal{I}}_{\rho\lambda,e} & \frac{T\text{tr}(\widehat{R}_{1,e})}{\widehat{\sigma}_{v,e}^2} & \frac{T\text{tr}(\widehat{R}_{1,e})}{\widehat{\sigma}_{v,e}^2} \\ \cdot & \cdot & T\text{tr}(\widehat{R}_{3,e}^2) + T\text{tr}(\widehat{R}_{4,e} \widehat{R}_{3,e} \widehat{R}_{4,e}^{-1} \widehat{R}'_{3,e}) + \widehat{\nu}_e & \frac{T\text{tr}(\widehat{R}_{3,e})}{\widehat{\sigma}_{v,e}^2} & \frac{T\text{tr}(\widehat{R}_{3,e})}{\widehat{\sigma}_{v,e}^2} \\ \cdot & \cdot & \cdot & \frac{NT^2}{2(\widehat{\sigma}_{v,e}^2)^2} & \frac{NT}{2(\widehat{\sigma}_{v,e}^2)^2} \\ \cdot & \cdot & \cdot & \cdot & \frac{NT}{2(\widehat{\sigma}_{v,e}^2)^2} \end{pmatrix},$$

where  $\widehat{y}_e = \widehat{B}_e^{-1} X \widehat{\beta}_e$ ,  $\widehat{\nu}_e = \widehat{y}'_e (I_T \otimes W') \widehat{A}'_e \widehat{A}_e (I_T \otimes W) \widehat{y}_e / \widehat{\sigma}_{v,e}^2$ , and  $\widehat{\mathcal{I}}_{\rho\rho,e}, \widehat{\mathcal{I}}_{\rho\lambda,e}, \widehat{R}_{1,e}, \widehat{R}_{3,e}, \widehat{R}_{4,e}$  are restricted MLEs of  $\mathcal{I}_{\rho\rho}, \mathcal{I}_{\rho\lambda}, R_{1,e}, R_{3,e}, R_{4,e}$  under  $H_0^e$ , respectively. In this case, straightforward calculation gives

$$\mathcal{J}_e = \begin{pmatrix} \widehat{\mathcal{I}}_{\rho\rho,e} & \widehat{\mathcal{I}}_{\rho\lambda,e} & \frac{T\text{tr}(\widehat{R}_{1,e})}{\widehat{\sigma}_{v,e}^2} & \frac{T\text{tr}(\widehat{R}_{1,e})}{\widehat{\sigma}_{v,e}^2} \\ \cdot & T\text{tr}(\widehat{R}_{3,e}^2) + T\text{tr}(\widehat{R}_{4,e} \widehat{R}_{3,e} \widehat{R}_{4,e}^{-1} \widehat{R}'_{3,e}) + \widehat{\omega}_e & \frac{T\text{tr}(\widehat{R}_{3,e})}{\widehat{\sigma}_{v,e}^2} & \frac{T\text{tr}(\widehat{R}_{3,e})}{\widehat{\sigma}_{v,e}^2} \\ \cdot & \cdot & \frac{NT^2}{2(\widehat{\sigma}_{v,e}^2)^2} & \frac{NT}{2(\widehat{\sigma}_{v,e}^2)^2} \\ \cdot & \cdot & \cdot & \frac{NT}{2(\widehat{\sigma}_{v,e}^2)^2} \end{pmatrix},$$

where  $\widehat{\omega}_e = \widehat{y}'_e (I_T \otimes W') \widehat{A}'_e [I_{NT} - \widehat{A}_e X (X' \widehat{A}'_e \widehat{A}_e X)^{-1} X' \widehat{A}'_e] \widehat{A}_e (I_T \otimes W) \widehat{y}_e / \widehat{\sigma}_{v,e}^2$ . After some straightforward but tedious calculation, we get that the term  $\mathcal{I}_{\sigma_\mu^2 \sigma_{\mu,e}^2}^\Delta$  is  $2(\widehat{\sigma}_{v,e}^2)^2/[NT(T-1)]$ . Finally, the LM test statistic in this case is given by

$$LM_e = \frac{T}{2N(T-1)} \widehat{z}_{\sigma_\mu^2,e}^2.$$

**B.4. Derivation of  $LM_f$**

The restricted MLE under  $H_0^f$  is essentially the OLS estimator. The score vector under  $H_0^f$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_f = \left( 0 \quad \widehat{z}_{\rho,a} \quad \widehat{z}_{\lambda,a} \quad 0 \right).$$

After a little calculation, the information matrix under  $H_0^f$  and evaluated at the restricted MLE is

$$\mathcal{I}_f = \begin{pmatrix} \frac{X'X}{\widehat{\sigma}_{v,a}^2} & 0 & \frac{X'(I_T \otimes W)\widehat{y}_a}{\widehat{\sigma}_{v,a}^2} & 0 \\ \cdot & Tb_1 & Tb_2 & 0 \\ \cdot & \cdot & Tb_3 + \widehat{v}_a & 0 \\ \cdot & \cdot & \cdot & \frac{NT}{2(\widehat{\sigma}_{v,a}^2)^2} \end{pmatrix}.$$

In this case, it is very easy to calculate that

$$\begin{pmatrix} \mathcal{I}_{\rho\rho,f}^\Delta & \mathcal{I}_{\rho\lambda,f}^\Delta \\ \cdot & \mathcal{I}_{\lambda\lambda,f}^\Delta \end{pmatrix} = \frac{1}{\widehat{\tau}_a} \begin{pmatrix} Tb_3 + \widehat{\omega}_a & -Tb_2 \\ -Tb_2 & Tb_1 \end{pmatrix}, \text{ where } \widehat{\tau}_a = T^2(b_1b_3 - b_2^2) + Tb_1\widehat{\omega}_a.$$

