JOURNAL OF REGIONAL SCIENCE

JOURNAL OF REGIONAL SCIENCE, VOL. 55, NO. 3, 2015, pp. 339–363

THE SLX MODEL*

Solmaria Halleck Vega

Faculty of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. E-mail: s.m.halleck.vega@rug.nl

J. Paul Elhorst

Faculty of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands. E-mail: j.p.elhorst@rug.nl

ABSTRACT. We provide a comprehensive overview of the strengths and weaknesses of different spatial econometric model specifications in terms of spillover effects. Based on this overview, we advocate taking the SLX model as point of departure in case a well-founded theory indicating which model is most appropriate is lacking. In contrast to other spatial econometric models, the SLX model also allows for the spatial weights matrix **W** to be parameterized and the application of standard econometric techniques to test for endogenous explanatory variables. This starkly contrasts commonly used spatial econometric specification strategies and is a complement to the critique of spatial econometrics raised in a special theme issue of the *Journal of Regional Science* (Volume 52, Issue 2). To illustrate the pitfalls of the standard spatial econometrics approach and the benefits of our proposed alternative approach in an empirical setting, the Baltagi and Li (2004) cigarette demand model is estimated.

1. INTRODUCTION

Spatial spillovers are a main interest in regional science. They can be defined as the impact of changes to explanatory variables in a particular unit i on the dependent variable values in other units j $(\neq i)$.[1] A valuable aspect of spatial econometric models is that the magnitude and significance of spatial spillovers can be empirically assessed. Improved accessibility to spatial panels and software, along with advances in the field, has increased the use of these methods over the past years.

Recently, spatial econometrics has been appraised in a special theme issue of the *Journal of Regional Science (JRS)*. Partridge et al. (2012) provide an overview of the three contributing papers of Gibbons and Overman (2012), McMillen (2012), and Corrado and Fingleton (2012). Gibbons and Overman's (2012) critique focuses on identification problems. For a better understanding, it is important to distinguish three different identification problems in spatial econometrics. The first is mentioned by McMillen (2012,

*The authors gratefully acknowledge James LeSage and Kelley Pace for their thoughtful review of a previous version in December 2012, as well as valuable comments from participants of the 59th Annual North American Meetings of the Regional Science Association International, Ottawa, November 2012, in particular Dennis Robinson and Oleg Smirnov, the Sorbonne University seminar in Paris, April 2013, the Econometrics of Social Interaction Symposium, University of York, May 2013, the 12th International Workshop Spatial Econometrics and Statistics, Université d'Orléans, June 2013, in particular Raymond Florax, the 53rd European Regional Science Association Congress, Palermo, August 2013. The authors also gratefully acknowledge the managing editor Mark Partridge and four anonymous referees for their numerous insightful comments and constructive suggestions that have helped to significantly improve their work. The authors are responsible for any errors that may remain and for the views expressed in the paper.

Received: February 2013; revised: December 2014; accepted: December 2014.

[1]Units could be firms, cities, regions, and so forth depending on the nature of the study.

 DOI: 10.1111/jors.12188

p. 196): "There are $N(N$-1) potential relationships among the observations, but only N data observations are available." This problem can be solved by imposing model structure at the start in the form of a spatial weights matrix **W**, which reduces the number of parameters to be estimated from $N(N$-1) to a number that corresponds to the number of spatial interaction effects that are considered in the spatial econometric model.[2] In this respect, spatial econometrics differs from the neighborhood effects literature. Whereas spatial econometric researchers can observe the geographical location of the units in the sample and utilize this information to impose structure on **W** by placing more weight on closer observations, neighborhood effects researchers often do not have this kind of information, mainly because group formation tends to be independent of geographical location. Often groups or cells of similar types of people within a neighborhood are formed based on gender, age, education, etc. The question which groups should be formed creates a new identification problem.

The second identification problem is stressed by Gibbons and Overman (2012) in section 2 of their paper, and also highlighted by McMillen (2012), Corrado and Fingleton (2012), and Partridge et al. (2012). Different spatial econometric models are generally impossible to distinguish without assuming prior knowledge about the true data-generating process, which is often not possessed in practice. The same applies to the **W** matrix. Although the proposition to place more weight on closer observations is widely accepted, the true **W** is generally unknown. McMillen (2012, p. 192) critiques the use of a pre-specified **W** because it demands a specific functional form, as well as the routine use of the spatial autoregressive model (SAR) and the spatial error model (SEM) as a quick fix for nearly any model misspecification issue related to space.[3] Elhorst (2010) confirms that up to 2007 spatial econometricians were mainly interested in the SAR and SEM models, and points out that the seminal book by Anselin (1988) and the testing procedure for these models based on robust Lagrange Multiplier tests developed by Anselin et al. (1996) may be considered as the main pillar behind this way of thinking. Due to this, still too many empirical papers consider only the SAR and SEM models based on one or more pre-specified **W** matrices. Corrado and Fingleton (2012) also acknowledge that too often applied studies have been driven by data-analytic considerations with an emphasis on diagnostics and empirical model validity. For this reason, they strongly argue for the use of more substantive theory in empirical spatial econometric modeling, especially regarding **W**.[4] While McMillen (2010, 2012) advocates the use of nonparametric and semiparametric methods as an alternative to simply imposing common specifications of **W** whose appeal seems to lie in the frequency of their use, Corrado and Fingleton (2012) propose a hierarchical model with a block-diagonal group interaction matrix taken from the neighborhood effects literature.

[2]An exception is if **W** is parameterized, which increases the number of parameters to be estimated. Note that there is a link with nonspatial regression models; the response coefficients of a simple linear regression model are also fundamentally unidentified if one does not make the basic assumption that they are the same for all observations in the sample.

[3]In this paper, we use the acronyms used in LeSage and Pace (2009). Figure 1 hereafter gives a full overview. As noted by a referee, although McMillen (2012) uses the SAR model for illustration purposes, his critique is applicable to other parametric model extensions. In general, the assumption of linear functional forms is questioned; in this respect, also see McMillen (2013).

[4]They also show that many studies are based on a well-founded theoretical background such as Fingleton and Lopez-Bazo (2006) and Ertur and Koch (2007) with specifications based on neoclassical growth theory. Another example they provide is the modeling of social networks using the SAR model where, e.g., a student's behavior is directly affected by the behavior of their friends. Another notable example, mentioned in McMillen (2012, 2013), is Brueckner (2006) who adopts the SAR model to empirically assess strategic interaction among local governments (see also Brueckner, 2003). Recently, Buonanno et al. (2012) provide a well-founded theoretical background of **W** as an exponential distance decay function to analyze crime and social sanction in Italy.

The third identification problem occurs when the unknown parameters of a model cannot be uniquely recovered from their reduced-from specification even if the spatial econometric model and **W** are correctly specified. This is the topic of section 3 in Gibbons and Overman's contribution. By now, many papers have shown that the parameters of the main spatial econometric models are formally identified and can be consistently estimated if correctly specified and **W** satisfies certain regularity conditions. Kelejian and Prucha (1998, 1999) consider the IV/GMM estimator of the SAR and the spatial autoregressive combined (SAC) model; Lee (2004) focus on the (quasi-)ML estimator of the SAR model; Bramoullé et al. (2009) study identification of the spatial Durbin model (SDM) in terms of instrumental variables for both the group interaction and arbitrary spatial weights matrices and prove that the matrices **I**, **W**, and $\mathbf{W}^2$ should be linearly independent, and that the parameters in the SDM model should not satisfy the common factor restriction derived by Burridge (1981).[5] They demonstrate that endogenous and exogenous interaction effects are separable, unless one of these two conditions breaks down. Liu and Lee (2013) consider the IV estimator of the SAR model and Drukker et al. (2013) the GMM estimator of the SAC model in the presence of additional endogenous regressors besides spatial interaction effects among the dependent variable. Generally, these studies find that the correlation between two units in the spatial weights matrix should converge to zero as the geographical or economic distance separating them increases to infinity, thereby corroborating the basic idea to place more weight on closer observations. The increased attention to endogenous regressors is laudable since researchers face uncertainty about the endogeneity of **X** in (nearly) all applications (Fingleton and Le Gallo, 2008, p. 320; Gibbons and Overman, 2012, p. 186). In sum, only the parameters of a spatial econometric model with all possible spatial interaction effects based on arbitrary spatial weights matrices have not been proved to be free of this type of identification problem, as we will also demonstrate empirically in this paper.

In conclusion, we can say that the basic identification problem in spatial econometrics is the difficulty to distinguish different models and different specifications of **W** from each other without reference to specific economic theories.

In view of these critical notes, it is clear that the way of thinking and the model selection strategy that are used to determine the structure of spatial processes need revision. Gibbons and Overman (2012) propose two solutions. One, which is their preferred approach, is the use of natural experiments and microeconomic data sets. The second solution is to take the spatial lag of **X** (SLX) model as point of departure (Gibbons and Overman, 2012, p. 183). We support and further work out this second proposal of Gibbons and Overman, but for additional reasons than pointed out in their paper.

Until recently, empirical studies used the coefficient estimates of a spatial econometric model to test the hypothesis as to whether or not spatial spillover effects exist. However, LeSage and Pace (2009) point out that a partial derivative interpretation of the impact from changes to the variables represents a more valid basis for testing this hypothesis. By considering these partial derivatives, we are able to show that some models are more flexible in modeling spatial spillover effects than others, and that the SLX model is the simplest one of those. Importantly, model selection strategies that have been developed in the literature so far generally focus on the SEM, SAR, SAC, and SDM models, whereas the SLX model is left out of the picture, even though this model has been considered in applied research before (see, e.g., Boarnet, 1994a, 1994b, 1998; Holtz-Eakin

[5] $\boldsymbol{\theta} + \rho\boldsymbol{\beta} = \mathbf{0}$ with $\lambda = 0$ in Equation (2).

and Schwartz, 1995; Dalenberg et al., 1998; Fischer et al., 2009).[6] This holds for the robust LM tests developed by Anselin et al. (1996) and for the recent interest in models containing more than just one spatial interaction effect, in particular pertaining to the SAC and SDM models. It also holds for studies comparing the general-to-specific and the specific-to-general approaches (Florax et al., 2003; Mur and Angulo, 2009),[7] the emerging literature on Bayesian posterior model probabilities (LeSage and Pace, 2009, chapter 6; Mur et al., 2013), and the J-test (Kelejian, 2008; Burridge and Fingleton, 2010; Burridge, 2012). In stark contrast to these standard and emerging spatial econometric specification strategies, we follow Gibbons and Overman's (2012) proposal to take the SLX model as point of departure, unless the researcher has an underlying theory or coherent economic argument pointing toward a different model.

Another reason, also not discussed in Gibbons and Overman (2012), is that in contrast to other spatial econometric models, the elements of $\mathbf{W}$ in the SLX model can be parameterized. This allows for greater flexibility in the specification of $\mathbf{W}$, which is an often criticized aspect of spatial econometric modeling (Corrado and Fingleton, 2012; McMillen, 2012). A final reason is that standard instrumental variables (IV) approaches, developed outside the spatial econometrics literature, can be used to investigate whether (part of) the $\mathbf{X}$ variables and their spatially lagged values, $\mathbf{WX}$, are endogenous.

