Geographical Analysis ISSN 0016-7363

A Generalized Randomization Approach to Local Measures of Spatial Association

Sang-Il Lee

Department of Geography Education, Seoul National University, Seoul, South Korea

This article establishes a unified randomization significance testing framework upon which various local measures of spatial association are commonly predicated. The generalized randomization approach presented is composed of two testing procedures, the extended Mantel test and the generalized vector randomization test. These two procedures employ different randomization assumptions, namely total and conditional randomization, according to the way in which they incorporate local measures. By properly specifying necessary matrices and vectors for a particular local measure of spatial association under a particular randomization assumption, the generalized randomization approach as a whole yields a reliable set of equations for expected values and variances, which then is confirmed by a Monte Carlo simulation utilizing random permutations.

Introduction

Local measures of spatial association have increasingly attracted considerable attention in a variety of academic fields dealing with geographically referenced data (for reviews, see Getis and Ord 1996; Fotheringham 1997; Unwin and Unwin 1998). Two sets of univariate measures have been proposed, Getis-Ord G_i and G_i^* (Getis and Ord 1992; Ord and Getis 1995) and local Moran's I_i and Geary's c_i (Anselin 1995), and collectively lead to the advent of a general class of local indicators of spatial association (LISA; Anselin 1995; Getis and Ord 1996). Boots (2003) seeks to develop a local measure of spatial association for categorical data. Significant efforts have been dedicated to provide viable significance testing methods for the local measures (Anselin 1995; Ord and Getis 1995; Bao and Henry 1996; Tiefelsdorf and Boots 1997; Sokal, Oden, and Thomson 1998; Tiefelsdorf 2000; Leung, Mei, and Zhang 2003). There has also been a growing interest in building a unified testing framework for both global and local measures

Correspondence: Sang-Il Lee, Department of Geography Education, Seoul National University, 599 Gwanak-ro, Gwanak-gu, Seoul 151-748, South Korea
e-mail: si_lee@snu.ac.kr

Submitted: September 3, 2006. Revised version accepted: May 23, 2008.

(Tiefelsdorf and Boots 1997; Tiefelsdorf 1998, 2000; Boots and Tiefelsdorf 2000; Ord and Getis 2001; Leung, Mei, and Zhang 2003).

Yet there remain many things to be done. First of all, it needs to be emphasized that little attention has been given to the distributional properties of *bivariate* measures of spatial association. Although a significance testing method for Cross-Moran (Wartenberg 1985) was proposed in the early 1990s (Czaplewski and Reich 1993; Reich, Czaplewski, and Bechtold 1994), it has little been known to geographers. Recently, Lee's L as a new bivariate measure of spatial association has been developed (Lee 2001a) and a significance testing method for the measure has been proposed to provide the equations for the first two moments under the randomization assumption (Lee 2004). The testing method can be regarded as a significant progress because it is derived from a unified framework termed *the extended Mantel test*, which can be applied to any global measures of spatial association, whether univariate or bivariate, with any form of spatial weights matrix, whether one with zero diagonal or nonzero diagonal. This kind of generalized approach has never been undertaken for bivariate local measures with few exceptions; for example, Tiefelsdorf (unpublished data) provides a valuable discussion on how to specify the spatial association between two residual vectors and proposes a set of new bivariate measures, global C and local C_i, along with a significance testing method. Yet we still need a generalized approach that inferential tests for different local measures, not only univariate but bivariate, are commonly predicated on.

The main objective of this article is to present a generalized randomization approach to local measures of spatial association in order to provide a reliable foundation that significance tests for the local measures are commonly based on. The generalized randomization approach proposed in this article is composed of two testing procedures, the *extended Mantel test* initially proposed for global measures of spatial association (Lee 2004) and the *generalized vector randomization test* presented by Hubert (1984, 1987), similar to one utilized by Sokal, Oden, and Thomson (1998). These two procedures are then assigned to two different randomization assumptions, total randomization and conditional randomization (Anselin 1995; Sokal, Oden, and Thomson 1998), in consideration of the different natures of various local measures. By properly specifying necessary matrices and vectors for a particular local measure of spatial association under a particular randomization assumption, the generalized randomization approach as a whole is expected to yield a reliable set of equations for the expected value and variance by taking into account all possible permutational situations occurring at and around a location.

Subsequently, I first conceptualize differences between total randomization and conditional randomization and demonstrate that the different randomization assumptions lead to different general forms of local measures resulting in different distributional properties. Second, for each randomization assumption, equations for the expected values and variances are presented for five different local measures of spatial association. Third, I will demonstrate the exactness of the proposed methods by conducting a simulation based on random permutations with a hypothetical data. Finally, I discuss some limitations of the randomization approach presented in this article.

A generalized randomization approach

Cliff and Ord (1981) contend that the randomization approach is preferable either (i) when we consider all possible permutations with a given data set or (ii) for any nonnormal population. The first argument indicates that the distributional properties are derived under a nonfree sampling situation where there is assumed to be no association between spatial locations and their numerical values. The second issue is more crucial because the variance computed under the set of random permutations provides an unbiased estimator for the variance of a statistic for any underlying distribution (Cliff and Ord 1981, p. 42). In this sense, the extended Mantel test (Lee 2004) provides a general foundation that global spatial association measures are commonly predicated on. These situations are more complicated for *local* measures of spatial association. Sokal, Oden, and Thomson (1998) argue that the Mantel test cannot apply to local measures when what they call the ''total randomization'' is intended, because different matrix settings for the Mantel test yield the same sets of local statistics with different reference distributions. Subsequently, I will demonstrate that those different matrix settings are based on different assumptions of randomization such that it is not unusual at all that different sampling distributions are obtained.

Conceptualizing different randomization assumptions

A general form of *global* measures of spatial association has been defined as follows (Anselin 1995; Lee 2004):

$$\Gamma = \sum_i \sum_j p_{ij} q_{ij} = \sum_{i,j} (\mathbf{P} \circ \mathbf{Q}) = \mathrm{tr}(\mathbf{P}^T\mathbf{Q}) = \mathrm{tr}(\mathbf{P}\mathbf{Q}^T) \tag{1}$$

where p_{ij} and q_{ij} are elements, respectively, in a *global spatial proximity matrix* **P** and a *global numeric similarity matrix* **Q**, and $\mathbf{P} \circ \mathbf{Q}$ denotes pairwise products between the two matrices. Accordingly, Anselin (1995, p. 98) formulates a general form of local measures of spatial association as

$$\Gamma_i = \sum_j p_{ij} q_{ij} \tag{2}$$

Now, a local measure of spatial association at an *i*th location is computed by summing up all the pairwise products between two *vectors* derived from the two matrices for its global counterparts.

However, this specification is somewhat misleading. As Sokal, Oden, and Thomson (1998) correctly point out, his specification leads to three different general forms for local spatial association:

$$\Gamma_i = \sum_{i,j} \left(\mathbf{P}^{(i)} \circ \mathbf{Q} \right) \tag{3}$$

$$\Gamma_i = \sum_{i,j} \left(\mathbf{P} \circ \mathbf{Q}^{(i)} \right) \tag{4}$$

$$\Gamma_i = \sum_{i,j} \left(\mathbf{P}^{(i)} \circ \mathbf{Q}^{(i)} \right) \tag{5}$$

where $\mathbf{P}^{(i)}$ and $\mathbf{Q}^{(i)}$ are certain forms of *local* spatial proximity matrix and *local* numeric similarity matrix for a location. These three different definitions yield an identical set of local measures but with different sampling distributions. From this observation, Sokal, Oden, and Thomson (1998) incorrectly conclude that ''when we permit only conditional permutations, all three Mantel versions of the LISA will give rise to the same (conditional) reference distribution'' (p. 335) and ''there is not a unique total reference distribution for a LISA, hence no unique set of total moments'' (p. 335). Conceptually, the three different specifications are associated with three different randomization assumptions, and thus it is not unusual to observe that they yield different sets of distributional properties.

First, one specified in (3) is related to what may be called the ''location-based total randomization'' that is identical to what Sokal, Oden, and Thomson (1998) calls the ''total randomization,'' where a local spatial configuration consisting of an *i*th *location* and its neighbors is fixed and different sets of values are permuted over there. A value being set to the *i*th location, all the other values are permuted to define a set of neighbors with resulting in a set of local statistics, and then another value is set to the *i*th location and the same procedure is undertaken, yielding another set of local statistics. In this randomization, the expected value is the average value of all possible local measures that can be given to the *i*th location. Subsequently, the term total randomization will be used to refer to this location-based total randomization.

Second, one specified in (4) is related to what may be called the ''value-based total randomization,'' where a numeric vector consisting of *n* observations are permuted over the given spatial configuration with an *i*th value always being positioned at the reference spot. An *i*th value being set to a location, all the other values are permuted to define a set of neighbors, and then the *i*th value moves on to another location and the same procedure is repeated. In this type of randomization, the expected value is the average value of all possible local measures that can be given to the *i*th value. Obviously, this randomization is deemed to be of little value for spatial statistics, because locations are lost.

Third, one specified in equation (5), in a conceptual sense, conforms to what has been called the ''conditional randomization'' (Anselin 1995; Sokal, Oden, and Thomson 1998), where an *i*th value is set on the corresponding *i*th location and all the other values are permuted to constitute a set of its neighbors over the given local spatial configuration. This more restrictive randomization scheme gives rise to another issue. The local measure may not have to be defined by matrices for certain measures. Rather, it is better seen as the sum of the pairwise products between two vectors that are drawn from the corresponding global matrices as seen in (2) (also see Sokal, Oden, and Thomson 1998). In other words, a local measure can be given as $\sum_i \left(\mathbf{p}^{(i)} \circ \mathbf{q}^{(i)}\right)$, rather than one in (5) (here $\mathbf{p}^{(i)}$ and $\mathbf{q}^{(i)}$ are certain forms of local vectors derived from their matrix counterparts, $\mathbf{P}^{(i)}$ and $\mathbf{Q}^{(i)}$). This implies that local measures of spatial association may be classified into two categories in terms of how to define a general form of local measures under the conditional randomization.

A *bivariate* situation requires a further clarification, because two elements in a pair should be specifically defined along with their relationships with spatial configurations. There might be three ways of defining the relationships among the three elements (two variables and a spatial setting). One way is first to assign a value in X to a location and permute other values to define its neighbors. For each local setting in X, values in Y are randomly permuted to define a local spatial setting in the other side. As in the univariate situation, then, another value in X is set to the location and all the permutation procedure is repeated. In this case, the link between x_i and y_i in the original variables does not have to be maintained. Second, the ith values in X and Y are bound to each other and are assigned to the original ith location. Their neighbors are then defined by permuting all the other values on each side. In this case, elements in a pair of neighbors, x_j and y_j, need not follow the geographical reference in the original vectors. In other words, once a pivot value at a location are determined from the original order, permutations for neighbors are independently conducted between two variables.

