PSJ Policy Studies Journal

Policy Studies Journal, Vol. 0, No. 0, 2019

Taking Time (and Space) Seriously: How Scholars Falsely Infer Policy Diffusion from Model Misspecification

Cody A. Drolc, Christopher Gandrud, and Laron K. Williams

Scholars have long been interested in how policies and ideas spread from one observation to another. Yet, the spatial and temporal dynamics of policy diffusion present unique challenges that empirical researchers often neglect. Scholars often use temporally lagged spatial lags (TLSL)—such as the number (or percentage) of prior adopters in a neighborhood—to test various mechanisms of delayed policy diffusion but are largely unaware of two under appreciated issues. First, the effects are not limited to one time period but persist over time by changing the future value of neighboring observations. Second, minor, yet common, choices in model specification—such as omitting spatially correlated and/or autoregressive covariates—can increase the risk of falsely inferring that the outcome is a result of spatial diffusion. Indeed, we offer two applications where small changes to the model specification of an otherwise well-specified model result in drastically different inferences about policy diffusion. We argue that scholars should avoid haphazardly including TLSLs without considerable theoretical justification, and we conclude on an optimistic note by offering straightforward solutions and new software to address these issues.

KEY WORDS: policy diffusion, spatial econometrics, model specification

长期以来，学者都对政策和观点如何从一个观察数据扩散到另一个表示兴趣。然而，政策扩散的时空动态所呈现的独特挑战却是实证研究者经常所忽略的。学者常使用随时间滞后而变化的空间滞后模型（*TLSL*）—例如某个邻区中先前采纳者的数量（或百分比）—来测试滞后政策扩散的不同机制，但学者在很大程度上没有意识到两个被轻视的问题。第一，效果不限于一个时间阶段，而是通过改变相邻观测数据的终值，随时间推移继续发挥作用。第二，模型设定中小众但常见的选择—例如省略与空间相关的和/或自回归的协变量—能增加错误推断的风险，这个错误推断即：结果是由空间扩散导致的。的确，我们提出两种（模型）应用，其中另一个设定完整的模型会因为模型设定发生的细小变化而导致极为不同的政策扩散推断。我们主张，学者应避免在没有充足理论依据的情况下无计划地应用*TLSLs*。我们的结论通过提出用于应对这些问题的简易解决措施和新的软件，传达了一个积极的信息。

关键词: 政策扩散, 空间计量经济学, 模型设定

Los académicos llevan mucho tiempo interesados en cómo las políticas e ideas se propagan de una observación a otra. Sin embargo, la dinámica espacial y temporal de la difusión de políticas presenta desafíos únicos que los investigadores empíricos a menudo descuidan. Los académicos a menudo usan retrasos espaciales temporalmente retrasados (TLSL), como

doi: 10.1111/psj.12374

el número (o porcentaje) de adoptantes anteriores en un vecindario, para probar varios mecanismos de difusión diferida de políticas, pero desconocen en gran medida dos problemas poco apreciados. Primero, los efectos no se limitan a un período de tiempo, sino que persisten en el tiempo al cambiar el valor futuro de las observaciones vecinas. En segundo lugar, las elecciones menores, aunque comunes, en la especificación del modelo, como la omisión de covariables espacialmente correlacionadas y/o autorregresivas, pueden aumentar el riesgo de inferir falsamente que el resultado es el resultado de la difusión espacial. De hecho, ofrecemos dos aplicaciones en las que pequeños cambios en la especificación del modelo de un modelo que de otro modo estaría bien especificado dan como resultado inferencias drásticamente diferentes sobre la difusión de políticas. Argumentamos que los académicos deben evitar incluir TLSL de manera fortuita sin una justificación teórica considerable, y concluimos con una nota optimista ofreciendo soluciones directas y un nuevo software para abordar estos problemas.

PALABRAS CLAVE: difusión de políticas, econometría espacial, especificación del modelo

Introduction

A common theme explored by scholars of public policy, international relations, and comparative politics is how policies spread from one government (e.g., country, state, county, city) to another (Gilardi, 2016; Graham, Shipan, & Volden, 2013; Shipan & Volden, 2012). Policy diffusion generally occurs through four primary mechanisms: learning, emulation, competition, and coercion (Graham et al., 2013). These mechanisms have been extensively tested in a variety of contexts including adoption of anti-smoking policies (Shipan & Volden, 2006, 2008), cigarette tax changes (Davis & Nicholson-Crotty, 2016), minimum wage increase (Whitaker, Herian, Larimer, & Lang, 2012), driving under the influence ignition (DUI) interlock laws (Sylvester & Haider-Markel, 2016), and border agreements (Clay & Owsiak, 2016). In general, these studies build on the classic view of diffusion that presumes similarity can be inferred from geography (Walker, 1969). Under this extensively tested perspective, geographically proximate governments learn, emulate, and compete with one another, driving policy change. Although geographic proximity dominates theories and empirical examinations of diffusion, policies and their features can spread across any connected network. Governments, for example, can learn from and emulate ideologically similar governments although they may not be geographically close (Desmarais, Harden, & Boehmke, 2015; Grossback, Nicholson-Crotty, & Peterson, 2004; Mallinson, 2019). In this way, the mechanisms of diffusion are implicitly spatial, but space encompasses more than geographic proximity (Beck, Gleditsch, & Beardsley, 2006).

In diffusion research, scholars often use the proportion of neighbors (geographic, ideological, economic) that previously adopted a policy to test the delayed effects of learning, emulation, or competition (Maggetti & Gilardi, 2016). This variable is what we call a temporally lagged spatial lag (TLSL), or the weighted sum (or average) of neighboring governments' outcomes in the previous period, and is intended to capture those processes where spatial diffusion occurs with a temporal lag. Although the mechanisms of diffusion are often tested and conceived of in spatial terms, scholars have yet to fully leverage and carefully consider the effects of space. Given a spatial process, certain assumptions about the data must match theoretical expectations in order to draw proper empirical inferences about policy diffusion. We build on recent methodological work that recognizes the spatial dimension of policy diffusion (Mitchell, 2018) by broadening the notion of space and carefully considering the effects of time. Just as deBoef and Keele (2008) offer guidelines for deriving inferences from time series data in their piece "Taking Time Seriously," we offer guidelines for specifying and interpreting models in the context of time and space.

We highlight two challenges that scholars face when using TLSLs. First, models with TLSLs present unique interpretive challenges because the spatial effects operate with a temporal lag. Scholars use TLSLs because their theories suggest that space matters for these outcomes, so it stands to reason that the effects of all the variables occur immediately *and* spatially with a temporal lag. As such, the coefficients themselves only represent the short-term direct effect on each observation. The true total effect of a variable is the sum of that short-term direct effect, and a spatial long-term effect (SLTE) that operates through the TLSL. Under typical conditions of policy diffusion (i.e., positive spatial dependence), this means that scholars who only interpret the coefficients are *understating* the true effects of explanatory variables by neglecting SLTEs. To aid the interpretation of diffusion models, we demonstrate how to calculate SLTE, as well as summary statistics that detail the average direct and indirect effects.

Second, improper use of TLSLs may lead to false inferences about the explanatory variables and the diffusion process itself. We offer evidence from Monte Carlo simulations showing how substantive conclusions change when models include irrelevant TLSLs that are not part of the data-generating process. When there are omitted autocorrelated and/or spatially clustered covariates—conditions endemic to the study of political phenomenon—the false identification rate of irrelevant TLSLs is much higher than expected.[1] Put simply, common errors in model specification are suggestive of patterns of policy diffusion that are not present (see also Boehmke, 2009a). In our experimental conditions, the false discovery rate ranges from 28 percent when an autoregressive covariate is missing to nearly 50 percent when a spatially clustered covariate is omitted. Excluding either an autoregressive or spatially clustered covariate leaves researchers at a high risk of discovering a spurious TLSL and thus incorrectly inferring the existence of a diffusion process. In doing so, we contribute to the rich literature that examines the consequences of incorrectly specifying models where temporal dynamics matter (Achen, 2000; Keele & Kelly, 2006).

We find evidence that these problems are widespread in the study of policy diffusion. Based on a survey of over 100 quantitative policy diffusion articles published from 2000 to early 2017, we identified several patterns regarding the use of TLSLs. The most concerning patterns were the lack of theoretical justification for including TLSLs, failure to account for temporal dependence, sporadic use of *a priori* diagnostics of spatial clustering to justify their inclusion, and a failure to explore spatial effects.

We offer a series of guidelines to avoid these problems: first, scholars should lean heavily on theory to eliminate alternative pathways that might otherwise cause a spurious policy diffusion effect, and second, scholars should conduct appropriate specification tests. In terms of the TLSL, we advocate testing for evidence of spatial clustering prior to specifying a model using standard tests such as Moran's *I*. If diagnostics consistently indicate the presence of spatial clustering, then researchers are justified in their inclusion of the TLSL. To this end, we introduce a package in R (R Core Team, 2019), called `spatialWeights`, that automates the creation of weights matrices and executes the Moran's *I* test for time series cross-sectional (TSCS) data. Finally, we use three real-world data examples to demonstrate that small and seemingly innocuous changes to an otherwise well-specified model can provide evidence of a diffusion process that does not exist.

Two Pictures of Spatial Diffusion

Policy diffusion scholars recognize that policymakers take cues from other governments when considering policy adoptions, innovations, or retrenchments. Walker's (1969) argument that policymakers look to "similar" governments permeates theories and empirical examinations of how policies move from one government to another. Geographic contiguity often acts as a proxy for similarity between governments. In this way, the physical space occupied by neighboring governments influences policy contagion. Although the exact diffusion mechanism can be difficult to disentangle, geographic proximity is often used to provide evidence for learning, emulation, or competition (Boehmke & Witmer, 2004; Maggetti & Gilardi, 2016; Volden, 2006; Weyland, 2005).

Attempts to explore the mechanisms by which policy diffuses across governments typically fall into one of two empirical approaches, which lead to two pictures of spatial diffusion. The first empirical approach uses event history analysis (EHA) to model the probability of some event occurring (typically either the adoption of a policy or policy convergence), given that it has not occurred up to that point (Berry & Berry, 1990). The *dyadic* component of EHA shifts the level of analysis from observations to *pairs* of observations, or pairs of senders (sources) and receivers (targets) (see Boehmke, 2009b; Gilardi, 2010; Gilardi & Fuglister, 2008; Volden, 2006). Dyadic EHA is appealing because it offers the possibility of exploring how characteristics of both governments, as well as how they relate to each other (in terms of similarities or differences), influence policy diffusion (Volden, 2006). While the discrete dependent variable lends itself well to measuring joint adoption of policies, dyadic EHA has been generalized to include any type of "increased similarity" between governments (Gilardi & Fuglister, 2008, p. 415).[2] This approach allows scholars to identify similar observations (or neighbors) in a variety of ways to assess how governments learn from and emulate each other. Overall, this approach improves upon prior studies of policy adoption because it offers "a vastly richer specification of the diffusion process between pairs of states and, consequently, a more precise comparison of the role of external forces with internal political and demographic characteristics" (Boehmke, 2009b, p. 1125).

