Cross-Section Spatial Regression Models

76

Julie Le Gallo

Contents

Abstract

This chapter provides a selective survey of specification issues in spatial econometrics. We first present the most commonly used spatial specifications in a cross-sectional setting in the form of linear regression models including a spatial lag and/or a spatial error term, heteroscedasticity or parameter instability. Second, we present a set of specification tests that allow checking

J. Le Gallo
CRESE, Université de Franche-Comté, Besançon, France
e-mail: jlegallo@univ-fcomte.fr

deviations from a standard, that is, nonspatial, regression model. An important space is devoted to unidirectional, multidirectional, and robust LM tests as they only require the estimation of the model under the null. Because of the complex links between spatial autocorrelation and spatial heterogeneity, we give some attention to the specifications incorporating both aspects and to the associated specification tests.

76.1   Introduction

In spatial regression models, the observations are collected from points or regions located in space. These models usually incorporate *spatial effects* that are commonly classified in two categories: *spatial autocorrelation* and *spatial heterogeneity*. On the one hand, spatial autocorrelation is a special case of cross-sectional dependence and refers to the coincidence of value similarity with locational similarity (Anselin and Bera 1998). Positive spatial autocorrelation means that observations from one location tend to exhibit values similar to those from nearby locations, while negative spatial autocorrelation points to the spatial clustering of dissimilar values. The typical characteristic of spatial autocorrelation is that it is two dimensional and multidirectional. On the other hand, spatial heterogeneity pertains to structural relations that vary over space, either in the form of nonconstant error variances in a regression model (heteroscedasticity) or in the form of spatially varying regression coefficients.

   In recent years, the interest in *spatial econometrics*, that is, the subset of econometric methods that deals with the analysis of spatial effects in regression analysis, has seen an exponential growth in social sciences, leading to the creation of the *Spatial Econometrics Association* in 2006 (Arbia 2011). The upsurge in spatial econometrics has been driven by the recognition of the role of space and spatial/social interactions in economic theory, the availability of datasets with georeferenced observations, and the development of geographic information systems and spatial data analysis softwares. This field has even reached a stage of maturity through general acceptance as a mainstream methodology, according to Anselin (2010).

   In this chapter, we provide a concise overview of the methodological issues related to the treatment of spatial effects in regression models. Attention here is given to specification issues, that is, how spatial correlation and spatial heterogeneity structures should be incorporated into a regression model and the implications for specification testing. We do not consider estimation issues, as this is the topic of other chapters in this volume (see Prucha and Jenish, ▶ Chap. 80, "Instrumental Variables/Method of Moments Estimation"; Mills and Parent, ▶ Chap. 79, "Bayesian MCMC Estimation" and Pace, ▶ Chap. 78, "Maximum Likelihood Estimation"). We have also limited the review to cross-sectional settings for linear regression models and do not consider spatial effects in space-time models (see Elhorst, ▶ Chap. 82, "Spatial Panel Models") nor models for limited dependent variables (see Wang, ▶ Chap. 81, "Limited and Censored Dependent Variable Models").

The chapter consists in two sections, starting with a presentation of the specification of spatial effects in cross-sectional linear regression models. Next, we consider specification tests that detect spatial autocorrelation and/or spatial heterogeneity. Most attention is devoted to spatial autocorrelation, the distinct nature of which requires a specialized set of techniques that are not a straightforward extension of time series methods to two dimensions. On the contrary, the treatment of spatial heterogeneity does not require specific econometric tools. However, we underline here the relationships between both effects. The chapter closes with some concluding remarks.

76.2   Spatial Effects in Cross-Sectional Models

Consider as a point of departure, the classical cross-sectional linear regression model:

$$y = X\beta + \varepsilon \qquad (76.1)$$

where N is the total number of observations, here geographical areas; K is the total number of unknown parameters to estimate; y is the $(N,1)$ vector of observations on the dependent variable; X is the (N,K) matrix of observations on the K explanatory variables; β is the $(N,1)$ vector of unknown parameters to be estimated; and ε is the $(N,1)$ vector of error terms. We also assume that X is a non-stochastic matrix of full rank $K < N$.

If the error terms are $iid(0, \sigma^2 I_N)$, where I_N is the identity matrix of order N, then the Ordinary Least Squares (OLS) estimator defined by $\tilde{\beta} = (X'X)^1 X'y$ is BLUE (Gauss-Markov theorem). However, the introduction of spatial effects in the linear regression model implies that some of these assumptions are not met. We first list the models incorporating some form of spatial autocorrelation and continue with models with spatial heterogeneity.

76.2.1 Forms of Spatial Autocorrelation in Regression Models

In the presence of spatial autocorrelation, the variance-covariance matrix in Eq. (76.1) $\Sigma = E(\varepsilon\varepsilon')$ contains N variances and $N(N-1)/2$ off-diagonal parameters following a spatial ordering. These cannot be estimated separately with a cross section of N observations. Hence, in order to incorporate spatial autocorrelation in regression models, several possibilities exist. Some aim at imposing some structure or constraints on the elements of Σ such that a finite number of parameters characterizing spatial autocorrelation can be estimated. Others remain nonparametric. We briefly review these options here.

First, a *stochastic process* may be specified that determines the form of the covariance structure. In doing this, spatial lags are incorporated in the regression model. Spatial lags are obtained as the product of a *spatial weights matrix W* with

the vector of observations on a random variable. This matrix is of dimension (N,N) and specifies the connectivity structure within the observations in the sample. It has nonzero elements w_{ij} in each row i for those columns j that are neighbors of location j. The elements on the diagonal are equal to 0. The notion of *neighbors* can be purely geographic, such as sharing a common border, or can be more general, such as neighbors in social network space. Spatial autocorrelation is then modeled by specifying various functional relationships between the vector of observations of the explained variable y and its spatial lag Wy, a spatially lagged error term $W\varepsilon$ and/or spatially lagged explanatory variables WX.

Second, the covariance between observations can be specified as a direct and continuous function of distance. Different specifications have been suggested.

Third, a nonparametric approach can be adopted where the functional form of the function of distance separating two equations is left unspecified. This can also accommodate heteroscedasticity of unknown form.

We detail these different possibilities below.

76.2.2 Spatial Lag Model

In this model, labeled SAR model, spatial autocorrelation is incorporated through a *spatial lag* of the *endogenous* variable. The *structural* model is written as

$$
\begin{aligned}
y &= \rho Wy + X\beta + \varepsilon \\
\varepsilon &\to iid\left(0, \sigma^2 I_N\right)
\end{aligned}
\tag{76.2}
$$

Wy is the endogenous lag variable for the spatial weights matrix W; ρ is the spatial autoregressive parameter that indicates the strength of interactions existing between the observations of y.

