The Multiplicative Interactions Framework for Spatial Regression Models: Theory, Bias, and Interpretation

Abstract: Both spatial regression models and network analysis use spatial relationships to predict meaningful social outcomes. But while network scholars view the connectivity matrix $\boldsymbol{W}$ as theoretically interesting in its own right, the interpretation of $\boldsymbol{W}$ as "weights" by spatial regression scholars leads to the dismissal of its predictive power outside of spatial lags. I argue that $\boldsymbol{W}$ should be seen as a set of conditioning variables and that spatial regression models are multiplicative interaction models with unique, though sometimes problematic, assumptions. Because the row-sum of $\boldsymbol{W}$ is not included as a covariate in these models, scholars unnecessarily risk multiple forms of endogeneity and threaten inference. But by recognizing these models as multiplicative interaction models, scholars can use marginal effects plots to complement existing modes of interpretation. I demonstrate my claims by using Monte Carlo analysis and replicating a pair of studies in the international relations and public administration literatures.

Word Count: 8415

1 Introduction

Scholars have long recognized the influence of social networks on human behavior, creating theoretical models of how diffusion, centrality, and related concepts impact outcomes for individuals and groups. To test these theories, scholars have developed two seemingly bifurcated approaches to analyzing networks. The first of these approaches are spatial regression models, which test theories of how individual outcomes spillover to affect others in close proximity (Ward and Gleditsch 2018). Similar to how dynamic regression models temporal autocorrelation, spatial regression models focus on autocorrelation between cross-sectional units due to the inferential problems autocorrelation can pose. More recent work in the field highlights the rich interpretive possibilities these models provide through the calculation of direct, indirect, and total marginal effects from variables in the model (Whitten, Williams, and Wimpy 2019).

The second approach is network analysis, which test theories of how social networks are structured (Ward, Stovel and Stokes 2011). Rather than focus on individual outcomes, the primary purpose of network analysis is to describe the overall structure of a network relative to possible, hypothetical structures for the data. By doing so, however, network analysis describes the location of individuals and groups within a structure, such as their centrality to the network. Scholars have often used these descriptive measures of network location to predict individual outcomes in more traditional regression analysis.

These two analytic approaches have striking similarities; both rely on an $n \times n$ connectivity matrix to represent the set of spatial relationships between observations in the data. Yet there are notable differences in the view of this matrix, which might explain their largely separate development. Spatial regression models view this matrix, known as the spatial weights matrix or simply $\boldsymbol{W}$ in applied work, as simply a set of weights used to created spatial lags of the dependent variable, independent variables, and the error term. Network analysis view the matrix, known as an adjacency matrix or a sociomatrix in applied work, as a theoretical quantity of interest. Of the two approaches, the latter comports more nicely with one of the broad objectives of social science: studying how the behavior of individuals and groups is influenced by interactions with others.

This paper joins the growing research joining spatial regression models to network analysis (Hays, Kachi, and Franzese 2010). I argue that the currently accepted view of $\boldsymbol{W}$ within the spatial regression model literature needs to shift from a set of spatial weights to a set of conditioning variables within the model. This is done by introducing a new framework for understanding spatial regression models as multiplicative interactions models with some unique - and sometimes problematic - assumptions about the coefficients. Specifically, a spatially lagged term in a regression model contains n-1 multiplicative interactions without their constitutive terms. When a set of the constitutive terms are omitted from the model, captured by the row-sum of $\boldsymbol{W}$ and known as the in-degree centrality of a unit in the network analysis literature, scholars risk both omitted variable bias and measurement error that affects estimates of spatial lags and subsequent marginal effects estimates. Including this term in spatial regression models is simple, however, involving trivial calculations and standard estimation strategies. Further, by recognizing spatial regression models as multiplicative interactions models, scholars can use conditional marginal effects plots to complement existing modes of interpretation for these models (Brambor, Clark, and Golder 2006). I demonstrate my claims through both Monte Carlo analysis and a pair of replications from the international relations and public administration literatures (Gleditsch and Ward 2001, Cook, An, and Favero Forthcoming). Finally, I discuss the core contributions of the study and areas of future research.

2 Spatial Regression Models and Network Analysis

At the heart of both spatial regression models and network analysis is a connectivity matrix, hereafter called $\boldsymbol{W}$ to comport with terminology in the spatial regression literature. $\boldsymbol{W}$ is a $n \times n$ matrix that, in the purely cross-sectional context, contains all of the n(n-1) relationships in the data in the off-diagonal elements: each element (or edge) w_{ij} represents the relative influence unit j has on unit i. More extreme values of w_{ij} indicate a stronger influence of unit j on unit i; likewise, a zero indicates that there is no influence. The diagonal of this matrix represents a unit's relationship with itself and its elements, w_{ii}, are normally given a uniform treatment; in the spatial regression model context, they are constrained to zero. And while $\boldsymbol{W}$ can be a symmetric matrix, it does not have to be; w_{ij} and w_{ji} can take on different values and, when they do so, units i and j have asymmetric influence on each other. In network analysis, such matrices are called directed networks.

When spatial regression models were first developed, $\boldsymbol{W}$ almost exclusively contained geographic relationships: the distance between cities, whether countries bordered one another or not, etc. But scholars quickly learned that the spatial relationships that can be contained in $\boldsymbol{W}$ could be more than just geography (Beck, Gleditsch, and Beardsley 2006). Political scientists commonly use ideological, financial, and organizational relationships when specifying patterns of spatial dependence. Indeed, these social relationships are what network analysis tools are used to study in the social sciences, again highlighting the similarity between the two approaches.

But for all their similarities, $\boldsymbol{W}$ is used very differently in spatial regression models and network analysis. Spatial regression models account for dependencies among observations in space. Similar to how dynamic regression models capture serial autocorrelation, spatial regression models can address cross-sectional autocorrelation using a spatial autoregressive model (SAR), in the independent variable(s) using spatial lags of $\boldsymbol{x}$ model (SLX), in the disturbance term using a spatial error model (SEM), or combinations of the above processes. They model these dependencies by creating spatial lags, a post-multiplication of $\boldsymbol{W}$ and one of the above variable(s) to achieve the desired model specification.

But while time series models are relatively easy to specify, given the universally agreed upon measurement units of time and its unidimensional nature, spatial regression models are more difficult: the differing units of space complicate measurement and the multidimensionality of space complicates complete model specification. Scholars using these models spend a great deal of time attempting to find a specification of $\boldsymbol{W}$ that models the autocorrelation within the data. Any particular specification of $\boldsymbol{W}$ can be derived using either theory or empirical analysis. For many scholars, however, Tobler's first law of geography drives the specification of $\boldsymbol{W}$: "everything is related to everything else, but near things are more related than distant things" (1970).

One classic example of spatial dependence from the international relations literature relates to the determinants of democratic institutions. Scholars have long argued for policy diffusion, by which political institutions learn from their peers and adopt new policies. Gleditsch and Ward (2001, see also Beck, Gleditsch and Beardsley 2006, Ward and Gleditsch 2018) argue that a country will adapt their political institutions to match those of their peers, becoming more or less democratic. To demonstrate this, they estimate a spatial autoregressive model of a country's POLITY score. They specify $\boldsymbol{W}$ as a binary contiguity matrix, where a 1 indicates that the minimum distance between the two countries is closer than some threshold and 0 otherwise; in essence, and 1 indicates that the two countries are neighbors. They find that a country's democratic institutions are strongly autoregressive; countries tend to be more democratic if their neighbors are also democratic and vice versa.

In contrast to spatial regression models, network analysis views social networks as interesting in their own regard. In the early days, network analysis was often descriptive, calculating unit-level measures of influence within a network. Eventually, network analysis began to focus on inference, describing characteristics of the network as a whole and the probability that these characteristics would occur due to chance alone. Some of the most common features of a network that scholars investigate are the relative density of a network, the relative amount of clustering in a network, and the amount of homophiliy in a network (or the extent to which similar individuals or groups form relationships with each other). Indeed, the study of networks as a whole represents a radically different approach to social scientific research than spatial regression models, though some scholars have suggested spatial regression models that account for larger network analysis concerns (Hays, Kachi, and Franzese 2010).

