

Elhorst, Paul; Vega, Solmaria Halleck

Conference Paper

On spatial econometric models, spillover effects, and W

53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy

Provided in Cooperation with:

European Regional Science Association (ERSA)

Suggested Citation: Elhorst, Paul; Vega, Solmaria Halleck (2013) : On spatial econometric models, spillover effects, and W, 53rd Congress of the European Regional Science Association: "Regional Integration: Europe, the Mediterranean and the World Economy", 27-31 August 2013, Palermo, Italy, European Regional Science Association (ERSA), Louvain-la-Neuve

This Version is available at:

<https://hdl.handle.net/10419/123888>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

On spatial econometric models, spillover effects, and W

Solmaria Halleck Vega[†] and J. Paul Elhorst[†]

February 2013

Abstract

Spatial econometrics has recently been appraised in a theme issue of the *Journal of Regional Science*. Partridge et al. (2012) provide an overview of the three contributing papers, the most critical being Gibbons and Overman (2012). Although some of the critiques raised are valid, they are issues that can be overcome by improving applied spatial econometric work. There has been excessive use of so-called global spillover models and too much emphasis on statistical testing procedures. Theory or the specific context of the empirical application should be the main guide for specifying a model. Especially the so-called SLX model merits more attention, as it produces local spillovers that are different for each explanatory variable in the equation and it allows for the parameterization of the spatial weights matrix W . The latter is a key contribution since a major concern of spatial econometric modeling, as is thoroughly discussed in Corrado and Fingleton (2012), is the a priori specification of W . This paper highlights these issues with an empirical application and recently proposed approaches for selecting a model specification, which are useful and promising steps forward for applied work involving spatial econometrics.

Key words: Spatial econometric models, spillover effects, spatial weights matrix

JEL classification: C01, C21, C23

To be presented at 53rd ERSA conference, Palermo, August 27-31, 2013

[†]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.p.elhorst@rug.nl. We would like to thank Dennis Robinson and Oleg Smirnov for their useful comments on a previous version presented at the 59th Annual North American Meetings of the Regional Science Association International, November 2012. We also gratefully acknowledge James LeSage and Kelley Pace for their thoughtful review of a revised version in December 2012. The present version of the paper has greatly benefited from their comments and suggestions. The views expressed herein, nevertheless, are those of the authors.

1. Introduction

Spatial spillovers are a main interest in regional science. In contrast to standard econometric models which restrict spillovers to be zero, a valuable aspect of spatial econometric models is that the magnitude and significance of spatial spillovers can be empirically assessed. This is why spatial econometric methods are extensively used in regional science research and have also seen increasing use in other social science fields. Recently, the *Journal of Regional Science* has published a much-discussed theme issue appraising spatial econometrics. Partridge et al. (2012) provide an overview of the three contributing papers. McMillen's (2012) critique mainly focuses on the limitations of the spatial lag model (SAR) and the spatial error model (SEM), and is partly based on previous work (McMillen, 2003, 2010).¹ 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.

In the last couple of years, however, there has been a growing interest in models containing more than just one spatial interaction effect. In particular, this pertains to the so-called SAC model that includes both a spatially lagged dependent variable and a spatially autocorrelated error term (based on Kelejian and Prucha, 1998 and related work) and the spatial Durbin model (SDM) that includes both a spatially lagged dependent variable and spatially lagged explanatory variables (based on LeSage and Pace, 2009). Especially the latter model is highly criticized in the contributing paper by Gibbons and Overman (2012) because of identification problems.² At most, they claim that the parameters of the SDM are only weakly identified in theory if the spatial weights matrix W is not idempotent, which still depends on the assumption that W is specified correctly. They conclude that it is therefore preferable to estimate the SLX model containing exogenous interaction effects rather than directly estimating the SAR or SDM models.³ Corrado and Fingleton (2012) are more positive, but nevertheless strongly argue for the use of more substantive theory in empirical spatial econometric modeling, especially regarding W .

In view of these critical notes it is clear that the modeling strategy to find the spatial econometric model that best describes the data needs revision. Instead of testing the OLS

¹In this paper, we use the acronyms most commonly used in the spatial econometrics literature to refer to the model specifications (see e.g., LeSage and Pace, 2009).

²This is related to Manski's (1993) reflection problem, that endogenous and exogenous interaction effects cannot be distinguished from each other.

³The label SLX (spatial lag of X) model is given by LeSage and Pace (2009).

model against the SAR and SEM models for an exogenously specified W , we propose to take the SLX model as point of departure using a W that is parameterized. Next, we propose to rely on theory or the specific context of the empirical application rather than statistical testing as the main guide to select either a so-called global or local spillover model. To explain this modeling strategy, we provide a concise overview in Section 2 of the spillover effects that result from linear spatial econometric models with all different combinations of interaction effects. In addition, we explain the distinction between local and global spillover models.

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, Gibbons and Overman (2012) do not discuss the issue of spatial spillover effects, while it is one of the reasons why we should follow them and take the SLX model as point of departure. Another reason, also not discussed in their paper, is that the elements of W can be parameterized. This is a significant contribution since an often criticized aspect of spatial econometric modeling is the a priori specification of W , which is a topic extensively discussed in Corrado and Fingleton (2012). This part is worked out in Section 3.

Using the well-known Baltagi and Li (2004) US state cigarette demand data set, Section 4 first illustrates the spatial spillovers resulting from the different model specifications when adopting an exogenously specified W and next the spatial spillovers that follow when adopting the revised modeling strategy with a parameterized W . This empirical application demonstrates that not parameterizing W has the effect that the researcher draws wrong conclusions. Section 5 concludes the paper with a summary of the main results.

