The SLX model: Extensions and the sensitivity of spatial spillovers to W

J. Paul Elhorst[a] and Solmaria Halleck Vega[a,b]

Elhorst J.P., Halleck Vega S.M (2017, English translation) The SLX model: Extensions and the sensitivity of spatial spillovers to W. *Papeles de Economía Española* 152: 34-50.

Abstract

The spatial lag of *X* model, a linear regression model extended to include explanatory variables observed on neighboring cross-sectional units, is the simplest spatial econometric model producing flexible spatial spillover effects and the easiest model to parameterize the spatial weights matrix, denoted by *W*, describing the spatial arrangement of the units in the sample. Nevertheless, it has received relatively little attention in the theoretical and applied spatial econometrics literature. This study fills this gap by considering several extensions of its basic form. It is found that the claim made in many empirical studies that their results are robust to the specification of *W* is not sufficiently substantiated. Especially the spatial spillover effects, often the main interest of spatial economic and econometric studies, turn out to be sensitive to the specification of *W*. In addition, it is found that the common practice to adopt the same *W* for every spatial lag should be rejected. These findings are illustrated using a cigarette demand model based on panel data of 46 U.S. states over the period 1963 to 1992.

Key words: Spatial econometric models, parameterizing *W*, spillovers, testing

JEL classification: C01, C21, C23, R15

---

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

[b]Paris School of Economics, Centre d'Economie de la Sorbonne, Université Paris 1 Panthéon-Sorbonne, 106/112 Blvd. de l'Hôpital, 75647 Paris Cedex 13, E-mail: solmaria.halleck@univ-paris1.fr

The authors gratefully acknowledge Jan Jacobs, Fei Jin and Alain Pirotte for their comments on a previous version of this paper.

1. Introduction

Two of the main interests in the spatial economics and spatial econometrics literature are spatial lags and spatial spillover effects. A spatial lag measures the impact of the dependent variable (Y), the explanatory variables (X) or the error term (u) observed in other cross-sectional units j than unit i on the dependent variable of unit i. A spatial spillover effect is defined as the marginal impact of a change to one explanatory variable in a particular cross-sectional unit on the dependent variable values in another unit, and is derived from the reduced form of a spatial econometric model. This spillover effect is a useful addition to the direct effect which measures the marginal impact of a change to one explanatory variable in a particular cross-sectional unit on the dependent variable of that unit itself.

The spatial econometrics literature distinguishes seven different types of static spatial econometric models. The differences between these models depend on the type and number of spatial lags that are included. Table 1 provides an overview, including their designations and abbreviations. The spatial weights matrix W symbolizes the spatial arrangement of the cross-sectional units in the sample. If there are N cross-sectional units, W is an $N \times N$ matrix describing which pairs of units are mutually connected to each other and which are not. If connected, the corresponding element is positive; if not, it is set to zero. Recently, Halleck Vega and Elhorst (2015) pointed out that the SAR, SEM and SAC models are of limited use in empirical research due to initial restrictions on the spillover effects they can potentially produce. In the SAR and SAC models the ratio between the spillover effect and the direct effect is the same for every explanatory variable, while in the SEM model the spillover effects are set to zero by construction. Only in the SLX, SDEM, SDM and GNS models can the spatial spillover effects take any value. Since the SLX model is the simplest one in this family of spatial econometric models, they recommend taking this model as point of departure when an empirical study focuses on spatial spillover effects and an underlying theory justifying any of the other models is lacking.

Insert Table 1

A concomitant advantage of the SLX model, in contrast to other spatial econometric models, is that W can be parameterized. A detailed explanation follows in Section 3 of this paper. Suppose a researcher wants to use a simple parametric approach applied to the elements of an inverse distance matrix $w_{ij} = 1/d^{\gamma}$, where γ is a parameter to be estimated, to obtain more information on the strength of the connections among the cross-sectional observations rather than to impose a certain specification of W in advance. A nonlinear, but straightforward estimation technique can then be used to estimate the parameters of the SLX model– the response parameters, given γ, and γ given the response parameters can be alternately estimated until convergence occurs. If γ appears to be small, observations at distant locations have relatively more impact than if this distance decay parameter appears to be large.

Notwithstanding these two advantages, there are several potential problems that may still cause misspecification. It might be that 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 spatial lags (WX). Instead of a parameterized inverse distance decay matrix, W may also be specified as a parameterized exponential distance decay matrix. Also, instead of one common γ for all spatially lagged explanatory variables, it is also possible to consider one separate γ_k for every variable WX_k (k=1,…,K), where K represents the number of X variables. Another option is to model the elements of W by a gravity type of function, i.e., including the size of units i and j in terms of population and/or gross product.

It may also be the case that the SLX model should still be extended to include a spatial lag among the dependent variable (WY) to obtain the SDM or a spatial lag among the error terms (Wu) to obtain the SDEM (see Table 1). If so, another interesting question is whether the spatial weights matrix of this additional spatial lag may be the same or should be different, and whether it can be parameterized as well. Finally, other potential forms of misspecification are erroneous omission of relevant explanatory variables from the model, heteroskedasticity, and endogeneity of some of the explanatory variables. The aim of this paper is to set out these extensions formally, discuss the rationale behind them, illustrate them in an empirical setting, and above all, to investigate the impact of these different modeling choices on the magnitude of spatial spillover effects. In view of this aim, the paper relates to the traditional theoretical econometric literature investigating the consequences of the omission of relevant explanatory variables, the introduction of irrelevant explanatory variables and the adoption of a misspecified W on the bias, consistency and efficiency of the parameter estimates, and related to this, the direct and spatial spillover effects. Many empirical spatial econometric studies consider only a few specifications of W and state that their results are robust to this specification. We will show that this claim is not sufficiently substantiated since spatial spillovers appear to be sensitive to the specification of W. Besides, we briefly discuss to which extent test procedures that have been suggested in the literature to deal with these specification issues are helpful.

The results derived in this study hold for both cross-sectional data and spatial panels, provided that the model specification is homogenous in its parameters. There is an upcoming literature of spatial econometric models in which the parameters are heterogeneous, i.e., are different for different units in the sample (Aquaro et al., 2015). Generally, the parameters of heterogeneous models can only be consistently estimated for spatial panels with a large time dimension (*T*). Whether the results derived in this paper also hold for these models is beyond the scope of the current paper, but note that most spatial panel data studies are characterized by a large number of observations in the cross-sectional domain (*N*) relative to *T*. Related to this, there is also an upcoming literature in which spatial econometric models are extended to include common factors (Bailey et al., 2016). The results derived in this study also hold for these type of models, since the common factors can be treated as additional control variables (Halleck Vega and Elhorst, 2016).

The setup of this paper is as follows. Section 2 sets out the SLX model and presents different parametric specifications of *W*. Even though the SLX model already covers *K* of the potential *K*+2 spatial lags, Section 3 deals with the question whether the SLX model should be further extended to contain one additional spatial lag, either among *WY* or *Wu*. Common test procedures developed in the spatial econometrics literature are considered: the (robust) Lagrange Multiplier (LM) tests, the general-to-specific approach, the J-test and the Bayesian comparison approach. In addition, Section 3 focuses on the question whether *W* of this additional spatial lag may be the same or should be different. Section 4 employs Baltagi and Li's (2004) cigarette demand model, covering 46 U.S. states over the period 1963-1992, to demonstrate the impact of these spatial weights matrices and potential extensions of the SLX model on the sign, magnitude and significance levels of spatial spillover effects in an empirical setting. Finally, Section 5 provides our main conclusions and suggests directions for further research.

