X Marks the Spot: Unlocking the Treasure of Spatial-X Models

Cameron Wimpy, Arkansas State University
Guy D. Whitten, Texas A&M University
Laron K. Williams, University of Missouri

In recent years, political scientists have made extensive use of spatial econometric models to test a wide range of theories. In a review of spatial papers, we find that a majority of these studies use the spatial autoregressive (SAR) model. Although this is a powerful method that reveals inferences about diffusion processes, it is also highly restrictive and makes assumptions that often are not appropriate given the expressed theories. We contend that spatial-X (SLX) models are a better reflection of typical theories about spatial processes. Our simulations demonstrate that SLX models consistently retrieve the direct and indirect effects of covariates when the true data-generating process reflects other spatial processes. SAR models, however, tend to find phantom higher-order effects that are not present in the data. We further demonstrate how SLX models reveal heterogeneity in patterns of spatial dependence in countries' defense burdens that SAR models cannot discover.

Political scientists have increasingly made use of spatial econometric models. The appeal of these models is that they relax rigid assumptions about the independence of observations across space. While the move toward these techniques is encouraging for the building of more realistic models of politics, there is frequently a substantial disjuncture between theoretical propositions, what is actually being tested, and how results are interpreted. This has particularly been the case when researchers rely only on the spatial autoregressive (SAR) model to test their theories.

The SAR model captures contemporaneous interdependence in outcomes—how the value of the dependent variable in one unit, y_i, affects the value of the dependent variable in another unit, y_j. Although the SAR model is the most popular spatial model among political scientists, it is a highly restrictive model that researchers should use with caution. We argue that, depending on the theory being tested and the pattern of spatial effects, the spatial-X (SLX) model may be preferable to the SAR. This is the case because the SLX model allows the spatial processes to influence the outcome through one or more independent variables. Examples of this include trash in a neighbor's yard affecting the value of your home and lax gun regulations in nearby states producing negative externalities (Knight 2013). Since there is no implicit endogeneity, both the estimation and the interpretation of SLX models are considerably easier than with SAR models.

Ideally, selecting the appropriate spatial econometric model should be driven by one's theory about the presence of spatial dependence in the outcomes, observables, or unobservables (Cook, Hays, and Franzese 2015). After selecting a theoretically grounded model, appropriate specification tests (such as a Lagrange multiplier test) should further buttress those choices (Darmofal 2015). In the absence of an appropriately specific theory, one at-first obvious approach would be to borrow a strategy from time series analyses (Hendry 1995): start with a general model that includes all three types of spatial dependence and then gradually pare down the model by testing restrictions (Vega and Elhorst 2015). A full interpretation of the various total, indirect, and direct effects and specification tests would then complete the process (LeSage and Pace 2009; Whitten, Williams, and Wimpy 2021). In practice, however, the estimation and interpretation of a general model that includes spatial relationships from all three sources is very difficult, and thus it is rarely done.

Cameron Wimpy (cwimpy@astate.edu) is an assistant professor, Department of Political Science, Arkansas State University. Guy D. Whitten (g-whitten@pols.tamu.edu) is a professor, Department of Political Science, Texas A&M University. Laron K. Williams (williamslaro@missouri.edu) is an associate professor, Department of Political Science, University of Missouri.

Data and supporting materials necessary to reproduce the numerical results in the article are available in the *JOP* Dataverse (https://dataverse.harvard.edu/dataverse/jop). An online appendix with supplementary material is available at https://doi.org/10.1086/710089.

The Journal of Politics, volume 83, number 2. Published online April 21, 2021. https://doi.org/10.1086/710089
© 2021 by the Southern Political Science Association. All rights reserved. 0022-3816/2021/8302-00XX$10.00

So what do political science researchers do? In our review of the literature, we found that the overwhelming majority of spatial publications in the discipline start and end with the SAR model. This is the case despite the fact that the underlying theories, as expressed by the authors, often do not match the assumptions that the SAR model implicitly imposes. We also find that researchers rarely interpret the variety of quantities of interest available in the SAR models that they estimate.

In this article we make a case for an approach to spatial model building that takes advantage of the flexibility and relative simplicity of the SLX model. In this approach, the theoretical implications of different model specifications are quite transparent, and the assumptions imposed by the SAR model can easily be turned into testable propositions. This approach simplifies the preliminary stages of model specification and minimizes inferential errors while maximizing the ease of interpretation. We offer these recommendations knowing full well that they are no substitute for careful theorizing and modeling of the data-generating process (DGP). At the same time, we think that this approach offers a better alternative to the current one that dominates the use of spatial econometrics in political science. By doing so, we hope to illuminate a path forward and provide a practical midway point between current unsatisfactory practice and the ideal modeling approaches advocated by Cook et al. (2015) and Vega and Elhorst (2015).

In the sections that follow, we begin with a brief overview of SAR and SLX models and point out how they differ in expectations about endogeneity, feedback, and higher-order effects. We demonstrate that SLX is more flexible in terms of producing an empirical test that closely matches underlying theory. We then identify a troubling pattern from our survey of the use of spatial econometric models: theories predicting spillovers among only first-order neighbors are often tested by SAR models that impose higher-order (and feedback) effects. In a series of Monte Carlo experiments, we explore which type of model is more robust to errors in expectations about the DGP; more specifically, what happens when the true model is an SLX, but we estimate an SAR and vice versa. These experiments show that SAR models perform poorly in terms of identifying the correct direction of spatial dependence, and this problem worsens as the degree of spatial heterogeneity increases. SLX models, however, do an adequate job of characterizing spatial effects of processes typically encountered in political science. We then provide an illustration of countries' defense burdens that shows how SLX models can reveal interesting patterns of spatial heterogeneity—in terms of both how the countries are connected and to what extent—that SAR models cannot. Most notably, we show that instability (in the form of interstate war) spills over into neighbors' defense burdens and that these effects are above and beyond what one might attribute to positive spatial dependence from an SAR model. We conclude with a discussion of the implications of our findings for future research along with some potential paths forward for model selection.

MODELING SPATIAL DEPENDENCE IN POLITICS

It has been a little more than a decade since political scientists wrote the first papers about the potential for spatial econometric models of political phenomena (Beck, Gleditsch, and Beardsley 2006; Franzese and Hays 2007; Ward and Gleditsch 2008). In the wake of these pathbreaking works, there has been a rapid increase in the employment of spatial econometric models by political scientists. Scholars have made theoretical arguments about how spatial relationships help to determine policies (Gray 1973; Neumayer, Plümper, and Epifanio 2014; Simmons and Elkins 2004), conflict (Buhaug and Gleditsch 2008; Garcia and Wimpy 2016), party competition (Williams and Whitten 2015; see also Bohmelt et al. 2016), terrorism (Midlarsky, Crenshaw, and Yoshida 1980; Neumayer and Plümper 2010a), and many other outcomes.

This section provides an overview of two of the most popular spatial econometric models: SAR and SLX. Spatial dependence occurs in both models, whether it is in the outcomes (SAR) or in the observables (SLX). The researcher specifies the manner in which all the observations are connected to one another via an $N \times N$ weights matrix ($\mathbf{W}$) in both models, so it is worthwhile spending some time on this critical task. For ease of presentation, we explore the basic mechanics of each model in the context of a simple contiguity or first-order weights matrix.[1] In figure 1 we show a toy example of how these types of matrices are constructed. On the far left side of this figure, we show a map of six units with arrows depicting first-order neighbors. For now, only consider the first-order contiguity weights matrix ($\mathbf{W}$), where each cell identifying a pair of neighboring units contains a one and all other cells contain zeros. Since being a neighbor is symmetrical, this matrix is symmetrical. And since a unit cannot be a neighbor of itself, the main diagonal contains all zeros. As we will see below, these features of $\mathbf{W}$ have meaningful consequences for substantive inferences.

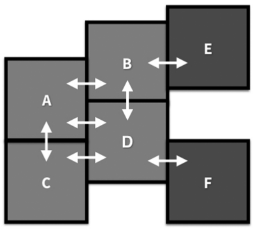
The spatial autoregressive model

The equation for a basic SAR model is

$$\mathbf{y} = \mathbf{X}\beta + \rho\mathbf{W}\mathbf{y} + \varepsilon, \tag{1}$$

1. In addition to making the presentation more clear, this is the most popular type of weights matrix used in political science research. In our survey of the literature (described below), we found that 72.3% of the models reported used a weights matrix specified on the basis of geography. Regardless, the same types of patterns and problems discussed in this article persist in SAR models regardless of the specification of $\mathbf{W}$.

Map



W

	A	B	C	D	E	F
A	0	1	1	1	0	0
B	1	0	0	1	1	0
C	1	0	0	1	0	0
D	1	1	1	0	0	1
E	0	1	0	0	0	0
F	0	0	0	1	0	0

First-Order Neighbors

W²

	A	B	C	D	E	F
A	3	1	1	2	1	1
B	1	3	2	1	0	1
C	1	2	2	1	0	1
D	2	2	1	4	1	0
E	1	0	0	1	1	0
F	1	1	1	0	0	1

First-Order Neighbors (Squared)

$\mathbf{W}_{2nd}$

	A	B	C	D	E	F
A	0	0	0	0	1	1
B	0	0	1	0	0	1
C	0	1	0	0	0	1
D	0	0	0	0	1	0
E	1	0	0	1	0	0
F	1	1	1	0	0	0

Second-Order Neighbors

Figure 1. Spatial arrangement with associated squared ($\mathbf{W}^2$) and second-order ($\mathbf{W}_{2nd}$) weights matrices

and the reduced equation (isolating **y** on the left side) is

$$\mathbf{y} = (\mathbf{I}_n - \rho\mathbf{W})^{-1}\mathbf{X}\beta + (\mathbf{I}_n - \rho\mathbf{W})^{-1}\varepsilon. \quad (2)$$

From the infinite series expansion of the spatial multiplier,

$$(\mathbf{I}_n - \rho\mathbf{W})^{-1} = (\mathbf{I}_n + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \rho^3\mathbf{W}^3 + \ldots), \quad (3)$$

we can see that higher-order and feedback effects, or *global effects*, are present in any and all SAR models. Consider the second weights matrix depicted in figure 1 ($\mathbf{W}^2$), which has two meaningful features. First, unlike **W**, the main diagonal elements—which convey the relationship of a unit with itself—do not equal zero. These values reflect feedback effects for which the effect of a unit on its neighbors comes back to affect the unit itself.