In the spatial lag model, observation y_i is, in part, explained by the values taken by y in neighboring observations: $(Wy)_i = \sum_{j \neq i} w_{ij} y_j$. Indeed, when W is standardized, each element $(Wy)_i$ is interpreted as a weighted average of the y values for i's neighbors. The introduction of Wy allows evaluating the degree of spatial dependence when the impact of other variables is controlled for. When Eq. (76.2) is the result of a theoretical modeling implying some process of social and spatial interaction, this parameter measures substantive spatial dependence, that is, the extent of spatial externalities or spatial diffusion.

Symmetrically, it allows controlling spatial dependence when evaluating the impact of other explanatory variables. In this case, particular care should be given to the interpretation of the coefficient estimates (see below).

LeSage and Pace (2009) provide several motivations for regression models that include a spatial lag. One is a time-dependence motivation: cross-sectional model relations with a spatial lag may come from economic agents considering past period behavior of neighboring agents. The presence of a spatial lag has also been justified with theoretical models involving diffusion, copycatting,

or spatial externalities. These are the cases of substantive spatial dependence. It is then the formal representation of the equilibrium outcome of spatial interaction processes.

Note that ρ is not a conventional correlation coefficient between vector y and its spatial lag Wy. Indeed, this parameter is not restricted to the range -1 to 1. From the DGP associated with the SAR model, the log-likelihood function involves a Jacobian term of the form $\ln|I_N - \rho W|$ that constrains the parameter ρ to be in the interval $[1/w_{\min}; 1/w_{\max}]$ where $w_{\min}$ and $w_{\max}$ are respectively the minimum and the maximum eigenvalues of W. If the latter is row standardized, then $w_{\max} = 1$.

When a spatial lag variable is ignored in the model specification, whereas it is present in the underlying data generating process, the OLS estimators in the spatial model Eq. (76.1) are biased and not consistent (omitted variable bias).

This specification has several properties:

76.2.2.1 Multiplier and Diffusion Effects

Assume that the matrix $(I_N - \rho W)$ is not singular. In this case, Eq. (76.2) can be rewritten in the following *reduced* form:

$$y = (I_N - \rho W)^{-1} X\beta + (I_N - \rho W)^{-1}\varepsilon \tag{76.3}$$

This model is nonlinear in ρ and β. It follows from Eq. (76.3) that $E(y) = (I_N - \rho W)^{-1} X\beta$. The matrix inverse $(I_N - \rho W)^{-1}$ is a full matrix and not triangular, as in the time series case where dependence is only one directional. When $|\rho| < 1$, this implies an infinite series, the *Leontief expansion*, involving the explanatory variables and the error term at all locations:

$$y = (I_N + \rho W + \rho^2 W^2 + \ldots) X\beta + (I_N + \rho W + \rho^2 W^2 + \ldots)\varepsilon \tag{76.4}$$

This expression allows defining two effects: a *multiplier* effect affecting the explanatory variables and a *spatial diffusion* effect affecting the error terms.

On the one hand, with respect to the explanatory variables, this expression means that in average, the value of y at one location i is not only explained by the values of the explanatory variables associated to this location but also by those associated to all the other locations (neighbors or not) *via* the inverse spatial transformation $(I_N - \rho W)^{-1}$. This spatial multiplier effect decreases with distance, that is, the powers of W in the series expansion of $(I_N - \rho W)^{-1}$.

On the other hand, with respect to the error process, this expression means that a random shock in a location i not only affects the value of y in this location but also has an impact on the values of y in all the other locations *via* the same spatial inverse transformation. This is the diffusion effect, which also declines with distance.

Both these effects are *global* in the sense that all locations in the system interact with each other (Anselin 2003).

From Eq. (76.3), it also follows that $E[(Wy)_i \varepsilon_i] = E\left[\left\{W(I_N - \rho W)^{-1}\varepsilon\right\}_i \varepsilon_j\right] \neq 0$. The spatial lag is therefore always endogenous, irrespective of the properties of ε, so that the estimation of model Eq. (76.2) cannot be based on OLS but should be performed using maximum likelihood (ML), instrumental variables (IV), or Bayesian methods.

76.2.2.2 Interpretation of Coefficient Estimates

A consequence of the multiplier effect in the spatial lag model is that particular care should be taken when interpreting the coefficient estimates (LeSage and Pace, ▶ Chap. 77, "Interpreting Spatial Econometric Models" for more details). Indeed, the impact of a marginal change in one variable X_k on $E(y)$ is not equivalent to the coefficient associated to X_k, noted β_k, as in the standard regression model. On the contrary, it follows from Eq. (76.3) that

$$\frac{\partial E(y_i)}{\partial X_{jk}} = S_k(W)_{ij} \tag{76.5}$$

where X_{jk} is the value of X_k at location j and $S_k(W)_{ij}$ is the ij^{th} element of the matrix $(I_N - \rho W)^{-1}\beta_k$. Hence, the impact of a change in an explanatory variable differs over all observations. Summary measures of these impacts are discussed in LeSage and Pace (2012).

76.2.2.3 Variance-Covariance Matrix

From Eq. (76.3), we derive the variance-covariance matrix of y:

$$E(yy') = (I_N - \rho W)^{-1}E(\varepsilon\varepsilon')(I_N - \rho W')^{-1} \tag{76.6a}$$

$$E(yy') = \sigma^2(I_N - \rho W)^{-1}(I_N - \rho W')^{-1} \tag{76.6b}$$

This variance-covariance matrix is full, which implies that each location is correlated with every other location in the system. However, this correlation decreases with distance.

76.2.2.4 Endogenous Spatial Lag and Heteroscedasticity

Note $u = (I_N - \rho W)^{-1}\varepsilon$. Its variance-covariance is written as

$$E(uu') = \sigma^2(I_N - \rho W)^{-1}(I_N - \rho W')^{-1} \tag{76.7}$$

Equation (76.7) shows that the covariance between each pair of error terms is not null and decreasing with the order of proximity. Moreover, the elements of the diagonal of $E(uu')$ are not constant. This implies error *heteroscedasticity* of u, whether or not ε is heteroscedastic.

76.2.3 Cross-Regressive Model: Lagged Exogenous Variable

Another possibility to incorporate spatial autocorrelation in a regression model is to include one or more *exogenous lagged variables* in Eq. (76.1):

$$y = \rho W y + W Z \delta + \varepsilon$$
$$\varepsilon \rightarrow iid\left(0, \sigma^2 I_N\right)$$

(76.8)

Z is a matrix of dimension (N,L) containing L variables that may or not correspond to the variables included in X; WZ is the matrix of observations for the exogenous lagged variables with weights matrix W, and δ is the $(L,1)$ vector of spatial parameters indicating the intensity of spatial correlation existing between the observations in y and those of Z.

In this model, the observation y_i is explained by the values taken by the variables in X in location i and by the variables in Z in neighboring regions. The interactions in the system hence remain *local*.

Contrary to the spatial lag model and the models with a spatial error autocorrelation (below), the estimation of the cross-regressive model can be based on OLS.

