



TI 2003-081/3

Tinbergen Institute Discussion Paper

# Misspecification in Linear Spatial Regression Models

*Raymond J.G.M. Florax<sup>1</sup>*

*Peter Nijkamp<sup>1,2,3</sup>*

*<sup>1</sup> Department of Spatial Economics, Faculty of Economics and Business Administration, Vrije Universiteit Amsterdam, <sup>2</sup> Regional Economics Applications Laboratory (REAL), University of Illinois, Urbana-Champaign, USA, <sup>3</sup> Tinbergen Institute.*

**Tinbergen Institute**

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

**Tinbergen Institute Amsterdam**

Roetersstraat 31

1018 WB Amsterdam

The Netherlands

Tel.: +31(0)20 551 3500

Fax: +31(0)20 551 3555

**Tinbergen Institute Rotterdam**

Burg. Oudlaan 50

3062 PA Rotterdam

The Netherlands

Tel.: +31(0)10 408 8900

Fax: +31(0)10 408 9031

Please send questions and/or remarks of non-scientific nature to [driessen@tinbergen.nl](mailto:driessen@tinbergen.nl).

Most TI discussion papers can be downloaded at <http://www.tinbergen.nl>.

# Misspecification in Linear Spatial Regression Models<sup>\*</sup>

Raymond J.G.M. Florax<sup>1</sup> and Peter Nijkamp<sup>2</sup>

<sup>1</sup> Department of Spatial Economics, Free University  
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands  
Phone: + 31 20 4446090  
Fax: + 31 20 4446004  
E-mail: rflorax@feweb.vu.nl

and

Regional Economics Applications Laboratory (REAL)  
University of Illinois, Urbana-Champaign, USA

<sup>2</sup> Department of Spatial Economics, Free University  
De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands  
Phone: + 31 20 4446090  
Fax: + 31 20 4446004  
E-mail: pnijkamp@feweb.vu.nl

and

Tinbergen Institute, Amsterdam, The Netherlands

**Abstract.** Spatial effects are endemic in models based on spatially referenced data. The increased awareness of the relevance of spatial interactions, spatial externalities and networking effects among actors, evoked the area of spatial econometrics. Spatial econometrics focuses on the specification and estimation of regression models explicitly incorporating such spatial effects. The multidimensionality of spatial effects calls for misspecification tests and estimators that are notably different from techniques designed for the analysis of time series. With that in mind, we introduce the notion of spatial effects, referring to both heterogeneity and interdependence of phenomena occurring in two-dimensional space. Spatial autocorrelation or dependence can be detected by means of cross-correlation statistics in univariate as well as multivariate data settings. We review tools for exploratory spatial data analysis and misspecification tests for spatial effects in linear regression models. A discussion of specification strategies and an overview of available software for spatial regression analysis, including their main functionalities, intend to give practitioners of spatial data analysis a head start.

<sup>\*</sup> This paper is forthcoming in the *Encyclopedia of Social Measurement*, San Diego: Academic Press, 2004, edited by K. Kempf-Leonard. The authors would like to thank Henri L.F. de Groot, Thomas de Graaff, and two anonymous reviewers for helpful comments.

## 1. Spatial effects

The awareness and incorporation of space is evidently fundamental to geography. In 1979, it spurs the geographer Waldo Tobler to formulate the first law of geography, stating “everything is related to everything else, but near things are more related than distant things.” The relevance of spatial effects extends, however, beyond geography, and is ubiquitous in many of the social sciences.

Cliff and Ord (1981) start-off their seminal book about models and applications of spatial processes with an example from epidemiology. They analyze the spatial mortality pattern from cholera in London, 1848–1849, and attribute the high incidence of the disease in the London metropolis to the high organic content of the Chelsea, Southwark and Vauxhall waters. Similarly, spatial analysis techniques are used to analyze diffusion–adoption patterns of innovations by farmers, to prescribe spatially differentiated fertilizer doses in precision agriculture, to explain network effects among individuals, and to investigate human–environment interactions in processes of environmental degradation.

Economists have traditionally been more reluctant to consider space as a relevant factor. In 1890, the economist Marshall acknowledges the role of space, maintaining that the working of the market depends “... chiefly on variations in the area of space, and the period of time over which the market in question extends; the influence of time being more fundamental than that of space.” It takes until the 1950s, however, before Walter Isard opposes this, what he calls, “Anglo-Saxon bias” that repudiates the factor of space, compresses everything within the economy to a point so that all spatial resistance disappears, and thus confines economic theory to “a wonderland of no spatial dimensions.” It is an ongoing debate ever since, whether space is merely a geographical facilitator or medium for movement, or whether space has an intrinsic explanatory function.

The Dixit-Stiglitz revolution in economic theory increases the awareness for imperfect competition and increasing returns to scale that is subsequently apparent in the New Economic Geography of Krugman, Fujita, and others. Nowadays, spatial dimensions are taken into account in the study of, for instance, economic growth, high-tech innovations, urban economics, public sector productivity, fiscal policy interdependence, and international trade.

### *1.1 Spatial dependence and spatial heterogeneity*

Spatial effects include spatial heterogeneity and spatial dependence. Spatial heterogeneity refers to structural relations that vary over space, either in a discrete or categorical fashion

(for instance, urban vs. rural, or according to an urban hierarchy), or in a continuous manner (such as on a trend surface). Spatial dependence points to systematic spatial variation that results in observable clusters or a systematic spatial pattern. These descriptions already show that in an observational sense, spatial dependence and spatial heterogeneity are not always easily discernable. The clustering of high values in, for instance, urban areas and urban fringes can be interpreted as spatial clustering of high values pertaining to urban areas and low values to rural areas, but it may as well be viewed as spatial heterogeneity distinguishing metropolitan areas from their hinterland.

The typical feature of spatial dependence or spatial autocorrelation is that it is two-dimensional and multidirectional. An observation of an attribute at one location can be correlated with the value of the same attribute at a different location, and vice versa, and the causation pattern can occur in different directions.

Figure 1 shows two identical (7×10) regular grid systems with distinct spatial distributions of the same values. The absolute location of the non-zero values is the same in both grids, but graph (a) shows a clustering of relatively low values on the left-hand side, and high values to the right. Graph (b) shows a much more random spatial allocation of values. In terms of spatial effects, we note that the distribution in graph (a) exhibits spatial dependence and spatial heterogeneity, whereas the graph (b) does not.

The occurrence of spatial heterogeneity does not necessarily have severe implications for the information that can be obtained from a spatial data series. Spatial autocorrelation does, however, because an observation is partly predictable from neighboring observations. A series of spatially dependent observations therefore contains less information. This is similar to the situation in time series analysis, where a forecast with respect to the future can be partly inferred from what happened in the past. The two-dimensional and multidirectional nature of spatial autocorrelation makes that the spatial case is more complex.

## *1.2 Spatial econometrics*

The history of the analysis of the spatial autocorrelation problem goes back to the work of statisticians such as Moran, Geary and Whittle in the late 1940s and early 50s. The development of the literature is rather slow until Cliff and Ord publish their seminal book about spatial autocorrelation in 1973. Their book focuses on the statistical analysis of spatial data series, although not exclusively from a spatial statistical point of view because there is also some attention for modeling. The modeling context is, however, much more pronounced in the efforts of the Dutch–Belgian economist Jean Paelinck, who coins the term ‘spatial

econometrics' in the early 1970s. Paelinck and Klaassen jointly write the first monograph on spatial econometrics in 1979, stressing the need to explicitly model spatial relations, epitomizing the asymmetry in spatial interrelations and the role of spatial interdependence. The edges of the field are in those days pushed ahead mainly by Dutch regional economists (Bartels, Brandsma, Hordijk, Ketellapper, and Nijkamp). Later, the center of activity shifts to the US, where both economists and geographers concentrate on introducing new statistical tests (particularly, tests developed in a maximum likelihood framework) and the specification and estimation of spatial regression models. During the late 1990s, these methodological developments are gradually being picked up in applied spatial research, among other things facilitated by the availability of spatial econometric software.