Thus the LM test statistic in this case is given by

$$LM_f = \frac{Tb_3 + \widehat{\omega}_a}{\widehat{\tau}_a} \widehat{z}_{\rho,a}^2 + \frac{Tb_1}{\widehat{\tau}_a} \widehat{z}_{\lambda,a}^2 - \frac{2Tb_2}{\widehat{\tau}_a} \widehat{z}_{\rho,a} \widehat{z}_{\lambda,a}.$$

**B.5. Derivation of  $LM_g$**

The score vector under  $H_0^g$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_g = \begin{pmatrix} 0 & \widehat{z}_{\rho,g} & \widehat{z}_{\lambda,g} & 0 & 0 \end{pmatrix},$$

where  $\widehat{z}_{\rho,g} = \widehat{e}'_g \widehat{\Omega}_g^{-1} (I_T \otimes M) \widehat{\epsilon}_g$ ,  $\widehat{z}_{\lambda,g} = \widehat{e}'_g \widehat{\Omega}_g^{-1} (I_T \otimes W) y$ ,  $\widehat{\epsilon}_g = y - X\widehat{\beta}_g$ , and  $\widehat{\beta}_g, \widehat{\Omega}_g^{-1}$  are restricted MLEs of  $\beta, \Omega^{-1}$  under  $H_0^g$ , respectively. After a little calculation, the information matrix under  $H_0^g$  and evaluated at the restricted MLE is

$$\mathcal{I}_g = \begin{pmatrix} X'\widehat{\Omega}_g^{-1}X & 0 & X'\widehat{\Omega}_g^{-1}(I_T \otimes W)\widehat{y}_g & 0 & 0 \\ \cdot & Tb_1 & Tb_2 & 0 & 0 \\ \cdot & \cdot & Tb_3 + \widehat{v}_g & 0 & 0 \\ \cdot & \cdot & \cdot & \frac{NT^2}{2(T\widehat{\sigma}_{\mu,g}^2 + \widehat{\sigma}_{v,g}^2)^2} & \frac{NT}{2(T\widehat{\sigma}_{\mu,g}^2 + \widehat{\sigma}_{v,g}^2)^2} \\ \cdot & \cdot & \cdot & \cdot & \frac{N}{2(T\widehat{\sigma}_{\mu,g}^2 + \widehat{\sigma}_{v,g}^2)^2} + \frac{N(T-1)}{2(\widehat{\sigma}_{v,g}^2)^2} \end{pmatrix},$$

where  $\widehat{v}_g = \widehat{y}'_g (I_T \otimes W') \widehat{\Omega}_g^{-1} (I_T \otimes W) \widehat{y}_g$ ,  $\widehat{y}_g = X\widehat{\beta}_g$ , and  $\widehat{\sigma}_{\mu,g}^2, \widehat{\sigma}_{v,g}^2$  are restricted MLEs of  $\sigma_\mu^2, \sigma_v^2$  under  $H_0^g$ , respectively. In this case,

$$\mathcal{J}_g = \begin{pmatrix} Tb_1 & Tb_2 & 0 & 0 \\ \cdot & Tb_3 + \widehat{\omega}_g & 0 & 0 \\ \cdot & \cdot & \frac{NT^2}{2(T\widehat{\sigma}_{\mu,g}^2 + \widehat{\sigma}_{v,g}^2)^2} & \frac{NT}{2(T\widehat{\sigma}_{\mu,g}^2 + \widehat{\sigma}_{v,g}^2)^2} \\ \cdot & \cdot & \cdot & \frac{N}{2(T\widehat{\sigma}_{\mu,g}^2 + \widehat{\sigma}_{v,g}^2)^2} + \frac{N(T-1)}{2(\widehat{\sigma}_{v,g}^2)^2} \end{pmatrix},$$

where  $\widehat{\omega}_g = \widehat{y}'_g (I_T \otimes W') \left[ \widehat{\Omega}_g^{-1} - \widehat{\Omega}_g^{-1} X (X' \widehat{\Omega}_g^{-1} X)^{-1} X' \widehat{\Omega}_g^{-1} \right] (I_T \otimes W) \widehat{y}_g$ . It is easy to calculate that

$$(\mathcal{K}_g)^{-1} = \frac{1}{\widehat{\tau}_g} \begin{pmatrix} Tb_3 + \widehat{\omega}_g & -Tb_2 \\ -Tb_2 & Tb_1 \end{pmatrix}, \text{ where } \widehat{\tau}_g = T^2(b_1b_3 - b_2^2) + Tb_1\widehat{\omega}_g.$$

Finally, the LM test statistic in this case is given by

$$LM_g = \frac{Tb_3 + \hat{\omega}_g}{\hat{\tau}_g} \hat{z}_{\rho,g}^2 + \frac{Tb_1}{\hat{\tau}_g} \hat{z}_{\lambda,g}^2 - \frac{2Tb_2}{\hat{\tau}_g} \hat{z}_{\rho,g} \hat{z}_{\lambda,g}.$$

**B.6. Derivation of LM<sub>h</sub>**

The restricted MLE under  $H_0^h$  is essentially the OLS estimator. The score vector under  $H_0^h$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} |_h = \begin{pmatrix} 0 & \hat{z}_{\rho,a} & 0 \end{pmatrix}.$$

The information matrix under  $H_0^h$  and evaluated at the restricted MLE is simply

$$\mathcal{I}_h = \begin{pmatrix} \frac{X'X}{\hat{\sigma}_{v,a}^2} & 0 & 0 \\ \cdot & Tb_1 & 0 \\ \cdot & \cdot & \frac{NT}{2(\hat{\sigma}_{v,a}^2)^2} \end{pmatrix},$$