The third way is to undertake the permutation procedure with all the values in X and Y being bound to each other according to the original order. For example, if x_j is chosen as x_i's neighbor in a given location, the corresponding y_j should be placed on that location as y_i's neighbor. Once the permutation is done for the ith pair, another pair of values is assigned to that location and all the permutation procedure is repeated. Among the three specifications, only the third way meets the randomization principles as can be seen from a Monte Carlo simulation conducted for Lee's L (Lee 2001a). This can be seen as a bivariate version of the total randomization. The conditional randomization for bivariate situations can be easily conceptualized. A permutation is conducted only on a location, not furthering onto other locations. That is, an ith pair, whose elements are bound to each other, is fixed at its original location, and all the other pairs, whose elements are also bound to each other, are permuted to define two sets of neighbors around the location.

Table 1 lists the five local measures of spatial association for which I will attempt to derive the first two moments based on the generalized randomization approach. Here, $\mathbf{z}_X$ and $\mathbf{z}_Y$ are the standardized vectors of variables X and Y (elements are subtracted by a mean and divided by a population standard deviation). The equation for local Geary's c_i is modified from Anselin's original

Table 1 Five Local Measures of Spatial Association

Dimensions	Measures	Summation notation	Matrix notation
Univariate	Local Moran's I_i	$\frac{n^2}{\sum_i \sum_j v_{ij}} \frac{\sum_j v_{ij}(x_i-\bar{x})(x_j-\bar{x})}{\sum_i (x_i-\bar{x})^2}$	$n\frac{(\mathbf{z}_X)^T \mathbf{V}_i \mathbf{z}_X}{\mathbf{1}^T \mathbf{V} \mathbf{1}}$
	Local Geary's c_i	$\frac{n(n-1)}{2\sum_i \sum_j v_{ij}} \frac{\sum_j v_{ij}(x_i-x_j)^2}{\sum_i (x_i-\bar{x})^2}$	$\frac{n-1}{2}\frac{(\mathbf{z}_X)^T [\Omega_i-(\mathbf{V}_i+\mathbf{V}_i^T)]\mathbf{z}_X}{\mathbf{1}^T \mathbf{V} \mathbf{1}}$
	Local Lee's S_i	$\frac{n^2}{\sum_i \left(\sum_j v_{ij}\right)^2} \frac{\left(\sum_j v_{ij}(x_j-\bar{x})\right)^2}{\sum_i (x_i-\bar{x})^2}$	$n\frac{(\mathbf{z}_X)^T (\mathbf{V}_i^T \mathbf{V}_i)\mathbf{z}_X}{\mathbf{1}^T (\mathbf{V}^T \mathbf{V})\mathbf{1}}$
Bivariate	Local Cross-Moran	$\frac{n^2}{\sum_i \sum_j v_{ij}} \frac{\sum_j v_{ij}(x_i-\bar{x})(y_j-\bar{y})}{\sqrt{\sum_i (x_i-\bar{x})^2}\sqrt{\sum_i (y_i-\bar{y})^2}}$	$n\frac{(\mathbf{z}_X)^T \mathbf{V}_i \mathbf{z}_Y}{\mathbf{1}^T \mathbf{V} \mathbf{1}}$
	Local Lee's L_i	$\frac{n^2}{\sum_i \left(\sum_j v_{ij}\right)^2} \frac{\left(\sum_j v_{ij}(x_j-\bar{x})\right)\left(\sum_j v_{ij}(y_j-\bar{y})\right)}{\sqrt{\sum_i (x_i-\bar{x})^2}\sqrt{\sum_i (y_i-\bar{y})^2}}$	$n\frac{(\mathbf{z}_X)^T (\mathbf{V}_i^T \mathbf{V}_i)\mathbf{z}_Y}{\mathbf{1}^T (\mathbf{V}^T \mathbf{V})\mathbf{1}}$

NOTE: All theses definitions satisfy the additivity requirement that the average value of local measures is equal to the corresponding global measure as presented in (6).

one in order to conform to a more restrictive additivity requirement that the average value of local measures is equal to the global measure as expressed by

$$\Gamma = \frac{\sum_i \Gamma_i}{n} \tag{6}$$

The matrix notation for local Geary's c_i was derived by Lee (2001b) for a quadratic form that will be elaborated. The local Cross-Moran equation is derived from Wartenberg's (1985) global measure, and local Lee's S_i and L_i follow Lee's definition (Lee 2001a). Getis-Ord's G statistic is not examined here partially because it does not have a corresponding global measure satisfying the additivity requirement and partially because local Lee's S_i is almost identical to a modified version of Getis-Ord G_i^* (Leung, Mei, and Zhang 2003, p. 730). In addition, one might easily recognize that L_i is the bivariate extension of S_i as local Cross-Moran is of local Moran's I_i.

The extended Mantel test and the generalized vector randomization test

Lee (2004) demonstrates that all the global measures of spatial association, univariate or bivariate, belong to the general class of the Mantel statistic and shows that his significance testing method (1967) and its extension (Heo and Gabriel 1998) provide a set of equations for the first two moments for all the measures with any kinds of spatial weights matrices, zero diagonal or not. This extended Mantel test is easily applicable to local measures as far as they conform to the general form of the Mantel statistic that is expressed a sum of pairwise products between two matrices, one associated with spatial proximity among locations and the other with numeric similarity among values at the locations.

If a spatial association measure is defined as the sum of pairwise products between two vectors rather than matrices, the extended Mantel test would be redundant. A much simpler randomization test is enough, which might be called the *generalized vector randomization test*. Hubert (1984, p. 453; 1987, p. 28) once provided the equations for expected values and variances.

These two methods will collectively provide a general foundation on which the distributional properties for all the local measures of spatial association are obtained when **P**, **Q**, **p**, and **q** for each measure are properly specified depending on a randomization assumption involved, total or conditional. The extended Mantel test is utilized for all the measures under the total randomization assumption, while the generalized vector randomization test is exploited under the conditional randomization, but only for three measures, local Moran I_i, local Geary c_i, and local Cross-Moran. Lee's measures need the extended Mantel test for both the assumptions, because none of the measures can be defined as a measure of vector comparison. This also implies that the general procedures should be tailorized when applied to a particular measure, which will be exhaustively discussed in the next section.

In summary, it was acknowledged in this section that, if a measure is defined as a sum of pairwise products either between two matrices or between two vectors under a particular randomization assumption, each of the two testing procedures, the extended Mantel test and the generalized vector randomization test, can be utilized to obtain the expectation and variance of the measure. In the next section, we will see how to formulate a local measure with relevant matrices and vectors under a particular randomization assumption and how to articulate the formulation to fit it into the specific situation of permutation dictated by a particular randomization assumption.

An application of the generalized randomization test to local measures of spatial association

Defining local matrices and vectors

It is necessary to define some local matrices and vectors in order subsequently to define general forms of spatial association measures based on the total and conditional randomization assumptions. A local measure of spatial association should satisfy either:

$$\Gamma_i = \sum_{i,j} \left(\mathbf{P}^{(i)} \circ \mathbf{Q} \right) = \sum_{i} \left(\mathbf{p}^{(i)} \circ \mathbf{q}^{(i)} \right) \text{ or} \tag{7}$$

$$\Gamma_i = \sum_{i,j} \left(\mathbf{P}^{(i)} \circ \mathbf{Q} \right) = \sum_{i,j} \left(\mathbf{P}^{(i)} \circ \mathbf{Q}^{(i)} \right) \tag{8}$$

where $\mathbf{P}^{(i)}$ is a *local spatial proximity matrix* at an ith location, $\mathbf{Q}$ is a global numeric similarity matrix among values at the locations, $\mathbf{p}^{(i)}$ and $\mathbf{q}^{(i)}$ are local vectors for the ith location derived from $\mathbf{P}^{(i)}$ and $\mathbf{Q}$, and $\circ$ denotes pairwise products between the two matrices or vectors. These two categories possess the common definition for the total randomization as seen in (3), but are differentiated in formulating a general form of local measures for the conditional randomization. As mentioned above, (7) works for local Moran's I_i, local Geary's c_i, and local Cross-Moran, while (8) does for Lee's S_i and L_i.

Tiefelsdorf (1998) defines a local spatial weights matrix in a symmetric star-shaped form. However, I here define a local spatial weights matrix $\mathbf{V}_i$ by assigning zeros to all the elements except for ones on an ith row in a global spatial weights matrix $\mathbf{V}$.

$$\mathbf{V}_i = \begin{bmatrix} & & \mathbf{0} & & \\ v_{i1} & \cdots & v_{ii} & \cdots & v_{in} \\ & & \mathbf{0} & & \end{bmatrix} \tag{9}$$

This nonsymmetric form of a local spatial weights matrix is preferred not only because the symmetricity required for the exact distribution approach can be preserved by a transformation function, $\mathbf{V}_i^S = \frac{1}{2}\left(\mathbf{V}_i + \mathbf{V}_i^T\right)$, but because the symmetric form does not work for bivariate measures. As a global spatial proximity matrix $\mathbf{P}$ is a normalized form of a global spatial weights matrix $\mathbf{V}$, a local spatial proximity matrix $\mathbf{P}^{(i)}$ for local Moran's I_i and local Cross-Moran is defined as a normalized form of $\mathbf{V}_i$ multiplied by n, and the associated spatial proximity vector, $\mathbf{p}^{(i)}$, is accordingly defined (refer to Table 2). It should be noted that the number of observations (n) should be multiplied for local measures in order to satisfy the more restrictive additivity requirement defined in (6).

The $\mathbf{P}^{(i)}$s for local Geary's c_i, Lee's S_i, and L_i, however, take more complicated forms. The matrix for local Geary's c_i is given as

$$\begin{aligned}\mathbf{P}^{(i)} &= \frac{n-1}{2\mathbf{1}^T\mathbf{V}\mathbf{1}}\left[\Omega_i - \left(\mathbf{V}_i + \mathbf{V}_i^T\right)\right] \\ &= \frac{n-1}{2\mathbf{1}^T\mathbf{V}\mathbf{1}}\begin{bmatrix} v_{i1} & \mathbf{0} & -v_{i1} & & \\ \mathbf{0} & \ddots & \vdots & \mathbf{0} & \\ -v_{i1} & \cdots & \sum_j v_{ij} - v_{ii} & \cdots & -v_{in} \\ & \mathbf{0} & \vdots & \ddots & \mathbf{0} \\ & & -v_{in} & \mathbf{0} & v_{in} \end{bmatrix}\end{aligned} \tag{10}$$

where $\mathbf{\Omega}_i$ is a diagonal matrix of $\{v_{i1}, \cdots, v_{ii}, \cdots v_{in}\}$ with v_{ii} being added by $\sum_j v_{ij}$, a row-sum at each location (Lee 2001b; also see Leung, Mei, and Zhang 2003). Unlike in local Moran's I_i and local Cross-Moran, the local spatial proximity vector,

Table 2 The Definition of Matrices and Vectors for Local Measures of Spatial Association