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 for reasons to be explained below. The simplest model considered in Figure 1 is the familiar linear regression model which takes the form

$$Y = \alpha \iota_N + X\beta + \varepsilon \quad (1)$$

where Y represents an $N \times 1$ vector consisting of one observation on the dependent variable for every unit in the sample ($i = 1, \dots, N$), ι_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, with the associated parameters β contained in a $K \times 1$ vector, and $\varepsilon = (\varepsilon_1, \dots, \varepsilon_N)^T$ is a vector of independently and identically distributed disturbance terms with zero mean and variance σ^2 .⁴ Since model (1) is commonly estimated by ordinary least squares (OLS), it is often referred to as the OLS model.

Insert Figure 1

Starting with the OLS model, the spatial econometrics literature has developed models that treat three different types of interaction effects among units: (1) endogenous interaction effects among the dependent variable, (2) exogenous interaction effects among the explanatory variables, and (3) interaction effects among the error terms.⁵

Unfortunately, there is large 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 and/or interaction effects among the error terms, such as the SAR, SEM, and SAC models, because of all the econometric problems accompanying the estimation of these models. The reason they do not focus on spatial econometric models with exogenous interaction effects is because the estimation of this model does not pose any econometric problems; standard estimation techniques suffice under these circumstances.⁶ Consequently, the SLX model is not part of the toolbox of researchers interested in the econometric theory of spatial models. By contrast, practitioners often take the SLX model as point of departure due to their main focus on spillovers. They tend to present their work at conferences organized by the Regional Science Association at different continents, whereas theoreticians visit econometric and spatial econometric conferences, as a result of which there is insufficient interaction. In this respect, the extension of Figure 1 with the SLX model can be seen as a first attempt to bridge this gap.

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

⁴The superscript T indicates the transpose of a vector or matrix.

⁵For a detailed description with examples of these different types of interaction effects refer to Elhorst (2013).

⁶By 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.

$$Y = \rho WY + \alpha \iota_N + X\beta + WX\theta + u, \quad u = \lambda Wu + \varepsilon \quad (2)$$

We will refer to model (2) as the general nesting spatial (GNS) model⁷ since it includes all types of interaction effects. 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 θ , like β , 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 and the SDEM are generally overlooked.

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.⁸ 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.

Spillover effects

The 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

⁷LeSage and Pace (2009) neither name nor assign an equation number to model (2), which reflects the fact that this model is typically not used in applied research.

⁸For 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.

spatial dependence in the disturbance process, it also provides no information about spillovers, as shown in Table 1. 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 applied researchers want to obtain inference on spillovers, alternative spatial econometric models need to be considered.

Insert Table 1

One such model that allows an empirical assessment of the magnitude and significance of spillover effects is the SAR model. This is clearly an advantage compared to the other widely used SEM model. If the SAR model (3) is rewritten to its reduced form (4), the direct and spillover effects can be obtained.

$$Y = \rho WY + \alpha \iota_N + X\beta + \varepsilon \quad (3)$$

$$Y = (I - \rho W)^{-1} \alpha \iota_N + (I - \rho W)^{-1} X\beta + (I - \rho W)^{-1} \varepsilon \quad (4)$$

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

$$\left[\frac{\partial E(Y)}{\partial x_{1k}} \dots \frac{\partial E(Y)}{\partial x_{Nk}} \right] = (I - \rho W)^{-1} \beta_k, \quad (5)$$

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

$$(I - \rho W)^{-1} = I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots \quad (6)$$

Since the non-diagonal elements of the first matrix term on the right-hand side (the identity matrix I) are zero, this term represents a direct effect of a change in X . Conversely, since the diagonal elements of the second matrix term on the right-hand side (ρW) are zero by

assumption, this term represents an indirect effect of a change in X . All other terms on the right-hand side represent second- and higher-order direct and spillover effects.⁹ From Table 1 it can also be noted that the SAC model shares the same direct and spillover effect properties as the SAR model.

An important characteristic of the spillovers produced by these models is that they are *global* in nature. Anselin (2003) describes the difference. A change in X at any location will be transmitted to all other locations following the matrix inverse in equation (6), also if two locations according to W are unconnected. In contrast, *local* 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. As will be discussed later, global spillovers are often more difficult to justify, which is an issue discussed in many studies, not only Gibbons and Overman (2012), but also Arbia and Fingleton (2008) and Lacombe and LeSage (2012).

In addition, a SAR model has several limitations. 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. 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 contrast to the models above, the SLX model contains spatially lagged explanatory variables, taking the following form

⁹A note regarding the direct and spillover effect estimates is how they can be reported. 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 $N \times N$ 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) 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. The total economy-wide effect is the sum of the direct and spillover effects. However, whether or not the researcher wants to apply this useful solution depends, of course, on the objective and nature of the study.

$$Y = \alpha_{1N} + X\beta + WX\theta + \varepsilon \quad (7)$$

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 non-spatial variables (β_k) and the spillover effects are those associated with the spatially lagged explanatory variables (θ_k). 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.

Furthermore, whereas endogenous interaction effects (WY) and interaction effects among the error terms (Wu) require conditions on W to obtain consistent parameter estimates (Kelejian and Prucha, 1998; Lee, 2004), no conditions are required with respect to W in the case where this matrix is used to model exogenous interaction effects (WX).¹⁰ One of the most important conditions that may be dropped is that W is exogenous and should be specified in advance. This opens up the opportunity to parameterize the elements of W . Like the SLX model, the direct and spillover effects of the SDEM are the vectors of the response parameters β and θ , respectively. Even though these models are easier to estimate and interpret and most importantly are useful for investigating local spillovers, they are not as commonly applied as global spillover specifications.