2. 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 (Corrado and Fingleton, 2012). 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 criterions such as log-likelihood function values, Bayesian posterior model probabilities, and

J-tests. However, parameterizing W would be a further step forward. Halleck Vega and Elhorst (2015) show that the SLX model offers that opportunity. This model takes the form

$$Y = \alpha\iota_N + X\beta + WX\theta + \varepsilon, \tag{1}$$

where Y represents an $N \times 1$ vector consisting of one observation on the dependent variable for every unit in the sample ($i = 1, \dots, N$), 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 associated with the $K \times 1$ parameter vector β, 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. Since W is $N \times N$ and X is $N \times K$, the WX matrix of exogenous spatial lags is also $N \times K$. Consequently, the vector of response parameters θ is just like β of order $K \times 1$. The spatial spillover effects of this model coincide with the parameter estimates θ of the WX variables, and the direct effects with the parameter estimates β of the X variables.

Most studies dealing with geographical units adopt a binary contiguity matrix with elements $w_{ij} = 1$ if two units share a common border and zero otherwise (see Table 2 below). Arguments in favor of a binary contiguity matrix are given by Stakhovych and Bijmolt (2009). They find that first-order contiguity matrices perform better in detecting the true model than other spatial weights matrices. It should be stressed, however, that they only test the SAR and SEM specifications against each other. Furthermore, the common practice to adopt one particular spatial weights matrix can be quite arbitrary and it is preferable to compare different specifications with each other and to estimate the degree of distance decay.

Inspired by the notion that that the connectivity between nearby units will be stronger than those further away, the elements of W can be defined based on inverse distances with a distance decay factor

$$w_{ij} = \frac{1}{{d_{ij}}^{\gamma}}, \tag{2}$$

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 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 adopted binary contiguity principle is not an accurate representation of the spatial dependence because contiguity can be seen as a restrictive distance measure confining interaction between units only to those units that share borders.

An alternative specification to model distance decay is the negative exponential function

$$w_{ij} = \exp(-\delta d_{ij}). \tag{3}$$

The higher the estimate of the parameter δ implies that units at greater distances influence each other less, i.e. the cut-off point of interaction becomes smaller.

Inspired by Newton's law of universal gravitation, another option is to model the elements of the spatial weights matrix by a gravity type of function. One reason to adopt this function is its popularity in the literature explaining flow data. Another reason is that the elements of a spatial weights matrix should reflect the degree of spatial interaction between units, which is synonymous to the intensity of flows. Examples of studies pointing this out are Bavaud (1998) and Corrado and Fingleton (2012, Section 3), among others. Another advantage of the gravity model is that it is well embedded in the economic-theoretical literature (see Behrens et al., 2012, among others), the lack of which is one of the objections put forward against spatial econometric model studies (Corrado and Fingleton, 2012). A gravity type of function can be specified as

$$w_{ij} = \frac{P_i^{\gamma_1} P_j^{\gamma_2}}{d_{ij}^{\gamma_3}}, \tag{4}$$

where P measures the size of units i and j in terms of population and/or gross product.

Many studies initially assume that the spatial weights matrix is different for different spatial lags, but eventually they impose the restriction that each W is exactly the same in either their Monte Carlo simulation experiment or their empirical analysis, which is another objection often put forward against spatial econometric model studies (McMillen, 2012). Instead of (2), the elements of the spatial weights matrix of every exogenous spatial lag $W_k X_k$ may also be modeled as

$$w_{ijk} = \frac{1}{d_{ij}^{\gamma_k}}, \tag{5}$$

to investigate whether this restriction makes sense.

It should be stressed that theory should, preferably, be the driving force that determines the specification of the functional form of W, as is discussed extensively in

Corrado and Fingleton (2012). However, 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.

3. The extension of SLX to SDM or SDEM

3.1 The specifications of SDM and SDEM

To find out whether the SLX model adequately describes the spatial relationships among the data, it may be further tested against the SDM and the SDEM. Table 1 shows that the former model extends the SLX model to contain an endogenous spatial lag among the dependent variables (*WY*) and the latter model to contain a spatial lag among the error terms (*Wu*). Taking equation (1) as point of departure, the first extension takes the form

$$Y = \rho WY + \alpha\iota_N + X\beta + WX\theta + \varepsilon, \tag{6}$$

whose spatial spillover effects are known to be the off-diagonal elements of the $N \times N$ matrix

$$(I - \rho W)^{-1}(I_N\beta + W\theta), \tag{7}$$

and its direct effects the diagonal elements of that matrix. The second extension takes the form

$$Y = \alpha\iota_N + X\beta + WX\theta + u \quad \text{and} \quad u = \lambda Wu + \varepsilon. \tag{8}$$

whose spatial spillover and direct effects are respectively θ and β, just as in the SLX model. Another difference is that the spatial spillovers in the SDM are global, while in the SDEM they are local in nature (see Halleck Vega and Elhorst, 2015). Global spillovers occur when a change in *X* of any spatial unit is transmitted to all other units, also if two units according to *W* are unconnected. This is due to the pre-multiplication of the response parameters by the spatial multiplier matrix $(I - \rho W)^{-1}$, provided that the spatial autoregressive coefficient of *WY* is non-zero ($\rho \neq 0$). By contrast, local spillovers are those that affect other units only if according to *W* they are connected to each other. This requires that the coefficient of *WY* is zero ($\rho = 0$) and that the coefficients of *WX* are non-zero ($\theta \neq 0$). Furthermore, it should be realized that the choice between global and local spillovers is related to the specification of *W*. A global spillover model with a sparse *W*, in which only a limited number of elements are non-zero such as a binary contiguity matrix, is more likely than with a dense *W*. Conversely, a local spillover model with a dense *W*, in which all off-diagonal elements are non-zero such as an inverse distance or exponential distance decay matrix, is more likely than with a sparse *W*.

3.2 Tests for the extension of SLX to SDM or SDEM

Although one might use the classic and robust LM tests proposed by Anselin (1988) and Anselin et al. (1996) to test for the extension with either *WY* or *Wu*, 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, i.e., the model without any *WX* variables, and not on those of the SLX model. Generally, the log-likelihood function values of the SDM (*WY*+*WX*) and SDEM (*WX*+*Wu*) models are closer to each other than those between the SAR (*WY*) and SEM (*Wu*) models, as a result of which one often can no longer reject one model against the other. Conversely, the estimation of the GNS model (*WY*+*WX*+*Wu*) is also not of much help due to overfitting. Even though the parameters of this latter model can be estimated, they have the tendency either to blow each other up or to become insignificant, as a result of which it does not perform any better than the SDM and SDEM models and so does not help to choose among simpler models with less spatial lags, known as the general-to-specific approach (see Burridge et al, 2016, for a similar finding).