In this case, we simply have  $\mathcal{I}_{\rho\rho,h}^\Delta = \frac{1}{Tb_1}$ , and thus the LM test statistic is given by

$$LM_h = \frac{1}{Tb_1} \hat{z}_{\rho,a}^2.$$

To derive  $LM_h^*$ , we make use of the Bera and Yoon (1993) Principle that  $LM_{\psi\phi} = LM_\phi + LM_\psi^*$ , where  $LM_{\psi\phi}$  denotes the joint LM test for  $\psi, \phi$ ,  $LM_\phi$  denotes the marginal LM test for  $\phi$ , while  $LM_\psi^*$  denotes the robust LM test for  $\psi$ . Thus it can be easily deduced that

$$LM_h^* = LM_f - LM_l = \frac{Tb_3 + \hat{\omega}_a}{\hat{\tau}_a} (\hat{z}_{\rho,a} - \frac{Tb_2}{Tb_3 + \hat{\omega}_a} \hat{z}_{\lambda,a})^2.$$

**B.7. Derivation of LM<sub>i</sub>**

The score vector under  $H_0^i$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} |_i = \begin{pmatrix} 0 & \hat{z}_{\rho,i} & 0 & 0 \end{pmatrix}, \text{ where } \hat{z}_{\rho,i} = \frac{1}{\hat{\sigma}_{v,i}^2} \tilde{\epsilon}'_i (I_T \otimes M) \hat{\epsilon}_i,$$

$\hat{\sigma}_{v,i}^2 = \tilde{\epsilon}'_i \hat{\epsilon}_i / (NT)$ ,  $\hat{\epsilon}_i = \hat{B}_i y - X \hat{\beta}_i$ , and  $\hat{\beta}_i, \hat{B}_i$  are restricted MLEs of  $\beta, B$  under  $H_0^i$ , respectively. The information matrix under  $H_0^i$  and evaluated at the restricted MLE is

$$\mathcal{I}_i = \begin{pmatrix} \frac{X'X}{\hat{\sigma}_{v,i}^2} & 0 & \frac{X'(I_T \otimes W) \hat{y}_i}{\hat{\sigma}_{v,i}^2} & 0 \\ \cdot & Tb_1 & T\hat{\vartheta}_{1,i} & 0 \\ \cdot & \cdot & T\hat{\vartheta}_{2,i} + \hat{v}_i & \frac{T\hat{\vartheta}_{3,i}}{\hat{\sigma}_{v,i}^2} \\ \cdot & \cdot & \cdot & \frac{NT}{2(\hat{\sigma}_{v,i}^2)^2} \end{pmatrix},$$

where  $\hat{y}_i = \hat{B}_i^{-1} X \hat{\beta}_i$ ,  $\hat{v}_i = \hat{y}'_i (I_T \otimes (W'W)) \hat{y}_i / \hat{\sigma}_{v,i}^2$ ,  $\hat{\vartheta}_{1,i}, \hat{\vartheta}_{2,i}, \hat{\vartheta}_{3,i}$  are  $\hat{\vartheta}_1, \hat{\vartheta}_2, \hat{\vartheta}_3$  evaluated at the restricted MLE under  $H_0^i$ , respectively. Straightforward but tedious calculation gives that

$$\mathcal{I}_{\rho\rho,i}^\Delta = \hat{\xi}_i \equiv \frac{NT\hat{\vartheta}_{2,i} + N\hat{\omega}_i - 2T\hat{\vartheta}_{3,i}^2}{Tb_1(NT\hat{\vartheta}_{2,i} + N\hat{\omega}_i - 2T\hat{\vartheta}_{3,i}^2) - N(T\hat{\vartheta}_{1,i})^2},$$

where  $\hat{\omega}_i = \hat{y}'_i (I_T \otimes W') [I_{NT} - X(X'X)^{-1}X'] (I_T \otimes W) \hat{y}_i / \hat{\sigma}_{v,i}^2$ . Finally, the LM test statistic in this case is given by

$$LM_i = \hat{\xi}_i \hat{z}_{\rho,i}^2.$$

### B.8. Derivation of LM<sub>j</sub>

The score vector under  $H_0^j$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_j = \begin{pmatrix} 0 & \hat{z}_{\rho,j} & 0 & 0 \end{pmatrix}, \text{ where } \hat{z}_{\rho,j} = \hat{\epsilon}'_j \hat{\Omega}_j^{-1} (I_T \otimes M) \hat{\epsilon}_j,$$

$\hat{\epsilon}_j = y - X \hat{\beta}_j$ , and  $\hat{\beta}_j, \hat{\Omega}_j^{-1}$  are restricted MLEs of  $\beta, \Omega^{-1}$  under  $H_0^j$ , respectively. The information matrix under  $H_0^j$  and evaluated at the restricted MLE is

$$\mathcal{I}_j = \begin{pmatrix} X' \hat{\Omega}_j^{-1} X & 0 & 0 & 0 \\ \cdot & Tb_1 & 0 & 0 \\ \cdot & \cdot & \frac{NT^2}{2(T\hat{\sigma}_{\mu,j}^2 + \hat{\sigma}_{v,j}^2)^2} & \frac{NT}{2(T\hat{\sigma}_{\mu,j}^2 + \hat{\sigma}_{v,j}^2)^2} \\ \cdot & \cdot & \cdot & \frac{N}{2(T\hat{\sigma}_{\mu,j}^2 + \hat{\sigma}_{v,j}^2)^2} + \frac{N(T-1)}{2(\hat{\sigma}_{v,j}^2)^2} \end{pmatrix}.$$

