

# SPATIAL ECONOMETRICS

**Luc Anselin**

**Bruton Center  
School of Social Sciences  
University of Texas at Dallas  
Richardson, TX 75083-0688**  
lanselin@utdallas.edu

April 26, 1999

**ABSTRACT:** Spatial econometric methods deal with the incorporation of spatial interaction and spatial structure into regression analysis. The field has seen a recent and rapid growth spurred both by theoretical concerns as well as by the need to be able to apply econometric models to emerging large geocoded data bases. The review presented in this chapter outlines the basic terminology and discusses in some detail the specification of spatial effects, estimation of spatial regression models, and specification tests for spatial effects.

# SPATIAL ECONOMETRICS

Luc Anselin

## 1. Introduction

Spatial econometrics is a subfield of econometrics that deals with the treatment of spatial interaction (*spatial autocorrelation*) and spatial structure (*spatial heterogeneity*) in regression models for cross-sectional and panel data [Paelinck and Klaassen (1979), Anselin (1988a)]. As such, the field is similar to geostatistics and spatial statistics, which form an important part of the methodological toolbox of the physical scientist [e.g., Ripley (1981), Cressie (1993)]. However, spatial econometrics is distinct from spatial statistics in the same sense as econometrics is distinct from statistics in general. This albeit subtle distinction boils down to the central role attributed to the theoretical model rather than the data when it comes to dictating the types of specifications that are of interest in spatial econometrics.

A focus on location and spatial interaction has recently gained a more central place not only in applied but also in theoretical econometrics.<sup>1</sup> In the past, models that explicitly incorporated “space” (or geography) and therefore applications of spatial econometrics were primarily found in specialized fields such as regional science, urban and real estate economics and economic geography [e.g., recent reviews in Anselin (1992a), Anselin and Florax (1995a), Anselin and Rey (1997), Pace et al. (1998)]. However, more recently, spatial econometric methods have increasingly been applied in a wide range of empirical investigations in more traditional fields of economics as well, including, among others, studies in demand analysis [Case (1991)], international economics [Aten (1996)], labor economics [Topa (1996)], public economics and local public finance [Case, Rosen and Hines (1993), Murdoch, Rahmatian and Thayer (1993), Holtz-Eakin (1994), Brueckner (1998)], and agricultural and environmental economics [Benirschka and Binkley (1994), Murdoch, Sandler and Sargent (1997), Nelson and Hellerstein (1997), Bell and Bockstael (1999)]. In addition, there is a growing methodological literature in econometrics that deals with alternative model specifications, estimators and test statistics geared to models that incorporate the “geography” of the data [recent examples include, among others, Conley (1996), Driscoll and Kraay (1998), Pinkse and Slade (1998), Kelejian and Prucha (1999a)].

---

1 The origins of this reaction to the so-called Anglo-Saxon focus in economics on the time domain can be traced to Isard (1956).

This new attention to specifying, estimating and testing for the presence of spatial interaction in the mainstream of applied and theoretical econometrics can be attributed to two major factors. One is a growing interest within theoretical economics in models that move away from the atomistic agent as a decision maker acting in isolation to an explicit accounting for the interaction of that agent (its preferences, utility, etc.) with other heterogeneous agents in the system. These new theoretical frameworks specify and study “direct” interaction between agents, in the form of social norms, neighborhood effects, copy-cattng and other peer group effects, and raise interesting questions about how the individual interactions can lead to emergent collective behavior and aggregate patterns. Examples of such models are found in the new macroeconomics of Aoki (1994, 1996), in theoretical models of social interaction [Brock and Durlauf (1995), Akerlof (1997)], interdependent preferences [Alessie and Kapteyn (1991)], models of evolving trading structures [Ioannides (1990, 1997)], neighborhood spillover effects [Durlauf (1994), Borjas (1995), Glaeser, Sacerdote and Scheinkman (1996)], and yardstick competition [Besley and Case (1995), Bivand and Szymanski (1997)]. These frameworks also form some of the underpinnings for empirical models that encompass strategic interaction between agents [e.g., Case, Rosen and Hines (1993), Murdoch, Sandler and Sargent (1997), Brueckner (1998)]. Much of this literature is inspired by principles developed in statistical mechanics, such as the study of interacting particle systems and random field models [for a review, see Durlauf (1997)], which, interestingly, also forms a basis for some of the spatial Markov field models developed in spatial statistics [e.g., Cressie (1993, Ch. 6)]. Related to this strand in the literature is the revived emphasis on the spatial aspects of Marshallian externalities, agglomeration economies and other spillovers that are central to the new economic geography reflected in the work of Arthur (1989), Krugman (1991a, 1991b, 1998), Glaeser et al. (1992) and others.

Paralleling and complementing this theoretical motivation, i.e., the need for a methodology to handle spatial *models*, there has also been a second driver behind the increase in demand for techniques to deal with spatial *data* from a practical, applied perspective. Irrespective of the nature of the model of interest (whether spatial or non-spatial), the explosive diffusion of geographic information systems (GIS) technology and the associated availability of geo-coded socioeconomic data sets (i.e., data sets that contain the location of the observational units) has created a need for specialized methods to deal with the distinguishing characteristics (primarily spatial autocorrelation) of such geographic data [Anselin (1989)]. The use of GIS in conjunction with spatial data analysis and modeling techniques is by now commonplace in applied economics and policy analysis, especially in real estate and housing economics [Anselin (1998a), Can

(1998)], environmental and resource economics [Bockstael (1996), Geoghegan, Wainger and Bockstael (1997)], and development economics [Nelson and Gray (1997)]. The recognition that standard econometric techniques often fail in the presence of spatial autocorrelation, which is commonplace in geographic (cross-sectional) data sets, is therefore a major second motivation for the growing interest in this field.

Historically, spatial econometrics originated as an identifiable field in Europe in the early 1970s because of the need to deal with sub-country data in regional econometric models [e.g., Hordijk and Paelinck (1976)]. In the first book-length treatment of the subject, Paelinck and Klaassen (1979, 5–11) defined the field in terms of five distinguishing characteristics: (a) the role of spatial interdependence in spatial models; (b) the asymmetry in spatial relations; (c) the importance of explanatory factors located in other spaces; (d) differentiation between ex-post and ex-ante interaction; and (e) explicit modeling of space. In Anselin (1988a, 7), spatial econometrics is defined as “the collection of techniques that deal with the peculiarities caused by space in the statistical analysis of regional science models.” In other words, spatial econometrics deals with methodological concerns that follow from the explicit consideration of *spatial effects*, such as spatial autocorrelation and spatial heterogeneity. This yields four broad areas of interest: (a) the formal *specification* of spatial effects in econometric models; (b) the *estimation* of models that incorporate spatial effects; (c) specification *tests* and diagnostics for the presence of spatial effects; and (d) spatial *prediction* (interpolation). In this brief review chapter, I will focus on the first three concerns, since they fall within the central preoccupation of econometric methodology.

The remainder of the chapter is organized as follows. In section 2, I outline some foundations and definitions. In section 3, the specification of spatial regression models is treated, including the incorporation of spatial dependence in panel data models and models with qualitative variables. Section 4 focuses on estimation and section 5 on specification testing. In section 7, some practical implementation and software issues are addressed. Concluding remarks are formulated in Section 8.

## 2. Foundations

### 2.1. Spatial Effects

In a regression context, spatial effects pertain to spatial dependence, or its weaker expression, spatial autocorrelation, and to *spatial heterogeneity*.<sup>2</sup> The latter is simply structural instability in the

2 In this chapter, I will use the terms spatial dependence and spatial autocorrelation interchangeably. Obviously, the two are not identical, but typically, the weaker form is used, in the sense of a moment of a joint distribution. Only seldom is the focus on the complete joint density [a recent exception can be found in Brett and Pinkse (1997)].

form of non-constant error variances (heteroskedasticity) or model coefficients (variable coefficients, spatial regimes) and can be tackled by means of the standard econometric toolbox. However, there are three reasons why it is important to consider spatial heterogeneity explicitly. First, the “structure” behind the instability is *spatial* (or geographic) in the sense that the location of the observations is crucial in determining the form of the instability. For example, groupwise heteroskedasticity could be modeled as different error variances for different compact geographic subsets of the data.

More formally, consider a set  $S$  of  $N$  geographical units (e.g., states, counties, census tracts), partitioned into  $R$  non-overlapping compact subsets  $S_r$  (with  $r = 1, 2, \dots, R$ ), such that for any  $r, s$  ( $r \neq s$ ),  $S_r \cap S_s = \emptyset$ , and  $\bigcup_{r=1}^R S_r = S$ . Spatial groupwise heteroskedasticity would then follow in the form of spatially clustered error variances for observation  $i$ ,  $\text{Var}[\varepsilon_i] = \sigma_r^2$  when  $i \in S_r$ . Similarly, variability in the regression coefficients could be specified to correspond to so-called spatial regimes, or geographic subsets  $S_r$  of the data where the model slope is different,  $\beta_i = \beta_r$ , for an observation  $i \in S_r$  [for further details, see Anselin (1988a, 1990a)]. Secondly, because the structure is spatial, heterogeneity often occurs jointly with spatial autocorrelation, and standard econometric techniques are no longer appropriate [e.g., tests for heteroskedasticity may be misleading, as illustrated in Anselin and Griffith (1988)]. Thirdly, in a single cross-section, spatial autocorrelation and spatial heterogeneity may be observationally equivalent.<sup>3</sup> For example, a spatial cluster (i.e., observed in locations that are in close proximity) of extreme residuals may be interpreted as due to spatial heterogeneity (e.g., groupwise heteroskedasticity) or as due to spatial autocorrelation (e.g., a spatial stochastic process yielding clustered values). This requires that both aspects of the problem be structured very carefully to obtain identifiability of the model parameters, and that one aspect can never be considered in isolation from the other. In the remainder of the chapter, however, the main focus of attention will be on spatial dependence.

Spatial autocorrelation, or the coincidence of value similarity with locational similarity, can be formally expressed by the moment condition

$$\text{Cov}[y_i, y_j] = E[y_i y_j] - E[y_i] \cdot E[y_j] \neq 0, \text{ for } i \neq j \quad (2.1)$$

where  $i, j$  refer to individual observations (locations) and  $y_{i(j)}$  is the value of a random variable of interest at that location. This covariance becomes meaningful from a spatial perspective when the particular configuration of nonzero  $i, j$  pairs has an interpretation in terms of spatial struc-

---

3 For similar concerns in a time series context, see, e.g., Heckman (1991).

ture, spatial interaction or the spatial arrangement of the observations.

## 2.2. *Spatial Weights and Spatial Lags*

In a cross-sectional setting with  $N$  observations, there is insufficient information to estimate the  $N$  by  $N$  covariance matrix (2.1) directly from the data. Even asymptotics are not helpful since the number of covariances increases with  $N^2$ , whereas the sample size only grows with  $N$ . In contrast, when repeated observations on the cross-section are available (as in a panel data setting), it may be possible to exploit the other dimension (i.e., asymptotics in  $T$ ) and obtain consistent nonparametric cross-sectional covariance matrix estimates.<sup>4</sup> In general however, it will be necessary to impose a structure on the covariance.