Second, we can see that $\mathbf{W}^2$ contains nonzero values for all cells containing second-order neighbors.[2] The result is that the effects extend beyond first- and second-order neighbors (because of **W** raised to increasingly higher values in eq. [3]), and they occur simultaneously at time t. We can thus label them as "global effects."

While interpreting the coefficients is a reasonable place to start, it is a bad place to stop in the interpretation of an SAR model. Given all of the different higher-order and feedback terms implicit in $\mathbf{W}^2$ and the higher-order terms in equation (3), a general interpretation of the estimated ρ masks substantial variation in the estimated effects across units. For instance, if we want to infer the effect of a single x on y, then one approach is to examine the partial derivatives matrix (LeSage and Pace 2009; Whitten et al. 2021):

$$\left[\frac{\partial \mathrm{E}(y)}{\partial x_1} \cdots \frac{\partial \mathrm{E}(y)}{\partial x_N}\right] = (\mathbf{I} - \hat{\rho}\mathbf{W})^{-1}\hat{\beta}, \quad (4)$$

where the resulting $N \times N$ matrix (N is the total number of observations) contains both the impacts of x_i on y_i, or direct effects (along the diagonal), and the impacts of x_i on y_j, or indirect effects (along the off-diagonal).

The spatial-X model

In contrast to the SAR model, the SLX model offers a framework in which researchers may choose whether to model local or global spatial relationships. In addition, the SLX model is relatively easier to estimate and interpret. In practice, the SLX model has been most often discussed as a modeling strategy for exogenous spatial effects from a direct neighbor. This effort has primarily been driven by the work of LeSage and Pace (2009) in which they suggest that the SLX model works best in the case of "externalities," or "local spillovers." Spillovers would appear to accurately characterize many political science phenomena; something happens to a neighbor that affects the outcome of interest.

The equation for a basic SLX model can we written as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{WZ}\boldsymbol{\theta} + \varepsilon, \quad (5)$$

where **Z** is the matrix of variables expected to exert spatial influence on **y** through a theoretically specified **W** matrix that connects observations to each other through a vector of spatial parameters $\boldsymbol{\theta}$.[3]

If we start with a simple SLX model with a single independent variable, x_1, such that $\mathbf{X} = \mathbf{Z} = x_1$, we calculate the effect of x_1 on y as

$$\left[\frac{\partial \mathrm{E}(y)}{\partial x_1}\right] = \hat{\beta}_1 + \hat{\theta}_1\mathbf{W}. \quad (6)$$

In contrast to the infinite series expansion in the SAR spatial multiplier, the indirect effect ($\hat{\theta}_1\mathbf{W}$)—or "neighbor effect"—is only present at the first order. In other words, with the specification presented in equation (5), the effect of x_1 is limited to only local effects; the effect does not continue to second-order neighbors, and there are no feedback effects.

2. The value in these cells reflects the number of paths through which each pair of units are second-order neighbors. So, e.g., $\text{cell}_{AE} = \text{cell}_{EA} = 1$ because A and E are only connected through B, but $\text{cell}_{BC} = \text{cell}_{CB} = 2$ because B and C are connected through both A and D.

3. We use **Z** to reflect the possibility that **Z** and **X** can have different contents.

However, as we discuss below, if we want to specify higher-order and feedback effects, we can incorporate them into an SLX model.

Without realizing it, scholars often use SLX models by including spatially weighted independent variables. In fact, any model that controls for the sum or average of neighbors' attributes reflects an SLX specification (Drolc, Gandrud, and Williams 2019). The most common example incorporates a temporally lagged spatial lag (TLSL), which tests a common argument in theories of policy diffusion that the diffusion process—or the influence of neighbors' covariates more generally—occurs with a temporal lag. This variant of the SLX equation can be written as

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{W}\mathbf{y}_{t-1}\theta + \varepsilon. \tag{7}$$

While the estimation is straightforward, the presence of the TLSL adds a temporal dimension to the quantities of interest.[4]

In the next section, we provide some guidelines as to when to use one model over the other, by highlighting each model's flexibility and ability to derive correct causal inferences.

THEORY, SPECIFICATION, AND ESTIMATION IN SAR AND SLX MODELS

The general motivating force behind spatial econometric models is to test theories about how y_i is a function of some aspect of unit j. If a researcher's expectations are that the spatial relationships are between y_i and y_j and that these relationships are global (including feedback from y_i to y_j and back to y_i) and occur immediately, then the SAR model is clearly the appropriate model to choose. Yet, if a researcher's expectations are that the spatial relationships in a study are between y_i and some variable z_j, and these relationships are local in nature, then the SLX model is clearly the appropriate model to choose. But what about a researcher whose expectations are less sharp? Under these circumstances, we echo the suggestion by Vega and Elhorst (2015) that the SLX model is a better place to start than the SAR. The advantages of the SLX model are fourfold: greater flexibility in specifying lower-order (local) versus higher-order (global) effects, the ability to relax and test common factor restrictions, the ease of estimation and interpretation, and more flexible and realistic specifications of temporal processes. We will explore each of these in turn.

First, the SLX model offers a more flexible approach to modeling lower-order versus higher-order spatial effects than the SAR model. As we outlined above, the SAR model implicitly imposes a number of strong assumptions that may not be immediately obvious to scholars and that often run contrary to their theoretical expectations about the spatial processes at work. In most incarnations of the SAR model, there is no way to limit spatial effects to only local first-order effects.[5] Although the SAR model certainly captures local effects, by construction it also finds higher-order effects. The infinite series expansion (eq. [3]) reveals that all observations are eventually affected by any change in x; this is the case even if there is no theoretical reason to expect that nth-order neighbors would be influenced (albeit in a small fashion) by a change.[6] The SAR model offers no flexibility in terms of how many orders of neighbors are affected. Feedback effects start in the second order (e.g., $\mathbf{W}^2$ in fig. 1). Although one may theorize that the total effect of a change in x_j for observation i arises only from first-order local effects, the SAR model forces there to be higher-order and feedback effects, whether this accurately reflects the true spatial process or not. The consequence is that SAR models will risk finding a global process even if one is not present.

In an SLX model, expectations about local first-order processes can be easily tested through the specification of a model like that in equation (5). But can SLX perform adequately when the true spatial process includes higher-order effects? The answer is yes. By altering the model specification, scholars can incorporate higher-order effects into their models of spatial processes. For example, imagine that we have an expectation of higher-order effects that stop at the third order. For ease of exposition, we limit our model to a single independent variable $\mathbf{x}$. In this case, we would estimate an SLX model specified as follows:[7]

$$\mathbf{y} = \beta\mathbf{x} + \theta_1\mathbf{W}\mathbf{x} + \theta_2\mathbf{W}^2\mathbf{x} + \theta_3\mathbf{W}^3\mathbf{x} + \varepsilon. \tag{8}$$

The total estimated effect of x on y would be

$$\left[\frac{\partial \mathrm{E}(y)}{\partial x_1} \cdots \frac{\partial \mathrm{E}(y)}{\partial x_N}\right] = \hat{\beta} + \hat{\theta}_1\mathbf{W} + \hat{\theta}_2\mathbf{W}^2 + \hat{\theta}_3\mathbf{W}^3, \tag{9}$$

4. A common misconception is that this is "an alternative specification of the spatial autoregressive model" (Beck et al. 2006, 40). However, the TLSL forces the causality to go in a single direction (i.e., from y_{t-1} to y_t), which eliminates the defining feature of SAR models: spatial endogeneity. Neumayer and Plümper (2010b, 158) note that a model with a temporal lag "is not strictly speaking a spatial lag model." Feedback for i from x_{t-1} can occur by first influencing i's neighbors at time t, which then influences j's neighbors (including i) at time $t + 1$. In the first section of our appendix (available online), we provide some more discussion of the features of the TLSL model and why it is more like an SLX model than an SAR model.

5. In the case of SAR models using directed dyadic data, one can limit the effects to the first order by distinguishing between the sources and targets of stimulus (e.g., Neumayer and Plümper 2010b, 152–54).

6. The one exception to this would be for a unit that has no connectivity with any other. In fig. 1 such a unit would appear as an island with values equal to zero for the relevant column and row in $\mathbf{W}$.

7. In this specification, $\mathbf{X} = \mathbf{x}$, a single vector containing the values for independent variable x, and $\mathbf{Z} = (\mathbf{x} \quad \mathbf{x} \quad \mathbf{x})$.

where $\hat{\beta}$ is the estimated direct (or zero-order) effect of x, $\hat{\theta}_1$ is the estimated first-order indirect effect, $\hat{\theta}_2$ is the estimated second-order indirect effect, and $\hat{\theta}_3$ is the estimated third-order indirect effect. It is worth noting that restrictions on the order of effects can be tested, so one strategy is to start with an n-order specification and then use hypothesis tests to pare down the model.[8]

With the model specification shown in equation (8), the SLX model mimics the SAR model in its estimation of feedback effects. However, there might be situations in which one expects higher-order neighbor effects but does not expect that the spatial process will have feedback effects. In these cases, one can exchange the terms with **W** squared, cubed, and raised to higher powers for n-order contiguity matrices such as $\mathbf{W}_{2\text{nd}}$ in figure 1 (as we demonstrate in the appendix). One can also test whether the spatial process exhibits feedback loops, by comparing the model fit (through, e.g., the Akaike information criterion and Bayesian information criterion) of an SLX specification with traditional squared and higher-order **W** terms versus that from one with nth-order contiguity matrices.

