Spatial Panel Data Analysis

J.Paul Elhorst
Faculty of Economics and Business, University of Groningen, the Netherlands

Elhorst, J.P. (2017) Spatial Panel Data Analysis. In: Shekhar S., Xiong H., Zhou X. (Eds.) *Encyclopedia of GIS*, 2nd edition, pp. 2050-2058. Springer International Publishing, Cham, Switzerland.

1. Historical background

Since the turn of the century, the spatial econometrics literature has shifted its interest from the specification and estimation of econometric relationships based on cross-sectional data to spatial panels. Spatial panels refer to georeferenced point data over time of individuals, households, firms, houses or public services such as universities and hospitals, or they refer to spatial units such as zip codes, neighborhoods, municipalities, counties, regions, jurisdictions, states or countries. One well-known example of a spatial panel that has been widely used for illustration purposes in many empirical studies is Baltagi and Li's (2004) dataset on cigarette demand in 46 American states over the period 1963-1992. In this study the dependent variable, real per capita sales of cigarettes measured in packs per person aged 14 years and older, is regressed on the average retail price of a pack of cigarettes and real per capita disposable income. The data is available at www.regroningen.nl/elhorst.

The main advantage of working with spatial panels is that one can control for spatial and time specific effects. Spatial units of observation are likely to differ in their background variables, which are usually space-specific time-invariant variables that do affect the dependent variable, but which are difficult to measure or hard to obtain. One unit is located at the seaside, the other just at the border; one is a rural area located in the periphery of a country, the other an urban area located in the center; norms and values regarding labor, crime and religion in one spatial unit might differ substantially from those in another unit, to mention just a few examples. Failing to account for these effects, as in a cross-sectional study, increases the risk of obtaining biased estimation results.

Similarly, the justification for time specific effects is that they control for all spatial-invariant variables whose omission could bias the estimates in a typical time-series study. One time period is marked by an economic recession, the other by a boom; changes in legislation or government policy can significantly affect the functioning of an economy, such that observations before these changes might be significantly different from those after it. Reasons to control for spatial and time-specific effects when explaining cigarette demand are provided by Baltagi and Levin (1986, 1992).

The main purpose of spatial econometric models is to test for the existence of spatial interaction effects, and related to that, spatial spillover effects. Spatial spillovers are a main interest in regional science, economic geography, and related fields. Many theories predict that changes to explanatory variables in a particular unit *i* impact the dependent variable not only in unit *i* itself, but also in other units *j* (≠ *i*). The main motivation to consider spatial interaction effects and to test for spatial spillover effects in the spatial panel on cigarette demand is the so-called bootlegging effect: do consumers purchase cigarettes in nearby states, legally or illegally, if there is a price advantage? Improved accessibility to spatial panels and software developed to deal with spatial panel data models has increased the use of spatial econometric models over the past decade. For a better understanding it is first demonstrated in the next section how to derive spillover effects from a spatial econometric model and what the differences are between spatial interaction and spillover effects. Then different types of spatial econometric models and modeling selection strategies are presented, as well as their limitations. Finally, two promising approaches are set out that have recently been developed to overcome these limitations.

2. Scientific fundamentals

A spatial econometric model is a linear regression model extended to include spatial interaction effects. A standard linear regression model for panel data without spatial interaction effects takes the form

$$Y_t = X_t\beta + \mu + \xi_t \iota_N + \varepsilon_t, \tag{1}$$

where Y_t denotes an N×1 vector consisting of one observation on the dependent variable for every unit in the sample (i=1,…,N) at time t (t=1,…,T), X_t denotes an N×K matrix of exogenous explanatory variables at time t associated with the K×1 vector β, and $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{Nt})'$ is a vector of disturbance terms, where ε_{it} are independently and identically distributed error terms for all *i* with zero mean and variance σ^2. $\mu = (\mu_1, \dots, \mu_N)'$ and ξ_t multiplied with ι_N denoting an N×1 vector of ones represent spatial and time specific effects, which are optional and may be treated as fixed effects or as random effects. In the fixed effects model, a dummy variable is introduced for each spatial unit and for each time period (except one to avoid perfect multicollinearity), while in the random effects model, μ_i and ξ_t are treated as random variables that are independently and identically distributed with zero mean and variance σ_μ^2 and σ_ξ^2, respectively. Furthermore, it is assumed that the random variables μ_i, ξ_t and ε_{it} are independent of each other.