There are three main approaches followed in the literature to address this issue: one based on the specification of a spatial stochastic process; a second on the direct parametric representation of the covariance structure; and a third where the covariance is left generally unspecified and treated in a non-parametric framework. I return to this issue in the next section. First, it is important to consider the important concepts of spatial weights and spatial lag operator.

In parallel to time series analysis, spatial stochastic processes are categorized as spatial autoregressive (SAR) and spatial moving average (SMA) processes, although there are several important differences between the cross-sectional and time series contexts.<sup>5</sup> Most importantly, in contrast to the unambiguous notion of a “shift” along the time axis, there is no corresponding concept in the spatial domain, especially when observations are located irregularly in space.<sup>6</sup> Instead of the notion of shift, a *spatial lag operator* is used, which is a weighted average of random variables at “neighboring” locations.<sup>7</sup> Essential in this concept is the definition of a neighborhood set for each location. This is obtained by specifying for each location  $i$  (as the row) the neighbors as the columns corresponding to non-zero elements  $w_{ij}$  in a fixed (non-stochastic) and positive  $N$  by  $N$  spatial weights matrix  $\mathbf{W}$ .<sup>8</sup> Formally, a spatial lag for  $y$  at  $i$  is then expressed as

$$[\mathbf{W}\mathbf{y}]_i = \sum_{j=1, \dots, N} w_{ij} \cdot y_j, \quad (2.2)$$

---

4 Recent reviews of the pertinent issues are given in Frees (1995) and Driscoll and Kraay (1998).

5 See Anselin and Bera (1998) for an extensive and technical discussion.

6 On a regular lattice (e.g., square grid), one can conceive of north, south, east, west shifts [e.g., Whittle (1954)]. Such a model is often used in statistical mechanics, but it is too restrictive for economic applications.

7 In Anselin (1988a), the term spatial lag is introduced to refer to this new variable, to emphasize the similarity to a distributed lag term rather than a spatial shift.

8 By convention,  $w_{ii} = 0$ , i.e., a location is never a neighbor of itself. This is arbitrary, but can be assumed without loss of generality. For a more extensive discussion of spatial weights, see Anselin (1988a, Ch. 3), Cliff and Ord (1981), Upton and Fingleton (1985). For a formal discussion of the related notion of a spatial markov process, see Ripley (1988), Cressie (1993, pp. 402–410).

or, in matrix form, as

$$\mathbf{W}\mathbf{y} \tag{2.3}$$

where  $\mathbf{y}$  is a  $N$  by 1 vector of observations on the random variable. Since for each  $i$  the matrix elements  $w_{ij}$  are only non-zero for those  $j \in S_i$  (where  $S_i$  is the neighborhood set), only the matching  $y_j$  are included in the lag. For ease of interpretation, the elements of the spatial weights matrix are typically row-standardized, such that for each  $i$ ,  $\sum_j w_{ij} = 1$ . Consequently, the spatial lag may be interpreted as a weighted average (with the  $w_{ij}$  being the weights) of the neighbors, or as a spatial smoother.

It is important to note that the elements of the weights matrix are non-stochastic and exogenous to the model. Typically, they are based on the geographic arrangement of the observations, or contiguity. Weights are non-zero when two locations share a common boundary, or are within a given distance of each other. However, this notion is perfectly general and alternative specifications of the spatial weights can be based on distance decay (inverse distance or inverse distance squared) [Anselin (1980, Ch. 8)], on the structure of a social network [Doreian (1980)], on economic distance [Case, Rosen and Hines (1993)],  $k$  nearest neighbors [Pinkse and Slade (1998)], or on empirical flow matrices [e.g., pollutant emission flows in Murdoch, Sandler and Sargent (1997), or trade-based interaction measures in Aten (1996, 1997)]. While this wide range of options may suggest that the specification of spatial weights is largely arbitrary, this is not the case. Due to the asymptotics required to obtain consistent and asymptotically normal estimators, the range of dependence allowed by the structure of  $\mathbf{W}$  must be constrained. Also, to avoid identification problems, the weights should truly be exogenous to the model [Manski (1993)]. In spite of their lesser theoretical appeal, this explains the popularity of geographically derived weights, since exogeneity is unambiguous.

### 2.3. *Formal Expression of Spatial Autocorrelation* *Spatial Stochastic Process Models*

The most often used approach to express spatial autocorrelation is through the specification of a spatial stochastic process, i.e., a functional relationship between a random variable at a given location and this same random variable at other locations. The covariance structure then follows from the nature of the process. For example, given the  $N$  by  $N$  weights matrix  $\mathbf{W}$ , a  $N$  by 1 vector of random variables  $\mathbf{y}$ , and a  $N$  by 1 vector of i.i.d. random errors  $\boldsymbol{\varepsilon}$ , a simultaneous spatial autoregressive process is defined as

$$(\mathbf{y} - \mu\mathbf{i}) = \rho\mathbf{W}(\mathbf{y} - \mu\mathbf{i}) + \varepsilon, \text{ or } (\mathbf{y} - \mu\mathbf{i}) = (\mathbf{I} - \rho\mathbf{W})^{-1}\varepsilon \quad (2.4)$$

and a spatial moving average process as

$$\mathbf{y} = \lambda\mathbf{W}\varepsilon + \varepsilon, \text{ or } \mathbf{y} = (\mathbf{I} + \lambda\mathbf{W})\varepsilon, \quad (2.5)$$

where  $\mathbf{I}$  is a  $N$  by  $N$  identity matrix,  $\mathbf{i}$  a  $N$  by 1 vector of ones,  $\mu$  is the common mean of the random variables  $y_i$ , the i.i.d. zero mean error terms  $\varepsilon_i$  have a common variance  $\sigma^2$ , and  $\rho$  and  $\lambda$  are respectively the autoregressive and moving average parameters.<sup>9</sup> As in time series analysis, it is also possible to write a spatial autoregressive process as an infinite spatial moving average, since the elements of  $\mathbf{W}$  are typically less than one and, in addition, in most cases  $|\rho| < 1$ , such that

$$(\mathbf{y} - \mu\mathbf{i}) = \mathbf{I}\varepsilon + \rho\mathbf{W}\varepsilon + \rho^2\mathbf{W}^2\varepsilon + \dots \quad (2.6)$$

The constraints imposed by the weights structure, together with the specific form of the spatial process determine the variance-covariance matrix for  $\mathbf{y}$  as a function of two parameters, the variance  $\sigma^2$  and the spatial coefficient,  $\rho$  or  $\lambda$ . For the SAR structure in (2.4), this yields (since  $E[\mathbf{y} - \mu\mathbf{i}] = 0$ )

$$\text{Cov}[(\mathbf{y} - \mu\mathbf{i}), (\mathbf{y} - \mu\mathbf{i})] = E[(\mathbf{y} - \mu\mathbf{i})(\mathbf{y} - \mu\mathbf{i})'] = \sigma^2[(\mathbf{I} - \rho\mathbf{W})'(\mathbf{I} - \rho\mathbf{W})]^{-1}. \quad (2.7)$$

This is a full matrix, which implies that shocks at any location affect all other locations, through a *spatial multiplier* effect (or, global interaction). In contrast, for an SMA process, the covariance is (since  $E[\mathbf{y}] = 0$ )

$$\text{Cov}[\mathbf{y}, \mathbf{y}] = E[\mathbf{y}\mathbf{y}'] = \sigma^2(\mathbf{I} + \rho\mathbf{W})(\mathbf{I} + \rho\mathbf{W})' = \sigma^2[\mathbf{I} + \rho(\mathbf{W} + \mathbf{W}') + \rho^2\mathbf{W}\mathbf{W}'], \quad (2.8)$$

which results in “local” interaction between a location and its first (through  $\mathbf{W}$ ) and second order neighbors (through  $\mathbf{W}\mathbf{W}'$ ), but does not yield a spatial multiplier as in the SAR model.

A major distinction between AR and MA processes in space compared to the time domain is that even with i.i.d. error terms  $\varepsilon_i$ , the diagonal elements in (2.7) and (2.8) are not constant. Furthermore, the heteroskedasticity depends on the neighborhood structure embedded in the spatial weights matrix  $\mathbf{W}$ . Consequently, the process in  $\mathbf{y}$  is not covariance-stationary. Stationarity is only obtained in very rare cases, for example on regular lattice structures when each observation has an identical weights structure, but this is of limited practical use. This lack of sta-

---

<sup>9</sup> In spatial econometrics the main focus is on processes for which the joint distribution is modeled, or, a simultaneous processes, in contrast to conditional processes. Simultaneity is more appropriate when the interest is in inference, whereas a conditional approach is natural when the interest is in prediction or spatial interpolation [see Anselin and Bera (1998) and Cressie (1993, p. 410) for technical details].

tionarity has important implications for the types of central limit theorems and laws of large numbers that need to be invoked to obtain asymptotic properties for estimators and specification test, a point that has not always been recognized in the literature.

A related specification, which also results in a heteroskedastic variance, is the spatial error components model outlined in Kelejian and Robinson (1993, 1995, 1997). This spatial process is very similar to a spatial MA process, but rather than pertaining to a single error term, it contains two uncorrelated zero mean error components  $\psi$  and  $\varepsilon$ :

$$\mathbf{y} = \mathbf{W}\psi + \varepsilon \quad (2.9)$$

where the term  $\mathbf{W}\psi$  implies a smoothing of neighboring values or “regional” effect, and the error term  $\varepsilon$  is location-specific. The resulting variance-covariance matrix is (since  $E[\mathbf{y}] = 0$ )

$$\text{Cov}[\mathbf{y}, \mathbf{y}] = E[\mathbf{y}\mathbf{y}'] = \sigma_\varepsilon^2 \mathbf{I} + \sigma_\psi^2 \mathbf{W}\mathbf{W}' \quad (2.10)$$

where  $\sigma_\varepsilon^2$  and  $\sigma_\psi^2$  are the variance components pertaining to the locational “innovation” and the regional effect respectively. The range of interaction implied by (2.10) is a subset of that in the spatial MA model, as seen from (2.8).