	Local Moran's I_i	Local Geary's c_i	Local Lee's S_i	Local Cross-Moran	Local Lee's L_i
$\mathbf{P}$	$\frac{\mathbf{V}}{\mathbf{1}^T\mathbf{V}\mathbf{1}}$	$\frac{n-1}{n}\frac{(\Omega-\mathbf{V})}{\mathbf{1}^T\mathbf{V}\mathbf{1}}$	$\frac{\mathbf{V}^T\mathbf{V}}{\mathbf{1}^T(\mathbf{V}^T\mathbf{V})\mathbf{1}}$	$\frac{\mathbf{V}}{\mathbf{1}^T\mathbf{V}\mathbf{1}}$	$\frac{\mathbf{V}^T\mathbf{V}}{\mathbf{1}^T(\mathbf{V}^T\mathbf{V})\mathbf{1}}$
$\mathbf{P}^{(i)}$	$n\frac{\mathbf{V}_i}{\mathbf{1}^T\mathbf{V}\mathbf{1}}$	$\frac{n-1}{2}\frac{\left[\Omega_i-\left(\mathbf{V}_i+\mathbf{V}_i^T\right)\right]}{\mathbf{1}^T\mathbf{V}\mathbf{1}}$	$n\frac{\mathbf{V}_i^T\mathbf{V}_i}{\mathbf{1}^T(\mathbf{V}^T\mathbf{V})\mathbf{1}}$	$n\frac{\mathbf{V}_i}{\mathbf{1}^T\mathbf{V}\mathbf{1}}$	$n\frac{\mathbf{V}_i^T\mathbf{V}_i}{\mathbf{1}^T(\mathbf{V}^T\mathbf{V})\mathbf{1}}$
$\mathbf{p}^{(i)}$	$\frac{n}{\mathbf{1}^T\mathbf{V}\mathbf{1}}(v_{i1},\cdots,v_{ii},\cdots,v_{in})^T$	$\frac{n-1}{2\mathbf{1}^T\mathbf{V}\mathbf{1}}(v_{i1},\cdots,-(v_{i.}-v_{ii}),\cdots,v_{in})^T$	Not defined	$\frac{n}{\mathbf{1}^T\mathbf{V}\mathbf{1}}(v_{i1},\cdots,v_{ii},\cdots,v_{in})^T$	Not defined
$\mathbf{Q}$	$\mathbf{z}_X(\mathbf{z}_X)^T$	$\mathbf{z}_X(\mathbf{z}_X)^T$	$\mathbf{z}_X(\mathbf{z}_X)^T$	$\mathbf{z}_X(\mathbf{z}_Y)^T$	$\mathbf{z}_X(\mathbf{z}_y)^T$
$\mathbf{Q}^{(i)}$	Not defined	Not defined	See (12)	Not defined	See (12)
$\mathbf{q}^{(i)}$	$z_{Xi}\mathbf{z}_X$	$\mathbf{z}_X^{(2)} - 2\ z_{Xi}\ \mathbf{z}_X$	Not defined	$z_{Xi}\ \mathbf{z}_Y$	Not defined

NOTE: $\mathbf{z}_X^{(2)} = \left(z_{Xii}^2,\cdots,z_{Xi}^2,\cdots,z_{Xn}^2\right)^T$ and, for the definition of the other quantities used in this table, refer to Appendix A.

$\mathbf{p}^{(i)}$, for local Geary's c_i is not defined in a successive manner. The specification shown in Table 2 first takes the diagonal element of $\mathbf{P}^{(i)}$ and then $p_i^{(i)}$ is multiplied by -1, making the sums of all the elements in each of $\mathbf{P}^{(i)}$ and $\mathbf{p}^{(i)}$ become equal. This is necessary for obtaining the distributional properties of the measure under the conditional randomization using the generalized vector randomization test.

The $\mathbf{P}^{(i)}$ matrix for S_i and L_i is given as

$$\mathbf{P}^{(i)} = \frac{n}{\mathbf{1}^T\left(\mathbf{V}^T\mathbf{V}\right)\mathbf{1}}\left[\mathbf{V}_i^T\mathbf{V}_i\right] = \frac{n}{\mathbf{1}^T\left(\mathbf{V}^T\mathbf{V}\right)\mathbf{1}}\begin{bmatrix} v_{i1}^2 & \cdots & v_{i1}v_{ii} & \cdots & v_{i1}v_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ v_{ii}v_{i1} & \cdots & v_{ii}^2 & \cdots & v_{ii}v_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ v_{in}v_{i1} & \cdots & v_{in}v_{ii} & \cdots & v_{in}^2 \end{bmatrix} \tag{11}$$

Unlike other measures, S_i and L_i do not have a local spatial proximity vector, thus working with (8), because there is no way to define the measures as sums of pairwise products between two vectors to satisfy (7). This implies that the extended Mantel test should be utilized for both assumptions of randomization for the measures.

Table 2 also shows that the $\mathbf{Q}$ matrix is defined for univariate and bivariate measures in the same way as in global measures (see Lee 2004, p. 1694, table 4). The local numeric similarity vector at an ith location, $\mathbf{q}^{(i)}$, for local Moran's I_i and local Cross-Moran is given, respectively, by $z_{Xi}\mathbf{z}_X$ and $z_{Xi}\mathbf{z}_Y$ where z_{Xi} is the ith element in a standardized form of variable X. It is also noted from Table 2 that the vector of $\mathbf{q}^{(i)}$ for local Geary's c_i again takes a more complicated form in order to satisfy (7).

Unlike the other measures, Lee's measures have a local numeric similarity *matrix*, $\mathbf{Q}^{(i)}$, to satisfy (8):

$$\mathbf{Q}^{(i)} = \begin{bmatrix} (q_{1,1}+q_{1,i}+q_{i,1}) & \cdots & q_{1,i-1} & 0 & q_{1,i+1} & \cdots & q_{1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{i-1,1} & \cdots & (q_{i-1,i-1}+q_{i-1,i}+q_{i,i-1}) & 0 & q_{i-1,i+1} & \cdots & q_{i-1,n} \\ 0 & \cdots & 0 & q_{i,i} & 0 & \cdots & 0 \\ q_{i+1,1} & \cdots & q_{i+1,i-1} & 0 & (q_{i+1,i+1}+q_{i+1,i}+q_{i,i+1}) & \cdots & q_{i+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ q_{n,1} & \cdots & q_{n,i-1} & 0 & q_{n,i+1} & \cdots & (q_{n,n}+q_{n,i}+q_{i,n}) \end{bmatrix} \quad (12)$$

This matrix is constructed by (i) replacing all the elements in the ith row and column of the global $\mathbf{Q}$ matrix with zeros except for q_{ii} and then (ii) moving the nullified values onto the diagonal for a summation. If a spatial weights matrix with a zero diagonal is concerned, the second part is not needed such that $\mathbf{Q}^{(i)}$ has the vector of $\{q_{11}, \cdots, q_{ii}, \cdots q_{nn}\}$ on its diagonal. This specification will be discussed in more detail where it is utilized for the conditional randomization of the measures.

The total randomization assumption

Following (7) and (8), the general form of local spatial association based on the total randomization assumption is given by

$$\Gamma_i = \sum_{i,j}\left(\mathbf{P}^{(i)} \circ \mathbf{Q}\right) = \mathrm{tr}\left(\mathbf{P}^{(i)^T}\mathbf{Q}\right) = \mathrm{tr}\left(\mathbf{P}^{(i)}\mathbf{Q}^T\right) \quad (13)$$

where $\mathbf{P}^{(i)}$ and $\mathbf{Q}$ for each local measure are seen in Table 2. Because the general form above is defined as some relationships between two matrices, the extended Mantel test presented by Lee (2004) applies to compute the expected value and variance. The difference between global and local measures in terms of the sampling distribution result only from differences between $\mathbf{P}$ and $\mathbf{P}^{(i)}$.

Table 3 lists some necessary quantities and the equations for the expected values for the measures. When the necessary quantities are computed, the expected value is obtained, as in global measures (Lee 2004, p. 1691), by

$$\begin{aligned} \mathrm{E}(\Gamma_i) &= \mathrm{E}\left(\Gamma_i^{\mathrm{off}}\right) + \mathrm{E}\left(\Gamma_i^{\mathrm{on}}\right) \\ &= \frac{\left[\left(\mathbf{1}^T\mathbf{P}^{(i)}\mathbf{1}\right) - \mathrm{tr}\left(\mathbf{P}^{(i)}\right)\right]\left[\left(\mathbf{1}^T\mathbf{Q}\mathbf{1}\right) - \mathrm{tr}(\mathbf{Q})\right]}{n(n-1)} + \frac{\mathrm{tr}\left(\mathbf{P}^{(i)}\right)\mathrm{tr}(\mathbf{Q})}{n} \end{aligned} \quad (14)$$

Regarding the computation of the quantities, there is one thing that should be cautiously acknowledged. In order to calculate necessary quantities for global measures, both matrices of $\mathbf{P}$ and $\mathbf{Q}$ should be symmetric (Lee 2004, pp. 1690–91). An asymmetric matrix (e.g., row-standardized spatial weights matrices) can be

Table 3 Expectations for Local Measures of Spatial Association Under the Total Randomization Assumption

	$\mathbf{P}^{(i)}$		$\mathbf{Q}$		Expectations		
Γ_i	$\mathbf{1}^T\mathbf{P}^{(i)}\mathbf{1}$	$\mathrm{tr}(\mathbf{P}^{(i)})$	$\mathbf{1}^T\mathbf{Q}\mathbf{1}$	$\mathrm{tr}(\mathbf{Q})$	$\mathrm{E}(\Gamma_i^{\mathrm{off}})$	$\mathrm{E}(\Gamma_i^{\mathrm{on}})$	$\mathrm{E}(\Gamma_i)$
Local Moran's I_i	$K_1 v_{i.}$	$K_1 v_{ii}$	0	n	$K_1 \frac{v_{ii}-v_{i.}}{n-1}$	$K_1 v_{ii}$	$K_1 \frac{nv_{ii}-v_{i.}}{n-1}$
Local Geary's c_i	0	$K_1 \frac{n-1}{n}(v_{i.}-v_{ii})$	0	n	$K_1 \frac{v_{i.}-v_{ii}}{n}$	$K_1 \frac{n-1}{n}(v_{i.}-v_{ii})$	$K_1(v_{i.}-v_{ii})$
Local Lee's S_i	$K_2 v_{i.}^2$	$K_2 v_{i.}^{(2)}$	0	n	$K_2 \frac{v_{i.}^{(2)}-v_{i.}^2}{n-1}$	$K_2 v_{i.}^{(2)}$	$K_2 \frac{nv_{i.}^{(2)}-v_{i.}^2}{n-1}$
Local Cross-Moran	$K_1 v_{i.}$	$K_1 v_{ii}$	0	nr_{XY}	$K_1 \frac{v_{ii}-v_{i.}}{n-1} r_{XY}$	$K_1 v_{ii} r_{XY}$	$K_1 \frac{nv_{ii}-v_{i.}}{n-1} r_{XY}$
Local Lee's L_i	$K_2 v_{i.}^2$	$K_2 v_{i.}^{(2)}$	0	nr_{XY}	$K_2 \frac{v_{i.}^{(2)}-v_{i.}^2}{n-1} r_{XY}$	$K_2 v_{i.}^{(2)} r_{XY}$	$K_2 \frac{nv_{i.}^{(2)}-v_{i.}^2}{n-1} r_{XY}$

NOTE: $K_1 = \frac{n}{\mathbf{1}^T\mathbf{V}\mathbf{1}} = n/\sum_i \sum_j v_{ij}$ and $K_2 = \frac{n}{\mathbf{1}^T(\mathbf{V}^T\mathbf{V})\mathbf{1}} = n/\sum_i \left(\sum_j v_{ij}\right)^2$. For the definition of the other quantities used in this table, refer to Appendix A.

rendered symmetric by an equation, $\frac{1}{2}(\mathbf{P}+\mathbf{P}^T)$. Although this requirement needs to be maintained for local measures so that both $\mathbf{P}^{(i)}$ and $\mathbf{Q}$ should be made symmetric, local Cross-Moran is exceptional such that its $\mathbf{Q}$ matrix should remain nonsymmetric in computing its necessary quantities.