2.1 Interaction and spillover effects

Three different types of spatial interaction effects may be considered. The first is an endogenous interaction effect, which measures whether the dependent variable of unit *i* depends on the dependent variables of other units *j (j≠ i)* and vice versa. This effect can be denoted by WY_t, where the spatial weights matrix *W* is a positive *N*×*N* matrix that describes the structure of dependence between the units in the sample. The second are exogenous interaction effects in that the dependent variable of unit *i* depends on the explanatory variables of other units *j (j ≠ i).* This effect can be denoted by WX_t. Note that if the number of explanatory variables is *K*, the maximum number of exogenous interaction effects is also

K. Finally, an interaction effect among the error terms might occur, denoted by Wu_t, indicating that units may behave similarly because they share the same unobserved characteristics or face similar unobserved environments. Summing up, a total of *K*+2 spatial interaction effects is possible.

If all spatial interaction effects would be added to the linear regression model for panel data, which is rarely done for reasons to be explained in the next section, one obtains:

$$Y_t = \rho W Y_t + X_t \beta + W X_t \theta + \mu + \xi_t \iota_N + u_t, \qquad u_t = \lambda W u_t + \varepsilon_t. \tag{2}$$

The scalar parameters ρ and λ and the $K \times 1$ vector of parameters θ measure the strength of spatial dependence between the units. Whether spatial and time specific effects need to be included and be treated as fixed or random effects in spatial econometric models can easily be tested using likelihood ratio (LR) tests and the Hausman test. The LR test is based on minus two times the difference between the log-likelihood function values in the restricted and the unrestricted model: $-2 \times (\log L_{\text{restricted}} - \log L_{\text{unrestricted}})$. The LR test of whether the spatial fixed effects are jointly insignificant and thus can be replaced by one single intercept, follows a chi-squared distribution with *N*-1 degrees of freedom. Similarly, the LR test of whether the time fixed effects are jointly insignificant, follows a chi-squared distribution with *T*-1 degrees of freedom. If these fixed effects appear to be jointly significant, the Hausman test can be used to investigate whether random effects can replace them. It tests whether the response parameters in the model when the spatial and time specific effects are treated as random are significantly different from those when the spatial and time specific effects are treated as fixed. If they are, the random effects model needs to be rejected in favor of the fixed effects model. The number of degrees of freedom is equal to the number of response parameters, which is *K* for the explanatory variables X_t plus 1 for WY_t if included and *K* for WX_t if included (see Elhorst, 2014a for further details).

If model (2) is rewritten to its reduced form (3),

$$Y_t = (I - \rho W)^{-1}(X_t \beta + W X_t \theta) + (I - \rho W)^{-1}(\mu + \xi_t \iota_N + (I - \lambda W)^{-1})\varepsilon_t), \tag{3}$$

the matrix of partial derivatives of the expectation of Y_t, $E(Y_t)$, with respect to the *k*th explanatory variable of X_t in unit 1 up to unit N is

$$\left[\frac{\partial E(Y_t)}{\partial x_{1kt}} \cdots \frac{\partial E(Y_t)}{\partial x_{Nkt}}\right] = (I - \rho W)^{-1}(\beta_k + W\theta_k). \tag{4}$$

whose diagonal elements represent direct effects and its off-diagonal elements indirect or spatial spillover effects and which are independent of *t*. Note that spatial and time specific effects, as well as the error term, drop out due to considering the expectation of Y_t.