#### *Direct Representation*

A second commonly used approach to the formal specification of spatial autocorrelation is to express the elements of the variance-covariance matrix in a parsimonious fashion as a “direct” function of a small number of parameters and one or more exogenous variables. Typically, this involves an inverse function of some distance metric, for example,

$$\text{Cov}[\varepsilon_i \varepsilon_j] = \sigma^2 f(d_{ij}, \varphi), \quad (2.11)$$

where  $\varepsilon_i$  and  $\varepsilon_j$  are regression disturbance terms,  $\sigma^2$  is the error variance,  $d_{ij}$  is the distance separating observations (locations)  $i$  and  $j$ , and  $f$  is a distance decay function such that  $\frac{\partial f}{\partial d} < 0$  and  $|f(d_{ij}, \varphi)| \leq 1$ , with  $\varphi \in \Phi$  as a  $p$  by 1 vector of parameters on an open subset  $\Phi$  of  $R^p$ . This form is closely related to the variogram model used in geostatistics, although with stricter assumptions regarding stationarity and isotropy.<sup>10</sup> Using (2.11) for individual elements, the full error covariance matrix follows as

---

10 The specification of spatial covariance functions is not arbitrary, and a number of conditions must be satisfied in order for the model to be “valid” [details are given in Cressie (1993, pp. 61–63, 67–68 and 84–86)]. For a stationary process, the covariance function must be positive definite and it is required that  $f \rightarrow 0$  as  $d_{ij} \rightarrow \infty$ , which is ensured by the conditions spelled out here. Furthermore, most models only consider positive spatial autocorrelation in this context. An exception is the so-called wave variogram, which allows both positive and negative correlation due to the periodicity of the process. This model has not seen application outside the physical sciences and is not considered here.

$$E[\varepsilon\varepsilon'] = \sigma^2\Omega(d_{ij}, \varphi) \quad (2.12)$$

where, because of the scaling factor  $\sigma^2$ , the matrix  $\Omega(d_{ij}, \varphi)$  must be a positive definite spatial correlation matrix, with  $\omega_{ii} = 1$  and  $|\omega_{ij}| \leq 1, \forall i, j$ .<sup>11</sup> Note that, in contrast to the variance for the spatial process models (2.7), (2.8) and (2.10), the direct representation model does not induce heteroskedasticity.

In spatial econometrics, models of this type have been used primarily in the analysis of urban housing markets, e.g., in Dubin (1988, 1992), Olmo (1995) and Basu and Thibodeau (1998). While this specification has a certain intuition, in the sense that it incorporates an explicit notion of spatial clustering as a function of the distance separating two observations (i.e., positive spatial correlation), it is also fraught with a number of estimation and identification problems [Anselin (1999)].

#### *Nonparametric Approaches*

A recently suggested alternative specification of the spatial covariance matrix is to estimate its elements non-parametrically, in the sense of not requiring an explicit spatial process or functional form for the distance decay. This is common in the case of panel data, when the time dimension is (considerably) greater than the cross-sectional dimension ( $T \gg N$ ) and the “spatial” covariance is estimated from the sample covariance for the residuals of each set of location pairs [e.g., in applications of Zellner’s SUR estimator, see Fiebig (1999)].

The new approaches are variants of the well known Newey-West (1987) heteroskedasticity and autocorrelation consistent covariance matrix and have been used in the context of generalized methods of moments (GMM) estimators of spatial regression models (see Section 4.3.). Conley (1996) suggested a covariance estimator based on a sequence of weighted averages of sample autocovariances computed for subsets of observation pairs that fall within a given distance band (or spatial window). Although not presented as such, this has a striking similarity to the non-parametric estimation of a semi-variogram in geostatistics [see, e.g., Cressie (1993), pp.69–70], but the assumptions of stationarity and isotropy required in the GMM approach are stricter than those needed in variogram estimation. In a panel data setting, Driscoll and Kraay (1998) use a similar idea, but avoid having to estimate the spatial covariances by distance bands. This is accomplished by using only the cross-sectional averages (for each time period) of the moment conditions, and by relying on asymptotics in the time dimension to yield an estimator for the spatial covariance structure.

---

11  $\omega_{ii} = 1$  is ensured by selecting a functional form for  $f$  such that  $f(d_{ij}, \varphi) = 1$  for  $d_{ij} = 0$ .

#### 2.4. *Asymptotics in Spatial Stochastic Processes*

As in time series analysis, the properties of estimators and tests for spatial series are derived from the asymptotics for stochastic processes. However, these properties are not simply extensions to two dimensions of the time series results. A number of complicating factors are present and to date some formal results for the spatial dependence case are still lacking. While an extensive treatment of this topic is beyond the scope of the current chapter, three general comments are in order. First, the intuition behind the asymptotics is fairly straightforward in that regularity conditions are needed to limit the extent of spatial dependence (memory) and heterogeneity of the spatial series in order to obtain appropriate (uniform) laws of large numbers and central limit theorems to establish consistency and asymptotic normality. In this context, it is important to keep in mind that both SAR and SMA processes yield heteroskedastic variances, so that the application of results for dependent *stationary* series are not appropriate.<sup>12</sup> In addition to the usual moment conditions that are similar in spirit to those for heterogeneous dependent processes in time [e.g., Pötscher and Prucha (1997)], specific spatial conditions will translate into constraints on the spatial weights and on the parameter space for the spatial coefficients [for some specific examples, see, e.g., Anselin and Kelejian (1997), Kelejian and Prucha (1998, 1999a, 1999b), Pinkse and Slade (1998), and Pinkse (1998, 1999)]. In practice, these conditions are likely satisfied by most spatial weights that are based on simple contiguity, but this is not necessarily the case for general weights based on economic distance.

A second distinguishing characteristic of asymptotics in space is that the limit may be approached in two different ways, referred to as *increasing domain* asymptotics and *infill* asymptotics.<sup>13</sup> The former consists of a sampling structure where new “observations” are added at the edges (boundary points), similar to the underlying asymptotics in time series analysis. Infill asymptotics are appropriate when the spatial domain is bounded, and new observations are added in between existing ones, generating an increasingly denser surface. Many results for increasing domain asymptotics are not directly applicable to infill asymptotics [Lahiri (1996)]. In most applications of spatial econometrics, the implied structure is that of an increasing domain.

Finally, for spatial processes that contain spatial weights, the asymptotics require the use of CLT and LLN for triangular arrays [Davidson (1994, Ch. 19, 24)]. This is caused by the fact that for the boundary elements the “sample” weights matrix changes as new data points are added

---

12 Specifically, this may limit the applicability of GMM estimators that are based on a central limit theorem for stationary mixing random fields such as the one by Bolthausen (1982), used by Conley (1996).

13 Cressie (1993, pp. 100–101).

[i.e., the new data points change the connectedness structure for existing data points].<sup>14</sup> Again, this is an additional degree of complexity, which is not found in time series models.

### 3. Spatial Regression Models

#### 3.1. Spatial Lag and Spatial Error Models

In the standard linear regression model, spatial dependence can be incorporated in two distinct ways: as an additional regressor in the form of a spatially lagged dependent variable ( $\mathbf{W}\mathbf{y}$ ), or in the error structure ( $E[\varepsilon_i\varepsilon_j] \neq 0$ ). The former is referred to as a *spatial lag* model and is appropriate when the focus of interest is the assessment of the existence and strength of spatial interaction. This is interpreted as substantive spatial dependence in the sense of being directly related to a *spatial model* [e.g., a model that incorporates spatial interaction, yardstick competition, etc.]. Spatial dependence in the regression disturbance term, or a *spatial error* model can take any of the forms outlined in Section 2.3. and is referred to as nuisance dependence. This is appropriate when the concern is with correcting for the potentially biasing influence of the spatial autocorrelation, due to the use of *spatial data* (irrespective of whether the model of interest is spatial or not).

Formally, a spatial lag model, or a mixed regressive, spatial autoregressive model is expressed as

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (3.1)$$

where  $\rho$  is a spatial autoregressive coefficient,  $\boldsymbol{\varepsilon}$  is a vector of error terms, and the other notation is as before.<sup>15</sup> Unlike what holds for the time series counterpart of this model, the spatial lag term  $\mathbf{W}\mathbf{y}$  is correlated with the disturbances, even when the latter are i.i.d. This can be seen from the reduced form of (3.1),

$$\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\varepsilon} \quad (3.2)$$

in which each inverse can be expanded into an infinite series as in (2.6), including both the explanatory variables and the error terms at all locations (the spatial multiplier). Consequently, the spatial lag term must be treated as an endogenous variable and proper estimation methods must account for this endogeneity (OLS will be biased and inconsistent due to the simultaneity bias).

A spatial error model is a special case of a regression with a non-spherical error term, in which the off-diagonal elements of the covariance matrix express the structure of spatial depen-

---

14 See Kelejian and Prucha (1999a, 1999b).

15 For ease of exposition, the error term is assumed to be i.i.d., although various forms of heteroskedasticity can be incorporated in a straightforward way [Anselin (1988a), Chapter 6].

dence. Consequently, OLS remains unbiased, but it is no longer efficient and the classical estimators for standard errors will be biased. As outlined in Section 2.3., the spatial structure can be specified in a number of different ways, and (except for the non-parametric approaches) results in a error variance covariance matrix of the form

$$E[\varepsilon\varepsilon'] = \Omega(\theta), \quad (3.3)$$

where  $\theta$  is a vector of parameters, such as the coefficients in a SAR or SMA error process. When the error process is SAR, the resulting model can also be expressed as a spatial lag specification, in the form of spatial Durbin or spatial common factor model [Anselin (1980)]. The SAR error model is

$$\mathbf{y} = \mathbf{X}\beta + \varepsilon \text{ and } \varepsilon = \lambda\mathbf{W}\varepsilon + \mathbf{u}. \quad (3.4)$$

Since  $\varepsilon = (\mathbf{I} - \lambda\mathbf{W})^{-1}\mathbf{u}$ , and thus  $\mathbf{y} = \mathbf{X}\beta + (\mathbf{I} - \lambda\mathbf{W})^{-1}\mathbf{u}$ , (3.4) is equivalent to

$$\mathbf{y} = \lambda\mathbf{W}\mathbf{y} + \mathbf{X}\beta - \lambda\mathbf{W}\mathbf{X}\beta + \varepsilon, \quad (3.5)$$

which is a spatial lag model with an additional set of spatially lagged exogenous variables ( $\mathbf{W}\mathbf{X}$ ) and a set of  $k$  nonlinear (common factor) constraints on the coefficients (the product of the spatial autoregressive coefficient with the regression coefficients  $\beta$  should equal the negative of the coefficients of  $\mathbf{W}\mathbf{X}$ ). The similarity between the error model (3.5) and the “pure” spatial lag model (3.1) will complicate specification testing in practice, since tests designed for a spatial lag alternative will also have power against a spatial error alternative, and vice versa.

Most spatial regression models used in practice are based on a single spatial weights matrix. However, in principle, higher order models are possible as well, such as higher order SAR models [e.g., Brandsma and Ketellapper (1979), Blommestein (1983, 1985)], spatial autoregressive, moving average specifications or SARMA [Huang (1984)], and models that include both a spatially lagged dependent variable as well as a SAR error process [Case (1992)]. In the specification of such higher order models, care must be taken to ensure that the weights are unique, orthogonal (non-overlapping) and that all coefficients are identified.<sup>16</sup>

### 3.2. *Spatial Dependence in Panel Data Models*

When observations are available across space as well as over time, the additional dimension allows the estimation of the full covariance of one type of association, using the other dimension to provide the asymptotics (e.g., in SUR models with  $N \ll T$ ). However, as in the pure cross-sectional case, there is insufficient information in the  $NT$  observations to estimate the complete

---

16 For a more extensive discussion, see Blommestein (1985), Anselin and Smirnov (1996), and Anselin and Bera (1998, pp. 251–252).