76.2.4 Models with Spatial Error Autocorrelation

Finally, spatial autocorrelation can be incorporated in a regression model by specifying a spatial process in the error terms. It is therefore a special form of a nonspherical error variance-covariance matrix with $E\left[\varepsilon_i \varepsilon_j\right] \neq 0$ for two locations $i \neq j$. As such, these models should be estimated using ML, generalized method of moments (GMM), or Bayesian methods. The different possibilities lead to different error spatial covariances that differ with respect to the range and extent of spatial interaction in the model.

76.2.4.1 Spatial Autoregressive Process
The most commonly used specification is a spatial autoregressive process in the error terms. The structural model can be then written as

$$y = X\beta + \varepsilon$$
$$\varepsilon = \lambda W \varepsilon + u$$

(76.9)

The parameter λ is the spatial autoregressive coefficient that reflects the interdependence between the regression residuals; u is the error term such as $u \rightarrow iid(0, \sigma^2 I_N)$. When spatial error autocorrelation is omitted, the OLS estimators are unbiased, but inefficient estimators and the statistical inference based on OLS are biased.

This specification has several properties:

Spatial Diffusion
First, if the matrix $(I_N - \lambda W)$ is not singular, then model Eq. (76.9) can be rewritten under the following reduced form:

$$y = X\beta + (I_N - \lambda W)^{-1}u \tag{76.10}$$

This expression leads to a global spatial diffusion effect as in model Eq. (76.3) but, as $E(y) = X\beta$, there is no spatial multiplier effect.

Variance-Covariance Matrix
From Eq. (76.10), we have

$$E(yy') = E(\varepsilon\varepsilon') = (I_N - \rho W)^{-1}E(\varepsilon\varepsilon')(I_N - \rho W')^{-1} \tag{76.11a}$$

$$E(yy') = E(\varepsilon\varepsilon') = \sigma^2(I_N - \rho W)^{-1}(I_N - \rho W')^{-1} \tag{76.11b}$$

Hence, we find, for ε and for y, a structure identical to that of the spatial lag model: this process leads to nonzero error covariance between each pair of observations, but these covariances decrease with distance. The spatial structure of the variance-covariance induced by the model with spatial error autocorrelation is therefore *global*, since it links all the locations of the system to all others.

Moreover, the error structure induces nonconstant elements of the diagonal of $E(\varepsilon\varepsilon')$, which implies heteroscedasticity of the errors ε, whether u is heteroscedastic or not.

Constrained Spatial Durbin Model
Model Eq. (76.9) can be rewritten in a form where both an endogenous spatial lag and all exogenous spatial lags appear. Indeed, by multiplying both sides of Eq. (76.10) by $(I_N - \lambda W)$ and moving the autoregressive term to the right, we obtain the constrained *spatial Durbin model*:

$$y = \lambda Wy + X\beta - \lambda WX\beta + u \tag{76.12}$$

This specification shows how the spatial error model is a special case of a spatial lag model, with additional nonlinear constraints on the parameters. This forms the basis of a specification test that will be presented below.

Several alternatives have been suggested in the literature even if their application is less frequent in the literature.

76.2.4.2 Spatial Moving-Average Process
The spatial *moving-average* process is specified as

$$\begin{aligned} y &= X\beta + \varepsilon \\ \varepsilon &= \gamma Wu + u \end{aligned} \tag{76.13}$$

where γ is the moving-average coefficient and u is the error term such as $u \rightarrow iid(0, \sigma^2 I_N)$. Contrary to the previous case, the reduced model does not contain any inverse matrices since Eq. (76.13) already corresponds to the reduced model. The variance-covariance matrix resulting from this process is

$$E(\varepsilon\varepsilon') = \sigma^2(I_N + \gamma W)(I_N + \gamma W') = \sigma^2[I_N + \gamma(W + W') + \gamma^2 WW'] \quad (76.14)$$

In contrast to the variance-covariance matrix associated with the autoregressive process, Eq. (76.14) is not a full matrix. The nonzero covariances only exist for first-order $(W + W')$ and second-order (WW') neighbors. This process therefore implies much less overall interaction than the autoregressive model, and the spatial structure of covariance induced by Eq. (76.14) is only *local* since it does link all the locations of system to each other.

Finally, as in the autoregressive case, the elements of the diagonal of Eq. (76.14) are not constant, implying, as in the previous model, heteroscedasticity in ε, irrespective of the nature of u.

76.2.4.3 Kelejian and Robinson Specification

Kelejian and Robinson (1995) suggest another specification in which the error term is the sum of two independent terms, one being a smoothing term of neighboring errors and the other being specific to the location:

$$\varepsilon = Wu + v \quad (76.15)$$

where u and v are supposed homoscedastic and independent. Then, the variance-covariance matrix of ε is

$$E(\varepsilon\varepsilon') = \sigma_v^2 I_N + \sigma_u^2 WW' = \sigma^2[I_N + \varphi WW'] \quad (76.16)$$

where σ_u^2 and σ_v^2 are the variance, respectively, associated with u and v, $\sigma^2 = \sigma_v^2 > 0$ and $\varphi = \sigma_u^2/\sigma_v^2$. The spatial interaction implied by Eq. (76.16) is more limited than in the moving-average model as it only concerns neighbors of the first and second order contained in the nonzero elements of WW'. Heteroscedasticity is also implied in this specification.

76.2.4.4 Direct Representation and Nonparametric Specifications

In this case, the covariance between each pair of error terms is directly specified as an inverse function of the distance between them: $\text{cov}(\varepsilon_i, \varepsilon_j) = \sigma^2 f(\theta, d_{ij})$ where d_{ij} is the distance between i and j, σ^2 is the error variance, and f is the distance function. This function is a distance decay function that should ensure definite-positive variance-covariance matrix. This imposes constraints on the functional form, the parameter space, the metric, and scale used for the distance measure. For instance, one might use a negative exponential distance decay function:

$$E(\varepsilon\varepsilon') = \sigma^2[I_N + \gamma\Psi] \quad (76.17)$$

where the off-diagonal elements of Ψ are given by $\Psi = e^{-\theta d_{ij}}$ where θ is a nonnegative scaling parameter. The diagonal elements of Ψ are set to zero.

Contrary to the previous specifications, the direct representation does not induce heteroscedasticity.

An alternative to parametric specifications is to leave the functional form unspecified: these are *nonparametric* models. We then have $\text{cov}(\varepsilon_i, \varepsilon_j) = f(d_{ij})$ where d_{ij} is a positive and symmetric distance metric. The regularity conditions on the distance metric have been derived by Conley (1999).

The presence of spatial error autocorrelation is often interpreted as a problem in the model specification, such as functional form problems or spatial autocorrelation resulting from a mismatch between the spatial scale of the phenomenon being studied and the spatial scale at which it is measured.

76.2.5 Spatial Durbin Model

An encompassing specification to the spatial lag model, the spatial cross-regressive model, and the spatial error model is the unconstrained spatial Durbin model. The latter contains a spatially lagged endogenous variable and all the spatially lagged exogenous variables. More specifically, it is written as

$$y = \lambda Wy + X\beta + \lambda WX\delta + u \tag{76.18}$$

The spatial lag model, the spatial cross-regressive model, and the spatial error model are found with the appropriate constraints on the parameters, respectively, $H_0 : \delta = 0$, $H_0 : \rho = \delta = 0$, and $H_0 : \lambda\beta + \delta = 0$.