The second empirical approach uses spatial econometric models to explicitly model how political, social, and economic processes diffuse across a network (Beck et al., 2006; Darmofal, 2015; Franzese & Hays, 2007; Ward & Gleditsch, 2008). To date, two prominent models dominate these spatial diffusion projects: the spatial autoregressive (SAR) and the spatial-X (SLX) models. As we demonstrate below, the difference in models is based on whether the outcome in one observation, y_i, responds to the outcomes in neighboring observations, y_j (SAR), or to covariates in neighboring observations, x_j (SLX; Cook, Hays, & Franzese, 2015). Diffusion in the SAR model occurs simultaneously and endogenously throughout all neighbors in the system; diffusion in the SLX model (via TLSLs) occurs with a user-specified temporal lag, one order of neighbors at a time. These notable differences across spatial models are often confused in studies of policy diffusion, so it is worth exploring them in more detail.

First, consider the SAR model in matrix form shown in equation (1):

$$\boldsymbol{y} = \mathbf{X}\beta + \rho\mathbf{W}\boldsymbol{y} + \epsilon \tag{1}$$

where $\mathbf{X}$ is an $N \times k$ matrix of $k-1$ independent variables and a constant, $\boldsymbol{y}$ is the outcome vector, ρ is the strength of spatial autocorrelation, $\mathbf{W}$ is an $N \times N$ matrix detailing how all N observations are connected to each other, and ϵ is a normally distributed error term with mean zero and constant variance (Darmofal, 2015, p. 97).[3] There are two features of the SAR model that shape inferences about processes of spatial diffusion. The first feature is that the outcomes in all observations are endogenous (because $\boldsymbol{y}$ is on both sides of equation [1]). As a result, we need to use the reduced form of equation (1) where $\mathbf{I}$ is an $N \times N$ identity matrix:[4]

$$\boldsymbol{y} = \left(\mathbf{I}_n - \rho\mathbf{W}\right)^{-1}\mathbf{X}\beta + \left(\mathbf{I}_n - \rho\mathbf{W}\right)^{-1}\epsilon \tag{2}$$

When we calculate the effects of x on y, we calculate the partial derivatives matrix (LeSage & Pace, 2009; Whitten, Williams, & Wimpy, forthcoming):

$$\left[\frac{\partial E(y)}{\partial x_1} \cdots \frac{\partial E(y)}{\partial x_N}\right] = (\mathbf{I}_n - \rho\mathbf{W})^{-1}\beta \tag{3}$$

and its infinite series expansion:

$$\left(\mathbf{I}_n - \rho\mathbf{W}\right)^{-1}\beta = \left(\mathbf{I}_n + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \rho^3\mathbf{W}^3 + \dots\right)\beta \tag{4}$$

From equation (4), the second feature of the SAR becomes clear. A change in x_i for any non-isolate observation will influence its own outcome, y_i, by β, its first-order neighbors, y_j, by $\beta\rho\mathbf{W}$, its second-order neighbors, y_k, by $\beta\rho^2\mathbf{W}^2$, and so on, and all non-isolate observations experience the impact. Thus, spatial diffusion processes in the SAR model occur simultaneously and endogenously via local (i.e., first-order) and global (i.e., higher-order and feedback) effects. In a scenario with changing tax rates, this picture of diffusion assumes that neighboring governments observe and adopt changes *concurrently*. Hence, it may not be appropriate for theories requiring a *temporal delay* between when the potential adopters observe the actions of neighbors and when they act.

Now, consider the SLX model in matrix form shown in equation (5):

$$\boldsymbol{y} = \mathbf{X}\beta + \theta\mathbf{WZ} + \epsilon \tag{5}$$

where $\boldsymbol{y}$ and $\mathbf{X}$ are defined above, $\mathbf{Z}$ is an $N \times k$ matrix of variables whose values influence $\boldsymbol{y}$ in neighboring observations (determined by $\mathbf{W}$). β represents the direct effect of changes in x_i on y_i while θ represents the indirect effect of x_j on y_i. The $\mathbf{X}$ and $\mathbf{Z}$ matrices do not have to be identical; a variable can have a direct effect on $\boldsymbol{y}$ (and be in $\mathbf{X}$) but not an indirect effect ($\mathbf{Z}$), the reverse, or both.[5] The $\mathbf{W}$ is intended to capture the influence of neighboring governments but is not limited to geographically contiguous or proximate observations (Shipan & Volden, 2012). For example, interdependence or connectivity between governments can be defined in terms of ideology (Grossback et al., 2004), trade (Shipan & Volden, 2008), or commuters (Gilardi & Wasserfallen, 2016).

A common empirical strategy is to use the lagged proportion of previous adopters overall or in a government's neighborhood (Berry & Berry, 1990; Fay & Wenger, 2016; Sylvester & Haider-Markel, 2016; Whitaker et al., 2012) to test spatial diffusion effects that occur with a temporal lag. TLSLs are useful for understanding diffusion processes because they can take the form of discrete or continuous outcomes like the adoption of a particular tax or convergence in the generosity of welfare payments (e.g., Berry, Fording, & Hanson, 2003; Davis & Nicholson-Crotty, 2016; Gilardi & Wasserfallen, 2016; Plümper, Troeger, & Winner, 2009). As an example, take three neighboring governments and the diffusion of income tax rates. When we consider discrete changes to the policy, equation (6) decomposes the commonly used lagged proportion of previous adopters into a simple un-row-standardized contiguity weights matrix, $\mathbf{W}$, where government 2 (second row and column) and 3 (third row and column) are neighbors with 1 (first row and column) and the adopters at $t-1$ are given by a column vector. For a contiguity weights matrix, 1 on the off-diagonal represents a connection. In this example, governments 2 and 3 changed their income tax policies in the previous period. The resulting TLSL (column vector on the far right) shows that the first government has two neighbors that adopted the tax while the other two governments have none.

$$\boldsymbol{x} = \mathbf{W}\boldsymbol{y}_{t-1} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix} \tag{6}$$

The same logic applies to *convergence* in the income tax rate. In equation (7), we change the values of the column vector to represent each government's income tax rate at $t-1$ where state 1 has no income tax. Once we multiply through, the TLSL is the weighted sum of the neighbors' tax rates for each government. For example, the weighted sum for the first government is 5 (based on its two neighbors' rates of 3 and 2), while the weighted sum for the other two government is 0 (because both governments are only neighbors with the first government, which did not have an income tax).

$$x = \mathbf{W}y_{t-1} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 5 \\ 0 \\ 0 \end{pmatrix} \tag{7}$$

An alternative specification choice is to row-standardize the $\mathbf{W}$.[6] Dividing each element of $\mathbf{W}$ by its row total (which forces each row to sum to 1) changes the value of $\mathbf{W}y_{t-1}$ in equations (6) and (7) to be the weighted *average* of previous adopters and tax rates, respectively. Moreover, these simple examples rely on a contiguity weights matrix, but researchers can easily define $\mathbf{W}$ using any theoretically motivated interconnectivity criterion. In fact, scholars can test theories of diffusion through multiple avenues by including additional TLSLs. For example, geographically close governments may compete with one another to attract workers, and ideologically similar governments may emulate tax structures. In this case, the first $\mathbf{W}$ could be defined by geographical proximity and the second $\mathbf{W}$ could be defined using a measure of governments' ideological similarity. Both matrices would weight the tax rates at $t-1$ in each government's neighborhood but test different diffusion pathways. Using theoretically motivated weights matrices allows scholars to define neighbors in a variety of ways and avoid inferential problems associated with incorrectly specifying the mechanisms of policy diffusion.[7]

Models including TLSLs—such as the weighted sum or average of previous adopters—therefore fall into the broader set of SLX models (LeSage & Pace, 2009, p. 192; Williams, 2015; Wimpy, Williams, & Whitten, forthcoming), where the $\mathbf{W}y_{t-1}$ vector contains the spatial lag of neighboring observations from the previous period, or

$$y = \mathbf{X}\beta + \theta\mathbf{W}y_{t-1} + \epsilon \tag{8}$$

This model specification offers a picture of diffusion that occurs with a temporal lag, influencing first-order neighbors at time $t + 1$, second-order neighbors at time $t + 2$, and so on. In the context of income tax rates, a change in government i influences the rates in their neighbors j at time $t + 1$, but the effects propagate to higher-order (the neighbors of government j) neighbors as time elapses. Both spatial models produce global effects; the SLX model captures this delay through the TLSL whereas the SAR model forces the diffusion process to occur all at time t. Theories of diffusion may suggest that the former approach is how the policy process operates with governments looking to the actions of neighbors and then responding with a delay (e.g., Berry & Berry, 1990; Böhmelt & Freyburg, 2015; Brooks, Cunha, & Mosley, 2015; Rogers, 2004).

For all the similarities of dyadic EHA and spatial models, the two empirical approaches have two surface-level differences and a larger, more substantive difference. Given that both approaches attempt to model how policies diffuse across governments, theoretical considerations are particularly useful in illuminating the appropriate modeling choice. At first, the biggest difference between approaches would appear to be the level of analysis; while spatial econometric models typically analyze governments, dyadic EHA analyzes *pairs* of governments. This difference, however, is somewhat superficial because there is nothing preventing scholars from using spatial models on directed dyadic datasets. In fact, Neumayer and Plümper (2010) show that directed dyadic data allow scholars to modify weights matrices to derive inferences about contagion across sources and targets. The second surface-level difference focuses on the outcome of interest; while dyadic EHA analyzes a binary event (such as mutual adoption or increased similarity), spatial econometric models typically analyze continuous events (such as tax rates). This difference is also a misconception, because recent advances in spatial econometric models have expanded into discrete, categorical, and count outcomes (Darmofal, 2015). The SLX in particular is quite flexible because its sole requirement is that the model must include an exogenous spatial lag (such as the proportion of previous adopters). Prominent diffusion examples abound of SLX variants of continuous models (e.g., Case, Rosen, & Hines, 1993), event count models (e.g., Boehmke & Witmer, 2004), and EHA models (e.g., Berry & Berry, 1990; Grossback et al., 2004).