$(NT)^2$  covariance matrix  $\text{Cov}[y_{it}y_{js}] \neq 0$ , (with  $i \neq j$  and  $t \neq s$ ) without imposing some structure. For small  $N$  and large  $T$ , the asymptotics in the time domain can be exploited to obtain a non-parametric estimate of cross-sectional dependence, while time dependence must be parameterized. Similarly, for large  $N$  and small  $T$ , the asymptotics in the spatial domain can be exploited to yield a non-parametric estimate of serial (time) dependence, while spatial dependence must be parameterized. As in the pure cross-sectional case, the latter requires the use of a spatial weights matrix. In each of these situations, asymptotics are only needed in one of the dimensions while the other can be treated as fixed.

When both spatial as well as serial dependence are parameterized, a range of specifications can be considered, allowing different combinations of the two. For ease of exposition, assume that the observations are stacked by time period, i.e., they can be considered as  $T$  time slices of  $N$  cross-sectional units. Restricting attention to “lag” dependence, and with  $f(z)$  as a generic designation for the regressors (which may be lagged in time and/or space), four types of models can be distinguished.

- (a) *pure space-recursive*, in which the dependence pertains to neighboring locations in a different period, or,

$$y_{it} = \gamma[\mathbf{W}\mathbf{y}_{t-1}]_i + f(z) + \varepsilon_{it}, \quad (3.6)$$

where, using the same notational convention as in (2.2),  $[\mathbf{W}\mathbf{y}_{t-1}]_i$  is the  $i$ -th element of the spatial lag vector applied to the observations on the dependent variable in the previous time period (using a  $N$  by  $N$  spatial weights matrix for the cross-sectional units).

- (b) *time-space recursive*, in which the dependence relates to the same location as well as the neighboring locations in another period, or,

$$y_{it} = \lambda y_{it-1} + \gamma[\mathbf{W}\mathbf{y}_{t-1}]_i + f(z) + \varepsilon_{it} \quad (3.7)$$

- (c) *time-space simultaneous*, with both a time-wise and a spatially lagged dependent variable, or,

$$y_{it} = \lambda y_{it-1} + \rho[\mathbf{W}\mathbf{y}_t]_i + f(z) + \varepsilon_{it} \quad (3.8)$$

where  $[\mathbf{W}\mathbf{y}_t]_i$  is the  $i$ -th element of the spatial lag vector in the same time period.

- (d) *time-space dynamic*, with all forms of dependence, or,

$$y_{it} = \lambda y_{it-1} + \rho[\mathbf{W}\mathbf{y}_t]_i + \gamma[\mathbf{W}\mathbf{y}_{t-1}]_i + f(z) + \varepsilon_{it}. \quad (3.9)$$

In order to estimate the parameters of the time-space simultaneous model, asymptotics are

needed in the cross-sectional dimension, while for the time-space dynamic model, asymptotics are needed in both dimensions. For the other models, the type of asymptotics required are determined by the dependence structure in the error terms. For example, the pure space-recursive model with i.i.d. errors satisfies the assumptions of the classical linear model and can be estimated by means of OLS.

Spatial lag and spatial error dependence can be introduced into the cross-sectional dimension of traditional panel data models in a straightforward way. For example, in a spatial SUR model, both autoregressive as well as regression parameters are allowed to vary by time period, in combination with a non-parametric serial covariance. The spatial lag formulation of such a model would be (in the same notation as before):

$$y_{it} = \rho_t[\mathbf{W}y_t]_i + x'_{it}\beta_t + \varepsilon_{it} \quad (3.10)$$

with  $\text{Var}[\varepsilon_{it}] = \sigma_t^2$  and  $E[\varepsilon_{it}\varepsilon_{is}] = \sigma_{ts}$ .<sup>17</sup>

An important issue to consider when incorporating spatial dependence in panel data models is the extent to which fixed effects may be allowed. Since the estimation of the spatial process models requires asymptotics in the cross-sectional domain ( $N \rightarrow \infty$ ), fixed effects (i.e., a dummy variable for each location) would suffer from the incidental parameter problem and no consistent estimator exists. Hence, fixed cross-sectional effects are incompatible with spatial processes and instead a random effects specification must be considered.

### 3.3. *Spatial Dependence in Models for Qualitative Data*

Empirical analysis of interacting agents requires models that incorporate spatial dependence for discrete dependent variables, such as counts or binary outcomes [Brock and Durlauf (1995)]. This turns out to be quite complex and continues to be an active area of research.<sup>18</sup> While an extensive discussion of the technical aspects associated with spatial discrete choice models is beyond the scope of the current chapter, the salient issues may be illustrated with a spatial version of the probit model, which has recently received considerable attention.<sup>19</sup>

The point of departure is the familiar expression for a linear model in a latent (unobserved) dependent variable  $y_i^*$

$$y_i^* = x_i'\beta + \varepsilon_i \quad (3.11)$$

---

17 For further details, see Anselin (1988a, 1988b). A recent application is Baltagi and Li (1999).

18 For an extensive treatment of these so-called auto-models, see Besag (1974), and also Cressie and Read (1989), Cressie (1993, pp. 427–428), and Kaiser and Cressie (1997).

19 Methodological issues associated with spatial probit models are considered in Case (1992), McMillen (1992, 1995), Bolduc, Fortin and Gordon (1997), Heagerty and Lele (1998), Pinkse and Slade (1998) and Beron and Vijverberg (1999).

where  $\varepsilon_i$  is a random variate for which a given distribution is assumed (e.g., the normal for the probit model). The realization of  $y_i^*$  is observed in the form of discrete events,  $y_i = 1$  for  $y_i^* \geq 0$ , and  $y_i = 0$  for  $y_i^* < 0$ . The discrete events are related to the underlying probability model through the error term, for example,  $y_i^* \geq 0$  implies  $-x_i'\beta < \varepsilon_i$ , and, therefore,

$$E[y_i] = P[y_i = 1] = \Phi[x_i'\beta] \quad (3.12)$$

where  $\Phi$  is the cumulative distribution function for the standard normal.

Spatial autocorrelation can be introduced into this model in the form of a spatial autoregressive process for the error term  $\varepsilon_i$  in (3.11), or

$$\varepsilon_i = \lambda \sum_j w_{ij} \varepsilon_j + u_i \quad (3.13)$$

where  $\lambda$  is an autoregressive parameter, the  $w_{ij}$  are the elements in the  $i$ -th row of a spatial weights matrix, and  $u_i$  may be assumed to be i.i.d. standard normal. As a consequence of the spatial multiplier in the autoregressive specification, the random error at each location now becomes a function of the random errors at all other locations as well. Its distribution is multivariate normal with  $N$  by  $N$  variance-covariance matrix

$$E[\varepsilon\varepsilon'] = [(\mathbf{I} - \lambda\mathbf{W})'(\mathbf{I} - \lambda\mathbf{W})]^{-1}, \quad (3.14)$$

As pointed out in Section 2.3., besides being non-diagonal, (3.14) is also heteroskedastic. Consequently, the usual inequality conditions that are at the basis of (3.12) no longer hold, and  $P[-x_i'\beta < \varepsilon_i]$  cannot be derived from the univariate standard normal distribution, but rather must be expressed explicitly as the marginal distribution of a  $N$ -dimensional multivariate normal vector, whose variance-covariance matrix contains the autoregressive parameter  $\lambda$ . This is non-standard and typically not analytically tractable, which greatly complicates estimation and specification testing. Similar issues are faced in the spatial lag model for a latent variable.<sup>20</sup>

When the dependent variable is a count or proportion, the latent variable model is not used, but instead the dependent variables is modeled directly as a random variable with a Poisson or binomial distribution. Introducing spatial dependence into these models is not trivial, since the multivariate versions of discrete distributions are not as tractable as the multivariate normal. This is typically handled by means of hierarchical Bayesian modeling and estimation is based on computation intensive methods such as Gibbs sampling and Markov Chain Monte Carlo.<sup>21</sup>

---

20 For an extensive discussion, see Beron and Vijverberg (1999).

21 For a recent review of the issues, see, e.g., Clayton and Kaldor (1987), Besag, York and Mollie (1991), Besag et al. (1995), Waller, Carlin and Xia (1997), Waller et al. (1997) and Diggle, Tawn and Moyeed (1998).

## 4. Estimation

### 4.1. Maximum Likelihood Estimation

Maximum likelihood (ML) estimation of spatial lag and spatial error regression models was first outlined by Ord (1975).<sup>22</sup> The point of departure is an assumption of normality for the error terms  $\varepsilon_i$  in (3.1) or  $u_i$  in (3.4). The joint likelihood then follows from the multivariate normal distribution for  $\mathbf{y}$ . Unlike what holds for the classic regression model, the joint log likelihood for a spatial regression does not equal the sum of the log likelihoods associated with the individual observations. This is due to the two-directional nature of the spatial dependence, which results in a Jacobian term that is the determinant of a full  $N$  by  $N$  matrix, e.g.,  $|\mathbf{I} - \rho\mathbf{W}|$  (spatial lag).

For the SAR error model, the log likelihood is based on the multivariate normal case, e.g., as used in the general treatments of Magnus (1978) and Breusch (1980). Since  $\varepsilon \sim \text{MVN}(0, \Sigma)$ , it follows that, with  $\varepsilon = \mathbf{y} - \mathbf{X}\beta$  and  $\Sigma = \sigma^2[(\mathbf{I} - \lambda\mathbf{W})'(\mathbf{I} - \lambda\mathbf{W})]^{-1}$ ,

$$\begin{aligned} \ln L = & -(N/2)\ln(2\pi) - (N/2)\ln\sigma^2 + \ln|\mathbf{I} - \lambda\mathbf{W}| \\ & - (1/2\sigma^2)(\mathbf{y} - \mathbf{X}\beta)'(\mathbf{I} - \lambda\mathbf{W})'(\mathbf{I} - \lambda\mathbf{W})(\mathbf{y} - \mathbf{X}\beta) \end{aligned} \quad (4.1)$$

Closer inspection of the last term in (4.1) reveals that, conditional upon  $\lambda$ , a maximization of the log likelihood is equivalent to the minimization of the sum of squared residuals in a regression of a spatially filtered dependent variable  $\mathbf{y}^* = \mathbf{y} - \lambda\mathbf{W}\mathbf{y}$  on a set of spatially filtered explanatory variables  $\mathbf{X}^* = \mathbf{X} - \lambda\mathbf{W}\mathbf{X}$ . The first order conditions for  $\hat{\beta}_{\text{ML}}$  indeed yield the familiar generalized least squares estimator:

$$\hat{\beta}_{\text{ML}} = [(\mathbf{X} - \lambda\mathbf{W}\mathbf{X})'(\mathbf{X} - \lambda\mathbf{W}\mathbf{X})]^{-1}(\mathbf{X} - \lambda\mathbf{W}\mathbf{X})'(\mathbf{y} - \lambda\mathbf{W}\mathbf{y}) \quad (4.2)$$

and, similarly, the ML estimator for  $\sigma^2$  follows as:

$$\hat{\sigma}_{\text{ML}}^2 = (\mathbf{e} - \lambda\mathbf{W}\mathbf{e})'(\mathbf{e} - \lambda\mathbf{W}\mathbf{e})/N \quad (4.3)$$

with  $\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\beta}_{\text{ML}}$ . However, unlike the time series case, a consistent estimator for  $\lambda$  cannot be obtained from the OLS residuals and therefore the standard two-step FGLS approach does not apply.<sup>23</sup> Instead, the estimator for  $\lambda$  must be obtained from an explicit maximization of a concentrated likelihood function [for details, see Anselin (1988a, Chapter 6) and Anselin and Bera (1998)]. Also, while the SAR error model estimator (4.2) yields an expression in spatially filtered

---

22 Other classic treatments of ML estimation in spatial models can be found in Whittle (1954), Besag (1974), Mardia and Marshall (1984), Warnes and Ripley (1987), and Mardia and Watkins (1989).