But network analysis can also complement more traditional regression analysis. When scholars calculate unit-levels measures of influence, these measures can be used as variables in regression analysis. The simplest of these measures is the degree centrality of a unit, which is the sum of units relationships with all other units in the data. When networks are directional, or when the elements of $\boldsymbol{W}$ are asymmetric, the degree of a single unit can be measured in two different ways. The in-degree centrality of a unit represents the sum of the relationships that all other units have with that unit; it is calculated as $\sum_{j \neq i} w_{ij}$. The out-degree centrality of a unit represents the sum of the relationships that a unit has with all other units in the data; it is calculated as $\sum_{j \neq i} w_{ji}$.

Scholars have long used degree centrality measures as variables in political science, often without realizing that such variables can be analyzed using network analysis. In international relations, for example, scholars have long used the number of geographic neighbors a country has as a predictor for both diplomacy and war (Starr and Most 1976); in network terminology, this is a degree measure based on the geographic network of country borders. But other scholars have developed degree-based measures explicitly from a network perspective. In dyadic conflict data, scholars have used the difference of degree centrality between two countries to predict the probability they will got to war, finding that countries are less likely to go to war if there is a large gap in their degree centrality (Hafner-Burton and Montgomery 2006).

There are many other measures that can be derived from network analysis. There are plenty of other measures of centrality that can be derived, including betweenness centrality and eigenvector centrality, which have been used to predict a wide variety of outcomes including the advocacy efforts of non-governmental organizations and the success of legislators in passing amendments in the U.S. Congress (Maoz et al 2006, Fowler 2014). There are also measures completely unrrelated to the centrality of units. Hafner-Burton and Montgomery (2006) find, for instance, that countries with similar sets of relationships in the international order are less likely to go to war. Yet the simplicity of degree centrality measures make them a common starting point for applying network analysis to regression.

Spatial regression models and network analysis use very different approaches to create variables from $\boldsymbol{W}$ for regression analysis. Despite there differences, however, there is nothing mutually exclusive about the approaches. Scholars can model both spillover effects among units and the effect of a unit's characteristics on outcomes of importance to scholars. Only utilizing one of these approaches given that both can be readily applied using the same data seems like a squandering of scholastic potential. But as will become evident in the following sections, the problems of using only a single approach are more severe for some scholars than others. In particular, spatial regression models that do not use tools from network analysis will leave models vulnerable to omitted variable bias.

3 Conditional Relationships and Spatial Regression Models

While scholars using spatial regression models recognize Tobler's first law as a cornerstone in the specification of $\boldsymbol{W}$, they rarely recognize that it is a conditional hypothesis. Tobler does not state that all observations in a particular geographic space are dependent upon one another; this would imply a $\boldsymbol{W}$ matrix that contains only ones on the non-diagonals. Rather, Tobler states that all observations are dependent upon one another conditional upon the distance between pairs of observations.

Such conditional relationships are usually modelled using multiplicative interactions (Brambor, Clark, and Golder 2006). But scholars usually do not approach spatial regression models from a multiplicative interactions framework. This can be attributed to the discussion of $\boldsymbol{W}$ as a "weights" matrix in the literature, where a unit's relationships with other units only affects the dependent variable through spatial lags. But the above references to the network analysis literature undermine this view, as sophisticated transformations of $\boldsymbol{W}$ are substantial predictors of important phenomena. Thus, the current view of the $\boldsymbol{W}$ matrix seems erroneous.

Virtually all of the social sciences do not view spatial regression models as multiplicative interactions. Neumayer and Plümper (2012) provide a notable exception. They argue that "most theories of spatial policy dependence either are already inherently conditional or, if not, should be, whereas empirical models with few exceptions estimate an unconditional spatial effect." While they make a number of notable contributions, the most important one for present purposes is their view of a spatial lag as a multiplicative interaction term between a row-normalized spatial lag and the row-sum of the spatial weights matrix; in network analysis, the row-sum of the spatial weights matrix is actually just the in-degree centrality of a unit. Thus a spatial autoregressive model without row-normalization is a special case of a more general model:

$$y_i = \alpha + \rho_1 \sum_{j \neq i} \left[\frac{w_{ij}}{\sum_{j \neq i} w_{ij}} y_j \right] + \rho_2 \sum_{j \neq i} \left[\frac{w_{ij}}{\sum_{j \neq i} w_{ij}} y_j \right] * \sum_{j \neq i} w_{ij} + \rho_3 \sum_{j \neq i} w_{ij} + \beta x_i + \epsilon_i \quad (1)$$

$$\Leftrightarrow \boldsymbol{y} = \boldsymbol{\iota}\alpha + \boldsymbol{W}^{Row}\boldsymbol{y}\rho_1 + \boldsymbol{W}\boldsymbol{y}\rho_2 + \boldsymbol{W}\boldsymbol{\iota}\rho_3 + \boldsymbol{x}\beta + \boldsymbol{\epsilon} \tag{2}$$

where $\boldsymbol{\iota}$ is a vector of ones and $\boldsymbol{W}^{Row}$ is the row-normalization of $\boldsymbol{W}$.

Neumayer and Plümper's approach represents a crucial, yet underappreciated, advance in our understanding of spatial regression models. But as it stands, their approach is limited in a number of respects. First, row-normalized spatial lags are not viewed as multiplicative interactions, limiting their approach to a subset of spatial regression models. Second, they discuss the interpretation of $\boldsymbol{W}\boldsymbol{\iota}$ in a manner that could mislead readers. They claim that $\boldsymbol{W}\boldsymbol{\iota}$ can be omitted from a model when a scholar thinks that "there is no independent effect of the row sum of weights on the dependent variable." Yet viewing $\boldsymbol{W}\boldsymbol{\iota}$ as an independent effect incorrectly implies that its coefficient can be interpreted in the same way as it can in an additive linear model. Rather, ρ_3 is the effect of $\boldsymbol{W}\boldsymbol{\iota}$ when $\boldsymbol{W}\boldsymbol{y}$ is equal to zero (Brambor, Clark, and Golder 2006). Third, they do not warn scholars of the dangers of incorrect model specification. While they state that "in general it will be better to free the coefficients ρ_1 and ρ_3 and to estimate [equation 2 rather than a more restricted model]," they do not mention that omitting $\boldsymbol{W}\boldsymbol{\iota}$ (or $\boldsymbol{W}^{Row}\boldsymbol{y}$) from the model when its coefficient is non-zero will result in omitted variable bias in the estimate of ρ_2.

4 The Multiplicative Interactions Framework of Spatial Regression Models

I argue that the conditional relationships modelled using spatial regression are inherently multiplicative interactions models with some restrictive assumptions. This is true no matter the particular spatial model estimated, although the details of the assumptions made when estimating each differ slightly in terms of implications and enormously in terms of mathematical proof. I demonstrate the logic of this worldview using the simplest spatial model possible, the SLX model with a single independent variable, leaving additional derivations to the reader. Consider the following model:

$$\boldsymbol{y} = \boldsymbol{\iota}\alpha + \boldsymbol{x}\beta + \boldsymbol{W}\boldsymbol{x}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{3}$$

To better see how multiplicative interactions work in these models, I express this model in expanded scalar notation:

$$y_i = \alpha + x_i\beta + \sum_{j \neq i} w_{ij}x_j\theta + \varepsilon_i \tag{4}$$

$$\Leftrightarrow y_i = \alpha + x_i\beta + w_{i1}x_1\theta + w_{i2}x_2\theta + ... + w_{in}x_n\theta + \varepsilon_i \tag{5}$$

I offer a view of spatial regression models that complements existing understanding, but relies on two key observations about the elements of $\boldsymbol{W}$. First, each x_j in the model is a variable. In a sample of cross-sectional data, each x_j is an observed value of variable $\boldsymbol{x}$ and is thus a constant. From a frequentist perspective, however, each x_j can also be thought of as an observation from a larger superpopulation of counterfactuals and is thus a variable. Extending beyond a single cross-section, each x_j also represents a variable in it of itself. For many data-generating processes, this can be as part of a time-series x_{jt} or larger cross-section x_{js}.

Second, each w_{ij} is also a variable. This claim is nothing novel in the network analysis literature, which has viewed relationships between units as variables to be modelled since at least the adoption of inferential approaches to networks (Ward, Stovel, and Sacks 2011). Those scholars who use spatial regression models may be initially skeptical of this view, seeing each w_{ij} as indexing the spatial relationships within the model. But a variable is simply a function that maps all possible outcomes of some event onto the real number line.

Each column of the spatial weights matrix w_{ij} meets this definition, where the event is the possible spatial relationships some observation j can have with other observations and the elements of column w_j contain the numerical representations of observed events.