The most comprehensive book of 'modern' developments is Anselin's 1988 book on methods and models of spatial econometrics, where he defines spatial econometrics as "the collection of techniques that deal with the peculiarities caused by space in the statistical analysis of regional science models." The modeling perspective, already pondered by Paelinck, distinguishes spatial econometrics from the broader field of spatial statistics (Cressie, 1993). Good theoretical overviews of spatial econometrics are available in Anselin and Bera (1998), and Anselin (2001). The tutorial of the SpaceStat software is available online (<http://www.terraseer.com/spacestat.html>) and provides a good introduction. LeSage's contribution to the Web book of regional science gives a more applied, hands-on introduction to spatial econometrics (<http://www.rri.wvu.edu>).

### *1.3 Article organization*

This article continues as follows. In Section 2, we discuss test statistics for spatial autocorrelation that have been developed, and their concurrent measurement level of the data. We cover various techniques for exploratory spatial data analysis, in particular those techniques that assist in determining the correct specification of regression models. Several frequently used models are introduced in Section 3, and in Section 4 we present misspecification tests for spatial dependence. It is common that in practical applications it is difficult to determine what the actual data generating process is on the basis of a series of misspecification tests. Little is known about this specific aspect of spatial econometric modeling. We review the limited knowledge about specification search strategies in Section 5. At the same time, we provide guidelines for practitioners of spatial data analysis by reviewing software that contains tools for misspecification testing and estimation of spatial process models. Section 6 concludes.

## 2. Testing for spatial autocorrelation

### 2.1 Definition spatial autocorrelation

The seminal work of Cliff and Ord (1973, 1981) has induced extensive attention for the statistical properties of spatial data, in particular for spatial autocorrelation or dependence. Statistical tests for spatial association or dependence are always based on the null hypothesis of spatial independence. The general notion of independence is easily formalized as:

$$\Pr(X_1 < x_1, \mathbf{K}, X_n < x_n) = \prod_{i=1}^n \Pr(X_i < x_i), \quad (1)$$

with  $i = 1, \dots, n$ . Consequently,

$$\Pr(X_i = a, X_j = b) = \Pr(X_i = a) \cdot \Pr(X_j = b). \quad (2)$$

For independent events, such as tossing an unbiased coin, the probability of the occurrence of a series of events (e.g., throwing 6 and subsequently 5) is determined multiplicatively by the chance of the occurrence of the individual events (i.e., throwing 6 and throwing 5). In the case of dependence, the above implication does not hold, because the chance of the second event occurring is somehow conditional on the occurrence of the first event.

The notion of spatial dependence is similar, but more specific in the sense that the dependence among events is mediated through space, for instance, either through distance or adjacency. We refer to spatial dependence using informal terms, such as spatial clustering of similar values, an organized pattern, or systematic spatial variation. Strictly speaking, spatial dependence is a characteristic of the joint probability density function. As such, it is only verifiable under simplifying conditions, such as normality. Spatial autocorrelation is simply a moment of that joint distribution. In this article, however, we follow common practice, and use spatial autocorrelation and spatial dependence interchangeably.

Spatial autocorrelation is easily observed in Figure 1. In the graph to the right, observations are randomly distributed over space, so the chances of observing a specific value are independently distributed. The left graph exhibits spatial autocorrelation or dependence, because the probability of a specific value occurring in a specific location depends on the value of neighboring locations.

## 2.2 Statistical tests for spatial autocorrelation

The common element in all tests for spatial autocorrelation is that they relate a vector of attribute values for various locations to all other locations through a matrix representing the structure of the spatial system. Most test statistics are special cases of the general cross-product statistic developed in various publications by Hubert, Golledge, Costanza, and Gale, given by:

$$\Gamma = \sum_{i=1}^n \sum_{j=1}^n w_{ij} y_{ij} , \quad (3)$$

where  $i$  and  $j$  ( $= 1, \dots, n$ ) index the locations,  $w_{ij}$  are the elements of a matrix of the spatial topology of the system, and  $y_{ij}$  the association between the values of a specific attribute at different locations. It is common to assume no self-association, which is easily incorporated in (3) by defining  $w_{ij} = 0, \forall i = j$ . Various metrics can be used to measure the association among attribute values  $y_{ij}$ , such as the product of the attribute values  $x_i \cdot x_j$ , or the squared difference  $(x_i - x_j)^2$ . Getis (1991) provides a good overview of spatial association statistics as well as a comparison to spatial interaction models.

Most statistics for spatial association are special cases of the above cross-product statistic. For spatial data measured on a *nominal* level (for example, unity for a location displaying the attribute, and zero otherwise, usually colored black and white on a map) so-called join-count statistics have been developed. They compute the observed frequency of black-black, black-white, and white-white joins. The comparison to what is expected by chance constitutes a formal test for spatial autocorrelation. The test statistics are asymptotically normally distributed and the moments, under free and non-free sampling, are derived in Cliff and Ord (1981). They also provide extensions to categorical data organized in multi-colored maps.

The best-known spatial association statistics for *ordinal* and *interval* data are Geary's  $c$  and Moran's  $I$ . These statistics are also special cases of the general cross-product statistic. The univariate Moran's  $I$  statistic is given by:

$$I = \frac{n}{S_0} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \quad (4)$$

or in matrix terms,

$$I = \frac{n}{S_0} \cdot \frac{\mathbf{x}'\mathbf{W}\mathbf{x}}{\mathbf{x}'\mathbf{x}}, \quad (5)$$

where  $\mathbf{x}$  is a  $(n \times 1)$  vector of observations  $x_i$  measured in deviations from the mean  $\bar{x}$ ,  $\mathbf{W}$  is a spatial weight matrix with  $(n \times n)$  elements  $w_{ij}$  representing the topology of the spatial system, and  $S_0$  is the sum of the elements of the spatial weights matrix. The weight matrix is typically exogenously determined, and can be defined using contiguity, distance, or more complicated specifications such as those based on the distance and length of the common border, referred to as ‘generalized weights’ (Cliff and Ord, 1981, pp. 17-18). As for all autocorrelation statistics, a slight rearrangement of the statistic shows that it is based on a comparison of the covariance of spatially connected observations to the variance of all observations.

Statistical inference can be based on the standardized or  $z$ -value of Moran’s  $I$ , as:

$$z_I = \frac{I - E(I)}{SD(I)}, \quad (6)$$

but an assumption about the distribution of the  $x$ -variates is needed. The moments of  $I$  can be derived analytically assuming that the  $x$  follow a normal distribution, or that the distribution is unknown but can be approximated in a nonparametric framework using a randomization approach. Alternatively, an estimated distribution function for Moran’s  $I$  can be empirically generated following a permutation approach (Cliff and Ord, 1981, pp. 42-46 and 63-65).

It is common practice to interpret Moran’s  $I$  as a correlation coefficient, although its value is strictly speaking not restricted to the  $[-1,+1]$  interval. High positive values signal the occurrence of similar attribute values over space (either high or low values), and hence spatial clustering. Negative values indicate the joint occurrence of high and low attribute values in nearby locations. A value close to the expected value of Moran’s  $I$  in the absence of

spatial correlation (which equals  $-1/(n-1)$  and therefore approaches zero with increasing sample size) can be taken as evidence of a random allocation of attribute values over space.

