

Testing for Spatial Correlations with Randomly Missing Observations in the Dependent Variable

Jing Gao¹, Wei Wang²

¹College of Sciences, Shanghai Institute of Technology, Shanghai, China

²Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China

Email: gaojane@sit.edu.cn, wangwei79@sjtu.edu.cn

Received 29 June 2014; revised 25 July 2014; accepted 26 August 2014

Copyright © 2014 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

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



Open Access

Abstract

We consider LM tests for spatial correlations in the spatial error model (SEM) and spatial autoregressive model (SAM) with randomly missing data in the dependent variable. We derive the formulas of the LM test statistics and provide finite sample performance of the LM tests through Monte Carlo experiments.

Keywords

LM Test, Spatial Correlations, Missing Data, Dependent Variable

1. Introduction

Spatial models have a long history in regional science and geography (see [1], for example). Recently, many economic processes that concern spatial correlations have been drawn more and more attention. Examples include housing decision, technology adoption, tax competition, welfare participation, and price decision. Therefore, spatial correlations are of much interest in the study of urban, environmental, labor, and developmental economics among others. Various spatial econometric models are currently being applied, among which the most popular ones are the spatial error model (SEM) and spatial autoregressive model (SAM). Before setting up a spatial econometric model and doing estimation, people tend to test the existence of the spatial correlations first. The LM tests for spatial correlations have already been developed by [2] and [1] for the SEM and the SAM. However, these tests are designed for models with fully observed data.

In practice, missing data are a common problem that researchers face. When there are missing data, the spatial econometric models will be difficult to handle due to the interdependence among the components of the error

term/dependent variable vector (see [3], for example). Therefore, the LM tests proposed by [2] and [1] will be no longer valid when missing data problem occurs. In this paper, we consider a case in which observations are randomly missing only from the dependent variable and study the LM tests for spatial correlations in this situation. This situation could be very common in regional studies, where exogenous variables may be available from different sources rather than from data available on a local government web site, but the dependent variable may have missing data. LeSage and Pace [4] and [3] [5] have considered this situation and study the estimations of the spatial econometric models¹. In this study, we focus on the tests of spatial correlations in both SEM and SAM.

The rest of the paper is organized as follows. Section 2 provides the SEM model specification with missing data in the dependent variable and LM test for the spatial correlation. We derive the formula of the LM test statistic, which is asymptotically $\chi^2(1)$. In Section 3, we study the SAM model and provide the LM test. Some Monte Carlo experiments are carried out in Section 4, and Section 5 concludes the paper.

2. LM Test for Spatial Correlation in the SEM

The Spatial Error Model is:

$$\begin{aligned} Y_n &= X_n \beta_0 + \varepsilon_n \\ \varepsilon_n &= \lambda_0 W_n \varepsilon_n + u_n, \end{aligned} \quad (1)$$

where Y_n is an $n \times 1$ vector of outcomes of n cross sectional units; X_n is an $n \times k$ matrix of exogenous variables representing the n units' exogenous characters; u_n is an $n \times 1$ vector of *i.i.d.* disturbances with zero mean and a finite variance σ_0^2 ; W_n is an $n \times n$ spatial weights matrix of known constants with a zero diagonal; and λ_0 is the spatial effect coefficient that measures the spatial autocorrelation on ε_n .

If the data are fully observed, we may test

$$H_0 : \lambda_0 = 0$$

for spatial autocorrelation. Burridge [2] and [1] derived the LM test statistic as

$$LM = \frac{1}{\text{tr}(W_n^s W_n)} \left(\frac{n e^T W_n e}{e^T e} \right)^2,$$

where $W_n^s = W_n + W_n^T$ and e is the OLS residual of model (1), *i.e.*, $e = M_n Y_n$ with $M_n = I_n - X_n (X_n^T X_n)^{-1} X_n^T$. However, if there are missing observations on Y_n , the above test statistic could not be computed.