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).

$$Y = \rho WY + \alpha_{1N} + X\beta + WX\theta + \varepsilon \quad (8)$$

$$Y = (I - \rho W)^{-1} \alpha_{1N} + (I - \rho W)^{-1} (X\beta + WX\theta) + (I - \rho W)^{-1} \varepsilon \quad (9)$$

From equation (9), the matrix of partial derivatives of $E(Y)$ with respect to the k th explanatory variable of X in unit 1 up to unit N is obtained

¹⁰ Lee (2004) shows that one of the following two conditions should be satisfied: (a) the row and column sums of the matrices W , $(I - \rho W)^{-1}$ and $(I - \rho W)^{-1} W$ before W is row-normalized should be uniformly bounded in absolute value as N goes to infinity, or (b) the row and column sums of W before W is row-normalized should not diverge to infinity at a rate equal to or faster than the rate of the sample size N . Condition (a) originates from Kelejian and Prucha (1998, 1999) and condition (b) from Lee (2004).

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

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 effects 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 parameter ρ and W , but also on the coefficient estimate θ_k (Elhorst, 2010). The same applies to the GNS model.

Even though taking the GNS model as point of departure to measure spillovers seems appealing since it contains all possible interaction effects, one major problem is that its parameters are weakly identified. Two simpler models, namely the SDM and SDEM, are already difficult to distinguish from each other.¹¹ This problem is strengthened when estimating the GNS model, often leading to a model that is overparameterized. Parameters have the tendency to become insignificant as a result of which this model does not outperform the SDM and SDEM models. The SDM and GNS models are global spillover specifications.

Global vs. local spillover specifications

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 for 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 not suitable to study spatial spillover effects since they impose restrictions on their magnitude in advance. Since Figure 1 shows that the SLX model is the simplest of these four models, it is recommendable to take that model as point of departure when having any empirical evidence that the observations in the sample are spatially dependent. The next question is whether the SLX model describes the spatial dependence structure adequately, and if not, how to decide whether to adopt a local or global spillover specification, respectively the SDEM or the SDM model.

¹¹Gibbons and Overman (2012, p. 178) emphasize the identification problem: “only the overall effect of neighbors’ characteristics is identified, not whether they work through exogenous or endogenous neighborhood effects.”

One might use the classic and robust LM-tests proposed by Anselin (1988) and Anselin et al. (1996) for this purpose, but it is not clear whether these tests are very powerful since their performance has only been investigated based on the residuals of the OLS model and not on those of the SLX model. Probably they are not since the difference between the log-likelihood function values of the SDEM and SDM models tends to be smaller than that between the SEM and SAR models. The empirical application in Section 4 will illustrate this. Furthermore, Gibbons and Overman (2012) criticize model specification testing, while LeSage and Pace (2009, p. 156) contrary to what Gibbons and Overman seem to believe, endorse this viewpoint: “there is too much emphasis in the spatial econometrics literature on use of statistical testing procedures to infer the appropriate model specification...”

The inclusion of spatial effects in applied econometric models is typically motivated either on theoretical grounds, following from the formal specification of spatial interaction in an economic model, or on practical grounds, due to peculiarities of the data used in an empirical analysis (Anselin, 2002). For example, Ertur and Koch (2007) adopt the SDM model to analyze economic growth based on a well-founded theoretical background.¹² One example based on the context of the study is Kirby and LeSage (2009), where commuting times for residents of Census Tracts is analyzed. A global spillover specification is appropriate because congestion on roads in one Census tract will spillover, impacting commuting times of residents in other Census tracts throughout the metropolitan area.

LeSage and Pace (2009, p.28) also provide an econometric-theoretical motivation for the SDM model under the following circumstances: (1) there is one (or there are more) potentially important variable(s) omitted from the model, (2) this variable is likely to be correlated with the explanatory variables included in the model; and (3) the disturbance process is likely to be spatially dependent. They show that the SDM model produces unbiased, though inefficient, parameter estimates under these circumstances, whereas the SLX, SEM and SDEM models do not. Gibbons and Overman (2012, appendix) confirm that this setup leads to the SDM model, but emphasize that this does not solve the problem whether the causal effect of the observed spatial patterns in the data is due to endogenous interaction effects or interaction effects among the error term. In the words of Corrado and Fingleton (2012), the coefficient estimate for the *WY* variable may be significant because it may be picking up the effects of omitted *WX* variables or nonlinearities in the *WX* variables if they are erroneously specified as being linear.

¹²Corrado and Fingleton (2012) provide more examples of studies where the spatial econometric model specification is theoretically justified.

The conclusion must be that global spillover specifications, unless theoretically motivated, are difficult to justify and have been overused in applied studies. This is recently confirmed by Lacombe and LeSage (2012, p. 9):

There are many regional science relationships where we should have a priori knowledge that spillovers are local in nature as opposed to global. Differences in state policies regarding taxes, minimum wages, public assistance programs and so on are likely to exert some spillover impacts on migration decisions of residents living in border counties. While it might be reasonable that these residents move to neighboring counties across the state line, it does not seem reasonable that these differences would set in motion a process leading to a new steady-state equilibrium having migration (or other) impacts on distant states.

3. The SLX model and the parameterization of W

An often criticized aspect of using spatial econometric models is that W is specified in advance instead of being estimated along with the parameters in the model. There have been many studies that attempt to investigate how robust results are to different specifications of W and to determine which one best fits the data using criteria such as the log-likelihood function values and Bayesian posterior model probabilities. However, parameterizing W is a further step forward. The SLX model offers that opportunity.

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 point of say m kilometers.¹³ 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, as critiqued by Partridge et al. (2012), this is often used too readily to justify spatial econometric specifications. Even if there are theoretical reasons indicating that distance matters, 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 can be quite arbitrary and it is preferable to be able to estimate the distance decay parameter.

¹³Alternative ways to specify W can be found in Corrado and Fingleton (2012), such as using economic variables. However, an advantage of specifying W based on location is that the elements are exogenous.