By contrast, LeSage (2014) demonstrates that a Bayesian comparison approach considerably simplifies the task of selecting an appropriate model specification, as well as the spatial weights matrix. This approach simultaneously determines the Bayesian posterior model probabilities of SDM and SDEM given a particular *W*, and the Bayesian posterior model probabilities of different *W* matrices given a particular model specification. These probabilities are based on the log marginal likelihood of the different model options, which are obtained by either integrating out the parameters of the model or by integrating out *W*. If the log marginal likelihood value of one model or of one *W* is higher than that of another model or another *W*, the Bayesian posterior model probability is also higher. Whereas the popular likelihood ratio, Wald and/or LM statistics only compare the performance of one model against another model based on specific parameter estimates within their parameter spaces, the strength of the Bayesian approach is that the posterior model probabilities are calculated over their entire parameter spaces, just because these parameters are integrated out. Inferences drawn on the log marginal likelihood function values are further justified because SDM and SDEM have the same set of explanatory variables ($[X\ WX]$) and are based on the same uniform prior for ρ or λ, related to *WY* or *Wu*. This prior takes the form $p(\rho) = p(\lambda) = 1/D$, where $D = 1/\omega_{max} - 1/\omega_{min}$ and ω_{max} and ω_{min} represent, respectively, the largest and the smallest (negative) eigenvalue of *W*. This prior requires no subjective information on the part of the practitioner as it relies on the parameter space $(1/\omega_{min}, 1/\omega_{max})$ on which ρ and λ are defined, where $\omega_{max} = 1$ if W is row-normalized or normalized by its largest eigenvalue. Full details on the choice of model can be found in LeSage (2014, 2015) and on the choice of *W* in LeSage and Pace (2009, chs. 5 and 6).

An alternative to the Bayesian comparison approach is the comparison approach based on J-tests. In their editorial to the first issue of the 11th volume of *Spatial Economic Analysis*, Elhorst et al. (2016) provide a brief overview of this relatively small but growing literature, comprising ten studies in total. However, they also conclude that the main bottleneck for practitioners is that up to now nobody has made their software code freely available. For this reason, the J-test is left aside in this study.

Another specification issue is that it is not clear whether the spatial weights matrix should be the same as the one used to model the exogenous spatial lags (*WX*). If the spatial weights matrix of these variables is specified as a binary contiguity matrix, it is common practice to use the same matrix when extending the model with either an endogenous spatial lag (*WY*) or a spatial lag among the error terms (*Wu)*. However, if *W* is parameterized, for example, with the distance decay parameter being estimated, it is more likely to also consider another distance parameter for *WY* than for *WX*, or for *Wu* than for *WX*. If this approach is followed, all the aforementioned tests fall short since they are based on the assumption that the *W* of *WY* or of *Wu* are exogenous and non-parameterized. Yu and Lee (2015) consider a spatial econometric model in which the elements of *W* are a function of explanatory variables which might be endogenous, but also this extension does not involve the estimation of unknown parameters which we use to parameterize *W*.

In sum, if it is found that the SLX model needs to be extended, there are two possible outcomes. Either the model needs to be extended with a spatial lag among the error terms, also known as spatial autocorrelation, or with an endogenous spatial lag. We discuss these two model extensions below.

3.3 SDEM: SLX plus spatial autocorrelation

If the SLX model needs to be extended with spatial autocorrelation, there is a strong methodological argument to adopt a different spatial weights matrix. Potentially, the *W* used to model the *WX* variables might still be different from the true matrix, say *W**. If so, this misspecification (*W*-*W**)*X* is transmitted to the error term specification, as a result of which it loses its property of being distributed with $Var(\varepsilon) = \sigma^2 I$. Instead, the error term specification will follow a spatial autoregressive process with spatial weights matrix *V* different from *W* and $Var(u) = \sigma^2[(I - \lambda V)^T(I - \lambda V)]^{-1}$. Further note that ignoring spatial autocorrelation, when relevant, only affects the efficiency of the parameter estimates and not their consistency, and that the W used to model spatial autocorrelation does not have any effect on the spatial spillovers derived from the reduced form of the spatial econometric model (see equation (8)). Consequently, the choice on which spatial weights matrix to adopt to model spatial autocorrelation is a less weighty decision. For this reason, we may follow Stakhovych and Bijmolt (2009) and initially consider a first-order contiguity matrix.

A Hausman test can be used whenever there are two consistent estimators, one of which is inefficient, while the other is efficient. Pace and LeSage (2008) develop this test for OLS and SEM estimates, which can also be used for comparing SLX and SDEM estimates. According to LeSage and Pace (2009, p. 62), rejection of the null hypothesis of equality in SLX and SDEM coefficient estimates can be useful in diagnosing the presence of omitted variables that are correlated with variables included in the model. The test statistic follows a chi-squared distribution with degrees of freedom equal to the number of regression parameters under test. Three different outcomes are possible. First, the SLX and SDEM coefficient estimates are not significantly different from each other and the spatial autocorrelation coefficient is not significant. When this occurs, the extension of the SLX model with spatial autocorrelation is not necessary and may be left aside. Second, the SLX and SDEM coefficient estimates are not significantly different from each other, but the spatial autocorrelation coefficient is significant. If this occurs, SDEM yields a significantly higher log-likelihood function value than SLX, as a result of which the conclusion must be that the spatial error term is still capturing the effect of omitted variables. However, since the null hypothesis cannot be rejected, it may also be concluded that these omitted variables are not correlated with the included variables and thus that the SDEM re-specification of the SLX model at the very most leads to an efficiency gain. Third, the SLX and SDEM coefficient estimates are significantly different from each other and the spatial autocorrelation coefficient is significant (the probability that the spatial autocorrelation coefficient will be insignificant here is negligible). This outcome points to serious misspecification problems due to omission of relevant explanatory variables.

One potential misspecification problem in the latter case might be that *W*, the binary contiguity matrix that was recommended as the first potential candidate, is too simple. To test for this, other specifications for *W* may be considered. Alternatively, one might parameterize this *W* as well. However, Halleck Vega and Elhorst (2015) point out that the parameterization of *W* using models other than the SLX model is hampered by the perfect solution problem.

They demonstrate that the solution $\rho = -1$, $\gamma = 0$, $\alpha = Y_1 + \cdots + Y_N$, and $\beta = 0$ fully fits the dependent variable in the SAR model, as well as the SEM and SDM models if in addition to this $\theta = 0$.[1] Since the log-likelihood of this perfect solution does not exist, it has been excluded formally. To prove consistency of the ML estimator of the SAR model, Lee (2004) shows that one of the following two conditions should be satisfied: (a) the row and column sums of the matrices W, $(I\text{-}\rho W)^{-1}$ and $(I\text{-}\lambda W)^{-1}$ 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) is originated by Kelejian and Prucha (1998, 1999).

If the elements of W take the form in (2), then $w_{ij} = 1$ provided that $\gamma = 0$. In that case all the off-diagonal elements of W equal one, 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*→1 as *N* goes to infinity. This implies that the spatial weights matrix that is obtained in case of the perfect solution should be formally excluded for reasons of consistency since neither condition (a) nor condition (b) is satisfied. However, to cover this problem when working with real data and fixed sample sizes, $N = \bar{N}$, numerically such that the optimum in computer software does not converge to the perfect solution is a challenge. The question is 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. Benjanuvatra and Burridge (2015) have investigated this, but they struggle with this problem too. Although they prove in the first theorem of their paper that the parameters, including the distance decay parameter, are identifiably unique if they are close to their true values provided that *N* goes to infinity, they face numerical problems to find these values computationally. In their Monte Carlo simulation experiments they report to find extreme, inaccurate or wild estimates due to the existence of local maxima. We therefore investigate this further for the SAR model below.