The second relative advantage of the SLX model is the ability to relax and test common factor restrictions. In the case of an SAR model, there is an imposed assumption that all global spatial dependence works through a single spatial parameter, ρ. For example, consider estimating an SAR model where we have two independent variables, x_1 and x_2, in **X**. Since there is only one estimated ρ in the model, the infinite series expansion of the spatial multiplier (found in eq. [3]) is identical, and the only difference in the estimated spatial effects comes from different coefficient estimates in $\hat{\beta}$. In other words, the strength of spatial dependence, and the manner in which higher-order effects reverberate throughout the system, is identical for both x_1 and x_2 by construction. Moreover, one spatial autocorrelation coefficient (ρ) is used to represent the declining spatial effects across all orders of neighbors, albeit at an order of magnitude smaller at each additional order of neighbors (ρ, ρ^2, ρ^3, etc.). Rather than this common factor restriction being imposed by the model (in the case of SAR), in the context of an SLX model it becomes a testable proposition; we can test the accuracy of these restrictions and modify the specification accordingly.[9]

It is not outside the realm of possibility to consider political science applications in which spatial relationships may vary across independent variables. In such cases of spatial heterogeneity, the SLX model offers a far more straightforward and accurate approach by allowing scholars to estimate different spatial parameters for the independent variables. Unlike the SAR model, the SLX model also provides the flexibility of spatially lagging only those variables from a neighboring unit that are theorized to affect y_i via some form of spatial dependence. In other words, **X** and **Z** do not have to be the same. Some covariates might only have direct effects (i.e., only be in **X**), and others might only have indirect effects (i.e., only be in **Z**). Furthermore, the **W**s can vary across the different z variables (if different interconnectivities are theorized to exist), and this model allows for spatial heterogeneity (e.g., when $\theta_1 \neq \theta_2$). As opposed to the complicated multiple SAR model, estimation for an SLX model with different interconnectivities is rather straightforward.

The third advantage of SLX models to applied researchers is ease of estimation and interpretation. Perhaps the most notable argument for the SLX model comes from the Gibbons and Overman (2012) critique of the general enterprise of spatial econometrics. These authors argue that analysts using the SAR model end up with models that are weakly identified because of the hurdles involved in estimation of these endogenous models. The authors further point out that the model being estimated rarely fits the expressed theory. As such, they make an argument for the far simpler, exogenous SLX model.

SLX models provide much more meaningful coefficients in terms of allowing scholars to quickly and correctly make inferences about the spatial processes at work. While the SAR model produces a global coefficient of average spatial dependence (ρ), getting substantive impacts and estimates of uncertainty from other coefficients of interest is notably difficult. Moreover, the ρ does not easily distinguish between local and higher-order connectivity, thus making it difficult to make substantive inferences about the degree of spatial dependence from one neighbor to another. Given the complexity of calculating quantities of interest from SAR models, it is perhaps understandable that few scholars move beyond simple interpretations of the sign and significance of ρ.

In contrast, the SLX model is easier to interpret. The biggest difference in interpretation is that in an SLX model, the respective $\hat{\beta}$ can be interpreted as is; $\hat{\beta}$ is the estimated change in y_i for a one-unit increase in x_i. Yet, in an SAR model, the $\hat{\beta}$ represents the estimated "direct effect" of x_i on y_i, which, because of spatial dependence in the outcomes, we never actually observe because they are merely the prespatial effects that have yet to be filtered through the spatial dependence. This is complicated by the fact that the effect of x_i on y_i also depends on the strength of feedback. The spatial coefficients from an SLX model, the $\hat{\theta}$ terms, are interpreted as the estimated impact of a one-unit increase in the relevant z variable on the

8. One word of caution is that if there are few neighboring observations in the weights matrix, then higher-order representations of **W** can exacerbate problems of multicollinearity and inflate standard errors.

9. For example, in Moore and Shellman's (2007) analysis on refugees' destination, they estimate different parameters for the spatial contiguity effects of a number of independent variables.

outcome variable in a neighboring unit.[10] In comparison, this effect is difficult, if not impossible, to ascertain from merely examining the $\hat{\rho}$ value from an SAR model.

The fourth advantage of SLX models is that they offer more flexible and realistic specifications of temporal processes. Up until now, we have ignored possible temporal variations in our data other than to note that by construction SAR models impose that all effects of y_j on y_i and feedback effects from y_i back on itself happen immediately. If our theory is that the effect of y_j on y_i and any associated feedback effects take time to happen, we can test this through an SLX model by including $\mathbf{y}_{t-1}$ as a term in our specification of $\mathbf{Z}$ (see eq. [7]).

In the next section, we explore patterns in how political scientists have selected and interpreted spatial econometric models.

SPATIAL DEPENDENCE IN PRACTICE

To provide a systematic overview of how researchers in political science have used spatial econometric models, we reviewed every work published through the end of 2015 that cited Beck et al. (2006), Franzese and Hays (2007), or both—two highly influential early papers promoting the use of spatial models.[11] We coded the use of these techniques in the main empirical model that was reported.

Table 1 shows a taxonomy of the use of spatial econometric models by political scientists in terms of the type of spatial theory expressed and the model estimated. There are three patterns worth noting in this table. First, although SAR models recover global effects—or those in which all nonisolate observations are influenced by changes in one observation—only 8.9% (5 of 94 studies) expressed theories that are global in nature. In those rare cases when scholars explicitly state global theories, however, they tend to appropriately test those theories with a model based on dependence in the outcomes (i.e., SAR or a more complex model that includes an SAR process). Second, when political science researchers have expressed a theory that is either local or not specific in terms of the types of spatial expectations, almost half the time (46.1%) they have estimated an SAR model that imposes global spatial effects.

Third, only 10.6% (10 out of 94 studies) of the publications we reviewed feature more than one pattern of spatial dependence. There is also generally little discussion of the process—if any—used by scholars to pare down more complex models

Table 1. Applications of Spatial Econometrics in Political Science

	Spatial Theory Type			
	Global		Local	
Model Used	Total	(%)	Total	(%)
SAR	4	80.0	41	46.1
SLX	0	.0	34	38.2
Complex	1	20.0	9	10.1
Other	0	.0	5	5.6

Note. Summary of the 94 studies citing Beck, Gleditsch, and Beardsley (2006), Franzese and Hays (2007), or both that reported results from at least one spatial econometric model.

to simpler models. Altogether, this is indicative of scholars beginning with a specific model in mind (usually the SAR), rather than a general-to-specfic approach. We return to this point in our discussion of paths forward below.

Among those papers in our survey that estimated an SAR model, the overwhelming majority offered a theory of indirect effects—the effect of x_i on y_j—that were local in nature rather than global. In fact, only 8.9% (4 out of 45 studies) of those that estimated an SAR model explicitly expressed a global theory. Essentially this means that scholars theorize that explanatory variables may have an effect on first-order neighbors (local) but are silent about potential impacts on higher-order neighbors (global; Elhorst 2014). This latter finding is striking because the SAR model imposes global relationships across all spatial units. In contrast, spatial relationships in SLX models can be specified as either local or global.

Another common pattern emerges in our in-depth exploration: scholars often estimate a series of SAR models, each with a different specification of the weights matrix (e.g., Flores 2011; Gassebner, Gaston, and Lamla 2011; Goldsmith 2007; Obinger and Schmitt 2011).[12] Recall that the SAR model imposes a common factor restriction that all of the spatial dependence operates through one parameter, ρ, connected through a properly specified $\mathbf{W}$. An SLX model, however, easily allows different variables to influence the outcome through multiple spatial dependence paths; the final model can then be pared down using traditional model selection criteria. The examples referenced above highlight at least three problems that arise from estimating

10. We are limiting the current discussion to cases in which $\mathbf{W}$ is a standard contiguity matrix with no row-standardization. With a more complicated $\mathbf{W}$ where the weights matrix contains values other than 0 and 1, the estimated effect of a one-unit increase in x_j on y_i is $\hat{\theta} \times w_{ij}$.

11. This initial search yielded 155 publications. We narrowed this down to 94 studies that reported results from at least one spatial econometric model.

12. These examples do not represent a comprehensive list. Instead, we identify them as representative examples of instances in which the inferences might be different with an SLX model, either because of methodological differences or because the SLX model represents a more appropriate way of testing the theoretical expectations.

separate SAR models: first, it severely limits whether authors can determine which model best approximates the DGP. Second, if multiple weighting schemes are operating simultaneously and are correlated, then any model including only one will be biased. Third, even if only one weights matrix correctly specifies the spatial process, if explanatory variables operate through the weights matrix in different ways, then the SAR model will not be able to disentangle that spatial heterogeneity.[13]

In the next section we present the results from two sets of Monte Carlo experiments in which we explore how SLX and SAR models perform under different DGPs.

EXPERIMENTS

In this section we provide two sets of Monte Carlo experiments to explore the robustness of both models to errors in expectations about the DGP. In the first set of experiments, we expect that SLX models will perform admirably in deriving inferences about the effects of explanatory variables, even though the true process is generated by an SAR. One of the advantages of SLX models that we described above is that they can mimic the higher-order effects of SAR models in situations that scholars are likely to encounter in practice. The opposite, however, is unlikely to be true. In the second set of experiments we expect that SAR models are too inflexible to effectively deal with spatial heterogeneity that often accompanies an SLX DGP.

SAR data-generating process

In the first set of experiments, we generate data using the reduced-form equation for the SAR model found in equation (2), with matrix $\mathbf{X}$ containing a single variable drawn from a uniform distribution, $x \in [-10, 10]$, and where $\beta = 1$. Matrix $\mathbf{W}$ is an $N \times N$ symmetric row-standardized contiguity weights matrix, where each element below the diagonal is randomly drawn from a Bernoulli distribution. We simulated 1,000 data sets at each of nine different scenarios defined by the strength of the spatial autocorrelation, $\rho \in \{-0.8, 0.8\}$.[14] In each panel of table 2, we display our findings from estimations using a different SLX specification.[15]

13. This is not to say that there are no good examples of scholars carefully selecting the appropriate empirical model and then painstakingly exploring the substantive meaning of those models. Freeman and Quinn (2012, 67), e.g., reveal how income inequality and financial integration influence democracy through multiple channels of spatial dependence.

14. It is worth noting, however, that values of ρ with an absolute value close to 1 are exceedingly rare in political science research.

15. In app. table 1, we present the same results for an SAR estimation. Not surprisingly, the SAR model does an excellent job of recovering the true DGP.

The first SLX model that we estimate includes one spatially lagged independent variable ($\mathbf{Wx}$) and is thus specified as

$$\mathbf{y} = \mathbf{x}\beta + \mathbf{Wx}\theta + \varepsilon. \qquad (10)$$

The βs are not directly comparable across SAR and SLX specifications; instead, with the use of the partial derivatives matrix one can easily compare the average direct, indirect, and total effects across models (LeSage and Pace 2009). Additionally, since these effects can be partitioned into n-order effects, we assess the performance of the model at different orders. Table 2 shows how often the estimated SLX model specification's 95% confidence intervals include the true effects given the DGP characterized in equation (2). Each panel of the table provides the recovery rates for a different SLX model specification (more on the others below), and each column represents a different strength of spatial lag coefficient (ρ in eq. [2]).