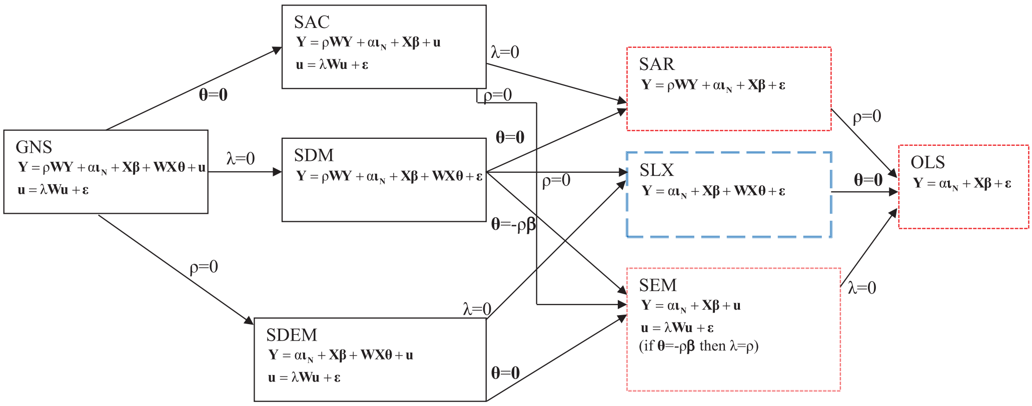
The setup of this paper is as follows. First, we provide a comprehensive overview of the spillover effects that result from linear spatial econometric models with all different combinations of interaction effects. By taking this perspective, we discuss the advantages and limitations of adopting different specifications. Our intention is to provide more guidance for researchers aiming to empirically assess spillover effects. Besides the advantage that the spillover effects using the SLX model are more straightforward, both in terms of estimation and interpretation, they are also more flexible than those from the commonly used SEM, SAR, and SAC models. Furthermore, instead of adopting the traditional binary contiguity matrix, in Section 3 we propose using a $\mathbf{W}$ that is parameterized. To compare both the standard spatial econometric approach and our proposed alternative approach in an empirical setting, the Baltagi and Li (2004) cigarette demand model is estimated. This empirical application demonstrates that the common approach leads to incorrect inferences and that ignoring potential endogeneity of regressors, and thereby, the application of IV estimators, may also lead to incorrect inferences. Finally, we provide conclusions and suggest directions for further research.

2. SPATIAL ECONOMETRIC MODELS AND CORRESPONDING DIRECT AND SPILLOVER EFFECTS

Figure 1 summarizes different spatial econometric models that have been considered in the literature. It extends the figure presented in Elhorst (2010) to include the SLX model. The simplest model considered in Figure 1 is the familiar linear regression model which takes the form

[6] Many of these studies consider explanatory variables measured in surrounding spatial units without explicitly labeling the model the SLX model, which is one of the reasons they are difficult to trace in the literature.

[7] One exception is Florax and Folmer (1992), who are among the first to compare the performance of three selection procedures to choose among SAR, SEM, SDM and, importantly, also the SLX model. They use the commonly applied LM-tests to test for either a spatial autoregressive or a spatial error term specification and propose an F-test to test for spatially lagged independent variables. Unfortunately, their Monte Carlo simulation experiment shows that the probability of finding the true model out of these four models is rather poor.



Note: GNS = general nesting spatial model, SAC = spatial autoregressive combined model, SDM = spatial Durbin model, SDEM = spatial Durbin error model, SAR = spatial autoregressive model, SLX = spatial lag of **X** model, SEM = spatial error model, OLS = ordinary least squares model.

FIGURE 1: Comparison of Different Spatial Econometric Model Specifications.

$$\mathbf{Y} = \alpha\boldsymbol{\iota}_N + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

where **Y** represents an $N \times 1$ vector consisting of one observation on the dependent variable for every unit in the sample ($i = 1, \ldots, N$), $\boldsymbol{\iota}_N$ is an $N \times 1$ vector of ones associated with the constant term parameter α, **X** denotes an $N \times K$ matrix of explanatory variables associated with the $K \times 1$ parameter vector $\boldsymbol{\beta}$, and $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_N)^T$ is a vector of independently and identically distributed disturbance terms with zero mean and variance σ^2.[8] Since model (1) is commonly estimated by ordinary least squares (OLS), it is often referred to as the OLS model.

Starting with the OLS model, the spatial econometrics literature has developed models that treat three different types of interaction effects among units: (i) endogenous interaction effects among the dependent variable, (ii) exogenous interaction effects among the explanatory variables, and (iii) interaction effects among the error terms.

Unfortunately, there is a gap in the level of interest in these interaction effects between econometric theoreticians and practitioners. Theoreticians are mainly interested in models containing endogenous interaction effects, interaction effects among the error terms or endogenous interaction effects in combination with either exogenous interaction effects or interaction effects among the error terms (i.e., the SAR, SEM, SDM, and SAC models, respectively), because of the econometric problems and often complicated regularity conditions accompanying the estimation of these models. The reason they do not focus on the spatial econometric model with only exogenous interaction effects is because the estimation of this model does not cause severe additional econometric problems, provided that the explanatory variables **X** are exogenous and the spatial weights matrix **W** is known and exogenous. Under these circumstances standard estimation techniques suffice.[9] Consequently, the SLX model is not part of the toolbox of researchers interested in the econometric theory behind spatial econometric models. We emphasize, however, that since the **X** variables are often not exogenous and the **W** matrix is generally

[8]The superscript T indicates the transpose of a vector or matrix.

[9]By replacing the argument **X** by **X** = [**X WX**] of routines that have been developed to estimate SAR, SEM, and SAC models, one can also estimate the SDM, SDEM, and GNS models.

unknown, econometric challenges remain.[10] The extension of Figure 1 with the SLX model is intended as a call for both theoretical and applied spatial econometric work to pay more attention to this model and these two issues.

The model in Figure 1 that includes all possible interaction effects takes the form:

$$\mathbf{Y} = \rho\mathbf{WY} + \alpha\boldsymbol{\iota}_N + \boldsymbol{X}\boldsymbol{\beta} + \mathbf{WX}\boldsymbol{\theta} + \mathbf{u}, \quad \mathbf{u} = \lambda\mathbf{Wu} + \boldsymbol{\varepsilon}. \tag{2}$$

We will refer to model (2) as the general nesting spatial (GNS) model since it includes all types of interaction effects.[11] The spatial weights matrix **W** is a positive $N \times N$ matrix that describes the structure of dependence between units in the sample. The variable **WY** denotes the endogenous interaction effects among the dependent variables, **WX** the exogenous interaction effects among the explanatory variables, and **Wu** the interaction effects among the disturbance terms of the different observations. The scalar parameters ρ and λ measure the strength of dependence between units, while $\boldsymbol{\theta}$, like $\boldsymbol{\beta}$, is a $K \times 1$ vector of response parameters. The other variables and parameters are defined as in model (1).

Since the GNS model incorporates all interaction effects, models that contain less interaction effects can be obtained by imposing restrictions on one or more of the parameters (shown next to the arrows in Figure 1). Both frequently used, but also largely neglected models are included. In particular, the SLX model is generally overlooked in the spatial econometrics literature.

Various methods can be applied to estimate spatial econometric models such as maximum likelihood (ML), instrumental variables or generalized method of moments (IV/GMM), and Bayesian methods. There is a large literature on how the coefficients of each of the interaction effects can be estimated.[12] Considerably less attention has been paid to the interpretation of these coefficients. Many empirical studies use the point estimates of the interaction effects to test the hypothesis as to whether or not spillovers exist. Only recently, thanks to the work of LeSage and Pace (2009), researchers started to realize that this may lead to erroneous conclusions, and that a partial derivative interpretation of the impact from changes to the variables of different model specifications represents a more valid basis for testing this hypothesis.

Direct and Spillover Effects

The direct and spillover effects corresponding to the different model specifications are reported in Table 1. By construction, the OLS model does not allow for spillovers since it makes the implicit assumption that outcomes for different units are independent of each other, which is restrictive especially when dealing with spatial data. Even though the SEM takes into account spatial dependence in the disturbance process, it also provides no information about spillovers, as shown in Table 1.[13] This is clearly a major limitation of the SEM if measuring the effects of spillovers is of great interest. The direct effect, i.e., the effect of a change of a particular explanatory variable in a particular unit on the dependent variable of the same unit, is the only information provided. Therefore, if

[10] The authors thank the referees for their comments on these issues and come back to them in Sections 3 and 4.

[11] LeSage and Pace (2009, p. 53) neither name nor assign an equation number to model (2), which reflects the fact that this model is typically not used in applied research.

[12] For example, LeSage and Pace (2009) provide details on the ML and Bayesian methods and Kelejian and Prucha (1998, 1999, 2010) and Kelejian et al. (2004) on IV/GMM estimators.

[13] Lacombe and Lesage (2013) also reserve the term “spillover effects” to refer to spillovers resulting from an observable explanatory variable. For the SEM model, the term “diffusion of shocks” is used.

TABLE 1: Direct and Spillover Effects Corresponding to Different Model Specifications

Model	Direct Effect	Spillover Effect
OLS/SEM	β_k	0
SAR/SAC	Diagonal elements of $(\mathbf{I}\text{-}\rho\mathbf{W})^{-1}\beta_k$	Off-diagonal elements of $(\mathbf{I}\text{-}\rho\mathbf{W})^{-1}\beta_k$
SLX/SDEM	β_k	θ_k
SDM/GNS	Diagonal elements of $(\mathbf{I}\text{-}\rho\mathbf{W})^{-1}[\beta_k+\mathbf{W}\theta_k]$	Off-diagonal elements of $(\mathbf{I}\text{-}\rho\mathbf{W})^{-1}[\beta_k+\mathbf{W}\theta_k]$

applied researchers want to obtain inference on spillovers, alternative spatial econometric models need to be considered.[14]

One such model that allows an empirical assessment of the magnitude and significance of spillover effects is the SAR model. If the SAR model (3) is rewritten to its reduced form (4), the direct and spillover effects can be obtained:

$$\mathbf{Y} = \rho\mathbf{WY} + \alpha\boldsymbol{\iota}_N + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{3}$$

$$\mathbf{Y} = (\mathbf{I} - \rho\mathbf{W})^{-1}\alpha\boldsymbol{\iota}_N + (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\varepsilon}. \tag{4}$$

The matrix of partial derivatives of the expectation of $\mathbf{Y}$, $E(\mathbf{Y})$, with respect to the kth explanatory variable of $\mathbf{X}$ in unit 1 up to unit N is

$$\left[\frac{\partial \mathrm{E}(\mathbf{Y})}{\partial x_{1k}} \cdots \frac{\partial \mathrm{E}(\mathbf{Y})}{\partial x_{Nk}}\right] = (\mathbf{I} - \boldsymbol{\rho}\mathbf{W})^{-1}\beta_k, \tag{5}$$

which is reported in Table 1. The diagonal elements of (5) represent direct effects, while the off-diagonal elements contain the spillover effects. To better understand the direct and spillover effects that follow from this model, the infinite series expansion of the spatial multiplier matrix is considered:

$$(\mathbf{I} - \rho\mathbf{W})^{-1} = \mathbf{I} + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \rho^3\mathbf{W}^3 + \cdots. \tag{6}$$

Since the off-diagonal elements of the first matrix term on the right-hand side (the identity matrix $\mathbf{I}$) are zero, this term represents a direct effect of a change in $\mathbf{X}$. Conversely, since the diagonal elements of the second matrix term on the right-hand side ($\rho\mathbf{W}$) are zero by assumption, this term represents an indirect effect of a change in $\mathbf{X}$. All other terms on the right-hand side represent second- and higher-order direct and spillover effects. Since both the direct and spillover effects vary for different units in the sample, the presentation of both effects can be challenging. With N units and K explanatory variables, it is possible to obtain K different NxN matrices of direct and spillover effects. Even if N and K are small, it may be difficult to compactly report the results. LeSage and Pace (2009) therefore propose to report one direct effect measured by the average of the diagonal elements and one spillover effect measured by the average row sums of the off-diagonal elements. From Table 1 it can also be seen that the SAC model shares the same direct and spillover effect properties as the SAR model.