To reduce the *K* different *N*x*N* matrices of direct and spillover effects to a manageable set of information, LeSage and Pace (2009) propose to report one direct effect measured by the average of the diagonal elements and one spillover effect measured by the average row or column sums of the off-diagonal elements. In addition to this, one can distinguish global and local spillover effects. Global spillovers occur when a change in X_t at any location will be transmitted to all other locations, also if two locations according to *W* are unconnected. This requires that $\rho \neq 0$. In contrast, local spillovers are those that occur at other locations only if according to W they are connected to each other. This requires that $\theta \neq 0$. Note that the choice between global and local spillovers is also related to the specification of W. A global spillover model with a spatial weights matrix that is sparse –a matrix in which only a limited number of elements is non-zero, such as a binary contiguity matrix– is more likely than with a dense matrix. Conversely, a local spillover model with a spatial weights matrix that is dense –a matrix in which all off-diagonal elements are non-zero, such as an inverse distance matrix– is more likely than with a sparse matrix.

Generally, it is harder to find empirical evidence in favor of significant spatial spillover effects than in favor of spatial interaction effects. This is because the former are composed of three parameters, among which two are spatial interaction effects. If already one of these three parameters happens to be insignificant, the spatial spillover effect also tends to become insignificant. For this reason, most empirical studies find only a fraction of their *K* explanatory variables to produce significant spatial spillover effects. This is not a weakness, but a validation that the hypothesis that a change in one of the determinants in one unit affects the dependent variable in another unit is strong.

2.2 Key spatial econometric models and their limitations

The full model with all possible spatial interaction effects, known as the general nesting spatial model (GNS), is seldom used in empirical research. There are two reasons for this (Elhorst, 2014a). First, a formal proof under which conditions the parameters of this model are identified is not available yet. Second, there is a problem of overfitting. Even if the parameters are not identified, they can be estimated, but have the tendency either to blow each other up or to become insignificant, as a result of which this model does not help to choose among simpler models with less spatial interaction effects. Another reason often mentioned to abandon this model is Manski's (1993) reflection problem, but this is based on a misconception. Manski demonstrated the failure of identification if interaction effects are just the means of all units belonging to a group, that is, if they are obtained by setting all elements of the W matrix equal to 1/*N*. However, if it is assumed that units cannot interact with themselves, which is a reasonable assumption when working with spatial data, the diagonal elements of the W matrix will be equal to zero, as a result of which the reflection problem no longer holds (Bramoullé et al., 2009).

Simpler models containing one type or two types of spatial interaction effects are known under different designations and abbreviations: (i) the spatial autoregressive (SAR) model containing the endogenous interaction effect WY_t, (ii) the spatial error model (SEM) containing the interaction effect (correlated effect) among the error terms Wu_t, (iii) the spatial lag of *X* model (SLX) containing the exogenous interaction effects WX_t, (iv) the spatial autoregressive combined (SAC) model containing both WY_t and Wu_t, also known as the

SARAR or Cliff-Ord type spatial model, (v) the spatial Durbin (SDM) model containing both WY_t and WX_t, and (vi) the spatial Durbin error (SDEM) model containing both WX_t and Wu_t. The parameters of these models have all been shown to be identified and to be free of overfitting. They can be estimated in Stata, R and Matlab. Sometimes it requires some additional programming in addition to standard routines, but in principle all software produce parameter estimates, direct and spillover effects, and significance levels.

By far the biggest problem in empirical research is to choose between these different models and different specifications of W, especially if no reference is made to specific economic theories. Consequently, too many empirical studies follow a statistical approach driven by data-analytic considerations and only consider the SAR and/or SEM model with one type of spatial interaction effect. Moreover, many of these studies are further limited to one or a few pre-specified W matrices. Other empirical studies go a step further by considering the SAC and SDM models with two types of spatial interaction effects, but again based on one or a few pre-specified W matrices. If these studies already provide a well-founded background for certain spatial interaction effects, they often lack guidance of how the spatial weights matrix should be specified. Most often, spatial weights matrices are used whose appeal seems to lie in the frequency of their use. For these reasons many empirical studies can easily be criticized, such as in the special theme issue of the Journal of Regional Science (Volume 52, Issue 2); see Partridge et al. (2012) for an overview of the contributing papers.