There are two clear patterns in panel A of table 2. First, the simple SLX model recovers the true average direct effect at an acceptably high rate (nearly 95%) for all values of ρ except 0.8.[16] The recovery rates for the zero-order direct effect (characterized by the β_{SLX}), however, dip below 95% at ρ values lower than -0.4 and higher than 0.4. This means that β_{SLX} is capturing the true average direct effect at common values of ρ, but it is not an accurate reflection of the zero-order direct effect (β_{SAR}). The second clear pattern is that the inclusion of one $\mathbf{W}x$ term is enough to recover the true first-order indirect effects nearly 95% of the time for all $\rho \in [-0.6, 0.6]$, but overall this model specification does a poor job of capturing the true total indirect effects. These two patterns are a consequence of an overly simplistic SLX model specification given the true DGP. The lack of higher-order $\mathbf{W}$ matrices means that there cannot be second-, third-, or higher-order indirect effects and that there can be no feedback effects. This is why the simple SLX in panel A of table 2 has a recovery rate for all higher-order direct and indirect effects that is 0% at all values of ρ (represented by ellipses).

These results from the experiments displayed in panel A of table 2 should certainly warrant caution from scholars. To what extent can we eliminate some of these concerns by modifying the SLX model specification? In the second set of experiments, we add a first-order weights matrix squared ($\mathbf{W}^2$) and estimate a second SLX coefficient (θ_2) with a model specified as

$$\mathbf{y} = \mathbf{x}\beta + \mathbf{Wx}\theta_1 + \mathbf{W}^2\mathbf{x}\theta_2 + \varepsilon. \qquad (11)$$

16. It is worth noting that negative values are quite rare and that we have never seen such a high value for ρ in a political science application.

Table 2. Recovery Rates for Various SLX Model Specifications: SAR Data-Generating Process

	−.8	−.6	−.4	−.2	0	.2	.4	.6	.8
	A. Simple SLX model: $\mathbf{y} = \mathbf{x}\beta + \mathbf{Wx}\theta + \varepsilon$								
Direct:									
Total	94.9	94.4	94.8	93.7	95.5	93.5	95.0	94.0	88.9
0 order	77.9	91.9	95.1	93.5	95.5	93.4	94.8	90.7	81.0
Second order	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
Third order	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
Indirect:									
Total	15.1	49.9	85.2	93.7	95.2	93.5	40.3	0	0
First order	91.4	93.8	95.5	93.1	95.2	94.5	94.4	93.8	93.1
Second order	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
Third order	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
	B. SLX model with a squared term: $\mathbf{y} = \mathbf{x}\beta + \mathbf{Wx}\theta_1 + \mathbf{W}^2\mathbf{x}\theta_2 + \varepsilon$								
Direct:									
Total	95.0	94.0	95.1	93.8	95.4	94.0	95.2	94.4	91.0
0 order	94.2	95.1	95.2	94.2	93.6	94.8	94.8	94.9	95.8
Second order	93.2	93.9	94.7	94.6	94.0	94.7	93.6	93.4	93.8
Third order	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
Indirect:									
Total	93.0	94.7	94.8	94.7	93.9	94.5	93.3	83.7	2.6
First order	93.0	94.9	95.9	94.0	94.9	94.9	94.4	93.2	92.6
Second order	93.2	93.9	94.7	94.6	94.0	94.7	93.6	93.4	93.8
	C. SLX model with squared and cubed terms: $\mathbf{y} = \mathbf{x}\beta + \mathbf{Wx}\theta_1 + \mathbf{W}^2\mathbf{x}\theta_2 + \mathbf{W}^3\mathbf{x}\theta_3 + \varepsilon$								
Direct:									
Total	95.0	94.2	94.6	94.3	96.1	94.4	94.5	93.7	93.4
0 order	94.2	95.1	95.3	94.1	93.9	94.9	94.3	94.9	95.9
Second order	93.9	94.0	94.8	95.2	94.2	95.2	93.8	93.2	93.3
Third order	94.2	93.6	94.7	93.5	95.0	94.4	95.0	93.9	93.4
Indirect:									
Total	95.5	95.1	95.6	94.4	95.2	94.5	93.9	92.6	87.0
First order	94.3	94.7	95.1	93.9	94.9	93.7	93.9	95.2	94.4
Second order	93.9	94.0	94.8	95.2	94.2	95.2	93.8	93.2	93.3
Third order	94.2	93.6	94.7	93.5	95.0	94.4	95.0	93.9	93.4

Recall that **W** matrices raised to higher orders allow for spatial effects that cycle through neighbors of neighbors and back to the originator (see the first-order squared matrix in fig. 1). This second model specification, therefore, should help address the poor performance of the first model in estimating the average indirect effects created in the DGP.

From table 2 panel B, we can see that the addition of the squared weights matrix improves the performance of the model in all of the weak spots detailed above. First, the zero-order direct effects are now recovered at 95% for nearly all the values of ρ, in addition to the second-order direct effects (feedback effects), which results in even more accurate estimates of the average direct effect. Second, the ability to model second-order indirect effects allows one to gain accurate inferences about not only those effects (nearly 95% recovery rates at all values) but also about the average indirect effects overall. Instead of poor performance at values outside of $\rho \in [-0.2, 0.2]$, the average indirect effects are well recovered at a much broader range of values ($\rho \in [-0.6, 0.4]$). Furthermore, in these experiments the improvement in model performance does not come at a cost of multicollinearity; on average the highest variance inflation factor is a little more than 2.[17]

As we discussed earlier, the infinite series expansion of the spatial multiplier in the SAR model (see eq. [3]) means that the indirect effects increase forever but at a declining rate. To mimic this sort of process with an SLX model, one would have to include a large number of higher-order weights matrices. The utility of this approach, however, would be limited, as the size of the indirect effects from higher-order effects is quite small, and this would likely exacerbate any issues of multicollinearity. The next step, then, is to see whether including a cubed weights matrix effectively recovers the direct and indirect effects by estimating a model specified as

$$\mathbf{y} = \mathbf{x}\beta + \mathbf{Wx}\theta_1 + \mathbf{W}^2\mathbf{x}\theta_2 + \mathbf{W}^3\mathbf{x}\theta_3 + \varepsilon. \tag{12}$$

The recovery rates presented in panel C of table 2 reveal the untapped promise of the SLX model even in those processes ripe with spatial diffusion. In all but the most extreme cases of positive spatial autocorrelation (i.e., when $\rho > 0.6$), the SLX model with both squared and cubed $\mathbf{W}$ terms recovers the average direct effect, average indirect effect, and by implication, the average total effect. Thus, in the vast majority of cases that scholars observe in practice, one could estimate an SLX model and make the same substantive inferences regarding the impact of a variable on the observation, its neighbors, and the other observations. The only drawback is higher multicollinearity (the average variance inflation factor score is now above 5), which leads to somewhat inflated standard errors. This is certainly not a reason to avoid estimating the SLX model and is a similar decision to those made by scholars choosing whether to include lower-order terms in interactive models.[18]

In the next section we explore the opposite scenario in which we have an SLX DGP but we estimate an SAR model.

SLX data-generating process

Our next set of experiments assesses the performance of SAR and SLX models when the spatial processes operate through dependence in the observables and there are no higher-order spatial effects. In other words, how do the models perform when the data are generated in a manner consistent with an SLX model? As before, we choose to compare the performance of SLX and SAR models on the basis of whether they can recover the true average direct and indirect effects, calculated via the partial derivatives approach.

We expect that SLX models—not surprisingly—will perform well.[19] Our expectations for SAR models, however, are mixed. We expect the SAR to do poorly in terms of recovering the true indirect effects because it will find evidence of higher-order and feedback effects in the data that we know do not exist. We also expect that situations of spatial heterogeneity—when the explanatory variables operate through different spatial processes—will be particularly problematic for recovering the correct estimates of direct and indirect effects. We generate data for our first set of SLX DGP Monte Carlo experiments as

$$\mathbf{y} = \mathbf{x}_1\beta_1 + \mathbf{x}_2\beta_2 + \mathbf{Wx}_1\theta_1 + \mathbf{Wx}_2\theta_2 + \varepsilon, \tag{13}$$

where x_1 and x_2 are drawn from uniform distributions, $x_1 \in [-10, 10]$ and $x_2 \in [-5, 5]$, and β_1 and β_2 are set to 1. Matrix $\mathbf{W}$ is an $N \times N$ symmetric row standardized contiguity weights matrix, where each element below the main diagonal is randomly drawn from a Bernoulli distribution. In order to assess how these models perform under varying degrees of spatial heterogeneity, we conduct 1,000 simulations of nine scenarios in which we vary the magnitude of the spatial coefficients (θ_1 and θ_2).

The first step is to determine whether the SAR can recover the true total effects of x_1 on $\mathbf{y}$. Figure 2 shows histograms for 1,000 estimated average total effects (with 95% confidence intervals via the percentile method depicted with vertical dashed lines), as we vary the values of θ_1 and θ_2. More specifically, the histograms are arranged to represent variation on two critical elements of these models. Moving vertically depicts varying levels of *spatial effects*, ranging from weak (bottom row) to moderate (middle row) to strong (top row). Moving horizontally depicts varying levels of *spatial heterogeneity*, ranging from none (left column) to moderate (middle column) and substantial (right column). In each experiment, the true total effect is the sum of the direct effect ($\beta_1 = 1$) and the indirect effect (θ_1) for x_1, or 0.8, 1.4, and 0.2 (going from the bottom row to the top row). If the 95% confidence intervals (dashed vertical lines) for the calculated total effects from the SAR include the true total effects from the DGP (solid vertical lines), then we can conclude that estimating the SAR model when the true DGP is an SLX will recover the substantive effects of x_1.