Inspired by Newton's law of universal gravitation, the distance decay can be formulated as a power function. For example, the elements of W can be defined based on inverse distances with a distance decay factor

$$w_{ij} = \frac{1}{d_{ij}^\gamma}, \quad (11)$$

where d_{ij} denotes the distance between observations i and j , and γ is the distance decay parameter. A nonlinear but straightforward estimation technique can be used to estimate the parameter γ , which provides more information about 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. However, there is still the possibility that interaction continues, depending on the degree of distance decay.

Insert Figure 2

We emphasize that theory should be the driving force that determines the specification of W , as is discussed extensively in Corrado and Fingleton (2012). If a substantive theoretical framework is lacking, then an option could be to compare the results using alternative functional forms of W . Although one might argue that there are still numerous functional forms that can be specified, of overriding importance is that by parameterizing W it is tested rather than assumed to which extent interaction decreases as the distance between units becomes greater. Fischer et al. (2006) and Fischer et al. (2009) estimate the distance decay parameter using an exponential function in empirical applications investigating knowledge spillovers. There have also been other studies that have parameterized W (see e.g., Burridge and Gordon, 1981; Pace et al., 1998; Kakamu, 2005).¹⁴

¹⁴ Burridge and Gordon (1981) employ an exponential distance decay function to investigate unemployment in British metropolitan labor areas. Pace et al. (1998) parameterized the weight given to neighbors in space and time in their spatiotemporal model and provide an empirical application using Fairfax County, Virginia housing prices data from 1969-1991. Kakamu (2005) proposes a distance functional weight matrix model where the parameter to be estimated reflects the intensity of spatial decay.

4. Spatial econometric model comparison: Empirical application

To demonstrate that the adoption of a non-parameterized W may lead to wrong inferences, even when applying sophisticated statistical testing procedures, a demand model for cigarettes is estimated based on panel data from 46 US states over the period 1963 to 1992. This data set is taken from Baltagi (2008) and has been used for illustration purposes in other studies as well.¹⁵ The dependent variable is real per capita sales of cigarettes, which is measured in packs per person aged 14 years and older. The explanatory variables are average retail price of a pack of cigarettes and real per capita disposable income. All variables are taken in logs, as is done in Baltagi and Li (2004). The bootlegging effect, i.e. consumers will purchase cigarettes in nearby states if there is a price advantage (Baltagi and Li, 2004), is the main motivation to consider an econometric model with spatial interaction effects.

We include state-specific and time-specific fixed effects.¹⁶ In view of the complaint raised by Partridge et al. (2012, p.168) that spatial econometrics does not always do a good job of differentiating between spatial correlation (perhaps due to common explanatory factors) and spatial causality, this control for time-period fixed effects is important. Using Monte Carlo simulation experiments, Lee and Yu (2010) report 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. Initially, the spatial weights matrix W is specified as a row-normalized binary contiguity matrix, with elements $w_{ij} = 1$ if two spatial units share a common border, and zero otherwise. It should be stressed that this specification of W is also used in Baltagi and Li (2004) and Elhorst (2012, 2013). Debarsy et al. (2012) specify a row-normalized W based on the binary contiguity principle and state border miles in common between the states and find that the results are similar regardless of the specification used when estimating a dynamic SDM.

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

¹⁵The data is available at www.wiley.co.uk/baltagi/. For an adapted version refer to www.regroningen.nl/elhorst. Except for Baltagi and Li (2004), this data set is also used in Elhorst (2012, 2013), Debarsy et al. (2012), and Kelejian and Piras (2012).

¹⁶In this way, specific state characteristics and e.g., policy changes that occurred during the period are controlled for. For more details on the reasons to include state and time specific effects, refer to Baltagi (2008). Elhorst (2012) 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.

ML, with the exception of the SLX model which is estimated using non-linear least squares. The coefficient estimates of the two explanatory variables, price and income, are statistically significant at the one 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.

Insert Table 2

Compared to the coefficient estimates of the explanatory variables, the coefficients of the spatial interaction effects show much more variation. The coefficient estimate of the spatially lagged dependent variable (WY) is positive, significant and almost similar in magnitude in the SAR and SDM models, whereas it is negative in the SAC model, and both negative and significant in the GNS model. The coefficient estimate of Wu is positive and significant in the SEM, SDEM, SAC and GNS models, though substantially higher in the latter model. The magnitude and significance level of the exogenous interaction effects (WX) are similar for the SLX and SDEM models, which is as anticipated since they differ only in that the latter model also accounts for spatial dependence in the disturbances (Wu). By contrast, the exogenous price interaction effect in the SDM model has a different sign than in the SLX, SDEM and GNS models, and the exogenous income interaction effect in the GNS model has a different sign than in the SLX, SDM and SDEM models.

Insert Table 3

Since for some of the model specifications the coefficient estimates cannot be interpreted directly, we turn to the direct and spillover effect estimates in Table 3 to better compare the different models.¹⁷ 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 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.

¹⁷To 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.

For example, in the SAR model the point estimate for the income variable is 0.465, whereas the direct effect is 0.461, amounting to a feedback effect of less than 1%. 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 what measure (coefficient estimate or direct effect) and model is used.

In contrast, and just like the coefficient estimates of the spatial interaction effects, the discrepancies between the spillover impacts are substantial (Table 3). The results show that the choice of model specification leads to 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 to the spillover impact corresponding to the SAR model, whereas a negative and significant 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 particular state will not only lead to reduced cigarette demand in the state itself, but also in nearby states.¹⁸ Importantly, this result is not in line with the bootlegging effect originally found by Baltagi and Levin (1992) and which was the main motivation to adopt a spatial econometric model.