LeSage and Pace (2009) provide several motivations for a spatial Durbin model. One is an *omitted variable* motivation. Indeed, they show that if the linear regression model Eq. (76.1) is affected by an omitted variables problem and if these omitted variables are spatially correlated and correlated with the included explanatory variables, then unbiased estimates of the coefficients associated with the endogenous variables X can still be obtained by fitting a spatial Durbin model. Other motivations detailed in LeSage and Pace (2009) are based on spatial heterogeneity and model uncertainty.

76.2.6 Higher-Order Spatial Models

In these models, multiple spatially lagged dependent variables and/or multiple spatially lagged error terms are included.

For instance, the spatial autoregressive, moving-average SARMA(p,q) process is as follows:

$$\begin{aligned} y &= X\beta + \rho_1 W_1 y + \rho_2 W_2 y + \ldots + \rho_p W_p y + \varepsilon \\ \varepsilon &= \lambda_1 W_1 u + \lambda_2 W_2 u + \ldots + \lambda_p W_p u + u \end{aligned} \tag{76.19}$$

In general, the weights W_i are associated to the ith order of contiguity. We could similarly consider a process where the errors follow a spatial autoregressive process of order q. However, in this case, identification issues may arise (Anselin 1988).

It may be that these *high-order processes* are the result of a poorly specified spatial weights matrix rather than a realistic data generating process (Anselin and Bera 1998). For instance, if the weights matrix of the model underestimates the real spatial interaction in the data, there will be residual spatial error autocorrelation. This can lead to the estimation of higher-order processes while only a well-specified weights matrix should be necessary. These higher-order models are in fact usually used as alternatives in diagnostic tests. Rejection of the null may then indicate that a different specification of the weights is necessary.

76.2.7 Heteroscedasticity

Until now, all specifications have assumed *iid* innovations. However, as we have seen, the sole presence of spatial autocorrelation induces *heteroscedasticity* in the models. In cross-sectional regression, additional heteroscedasticity is also frequently present. For instance, in the spatial autoregressive error model, we can have

$$
\begin{aligned}
y &= X\beta + \varepsilon \\
\varepsilon &= \lambda W \varepsilon + u \\
u &\rightarrow iii(0, \Omega)
\end{aligned}
\tag{76.20}
$$

In this case, the variance-covariance matrix of ε is

$$
E(\varepsilon\varepsilon') = (I_N - \rho W)^{-1} \Omega (I_N - \rho W')^{-1}
\tag{76.21}
$$

Several specifications have been used for Ω. In a spatial context, a useful one is that of *groupwise heteroscedasticity*. When the data are organized into spatial regimes, one variance is estimated for each regime so that Ω has a block-diagonal structure:

$$
\Omega = \begin{bmatrix} \sigma_1^2 I_{N_1} & 0 & \cdots & 0 \\ 0 & \sigma_2^2 I_{N_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_2^2 I_{N_2} \end{bmatrix}
\tag{76.22}
$$

where L is the number of regimes, $N_l, \ l = 1 \dots L$ is the number of observations in regime l, and $I_{N_l}, \ l = 1 \dots L$ is the identity matrix of dimension N_l.

The variance can also be specified as a function of variables:

$$
\sigma_i^2 = \sigma^2 f(z_i' \alpha)
\tag{76.23}
$$

where σ^2 is a scale parameter, f is some functional form, and z_i is a $(P, 1)$ vector of variables and α_i, $i = 1 \ldots P$ are unknown parameters to estimate. For instance, in a spatial context, Casetti and Can (1999) suggest the DARP (Drift Analysis of Regression Parameters) model: the variance of the error terms is expanded into a monotonic function of the observations' distance from a reference point in an expansion space:

$$\sigma_i^2 = e^{\gamma_0 + \gamma_1 h_i} \tag{76.24}$$

where h_i is the square of the distance between the ith observation and one reference point (such as the Central Business District in a city).

The variance-covariance matrix can also be left unspecified as in the nonparametric approach. For instance, Kelejian and Prucha (2007) suggest a nonparametric heteroscedasticity- and autocorrelation-consistent (HAC) estimator of the variance-covariance matrix in a spatial context, that is, a SHAC procedure. They assume that the $(N,1)$ disturbance vectors ε of model Eq. (76.1) are generated as follows: $\varepsilon = R\xi$ where R is a (N,N) non-stochastic matrix whose elements are not known. This disturbance process allows for general patterns of correlation and heteroscedasticity. The asymptotic distribution of the corresponding OLS or instrumental variable (IV) estimators implies the variance-covariance matrix $\psi = N^{-1}Z'\Sigma Z$, where $\Sigma = (\sigma_{ij})$ denotes the variance-covariance matrix of ε. Kelejian and Prucha (2007) show that the SHAC estimator for its $(r,s)^{\text{th}}$ element is

$$\hat{\psi}_{rs} = N^{-1} \sum_{i=1}^{N} \sum_{j=1}^{N} x_{ir} x_{js} \hat{\varepsilon}_i \hat{\varepsilon}_j K\left(d_{ij}/d_n\right) \tag{76.25}$$

where x_{ir} is the ith element of the rth explanatory variable, $\hat{\varepsilon}_i$ is the ith element of the OLS or IV residual vector, d_{ij} is the distance between unit i and unit j, d_n is the bandwidth, and $K(.)$ is the kernel function with the usual properties.

76.2.8 Parameter Instability

Spatial heterogeneity can also manifest by *parameter instability*, that is, the lack of constancy in some, or all, of the parameters in the regression model. This instability has a spatial dimension: the regression coefficients correspond to a number of distinct spatial regimes. The spatial variability of the coefficients can be discrete, if systematic differences between regimes are observed. In this case, model coefficients are allowed to vary between regimes. It can also be continuous over space.

In the absence of spatial autocorrelation, the case of discrete spatial heterogeneity can be readily treated with standard tools such as dummy variables, ANOVA, or spline functions. Recently, some authors have investigated the possibility of spatial heterogeneity affecting the spatial lag or spatial error coefficients. In this

case, the methodology consists in estimating higher-order models where the spatial matrices pertain to different spatial regimes rather than different order of contiguities.