The extended Mantel test can be utilized for any kinds of spatial weighting schemes, including nonzero diagonal ones. When a row-standardized matrix ($\mathbf{W}$) with a zero diagonal is concerned, the equations are more simplified. From Table 3, one can easily recognize that expected values are reduced to $-1/(n-1)$, 1, and $-r_{X,Y}/(n-1)$, respectively, for local Moran's I_i, local Geary's c_i, and local Cross-Moran and understand that they are identical to those for the corresponding global measures (see Lee 2004, p. 1696, table 4). The expected value equations for Lee's S_i and L_i are hardly simplified mainly because the measures may work better with a spatial weights matrix with a nonzero diagonal. As Lee (2004) explains, both measures are predicated on gauging a representative value for an overall focal set composed of a reference area and one's neighbors, rather than comparing the former with the latter; hence, it would be necessary to utilize a spatial weights matrix with a nonzero diagonal, such as $\mathbf{C}^*$ with 1s on the diagonal of a binary connectivity matrix $\mathbf{C}$ or $\mathbf{W}^*$ as a row-standardized version of $\mathbf{C}^*$ (see Lee 2004, p. 1696). When the number of neighbors of an ith location in $\mathbf{W}^*$ is denoted by n_i^*, the expected values for both measure will be reduced, respectively, to

$$\mathrm{E}(S_i)=\frac{n-n_i^*}{(n-1)n_i^*}\ \text{and}\ \mathrm{E}(L_i)=\frac{n-n_i^*}{(n-1)n_i^*}r_{XY} \tag{15}$$

Note that n_i^* is always one degree larger than n_i for $\mathbf{W}$ or $\mathbf{C}$, because a reference area oneself is regarded as a neighbor in $\mathbf{W}^*$.

The equations for variances are also obtained according to the general procedure provided by Lee (2004, p. 1691). Table 4 lists the variance equations for the three measures, and those for Lee's measures are given in Appendix A. The equations in Table 4 will be simplified by putting $v_{ii}=0$ when a zero-diagonal spatial weights matrix is concerned.

The conditional randomization assumption

The general form of local spatial association based on the conditional randomization is not universally defined. As already seen from the distinction between (7) and (8), there are two forms: one defined as the sum of pairwise products between two vectors (local Moran's I_i, local Geary's c_i, and local Cross-Moran) and the other as the sum of pairwise products between two matrices (Lee's S_i and L_i). Thus, it is obvious that the former utilizes the generalized vector randomization test whereas the latter should depend on the extended Mantel test again. It should be acknowledged that a reference location is dropped in a permutation process by the definition.

With reference to these permutation principles, the general form for local Moran's I_i, Geary'c_i, and Cross-Moran based on the conditional randomization assumption is given as

Table 4 Variances for Local Measures of Spatial Association Under the Total Randomization Assumption

Γ_i	Variances
Local Moran's I_i	$\frac{K_1^2}{(n-1)(n-2)}\left\{\begin{array}{l}\left[(n-1)\left(v_{i.}^{(2)}-v_{ii}^2\right)-(v_{i.}-v_{ii})^2\right]n-\left[n\left(v_{i.}^{(2)}-v_{ii}^2\right)-2(v_{i.}-v_{ii})^2\right]b_2\\+(n-2)v_{ii}[(n+1)v_{ii}-2v_{i.}](b_2-1)\end{array}\right\}$ $-\left(K_1\frac{v_{ii}-v_{i.}}{n-1}\right)^2$
Local Geary's c_i	$K_1^2\left\{\frac{\left[\left(v_{i.}^{(2)}-v_{ii}^2\right)+(v_{i.}-v_{ii})^2\right](n-1)(b_2+3)}{4n}\right\}-\left[K_1(v_{i.}-v_{ii})\right]^2$
Local Lee's S_i	Given in Appendix A
Local Cross-Moran	$\frac{K_1^2}{(n-1)(n-2)}\left\{\begin{array}{l}\left[(n-1)\left(v_{i.}^{(2)}-v_{ii}^2\right)-(v_{i.}-v_{ii})^2\right]n-\left[n\left(v_{i.}^{(2)}-v_{ii}^2\right)-2(v_{i.}-v_{ii})^2\right]b_1^{XY}r_{XY}^2\\+(n-2)v_{ii}[(n+1)v_{ii}-2v_{i.}]r_{XY}^2\left(b_1^{XY}-1\right)\end{array}\right\}$ $-\left(K_1\frac{v_{ii}-v_{i.}}{n-1}r_{XY}\right)^2$
Local Lee's L_i	Given in Appendix A

NOTE: $K_1=\frac{n}{\mathbf{1}^T\mathbf{V}\mathbf{1}}=n/\sum_i\sum_j v_{ij}$ and, for the definition of the other quantities used in this table, refer to Appendix A.

$$\begin{aligned}\Gamma_i&=\sum_i\left(\mathbf{p}^{(i)}\circ\mathbf{q}^{(i)}\right)\\&=p_i^{(i)}q_i^{(i)}+\sum_k\left(\mathbf{p}^{(-i)}\circ\mathbf{q}^{(-i)}\right)\end{aligned}\tag{16}$$

where $p_i^{(i)}$ and $q_i^{(i)}$ are the ith entries in $\mathbf{p}^{(i)}$ and $\mathbf{q}^{(i)}$ as defined in Table 2, and $\mathbf{p}^{(-i)}$ and $\mathbf{q}^{(-i)}$ are $(n-1)$-by-1 vectors derived from $\mathbf{p}^{(i)}$ and $\mathbf{q}^{(i)}$ by eliminating the ith elements. The specification in the lower part in (16) is needed to articulate the general form in the upper part to fit it into the simulation situation dictated by the conditional randomization assumption. Now, the sampling distribution of local measures is determined by a permutation between two vectors, $\mathbf{p}^{(-i)}$ and $\mathbf{q}^{(-i)}$. By slightly modifying the original equations provided by Hubert (1984, 1987), the expected value and variance for the local measures are given, respectively, by

$$\mathrm{E}(\Gamma_i)=p_i^{(i)}q_i^{(i)}+\frac{1}{n-1}\sum_k p_k^{(-i)}\sum_k q_k^{(-i)}\tag{17}$$

$$\mathrm{Var}(\Gamma_i)=\frac{1}{n-2}\sum_k\left(p_k^{(-i)}-\bar{p}^{(-i)}\right)^2\sum_k\left(q_k^{(-i)}-\bar{q}^{(-i)}\right)^2\tag{18}$$

where $p_k^{(-i)}$ and $q_k^{(-i)}$ are entries in $\mathbf{p}^{(-i)}$ and $\mathbf{q}^{(-i)}$, and $\bar{p}^{(-i)}$ and $\bar{q}^{(-i)}$ are the mean values of those vectors. Note that n in the original equations must be replaced by $(n-1)$.

Table 5 lists the equations for the expected values along with the necessary quantities for local Moran's I_i, Geary's c_i, and local Cross-Moran. If a row-standardized spatial weights matrix with zeros in its diagonal ($\mathbf{W}$) is concerned,

Table 5 Expectations for Local Moran's I_i, Local Geary's c_i, and Local Cross-Moran Under the Conditional Randomization Assumption

	p vector		**q** vector		
Γ_I	$p_i^{(i)}$	$\sum_k p_k^{(-i)}$	$q_i^{(i)}$	$\sum_k q_k^{(-i)}$	$\mathrm{E}(\Gamma_i)$
Local Moran's I_i	$K_1 v_{ii}$	$K_1(v_{i.} - v_{ii})$	z_{Xi}^2	$-z_{Xi}^2$	$K_1 \frac{nv_{ii}-v_{i.}}{n-1} z_{Xi}^2$
Local Geary's c_i	$-K_1 \frac{n-1}{2n}(v_{i.} - v_{ii})$	$K_1 \frac{n-1}{2n}(v_{i.} - v_{ii})$	$-z_{Xi}^2$	$n+z_{Xi}^2$	$K_1 \frac{v_{i.}-v_{ii}}{2}\left(1 + z_{Xi}^2\right)$
Local Cross-Moran	$K_1 v_{ii}$	$K_1(v_{i.} - v_{ii})$	$z_{Xi}z_{Yi}$	$-z_{Xi}z_{Yi}$	$K_1 \frac{nv_{ii}-v_{i.}}{n-1} z_{Xi}z_{Yi}$

NOTE: $K_1 = \frac{n}{\mathbf{1}^T\mathbf{V1}} = n/\sum_i \sum_j v_{ij}$ and, for the definition of the other quantities used in this table, refer to Appendix A.

the expected value for the measures are computed, respectively, by $-z_{Xi}^2/(n-1)$, $(z_{Xi}^2+1)/2$, and $-z_{Xi}^2 z_{Yi}^2/(n-1)$. Whereas the expected values under the total randomization are identical for all the locations with **W** (refer to Table 3), ones under the conditional randomization are dependent on the value on each location.

An inferential test for S_i and L_i based on the conditional randomization is much more complicated. With reference to the permutation principles described above, the general form for both measures is given by

$$\begin{aligned}\Gamma_i &= \sum_{i,j}\left(\mathbf{P}^{(i)} \circ \mathbf{Q}^{(i)}\right) \\ &= p_{ii}^{(i)} q_{ii}^{(i)} + \sum_{k,l}\left(\mathbf{P}^{(-i)} \circ \mathbf{Q}^{(-i)}\right)\end{aligned} \tag{19}$$

where $p_{ii}^{(i)}$ and $q_{ii}^{(i)}$ are diagonal entries in $\mathbf{P}^{(i)}$ and $\mathbf{Q}^{(i)}$ as given in Table 2, and $\mathbf{P}^{(-i)}$ and $\mathbf{Q}^{(-i)}$ are $(n-1)$-by-$(n-1)$ matrices obtained by eliminating the ith row and column from $\mathbf{P}^{(i)}$ and $\mathbf{Q}^{(i)}$. Here $\mathbf{Q}^{(-i)}$ needs a further elaboration. When the ith row and column of the matrix defined in (12) are eliminated, $\mathbf{Q}^{(-i)}$s for S_i and L_i are, respectively, given:

$$\mathbf{Q}^{(-i)} = \mathbf{z}_X^{(-i)}\left(\mathbf{z}_X^{(-i)}\right)^T + 2\mathrm{diag}\left(z_{Xi}\mathbf{z}_X^{(-i)}\right) \text{ and} \tag{20}$$

$$\mathbf{Q}^{(-i)} = \mathbf{z}_X^{(-i)}\left(\mathbf{z}_Y^{(-i)}\right)^T + \mathrm{diag}\left(z_{Xi}\mathbf{z}_Y^{(-i)}\right) + \mathrm{diag}\left(z_{Yi}\mathbf{z}_X^{(-i)}\right) \tag{21}$$

This specification is necessary to ensure that the ith elements themselves in X and Y are not involved in the permutation process but their associations with other values resulting from nonzero-diagonal elements in **V** are maintained in the permutation process. When a spatial weights matrix with a *zero* diagonal is concerned, (20) and (21) will be simplified, respectively, as $\mathbf{Q}^{(-i)} = \mathbf{z}_X^{(-i)}\left(\mathbf{z}_X^{(-i)}\right)^T$ and $\mathbf{Q}^{(-i)} = \mathbf{z}_X^{(-i)}\left(\mathbf{z}_Y^{(-i)}\right)^T$.

