SMU ECONOMICS & STATISTICS
WORKING PAPER SERIES

A Robust LM Test for Spatial Error Components

Zhenlin Yang
January 2009

Paper No. 04-2009

A Robust LM Test for Spatial Error Components[1]

Zhenlin Yang

School of Economics, Singapore Management University

email: zlyang@smu.edu.sg

January, 2009

Abstract

This paper presents a modified LM test of spatial error components, which is shown to be robust against distributional misspecifications and spatial layouts. The proposed test differs from the LM test of Anselin (2001) by a term in the denominators of the test statistics. This term disappears when either the errors are normal, or the variance of the diagonal elements of the product of spatial weights matrix and its transpose is zero or approaching to zero as sample size goes large. When neither is true, as is often the case in practice, the effect of this term can be significant even when sample size is large. As a result, there can be severe size distortions of the Anselin's LM test, a phenomenon revealed by the Monte Carlo results of Anselin and Moreno (2003) and further confirmed by the Monte Carlo results presented in this paper. Our Monte Carlo results also show that the proposed test performs well in general.

Key Words: Distributional misspecification; Robustness, Spatial layouts; Spatial error components; LM tests.

JEL Classification: C23, C5

---

[1] I would like to thank Ingmar Prucha, Steven Farber, and the conference participants of II World Conference of the Spatial Econometrics Association, Nov. 17-19, 2008, New York City, for the helpful comments. I am grateful to the support from a research grant (Grant number: C208/MS63E046) from Singapore Management University.

1 Introduction.

The spatial error components model proposed by Kelejian and Robinson (1995) provides a useful alternative to the traditional spatial models with a spatial autoregressive (SAR) error process or a spatial moving average (SMA) error process, in particular in the situation where the range of spatial autocorrelation is constrained to close neighbors, e.g., spatial spillovers in the productivity of infrastructure investments (Kelejian and Robinson, 1997; Anselin and Moreno, 2003). Anselin (2001) derived an LM test for spatial error components based on the assumptions that the errors are normally distributed. Anselin and Moreno (2003) conducted Monte Carlo experiments to assess the finite sample behavior of Anselin's test and to compare it with other tests such as the GMM test of Kelejian and Robinson (1995) and Moran's (1950) I test, and found that none seems to perform satisfactorily in general. While Anselin and Moreno (2003) recognized that the LM test for spatial error components of Anselin (2001) is sensitive to distributional misspecifications and the spatial layouts, it is generally unclear on the exact cause of it and how this normal-theory based test performs under alternative distributions for the errors and under different spatial layouts.

In this paper, we present a modified LM test of spatial error components, which is shown to be robust against distributional misspecifications and spatial layouts. We show that the proposed test differs from the LM test of Anselin (2001) by a term in the denominators of the test statistics. This term disappears when either the errors are normal, or the variance of the diagonal elements of the product of spatial weights matrix and its transpose is zero or approaching zero as sample size goes large. When neither is true, as is often the case in practice, we show that (i) if the elements of the weights matrix are fixed, this term poses a large sample effect in the sense that without this term Anselin's LM test does not converge to a correct level as sample size goes large; and (ii) if the elements of the weights matrix depend on sample size, this term poses a significant finite sample effect in the sense that without this term Anselin's LM test can have a large size distortion which gets smaller very slowly as sample size gets large.

Anselin and Bera (1998), Anselin (2001) and Florax and de Graaff (2004) provide excellent reviews on tests of spatial dependence in linear models. Section 2 introduces the spatial error components model and the describes the existing test. Section 3 introduces a robust

LM test for spatial error components. Section 4 presents Monte Carlo results. Section 5 concludes the paper.

2 The Spatial Error Components Model

The spatial error components (SEC) model proposed by Kelejian and Robinson (1995) takes the following form:

$$Y_n = X_n\beta + u_n \quad \text{with} \;\; u_n = W_n\nu_n + \varepsilon_n \tag{1}$$

where Y_n is an $n \times 1$ vector of observations on the response variable, X_n is an $n \times k$ matrix containing the values of explanatory (exogenous) variables, β is a $k \times 1$ vector of regression coefficients, W_n is an $n \times n$ spatial weights matrix, ν_n is an $n \times 1$ vector of errors that together with W_n incorporates the spatial dependence, and ε is an $n \times 1$ vector of location specific disturbance terms. The error components ν_n and ε_n are assumed to be independent, with independent and identically distributed (iid) elements of mean zero and variances σ_ν^2 and σ_ε^2, respectively. So, in this model the null hypothesis of no spatial effect can be either $H_0 : \sigma_\nu^2 = 0$, or $\theta = \sigma_\nu^2/\sigma_\varepsilon^2 = 0$. The alternative hypothesis can only be one-sided as σ_ν^2 is non-negative, i.e., $H_a : \sigma_\nu^2 > 0$, or $\theta > 0$. Anselin (2001) derived an LM test based on the assumptions that errors are normally distributed. The test is of the form

$$\text{LM}_{\text{SEC}} = \frac{\tilde{u}_n' W_n W_n' \tilde{u}_n / \tilde{\sigma}_\varepsilon^2 - T_{1n}}{(2T_{2n} - \frac{2}{N}T_{1n}^2)^{\frac{1}{2}}} \tag{2}$$

where $\tilde{\sigma}_\varepsilon^2 = \frac{1}{n}\tilde{u}_n'\tilde{u}_n$, $\tilde{u}_n$ is the vector of OLS residuals, $T_{1n} = \text{tr}(W_n W_n')$ and $T_{2n} = \text{tr}(W_n W_n' W_n W_n')$. Under H_0, the positive part of LM_{SEC} converges to that of $N(0,1)$. This means that the above one sided test can be carried out as per normal. Alternatively, if the squared version LM_{SEC}^2 is used, the reference null distribution of the test statistic for testing this one sided test is a chi-square mixture. See Verbeke and Molenberghs (2003) for a detailed discussion on tests where the parameter value under the null hypothesis falls on the boundary of parameter space.

Anselin and Moreno (2003) provide Monte Carlo evidence for the finite sample performance of LM_{SEC} and find that LM_{SEC} can be sensitive to distributional misspecifications and spatial layouts. Our Monte Carlo results given in Section 5 reinforce this point. However, the exact cause of this sensitive is not clear, and a robust test is not available.

3 Robust LM Test for Spatial Error Components

We now present a robustified version of the above LM test statistic. The following basic regularity conditions are necessary for studying the asymptotic behavior of the test statistics.

Assumption 1: *The innovations* $\{\varepsilon_i\}$ *are iid with mean zero, variance* σ_ε^2*, and excess kurtosis* κ_ε*. Also, the moment* $E|\varepsilon_i|^{4+\eta}$ *exists for some* $\eta > 0$*.*

Assumption 2: *The elements* $w_{n,ij}$ *of* W_n *are at most of order* h_n^{-1} *uniformly for all i and j, with the rate sequence* $\{h_n\}$*, bounded or divergent but satisfying* $h_n/n \to 0$ *as* $n \to \infty$*. The sequence* $\{W_n\}$ *are uniformly bounded in both row and column sums. As normalizations, the diagonal elements* $w_{n,ii} = 0$*, and* $\sum_j w_{n,ij} = 1$ *for all i.*

Assumption 3: *The elements of the* $n \times k$ *matrix* X_n *are uniformly bounded for all n, and* $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ *exists and is nonsingular.*

The Assumption 1 corresponds to one assumption of Kelejian and Prucha (2001) for their central limit theorem of linear-quadratic forms. Assumption 2 corresponds to one assumption in Lee (2004a) which identifies the different types of spatial dependence. Typically, one type of spatial dependence corresponds to the case where each unit has fixed number of neighbors such as Rook contiguity and in this case h_n is bounded, and the other type of spatial dependence corresponds to the case where the number of neighbors of each spatial unit grows as n goes to infinity such as the case of group interaction and in this case h_n is divergent. To limit the spatial dependence to a manageable degree, it is thus required that $h_n/n \to 0$ as $n \to \infty$.