An important characteristic of the spillovers produced by the SAR and SAC models is that they are *global* in nature. Anselin (2003b) describes the difference. A change in $\mathbf{X}$ at any location will be transmitted to all other locations following the matrix inverse in Equation (6), also if two locations according to $\mathbf{W}$ are unconnected. In contrast, *local*

[14]The SEM model might still be relevant when empirical evidence in favor of spillover effects cannot be found.

spillovers are those that occur at other locations without involving an inverse matrix, i.e., only those locations that according to **W** are connected to each other. According to LeSage and Pace (2011) another distinction between the two is that global spillovers include feedback effects that arise as a result of impacts passing through neighboring units (e.g., from region i to j to k) and back to the unit that the change originated from (region i), whereas local spillovers do not.

Despite its popularity, the SAR model has many serious limitations. Elhorst (2010) demonstrates that the ratio between the spillover effect and direct effect of an explanatory variable is independent of β_k. The implication is that the ratio between the spillover and direct effects is the same for every explanatory variable, which is unlikely to be the case in many empirical studies. Pace and Zhu (2012) point out that the parameter ρ affects both the estimation of spillovers and the estimation of spatial disturbances. This implies that if the degree of spatial dependence in the error terms is different from that in the spillovers, then it can be the case that both are estimated incorrectly. In this respect, Pinkse and Slade (2010, p. 106) criticize the SAR model for the laughable notion that the entire spatial dependence structure is reduced to one single unknown coefficient. Corrado and Fingleton (2012) point out and demonstrate by using a simple Monte Carlo simulation experiment that the coefficient estimate for the **WY** variable may be significant because it could be picking up the effects of omitted **WX** variables or nonlinearities in the **WX** variables if they are erroneously specified as being linear. This makes the interpretation of a causal (spillover) effect difficult and what we considered to be the basic identification problem in the introduction. That is, the issue is if it is possible for a researcher to discern whether the significant coefficient of the **WY** variable is due to omitted variables or due to a causal effect of **WY**. An issue that has recently gained more attention in this respect is that global spillovers are often more difficult to justify (see, e.g., Arbia and Fingleton, 2008; Corrado and Fingleton, 2012; Gibbons and Overman, 2012; Partridge et al., 2012; Lacombe and LeSage, 2013). According to Pinkse and Slade (2010), this is also a primary criticism of standard spatial econometrics; researchers try to fit their preferred model (usually a SAR model) onto every empirical problem rather than having the nature of the empirical problem inform which particular model best answers the question. We come back to this issue in the empirical application.

In contrast to the models above, the SLX model contains spatially lagged explanatory variables, taking the following form:

$$\mathbf{Y} = \alpha \boldsymbol{\iota}_N + \mathbf{X}\boldsymbol{\beta} + \mathbf{W}\mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon}. \tag{7}$$

The direct and spillover effects do not require further calculation compared to other models such as the SAR model. As reported in Table 1, the direct effects are the coefficient estimates of the nonspatial variables (β_k) and the spillover effects are those associated with the spatially lagged explanatory variables (θ_k).[15] Therefore, a strong aspect is that there are no prior restrictions imposed on the ratio between the direct effects and spillover effects, which was a limitation of the SAR and SAC models. Like the SLX model, the direct and spillover effects of the spatial Durbin error model (SDEM) are the vectors of the response parameters $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, respectively. Even though these models are more straightforward in terms of estimation and interpretation and, most importantly, are useful for investigating local spillovers, they are not as commonly applied as global spillover specifications.[16]

[15] See also LeSage and Pace (2011, p. 22).

[16] The urban economics literature is an exception; if interaction effects are incorporated, exogenous interaction effects are usually preferred.

The SDM model, which has recently become more widely used in applied research, includes both endogenous and exogenous interaction effects (LeSage and Pace, 2009; Elhorst, 2010). To obtain the direct and spillover effects shown in Table 1, the SDM (8) can be expressed in its reduced form (9):

$$\mathbf{Y} = \rho\mathbf{WY} + \alpha\boldsymbol{\iota}_N + \mathbf{X}\boldsymbol{\beta} + \mathbf{WX}\boldsymbol{\theta} + \boldsymbol{\varepsilon}, \tag{8}$$

$$\mathbf{Y} = (\mathbf{I} - \rho\mathbf{W})^{-1}\alpha\boldsymbol{\iota}_N + (\mathbf{I} - \rho\mathbf{W})^{-1}(\mathbf{X}\boldsymbol{\beta} + \mathbf{WX}\boldsymbol{\theta}) + (\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\varepsilon}. \tag{9}$$

From Equation (9), the matrix of partial derivatives of $E(\mathbf{Y})$ with respect to the kth explanatory variable of $\mathbf{X}$ in unit 1 up to unit N is obtained

$$\left[\frac{\partial \mathrm{E}(\mathbf{Y})}{\partial x_{1k}} \cdots \frac{\partial \mathrm{E}(\mathbf{Y})}{\partial x_{Nk}}\right] = (\mathbf{I} - \rho\mathbf{W})^{-1}\left[\mathbf{I}\beta_k + \mathbf{W}\theta_k\right]. \tag{10}$$

As reported in Table 1, the diagonal elements of the matrix represent the direct effects and the off-diagonal elements, the spillover effects. Just as for the SLX and the SDEM models, there are no prior restrictions imposed on the ratio between the direct and spillover effects. Table 1 shows that this is due to the fact that both the direct effect and the spillover effect of an explanatory variable depends not only on the parameters ρ and β_k, and the matrix $\mathbf{W}$, but also on the coefficient estimate θ_k (Elhorst, 2010). LeSage and Pace (2009, p. 28) and Lacombe and LeSage (2013) provide an econometric-theoretical motivation in favor of the SDM model if (i) the true model is SEM, (ii) there is at least one potentially important variable omitted from the model, and (iii) this variable is likely to be correlated with the independent variables included in the model. It is shown that only the SDM model produces unbiased, though inefficient, parameter estimates under these circumstances. They also show that one can determine the strength of correlation between included and excluded independent variables to test whether assumptions (ii) and (iii) are true. Unfortunately, this test statistic is not valid when assumption (i) is violated, i.e., if for example SDEM is the true model. Furthermore, although Gibbons and Overman (2012, appendix) confirm that this setup leads to the SDM model, they also emphasize that this does not solve the problem of whether the causal effect of the observed spatial patterns in the data is due to endogenous interaction effects or interaction effects among the error terms.

Finally, as can be noted in Table 1, the GNS model shares the same spillover properties as the SDM. Even though taking the GNS model as point of departure to measure spillovers seems appealing since it contains all possible interaction effects, two major issues are that a formal proof under which conditions the parameters of this model are identified is not available yet (see Section 1) and the problem of overfitting. Even though the parameters are not identified, they can still be estimated. However, they 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 the SDM and SDEM models. We come back to this issue in the empirical application.

Our overview of spatial econometric models with all conceivable combinations of different types of interaction effects makes clear that four models are able to produce spillover effects that in relation to their corresponding direct effects may be different from one explanatory variable to another. It concerns the SLX, SDEM, SDM, and GNS models. The other models, although interesting from an econometric-theoretical viewpoint, are less flexible since they impose restrictions on the magnitude of spillover effects in advance. Since Figure 1 shows that the SLX model is the simplest of these four more flexible models, it is recommendable to take this model as point of departure when having any empirical evidence that the observations in the sample are spatially dependent.[17]

[17]As mentioned before, an exception is if the researcher has a theoretical framework or coherent economic argument for another type of model.

3. THE SLX MODEL AND PARAMETERIZING W

Studies dealing with geographical units often adopt a binary contiguity matrix with elements $w_{ij} = 1$ if two units share a common border and zero otherwise, an inverse distance matrix, or an inverse distance matrix with a cut-off threshold distance of for example, m miles.[18] If in a particular study theory predicts that the connectivity between nearby units will be stronger than those further away, this is related to the well-known first law of geography: "Everything is related to everything else, but near things are more related than distant things" (Tobler, 1970, p. 236). However, this is often used too readily to justify spatial econometric specifications, including **W** (Partridge et al., 2012; Neumayer and Plümper, 2013). Even if there are theoretical reasons indicating that spatial interaction effects are related to distance, it is often not clear from the theory the degree at which the spatial dependence between units diminishes as distance increases. The common practice to adopt one of the spatial weight matrices mentioned above (or variants of these) can be quite arbitrary, also if alternative specifications are considered to check whether the results are robust.

For this reason, the literature on the specification of **W** is extensive and can be divided into five main approaches. Researchers may adopt an: (i) exogenous[19] pre-specified **W** with fixed weights, (ii) exogenous pre-specified but parameterized **W** (where one or more parameters are estimated), (iii) endogenous pre-specified **W**, (iv) exogenous unknown **W** that is estimated, or (v) endogenous unknown **W** that is estimated. Furthermore, a distinction can be made between using parametric methods discussed in Corrado and Fingleton (2012), and semiparametric and nonparametric methods discussed in McMillen (2010, 2012, 2013). These latter methods have been used to estimate **W**, and as an alternative to using **W**. An advantage of semiparametric and nonparametric methods is that a specific functional form is not assumed. A disadvantage is that spatial dependence is often taken to be a nuisance phenomenon and thus does not account for spatial spillovers (Corrado and Fingleton, 2012, section 6).[20]

The most widely used approach is (i) within a linear regression framework, followed by approach (ii) which has the advantage that it can provide more flexibility if it is not clear how sensitive interaction is to distance.[21] Approaches (iii) and (iv) are receiving more attention, while (v) is yet to be explored to the best of our knowledge (cf. Harris et al., 2011;

[18]Alternative ways to specify **W** can be found in Corrado and Fingleton (2012), such as using economic variables. One potential problem of specifying **W** based on economic variables is endogeneity because exogeneity of **W** is one of the regularity conditions for formal identification.

[19]Typically, weights matrices are treated as being exogenous, although there are cases where this assumption may be inappropriate (see, e.g., Kelejian and Piras, 2014).