Concurrent results for Geary's  $c$  are developed in a similar fashion. The statistic is based on a different metric, and reads as:

$$I = \frac{n-1}{2S_0} \cdot \frac{\sum_{i=1}^n \sum_{j=1}^n w_{ij} (x_i - x_j)^2}{\sum_{i=1}^n (x_i - \bar{x})^2}. \quad (7)$$

Cliff and Ord (1981) also derive the moments for the Geary statistic, under different assumptions. The expected value of Geary's  $c$  is one under the null hypothesis of no spatial autocorrelation, and the variance again depends on the stochastic assumption (normality or randomization). Statistical inference is straightforwardly based on the corresponding  $z$ -value of the statistic. Somewhat counterintuitive, a negative (positive) of Geary's  $c$  indicates positive (negative) spatial autocorrelation. Cliff and Ord (1981) show that the asymptotic relative efficiency is generally better for the  $I$ -test than for the  $c$ -test, and it is therefore common practice to use Moran's  $I$ , especially in a regression context. The variance of Geary's  $c$  is also more sensitive to the distribution of the sample data (see Cliff and Ord, 1981).

Note that the above statistics and their asymptotic properties refer to 'raw data.' The moments for regression residuals are different, because residuals are correlated by definition – even if the errors are independent. The asymptotic properties of Moran's  $I$  for regression residuals have been derived (see below), whereas those for Geary's  $c$  are not available.

The above autocorrelation statistics have a series of important limitations (Cliff and Ord, 1981; Anselin, 1988). First, for most straightforward specifications of the weight matrix, the statistics are topologically invariant, implying that changes in the size, shape, and relative strength of the link between the areal units are ignored. Second, in the connection or weights matrix all links are given equal weight, which does not necessarily confirm to an appropriate perception of space. Third, the level of aggregation of the data and the spatial arrangement is usually rather arbitrary, but it affects the magnitude of the statistics. In order to ensure identification, the weights have to be defined a priori and exogenously, but this implies that for irregularly spaced areal units no analytical mathematical properties can be derived. Finally, the specification of the weight matrix is such that out-of-sample observations are

simply ignored. This leads to edge effects that are difficult to assess and it is cumbersome, or even impossible, to alleviate their disturbing influence.

### 2.3 Local statistics and the Moran scatterplot

During the 1990s, two extensions of the above overall or ‘global’ statistics are developed to further intensify exploratory spatial data analysis (ESDA for short). One extension centers on disaggregating the overall pattern of spatial variation in local spatial patterns, and is caught in so-called local indicators of spatial association (commonly referred to as LISA).

Exploratory spatial data analysis focuses on identifying spatial dependence patterns and spatial heterogeneity, either as an independent analytical tool or as a tool that assists in identifying the specification of a spatial regression model. The cross-product statistics discussed above are ‘global’ in nature in that they identify clusters of spatial association pertaining to the entire spatial system. However, spatial processes may be instable or non-stationary. Within an overall autocorrelated sample, local pockets of randomness may exist. Alternatively, localized clusters of spatial association, sometimes referred to as ‘hot spots,’ may be observed within an overall randomly distributed sample of observations. The generalized cross-product statistic, and variants such as Moran’s  $I$  can be disaggregated into local statistics. For instance, for a standardized weights matrix (i.e., the row sums are scaled to unity) the local Moran is given by:

$$I_i = \frac{(x_i - \bar{x}) \sum_{j=1}^n w_{ij} (x_j - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2 / n}, \quad (8)$$

which clearly shows that the local Moran is proportional to the global Moran’s  $I$  up to a scaling constant. Anselin (1995) shows that the null distribution of the local statistic should not be approximated by the normal, but instead a randomization approach should be used for statistical inference.

The proportionality property distinguishes LISA from other local statistics, notably the Getis–Ord statistics. The latter compare the sum of values in neighboring locations within a certain distance band to the sum over all locations, and have a slightly different interpretation (Getis and Ord, 1992). The LISA and the Getis–Ord statistic can both be used

to identify local pockets of non-stationarity, or to evaluate the contribution of individual areal units to the global spatial pattern.

A second extension of the overall or global statistics builds on the general property that any statistic taking the form of a ratio of a quadratic form and its sum of squares can be interpreted as a bivariate regression coefficient. Equations (4) and (5) show that Moran's  $I$  has this form, and it is thus equivalent to the estimated parameter of a regression of  $\mathbf{W}\mathbf{x}$  on  $\mathbf{x}$ . Rearranging gives:

$$I = \mathbf{x}'(\mathbf{x}'\mathbf{x})^{-1} \mathbf{W}\mathbf{x}, \quad (9)$$

omitting the standardization term that is unity for a standardized weights matrix. This result can be visualized in a Moran scatterplot that can be used to visually identify spatial clusters, outliers, and local non-stationarity (Anselin 1995, 1996). The Moran scatterplot is similar to the spatial lag scatterplot in exploratory geostatistics (Cressie, 1991, pp. 38-39). When the  $x$ -variable is standardized (i.e., expressed in deviations from the sample mean, and divided by the standard deviation), and the value of  $x$  is plotted on the horizontal axis and the weighted average value of  $x$  for the neighbors on the vertical axis, the scatterplot contains four quadrants (see Figure 3, for an example). The quadrants possibly show clusters of high-high values (top-right quadrant), low-low values (bottom-left), and low-high and high-low values (top-left and bottom-right, respectively). Because the  $x$ -variable is standardized, exploratory tools such as the  $2\sigma$ -rule for outliers can be easily assessed. The Moran scatterplot can be used for any variable, including residuals of a regression model.

#### *2.4 Illustration of exploratory spatial data analysis*

We provide a simple illustration from neoclassical economic growth theory that we will also use for modeling purposes below. In their path-breaking article Mankiw, Romer and Weil (1992, *Quarterly Journal of Economics*; MRW for short) analyze the economic growth performance of 98 non-oil countries, over the period 1960–1985. The data and the model presented here are identical to those used by MRW, except that we explicitly consider spatial effects. The data can be found online at <http://post.economics.harvard.edu/faculty/mankiw/mankiw.html>. The weight matrix has been compiled using the centroid function in ArcView (see <http://www.esri.com>), and SpaceStat (see <http://www.terraser.com/spacestat.html>) to compute spherical distances between the countries' centroids.

Figure 2 shows the spatial distribution of per capita income in 1960 (a) and 1985 (b), and the growth rate of per capita GDP over the period 1960–1985 (c). The maps reveal that the spatial distribution of the level of GDP per capita is highly persistent over time. The richest countries are clustered in North America and Western Europe, and Australia and New Zealand have high GDP per capita levels as well, both in 1960 and in 1985. The spatial distribution of economic growth seems much more random although, in accordance with the theoretical prediction of neoclassical growth theory, the fast-growing economies are typically located outside the abovementioned area with high GDP-level countries.

This pattern of spatial clustering is corroborated in the statistical analysis. We specify the elements of the spatial weights matrix as  $1/d_{ij}$ , where  $d_{ij}$  is the spherical distance of the geographical midpoints of the various countries, and we put zeros on the diagonal. The weights matrix is standardized, implying that all elements are scaled such that each row sum equals one, in order to avoid scale effects. Table 1 shows that GDP per capita levels change from approximately US\$ 3,000 in 1960 to US\$ 5,300 in 1985 (in 1985 prices), and the growth rate of per capita GDP is approximately 2 percent. The normality test is rejected for the levels, but not for the growth rate, so we use the normality assumption for Moran's  $I$  in the latter case and apply the randomization approach for the levels. The Moran's  $I$  tests show that there is significant spatial clustering of similar values for GDP per capita levels. The same is true for GDP per capita growth rates, although at a considerably lower level.

Figure 3 shows a Moran scatterplot of economic growth (a), and a mapping of the four types of countries identified in the Moran scatterplot. The former reveals the equivalence of Moran's  $I$  to the estimated regression coefficient of the linear trend line. The map shows that low-growth countries tend to cluster together (especially in Africa), but high-growth countries are either surrounded by other high-growth countries (in Europe) or by countries that exhibit considerably lower growth rates (in North and South America, and in Africa). This is a first visual indication that spatial heterogeneity may be relevant for the regression model, either by varying the coefficients over different types of regions and/or allowing for groupwise heteroskedasticity.