Theorem 1: *If* W_n*,* $\{\varepsilon_i\}$ *and* X_n *of Model (1) satisfy the Assumptions A1-A3, then a robust LM test statistic for testing* $H_0 : \sigma_\nu^2 = 0$ *vs* $H_a : \sigma_\nu^2 > 0$ *takes the form*

$$\mathrm{LM}_{\mathrm{SEC}}^* = \frac{\tilde{u}_n' W_n W_n' \tilde{u}_n / \tilde{\sigma}_\varepsilon^2 - S_{1n}}{(\tilde{\kappa}_\varepsilon S_{2n} + S_{3n})^{\frac{1}{2}}} \tag{3}$$

where $S_{1n} = \frac{n}{n-k}\mathrm{tr}(W_n W_n' M_n)$, $S_{2n} = \sum_i a_{n,ii}^2$ *with* $\{a_{n,ii}\}$ *being the diagonal elements of* $A_n = M_n(W_n W_n' - \frac{1}{n} S_{1n} I_n) M_n$*,* $S_{3n} = 2\mathrm{tr}(A_n^2)$*, and* $\tilde{\kappa}_\varepsilon$ *is the excess sample kurtosis of* $\tilde{u}_n$*. Under* H_0*, (i) the positive part of* LM_{SEC}^* *converges to that of* $N(0,1)$*, and (ii)* $\mathrm{LM}_{\mathrm{SEC}}^*$ *is asymptotically equivalent to* $\mathrm{LM}_{\mathrm{SEC}}$ *when* $\kappa_\varepsilon = 0$*.*

Proof: Proof of the theorem needs four lemmas given in Appendix. We have

$$\mathrm{LM}^*_{\mathrm{SEC}} = \frac{\tilde{u}_n' W_n W_n' \tilde{u}_n / \tilde{\sigma}_\varepsilon^2 - S_{1n}}{(\tilde{\kappa}_\varepsilon S_{2n} + S_{3n})^{\frac{1}{2}}} = \frac{\tilde{u}_n' W_n W_n' \tilde{u}_n - \tilde{\sigma}_\varepsilon^2 S_{1n}}{\sigma_\varepsilon^2(\kappa_\varepsilon S_{2n} + S_{3n})^{\frac{1}{2}}} \cdot \frac{\sigma_\varepsilon^2(\kappa_\varepsilon S_{2n} + S_{3n})^{\frac{1}{2}}}{\tilde{\sigma}_\varepsilon^2(\tilde{\kappa}_\varepsilon S_{2n} + S_{3n})^{\frac{1}{2}}}$$

and $\tilde{u}_n' W_n W_n' \tilde{u}_n - \tilde{\sigma}_\varepsilon^2 S_{1n} = \tilde{u}_n'(W_n W_n' - \frac{1}{n} S_{1n} I_n)\tilde{u}_n = u_n' M_n (W_n W_n' - \frac{1}{n} S_{1n} I_n) M_n u_n = u_n' A_n u_n$. Under H_0, the elements of u_n are iid, we have $\mathrm{E}(u_n' A_n u_n) = \sigma_\varepsilon^2 \mathrm{tr}(A_n) = 0$. By Assumption 1 and Lemma A.1 in Appendix and noticing that the matrix A_n is symmetric, we have $\mathrm{Var}(u_n' A_n u_n) = \sigma_\varepsilon^4(\kappa_\varepsilon S_{2n} + S_{3n})$. Now, Assumption 2 ensures that the elements of $W_n W_n'$ are uniformly of order h_n^{-1}, and Lemma A.2 shows that M_n is uniformly bounded in both row and column sums. It follows from Lemma A.4(iii) that $\frac{1}{n} S_{1n} = O(h_n^{-1})$. Assumption 2 and Lemma A.4 lead to that $\{W_n W_n'\}$ are uniformly bounded in both row and column sums. Thus, $\{W_n W_n' - \frac{1}{n} S_{1n} I_n\}$ are uniformly bounded in both row and column sums. Finally, Lemma A.4(i) gives that $\{A_n\}$ are uniformly bounded in both row and column sums. It follows that the central limit theorem of linear-quadratic forms of Kelejian and Prucha (2001) is applicable, which gives

$$\frac{\tilde{u}_n' W_n W_n' \tilde{u}_n - \tilde{\sigma}_\varepsilon^2 S_{1n}}{\sigma_\varepsilon^2(\kappa_\varepsilon S_{2n} + S_{3n})^{\frac{1}{2}}} \xrightarrow{D} N(0,1).$$

Now, it is easy to show that $\tilde{\sigma}_\varepsilon^2 \xrightarrow{p} \sigma_\varepsilon^2$ and that $\tilde{\kappa}_\varepsilon \xrightarrow{p} \kappa_\varepsilon$. Thus,

$$\frac{\sigma_\varepsilon^2(\kappa_\varepsilon S_{2n} + S_{3n})^{\frac{1}{2}}}{\tilde{\sigma}_\varepsilon^2(\tilde{\kappa}_\varepsilon S_{2n} + S_{3n})^{\frac{1}{2}}} \xrightarrow{p} 1.$$

This finishes the proof of Part (i).

For Part (ii), it suffices to show that $S_{1n} \sim T_{1n}$ and $S_{3n} \sim 2T_{2n} - \frac{2}{n} T_{1n}^2$, where '$\sim$' stands for 'asymptotic equivalence'. The former follows from Lemma A.3(i), i.e., $\mathrm{tr}(W_n W_n') + O(1)$. For the latter, write $A_n = M_n A_n^o M_n$, where $A_n^o = W_n W_n' - \frac{1}{n} S_{1n} I_n$. We have

$$\begin{aligned}
\mathrm{tr}(A_n^2) &= \mathrm{tr}(M_n A_n^o M_n M_n A_n^o M_n) \\
&= \mathrm{tr}[(A_n^o M_n)^2] \\
&= \mathrm{tr}((A_n^o)^2) + O(1) \ \text{ by Lemma 3(iii)} \\
&= \mathrm{tr}[(W_n W_n' - \frac{1}{n} S_{1n} I_n)(W_n W_n' - \frac{1}{n} S_{1n} I_n)] \\
&= \mathrm{tr}(W_n W_n' W_n W_n') - \frac{2}{n} S_{1n} \mathrm{tr}(W_n W_n') + \frac{1}{n^2} S_{1n}^2 \mathrm{tr}(I_n) + O(1) \\
&= T_{2n} - \frac{1}{n} T_{1n}^2 + O(1),
\end{aligned}$$

which shows that $S_{3n} = 2\text{tr}(A_n^2) = 2T_2 - \frac{2}{n}T_1^2 + O(1)$. *Q.E.D.*

From Theorem 1 we see that LM^*_{SEC} differs from LM_{SEC} essentially in the denominators and by a term $\tilde{\kappa}_\varepsilon S_{n2}$. This term becomes (asymptotically) negligible when $\kappa_\varepsilon = 0$, which occurs when ε is normal. This is because $S_{2n} = \sum_{i=1}^n a_{ii}^2 \le \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 = \frac{1}{2}S_{3n}$. When $\kappa_\varepsilon \neq 0$, which typically occurs when ε is non-normal, $\tilde{\kappa}_\varepsilon = O_p(1)$. In this case it becomes unclear whether $\tilde{\kappa}_\varepsilon S_{2n}$ is also asymptotically negligible. The key is the relative magnitudes of S_{2n} and S_{3n}, which depend on many factors. The following corollary summarize the detailed results.