Because S_i and L_i based on the conditional randomization are defined by some relationships between two matrices, the extended Mantel test again applies to generate the equations for the expected value and variance. By slightly modifying (14), we have

$$\begin{aligned}\mathrm{E}(\Gamma_i) =& p_{ii}^{(i)} q_{ii}^{(i)} \\ &+ \frac{\left[\left(\mathbf{1}^T\mathbf{P}^{(-i)}\mathbf{1}\right) - \mathrm{tr}\left(\mathbf{P}^{(-i)}\right)\right]\left[\left(\mathbf{1}^T\mathbf{Q}^{(-i)}\mathbf{1}\right) - \mathrm{tr}\left(\mathbf{Q}^{(-i)}\right)\right]}{(n-1)(n-2)} \\ &+ \frac{\mathrm{tr}\left(\mathbf{P}^{(-i)}\right)\mathrm{tr}\left(\mathbf{Q}^{(-i)}\right)}{n-1}\end{aligned} \tag{22}$$

Note that n in (14) must be replaced by $(n-1)$ in the computation of the necessary quantities. It is shown that the overall expected value is divided into three parts: one for the reference area that is not involved in the permutation process, another for the off-diagonal elements, and the other for the on-diagonal. Table 6 lists the equations of the expected values for the measures along with the three parts

Table 6 Expectations for Local Lee's S_i and L_i Under the Conditional Randomization Assumption

Measures	Expectations
Local Lee's S_i	
$\mathrm{E}(S_i)$	$K_2 v_{ii}^2 z_{Xi}^2 + \frac{K_2}{(n-1)(n-2)} \left\{ \begin{array}{l} \left[(n-1)\left(v_{i.}^{(2)} - v_{ii}^2\right) - (v_{i.} - v_{ii})^2\right] n \\ -\left[(3n-4)\left(v_{i.}^{(2)} - v_{ii}^2\right) - 2(v_{i.} - v_{ii})^2\right] z_{Xi}^2 \end{array} \right\}$
Reference location	$K_2 v_{ii}^2 z_{Xi}^2$
$\mathrm{E}(S_i^{\mathrm{off}})$	$K_2 \frac{\left[\left(v_{i.}^{(2)} - v_{ii}^2\right) - (v_{i.} - v_{ii})^2\right]\left(n - 2z_{Xi}^2\right)}{(n-1)(n-2)}$
$\mathrm{E}(S_i^{\mathrm{on}})$	$K_2 \frac{\left(v_{i.}^{(2)} - v_{ii}^2\right)\left(n - 3z_{Xi}^2\right)}{n-1}$ (for a zero-diagonal, $K_2 \frac{v_{i.}^{(2)}\left(n - z_{Xi}^2\right)}{n-1}$)
Local Lee's L_i	
$\mathrm{E}(L_i)$	$K_2 v_{ii}^2 z_{Xi} z_{Yi} + \frac{K_2}{(n-1)(n-2)} \left\{ \begin{array}{l} \left[(n-1)\left(v_{i.}^{(2)} - v_{ii}^2\right) - (v_{i.} - v_{ii})^2\right] n r_{XY} \\ -\left[(3n-4)\left(v_{i.}^{(2)} - v_{ii}^2\right) - 2(v_{i.} - v_{ii})^2\right] z_{Xi} z_{Yi} \end{array} \right\}$
Reference location	$K_2 v_{ii}^2 z_{Xi} z_{Yi}$
$\mathrm{E}(L_i^{\mathrm{off}})$	$K_2 \frac{\left[\left(v_{i.}^{(2)} - v_{ii}^2\right) - (v_{i.} - v_{ii})^2\right](n r_{XY} - 2 z_{Xi} z_{Yi})}{(n-1)(n-2)}$
$\mathrm{E}(L_i^{\mathrm{on}})$	$K_2 \frac{\left(v_{i.}^{(2)} - v_{ii}^2\right)(n r_{XY} - 3 z_{Xi} z_{Yi})}{n-1}$ (for a zero-diagonal, $K_2 \frac{v_{i.}^{(2)}(n r_{XY} - z_{Xi} z_{Yi})}{n-1}$)

NOTE: $K_2 = \frac{n}{\mathbf{1}^T(\mathbf{V}^T\mathbf{V})\mathbf{1}} = n/\sum_i \left(\sum_j v_{ij}\right)^2$ and, for the definition of the other quantities used in this table, refer to Appendix A.

that will be needed to compute the variances in accordance with the equations given in Appendix A.

It should be acknowledged from Table 6 that the overall expectation for a *zero*-diagonal situation cannot be obtained by simply substituting zero for v_{ii} in the equations. That works for the reference location value and the expected value for off-diagonal elements, but not for the expected value for on-diagonal elements. This is simply because $\mathbf{Q}^{(-i)}$ matrix is differently defined between the nonzero and zero-diagonal situations as discussed above. Thus the equation for the expected value in a zero-diagonal situation can be generated by summing up $\mathrm{E}(\Gamma_i^{\mathrm{off}})$ with $v_{ii} = 0$ and $\mathrm{E}(\Gamma_i^{\mathrm{on}})$ presented in parentheses in Table 6 (the reference location value will be zero, so it drops out). When the number of neighbors of an *i*th observation in $\mathbf{W}^*$ is denoted by n_i^*, the expected values for S_i and L_i will be reduced, respectively, to

$$\mathrm{E}(S_i) = \frac{\left\{(n-n_i^*)(n_i^*-1)n + \left[(n-n_i^*)(n-2n_i^*)+2\right]z_{Xi}^2\right\}}{(n-1)(n-2)n_i^{*2}} \tag{23}$$

$$\mathrm{E}(L_i) = \frac{\left\{(n-n_i^*)(n_i^*-1)nr_{XY} + \left[(n-n_i^*)(n-2n_i^*)+2\right]z_{Xi}z_{Yi}\right\}}{(n-1)(n-2)\cdot n_i^{*2}} \tag{24}$$

Note again that n_i^* is always one degree larger than n_i for $\mathbf{W}$. A more comprehensive investigation on the distributional aspects of Lee's L_i will be undertaken elsewhere. The equations of variance for local Moran's I_i, local Geary's c_i, and local Cross-Moran are given in Table 7 and ones for S_i and L_i are given in Appendix A. Again, the equations in Table 7 will be simplified by substituting zero for v_{ii} in the equations when a zero-diagonal spatial weights matrix is concerned.

In summary, it was acknowledged in this section: (i) under the total randomization assumption, all five measures are defined as the sum of pairwise products

Table 7 Variances for Local Measures of Spatial Association Under the Conditional Randomization Assumption

Γ_i	Variances
Local Moran's I_i	$K_1^2 \frac{n}{(n-1)^2(n-2)}\left[(n-1)\left(v_{i.}^{(2)} - v_{ii}^2\right) - (v_{i.} - v_{ii})^2\right](n-1-z_{Xi}^2)z_{Xi}^2$
Local Geary's c_i	$\frac{K_1^2}{4n^2(n-2)}\left[(n-1)\left(v_{i.}^{(2)} - v_{ii}^2\right) - (v_{i.} - v_{ii})^2\right]$ $\left[(n-1)(nb_2 - 4nb_1 z_{Xi} + 4nz_{Xi}^2 - z_{Xi}^4) - (n + z_{Xi}^2)^2\right]$
Local Lee's S_i	Given in Appendix A
Local Cross-Moran	$K_1^2 \frac{n}{(n-1)^2(n-2)}\left[(n-1)\left(v_{i.}^{(2)} - v_{ii}^2\right) - (v_{i.} - v_{ii})^2\right](n-1-z_{Yi}^2)z_{Xi}^2$
Local Lee's L_i	Given in Appendix A

NOTE: $K_1 = \frac{n}{\mathbf{1}^T\mathbf{V}\mathbf{1}} = n/\sum_i\sum_j v_{ij}$ and, for the definition of the other quantities used in this table, refer to Appendix A.

between two matrices such that all of them utilize the extended Mantel test; (ii) there are two categories of the measures in terms of how to formulate them under the conditional randomization, one defined as the sum of pairwise products between two matrices and the other as sum of pairwise products between two vectors; (iii) the former still depends on the extended Mantel test, but the latter should utilize the generalized vector randomization test; and (iv) the testing procedures need to be tailorized in order to embrace the different nature of the measures and the particular situation of permutation. In the next section, I will investigate the validity of the testing procedure by conducting a Monte Carlo simulation.

An illustration

For an experiment, I designed two different spatial patterns on a hypothetical space that is composed of 37 hexagons (Fig. 1). The two spatial patterns have the same mean (1.730) and variance (0.576) and the Pearson's correlation coefficient between them is 0.530. Pattern A is used for univariate measures, and the relation between Patterns A and B is utilized for bivariate measures. I choose three hexagons labeled, respectively, a, b, and c, each of which has different linkage degrees (six neighbors for a, four for b, and three for c) and different values (three for a, two for b, and one for c). Two different spatial weights matrices are built; a binary

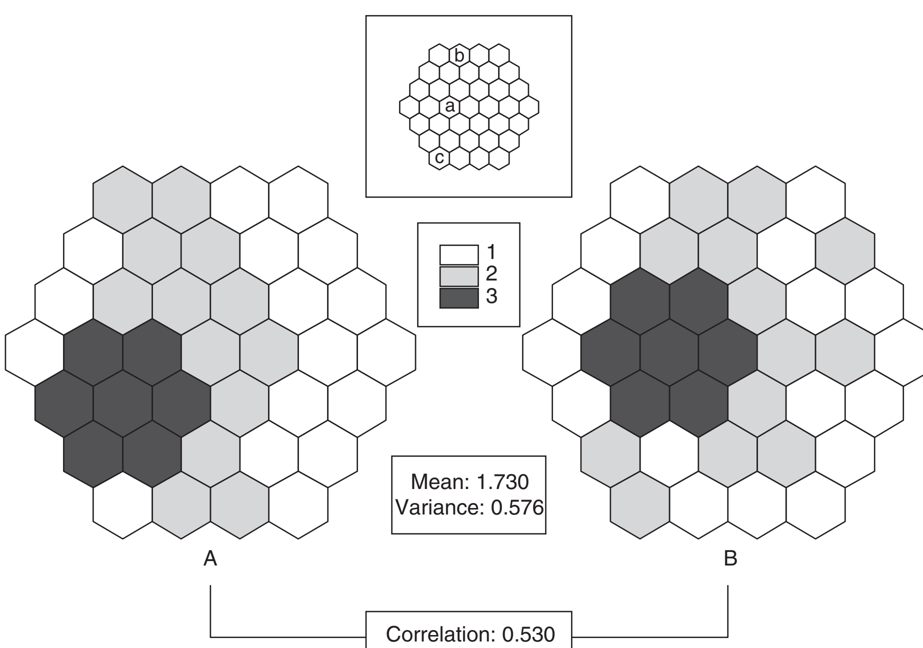

Figure 1. Hypothetical spatial patterns.

connectivity matrix **C** for local Moran's I_i, local Geary's c_i, local Cross-Moran, and a row-standardized version of $\mathbf{C}^*$, that is, $\mathbf{W}^*$ for Lee's S_i and L_i.