With these two observations explicitly made, my larger view of spatial regression models is that each spatial component in the model contains n-1 multiplicative interactions, in which the effect of each x_j on y_i is conditioned by the spatial relationship between the two observations (as defined by w_{ij}).[1] The terms are subject to the linear restriction that they all have the same coefficient, θ. This restriction, combined with the assumption that each w_{ij} accurately reflects the spatial relationships in the data, helps scholars confront the inherent identification problem within spatial models that the data contains n(n-1) potential spatial relationships while providing only n observations (Halleck Vega and Elhorst 2015).

How does this interpretation fit an applied example? As discussed earlier, Gleditsch and Ward propose a spatial autoregressive model for a country's democratic institutions using a binary contiguity matrix. In essence, they argue that a country's level of democracy is a function of the level of democracy in other countries conditional upon two countries being a neighbor. France's level of democracy will be influenced by the democratic institutions in Germany because the two are neighbors; it will not be influenced by the level of democracy in Japan, however, because the two are not neighbors.

But while spatial models are inherently multiplicative interactions models, they omit the constitutive terms w_{ij} and x_j from the model. This functionally constrains the effect of these

[1]Spatial regression models do not contain the relationship $w_{ii}x_i$ by design, as the spatial relationship between an observation and itself is infinite in comparison to other spatial relationships. Instead, the relationship between x_i and y_i is captured with β, and all elements w_{ii} are forced to be zero.

variables to zero. This is a dangerous constraint, as bias will permeate the interaction terms $w_{ij}x_j$ if this constraint is incorrect (Brambor, Clark, and Golder 2006). For the simple SLX model being examined, the true spatial model being estimated is:

$$y_i = \alpha + x_i\beta + w_{i1}0 + x_10 + w_{i1}x_1\theta + w_{i2}0 + x_20 + w_{i2}x_2\theta + ... + w_{in}0 + x_n0 + w_{in}x_n\theta + \varepsilon_i \quad (6)$$

Of course, omitting variables from a model does not necessarily lead to bias. Examining whether the constitutive terms are correlated with the dependent variable and the interaction terms will be the subject of the next sections. But assume for the moment that a scholar wanted to include these constitutive terms in the model. To do so, the scholar could extend the identification strategy for the multiplicative interaction terms to the constitutive terms by subjecting them the linear restriction that each set of terms are equal. For the SLX model being examined, group the like terms together and factor them:

$$y_i = \alpha + x_i\beta + w_{i1}0 + w_{i2}0 + ... + w_{in}0 + x_10 + x_20 + ... + x_n0 + w_{i1}x_1\theta + w_{i2}x_2\theta + ... + w_{in}x_n\theta + \varepsilon_i \quad (7)$$

$$\Leftrightarrow y_i = \alpha + x_i\beta + \sum_{j \neq i} w_{ij}0 + \sum_{j \neq i} x_j0 + \sum_{j \neq i} w_{ij}x_j\theta + \varepsilon_i \quad (8)$$

and then instead of constraining the two variables to be zero, allow them to have their own coefficient:

$$y_i = \alpha + x_i\beta + \sum_{j \neq i} w_{ij}\theta_1 + \sum_{j \neq i} x_j\theta_2 + \sum_{j \neq i} w_{ij}x_j\theta_3 + \varepsilon_i \quad (9)$$

$$\Leftrightarrow \boldsymbol{y} = \boldsymbol{\iota}\alpha + \boldsymbol{x}\beta + \boldsymbol{W}\boldsymbol{\iota}\theta_1 + \boldsymbol{K}\boldsymbol{x}\theta_2 + \boldsymbol{W}\boldsymbol{x}\theta_3 + \boldsymbol{\varepsilon} \quad (10)$$

where $\boldsymbol{K}$ is a spatial weights matrix in which all of its non-diagonal elements are equal to 1.

5 The Necessity of Including the Row-Sum of W as a Variable

In order to determine whether scholars can safely omit these constitutive terms, one must determine their substantive meaning. I do this with each set of terms in turn. In the general SLX model in equations 9 and 10, $\boldsymbol{Kx}$ is a spatial lag of $\boldsymbol{x}$ with all non-diagonal elements equal to one. Its coefficient can be interpreted as the effect of x_j on y_i when w_{ij} equals zero. As stated earlier, scholars specify a given w_{ij} to equal zero when they expect x_j to not have an effect on y_i. Thus, the omission of $\boldsymbol{Kx}$ from the model seems to be a benign one when $\boldsymbol{W}$ is correctly specified since there is a theoretical reason to expect its coefficient to equal zero.

Indeed, Vande Kamp (Forthcoming) shows that $\boldsymbol{Kx}$ functions as a specification test of $\boldsymbol{W}$. Deriving his conclusions from a measurement error perspective, he shows that the coefficient of $\boldsymbol{Kx}$ will be statistically indistinguishable from zero when $\boldsymbol{W}$ is correctly specified but will be statistically distinguishable from zero when the non-diagonal elements of $\boldsymbol{W}$ are inflated by a constant. While this "K test" is still the subject of ongoing research, I direct interested readers to that manuscript. But for present purposes, $\boldsymbol{Kx}$ can be safely omitted from the model if $\boldsymbol{W}$ is correctly specified.

Unlike with $\boldsymbol{Kx}$ and its coefficient, there is no theoretical reason to believe that the coefficient of $\boldsymbol{W\iota}$ is equal to zero. As mentioned earlier, previous studies using degree centrality as a variable shows that the parameter has a straightforward interpretation when used in an additive model. When used in a multiplicative interaction, there is no reason to believe its coefficient will be equal to zero. Excluding $\boldsymbol{W\iota}$ would risk omitted variable bias given the well demonstrated correlation between multiplicative interactions and their constitutive terms (Braumoeller 2004, Brambor, Clark, and Golder 2006).

Beyond the substantive reasons to believe that the coefficient of $\boldsymbol{W\iota}$ will be non-zero, the risk of measurement error in variables being spatially lagged also necessitates its inclusion. Suppose that a scholar wants to estimate the SLX model from equation 10 but cannot observe $\boldsymbol{x}$. Rather, the scholar observes $\boldsymbol{x*}$ which is the variable $\boldsymbol{x}$ inflated by a constant c:

$$\boldsymbol{x}* = \boldsymbol{x} + \boldsymbol{\iota} c \tag{11}$$

If the scholar attempts to estimate the model, it results in the following endogeneity:

$$\boldsymbol{y} = \boldsymbol{\iota}\alpha + (\boldsymbol{x} + \boldsymbol{\iota}c)\beta + \boldsymbol{W}(\boldsymbol{x} + \boldsymbol{\iota}c)\theta + \boldsymbol{\varepsilon} \tag{12}$$

$$\Leftrightarrow \boldsymbol{y} = \boldsymbol{\iota}\alpha + \boldsymbol{x}\beta + \boldsymbol{\iota}c\beta + \boldsymbol{W}\boldsymbol{x}\theta + \boldsymbol{W}\boldsymbol{\iota}c\theta + \boldsymbol{\varepsilon} \tag{13}$$

$$\Leftrightarrow \boldsymbol{y} = \boldsymbol{\iota}(\alpha + c\beta) + \boldsymbol{x}\beta + \boldsymbol{W}\boldsymbol{x}\theta + (\boldsymbol{\varepsilon} + \boldsymbol{W}\boldsymbol{\iota}c\theta) \tag{14}$$

Here, we see that part of the endogeneity from using a measure of $\boldsymbol{x}$ inflated by a constant is captured in the intercept term α. But an additionally endogeneity, $\boldsymbol{W\iota}c\theta$ is stuck in the error term and will almost certainly be correlated with the spatial lag. This endogneity, if left unchecked, will lead to inconsistent estimates and can hamper inferences. If, however, $\boldsymbol{W\iota}$ were also included in the regression model, then the endogeneity would be transferred to its coefficient:

$$\boldsymbol{y} = \boldsymbol{\iota}(\alpha + c\beta) + \boldsymbol{x}\beta + \boldsymbol{W}\boldsymbol{\iota}\theta_1 + \boldsymbol{W}\boldsymbol{x}\theta_3 + (\boldsymbol{\varepsilon} + \boldsymbol{W}\boldsymbol{\iota}c\theta) \tag{15}$$

$$\Leftrightarrow \boldsymbol{y} = \boldsymbol{\iota}(\alpha + c\beta) + \boldsymbol{x}\beta + \boldsymbol{W}\boldsymbol{\iota}(\theta_1 + c\theta_3) + \boldsymbol{W}\boldsymbol{x}\theta_3 + \boldsymbol{\varepsilon} \tag{16}$$

Indeed, this proof mirrors the one given by Braumoeller (2004) when advocating that constitutive terms should be included in regression models even when there is a theoretical reason to believe they are zero. Omitting $\boldsymbol{W\iota}$ in spatial regression models, then, risks endogeniety from both omitted variable bias and measurement error.[2] But such risks are unnecessary, as calculating $\boldsymbol{W\iota}$ and its incorporation into spatial regression models presents no econometric difficulties.