One limitation of the SAR, SAC and SDM models is that the spillover effects are global by construction ($\rho \neq 0$), while global spillovers are often more difficult to justify than local spillovers (see Halleck Vega and Elhorst, 2015; and the references therein). In this respect, the SLX and SDEM models whose spillover effects are local ($\rho = 0, \theta \neq 0$) are generally overlooked. Another limitation of the SAR and SAC models is that that the ratio between the spillover effect and direct effect is the same for every explanatory variable, which is unlikely to be the case in many empirical studies. One limitation of the SEM model is that the spillover effects are set to zero by construction ($\rho = 0, \theta = 0$). The direct effect, i.e. the effect of a change of a particular explanatory variable in one unit on the dependent variable of that unit, is the only information provided.

In view of this it should be clear that the way of thinking and the model selection strategies that are used in most empirical studies to determine the structure of spatial processes need revision. Two approaches have recently been developed that are promising. The first is developed by LeSage (2014) and based on Bayesian comparison methods, and the second by Halleck Vega and Elhorst (2015) and based on taking the SLX model as point of departure. These approaches are set out in the next section.

3. Future directions

3.1 The Bayesian comparison approach

According to LeSage (2014), there are only two spatial econometric models that need to be considered: the spatial Durbin model (SDM) and the spatial Durbin error model (SDEM). The first model implies that spillover effects are global and the second that they are local. If it can be argued on theoretical or substantive aspects of the problem that one type of spillover effects is more likely than the other, the corresponding model can be taken as point of departure. If this is not possible or if two theories circulate of which one implies the SDM and the other SDEM, these models are better tested against each other.

To test whether the SAR or SEM model is more appropriate to describe the data than a model without any spatial interaction effects, researchers tended to use (robust) Lagrange Multiplier (LM) tests for interaction effects among the dependent variable or among the error terms. These tests, which have been developed by Anselin et al. (1996) in a cross-sectional setting and by Elhorst (2010) in a spatial panel data setting, have become very popular in empirical research. Unfortunately, they are not very helpful in finding the right model. The same applies to the emerging literature on the J-test (Kelejian and Piras, 2016; but note a strong point of this paper is that their J-test also allows for endogeneity of the regressors, see the next section). One reason is that these tests do not account for exogenous interaction effects. Consequently, it might be due to omitting WX_t variables that strong evidence is found in favor of the interaction effect among the dependent variable or among the error term. The second reason is that that these LM and J-tests suffer from low(er) power once exogenous interaction effects are also controlled for. This is because the log-likelihood function values of the SDM and SDEM models are generally much closer to each other than those of the SAR and SEM models, and because the spatial spillover effects are often comparable numerically, even though their interpretation (global vs. local) is completely different (Elhorst, 2014a). The explanation is that the point estimates of ρ in the SDM and λ in the SDEM provide too little statistical information to choose between these models. If the spatial interaction effect among the dependent variable is ignored while it should be in, the interaction effect among the error term may counterbalance this error, and vice versa.

LeSage (2014) demonstrates that a Bayesian comparison approach considerably simplifies the task of selecting an appropriate model. This approach determines the Bayesian posterior model probabilities of SDM and SDEM given a particular spatial weights matrix, as well as the Bayesian posterior model probabilities of different W matrices given a particular model specification. Typically, these sets of probabilities take the form as illustrated in Table 1.

Table 1. Comparison of model specifications and spatial weights matrices

W Matrix	Statistics	SDM	SDEM
Binary Contiguity	log marginal likelihood	3616.03	3611.80
	model probabilities	0.9855	0.0145
Inverse distance	log marginal likelihood	3444.87	3455.44
	model probabilities	0.0000	1.0000
K=6 nearest neighbors	log marginal likelihood	3613.06	3613.60
	model probabilities	0.3676	0.6324

Source: Firmino et al. (2014)