The substantive difference between these two empirical approaches arises in the picture of spatial policy diffusion that they depict. The dyadic EHA model examines pairs in isolation; if a variable (for example, previous policy success in government *A*) has a positive coefficient, then we can infer that policy success in government *A* influences the probability of the event to the same extent in all pairs. Since the connections between *all* governments (**W** in the spatial models) are ignored in favor of dyadic connections, a significant predictor will influence all pairs of governments similarly, regardless of the number of other neighbors or the strength of those connections. It might be the case that having more neighbors increases the degree to which an observation is influenced by other observations; it is also possible that the effect of one observation on another declines as the number of neighbors increases. This is a question that is best answered with theory,[8] but the dyadic EHA imposes the former on the diffusion process. Furthermore, in a dyadic EHA model, the effects of some characteristic in government *A* only influence the similarity of governments *A* and *B*; in a spatial econometric model, the effects of a characteristic in government *A* go beyond the local or first-order effects of influencing government *B* to include higher-order and feedback effects (though the SAR and SLX differ over the temporal sequence of these effects). These two empirical approaches clearly offer different pictures of spatial diffusion; so, it is critical to have theory guide these choices.

With these differences in mind, we focus our efforts on exploring issues related to model specification and interpretation in spatial econometric models with TLSLs and continuous outcomes.

Quantities of Interest in the Context of TLSLs

Even though scholars have used TLSLs (and SLX models more broadly) to test theories of delayed diffusion processes, scholars have yet to fully appreciate that this feature changes the inferences they draw about diffusion processes. To see how this complicates interpretation, assume that we have a simple model employing a TLSL (equation [8]). The short-term direct effect of X_t on y_t for all i observations is $\Delta y_t \mid X_t = \beta$. Each covariate also has a *spatial long-term effect* (SLTE) that arises through changing the values of y_t, which then influences y_{t+1}, and so on, through $\theta\mathbf{W}$. The spatial long-term effect relies on the effect size (β), the strength of the temporally lagged diffusion parameter (θ), and the spatial distribution of observations ($\mathbf{W}$). Exploring the SLTE allows us to better understand this process so that the empirical results shed light on the theoretical expectations.

Consider a model of income tax rates with the weighted sum of neighbors' tax rates as a control variable.[9] We calculate the spatial long-term effects of a covariate, say citizen ideology, on income tax rates based on the following illustration: First assume the coefficient for citizen ideology is $\beta = 1$ and the coefficient for the weighted average tax rate is $\theta = 0.25$, and that the governments are connected through an un-row-standardized contiguity weights matrix (where governments 2 and 3 are neighbors with 1):

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}. \tag{9}$$

In general, we calculate the value of y at any time, $t + s$, based on the change in the previous period, $t + s-1$:

$$\Delta y_{t+s} | X_t = \theta\mathbf{W}\left(\Delta y_{t+s-1} | X_t\right) \tag{10}$$

For example, the change in y at time $t + 1$ given an increase in X_t, or $\Delta y_{t+1} \mid X_t$.

$$\Delta y_{t+1} | X_t = \theta\mathbf{W}\Delta y_t | X_t \tag{11}$$

If we substitute $X\beta$ for $\Delta y_t \mid X_t$, we get the following change for time $t + 1$:

$$\Delta y_{t+1} | X_t = \theta\mathbf{W}X\beta \tag{12}$$

A 1-unit increase in X results in an $N \times N$ matrix (consistent with the partial derivatives interpretation approach; see LeSage & Pace, 2009; Whitten et al., forthcoming) with the following values:

$$\Delta y_{t+1}|X_t=\theta \mathbf{W}X\beta=\left[\frac{\partial \mathrm{E}\left(y_{t+1}\right)}{\partial X_{t,1}}\cdots\frac{\partial \mathrm{E}\left(y_{t+1}\right)}{\partial X_{t,N}}\right]=\begin{pmatrix} 0 & 0.25 & 0.25 \\ 0.25 & 0 & 0 \\ 0.25 & 0 & 0 \end{pmatrix}. \tag{13}$$

The partial derivatives matrix provides the complete effects—including direct and indirect effects—of X on y_{t+1} for every single observation in the data. The values along the diagonal provide the direct effect (or feedback effect) of a change in X_t on y_{t+1} for each government. Since the diagonal of **W** contains only 0s, by construction, there are no feedback effects at time t + 1. The values along the off-diagonal provide the indirect effects of changes in one government on other governments. For example, the values in the first column (going from top row to bottom) represent the effects of increasing $X_{1,t}$ by 1-unit on itself (0), government 2 (0.25), and government 3 (0.25). It is easy to see that the effect at time t + 1 on any neighbor is $\theta w_{ij}X\beta$.

Because of the inclusion of a TLSL in the model specification, the change in $X_{i,t}$ has a spatial long-term effect that declines with t. At each additional time period, we calculate the effect by incorporating the change in y from the previous period. For example, the effects at time t + 2 are calculated by substituting $\theta\mathbf{W}X\beta$ for $\Delta y_{t+s-1}\,|\,X_t$ in equation (10):

$$\begin{aligned}\Delta y_{t+2}|X_t=\theta\mathbf{W}(\theta\mathbf{W}X\beta) &=\begin{pmatrix} 0 & 0.25 & 0.25 \\ 0.25 & 0 & 0 \\ 0.25 & 0 & 0 \end{pmatrix}\times\begin{pmatrix} 0 & 0.25 & 0.25 \\ 0.25 & 0 & 0 \\ 0.25 & 0 & 0 \end{pmatrix} \\ &=\begin{pmatrix} 0.125 & 0 & 0 \\ 0 & 0.0625 & 0.0625 \\ 0 & 0.0625 & 0.0625 \end{pmatrix}.\end{aligned} \tag{14}$$

The direct effects along the diagonal of equation (14) show that each government gets additional feedback due to changing its neighbors' values of y_{t+1}; government 1 has the largest effect because it has two neighbors while governments 2 and 3 only have one. At each value of t, we can calculate the total effects of X_t on y by summing the row values in each partial derivatives matrix. For example, the total effect of X_t on y_{t+2} is 0.125 for each government.

The formula for the spatial long-term effects in the third period, $\Delta y_{t+3}\,|\,X_t = \theta\mathbf{W}(\theta\mathbf{W}(\theta\mathbf{W}X\beta))$, demonstrates that the effects decline with each additional time period. The spatial long-term effect for periods t + 1 to t + S therefore simplifies to the following:

$$\mathrm{SLTE}=\sum\nolimits_{s}^{S}\theta W\left(\Delta y_{t+s-1}|X_t\right) \tag{15}$$

Thus, while the short-term direct effect of a change in X_t for observation i is β, the long-term spatial effect on all observations can be found in the partial derivatives matrix in equation (15). Note that the partial derivatives matrix depicts the direct and indirect effects for all N observations in the data. Since this matrix becomes unwieldy at high values of N, we offer three variations of summary measures suggested in LeSage and Pace (2009). The average total spatial long-term effect (AT-SLTE) is $\frac{\sum_j^I \sum_i^I \text{SLTE}}{N}$ and summarizes the average total effect of a 1-unit change in X_t on y. The average direct spatial long-term effect (AD-SLTE) is $\frac{\text{trace(SLTE)}}{N}$, and the average indirect spatial long-term effect (AI-SLTE) is AT-SLTE minus AD-SLTE. The latter two summary measures give an average of how much a change in each government affects itself (through feedback) and other governments, respectively. Since all of these quantities of interest are based on estimates, we would encourage scholars to provide measures of uncertainty (such as confidence intervals derived from simulation methods) to aid in their hypothesis tests (Carsey & Harden, 2013; Whitten et al., forthcoming).

Now, consider the typical approach to interpreting effects in models with TLSLs, as demonstrated by Brooks and Kurtz (2012). In this study, the authors examine the diffusion of capital account policy liberalization through a variety of diffusion pathways including geographic neighborhoods. While they theorize about the spatial and temporal nature of diffusion processes, their interpretations are limited to the reported coefficients and not the long-term or spatial effects. Yet, we know from decades of policy diffusion research that these effects last beyond time t and instead linger over time (Gilardi, 2016; Maggetti & Gilardi, 2016; Rogers, 2004; Shipan & Volden, 2012). By calculating SLTE with the partial derivatives approach, scholars have access to a wider variety of information about the effects of covariates on spatial policy diffusion over time. If one's substantive focus is on a particular subset of governments (as in a neighborhood or region), then this interpretation approach generates meaningful effects that reflect the distinct patterns of connections. If one's substantive focus is on the average effects overall, then the summary statistics described above represent an accurate picture of the effects of covariates. Whatever one's objective in interpretation, the SLTE reflects how policy diffuses to other governments over time.

Common Practices in TLSL Model Specification

In order to get a sense of common modeling practices, we conducted a survey of all *Web of Science* political science citations from 2000 to early 2017 mentioning "policy diffusion." We identified 226 articles addressing policy diffusion, of which 105 were quantitative and modeled a diffusion process (i.e., learning, emulation, competition, or coercion). These citations include observational studies of the American states, diffusion in other federal systems, and international policy diffusion. Overall, TLSLs are quite common, as nearly 65 percent of studies use some form of a TLSL to model a diffusion process or to control for neighborhood effects. We summarize the results of this survey for articles using a TLSL (N = 68) in Table 1.

We identified five patterns concerning the use of TLSLs that, when viewed together, raise serious concerns about their effectiveness in helping scholars make

Table 1. Summary of Policy Diffusion Survey

	Yes (%)	No (%)
Among articles studying policy diffusion (N = 105)		
Use TLSL	64.8	35.2
Among articles with a TLSL (N = 68)		
Theoretical or methodological justification	36.8	63.2
Address temporal dependence	70.6	29.4
Spatial autocorrelation diagnostics	7.4	92.6
Spatial long-term effects	5.9	94.1
Incorrect notation	27.9	72.1

accurate inferences. First, theoretical or methodological justifications for including the TLSL are sparse. "Common practice" or "traditional fashion" in diffusion studies are popular justifications for including the TLSL (e.g., Fay & Wenger, 2016; Karch, Nicholson-Crotty, Woods, & Bowman, 2016; Makse & Volden, 2011). There are some notable exceptions, however, of scholars who connect the model to theoretical expectations, especially in cases of policy learning and emulation where adoptions by neighbors partially determine subsequent adoptions. Jacob, Scherpereel, and Adams (2014), for example, explain that the TLSL is appropriate because diffusion takes time to occur and reduces endogeneity concerns as it has a strict temporal order (see also Böhmelt, Ezrow, Lehrer, & Ward, 2016; Butz, Fix, & Mitchell, 2015; Sugiyama, 2008).

Second, a concerning number of scholars do not properly account for temporal dependence or autocorrelation. In studies of diffusion—and indeed, most studies of political phenomena—outcomes in one period depend heavily on outcomes in prior periods. Yet, more than 29 percent of studies using a TLSL did not address temporal dependence. Third, few articles provide *a priori* diagnostics for spatial autocorrelation before including a TLSL in their model.[10] In our survey, only 7.4 percent of articles with a TLSL reported spatial autocorrelation tests such as Moran's I or Geary's c. Without testing for spatial autocorrelation first, some may improperly include a TLSL when other measures are more appropriate. An exception is Callen (2011), who conducts Moran's I tests at multiple time periods to ensure the entire series exhibits spatial autocorrelation.