In order to investigate the exactness of the computation of the first two moments offered by the generalized randomization test for each measure, I conduct 10,000 random permutations for each of the three locations, whose results are shown in Table 8. The computation of expected values and variances based on the equations drawn from the proposed randomization tests appears highly reasonable for the first two moments when compared with the random permutation results.

Table 8 provides some additional information on other aspects of the distributional properties of the measures. First, for the total randomization, the magnitude of variances are positively related to local linkage degrees at locations with **C** while negatively related with $\mathbf{W}^*$, which is correspondent to findings by Tiefelsdorf, Griffith, and Boots (1999). Second, for the conditional randomization, there is no direct relationship between local linkage degrees and the magnitude of variances, mainly because z-standardized values on the reference locations are involved in the variance computation. A higher z-standardized value at a reference location, when squared, results in a higher variance for local Moran's I_i, Geary's c_i, and Cross-Moran. However, Lee's S_i and L_i still have a relatively high positive relationship between local linkage degrees and the magnitude of variances as in the total randomization.

Discussion and conclusions

This article shows that the generalized randomization approach including the extended Mantel test and the generalized vector randomization test succeeds in providing the first two moments for local measures of spatial association under different randomization assumptions. Specifically, it has been proven that: (i) the extended Mantel test can be applied to all of the local measures under the total randomization; (ii) it can also be applied to Lee's S_i and L_i under the conditional randomization; and (iii) the generalized vector randomization test can be used for local Moran's I_i, local Geary's c_i, and local Cross-Moran under the conditional randomization. The testing methods presented in this article have been proven in comparison with the simulation results not only to conform to the earlier findings for univariate local measures such as Moran's I_i and Geary's c_i (Anselin 1995; Sokal, Oden, and Thomson 1998), but successfully to be extended to a new univariate measure, S_i, and to bivariate local measures such as local Cross-Moran and L_i. The testing procedure presented in this article is also extendable to some newly devised measures such that it could provide a reliable platform of hypothesis testing for spatial analysts who come up with new measures but suffer from inability to offer the equations for the first two moments, being forced to depend on a certain form of simulation.

This approach, however, has some drawbacks. First, it does not solve the two problems that have been regarded as crucial for the significance testing for local

Table 8 Expectations and Variances of Local Measures of Spatial Association Under the Two Randomization Assumptions, Compared with the Results from 10,000 Random Permutations

			Randomization assumptions							
			Total randomization				Conditional randomization			
			Computation		Permutations		Computation		Permutations	
Local measures	Area ID	Value	$E(\Gamma_i)$	$Var(\Gamma_i)$	Mean	Variance	$E(\Gamma_i)$	$Var(\Gamma_i)$	Mean	Variance
Univariate										
Moran's I_i	a	2.0965	−0.0343	0.2127	−0.0318	0.2154	−0.0960	0.5773	−0.1005	0.5752
	b	0.0078	−0.0228	0.1510	−0.0266	0.1511	−0.0029	0.0201	−0.0019	0.0197
	c	−0.7325	−0.0171	0.1167	−0.0143	0.1198	−0.0158	0.1107	−0.0173	0.1113
Geary's c_i	a	0.5212	1.2333	0.5879	1.2443	0.5874	2.3454	0.4430	2.3482	0.4459
	b	0.1737	0.8222	0.3282	0.8183	0.3236	0.4633	0.0245	0.4649	0.0246
	c	1.5636	0.6167	0.2223	0.6116	0.2194	0.5936	0.1857	0.5961	0.1894
Lee's S_i	a	1.2308	0.1190	0.0258	0.1191	0.0261	0.1392	0.0379	0.1385	0.0365
	b	0.0086	0.1778	0.0557	0.1827	0.0595	0.1538	0.0439	0.1528	0.0433
	c	0.4703	0.2292	0.0897	0.2285	0.0905	0.2256	0.0693	0.2244	0.0683
Bivariate										
Cross-Moran$_i$	a	3.4574	−0.0182	0.2163	−0.0160	0.2163	−0.0960	0.5773	−0.0911	0.5783
	b	0.0078	−0.0121	0.1535	−0.0133	0.1527	−0.0029	0.0201	−0.0046	0.0199
	c	0.3099	−0.0091	0.1186	−0.0101	0.1194	0.0059	0.1132	0.0036	0.1113
Lee's L_i	a	1.8575	0.0631	0.0170	0.0613	0.0163	0.0886	0.0253	0.0893	0.0254
	b	0.0086	0.0943	0.0375	0.0917	0.0359	0.0832	0.0301	0.0843	0.0314
	c	−0.2076	0.1216	0.0619	0.1255	0.0665	0.0801	0.0377	0.0834	0.0391

NOTE: **C** applies to Moran, Geary, and Cross-Moran measures and $\mathbf{W}^*$ applies to Lee's measures.

measures: (i) the relationship between overall and individual α levels and (ii) the presence of global spatial dependence. The former problem results from the fact that an α level for an overall set should be lowered when it applies to its individuals (Getis and Ord 1992; Anselin 1995). Even though some procedures such as the Bonferroni bounds procedure have been proposed, this problem has never been solved yet. The second problem arises because significance-testing procedures for local spatial associations are indifferent to the global level of spatial dependence; the distributional properties of local statistics should change as the levels of global spatial dependence change (see Anselin 1995; Ord and Getis 1995, 2001; Tiefelsdorf 1998). To date, only the exact distribution approach can solve this problem but only for univariate measures (for local Moran's I_i, see Tiefelsdorf 1998, 2000). The two problems together dictate a restriction on the use of local measures that they should be used in an exploratory manner, not in a confirmatory manner (Sokal, Oden, and Thomson 1998).

The second drawback of the approach presented here is that it does not cope with nonnormality of sampling distributions as discussed for the extended Mantel test (Lee 2004). As Boots and Tiefelsdorf (2000) demonstrate, local measures are far from normally distributed regardless of the number of neighbors at a location especially due to extremely high kurtosis. Thus, devising a way of extracting higher moments is much more needed for local measures than global ones. Thus, the next step will be to extract easily understandable sets of equations for higher moments with reference to some works done by Siemiatycki (1978), Mielke (1979), and Hubert (1984, 1987). When skewness is solely known, a Pearson Type III (gamma) function can be utilized (Costanzo, Hubert, and Golledge 1983), and when both skewness and kurtosis are known, a beta distribution can be applied for a more reliable inferential test (Hepple 1998).

It has been suggested that the conditional randomization is superior to the total randomization (Anselin 1995; Sokal, Oden, and Thomson 1998). Anselin (1995, p. 112) contends that the conditional randomization approach could provide a reliable basis even in the presence of global spatial autocorrelation. Sokal, Oden, and Thomson (1998, p. 349) find that skewness and kurtosis are less pronounced in the conditional randomization than in the total randomization. Despite some advantages of the conditional randomization over the total randomization, however, it significantly erodes the nature of random variables by arbitrarily fixing a reference value on a particular location (from a discussion with Tiefelsdorf).

In spite of some drawbacks, the randomization approach as presented here should be appreciated in terms of its generality: (i) it is a distribution-free method, enabling to escape from the normal distribution assumption in a population; (ii) it applies to virtually all kinds of spatial association measures, whether global or local, or whether contiguity based or distance based; and (iii) it can handle spatial weights matrices with nonzero entries on their diagonals. Further, as long as we aim to gauge each location's relative contribution to the global spatial dependence with local measures and to obtain an exploratory device of examining spatial heterogeneity by utilizing them, the randomization approach as a way of giving probabilistic weights to local statistics may provide researchers with an initial and reliable guideline in a pattern detection process. This could set the pace for the ''local turn'' occurring in a wide variety of subfields in spatial data analysis (see Fotheringham and Brunsdon 1999).

Appendix A. Equations of variance elements for local Lee's S_i and L_i under the total and conditional randomization assumptions

For both measures, the usage of a spatial weights matrix with a nonzero diagonal is assumed. The equations for the expected values of off- and on-diagonal elements ($E(\Gamma_i^{off})$ and $E(\Gamma_i^{on})$) are referenced by Tables 3 and 6. For Lee's S_i, the following quantities are needed.

$$K_2 = n/\sum_i\left(\sum_j v_{ij}\right)^2,\ v_{i.} = \sum_j v_{ij},\ v_{i.}^{(2)} = \sum_j v_{ij}^2,\ v_{i.}^2 = \left(\sum_j v_{ij}\right)^2,$$

$$v_{i.}^{(3)} = \sum_j v_{ij}^3,\ v_{i.}^{(4)} = \sum_j v_{ij}^4,\ z_{Xi} = \frac{x_i - \bar{x}}{\sqrt{\sum_i (x_i-\bar{x})^2/n}},\ z_{Yi} = \frac{y_i - \bar{y}}{\sqrt{\sum_i (y_i-\bar{y})^2/n}},$$

$$b_1 = \frac{m_3}{(m_2)^{\frac{3}{2}}},\ b_2 = \frac{m_4}{m_2^2},\ m_2 = \frac{\sum_i (x_i-\bar{x})^2}{n},\ m_3 = \frac{\sum_i (x_i-\bar{x})^3}{n},\ m_4 = \frac{\sum_i (x_i-\bar{x})^4}{n}$$

For Lee's L_i, the following quantities are additionally needed.

$$r_{XY} = \frac{\sum_i (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_i (x_i-\bar{x})^2}\sqrt{\sum_i (y_i-\bar{y})^2}},\ b_1^{XY} = \frac{m_2^{XY}}{(m_1^{XY})^2},\ b_{12}^{XY} = \frac{m_{12}^{XY}}{(m_2^Y)^{\frac{1}{2}} m_1^{XY}},\ b_{21}^{XY} = \frac{m_{21}^{XY}}{(m_2^X)^{\frac{1}{2}} m_1^{XY}},$$

$$m_2^X = \frac{\sum_i (x_i-\bar{x})^2}{n},\ m_2^Y = \frac{\sum_i (y_i-\bar{y})^2}{n},\ m_{12}^{XY} = \frac{\sum_i (x_i-\bar{x})(y_i-\bar{y})^2}{n},\ m_{21}^{XY} = \frac{\sum_i (x_i-\bar{x})^2(y_i-\bar{y})}{n},$$

$$m_1^{XY} = \frac{\sum_i (x_i-\bar{x})(y_i-\bar{y})}{n},\ m_2^{XY} = \frac{\sum_i (x_i-\bar{x})^2(y_i-\bar{y})^2}{n}$$