Heterogeneity can also be continuous. In this case, rather than partitioning the cross-sectional sample into regimes, we assume that parameter heterogeneity is location specific. One possibility is to use *geographically weighted regression*, labeled GWR (Fotheringham et al. 2004), which is a locally linear, nonparametric estimation method. The base model for one location i is

$$y_i = \sum_{k=1}^{K} \beta_{ki} x_{ki} + \varepsilon_i \qquad (76.26)$$

A different set of parameters is estimated for each observation by using the values of the characteristics taken by neighboring observations. With respect to spatial autocorrelation, Pace and LeSage (2004) have pointed out that if spatial autocorrelation only arises due to inadequately modeled spatial heterogeneity, GWR can potentially eliminate the problem. However, this is not necessarily the case when substantive interactions coexist with parameter heterogeneity. Therefore, Pace and LeSage (2004) have generalized GWR to allow simultaneously for spatial parameter heterogeneity and spatial autocorrelation: the *spatial autoregressive local estimation* (SALE):

$$U(i)y = \rho_i U(i)Wy + U(i)X\beta_i + U(i)\varepsilon \qquad (76.27)$$

where $U(i)$ represents a (N,N) diagonal matrix containing distance-based weights for observation i that assigns the weights of one to the m nearest neighbors to observation i and weights of zero to all the other observations. The product $U(i)y$ then represents a ($m,1$) subsample of observations on the explained variables associated with the m observations nearest in location to observation i. The other products are interpreted in a similar fashion. As $m \rightarrow N$, $U(i) \rightarrow I_N$, the local estimates approach the global estimates from the SAR model as the subsample increases.

76.3 Specification Tests in Spatial Cross-Sectional Models

Ignoring spatial effects when it is present have various effects on the estimates' properties. It may lead to biased and inconsistent estimates of the model parameters for an omitted spatial lag or inefficient estimated and biased inference for omitted spatial error autocorrelation and/or omitted heteroscedasticity. Hence, specification testing is therefore relevant in applied work and constitutes the topic of this section.

We first present Moran's *I* test, where the alternative is an unspecified form of spatial autocorrelation. Second, we detail the most commonly used tests of spatial autocorrelation based on *maximum likelihood*: tests of a single alternative,

conditional tests, and robust tests. Indeed, as featured in ▶ Chap. 80, "Instrumental Variables/Method of Moments Estimation" and ▶ Chap. 78, "Maximum Likelihood Estimation", there might be some complexities involved in the estimation of spatial processes, based on nonlinear optimization (maximum likelihood or generalized methods of moments). Consequently, tests based on the *Lagrange multiplier* (LM) principle (or score test) have been extensively used in specification testing. Contrary to *Wald* (W) or *likelihood ratio* (LR) tests, they only necessitate the estimation of the model under the null hypothesis, typically the simple regression model as in Eq. (76.1). We also briefly present tests based on alternative principles. Third, some strategies aimed at finding the best specification have been devised, when the researcher does not have an a priori of the form taken by spatial autocorrelation. Finally, we outline the complex interactions between spatial autocorrelation and spatial heterogeneity and present how spatial heterogeneity can be tested.

76.3.1 Moran's *I* Test

Moran's *I* test is a *diffuse test* as the alternative is not a specified form of spatial autocorrelation. It is the two-dimensional analog of the test of temporal correlation in univariate time series for regression residuals (Moran 1950). In matrix notations, it is formally written as

$$I = \frac{N}{S_0} \left(\frac{e'We}{e'e} \right) \tag{76.28}$$

where $e = y - X\tilde{\beta}$ is the vector of OLS regression residuals, W is the spatial weights matrix, and S_0 is a standardization factor equal to the sum of all elements of W. For a row-standardized weights matrix W, this element simplifies to 1. The first two moments under the null were derived by Cliff and Ord (1972):

$$E(I) = \frac{tr(MW)}{N - K} \tag{76.29}$$

$$V(I) = \frac{tr(MWMW') + tr(MW)^2 + \left\{ tr(MW)^2 \right\}}{(N - K)(N - K + 2)} - [E(I)]^2 \tag{76.30}$$

where M is the usual symmetric and idempotent matrix : $M = I_N - X(X'X)^{-1}X'$. Inference is then based on the standardized value: $Z(I) = [I - E(I)]/V(I)$. For normally distributed residuals, $Z(I)$ asymptotically follows a centered normal distribution. Under the null assumption of spatial independence, Moran's *I* test is locally best invariant and is also asymptotically equivalent to a likelihood ratio of $H_0 : \lambda = 0$ in Eq. (76.9) or of $H_0 : \gamma = 0$ in Eq. (76.13); it therefore shares the asymptotic properties of these statistics. Moreover, Moran's *I* has power against any alternative of spatial correlation, including a spatial lag alternative.

In the remainder of the section, we consider tests with a specific alternative, that is, focused tests, and concentrate on Lagrange multiplier tests that only require the estimation of the model under the null hypothesis. Some of these tests are *unidirectional* when the alternative deals with one specific misspecification; others are *multidirectional* when the alternative comprises various misspecifications.

76.3.2 Tests of a Single Assumption

76.3.2.1 Spatial Error Autocorrelation

First, consider the case where the error terms follow a spatial autoregressive model Eq. (76.9): $\varepsilon = \lambda W \varepsilon + u$. We test $H_0 : \lambda = 0$. The null corresponds to the linear classical model Eq. (76.1). The multiplier Lagrange statistic can be written the following way (Anselin 1988):

$$LM_{ERR} = \frac{[e'We/(e'e/N)]^2}{T} \qquad (76.31)$$

where $T = tr[(W' + W)W]$, tr is the trace operator, and e is the vector of OLS regression residuals. This is equivalent to a scaled Moran coefficient. Since there is only one constraint, under the null, this statistic is asymptotically distributed as a $\chi^2(1)$.

The test statistic is the same if we specify as alternative assumption the moving-average process Eq. (76.13) with the test $H_0 : \gamma = 0$. LM_{ERR} is therefore locally optimal for the two alternatives (autoregressive and moving average). Consequently, when the null is rejected, the test does not provide any indications with respect to the form of the error process.

Pace and LeSage (2008) argue that the test of spatial error autocorrelation can be performed using a Hausman test, since under the null (model 1), there are two consistent estimators differing in efficiency (OLS and ML), and under the alternative (model 2) only one estimator is efficient (ML).

76.3.2.2 Kelejian-Robinson Specification

For the specification of the error suggested by Kelejian and Robinson (1995), a Lagrange multiplier test can also be derived following the same principle. Using notations of model Eq. (76.15), testing the null $H_0 : \varphi = 0$ yields a statistic of the form (Anselin 2001)

$$KR = \left[\frac{e'W'We}{e'e/N} - T_1\right]^2 \Big/ 2\left[T_2 - \frac{T_1^2}{N}\right] \qquad (76.32)$$

where $T_1 = tr(WW^2)$ and $T_2 = tr(WW'WW')$. Under the null, this statistic is asymptotically distributed as a $\chi^2(1)$.

76.3.2.3 Common Factor Test

The *common factor test* allows choosing between a model with spatial error autocorrelation and a spatial Durbin model. The unconstrained spatial Durbin model in Eq. (76.18) and the spatial error model in Eq. (76.9) are equivalent if $H_0 : \lambda\beta + \delta = 0$. This test can be performed with the Lagrange multiplier principle. The corresponding statistic is asymptotically distributed as a $\chi^2(K-1)$.