3.4 SDM: SLX plus endogenous spatial lag with parameterized *W*

Consider the distance decay specification of the spatial weights matrix in equation (2) and let γ_{WY} denote the distance decay parameter of the endogenous spatial lag and γ_{WX} of the exogenous spatial lags. The first question that needs to be answered is on which interval γ_{WY} is defined. The following exposition helps to define this interval based on the two

[1] The SAR model is a special case of the SDM specified in equation (6), obtained by setting $\theta = 0$. The SEM model can be rewritten as a constrained SDM model (θ=-$\rho\beta$). This explains why the perfect solution of the SDM also holds for the SEM.

aforementioned conditions (a) and (b), of which one needs to be satisfied to obtain a consistent ML estimator of the parameters of the model.

Consider an infinite number of spatial units that are arranged linearly. The distance of each spatial unit to its first left- and right-hand neighbor is d; to its second left- and right-hand neighbor, the distance is $2d$; and so on. Since the off-diagonal elements of W are of the form $1/d_{ij}^{\gamma_{WY}}$, where $d_{ij}=d$ is the distance between two spatial units i and j, each row sum is $2\times[1/d^{\gamma_{WY}}+1/(2d)^{\gamma_{WY}}+1/(3d)^{\gamma_{WY}}+\cdots]$, representing a series that is finite if $\gamma_{WY}>1$ (see Salas and Hille, 1990, p. 602).[2] This implies that condition (a) is satisfied for values of the distance decay parameter greater than unity and that the ML estimator will give consistent parameter estimates under these circumstances. Condition (b) is satisfied if the ratio between the row sum and N is smaller than unity if N goes to infinity. This ratio can be represented by $2\times[1/d^{\gamma_{WY}+1}+1/(2d)^{\gamma_{WY}+1}+1/(3d)^{\gamma_{WY}+1}+\cdots]$, which is finite for $\gamma_{WY}>0$.[3]

4. Spatial econometric model comparison: Empirical illustrations

4.1 The cigarette demand model

For empirical illustration, we use the well-known data set on cigarette demand of Baltagi and Li (2004). This data set has been used in other spatial econometric studies as well. It was used for the first time by Baltagi and Levin (1986, 1992), but then respectively over the periods 1963-1980 and 1963-1988. Table 2 provides an overview of different studies taken from Elhorst (2016) and shows the progress that has been made over the years. This table is extended with a recent study of He and Lin (2015). Most studies now control for spatial and time period fixed effects. Elhorst (2014) explicitly tests for these controls and finds that this model specification outperforms its counterparts without spatial and/or time fixed effects, as well as the random effects model. Many studies also include the dependent variable lagged in time to control for habit persistence. In that case one can distinguish both short-term and long-term direct and spatial spillover effects; Elhorst (2013) and Debarsy et al. (2014) provide the mathematical formulas of these effects. Most studies also share the view that exogenous spatial lags should be included, but whether it is the SDM or the SDEM specification that best describes the data is still unclear. Halleck Vega and Elhorst (2015) argue that including an endogenous spatial lag is

[2] In a continuous space, the row sum can be represented by the surface below the graph $f(d)=1/d^{\gamma}$ for $1\leq d<N$, where N goes to infinity. By calculating the integral of the $f(d)$-function over this interval, one obtains $1/(\gamma-1)\,(1-1/(N^{\gamma-1}))$, which is finite for $\gamma>1$ as N goes to infinity.

[3] The division of this row sum by N is equivalent to a division of $f(d)=1/d^{\gamma}$ by d, which can be represented by $g(x)=1/d^{(\gamma+1)}$. The surface below this graph is $1/\gamma\,(1-1/N^{\gamma})$, which is finite for $\gamma>0$ as N goes to infinity.

difficult to justify since it would mean that a change in price or income in a particular state potentially impacts consumption in all states, including states that according to W (such as California and Illinois) are unconnected. Finally, almost all studies adopt a row-normalized binary contiguity matrix. One exception is Debarsy et al. (2014) who also consider a row-normalized matrix based on state border miles in common between states. Recently, Kelejian and Piras (2014) and Halleck Vega and Elhorst (2015) are among the first going beyond an exogenous pre-specified spatial weights matrix with fixed weights.

Insert Table 2

In the cigarette demand model, real per capita sales of cigarettes (C_{it}, where i denotes one of 48 U.S. states and t =1963,…,1992) is regressed on the average retail price of a pack of cigarettes (P_{it}) and real per capita disposable income (I_{it}). In addition, variables are in logs. This rather basic equation can be obtained from maximizing a utility function depending on cigarettes and other consumer goods subject to a budget constraint (Chintagunta and Nair, 2011). The model is aggregated over individuals since the objective is to explain sales in a particular state. If the purpose would be to model individual behavior (e.g., the reduction in the number of smokers or teenage smoking behavior) then this is better studied using micro data. 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. Spatial and time fixed effects are controlled for based on test results in Elhorst (2014). The reason to consider this data set is because consumers may be expected to purchase cigarettes in nearby states, both legally and illegally, if there is a price advantage, known as the bootlegging effect.

4.2 Estimation results of the SLX model

The first column of Table 3 reports the estimation results when the SLX model is taken as point of departure and when using the row-normalized binary contiguity matrix, the second column when using the inverse distance with the distance decay parameter γ set equal to one in advance, the third column when the distance decay parameter is estimated, and the fourth column when the distance decay parameter is estimated but the distance decay function is not specified as an inverse but as a negative exponential function.

Row-normalizing a weights matrix based on inverse or exponential 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 between the elements in the different rows of *W* gets lost. We therefore scale the elements of *W* matrices based on distance by the maximum eigenvalue. Since the SLX model does not contain the spatial lag *WY*, the direct effects are similar to 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).

Insert Table 3

Whereas the direct effects of price and income across these four different specifications show a stable pattern, the spatial spillover effects show a changeable pattern. The direct effect of the price of cigarettes fluctuates closely around -1, and of income around 0.6. By contrast, whereas the price spillover effect is negative and strongly significant when adopting the binary contiguity matrix (-0.220, t-value -2.80), it becomes insignificant when adopting the inverse distance matrix (-0.021, t-value -0.34), and it changes sign, becomes significant and, above all, consistent with the bootlegging effect hypothesis when the distance decay parameter is estimated (0.254, t-value 3.08). The estimate of this distance decay parameter amounts to 2.938 and is also significant (t-value 16.48). This makes sense because only 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, whereas people living further from their state borders 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. This is corroborated by the R^2, which increases respectively from 0.897 to 0.899 and to 0.916, and the log-likelihood function value, which increases from 1668.2 to 1689.8 and to 1812.9. If the exponential rather than the inverse distance decay function is adopted, the picture changes again. Whereas the spatial price spillover effect and the distance decay parameter remain significant, the magnitude of the price spillover effect more than halves (0.129, t-value 1.80).