Thus the LM test statistic in this case is trivially given by

$$LM_j = \frac{1}{Tb_1} \hat{z}_{\rho,j}^2.$$

To derive  $LM_j^*$ , we again make use of the Bera and Yoon (1993) [19] Principle and it can be easily deduced that

$$LM_j^* = LM_g - LM_n = \frac{Tb_3 + \hat{\omega}_j}{\hat{\tau}_j} \left( \hat{z}_{\rho,j} - \frac{Tb_2}{Tb_3 + \hat{\omega}_j} \hat{z}_{\lambda,j} \right)^2.$$

### B.9. Derivation of LM<sub>k</sub>

The score vector under  $H_0^k$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_k = \begin{pmatrix} 0 & \hat{z}_{\rho,k} & 0 & 0 & 0 \end{pmatrix}, \text{ where } \hat{z}_{\rho,k} = \hat{\epsilon}'_k \hat{\Omega}_k^{-1} (I_T \otimes M) \hat{\epsilon}_k,$$

$\hat{\epsilon}_k = \hat{B}_k y - X \hat{\beta}_k$ ,  $\hat{\beta}_k, \hat{B}_k, \hat{\Omega}_k^{-1}$  are restricted MLEs of  $\beta, B, \Omega^{-1}$  under  $H_0^k$ , respectively. The information matrix under  $H_0^k$  and evaluated at the restricted MLE is

$$\mathcal{I}_k = \begin{pmatrix} X' \hat{\Omega}_k^{-1} X & 0 & X' \hat{\Omega}_k^{-1} (I_T \otimes W) \hat{y}_k & 0 & 0 \\ \cdot & T b_1 & T \hat{\vartheta}_{1,k} & 0 & 0 \\ \cdot & \cdot & T \hat{\vartheta}_{2,k} + \hat{v}_k & \frac{T \hat{\vartheta}_{3,k}}{T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2} & \frac{T[(T-1)\hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2] \hat{\vartheta}_{3,k}}{(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2) \hat{\sigma}_{v,k}^2} \\ \cdot & \cdot & \cdot & \frac{NT^2}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} & \frac{NT}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} \\ \cdot & \cdot & \cdot & \cdot & \frac{N(T-1)}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} + \frac{N(T-1)}{2(\hat{\sigma}_{v,k}^2)^2} \end{pmatrix},$$

where  $\hat{y}_k = \hat{B}_k^{-1} X \hat{\beta}_k$ ,  $\hat{v}_k = \hat{y}'_k (I_T \otimes W') \hat{\Omega}_k^{-1} (I_T \otimes W) \hat{y}_k$ ,  $\hat{\sigma}_{\mu,k}, \hat{\sigma}_{v,k}$  are restricted MLEs of  $\sigma_\mu^2, \sigma_v^2$  under  $H_0^k$ , and  $\hat{\vartheta}_{1,k}, \hat{\vartheta}_{2,k}, \hat{\vartheta}_{3,k}$  are  $\hat{\vartheta}_1, \hat{\vartheta}_2, \hat{\vartheta}_3$  evaluated at the restricted MLE under  $H_0^k$ , respectively. Straightforward calculation gives

$$\mathcal{J}_k = \begin{pmatrix} T b_1 & T \hat{\vartheta}_{1,k} & 0 & 0 \\ \cdot & T \hat{\vartheta}_{2,k} + \hat{\omega}_k & \frac{T \hat{\vartheta}_{3,k}}{T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2} & \frac{T[(T-1)\hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2] \hat{\vartheta}_{3,k}}{(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2) \hat{\sigma}_{v,k}^2} \\ \cdot & \cdot & \frac{NT^2}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} & \frac{NT}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} \\ \cdot & \cdot & \cdot & \frac{N(T-1)}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} + \frac{N(T-1)}{2(\hat{\sigma}_{v,k}^2)^2} \end{pmatrix},$$

where  $\hat{\omega}_k = \hat{y}'_k (I_T \otimes W') [\hat{\Omega}_k^{-1} - \hat{\Omega}_k^{-1} X (X' \hat{\Omega}_k^{-1} X)^{-1} X' \hat{\Omega}_k^{-1}] (I_T \otimes W) \hat{y}_k$ . Next, we need to calculate the (1, 1)th element of  $(\mathcal{K}_k)^{-1}$ . Straightforward calculation yields

$$\mathcal{K}_k = \begin{pmatrix} T b_1 & T \hat{\vartheta}_{1,k} \\ \cdot & \hat{\varkappa}_k \end{pmatrix},$$

where

$$\hat{\varkappa}_k = T \hat{\vartheta}_{2,k} + \hat{\omega}_k - \frac{T^2 (T-1)^2 \hat{\vartheta}_{3,k}^2 \hat{\eta}_{3,k} (\hat{\sigma}_{\mu,k}^2)^2}{(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2 (\hat{\sigma}_{v,k}^2)^2} - \frac{2T^2 (T-1) \hat{\vartheta}_{3,k}^2 (\hat{\eta}_{2,k} + \hat{\eta}_{3,k}) \hat{\sigma}_{\mu,k}^2}{(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2 \hat{\sigma}_{v,k}^2} - \frac{T^2 \hat{\vartheta}_{3,k}^2 (\hat{\eta}_{1,k} + 2\hat{\eta}_{2,k} + \hat{\eta}_{3,k})}{(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2},$$

and

$$\begin{pmatrix} \hat{\eta}_{1,k} & \hat{\eta}_{2,k} \\ \hat{\eta}_{2,k} & \hat{\eta}_{3,k} \end{pmatrix} = \begin{pmatrix} \frac{NT^2}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} & \frac{NT}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} \\ \frac{NT}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} & \frac{N(T-1)}{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2} + \frac{N(T-1)}{2(\hat{\sigma}_{v,k}^2)^2} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{2(\hat{\sigma}_{v,k}^2)^2}{NT^2(T-1)} + \frac{2(T \hat{\sigma}_{\mu,k}^2 + \hat{\sigma}_{v,k}^2)^2}{NT^2} & -\frac{2(\hat{\sigma}_{v,k}^2)^2}{NT(T-1)} \\ -\frac{2(\hat{\sigma}_{v,k}^2)^2}{NT(T-1)} & \frac{2(\hat{\sigma}_{v,k}^2)^2}{N(T-1)} \end{pmatrix}.$$