Fourth, few studies demonstrate how the effects diffuse across space and/or erode over time (see Böhmelt et al., 2016; Böhmelt & Freyburg, 2015; Lopez-Cariboni & Cao, 2015; Rogers, 2004). In models with TLSLs, the long-term effect of a change in $X_{i,t}$ decays with t. Yet, most studies only explore the *immediate* impact of the TLSL. Brooks et al. (2015) is a notable exception that follows the procedure outlined by Williams and Whitten (2012) to plot the long-term impact of peer diffusion. Doing so demonstrates the rapid decay of the diffusion effect due to a hypothetical price shock in one country as it spreads to its peers, an effect that cannot be captured by solely examining the TLSL coefficient.

Fifth, scholars often confuse the TLSL specification with SAR. The confusion is evident by incorrect model notation. Since TLSLs are exogenous to the outcome at time t, these models should reflect the notation from SLX ($\theta \mathbf{W} y_{t-1}$) rather than

SAR ($\rho\mathbf{W}y$). Of those estimating a TLSL, the term is often reported as $\rho\mathbf{W}y_{t-1}$ (e.g., Cao, 2010; Jordana, Levi-Faur, & Marin, 2011; Linos, 2011). As a consequence of this notational confusion, among the authors that estimate a TLSL (with no concurrent spatial lag, or $\rho\mathbf{W}y$), nearly 28 percent incorrectly refer to their specification as SAR or m-STAR. This is a problem that extends beyond simple notational consistency and results in confusion about whether the spatial processes are simultaneous (SAR) or delayed (SLX).

Spurious Temporally Lagged Spatial Lags

Of the two spatial diffusion processes highlighted above, including a TLSL is appropriate when theory suggests that policy diffusion occurs with a temporal lag. We illustrate this assumed relationship between a TLSL and the outcome by the directed acyclic graph (DAG) shown in Figure 1a. The TLSL, denoted by $\mathbf{W}\boldsymbol{y}_{t-1}$, has an effect on the outcome $\boldsymbol{y}_t$ based on θ. Unfortunately, there are a number of causal pathways that could result in a spurious TLSL that is *not* part of a process generating y_t. In these circumstances, a slight error in model specification might lead to an incorrect conclusion regarding temporally lagged spatial diffusion. In this section, we lay out a number of these common pathways (Figure 1b–e) to demonstrate the inferential risks associated with including irrelevant TLSLs.

Figure 1b shows that the effect of an autoregressive covariate $\boldsymbol{x}$ starting from time t–2 on $\boldsymbol{y}_t$. $\boldsymbol{y}_t$ is impacted by $\boldsymbol{x}_{t-1}$ via the latter's effect on $\boldsymbol{x}_t$. This is represented by an autoregressive parameter $\boldsymbol{\phi}$ shared across all x of $\boldsymbol{x}$. $\boldsymbol{x}_{t-1}$ also affects the TLSL via its impact on $\boldsymbol{y}_{t-1}$. Figure 1c shows how spatial clustering in the previous period $\boldsymbol{y}_{t-1}$ represented by a weights matrix $\mathbf{W}$ and non–spatially clustered component $\boldsymbol{z}$ impact $\boldsymbol{x}_t$ via a parameter θ. $\boldsymbol{x}_t$ in turn impacts $\boldsymbol{y}_t$. A TLSL created with the same matrix would share characteristics of $\boldsymbol{x}_t$ without being part of the data-generating process (DGP) for $\boldsymbol{y}_t$. It does so through both the $\mathbf{W}$ and the effect of $\boldsymbol{x}_{t-1}$ on $\boldsymbol{y}_{t-1}$ which is related to $\mathbf{W}\boldsymbol{z}$ via θ. Figure 1d shows the spurious relationship between a TLSL and $\boldsymbol{y}_t$ when the dependent variable is autoregressive. $\boldsymbol{\phi y}_{t-1}$ affects $\boldsymbol{y}_t$, which also is part of the TLSL. $\boldsymbol{\phi y}_{t-2}$ then affects $\boldsymbol{y}_{t-1}$. Finally, Figure 1e shows a DGP with both a spatially clustered covariate $\boldsymbol{y}_{t-1}$ and an autoregressive dependent variable. This creates a spurious TLSL via sharing a weighting matrix and with $\boldsymbol{x}_t$, the θ effect of the matrix on $\boldsymbol{x}_{t-1}$, and the autoregressive effect of $\boldsymbol{y}_{t-2}$ on $\boldsymbol{y}_{t-1}$.

The last four causal pathways present risks that scholars would find false evidence of temporally lagged diffusion as a result of errors in model specification. In the next section, we provide a series of Monte Carlo experiments to determine how model misspecification in the context of TLSLs influences inferences about policy diffusion.

Monte Carlo Experiments

Our goal in this section is to determine the consequences of including an *irrelevant* TLSL—i.e., one that is not part of the data-generating process—in linear regression models. The following Monte Carlo experiments explore if and to what extent

(a)
$$\mathbf{W}\mathbf{y}_{t-1} \xrightarrow{\theta} \mathbf{y}_t$$

(b)
$$\begin{tikzcd}
\mathbf{x}_{t-2} \arrow[r, "\phi"] & \mathbf{x}_{t-1} \arrow[r, "\phi"] \arrow[d] & \mathbf{x}_t \arrow[r] & \mathbf{y}_t \\
& \mathbf{y}_{t-1} \arrow[r] & \mathbf{W}\mathbf{y}_{t-1} &
\end{tikzcd}$$

(c)
$$\begin{tikzcd}
\mathbf{W}\mathbf{z}_t \arrow[r, "\theta"] \arrow[d, "\theta"'] & \mathbf{x}_t \arrow[r] & \mathbf{y}_t \\
\mathbf{x}_{t-1} \arrow[r] & \mathbf{y}_{t-1} \arrow[r] & \mathbf{W}\mathbf{y}_{t-1}
\end{tikzcd}$$

(d)
$$\begin{tikzcd}
\mathbf{y}_{t-2} \arrow[r, "\phi"] & \mathbf{y}_{t-1} \arrow[r, "\phi"] \arrow[d] & \mathbf{y}_t \\
& \mathbf{W}\mathbf{y}_{t-1} &
\end{tikzcd}$$

(e)
$$\begin{tikzcd}
\mathbf{W}\mathbf{z}_t \arrow[r, "\theta"] \arrow[d, "\theta"'] & \mathbf{x}_t \arrow[r] & \mathbf{y}_t \\
\mathbf{x}_{t-1} \arrow[r] & \mathbf{y}_{t-1} \arrow[r] \arrow[ur, "\phi"] & \mathbf{W}\mathbf{y}_{t-1} \\
& \mathbf{y}_{t-2} \arrow[u, "\phi"] &
\end{tikzcd}$$

Figure 1. Directed Acyclic Graph for Temporally Lagged Spatial Lags.

including an unnecessary TLSL can lead to incorrect inferences about policy diffusion. We generate data to simulate the five causal pathways depicted in Figure 1. For each scenario, we estimate a model where the true data-generating process (shown in Table 2) does not include a temporally lagged diffusion process. We then assess the percentage of simulations where scholars would falsely conclude that a spatial diffusion process shapes the outcome.

Table 2. Monte Carlo Evidence of TLSL False Discovery Rates (FDR) Across Five Scenarios

				False Discovery Rate	
Scenario Number	Description	Data-Generating Process	Estimated Models	Under	Over
1	Control	$Y_t = \alpha + \beta_1 X_1 + \beta_2 X_2 + \epsilon$	$Y_t = \alpha + \beta_1 X_1 + \theta W Y_{t-1}$	0.05	0.05
2	Omitted autoregressive covariate	$Y_t = \alpha + \beta_1 X_1 + \beta_2 X_{AR} + \epsilon$	$Y_t = \alpha + \beta_1 X_1 + \theta W Y_{t-1}$	**0.28**	0.05
3	Omitted spatial-x (X)	$Y_t = \alpha + \beta_1 X_1 + \theta_{WZ} X_{WZ} + \epsilon$	$Y_t = \alpha + \beta_1 X_1 + \theta W Y_{t-1}$	**0.49**	0.04
4	Omitted autoregressive DV	$Y_t = \alpha + \beta_1 X_1 + \beta_2 X_2 + \phi Y_{t-1} + \epsilon$	$Y_t = \alpha + \beta_1 X_1 + \beta_2 X_2 + \theta W Y_{t-1}$	**0.33**	0.06
5	Omitted AR DV, spatial-x (X)	$Y_t = \alpha + \beta_1 X_1 + \theta_{WZ} X_{WZ} + \phi Y_{t-1} + \epsilon$	$Y_t = \alpha + \beta_1 X_1 + \theta W Y_{t-1}$	**0.39**	0.04

Note: Bold type highlights results that substantially exceed the expected FDR of 5 percent.

The first scenario in Table 2 can be thought of as a control in that it examines the effect of including a TLSL when the DGP does not include autoregressive or spatially clustered variables. Since scholars very rarely *knowingly* omit relevant covariates, this is often the DGP that researchers *implicitly* assume when using a TLSL variable. Scenarios two through five in Table 2 correspond to the DGPs in Figure 1. They reveal the consequences of omitting autoregressive and/or spatially clustered variables in models with unnecessary TLSLs.[11] The intuition behind these inferential problems is straight forward: imagine some policy where all observations independently alter the policy in an incremental and positive manner; if scholars neglect this incrementalism by omitting the lagged dependent variable, then controlling for the weighted average of neighbors' policies would give the false impression of a diffusion process. In reality, they are all acting independently based on their own policy histories. Likewise, neglecting any shock that influences all the observations in a neighborhood similarly will give the impression of policy diffusion. Keep in mind that the TLSL is not part of the DGP for any of the five scenarios, so retrieving a statistically significant coefficient is evidence of a spurious TLSL.

All of the models include an intercept with $\alpha = 1$. The covariates are drawn from uniform distributions (X_1), normal distributions (X_2 and Z), or based on an autoregressive process (X_{AR}), and the error term is normally distributed with mean 0 and variance equal to 1. Z is used in Scenarios 3 and 5 to generate a spatially clustered covariate X_{WZ}, specifically:

$$X_{WZ} = \sum_j w_{ij} z_{jt} + e \tag{16}$$

where Z is a vector of individual, non–spatially clustered components for each time point t and e is an error term. For each observation i and other observations j, w_{ij} is from $\mathbf{W}$, an $N \times N$ matrix of Euclidean distances between all ij dyads based on each observation's "location." Locations were drawn from $\mathcal{N}(0,1)$. The connectivity between the i and j is treated as symmetrical and undirected (see Neumayer & Plümper, 2010, for more details) and is not row-standardized. We use θ_{WZ} to denote the coefficient estimated for this variable. The parameters for the various DGPs are initially set to $\beta_1 = 2$, $\beta_2 = 3$, $\phi = 0.6$, and $\theta_{WZ} = 0.001$. Later we vary ϕ and θ_{WZ} to examine how the strength of the autoregression and spatial clustering impacts TLSL false identification.