Just as for price, the income spatial spillover effect shows a changeable pattern. This effect increases in magnitude when moving from the binary contiguity matrix (-0.219, t-value

-2.80) to the inverse distance matrix (-0.314, t-value (-6.63), and next to the parameterized inverse distance matrix (-0.815, t-value -4.76), to change sign and to become weakly significant (90%) when adopting the exponential distance decay matrix (0.129, t-value 1.80). However, since the R^2 and the log-likelihood function value of the exponential distance decay matrix are substantially lower, this latter specification of the spatial weights matrix and this positive sign of the income spatial spillover effect should be rejected. Note that a negative income spatial spillover effect implies that an increase in own-state per capita income decreases cigarette sales in neighboring states. An explanation is that higher income levels reduce the necessity or incentive to purchase less expensive cigarettes elsewhere.

In view of the changeable patterns in the spatial spillover effects, it is interesting to consider the impact of other misspecification issues, where the main focus will be on the price spillover effect. If heteroskedasticity, as in column (5), endogeneity, as in column (6), are controlled for, or if one separate γ is considered for each explanatory variable, as in column (7), the sign and significance levels of the price and income spillover effects remain more or less similar to those for the parameterized inverse distance matrix, as in column (3), but their magnitudes still vary. We continue with a more detailed explanation below.

Anselin (1988) advocated the incorporation of heteroskedastic disturbances in spatial econometric models over twenty-five years ago. Since cross-border consumption might be affected by differences in population size among states, we control for heteroskedasticity of known functional form.[4] Following Anselin (1988, pp. 34-35), we specify $\sigma_{it} = \alpha_1 + \alpha_2(1/\text{Population}_{it})$ and estimated this model by an iterative two-step procedure in which the αs and γ on the one hand and the βs on the other hand are alternately estimated until convergence occurs. These results are reported in column (5) of Table 3. A test for reduction to homoskedasticity is a test of the hypothesis that α_2=0, and therefore has one degree of freedom. The likelihood ratio test statistic is equal to 2(1841.3-1812.9)=56.8, which is highly significant if treated as χ_1^2 under the null. It signals that spatial econometricians should devote more attention to accounting for heteroskedasticity. When doing so, the price spillover effect decreases from 0.254 to 0.182.

Another issue is whether or not cigarette prices are endogenous. For this purpose, the Hausman test for endogeneity in combination with tests for the validity of the instruments are performed, namely to assess if they satisfy the relevance and exogeneity criterions. As instrumental variables, the cigarette excise tax rate, state compensation per employee in

[4] Kelejian and Prucha (2010) consider a nonparametric approach to account for heteroskedasticity.

neighboring states, and income and income in neighboring states (the latter two which are already part of the SLX model) are used. Details can be found in Halleck Vega and Elhorst (2015), who find that the price of cigarettes observed in neighboring states may be used as an exogenous determinant of cigarette demand in the U.S., whereas the price of cigarettes observed in the own state may not. Apparently, consumption has feedback effects on the price in the own state, but if consumers decide to buy more cigarettes in neighboring states due to a price increase in their own state this has no significant feedback effects on prices there too. When accounting for endogeneity, the price spillover effect decreases from 0.254 to 0.192.

Finally, when instead of one common distance decay parameter holding for all explanatory variables, a separate distance decay parameter is estimated for every explanatory variable, the price spillover effect increases from 0.254 to 0.385.

In contrast to these three model extensions above, the sensitivity of the price and income spatial spillover effects increases again when the elements of the spatial weights matrix are parameterized by the gravity model (see column (8)). Although the spatial spillover effect of the price of cigarettes remains positive, it is no longer significant, whereas the spatial spillover effect of income compared to the previous models in columns (3), (5) and (6), though still negative and significant, more than halves.

4.3 Testing for the extension of SLX to SDM or SDEM

Table 3 also reports the results of LeSage's (2014) Bayesian comparison approach and Anselin et al.'s (1996) robust LM tests to investigate statistically whether a local spillover model specification is more likely than a global one. We also consider two versions of the robust LM tests[5] to test for the extension of the SLX model with an endogenous spatial lag WY and for a spatial lag among the error terms Wu; one in which W is the same as the W used to model the exogenous spatial lags, and one where this W is specified as a binary contiguity matrix. In Section 3 we provided arguments in favor of the last option. The robust tests are based on the residuals of the SLX model and follow a chi-squared distribution with one degree of freedom; the critical value amounts to 3.84 at the five percent significance level.

Judging by the Bayesian comparison approach, the global spillover specification (SDM) initially seems to be more likely than the local one when adopting the rejected binary contiguity specification of W; the corresponding Bayesian posterior model probabilities of this

[5] Besides the robust, we also calculated the classic LM test statistics. Since the latter pointed to both extensions rather than one, they are not helpful in choosing between them. For this reason, they are not reported. The employed tests are called robust because they test for the existence of one type of spatial dependence in the local presence of the other.

sparse matrix are 0.55 to 0.45. However, when the pre-specified or parameterized inverse distance matrix is adopted, which are both dense matrices, this proportion changes into 0 to 1. Finally, when adopting the rejected parameterized exponential distance matrix, this proportion reduces to 0.35 to 0.65. Hence, the SDEM seems the most appropriate extension of the SLX model.

This picture changes again when employing the robust LM tests. When the rejected binary contiguity matrix is adopted, the LM tests do not indicate which extension is more appropriate since both test statistics appear to be insignificant. When the inverse or parameterized exponential distance decay matrices are adopted, the LM tests point to the SDEM specification, just as the Bayesian comparison approach. However, when the parameterized inverse distance matrix is taken as point of departure, a different result is obtained. The Bayesian comparison approach points to SDEM, while the LM tests to SDM. For this reason, we consider both model extensions below.

4.4 Estimation results of SDEM

Column (1) of Table 4 reports the results of the SLX model extended to include the spatial lag among the error terms. For reasons pointed out in Section 2, the binary contiguity matrix is used for this purpose rather than the parameterized inverse distance matrix which is used to model the exogenous spatial lags. The Hausman test statistic based on Pace and LeSage (2009, section 3.3.1) 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. Since the Hausman test whether the SLX and SDEM estimates are the same cannot be rejected, but the spatial autocorrelation coefficient nonetheless appears to be significant (0.164 with t-value 4.58), the conclusion must be that the SDEM re-specification of the SLX model leads (at the very most) to an efficiency gain.

Insert Table 4

4.5 Estimation results of SDM

Column (2) of Table 4 reports the results of the SLX model extended to contain the spatial lag among the dependent variables, where W just as in the SDEM specification is specified as a binary contiguity matrix. The log-likelihood function value of this model (1821.7) is somewhat greater than that of its counterpart the SDEM specification (1819.2), which is in line with the LM test results reported in column (3) of Table 3.