Then the (1, 1)th element of  $(\mathcal{K}_k)^{-1}$  can be easily calculated as

$$\widehat{\xi}_k = \frac{\widehat{\varkappa}_k}{Tb_1\widehat{\varkappa}_k - (T\widehat{\vartheta}_{1,k})^2} = \frac{NT\widehat{\vartheta}_{2,k} + N\widehat{\omega}_k - 2T\widehat{\vartheta}_{3,k}^2}{Tb_1(NT\widehat{\vartheta}_{2,k} + N\widehat{\omega}_k - 2T\widehat{\vartheta}_{3,k}^2) - N(T\widehat{\vartheta}_{1,k})^2}.$$

Finally, the LM test statistic in this case is given by

$$LM_k = \widehat{\xi}_k \widehat{z}_{\rho,k}^2.$$

**B.10. Derivation of  $LM_l$**

The restricted MLE under  $H_0^l$  is essentially the OLS estimator. The score vector under  $H_0^l$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_l = \begin{pmatrix} 0 & \widehat{z}_{\lambda,a} & 0 \end{pmatrix}.$$

The information matrix under  $H_0^l$  and evaluated at the restricted MLE is

$$\mathcal{I}_l = \begin{pmatrix} \frac{X'X}{\widehat{\sigma}_{v,a}^2} & \frac{X'(I_T \otimes W)\widehat{y}_a}{\widehat{\sigma}_{v,a}^2} & 0 \\ \cdot & Tb_3 + \widehat{v}_a & 0 \\ \cdot & \cdot & \frac{NT}{2(\widehat{\sigma}_{v,a}^2)^2} \end{pmatrix}.$$

Straightforward calculation gives that  $\mathcal{I}_{\lambda\lambda,l}^\Delta = 1/(Tb_3 + \widehat{\omega}_a)$ , thus the LM test statistic corresponding to  $H_0^l$  is given by

$$LM_l = \frac{1}{Tb_3 + \widehat{\omega}_a} \widehat{z}_{\lambda,a}^2.$$

To derive  $LM_l^*$ , we again make use of the Bera and Yoon (1993) [19] Principle and it can be easily deduced that

$$LM_l^* = LM_f - LM_h = \frac{Tb_1}{\widehat{\tau}_a} \left( \widehat{z}_{\lambda,a} - \frac{b_2}{b_1} \widehat{z}_{\rho,a} \right)^2.$$

**B.11. Derivation of  $LM_m$**

The score vector under  $H_0^m$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_m = \begin{pmatrix} 0 & 0 & \widehat{z}_{\lambda,m} & 0 \end{pmatrix}, \text{ where } \widehat{z}_{\lambda,m} = \frac{1}{\widehat{\sigma}_{v,m}^2} \widetilde{\epsilon}'_m \widehat{A}'_m \widehat{A}_m (I_T \otimes W) \widehat{y}_m,$$

$\widehat{\sigma}_{v,m}^2 = \widetilde{\epsilon}'_m \widehat{A}'_m \widehat{A}_m \widehat{\epsilon}_m / (NT)$ ,  $\widehat{\epsilon}_m = y - \widehat{y}_m$ ,  $\widehat{y}_m = X\widehat{\beta}_m$ , and  $\widehat{\beta}_m, \widehat{A}_m$  are restricted MLEs of  $\beta, A$  under  $H_0^m$ , respectively. The information matrix under  $H_0^m$  and evaluated at the restricted MLE is

$$\mathcal{I}_m = \begin{pmatrix} \frac{X'\widehat{A}'_m \widehat{A}_m X}{\widehat{\sigma}_{v,m}^2} & 0 & \frac{X'\widehat{A}'_m \widehat{A}_m (I_T \otimes W) \widehat{y}_m}{\widehat{\sigma}_{v,m}^2} & 0 \\ \cdot & T\widehat{\theta}_{1,m} & T\widehat{\theta}_{2,m} & \frac{T\widehat{\theta}_{3,m}}{\widehat{\sigma}_{v,m}^2} \\ \cdot & \cdot & T\widehat{\theta}_{4,m} + \widehat{v}_m & 0 \\ \cdot & \cdot & \cdot & \frac{NT}{2(\widehat{\sigma}_{v,m}^2)^2} \end{pmatrix},$$

where  $\hat{v}_m = \hat{y}'_m(I_T \otimes W')\hat{A}'_m\hat{A}_m(I_T \otimes W)\hat{y}_m/\hat{\sigma}_{v,m}^2$ , and  $\hat{\theta}_{1,m}, \hat{\theta}_{2,m}, \hat{\theta}_{3,m}, \hat{\theta}_{4,m}$  are  $\hat{\theta}_1, \hat{\theta}_2, \hat{\theta}_3, \hat{\theta}_4$  evaluated at the restricted MLE under  $H_0^m$ , respectively. Let

$$\mathcal{I}_m = \begin{pmatrix} \mathcal{I}_{11,m} & \mathcal{I}_{12,m} \\ \mathcal{I}_{21,m} & \mathcal{I}_{22,m} \end{pmatrix}, \text{ where } \mathcal{I}_{11,m} = \frac{X'\hat{A}'_m\hat{A}_mX}{\hat{\sigma}_{v,m}^2}.$$