23 For a formal demonstration, see Anselin (1988a). In Kelejian and Prucha (1997) it is also shown that an approach based on the spatial Durbin specification yields an inconsistent estimator.

variables, this is not the case for the SMA error model, since its variance (2.8) requires the inverse of a  $N$  by  $N$  matrix in order to obtain a FGLS expression. This model has therefore not seen much application in empirical practice [see, e.g., Sneek and Rietveld (1999)].<sup>24</sup>

The log-likelihood for the spatial lag model is obtained using the same general principles [see Anselin (1988, Chapter 6) for details] and takes the form

$$\begin{aligned} \ln L = & -(N/2)\ln(2\pi) - (N/2)\ln\sigma^2 + \ln|\mathbf{I} - \rho\mathbf{W}| \\ & - (1/2\sigma^2)(\mathbf{y} - \rho\mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta})'(\mathbf{y} - \rho\mathbf{W}\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \end{aligned} \quad (4.4)$$

The minimization of the last term in (4.4) corresponds to OLS, but since this ignores the log Jacobian  $\ln|\mathbf{I} - \rho\mathbf{W}|$ , OLS is not a consistent estimator in this model. As in the spatial error model, there is no satisfactory two-step procedure and estimators for the parameters must be obtained from an explicit maximization of the likelihood. This is greatly simplified since both  $\hat{\boldsymbol{\beta}}_{\text{ML}}$  and  $\hat{\sigma}_{\text{ML}}^2$  can be obtained conditional upon  $\rho$  from the first order conditions:

$$\hat{\boldsymbol{\beta}}_{\text{ML}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{y} - \lambda\mathbf{W}\mathbf{y}) \quad (4.5)$$

or, with  $\hat{\boldsymbol{\beta}}_0 = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$ ,  $\mathbf{e}_0 = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_0$ ,  $\hat{\boldsymbol{\beta}}_L = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{W}\mathbf{y}$ ,  $\mathbf{e}_L = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}}_L$ ,

$$\hat{\boldsymbol{\beta}}_{\text{ML}} = \hat{\boldsymbol{\beta}}_0 - \rho\hat{\boldsymbol{\beta}}_L \quad (4.6)$$

and

$$\hat{\sigma}_{\text{ML}}^2 = (\mathbf{e}_0 - \rho\mathbf{e}_L)'(\mathbf{e}_0 - \rho\mathbf{e}_L)/N \quad (4.7)$$

This yields a concentrated log-likelihood in a single parameter, which is straightforward to optimize by means of direct search techniques [see Anselin (1980, 1988a) for derivations and details].

Both spatial lag and spatial error models are special cases of a more general specification that may include forms of heteroskedasticity as well. This also provides the basis for ML estimation of spatial SUR models with spatial lag or spatial error terms [Anselin (1980, Ch. 10)]. Similarly, ML estimation of error components models with spatial lag or spatial error terms can be implemented as well. Spatial models with discrete dependent variables are typically not estimated by means of ML, given the prohibitive nature of evaluating multiple integrals to determine the relevant marginal distributions.<sup>25</sup>

---

24 Note that the spatial error components specification (2.10) and the direct representation models (2.12) also suffer from this problem.

25 For details, see, e.g., McMillen (1992), Pinkse and Slade (1998), Beron and Vijverberg (1999), and also, for general principles, Poirier and Ruud (1988).

Finally, it is important to note that models with spatial dependence do not fit the classical framework [e.g., as outlined in Rao (1973)] under which the optimal properties (consistency, asymptotic efficiency, asymptotic normality) of ML estimators are established. This implies that these properties do not necessarily hold and that careful consideration must be given to the explicit formulation of regularity conditions. In general terms, aside from the usual restrictions on the variance and higher moments of the model variables, these conditions boil down to constraints on the range of dependence embodied in the spatial weights matrix.<sup>26</sup> In addition, to avoid singularity or explosive processes, the parameter space for the coefficient in a spatial process model is restricted to an interval other than the familiar  $-1, +1$ . For example, for a SAR process, the parameter space is  $1/\omega_{\min} < \rho < 1/\omega_{\max}$ , where  $\omega_{\min}$  and  $\omega_{\max}$  are the smallest (on the real line) and largest eigenvalues of the spatial weights matrix  $\mathbf{W}$ . For row-standardized weights,  $\omega_{\max} = 1$ , but  $\omega_{\min} > -1$ , such that the lower bound on the parameter space is less than  $-1$  [Anselin (1980)]. This must be taken into account in practical implementations of estimation routines.

#### 4.2. *Spatial Two Stage Least Squares*

The endogeneity of the spatially lagged dependent variable can also be addressed by means of an instrumental variables or two stage least squares approach [Anselin (1980, 1988a, 1990b), Land and Deane (1992), Kelejian and Robinson (1993), Kelejian and Prucha (1998)]. As demonstrated in Kelejian and Robinson (1993), the choice of an instrument for  $\mathbf{W}\mathbf{y}$  follows from the conditional expectation in the reduced form (3.2),

$$E[\mathbf{y} | \mathbf{X}] = (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X}\beta = \mathbf{X}\beta + \rho \mathbf{W}\mathbf{X}\beta + \rho^2 \mathbf{W}^2 \mathbf{X}\beta + \dots \quad (4.8)$$

Apart from the exogenous variables  $\mathbf{X}$  (which are always instruments), this includes their spatial lags as well, suggesting  $\mathbf{W}\mathbf{X}$  as a set of instruments.<sup>27</sup>

Under a set of reasonable assumptions that are easily satisfied when the spatial weights are based on contiguity, the spatial two stage least squares estimator achieves the consistency and asymptotic normality properties of the standard 2SLS [see, e.g., the theorems spelled out in Schmidt (1976)].<sup>28</sup> A straightforward extension is the application of 3SLS to the spatial SUR model with a spatial lag [Anselin (1988a, Chapter 10)]. However, 2SLS is not appropriate to obtain a consistent estimator for the nuisance parameter in a spatial error model, for example based on the spatial Durbin formulation [(3.5)], as demonstrated by Kelejian and Prucha (1997).

26 For a careful consideration of these issues, see Kelejian and Prucha (1999a).

27 Higher lags may be included as well to improve precision of the estimator, although this may also increase its bias. When other (i.e., non-spatial) endogenous variables are included in the specification, they will require instruments as well, in the standard fashion.

28 For technical details, see, e.g., Kelejian and Robinson (1993), Kelejian and Prucha (1998).

The spatial 2SLS model is extended to include spatial error components in Kelejian and Robinson (1993) and SAR errors in Kelejian and Prucha (1998).

#### 4.3. *Method of Moments Estimators*

Recently, a number of approaches have been outlined to estimate the coefficients in a spatial error model as an application of general principles underlying the method of moments. Kelejian and Prucha (1999a) develop a set of moment conditions that yield estimation equations for the parameter of a SAR error model. Specifically, assuming an i.i.d. error vector for  $\mathbf{u}$  in (3.4), the following three conditions readily follow

$$\begin{aligned} E[\mathbf{u}'\mathbf{u}/N] &= \sigma^2 \\ E[\mathbf{u}'\mathbf{W}'\mathbf{W}\mathbf{u}/N] &= \sigma^2(1/N)\text{tr}(\mathbf{W}'\mathbf{W}) \\ E[\mathbf{u}'\mathbf{W}\mathbf{u}/N] &= 0 \end{aligned} \tag{4.9}$$

where  $\text{tr}$  is the matrix trace operator. Replacing  $\mathbf{u}$  by  $\mathbf{e} - \lambda\mathbf{W}\mathbf{e}$  (with  $\mathbf{e}$  as the vector of OLS residuals) in (4.9) yields a system of three equations in the parameters  $\lambda$ ,  $\lambda^2$  and  $\sigma^2$ . Kelejian and Prucha (1999a) suggest the use of nonlinear least squares to obtain a consistent generalized moment estimator for  $\lambda$  from this system, which can then be used to obtain consistent estimators for the  $\beta$  in a FGLS approach. Since the  $\lambda$  is considered as a nuisance parameter, its significance (as a test for spatial autocorrelation) cannot be assessed, but its role is to provide a consistent estimator for the regression coefficients.<sup>29</sup>

A different approach is taken in the application of Hansen's (1982) generalized method of moments estimator (GMM) to spatial error autocorrelation in Conley (1996). This estimator is the standard minimizer of a quadratic form in the sample moment conditions, where the covariance matrix is obtained in non-parametric form as an application of the ideas of Newey and West (1987). Specifically, the spatial covariances are estimated from weighted averages of sample covariances for pairs of observations that are within a given distance band from each other. Note that this approach requires covariance stationarity, which is only satisfied for a restricted set of spatial processes (e.g., it does not apply to SAR error models).

Pinkse and Slade (1998) use a set of moment conditions to estimate a probit model with SAR errors. However, they focus on the induced heteroskedasticity of the process and do not explicitly deal with the spatial covariance structure.<sup>30</sup>

---

29 A recent application of this method is given in Bell and Bockstael (1999). An extension of this idea to the residuals of a spatial 2SLS estimation is provided in Kelejian and Prucha (1998).

30 See also Case (1992) and McMillen (1992) for a similar focus on heteroskedasticity in the spatial probit model.

#### 4.4. Other Estimation Methods

A number of other approaches have been suggested to deal with the estimation of spatial regression models. An early technique is the so-called coding method, originally examined in Besag and Moran (1975) and Besag (1977).<sup>31</sup> This approach consists of selecting a subsample from the data such that the relevant neighbors are removed (a non-contiguous subsample). This in effect eliminates the simultaneity bias in the spatial lag model, but at the cost of converting the model to a conditional one and with a considerable reduction of the sample size (down to 20% of the original sample for irregular lattice data). The advantage of this approach is that standard methods may be applied (e.g., for discrete choice models). However, it is not an efficient procedure and considerable arbitrariness is involved in the selection of the coding scheme.

Another increasingly common approach consists of the application of computational estimators to spatial models. A recent example is the recursive importance sampling (RIS) estimator [Vijverberg (1997)] applied to the spatial probit model in Beron and Vijverberg (1999).