We consider the case where some of the observations in the outcome vector are unavailable. Without loss of generality, we assume that the outcomes of the last n_1 units are missing, where $0 < n_1 < n$. Therefore, we can write

$$Y_n = \begin{pmatrix} Y_n^{(o)} \\ Y_n^{(u)} \end{pmatrix},$$

where $Y_n^{(o)}$ is the $n_2 \times 1$ subvector of observed outcomes, where $n_2 = n - n_1$, and $Y_n^{(u)}$ is the remaining $n_1 \times 1$ subvector of unobserved (missing) outcomes. So the (population) system under consideration is

$$\begin{aligned} \begin{pmatrix} Y_n^{(o)} \\ Y_n^{(u)} \end{pmatrix} &= X_n \beta_0 + \varepsilon_n \\ \varepsilon_n &= \lambda_0 W_n \varepsilon_n + u_n. \end{aligned} \quad (2)$$

Note that $Y_n^{(u)} = J_n Y_n$, where $J_n = [0_{n_1 \times n_2}, I_{n_1 \times n_1}]$ is a selection matrix which picks the unobserved elements from the whole vector Y_n . Similarly, $Y_n^{(o)} = J_n^{(o)} Y_n$ where $J_n^{(o)} = [I_{n_2 \times n_2}, 0_{n_2 \times n_1}]$. To simplify some of the notations, denote $S_n = I_n - \lambda_0 W_n$. Then we can write $Y_n^{(o)}$ as

¹LeSage and Pace [4] consider an example of housing prices, where the unsold properties have known characteristics. Examples of Wang and Lee [3] [5] include censuses that provide regional demographic data, which can be aggregated to regional-level data.

$$Y_n^{(o)} = J_n^{(o)} X_n \beta_0 + J_n^{(o)} \varepsilon_n = J_n^{(o)} X_n \beta_0 + J_n^{(o)} S_n^{-1} u_n. \quad (3)$$

The maximum likelihood (ML) approach can be based on the above equation. Let $v_n = J_n^{(o)} S_n^{-1} u_n$, then $Var(v_n) = \sigma_0^2 \Sigma_{v,n}$, where $\Sigma_{v,n} = B_n B_n^T$ with $B_n = J_n^{(o)} S_n^{-1}$. Let $\theta = (\lambda, \beta^T, \sigma^2)^T$, with θ_0 being the true parameter value. Under normality, the log likelihood function is

$$\begin{aligned} \ln L_n(\theta) = & -\frac{n_2}{2} \ln(2\pi) - \frac{n_2}{2} \ln \sigma^2 - \frac{1}{2} \ln |\Sigma_{v,n}(\lambda)| \\ & - \frac{1}{2\sigma^2} [Y_n^{(o)} - J_n^{(o)} X_n \beta]^T \Sigma_{v,n}^{-1}(\lambda) [Y_n^{(o)} - J_n^{(o)} X_n \beta], \end{aligned}$$

where $\Sigma_{v,n} = B_n(\lambda) B_n^T(\lambda)$ with $B_n(\lambda) = J_n^{(o)} (I_n - \lambda W_n)^{-1}$. The expressions for the elements of the score vector are:

$$\begin{aligned} \frac{\partial \ln L_n(\theta)}{\partial \lambda} = & -\frac{1}{2} \frac{\partial \ln |\Sigma_{v,n}(\lambda)|}{\partial \lambda} - \frac{1}{2\sigma^2} [Y_n^{(o)} - J_n^{(o)} X_n \beta]^T \frac{\partial \Sigma_{v,n}^{-1}(\lambda)}{\partial \lambda} [Y_n^{(o)} - J_n^{(o)} X_n \beta] \\ = & \frac{1}{2\sigma^2} \left\{ v_n^T(\beta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\beta) \right. \\ & \left. - \sigma^2 \text{tr} \left(B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] \right) \right\}, \end{aligned} \quad (4)$$

$$\frac{\partial \ln L_n(\theta)}{\partial \sigma^2} = \frac{1}{2\sigma^4} \left\{ v_n^T(\beta) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\beta) - n_2 \sigma^2 \right\}, \quad (5)$$

and

$$\frac{\partial \ln L_n(\theta)}{\partial \beta^T} = \frac{1}{\sigma^2} v_n^T(\beta) [B_n(\lambda) B_n^T(\lambda)]^{-1} J_n^{(o)} X_n, \quad (6)$$

where $v_n(\beta) = Y_n^{(o)} - J_n^{(o)} X_n \beta$, and $G_n(\lambda) = W_n (I_n - \lambda W_n)^{-1}$. The second order derivatives are, for the relevant combinations of parameters, Equations (A.1)-(A.6) in the Appendix. Thus the elements of the information matrix are,

$$\begin{aligned} I_{\lambda\lambda} = & \text{tr} \left\{ B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} \right\} \\ & - \text{tr} \left\{ B_n(\lambda) G_n(\lambda) G_n^T(\lambda) B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} \right\} + \text{tr} \left\{ [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) G_n^T(\lambda) B_n^T(\lambda) \right\} \\ & - \text{tr} \left\{ [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) B_n^T(\lambda) \right\}, \end{aligned} \quad (7)$$

$$I_{\lambda\sigma^2} = \frac{1}{\sigma^2} \text{tr} \left\{ B_n(\lambda) G_n(\lambda) B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} \right\}, \quad I_{\lambda\beta} = 0, \quad I_{\sigma^2\sigma^2} = \frac{n_2}{2\sigma^4}, \quad I_{\sigma^2\beta} = 0, \quad (8)$$

$$I_{\beta\beta} = \frac{1}{\sigma^2} X_n^T J_n^{(o)T} [B_n(\lambda) B_n^T(\lambda)]^{-1} J_n^{(o)} X_n. \quad (9)$$