The log marginal likelihood of a model reported in this table is obtained by integrating out all parameters of the model over the entire parameter space on which they are defined. If the log marginal likelihood value of one model is higher than that of another model, the Bayesian posterior model probability is also higher. It should be stressed that the model parameters are not estimated and so cannot be reported when applying the Bayesian comparison approach. This is the main strength of this approach. Whereas the popular likelihood ratio, Wald and/or Lagrange multiplier statistics compare the performance of one model against another model based on specific parameter estimates within the parameter space, the Bayesian approach compares the performance of one model against another model, in this case SDM against SDEM, on their entire parameter space. Inferences drawn on the log marginal likelihood function values for the SDM and SDEM model are further justified because they have the same set of explanatory variables ($[X_t \; WX_t]$) and are based on the same uniform prior for ρ and λ. This prior takes the form $p(\rho) = p(\lambda) = 1/D$, where $D = 1/\omega_{max} - 1/\omega_{min}$ and ω_{max} and ω_{min} represent respectively the largest and the smallest (negative) eigenvalue of the spatial weights matrix W. This prior requires no subjective information on the part of the practitioner as it relies on the parameter space ($1/\omega_{min}$, $1/\omega_{max}$) on which ρ and λ are defined, where $\omega_{max} = 1$ if W is row-normalized or normalized by its largest eigenvalue. Full details regarding the choice of model can be found in LeSage (2014) and regarding the choice of W in LeSage and Pace (2009, chs. 5 and 6).

The results reported in Table 1 show that in this particular case the first-order binary contiguity matrix in combination with the SDM model gives the best performance. If the spatial weights matrix would be specified as an inverse distance matrix, the SDEM model is a better choice. In other words, it is found that the global spillover model in combination with a sparse matrix outperforms the local spillover model in combination with a dense matrix.

3.2 The SLX approach

According to Halleck Vega and Elhorst (2015), the SLX model can be best taken as point of departure when an underlying theory is lacking. It is the simplest spatial econometric model producing flexible spatial spillover effects. In contrast to other spatial econometric models, the spatial weights matrix W in the SLX model can be parameterized. Moreover, standard instrumental variables (IV) approaches can be used to investigate whether (part of) the X_t variables and their spatially lagged values WX_t are endogenous.

From the overview in section 3 we learned that the SAR, SAC and SEM models are of limited use in empirical research due to initial restrictions on the spillover effects they can potentially produce. By contrast, the spillover effects produced by the SLX, SDM and SDEM models are flexible, as they can take any value. Since both SDM and SDEM are extensions of the SLX model, the latter is the simplest one of this family of models.

Suppose a researcher wants to use a simple parametric approach applied to the elements of an inverse distance matrix $w_{ij} = 1/d^{\gamma}$, where γ is a parameter to be estimated, to obtain more information on the strength of interdependencies among the cross-sectional observations at each point in time t, rather than to impose a certain specification of the spatial weights matrix in advance. A nonlinear, but straightforward estimation technique –the parameter vectors β and θ, given γ, and γ given β and θ can be alternately estimated until convergence occurs– can then be used to estimate the parameters of the SLX model. If γ appears to be small, observations at distant locations have relatively more impact than if this distance decay parameter appears to be large. Next, the Bayesian comparison approach can be used to find out whether the SLX model needs to be extended to the SDM or SDEM model specification, keeping γ fixed. Unfortunately, due to computational problems, a numerical procedure to estimate γ also within these models is not available yet; see the perfect solution problem discussed in Halleck Vega and Elhorst (2015). Possible and useful extensions of parameterizing the spatial weights matrix in the SLX model using $w_{ij} = 1/d^{\gamma}$ are one γ_k (k=1,…,K) for every single exogenous interaction effect, more variables (V) determining the degree of interaction among spatial units, each with their own coefficient (δ), $w_{ij} = V_i^{\delta_i} V_j^{\delta_j} / d_{ij}^{\gamma}$, as in the popular gravity model, or alternatively, semi-parametric and non-parametric approaches (McMillen, 2013).