After some calculation, we get

$$\mathcal{I}_{22,m} - \mathcal{I}_{21,m}(\mathcal{I}_{11,m})^{-1}\mathcal{I}_{12,m} = \begin{pmatrix} T\hat{\theta}_{1,m} & T\hat{\theta}_{2,m} & \frac{T\hat{\theta}_{3,m}}{\hat{\sigma}_{v,m}^2} \\ \cdot & T\hat{\theta}_{4,m} + \hat{\omega}_m & 0 \\ \cdot & \cdot & \frac{NT}{2(\hat{\sigma}_{v,m}^2)^2} \end{pmatrix},$$

where  $\hat{\omega}_m = \hat{y}'_m(I_T \otimes W')\hat{A}'_m[I_{NT} - \hat{A}_mX(X'\hat{A}'_m\hat{A}_mX)^{-1}X'\hat{A}'_m]\hat{A}_m(I_T \otimes W)\hat{y}_m/\hat{\sigma}_{v,m}^2$ . Straightforward calculation gives that

$$\mathcal{I}_{\lambda\lambda,m}^\Delta = \hat{\zeta}_m \equiv \frac{N\hat{\theta}_{1,m} - 2\hat{\theta}_{3,m}^2}{(N\hat{\theta}_{1,m} - 2\hat{\theta}_{3,m}^2)(T\hat{\theta}_{4,m} + \hat{\omega}_m) - NT\hat{\theta}_{2,m}^2}.$$

Finally, the LM test statistic in this case is given by

$$LM_m = \hat{\zeta}_m \hat{z}_{\lambda,m}^2.$$

### B.12. Derivation of $LM_n$

The score vector under  $H_0^n$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_n = \begin{pmatrix} 0 & \hat{z}_{\lambda,n} & 0 & 0 \end{pmatrix}, \text{ where } \hat{z}_{\lambda,n} = \hat{\epsilon}'_n \hat{\Omega}_n^{-1} (I_T \otimes W)y,$$

$\hat{\epsilon}_n = y - X\hat{\beta}_n$ , and  $\hat{\beta}_n, \hat{\Omega}_n^{-1}$  are the restricted MLEs of  $\beta, \Omega^{-1}$  under  $H_0^n$ , respectively. The information matrix under  $H_0^n$  and evaluated at the restricted MLE is

$$\mathcal{I}_n = \begin{pmatrix} X'\hat{\Omega}_n^{-1}X & X'\hat{\Omega}_n^{-1}(I_T \otimes W)\hat{y}_n & 0 & 0 \\ \cdot & Tb_3 + \hat{v}_n & 0 & 0 \\ \cdot & \cdot & \frac{NT^2}{2(T\hat{\sigma}_{\mu,n}^2 + \hat{\sigma}_{v,n}^2)^2} & \frac{NT}{2(T\hat{\sigma}_{\mu,n}^2 + \hat{\sigma}_{v,n}^2)^2} \\ \cdot & \cdot & \cdot & \frac{N}{2(T\hat{\sigma}_{\mu,n}^2 + \hat{\sigma}_{v,n}^2)^2} + \frac{N(T-1)}{2(\hat{\sigma}_{v,n}^2)^2} \end{pmatrix},$$

where  $\hat{y}_n = X\hat{\beta}_n, \hat{v}_n = \hat{y}'_n(I_T \otimes W')\hat{\Omega}_n^{-1}(I_T \otimes W)\hat{y}_n$ . The LM test statistic in this case can be easily calculated as

$$LM_n = \frac{1}{Tb_3 + \hat{\omega}_n} \hat{z}_{\lambda,n}^2,$$

where

$$\hat{\omega}_n = \hat{y}'_n(I_T \otimes W')[\hat{\Omega}_n^{-1} - \hat{\Omega}_n^{-1}X(X'\hat{\Omega}_n^{-1}X)^{-1}X'\hat{\Omega}_n^{-1}](I_T \otimes W)\hat{y}_n.$$

To derive  $LM_n^*$ , we again make use of the Bera and Yoon (1993) [19] Principle and it can be easily deduced that

$$LM_n^* = LM_g - LM_j = \frac{Tb_1}{\hat{\tau}_n} \left( \hat{z}_{\lambda,n} - \frac{b_2}{b_1} \hat{z}_{\rho,n} \right)^2.$$

**B.13. Derivation of LM<sub>o</sub>**

The score vector under  $H_0^o$  and evaluated at the restricted MLE is

$$\frac{\partial L(\delta)}{\partial \delta} \Big|_o = \begin{pmatrix} 0 & \widehat{z}_{\lambda,o} & 0 & 0 & 0 \end{pmatrix}, \text{ where } \widehat{z}_{\lambda,o} = \widehat{\epsilon}'_o \widehat{A}'_o \widehat{\Omega}_o^{-1} \widehat{A}_o (I_T \otimes W) y,$$