To perform the LM test, expressions (4)-(6) and (7)-(9) need to be evaluated under constrained estimation, *i.e.*, with the parameter values included in the null hypothesis set to zero (namely, $\hat{\lambda}_R = 0$), and with the other parameters set to their ordinary-least-squares estimates, *i.e.*, $\hat{\beta}_R = (X_n^T J_n^{(o)T} J_n^{(o)} X_n)^{-1} X_n^T J_n^{(o)T} Y_n^{(o)}$, and $\hat{\sigma}_R^2 = (1/n_2) Y_n^{(o)T} M_n^{(o)} Y_n^{(o)}$, with $M_n^{(o)} = I_{n_2} - J_n^{(o)} X_n (X_n^T J_n^{(o)T} J_n^{(o)} X_n)^{-1} X_n^T J_n^{(o)T}$.

Note that $B_n(0) = J_n^{(o)}$, $G_n(0) = W_n$ and

$$\text{tr}\left\{B_n^T(0)\left[B_n(0)B_n^T(0)\right]^{-1}B_n(0)\left[G_n(0)+G_n^T(0)\right]\right\}=\text{tr}\left\{J_n^{(o)T}J_n^{(o)}\left[W_n+W_n^T\right]\right\}=0$$

because W_n has a zero diagonal. Thus we have the score vector as follows

$$\frac{\partial \ln L_n(\hat{\theta}_R)}{\partial \theta} = \begin{bmatrix} e^T J_n^{(o)} (W_n + W_n^T) J_n^{(o)T} e / 2\hat{\sigma}_R^2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} n_2 e^T J_n^{(o)} W_n J_n^{(o)T} e / e^T e \\ 0 \\ 0 \end{bmatrix},$$

where $e = M_n^{(o)} Y_n^{(o)}$ is the ordinary-least-squares residual of model (3). And the estimated information matrix is

$$I(\hat{\theta}_R) = \begin{bmatrix} \text{tr}\left[J_n^{(o)T} J_n^{(o)} (W_n + W_n^T) J_n^{(o)T} J_n^{(o)} W_n\right] & 0 & 0 \\ 0 & \frac{n_2}{2\hat{\sigma}_R^4} & 0 \\ 0 & 0 & \frac{1}{\hat{\sigma}_R^2} X_n^T J_n^{(o)T} J_n^{(o)} X_n \end{bmatrix}.$$

Therefore, the LM test statistic is

$$LM = \left[\frac{\partial \ln L_n(\hat{\theta}_R)}{\partial \hat{\theta}_R} \right]^T \left[I(\hat{\theta}_R) \right]^{-1} \left[\frac{\partial \ln L_n(\hat{\theta}_R)}{\partial \hat{\theta}_R} \right] = \frac{1}{\text{tr}\left[J_n^{(o)T} J_n^{(o)} W_n^s J_n^{(o)T} J_n^{(o)} W_n\right]} \left(\frac{n_2 e^T J_n^{(o)} W_n J_n^{(o)T} e}{e^T e} \right)^2,$$

with $e = M_n^{(o)} Y_n^{(o)} = \left(I_{n_2} - J_n^{(o)} X_n \left(X_n^T J_n^{(o)T} J_n^{(o)} X_n \right)^{-1} X_n^T J_n^{(o)T} \right) Y_n^{(o)}$ and $W_n^s = W_n + W_n^T$.

Under the null, we have $LM \rightarrow \chi^2(1)$.

3. LM Test for Spatial Correlation in the SAM

The Spatial Autoregressive Model is:

$$Y_n = \lambda_0 W_n Y_n X_n \beta_0 + \varepsilon_n \quad (10)$$

where all the notations have same meanings as those in the previous section, except that ε_n now is an $n \times 1$ vector of *i.i.d.* disturbances with zero mean and a finite variance σ_0^2 . We can see that in this model, the spatial correlations exist among the components of Y_n instead of ε_n , compared with SEM.