[20]This was also discussed in the first two sections. It is stressed that in general, and in a marked departure from standard spatial modeling, McMillen advocates the use of more flexible semiparametric and nonparametric approaches to **W**. However, if the objective is to directly test for causal (spillover) effects based on a rigorous theoretical foundation, it is reasonable to impose structure to identify the model (McMillen, 2012, section 5), as in the form of **W**.

[21]See among others, Burridge and Gordon, 1981; Boarnet, 1992, 1994a, 1994b; Song, 1996; Pace et al., 1998; Boarnet et al., 2005; Fischer et al., 2009. Most studies use grid search to scan over a range of values of the nonlinear parameter(s), among which is Song (1996) who also considers alternative distance decay functions. In contrast to the other studies, Fischer et al. (2009) estimate the decay parameter simultaneously along with the other parameters in their SLX specification using a direct search procedure. There is also previous research focusing on spatial autocorrelation (Cliff and Ord, 1981, ch. 5; Dubin, 1988, 1992), where it is shown that instead of **W**, it is also possible to directly parameterize the variance-covariance matrix; Anselin (2003a) shows that this approach is problematic since the null hypothesis of no spatial autocorrelation does not correspond to an interior point of the parameter space and hence does not satisfy the regularity conditions for estimation.

Kelejian and Piras, 2014; Qu and Lee, 2015). The latter study distinguishes two general ways to estimate a spatial weights matrix: letting the data determine **W** using geostatistical modeling techniques (e.g., Getis and Aldstadt, 2004) or regressing the residuals for each unit on the residuals of all other units, a technique originally proposed by Meen (1996) that is viable in long panels (Beenstock and Felsenstein, 2012).[22] Unfortunately, it is commonly the case that researchers are faced with situations where $N > T$ rather than $T > N$, making this method unfeasible due to the large amount of parameters that need to be estimated. Other potential limitations of estimating **W** are discussed in Harris et al. (2011, p. 254), Corrado and Fingleton (2012, p. 14), and Neumayer and Plümper (2013, p. 5), with the last two particularly stressing the atheoretical nature of data-driven approaches.

Although interaction between units can be independent from geographical proximity (see Neumayer and Plümper, 2013 for examples), for those cases in which researchers can place more weight on closer observations, we demonstrate the flexibility that occurs if they take one step forward by using a simple parametric approach applied to the elements of an inverse distance matrix

$$w_{ij} = \frac{1}{d_{ij}^{\gamma}}, \tag{11}$$

where d_{ij} denotes the distance between observations i and j, and γ is the distance decay parameter to be estimated. We hasten to point out that this is one possible approach and that, in this respect, a literature overview and a comparison of the pros and cons of the alternative approaches is an important topic for further research.[23]

A nonlinear but straightforward estimation technique[24] can be used to estimate the parameter γ, providing more information on the nature of the interdependencies of the observations in the sample. For example, if the estimate of γ is small this is an indication that the commonly applied binary contiguity principle is not an accurate representation of the spatial dependence. This is because contiguity can be thought of as a restrictive distance measure where interaction between units is confined only to those units that share borders. This is visually depicted in Figure 2 where the vertical line (BC) shows how a binary contiguity specification would cut-off interaction between units.

One unexplored advantage of parameterizing **W** in the SLX model is that unlike the other spatial econometric models in Figure 1, it is not hampered by the perfect solution problem. The solution $\rho = -1$, $\gamma = 0$, $\alpha = Y_1 + \cdots + Y_N$, and $\boldsymbol{\beta} = \mathbf{0}$ in Equations (3) and (11) would perfectly fit the dependent variable in the SAR model, as well as the SEM and SDM models (together with $\theta = 0$).[25] Formally, this perfect solution has been excluded. To prove consistency of the ML estimator of the SAR model, Lee (2004) shows that one of the following two regularity conditions should be satisfied: (a) the row and column sums of the matrices **W** and $(\mathbf{I}-\rho\mathbf{W})^{-1}$ before **W** is row-normalized should be uniformly bounded

[22]Recent studies building upon this approach in a spatial panel data framework using nonparametric methods include Beenstock and Felsenstein (2012) and Bhattacharjee and Jensen-Butler (2013).

[23]We also considered and estimated an exponential distance decay function, a specification with a different γ for every explanatory variable, and a gravity type of specification with parameterized power functions of the population size of states i and j, but these are all variants of the basic idea and therefore might be investigated and compared in future research.

[24]If the SLX model in Equation (7) is taken as point of departure and the elements of **W** are specified as in Equation (11), the scalar α and the parameter vectors $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$, given γ, and γ given α, $\boldsymbol{\beta}$ and $\boldsymbol{\theta}$ can be alternately estimated until convergence occurs. Matlab code of this routine is provided at the Web site of the second author.

[25]Note that the SEM model can be rewritten as a constrained SDM model ($\boldsymbol{\theta} = -\rho\boldsymbol{\beta}$). This explains why the perfect solution of the SDM also holds for the SEM.

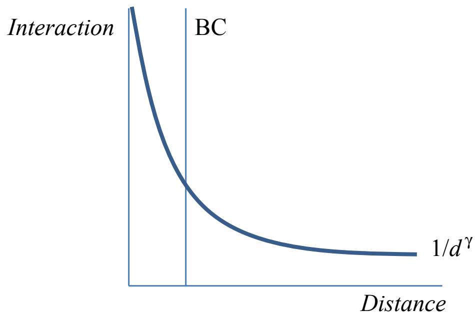


FIGURE 2: Distance Decay.

in absolute value as N goes to infinity, or (b) the row and column sums of $\mathbf{W}$ before $\mathbf{W}$ is row-normalized should not diverge to infinity at a rate equal to or faster than the rate of the sample size N.[26] If the elements of $\mathbf{W}$ take the form in (11) and $\gamma = 0$, we have $w_{ij} = 1$ for all off-diagonal elements of $\mathbf{W}$, as a result of which the row and column sums are N-1, which diverge to infinity as N goes to infinity. Furthermore, we have $(N - 1)/N \rightarrow 1$ as N goes to infinity. Since neither condition (a) nor condition (b) is satisfied, this perfect solution problem has been excluded formally. However, computer software working with real data and fixed sample sizes, $N = \bar{N}$, cannot simply rule out these parameter values for ρ and γ, as a result of which it often converges to the perfect solution computationally. The question whether a local next to this invalid global optimum exists and whether it is possible to program the SAR, SEM or SDM models such that this local optimum can be obtained is an interesting topic for further research. Importantly, when the SLX model specification is taken as point of departure, the perfect solution problem is not relevant, which simplifies the analysis.

4. EMPIRICAL APPLICATON

State Cigarette Demand Model

To compare the standard approach and our proposed alternative approach in an empirical setting, we consider the following cigarette demand model taken from Baltagi and Li (2004) as a candidate to include spatial interaction effects

$$\ln(C_{it}) = \alpha + \beta_1 \ln(P_{it}) + \beta_2 \ln(I_{it}) + \mu_i + \lambda_t + \varepsilon_{it}, \tag{12}$$

where the subscript i denotes states ($i = 1, \ldots, 46$) and the subscript t denotes time periods ($t = 1, \ldots, 30$). C_{it} is real per capita sales of cigarettes, which is measured in packs per person aged 14 years and older. P_{it} is the average retail price of a pack of cigarettes and I_{it} is real per capita disposable income. The equation is obtained from maximizing a utility function depending on cigarettes and other consumer goods subject to a budget constraint.[27] The model is aggregated over individuals since the objective is to explain sales in a particular state, as in Baltagi and Levin (1986, 1992), Baltagi and Li (2004), Debarsy et al. (2012), and Elhorst (2014), among others. If the purpose, on the other hand, is to model individual behavior (e.g., the reduction in the number of smokers or teenage

[26] Condition (a) originates from Kelejian and Prucha (1998, 1999).

[27] For details on demand models derived from an underlying theory of consumer utility maximization see, e.g., Chung (1994) and Chintagunta and Nair (2011).

smoking behavior) then this is better studied using micro data (see, e.g., Lewit and Coate, 1982; Lewit et al., 1987; Wasserman et al., 1991; Soetevent and Kooreman, 2007).[28]

The data consist of a panel of 46 U.S. states over the period 1963–1992.[29] This data set has also been used for illustrative purposes in other spatial econometric studies and is also widely used for pedagogical motives. We therefore find it conducive to use it to illustrate the points raised in this paper. Moreover, since the focus is on spillovers it is particularly relevant that the empirical application encompasses a well-founded motivation for spillover effects, which we discuss shortly. All the variables are log-transformed as in previous studies, and thus the interpretation of the estimates can be in elasticity terms.

The model specification includes state-specific fixed effects, μ_i, and time-specific fixed effects, λ_t. In this way, specific state characteristics (e.g., states with tax exempt military bases) and omitted effects that are common across all states that occurred during the period (e.g., policy changes and health warnings) are controlled for.[30] Using Monte Carlo simulation experiments, Lee and Yu (2010) show that ignoring time-period fixed effects may lead to large upward biases (up to 0.45) in the coefficient of the spatial lag. The explanation is that most variables tend to increase and decrease together in different spatial units over time (e.g., along the business cycle). If this common effect is not taken into account and thus not separated from the interaction effect among units, the latter effect might be overestimated. This is a typical example of one of the shortcomings of the SAR model discussed in Section 2: the spatial lag picking up the effects of omitted variables, in this case of time period fixed effects when they are not controlled for.

The main motivation to extend the basic model to include spatial interaction effects is the so-called bootlegging effect; consumers are expected to purchase cigarettes in nearby states if there is a price advantage. This smuggling behavior is a result of significant price variation in cigarettes across U.S. states and partly due to the disparities in state cigarette tax rates.[31] Baltagi and Levin (1986, 1992) incorporate the minimum real price of cigarettes in any neighboring state as a proxy for the bootlegging effect. A limitation is that this proxy does not account for cross-border shopping that may take place between other states than the minimum-price neighboring state (Baltagi and Levin, 1986).[32] This can be due to smuggling taking place over longer distances by trucks since cigarettes can be stored and are easy to transport (Baltagi and Levin, 1992) or due to geographically large states where cross-border shopping may occur in different neighboring states. To

[28]Blundell and Stoker (2007) provide a review and propositions to bridge the gap between micro and macro level research and point out that both approaches have a role to play. We use state-level data mainly due to our illustration purposes. In addition, even though cigarette consumption is observed at the individual level, cigarette price is not, which can result in severely downward biased estimates of the standard errors (Moulton, 1990). Another issue is zero observations requiring the use of, e.g., a Tobit model, which raises additional questions that would need to be addressed such as whether the zero values are a result of nonconsumption or corner solutions (Yen and Huang, 1996).

[29]The data can be accessed at the Baltagi (2008) companion Web site: www.wileyeurope.com/college/baltagi. An adapted version is available at: www.regroningen.nl/elhorst.

[30]For more details on reasons to include state and time specific effects, refer to Baltagi (2008). Elhorst (2014) found that the model specification with spatial and time-period fixed effects outperforms its counterparts without spatial and/or time-period fixed effects, as well as the random effects model.

[31]For example, in Massachusetts the cigarette tax rate is around double that of New Hampshire. Baltagi (2008) and references therein provide more details.