Corollary 1: *Under the assumptions of Theorem 1, we have*

(i) If h_n is bounded, then $S_{2n} \sim S_{3n}$, where $\sim$ stands for asymptotic equivalence;

(ii) If $h_n \to \infty$ as $n \to \infty$, then $S_{2n} = O(n/h_n^2)$ and $S_{3n} = O(n/h_n)$;

(iii) $\frac{1}{n}S_{2n} \sim \sigma_n^2(r) \equiv \frac{1}{n}\sum_{i=1}^n (r_{n,i} - \bar{r}_n)^2$, where $\bar{r}_n = \sum_{i=1}^n r_{n,i}$ with $\{r_{n,i}, i = 1, \cdots, n\}$ being the diagonal elements of $W_n W_n'$.

Proof: For (i) and (ii), note that the elements of $W_n W_n' - \frac{1}{n}S_{1n}I_n$ are at most of order $O(h_n^{-1})$ uniformly. By Lemma A.4 (iii), the elements of A_n are also at most of order $O(h_n^{-1})$ uniformly as $\{M_n\}$ are uniformly bounded in both row and column sums. This leads to $S_{2n} = O(n/h_n^2)$. Furthermore, as A_n itself is uniformly bounded in both row and column sums (see the proof of Theorem 1), Lemma A.4 (iii) shows that the elements of A_n^2 are at most of order $O(h_n^{-1})$ uniformly. This shows that $S_{3n} = O(n/h_n)$.

To prove (iii), Note that W_n is row normalized. We have $\frac{1}{n}S_{1n} \sim \frac{1}{n}T_{1n} = \frac{1}{n}\text{tr}(W_n W_n') = \frac{1}{n}\sum_{i=1}^n r_{n,i}$. From Lemma A.3 (v) we have

$$\frac{1}{n}S_{2n} = \frac{1}{n}\sum_{i=1}^n a_{ii}^2 \sim \frac{1}{n}\sum_{i=1}^n \Big((W_n W_n' - \frac{1}{n}S_{1n}I_n)_{ii}\Big)^2 \sim \frac{1}{n}\sum_{i=1}^n (r_{n,i} - \bar{r}_n)^2 \equiv \sigma_n^2(r), \quad (4)$$

which completes the proof of Corollary 1.

The results of Corollary 1 lead to some important conclusions. Firstly, when h_n is bounded, $S_{2n} \sim S_{3n}$. Hence, if $\kappa_\varepsilon \neq 0$, the asymptotic variance of LM_{SEC} will be larger than 1, leading to a test that over-rejects the null hypothesis when errors are nonnormal. This point is confirmed by the Monte Carlo results given in Table 1, where we see that the empirical coverage of the LM_{SEC} test under non-normal errors increases with n, reaching to around {15%, 9.4%, 3.8%} when $n = 1500$ (last line of Table 1) for tests of nominal levels {10%, 5%, 1%}. In contrast, the LM^*_{SEC} test performs very well in general.

Secondly, when h_n increases with n, S_{3n} is generally of higher oder in magnitude than S_{2n}. Hence, as n increases S_{3n} eventually becomes the dominate term in the denominator of the test statistic LM^*_{SEC}, and LM^*_{SEC} would eventually behave like LM_{SEC} even when there exists excess kurtosis or non-normality in general. However, a detailed examination shows that the finite sample difference between LM^*_{SEC} and LM_{SEC} could still be large even when the sample size is very large. Take, for example, $h_n = n^{0.2}$. We have $S_{2n} = O(n^{0.6})$ and $S_{3n} = O(n^{0.8})$. It follows that with $\tilde{\kappa}_\varepsilon$ being $O_p(1)$ the excess kurtosis may have significant impact on the variance and hence on the test statistic even when n is very large. The Monte Carlo simulation results under $h_n = O(n^{0.25})$ given in Section 5 (Table 4, Panels 2 & 3) indeed confirm this point, where we see huge size distortions of the LM_{SEC} test in the cases of non-normal errors. Although the magnitude of size distortion seems decrease as n increases the empirical sizes of the test LM_{SEC} are still around {17%, 11%, 6%} corresponding to nominal sizes {10%, 5%, 1%} even when n is 1500. Take another example with $h_n = n^{0.8}$. We have $S_{2n} = O(n^{-0.6})$ and $S_{3n} = O(n^{0.2})$. Apparently under this situation, the impact of the $\tilde{\kappa}_\varepsilon S_{2n}$ term is negligible. Once again our Monte Carlo simulation confirms this observation (Table 4, Panels 5 & 6).[2]

Thirdly and perhaps more importantly, the result (iii) of Corollary 1 shows that the variance $\sigma_n^2(r)$ (variability in general) of the diagonal elements of W_nW_n' plays a key role in the behavior of the test statistics. When $\sigma_n^2(r) = 0$ or $\sigma_n^2(r) \to 0$ as $n \to \infty$, $\text{LM}_{\text{SEC}} \sim \text{LM}^*_{\text{SEC}}$. When $\sigma_n^2(r) \neq 0$ even when n is large, LM_{SEC} may differ from LM^*_{SEC}, and as n goes large the difference may grow (as in the case where h_n is bounded and errors are nonnormal), or may shrink (as in the case where h_n is unbounded and the errors are nonnormal). It is interesting to note that in the framework of spatial contiguity the ith diagonal element $r_{n,i}$ of W_nW_n' is the reciprocal of the number of neighbors that the ith spatial unit has; and that in the framework of group interaction, $r_{n,i}$ is the reciprocal of

[2]Note that one spatial layout leading to $h_n = n^\delta, 0 < \delta < 1$, is the so-called group interaction (see, e.g., Lee, 2007). In this case h_n corresponds to the average group size. If $\delta = 0.2$, for example, then there are many groups but each group contains only a few members although the number of units in each group grows with n. If $\delta = 0.8$, however, then there are a few groups, but each group contains many members.

the size of the ith group. Hence, in these situations, the variability of $\{r_{n,i}\}$ boils down to whether the number of neighbors or whether the group size varies across the spatial units, and whether these variations disappear as the sample size goes large.

4 Monte Carlo Results

The finite sample performance of the test statistics introduced in this paper is evaluated based on a series of Monte Carlo experiments under a number of different error distributions and a number of different spatial layouts. Comparisons are made between the newly introduced test LM^*_{SEC} and the existing LM_{SEC} of Anselin (2001) to see the improvement of the new tests in the situations where there is a distributional misspecification. The Monte Carlo experiments are carried out based on the following data generating process:

$$Y_i = \beta_0 + X_{1i}\beta_1 + X_{2i}\beta_2 + u_i,$$

where X_{1i}'s are drawn from $10U(0,1)$ and X_{2i}'s are drawn from $5N(0,1)+5$. Both are treated as fixed in the experiments. The parameters $\beta = \{5, 1, 0.5\}'$ and $\sigma = 1$. Seven different sample sizes are considered for each combination of error distributions and spatial layouts. Each set of Monte Carlo results (each row in the table) is based on 10,000 samples.

Three general spatial layouts are considered in the Monte Carlo experiments. The first is based on Rook contiguity, the second is based on Queen contiguity and the third is based on the notion of group interactions.