$\widehat{\epsilon}_o = y - X\widehat{\beta}_o$ , and  $\widehat{\beta}_o, \widehat{A}_o, \widehat{\Omega}_o^{-1}$  are restricted MLEs of  $\beta, A, \Omega^{-1}$  under  $H_0^o$ , respectively. The information matrix under  $H_0^o$  and evaluated at the restricted MLE is

$$\mathcal{I}_o = \begin{pmatrix} X' \widehat{A}'_o \widehat{\Omega}_o^{-1} \widehat{A}_o X & 0 & X' \widehat{A}'_o \widehat{\Omega}_o^{-1} \widehat{A}_o (I_T \otimes W) \widehat{y}_o & 0 & 0 \\ \cdot & T\widehat{\theta}_{1,o} & T\widehat{\theta}_{2,o} & \frac{T\widehat{\theta}_{3,o}}{T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2} & \frac{T[(T-1)\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2]\widehat{\theta}_{3,o}}{(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)\widehat{\sigma}_{v,o}^2} \\ \cdot & \cdot & T\widehat{\theta}_{4,o} + \widehat{v}_o & 0 & 0 \\ \cdot & \cdot & \cdot & \frac{NT^2}{2(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2} & \frac{NT}{2(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2} \\ \cdot & \cdot & \cdot & \cdot & \frac{N}{2(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2} + \frac{N(T-1)}{2(\widehat{\sigma}_{v,o}^2)^2} \end{pmatrix},$$

where  $\widehat{y}_o = X\widehat{\beta}_o$ ,  $\widehat{v}_o = \widehat{y}'_o (I_T \otimes W') \widehat{A}'_o \widehat{\Omega}_o^{-1} \widehat{A}_o (I_T \otimes W) \widehat{y}_o$ ,  $\widehat{\sigma}_{\mu,o}^2, \widehat{\sigma}_{v,o}^2$  are restricted MLEs of  $\sigma_{\mu}^2, \sigma_v^2$  under  $H_0^o$ , and  $\widehat{\theta}_{1,o}, \widehat{\theta}_{2,o}, \widehat{\theta}_{3,o}, \widehat{\theta}_{4,o}$  are  $\widehat{\theta}_1, \widehat{\theta}_2, \widehat{\theta}_3, \widehat{\theta}_4$  evaluated at the restricted MLE under  $H_0^o$ , respectively.

A little calculation gives

$$\mathcal{J}_o = \begin{pmatrix} T\widehat{\theta}_{1,o} & T\widehat{\theta}_{2,o} & \frac{T\widehat{\theta}_{3,o}}{T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2} & \frac{T[(T-1)\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2]\widehat{\theta}_{3,o}}{(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)\widehat{\sigma}_{v,o}^2} \\ \cdot & T\widehat{\theta}_{4,o} + \widehat{\omega}_o & 0 & 0 \\ \cdot & \cdot & \frac{NT^2}{2(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2} & \frac{NT}{2(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2} \\ \cdot & \cdot & \cdot & \frac{N}{2(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2} + \frac{N(T-1)}{2(\widehat{\sigma}_{v,o}^2)^2} \end{pmatrix},$$

where  $\widehat{\omega}_o = \widehat{y}'_o (I_T \otimes W') \widehat{A}'_o [\widehat{\Omega}_o^{-1} - \widehat{\Omega}_o^{-1} \widehat{A}_o X (X' \widehat{A}'_o \widehat{\Omega}_o^{-1} \widehat{A}_o X)^{-1} X' \widehat{A}'_o \widehat{\Omega}_o^{-1}] \widehat{A}_o (I_T \otimes W) \widehat{y}_o$ . Next, we need to calculate the (2, 2)th element of  $(\mathcal{K}_o)^{-1}$ . Straightforward calculation yields

$$\mathcal{K}_o = \begin{pmatrix} T\widehat{\theta}_{1,o} - \frac{(T\widehat{\theta}_{3,o})^2 \widehat{\eta}_{1,o}}{(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2} - \frac{2T^2[(T-1)\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2]\widehat{\theta}_{3,o} \widehat{\eta}_{2,o}}{(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2 \widehat{\sigma}_{v,o}^2} - \frac{T^2[(T-1)\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2]\widehat{\theta}_{3,o}^2 \widehat{\eta}_{3,o}}{(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2 (\widehat{\sigma}_{v,o}^2)^2} & T\widehat{\theta}_{2,o} \\ \cdot & T\widehat{\theta}_{4,o} + \widehat{\omega}_o \end{pmatrix},$$

where

$$\begin{pmatrix} \widehat{\eta}_{1,o} & \widehat{\eta}_{2,o} \\ \widehat{\eta}_{2,o} & \widehat{\eta}_{3,o} \end{pmatrix} = \begin{pmatrix} \frac{2(\widehat{\sigma}_{v,o}^2)^2}{NT^2(T-1)} + \frac{2(T\widehat{\sigma}_{\mu,o}^2 + \widehat{\sigma}_{v,o}^2)^2}{NT^2} & -\frac{2(\widehat{\sigma}_{v,o}^2)^2}{NT(T-1)} \\ -\frac{2(\widehat{\sigma}_{v,o}^2)^2}{NT(T-1)} & \frac{2(\widehat{\sigma}_{v,o}^2)^2}{N(T-1)} \end{pmatrix}.$$

Then straightforward calculation gives the (2, 2)th element of  $(\mathcal{K}_o)^{-1}$  as

$$\widehat{\zeta}_o = \frac{NT\widehat{\theta}_{1,o} - 2T\widehat{\theta}_{3,o}^2}{(NT\widehat{\theta}_{1,o} - 2T\widehat{\theta}_{3,o}^2)(T\widehat{\theta}_{4,o} + \widehat{\omega}_o) - N(T\widehat{\theta}_{2,o})^2}.$$