Column (3) of Table 4 again reports results of the SLX model extended to include the spatial lag among the dependent variables. However, instead of *a priori* imposing a binary contiguity matrix for the *W* of *WY*, in the third column we also specify this *W* using inverse distances between the states and estimate the decay parameter γ_{WY}. The parameterization of *W* using models with an endogenous spatial lag is hampered by the perfection solution problem, as mentioned in Section 3. Figure 1 shows the log-likelihood function for the distance decay parameter pertaining to the *W* modeling the endogenous spatial lag, where γ_{WY} takes values in the interval (0, 10] using increments of 0.1. The perfect solution problem can be observed when the decay parameter approaches the value γ_{WY}=0. Another important observation from Figure 1 is that the function is not concave, also when considering values of γ_{WY}>1 only. Although it has a maximum of approximately 5.54, there are many other possible values of γ_{WY} that are near this maximum. We come back to this significant finding shortly.

Insert Figure 1

Comparing the SDEM and SDM(a), there are large differences in the spillover effects. In fact, although all the spillover effects are highly significant under both specifications, they have opposite signs! Whereas the price spillover in the SLX model is positive corroborating the bootlegging effect and the income effect is negative, these spillover effects change sign in the SDM(a) model if the *W* of *WY* is specified as a binary contiguity matrix. In other words, although the performance of the model improves statistically in terms of the R^2 and the log-likelihood function value, it is not an improvement from an economic-theoretical viewpoint. A better performance, in both domains, is obtained in the case of the SDM(b) model. The spillover effects, reported under (b1) in column 3 of Table 4, have the same sign as those in the SDEM model and the significance levels are comparable One difference is that the magnitude of the income spillover in the SDM(b) model is much lower (in absolute value) than in the SDEM model, -0.125 versus -0.372, which in turn are much lower than the value of -0.815 in the SLX model. Finally, the point estimate of ρ in the SDM(b) model is negative and highly significant, thereby providing additional evidence in favor of competition effects, next to price and income level. Overall, these results again confirm that the search for the right spatial econometric model and the right functional form of *W* may have significant effects on the magnitude and significance levels of the spillover effects.

It should be emphasized that the spillover effects resulting from the SDM(b) model are conditional on the exogeneity of *W*, i.e. the effects estimates are calculated with a fixed value for the decay parameter γ_{WY}, taking a value of 5.54 in our application. This value is much higher than that for γ_{WX}, while the t-statistic of 0.971 is much lower. Yet, as can be observed in Figure 1, there is large uncertainty on how the *W* of *WY* should be specified. This uncertainty can be a reason why many applied studies find their results to be robust to changes in *W*, as illustrated in Figure 1 where the log-likelihood function hardly changes value if $\gamma_{WY} > 4$. However, what is often taken as a strong point may actually be a weak point. By showing that the results- generally the point estimates of the response parameters- are robust to different specifications of *W*, the researcher is more or less saying that they are uncertain about the true specification of *W*, while we have seen that the impact of different specifications of *W* on the spillover effects might be substantial.

To demonstrate the importance of the uncertainty about specifying the appropriate *W* of *WY*, we simulated the effects estimates of SDM(b), but then also accounting for the variance in γ_{WY} rather than treating it as fixed. These effects estimates are reported under (b2) in column 3 of Table 4. Since the parameter estimates are the same as in SDM(b), these results are not repeated. The direct effects are again relatively constant. The key difference is observed in the spillover impacts (both for the price and income variables), which are small and not different from zero when the distance decay parameter is not taken as fixed. Therefore, what is observed from this application to state-level cigarette demand is that there are significant spillover effects, but conditional on the exogeneity of *W*. That is, we can only claim there is evidence of spillovers (e.g., bootlegging behavior) if we are willing to accept that *W* is exogenous. The fact that the log-likelihood function for γ_{WY} (Figure 1) is not strictly concave and rather flat points to identification problems since there is not a clear unique global maximum. It thus seems that spatial econometric models containing an endogenous spatial lag (i.e. global spillovers) are more difficult to identify.

Finally, Figure 1 shows that limiting the interval of the distance decay parameter to for example, [1, 2] or [1, 4] can be quite restrictive. The true decay parameter may be outside the fixed interval defined by the researcher, as in the cigarette demand results where the estimated parameter is quite large compared to values that are normally used in applied research.

5. Conclusion

Recently, Halleck Vega and Elhorst (2015) recommended taking the SLX model as point of departure when a study focuses on spatial spillover effects and an underlying theory justifying alternative models is lacking. It is the simplest spatial econometric model producing flexible spillover effects and, in contrast to other spatial econometric models, the spatial weights matrix *W* in the SLX model can easily be parameterized.

More than a decade ago, Leenders (2002) demonstrated that the chosen specification of *W* is of vital importance when estimating the SAR or SEM models since both the value and the significance level of the spatial lag parameter of *WY* or *Wu* depends on the specification of *W*. This study corroborates Leender's conclusions, but then according to latest insights. Instead of focusing on models with just one spatial lag, such as SAR or SEM, present studies adopt models with multiple spatial lags, among which is the SLX model. Rather than addressing the parameter estimates of the spatial lags, present studies are considering spatial spillover effects. The SLX model has the property that the parameter estimates of the exogenous spatial lags (*WX*) coincide with the spatial spillover effects, a property which may be another reason to take this model as point of departure in an empirical study.

In this paper, we have seen that the sign, magnitude, and significance level of the spillover effects are sensitive to both the specification of *W* and the spatial econometric model specification using the well-known Baltagi and Li (2004) cigarette demand model based on panel data of 46 U.S. states over the period 1963-1992. The first has been tested by estimating the SLX model for six different specifications of *W*, varying from a binary contiguity matrix to a parameterized gravity type of model reflecting the intensity of flows among spatial units. The second has been tested by controlling for heteroskedasticity and endogeneity, and by extending the model to include either a spatial lag among the error terms or among the dependent variables. The price spillover effect ranged from a negative and significant value of -0.102 to a positive and significant value of 0.417. Similarly, the income spillover effect ranged from a negative and significant value of -0.901 to a positive and weakly significant (90%) value of 0.129. The claim made in many empirical studies that their results are robust to the specification of *W* should thus be more sufficiently substantiated. It might be that these studies mainly focus on the direct effects. It has been found that the direct effects of price and income fluctuate around respectively -1 and 0.6, no matter how the spatial weights matrix and the spatial econometric model are specified. It might also be that the log-likelihood function value hardly changes for a specific range of the distance decay parameter when *W* is parameterized, as shown in the last part of Section 4. The conclusion must be that studies checking whether their results are robust to the specification of the spatial weights matrix should better focus on the spillover effects and, in addition, better consider spatial weights matrices that are different in nature. Spatial weights matrices that are similar in nature, such as an inverse distance matrix for different values of the distance decay parameter or *k*-nearest neighbor matrices for different values of *k*, are less useful.

Test statistics that help to find out whether the SLX model should be further extended to SDM or SDEM are still in its infancy because the SLX model has received relatively little attention in both the theoretical and applied spatial econometrics literature. The performance of the well-known and widely used (robust) LM tests has only been investigated starting from the OLS model, but not from the SLX model. It might explain why the robust LM tests, in contrast to the Bayesian comparison approach, points to the SDM model in our empirical application rather than the SDEM model when taking the SLX model with the parameterized inverse distance matrix as point of departure. However, when this SDM model is subsequently estimated, it turns out to be no improvement from an economic viewpoint, only from a statistical viewpoint. One explanation why the Bayesian comparison approach points to the SDEM model and performs better in this respect is that it compares the performance of one model against another on the entire parameter space (since these parameters are integrated out), while the popular (robust) LM tests only compare the performance of one model against another model based on specific parameter estimates within the parameter space. Notwithstanding this advantage, there is room for improvement; the Bayesian comparison approach assumes that the spatial weights matrix is the same for every spatial lag, while we found empirical evidence in favor of the proposition that these matrices should be different. In sum, these findings point to interesting directions for future research.