### **3. Spatial regression models**

The specification of spatial dependence in (linear) regression models is usually distinguished according to two types of representation. One is to directly model the dependence, for instance, by means of a distance deterrence function. If a sufficient number of data points is

available the influence of spatial impediments can be estimated directly. Another approach is to provide exogenous structure by specifying the spatial topology and interaction structure in a spatial weights matrix in order to avoid the incidental parameter problem that would arise when the influence of neighboring locations is estimated for each region separately. We restrict the discussion to the latter category of so-called spatial process models. There is a third type of model, the so-called conditional autoregressive models, but for these types of models, it is rather cumbersome to attain appropriate estimates (Cliff and Ord, 1981).

### 3.1 Spatial process models

A general specification for a spatial regression model is the specification combining a spatially autoregressive dependent variable among the set of explanatory variables and spatially autoregressive disturbances. For higher-order processes this ARAR( $i,j$ ) model is given by:

$$\begin{aligned} \mathbf{y} &= (\rho_1 \mathbf{W}_1 + \Lambda + \rho_i \mathbf{W}_i) \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \\ \boldsymbol{\varepsilon} &= (\lambda_1 \mathbf{W}_1 + \Lambda + \lambda_j \mathbf{W}_j) \boldsymbol{\varepsilon} + \boldsymbol{\mu}. \end{aligned} \quad (10)$$

For a first-order process, this reduces to the ARAR(1,1) model:

$$\begin{aligned} \mathbf{y} &= \rho \mathbf{W}_1 \mathbf{y} + \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\varepsilon}, \\ \boldsymbol{\varepsilon} &= \lambda \mathbf{W}_1 \boldsymbol{\varepsilon} + \boldsymbol{\mu}, \end{aligned} \quad (11)$$

where  $\mathbf{y}$  is the  $(n \times 1)$  vector with observations on the dependent variable,  $\mathbf{X}$  the  $(n \times k)$  design matrix containing the explanatory variables,  $\boldsymbol{\beta}$  the  $(k \times 1)$  vector with parameters,  $\rho$  the scalar spatially autoregressive parameter,  $\lambda$  the scalar spatial autoregressive disturbance parameter, and  $\boldsymbol{\mu}$  is an  $(n \times 1)$  independently and identically distributed vector of error terms. The spatial interaction pattern can be restricted to first-order neighbors (i.e., neighbors that have a common border) as in equation (11), or extended to higher-order neighbors (i.e., neighbors of the neighbors, etc.) as in equation (10). Alternatively, the weight matrix can be specified using inverse distances, eventually with a cut-off point. In the latter case, an interpretation similar to the higher-order interpretation is feasible, using distance bands. The weights matrices for the spatially lagged dependent variable and in the error term do not necessarily have to be identical, but they are assumed the same here for reasons of simplicity.

Spatial stationarity is assured by constraining the spatial autoregressive parameters to lie in the interval  $1/\omega_{\min}$  and  $1/\omega_{\max}$ , where  $\omega_{\min}$  and  $\omega_{\max}$  are respectively the smallest and the largest eigenvalues (on the real line) of the spatial weights matrix. For a row-standardized spatial weights matrix (i.e., the row sums are scaled to unity), the largest eigenvalue is always +1. The model contains a spatially homogeneous set of coefficients but can trivially be extended to allow for spatially varying coefficients and/or heteroskedasticity. A related model is the ARMA(1,1) variant that, for the first-order case, reads as  $\mathbf{y} = \rho\mathbf{W}_1\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} + \lambda\mathbf{W}_1)\boldsymbol{\mu}$ .

The most frequently used models are two special cases of the ARAR(1,1) model, the spatial autoregressive error model or AR(1) model obtained for  $\rho = 0$ , and the spatial lag model when  $\lambda = 0$ . For ease of notation, we use one weight matrix from here on, and omit the subscript. Extensions to higher-order cases are trivial. The spatial (autoregressive) error model reads as:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \lambda\mathbf{W})^{-1}\boldsymbol{\mu}, \quad (12)$$

The error covariance matrix is given by  $\sigma^2[(\mathbf{I} - \lambda\mathbf{W})'(\mathbf{I} - \lambda\mathbf{W})]^{-1}$ , showing that heteroskedasticity is present even if the error term is homoskedastic. This applies likewise to the spatial lag model,

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\mu}, \quad (13)$$

for which the error covariance is the same as in the AR(1) model assuming otherwise homoskedastic errors.

As a special case of the spatial ARMA model one can easily derive the spatial moving average model as  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + (\mathbf{I} + \lambda\mathbf{W})\boldsymbol{\mu}$ , with error covariance matrix  $\sigma^2(\mathbf{I} + \lambda\mathbf{W})(\mathbf{I} + \lambda\mathbf{W}')$ . Although the AR and MA specification appear to be similar, the difference in the error covariance shows that the spatial dependence in the MA model is restricted to first and second-order neighbors, whereas in the AR model the dependence extends throughout the spatial system.

### 3.2 The interpretation of spatial effects

The standard spatial process models introduced above illustrate two quite different interpretations of spatial dependence. Dependence can either be modeled as a substantive process or as a nuisance (Anselin and Rey, 1991). In the spatial lag model, a substantive theoretical interpretation can be given to the spatial interaction. In the spatial error model, spatial dependence is caused either by (erroneously) omitted spatially correlated variables or it is caused by boundaries of regions that do not coincide with actual behavioral units. The different specifications of spatial dependence have divergent implications for estimation and statistical inference. The spatial error model is an example of the more general class of models with a non-spherical variance-covariance matrix, although due to the multidirectional nature of spatial dependence the estimation is more difficult than for time series (in particular, estimated generalized least squares estimators are inconsistent). The spatial lag model exhibits endogeneity that can be taken into account by instrumental variable or general methods of moments techniques, but should preferably be solved using an appropriate maximum likelihood estimator (see Anselin, 1988, for details). In contrast to the time series case, where OLS remains consistent if the errors are not serially correlated and its use is therefore asymptotically warranted, OLS estimators for the spatial lag model are biased and inconsistent, irrespective of the properties of the error term. Nuisance dependence in the error term is less serious because OLS remains unbiased, but it is inefficient.

Rearranging equation (12) shows that the spatial error model is equivalent to an extended spatial lag model comprising both the spatially lagged dependent and spatially lagged exogenous variables:

$$\mathbf{y} = \lambda \mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} - \mathbf{W}\overset{\lambda}{\mathbf{X}}\boldsymbol{\gamma} + \boldsymbol{\mu}, \quad (14)$$

where  $\overset{\lambda}{\mathbf{X}}$  is the original design matrix  $\mathbf{X}$ , except for the constant. The formal equivalence only holds if  $(k - 1)$  nonlinear constraints are satisfied, specifically  $\lambda\boldsymbol{\beta} = -\boldsymbol{\gamma}$ . This model is generally referred to as the ‘spatial Durbin’ or ‘common factor’ model. The similarity between the spatial error and the common factor model shows that tests with either the spatial error or the spatial lag model as the alternative hypothesis are likely to have power against the other alternative as well (Anselin, 2001).