In all of the experiments, we included a TLSL in the regressions that was not part of the data-generating process. The TLSL was created for each observation i given other observation j at time t:

$$\mathbf{W}Y_{t-1} = \sum_j w_{ij} y_{jt-1} \tag{17}$$

We used the same $\mathbf{W}$ weighting matrix to create the TLSL as we used in the DGP for X_{WZ}. In each scenario, we made 100 draws for 100 units over 100 time points per unit following a two-period burn-in.[12]

We first look for a specific Type-I error, which in this case is evidence in favor of rejecting the null hypothesis of no temporally lagged spatial diffusion process (the null hypothesis is actually true). We define the false discovery rate (FDR) as the proportion of simulations where the *p* value for the estimated TLSL θ_{WYt-1} coefficient was $p < 0.05$. The second-to-last column of Table 2 shows the FDR of the TLSL in the Monte Carlo experiments for the underfitted model in all five scenarios. A clear pattern arises from these experiments that should be concerning for scholars studying spatial processes; in all of the models that omitted autoregressive and/or spatially clustered covariates that were part of the DGP (Scenarios 2–5), the TLSL false discovery rate was much higher than expected. In these scenarios, the FDR ranges from about 0.28 for the scenario with an omitted autoregressive covariate to almost 0.50 for the scenario with an omitted spatially clustered covariate. If the researcher has omitted an autoregressive or spatially clustered covariate, there is a reasonably high risk of discovering a spurious TLSL.

As revealing as those FDR patterns are in Table 2, they are specific to a unique set of parameters and conditions. The DAGs in Figure 1 suggest that the degree of autocorrelation and spatial clustering in the covariate that is part of the DGP could inflate the TLSL false discovery rate. To explore this, we ran further simulations for Scenarios 2 and 3[13] varying the autoregressive term (ϕ) and coefficient for the spatially clustered variable (θ_{WZ}) and show the results in Figure 2. For now, focus on the dashed line which shows the TLSL false discovery rate for Scenarios 2 and 3 (we describe the solid line below).

The TLSL false discovery rate for these models is much higher than indicated by its *p*-value, and the problem is exacerbated when the omitted covariate's level of autoregression (ϕ in Scenario 2) or spatial clustering (θ_{WZ} in Scenario 3) increases. The results reveal that scholars should be wary of the widespread use of TLSLs in models prone to misspecification. In the next section, we provide guidelines for preventing these inferential errors by relying heavily on strong theoretical justification and *a priori* diagnostics.

Guidelines for Preventing TLSL Type-I Errors

The Monte Carlo experiments revealed a couple of troubling trends about model misspecification. When a true temporally lagged policy diffusion does not exist, nearly any type of omitted relevant variable will lead to false inferences about diffusion. Indeed, the false discovery rate is unacceptably high across all four scenarios, and that rate worsens as the strength of spatial and temporal autocorrelation increases. It is clear that TLSLs should not be included in these situations. However, if one can correctly specify the rest of the model, then is it harmful to include a possibly irrelevant TLSL? The experiments suggest that the false discovery rate returns to acceptable rates,[14] so is it advisable to include a TLSL if in doubt? We would caution against this approach (see also Gujarati, 2003, p. 514). Additional Monte Carlo experiments of Scenarios 2 and 3 (explored more in the supporting information) reveal that including a TLSL—if not in the DGP—produces bias in the estimates of ϕ

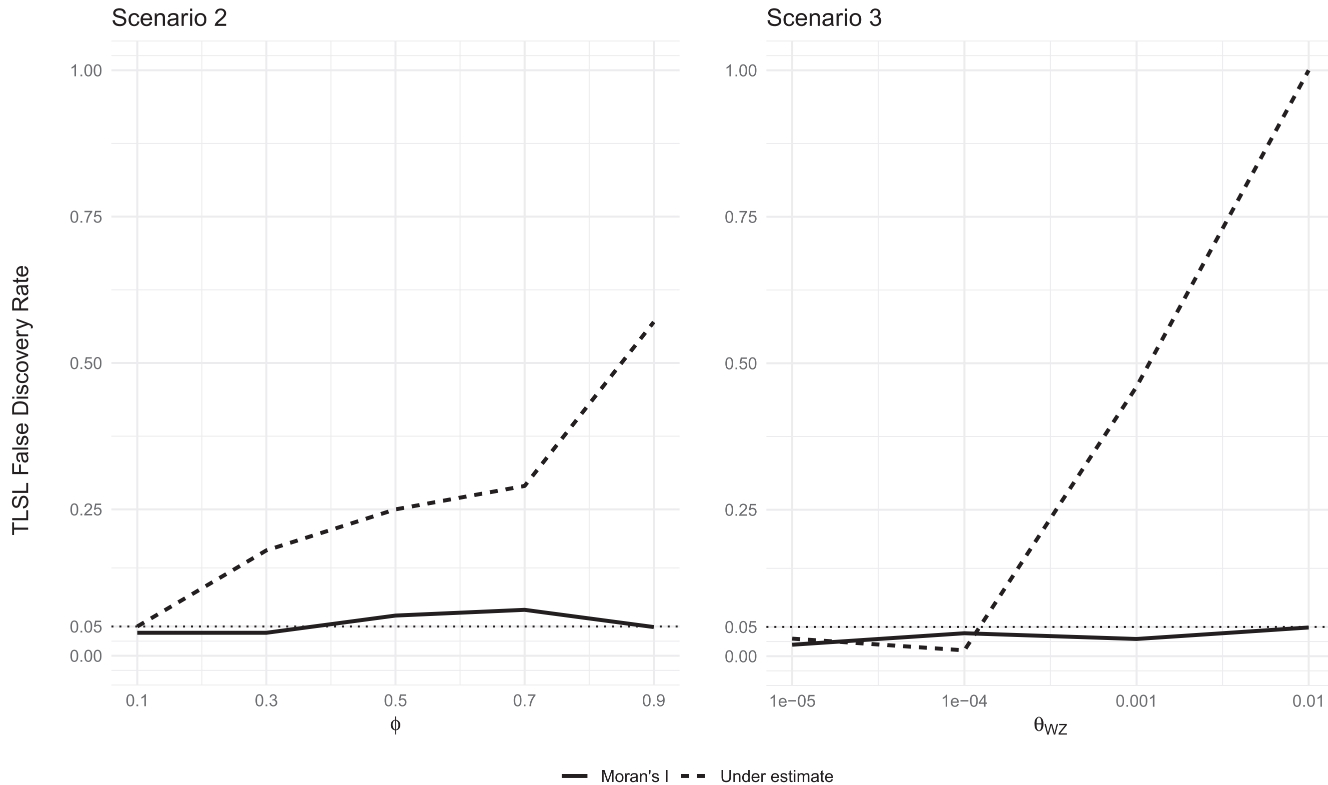


Figure 2. TLSL False Discovery Rate across the Omitted Covariate's Degree of Autoregression (ϕ in Scenario 2) and Spatial Clustering (θ_{WZ} in Scenario 3), with and without Moran's *I*

(Scenario 2) and θ (Scenario 3). Furthermore, there is also the risk of bias to the other covariates depending on the particular scenario. It is clear, then, that there is an inferential penalty associated with this attempt to prevent false TLSLs.

The simplest solution is also perhaps the least helpful in practice: specify the correct model. This guidance is of little practical use because in reality researchers do not know the data-generating process *ex ante*. Even under the best possible conditions where scholars know which variables to include, it may not be practical or possible to observe all relevant autocorrelated and spatially clustered variables. Thus, we are left with two troubling facts about applied policy diffusion research. First, the conditions that lead to falsely identifying TLSL at high rates are endemic to political science research and second, they are not easily addressed within standard regression modeling frameworks. How can researchers avoid false evidence of diffusion (by omitting a relevant variable), and avoid biasing the other coefficients (by including an irrelevant TLSL to an otherwise well-specified model)?

We offer guidelines for scholars to use when confronting the inherent difficulty in properly specifying policy processes across time and space. The first guideline is that theory should motivate all of the decisions about model specification. In the study of policy diffusion, scholars typically want to explain why some governments' policies converge while others do not. A reasonable starting place is to posit a diffusion process where governments learn from and emulate others, so that over time their policies converge. In their efforts to hone in on diffusion, there is a chance that scholars neglect the multitude of alternative pathways (as shown in Figure 1) that could also produce policy convergence. Scholars must therefore think carefully about these pathways and lean heavily on their theories to eliminate possibilities. If they are unable to envision a diffusion process operating outside of a government's own policy trajectory, common shocks in a region, or some combination of both, then the TLSL might be unnecessary. Careful theorizing can therefore help scholars steer clear of a variety of specification issues.

Even if the empirical model is on solid theoretical foundations, the possibility of the types of model misspecification described above remain. Our second guideline is to follow Darmofal's (2015, p. 69) advice and trust but verify. In this context, it means using one's theory as a starting point to gain insight about the preliminary model but then determining the appropriate specification through diagnostic tests. Most econometrics textbooks devote considerable space to the sources, consequences of, and tests for model specification errors (e.g., Gujarati, 2003, Chapter 13), and others have explored these issues in time series (Achen, 2000; deBoef & Keele, 2008; Keele & Kelly, 2006) and spatial models (Darmofal, 2015; Neumayer & Plümper, 2010), so we focus on specification issues related to models with TLSLs. Following Darmofal (2015), we advocate a simple test (such as Moran's *I*, Geary's *c*, or join count analysis[15]) to determine if there is enough evidence of spatial clustering to require a TLSL. For time series cross-sectional data, one can test for spatial autocorrelation at each observed time point. It may be that a series is spatially correlated for one set of time points but not for another. For example, falling transportation costs over an observation period could make contact between governments more frequent later in the period and spread an innovation across space more frequently. Such a pattern would result in spatial autocorrelation in the adoption of the innovation later in the observation period, where there had not been one earlier. This could easily be accounted for in a regression analysis by interacting a TLSL by some indicator of time that was supported by the repeated Moran's *I* test. If tests of spatial clustering do not indicate that a TLSL is spatially autocorrelated, then there is not a justification for including the TLSL in the regression model and including it might lead to incorrect conclusions.