76.3.2.4 Test of an Endogenous Spatial Lag

In this case, the null hypothesis is $H_0 : \rho = 0$ in Eq. (76.2). The test statistic is (Anselin 1988)

$$LM_{LAG} = \frac{[e'Wy/(e'e/N)]^2}{D} \tag{76.33}$$

with $D = (WX\tilde{\beta})' M (WX\tilde{\beta})/\tilde{\sigma}^2 + tr(W'W + WW)$ where $\tilde{\beta}$ and $\tilde{\sigma}^2$ are the OLS estimates. This statistic is asymptotically distributed as a $\chi^2(1)$.

76.3.3 Tests in Presence of Spatial Autocorrelation or Spatial Lag

In specification testing, it is useful to know if the model contains both a spatial error autocorrelation and an endogenous spatial lag. In this respect, Anselin et al. (1996) note that LM_{ERR} is the test statistic corresponding to $H_0 : \lambda = 0$ when assuming a correct specification for the rest of the model, that is, $\rho = 0$. However, if $\rho \neq 0$, this test is not valid anymore, even asymptotically as it is not distributed as a centered χ^2. Hence, valid statistical inference necessitates taking account of a possible endogenous variable when testing spatial error autocorrelation and vice versa.

Facing this problem, three strategies are possible. First, one can perform a *joint test* of the presence of an endogenous spatial lag and a spatial error autocorrelation. However, if the null is rejected, the exact nature of spatial dependence is not known. Second, another solution consists in estimating a model with an endogenous spatial lag and then tests for residual spatial autocorrelation and vice versa. Third, Anselin et al. (1996) suggest *robust tests* based on OLS residuals in the simple model but that are capable of taking account a spatial error autocorrelation when testing endogenous spatial lag and vice versa.

76.3.3.1 Joint Test

The first approach is the test of the joint null hypothesis $H_0 : \lambda = \rho = 0$ in a model containing both a spatial lag and a spatial error:

$$\begin{aligned} y &= \rho W_1 y + X\beta + \varepsilon \\ \varepsilon &= \lambda W_2 \varepsilon + u \end{aligned} \tag{76.34}$$

The Lagrange multiplier test is based on the OLS residuals. The test statistic is (Anselin 1988)

$$SARMA = \frac{\left[\left(\tilde{d}_\lambda\right)^2 D + \left(\tilde{d}_\rho\right)^2 T_{22} - 2\tilde{d}_\lambda \tilde{d}_\rho T_{12}\right]}{DT_{22} - T_{12}^2} \tag{76.35a}$$

or

$$SARMA = \frac{\tilde{d}_\lambda}{T} + \frac{\left(\tilde{d}_\lambda - \tilde{d}_\rho\right)^2}{D - T} \text{ if } W_1 = W_2 \tag{76.35b}$$

where $\tilde{d}_\lambda = (e'We)/(e'e/n)$, $\tilde{d}_\rho = (e'Wy)/(e'e/n)$, and $T_{ij} = tr\left[W_i W_j + W'_j W_j\right]$. Under the null, $SARMA$ is asymptotically distributed as a $\chi^2(2)$. If the null is rejected, the exact nature of spatial dependence is not known. Extensions of these principles to joint tests in $SARMA$ (p,q) models are derived in Anselin (2001).

76.3.3.2 Conditional Tests

This approach consists in performing a Lagrange multiplier test for a form of spatial dependence when the other form is not constrained. For instance, we test $H_0 : \lambda = 0$ in presence of ρ. The null corresponds to the spatial lag model, whereas the alternative corresponds to Eq. (76.31). The test is then based on the residuals of model Eq. (76.2) estimated by maximum likelihood. The test statistic is as follows (Anselin 1988):

$$LM^*_{ERR} = \frac{\hat{d}_\rho^2}{T_{22} - (T_{21A})^2 \hat{V}(\hat{\rho})} \tag{76.36}$$

where $T_{21A} = tr\left[W_2 W_1 A^{-1} + W'_2 W_1 A^{-1}\right]$, $A = I_N - \hat{\rho} W_1$, $\hat{\rho}$ is the maximum likelihood estimator of ρ, and $\hat{V}(\hat{\rho})$ is the estimated variance of $\hat{\rho}$ in model Eq. (76.2). Under the null, this statistic is asymptotically distributed as a $\chi^2(1)$.

Conversely, we can also $H_0 : \rho = 0$ in presence of λ; the test is then based on the maximum likelihood $\hat{\varepsilon}$ in the spatial error model Eq. (76.9). The statistic is (Anselin 1988)

$$LM^*_{LAG} = \frac{(\hat{e}B'BW_1 y)^2}{H_\rho - H_{\theta\rho}\hat{V}(\hat{\theta})H'_{\theta\rho}} \tag{76.37}$$

where $\theta = (\beta', \lambda, \sigma^2)$, $\hat{\theta}$ is the maximum likelihood estimator of θ, $B = I_N - \hat{\lambda}W_1$, and $\hat{V}(\hat{\theta})$ is the estimated variance-covariance matrix of $\hat{\theta}$ in model Eq. (76.9). The other terms are

$$H_\rho = tr\left(W_1^2\right) + tr\left(BW_1 B^{-1}\right) + \frac{\left(BW_1 X\hat{\beta}\right)'\left(BW_1 X\hat{\beta}\right)}{\hat{\sigma}^2} \tag{76.38}$$

$$H_{\theta\rho} = tr \left[\begin{array}{c} \frac{(BX)'BW_1X\hat{\beta}}{\hat{\sigma}^2} \\ tr(W_2B^{-1})BW_1B^{-1} + tr(W_2W_1B^{-1}) \\ 0 \end{array} \right] \tag{76.39}$$

Under the null, this statistic is asymptotically distributed as a $\chi^2(1)$.

76.3.3.3 Robust Tests

The third approach, suggested by Anselin et al. (1996), consists in using robust tests to a local misspecification. For instance, LM_{ERR} is adjusted so that its asymptotic distribution remains a centered $\chi^2(1)$, even in local presence of ρ. This test can be done using the OLS residuals of the simple model Eq. (76.1). Assuming $W_1 = W_2$, the modified statistic for the test $H_0 : \lambda = 0$ is

$$RLM_{ERR} = \frac{\left(\tilde{d}_\lambda - TD^{-1}\tilde{d}_\rho\right)^2}{[T(1 - TD)]} \tag{76.40}$$

Similarly, the test statistic of $H_0 : \rho = 0$ in local presence of λ is

$$RLM_{LAG} = \frac{\left(\tilde{d}_\lambda - \tilde{d}_\rho\right)^2}{D - T} \tag{76.41}$$

76.3.4 Specification Search Strategies

Tests based on Lagrange multiplier have been very popular in applied spatial econometrics in *specification search*, as they only require the estimation of the model under the null, typically, the simple model estimated by OLS. They can be combined to develop a *specific-to-general* sequential specification search strategy, that is, a forward stepwise specification search, whenever no a priori spatial specification has been chosen.