Figure 2 shows that both spatial effect size and spatial heterogeneity have meaningful influences on the degree of bias for the SAR under these SLX DGPs. In nearly all of the scenarios, the true total effect appears far away from the bulk of the estimated total effects from the simulations. As one goes from no spatial heterogeneity (first column) to moderate and substantial

17. Of course, scholars must make their own decisions as to whether the increased multicollinearity in their model is worth being able to derive more accurate inferences about spatial processes.

18. In the appendix, we also present the performance of an SLX model with a matrix specified as $\mathbf{W}_{2nd}$ in addition to $\mathbf{W}$.

19. Results from these estimates are shown in app. table 2.

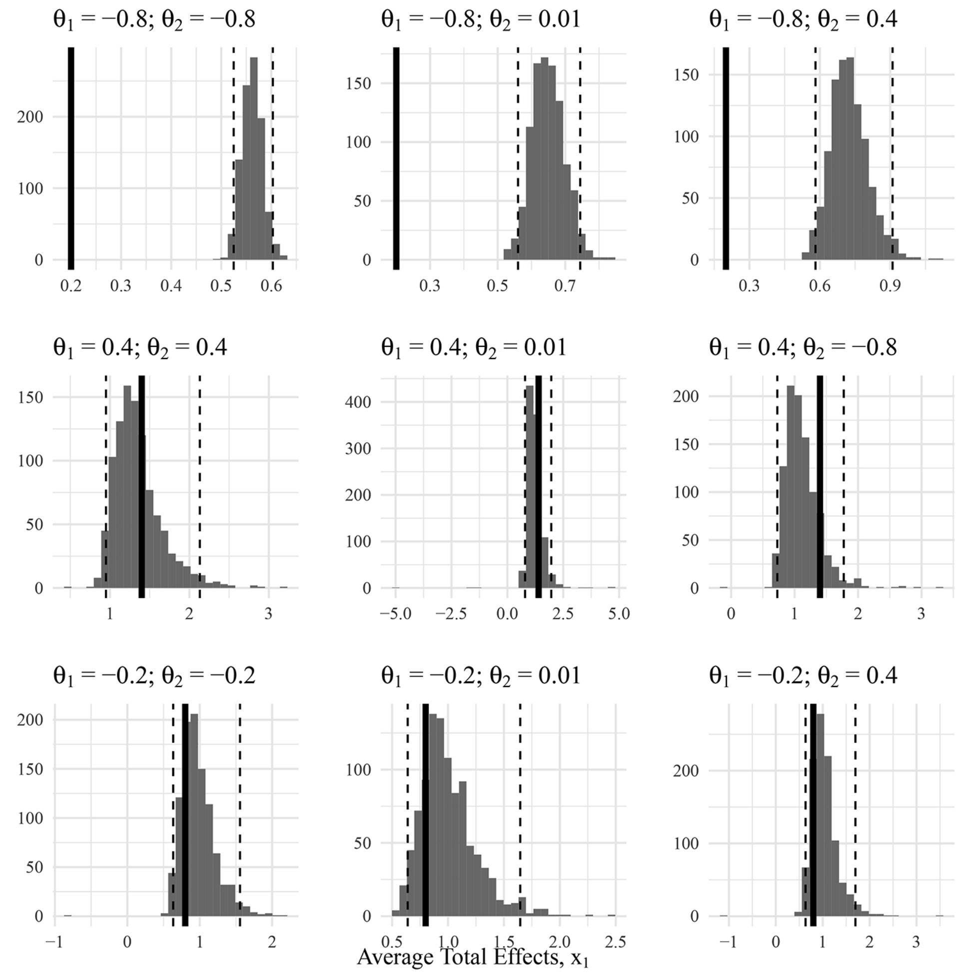


Figure 2. Recovery of the true total effects of x_1 in the SAR model across strength of spatial autocorrelation and spatial heterogeneity. The θ values refer to those from the SLX Monte Carlo simulations. Vertical lines represent 95% confidence intervals (*dashed*) and the true total effect of x_1 (*solid*).

heterogeneity (third column), the true total effect is less likely to be located near the center of the 95% confidence interval. Thus, when there are multiple patterns of spatial dependence in one model, estimating an SAR model (with one **W**) is likely to provide incorrect inferences about x_1.

Moving up the rows (from weak to strong spatial effect size), the consequences for inferences become even more severe; while the SAR confidence intervals include the true total effect under weak and moderate (with the true total effect falling near the edges of the intervals for substantial heterogeneity) spatial effects, the true total effect is much smaller than those predicted by the SAR model under strong spatial effects. This should cause scholars to be cautious about using SAR models when they expect either strong spatial dependence or that multiple spatial processes are at work.

While figure 2 provides clear evidence of overestimation of the total effects, it does not identify whether the source of overestimation is related to the direct effects or indirect effects. Our expectation was that it would be caused by the latter, since the estimation of the direct effect is rather straightforward in both types of models. Indeed, across all scenarios, the SAR model is able to recover the true average direct effects in at least 95% of the simulations.

Another source of bias in the SAR's calculation of indirect effects is the common factor restriction that forces all of the independent variables to influence the dependent variable via the same spatial multiplier, whether this makes sense or not. We can imagine circumstances in which there are different causal processes (or at least different functional forms) operating across the explanatory variables. Figure 3 confirms our suspicions about the dangers of the common factor restriction in circumstances of substantial heterogeneity. The 95% confidence interval only includes the true indirect effect for x_2 (in this case, θ_2) in five of the nine scenarios. More troubling is the fact that under substantial spatial heterogeneity (last column), scholars would conclude that the indirect effects of x_2 are negative (positive), when in fact the true spatial pattern is positive (negative) spatial dependence. Because there are both positive and negative spatial patterns present in the DGP that the SAR model cannot disentangle,

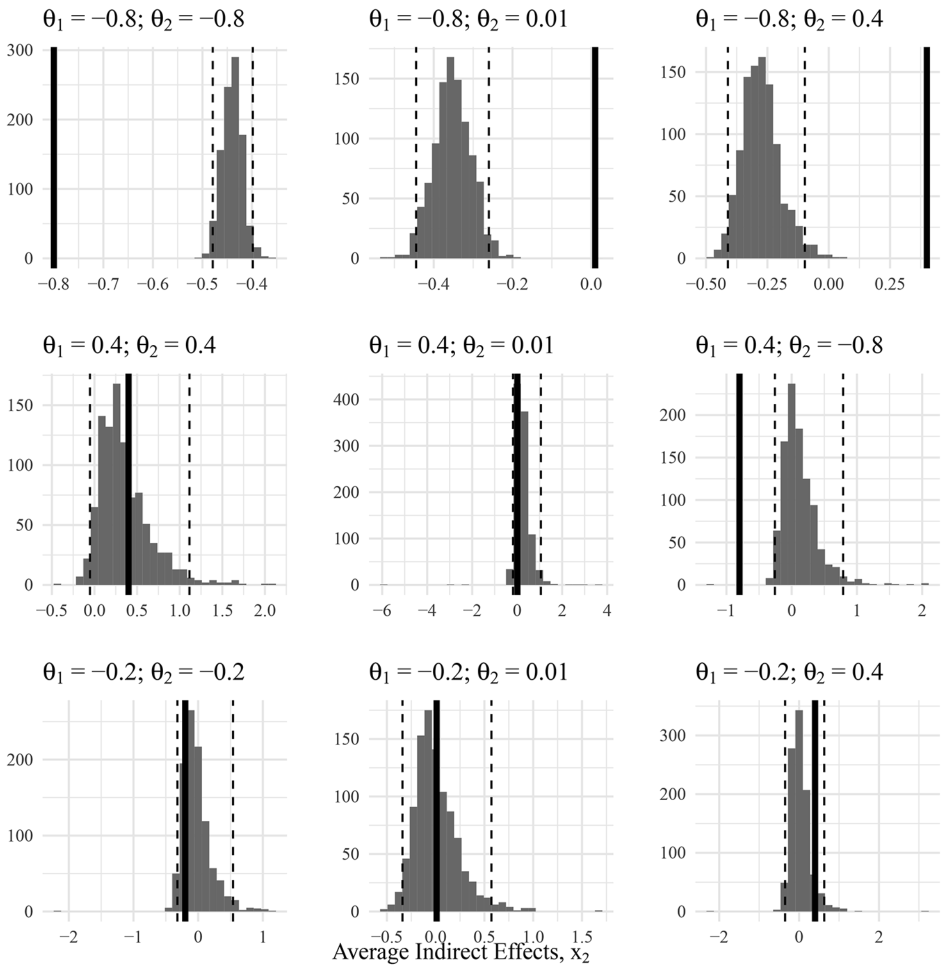


Figure 3. Recovery of the true indirect effects of x_2 in the SAR model across strength of spatial autocorrelation and spatial heterogeneity. The θ values refer to those from the SLX Monte Carlo simulations. Vertical lines represent 95% confidence intervals (*dashed*) and the true indirect effect for x_2 (*solid*).

there is a risk that scholars will falsely conclude that a spatial pattern exists that is completely opposite of the true pattern.

We suspect that the inability of the SAR model to accurately predict indirect effects in some cases is due to the "phantom" higher-order and feedback effects found by the SAR model. Recall that our experiments are structured so that there are no higher-order effects; by only including one θ for each variable, the indirect effects are limited to being first-order effects (i.e., from first-order neighbors of *i*). As a result of the SAR model's infinite expansion of the spatial multiplier and the fact that weights matrices are treated as polynomials rather than higher-order contiguity matrices, the total effects of x on y will be distributed across orders of neighbors. Thus, the SAR model may find higher-order effects, whether they are there or not.

Figure 4 shows the average estimated second- (top row) and third-order (bottom row) indirect effects for x_1 from the SAR model. The three columns reflect three scenarios of weak, moderate, and strong spatial effects (under no spatial heterogeneity), and the solid vertical line shows the true higher-order effects (0). The histograms in figure 4 demonstrate that the SAR model will often incorrectly discover higher-order (and feedback) effects when they are not actually present (with the lone exception being the third-order effects under weak spatial effects). Indeed, the size of these phantom effects increases with the strength of spatial dependence. The issue is twofold. First, the DGP is strictly exogenous, so there is no reason for unit *i*'s neighbors to simultaneously influence unit *i*. Second, there are no higher-order effects within the DGP. The SAR model thus produces inferences about feedback and global effects that are not warranted by the true DGP. It is clear that there is a substantial inferential penalty for estimating an model on an SLX DGP.