The detail for generating the W_n matrix under rook contiguity is as follows: (i) index the n spatial units by $\{1, 2, \cdots, n\}$, randomly permute these indices and then allocate them into a lattice of $k \times m (\geq n)$ squares, (ii) let $W_{ij} = 1$ if the index j is in a square which is on immediate left, or right, or above, or below the square which contains the index i, otherwise $W_{ij} = 0, i, j = 1, \cdots, n$, to form an $n \times n$ matrix, and (iii) divide each element of this matrix by its row sum to give W_n. So, under Rook contiguity there are 4 neighbors for each of the inner units, 3 for a unit on the edge, and 2 for a corner unit. The W_n matrix under Queen contiguity can be generated in a similar way as that under rook contiguity, but with additional neighbors which share a common vertex with the unit of interest. In this case a inner unit has 8 neighbors, an edge unit has 5, and a corner unit has 3. Thus the variability of the number of neighbors is greater under Queen than under Rook contiguity. For irregular spatial contiguity, the variation in number of neighbors is greater.

For both regular Rook and Queen spatial layouts, weather k is fixed makes a difference. Thus, we consider two cases: (i) $k = 5$, and (ii) $k = m$. It is easy to show that for spatial units arranged in a regular $k \times m$ lattice, Rook contiguity leads to $\sigma_n^2(r)$ defined in (4) as

$$n\sigma_n^2(r) = \left(\frac{k+m+2}{6km}\right)^2 + 2\left(\frac{1}{12} - \frac{m+k+2}{6km}\right)^2 (k+m-4) + 4\left(\frac{1}{4} - \frac{m+k+2}{6km}\right). \quad (5)$$

With $n = km$, it is easy to see that, if k is fixed, then $m = O(n)$ and $\sigma_n^2(r) = O(1)$ for $k > 2$; if both k and m go large as $n \to \infty$, then $\sigma_n^2(r) = o(1)$. Thus the case of either $k > 2$ or $m > 2$ fixed leads to a permanent variablity in $\{r_{n,i}\}$, whereas the case of neither k nor m fixed leads to a temporary or finite sample variability in $\{r_{n,i}\}$ which disappears as $n \to \infty$. Similarly, under Queen contiguity, we have

$$\begin{aligned} n\sigma_n^2(r) &= \left(\frac{9(k+m)-14}{60km}\right)^2 (k-2)(m-2) + 2\left(\frac{3}{40} - \frac{9(k+m)-14}{60km}\right)^2 (k+m-4) \\ &\quad +4\left(\frac{5}{24} - \frac{9(k+m)-14}{60km}\right), \qquad (6) \end{aligned}$$

which gives $\sigma_n^2(r) = O(1)$ when either $k > 2$ or $m > 2$ is fixed, and $\sigma_n^2(r) = o(1)$ when neither k nor m is fixed.

To generate the W_n matrix according to the group interaction scheme, suppose we have k groups of sizes $\{m_1, m_2, \cdots, m_k\}$. Define $W_n = \text{diag}\{W_j/(m_j - 1), j = 1, \cdots, k\}$, a matrix formed by placing the submatrices W_j along the diagonal direction, where W_j is an $m_j \times m_j$ matrix with ones on the off-diagonal positions and zeros on the diagonal positions. Note that $n = \sum_{j=1}^{k} m_j$. We consider three different ways of generating the group sizes.

The first is that the group size is a constant across the groups and with respect to the sample size n, i.e., $m_1 = m_2 = \cdots = m_k = m$, where m is free of n. In this case increasing n means having more groups of the same size m. The second is that the group size changes across the groups but not with respect to the sample size n. In this case, increasing n means having more groups of sizes m_1 or $m_2, \cdots$. The third method is the most complicated one and is described as follows: (i) calculate the number of groups according to $k =$Round(n^ϵ), and the approximate average group size $m = n/k$, (ii) generate the group sizes $(m_1, m_2, \cdots, m_k)$ according to a discrete uniform distribution from $m/2$ to $3m/2$, (iii) adjust the group sizes so that $\sum_{j=1}^{k} m_j = n$, and (iv) define $W = \text{diag}\{W_j/(m_j-1), j = 1, \cdots, k\}$ as described above. In our Monte Carlo experiments, we choose $\epsilon = 0.25, 0.50$, and 0.75, representing respectively the situations where (i) there are few groups and many spatial units in a group, (ii) the number of groups and the sizes of the groups are of the same magnitude, and (iii) there are many groups of few elements in each. Clearly, the first method leads to $\sigma_n^2(r) = 0$, the second method leads to $\sigma_n^2(r) \neq 0$, and the third method leads to $\sigma_n^2(r) \to 0$ as $n \to \infty$. In all spatial layouts described above, only the last one gives h_n unbounded with $h_n = n^{1-\epsilon}$.

For the error distributions, the reported Monte Carlo results correspond to the following three: (i) standard normal, (ii) mixture normal, (iii) log-normal, and (iv) chi-square, where (ii)-(iv) are all standardized to have mean zero and variance one. The standardized normal-mixture variates are generated according to

$$u_i = ((1-\xi_i)Z_i + \xi_i \sigma Z_i)/(1-p+p*\sigma^2)^{0.5},$$

where ξ is a Bernoulli random variable with probability of success p and Z_i is standard normal independent of ξ. The parameter p in this case also represents the proportion of mixing the two normal populations. In our experiments, we choose $p = 0.05$, meaning that 95% of the random variates are from standard normal and the remaining 5% are from another normal population with standard deviation σ. We choose $\sigma = 10$ to simulate the situation where there are gross errors in the data. The standardized lognormal random variates are generated according to

$$u_i = [\exp(Z_i) - \exp(0.5)]/[\exp(2) - \exp(1)]^{0.5}.$$

This gives an error distribution that is both skewed and leptokurtic. The normal mixture gives an error distribution that is still symmetric like normal but leptokurtic. The standardized chi-square random variates are generated in a similar fashion. Other non-normal distributions, such as normal-gamma mixture, are also considered and the results are available from the author upon request.

Selected Monte Carlo results are summarized in Tables 1-4. The results in Table 1 correspond to Rook or Queen contiguity with k fixed at 5. In this case, h_n is bounded, $\sigma_n^2(r) = O(1)$, and $\tilde{\kappa} = O_p(1)$ if errors are nonnormal. Hence $\tilde{\kappa} S_{2n} \sim S_{3n}$, which means that the difference between LM_{SEC} and LM_{SEC}^* does not vanish as n goes large. Monte Carlo results in Table 1 indeed show that when errors are not normal the sizes of the $\mathrm{LM}_{\mathrm{SEC}}$ test get larger as n increases. This gets more severe under Queen contiguity as in this case $\sigma_n^2(r)$ is larger. In contrast, the sizes of the $\mathrm{LM}_{\mathrm{SEC}}$ test are all quite close to their nominal levels irrespective of error distributions and spatial layouts. It is interesting to note that $\mathrm{LM}^*_{\mathrm{SEC}}$ seems perform better than $\mathrm{LM}_{\mathrm{SEC}}$ even when the errors are drawn from a normal population.

The results reported in Table 2 also correspond to Rook or Queen contiguity but with $k = m = \sqrt{n}$. In this case $\sigma_n^2(r) = o(1)$, and hence the term $\tilde{\kappa} S_{2n}$ is negligible relative to S_{3n} when n is large, and the two statistics should behave similarly. The results confirm this theoretical finding although the comparative advantage seems go to the new statistic.