The first step consists in estimating the simple model Eq. (76.1) by means of OLS and in performing Moran's *I* test and the *SARMA* test. The rejection of the null in both cases indicates omitted spatial autocorrelation but not the form taken by this autocorrelation.

If the null hypothesis is rejected, it may be a sign of model misspecification. For instance, using a Monte Carlo experiment, McMillen (2003) shows that incorrect functional forms or omitted variables that are correlated over space might produce spurious spatial autocorrelation. It may therefore be useful to include in the model, if possible, additional variables. It can be exogenous additional variables that may eliminate or reduce spatial dependence, or exogenous spatial lags, corresponding in total or in part to the initial explanatory variables.

If the addition of exogenous variables has not eliminated spatial autocorrelation, a model incorporating a spatial lag and/or a spatial error must be estimated.

The choice between these two forms of spatial dependence can be done by comparing the significance levels of LM_{ERR} Eq. (76.31) and LM_{LAG} Eq. (76.33) and their robust versions RLM_{ERR} Eq. (76.40) and RLM_{LAG} Eq. (76.41): if LM_{LAG} (resp. LM_{ERR}) is more significant than LM_{ERR} (resp. LM_{LAG}) and RLM_{LAG} (resp. RLM_{ERR}) is significant but not RLM_{ERR} (resp. RLM_{LAG}), a spatial lag (resp. a spatial error) must be included in the regression model (Anselin and Florax 1995).

Once the spatial lag or the spatial error model has been estimated, three additional tests can be implemented. On the one hand, for a spatial lag model, LM^*_{ERR} allows checking whether an additional spatial error is still necessary. On the other hand, for a spatial error model, LM^*_{LAG} allows checking whether an additional spatial lag is still necessary. The common factor test allows checking whether the restriction $H_0 : \lambda\beta + \delta = 0$ is rejected or not. If not, Eq. (76.18) reduces to the spatial error model Eq. (76.9).

There are several drawbacks with this classical specific-to-general approach. First, the significance levels of the sequence of tests are unknown. Second, every test is conditional on arbitrary assumptions that may be tested later. The inference is then invalid if these assumptions are indeed rejected. As a consequence, the results of this approach is subject to the order in which the tests are carried out and whether or not adjustments are made in the significance levels of the sequence of tests.

Alternatively, a *general-to-specific* search strategy, that is, a forward stepwise specification search, can be implemented based on the spatial Durbin model Eq. (76.18) as it encompasses most spatial specifications. Model Eq. (76.18) is estimated, and testing is performed using Wald statistics or likelihood ratio statistics. Then, the failure to reject the common factor constraints suggests a spatial error model, while rejection of these constraints suggests a spatial lag model. In the first case, the significance of the spatial error coefficient is tested; if it is significant, the final specification is the error model Eq. (76.9); if it is not, the final model is the simple model Eq. (76.1). Likewise, in the second case, the significance of the spatial lag coefficient is tested; if it is not significant, the final model selection is the standard regression model. Simulation experiments performed by Florax et al. (2003) compare the specific-to-general and the general-to-specific strategies and provide some evidence of better performances of the forward strategy, in terms of power and accuracy.

76.3.5   Non-nested Tests

The basis of these specification search strategies above is that the competing models are nested within a more general model (spatial Durbin model). However, for non-nested alternatives, other strategies must be devised. For instance, Kelejian and Piras (2011) have extended the *J*-test procedure to a spatial framework. The null hypothesis corresponds to a spatial error-spatial lag model as in Eq. (76.34) with similar weights, while the alternative hypothesis corresponds to a set of G models that differ with the model in H_0 with respect to the specification of the regressor matrix, the weighting matrix, the disturbance term, or a combination of these three.

76.3.6 Spatial Autocorrelation and Spatial Heterogeneity

Spatial autocorrelation and spatial heterogeneity are often both present in regressions. We have already underlined that heteroscedasticity is implied by the presence of a spatial lag or a spatial error term. More generally, these two effects entertain complex links. First, there may be observational equivalence between these two effects in a cross section (Anselin and Bera 1998). Secondly, heteroscedasticity and structural instability tests are not reliable in the presence of spatial autocorrelation. Conversely, spatial autocorrelation tests are affected by heteroskedasticity. Thirdly, spatial autocorrelation is sometimes the result of unmodeled parameter instability. In other words, if space-varying relationships are modeled within a global regression, the error terms may be spatially autocorrelated. All these elements suggest that both aspects cannot be considered separately. We briefly review here some tests that have tackled this issue.

76.3.6.1 Spatial Autocorrelation and Heteroscedasticity

First, a joint test of spatial error autocorrelation and heteroscedasticity consists in the sum of a Breusch-Pagan test and the LM_{ERR} (Anselin 1988). The resulting statistic is asymptotically distributed as a $\chi^2(P)$, where P is the number of variables that affect the variance (Eq. 76.23). Alternatively, Kelejian and Robinson (1998) derive a joint test for spatial autocorrelation and heteroscedasticity that does not require the normality assumption for the error terms and the regression model to be linear.

Conditional tests may also be performed. On the one hand, a Lagrange multiplier test of spatial autocorrelation in a regression with heteroscedastic error terms may be derived. Let $\hat{\Omega}$ be the estimated diagonal variance-covariance matrix, then the heteroscedastic LM statistics becomes (Anselin 1988):

$$LM = \frac{\left(e'\hat{\Omega}^{-1}We\right)^2}{tr\left(WW + W'\hat{\Omega}^{-1}W\hat{\Omega}\right)'} \tag{76.42}$$

where e is the vector of residuals in the heteroscedastic regression. This statistic is asymptotically distributed as a $\chi^2(1)$.

On the other hand, a test of heteroscedasticity in a spatial lag model or a spatial error model can be performed. In the first case, a Breusch-Pagan statistic is computed on the ML residuals, while in the second case, it is performed on spatially filtered residuals in the ML estimation.

76.3.6.2 Spatial Autocorrelation and Parameter Instability

In the case of discrete parameter heterogeneity under the form of spatial regimes in a homoscedastic model, a test of equality of some or all parameters between regimes can be performed using a standard Chow test. However, when error spatial

autocorrelation and/or heteroscedastic is present, this must be adjusted. Formally, without loss of generality, consider a model with two regimes:

$$
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & X_2 \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \end{bmatrix}
\tag{76.43}
$$

Let $\varepsilon = [\varepsilon_1' \ \varepsilon_2']$ and the variance-covariance matrix: $\Psi = E(\varepsilon\varepsilon')$. The test of parameter stability is $H_0 : \beta_1 = \beta_2$.

When $\Psi = \sigma^2 \Omega$, then the test statistic is (Anselin 1988)

$$
C_G = \frac{\hat{e}'_c \hat{\Omega}^{-1} \hat{e}_c - \hat{e}'_L \hat{\Omega}^{-1} \hat{e}_L}{\hat{\sigma}^2}
\tag{76.44}
$$

where $\hat{e}_c$ is the vector of estimated residuals of the constrained model and $\hat{e}_L$ the vector of estimated residuals of the unconstrained residuals. This statistic is asymptotically distributed as a $\chi^2(K)$, where K is the number of explanatory variables in the model.