We consider testing the spatial lag dependence of the model, namely testing the null hypothesis

$$H_0: \lambda_0 = 0.$$

If the data are fully observed, by using the likelihood function [1], derived the LM test statistic explicitly as

$$LM = \frac{1}{T} \left(\frac{ne^T W_n y}{e^T e} \right)^2,$$

where e is the OLS residual of model (10) under the null, and

$$T = \left\{ n \left[W_n X_n \left(X_n^T X_n \right)^{-1} X_n^T Y_n \right]^T M_n W_n X_n \left(X_n^T X_n \right)^{-1} X_n^T Y_n + \text{tr}\left[\left(W_n + W_n^T\right)W_n\right] e^T e \right\} / e^T e.$$

We consider the case where some of the observations in the outcome vector are unavailable. By adopting the same notations as those in the previous section, the (population) system under consideration can be written as

$$\begin{pmatrix} Y_n^{(o)} \\ Y_n^{(u)} \end{pmatrix} = \lambda_0 W \begin{pmatrix} Y_n^{(o)} \\ Y_n^{(u)} \end{pmatrix} + X_n \beta_0 + \varepsilon_n. \quad (11)$$

The reduced form Equation of (11) for write $Y_n^{(o)}$ is

$$Y_n^{(o)} = J_n^{(o)} S_n^{-1} X_n \beta_0 + J_n^{(o)} S_n^{-1} \varepsilon_n, \quad (12)$$

and therefore, the ML approach based on this reduced form equation can be applied. Under normality, the log likelihood function is

$$\begin{aligned} \ln L_n(\theta) = & -\frac{n_2}{2} \ln(2\pi) - \frac{n_2}{2} \ln \sigma^2 - \frac{1}{2} \ln |\Sigma_{v,n}(\lambda)| \\ & - \frac{1}{2\sigma^2} [Y_n^{(o)} - B_n(\lambda) X_n \beta]^T \Sigma_{v,n}^{-1}(\lambda) [Y_n^{(o)} - B_n(\lambda) X_n \beta]. \end{aligned}$$

The expressions for the elements of the score vector are

$$\begin{aligned} \frac{\partial \ln L_n(\theta)}{\partial \lambda} = & \frac{1}{2\sigma^2} \left\{ v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta) \right. \\ & \left. - \sigma^2 \text{tr} \left(B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] \right) \right\} \\ & + \frac{1}{\sigma^2} [B_n(\lambda) G_n(\lambda) X_n \beta]^T [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta), \end{aligned} \quad (13)$$

$$\frac{\partial \ln L_n(\theta)}{\partial \sigma^2} = \frac{1}{2\sigma^4} \left\{ v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta) - n_2 \sigma^2 \right\}, \quad (14)$$

and

$$\frac{\partial \ln L_n(\theta)}{\partial \beta^T} = \frac{1}{\sigma^2} v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) X_n, \quad (15)$$

where $v_n(\theta) = Y_n^{(o)} - B_n(\lambda) X_n \beta$. The second order derivatives are, for the relevant combinations of parameters, Equations (A.7)-(A.12) in the Appendix. Thus the elements of the information matrix are,

$$\begin{aligned} I_{\lambda\lambda} = & \text{tr} \left\{ B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} \right\} \\ & - \text{tr} \left\{ B_n(\lambda) G_n(\lambda) G_n^T(\lambda) B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} \right\} + \text{tr} \left\{ [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) G_n^T(\lambda) B_n^T(\lambda) \right\} \\ & - \text{tr} \left\{ [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) B_n^T(\lambda) \right\} \\ & + \frac{1}{\sigma^2} [B_n(\lambda) G_n(\lambda) X_n \beta]^T [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) X_n \beta, \end{aligned} \quad (16)$$

$$I_{\lambda\sigma^2} = \frac{1}{\sigma^2} \text{tr} \left\{ B_n(\lambda) G_n(\lambda) B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} \right\}, \quad (17)$$

$$I_{\lambda\beta} = \frac{1}{\sigma^2} [B_n(\lambda) G_n(\lambda) X_n \beta]^T [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) X_n \beta, \quad (18)$$

$$I_{\sigma^2\sigma^2} = \frac{n_2}{2\sigma^4}, \quad I_{\sigma^2\beta} = 0, \quad (19)$$

$$I_{\beta\beta} = \frac{1}{\sigma^2} X_n^T B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) X_n. \quad (20)$$