The results reported in Table 3 correspond to group interaction spatial layout with fixed group sizes $\{2, 3, 4, 5, 6, 7\}$ for upper three panels, and a fixed group size 5 for the lower three panels. The first case gives $\sigma_n^2(r) = O(1)$ and the second gives $\sigma_n^2(r) = 0$. The results in the upper three panels show that when errors are not normal, the $\mathrm{LM}_{\mathrm{SEC}}$ test can perform quite badly with the empirical sizes far above their nominal levels. In contrast, the $\mathrm{LM}^*_{\mathrm{SEC}}$ test performs very well in all situations.

The results given in Table 4 correspond to group interaction again, but this time the group size varies across groups and increases with n, which results in an unbounded h_n. Although the theory predicts that the two tests should behave similarly when n is large and the results indeed show some signs of convergence in size, the $\mathrm{LM}_{\mathrm{SEC}}$ can still perform badly even when $n = 1500$, when there are many small groups (upper three panels where $k = n^{0.75}$). Again, the $\mathrm{LM}^*_{\mathrm{SEC}}$ test performs very well in general.

Table 1. Empirical Means, SDs and Sizes for One-sided Tests, Rook or Queen, $k = 5^*$

	Anselin's Test					Proposed Test				
n	Mean	SD	10%	5%	1%	Mean	SD	10%	5%	1%
	Rook Contiguity, Normal Error									
20	-0.2010	0.9273	.0673	.0325	.0084	-0.0168	1.1201	.1332	.0831	.0279
50	-0.1823	0.9564	.0712	.0366	.0092	0.0112	1.0435	.1148	.0675	.0203
100	-0.1516	0.9909	.0817	.0406	.0101	-0.0013	1.0376	.1123	.0629	.0167
200	-0.1171	0.9904	.0841	.0443	.0092	-0.0006	1.0145	.1066	.0581	.0134
500	-0.0603	0.9951	.0916	.0461	.0099	0.0174	1.0051	.1095	.0556	.0130
1000	-0.0446	0.9938	.0941	.0463	.0093	0.0056	0.9986	.1049	.0532	.0112
1500	-0.0459	1.0045	.0972	.0501	.0099	-0.0010	1.0077	.1049	.0555	.0115
	Rook Contiguity, Normal-Mixture, $\tau = 10, p = 0.05$									
20	-0.1899	0.9237	.0663	.0326	.0061	-0.0033	1.0526	.1209	.0678	.0204
50	-0.1941	0.9862	.0758	.0480	.0158	-0.0006	0.9570	.0949	.0594	.0189
100	-0.1414	1.0721	.0883	.0529	.0216	0.0072	0.9485	.0886	.0508	.0165
200	-0.1139	1.1124	.0983	.0582	.0215	0.0021	0.9595	.0894	.0476	.0141
500	-0.0782	1.1436	.1107	.0663	.0212	-0.0005	0.9814	.0919	.0479	.0121
1000	-0.0443	1.1576	.1206	.0726	.0226	0.0052	0.9930	.0969	.0503	.0115
1500	-0.0496	1.1506	.1214	.0699	.0210	-0.0043	0.9872	.0975	.0483	.0109
	Rook Contiguity, Lognormal									
20	-0.1866	0.9192	.0685	.0342	.0083	0.0005	1.0527	.1211	.0712	.0223
50	-0.2102	0.9740	.0775	.0446	.0132	-0.0182	0.9815	.1041	.0611	.0198
100	-0.1424	1.0493	.0903	.0538	.0207	0.0083	0.9843	.1019	.0587	.0196
200	-0.1142	1.0864	.1011	.0588	.0234	0.0017	0.9827	.0999	.0545	.0189
500	-0.0820	1.1306	.1128	.0709	.0259	-0.0037	0.9747	.0956	.0535	.0156
1000	-0.0561	1.1976	.1345	.0811	.0289	-0.0070	0.9922	.1031	.0538	.0137
1500	-0.0331	1.2093	.1390	.0898	.0309	0.0089	0.9885	.1066	.0544	.0139
	Queen Contiguity, Normal Error									
20	-0.6166	0.6046	.0073	.0027	.0001	-0.0301	1.0992	.1223	.0802	.0328
50	-0.2943	0.8860	.0546	.0304	.0098	0.0277	1.0541	.1178	.0756	.0311
100	-0.2583	0.9216	.0597	.0336	.0089	-0.0114	1.0123	.1070	.0644	.0215
200	-0.2125	0.9695	.0714	.0385	.0102	-0.0215	1.0158	.1063	.0599	.0200
500	-0.1087	0.9763	.0840	.0448	.0124	0.0105	0.9947	.1042	.0590	.0161
1000	-0.1011	1.0008	.0860	.0431	.0115	-0.0126	1.0107	.1018	.0544	.0141
1500	-0.0859	0.9809	.0855	.0439	.0099	-0.0150	0.9873	.0959	.0502	.0121
	Queen Contiguity, Normal-Mixture, $\tau = 10, p = 0.05$									
20	-0.6033	0.6535	.0071	.0019	.0001	-0.0073	1.0820	.1246	.0806	.0321
50	-0.3206	0.9617	.0662	.0424	.0160	-0.0020	0.9774	.1043	.0669	.0249
100	-0.2399	1.0861	.0837	.0517	.0221	0.0059	0.9666	.0958	.0564	.0220
200	-0.1877	1.1425	.0973	.0579	.0222	0.0031	0.9605	.0920	.0495	.0160
500	-0.1341	1.1929	.1160	.0733	.0267	-0.0128	0.9809	.0966	.0500	.0154
1000	-0.0840	1.2116	.1317	.0796	.0268	0.0039	0.9899	.1024	.0521	.0132
1500	-0.0755	1.2058	.1320	.0808	.0306	-0.0039	0.9845	.0980	.0533	.0126
	Queen Contiguity, Lognormal									
20	-0.6127	0.6418	.0078	.0022	.0003	-0.0214	1.0747	.1216	.0798	.0300
50	-0.3231	0.9232	.0573	.0353	.0134	-0.0067	0.9813	.0990	.0621	.0242
100	-0.2604	1.0219	.0738	.0460	.0165	-0.0121	0.9754	.0966	.0597	.0205
200	-0.1994	1.1125	.0902	.0559	.0223	-0.0076	0.9826	.0959	.0558	.0195
500	-0.1397	1.1957	.1179	.0751	.0292	-0.0172	0.9944	.0976	.0551	.0185
1000	-0.0783	1.2472	.1375	.0898	.0371	0.0061	0.9873	.1034	.0580	.0148
1500	-0.0715	1.2921	.1473	.0942	.0379	0.0007	0.9968	.1033	.0551	.0144

*The n spatial units are randomly placed on a lattice of $k \times m$ squares.