6 Interpretational Gains from a Multiplicative Interactions Framework

Thus far, the multiplicative interactions framework of spatial regression models has highlighted the need to include $\boldsymbol{W\iota}$ as a variable. This was motivated from endogeneity concerns. But the framework also allows new possibilities for interpreting spatial regression models that complement existing techniques. A non-spatial regression model assumes that a given unit's observation x_i only has a direct effect on that unit's dependent variable y_i, captured in the coefficient β. A one-unit increase in x_i will cause a β unit increase in y_i. This interpretation remains the same in SEM models.

Spatial regression models allow the possibility, however, that x_i has an indirect effect on other units y_j, which is only possible through the creation of spatial lags. A summary of these direct and indirect effects are given in Table 1. In a SLX model, hypothesis testing is relatively straightforward: the direct effect of x_i on y_i is captured by its coefficient, β, while the indirect effect is captured by the coefficient of the spatial lag of x_i, θ. But while β

[2] The one exception is for spatial regression models that specify $\boldsymbol{W}$ using a K-Nearest Neighbor specification. In these circumstances, $\boldsymbol{W\iota}$ will result in a constant that should already be included in the model.

Table 1: Direct and Indirect Effects of a Variable Using Different Spatial Regression Model Specifications

Model	Direct Effect	Indirect Effect
OLS/SEM	β	0
SLX	β	θ
SAR	Diagonal elements of $(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\beta$	Non-diagonal elements of $(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\beta$

Reproduced from Halleck Vega and Elhorst (2015). Specifications involving permutations of these models are omitted for simplicity.

has a clear substantive interpretation, direct interpretation of the SLX coefficient is difficult. A one-unit increase in a spatial lag of x_i can result from a change in one or more x_j, one or more w_{ij}, or a combination of the two. This contrasts with the standard regression context, in which a one-unit increase in an independent variable often has a clean substantive interpretation.

In a SAR model, interpretation becomes considerably more complicated. This complication arises from the nature of SAR models, which allow for global spillovers of effects. Suppose a scholar wanted to interpret the following (restricted) SAR model:

$$\boldsymbol{y} = \boldsymbol{\iota}\alpha + \boldsymbol{x}\beta + \boldsymbol{W}\boldsymbol{y}\boldsymbol{\rho} + \boldsymbol{\varepsilon} \tag{17}$$

Because $\boldsymbol{W}\boldsymbol{y}$ necessarily contains $\boldsymbol{W}\boldsymbol{x}$, the effect of the independent variable $\boldsymbol{x}$ on $\boldsymbol{y}$ is unclear. Thus scholars use this transform this equation to isolate $\boldsymbol{y}$ on the right hand side which, in turn, produces the core matrix of interest from Table 1:

$$\boldsymbol{y} - \boldsymbol{W}\boldsymbol{y}\boldsymbol{\rho} = \boldsymbol{\iota}\alpha + \boldsymbol{x}\beta + \boldsymbol{\varepsilon} \tag{18}$$

$$(\boldsymbol{I} - \rho\boldsymbol{W})\boldsymbol{y} = \boldsymbol{\iota}\alpha + \boldsymbol{x}\beta + \boldsymbol{\varepsilon} \tag{19}$$

$$\boldsymbol{y} = (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{\iota}\alpha + (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{x}\beta + (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{\varepsilon} \tag{20}$$

The matrix $(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}$ is called the spatial multiplier matrix. This matrix contains an infinitely repeating series of $\boldsymbol{W}$ matrices of the following form:

$$(\boldsymbol{I} - \rho \boldsymbol{W})^{-1} = \boldsymbol{I} + \rho \boldsymbol{W} + \rho^2 \boldsymbol{W}^2 + \rho^3 \boldsymbol{W}^3 + ... \tag{21}$$

This expansion demonstrates the difficulties with interpreting SAR models. $\boldsymbol{I}$ is an identity matrix and, because it has zeroes along the non-diagonal elements, only returns the direct effect of x_i on y_i when post-multiplied by $\boldsymbol{x}\beta$. Similarly, when $\rho\boldsymbol{W}$ is post-multiplied by $\boldsymbol{x}$, it only returns the indirect effects of x_j on y_i because $\boldsymbol{W}$ has zeroes along the diagonal. The higher order $\boldsymbol{W}$ matrices, however, contain both direct and indirect effects due to the feedback present in spatial terms. For the second order $\boldsymbol{W}$ matrix, for example, unit i affects its neighbors, including unit j, which then affects unit j's neighbors, which also include unit i. This feedback makes interpretation of direct and indirect effects in SAR models extremely difficult and, as a result, each unit has its own unique direct effect as well as a unique indirect effect on every other unit in the sample. Of course, $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1}\boldsymbol{x}\beta$ contains an extraordinarily large amount of information; this has lead many scholars to call for the presentation of average direct and indirect effect estimates (Whitten, Williams, and Wimpy 2019). Standard errors for all of these marginal effects have to be approximated.

These difficulties in interpreting spatial regression models are unlikely to go away, particularly for the SAR specification. But the multiplicative interactions framework for these models also opens up additional interpretive possibilities that may be attractive to scholars. Recall that while spatial regression models contain n-1 multiplicative interactions, they are linearly restricted to have the same coefficient(s). Thus, one can find the marginal effect an increase in any w_{ij} on y_i in an SLX model, for example, by taking the partial derivative of equation 9 with respect to w_{ij}:

$$\frac{\partial y_i}{\partial w_{ij}} = \theta_1 + \theta_3 x_j \tag{22}$$

This conditional marginal effect is relatively straightforward to interpret using marginal effects plots (Brambor, Clark, and Golder 2006). Indeed, such a plot would concisely display a marginal effect that is generalizable to a change in the spatial relationship between any two units in the data. While certainly useful in the SLX context, such a presentation is a notable simplification over current SAR interpretations using the spatial multiplier. A marginal effect plot avoids the need for presenting either the enormous amount of information in $(\boldsymbol{I}-\rho\boldsymbol{W})^{-1}\boldsymbol{x}\beta$ or the overly simplified information contained in averages of direct and indirect effects. Indeed, this approach also includes a straightforward way to calculate standard errors for marginal effects that avoids the need for simulation. The marginal effect of x_j on y_i can also be presented using marginal effects plots. By doing so, it allows for a straightforward, substantive interpretation of the coefficients of spatial lags that is not currently possible, as described above.

Of course, using marginal effect plots will not totally simplify the interpretation of SAR models. The use of marginal effect plots is necessarily restricted to the effect of w_{ij} and y_j on y_i. For other variables in the model, the use of the spatial multiplier is still necessary in order to draw substantive conclusions about the direct and indirect effects of a variable. The advances described here are a complement, not a substitute, for current practices in the literature.

One might express hesitation with the use of marginal effects plots for spatial regression models due to the difficulty in imagining a counterfactual. Spatial regression models commonly employ some function of geographic space when specifying $\boldsymbol{W}$, and natural geography can be highly autoregressive. Rivers change course over hundreds of years, and the shifting of tectonic plates happens over the course of millions. In reference to our applied example, Russia will never border Ireland, so it is foolish to imagine a counterfactual in which they are neighbors. But given this difficulty, there are at least two responses that justify the use of marginal effects plots for spatial regression models.

First, and as noted earlier, space is more than geography; scholars have applied spatial regression models to contexts beyond physical space (Beck, Gleditsch, and Beardsley 2006). Referencing our applied example of democratic institutions, Russia might not border Ireland geographically, but it does trade with them and could potentially enter into an alliance with them. It is pretty easy to imagine a world in which Ireland's political institutions were conditional upon Russia's, even if not in the context of political geography. Criticizing the multiplicative interactions framework due to one class of spatial regression models seems overly limiting.