A final advantage of the SLX model over other spatial econometric models is that nonspatial econometric techniques can be used to test for endogeneity among the explanatory variables $[X_t \; WX_t]$. It concerns the Hausman test for endogeneity in combination with tests for the validity of the instruments to assess whether they satisfy the relevance and exogeneity criterions. The attention for endogenous regressors (other than the endogenous interaction effects WY_t) is important since researchers face uncertainty about the endogeneity not only of the explanatory variables X_t themselves, but also of their exogenous interaction effects WX_t. Halleck Vega and Elhorst (2015) find that the price of cigarettes observed in neighboring states may be used as an exogenous determinant of cigarette demand in the U.S., whereas the price of cigarettes observed in the own state may not. Apparently, consumption has feedback effects on the price in the own state, but if consumers decide to buy more cigarettes in neighboring states due to a price increase in their own state this has no significant feedback effects on prices there too. Today endogeneity of the regressors is getting more attention in the spatial econometrics literature (Drukker et al., 2013); Halleck and Vega (2015) is one of the first studies making a distinction between regular explanatory variables and interaction effects among these explanatory variables.

4. Empirical applications

Since there are numerous applications of spatial panel data models in the literature, we limit this overview to those studies that are based on Baltagi and Li's (2004) spatial panel of cigarette demand. This dataset was used for the first time by Baltagi and Levin (1986, 1992), but then respectively over the periods 1963-1980 and 1963-1988. Table 2 provides an overview of the different studies and shows the progress that has been made over the years. Most studies now control for spatial and time period fixed effects. Elhorst (2014b) explicitly tests for these controls and finds that this model specification outperforms its counterparts without spatial and/or time fixed effects, as well as the random effects model. Many studies also include the dependent variable lagged in time so as to control for habit persistence. In that case one can distinguish both short-term and long-term direct and spatial spillover effects; Elhorst (2013) and Debarsy et al. (2014) provide the mathematical formulas of these effects. Most studies also share the view that exogenous interaction effects should be included, but whether it is the SDM or the SDEM specification that best describes the data is still unclear. Halleck Vega and Elhorst (2015) argue that including endogenous interaction effects is difficult to justify, since it would mean that a change in price or income in a particular state potentially impacts consumption in all states, including states that according to W (such as California and Illinois) are unconnected. Finally, most studies adopt a row-normalized binary contiguity matrix. One exception is Debarsy et al. (2014) who also consider a row-normalized matrix based on state border miles in common between states. More importantly, Kelejian and Piras (2014) and Halleck and Vega (2015) are among the first going beyond an exogenous pre-specified spatial weights matrix with fixed weights.

Table 2. Spatial panel data studies on cigarette demand

Study	Panel	Dynamic	Spatial	W
Baltagi and Levin (1986)	TFE or TRE	+	SLX, price	-
Baltagi and Levin (1992)	SFE or SRE + TFE	+	SLX, price	-
Baltagi and Li (2004)	SFE or SRE	-	SEM	BC
Elhorst (2005)	SFE + TFE	+	SDEM	BC
Elhorst (2013)	SFE + TFE	+	SDM	BC
Debarsy et al. (2014)	SRE	+	SDM	BC, border lengths
Kelejian and Piras (2014)	SFE + TFE	-	SAR	Endogenous
Elhorst (2014b)	SFE + TFE	-	SDM	BC
Halleck Vega and Elhorst (2015)	SFE + TFE	-	All SLX	BC parameterized IV

Panel: SFE = spatial fixed effects, SRE = spatial random effects, TFE = time fixed effects, TRE = time random effects; Dynamic: + = Y_{t-1} included; Spatial: See main text for abbreviations, All = SAR, SEM, SLX, SAC, SDM, SDEM, GNS; W: BC = binary contiguity matrix, IV = inverse distance matrix.

5. Conclusion

Today a (spatial) econometric practitioner has the choice of many models. It should be investigated whether or not spatial and/or time specific effects should be accounted for and, if so, whether they should be treated as fixed or as random effects. It should be investigated which type of spatial interaction effects should be accounted for: (1) an endogenous spatial interaction effect, (2) exogenous spatial interaction effects, (3) an interaction effect among the error terms, or (4) a combination of these. Different specifications of the spatial weights matrix should be tested against each other. Finally, it should be tested whether one or more explanatory variables are endogenous. A systematic procedure that works under all circumstances does not exist, but recently two promising approaches have been developed that throw more light on many of these model choices.