To perform the LM test, expressions (13)-(15) and (16)-(20) need to be evaluated under constrained estimation, *i.e.*, with the parameter values included in the null hypothesis set to zero (namely, $\hat{\lambda}_R = 0$), and with the other parameters set to their ordinary-least-squares estimates, *i.e.*, $\hat{\beta}_R = (X_n^T J_n^{(o)T} J_n^{(o)} X_n)^{-1} X_n^T J_n^{(o)T} Y_n^{(o)}$, and $\hat{\sigma}_R^2 = (1/n_2) Y_n^{(o)T} M_n^{(o)} Y_n^{(o)}$, with $M_n^{(o)} = I_{n_2} - J_n^{(o)} X_n (X_n^T J_n^{(o)T} J_n^{(o)} X_n)^{-1} X_n^T J_n^{(o)T}$.

Note that $B_n(0) = J_n^{(o)}$, $G_n(0) = W_n$ and

$$\text{tr}\left\{B_n^T(0)\left[B_n(0)B_n^T(0)\right]^{-1}B_n(0)\left[G_n(0)+G_n^T(0)\right]\right\} = \text{tr}\left\{J_n^{(o)T}J_n^{(o)}\left[W_n+W_n^T\right]\right\} = 0$$

because W_n has a zero diagonal. Thus we have the score vector as follows

$$\frac{\partial \ln L_n(\hat{\theta}_R)}{\partial \theta} = \begin{bmatrix} n_2 e^T J_n^{(o)} W_n \left[X_n \hat{\beta}_R + J_n^{(o)T} e \right] / e^T e \\ 0 \\ 0 \end{bmatrix},$$

where $e = M_n^{(o)} Y_n^{(o)}$ is the ordinary-least-squares residual of model (12) with $\hat{\lambda}_R = 0$. And the estimated information matrix is

$$I(\hat{\theta}_R) = \begin{bmatrix} \text{tr}\left[J_n^{(o)T}J_n^{(o)}(W_n+W_n^T)J_n^{(o)T}J_n^{(o)}W_n\right] + \frac{1}{\hat{\sigma}_R^2}\left[J_n^{(o)}W_nX_n\hat{\beta}_R\right]^T J_n^{(o)}W_nX_n\hat{\beta}_R & 0 & \frac{1}{\hat{\sigma}_R^2}\left[J_n^{(o)}W_nX_n\hat{\beta}_R\right]^T J_n^{(o)}X_n \\ 0 & \frac{n_2}{2\hat{\sigma}_R^4} & 0 \\ \frac{1}{\hat{\sigma}_R^2}\left[J_n^{(o)}X_n\right]^T J_n^{(o)}W_nX_n\hat{\beta}_R & 0 & \frac{1}{\hat{\sigma}_R^2}X_n^T J_n^{(o)T}J_n^{(o)}X_n \end{bmatrix}.$$

Using the formula of the inverse of a partitioned matrix, we have

$$\left[I(\hat{\theta}_R)\right]_{\lambda\lambda}^{-1} = \text{tr}\left[J_n^{(o)T}J_n^{(o)}(W_n+W_n^T)J_n^{(o)T}J_n^{(o)}W_n\right] + \frac{1}{\hat{\sigma}_R^2}\left[J_n^{(o)}W_nX_n\hat{\beta}_R\right]^T M_n^{(o)}J_n^{(o)}W_nX_n\hat{\beta}_R.$$

Therefore, the LM test statistic is

$$LM = \left[I(\hat{\theta}_R)\right]_{\lambda\lambda}^{-1} \left[\frac{\partial \ln L_n(\hat{\theta}_R)}{\partial \lambda} \right]^2 = \frac{1}{T^{(o)}} \left(\frac{n_2 e^T J_n^{(o)} W_n \left[X_n \left(X_n^T J_n^{(o)T} J_n^{(o)} X_n \right)^{-1} X_n^T J_n^{(o)T} Y_n^{(o)} + J_n^{(o)T} e \right]}{e^T e} \right)^2,$$

where

$$T^{(o)} = \left\{ n_2 \left[J_n^{(o)} W_n X_n \left(X_n^T J_n^{(o)T} J_n^{(o)} X_n \right)^{-1} X_n^T J_n^{(o)T} Y_n^{(o)} \right]^T M_n^{(o)} J_n^{(o)} W_n X_n \left(X_n^T J_n^{(o)T} J_n^{(o)} X_n \right)^{-1} X_n^T J_n^{(o)T} Y_n^{(o)} \right. \\ \left. + \text{tr}\left[J_n^{(o)T}J_n^{(o)}W_n^s J_n^{(o)T}J_n^{(o)}W_n\right] e^T e \right\} / e^T e,$$

with $e = M_n^{(o)} Y_n^{(o)} = \left(I_{n_2} - J_n^{(o)} X_n \left(X_n^T J_n^{(o)T} J_n^{(o)} X_n \right)^{-1} X_n^T J_n^{(o)T} \right) Y_n^{(o)}$ and $W_n^s = W_n + W_n^T$.