The explanation for this contradictory finding is not so much that the spatial econometric model being adopted is too limited. From a statistical viewpoint, it is true that the flexible models (SLX, SDEM, SDM, and GNS) produce price and income spillover effects that are comparable in terms of sign, magnitude and significance levels, whereas the non-flexible models produce erratic results and for this reason should be rejected. Furthermore, since the most general model in the set, the GNS, reduces to the next-most general, either the SDM if $\lambda=0$ or the SDEM if $\rho=0$, but these two parameters in the GNS model appear to be significant, these model reductions need to be rejected too. Nevertheless, even this model with all possible spatial interaction effects is still not able to capture the bootlegging effect, indicating that it might still be misspecified. A first indication for this is

¹⁸The 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.

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.

Additionally, there is the problem that the type of interaction effect causing the spillover effects, i.e. the mechanism through which observations at other locations is affected, is very different across the SLX, SDEM, SDM, and GNS models. In the SDM and GNS model the spillovers work through both endogenous interaction effects, whereas in the SLX and SDEM models they only work through exogenous interaction effects. It is therefore important to pay more attention to the context of the empirical application, a point we emphasized previously. In particular, the nature of the spillovers, either local or global in nature should be given more careful thought. In the case of cigarette demand, a local spillover model is probably more appropriate since cross-border shopping is not likely to influence distant states accompanied by feedback effects. Another problem is that the binary contiguity matrix limits cross-border shopping to only adjacent states, while in reality people may also benefit from lower prices if they visit states for other purposes. In the next section we therefore investigate whether the results improve when adopting a parameterized inverse distance matrix.

5. Empirical application using the SLX model with parameterized W

Table 4 reports the estimation results explaining cigarette demand for the SLX model using different functional forms of W and estimating the distance decay parameter, as was described in Section 3. The first column shows the results using the row-normalized binary contiguity matrix.¹⁹ The second column reports the estimation results using inverse distance with the distance decay parameter γ set equal to one in advance. Row-normalizing a weights matrix based on inverse distance causes its economic interpretation in terms of distance decay to no longer be valid (Kelejian and Prucha, 2010). 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

¹⁹ These are the same as the estimation results reported under the SLX model column in Table 2, which are repeated in Table 4 to facilitate comparison with the other functional forms.

between the elements in the different rows of W gets lost. We therefore scale the elements of W based on inverse distance by the maximum eigenvalue.

Insert Table 4

The direct effect estimates are similar across all different functional forms of W , although the estimate is a bit lower for price for the parameterized inverse distance specification. However, the differences in the direct effects are miniscule compared to the substantial differences in the spillover effect estimates. We first draw attention to the results of the price spillover effects, which are quite remarkable. In the first specification using the binary contiguity matrix, the spillover effect is strongly significant and negative, 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.

In the second column using specification (11) with $\gamma = 1$ for W , the price spillover is still negative, but statistically insignificant. Although the results in the third column are also based on the inverse distance matrix, when the distance decay parameter is estimated, the result is completely different. The price spillover effect is positive with an elasticity of 0.254 and significant (t-statistic = 3.083). 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 and it is notable that this specification captures bootlegging even without making the model dynamic, as in Debarsy et al. (2012), Kelejian and Piras (2012), and Elhorst (2013).²⁰ In this respect it should be noted that some researchers prefer simpler models to more complex ones (Occam's razor). One problem of complex models is overfitting, the fact that excessively complex models are affected by statistical noise, whereas simpler models may capture the underlying process better and may thus have better predictive performance. The estimate of the distance

²⁰The first two studies adopt a dynamic SDM model, and the last study a dynamic SAR model. The reason Kelejian and Piras (2012) chose the SAR model, despite the disadvantages of this model spelled out in Section 2, is probably again that this model is more interesting from an econometric-theoretic point of view.

decay parameter is 2.938 and highly significant. This makes sense because only (the few) people living at the border of a state are able to benefit from lower prices in a neighboring state on a daily or weekly basis. People living further from their state borders can only benefit from lower prices if they visit states for other purposes. 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 all different functional forms. The main difference is that under the third column, the estimate 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 makes sense since higher own-state prices will motivate people to search elsewhere. However, increases in income have the opposite effect since there will be less motivation to make the effort to travel across the border even if there is a relative price advantage.

To test whether specification problems are still present in the SLX model with parameterized W , it is re-estimated extended to include spatial autocorrelation. There are two potential problems that may cause misspecification. Relevant explanatory variables may have erroneously been omitted from the model and the true spatial weights matrix W^* that generated the data may be different from the spatial weights matrix W that is used to model exogenous interaction effects. The re-estimated model may be labeled an SDEM model, the difference being that the spatial weights matrix used to model spatial dependence in the disturbances is a binary contiguity matrix and not the parameterized inverse distance matrix used to model the exogenous interaction effects. The main reason to adopt a different matrix is methodological. If the spatial weight matrix used to model the exogenous interaction effects is still different from the true matrix, this misspecification in the exogenous interaction effects $(W-W^*)X$ is transmitted to the error term specification, as a result of which it loses its property of being distributed with $Var(\epsilon) = \sigma^2 I$. Instead, the error term specification will follow a spatial autoregressive process with spatial weights matrix V different from W and $Var(u) = \sigma^2 [(I - \lambda V)^T (I - \lambda V)]^{-1}$. Under the null of correct

specification, the coefficient estimates should be similar between the SLX and SDEM model. This can be tested using a Hausman test since both estimators of the response coefficients in the SLX and SDEM models are consistent but differ in efficiency. Pace and LeSage (2008) develop a Hausman test for OLS and SEM estimates, which can also be used for comparing SLX and SDEM estimates. We find that the test statistic amounts to 0.9267, which follows a chi-squared distribution with 4 degrees of freedom equal to the number of regression parameters under test. The corresponding p-value is 0.9207. Even though the spatial autocorrelation coefficient appears to be significant and the SDEM yields a significantly higher log-likelihood function value than SLX, the latter result implies that the null hypothesis of equality in SLX and SDEM coefficient estimates cannot be rejected. This indicates that a misspecified W or spatially dependent omitted variables do not represent a serious problem.