Finally, the LM test statistic in this case is given by

$$LM_o = \widehat{\zeta}_o \widehat{z}_{\lambda,o}^2.$$

## References

1. Cliff, A.; Ord, J. *Spatial Autocorrelation*; Pion: London, UK, 1973.
2. Cliff, A.; Ord, J. *Spatial Processes, Models and Applications*; Pion: London, UK, 1981.
3. Anselin, L. *Spatial Econometrics: Methods and Models*; Kluwer Academic Publishers: Dordrecht, The Netherlands, 1988.
4. Anselin, L. Lagrange Multiplier Test Diagnostics for Spatial Dependence and Spatial Heterogeneity. *Geogr. Anal.* **1988**, *20*, 1–17.
5. Anselin, L.; Bera, A.K. Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics. In *Handbook of Applied Economic Statistics*; Ullah, A., Giles, D.E.A., Eds.; Marcel Dekker: New York, NY, USA, 1998.
6. Anselin, L. Rao's Score Tests in Spatial Econometrics. *J. Stat. Plan. Inference* **2001**, *97*, 113–139.
7. Baltagi, B.H.; Song, S.; Koh, W. Testing Panel Data Regression Models with Spatial Error Correlation. *J. Econom.* **2003**, *117*, 123–150.
8. Baltagi, B.H.; Liu, L. Testing for Random Effects and Spatial Lag Dependence in Panel Data Models. *Stat. Probab. Lett.* **2008**, *78*, 3304–3306.
9. Debarsy, N.; Ertur, C. Testing for Spatial Autocorrelation in a Fixed Effects Panel Data Model. *Reg. Sci. Urban Econ.* **2010**, *40*, 453–470.
10. Lee, L.F.; Yu, J. Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects. *J. Econom.* **2010**, *154*, 165–185.
11. Qu, X.; Lee, L.F. LM Tests for Spatial Correlation in Spatial Models with Limited Dependent Variables. *Reg. Sci. Urban Econ.* **2012**, *42*, 430–445.
12. Baltagi, B.H.; Egger, P.; Pfaffermayr, M. A Generalized Spatial Panel Data Model with Random Effects. *Econom. Rev.* **2013**, *32*, 650–685.
13. Kapoor, M.; Kelejian, H.H.; Prucha, I.R. Panel Data Models with Spatially Correlated Error Components. *J. Econom.* **2007**, *140*, 97–130.
14. Yang, Z.L. A Robust LM Test for Spatial Error Components. *Reg. Sci. Urban Econ.* **2010**, *40*, 299–310.
15. Baltagi, B.H.; Yang, Z.L. Standardized LM Tests for Spatial Error Dependence in Linear or Panel Regressions. *Econom. J.* **2013**, *16*, 103–134.
16. Born, B.; Breitung, J. Simple Regression-Based Tests for Spatial Dependence. *Econom. J.* **2011**, *14*, 330–342.
17. Baltagi, B.H.; Yang, Z.L. Heteroskedasticity and Non-Normality Robust LM Tests for Spatial Dependence. *Reg. Sci. Urban Econ.* **2013**, *43*, 725–739.
18. Yang, Z.L. LM Tests of Spatial Dependence Based on Bootstrap Critical Values. *J. Econom.* **2015**, *185*, 33–59.
19. Bera, A.K.; Yoon, M. Specification Testing with Locally Misspecified Alternatives. *Econom. Theory* **1993**, *9*, 649–658.

20. Anselin, L.; Bera, A.K.; Florax, R.; Yoon, M.J. Simple Diagnostic Tests for Spatial Dependence. *Reg. Sci. Urban Econ.* **1996**, *26*, 77–104.
21. Bera, A.K.; Sosa-Escudero, W.; Yoon, M. Tests for the Error Component Model in the Presence of Local Misspecification. *J. Econom.* **2001**, *101*, 1–23.
22. Bera, A.K.; Montes-Rojas, G.; Sosa-Escudero, W. Testing Under Local Misspecification and Artificial Regression. *Econ. Lett.* **2009**, *150*, 66–68.
23. Bera, A.K.; Montes-Rojas, G.; Sosa-Escudero, W. General Specification Testing with Locally Misspecified Models. *Econom. Theory* **2010**, *26*, 1838–1845.
24. He, M.; Lin, K.P. Locally Adjusted LM Test for Spatial Dependence in Fixed Effects Panel Data Models. *Econ. Lett.* **2013**, *121*, 59–63.
25. Elhorst, J.P. *Spatial Econometrics from Cross-Sectional Data to Spatial Panels*; Springer: New York, NY, USA, 2014.
26. Lee, L.F.; Yu, J. Spatial Panels: Random Components versus Fixed Effects. *Int. Econ. Rev.* **2012**, *53*, 1369–1412.
27. Magnus, J.R. Multivariate Error Components Analysis of Linear and Nonlinear Regression Models by Maximum Likelihood. *J. Econom.* **1982**, *19*, 239–285.
28. Kelejian, H.H.; Prucha, I.R. On the Asymptotic Distribution of the Moran's I Test Statistic with Applications. *J. Econom.* **2001**, *104*, 219–257.
29. Kelejian, H.H.; Prucha, I.R. Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances. *J. Econom.* **2010**, *157*, 53–67.
30. Honda, Y. Testing the Error Components Model with Non-Normal Disturbances. *Rev. Econ. Stud.* **1985**, *52*, 681–690.
31. Honda, Y. A Standardized Test for the Error Components Model with the Two-Way Layout. *Econ. Lett.* **1991**, *37*, 125–128.
32. Baltagi, B.H.; Levin, D. Cigarette Taxation: Raising Revenues and Reducing Consumption. *Struct. Chang. Econ. Dyn.* **1992**, *3*, 321–335.
33. Debarsy, N.; Ertur, C.; LeSage, J. Interpreting Dynamic Space-Time Panel Data Models. *Stat. Methodol.* **2012**, *9*, 158–171.
34. Holmes, T. The State Border Data Set. Available online: <http://www.econ.umn.edu/~holmes/data/BorderData.html> (accessed on 11 September 2015).