Despite these uncertainties, we may conclude to have found empirical evidence in favor of the bootlegging effect. A price increase of one percent in one state causes a shift in consumption to other states of approximately 0.1-0.3 percent. In addition, we may say that an income increase of one percent in one state diminishes consumption in neighboring states in the range of 0.2-0.9 percent. These ranges are smaller than the full ranges reported above since the outcomes found outside these two ranges could be rejected largely on statistical grounds and partly on economic grounds.

References

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

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.

Aquaro, M., N Bailey and M.H. Pesaran (2015) *Quasi maximum likelihood estimation of spatial models with heterogeneous coefficients*, USC-INET Research Paper No. 15-17. Available at SSRN: http://ssrn.com/abstract=2623192, http://dx.doi.org/10.2139/ssrn.2623192.

Bailey, N., S. Holly and M.H. Pesaran (2016), "A two-stage approach to spatio-temporal analysis with strong and weak cross-sectional dependence", *Journal of Applied Econometrics*, 31, 249-280.
Baltagi, B.H. and D. 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.
Baltagi, B.H. and D. Levin (1992), "Cigarette taxation: raising revenues and reducing consumption", *Structural Change and Economic Dynamics*, 3, 321-335.
Baltagi, B.H. and D. Li (2004), "Prediction in the panel data model with spatial autocorrelation", in: Anselin, L., R. Florax, and S.J. Rey (eds.), *Advances in spatial econometrics: methodology, tools, and applications*. Berlin, Springer, 283-295.
Bavaud, F. (1998), "Models for spatial weights: a systematic look", *Geographical Analysis*, 30, 153-171.
Behrens, K., C. Ertur and W. Koch (2012), "'Dual' gravity: using spatial econometrics to control for multilateral resistance", *Journal of Applied Econometrics*, 27, 773-794.
Benjanuvatra, S. and P. Burridge (2015) *QML estimation of the spatial weight matrix in the MR-SAR model*. York, DERS University of York working paper.
Blundell, R. and T.M. Stoker (2007), "Models of aggregate economic relationships that account for heterogeneity", in: Heckman, J.J. and E. Leamer (eds.), *Handbook of econometrics*, Vol. 6A. Amsterdam, Elsevier, 4609-4663.
Burridge, P., J.P. Elhorst and K. Zigova (2016), "Group interaction in research and the use of general nesting spatial models", in: Pace, K., J.P. LeSage and B. Baltagi (eds.) *Advances in econometrics, volume 37, Spatial and Spatiotemporal Econometrics.* Amsterdam, Elsevier, Forthcoming.
Chintagunta, P.K. and H.S. Nair (2011), "Discrete-choice models of consumer demand in marketing", *Marketing Science* 30, 977–996.
Corrado, L. and B. Fingleton (2012), "Where is the economics in spatial econometrics?", *Journal of Regional Science*, 52, 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. (2005), "Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels", *Geographical Analysis*, 37, 85-106.
Elhorst, J.P. (2013), "Spatial panel models", in: Fischer, M.M and P. Nijkamp (eds.) *Handbook of regional science.* Berlin, Springer, 1637-1652.
Elhorst, J.P. (2014), "Matlab software for spatial panels", *International Regional Science Review*, 37, 389–405.
Elhorst, J.P. (2016), "Spatial Panels", in: Shekhar, S. and H. Xiong (eds.) E*ncyclopedia of GIS*, 2nd Edition. Berlin, Springer. Forthcoming.
Elhorst, J.P., M. Abreu, P. Amaral, A. Bhattacharjee, L. Corrado, B. Fingleton, F. Fuerst, H. Garretsen, D. Igliori, J. Le Gallo, P. McCann, V. Monastiriotis, G. Pryce and J. Yu (2016), "Editorial: raising the bar (1)", *Spatial Economics Analysis*, 11, 1-6.
Halleck Vega, S. and J.P. Elhorst (2015), "The SLX model", *Journal of Regional Science*, 55, 339-363.
Halleck Vega, S.M., Elhorst J.P. (2016), "A regional unemployment model simultaneously accounting for serial dynamics, spatial dependence and common factors", *Regional Science and Urban Economics*, forthcoming.
He, M. and K.-P. Lin (2015), "Testing random effects panel data models with spatially autocorrelated error components and spatially lagged dependent variables", *Econometrics* (open access journal) 3, 761-796.

Kelejian, H.H. and G. 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, 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.

Lee, L.-f. (2004), “Asymptotic distribution of quasi-maximum likelihood estimators for spatial autoregressive models”, *Econometrica*, 72, 1899–1925.

Leenders, R.T.A.J. (2002), “Modeling social influence through network autocorrelation: Constructing the weight matrix”, *Social Networks*, 24, 21-47.

LeSage, J.P. (2014), “Spatial econometric panel data model specification: a Bayesian approach”, *Spatial Statistics*, 9, 122-145.

LeSage, J.P. (2015), “Software for Bayesian cross section and panel spatial model comparison”, *Journal of Geographical Systems*, 17, 297-310.

LeSage, J.P. and R.K. Pace (2009) *Introduction to spatial econometrics*. Boca Raton, FL, Taylor and Francis.

McMillen, D.P. (2012), “Perspectives on spatial econometrics: linear smoothing with structured models”, *Journal of Regional Science*, 52, 192–209.

Pace, R.K. and J.P. LeSage (2008), “A spatial Hausman test”, *Economics Letters*, 101, 282-284.

Salas, S.L. and E. Hille (1990) *Calculus*, 6th edition. New York, Wiley.

Stakhovych, S. and T.H.A. Bijmolt (2009), “Specification of spatial models: a simulation study on weights matrices”, *Papers in Regional Science*, 88, 389-408.

Yu, J. and L.F. Lee (2015), “Estimating a spatial autoregressive model with an endogenous spatial weight matrix”, *Journal of Econometrics*, 184, 209-232.

Table 1 Spatial econometric models with different combinations of spatial lags and their flexibility regarding spatial spillovers

Type of model	Spatial lag(s)	Flexibility spatial spillovers
SAR, Spatial autoregressive model*	*WY*	Constant ratios
SEM, Spatial error model	*Wu*	Zero by construction
SLX, Spatial lag of *X* model	*WX*	Fully flexible
SAC, Spatial autoregressive combined model**	*WY*, *Wu*	Constant ratios
SDM, Spatial Durbin model	*WY*, *WX*	Fully flexible
SDEM, Spatial Durbin error model	*WX*, *Wu*	Fully flexible
GNS, General nesting spatial model	*WY*, *WX*, *Wu*	Fully flexible

* Also known as the spatial lag model, ** Also known as the SARAR or Cliff-Ord model

Table 2 Spatial panel data studies on cigarette demand

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

Panel: SFE = spatial fixed effects, SRE = spatial random effects, TFE = time fixed effects, TRE = time random effects; Dynamic: + = Y_{t-1} included; Spatial: see Table 1 for abbreviations, All = SAR, SEM, SLX, SAC, SDM, SDEM, GNS; *W*: BC = binary contiguity matrix, ID = inverse distance matrix. Source: Elhorst (2016).