#### 4. Testing for spatial effects in spatial process models

Apart from focused tests with an informative alternative hypothesis that points to, for instance, the spatial lag or the spatial error model, various diffuse tests merely reflecting whether the residuals are spatially correlated have been developed. The oldest and best known is Moran's  $I$  for regression residuals, given by:

$$I = \frac{n}{S_0} \cdot \frac{\mathbf{u}'\mathbf{W}\mathbf{u}}{\mathbf{u}'\mathbf{u}}, \quad (15)$$

where  $\mathbf{u}$  is the  $(n \times 1)$  vector of OLS residuals. As with the Moran's  $I$  for raw data given in Equation (5), statistical inference can be based on the assumption of asymptotic normality, but an exact approach depending on the matrix  $\mathbf{X}$  is available too, although rather cumbersome to apply. Moran's  $I$  for regression residuals is a locally best invariant test, and moments and estimation details are given in Cliff and Ord (1981), and Anselin (1988). An alternative large sample test derived under less restrictive assumptions has been derived in Kelejian and Robinson (1992).

##### 4.1 Unidirectional and multidirectional tests for spatial dependence

Focused tests have a clear alternative hypothesis and are developed in a maximum likelihood framework. They either refer to a *unidirectional* alternative hypothesis dealing with one specific misspecification, a *multidirectional* alternative comprising various misspecifications, or they are robust in the sense that the test allows for the potential presence of a second type of misspecification. The Lagrange Multiplier tests  $LM_\lambda$  and  $LM_\rho$  are unidirectional tests with the spatial error and the spatial lag model as their respective alternative hypothesis. The  $LM$ -error test is identical to a scaled Moran coefficient (for row-standardized weights), and reads as:

$$LM_\lambda = \frac{1}{T} \left( \frac{\mathbf{u}'\mathbf{W}\mathbf{u}}{s^2} \right)^2, \quad (16)$$

where  $s^2$  is the maximum likelihood variance  $\mathbf{u}'\mathbf{u}/n$ ,  $T$  is a scalar computed as the trace of a quadratic expression in the weight matrix,  $T = \text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)$ , and the test asymptotically follows a  $\chi^2$  distribution with one degree of freedom.

The  $LM$ -lag test has the same asymptotic distribution, and looks similar:

$$LM_{\rho} = \frac{1}{nJ_{\rho,\beta}} \left( \frac{\mathbf{u}'\mathbf{W}\mathbf{y}}{s^2} \right)^2, \quad (17)$$

where  $J_{\rho,\beta} = [(\mathbf{W}\mathbf{X}\mathbf{b})'\mathbf{M}(\mathbf{W}\mathbf{X}\mathbf{b}) + Ts^2]/ns^2$  is a part of the estimated information matrix,  $\mathbf{b}$  the OLS estimated parameter vector, and  $\mathbf{M}$  the projection matrix  $\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$ .

Unfortunately, the  $LM$ -error test cannot be used to distinguish between the spatial autoregressive and the moving average processes. The same holds for multidirectional tests, such as the test against the ARAR and the ARMA model, which is similar but not equal to the sum of the error and the lag test:

$$LM_{\rho\lambda} = \frac{(\mathbf{u}'\mathbf{W}\mathbf{y}/s^2 - \mathbf{u}'\mathbf{W}\mathbf{u}/s^2)^2}{RJ_{\rho,\beta} - T} + \frac{(\mathbf{u}'\mathbf{W}\mathbf{u}/s^2)^2}{T}, \quad (18)$$

and follows a  $\chi_{(2)}^2$  distribution.

Finally, a multidirectional test against spatial error dependence and heteroskedasticity does equal the sum of the  $LM$ -error test and the familiar  $LM$  test for heteroskedasticity due to Breusch and Pagan:

$$BP = \frac{1}{2} \mathbf{f}'\mathbf{Z}(\mathbf{Z}'\mathbf{Z})^{-1}\mathbf{Z}'\mathbf{f}, \quad (19)$$

where the elements of  $\mathbf{f}$  are defined as  $f_i = (u_i/s)^2 - 1$ , and  $\mathbf{Z}$  is a  $(n \times k)$  matrix containing the variables causing the heteroskedasticity. The  $BP$  test is asymptotically distributed as  $\chi^2$  with  $k$  degrees of freedom, and hence this joint test follows a  $\chi_{(k+1)}^2$  distribution.

Rather than simple spatial heteroskedasticity, increasingly complex forms of spatial heterogeneity are possible as well. They occur in the case of discrete or continuous spatial variation, and they can be restricted to the means of homogeneous subgroups (spatial ANOVA or a trend surface) or extend to all coefficients (spatial regimes, and error component or spatial expansion models). Anselin (1988) provides details and appropriate test statistics.

#### 4.2 Robust tests for spatial dependence

The derivation of specification tests with locally misspecified alternatives makes it possible to more clearly distinguish between spatial error and spatial lag dependence. These tests are labeled robust because the test statistics account for the potential presence of a spatial lag or spatially correlated errors when testing for the presence of spatially correlated errors or a spatial lag, respectively (Anselin *et al.*, 1996). The test for a spatial error process robust to the local presence of a spatial lag is:

$$LM_{\lambda}^* = \frac{1}{T - T^2(nJ_{\rho,\beta})^{-1}} \left( \frac{\mathbf{u}'\mathbf{W}\mathbf{u}}{s^2} - T(nJ_{\rho,\beta})^{-1} \frac{\mathbf{u}'\mathbf{W}\mathbf{y}}{s^2} \right)^2. \quad (20)$$

This clearly shows the subtraction of a correction factor that accounts for the local misspecification of a spatial lag process. The test for a spatial lag process robust to the local presence of a spatial error is given by:

$$LM_{\rho}^* = \frac{1}{nJ_{\rho,\beta} - T} \left( \frac{\mathbf{u}'\mathbf{W}\mathbf{y}}{s^2} - \frac{\mathbf{u}'\mathbf{W}\mathbf{u}}{s^2} \right)^2. \quad (21)$$

Both tests asymptotically follow a  $\chi_{(1)}^2$  distribution.

It would be fruitful if one could also test for a spatial error or a spatial lag model assuming local misspecification in the form of heteroskedasticity. This approach, however, leads to a highly non-linear problem (see Anselin, 1988, pp. 112-115).

### 5. Small samples, specification strategies and software

In the spatial econometrics literature, the design of misspecification tests has predominantly focused on continuous regression models, although some recent results are also concerned with categorical models (Kelejian and Prucha, 2001), and on the use of Lagrange Multiplier tests. Although *LM* tests are asymptotically equivalent to Wald and Likelihood Ratio tests, *LM* tests are computationally convenient because they are based on parameter estimates under the null hypothesis. Although the increasing availability of large data samples alleviates the problem of the small sample distribution of the tests being markedly different

from their asymptotic distribution, substantial efforts have been put into simulating the performance of the tests in small sample situations.

### *5.1 Small sample performance and specification strategies*

The finite sample performance of spatial dependence tests across various ‘true’ models is actually quite well documented (see Florax and De Graaff, 2004, for a comprehensive overview). For diffuse misspecification tests, it has been shown that the power of Moran’s  $I$  is greater than the Kelejian-Robinson test. The latter is also more sensitive to departures of the normality of the error distribution, although on the positive side, it has power against spatial dependence as well as heteroskedasticity. As one would expect, the power of focused  $LM$  tests against their correct alternative is very high. In general, the power of the tests in small samples is very satisfactory. Asymptotic results are approached for moderately sized samples of approximately 100 observations.

The power of the Lagrange Multiplier spatial error (lag) test against a spatial lag (error) model seriously complicates the work of the practitioner of spatial modeling techniques. If both tests are significant, which underlying model is then correct? This issue is addressed in research on so-called specification strategies. Some researchers (for instance, Haining, Getis, Griffith, and Tiefelsdorf) advocate avoiding the search for the proper model specification altogether. They propose to ‘pre-whiten’ or ‘filter’ the variables in such a way that they are free of spatial dependence, and subsequently one can obviously resort to simple OLS for estimation purposes. As an alternative, one can either follow a ‘classical’ approach (start with simple OLS, perform  $LM$  tests, and select the model corresponding to the test with the highest value) or an approach inspired by Hendry’s work (start with a very general model including spatial effects, and reduce the model on the basis of significance tests). It has been shown, by means of Monte Carlo simulation, that the classical approach outperforms the Hendry approach (Florax *et al.*, 2003), but a comparison to the filtering approaches is not yet available. The latter, as well as a specification strategy centering on how to deal with the joint occurrence of heteroskedasticity (or spatial heterogeneity) and spatial dependence, would be extremely relevant for practitioners of spatial econometric modeling.