Under the null, we have $LM \rightarrow \chi^2(1)$.

4. Monte Carlo Experiments

To investigate the finite sample performance of the LM tests, we conduct Monte Carlo experiments, designed as follows.

4.1. LM Tests in SEM

The model has two regressors x_{i1} and x_{i2} . The (x_{i1}, x_{i2}) 's are independent for all i . The true slope parameters

are $\beta_{10} = 1$ and $\beta_{20} = 1$. The u_i 's are independently drawn from $N(0,1)$ and are independent of x_{i1} and x_{i2} . x_{i1} and x_{i2} are generated from $N(0,1)$.

For weights matrix W_n , we follow the design of [6] and [3], which is referred to as the “circular world matrix.” The weights matrix is designed as follows. For the $n \times n$ weights matrix, the first $n/3$ rows (except for the first row) have zeroes everywhere, except for the elements in positions $(i, i+1)$ and $(i, i-1)$. In the first row, the non-zero elements are in positions $(1, 2)$ and $(1, n)$ so that it relates to a circular world. The nonzero elements in the first $n/3$ rows are all random draws from $U(0,1)$; i.e., we allow the neighbors to asymmetrically affect one another². Then, these rows are row normalized, so that the sum of each row is equal to 1. The next $n/3$ rows (say, $j = n/3+1, \dots, 2n/3$) have zeroes everywhere, except in positions $(j, j \pm r)$, where $r = 1, 2, \dots, 5$. The nonzero elements are designed in the same fashion as those in the first $n/3$ rows. The last $n/3$ rows are defined in a similar manner to the first $n/3$ rows. Specifically, the nonzero elements in rows $j = 2n/3+1, \dots, n-1$ are in positions $(j, j+1)$ and $(j, j-1)$; in the last row, the nonzero elements are in positions $(n, 1)$ and $(n, n-1)$. The nonzero elements in these rows are also designed in the same fashion as those in the first $2n/3$ rows. The weights matrix is a sparse weights matrix, with each individual having only several “neighbors.” The number of neighbors differs for each individual, depending on its position.

For sample sizes, we set n from “small”, $n = 60$ and “moderate”, $n = 180$, to “large”, $n = 540$. For missing observations, the y_i 's of the first α percent of the n individuals are unobserved for each sample size n , where α is set as 10, 25, and 50.

For each n and α (percentage of missing) combination, we report the percentages of rejecting the null hypothesis in all the 1000 Monte Carlo replications, for different nominal sizes 1%, 5% and 10%. The first row shows the results for the $\lambda_0 = 0$ case, and the second and third row show those for $\lambda_0 = 0.2$ and 0.5, respectively.

Tables 1-3 below show the finite performance of the LM test in the SEM. The empirical levels (first row) of the LM test are close to the theoretical ones. But for the powers (second and third row), they depend on the sample sizes and the value of λ_0 . For small value of λ_0 , the powers are poor, especially for small n . For larger λ_0 , the powers are good, except for small n .

Table 1. SEM: 10% missing data.

n	60			180			540		
Nominal sizes	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\lambda_0 = 0$	0.6	3.7	8.1	1.3	4.9	9.5	1.2	4.6	10.1
$\lambda_0 = 0.2$	9.8	24.1	33.9	37.5	61.7	73.0	91.5	96.5	97.9
$\lambda_0 = 0.5$	78.7	91.4	95.9	99.9	100	100	100	100	100

Table 2. SEM: 25% missing data.

n	60			180			540		
Nominal sizes	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\lambda_0 = 0$	0.8	5.2	9.4	1.6	4.8	10.1	0.8	5.2	9.9
$\lambda_0 = 0.2$	6.4	18.9	28.0	28.9	49.7	62.0	81.6	94.3	97.4
$\lambda_0 = 0.5$	64.1	82.1	87.4	99.9	99.9	99.9	100	100	100

Table 3. SEM: 50% missing data.

n	60			180			540		
Nominal sizes	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\lambda_0 = 0$	1.3	4.5	8.3	1.3	6.4	10.9	0.8	5.0	9.9
$\lambda_0 = 0.2$	4.3	14.1	21.5	19.2	39.9	54.2	66.6	84.2	90.8
$\lambda_0 = 0.5$	45.2	68.3	76.8	97.1	99.1	99.6	100	100	100

²Wang and Lee [3] generate the symmetric settings in [6] to allow for asymmetry.