Table 3 SLX model estimation results explaining cigarette demand and the parameterization of *W*

	BC	ID (γ=1)	ID	ED	ID Hetero-skedasticity*	2SLS	ID γ_k's	ID Gravity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Price	-1.017	-1.013	-0.908	-1.046	-0.884	-1.246	-0.903	-0.841
	(-24.77)	(-25.28)	(-24.43)	(-29.58)	(-24.93)	(-16.32)	(-24.49)	(-23.03)
Income	0.608	0.658	0.654	0.560	0.716	0.591	0.667	0.641
	(10.38)	(13.73)	(15.39)	(15.44)	(17.03)	(13.34)	(15.76)	(15.16)
W×Price	-0.220	-0.021	0.254	0.108	0.182	0.192	0.385	0.041
	(-2.95)	(-0.34)	(3.08)	(2.08)	(3.02)	(3.00)	(1.81)	(0.87)
W×Income	-0.219	-0.314	-0.815	0.129	-0.728	-0.750	-0.838	-0.372
	(-2.80)	(-6.63)	(-4.76)	(1.80)	(-13.21)	(-14.14)	(-5.21)	(-4.97)
$\gamma_{distance}$			2.938	0.467	2.966	3.141		2.986
			(16.48)	(9.99)	(24.39)	(11.11)		(11.50)
$\gamma_{distance\ price}$ in col.(7) and $\gamma_{own\ population}$ in col. (8)							5.986	-0.018
							(8.86)	(-0.41)
$\gamma_{distance\ income}$ in col.(7) and $\gamma_{population\ neighbors}$ in col.(8)							2.938	0.340
							(17.70)	(2.63)
R^2	0.897	0.899	0.916	0.896	0.906	0.484	0.917	0.923
LogL	1668.4	1689.8	1812.9	1666.9	1841.3		1818.4	1868.0
LM test WY, W§	-	0.27	27.56	0.47				
LM test Wu, W§	-	12.39	4.16	0.22				
LM test WY, W=BC	0.30	0.40	5.13	0.72				
LM test Wu, W=BC	0.01	12.67	0.14	11.59				
Prob. SDM	0.5502	0.0000	0.0000	0.3536				
Prob. SDEM	0.4498	1.0000	1.0000	0.6464				

Notes: State and time fixed effects are controlled for in all specifications; t-statistics in parentheses; coefficient estimates of *WX* variables in the SLX represent spillover effects.

*We specified $\sigma_{it} = \alpha_1 + \alpha_2(1/Population_{it})$ and found α_1=0.003 (t-value 14.14) and α_2=1.971 (t-value 7.11).

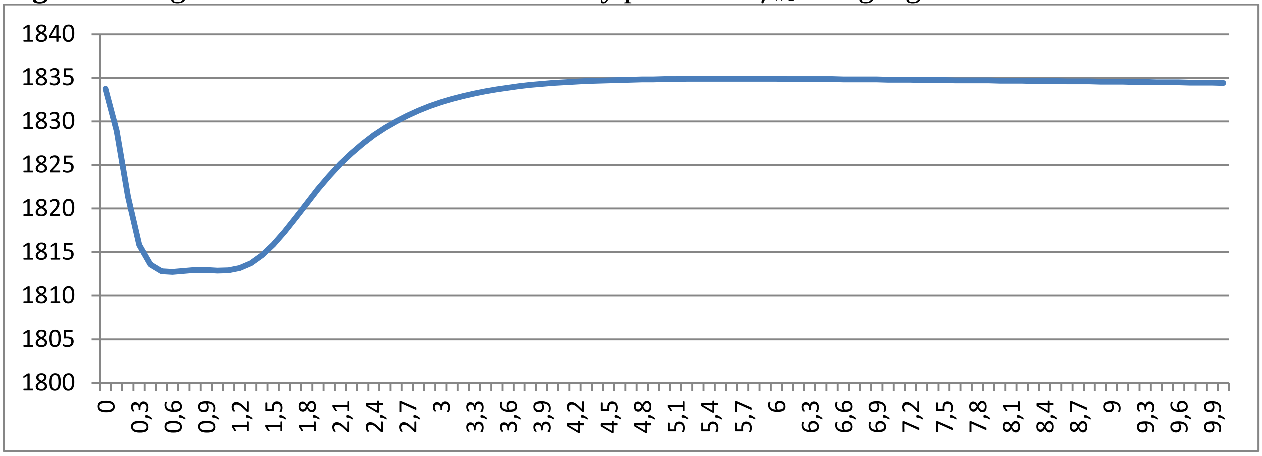
§*W* matrix similar to the one used to model exogenous spatial lags *WX*.

Table 4 Beyond the SLX model with parameterized W: SDEM and SDM

	SDEM	SDM(a)	SDM(b)	
	ID+$\lambda W_{BC}u$	ID+$\rho W_{BC}Y$	ID+ $\rho W_{ID}(\gamma_{WY})Y$	
	(1)	(2)	(3)	
Point estimates				
Price	-0.841	-0.881	-0.894	
	(-23.03)	(-23.13)	(-24.60)	
Income	0.641	0.588	0.665	
	(15.16)	(13.35)	(16.27)	
W×Price	0.041	0.288	0.076	
	(0.87)	(4.68)	(1.30)	
W×Income	-0.372	-0.804	-0.932	
	(-4.97)	(-16.04)	(-17.54)	
W×u	0.164	-	-	
	(4.58)			
W×Y	-	0.143	-0.208	
		(5.19)	(-3.78)	
γ_{WX}	2.904	3.134	3.035	
	(21.36)	(12.61)	(15.43)	
γ_{WY}	-	-	5.540	
			(0.971)	
Direct effects			(b1)	(b2)
Price	-0.841	-0.885	-0.903	-0.896
	(-23.03)	(-23.60)	(-23.87)	(-22.77)
Income	0.641	0.592	0.671	0.668
	(15.16)	(13.66)	(16.18)	(15.36)
Spillover effects				
Price	0.041	-0.102	0.162	0.028
	(0.87)	(4.43)	(4.32)	(0.24)
Income	-0.372	-0.019	-0.125	-0.025
	(-4.97)	(-4.37)	(-4.40)	(-0.29)
R^2	0.916	0.918	0.918	
LogL	1819.2	1821.7	1868.0	

Notes: State and time fixed effects are controlled for in all specifications; t-statistics are reported in parentheses. W of WX variables parameterized by $w_{ij}=1/d_{ij}^{\gamma}$ with γ estimated. 1) SDEM=SLX extended with Wu where W is binary contiguity matrix. 2) SDM=SLX extended with WY where W is binary contiguity matrix. 3) SDM=SLX extended with WY where W is parameterized inverse distance matrix, $w_{ij}=1/d_{ij}^{\gamma}$, and different γ parameters for WY and WX (see Section 4 for details).

Figure 1. Log-likelihood function for the decay parameter γ_{WY} using cigarette demand data set



Note: Values of γ_{WY} are in the interval (0, 10] using increments of 0.1.