How well does our proposed solution do in preventing the false discovery of diffusion processes? The solid line in Figure 2 shows the false discovery rate for spatial clustering in the TLSL for Scenarios 2 and 3 across the strength of temporal (ϕ) and spatial autocorrelation (θ). Recall that the false discovery rate is defined as the proportion of simulations with *p*-values for Moran's *I* less than 0.05. In both cases, Moran's *I* consistently has an expected false discovery rate of approximately 5 percent across values of temporal and spatial autocorrelation. This buttresses our claim that the Moran's *I* can be used to determine if there is spatial autocorrelation in the TLSL and that the absence of evidence for spatial clustering should dissuade scholars from its inclusion. We therefore advocate a two-step process based on careful theorizing and specification diagnostics.

Applications

This section offers a departure from traditional methodological studies. Common practice would be to identify a prominent study or two that resulted in *inappropriate* uses of TLSLs. As a practical consideration, this is difficult because scholars rarely provide the materials needed to identify inappropriate uses of TLSLs. Replication materials are still not required at some journals, and certainly scholars have not shown a willingness to provide the data and script files used to *generate* TLSLs (most notably, the weights matrix), which are necessary for the tests described above. The result is a systematic lack of transparency needed for third parties to conduct stress tests on models of policy diffusion.

Instead, we highlight three *appropriate* uses (or non-uses) of TLSLs. The first two examples show how easily one would find a spurious TLSL in real-world data with some minor changes to a well-specified model. The third example demonstrates how scholars choose to include TLSLs in their model in practice, which we extend with an example for presenting spatial long-term effects. The insights from three examples echo the conclusions of the simulations and show that the risk of spurious TLSLs is quite high in all but the most completely specified models.

Lipsmeyer and Zhu (2011)

The first example comes from Lipsmeyer and Zhu's (2011) analysis of the effects of immigration on unemployment benefits in 15 EU member states from 1971 to 2007, and more specifically, how domestic political institutions and labor market integration condition those effects. The authors include three interactive variables to measure how changes in economic integration, union density, and the percentage of left-wing parliamentary seats condition states' responses to immigration flows. In addition to some economic control variables (such as changes in GDP, unemployment, trade, and FDI), the authors include the lagged value of unemployment entitlement (i.e., a lagged dependent variable).[16] In Table 3, we successfully replicate the authors' model (Table 1, page 654). The empirical results "suggest that EU integrative forces demonstrate less of an impact on unemployment entitlements than domestic political forces" (Lipsmeyer & Zhu, 2011, p. 660).

For all intents and purposes, this appears to be a well-specified model with no obvious omitted variables. What happens if we slightly alter the model specification by first, including an unnecessary TLSL, and then second, omitting the relevant lagged dependent variable? The TLSL coefficient in Model 1 (Table 3) is not statistically different from zero, which suggests that there are no spatial spillovers among EU member nations.[17] In Model 2, we purposely exclude the lagged

Table 3. Replication of Lipsmeyer and Zhu (2011)

	Replication	Model 1	Model 2
Entitlement$_{t-1}$ (ϕ)	0.890**	0.899**	
	(0.019)	(0.021)	
TLSL (θ)		−0.025	0.277**
		(0.027)	(0.083)
Δ Immigration	−0.250	−0.257	−0.156
	(0.169)	(0.171)	(0.197)
Δ Integration	1.921	2.063	−0.072
	(1.984)	(2.012)	(2.608)
Union density$_{t-1}$	−2.612	−2.928*	−12.393**
	(1.653)	(1.654)	(4.773)
Left seats$_{t-1}$	0.011	0.012	−0.02
	(0.012)	(0.012)	(0.024)
Δ Immigration × Δ Integration	−0.417	−0.418	0.563
	(0.941)	(0.948)	(1.187)
Immigration × Union density$_{t-1}$	0.385	0.388	0.231
	(0.260)	(0.262)	(0.287)
Δ Immigration × Left seats$_{t-1}$	0.004*	0.004*	0.003
	(0.002)	(0.002)	(0.003)
Δ GDP	−0.0001	−0.0001	−0.00001
	(0.0002)	(0.0002)	(0.0003)
Δ Unemployment	−0.125	−0.122	−0.163
	(0.083)	(0.082)	(0.123)
Δ Trade	0.004	0.003	−0.0003
	(0.014)	(0.014)	(0.02)
Δ FDI	−0.071*	−0.072	−0.036
	(0.043)	(0.044)	(0.063)
Intercept	3.988**	4.487**	25.832**
	(1.100)	(1.141)	(3.701)
N	496	496	496
ρ	0.414	0.389	0.808
R^2	0.961	0.964	0.569

Note: OLS regression with panel-corrected standard errors and an AR(1) correction.
$^{*}p < 0.10$, $^{**}p < 0.05$.

dependent variable to observe how our inferences regarding the TLSL change. The θ coefficient for the TLSL is now highly significant and positive, which suggests that countries respond positively to other EU member nations' unemployment benefits. Additionally, the key finding of interactive effects between changes in immigration and left-wing seats disappears.

Making the reasonable—but completely unnecessary—change to the model specification in the form of removing the lagged dependent variable in exchange for the TLSL results in incorrect inferences regarding spatial diffusion. Model 2 essentially trades long-term direct effects for long-term indirect effects. In the replication model, the covariates operate through the ϕ; a change in x_i has a short- and long-term effect on y_i (as shown on page 660). In Model 2, the inclusion of the TLSL means that the covariate operates through both time and space; a change in x_i has only a short-term direct effect on y_i (depicted by the coefficient for x), but has a spatial effect on y_j (i's first-order neighbors) at time $t + 1$, another spatial effect on y_k (j's first-order neighbors, or i's second-order neighbors) at time $t + 2$, and so on. Thus, the difference between models with a lagged dependent variable (replication model) and a TLSL (Model 2) is quite large and has meaningful consequences for the inferences derived from the model. In this example, tests suggest that the authors' specification is more appropriate, which points to the spatial long-term effects found in Model 2 being the result of a spurious TLSL.

Hollyer, Rosendorff, and Vreeland (2011)

Hollyer, Rosendorff, and Vreeland (2011) examine whether electoral politics provide incentives for leaders to be more transparent. The authors develop a new measure of transparency for 188 countries from 1961 to 2007 based on "a government's willingness to disseminate policy-relevant data" and demonstrate that democracies are in fact more transparent. The authors' model is parsimonious and is intended to control for the confounding effects of wealth and participation in IMF programs (1199) and possible concerns about unmodeled country-specific variation (with fixed effects) and the passage of time (with cubic polynomials). In Table 4, we successfully replicate Model 6 (Table 3, p. 1201). As theorized, the measure of democracy based on the Polity index is statistically significant and positive, indicating that democracies provide more data.

In the next specification, we incorporate a TLSL based on a row-standardized binary weights matrix where every observation is coded as a neighbor. There are no obvious reasons why governments' decisions about making data available would be spatially clustered. Indeed, in Model 1, the coefficient for the TLSL is not statistically significant at conventional levels, indicating that there is no evidence of a temporally lagged diffusion process.

The next two models, however, demonstrate that minor changes to the model specification can change this conclusion through spurious TLSLs. In Model 2, we exclude the cubic polynomials that control for the trend toward increased transparency over time. Since these are simple counters that are based on the number of years

Table 4. Replication of Hollyer et al. (2011)

	Replication	Model 1	Model 2	Model 3
TLSL (θ)		0.215	0.892***	0.237*
		(0.133)	(0.041)	(0.134)
Polity 2	0.003**	0.003**	0.002	0.004**
	(0.002)	(0.002)	(0.002)	(0.002)
GDP per capita	0.0002	0.0003	0.0002	
	(0.001)	(0.001)	(0.001)	
Under IMF	0.054***	0.053***	0.052***	0.058***
	(0.0128)	(0.0129)	(0.013)	(0.014)
Time	0.029***	0.028***		0.026***
	(0.005)	(0.006)		(0.005)
Time^2	−0.0003*	−0.0005**		−0.0004**
	(0.0002)	(0.0002)		(0.0002)
Time^3	−0.0000	0.0000		0.0000
	(0.0000)	(0.0000)		(0.0000)
Intercept	0.266***	0.183***	0.104***	0.187***
	(0.028)	(0.049)	(0.027)	(0.047)
N	5,566	5,451	5,451	5,610
R^2	0.329	0.318	0.305	0.367

Note: OLS regression with standard errors clustered by country.
*$p < 0.10$,**$p < 0.05$,***$p < 0.01$.

since 1958, it is clear that they are representative of Scenario 2 in the Monte Carlo experiments (omitted autoregressive covariate). In Model 3, we exclude the measure of wealth (*GDP per capita*). We know that countries' levels of wealth are likely to be spatially clustered for a variety of reasons (e.g., Krieckhaus, 2006), so this situation mimics Scenario 3 in the Monte Carlo experiments (omitted spatial-X). In both models, the TLSL coefficient is positive and statistically significant, which is consistent with patterns of spurious TLSLs in the Monte Carlo simulations. Essentially, one could falsely infer from these models that data dissemination policies are driven by neighboring countries' policies.

In addition to the shift in the meaning of the effects detailed in the previous example, there is a substantial risk of falsely concluding that democracy has no effect on transparency. In Model 2, the coefficient for *democracy* is closer to zero and not statistically significant. Moreover, based on the statistically significant coefficient for the TLSL, scholars are likely to incorrectly infer that the process of disseminating data contains a positive spatial pattern. Keep in mind that the two modifications to model specification that produced these changes at first appeared minor and relatively innocuous, but their inferential consequences are quite severe.

Gilardi and Wasserfallen (2016)

In the final example, we identified a notable study of policy diffusion that proceeded in a manner consistent with our guidelines by testing for spatial autocorrelation with Moran's *I* and then presenting a fully specified model that takes into account spatial and temporal dynamics. Gilardi and Wasserfallen's (2016) analysis of the effects of socialization on tax competition in the 16 Swiss cantons between 1990 and 2007 is a perfect example of the thoughtful use of TLSLs. The authors start with simple geographic specifications of the interconnectivities within the Swiss federal system (contiguity) and then proceed to more nuanced specifications (commuters, membership in regional conferences). They then multiply these weights matrices by tax rates in the previous year to test their conditional hypotheses of socialization of tax competition.

Their project offers some guidance as to how to proceed with TLSLs in a variety of respects. First, the authors derive clear expectations for why tax competition occurs with a one-year lag (thus justifying the TLSL instead of a concurrent lag) based on important anecdotal examples (Gilardi & Wasserfallen, 2016, pp. 48–49). Second, the authors use Moran's I as a starting point to illustrate how the degree of spatial autocorrelation varies across cantons by tax rate. Deeper analysis of all models reveals clear patterns of spatial autocorrelation.[18] Third, the authors present a complete model specification that incorporates unobserved heterogeneity at both the government level (via canton- and region-specific fixed effects) and across time (via year fixed effects).