Table 4. SAM: 10% missing data.

n	60			180			540		
Nominal sizes	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\lambda_0 = 0$	1.1	4.9	10.3	0.8	5.0	10.0	0.5	5.5	11.4
$\lambda_0 = 0.2$	26.8	47.7	58.7	77.6	91.3	95.3	100	100	100
$\lambda_0 = 0.5$	98.2	99.5	99.7	100	100	100	100	100	100

Table 5. SAM: 25% missing data.

n	60			180			540		
Nominal sizes	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\lambda_0 = 0$	1.8	5.7	10.5	1.6	5.5	10.1	0.6	4.5	10.3
$\lambda_0 = 0.2$	16.7	33.9	47.5	63.9	81.2	87.3	99.3	99.9	99.9
$\lambda_0 = 0.5$	93.9	97.9	98.9	100	100	100	100	100	100

Table 6. SAM: 50% missing data.

n	60			180			540		
Nominal sizes	1%	5%	10%	1%	5%	10%	1%	5%	10%
$\lambda_0 = 0$	1.1	5.8	10.7	0.9	4.5	9.0	1.1	4.7	10.6
$\lambda_0 = 0.2$	12.1	25.9	38.8	47.7	69.5	79.1	95.1	98.9	99.5
$\lambda_0 = 0.5$	78.9	91.1	95.6	100	100	100	100	100	100

4.2. LM Tests in SAM

In the SAM, we generate ε_i 's independently from $N(0,1)$ and independent of x_i 's. All other designs are the same as those in the previous subsection. **Tables 4-6** show the results of the LM test. The empirical levels are all close to the theoretical ones. But for the powers, they are not good for small value of λ_0 when sample sizes are small. For large value of λ_0 and larger sample sizes, the powers are good.

5. Conclusion

In this paper, we extend the LM tests for spatial correlations to the case where there are missing data in the dependent variable. We considered the spatial error model as well as the spatial autoregressive model and derived the formulas of the LM test statistics in both models. Monte Carlo experiments show good finite sample performance of the tests. The empirical levels of the LM tests are close to the theoretical ones and the powers are good for large sample sizes.

References

- [1] Anselin, L. (1988) Spatial Econometrics: Methods and Models. Kluwer, Dordrecht. <http://dx.doi.org/10.1007/978-94-015-7799-1>
- [2] Burridge, P. (1980) On the Cliff-Ord Test for Spatial Autocorrelation. *Journal of the Royal Statistical Society*, **42**, 107-108.
- [3] Wang, W. and Lee, L. (2013a) Estimation of Spatial Autoregressive Models with Randomly Missing Data in the Dependent Variable. *Econometrics Journal*, **16**, 73-102. <http://dx.doi.org/10.1111/j.1368-423X.2012.00388.x>
- [4] LeSage, J. and Pace, R.K. (2004) Models for Spatially Dependent Missing Data. *Journal of Real Estate Finance and Economics*, **29**, 233-254. <http://dx.doi.org/10.1023/B:REAL.0000035312.82241.e4>
- [5] Wang, W. and Lee, L. (2013) Estimation of Spatial Panel Data Models with Randomly Missing Data in the Dependent Variable. *Regional Science and Urban Economics*, **43**, 521-538. <http://dx.doi.org/10.1016/j.regsciurbeco.2013.02.001>

- [6] Arraiz, I., Drukker, D., Kelejian, H. and Prucha, I. (2008) A Spatial Cliff-Ord-Type Model with Heteroskedastic Innovations: Small and Large Sample Results. CESIFO Working Paper No. 2485.