APPLICATION

If scholars believe that strategic responses occupy a central place in almost any explanation of military expenditures (e.g., Palmer 1990; Powell 1999; Richardson 1960), then spatial econometric models are appropriate. The advancement of spatial econometric models means that scholars can model

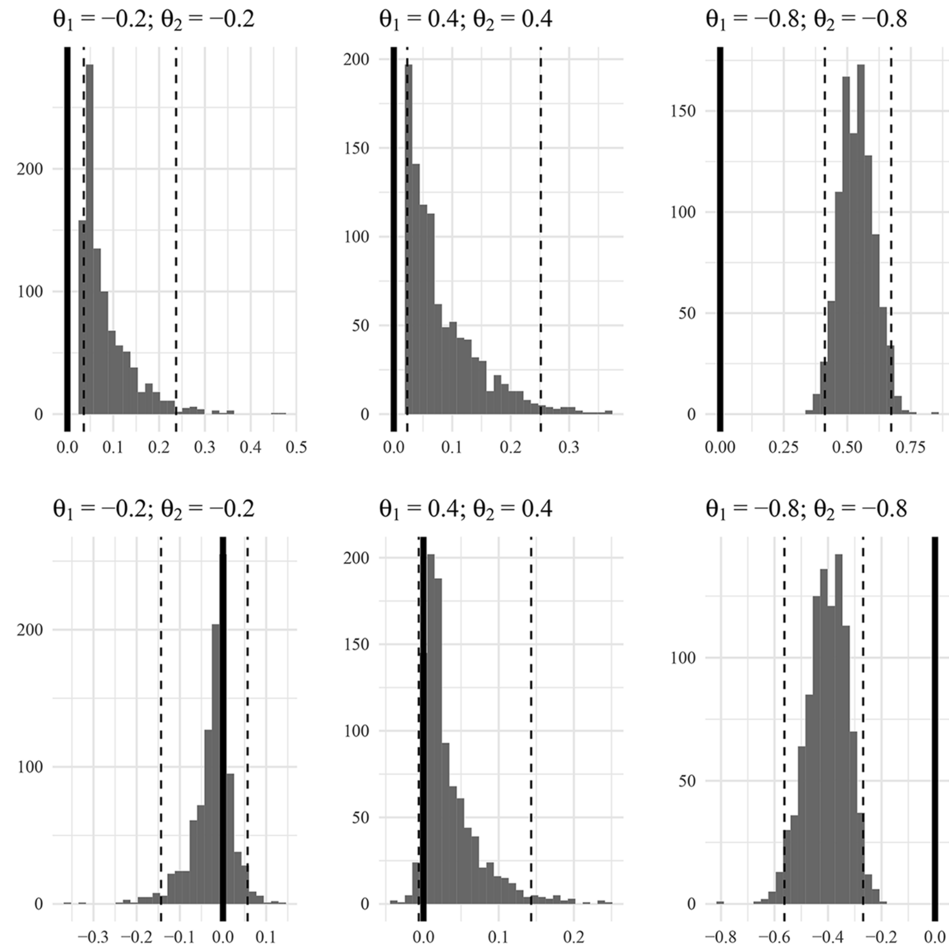


Figure 4. SAR model indicates the presence of higher-order indirect effects for x_1 when they are not there. The θ values refer to those from the SLX Monte Carlo simulations. Vertical lines represent 95% confidence intervals (*dashed*) and the true higher-order effects for x_1 (*solid*).

the precise manners in which choices (such as military expenditures) or conditions (such as instability) spill over into neighboring countries (see, e.g., Flores 2011; Plümper and Neumayer 2015; Shin and Ward 1999). Of course, theory might suggest that defense burdens in *i* directly influence defense burdens in *i*'s neighbors through a single global autocorrelation coefficient (SAR), or defense burdens in *i* influence defense burdens in *i*'s neighbors through multiple patterns of connectivities or with a temporal delay (SLX), or defense burdens in *i* are unrelated to defense burdens in *i*'s neighbors (nonspatial ordinary least squares [OLS]). Although theory can guide this decision, the true process driving military spending is unknown. We argue that a reasonable starting place given this inherent uncertainty is the SLX model since it provides a more flexible specification for estimating spatial heterogeneity and because it avoids some of the problematic assumptions hidden in SAR models.

To provide an applied example of the modeling choices that we discussed and simulated above, we collected data from 1953 to 2008 in 193 countries, both developed and developing, democratic and authoritarian. We measured defense spending in a manner similar to other scholars (Phillips 2015; Whitten and Williams 2011) as a country's defense burden, or military expenditures as a percentage of gross domestic product (GDP). We use the measure developed by Phillips (2015), who uses expenditure data from both the Stockholm International Peace Research Institute and the Correlates of War national material capabilities data (Singer and Small 1972) divided by GDP data from the Penn World Table version 6.3 (Heston, Summers, and Aten 2009). Both theories of budgeting and previous research tell us that defense burdens are highly autoregressive; the budget in one year is highly dependent on the previous year plus responses to short-term fluctuations.[20] As a result, the Defense Burden variable is nonstationary in most

20. This is a general feature of budgeting processes in organizations. As Ostrom and Marra (1986, 822) note, "given the complexity of the behavior under examination and the cognitive limitations of decision makers, it is highly unlikely that the budget decisions of any of the organizations are rebuilt from zero each year."

countries.[21] This poses a substantial problem for inferences since it increases the risk of spurious results (Granger and Newbold 1974). Since defense burdens are first-order integrated, taking the first difference makes the series stationary. To further control for the dynamics of defense burden, we include the level in the previous year (Defense Burden$_{t-1}$) and the previous change (ΔDefense Burden$_{t-1}$).[22]

For the purposes of comparing different ways of estimating spatial econometric models, we begin with a simple model of defense burden. We include Total Population (logged) to account for larger countries having larger defense burdens. We also take into account the impact of domestic and international instability by including a dichotomous variable representing the presence of a Civil War that is an armed conflict resulting in at least 25 battle-related deaths in a year (Allansson, Melander, and Themner 2017) and Interstate War (coded 1 if the hostility level of the militarized interstate dispute reaches 5). We expect that both interstate wars and civil wars will have a meaningful impact on defense burdens. A large portion of the time period under examination occurs in the Cold War, where military spending is driven by alliance commitments and actions of the two superpowers (for a review of this literature, see Goldsmith [2003], 557–59). We thus believe that the degree of responsiveness to the superpowers' changes in spending will be conditioned by the presence of an alliance with the United States (Plümper and Neumayer 2015). To take this into account, we include two interaction variables (and their lower-order terms) made up of Alliance with US (coded 1 if the state has an alliance with the United States according to the Correlates of War), ΔUS Military Expenditures, and ΔUSSR/Russian Military Expenditures. Finally, Plümper and Neumayer (2010) demonstrate that unmodeled spatial and temporal shocks can falsely suggest spatial dependence. To guard against this possibility, we include Trend (which is coded 1 for 1950 and increases each year), a dichotomous variable for 1992, and regional variables.[23] Although there are a variety of ways of specifying the connections between countries (for some alternatives, see below), for the simple preliminary analysis we use a binary, un-row-standardized contiguity weights matrix.[24]

We first compare the results of the nonspatial OLS model (model 1) to the SAR model (model 2) in table 3. The results clearly show that ignoring spatial dependence forces OLS to overestimate the immediate impacts of the covariates; almost all of the OLS coefficients are larger (either more negative or more positive) than their SAR counterparts (Ward and Gleditsch 2008, 68–69). Most notably, the nonspatial OLS would lead scholars to falsely conclude that there is no spatial dependence in states' defense burdens. The SAR model corrects this mistake. As expected, our estimate of the global spatial autocorrelation coefficient, ρ, is statistically significant and positive, indicating that the defense burdens of contiguous states are positively correlated. In other words, we have found evidence that increases to one state's defense burden simultaneously induce increases in its neighbors' defense burdens (see also Flores 2011). The result is that covariates, such as civil and interstate wars, influence defense burdens directly (via β) and indirectly through spatial diffusion (via ρ).[25]

However, as we discussed above, the various assumptions underpinning SAR models may not be consistent with how these processes actually work. Most notably, it is possible that to the extent that spillovers occur, they occur with a delay. Indeed, because of incrementalism in budgeting, defense outlays are likely to be highly path dependent and thus responsive to other countries' outlays with a temporal lag. In other words, if states are responsive to the spending patterns by other states, responses will occur one year later rather than simultaneously. Given this clear distinction, there is ample justification for estimating an SLX model with temporal lags rather than with an endogenous SAR model in which spatial effects are instantaneous.

In the case of defense spending, we believe that the SLX model offers the added benefit of allowing one to properly model spatial heterogeneity. More specifically, as outlined above, the SLX model offers a great deal of flexibility in specifying the particular spatial patterns (if any) that are theorized to govern different variables. This heterogeneity can be accomplished by determining which other countries are important spatially (via the **W**) and by allowing the strength (and sign) of the spatial clustering to vary by variable. This flexibility is particularly helpful when it comes to modeling conditional patterns of spatial dependence and varying degrees of higher-order effects. We think this is more appropriate than estimating

21. We estimate separate augmented Dickey-Fuller tests for each country with a long enough time series to get stable test statistics. In the original Defense Burden variable, we can reject the null hypothesis of a unit root at the 90% confidence level in only 29% of the countries. We reject the null hypothesis in 99% of the countries for the ΔDefense Burden$_{t-1}$ series.

22. There is a considerable debate about the casual use of lagged dependent variables to control for temporal dependence (see Wilkins [2018] for an overview of this debate).

23. In a model with annual fixed effects, 1992 is the only year that is statistically different from the others.

24. As Neumayer and Plümper (2016) note, the decision to row standardize the weights matrix should follow closely from one's theory. In the case of defense burdens, we believe that all countries are not equally influenced by their neighbors' burdens. Instead, the degree to which one state is influenced is based partly on the number of neighbors. Only those states that are contiguous by land or river are coded as neighbors, according to the Correlates of War project.

25. In the appendix, we move beyond the coefficients to provide a more in-depth exploration of total, direct, and indirect effects.