With these quantities, the variance for each measure is computed by (Lee, 2004, p. 1690):

$$\mathrm{Var}(\Gamma_i) = \mathrm{Var}\left(\Gamma_i^{off}\right) + \mathrm{Var}\left(\Gamma_i^{on}\right) + 2\mathrm{Cov}\left(\Gamma_i^{off}, \Gamma_i^{on}\right)$$

1. Equations of variance quantities for local Lee's S_i under the total randomization assumption

$$\mathrm{Var}\left(S_i^{\mathrm{off}}\right) = K_2^2 \left\{ \frac{2\left[\left(v_{i.}^{(2)}\right)^2 - v_{i.}^{(4)}\right](n-b_2)}{n-1} - \frac{4\left[2v_{i.}^{(4)} - 2v_{i.}v_{i.}^{(3)} + \left(v_{i.}^2 - v_{i.}^{(2)}\right)v_{i.}^{(2)}\right](n-2b_2)}{(n-1)(n-2)} + \frac{3\left[v_{i.}^4 - 6v_{i.}^{(4)} + 8v_{i.}v_{i.}^{(3)} + 3\left(v_{i.}^{(2)} - 2v_{i.}^2\right)v_{i.}^{(2)}\right](n-2b_2)}{(n-1)(n-2)(n-3)} \right\} - \left[\mathrm{E}\left(S_i^{\mathrm{off}}\right)\right]^2$$

$$\mathrm{Var}\left(S_i^{\mathrm{on}}\right) = K_2^2 \left\{ \frac{\left[\left(v_{i.}^{(2)}\right)^2 - v_{i.}^{(4)}\right]n - \left[\left(v_{i.}^{(2)}\right)^2 - nv_{i.}^{(4)}\right]b_2}{n-1} \right\} - \left[\mathrm{E}\left(S_i^{\mathrm{on}}\right)\right]^2 = K_2^2 \frac{\left[v_{i.}^{(4)}n - \left(v_{i.}^{(2)}\right)^2\right](b_2-1)}{n-1}$$

$$\mathrm{Cov}\left(S_i^{\mathrm{off}}, S_i^{\mathrm{on}}\right) = K_2^2 \left\{ \frac{2\left(v_{i.}^{(4)} - v_{i.}v_{i.}^{(3)}\right)b_2}{n-1} - \frac{\left[\left(v_{i.}^2 - v_{i.}^{(2)}\right)v_{i.}^{(2)} + 2\left(v_{i.}^{(4)} - v_{i.}v_{i.}^{(3)}\right)\right](n-2b_2)}{(n-1)(n-2)} \right\} - \mathrm{E}\left(S_i^{\mathrm{off}}\right)\mathrm{E}\left(S_i^{\mathrm{on}}\right) = K_2^2 \frac{2\left[\left(v_{i.}^{(4)} - v_{i.}v_{i.}^{(3)}\right)n + \left(v_{i.}^2 - v_{i.}^{(2)}\right)v_{i.}^{(2)}\right](b_2-1)}{(n-1)(n-2)}$$

2. Equations of variance quantities for local Lee's S_i under the conditional randomization assumption

$$\mathrm{Var}\left(S_i^{\mathrm{off}}\right) = K_2^2 \left\{ \frac{2\left[\left(v_{i.}^{(2)} - v_{ii}^2\right)^2 - \left(v_{i.}^{(4)} - v_{ii}^4\right)\right]\left[(n-b_2)n - 2\left(n - z_{Xi}^2\right)z_{Xi}^2\right]}{(n-1)(n-2)} \right.$$

$$+ \frac{4\left\{\left[(v_{i.} - v_{ii})^2 - \left(v_{i.}^{(2)} - v_{ii}^2\right)\right]\left(v_{i.}^{(2)} - v_{ii}^2\right) - 2(v_{i.} - v_{ii})\left(v_{i.}^{(3)} - v_{ii}^3\right) + 2\left(v_{i.}^{(4)} - v_{ii}^4\right)\right\}\left[(2b_2 - n + 2b_1 z_{Xi})n + 3\left(n - 2z_{Xi}^2\right)z_{Xi}^2\right]}{(n-1)(n-2)(n-3)}$$

$$\left. + \frac{\left[(v_{i.} - v_{ii})^4 + 3\left(v_{i.}^{(2)} - v_{ii}^2\right)^2 - 6(v_{i.} - v_{ii})^2\left(v_{i.}^{(2)} - v_{ii}^2\right) - 6\left(v_{i.}^{(4)} - v_{ii}^4\right) + 8(v_{i.} - v_{ii})\left(v_{i.}^{(3)} - v_{ii}^3\right)\right]\left[(3n - 6b_2 - 8b_1 z_{Xi})n - 12\left(n - 2z_{Xi}^2\right)z_{Xi}^2\right]}{(n-1)(n-2)(n-3)(n-4)} \right\}$$

$$- \left[\mathrm{E}\left(S_i^{\mathrm{off}}\right)\right]^2$$

$$\mathrm{Var}\left(S_i^{\mathrm{on}}\right) = K_2^2 \left\{ \frac{\left(v_{i.}^{(4)} - v_{ii}^4\right)\left[(b_2 + 4b_1 z_{Xi})n + \left(4n - 9z_{Xi}^2\right)z_{Xi}^2\right]}{n-1} \right.$$

$$\left. + \frac{\left[\left(v_{i.}^{(2)} - v_{ii}^2\right)^2 - \left(v_{i.}^{(4)} - v_{ii}^4\right)\right]\left[(n - b_2 - 4b_1 z_{Xi})n - 2\left(5n - 9z_{Xi}^2\right)z_{Xi}^2\right]}{(n-1)(n-2)} \right\}$$

$$- \left[\mathrm{E}\left(S_i^{\mathrm{on}}\right)\right]^2$$

$$\mathrm{Cov}\left(S_i^{\mathrm{off}}, S_i^{\mathrm{on}}\right) =$$

$$K_2^2 \left\{ \frac{2\left[\left(v_{i.}^{(4)} - v_{ii}^4\right) - (v_{i.} - v_{ii})\left(v_{i.}^{(3)} - v_{ii}^3\right)\right]\left[(b_2 + 3b_1 z_{Xi})n + 2\left(n - 3z_{Xi}^2\right)z_{Xi}^2\right]}{(n-1)(n-2)} \right.$$

$$\left. + \frac{\left[\left(v_{i.}^{(2)} - v_{ii}^2\right)^2 - (v_{i.} - v_{ii})^2\left(v_{i.}^{(2)} - v_{ii}^2\right) - 2\left(v_{i.}^{(4)} - v_{ii}^4\right) + 2(v_{i.} - v_{ii})\left(v_{i.}^{(3)} - v_{ii}^3\right)\right]\left[(n - 2b_2 - 6b_1 z_{Xi})n - 9\left(n - 2z_{Xi}^2\right)z_{Xi}^2\right]}{(n-1)(n-2)(n-3)} \right\}$$

$$- \left[\mathrm{E}\left(S_i^{\mathrm{off}}\right)\mathrm{E}\left(S_i^{\mathrm{on}}\right)\right]$$

3. Equations of variance quantities for local Lee's L_i under the total randomization assumption

$$\mathrm{Var}\left(L_i^{\mathrm{off}}\right) = K_2^2 \left\{ \frac{\left[\left(v_{i.}^{(2)}\right)^2 - v_{i.}^{(4)}\right]\left[n - \left(2b_1^{XY} - n\right) r_{XY}^2\right]}{n-1} \right.$$

$$- \frac{2\left[2v_{i.}^{(4)} - 2v_{i.}v_{i.}^{(3)} + \left(v_{i.}^2 - v_{i.}^{(2)}\right)v_{i.}^{(2)}\right]\left[n - \left(4b_1^{XY} - n\right) r_{XY}^2\right]}{(n-1)(n-2)}$$

$$\left. + \frac{\left[v_{i.}^4 - 6v_{i.}^{(4)} + 8v_{i.}v_{i.}^{(3)} + 3\left(v_{i.}^{(2)} - 2v_{i.}^2\right)v_{i.}^{(2)}\right]\left[n - 2\left(3b_1^{XY} - n\right) r_{XY}^2\right]}{(n-1)(n-2)(n-3)} \right\}$$

$$- \left[\mathrm{E}\left(L_i^{\mathrm{off}}\right)\right]^2$$

$$\mathrm{Var}\left(L_i^{\mathrm{on}}\right) = K_2^2 \left\{ \frac{\left[\left(v_{i.}^{(2)}\right)^2 - v_{i.}^{(4)}\right] n - \left[\left(v_{i.}^{(2)}\right)^2 - n v_{i.}^{(4)}\right] b_1^{XY} r_{XY}^2}{n-1} \right\} - \left[\mathrm{E}\left(L_i^{\mathrm{on}}\right)\right]^2$$

$$= K_2^2 \frac{\left[v_{i.}^{(4)} n - \left(v_{i.}^{(2)}\right)^2\right] r_{XY}^2 \left(b_1^{XY} - 1\right)}{n-1}$$

$$\mathrm{Cov}\left(L_i^{\mathrm{off}}, L_i^{\mathrm{on}}\right) = K_2^2 \left\{ \frac{2\left(v_{i.}^{(4)} - v_{i.}v_{i.}^{(3)}\right) b_1^{XY} r_{XY}^2}{n-1} \right.$$

$$\left. - \frac{\left[\left(v_{i.}^2 - v_{i.}^{(2)}\right)v_{i.}^{(2)} + 2\left(v_{i.}^{(4)} - v_{i.}v_{i.}^{(3)}\right)\right]\left(n - 2b_1^{XY} r_{XY}^2\right)}{(n-1)(n-2)} \right\} - \mathrm{E}\left(L_i^{\mathrm{off}}\right)\mathrm{E}\left(L_i^{\mathrm{on}}\right)$$

$$= K^2 \frac{2\left[\left(v_{i.}^{(4)} - v_{i.}v_{i.}^{(3)}\right) n + \left(v_{i.}^2 - v_{i.}^{(2)}\right) v_{i.}^{(2)}\right] r_{XY}^2 \left(b_1^{XY} - 1\right)}{(n-1)(n-2)}$$