If there is one weakness of the manuscript, it is in the form of a missed opportunity. Because there are no quantities of interest provided, the readers are unable to explore how the explanatory variables influence tax rates through spatial long-term effects. We now demonstrate two quantities of interest from Models 1–5 in Gilardi and Wasserfallen (2016, pp. 48–49) that are easy to calculate and quite revealing. Table 5 calculates the average total, direct, and indirect spatial long-term effects (based on equation [15]) for a one standard deviation increase in the two variables that are statistically significant in some models: *deficit per capita* and *unemployment rate*. Indirect effects represent the effect of the increase in observation i on all other observations, and direct effects represent feedback effects. All observations have the identical direct effect at time t; since this is β for all i, the direct effects in Table 5 only depict the feedback effects. We calculate 95 percent confidence intervals using the percentile method based on 1,000 draws from the multivariate normal distribution implied by the model.

Displaying the average direct and indirect spatial long-term effects in this manner can help illuminate the amount of the total effect that arises due to feedback. For both variables in most models, feedback effects make up less than a tenth of the total spatial long-term effects. Table 5 displays additional evidence in favor of carefully specifying weights matrices based on theoretical motivations (see Neumayer & Plümper, 2016). The size and statistical significance of the effect sizes varies a great deal across the five models. Most notably, the average total spatial long-term effect of the *unemployment rate* is nearly 10 times as large in Model 5 (−0.208) compared to Model 4 (−0.021), though it is not statistically significant in Model 4. The various weights matrices used by Gilardi and Wasserfallen (2016) are all quite reasonable yet lead to substantially different effect sizes. Without carefully considering the specification, scholars are liable to make different inferences due to small changes in the weights matrix.

While the spatial long-term effects shown in Table 5 provide a sense of the *average* total effect, it ignores a great deal of variation that may be of interest to scholars.

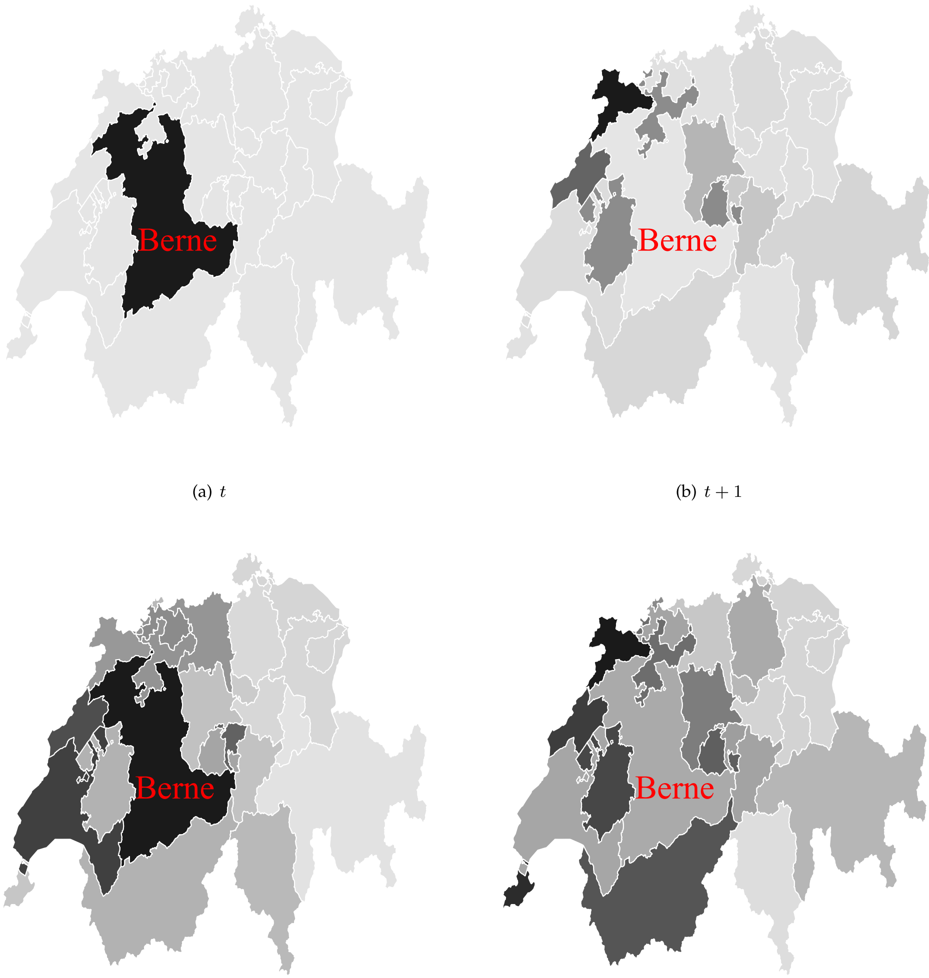
Table 5. Average Total, Direct, and Indirect Spatial Long-Term Effects (t = 15) in Models 1-5 of Gilardi and Wasserfallen (2016)

	Total	Direct	Indirect
Model 1: Neighbors			
Deficit per capita	0.022	0.001	0.020
Δ = +0.057	[−0.006, 0.067]	[−0.0004, 0.005]	[−0.005, 0.062]
Unemployment	−0.062	−0.004	−0.058
Δ = +1.738	[−0.185, 0.024]	[−0.014, 0.002]	[−0.171, 0.023]
Model 2: Commuters			
Deficit per capita	0.017*	0.001*	0.016*
Δ = +0.057	[−0.0008, 0.052]	[−0.00003, 0.003]	[−0.0008, 0.049]
Unemployment	−0.045	−0.002	−0.043
Δ = +1.738	[−0.150, 0.013]	[−0.010, 0.0007]	[−0.140, 0.012]
Model 3: Cantons not in same conference			
Deficit per capita	0.043**	0.002**	0.041**
Δ = +0.057	[0.002, 0.129]	[0.0001, 0.007]	[0.002, 0.122]
Unemployment	−0.110*	−0.005*	−0.105*
Δ = +1.738	[−0.358, 0.016]	[−0.019, 0.0008]	[−0.339, 0.016]
Model 4: Cantons in same conference			
Deficit per capita	0.008	0.0003	0.007
Δ = +0.057	[−0.002, 0.028]	[−0.00002, 0.002]	[−0.002, 0.026]
Unemployment	−0.021	−0.001	−0.020
Δ = +1.738	[−0.081, 0.009]	[−0.005, 0.0003]	[−0.076, 0.009]
Model 5: Cantons in and not in same conference			
Deficit per capita	0.070**	0.004**	0.066*
Δ = +0.057	[0.0002, 0.231]	[0.00001, 0.014]	[0.0002, 0.217]
Unemployment	−0.208*	−0.013*	−0.196*
Δ = +1.738	[−0.582, −0.021]	[−0.044, 0.0003]	[−0.702, 0.005]

Note: Changes reflect one-standard deviation increase; *W*s in models 2–5 are from 2000 (see Gilardi & Wasserfallen, 2016, p. 56).
*$p < 0.1$, **$p < 0.05$.

In Figure 3, we demonstrate how a shock to the deficit (a one standard deviation increase to *deficit per capita*) in Berne diffuses over time to influence the other cantons, as well as itself (based on the "not in the same conference" weights matrix in Model 5).

At time t, a .057 increase *deficit per capita* leads to a .049 increase in the tax rate for an annual income of CHF 150,000 in Berne (i.e., $\beta = 0.867$, so $0.867 \times 0.057 = 0.049$). In the next year ($t + 1$), those cantons that are not in the same regional conference (i.e., first-order neighbors) respond positively to Berne's increase in the deficit. The size of the responses (darker shades mean larger effects) is a function of how many other neighbors each canton has; the canton of Jura in the northeast experiences the largest impulse at time $t + 1$. In the second year ($t + 2$), Berne experiences the largest effect because it now competes with its first-order neighbors in the previous year. These are the feedback effects that are in the "Direct" column of Table 5. The final panel ($t + 3$) shows that Berne's deficit shock expands out to almost all of the cantons within the network through higher-order connections, though with the caveat that the effects quickly become quite small. These two methods demonstrate how to make inferences about the spatial and temporal dynamics at work in models with TLSLs, both on average (see Table 5) and as counterfactual scenarios (see Figure 3).



(a) t (b) $t+1$

(c) $t+2$ (d) $t+3$

Figure 3. Spatial Long Term Effects of a One Standard Deviation Increase in Deficits Per Capita in Berne at Time t on Other Swiss Cantons (Model 5).
Note: Estimates are based on the weights matrix capturing cantons *not* in the same regional conference. Darker shades represent larger effects.

Conclusion

Policy diffusion scholars interested in how governments learn from their neighbors often examine an implicitly spatial process without recognizing it. To empirically model diffusion, scholars often utilize the weighted sum (or average) of the outcomes in a government's neighborhood. The intuition is that this temporally lagged spatial lag, or TLSL, will capture spatial diffusion processes that occur over time. In practice, however, the benefits of TLSLs come with the costs of greater confusion about the exact process being modeled, difficulty in interpreting effects, and a greater risk of model misspecification. We offer insight to guide scholars so that they select the appropriate empirical approach for their theory, hone in on the correct model specification, and fully interpret the effects across time and space.

First, we argue that dyadic EHA and spatial models offer two pictures of policy diffusion processes. Though both share some critical features (such as linking the observations across space), they depict substantively different spatial diffusion processes. We then provided guidance as to when one approach is more appropriate than the other. Unfortunately, our survey of the policy diffusion literature suggested that there is often a large disconnect between one's theory and the empirical tests and a considerable amount of notational confusion. Indeed, the survey suggests that scholars should be more diligent in their efforts to select an empirical model that is consistent with their theory.

Second, this project is the first that we know of to demonstrate the correct interpretation of the other covariates in the context of TLSLs. Scholars estimate models with TLSLs as a way of controlling for other observations' outcomes in the past, so it makes sense that any variable that affects these outcomes will also have a lasting effect. We call these effects *spatial long-term effects* and demonstrate how to calculate them as well as simple summary statistics. We hope that these tools will empower scholars to explore their diffusion models in deeper and more thought-provoking ways.

Third, our experiments have shown that there are multiple conditions that can lead to false inferences of spatial diffusion in TLSL models. These conditions are widespread, and occur when the model omits an autoregressive covariate, a spatially autocorrelated covariate, an autoregressive lagged dependent variable, or any combination thereof. When the omitted variables are included in the model alongside an irrelevant TLSL, the false discovery rate is reduced to the proper level, but bias can be introduced into the coefficients. To avoid these problems, we offer simple guidelines for specifying models with TLSLs: first, use theory to eliminate alternative pathways, and if necessary, control for them in the model; and second, use the appropriate specification diagnostics to confirm your choice of specification. Indeed, our findings strongly support the conclusion that researchers should only include TLSLs when there is evidence of spatial clustering, as given by a test such as Moran's *I*. We are introducing a new software package for the R programming language–`spatialWeights`[19]–to make it easy for researchers to both create spatial weights and test for spatial clustering in TSCS data. We describe this package and provide an example in the supporting information.