Whenever the break affects the spatial coefficient, Mur et al. (2010) suggest LM tests. For instance, assume a spatial lag model where a simple break (such a center vs. periphery) only affects the parameter of spatial dependence:

$$
\begin{aligned}
y &= \rho_0 W y + \rho_1 W^* y + X\beta + \varepsilon \\
\varepsilon &\to iid\left(0, \sigma^2 I_N\right)
\end{aligned}
\tag{76.45}
$$

where ρ_0 is the spatial lag coefficient pertaining to the second regime, ρ_1 represents the difference between the first regime and the second regime, and W^* is a weights matrix defined as $w^*_{ij} = w_{ij}$ if location i or location j belongs to the first regime and $w^*_{ij} = 0$ otherwise. Then the LM statistic for the test $H_0 : \rho_1 = 0$ is

$$
LM_{LAG}^{BREAK} = \frac{\left[\frac{y'W^*\tilde{\varepsilon}}{\hat{\sigma}^2} - tr\tilde{A}^{-1}W^*\right]^2}{\hat{\sigma}^2}
\tag{76.46}
$$

where $\tilde{\varepsilon}$ is the vector of residuals of the ML estimation of Eq. (76.2), $\tilde{\sigma}^2$ is the corresponding estimated variance, $\tilde{A} = I_N - \tilde{\rho}W$ where $\tilde{\rho}$ is the ML estimation in Eq. (76.2), and $\hat{\sigma}^2$ is the ML estimated variance corresponding to the linear restriction of the null. This statistic is asymptotically distributed as a $\chi^2(1)$.

A spatial error model with a structural break affecting the spatial error parameter is

$$
\begin{aligned}
y &= X\beta + \varepsilon \\
\varepsilon &= \lambda_0 W\varepsilon + \lambda_1 W^*\varepsilon + u \\
u &\to iid\left(0, \sigma^2 I_N\right)
\end{aligned}
\tag{76.47}
$$

The LM statistic for the test $H_0 : \lambda_1 = 0$ is as follows:

$$LM_{LAG}^{BREAK} = \frac{\left[\frac{\tilde{\varepsilon}' W^* \tilde{B} \tilde{\varepsilon}}{\tilde{\sigma}^2} - tr\tilde{B}^{-1} W^*\right]^2}{\hat{\sigma}^2} \tag{76.48}$$

where $\tilde{\varepsilon}$ is the vector of residuals of the ML estimation of Eq. (76.9), $\tilde{\sigma}^2$ is the corresponding estimated variance, $\tilde{B} = I_N - \tilde{\lambda}W$ where $\tilde{\lambda}$ is the ML estimation in Eq. (76.9), $\hat{\sigma}^2$ is the ML estimated variance corresponding to the linear restriction of the null. This statistic is asymptotically distributed as a $\chi^2(1)$.

76.4 Conclusion

The objective of this chapter was to provide a concise review of specification issues in spatial econometrics. We focused on the way spatial effects may be incorporated into regression models and on specification testing. We first presented the most commonly used spatial specifications in a cross-sectional setting in the form of linear regression models including a spatial lag and/or a spatial error term, heteroscedasticity, or parameter instability. Second, we presented a set of specification tests that allow checking deviations from a standard, that is, nonspatial, regression model. An important space has been devoted to LM tests as they only require the estimation of the model under the null. Unidirectional, multidirectional, and robust LM tests are now in the standard toolbox of spatial econometrics. They are still frequently used in applied work, even though the technical/numerical difficulties associated to the estimation of spatial models have become much more tractable, even for very large samples. Because of the complex links between spatial autocorrelation and spatial heterogeneity, we have given some attention to the specifications incorporating both aspects and to the associated specification tests.

References

Anselin L (1988) Spatial econometrics, methods and models. Kluwer, Dordrecht

Anselin L (2001) Rao's score test in spatial econometrics. J Stat Plan Infer 97:113–139

Anselin L (2003) Spatial externalities, spatial multipliers and spatial econometrics. Int Reg Sci Rev 26:153–166

Anselin L (2010) Thirty years of spatial econometrics. Pap Reg Sci 89:3–25

Anselin L, Bera AK (1998) Spatial dependence in linear regression models with an application to spatial econometrics. In: Ullah A, Giles DEA (eds) Handbook of applied economics statistics. Springer, Berlin

Anselin L, Florax RGJM (1995) Small sample properties of tests for spatial dependence in regression models: some further results. In: Anselin L, Florax RJGM (eds) New directions in spatial econometrics. Springer, Berlin

Anselin L, Bera A, Florax RGJM, Yoon M (1996) Simple diagnostic test for spatial dependence. Reg Sci Urban Econ 26:77–104

Arbia G (2011) A lustrum of SEA: recent research trends flowing the creation of the spatial econometrics association (2007–2011). Spat Econ Anal 6:377–395

Casetti E, Can A (1999) The econometric estimation and testing of DARP models. J Geogr Syst 1:91–106

Cliff A, Ord JK (1972) Testing for spatial autocorrelation among regression residuals. Geogr Anal 4:267–284

Conley TG (1999) GMM estimation with cross-sectional dependence. J Econom 92:1–44

Florax RJGM, Folmer H, Rey SJ (2003) Specification searches in spatial econometrics: the relevance of Hendry's methodology. Reg Sci Urban Econ 33:557–579

Fotheringham AS, Brundson C, Charlton M (2004) Geographically weighted regression: the analysis of spatially varying relationships. Wiley, Chichester

Kelejian HH, Piras G (2011) An extension of Kelejian's J-test for non-nested spatial models. Reg Sci Urban Econ 41:281–292

Kelejian HH, Prucha I (2007) HAC estimation in a spatial framework. J Econom 140:131–154

Kelejian HH, Robinson DP (1995) Spatial correlation: a suggested alternative to the autoregressive model. In: Anselin L, Florax RJGM (eds) Advances in spatial econometrics. Springer, Heidelberg

Kelejian HH, Robinson DP (1998) A suggested test for spatial autocorrelation and/or heteroskedasticity and corresponding Monte-Carlo results. Reg Sci Urban Econ 28:389–417

LeSage J, Pace KP (2009) Introduction to spatial econometrics. CRC Press, Boca Raton

McMillen DP (2003) Spatial autocorrelation or model misspecification? Int Reg Sci Rev 26:208–217

Moran P (1950) A test for the serial dependence of the residuals. Biometrika 35:255–260

Mur J, Lopez F, Angulo A (2010) Instability in spatial error models: an application to the hypothesis of convergence in the European case. J Geogr Syst 12:259–280

Pace RK, LeSage JP (2008) A spatial Hausman test. Econ Lett 101:282–284

Pace RK, LeSage JP (2004) Spatial autoregressive local estimation. In: Getis A, Mur J, Zoller H (eds) Spatial statistics and spatial econometrics. Palgrave MacMillan