Second, scholars rarely use natural geography when specifying spatial regression models. Instead, scholars usually employ political geography, or the spatial structures that govern politics. For example, studies of countries in the international relations and comparative politics literatures commonly specify $\boldsymbol{W}$ to be a contiguity matrix of whether two countries border each other. Borders are political constructs. While borders can be influenced by natural geography, it is not the exclusive determinant of borders; war, rebellion, and diplomacy can change the political landscape and influence whether two countries border each other (Starr and Most 1976). Similarly, countries may divide and redivide their territory into different subnational political units in order to facilitate governance. For example, the American states regularly engage in redistricting of national congressional districts following the country's decennial census and subsequent reapportionment. While natural geography is a concern, it is far from the only one. Such border changes affect not only the number of neighbors a political unit has, but also the length of shared borders, the distance between the borders of political units, and the centroids of political units; all of these factors can influence the construction of $\boldsymbol{W}$.[3]

[3] One can imagine similar processes for capital cities and other types of political geography.

To expand upon this point, I reference the borders of Poland in the first half of the twentieth century. In the aftermath of World War I, the Greater Polish Uprising helped Poland secure independence from Russia. When conflict finally ceased in 1921, the officially recognized Polish state bordered six countries: Romania, Czechoslovakia, Germany, Lithuania, Latvia, and the soon to be Soviet Union. During the post-war period, Poland was involved in a number of border skirmishes that changed its territorial borders but did not change its number of neighbors. At the start of World War II, however, Hungary annexed the easternmost part of Czechoslovakia in early 1939, becoming a new neighbor of Poland and changing their count to seven. Later that year, Germany and the Soviet Union successfully invaded Poland, de facto eliminating its existence as a state. When the Allies won the War in Europe in 1945, Poland was restored as an independent state. The number of states it neighbored dropped tremendously, however, as the Soviet Union annexed Lithuania, Latvia, and portions of Southern Poland that included the country's border with both Hungary and Romania. During the beginning of the Cold War, then, Poland had only three neighbors. While the evolution of Poland's borders are not typical for states in the twenty first century, particularly as interstate war declined and most former colonies secured independence from their occupiers, it demonstrates that political geography can and does change. After all, Poland's borders today are substantially different from what they were during the Cold War. Borders and other political geographic constructs, then, should be seen as any other variable within the political science literature.

7 Considerations for Normalization Procedures

Thus far, the discussion of spatial regression models as multiplicative interactions has not made any restrictions on the $\boldsymbol{W}$ matrix used. Oftentimes, however, scholars normalize their $\boldsymbol{W}$ matrices in order to aid in the diagnosis of stationarity or non-stationarity in the data. These normalizations can complicate the interpretation of spatial regression models as multiplicative interactions. Therefore, I consider the two most common normalizations now.

Spectral normalization dividing the elements of $\boldsymbol{W}$ by the largest eigenvalue of $\boldsymbol{W}$. This bounds the upper limit of the spatial lag's coefficient to one; a spatial lag indistinguishable from one is non-stationary. Because $\boldsymbol{W}$ is normalized by a scalar, the subsequent coefficients are proportional to the true coefficients of the data-generating process. Importantly, however, is that the normalization process means that marginal effects can only be derived after a small amount of work from the scholar. In order to recover the coefficients from the data-generating process, a scholar has two options. The first is to estimate the model twice: once using the normalized matrix to check for nonstationarity and once using the original matrix in order to generate marginal effects. The other is to only estimate a model using spectral normalization and simply divide the coefficients of the model by the maximum eigenvalue used to normalize $\boldsymbol{W}$ and then calculate marginal effects. This second solution only works because $\boldsymbol{W}$ is normalized by a scalar.

Row normalization divides each element of $\boldsymbol{W}$ by its row sum. This bounds the upper limit and the lower limit of the spatial lag's coefficient to one and negative one, respectively; these values have the same interpretation with regard to nonstationarity. Unlike with spectral normalization, however, row normalization fundamentally changes the resulting of the matrix because it is not adjusted by a scalar but, instead, is adjusted by the vector of row sums of the matrix (Neumayer and Plümper 2016). This leads to two considerations for scholars.

First, row normalizing $\boldsymbol{W}$ may result in $\boldsymbol{W\iota}$ being a constant. After row normalization, $\boldsymbol{W\iota}$ will equal one for any unit that is not an island; that is, for any unit that has at least one non-zero spatial relationship with another unit of analysis and does not have zeroes across its entire row of $\boldsymbol{W}$. As a corollary, $\boldsymbol{W\iota}$ will equal zero for any unit that is an island. Thus for samples with islands, $\boldsymbol{W\iota}$ will be equivalent to a dummy variable in which a zero indicates that a unit is an island. For samples without islands, however, $\boldsymbol{W\iota}$ will result in a vector of ones. In these situations, the inclusion of $\boldsymbol{W\iota}$ should already occur from including the constant and scholars should not worry about bias from its exclusion.

Regardless of whether $\boldsymbol{W\iota}$ is already included in the model or not, the interpretation of spatial regression models in the multiplicative interactions framework is severely limited when row normalizing. The individual elements of $\boldsymbol{W}$ are considerably more difficult to interpret from this perspective, as a larger value of any w_{ij} may be indicative of a larger w_{ij} prior to row normalization, a smaller row sum $\sum_{\forall i} w_{ij}$, or both. While marginal effects plots could still be presented, they would not have as natural an interpretation as spectral normalized matrices. This presents out another limitation of row normalization that scholars should consider when specifying their models.

8 Monte Carlo Design

I have advanced two interrelated arguments concerning the estimation of spatial regression models. First, the information in $\boldsymbol{W}$ may be predict important phenomena outside of their role in creating spatial lags and warrants the inclusion of $\boldsymbol{W\iota}$ as an independent variable in spatial regression models. Second, omitting $\boldsymbol{W\iota}$ when it is part of the data-generating process of $\boldsymbol{y}$ risks omitted variable bias in spatial lags because the two are necessarily correlated. If these two arguments are true, then spatial regression models are inherently multiplicative interactions models, a framework which allows for new methods of interpreting spatial regression models by deriving conditional marginal effects and plotting them.

The validation of these two different points requires two different approaches. The question of whether $\boldsymbol{W\iota}$ predicts important social phenomena is an empirical one. To that end, a pair of empirical applications is presented towards the end of the paper to demonstrate the necessity of this multiplicative interactions framework within social science. In contrast, the question of whether omitting $\boldsymbol{W\iota}$ when its coefficient in the data-generating process is non-zero causes bias is an analytic one. To demonstrate that this is true, I rely on Monte Carlo simulations.

The design of the Monte Carlo analysis is important in demonstrating my point. First, the Monte Carlo analysis must feature the SLX model. Spatial lags of the dependent variable are necessarily correlated with all of the other independent variables in the model. If $\boldsymbol{W\iota}$ is part of a variable's data-generating process, it will be correlated with $\boldsymbol{Wy}$ even if the multiplicative interactions framework is incorrect; any simulations featuring SAR models will therefore be unconvincing. In contrast, spatial lags of independent variables are not necessarily with all other variables in the model. Thus, demonstrating with an SLX model that $\boldsymbol{W\iota}$ and $\boldsymbol{Wx}$ are necessarily correlated will strongly support the multiplicative interactions framework.

The design of the Monte Carlo analysis also must demonstrate that all spatial regression models are multiplicative interactions models. This stands in contrast to the approach of Neumayer and Plümper (2012), who argue spatial regression models that do not row-normalize $\boldsymbol{W}$ are multiplicative interactions models (and no further). The Monte Carlo analysis presented here use both untransformed $\boldsymbol{W}$ matrices as well as row-normalized matrices with islands. This is necessary to differentiate my argument from Neumayer and Plümper; I propose that even row-normalized spatial lags should be viewed as multiplicative interaction models, while Neumayer and Plümper make no such proposition.

My Monte Carlo analysis demonstrates that $\boldsymbol{W\iota}$ is necessarily correlated with $\boldsymbol{Wx}$, where $\boldsymbol{x}$ is a random variable that is not spatially autoregressive. I generate one thousand observations of $\boldsymbol{x}$ from a uniform distribution, $U \sim (-1, 1)$ and, after generating $\boldsymbol{W}$, calculate both $\boldsymbol{Wx}$ and $\boldsymbol{W\iota}$. I then calculate the correlation between the two variables and its 95% confidence interval, recording each. This process is repeated one thousand times, with the average correlation and the proportion of correlations statistically distinguishable from zero presented in the paper. If spatial regression models are multiplicative interactions models, then the correlation between $\boldsymbol{Wx}$ and $\boldsymbol{W\iota}$ should be statistically distinguishable from zero at a rate greater than chance (or five percent).