A considerable literature also exists on Bayesian estimation of spatial models, but a detailed treatment of this is beyond the current scope.<sup>32</sup>

## 5. Specification Tests

### 5.1. Moran's $I$

The most commonly used specification test for spatial autocorrelation is derived from a statistic developed by Moran (1948, 1950a, 1950b) as the two-dimensional analog of a test for univariate time series correlation [see also Cliff and Ord (1972, 1973)]. In matrix notation, Moran's  $I$  statistic is

$$I = (N/S_0)(\mathbf{e}'\mathbf{W}\mathbf{e}/\mathbf{e}'\mathbf{e}) \quad (5.1)$$

with  $\mathbf{e}$  as a vector of OLS residuals and  $S_0 = \sum_i \sum_j w_{ij}$ , a standardization factor that corresponds to the sum of the weights for the non-zero cross-products. The statistic shows a striking similarity to the familiar Durbin-Watson test.<sup>33</sup> Inference for Moran's  $I$  is based on a normal approximation, using a standardized z-value obtained from expressions for the mean and variance of the statistic [Cliff and Ord (1972, 1981)]. Alternatively, as shown in Tiefelsdorf and Boots (1995), an exact test can be constructed based on numerical integration and the principles outlined in Imhof (1961) and Koerts and Abrahamse (1968).

---

31 See also the discussion in Haining (1990, pp. 131–133).

32 For a review, see the references in footnote 21, and also Hepple (1995) and LeSage (1997).

33 For example, for row-standardized weights,  $S_0 = N$ , and  $I = \mathbf{e}'\mathbf{W}\mathbf{e}/\mathbf{e}'\mathbf{e}$ . See Anselin and Bera (1998) for an extensive discussion.

Moran's  $I$  test has been shown to be locally best invariant [King (1981)] and consistently outperforms other tests in terms of power in simulation experiments [see, e.g., Bartels and Hordijk (1977), Anselin and Rey (1991), Anselin and Florax (1995b), Kelejian and Robinson (1998)]. Its application has been extended to residuals in 2SLS regression in Anselin and Kelejian (1997), and to generalized residuals in probit models in Pinkse (1998, 1999). General formal conditions and proofs for the asymptotic normality of Moran's  $I$  in a wide range of regression models are given in Pinkse (1998) and Kelejian and Prucha (1999b). The consideration of Moran's  $I$  in conjunction with spatial heteroskedasticity is covered in Kelejian and Robinson (1998, 1999).<sup>34</sup>

### 5.2. *ML Based tests*

When spatial regression models are estimated by maximum likelihood, inference on the spatial autoregressive coefficients may be based on a Wald or asymptotic t-test (from the asymptotic variance matrix) or on a likelihood ratio test [see Anselin (1988, Chapter 6), Anselin and Bera (1998)]. Both approaches require that the alternative model (i.e., the spatial model) be estimated. In contrast, a series of test statistics based on the Lagrange Multiplier (LM) or Rao Score (RS) principle only require estimation of the model under the null. The LM/RS tests also allow for the distinction between a spatial error and a spatial lag alternative.<sup>35</sup>

A LM/RS test against a spatial error alternative was originally suggested by Burridge (1980) and takes the form

$$LM_{\text{err}} = [\mathbf{e}'\mathbf{W}\mathbf{e}/(\mathbf{e}'\mathbf{e}/N)]^2 / [\text{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W})] \quad (5.2)$$

This statistic has an asymptotic  $\chi^2(1)$  distribution and, apart from a scaling factor, corresponds to the square of Moran's  $I$ .<sup>36</sup> From several simulation experiments [Anselin and Rey (1991), Anselin and Florax (1995b)] it follows that Moran's  $I$  has slightly better power than the  $LM_{\text{err}}$  test in small samples, but the performance of both tests becomes indistinguishable in medium and large size samples. The LM/RS test against a spatial lag alternative was outlined in Anselin (1988c) and takes the form

$$LM_{\text{lag}} = [\mathbf{e}'\mathbf{W}\mathbf{y}/(\mathbf{e}'\mathbf{e}/N)]^2 / D \quad (5.3)$$

where  $D = [(\mathbf{W}\mathbf{X}\boldsymbol{\beta})'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')(\mathbf{W}\mathbf{X}\boldsymbol{\beta})/\sigma^2] + \text{tr}(\mathbf{W}^2 + \mathbf{W}'\mathbf{W})$ . This statistic also has an asymptotic  $\chi^2(1)$  distribution.

34 See also Kelejian and Robinson (1992) for an alternative test statistic.

35 Moran's  $I$  is not based on an explicit alternative and has power against both [see Anselin and Rey (1991)].

36 As shown in Anselin and Kelejian (1997) these tests are asymptotically equivalent.

Since both tests have power against the other alternative, it is important to take account of possible lag dependence when testing for error dependence and vice versa. This can be implemented by means of a joint test [Anselin (1988c)] or by constructing tests that are robust to the presence of local misspecification of the other form [Bera and Yoon (1993), Anselin et al. (1996)].

The LM/RS principle can also be extended to more complex spatial alternatives, such as higher order processes, spatial error components and direct representation models [Anselin (1999)], to panel data settings [Anselin (1988b)], and to probit models [Pinkse (1998, 1999), Pinkse and Slade (1998)]. A common characteristics of the LM/RS tests against spatial alternatives is that they do not lend themselves readily to a formulation as a  $NR^2$  expression based on an auxiliary regression. However, as recently shown in Baltagi and Li (1998), it is possible to obtain tests for spatial lag and spatial error dependence in a linear regression model by means of Davidson and MacKinnon's (1984, 1988) double length artificial regression approach.

## 6. Implementation Issues

To date, spatial econometric methods are not found in the main commercial econometric and statistical software packages, although macro and scripting facilities may be used to implement some estimators [Anselin and Hudak (1992)]. The only comprehensive software to handle both estimation and specification testing of spatial regression models is the special-purpose SpaceStat package [Anselin (1992b, 1998b)]. Maximum likelihood estimation of spatial models is also included in the Matlab routines of Pace and Barry (1998) and estimation of spatial error models is part of the S+Spatialstats add-on to S-Plus [MathSoft (1996)].<sup>37</sup>

In contrast to maximum likelihood estimation, method of moments and 2SLS can easily be implemented with standard software, provided that spatial lags can be computed. This requires the construction of a spatial weights matrix which must often be derived from information in a geographic information system [e.g., Can (1996)]. Similarly, once a spatial lag can be computed, the LM/RS statistics are straightforward to implement.

The main practical problem is encountered in maximum likelihood estimation where the Jacobian determinant must be evaluated for every iteration in a nonlinear optimization procedure. The original solution to this problem was suggested by Ord (1975), who showed how the log Jacobian can be decomposed in terms that contain the eigenvalues of the weights matrix  $\omega_j$ ,

---

37 Neither of these toolboxes include specification tests and S+Spatialstats has no routines to handle the spatial lag model.

$$\ln|\mathbf{I} - \rho\mathbf{W}| = \sum_{i=1}^n \ln(1 - \rho\omega_i). \quad (6.1)$$

This is easy to implement in a standard optimization routine by treating the individual elements in the sum as observations on an auxiliary term in the log-likelihood [see Anselin and Hudak (1992)]. However, the computation of the eigenvalues quickly becomes numerically unstable for matrices of more than 1,000 observations. In addition, for large data sets this approach is inefficient in that it does not exploit the high degree of sparsity of the spatial weights matrix. Recently suggested solutions to this problem fall into two categories. Approximate solutions avoid the computation of the Jacobian determinant, but instead approximate it by a polynomial function or by means of simulation methods [e.g., Griffith (1992), Griffith and Sone (1995), Martin (1993), Barry and Pace (1999)]. Exact solutions are based on Cholesky or LU decomposition methods that exploit the sparsity of the weights [Pace (1997), Pace and Barry (1997a, 1997b)], or use a characteristic polynomial approach [Smirnov and Anselin (1999)]. While much progress has been made, considerable work remains to be done to develop efficient algorithms and data structures to allow for the analysis of very large spatial data sets.

## 7. Concluding Remarks

This review chapter has been an attempt to present the salient issues pertaining to the methodology of spatial econometrics. It is by no means complete, but it is hoped that sufficient guidance is provided to pursue interesting research directions. Many challenging problems remain, both methodological in nature as well as in terms of applying the new techniques to meaningful empirical problems. Particularly in dealing with spatial effects in models other than the standard linear regression, much needs to be done to complete the spatial econometric toolbox. It is hoped that the review presented here will stimulate statisticians and econometricians to tackle these interesting and challenging problems.

## References

- Akerlof, George A. (1997). Social distance and social decisions. *Econometrica* 65, 1005–1027.
- Alessie, Rob and Arie Kapteyn (1991). Habit formation, interdependent preferences and demographic effects in the almost ideal demand system. *Economic Journal* 101, 404–419.
- Anselin, L. (1980). *Estimation methods for spatial autoregressive structures*. Regional Science Dissertation and Monograph Series 8. Field of Regional Science, Cornell University, Ithaca, N.Y.
- Anselin, L. (1988a). *Spatial Econometrics: Methods and Models*. Kluwer Academic, Dordrecht.
- Anselin, L. (1988b). A test for spatial autocorrelation in seemingly unrelated regressions. *Economics Letters* 28, 335–341.
- Anselin, L. (1988c). Lagrange multiplier test diagnostics for spatial dependence and spatial hetero-

- ogeneity. *Geographical Analysis* 20, 1–17.
- Anselin, L. (1989). What is special about spatial data? Alternative perspectives on spatial data analysis. Technical Report 89–4, National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA [contained on CD-Rom “Fundamental research in geographic information and analysis”, Santa Barbara, Ca: NCGIA]
- Anselin, L. (1990a). Spatial dependence and spatial structural instability in applied regression analysis. *Journal of Regional Science* 30, 185–207.
- Anselin, L. (1990b). Some robust approaches to testing and estimation in spatial econometrics. *Regional Science and Urban Economics* 20, 141–163.
- Anselin, L. (1992a). Space and applied econometrics. Special Issue, *Regional Science and Urban Economics* 22.
- Anselin, L. (1992b). *SpaceStat, a Software Program for the Analysis of Spatial Data*. National Center for Geographic Information and Analysis, University of California, Santa Barbara, CA.
- Anselin, L. (1998a). GIS research infrastructure for spatial analysis of real estate markets. *Journal of Housing Research* 9, 113–33.
- Anselin, L. (1998b). *SpaceStat Version 1.90*. <http://www.spacestat.com>.
- Anselin, L. (1999). Rao’s score test in spatial econometrics. *Journal of Statistical Planning and Inference* (forthcoming).
- Anselin, L. and A. Bera (1998). Spatial dependence in linear regression models with an introduction to spatial econometrics. In: A. Ullah and D. E. A. Giles, Eds., *Handbook of Applied Economic Statistics*, pp. 237–289. New York: Marcel Dekker.
- Anselin, L. and R. Florax (1995a). Introduction. In L. Anselin and R. Florax (Eds.), *New Directions in Spatial Econometrics*, pp. 3–18. Berlin: Springer-Verlag.
- Anselin, L. and R. Florax (1995b). Small sample properties of tests for spatial dependence in regression models: some further results. In L. Anselin and R. Florax (Eds.), *New Directions in Spatial Econometrics*, pp.21–74. Berlin: Springer-Verlag.
- Anselin, L. and D. Griffith (1988). Do spatial effects really matter in regression analysis. *Papers, Regional Science Association* 65, 11–34.
- Anselin, L. and S. Hudak (1992). Spatial econometrics in practice, a review of software options. *Regional Science and Urban Economics* 22, 509–536.
- Anselin, L. and H. H. Kelejian (1997). Testing for spatial error autocorrelation in the presence of endogenous regressors. *International Regional Science Review* 20, 153–182.
- Anselin, L. and S. Rey (1991). Properties of tests for spatial dependence in linear regression models. *Geographical Analysis* 23, 112–131.
- Anselin, L. and S. Rey (1997). Introduction to the special issue on spatial econometrics. *International Regional Science Review* 20, 1–7.
- Anselin, L. and O. Smirnov (1996). Efficient algorithms for constructing proper higher order spatial lag operators. *Journal of Regional Science* 36, 67–89.
- Anselin, L., A. Bera, R. Florax and M. Yoon (1996). Simple diagnostic tests for spatial dependence. *Regional Science and Urban Economics* 26, 77–104.
- Aoki, Masanao (1994). New macroeconomic modelling approaches. Hierarchical dynamics and mean-field approximation. *Journal of Economic Dynamics and Control* 18, 865–77.
- Aoki, Masanao (1996). *New Approaches to Macroeconomic Modeling*. Cambridge: Cambridge University Press.
- Arthur, W. Brian (1989). Increasing returns, competing technologies and lock-in by historical small events: the dynamics of allocation under increasing returns to scale. *Economic Journal* 99, 116–31.
- Aten, Bettina (1996). Evidence of spatial autocorrelation in international prices. *Review of Income and Wealth* 42, 149–63.
- Aten, Bettina (1997). Does space matter? International comparisons of the prices of tradables and nontradables. *International Regional Science Review* 20, 35–52.