Apart from further research on the small sample performance and specification strategies, there are several other unresolved questions or points that deserve further attention. For instance, mainstream econometrics is reluctant to accept the a priori imposition of the elements of the weights matrix, which can be avoided by using so-called direct representation models. Moreover, further advances in the treatment of spatial effects in a panel data setting

are required (see Baltagi *et al.*, 2003), and computational problems when dealing with large data sets need to be further explored.

### *5.2 Software for spatial exploratory and explanatory analysis*

In recent years, there is an upsurge in the development and availability of software for spatial statistics and spatial econometrics. The two most complete software packages are SpaceStat (Anselin, 1992; see <http://www.terraseer.com/spacestat.html>), and James LeSage's extensive spatial econometrics toolkit of routines for the estimation of spatial econometric models developed in MATLAB (see <http://www.spatial-econometrics.com>). Anselin's SpaceStat package combines both exploratory tools and an extensive array of 'classical' misspecification tests and estimators for spatial models. The package can be complemented with a free downloadable tool for mapping and exploratory spatial data analysis called GeoDa (see <http://sal.agecon.uiuc.edu/csiss/index.html#geoda>). This tool also facilitates the building of weighting matrices, and the output can be easily transferred to SpaceStat. LeSage's toolkit is slightly more oriented towards modeling, and contains an array of Bayesian routines in addition to the classical tests and estimators. Several other online software sources are available. In the context of spatial econometric modeling the most relevant are:

- Kelly Pace's spatial statistics toolbox (see [http://spatial-statistics.com/software\\_index.htm](http://spatial-statistics.com/software_index.htm)), containing routines for MATLAB and programs to estimate spatial models for large samples;
- Roger Bivand's SPDEP Spatial Analysis Tools (see <http://cran.r-project.org/src/contrib/PACKAGES.html#spdep>), which gives programs for spatial autocorrelation and regression analysis written in R;
- S+ SpatialStats (see <http://www.insightful.com/products/product.asp?PID=17>), which is a complementary tool to the S-Plus statistical package that contains some spatial regression routines; and
- Winbugs–Geobugs (see <http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/geobugs.shtml>) that has routines for spatial models based on Gibbs sampler and Markov Chain Monte Carlo (MCMC) estimation.

Two developments are particularly noteworthy and promising. One is the ongoing development of software, and the boost to this development that can be expected to result from so-called 'open–source–projects,' which are typically Web based and essentially

constitute a platform and a common language for anybody who can credibly contribute to the development of software tools. The other development relates to the available information, the integration and the accessibility of knowledge regarding spatial analysis. The US National Science Foundation has recently funded a large-scale project to establish a Center for Spatially Integrated Social Science (CSISS; see <http://www.csiss.org>), that is intended to develop a successful clearinghouse for the tools, case studies, educational opportunities, and other resources needed by a spatially integrated social science approach. On the CSISS website Luc Anselin has developed a spatial tools search engine that indexes a list of URLs and software archives of over 700 individual software titles. A spatial tool listing of links to portals (i.e., collections of links) is available as well.

### *5.3 Illustration of the spatial modeling approach*

In order to illustrate the workings of the spatial econometric modeling approach we continue the economic growth example that we started in Section II. Table 2 presents the results for the Mankiw–Romer–Weil (MRW) analysis in the first column. Economic growth over the period 1960–1985 is regressed on the GDP level in 1960, and three variables (investment share, population growth, and human capital) accounting for differences across countries. The implication of the estimation findings is that the 98 non-oil countries are converging towards each other (or actually their so-called ‘steady state’) at a rate of 1.4% per year. The misspecification tests in column (1) show that the estimated standard errors are not accurate because of heteroskedasticity, and – more important – the estimates are biased because of the erroneously omitted spatially lagged dependent variable. We therefore estimate the spatial lag model, with results being reported in column (2). The misspecification tests for this model indicate that there is no remaining spatial error correlation, but the model still suffers from heteroskedasticity. We proceed by allowing the error variance to differ between slow and fast growing economies, and present the results in column (3). Relaxing the homoskedasticity restriction is, however, a rather mechanical solution to the potential relevance of spatial heterogeneity. It does not have a substantive interpretation. We therefore introduce the possibility of two spatial regimes, allowing the coefficients and the variances to differ between slow and fast growing economies. Columns (4a and b) show that this is an appropriate model. There is no heteroskedasticity when the regression parameters are allowed to vary between two different ‘clubs’ of countries; the Chow test indicates that overall the estimated regression parameters for the two types of countries are not identical; and the spatially lagged dependent variable is highly significant and is large in magnitude. From the

‘reduced’ form of the spatial lag model,  $\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1} (\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$ , it is easily seen that the significance of the spatial lag variable can be caused by the omission of spatially correlated exogenous variables and/or by spatially correlated factors within each country that are not considered in the specification. The modeling results reveal that instead of MRW’s 1.4% convergence rate, a differentiation among slow growing economies converging at a 0.3% rate, and fast growing economies converging at a 1.2% rate, is more in accordance with reality. The ultimate model specification allows for spatial heterogeneity and accounts for the intrinsic spatial interrelatedness of countries’ economies.

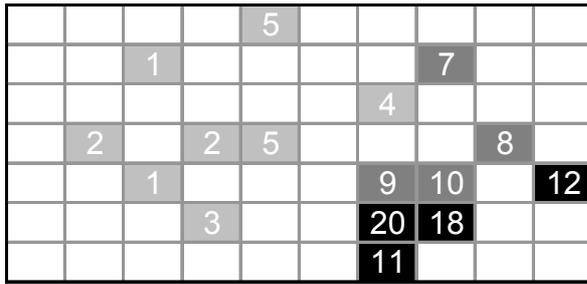
## 6. Conclusion

The use of appropriate spatial statistical and spatial econometric techniques is imperative. This applies to all analyses in the social sciences that are based, or should be based, on spatially referenced data. Such analyses are not necessarily restricted to cases in which data for a specific spatial area (cities, counties, states, or countries) are employed. For any analysis that is intrinsically spatial, whether it is operationalized using individual data, institutional data (firms, hospitals, and the like), or spatially demarcated data, the use of spatial analysis techniques is a *conditio sine qua non*. The growing availability of easy-to-use software for spatial analysis makes it tempting to simply ‘try’ all kinds of different spatial models and techniques, but it is highly desirable that this blends in with a careful consideration of the inherent subtleties underlying different spatial modeling techniques, and theoretical considerations with respect to the topic at hand. Careful consideration of the works mentioned in the bibliography may assist the reader in doing so.