I generate a number of different $\boldsymbol{W}$ specifications in order to determine whether the correlation of $\boldsymbol{W\iota}$ and $\boldsymbol{Wx}$ is dependent upon the structure of $\boldsymbol{W}$. I generate specifications that are both likely to be encountered by scholars and that push the bounds of what are theoretically possible specifications of $\boldsymbol{W}$ (Neumayer and Plümper 2016). Most specifications rely on each observation of $\boldsymbol{x}$ being randomly located in a bounded two-dimensional space, with $\boldsymbol{W}$ being calculated based on the Euclidean distance between the two variables; this results in a series of symmetric $\boldsymbol{W}$ matrices. I create matrices based on linear distance, inverse distance, a small sphere of influence, and a large sphere of influence, leading to a combination of binary contiguity matrices and more general matrices. Additionally, non-symmetric $\boldsymbol{W}$ matrices are generated by randomly generating the non-diagonal elements of $\boldsymbol{W}$ from a uniform distribution. I compute two classes of non-symmetric $\boldsymbol{W}$: those with only positive values along the non-diagonal elements, $U \sim (0, 1)$, and those with positive and negative values, $U \sim (-0.25, 1.25)$. For all specification types, a new $\boldsymbol{W}$ is generated for each set of observations generated. The Monte Carlo results, then, are not dependent on a few idiosyncratic specifications of $\boldsymbol{W}$. For this set of $\boldsymbol{W}$ specifications, I perform two sets of analysis: one using the untransformed matrices and one using row-normalized matrices. With the row-normalized matrices, one hundred of the observations are islands.

9 Monte Carlo Analysis

The results of the simulations are presented in Tables 2 and 3. The average correlations using untransformed matrices between $\boldsymbol{Wx}$ and $\boldsymbol{W\iota}$, reported in Table 2, are functionally zero in all sets of analysis, likely due to the face that both $\boldsymbol{x}$ and $\boldsymbol{W}$ are randomly generated for each set of observations. Importantly, the proportion of cases statistically distinguishable from zero are high. Every time a symmetric $\boldsymbol{W}$ is used, the proportion is between 70% and 90%. This provides strong evidence that $\boldsymbol{Wx}$ and $\boldsymbol{W\iota}$ are correlated. For non-symmetric matrices, the proportion of statistically significant cases is much lower: about 20% for each one. Still, this rate is four times greater than we expect to observe from chance alone. Even in these general cases, there is evidence that $\boldsymbol{Wx}$ and $\boldsymbol{W\iota}$ may be correlated. This provides evidence for my argument that spatial regression models are multiplicative interactions models.

Table 2: Average Correlation Between $\boldsymbol{Wx}$ and $\boldsymbol{W\iota}$

$\boldsymbol{W}$	Average Correlation	Proportion Statistically Significant
Linear Distance	-0.01	0.90
Inverse Distance	0.00	0.71
Small Sphere of Influence	0.00	0.84
Large Sphere of Influence	0.01	0.89
Random $\boldsymbol{W}$, + Values	-0.00	0.19
Random $\boldsymbol{W}$, +/- Values	0.00	0.18

For the set of correlations using a row-normalized weights matrix, listed in Table 3, the average correlations are again functionally zero. The proportion of statistically significant cases, however, are much higher. While the majority of correlations are statistically significant for the symmetric matrices, roughly 90% of the correlation are significant for non-symmetric matrices. This provides evidence that even row-normalized spatial lags should be viewed as multiplicative interactions models.

Table 3: Average Correlation Between $\boldsymbol{Wx}$ and $\boldsymbol{W\iota}$ - Row-Normalized with Islands

$\boldsymbol{W}$	Average Correlation	Proportion Statistically Significant
Linear Distance	-0.01	0.97
Inverse Distance	0.00	0.72
Small Sphere of Influence	0.00	0.61
Large Sphere of Influence	-0.01	0.66
Random $\boldsymbol{W}$, + Values	-0.00	0.92
Random $\boldsymbol{W}$, +/- Values	0.00	0.88

10 Empirical Example 1: Determinants of Democracy

To demonstrate my proposed methods, I replicate a pair of studies using the spatial analysis suite in Stata 15. The first study is our running example of the determinants of democratic institutions, by Gleditsch and Ward (2001). The model has been widely used in advances of our understanding of spatial regression models and as a teaching device (Beck, Gleditsch and Beardsley 2006, Ward and Gleditsch 2018). To replicate their study, I use the data files from their instructional text on spatial regression models (Ward and Gleditsch 2018). They regress a country's POLITY score in 2014 on logged GDP per capita and a spatial autoregressive parameter.

They specify $\boldsymbol{W}$ as a binary contiguity matrix; any two countries whose borders are 400 kilometers apart or less are coded as a 1, all other spatial relationships are coded as a zero. They exclude islands, or countries with no borders, from their analysis. They also row-normalize their weights matrix. Instead of row-normalizing, I choose to spectral normalize the weights matrix for two distinct reasons. Theoretically, row-normalization changes the interpretation of the weights matrix in a way that might bias results and harm inferences (Neumayer and Plümper 2016). Methodologically, $\boldsymbol{W\iota}$ is a constant when estimated using a row-normalized weights matrix without island observations, as is the case here.[4]

Table 4 contains a replication of their analysis along with a model including $\boldsymbol{W\iota}$, as well as including estimates of the average direct and average indirect effect of Logged GDP. As one can see, the inclusion of $\boldsymbol{W\iota}$ is important; it is the model's largest coefficient and statistically significant (p<0.000). Its inclusion also changes the estimate of the spatial lag's coefficient. While the original model indicated that democratic institutions were spatially autoregressive, the exclusion of $\boldsymbol{W\iota}$ dampened the effect; its inclusion strengthened the autoregressive pattern in the data.

[4] The exclusion of islands from the data has little justification and might bias inferences.

Table 4: Determinants of Polity in 2014, Including Average Direct and Indirect Effect Estimates of Logged GDP (Standard Errors in Parentheses)

Wy	0.71	0.92
	(0.12)	(0.07)
$W\iota$		-5.64
		(1.31)
Logged GDP	0.79	0.79
	(0.30)	(0.27)
Constant	-4.50	-1.94
	(2.54)	(2.43)
Average Direct Effect	0.82	0.86
	(0.30)	(0.29)
Average Indirect Effect	0.67	2.10
	(0.33)	(1.70)

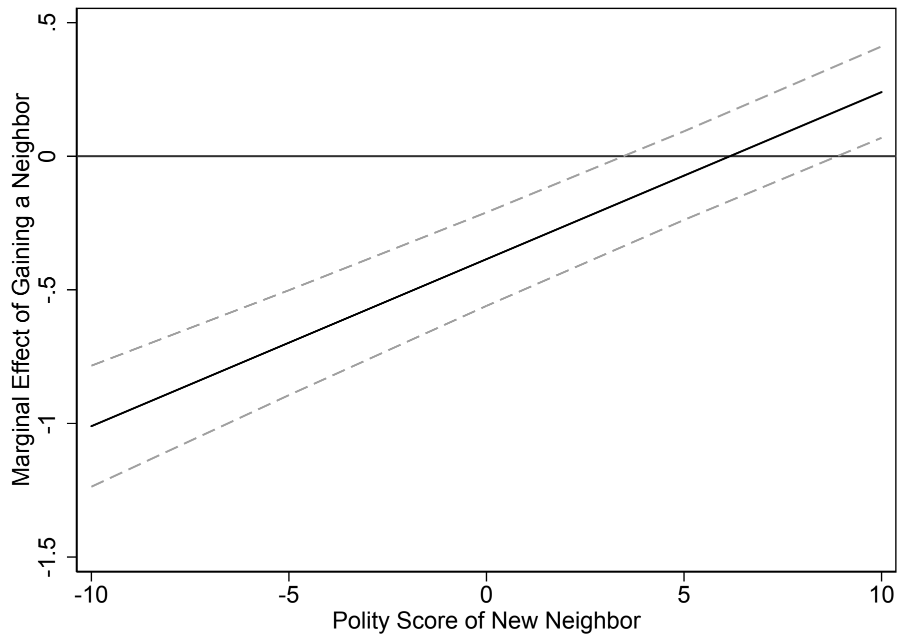
The inclusion of $\boldsymbol{W\iota}$ also makes a large difference on the average marginal effects of Logged GDP. The average direct effect between the two models changes little difference in the estimate. The average indirect effect, however, triples in magnitude. Of course, the standard error also increases to the point where the effect becomes statistically indistinguishable from zero (and likely indistinguishable from the original estimate). Yet the difference is striking enough to serve as a warning to scholars who exclude $\boldsymbol{W\iota}$ from a spatial regression model.