- Baltagi, Badi and Dong Li (1998). Double length artificial regressions for testing spatial dependence. Working Paper, Department of Economics, Texas A&M University, College Station, TX.
- Baltagi, Badi and Dong Li (1999). Prediction in the panel data model with spatial correlation. In L. Anselin and R. Florax (Eds.), *Advances in Spatial Econometrics*. Heidelberg: Springer-Verlag (forthcoming).
- Barry, R.P. and R. K. Pace (1999). Monte Carlo estimates of the log determinant of large sparse matrices. *Linear Algebra and its Applications* 289, 41–54.
- Bartels, C.P.A. and L. Hordijk (1977). On the power of the generalized Moran contiguity coefficient in testing for spatial autocorrelation among regression disturbances. *Regional Science and Urban Economics* 7, 83–101.
- Basu, Sabyasachi and Thomas G. Thibodeau (1998). Analysis of spatial autocorrelation in housing prices. *The Journal of Real Estate Finance and Economics* 17, 61–85.
- Bell, Kathleen P. and Nancy E. Bockstael (1999). Applying the generalized moments estimation approach to spatial problems involving micro-level data. *The Review of Economics and Statistics* 81 (in press)
- Benirschka, M. and J.K. Binkley (1994). Land price volatility in a geographically dispersed market. *American Journal of Agricultural Economics* 76, 185–95.
- Bera, A. and M. J. Yoon (1993). Specification testing with misspecified alternatives. *Econometric Theory* 9, 649–658.
- Beron, Kurt J. and Wim P.M. Vijverberg (1999). Probit in a spatial context: a Monte Carlo approach. In L. Anselin and R. Florax (Eds.), *Advances in Spatial Econometrics*. Heidelberg: Springer-Verlag (forthcoming).
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society B* 36, 192–225.
- Besag, J. (1977). Efficiency of pseudo-likelihood estimators for simple Gaussian fields. *Biometrika* 64, 616–618.
- Besag, J. and P.A.P. Moran (1975). On the estimation and testing of spatial interaction in Gaussian lattice processes. *Biometrika* 62, 555–562.
- Besag, J., J. York and A. Mollie (1991). Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics* 43, 1–59.
- Besag, J., P. Green, D. Higdon and K. Mengersen (1995). Bayesian computation and stochastic systems. *Statistical Science* 10, 3–66.
- Besley, T. and A. Case (1995). Incumbent behavior: vote-seeking, tax-setting and yardstick competition. *American Economic Review* 85, 25–45.
- Bivand, R. and S. Szymanski (1997). Spatial dependence through local yardstick competition: theory and testing. *Economics Letters* 55, 257–265.
- Blommestein, H. (1983). Specification and estimation of spatial econometric models: a discussion of alternative strategies for spatial economic modelling. *Regional Science and Urban Economics* 13, 250–271.
- Blommestein, H. (1985). Elimination of circular routes in spatial dynamic regression equations. *Regional Science and Urban Economics* 15, 121–130.
- Bockstael, Nancy E. (1996). Modeling economics and ecology: the importance of a spatial perspective. *American Journal of Agricultural Economics* 78, 1168–1180.
- Bolduc, Denis, Bernard Fortin and Stephen Gordon (1997). Multinomial probit estimation of spatially interdependent choices: an empirical comparison of two new techniques. *International Regional Science Review* 20, 77–101.
- Bolthausen, E. (1982). On the central limit theorem for stationary mixing random fields. *The Annals of Probability* 10, 1047–1050.
- Borjas, George J. (1995). Ethnicity, neighborhoods, and human-capital externalities. *American Economic Review* 85, 365–90.

- Brandsma, A.S. and R.H. Ketellapper (1979). A biparametric approach to spatial autocorrelation. *Environment and Planning A* 11, 51–58.
- Brett, Craig and Joris Pinkse (1997). Those taxes all over the map! A test for spatial independence of municipal tax rates in British Columbia. *International Regional Science Review* 20, 131–51.
- Breusch, T. (1980). Useful invariance results for generalized regression models. *Journal of Econometrics* 13, 327–340.
- Brock, William A. and Steven N. Durlauf (1995). Discrete choice with social interactions I: Theory. NBER Working Paper No. W5291. Cambridge, MA: National Bureau of Economic Research.
- Brueckner, Jan K. (1998). Testing for strategic interaction among local governments: the case of growth controls. *Journal of Urban Economics* 44, 438–67.
- Burrige, P. (1980). On the Cliff-Ord test for spatial autocorrelation. *Journal of the Royal Statistical Society B* 42, 107–108.
- Can, A. (1996). Weight matrices and spatial autocorrelation statistics using a topological vector data model. *International Journal of Geographical Information Systems* 10, 1009–1017.
- Can, A. (1998). GIS and spatial analysis of housing and mortgage markets. *Journal of Housing Research* 9, 61–86.
- Case, A. (1991). Spatial patterns in household demand. *Econometrica* 59, 953–965.
- Case, A. (1992). Neighborhood influence and technological change. *Regional Science and Urban Economics* 22, 491–508.
- Case, A., H.S. Rosen and J.R. Hines (1993). Budget spillovers and fiscal policy interdependence: evidence from the States. *Journal of Public Economics* 52, 285–307.
- Clayton, D.G. and J. Kaldor (1987). Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 43, 671–681.
- Cliff, A. and J. K. Ord (1972). Testing for spatial autocorrelation among regression residuals. *Geographical Analysis* 4, 267–284.
- Cliff, A. and J. K. Ord (1973). *Spatial Autocorrelation*. London: Pion.
- Cliff, A. and J. K. Ord (1981). *Spatial Processes: Models and Applications*. London: Pion.
- Conley, T.G. (1996). *Econometric modelling of cross-sectional dependence*. Ph.D. Dissertation. Department of Economics, University of Chicago, Chicago, IL.
- Cressie, N. (1993). *Statistics for Spatial Data*. Wiley, New York.
- Cressie, N. and T. Read (1989). Spatial data analysis of regional counts. *Biometrical Journal* 31, 699–719.
- Davidson, J. (1994). *Stochastic Limit Theory*. Oxford: Oxford University Press.
- Davidson, R. and J.G. MacKinnon (1984). Model specification tests based on artificial regressions. *International Economic Review* 25, 485–502.
- Davidson, R. and J.G. MacKinnon (1988). Double-length artificial regressions. *Oxford Bulletin of Economics and Statistics* 50, 203–217.
- Diggle, P., J.A. Tawn and R.A. Moyeed (1998). Model-based geostatistics. *Applied Statistics* 47, 299–350.
- Doreian, P. (1980). Linear models with spatially distributed data, spatial disturbances or spatial effects. *Sociological Methods and Research* 9, 29–60.
- Driscoll, John C. and Aart C. Kraay (1998). Consistent covariance matrix estimation with spatially dependent panel data. *The Review of Economics and Statistics* 80, 549–60.
- Dubin, R. (1988). Estimation of regression coefficients in the presence of spatially autocorrelated error terms. *Review of Economics and Statistics* 70, 466–474.
- Dubin, R. (1992). Spatial autocorrelation and neighborhood quality. *Regional Science and Urban Economics* 22, 433–452.
- Durlauf, Steven N. (1994). Spillovers, stratification and inequality. *European Economic Review* 38, 836–845.
- Durlauf, Steven N. (1997). Statistical mechanics approaches to socioeconomic behavior. In W.

- Brian Arthur, Steven N. Durlauf and David A. Lane (Eds.), *The economy as an evolving complex system II*, pp. 81–104. Reading, MA: Addison-Wesley.
- Fiebig, Denzil G. (1999). Seemingly unrelated regression. In Baltagi (Ed.), *Companion in Theoretical Econometrics*. Oxford: Basil Blackwell.
- Frees, Edward W. (1995). Assessing cross-sectional correlation in panel data. *Journal of Econometrics* 69, 393–414.
- Geoghegan, J., L. Wainger and N. Bockstael (1997). Spatial landscape indices in a hedonic framework: an ecological economics analysis using GIS. *Ecological Economics* 23, 251–64.
- Glaeser, E. L., H. D. Kallal, J. A. Scheinkman and A. Schleifer (1992). Growth in cities. *Journal of Political Economy* 100, 1126–52.
- Glaeser, E. L., B. Sacerdote and J. A. Scheinkman (1996). Crime and social interactions. *Quarterly Journal of Economics* 111, 507–48.
- Griffith, D. (1992). Simplifying the normalizing factor in spatial autoregressions for irregular lattices. *Papers in Regional Science* 71, 71–86.
- Griffith, D. and A. Sone. (1995). Trade-offs associated with normalizing constant computational simplifications for estimating spatial statistical models. *Journal of Statistical Computation and Simulation* 51, 165–183.
- Haining, R. (1990). *Spatial Data Analysis in the Social and Environmental Sciences*. Cambridge University Press, Cambridge.
- Hansen, Lars P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.
- Heagerty, Patrick J. and Subhash R. Lele (1998). A composite likelihood approach to binary spatial data. *Journal of the American Statistical Association* 93, 1099–1111.
- Heckman, J. (1991). Identifying the end of past: distinguishing state dependence from heterogeneity. *American Economic Review* 81, 75–79.
- Hepple, L. W. (1995). Bayesian techniques in spatial and network econometrics: 2. Computational methods and algorithms. *Environment and Planning A* 27, 447–469.
- Holtz-Eakin, D. (1994). Public-sector capital and the productivity puzzle. *Review of Economics and Statistics* 76, 12–21.
- Hordijk, L. and J. Paelinck (1976). Some principles and results in spatial econometrics. *Recherches Economiques de Louvain* 42, 175–97.
- Huang, J. S. (1984). The autoregressive moving average model for spatial analysis. *Australian Journal of Statistics* 26, 169–178.
- Imhof, J.P. (1961). Computing the distribution of quadratic forms in normal variables. *Biometrika* 48, 419–426.
- Ioannides, Y. M. (1990). Trading uncertainty and market form. *International Economic Review* 31, 619–38.
- Ioannides, Y. M. (1997). Evolution of trading structures. In W. Brian Arthur, Steven N. Durlauf and David A. Lane (Eds.), *The economy as an evolving complex system II*, pp. 129–167. Reading, MA: Addison-Wesley.
- Isard, Walter (1956). *Location and Space-Economy*. Cambridge: MIT Press.
- Kaiser, Mark S. and Noel Cressie (1997). Modeling Poisson variables with positive spatial dependence. *Statistics and Probability Letters* 35, 423–432.
- Kelejian, H. and I. Prucha (1997). Estimation of spatial regression models with autoregressive errors by two stage least squares procedures: a serious problem. *International Regional Science Review* 20, 103–111.
- Kelejian, H. and I. Prucha (1998). A generalized spatial two stage least squares procedure for estimating a spatial autoregressive model with autoregressive disturbances. *Journal of Real Estate Finance and Economics* 17, 99–121.
- Kelejian, H. and I. Prucha (1999a). A generalized moments estimator for the autoregressive