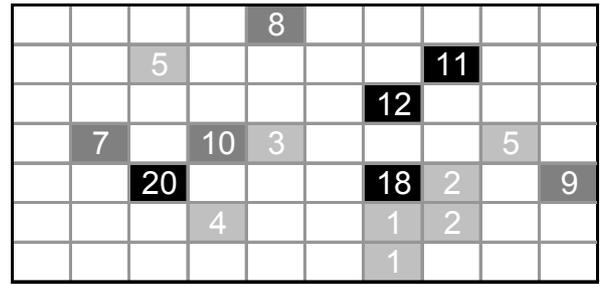
## Bibliography

- Anselin, L. (1988). *Spatial econometrics: Methods and models*. Kluwer, Dordrecht.
- Anselin, L. (1992). *SpaceStat: A program for the analysis of spatial data*. National Center for Geographic Information and Analysis, University of California, Santa Barbara.
- Anselin, L. (1995). Local indicators of spatial association – LISA. *Geograph. Anal.* **27**, 93–115.
- Anselin, L. (1996). The Moran scatterplot as an ESDA tool to assess local instability in spatial association. In *Spatial Analytical Perspectives on GIS* (M. Fischer, H. Scholten, and D. Unwin, Eds.). Taylor and Francis, London.
- Anselin, L. (2001). Spatial econometrics. In *Companion to Econometrics* (B. Baltagi, Ed.). Basil Blackwell, Oxford.
- 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* (A. Ullah, and D. Giles, Eds.). Marcel Dekker, New York.

- Anselin, L., and Rey, S. (1991). Properties of tests for spatial dependence in linear regression models. *Geograph. Anal.* **23**, 112–131.
- Anselin, L., Bera, A. K., Florax, R. J. G. M., and Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. *Reg. Sci. Urban Econ.* **26**, 77–104.
- Baltagi, B. H., Song, S. H., and Koh, W. (2003). Testing panel data regression models with spatial error correlation. *J. Econometrics* **117**, 123–150.
- Cliff, A.D., and Ord, J.K. (1981; first edition 1973). *Spatial Processes: Models & Applications*. Pion, London.
- Cressie, N. (1993; first edition 1991). *Statistics for Spatial Data*. Wiley, New York.
- Florax, R. J. G. M., and De Graaff, T. (2004). The performance of diagnostics tests for spatial dependence in linear regression models: A meta-analysis of simulation studies. In *Advances in Spatial Econometrics: Methodology, Tools and Applications* (L. Anselin, R. J. G. M. Florax, and S. J. Rey, Eds.). Springer Verlag, Berlin.
- Florax, R. J. G. M., Folmer, H., and Rey, S. (2003). Specification searches in spatial econometrics: The relevance of Hendry's methodology. *Reg. Sci. Urban Econ.* **33**, 557–579.
- Getis, A. (1991). Spatial interaction and spatial autocorrelation: A cross-product approach. *Environ. Planning A* **23**, 1269–1277.
- Getis, A., and Ord, K. J. (1992). The analysis of spatial association by means of distance statistics. *Geograph. Anal.* **24**, 189–206.
- Kelejian, H. H., and Prucha, I. R. (2001). On the asymptotic distribution of the Moran I test statistic with applications. *J. Econometrics* **104**, 219–257.
- Kelejian, H. H., and Robinson, D. P. (1992). Spatial autocorrelation: A new computationally simple test with an application to per capita county policy expenditures. *Reg. Sci. Urban Econ.* **22**, 317–331
- Paelinck, J. H. P., and Klaassen, L. H. (1979). *Spatial Econometrics*. Saxon House, Farnborough.
- Upton, G., and Fingleton, B. (1985). *Spatial Data Analysis by Example*. Wiley: New York.



(a)



(b)

FIGURE 1 Hypothetical example of a spatial distribution of values on a (7×10) regular grid. Source: with modifications adopted from Upton and Fingleton (1985, p. 153).

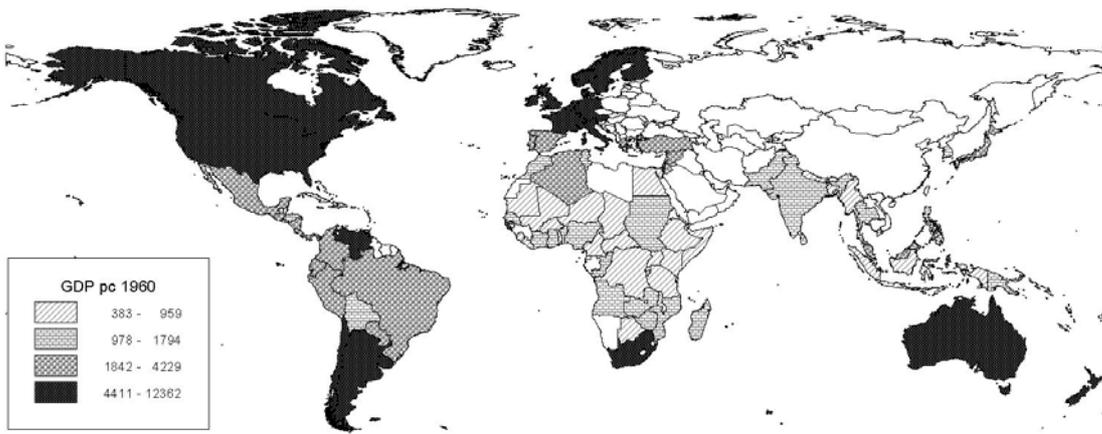


FIGURE 2(a) Map of GDP per capita 1960 for the MRW sample of 98 non-oil countries.

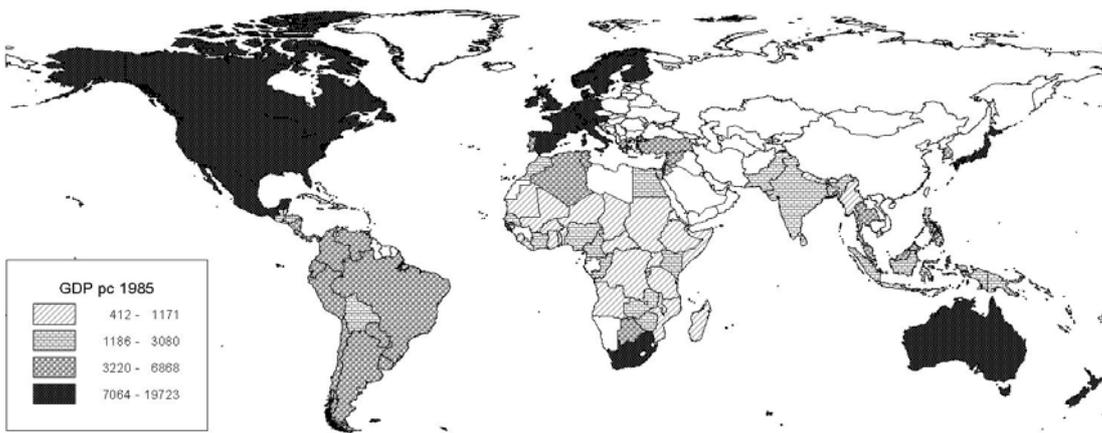


FIGURE 2(b) Map of GDP per capita 1985 for the MRW sample of 98 non-oil countries.

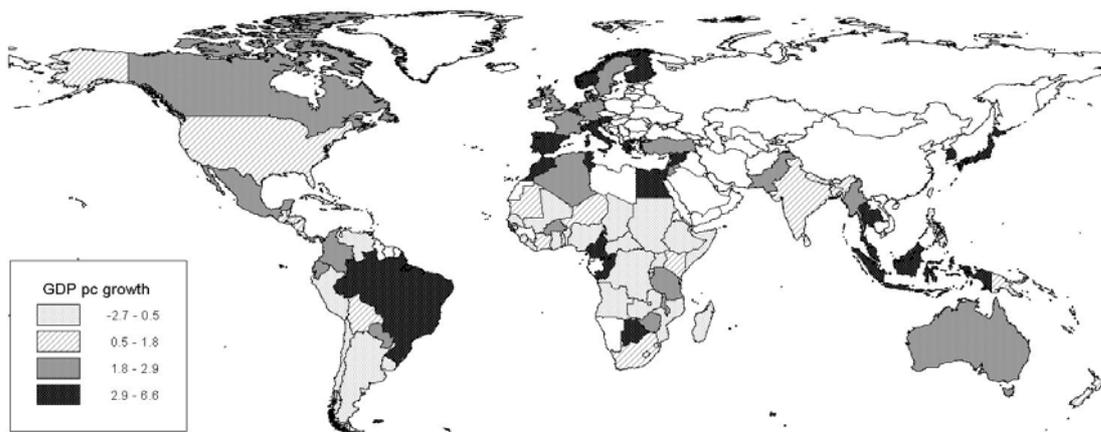


FIGURE 2(c) Map of annual GDP growth over the period 1960–1985 for the MRW sample of 98 non–oil countries.