Beyond the correct modelling of spatial dependencies, the inclusion of $\boldsymbol{W\iota}$ allows for new possibilities of interpretation. Using the coefficients of the SAR model from Table 5, I construct a marginal effects plot of the impact of w_{ij} on y_i.[5] I plot the effect of a country gaining a new neighbor on that country's POLITY score across the empirical range of the new neighbor's POLITY score. The results are in Figure 1.

As one can see, a country gaining a new neighbor is largely an undemocratic affair. If a new neighbor is completely autocratic with a score of -10, like North Korea, this predicts a one-point decline in the original country's POLITY score. As the POLITY score increase, the marginal effect of a new neighbor is still negative even to more neutral values. Only

[5]Due to the innovations in this paper, this task requires original code.

Figure 1: Conditional Marginal Effect of Gaining a New Neighbor, by the POLITY score of the New Neighbor - Adjusted for Normalization (95% Confidence Intervals)



when a country achieves a moderately democratic score of 5, like Ecuador, does the effect of a new neighbor become negligible. At maximum levels of democracy, like the U.S., a new neighbor actually leads to a small increase in a country's POLITY score.

11 Empirical Example 2: Determinants of Insurance Coverage

For my second replication, I turn to a study by Cook, An, and Favero (Forthcoming). They introduce spatial regression models to the public administration subfield, arguing that a number of theories proposed in that literature would be best tested by using spatial lags. They demonstrate the techniques usefulness in two different applications, while also commenting on best practices when using this class of models. The study is a robust, streamlined presentation of a variety of spatial regression models, making it a perfect candidate to replicate and demonstrate the multiplicative interactions framework.

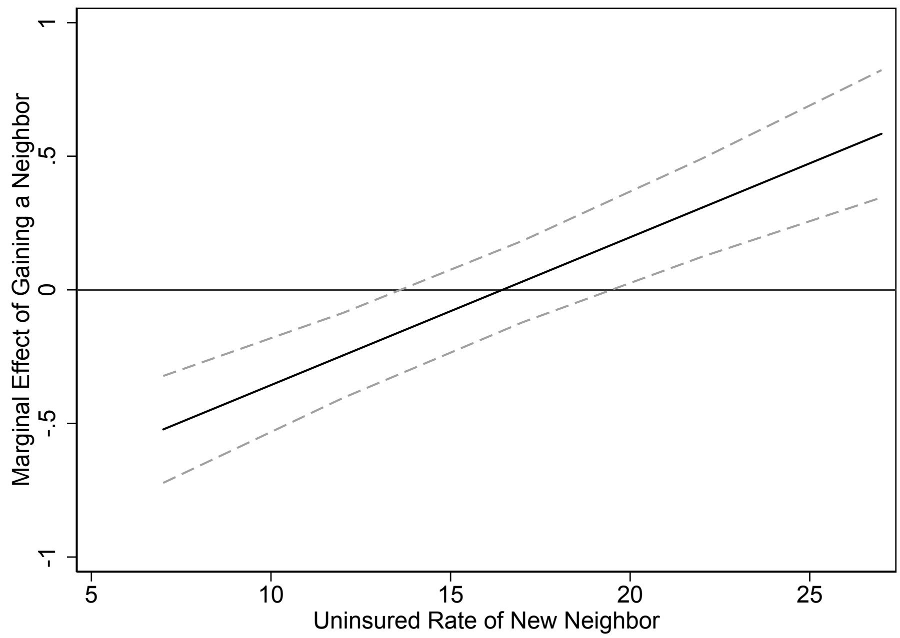
Specifically, I replicate a panel analysis on healthcare administration in the American states. The dependent variable is state uninsured rates from 1990-2006. A variety of predictors are included, including institutional, economic, demographic, and health variables. They estimate their models with random effects and consider both static and dynamic specifications; I replicate the dynamic models due to their superior model fit. Importantly, the authors estimate a series of spatial regression models to determine if they improve upon standard OLS techniques. They specifically test a linear regression, a SEM model, a SAR model, and a spatial autocorrelation model (SAC), which combines both SEM and SAR models into a single model. They find that a SEM model best fits the data, a conclusion they reach after estimating the SAC model and only finding the spatial error term statistically significant.

[Table 5 Here]

I replicate their analysis exactly and also reestimate their models while including $\boldsymbol{W\iota}$. The results are in Table 5. We see that $\boldsymbol{W\iota}$ is negative but statistically insignificant in the OLS and SEM models in columns 1 through 4. Their inclusion does not seem to change much in the estimation of either model, as the models' coefficients remain largely the same. In the SAR models in columns 5 and 6, however, $\boldsymbol{W\iota}$ has an enormous impact. Its coefficient is negative and statistically significant. More importantly, however, the spatial lag of y, $\boldsymbol{Wy}$, changes values after including $\boldsymbol{W\iota}$: the already positive coefficient triples in magnitude, indicating that the failure to include $\boldsymbol{W\iota}$ biased the results of the previous model.

In columns 7 and 8, we see the most compelling results for the inclusion of $\boldsymbol{W\iota}$. In this general model, $\boldsymbol{W\iota}$ is again negative and statistically significant. What is more, the coefficients of the spatially lagged terms change drastically after its inclusion. Prior to the inclusion of $\boldsymbol{W\iota}$, $\boldsymbol{Wy}$ was close to zero and statistically insignificant; afterwards, $\boldsymbol{Wy}$ matches

Figure 2: Conditional Marginal Effect of Gaining a New Neighbor, by the Uninsured Rate of the New Neighbor - Adjusted for Normalization (95% Confidence Intervals)



its values in column 6 of being positive and statistically significant. In contrast, $\boldsymbol{W\varepsilon}$ was positive and statistically significant without $\boldsymbol{W\iota}$; after its inclusion, $\boldsymbol{W\varepsilon}$ becomes close to zero and statistically insignificant. By including $\boldsymbol{W\iota}$, our inferences fundamentally change from the data being an SEM model with positive autocorrelation to an SAR model with positive autoregression. This new understanding of uninsured rates as a spatially autoregressive process is only possible by including $\boldsymbol{W\iota}$ in the model.

Figure 2 contains the conditional marginal effects plot of the fully specified spatial autoregressive model in column 6. The plot contains the empirical range of state uninsured rates in this time period, from a minimum of 7 percent to a maximum of 28 percent. Starting at the lowest observed value, if a state's new neighbor has a 7 percent uninsured rate, this leads to a subsequent decrease in the uninsured rate of the original state of about half a percent. As the new neighbor's uninsured rate increases, however, this effect becomes more and more positive until there is a statistically significant effect in the other direction: if a state's new neighbor has the highest uninsured rate observed, 28 percent, this leads to an increase in the uninsured rate of the original state by half a percent.

12 Discussion

I propose a reinterpretation of the spatial weights matrix, $\boldsymbol{W}$, in spatial regression models. Rather than being a set of weights, I argue that scholars should take a cue from network analysis and view the elements of $\boldsymbol{W}$ as themselves variables; more specifically, that $\boldsymbol{W}$ contains a set of conditioning variables. I further argue that all spatial regression models contain a series of n-1 multiplicative interaction terms with the same coefficient and, as currently estimated, do not have their constitutive terms. I argue that one of these terms - the row-sum of $\boldsymbol{W}$, $\boldsymbol{W\iota}$ - should be included in the model or scholars unnecessarily risk endogeneity from both omitted variable bias and measurement error. I demonstrate this through both Monte Carlo Analysis and a pair of replication studies.

The study makes three contributions to the literature. First, it jointly discusses spatial regression models and network analysis in a way rarely done in the literature. Because spatial regression models and network analysis rely on the same data sources to do their analysis, scholars can use both in regression models. From a substantive viewpoint, the failure to utilize both approaches leaves plenty of interesting theoretical relationships undiscussed and untested. From a methodological view, failure to include both can lead to bias in spatial regression models as demonstrated in the paper.