7. Conclusions

The *Journal of Regional Science* has recently published a much-discussed theme issue appraising spatial econometrics. Although some of the critiques raised are valid, they are issues that can be overcome by improving applied spatial econometric work. This paper points out that there should be a shift of focus in the modeling strategy for specifying a spatial econometric model. The commonly adopted procedure is to test the OLS model against the SAR and SEM models for an exogenously specified spatial weights matrix W . Instead, we propose that taking the SLX model using a W that is parameterized represents a better point of departure. This paper also emphasizes that there has been excessive reliance on statistical tests to infer the appropriate model specification and that instead, theory or the specific context of the empirical application should be the most important criterion for selecting either a local or global spillover model.

To explain the important distinction between local and global spillovers and the proposed modeling strategy, a concise overview of the spillovers corresponding to spatial econometric models with all different combinations of interaction effects is provided. By considering the partial derivatives, which as pointed out by LeSage and Pace (2009) represent a more valid basis for examining spatial spillovers, we show that some models are more flexible in modeling spillovers than others. Even though most applied work has relied on global spillover specifications, unless they are based on substantive theoretical arguments, they are more difficult to justify, as is also discussed in other studies (see e.g., Arbia and

Fingleton, 2008; Gibbons and Overman, 2012; Lacombe and LeSage, 2012). Although the SAR and SAC models are interesting from an econometric viewpoint, for empirical studies they are not suitable for investigating spatial spillovers since they impose restrictions on their magnitude in advance. Taking the SLX model as point of departure has the benefit that it produces spillover effects that are different for different explanatory variables and it allows for the parameterization of the spatial weights matrix W . This is a key contribution since one of the main criticisms of spatial econometrics is the a priori specification of W .

As an empirical application we use the well-known Baltagi and Li (2004) US state cigarette demand data to demonstrate the spatial spillovers resulting from different model specifications adopting an exogenously specified W . Although other studies have used this data set for illustrative purposes, no previous study has considered modeling only exogenous interaction effects. We also illustrate the proposed modeling strategy with a parameterized W , demonstrating that adopting a non-parameterized W may lead to wrong inferences. A notable result from the SLX estimation results is that when W is specified using the binary contiguity matrix and the inverse distance specification *without* parameterizing W , the price spillover effect estimate does not corroborate the existence of bootlegging. However, when W is specified using the *parameterized* inverse distance specification, there is significant evidence of the bootlegging effect. Furthermore, we test for remaining specification problems in the SLX model by applying a spatial Hausman test on the SLX and SDEM coefficient estimates. This test provides evidence that the SLX model is appropriate.

References

- Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Kluwer, Dordrecht.
- Anselin, L. (2002), "Under the Hood. Issues in the Specification and Interpretation of Spatial Regression Models," *Agricultural Economics* 27, 247-267.
- Anselin, L. (2003), "Spatial Externalities, Spatial Multipliers, and Spatial Econometrics," *International Regional Science Review*, 26, 153-166.
- Arbia, G. and B. Fingleton (2008), "New Spatial Econometric Techniques and Applications in Regional Science," *Papers in Regional Science*, 87, 311–317.
- Anselin, L., A.K. Bera, R. Florax, and M.J. Yoon (1996), "Simple Diagnostic Tests for Spatial Dependence," *Regional Science and Urban Economics*, 27, 77-104.
- Baltagi, B.H. (2008), *Econometric Analysis of Panel Data*, 4th edition, Wiley, Chichester.

- Baltagi, B.H. and D. Levin (1992), "Cigarette Taxation: Raising Revenues and Reducing Consumption," *Structural Change and Economic Dynamics*, 3(2), 321-335.
- Baltagi, B.H. and D. Li (2004), "Prediction in the Panel Data Model with Spatial Autocorrelation," In: Anselin, L., R.J.G.M Florax, and S.J. Rey (eds.), *Advances in Spatial Econometrics: Methodology, Tools, and Applications*, Springer, Berlin, 283-295.
- Burridge, P. and I. Gordon (1981), "Unemployment in the British Metropolitan Labour Areas," *Oxford Economic Papers*, 33(2), 274-297.
- Corrado, L. and B. Fingleton (2012), "Where is the Economics in Spatial Econometrics?" *Journal of Regional Science*, 52(2), 210-239.
- Debarsy, N., C. Ertur, and J.P. LeSage (2012), "Interpreting Dynamic Space-time Panel Data Models," *Statistical Methodology*, 9, 158-171.
- Elhorst, J.P. (2010), "Applied Spatial Econometrics: Raising the Bar," *Spatial Economic Analysis*, 5(1), 9-28.
- Elhorst, J.P. (2012), "Matlab Software for Spatial Panels," *International Regional Science Review* DOI: 10.1177/0160017612452429.
- Elhorst, J.P. (2013), "Spatial Panel Models," *Handbook of Regional Science*. Forthcoming.
- Ertur, C. and W. Koch (2007), "Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence," *Journal of Applied Econometrics*, 22, 1033-1062.
- Fischer, M.M., T. Scherngell, and E. Jansenberger (2006), "The Geography of Knowledge Spillovers Between High-Technology Firms in Europe: Evidence from a Spatial Interaction Modelling Perspective," *Geographical Analysis*, 38(3), 288-309.
- Fischer, M.M., T. Scherngell, and M. Reismann (2009), "Knowledge Spillovers and Total Factor Productivity: Evidence Using a Spatial Panel Data Model," *Geographical Analysis*, 41(2), 204-220.
- Gibbons, S. and H.G. Overman (2012), "Mostly Pointless Spatial Econometrics?" *Journal of Regional Science*, 52(2), 172-191.
- Kakamu, K. (2005), "Bayesian Estimation of a Distance Functional Weight Matrix Model," *Economics Bulletin*, 3(57), 1-6.
- Kirby, D.K. and J.P. LeSage (2009), "Changes in Commuting to Work Times over the 1990 to 2000 Period," *Regional Science and Urban Economics*, 39(4), 460-471.
- Kelejian, H.H. and G. Piras (2012), "Estimation of Spatial Models with Endogenous Weighting Matrices and an Application to a Demand Model for Cigarettes." Paper presented at the 59th North American Meetings of the RSAI, 2012, Ottawa, Canada.