Table 2. Empirical Means, SDs and Sizes for One-sided Tests, Rook or Queen, $k = m^*$

n	Anselin's Test Mean	SD	10%	5%	1%	Proposed Test Mean	SD	10%	5%	1%
	Rook Contiguity, Normal Error									
5^2	-0.2445	0.9219	.0584	.0305	.0064	0.0008	1.0915	.1249	.0752	.0261
7^2	-0.2366	0.9405	.0624	.0315	.0065	0.0091	1.0525	.1165	.0689	.0220
10^2	-0.1817	0.9560	.0674	.0345	.0066	-0.0007	1.0063	.1038	.0547	.0141
15^2	-0.1050	0.9848	.0840	.0415	.0100	0.0126	1.0074	.1057	.0579	.0139
23^2	-0.0873	0.9996	.0865	.0443	.0120	-0.0083	1.0096	.1019	.0534	.0142
32^2	-0.0723	0.9943	.0879	.0420	.0092	-0.0154	0.9994	.0995	.0483	.0110
39^2	-0.0536	0.9978	.0917	.0451	.0089	-0.0059	1.0014	.1011	.0506	.0102
	Rook Contiguity, Normal-Mixture, $\tau = 10, p = 0.05$									
5^2	-0.2374	0.9456	.0722	.0412	.0099	0.0081	1.0307	.1208	.0732	.0247
7^2	-0.2574	0.9919	.0751	.0424	.0134	-0.0129	0.9775	.1044	.0616	.0170
10^2	-0.1842	1.0612	.0818	.0502	.0223	-0.0018	0.9573	.0877	.0523	.0206
15^2	-0.1194	1.1046	.0957	.0589	.0232	-0.0017	0.9625	.0878	.0502	.0166
23^2	-0.0925	1.0985	.1023	.0594	.0216	-0.0123	0.9798	.0921	.0493	.0159
32^2	-0.0311	1.0864	.1083	.0635	.0213	0.0238	0.9910	.0995	.0529	.0163
39^2	-0.0645	1.0562	.0997	.0564	.0157	-0.0157	0.9761	.0917	.0481	.0118
	Rook Contiguity, Logormal									
5^2	-0.2428	0.9405	.0718	.0370	.0092	0.0023	1.0382	.1213	.0754	.0234
7^2	-0.2479	0.9542	.0661	.0383	.0133	-0.0050	0.9788	.1004	.0594	.0191
10^2	-0.1829	1.0234	.0818	.0497	.0185	-0.0027	0.9714	.0982	.0568	.0212
15^2	-0.1088	1.0942	.0987	.0629	.0249	0.0080	0.9898	.0986	.0611	.0226
23^2	-0.0762	1.0997	.1084	.0682	.0265	0.0022	0.9836	.1012	.0595	.0185
32^2	-0.0379	1.1008	.1176	.0712	.0261	0.0189	0.9881	.1075	.0611	.0187
39^2	-0.0584	1.0896	.1067	.0645	.0227	-0.0088	0.9835	.0968	.0557	.0166
	Queen Contiguity, Normal Error									
5^2	-0.5225	0.7358	.0238	.0123	.0033	-0.0105	1.0931	.1226	.0826	.0385
7^2	-0.3269	0.8596	.0506	.0256	.0078	0.0033	1.0357	.1157	.0730	.0275
10^2	-0.2960	0.9100	.0536	.0271	.0093	-0.0133	1.0149	.1066	.0620	.0212
15^2	-0.1800	0.9562	.0738	.0399	.0099	0.0106	1.0013	.1094	.0618	.0190
23^2	-0.1276	0.9776	.0785	.0420	.0096	0.0016	0.9979	.1018	.0555	.0146
32^2	-0.0924	1.0037	.0899	.0478	.0118	0.0026	1.0149	.1089	.0581	.0160
39^2	-0.0721	0.9906	.0931	.0479	.0109	0.0048	0.9980	.1061	.0573	.0139
	Queen Contiguity, Normal-Mixture, $\tau = 10, p = 0.05$									
5^2	-0.5254	0.7444	.0244	.0117	.0016	-0.0147	1.0131	.1102	.0707	.0272
7^2	-0.3285	0.9625	.0660	.0418	.0158	0.0010	0.9760	.1022	.0666	.0263
10^2	-0.2933	1.0668	.0772	.0499	.0216	-0.0089	0.9668	.0933	.0588	.0233
15^2	-0.2062	1.1313	.0997	.0621	.0253	-0.0133	0.9558	.0931	.0530	.0188
23^2	-0.1173	1.1626	.1119	.0680	.0274	0.0105	0.9927	.1032	.0569	.0189
32^2	-0.0760	1.1260	.1115	.0658	.0249	0.0172	0.9883	.0984	.0556	.0172
39^2	-0.0675	1.1151	.1105	.0669	.0240	0.0084	0.9932	.0973	.0548	.0158
	Queen Contiguity, Lognormal									
5^2	-0.5093	0.7573	.0257	.0122	.0028	0.0073	1.0434	.1203	.0789	.0305
7^2	-0.3285	0.9280	.0617	.0364	.0122	0.0002	0.9883	.1041	.0678	.0261
10^2	-0.2768	1.0158	.0725	.0459	.0181	0.0090	0.9868	.1032	.0637	.0257
15^2	-0.1940	1.0903	.0931	.0557	.0233	-0.0041	0.9652	.0966	.0537	.0206
23^2	-0.1260	1.1619	.1118	.0710	.0316	0.0034	0.9958	.1058	.0623	.0222
32^2	-0.0767	1.1719	.1144	.0729	.0317	0.0154	1.0069	.1029	.0621	.0218
39^2	-0.0792	1.1359	.1089	.0704	.0290	-0.0014	0.9813	.0979	.0576	.0193

*The n spatial units are randomly placed on a lattice of $k \times m$ squares.