[32]Note that even if individual level data are used, capturing the bootlegging effect also requires an *a priori* specification. For example, Lewit and Coate (1982) and Wasserman et al. (1991) excluded individuals who lived in communities where the price of cigarettes exceeded another price found within a 20-mile wide band around their place of residence from the analysis. Wasserman et al. (1991, p. 46, footnote 4) admit that the choice of an appropriate band is somewhat arbitrary.

take this into account, other studies have extended the model to explicitly incorporate spatial interaction effects. However, while the specification originally adopted by Baltagi and Levin (1992) resembles the SLX model but then with only one exogenous interaction effect (price), applied spatial econometric studies have either included: (i) endogenous interaction effects, (ii) interaction effects among the error terms or (iii) a combination of endogenous and exogenous interaction effects. It reflects the fact that the SLX model is overlooked in the applied spatial econometrics literature.

In this case including endogenous interaction effects implies that state cigarette sales directly affect one another, which is difficult to justify.[33] The resulting global spillovers (discussed in Section 2) 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** are unconnected.[34] Pinkse and Slade (2010, p. 115) argue that an empirical problem like this is insightful precisely because it is difficult to form a reasonable argument to include endogenous interaction effects even though they are easily found statistically. Given the research question of whether consumers purchase cigarettes in nearby states if there is a price advantage, this example points toward a local spillover specification such as the SLX model rather than a global spillover specification.

Standard Approach

In this section, we first take a more standard spatial econometric approach and estimate the models discussed in Section 2, especially focusing on the spillover results and identification issues. The basic model (12) is therefore extended to include different (combinations of) interaction effects following Figure 1. Even though models with a spatially lagged dependent variable (**WY**) have just been argued to represent a misspecification, they are included to further feed the discussion on their limitations. It is also shown that there is relatively little guidance into which specification is best, confirming the critique by Gibbons and Overman (2012), McMillen (2012), and Corrado and Fingleton (2012) that identification is a crucial issue and that a statistical approach driven by data-analytic considerations may lead to erroneous conclusions. Just as in many other studies, the spatial weights matrix **W** is initially specified as a row-normalized binary contiguity matrix, with elements $w_{ij} = 1$ if two states share a common border, and zero otherwise. This specification of **W** was also used by Baltagi and Li (2004), Elhorst (2013, 2014), and Debarsy et al. (2012).[35] In subsection "The SLX Approach" we demonstrate our proposed alternative approach using the SLX model with a parameterized distance based **W**, and in subsection "Endogenous Regressors" we apply instrumental variables techniques to test for potential endogeneity of the price of cigarettes in the own and in neighboring states and adjust the results accordingly.

Table 2 reports the estimation results explaining cigarette demand for the different spatial econometric models, as well as the OLS model. The spatial models are estimated by ML, with the exception of the SLX model which is estimated by nonlinear OLS (see footnote 24). The coefficient estimates of the two explanatory variables, price and income,

[33]The geographical scale is large; see Partridge et al. (2012, p. 169) for a similar argument pertaining to neighboring county poverty rates. In contrast, if teen smoking behavior is being analyzed then it would be sound to argue that an individual's propensity to smoke is directly influenced by the smoking behavior of their friends.

[34]This implies that, e.g., price changes in California would exert an impact on cigarette consumption even in states as distant as Illinois or Wisconsin.

[35]The latter study also specifies a row-normalized **W** based on state border miles in common between states and find that the results are similar.

TABLE 2: Model Comparison of the Estimation Results Explaining Cigarette Demand, **W** = Pre-Specified Binary Contiguity Matrix

	OLS	SAR	SEM	SLX	SAC	SDM	SDEM	GNS	GNS2
ln(**P**)	−1.035	−0.993	−1.005	−1.017	−1.004	−1.003	−1.011	−1.020	−1.017
	(−25.63)	(−24.48)	(−24.68)	(−24.77)	(−24.49)	(−24.60)	(−24.88)	(−25.40)	
ln(**I**)	0.529	0.461	0.554	0.608	0.557	0.601	0.588	0.574	0.575
	(11.67)	(9.86)	(11.07)	(10.38)	(10.51)	(10.33)	(10.57)	(11.02)	
W × ln(**C**)		0.195			−0.013	0.225		−0.481	−0.400
		(6.79)			(−0.22)	(6.85)		(−7.01)	
W × ln(**P**)				−0.220		0.051	−0.177	−0.645	−0.555
				(−2.95)		(0.62)	(−2.24)	(−5.97)	
W × ln(**I**)				−0.219		−0.293	−0.168	0.079	0.053
				(−2.80)		(−3.70)	(−2.12)	(0.85)	
W × **u**			0.238		0.292		0.229	0.628	0.550
			(7.26)		(4.73)		(6.95)	(14.60)	
R^2	0.896	0.900	0.895	0.897	0.895	0.901	0.897	0.873	
Log-likelihood	1,661.7	1,683.5	1,687.2	1,668.4	1,687.2	1,691.4	1,691.2	1,695.1	

Note: t-values are reported in parentheses; state and time-period fixed effects are included in every model.

TABLE 3: Model Comparison of the Estimated Direct and Spillover Effects on Cigarette Demand, **W** = Pre-Specified Binary Contiguity Matrix

	OLS	SAR	SEM	SLX	SAC	SDM	SDEM	GNS
Direct effects								
ln(**P**)	−1.035	−1.003	−1.005	−1.017	−1.004	−1.016	−1.011	−0.999
	(25.63)	(−25.10)	(−24.68)	(−24.77)	(−24.47)	(−24.84)	(−24.88)	(−25.43)
ln(**I**)	0.529	0.465	0.554	0.608	0.556	0.594	0.588	0.594
	(11.67)	(10.18)	(11.07)	(10.38)	(10.56)	(10.88)	(10.57)	(10.35)
Spillover effects								
ln(**P**)		−0.232		−0.220	0.010	−0.215	−0.177	−0.122
		(−5.63)		(−2.95)	(0.17)	(−2.39)	(−2.24)	(−1.89)
ln(**I**)		0.107		−0.219	−0.006	−0.200	−0.168	−0.155
		(5.51)		(−2.80)	(−0.20)	(−2.30)	(−2.12)	(−2.16)

Note: See notes to Table 2.

are statistically significant at the 1 percent level for all eight model specifications. The magnitudes and signs are as expected. There is a negative price effect and a positive income effect on cigarette sales, which is consistent with economic theory and previous studies.

Since the coefficient estimates of the global specifications cannot be compared with each other and with those of the local specifications, we immediately turn to the direct and spillover effects reported in Table 3 derived from these coefficient estimates.[36] If the spatial econometric model contains endogenous interaction effects (**WY**), the direct effect estimates include feedback effects that arise as a result of impacts passing through neighboring states and back to the state where the change instigated. This is the reason that there are differences between the direct effects (Table 3) and point estimates of the

[36]To draw inferences regarding the statistical significance of the effects estimates, LeSage and Pace (2009, p. 39) suggest simulating the distribution of the direct and indirect effects using the variance-covariance matrix implied by the maximum likelihood estimates. We use the variation of 1,000 simulated parameter combinations drawn from the multivariate normal distribution implied by the ML estimates.

explanatory variables (Table 2) for the SAR, SAC, SDM and GNS models, but not for the OLS, SEM, SLX, and SDEM models (see also Table 1). On the other hand, the feedback effects in the first group of models appear to be relatively small. For example, the greatest difference between the direct effect and the point estimate for the income variable is found in the GNS model, 0.588–0.574 = 0.014, amounting to a feedback effect of only 2.4 percent. Overall, the impact of a change in income or a change in price in a particular state on cigarette demand in that state has almost a similar estimate and inference regardless of which measure (coefficient estimate or direct effect) and model is used.

In contrast, the discrepancies between the spillover impacts are substantial (Table 3). The results show that the choice of model specification leads to completely different conclusions. In particular, both the price and income spillover effects in the SAC model are almost zero and statistically insignificant. This is due to the point estimate of the spatially lagged dependent variable (**WY**) being close to zero and reflects the fact that this model is similar to the SEM, which by construction does not allow for spillover effects. Another noticeable difference is that the spillover effect of income corresponding to the SAR model is positive, whereas in the other models it is negative. Thus, an increase in per capita income in a particular state leads to increased cigarette sales in neighboring states according the SAR model, whereas a negative and significant spillover effect is found in the SLX, SDEM, SDM, and GNS models. The spillover effect of the price variable is negative for all models, which implies that an increase in the price of a pack of cigarettes in a state will not only lead to reduced cigarette demand in the state itself, but also in nearby states.[37] This result is clearly not in line with the bootlegging effect originally found by Baltagi and Levin (1986, 1992) and which was the main motivation to adopt a spatial econometric model.

The conclusions from these findings are as follows. The flexible models (SLX, SDEM, SDM, and GNS) produce price and income spillover effects that are mutually comparable in terms of sign, magnitude and significance levels (although with different mathematical formulas) but significantly different from those produced by the nonflexible models (SAR, SEM and SAC). Yet, these flexible models do not provide empirical evidence in favor of the bootlegging effect. The explanation for this counterintuitive finding can be traced back to the identification issues discussed in the introduction. The first identification problem is that the true **W** is unknown. Although we used information about the spatial arrangement of the states in the sample to place more weight on closer observations, the pre-specified binary contiguity matrix has been taken for granted rather than tested. One notable objection to the binary contiguity matrix is that it limits cross-border shopping to only adjacent states, while in reality people may also benefit from lower prices if they visit distant states for purposes other than just buying cigarettes, as well as when smuggling takes place over longer distances by trucks (Baltagi and Levin, 1992).

The second identification problem is that different spatial econometric models are difficult to distinguish. The most significant identification problem of our empirical example pertains to the type of interaction effects causing the spillover effects. Even though the flexible models produce price and income spillover effects that are comparable in terms of sign, magnitude and significance levels, the mechanism through which observations at other locations is affected is very different across these models. In the SDM and GNS model the spillovers work through both endogenous and exogenous interaction effects, whereas in the SLX and SDEM models they work through exogenous interaction effects.

[37]The exceptions are the SAC model which has a positive estimate, although it is almost zero and insignificant as was mentioned previously, and the SEM and OLS model which do not allow for the quantification of spillover impacts.

The third identification problem is that the unknown parameters of the GNS model are not formally identified even if **W** is correctly specified. To examine this, we performed an empirical Monte Carlo experiment, which is termed empirical since it is design dependent (Huber et al., 2013). We generated the dependent variable of each of the spatial econometric models 1,000 times based on the independent variables and their coefficient estimates reported in Table 2, taking a random draw from the normal distribution based on the parameter estimate of σ^2. On average, these results should be similar to those of the original coefficient estimates if the model is identified. This Monte Carlo experiment demonstrated that the biases in the parameter estimates are negligible, except for the GNS model. The last column in Table 2 (labeled "GNS2") reports the simulated parameter values averaged over 1,000 replications for the GNS model.[38] The largest biases are found in the coefficient estimates of the spatially lagged price variable ($\mathbf{W}\times\ln(\mathbf{P})$), the spatially lagged dependent variable ($\mathbf{W}\times\ln(\mathbf{C})$), and the spatially autocorrelated error term ($\mathbf{W}\times\mathbf{u}$).[39] The reason for these biases is that in 7 percent of the replications the parameter estimates of the spatially lagged dependent variable (ρ) and the spatially autocorrelated error term (λ) appeared to be interchanged.