Kelejian, H.H. and I.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.

Kelejian, H.H. and I.R. Prucha (1999), "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model," *International Economic Review*, 40, 509-533.

Kelejian, H.H. and I.R. Prucha (2010), "Specification and Estimation of Spatial Autoregressive Models with Autoregressive and Heteroskedastic Disturbances," *Journal of Econometrics*, 157, 53-67.

Kelejian, H.H., I.R. Prucha, and Y. Yuzefovich (2004), "Instrumental Variable Estimation of a Spatial Autoregressive Model with Autoregressive Disturbances: Large and Small Sample Results," In: LeSage, J.P. and K. Pace (eds.), *Spatial and Spatiotemporal Econometrics*, Amsterdam, Elsevier, 163-198.

Lacombe, D.J. and J.P. LeSage (2012), "Using Bayesian Posterior Model Probabilities to Identify Omitted Variables in Spatial Regression Models," Available at SSRN: <http://ssrn.com/abstract=2174610>.

Lee, L.F. (2004), "Asymptotic Distribution of Quasi-maximum Likelihood Estimators for Spatial Autoregressive Models," *Econometrica*, 72, 1899-1925.

Lee, L.-F., and J. Yu (2010), "Some Recent Developments in Spatial Panel Data Models," *Regional Science and Urban Economics*, 40, 255-271.

LeSage, J.P. and R.K. Pace (2009), *Introduction to Spatial Econometrics*, Taylor & Francis CRC Press, Boca Raton.

LeSage, J.P. and R.K. Pace (2011), "Pitfalls in Higher Order Model Extensions of Basic Spatial Regression Methodology," *The Review of Regional Studies*, 41(1), 13-26.

Manski, C.F. (1993), "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 60(3), 531-542.

McMillen, D.P. (2003), "Spatial Autocorrelation or Model Misspecification?" *International Regional Science Review*, 26(2), 208-217.

McMillen, D.P. (2010), "Issues in Spatial Data Analysis," *Journal of Regional Science*, 50, 119-141.

McMillen, D.P. (2012), "Perspectives on Spatial Econometrics: Linear Smoothing with Structured Models," *Journal of Regional Science*, 52(2), 192-209.

Pace, R.K., R. Barry, J.M. Clapp, and M. Rodriguez (1998), "Spatiotemporal Autoregressive Models of Neighborhood Effects," *Journal of Real Estate Finance and Economics*, 17(1), 15-33.

Pace, R.K. and J.P. LeSage (2008), "A Spatial Hausman Test," *Economics Letters*, 101, 282-284.

Pace, R.K. and S. Zhu (2012), "Separable Spatial Modeling of Spillovers and Disturbances," *Journal of Geographical Systems*, 14(1), 75-90.

Partridge, M.D., M. Boarnet, S. Brakman, and G. Ottaviano (2012), "Introduction: Whither Spatial Econometrics?" *Journal of Regional Science*, 52(2), 167-171.

Pinkse, J., and M.A. Slade (2010), "The Future of Spatial Econometrics," *Journal of Regional Science*, 50(1), 103-117.

Tobler, W.R. (1970), "A Computer Movie Simulating Urban Growth in the Detroit Region," *Economic Geography*, 46, 234-240.