Table 3. Nonspatial OLS, SAR, and SLX Models of Neighborhood Effects on Defense Burdens

	OLS Model 1	SAR Model 2	SLX Model 3
Spatial estimates (ρ and θ):			
ρ		.07***	
		(.004)	
Contiguity × Civil War$_{t-1}$			−.11*
			(.07)
Contiguity × Interstate War$_{t-1}$			.18***
			(.07)
Ally × Interstate War$_{t-1}$			−.006
			(.03)
Contiguity × Defense Burden$_{t-1}$			.006***
			(.002)
Defense Pact × Defense Burden$_{t-1}$			.003***
			(.001)
Nonspatial estimates (β):			
Civil War$_{t-1}$	−.47***	−.45***	−.39***
	(.17)	(.16)	(.15)
Interstate War$_{t-1}$	.46***	.45***	.37**
	(.16)	(.16)	(.16)
Total Population (Logged)$_{t-1}$	.03	.03	.02
	(.02)	(.02)	(.02)
Alliance with US	−.21**	−.22**	−.15*
	(.10)	(.09)	(.08)
ΔUS Defense Burden	−.04	−.03	−.05
	(.07)	(.07)	(.08)
ΔUSSR/Russia Defense Burden	.006	−.01	−.05***
	(.02)	(.02)	(.02)
US Ally × ΔUS Defense Burden	.14	.11	.21
	(.12)	(.11)	(.13)
US Ally × ΔUSSR Defense Burden	.005	.01	.02
	(.03)	(.03)	(.02)
Annual Trend	−.007***	−.006***	−.004**
	(.002)	(.002)	(.002)
1992	1.13***	.74***	−1.26***
	(.28)	(.27)	(.21)
Defense Burden$_{t-1}$	−.12***	−.12***	−.14***
	(.006)	(.01)	(.006)
ΔDefense Burden$_{t-1}$	−.14***	−.13***	−.14***
	(.01)	(.02)	(.01)
Constant	.46**	.41*	.22
	(.22)	(.21)	(.17)
N	6,328	6,328	7,266

Note. Models include regional fixed effects. The SAR model excludes isolates.
* $p < .1$.
** $p < .05$.
*** $p < .01$.

one global spatial autocorrelation coefficient (ρ) that dictates the same spatial process for all the variables.

To demonstrate the flexibility in modeling various spatial patterns, we add the SLX variables derived from three weights matrices (contiguity, alliance commitments, and defense pacts) to reflect the following expectations:

1. We expect that civil wars in neighboring states will have an impact on states' defense burdens, but there are possibly conflicting reasons for the impact (Phillips 2015). To account for these spatial patterns, we include Contiguity × Civil War$_{t-1}$.
2. States are likely to increase their defense burdens in response to interstate wars that either increase their own instability or demand that they fulfill alliance commitments. To account for these expectations, we include Contiguity × Interstate War$_{t-1}$ and Alliance × Interstate War$_{t-1}$.
3. TLSL variables measure the influence of neighbors' defense burdens in the previous year. We believe that states will respond positively to prior military spending patterns by both contiguous states and those states with which they share a defense pact. We include Contiguity × Defense Burden$_{t-1}$ and Defense Pact × Defense Burden$_{t-1}$.

Results from the SLX model (model 2 in table 3) show that a flexible estimation technique (such as SLX) is necessary to derive accurate inferences about spatial patterns. Recall that the SAR model is limited to one specification of the neighbors' connections (via **W**) and one global spatial autocorrelation parameter (ρ).[26] The SLX model shows that both of these constraints are questionable at best. It reveals that countries' defense burdens are connected in more than one way; a single covariate (such as Defense Burden$_{t-1}$) can exhibit spatial dependence through multiple weights matrices (in this case, contiguity and defense pacts), and the nature of this connectivity may differ across covariates (e.g., Interstate War$_{t-1}$ and Defense Burden$_{t-1}$).

The second constraint of the SAR model—that the single ρ parameter accurately reflects the patterns of spatial dependence—is also potentially damaging. The coefficients for the spatial estimates (θ) are statistically different (at the 95% confidence level) in some cases, which means that, for example, the effects of contiguity are more negative for civil wars than interstate wars, and interstate war can influence defense burdens differently on the basis of either contiguity or alliance patterns.[27] Since the coefficients for the two SLX variables for Defense Burden$_{t-1}$ are in the opposite direction as the direct effect, it points to a different spillover story than the one from the SAR model. In both models, increases in i's Defense Burden at $t-1$ lead to decreases at time t for i ($\beta = -0.14$ in the SLX model, and $\beta = -0.12$ in the SAR model). In the SLX model, an increase in i's defense burden at $t-1$ positively spills over in the defense burdens of i's neighbors at t ($\theta_1 = 0.006$, and $\theta_2 = 0.003$); in the SAR model, an increase in i's defense burden at $t-1$ decreases its defense burden at t, which leads to decreases in i's neighbors' burdens because of positive spatial dependence ($\rho = 0.07$). The two models therefore give contrasting explanations of the influence of Defense Burden in neighbors, depending on whether y_j is influenced by x_i (in the SLX) or y_i (in the SAR). The fact that the direct and indirect coefficients are different signs in the SLX model implies that the SAR model (with its global autocorrelation coefficient) cannot accurately reflect the spatial patterns present in Defense Burden$_{t-1}$.

This exploration of defense spending highlights a significant advantage of SLX models related to spatial heterogeneity. In this case, states' defense burdens are spatially dependent in more than one way, and the strength of spatial dependence varies across connection type and covariate. Both of these—combined with the ability to model higher-order and conditional spatial effects—suggest that SLX models offer the key to unlocking a treasure trove of spatial inferences.[28]

PATHS FORWARD

From our discussion and findings above, it is clear that we are advocates of the simplicity offered by the SLX approach. Our goal, however, is not to suggest that the SLX is a superior model for all spatial econometric enterprises. Rather, we advocate for this approach on the basis of several important realities when it comes to the adoption of spatial models in political science. First, many applied researchers are not exposed to spatial models in a way that is complementary to their training (i.e., building from basic linear models). Likewise, many researchers interested in spatial models likely find the

26. This is the same obstacle that Flores (2011) faces. He makes a number of statements suggesting that countries' responses to their neighbors' defense burdens are conditioned by circumstances such as alliance commitments and neighborhood size. These patterns of conditional spatial dependence are quite reasonable, but Flores (2011) is unable to make those inferences from SAR models with one global autocorrelation parameter.

27. In the former case, F-tests suggest that we can reject the null of equal coefficients at the 99% confidence level; in the latter case, we can reject the null of equal coefficients at the 95% confidence level.

28. In the appendix, we first reveal that the extent to which civil wars drive military spending in nearby states depends on the particular region, and then we demonstrate how to incorporate higher-order effects in the SLX model.

current offerings of software tools insurmountable to getting started.[29] Without more accessibility, widespread adoption is less likely. How then, should applied researchers proceed when choosing the spatial model that best fits their theory?

A variety of typologies exist for categorizing spatial processes (e.g., Darmofal 2015; Vega and Elhorst 2015), but a particularly elegant typology is demonstrated by Cook et al. (2015). The authors note that spatial dependence can arise from spatial clustering in the outcomes (e.g., $\rho \mathbf{W}y$), spatial clustering in the observables (e.g., $\mathbf{WZ}\boldsymbol{\theta}$), or spatial clustering in the unobservables (e.g., $\lambda \mathbf{W}\mu$). These typologies demonstrate, among other things, that there are more than just two models to consider on the spatial econometric menu. Without repeating the full listing, several of these models have particular relevance for our discussion in this article—especially when it comes to model choice. The so-called spatial Durbin model (SDM) is in some ways a middle ground between an SLX and SAR because it allows for modeling spatial dependence both between outcomes (e.g., y_j, y_i) and between predictors and outcomes (e.g., x_j, y_i; Elhorst 2014, 9):

$$\mathbf{y} = \mathbf{X}\beta + \rho\mathbf{W}\mathbf{y} + \mathbf{WZ}\boldsymbol{\theta} + \varepsilon. \tag{14}$$

Cook et al. (2015) suggest starting with a more general spatial model and then testing the restrictions to pare down the model of unnecessary spatial components. Since it is difficult to recover estimates of the general nesting spatial model (i.e., one with estimates of all three spatial parameters) because of weak identification, Cook et al. (2015) suggest two strategies depending on the goal of the project. If there is concern about accurately characterizing the spatial process, then a spatial autocorrelation model is appropriate; if there is concern about getting accurate estimates of the effects of covariates, then the SDM is appropriate. Our approach is similar in that we advocate thinking carefully about theory and whether there are any expectations for higher-order effects, specifying the theoretically motivated model, and then testing the accuracy of those model specifications. If any differences between our approaches arise, it is the point of departure—which we advocate in most political science applications is more closely approximated by the SLX model. This is especially the case in the context of the SAR being the only relatable alternative.

Our advice for the applied researcher is to first follow theory and apply the most analogous spatial model. If, for example, the goal is to account for all spatial dependence, then the SDM or other alternatives (e.g., spatial error models) that capture and eliminate spatial noise in the unobservables may be more useful. It is often the case, however, that one's theory does not provide clear guidance and available diagnostics are uncertain. Which of the two prominent models is less prone to inferential errors as a result of misspecifying the type of dependence? In the appendix, we provide evidence that the SLX model is more robust to misspecifying the source of the spatial dependence. Additional Monte Carlo analysis based on a DGP with dependence in the error terms (one that is consistent with a spatial error model) shows that simple SLX specifications correctly recover true average effects of variables at expected rates. SAR models, however, have unsatisfactory recovery rates (i.e., much lower than 95%) for first-order indirect effects and, as a result, the total effects. The SAR is only really appropriate for those theories that specify global dependence in the outcomes among all units—something that we demonstrate is relatively rare in political science. When in doubt, political scientists are well served to start with an SLX.

CONCLUSION

The increased use of spatial models in political science is a welcome trend that involves the relaxation of the incredible assumption that there are no contagion or spillover effects across our observations. However, as our review of this literature demonstrates, there is often a substantial disjuncture between expressed theories of spatial relationships and the models that political scientists estimate. This disjuncture has been particularly prevalent in projects that use the SAR model. We have shown that the SLX model is a good starting point for spatial econometrics in political science because of its flexibility in model specification, ease in estimation, and simplicity in interpretation.