Another problem is that the coefficient estimates of the spatial interaction effects in the GNS model have a tendency to blow each other up (in absolute value). The spatial autoregressive coefficient equals –0.481 and the spatial autocorrelation coefficient 0.628. Although the net effect of these two coefficients, 0.147, is close to the spatial autoregressive coefficient of 0.225 in the SDM model and the spatial autocorrelation coefficient of 0.229 in the SDEM model, these two individual coefficients are difficult to interpret, both in terms of sign and magnitude.[40] For this reason, the GNS model also does not help us to choose between the next-most general models, either the SDM if $\lambda = 0$ or the SDEM if $\rho = 0$, since both parameters in this model appear to be significant. The conclusion from both findings is that the GNS model is overfitted and that Occam's razor applies. Some researchers prefer simpler models to more complex ones due to overfitting; excessively complex models are affected by statistical noise, whereas simpler models may capture the underlying process better and thus have better predictive performance.

To overcome these identification problems, we follow our proposed alternative strategy taking the SLX model as point of departure and using a parameterized distance based weights matrix instead of one based on contingency to estimate the extent at which interaction dampens as the distance between units increase. In addition, we test for endogenous regressors.

The SLX Approach

Table 4 reports the estimation results explaining cigarette demand for the SLX model; column (1) repeats the results using the row-normalized binary contiguity matrix and column (2) shows the results using the parameterized inverse distance matrix. Row-normalizing a weights matrix based on inverse distance causes its economic interpretation in terms of distance decay to no longer be valid (Anselin, 1988, pp. 23–24; Kelejian and

[38]The value of σ^2 for the GNS model is 0.00445. Results for the other models are available upon request from the authors.

[39]The original estimate for $\mathbf{W}\times\ln(\mathbf{P})$ is –0.645, while the simulated coefficient estimate is –0.555, representing a bias of 0.090 or around –13.95 percent of the original parameter value. For the other two parameter estimates, the biases are 0.081 and –0.078 (or –16.67 percent and –12.42 percent of the original parameter values), respectively.

[40]A recent interesting empirical study related to this point can be found in Bivand (2012), where questions are raised concerning the numerical issues involved in fitting the GNS model (termed SAC Durbin model in his paper) and how one should interpret the output.

TABLE 4: SLX Model Estimation Results for Pre-Specified and Parameterized **W**, for All Regressors Treated as Exogenous, for ln(**P**) Treated as Endogenous, and for Both ln(**P**) and **W** × ln(**P**) Treated as Endogenous

	OLS, **W** = BC (1)	Nonl. OLS, **W** = $1/d^{\gamma}$ (2)	2SLS, **W** = BC, ln(**P**) endogenous[a] (3)	2SLS, **W** = BC, ln(**P**), **W** × ln(**P**) endogenous[b] (4)	2SLS, **W** = $1/d^{\gamma}$, ln(**P**) endogenous[c] (5)	2SLS, **W** = $1/d^{\gamma}$, ln(**P**), **W** × ln(**P**) endogenous[d] (6)
ln(**P**)	−1.017	−0.908	−1.334	−0.785	−1.246	−1.273
	(−24.77)	(−24.43)	(−16.63)	(−3.69)	(−16.32)	(−15.40)
ln(**I**)	0.608	0.654	0.579	0.576	0.591	0.502
	(10.38)	(15.39)	(9.63)	(6.81)	(13.34)	(10.59)
W × ln(**P**)	−0.220	0.254	−0.109	−3.067	0.192	0.898
	(−2.95)	(3.08)	(−1.36)	(−3.59)	(3.00)	(6.25)
W × ln(**I**)	−0.219	−0.815	−0.230	−0.901	−0.750	−1.068
	(−2.80)	(−4.76)	(−2.89)	(−4.09)	(−14.14)	(−12.79)
γ		2.938			3.141	3.322
		(16.48)			(11.11)	(15.24)
R^2	0.897	0.916	0.374	<0	0.484	0.421
Log-Likelihood	1,668.4	1,812.9				
F-test instruments ln(**P**)			100.54	102.60	110.13	106.77
			[0.00]	[0.00]	[0.00]	[0.00]
F-test instruments **W** × ln(**P**)				46.07		156.09
				[0.00]		[0.00]
χ^2-test exogeneity instruments			0.087	3.84	0.112	0.907
			[0.99]	[0.28]	[0.99]	[0.82]
t-test ln(**P**) residual			4.63	−0.67	5.14	4.50
t-test **W** × ln(**P**) residual				2.94		−1.33

Note: See note to Table 2; coefficient estimates of **W** × ln(**P**) and **W** × ln(**I**) represent spillover effects. p-values of test statistics in squared brackets. Degrees of freedom of the F-test is (i) number of instruments and (ii) number of observations minus number of instruments and number of fixed effects. Degrees of freedom of χ^2-test is number of surplus instruments.

[a]Instruments (+exog.var. in eq.): **W** × Population, Tax, **W** × Tax + ln(**I**), **W** × ln(**P**), **W** × ln(**I**).

[b]Instruments (+exog.var. in eq.): **W** × Population, Tax, **W** × Compensation + ln(**I**), **W** × ln(**I**).

[c]Instruments (+exog.var. in eq.): Tax, **W** × Compensation + ln(**I**), **W** × ln(**P**), **W** × ln(**I**).

[d]Instruments (+exog.var. in eq.): **W** × Population,Tax, **W** × Compensation + ln(**I**), **W** × ln(**I**).

Prucha, 2010).[41] For example, the impact of unit i on unit j is not the same as that of unit j on unit i, and the information about the mutual proportions between the elements in the different rows of **W** gets lost. We therefore scale the elements of **W** based on inverse distance by its maximum eigenvalue.

The direct effect estimates change slightly when adopting the parameterized inverse distance specification, whereas the differences in the spillover effect estimates are substantial. We first draw attention to the results of the price spillover effects. In the first specification using the binary contiguity matrix, the spillover effect is negative and significant, which was discussed previously when comparing the different spatial econometric models. In fact, all the models allowing for the quantification of spillover effects resulted in negative price spillovers, which is not consistent with the bootlegging effect. In other words, these specifications do not confirm that consumers near state borders will purchase cigarettes in neighboring states if they are cheaper relative to prices in their own state.[42] By contrast, in the second specification using the parameterized inverse distance matrix, the price spillover effect is positive with an elasticity of 0.254 and significant (t-value = 3.08). The interpretation of this latter estimate is that a positive increase in own-state prices leads to increased sales of cigarette packs in neighboring states, which corroborates the existence of bootlegging behavior. Previous studies have used different specifications to capture this bootlegging effect and have mostly found evidence for it. However, no previous study has considered the SLX model and parameterizing **W**.

The estimate of the distance decay parameter is 2.938 and also highly significant. This makes sense because only people living near the border of a state are able to benefit from lower prices in a neighboring state on a daily or weekly basis. If the distance decay effect at five miles from the border is set to 1, it falls to 0.130 at 10 miles, 0.040 at 15 miles, and 0.017 at 20 miles. People living further from the border can only benefit from lower prices if they visit states for other purposes or if smuggling takes place by trucks over longer distances. It explains why the parameterized inverse distance matrix gives a much better fit than the binary contiguity matrix; the degree of spatial interaction on shorter distances falls much faster and on longer distances more gradually than according to the binary contiguity principle (see Figure 2). This is corroborated by the R^2, which increases from 0.897 to 0.916, and the log-likelihood function value, which increases from 1668.2 to 1812.9.

Turning to the income spillover effects in Table 4, the estimates are negative and highly significant across both spatial weights matrices. The main difference is that under the second column, the income elasticity is higher. These results indicate that increases in own-state per capita income decrease cigarette sales in neighboring states. An explanation could be that higher income levels reduce the necessity or incentive to purchase less expensive cigarettes elsewhere. In sum, the results suggest two forces at work that influence bootlegging behavior. There is a positive price effect which is reasonable since higher own-state prices will motivate people to search elsewhere, i.e., a substitution effect. However, increases in income have the opposite effect since there will be less motivation to make the effort of travelling across the border even if there is a relative price advantage, i.e., an income effect.

[41] For further discussion on the consequences of row-normalization, see Neumayer and Plümper (2012, 2013).

[42] We also estimated the model with the decay parameter set to one in advance ($\gamma = 1$ in Equation (11) and find that the price spillover is also negative, but statistically insignificant; results are available upon request.

Endogenous Regressors

Another important issue to address is whether or not cigarette prices are endogenous. Except for Kelejian and Piras (2014), previous spatial econometric studies based on Baltagi and Li's cigarette demand model did not treat price as being potentially endogenous. Although these studies argue or assume that price differences across states are largely due to state tax differences which are exogenously set by state legislatures, it is likely that demand has a feedback effect on price. Therefore, we formally test whether price and prices observed in neighboring states may be considered exogenous. The advantage of the SLX model over other spatial econometric models is that nonspatial econometric techniques can be used for this purpose. 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 methodology behind these tests is explained in many econometric textbooks; we used Hill et al. (2012, pp. 419–422).

As instrumental variables we initially exploited variables that can be taken from the Baltagi and Li (2004) cigarette demand data set. It concerns population size (16 years and over) in each state and the weighted average in neighboring states, as well as consumption and both own-state prices and neighboring state prices in the previous year. Regarding population size, it is important to note that the dependent variable is measured as per capita sales of cigarettes. Consumer demand for a product (y) depends on the market (x) in which a consumer operates (e.g., the geographical area or time period) and the equilibrium price in the market ($p(x)$). This can be formulated as: $\mathrm{E}(y\,|\,x,z) = D(p(x),z)$, where z are consumer attributes (in our case: income) and D reflects mean demand dependent on (x,z). The market equilibrium price can be seen to be: $p(x) = \pi[\mathrm{E}(y\,|\,x,\mathrm{z})^{*}m(x),s(x)]$, where $m(x)$ is population size in x and $s(x)$ denotes supply. In other words, although population size does not directly influence an individual's cigarette consumption, it may affect the price of cigarettes. Therefore, in addition to income and income in neighboring states which are already part of the SLX model, five potential instrumental variables were examined, among which is population size (note that state and time period fixed effects were also accounted for). As expected, the strongest instruments for the own-state and neighboring state prices appeared to be their respective values in the previous year. However, these instruments also appeared to be invalid due to serial correlation. Conversely, if this instrument set of lagged prices is left aside, we ended up with a weak set of instruments (with the first-stage F-statistic < 10). If instruments are weak, the 2SLS estimator can suffer large biases.[43]

For this reason, we decided to expand the Baltagi and Li cigarette demand data set with additional variables that may potentially serve as instruments for cigarette prices. After examining the literature and data availability for the whole sample covering 1963–1992, we included state compensation per employee and cigarette excise tax rates.[44] The former may affect the supply curve as it reflects labor costs, and thus the price as can be seen in the formulation above.[45] The latter is expected to be relevant since the price per pack of cigarettes is composed of the excise tax rate; the correlation coefficient between

[43]For a discussion on problems related to weak instruments, such as small sample over-fitting bias see, e.g., Staiger and Stock (1997).