Second, it highlights two sources of bias in spatial regression models in a single framework. In this study, I focus on how omitting $\boldsymbol{W\iota}$ will usually result in bias in spatial regression models. Additionally, Vande Kamp (Forthcoming) shows that omitting a spatial lag of a variable where $\boldsymbol{W}$ is entirely composed of ones - whether $\boldsymbol{Ky}$, $\boldsymbol{Kx}$, or $\boldsymbol{K\varepsilon}$ - will normally result in bias if $\boldsymbol{W}$ is misspecified. This study demonstrates that both biases are predictable when approaching spatial regression models from a multiplicative interactions framework, as each of these two types of covariates represent the two sets of constitutive terms omitted by scholars when estimating spatial regression models in the status quo. This is a notable improvement in generalizing our understanding of bias in spatial regression models.

Third, this study stengthens efforts to communicate meaningful effects from spatial regression models. While spatial regression models are well understood as a whole, interpreting spatial coefficients directly is difficult and has triggered continued research on the interpretation of these models (Whitten, Williams, and Wimpy 2019). The multiplicative interactions framework contributes to interpretation in two ways. First, it demonstrates that commonly calculated quantitites of interest, like indirect effects and total effects, can be biased when $\boldsymbol{W\iota}$ is excluded from the model. Second, it provides a ready interpretation of both the spatial lag and $\boldsymbol{W\iota}$ by using marginal effects plots, as made popular by Brambor, Clark, and Golder (2006). I demonstrate this approach in my replication.

This study is not without limitations. I assume that, apart from the omission of $\boldsymbol{W\iota}$, spatial regression models are otherwise well-specified. This is likely not the case in most empirical applications. It is possible that scholars only model one of many spatial dependencies in the data. Scholars might omit variables that are spatially distributed or correlated with the in-degree centrality of a unit. And in cases beyond a single cross-section, temporal dynamics might not be adequately modelled. In these circumstances, it is not entirely clear how including $\boldsymbol{W\iota}$ will affect our inferences; future research should consider these possibilities.

In addition, I largely restrict my analysis to $\boldsymbol{W\iota}$, which in network terms is the in-degree centrality of a unit. There are many other measures of a unit's position in a network that can be derived from $\boldsymbol{W}$. While I motivate my focus on $\boldsymbol{W\iota}$ due to its relationship with spatial regression models, it is entirely possible that other network measures will have interesting interactions with spatial regression models. Given this possibility, scholars should continue researching the relationship between spatial regression models and network analysis.

References

[1] Beck, Nathaniel, Kristian S. 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.

[2] Brambor, Thomas, William R. Clark, and Matt Golder. 2006. "Understanding interaction models: Improving empirical analyses." *Political Analysis* 14(1): 63-82.

[3] Braumoeller, Bear F. 2004. "Hypothesis testing and multiplicative interaction terms." *International Organization* 58(4): 807-820.

[4] Cook, Scott J., Seung-Ho An, and Nathan Favero. Forthcoming. "Beyond Policy Diffusion: Spatial Econometric Models of Public Administration." *Journal of Public Administration Research and Theory*

[5] Fowler, James H. 2006. "Connecting the Congress: A Study of Cosponsorship Networks." *Political Analysis* 14:456–487.

[6] Gleditsch, Kristian S. and Michael D. Ward. 2001. "Measuring Space: A Minimum-Distance Database and Applications to International Studies." *Journal of Peace Research* 38(6): 739–758.

[7] Hafner-Burton, Emilie M. and Alexander H. Montgomery. 2006. "International Organizations, Social Networks, and Conflict." *Journal of Conflict Resolution* 50(1):3-27.

[8] Halleck Vega, Solmaria, and J. Paul Elhorst. 2015. "The SLX model." *Journal of Regional Science* 55(3): 339-363.

[9] Hays, Jude C., Aya Kachi, and Robert J. Franzese Jr. "A spatial model incorporating

dynamic, endogenous network interdependence: A political science application." *Statistical Methodology* 7(3): 406-428.

[10] Maoz, Zeev, Kuperman, Ranan D., Terris, Leslie., and Ilan Talmud. 2006. "Structural equivalence and international conflict: A social networks analysis." *Journal of Conflict Resolution* 50(5): 664-689.

[11] Neumayer, Eric, and Thomas Plümper. 2012. "Conditional Spatial Policy Dependence: Theory and Model Specification" *Comparative Political Studies* 45(7): 819-849.

[12] Neumayer, Eric and Thomas Plümper. 2016. "W." *Political Science Research and Methods* 4(1):175-193.

[13] Starr, Harvey and Benjamin A. Most. 1976. "The Substance and Study of Borders in International Relations Research" *International Studies Quarterly* 20(4), 581-620.

[14] Tobler, W. R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46: pp. 234-240.

[15] Vande Kamp, Garrett N. Forthcoming. "Measurement Error and the Specification of the Weights Matrix in Spatial Regression Models." *Political Analysis*.

[16] Ward, Michael D., Katherine Stovel, and Audrey Sacks. 2011. "Network Analysis and Political Science" *Annual Review of Political Science* 14: 245-264.

[17] Ward, Michael D. and Kristian S. Gleditsch. 2018. *Spatial Regression Models* Second Edition. Thousand Oaks, CA: SAGE Publishing.

[18] Whitten, Guy D., Laron K. Williams, and Cameron Wimpy. 2019. "Interpretation: the final spatial frontier." *Political Science Research and Methods*

Table 5: Determinants of State-Level Uninsured Rates, 1992 to 2006 - Dynamic & RE (Standard Errors in Parentheses)

% Uninsured	OLS		SEM		SAR		SAC	
	1	2	3	4	5	6	7	8
$W\varepsilon$			0.325	0.323			0.305	-0.078
			(0.057)	(0.057)			(0.063)	(0.121)
Wy					0.086	0.303	0.022	0.345
					(0.034)	(0.044)	(0.029)	(0.077)
$W\iota$		-0.248		-0.225		-4.984		-5.643
		(0.231)		(0.467)		(0.799)		(1.280)
Medicaid eligibility	-0.004	-0.004	-0.006	-0.006	-0.005	-0.006	-0.006	-0.006
	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Public finance$_{t-2}$	-0.026	-0.025	0.000	0.001	0.031	0.031	0.002	0.037
	(0.022)	(0.022)	(0.035)	(0.035)	(0.037)	(0.032)	(0.035)	(0.033)
Public ownership	0.002	0.002	0.007	0.007	0.008	0.006	0.007	0.006
	(0.004)	(0.004)	(0.008)	(0.008)	(0.009)	(0.008)	(0.008)	(0.008)
State liberalism	-0.004	-0.005	-0.002	-0.003	-0.003	-0.005	-0.002	-0.005
	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)	(0.003)
Unemployment	0.182	0.170	0.191	0.186	0.222	0.195	0.198	0.195
	(0.061)	(0.062)	(0.076)	(0.076)	(0.070)	(0.067)	(0.076)	(0.065)
Poverty	0.179	0.184	0.219	0.221	0.208	0.195	0.214	0.192
	(0.033)	(0.033)	(0.041)	(0.041)	(0.042)	(0.040)	(0.041)	(0.040)
Black population	0.011	0.010	0.028	0.028	0.030	0.021	0.028	0.020
	(0.009)	(0.009)	(0.018)	(0.018)	(0.018)	(0.016)	(0.018)	(0.016)
Latino population	0.072	0.073	0.114	0.114	0.119	0.102	0.113	0.100
	(0.010)	(0.011)	(0.020)	(0.020)	(0.021)	(0.018)	(0.020)	(0.018)
Aged population	-0.084	-0.083	-0.050	-0.048	-0.051	-0.003	-0.050	0.007
	(0.043)	(0.043)	(0.078)	(0.077)	(0.086)	(0.075)	(0.078)	(0.078)
Obesity rate	-0.004	-0.006	-0.016	-0.016	-0.019	-0.013	-0.015	-0.013
	(0.019)	(0.019)	(0.025)	(0.025)	(0.022)	(0.020)	(0.025)	(0.020)
Perceived poor health	0.018	0.028	0.014	0.016	-0.024	-0.003	0.009	-0.007
	(0.059)	(0.059)	(0.061)	(0.061)	(0.061)	(0.059)	(0.061)	(0.059)
% Uninsured$_{t-1}$	0.678	0.673	0.484	0.485	0.475	0.452	0.483	0.440
	(0.026)	(0.026)	(0.041)	(0.041)	(0.041)	(0.037)	(0.041)	(0.041)
Constant	3.233	3.507	4.607	4.768	3.287	4.775	4.340	4.879
	(0.814)	(0.853)	(1.335)	(1.373)	(1.444)	(1.306)	(1.381)	(1.327)