6. Further reading

Recommended introductory textbooks in spatial econometrics are Anselin (1988) and LeSage and Pace (2009), and in spatial panel data econometrics, Anselin et al. (2008) and Elhorst (2010, 2014a). Recommended journal articles or book chapters providing a good overview of the field are Anselin (2010), Lee and Yu (2010), and different contributions to the Handbook of Regional Science edited by Fischer and Nijkamp (2013).

References

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

Anselin L (2010) Thirty years of spatial econometrics. Papers in Regional Science 89: 3-25

Anselin L, Bera AK, Florax R, Yoon MJ (1996) Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics 27: 77–104

Anselin L, LeGallo J, Jayet H (2008) Spatial panel econometrics. In: Matyas L, Sevestre P (eds) The econometrics of panel data, fundamentals and recent developments in theory and practice, 3rd edn. Kluwer, Dordrecht, p 901-969

Baltagi BH, Levin D (1986) Estimating dynamic demand for cigarettes using panel data: the effects of bootlegging, taxation and advertising reconsidered. The Review of Economics and Statistics 68: 148–155

Baltagi BH, Levin D (1992) Cigarette taxation: raising revenues and reducing consumption. Structural Change and Economic Dynamics 3: 321-335

Baltagi BH, Li D (2004) Prediction in the panel data model with spatial autocorrelation. In: Anselin L, Florax R, Rey SJ (eds) Advances in spatial econometrics: methodology, tools, and applications. Springer, Berlin, p 283–295

Bramoullé Y, Djebbari H, Fortin B (2009) Identification of Peer Effects through Social Networks. Journal of Econometrics 150: 41-55

Debarsy N, Ertur C, LeSage JP (2012) Interpreting dynamic space-time panel data models. Statistical Methodology 9: 158–171

Drukker DM, Egger P, Prucha IR (2013) On two-step estimation of a spatial autoregressive model with autoregressive disturbances and endogenous regressors. Econometric Reviews 32: 686-733

Elhorst JP (2005) Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels. Geographical Analysis 37: 85-106

Elhorst JP (2010) Spatial panel data models. In: Fischer MM, Getis A (eds) Handbook of applied spatial analysis. Springer, Berlin, p 377-407

Elhorst JP (2013) Spatial panel models. In: Fischer MM, Nijkamp P (eds) Handbook of regional science. Springer, Berlin, p 1637-1652

Elhorst JP (2014a) Spatial econometrics: from cross-sectional data to spatial panels. Springer, Heidelberg

Elhorst JP (2014b) Matlab software for spatial panels. International Regional Science Review 37: 385-401

Firmino DC da S, Elhorst JP, Neto R da MS (2016) Urban and rural population growth in a spatial panel of municipalities. Regional Studies. Early view doi:10.1080/00343404.2016.1144922

Fischer MM, Nijkamp P (2013) Handbook of regional science. Springer, Berlin Heidelberg

Halleck Vega S, Elhorst JP (2015) The SLX model. Journal of Regional Science 55: 339-363.

Kelejian HH, Piras G (2014) Estimation of spatial models with endogenous weighting matrices, and an application to a demand model for cigarettes. Regional Science and Urban Economics 46: 140–149

Kelejian HH, Piras G (2016) An extension of the J-test to a spatial panel data framework. Journal of Applied Econometrics 31: 387–402

Lee LF, Yu J (2010) Some recent developments in spatial panel data models. Regional Science and Urban Economics 40: 255-271

LeSage JP (2014) Spatial econometric panel data model specification: a Bayesian approach. Spatial Statistics 9: 122-145

LeSage JP, Pace RK (2009) Introduction to spatial econometrics. Taylor and Francis, Boca Raton

Partridge MD, Boarnet MG, Brakman S, Ottaviano G (2012) Introduction: whither spatial econometrics? Journal of Regional Science 52: 167–171

Manski CF (1993) Identification of endogenous social effects: the reflection problem. Review of Economic Studies 60: 531-542

McMillen DP (2013) Quantile Regression for Spatial Data. Springer, Heidelberg