[44]Both the expanded data set and the developed Matlab routines will be made available at the Web site of the second author. The data was taken from Orzechowski and Walker (2012) and the U.S. Bureau of Economic Analysis.

[45]A relevant point is raised by reviewers about the validity of using this variable as an instrument since the model includes income. Although Kelejian and Piras (2014) do not provide motivation for including this variable as an instrument, as we mentioned above, we include it because it can influence supply and thus price. Nevertheless, it is important to formally test for validity, which is discussed shortly.

these two variables is 0.60.[46] Although a rationale is not provided, Kelejian and Piras (2014) also include these measures as instrumental variables. Unfortunately, they do not carry out standard tests for weak instruments and overidentifying restrictions; their argument is that these tests first need to be extended to situations of spatial models dealing with an endogenous **W** (p. 147). Several nonspatial econometric studies have also used state cigarette excise taxes as an instrument for cigarette price (e.g., Keeler et al., 1993; Lovenheim, 2008). Although these studies, as well as a related study of Keeler et al. (1996), argue that cigarette excise tax rates have an exogenous quality, formal tests of the validity of the instruments are again not provided. Golden et al. (2014) investigate the determinants of cigarette tax rates using a panel of all U.S. states from 1981 to 2011, and find that these tax rates are largely driven by political characteristics (e.g., political party control) and also whether a state is an agricultural producer of tobacco. In addition, citizens' attitudes toward taxes and tobacco control, as well as cigarette tax rates in neighboring states are found to be significant. This suggests that the tax rate is not driven by cigarette demand. However, it is better to formally test whether these two potential instruments are exogenous and so we proceed.

The results reported in columns (3)–(6) cover two sets of two possibilities based on this extended data set. All models are estimated by 2SLS. The first set is used to investigate whether only $\ln(\mathbf{P})$ or both $\ln(\mathbf{P})$ and $\mathbf{W}\times\ln(\mathbf{P})$ need to be treated as endogenous regressors. The second set is used to find out whether empirical evidence is obtained in favor of the bootlegging effect if we take a step back and start from the pre-specified binary contiguity matrix, as in section "Standard Approach," or if we stick to the parameterized inverse distance matrix, as in section "The SLX Approach." The notes to Table 4 show that we ended up with a different set of instrumental variables in each case. The main reason is that $\mathbf{W}\times\ln(\mathbf{P})$ is used as an instrument for $\ln(\mathbf{P})$ and does not need to be instrumented if it is exogenous. Each set passed the *F*-test for strong enough instruments and the χ^2-test for exogenous instruments (also known as Sargan's overidentification test). The tax variables pass the test, but population size and compensation per employee measured in the own-state did not, but rather the neighboring values.

The *t*-tests on the residuals of the price variables taken from the first-stage regressions in the original SLX model (also known as the Hausman test) point to endogeneity of the price observed in the own state, but not of prices observed in neighboring states, when adopting the parameterized inverse distance matrix (columns 5 and 6). The *t*-value of the residual of the first-stage regression of the own-state price amounts to 4.50 and of prices in neighboring states to -1.33. 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. This implies that the results reported in column (5) outperform the results in columns (2) and (6).[47] Although the results in column (5) compared to those reported in column (2) change, they are relatively small compared to the changes we have

[46]These excise taxes are set by legislation and are usually collected before the point of sale from manufacturers, distributors, or wholesalers and often are denoted by a tax stamp; thus, unlike other sales taxes, they are usually included in the price of the item (CDC, 2012).

[47]Importantly, we not only took into account that the **W** matrix determining the exogenous interaction effects in the SLX model depend on the estimated distance decay parameter γ in the second stage of the 2SLS estimation procedure, but also that the **W** matrix of lagged prices in neighboring states depends on γ in the first stage of the 2SLS estimation procedure. This explains why the parameter estimate of γ in this model also changes.

seen among columns (1) and (2). The price spillover effect falls from 0.254 to 0.192, but remains significant (t-value = 3.00).

When the pre-specified binary contiguity matrix is adopted and ln(**P**) is treated as endogenous (column 3), which it should according to the t-test of 4.63 on the residual of the first-stage regression, the price spillover effect is negative rather than positive, just as in its counterpart estimated by OLS (column 1), although no longer significant. These things worsen when treating both ln(**P**) and **W**×ln(**P**) as endogenous (column 4), but this model variant appears to be less relevant. Based on the t-test statistics **W**×ln(**P**) should be treated as endogenous, whereas ln(**P**) should not, which is illogical. Furthermore, the R^2 falls below zero. Although mathematically possible when applying 2SLS, this is not very promising. Since the R^2's of the 2SLS regressions based on the parameterized inverse distance matrix also outperform their counterparts based on the binary contiguity matrix, the conclusion must be that the parameterization of the spatial weights matrix remains a necessary step to obtain plausible results.

5. CONCLUSIONS

This paper proposes a guide for researchers interested in measuring spillover effects. We first provide a comprehensive overview of what different spatial econometric model specifications imply about spillover effects. We show that the more commonly used models, especially the SAR model, are less flexible in their ability to measure spillover effects. Unless there is a well-founded theoretical argument pointing toward a model with endogenous interaction effects, this model is hard to justify. This point is also emphasized in the recent special theme issue of the *Journal of Regional Science* (Volume 52, Issue 2) appraising spatial econometrics. Therefore, instead of taking the more standard route of model selection in the spatial econometrics literature, we recommend taking the SLX model as point of departure, since it is the simplest model of all models producing flexible spillovers. Furthermore, we illustrate that adopting a parameterized instead of a pre-specified **W** allows for even more flexibility. It should be stressed that this is not a general result; alternatives, such as semiparametric and nonparametric approaches, might also be used for this purpose. A comparison of the pros and cons of these different approaches is an important topic for further research.

To compare the standard approach with our proposed alternative approach in an empirical setting, the Baltagi and Li (2004) cigarette demand model is estimated. Although other spatial econometric studies have used this data set for illustrative purposes, no previous study has considered the SLX model. A notable result from the SLX estimation results is that when employing the commonly used binary contiguity matrix, the price spillover effect estimate does not corroborate the existence of bootlegging. By contrast, when **W** is specified using the parameterized inverse distance specification, we do find significant evidence in favor of the bootlegging effect.

The SLX model is also useful to test for endogenous regressors since nonspatial econometric techniques can be used for this purpose. We find empirical evidence in favor of endogeneity of the price observed in the own state, but not of prices observed in neighboring states. Using the SLX model, a parameterized inverse distance matrix, and treating price in the state as endogenous, we find a significant price spillover effect of 0.192, indicating that if prices of cigarettes rise by 1 percent, cigarette demand in neighboring states will increase by 0.192 percent. In addition, we find a strong distance decay effect; at 20 miles from the border the impact of this bootlegging effect is already more than 50 times smaller than at 5 miles.

REFERENCES

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.

———. 2003a. "Rao's Score Test in Spatial Econometrics," *Journal of Statistical Planning and Inference*, 97, 113–139.

———. 2003b. "Spatial Externalities, Spatial Multipliers, and Spatial Econometrics," *International Regional Science Review*, 26, 153–166.

Anselin, Luc, Anil K. Bera, Raymond Florax, and Mann J. Yoon. 1996. "Simple Diagnostic Tests for Spatial Dependence," *Regional Science and Urban Economics*, 27, 77–104.

Arbia, Giuseppe and Bernard Fingleton. 2008. "New Spatial Econometric Techniques and Applications in Regional Science," *Papers in Regional Science*, 87, 311–317.

Baltagi, Badi H. 2008. *Econometric Analysis of Panel Data*. Chichester: Wiley.

Baltagi, Badi H. and Dan Levin. 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.

———. 1992. "Cigarette Taxation: Raising Revenues and Reducing Consumption," *Structural Change and Economic Dynamics*, 3(2), 321–335.

Baltagi, Badi H. and Dong Li. 2004. "Prediction in the Panel Data Model with Spatial Autocorrelation," in L. Anselin, R. Florax, and S.J. Rey (eds.), *Advances in Spatial Econometrics: Methodology, Tools, and Applications*. Berlin: Springer, pp. 283–295.

Beenstock, Michael and Daniel Felsenstein. 2012. "Nonparametric Estimation of the Spatial Connectivity Matrix Using Spatial Panel Data," *Geographical Analysis*, 44(4), 386–397.

Bhattacharjee, Arnab and Chris Jensen-Butler. 2013. "Estimation of the Spatial Weights Matrix Under Structural Constraints," *Regional Science and Urban Economics*, 43, 617–634.

Bivand, Roger S. 2012. "After 'Raising the Bar': Applied Maximum Likelihood Estimation of Families of Models in Spatial Econometrics," *Estadística Española*, 54(177), 71–88.

Blundell, Richard and Thomas M. Stoker. 2007. "Models of Aggregate Economic Relationships that Account for Heterogeneity," in J.J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*, Volume 6A. Amsterdam: Elsevier, pp. 4609–4663.

Boarnet, Marlon G. 1992. "Intra-metropolitan Growth Patterns: The Nature and Causes of Population and Employment Changes within an Urban Area," Ph.D. dissertation, Princeton University.

———. 1994a. "An Empirical Model of Intrametropolitan Population and Employment Growth," *Papers in Regional Science*, 73(2), 135–152.

———. 1994b. "The Monocentric Model and Employment Location," *Journal of Urban Economics*, 36, 79–97.

———. 1998. "Spillovers and the Locational Effects of Public Infrastructure," *Journal of Regional Science*, 38(3), 381–400.

Boarnet, Marlon G., Saksith Chalermpong, and Elizabeth Geho. 2005. "Specification Issues in Models of Population and Employment Growth," *Papers in Regional Science*, 84(1), 21–46.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. 2009. "Identification of Peer Effects through Social Networks," *Journal of Econometrics*, 150, 41–55.

Brueckner, Jan K. 2003. "Strategic Interaction among Governments: An Overview of Empirical Studies," *International Regional Science Review*, 26(2), 175–188.

———. 2006. "Strategic Interactions among Governments," in R.J. Arnott and D.P. McMillen (eds.), *A Companion to Urban Economics*. Malden, MA: Blackwell, pp. 332–347.

Buonanno, Paolo, Giacomo Pasini, and Paolo Vanin. 2012. "Crime and Social Sanction," *Papers in Regional Science*, 91, 193–218.

Burridge, Peter. 1981. "Testing for a Common Factor in a Spatial Autoregression Model," *Environment and Planning A*, 13, 795–400.

———. 2012. "Improving the J Test in the SARAR Model by Likelihood-based Estimation," *Spatial Economic Analysis*, 7(1), 75–107.

Burridge, Peter and Bernard Fingleton. 2010. "Bootstrap Inference in Spatial Econometrics: The J-Test," *Spatial Economic Analysis*, 5, 93–119.

Burridge, Peter and Ian Gordon (1981), "Unemployment in the British Metropolitan Labour Areas," *Oxford Economic Papers*, 33(2), 274–297.

CDC. 2012. (Centers for Disease Control and Prevention), *Morbidity and Mortality Weekly Report*, 61(12), 201–204.