Table 3. Empirical Means, SDs and Sizes for One-sided Tests, Group Interaction

	Anselin's Test					Proposed Test				
m	Mean	SD	10%	5%	1%	Mean	SD	10%	5%	1%
	Group sizes = $\{2,3,4,5,6,7\}^*$ repeated m times, Normal Error									
1	-0.3336	0.9075	.0547	.0293	.0065	-0.0007	1.0951	.1279	.0825	.0320
2	-0.1264	0.9809	.0861	.0463	.0124	-0.0004	1.0458	.1185	.0693	.0217
4	-0.1178	0.9873	.0835	.0471	.0113	-0.0047	1.0218	.1089	.0615	.0164
8	-0.0938	0.9996	.0891	.0457	.0117	-0.0126	1.0171	.1080	.0573	.0143
19	-0.0548	0.9965	.0905	.0462	.0119	-0.0022	1.0035	.1022	.0528	.0135
37	-0.0464	0.9975	.0953	.0481	.0115	-0.0096	1.0015	.1019	.0531	.0130
56	-0.0201	0.9946	.0977	.0484	.0088	0.0101	0.9971	.1029	.0528	.0105
	Group sizes = $\{2,3,4,5,6,7\}$ repeated m times, Normal Mixture, $\tau = 5, p = 0.05$									
1	-0.3616	1.0951	.0759	.0507	.0244	-0.0246	1.1121	.1241	.0853	.0419
2	-0.1429	1.3396	.1237	.0851	.0488	-0.0119	1.0576	.1149	.0791	.0396
4	-0.0806	1.5452	.1529	.1176	.0732	0.0198	1.0376	.1160	.0818	.0393
8	-0.0885	1.6708	.1720	.1334	.0843	-0.0060	1.0153	.1108	.0745	.0294
19	-0.0603	1.7699	.1962	.1538	.0950	-0.0065	1.0100	.1073	.0686	.0233
37	-0.0260	1.8042	.2152	.1663	.1010	0.0051	1.0007	.1098	.0664	.0216
56	-0.0316	1.8223	.2215	.1691	.1025	-0.0023	1.0062	.1069	.0633	.0186
	Group sizes = $\{2,3,4,5,6,7\}$ repeated m times, Chi-Square with df = 3									
1	-0.3270	1.0297	.0783	.0503	.0200	0.0058	1.1219	.1325	.0934	.0433
2	-0.1261	1.1192	.1080	.0702	.0311	0.0022	1.0433	.1197	.0749	.0289
4	-0.0937	1.1844	.1235	.0800	.0346	0.0164	1.0313	.1190	.0704	.0242
8	-0.0943	1.2152	.1230	.0820	.0352	-0.0090	1.0179	.1051	.0630	.0212
19	-0.0460	1.2414	.1404	.0901	.0382	0.0048	1.0127	.1077	.0626	.0197
37	-0.0329	1.2334	.1426	.0920	.0368	0.0035	0.9976	.1048	.0585	.0169
56	-0.0283	1.2400	.1458	.0909	.0366	0.0009	0.9958	.1020	.0557	.0150
	Group sizes = 5^* repeated m times, Normal Error									
5	-0.3917	0.8607	.0468	.0256	.0077	-0.0117	1.0770	.1245	.0804	.0336
10	-0.2553	0.9431	.0676	.0400	.0100	0.0004	1.0428	.1193	.0726	.0263
20	-0.1217	0.9845	.0838	.0475	.0139	0.0078	1.0225	.1081	.0650	.0214
40	-0.0904	1.0024	.0902	.0528	.0128	0.0144	1.0235	.1087	.0649	.0175
100	-0.0653	0.9883	.0911	.0454	.0100	-0.0012	0.9964	.1036	.0545	.0117
200	-0.0392	1.0063	.0974	.0490	.0111	0.0034	1.0102	.1064	.0545	.0126
300	-0.0290	0.9954	.0962	.0519	.0121	0.0088	0.9982	.1032	.0547	.0131
	Group sizes = 5 repeated m times, Normal Mixture, $\tau = 5, p = 0.05$									
5	-0.3917	0.8169	.0393	.0199	.0054	-0.0117	1.0148	.1111	.0711	.0275
10	-0.2580	0.8821	.0552	.0275	.0069	-0.0025	0.9709	.1019	.0595	.0183
20	-0.1300	0.9263	.0710	.0373	.0085	-0.0008	0.9609	.0960	.0524	.0142
40	-0.0986	0.9412	.0742	.0383	.0095	0.0061	0.9605	.0936	.0518	.0129
100	-0.0599	0.9807	.0871	.0432	.0095	0.0042	0.9886	.0979	.0498	.0117
200	-0.0405	0.9772	.0894	.0474	.0093	0.0021	0.9810	.0976	.0510	.0111
300	-0.0399	0.9946	.0908	.0453	.0095	-0.0021	0.9973	.0994	.0497	.0111
	Group sizes = 5 repeated m times, Chi-Square with df = 3									
5	-0.3669	0.8508	.0465	.0259	.0080	0.0192	1.0603	.1229	.0810	.0343
10	-0.2461	0.9321	.0640	.0382	.0131	0.0105	1.0285	.1114	.0687	.0285
20	-0.1439	0.9663	.0817	.0434	.0116	-0.0153	1.0033	.1083	.0600	.0190
40	-0.0948	0.9852	.0911	.0479	.0120	0.0099	1.0058	.1106	.0620	.0167
100	-0.0674	0.9906	.0905	.0485	.0127	-0.0033	0.9986	.1013	.0565	.0152
200	-0.0549	0.9947	.0908	.0482	.0121	-0.0123	0.9986	.0968	.0528	.0133
300	-0.0434	0.9940	.0909	.0492	.0121	-0.0057	0.9968	.0966	.0537	.0131

* For group sizes = $\{2,3,4,5,6,7\}$, $n = 27m$; for group size =5, $n = 5m$.

Table 4. Empirical Means, SDs and Sizes for One-sided Tests, Group Interaction

	Anselin's Test					Proposed Test				
n	Mean	SD	10%	5%	1%	Mean	SD	10%	5%	1%
	$k = n^{.75}$, group sizes $\sim U(.5n^{.25}, 1.5n^{.25})$, Normal Error*									
20	-0.1115	0.9529	.0758	.0308	.0018	-0.0015	1.1489	.1405	.0809	.0150
50	-0.0878	0.9946	.0865	.0369	.0056	0.0057	1.0608	.1169	.0608	.0115
100	-0.0758	0.9849	.0879	.0429	.0078	0.0074	1.0176	.1102	.0551	.0120
200	-0.0880	0.9975	.0870	.0449	.0090	-0.0166	1.0148	.1029	.0547	.0127
500	-0.0669	0.9990	.0958	.0489	.0109	-0.0050	1.0071	.1079	.0581	.0135
1000	-0.0597	0.9954	.0890	.0440	.0116	-0.0114	0.9998	.0983	.0493	.0136
1500	-0.0379	0.9968	.0941	.0489	.0106	0.0024	0.9999	.1016	.0544	.0122
	$k = n^{.75}$, group sizes $\sim U(.5n^{.25}, 1.5n^{.25})$, Normal Mixture									
20	-0.1230	1.3587	.1569	.1167	.0582	-0.0150	1.1944	.1859	.1121	.0108
50	-0.1199	1.9904	.2418	.2051	.1472	-0.0074	1.0853	.1799	.0704	.0048
100	-0.1266	2.3663	.2421	.2097	.1591	-0.0166	1.0398	.1502	.0923	.0182
200	-0.0613	2.1549	.2092	.1708	.1201	0.0037	1.0275	.1119	.0836	.0373
500	-0.0680	1.7972	.1954	.1519	.0918	-0.0023	0.9985	.0988	.0608	.0242
1000	-0.0728	1.4726	.1750	.1195	.0571	-0.0156	0.9886	.0972	.0535	.0145
1500	-0.0415	1.4200	.1697	.1145	.0541	-0.0011	1.0026	.1008	.0573	.0178
	$k = n^{.75}$, group sizes $\sim U(.5n^{.25}, 1.5n^{.25})$, Lognormal									
20	-0.1327	1.3252	.1606	.1144	.0381	-0.0249	1.2033	.1839	.1038	.0108
50	-0.0631	1.7203	.2239	.1778	.1064	0.0235	1.0846	.1698	.0775	.0068
100	-0.0699	2.0036	.2207	.1838	.1239	0.0064	1.0500	.1504	.0921	.0155
200	-0.0738	1.8680	.1814	.1438	.0851	-0.0016	1.0179	.1084	.0728	.0360
500	-0.0586	1.7801	.1731	.1337	.0779	0.0004	1.0056	.1006	.0604	.0242
1000	-0.0394	1.5745	.1695	.1234	.0676	0.0027	0.9955	.1065	.0622	.0236
1500	-0.0426	1.5613	.1634	.1148	.0640	-0.0037	1.0063	.0998	.0612	.0243
	$k = n^{.5}$, group sizes $\sim U(.5n^{.5}, 1.5n^{.5})$, Normal Error									
20	-0.2132	0.9496	.0794	.0495	.0152	-0.0085	1.1136	.1320	.0894	.0396
50	-0.2461	0.9408	.0727	.0429	.0125	0.0047	1.0487	.1204	.0793	.0312
100	-0.2265	0.9570	.0737	.0417	.0142	0.0002	1.0302	.1172	.0717	.0265
200	-0.1998	0.9649	.0759	.0442	.0152	-0.0211	1.0078	.1041	.0653	.0242
500	-0.1549	0.9902	.0818	.0461	.0143	-0.0055	1.0172	.1088	.0624	.0222
1000	-0.1191	0.9767	.0853	.0443	.0119	0.0060	0.9943	.1074	.0581	.0178
1500	-0.0874	0.9891	.0953	.0503	.0131	0.0238	1.0028	.1139	.0621	.0177
	$k = n^{.5}$, group sizes $\sim U(.5n^{.5}, 1.5n^{.5})$, Normal Mixture									
20	-0.2098	0.9573	.0920	.0530	.0136	-0.0029	1.0531	.1358	.0876	.0291
50	-0.2373	0.9128	.0605	.0336	.0112	0.0130	0.9410	.0954	.0533	.0199
100	-0.2362	0.9538	.0702	.0411	.0147	-0.0090	0.9265	.0922	.0538	.0190
200	-0.1824	1.0352	.0869	.0525	.0199	-0.0019	0.9476	.0936	.0548	.0192
500	-0.1611	1.0072	.0862	.0482	.0152	-0.0112	0.9797	.0993	.0582	.0168
1000	-0.1159	0.9939	.0866	.0481	.0144	0.0090	0.9819	.1018	.0567	.0173
1500	-0.1197	1.0109	.0885	.0508	.0151	-0.0086	0.9911	.1014	.0582	.0170
	$k = n^{.5}$, group sizes $\sim U(.5n^{.5}, 1.5n^{.5})$, Lognormal									
20	-0.2075	0.9667	.0883	.0525	.0188	-0.0001	1.0721	.1279	.0853	.0354
50	-0.2366	0.9137	.0660	.0394	.0132	0.0140	0.9656	.1010	.0635	.0235
100	-0.2268	0.9661	.0722	.0454	.0182	-0.0003	0.9732	.0983	.0610	.0254
200	-0.1921	1.0295	.0853	.0553	.0245	-0.0120	0.9715	.0960	.0608	.0261
500	-0.1456	1.0137	.0838	.0519	.0211	0.0026	0.9865	.1006	.0611	.0233
1000	-0.1130	1.0002	.0877	.0535	.0187	0.0119	0.9832	.1016	.0612	.0212
1500	-0.1116	1.0249	.0929	.0561	.0190	-0.0014	0.9928	.1043	.0616	.0198