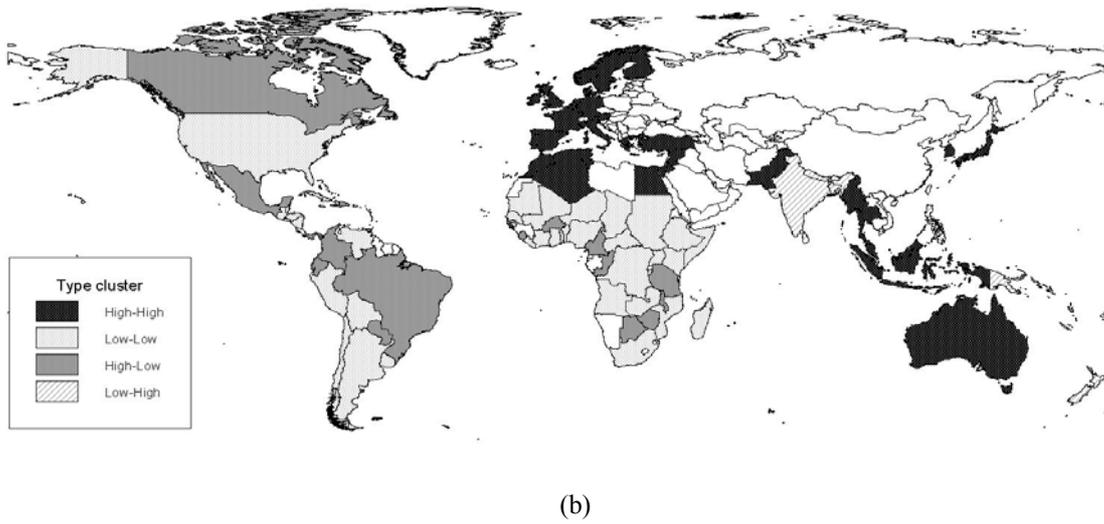
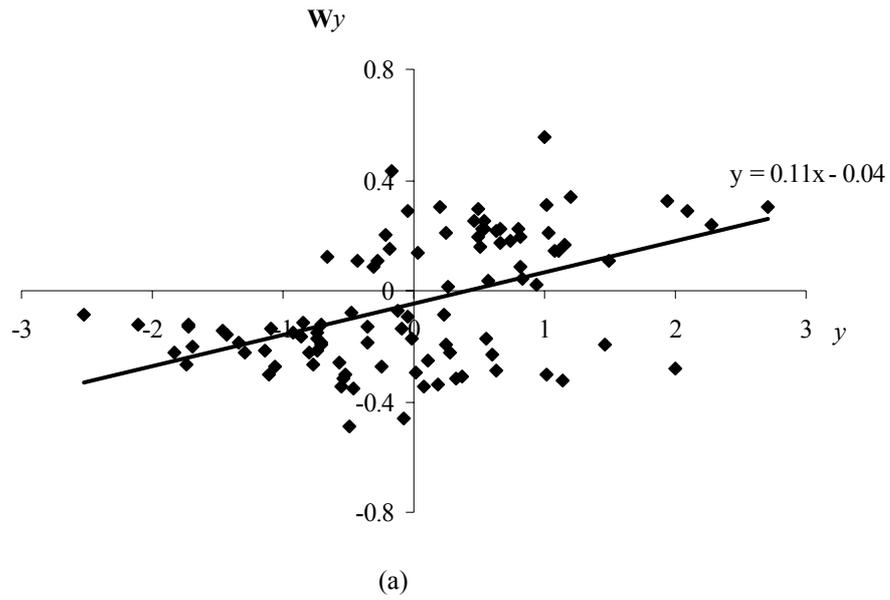


FIGURE 3 Moran scatterplot of economic growth (a), and a map differentiating between four types of countries based on the quadrants in the Moran scatterplot (b).

TABLE 1 Exploratory statistics for GDP per capita levels in 1985 US\$, and annual growth rate of GDP per capita, of the MRW sample comprising 98 non-oil countries, 1960–1985.<sup>a</sup>

	Average	Wald test for normality	Moran's <i>I</i>
GDP per capita 1960	2994.90	38.18*** (0.00)	0.20*** <sup>b</sup> (11.44)
GDP per capita 1985	5309.77	21.79*** (0.00)	0.24*** <sup>b</sup> (13.31)
GDP per capita, annual growth rate	0.02	0.06 (0.97)	0.11*** <sup>c</sup> (6.69)

<sup>a</sup> In parentheses we present probability levels for the Wald test and *z*-values for Moran's *I*.

<sup>b</sup> Using the randomization approach.

<sup>c</sup> Using the normality assumption.

TABLE 2 MRW specification with diagnostics for spatial effects, and spatial process models allowing for spatial dependence and heterogeneity.

Variables	OLS	LAG	LAG-HET	LAG-REG	
				Slow (4a)	Fast (4b)
<i>Constant</i>	3.02*** (3.65)	3.03*** (3.95)	3.81*** (5.65)	1.16 (1.17)	3.88*** (5.39)
<i>GDP level 1960</i>	-0.29*** (-4.68)	-0.28*** (-4.86)	-0.31*** (-5.70)	-0.08 (-1.30)	-0.26*** (-3.92)
<i>Investment share</i>	0.52*** (6.03)	0.52*** (6.44)	0.56*** (6.87)	0.21*** (2.64)	0.41*** (3.62)
<i>Population growth</i>	-0.51* (-1.75)	-0.29 (-1.09)	-0.17 (-0.73)	0.02 (0.07)	0.07 (0.27)
<i>Human capital</i>	0.23*** (3.89)	0.19*** (3.41)	0.20*** (3.73)	0.06 (1.08)	0.17** (2.50)
<i>Spatial lag GDP growth</i>		0.74*** (4.73)	0.74*** (4.99)		0.60*** (3.16)
Convergence rate <sup>f</sup>	1.4	1.3	1.5	0.3	1.2
$R^2$ -adjusted	0.46	0.50	0.56		0.75
<i>LIK</i>	-26.95	-23.07	-20.18		9.50
<i>BP</i>	10.26**	10.93** <sup>a</sup>	5.78** <sup>d</sup>		0.83 <sup>d</sup>
<i>Chow-Wald</i>					86.85*** <sup>c</sup>
<i>I</i>	3.12***				
<i>LM-error</i>	2.24	0.06 <sup>b</sup>			
Robust <i>LM-error</i>	5.08**				
<i>LM-lag</i>	10.34***	7.76*** <sup>c</sup>			
Robust <i>LM-lag</i>	13.18***				

<sup>a</sup> Test on random coefficients. The *BP* test for two regimes, groupings of slow and fast growing economies, is 1.97 (with probability 0.16) for the spatial lag model.

<sup>b</sup> *LM* test on remaining spatial error correlation.

<sup>c</sup> Likelihood Ratio test on spatial lag.

<sup>d</sup> Likelihood Ratio test on groupwise heteroskedasticity (using slow and fast growing economies as groups).

<sup>e</sup> For the tests on equality of individual coefficients for the different groups only the constant and the GDP level are significant (4.90 with probability 0.03, and 4.17 with probability 0.04, respectively).

<sup>f</sup> In percents per year. The convergence rate is equal to  $100 \times (\ln(\beta+1))/-T$ .