Figure 1. Comparison of different spatial econometric model specifications

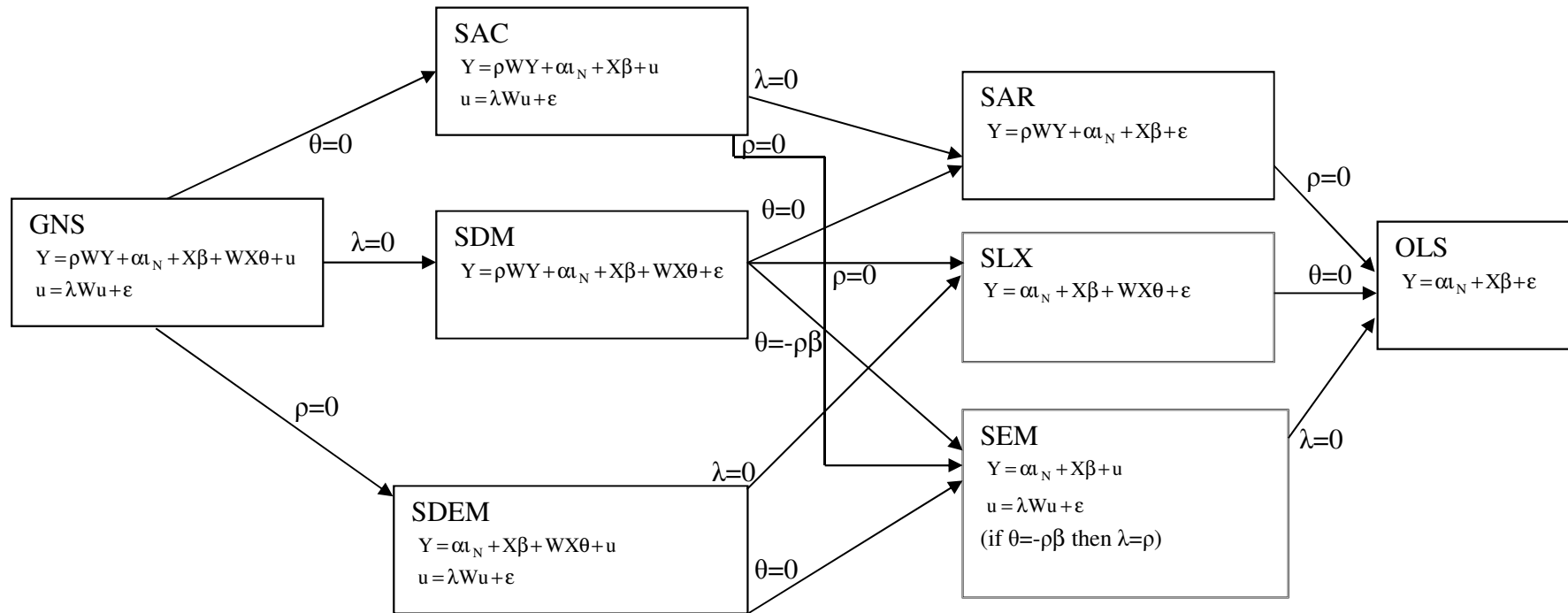


Table 1. Direct and spillover effects corresponding to different model specifications

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

Figure 2. Distance decay

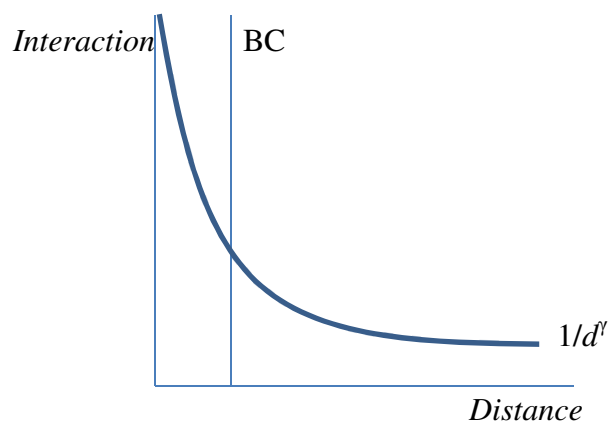


Table 2. Model comparison of the estimation results explaining cigarette demand

	OLS	SAR	SEM	SLX	SAC	SDM	SDEM	GNS
<i>Price</i>	-1.035 (-25.633)	-0.993 (-24.483)	-1.005 (-24.684)	-1.017 (-24.770)	-1.004 (-24.494)	-1.003 (-24.597)	-1.011 (-24.879)	-1.020 (-25.404)
<i>Income</i>	0.529 (11.668)	0.461 (9.857)	0.554 (11.069)	0.608 (10.381)	0.557 (10.506)	0.601 (10.329)	0.588 (10.568)	0.574 (11.018)
<i>W x Y</i>		0.195 (6.791)			-0.013 (-0.219)	0.225 (6.849)		-0.481 (-7.007)
<i>W x price</i>				-0.220 (-2.948)		0.051 (0.622)	-0.177 (-2.244)	-0.645 (-5.968)
<i>W x income</i>				-0.219 (-2.797)		-0.293 (-3.696)	-0.168 (-2.120)	0.079 (0.852)
<i>W x u</i>			0.238 (7.263)		0.292 (4.732)		0.229 (6.953)	0.628 (14.599)
R^2	0.896	0.900	0.895	0.897	0.895	0.901	0.897	0.873
Log-likelihood	1661.700	1683.474	1687.236	1668.400	1687.150	1691.373	1691.203	1695.123

Note: t-statistics are reported in parentheses.

Table 3. Model comparison of the estimated direct and spillover effects on cigarette demand

	OLS	SAR	SEM	SLX	SAC	SDM	SDEM	GNS
Direct effects								
<i>Price</i>	-1.035 (-25.633)	-1.003 (-25.099)	-1.005 (-24.684)	-1.017 (-24.770)	-1.004 (-24.469)	-1.016 (-24.841)	-1.011 (-24.879)	-0.999 (-25.433)
<i>Income</i>	0.529 (11.668)	0.465 (10.175)	0.554 (11.069)	0.608 (10.381)	0.556 (10.563)	0.594 (10.884)	0.588 (10.568)	0.594 (10.346)
Spillover effects								
<i>Price</i>		-0.232 (-5.626)		-0.220 (-2.948)	0.010 (0.172)	-0.215 (-2.388)	-0.177 (-2.244)	-0.122 (-1.894)
<i>Income</i>		0.107 (5.511)		-0.219 (-2.797)	-0.006 (-0.197)	-0.200 (-2.304)	-0.168 (-2.120)	-0.155 (-2.162)

Note: See note to Table 2.

Table 4. SLX model estimation results explaining cigarette demand and parameterization of W

	BC	ID ($\gamma=1$)	ID	ID + $\lambda V_{BC} u$
<i>Price</i>	-1.017 (-24.770)	-1.013 (-25.282)	-0.908 (-24.427)	-0.902 (-24.218)
<i>Income</i>	0.608 (10.381)	0.658 (13.726)	0.654 (15.392)	0.645 (14.697)
<i>W x price</i>	-0.220 (-2.948)	-0.021 (-0.335)	0.254 (3.083)	0.298 (3.943)
<i>W x income</i>	-0.219 (-2.797)	-0.314 (-6.627)	-0.815 (-4.758)	-0.819 (-6.567)
$V_{BC} x u$				0.164 (4.584)
γ			2.938 (16.478)	2.904 (21.361)
R^2	0.897	0.899	0.916	0.916
Log-likelihood	1668.4	1689.8	1812.9	1819.2

Notes: t-statistics are reported in parentheses and take into account the uncertainty in the γ estimate under the ID column; coefficient estimates of WX variables also denote spillover effects.