* k is the number of groups. Average group size $= n/k$

5 Conclusions

The proposed modified LM test performs very well in general, irrespective of error distributions and spatial layouts. Variations in the diagonal elements of $W_n W_n'$ play a key role in the robustness of Anselin's (2001) LM test. When there is a variation and this variation does not vanish as n goes large, Anselin's LM test is not robust against non-normality, and the size of the test does not converge to its nominal value. When there is a variation but it vanishes as n goes large, there can be huge finite sample size distortions for Anselin's LM test when errors are nonnormal, even when n is fairly large. When the elements of W_n is of order h_n^{-1} with $h_n \to \infty$ as $n \to \infty$, our theory shows that the proposed LM test and Anselin's LM test are asymptotically equivalent, but it may require a very large n for Anselin's LM test to behave properly.

Appendix: Some Useful Lemmas

Lemma A.1 (Lee, 2004a, p. 1918): *Let v_n be an $n \times 1$ random vector of iid elements with mean zero, variance σ_v^2, and finite excess kurtosis κ_v. Let A_n be an n dimensional square matrix. Then* $\mathrm{E}(v_n' A_n v_n) = \sigma_v^2 \mathrm{tr}(A_n)$ *and* $\mathrm{Var}(v_n' A_n v_n) = \sigma_v^4 \kappa_v \sum_{i=1}^n a_{n,ii}^2 + \sigma_v^4 \mathrm{tr}(A_n A_n' + A_n^2)$*, where $\{a_{n,ii}\}$ are the diagonal elements of A_n.*

Lemma A.2 (Lee, 2004a, p. 1918): *Suppose that the elements of the $n \times k$ matrix X_n are uniformly bounded; and $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ exists and is nonsingular. Then the projectors $P_n = X_n(X_n'X_n)^{-1}X_n'$ and $M_n = I_n - X_n(X_n'X_n)^{-1}X_n'$ are uniformly bounded in both row and column sums.*

Lemma A.3 (Lemma A.9, Lee, 2004b): *Suppose that A_n represents a sequence of $n \times n$ matrices that are uniformly bounded in both row and column sums. The elements of the $n \times k$ matrix X_n are uniformly bounded; and $\lim_{n\to\infty} \frac{1}{n} X_n' X_n$ exists and is nonsingular. Let $M_n = I_n - X_n(X_n'X_n)^{-1}X_n'$. Then*

(i) $\mathrm{tr}(M_n A_n) = \mathrm{tr}(A_n) + O(1)$

(ii) $\mathrm{tr}(A_n' M_n A_n) = \mathrm{tr}(A_n' A_n) + O(1)$

(iii) $\mathrm{tr}[(M_n A_n)^2] = \mathrm{tr}(A_n^2) + O(1)$*, and*

(iv) $\mathrm{tr}[(A_n' M_n A_n)^2] = \mathrm{tr}[(M_n A_n' A_n)^2] = \mathrm{tr}[A_n' A_n)^2] + O(1)$

Furthermore, if $a_{n,ij} = O(h_n^{-1})$ for all i and j, then

(v) $\mathrm{tr}^2(M_n A_n) = \mathrm{tr}^2(A_n) + O(\frac{n}{h_n})$*, and*

(vi) $\sum_{i=1}^n [(M_n A_n)_{ii}]^2 = \sum_{i=1}^n (a_{n,ii})^2 + O(h_n^{-1})$*,*

where $(M_n A_n)_{ii}$ are the diagonal elements of $M_n A_n$, and $a_{n,ii}$ the diagonal elements of A_n.

Lemma A.4 (Kelejian and Prucha, 1999; Lee, 2002): *Let $\{A_n\}$ and $\{B_n\}$ be two sequences of $n \times n$ matrices that are uniformly bounded in both row and column sums. Let C_n be a sequence of confirmable matrices whose elements are uniformly $O(h_n^{-1})$. Then*

(i) the sequence $\{A_n B_n\}$ are uniformly bounded in both row and column sums,

(ii) the elements of A_n are uniformly bounded and tr$(A_n) = O(n)$, and

(iii) the elements of $A_n C_n$ and $C_n A_n$ are uniformly $O(h_n^{-1})$.

References

[1] Anselin L. (2001). Rao's score test in spatial econometrics. *Journal of Statistical Planning and Inferences* **97**, 113-139.

[2] Anselin L. and Bera, A. K. (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In: *Handbook of Applied Economic Statistics, Edited by Aman Ullah and David E. A. Giles*. New York: Marcel Dekker.

[3] Anselin L. and Moreno, R. (2003). Properties of tests for spatial error components. *Regional Science and Urban Economics* **33**, 595-618.

[4] Florax, R. J. G. M. and de Graaff, T. (2004). The performance of diagnostic tests for spatial dependence in linear regression models: a meta-analysis of simulation studies. In: *Advances in Spatial Econometrics, Edited by L. Anselin, R. J. G. M. Florax and S. J. Rey*. Berlin: Springer-Verlag.

[5] Kelejian, H. H. and Robinson, D. P. (1995). Spatial correlation: a suggested alternative to the autoregressive models. In: *New Directions in Spatial Econometrics, Edited by L. Anselin and R. J. G. M. Florax*. Berlin: Springer-Verlag.

[6] Kelejian H. H. and Prucha, I. R. (2001). On the asymptotic distribution of the Moran I test statistic with applications. *Journal of Econometrics* **104**, 219-257.

[7] Lee, L. F. (2002). Consistency and efficiency of least squares estimation for mixed regressive, spatial autoregressive models. *Econometric Theory* **18**, 252-277.

[8] Lee, L. F. (2004a). Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models. *Econometrica* **72**, 1899-1925.

[9] Lee, L. F. (2004b). A supplement to 'Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models'. Working paper, Department of Economics, Ohio State University.

[10] Lee, L. F. (2007). Identification and estimation of econometric models with group interaction, contextual factors and fixed effects. *Journal of Econometrics* **140**, 333-374.

[11] Moran, P.A.P. (1950). Notes on continuous stochastic phenomena. *Biometrika* **37**, 17-33.

[12] Verbeke, G. and Molenberghs, G. (2003). The use of score test for inference on variance components. *Biometrics* **59**, 254-262.