Chintagunta, Pradeep K. and Harikesh S. Nair. 2011. "Discrete-Choice Models of Consumer Demand in Marketing," *Marketing Science*, 30(6), 977–996.

Chung, Jae Wan. 1994. *Utility and Production Functions*. Cambridge: Blackwell Publishing.

Cliff, Andrew D. and J. Keith Ord. 1981. *Spatial Processes*. London: Pion.

Corrado, Luisa and Bernard Fingleton. 2012. "Where is the Economics in Spatial Econometrics?" *Journal of Regional Science*, 52(2), 210–239.

Dalenberg, Douglas R., Mark D. Partridge, and Dan S. Rickman. 1998. “Public Infrastructure: Pork or Jobs Creator,” *Public Finance Review*, 26(1), 24–52.

Debarsy, Nicolas, Cem Ertur, and James P. LeSage. 2012. “Interpreting Dynamic Space-time Panel Data Models,” *Statistical Methodology*, 9, 158–171.

Drukker, David M., Peter Egger, and Ingmar R. Prucha. 2013. “On Two-Step Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances and Endogenous Regressors,” *Econometric Reviews*, 32(5–6), 686–733.

Dubin, Robin A. 1988. “Spatial Autocorrelation,” *Review of Economics and Statistics*, 70, 466–474.

———. 1992. “Spatial Autocorrelation and Neighborhood Quality,” *Regional Science and Urban Economics*, 22, 433–452.

Elhorst, J. Paul. 2010. “Applied Spatial Econometrics: Raising the Bar,” *Spatial Economic Analysis*, 5(1), 9–28.

———. 2013. “Spatial Panel Models,” in M.M. Fischer and P. Nijkamp (eds.), *Handbook of Regional Science*. Berlin: Springer, pp. 1637–1652.

———. 2014. “Matlab Software for Spatial Panels,” *International Regional Science Review*, 37(3), 389–405.

Ertur, Cem and Wilfried Koch. 2007. “Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence,” *Journal of Applied Econometrics*, 22, 1033–1062.

Fingleton, Bernard and Julie Le Gallo. 2008. “Estimating Spatial Models with Endogenous Variables, a Spatial Lag and Spatially Dependent Disturbances: Finite Sample Properties,” *Papers in Regional Science*, 87(3), 319–339.

Fingleton, Bernard and Enrique Lopez-Bazo. 2006. “Empirical Growth Models with Spatial Effects,” *Papers in Regional Science*, 85(2), 177–198.

Fischer, Manfred M., Thomas Scherngell, and Martin Reismann. 2009. “Knowledge Spillovers and Total Factor Productivity: Evidence Using a Spatial Panel Data Model,” *Geographical Analysis*, 41(2), 204–220.

Florax, Raymond and Henk Folmer. 1992. “Specification and Estimation of Spatial Linear Regression Models: Monte Carlo Evaluation of Pre-test Estimators,” *Regional Science and Urban Economics*, 22, 405–432.

Florax, Raymond, Henk Folmer, and Sergio J. Rey. 2003. “Specification Searches in Spatial Econometrics: The Relevance of Hendry’s Methodology,” *Regional Science and Urban Economics*, 33, 557–579.

Getis, Arthur and Jared Aldstadt. 2004. “Constructing the Spatial Weights Matrix Using a Local Statistic,” *Geographical Analysis*, 36(2), 90–104.

Gibbons, Stephen and Henry G. Overman. 2012. “Mostly Pointless Spatial Econometrics?” *Journal of Regional Science*, 52(2), 172–191.

Golden, Shelley D., Kurt M. Ribisl, and Krista M. Perreira. 2014. “Economic and Political Influence on Tobacco Tax Rates: A Nationwide Analysis of 31 Years of State Data,” *American Journal of Public Health*, 104(2), 350–357.

Harris, Richard, John Moffat, and Victoria Kravtsova. 2011. “In Search of W,” *Spatial Economic Analysis*, 6(3), 249–270.

Hill, R. Carter, William E. Griffiths, and Guay C. Lim. 2012. *Principles of Econometrics*. Asia: Wiley.

Holtz-Eakin, Douglas and Amy Ellen Schwartz. 1995. “Spatial Productivity Spillovers from Public Infrastructure: Evidence from State Highways,” *International Tax and Public Finance*, 2, 459–468.

Huber, Martin, Michael Lechner, and Conny Wunsch. 2013. “The Performance of Estimators Based on the Propensity Score,” *Journal of Econometrics*, 175(1), 1–12.

Keeler, Theodore E., Teh-wei Hu, Paul G. Barnett, and Willard G. Manning. 1993. “Taxation, Regulation, and Addiction: A Demand Function for Cigarettes Based on Time-series Evidence,” *Journal of Health Economics*, 12, 1–18.

Keeler, Theodore E., Teh-Wei Hu, Paul G. Barnett, Willard G. Manning, Hai-Yen Sung. (1996), “Do Cigarette Producers Price-Discriminate by State? An Empirical Analysis of Local Cigarette Pricing and Taxation,” *Journal of Health Economics*, 15, 499–512.

Kelejian, Harry H. 2008. “A Spatial J-Test for Model Specification against a Single or a Set of Non-Nested Alternatives,” *Letters in Spatial and Resource Sciences*, 1(1), 3–11.

Kelejian, Harry H. and Gianfranco Piras. 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, Harry H. and Ingmar R. 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.

———. 1999. “A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model,” *International Economic Review*, 40, 509–533.

———. 2010. “Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances,” *Journal of Econometrics*, 157, 53–67.

Kelejian, Harry H., Ingmar R. Prucha, and Yevgeny Yuzefovich. 2004. “Instrumental Variable Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances: Large and Small Sample Results,” in J.P. LeSage and R.K. Pace (eds.), *Spatial and Spatiotemporal Econometrics*. Amsterdam: Elsevier, pp. 163–198.

Lacombe, Donald J. and James P. LeSage. 2013. “Using Bayesian Posterior Model Probabilities to Identify Omitted Variables in Spatial Regression Models,” *Papers in Regional Science*, DOI: 10.1111/pirs.12070.

Lee, Lung-Fei. 2004. “Asymptotic Distribution of Quasi-maximum Likelihood Estimators for Spatial Autoregressive Models,” *Econometrica*, 72, 1899–1925.

Lee, Lung-Fei and Jihai Yu. 2010. “Some Recent Developments in Spatial Panel Data Models,” *Regional Science and Urban Economics*, 40, 255–271.

LeSage, James P. and R. Kelley Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton, FL: Taylor and Francis.

———. 2011. “Pitfalls in Higher Order Model Extensions of Basic Spatial Regression Methodology,” *The Review of Regional Studies*, 41(1), 13–26.

Lewit, Eugene M. and Douglas Coate. 1982. “The Potential of Using Excise Taxes to Reduce Smoking,” *Journal of Health Economics*, 1, 121–145.

Lewit, Eugene M., Douglas Coate, and Michael Grossman. 1987. “The Effects of Government Regulation on Teenage Smoking,” *Journal of Law and Economics*, 24, 545–569.

Liu, Xiaodong and Lung-Fei Lee. 2013. “Two-Stage Least Squares Estimation of Spatial Autoregressive Models with Endogenous Regressors and Many Instruments,” *Econometric Reviews*, 32(5–6), 734–753.

Lovenheim, Michael F. 2008. “How Far to the Border?: The Extent and Impact of Cross-Border Casual Cigarette Smuggling,” *National Tax Journal*, 56(1), 7–33.

Meen, Geoff. 1996. “Spatial Aggregation, Spatial Dependence and Predictability in the UK Housing Market,” *Housing Studies*, 11, 345–372.

McMillen, Daniel P. 2010. “Issues in Spatial Data Analysis,” *Journal of Regional Science*, 50(1), 119–141.

———. 2012. “Perspectives on Spatial Econometrics: Linear Smoothing with Structured Models,” *Journal of Regional Science*, 52(2), 192–209.

———. 2013. *Quantile Regression for Spatial Data*. Heidelberg, New York, Dordrecht, London: Springer.

Moulton, Brent R. 1990. “An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units,” *Review of Economics and Statistics*, 72, 334–338.

Mur, Jesus and Ana Angulo. 2009. “Model Selection Strategies in a Spatial Setting: Some Additional Results,” *Regional Science and Urban Economics*, 39, 200–213.

Mur, Jesus, Marcos Herrera, and Manuel Ruiz. 2013. “Selecting the W Matrix: Parametric vs. Non Parametric Approaches,” paper presented at the Econometrics of Social Interaction Symposium, University of York.

Neumayer, Eric and Thomas Plümper. 2012. “Conditional Spatial Policy Dependence: Theory and Model Specification,” *Comparative Political Studies*, 45(7), 819–849.

———. 2013. “W,” available on SSRN at http://ssrn.com/abstract=2208706. Last accessed January 21, 2015.

Orzechowski and Walker. 2012. *The Tax Burden on Tobacco, Historical Compilation, Volume* 47, 2012. Arlington, Virginia: Orzechowski and Walker.

Pace, R. Kelley, Ronald Barry, John M. Clapp, and Mauricio Rodriguez. 1998. “Spatiotemporal Autoregressive Models of Neighborhood Effects,” *Journal of Real Estate Finance and Economics*, 17(1), 15–33.

Pace, R. Kelley and Shuang Zhu. 2012. “Separable Spatial Modeling of Spillovers and Disturbances,” *Journal of Geographical Systems*, 14(1), 75–90.

Partridge, Mark D., Marlon G. Boarnet, Steven Brakman, and Gianmarco Ottaviano. 2012. “Introduction: Whither Spatial Econometrics?” *Journal of Regional Science*, 52(2), 167–171.

Pinkse, Joris and Margaret A. Slade. 2010. “The Future of Spatial Econometrics,” *Journal of Regional Science*, 50(1), 103–117.

Qu, Xi and Lung-Fei Lee. 2015. “Estimating a Spatial Autoregressive Model with an Endogenous Spatial Weight Matrix,” *Journal of Econometrics*, 184, 209–232.

Soetevent, Adriaan R. and Peter Kooreman. 2007. “A Discrete Choice Model with Social Interactions: with an Application to High School Teen Behavior,” *Journal of Applied Econometrics*, 22, 599–624.

Song, Shunfeng. 1996. “Some Tests of Alternative Accessibility Measures: A Population Density Approach,” *Land Economics*, 72(4), 474–482.

Staiger, Douglas and James H. Stock. 1997. “Instrumental Variables Regression with Weak Instruments,” *Econometrica*, 65(3), 557–586.

Tobler, Waldo R. 1970. “A Computer Movie Simulating Urban Growth in the Detroit Region,” *Economic Geography*, 46(2), 234–240.

Wasserman, Jeffrey, Willard G. Manning, Joseph P. Newhouse, and John D. Winkler. 1991. “The Effects of Excise Taxes and Regulations on Cigarette Smoking,” *Journal of Health Economics*, 10, 43–64.

Yen, Steven T. and Chung L. Huang. 1996. “Household Demand for Finfish: A Generalized Double-hurdle Model,” *Journal of Agricultural and Resource Economics*, 21(2), 220–234.