Building on previous works, we provide a comprehensive comparison of the limits of both the SAR and SLX models using simulations. Our Monte Carlo analyses indicate that, even if the true DGP is SAR, the SLX performs quite well at detecting spatial relationships at typically observed levels. The same cannot be said of the SAR when the true DGP is SLX. This is especially the case if there are even relatively small amounts of heterogeneity in terms of the spatial effects across independent variables in the DGP. Our suggested approach places the power of specification in the researcher's hands and the ability to turn unrealistically imposed assumptions from SAR models into testable propositions.

Our findings in this study lead us to a set of four considerations that researchers should keep in mind when employing spatial models to test their theories. First, researchers should

29. It is worth pointing out that this is changing with more recent releases of Stata and continual updating of the available R packages. Yet, estimation and interpretation are still far less straightforward than competing approaches of dealing with spatial heterogeneity—many of which aim to treat it is as nuisance rather than substance.

carefully consider the nature of spatial dependence in their theories and choose a model that best reflects their expectations of how these processes work. Second, unless their theories are about global dependence among outcomes, they should start with a simple SLX model and work from there toward more complex models and specifications as appropriate. Third, when they suspect that the spatial diffusion processes are heterogeneous, researchers should employ the SLX with varied specifications of matrices based on these expectations. And fourth, researchers should move beyond a cursory discussion of the direction and statistical significance of the spatial parameters to an interpretation of all the estimated effects (i.e., direct, indirect, and total).

Although we have critiqued the choice of the SAR model for many political science applications, we do not begrudge its use under the proper circumstances—theorized endogenous and global dependence among outcomes. In this article, we aimed to move the spatial revolution in political science to a place where researchers are more easily able to implement, interpret, and understand the applications. The SLX model fills this void by providing a more flexible mapping onto most political science theoretical expectations and by presenting fewer challenges to specification and interpretation.

ACKNOWLEDGMENTS

We thank Thiago Silva and Clint Swift for able research assistance. A previous version of this article was presented to the Modeling Politics and Policy Conference at Texas A&M University in May 2017 and the 1st Conference on Robustness Tests and the Empirical Analysis of Observational Data in the Social Sciences, University of Cologne Center for Comparative Politics. The authors would like to thank Eric Neumayer, Thomas Plümper, and Vera Troeger for their comments on previous versions.

REFERENCES

Allansson, Marie, Erik Melander, and Lotta Themner. 2017. "Organized Violence, 1989–2016." *Journal of Peace Research* 54 (4): 574–87.

Beck, Nathaniel, Kristian Skrede Gleditsch, and Kyle Beardsley. 2006. "Space Is More than Geography: Using Spatial Econometrics in the Study of Political Economy." *International Studies Quarterly* 50 (1): 27–44.

Bohmelt, Tobias, Lawrence Ezrow, Roni Lehrer, and Hugh Ward. 2016. "Party Policy Diffusion." *American Political Science Review* 110 (2): 397–410.

Buhaug, Halvard, and Kristian Skrede Gleditsch. 2008. "Contagion or Confusion? Why Conflicts Cluster in Space." *International Studies Quarterly* 52 (2): 215–33.

Cook, Scott J., Jude C. Hays, and Robert J. Franzese. 2015. "Model Specification and Spatial Interdependence." American Political Science Association conference paper. http://www.sas.rochester.edu/psc/polmeth/papers/Cook_Hays_Franzese.pdf.

Darmofal, David. 2015. *Spatial Analysis for the Social Sciences*. Cambridge: Cambridge University Press.

Drolc, Cody, Christopher Gandrud, and Laron K. Williams. 2019. "Taking Time (and Space) Seriously: How Scholars Falsely Infer Policy Diffusion from Model Misspecification." *Policy Studies Journal* (forthcoming).

Elhorst, J. Paul. 2014. "Linear Spatial Dependence Models for Cross-Section Data." In *Spatial Econometrics*. Dordrecht: Springer, 5–36.

Flores, Alejandro Quiroz. 2011. "Alliances as Contiguity in Spatial Models of Military Expenditures." *Conflict Management and Peace Science* 28 (4): 402–18.

Franzese, Robert J., Jr., and Jude C. Hays. 2007. "Spatial Econometric Models of Cross-Sectional Interdependence in Political Science Panel and Time-Series-Cross-Section Data." *Political Analysis* 15 (2): 140–64.

Freeman, John R., and Dennis P. Quinn. 2012. "The Economic Origins of Democracy Reconsidered." *American Political Science Review* 106 (1): 58–80.

Garcia, Blake E., and Cameron Wimpy. 2016. "Does Information Lead to Emulation? Spatial Dependence in Anti-government Violence." *Political Science Research and Methods* 4 (1): 27–46.

Gassebner, Martin, Noel Gaston, and Michael J. Lamla. 2011. "The Inverse Domino Effect: Are Economic Reforms Contagious?" *International Economic Review* 52 (1): 183–200.

Gibbons, Stephen, and Henry G. Overman. 2012. "Mostly Pointless Spatial Econometrics." *Journal of Regional Science* 52 (2): 172–91.

Goldsmith, Benjamin E. 2003. "Bearing the Defense Burden, 1886–1989: Why Spend More?" *Journal of Conflict Resolution* 47 (5): 551–73.

Goldsmith, Benjamin E. 2007. "Arms Racing in 'Space': Spatial Modelling of Military Spending around the World." *Australian Journal of Political Science* 42 (3): 419–40.

Granger, Clive W. J., and Paul Newbold. 1974. "Spurious Regressions in Econometrics." *Journal of Econometrics* 2 (2): 111–20.

Gray, Virginia. 1973. "Innovation in the States: A Diffusion Study." *American Political Science Review* 67 (4): 1174–85.

Hendry, David F. 1995. *Dynamic Econometrics*. Oxford: Oxford University Press.

Heston, Alan, Robert Summers, and Bettina Aten. 2009. Penn World Table version 6.3. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania. http://datacentre2.chass.utoronto.ca/pwt/.

Knight, Brian. 2013. "State Gun Policy and Cross-State Externalities: Evidence from Crime Gun Tracing." *American Economic Journal: Economic Policy* 5 (4): 200–229.

LeSage, James P., and R. Kelley Pace. 2009. *Introduction to Spatial Econometrics*. Boca Raton, FL: Chapman & Hall/CRC.

Midlarsky, Manus I., Martha Crenshaw, and Fumihiko Yoshida. 1980. "Why Violence Spreads: The Contagion of International Terrorism." *International Studies Quarterly* 24 (2): 262–98.

Moore, Will H., and Stephen M. Shellman. 2007. "Whither Will They Go? A Global Study of Refugees' Destinations, 1965–1995." *International Studies Quarterly* 51 (4): 811–34.

Neumayer, Eric, and Thomas Plümper. 2010a. "Galton's Problem and Contagion in International Terrorism along Civilizational Lines." *Conflict Management and Peace Science* 27 (4): 308–25.

Neumayer, Eric, and Thomas Plümper. 2010b. "Spatial Effects in Dyadic Data." *International Organization* 64 (1): 145–66.

Neumayer, Eric, and Thomas Plümper. 2016. "W." *Political Science Research and Methods* 4 (1): 175–93.

Neumayer, Eric, Thomas Plümper, and Mariaelisa Epifanio. 2014. "The 'Peer-Effect' in Counterterrorist Policies." *International Organization* 68 (1): 211–34.

Obinger, Herbert, and Carina Schmitt. 2011. "Guns and Butter? Regime Competition and the Welfare State during the Cold War." *World Politics* 63 (2): 246–70.

Ostrom, Charles W., and Robin F. Marra. 1986. "US Defense Spending and the Soviet Estimate." *American Political Science Review* 80 (3): 819–42.

Palmer, Glenn. 1990. "Alliance Politics and Issue Areas: Determinants of Defense Spending." *American Journal of Political Science* 34:190–211.

Phillips, Brian J. 2015. "Civil War, Spillover and Neighbors' Military Spending." *Conflict Management and Peace Science* 32 (4): 425–42.

Plümper, Thomas, and Eric Neumayer. 2010. "Model Specification in the Analysis of Spatial Dependence." *European Journal of Political Research* 49 (3): 418–42.

Plümper, Thomas, and Eric Neumayer. 2015. "Free-Riding in Alliances Testing: An Old Theory with a New Method." *Conflict Management and Peace Science* 32 (3): 247–68.

Powell, Robert. 1999. *In the Shadow of Power: States and Strategies in International Politics*. Princeton, NJ: Princeton University Press.

Richardson, Lewis F. 1960. *Arms and Insecurity*. Pittsburgh: Boxwood.

Shin, Michael, and Michael D. Ward. 1999. "Lost in Space: Political Geography and the Defense-Growth Trade-Off." *Journal of Conflict Resolution* 43 (6): 793–817.

Simmons, Beth A., and Zachary Elkins. 2004. "The Globalization of Liberalization: Policy Diffusion in the International Political Economy." *American Political Science Review* 98 (1): 171–89.

Singer, J. David, and Melvin Small. 1972. *The Wages of War, 1816–1965: A Statistical Handbook*. New York: Wiley.

Vega, Solmaria Halleck, and J. Paul Elhorst. 2015. "The SLX Model." *Journal of Regional Science* 55 (3): 339–63.

Ward, Michael D., and Kristian Skrede Gleditsch. 2008. *Spatial Regression Models*. Thousand Oaks, CA: Sage.

Whitten, Guy D., and Laron K. Williams. 2011. "Buttery Guns and Welfare Hawks: The Politics of Defense Spending in Advanced Industrial Democracies." *American Journal of Political Science* 55 (1): 117–34.

Whitten, Guy D., Laron K. Williams, and Cameron Wimpy. 2021. "Interpretation: The Final *Spatial* Frontier." *Political Science Research and Methods* 9 (1): 140–56.

Wilkins, Arjun S. 2018. "To Lag or Not to Lag? Re-evaluating the Use of Lagged Dependent Variables in Regression Analysis." *Political Science Research and Methods* 6 (2): 393–411.

Williams, Laron K., and Guy D. Whitten. 2015. "Don't Stand So Close to Me: Spatial Contagion Effects and Party Competition." *American Journal of Political Science* 59 (2): 309–25.