Appendix

The second order derivatives are, for the relevant combinations of parameters of the log likelihood function for the SEM,

$$\begin{aligned}
 -\frac{\partial^2 \ln L_n(\theta)}{\partial \lambda^2} &= \frac{1}{\sigma^2} \left\{ v_n^T(\beta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] \right. \\
 &\quad \times B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\beta) \\
 &\quad \left. - v_n^T(\beta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) G_n^T(\lambda) B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\beta) \right\} \\
 &\quad - \frac{1}{2} \text{tr} \left\{ [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} \right. \\
 &\quad \left. \times B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) \right\} + \text{tr} \left\{ [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) G_n^T(\lambda) B_n^T(\lambda) \right\},
 \end{aligned} \quad (\text{A.1})$$

$$-\frac{\partial^2 \ln L_n(\theta)}{\partial \lambda \partial \sigma^2} = \frac{1}{2\sigma^4} v_n^T(\beta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\beta), \quad (\text{A.2})$$

$$-\frac{\partial^2 \ln L_n(\theta)}{\partial \lambda \partial \beta^T} = \frac{1}{\sigma^2} v_n^T(\beta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} J_n^{(o)} X_n, \quad (\text{A.3})$$

$$-\frac{\partial^2 \ln L_n(\theta)}{\partial (\sigma^2)^2} = \frac{1}{\sigma^6} v_n^T(\beta) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\beta) - \frac{n_2}{2\sigma^4}, \quad (\text{A.4})$$

$$-\frac{\partial^2 \ln L_n(\theta)}{\partial \sigma^2 \partial \beta^T} = \frac{1}{\sigma^4} v_n^T(\beta) [B_n(\lambda) B_n^T(\lambda)]^{-1} J_n^{(o)} X_n, \quad (\text{A.5})$$

and

$$-\frac{\partial^2 \ln L_n(\theta)}{\partial \beta \partial \beta^T} = \frac{1}{\sigma^2} X_n^T J_n^{(o)T} [B_n(\lambda) B_n^T(\lambda)]^{-1} J_n^{(o)} X_n. \quad (\text{A.6})$$

The second order derivatives are, for the relevant combinations of parameters of the log likelihood function for the SAM,

$$\begin{aligned}
 -\frac{\partial^2 \ln L_n(\theta)}{\partial \lambda^2} &= \frac{1}{\sigma^2} \left\{ v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} \right. \\
 &\quad \times B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta) - v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} \\
 &\quad \times B_n(\lambda) G_n(\lambda) G_n^T(\lambda) B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta) \left. \right\} + \frac{1}{\sigma^2} v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} \\
 &\quad \times B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) X_n \beta \\
 &\quad - \frac{1}{2} \text{tr} \left\{ [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} \right. \\
 &\quad \times B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) \left. \right\} + \text{tr} \left\{ [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) G_n^T(\lambda) B_n^T(\lambda) \right\} \\
 &\quad + \frac{1}{\sigma^2} \left\{ [B_n(\lambda) G_n(\lambda) X_n \beta]^T [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) G_n(\lambda) X_n \beta \right. \\
 &\quad \left. - ([B_n(\lambda) - G_n(\lambda)] G_n(\lambda) X_n \beta)^T [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta) + [B_n(\lambda) G_n(\lambda) X_n \beta]^T \right. \\
 &\quad \left. \times [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta) \right\},
 \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned}
-\frac{\partial^2 \ln L_n(\theta)}{\partial \lambda \partial \sigma^2} &= \frac{1}{2\sigma^4} v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta) \\
&\quad + \frac{1}{\sigma^4} [B_n(\lambda) G_n(\lambda) X_n]^T [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta),
\end{aligned} \tag{A.8}$$

$$\begin{aligned}
-\frac{\partial^2 \ln L_n(\theta)}{\partial \lambda \partial \beta^T} &= \frac{1}{\sigma^2} v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) [G_n(\lambda) + G_n^T(\lambda)] B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) X_n \\
&\quad - \frac{1}{\sigma^2} [B_n(\lambda) G_n(\lambda) X_n]^T [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta) + \frac{1}{\sigma^2} [B_n(\lambda) G_n(\lambda) X_n]^T [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) X_n,
\end{aligned} \tag{A.9}$$

$$-\frac{\partial^2 \ln L_n(\theta)}{\partial (\sigma^2)^2} = \frac{1}{\sigma^6} v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} v_n(\theta) - \frac{n_2}{2\sigma^4}, \tag{A.10}$$

$$-\frac{\partial^2 \ln L_n(\theta)}{\partial \sigma^2 \partial \beta^T} = \frac{1}{\sigma^4} v_n^T(\theta) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) X_n, \tag{A.11}$$

and

$$-\frac{\partial^2 \ln L_n(\theta)}{\partial \beta \partial \beta^T} = \frac{1}{\sigma^2} X_n^T B_n^T(\lambda) [B_n(\lambda) B_n^T(\lambda)]^{-1} B_n(\lambda) X_n. \tag{A.12}$$

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

Other selected journals from SCIRP are listed as below. Submit your manuscript to us via either submit@scirp.org or [Online Submission Portal](#).