4. Equations of variance quantities for local Lee's L_i under the conditional randomization assumption

$$\mathrm{Var}\left(L_i^{\mathrm{off}}\right) = K_2^2 \left\{ \frac{\left[\left(v_{i.}^{(2)} - v_{ii}^2\right)^2 - \left(v_{i.}^{(4)} - v_{ii}^4\right)\right]\left[n^2 - \left(z_{Xi}^2 + z_{Yi}^2\right)n + 4z_{Xi}^2 z_{Yi}^2 + \left(nr_{XY} - 2b_1^{XY} r_{XY} - 2z_{Xi}z_{Yi}\right)nr_{XY}\right]}{(n-1)(n-2)} \right.$$

$$+ \frac{\left\{\left[(v_{i.} - v_{ii})^2 - \left(v_{i.}^{(2)} - v_{ii}^2\right)\right]\left(v_{i.}^{(2)} - v_{ii}^2\right) - 2(v_{i.} - v_{ii})\left(v_{i.}^{(3)} - v_{ii}^3\right) + 2\left(v_{i.}^{(4)} - v_{ii}^4\right)\right\} \times \left[3\left(z_{Xi}^2 + z_{Yi}^2\right)n - 2n^2 - 24z_{Xi}^2 z_{Yi}^2 + 2\left(4b_1^{XY} r_{XY} + 3z_{Xi}z_{Yi} - nr_{XY} + 2z_{Xi}b_{12}^{XY} + 2z_{Yi}b_{21}^{XY}\right)nr_{XY}\right]}{(n-1)(n-2)(n-3)}$$

$$\left. + \frac{\left[(v_{i.} - v_{ii})^4 + 3\left(v_{i.}^{(2)} - v_{ii}^2\right)^2 - 6(v_{i.} - v_{ii})^2\left(v_{i.}^{(2)} - v_{ii}^2\right) - 6\left(v_{i.}^{(4)} - v_{ii}^4\right) + 8(v_{i.} - v_{ii})\left(v_{i.}^{(3)} - v_{ii}^3\right)\right] \times \left[n^2 - 2\left(z_{Xi}^2 + z_{Yi}^2\right)n + 24z_{Xi}^2 z_{Yi}^2 + 2\left(nr_{XY} - 4z_{Xi}z_{Yi} - 3b_1^{XY} r_{XY} - 2z_{Xi}b_{12}^{XY} - 2z_{Yi}b_{21}^{XY}\right)nr_{XY}\right]}{(n-1)(n-2)(n-3)(n-4)} \right\}$$

$$- \left[\mathrm{E}\left(L_i^{\mathrm{off}}\right)\right]^2$$

$$\mathrm{Var}\left(L_i^{\mathrm{on}}\right) = K_2^2 \left\{ \frac{\left(v_{i.}^{(4)} - v_{ii}^4\right)\left[\left(z_{Xi}^2 + z_{Yi}^2\right)n - 9z_{Xi}^2 z_{Yi}^2 + \left(2z_{Xi}z_{Yi} + b_1^{XY} r_{XY} + 2z_{Xi}b_{12}^{XY} + 2z_{Yi}b_{21}^{XY}\right)nr_{XY}\right]}{n-1} \right.$$

$$\left. + \frac{\left[\left(v_{i.}^{(2)} - v_{ii}^2\right)^2 - \left(v_{i.}^{(4)} - v_{ii}^4\right)\right]\left[18z_{Xi}^2 z_{Yi}^2 - \left(z_{Xi}^2 + z_{Yi}^2\right)n - \left(8z_{Xi}z_{Yi} + b_1^{XY} r_{XY} + 2z_{Xi}b_{12}^{XY} + 2z_{Yi}b_{21}^{XY} - nr_{XY}\right)nr_{XY}\right]}{(n-1)(n-2)} \right\}$$

$$- \left[\mathrm{E}\left(L_i^{\mathrm{on}}\right)\right]^2$$

$$\operatorname{Cov}\left(L_i^{\text{off}}, L_i^{\text{on}}\right) = K_2^2 \left\{ \frac{\left[\left(v_{i.}^{(4)} - v_{ii}^4\right) - (v_{i.} - v_{ii})\left(v_{i.}^{(3)} - v_{ii}^3\right)\right] \times \left[(z_{Xi} + z_{Yi})n - 12 z_{Xi}^2 z_{Yi}^2 + \left(2 z_{Xi} z_{Yi} + 2 b_1^{XY} r_{XY} + 3 z_{Xi} b_{12}^{XY} + 3 z_{Yi} b_{21}^{XY}\right) n r_{XY}\right]}{(n-1)(n-2)} \right.$$

$$\left. + \frac{\left[\left(v_{i.}^{(2)} - v_{ii}^2\right)^2 - (v_{i.} - v_{ii})^2 \left(v_{i.}^{(2)} - v_{ii}^2\right) - 2\left(v_{i.}^{(4)} - v_{ii}^4\right) + 2(v_{i.} - v_{ii})\left(v_{i.}^{(3)} - v_{ii}^3\right)\right] \times \left[18 z_{Xi}^2 z_{Yi}^2 - (z_{Xi} + z_{Yi})n - \left(7 z_{Xi} z_{Yi} - n r_{XY} + 2 b_1^{XY} r_{XY} + 3 z_{Xi} b_{12}^{XY} + 3 z_{Yi} b_{21}^{XY}\right) n r_{XY}\right]}{(n-1)(n-2)(n-3)} \right\}$$

$$- \left[\operatorname{E}\left(L_i^{\text{off}}\right) \operatorname{E}\left(L_i^{\text{on}}\right)\right]$$

References

Anselin, L. (1995). ''Local Indicators of Spatial Association: LISA.'' *Geographical Analysis* 27(2), 93–115.

Bao, S., and M. Henry. (1996). ''Heterogeneity Issues in Local Measurements of Spatial Association.'' *Geographical Systems* 3(1), 1–13.

Boots, B. N. (2003). ''Developing Local Measures of Spatial Association for Categorical Data.'' *Journal of Geographical Systems* 5(2), 139–60.

Boots, B. N., and M. Tiefelsdorf. (2000). ''Global and Local Spatial Autocorrelation in Bounded Regular Tessellations.'' *Journal of Geographical Systems* 2(4), 319–48.

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

Costanzo, M., L. J. Hubert, and R. G. Golledge. (1983). ''A Higher Moments for Spatial Statistics.'' *Geographical Analysis* 15(4), 347–51.

Czaplewski, R. L., and R. M. Reich. (1993). ''Expected Value and Variance of Moran's Bivariate Spatial Autocorrelation Statistic for a Permutation Test.'' Research Paper RM-309. U.S. Department of Agriculture, Rocky Mountain Forest and Range Experiment Station, Fort Collins, CO.

Fotheringham, A. S. (1997). ''Trend in Quantitative Methods I: Stressing the Local.'' *Progress in Human Geography* 21(1), 88–96.

Fotheringham, A. S., and C. Brunsdon. (1999). ''Local Forms of Spatial Analysis.'' *Geographical Analysis* 31(4), 340–58.

Getis, A., and J. K. Ord. (1992). ''The Analysis of Spatial Association by Use of Distance Statistics.'' *Geographical Analysis* 24(3), 189–206.

Getis, A., and J. K. Ord. (1996). ''Local Spatial Statistics: An Overview.'' In *Spatial Analysis: Modelling in a GIS Environment*, 261–77, edited by P. Longley and M. Batty. Cambridge, UK: GeoInformation International.

Heo, M., and K. R. Gabriel. (1998). ''A Permutation Test of Association Between Configurations by Means of the RV Coefficient.'' *Communications in Statistics: Simulation and Computation* 27(3), 843–56.

Hepple, L. W. (1998). ''Exact Testing for Spatial Autocorrelation Among Regression Residuals.'' *Environment and Planning A* 30(1), 85–108.

Hubert, L. J. (1984). ''Statistical Applications of Linear Assignment.'' *Psychometrika* 49(4), 449–73.

Hubert, L. J. (1987). *Assignment Methods in Combinatorial Data Analysis*. New York: Marcel Dekker.

Lee, S.-I. (2001a). ''Developing a Bivariate Spatial Association Measure: An Integration of Pearson's *r* and Moran's *I*.'' *Journal of Geographical Systems* 3(4), 369–85.

Lee, S.-I. (2001b). ''Spatial Association Measures for an ESDA-GIS Framework: Developments, Significance Tests, and Applications to Spatio-Temporal Income Dynamics of U.S. Labor Market Areas, 1969–1999.'' Unpublished Ph.D. Dissertation, Department of Geography, The Ohio State University.

Lee, S.-I. (2004). ''A Generalized Significance Testing Method for Global Measures of Spatial Association: An Extension of the Mantel Test.'' *Environment and Planning A* 36(9), 1687–703.

Leung, Y., C.-L. Mei, and W.-X. Zhang. (2003). ''Statistical Test for Local Patterns of Spatial Association.'' *Environment and Planning A* 35(4), 725–44.

Mantel, N. (1967). ''The Detection of Disease Clustering and a Generalized Regression Approach.'' *Cancer Research* 27(2), 209–20.

Mielke, P. W. (1979). ''On Asymptotic Non-normality of Null Distributions of MRPP Statistics.'' *Communications in Statistics: Theory and Methods* A8(15), 1541–50; 1981, Erratum A10(17) 1795; Erratum 11(7) 847.

Ord, J. K., and A. Getis. (1995). ''Local Spatial Autocorrelation Statistics: Distributional Issues and an Application.'' *Geographical Analysis* 27(4), 286–306.

Ord, J. K., and A. Getis. (2001). ''Testing for Local Spatial Autocorrelation in the Presence of Global Autocorrelation.'' *Journal of Regional Science* 41(3), 411–32.

Reich, R., R. L. Czaplewski, and W. A. Bechtold. (1994). ''Spatial Cross-Correlation of Undisturbed, Natural Shortleaf Pine Stands in Northern Georgia.'' *Environmental and Ecological Statistics* 1(3), 201–17.

Siemiatycki, J. (1978). ''Mantel's Space–Time Clustering Statistic: Computing Higher Moments and a Composition of Various Data Transforms.'' *Journal of Statistical Computation and Simulation* 7(1), 13–31.

Sokal, R. R., N. L. Oden, and B. A. Thomson. (1998). ''Local Spatial Autocorrelation in a Biological Model.'' *Geographical Analysis* 30(4), 331–54.

Tiefelsdorf, M. (1998). ''Some Practical Applications of Moran's *I*'s Exact Conditional Distribution.'' *Papers in Regional Science* 77(2), 101–29.

Tiefelsdorf, M. (2000). *Modelling Spatial Processes: The Identification and Analysis of Spatial Relationships in Regression Residuals by Means of Moran's I*. Berlin, Germany: Springer.

Tiefelsdorf, M., and B. Boots. (1997). ''A Note on the Extremities of Local Moran's I_is and Their Impact on Global Moran's *I*.'' *Geographical Analysis* 29(3), 248–57.

Tiefelsdorf, M., D. A. Griffith, and B. Boots. (1999). ''A Variance-Stabilizing Coding Scheme for Spatial Link Matrices.'' *Environment and Planning A* 31(1), 165–80.

Unwin, A., and D. J. Unwin. (1998). ''Exploratory Spatial Data Analysis with Local Statistics.'' *Journal of the Royal Statistical Society D: The Statistician* 47(3), 415–21.
Wartenberg, D. (1985). ''Multivariate Spatial Correlation: A Model for Exploratory Geographical Analysis.'' *Geographical Analysis* 17(4), 263–83.