- parameter in a spatial model. *International Economic Review* (forthcoming).
- Kelejian, H.H. and I. Prucha (1999b). On the asymptotic distribution of the Moran I test statistic with applications. Working Paper, Department of Economics, University of Maryland, College Park, MD.
- Kelejian, H.H. and D.P. Robinson (1992). Spatial autocorrelation: a new computationally simple test with an application to per capita country police expenditures. *Regional Science and Urban Economics* 22, 317–333.
- Kelejian, H.H. and D.P. Robinson (1993). A suggested method of estimation for spatial interdependent models with autocorrelated errors, and an application to a county expenditure model. *Papers in Regional Science* 72, 297–312.
- Kelejian, H.H. and D.P. Robinson (1995). Spatial correlation: a suggested alternative to the autoregressive model. In L. Anselin and R. Florax (Eds.), *New Directions in Spatial Econometrics*, pp. 75–95. Berlin: Springer-Verlag.
- Kelejian, H.H. and D.P. Robinson (1997). Infrastructure productivity estimation and its underlying econometric specifications: a sensitivity analysis. *Papers in Regional Science* 76, 115–131.
- Kelejian, H.H. and D.P. Robinson (1998). A suggested test for spatial autocorrelation and/or heteroskedasticity and corresponding Monte Carlo results. *Regional Science and Urban Economics* 28, 389–417.
- Kelejian, H.H. and D.P. Robinson (1999). The influence of spatially correlated heteroskedasticity on tests for spatial correlation. In L. Anselin and R. Florax (Eds.), *Advances in Spatial Econometrics*. Heidelberg: Springer-Verlag (forthcoming).
- King, M. (1981). A small sample property of the Cliff-Ord test for spatial correlation. *Journal of the Royal Statistical Society B* 43, 263–264.
- Koerts, J. and A.P.I. Abrahamse (1968). On the power of the BLUS procedure. *Journal of the American Statistical Association* 63, 1227–1236.
- Krugman, Paul (1991a). Increasing returns and economic geography. *Journal of Political Economy* 99, 438–99.
- Krugman, Paul (1991b). *Geography and trade*. Cambridge, MA: MIT Press.
- Krugman, Paul (1998). Space: the final frontier. *Journal of Economic Perspectives* 12, 161–174.
- Lahiri, S.N. (1996). On the inconsistency of estimators under infill asymptotics for spatial data. *Sankhya A* 58, 403–417.
- Land, K. and G. Deane (1992). On the large-sample estimation of regression models with spatial- or network-effects terms: a two stage least squares approach. In P. Marsden (Ed.), *Sociological Methodology*, pp. 221–248. San Francisco: Jossey-Bass.
- LeSage, James P. (1997). Bayesian estimation of spatial autoregressive models. *International Regional Science Review* 20, 113–129.
- Magnus, J. (1978). Maximum likelihood estimation of the GLS model with unknown parameters in the disturbance covariance matrix. *Journal of Econometrics* 7, 281–312. (Corrigenda, *Journal of Econometrics* 10, 261).
- Manski, C. F. (1993). Identification of endogenous social effects: the reflection problem. *Review of Economic Studies* 60, 531–542.
- Mardia, K.V. and R.J. Marshall (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika* 71, 135–146.
- Mardia, K.V. and A.J. Watkins (1989). On multimodality of the likelihood for the spatial linear model. *Biometrika* 76, 289–295.
- Martin, R. J. (1993). Approximations to the determinant term in Gaussian maximum likelihood estimation of some spatial models. *Communications in Statistics: Theory and Methods* 22, 189–205.
- MathSoft (1996). *S+SpatialStats User's Manual for Windows and Unix*. Seattle, WA: MathSoft, Inc.
- McMillen, D. P. (1992). Probit with spatial autocorrelation. *Journal of Regional Science* 32, 335–48.
- McMillen, D. P. (1995). Spatial effects in probit models: a Monte Carlo investigation. In L. Anselin and

- R. Florax (Eds.), *New Directions in Spatial Econometrics*, pp. 189–228. Heidelberg: Springer-Verlag.
- Moran, P. A. P. (1948). The interpretation of statistical maps. *Biometrika* 35, 255–260.
- Moran, P. A. P. (1950a). Notes on continuous stochastic phenomena. *Biometrika* 37, 17–23.
- Moran, P. A. P. (1950b). A test for the serial dependence of residuals. *Biometrika* 37, 178–181.
- Murdoch, J.C., M. Rahmatian and M. A. Thayer (1993). A spatially autoregressive median voter model of recreation expenditures. *Public Finance Quarterly* 21, 334–350.
- Murdoch, James C., Todd Sandler and Keith Sargent (1997). A tale of two collectives: sulphur versus nitrogen oxides emission reduction in Europe. *Economica* 64, 281–301.
- Nelson, Gerald C. and David A. Gray (1997). *The use of spatial analysis at the World Bank*. Environmental Economics Series Paper No. 56. The World Bank, Washington DC.
- Nelson, Gerald C. and Daniel Hellerstein (1997). Do roads cause deforestation? Using satellite images in econometric analysis of land use. *American Journal of Agricultural Economics* 79, 80–88.
- Newey, Whitney K. and Kenneth D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55, 703–708.
- Olmo, J. C. (1995). Spatial estimation of housing prices and locational rents. *Urban Studies* 32, 1331–1344.
- Ord, J.K. (1975). Estimation methods for models of spatial interaction. *Journal of the American Statistical Association* 70, 120–126.
- Pace, R. K. (1997). Performing large spatial regressions and autoregressions. *Economics Letters* 54, 283–291.
- Pace, R. K. and R. Barry (1997a). Sparse spatial autoregressions. *Statistics and Probability Letters* 33, 291–297.
- Pace, R. K. and R. Barry (1997b). Quick computation of spatial autoregressive estimators. *Geographical Analysis* 29, 232–246.
- Pace, R. K. and R. Barry (1998). *Spatial Statistics Toolbox 1.0*. Real Estate Research Institute, Louisiana State University, Baton Rouge, LA.
- Pace, R. Kelley, Ronald Barry, C.F. Sirmans (1998). Spatial statistics and real estate. *Journal of Real Estate Finance and Economics* 17, 5–13.
- Paelinck, J. and L. Klaassen (1979). *Spatial Econometrics*. Saxon House, Farnborough.
- Pinkse, Joris (1998). Asymptotic properties of the Moran and related tests and a test for spatial correlation in probit models. Working Paper, Department of Economics, University of British Columbia, Vancouver, BC.
- Pinkse, Joris (1999). Moran-flavored tests with nuisance parameters. In L. Anselin and R. Florax (Eds.), *Advances in Spatial Econometrics*. Heidelberg: Springer-Verlag (forthcoming).
- Pinkse, Joris and Margaret E. Slade (1998). Contracting in space: an application of spatial statistics to discrete-choice models. *Journal of Econometrics* 85, 125–54.
- Poirier, Dale J. and Paul A. Ruud (1988). Probit with dependent observations. *Review of Economic Studies* 55, 593–614.
- Pötscher, B.M. and I. R. Prucha (1997). *Dynamic Nonlinear Econometric Models*. Springer, Berlin.
- Rao, C. R. (1973). *Linear Statistical Inference and its Applications (2nd Ed.)*. Wiley, New York.
- Ripley, B. D. (1981). *Spatial Statistics*. Wiley, New York.
- Ripley, B. D. (1988). *Statistical Inference for Spatial Processes*. Cambridge: Cambridge University Press.
- Schmidt, P. (1976). *Econometrics*. New York: Marcel Dekker.
- Smirnov, O. and L. Anselin (1999). Fast maximum likelihood estimation of very large spatial autoregressive models: a characteristic polynomial approach. Working Paper, Bruton Center, University of Texas at Dallas, Richardson, TX.
- Sneek, J.M. and P. Rietveld (1999). On the estimation of the spatial moving average model. In L. Anselin and R. Florax (Eds.), *Advances in Spatial Econometrics*. Heidelberg: Springer-Verlag (forthcoming).
- Tiefelsdorf, M. and B. Boots (1995). The exact distribution of Moran's *I*. *Environment and Planning*

A 27, 985–999.

- Topa, Giorgio (1996). Social interactions, local spillovers and unemployment. Ph.D. Dissertation, Department of Economics, University of Chicago.
- Upton, G. J. and B. Fingleton (1985). *Spatial Data Analysis by Example. Volume 1: Point Pattern and Quantitative Data*. Wiley, New York.
- Vijverberg, Wim (1997). Monte Carlo evaluation of multivariate normal probabilities. *Journal of Econometrics* 76, 281–307.
- Waller, Lance, Bradley Carlin and Hong Xia (1997). Structuring correlation within hierarchical spatio-temporal models for disease rates. In T. Gregoire, D. Brillinger, P. Diggle, E. Russek-Cohen, W. Warren and R. Wolfinger (Eds.), *Modelling Longitudinal and Spatially Correlated Data*, pp. 309–319. New York: Springer-Verlag.
- Waller, Lance, Bradley Carlin, Hong Xia and Alan Gelfand (1997). Hierarchical spatio-temporal mapping of disease rates. *Journal of the American Statistical Association* 92, 607–617.
- Warnes, J.J. and B.D. Ripley (1987). Problems with likelihood estimation of covariance functions of spatial Gaussian processes. *Biometrika* 74, 640–642.
- Whittle, P. (1954). On stationary processes in the plane. *Biometrika* 41, 434–449.