While we draw insights from the policy diffusion literature, our methodological recommendations apply to other diffusion process where TLSLs are theoretically appropriate, such as the dispersion of governing structures (e.g., Houle, Kayser, & Xiang, 2016; Miller, 2016), tax and regulatory structures (e.g., Baccini & Koenig-Archibugi, 2014; Fink, 2011; Jordana et al., 2011), or conflict (e.g., Buhaug & Gleditsch, 2008; Siverson & Starr, 1990). Moving forward, scholars should pay careful attention to how both temporal and spatial processes influence diffusion.

Cody A. Drolc is a Ph.D. candidate at the University of Missouri. His research focuses on accountability, oversight, and the politics of policy implementation and formation.

Christopher Gandrud, Ph.D., is head of economics and experimentation at Zalando SE. His research focuses on causal inference methods, particularly for large scale organizations.

Laron K. Williams is an associate professor of political science at the University of Missouri. His research focuses on spatial econometrics, party competition, and voting behavior.

Notes

The authors would like to thank participants of the 2017 Modeling Politics and Policy Conference at Texas A&M University, the four anonymous reviewers, and editors for their helpful feedback and invaluable comments.

1. This finding builds upon those in Bellemare, Masaki, and Pepinsky (2017). They examined the effects of lagged, but not spatially weighted, dependent variables in regression models when there are causal dynamics among unobservables.
2. Scholars should be aware that the relational nature of the dependent variable raises the possibility of "emulation bias" (Boehmke, 2009b).
3. Of course, some observations may not be connected to any other observations. These are called isolates and they are completely unaffected by spatial dependence.
4. We end up at equation (2) by moving all the terms that involve $\boldsymbol{y}$ to the left-hand side and then solving for $\boldsymbol{y}$ (see Ward & Gleditsch, 2008, pp. 44–45 for more detail).
5. The SLX is much more flexible than the SAR model in specifying conditional spatial dependence and spatial heterogeneity (Wimpy et al., forthcoming). One can specify different patterns of connections (**W**) for different variables, and easily produce interactions to test expectations of conditional spatial dependence.
6. For an in-depth discussion of the consequences of row-standardizing weights matrices, see Neumayer and Plümper (2016).
7. In the supporting information Appendix, we provide additional Monte Carlo experiments that explore the consequences of two types of model misspecification when the true process operates through multiple avenues: first, omitting a relevant avenue, and second, incorrectly specifying the diffusion avenue. The first type will most likely lead to bias in the estimate of the included diffusion avenue (since it will probably be correlated with the omitted avenue), and the second type produces unacceptably high false discovery rates. These results show that scholars must carefully consider whether policy convergence is the result of multiple avenues of spatial diffusion, and if so, properly model those avenues.
8. This is analogous to the debate over whether to row-standardize the weights matrix in spatial econometric models (Neumayer & Plümper, 2010).
9. If the dependent variable is a discrete event (such as adoption), then the spatial long-term effects are probabilistic. Changing future values of the percentage of previous adopters involves classifying the outcome based on the probability that $y_t = 1$ (see Williams, 2016).
10. See Darmofal (2015) and the following sections for discussion of why this is important.
11. Gujarati (2003) would refer to this specification problem as an "underfitted" model that also includes an irrelevant variable (TLSL).
12. We found that the results did not substantively change with more simulations.
13. The other TLSL scenarios are variations on these two.

14. In the last column of Table 2, we show that correctly specifying the other parts of the model ("Over") returns the FDR to acceptable levels.
15. Moran's I is generally preferred as Geary's c gives greater weight to extreme values (see Cliff & Ord, 1981, pp. 14–15). Moran's I is a measure of spatial (dis)similarity $\mathrm{I} = \frac{N}{S}\frac{\Sigma_i \Sigma_j w_{ij}\left(y_i - \bar{y}\right)\left(y_j - \bar{y}\right)}{\Sigma_i\left(y_i - \bar{y}\right)^2}$, where N is the number of observations, S is the sum of the weights, w_{ij} is an element of $\mathbf{W}$. y_i and y_j are the values of the random variable at locations i and j. $\bar{y}$ is the mean y. A Moran's I test statistic with a p value below an accepted value, such as $p < 0.05$, suggests evidence against the null hypothesis of spatial randomness. Join count analysis can be used when the data are dichotomous (see Chapter 4 of Darmofal, 2015, for more details).
16. The use of a lagged dependent variable is quite common in public policy research and makes sense here because the authors "assume a path-dependent process whereby the lagged dependent variable controls for the level of entitlements in the previous year" (Lipsmeyer & Zhu, 2011, p. 654).
17. We use a row-standardized binary weights matrix where all EU member nations are neighbors.
18. With few exceptions, the Moran's I is statistically significant (at least at the 90 percent confidence level) for each year and each model presented in Table 3 (Gilardi & Wasserfallen, 2016, p. 58).
19. Available for download from: https://github.com/christophergandrud/spatialWeights.

References

Achen, Christopher H. 2000. "Why Lagged Dependent Variables Can Supress the Explanatory Power of Other Independent Variables." Presented at the Annual Meeting of Political Methodology, Los Angeles, CA, July 20–22.

Baccini, Leonardo, and Mathias Koenig-Archibugi. 2014. "Why Do States Commit to International Labor Standards? Interdependent Ratification of Core ILO Conventions, 1948–2009." *World Politics* 66 (3): 446–90.

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.

Bellemare, Marc F., Takaaki Masaki, and Thomas B. Pepinsky. 2017. "Lagged Explanatory Variables and the Estimation of Causal Effect." *The Journal of Politics* 79 (3): 949–63.

Berry, Frances Stokes, and William D. Berry. 1990. "State Lottery Adoptions as Policy Innovations: An Event History Analysis." *American Political Science Review* 84 (2): 395–415.

Berry, William D., Richard C. Fording, and Russell L. Hanson. 2003. "Reassessing the "Race to the Bottom" in State Welfare Policy." *The Journal of Politics* 65 (2): 327–49.

Boehmke, Frederick J. 2009a. "Approaches to Modeling the Adoption and Diffusion of Policies with Multiple Components." *State Politics & Policy Quarterly* 9 (2): 229–52.

———. 2009b. "Policy Emulation or Policy Convergence? Potential Ambiguities in the Dyadic Event History Approach to State Policy Emulation." *The Journal of Politics* 71 (3): 1125–40.

Boehmke, Frederick J., and Richard Witmer. 2004. "Disentangling Diffusion: The Effects of Social Learning and Economic Competition on State Policy Innovation and Expansion." *Political Research Quarterly* 57 (1): 39–51.

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

Böhmelt, Tobias, and Tina Freyburg. 2015. "Diffusion of Compliance in the 'Race towards Brussels?' A Spatial Approach to EU Accession Conditionality." *West European Politics* 38 (3): 601–26.

Brooks, Sarah M., Raphael Cunha, and Layna Mosley. 2015. "Categories, Creditworthiness, and Contagion: How Investors' Shortcuts Affect Sovereign Debt Markets." *International Studies Quarterly* 59 (3): 587–601.

Brooks, Sarah M., and Marcus J. Kurtz. 2012. "Paths to Financial Policy Diffusion: Statist Legacies in Latin America's Globalization." *International Organization* 66 (1): 95–128.

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

Butz, Adam M., Michael P. Fix, and Joshua L. Mitchell. 2015. "Policy Learning and the Diffusion of Stand-Your-Ground Laws." *Politics & Policy* 43 (3): 347–77.

Callen, Zachary. 2011. "Local Rail Innovations: Antebellum States and Policy Diffusion." *Studies in American Political Development* 25 (2): 117–42.

Cao, Xun. 2010. "Networks as Channels of Policy Diffusion: Explaining Worldwide Changes in Capital Taxation, 1998–2006." *International Studies Quarterly* 54 (3): 823–54.

Carsey, Thomas M., and Jeffrey J. Harden. 2013. *Monte Carlo Simulation and Resampling Methods for Social Science*. Thousand Oaks, CA: Sage.

Case, Anne C., Harvey S. Rosen, and James R. Hines. 1993. "Budget Spillovers and Fiscal Policy Interdependence: Evidence from the States." *Journal of Public Economics* 52 (3): 285–307.

Clay, K. Chad, and Andrew P. Owsiak. 2016. "The Diffusion of International Border Agreements." *The Journal of Politics* 78 (2): 427–42.

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

Cook, Scott J., Jude C. Hays, and Robert J. Franzese. 2015. "Model Specification and Spatial Interdependence." *APSA Conference Paper* [Online]. http://www.sas.rochester.edu/psc/polmeth/papers/Cook Hays Franzese.pdf. Accessed November 29, 2019.

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

Davis, John C., and Sean Nicholson-Crotty. 2016. "Partisanship and Tax Competition in the American States." *Journal of Public Policy* 36 (3): 437–56.

deBoef, Suzanna, and Luke Keele. 2008. "Taking Time Seriously." *American Journal of Political Science* 52 (1): 184–200.

Desmarais, Bruce A., Jeffrey J. Harden, and Frederick J. Boehmke. 2015. "Persistent Policy Pathways: Inferring Diffusion Networks in the American States." *American Political Science Review* 109 (2): 392–406.

Fay, Daniel L., and Jeffrey B. Wenger. 2016. "The Political Structure of Policy Diffusion." *Policy Studies Journal* 44 (3): 349–65.

Fink, Simon. 2011. "A Contagious Concept: Explaining the Spread of Privatization in the Telecommunications Sector." *Governance* 24 (1): 111–39.

Franzese, Robert J., 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.

Gilardi, Fabrizio. 2010. "Who Learns from What in Policy Diffusion Processes?" *American Journal of Political Science* 54 (3): 650–66.

———. 2016. "Four Ways We Can Improve Policy Diffusion Research." *State Politics & Policy Quarterly* 16 (1): 8–21.

Gilardi, Fabrizio, and Katharina Fuglister. 2008. "Empirical Modeling of Policy Diffusion in Federal States: The Dyadic Approach." *Swiss Political Science Review* 14 (3): 413–50.

Gilardi, Fabrizio, and Fabio Wasserfallen. 2016. "How Socialization Attenuates Tax Competition." *British Journal of Political Science* 46 (1): 45–65.

Graham, Erin R., Charles R. Shipan, and Craig Volden. 2013. "The Diffusion of Policy Diffusion Research in Political Science." *British Journal of Political Science* 43 (3): 673–701.

Grossback, Lawrence J., Sean Nicholson-Crotty, and David A. Peterson. 2004. "Ideology and Learning in Policy Diffusion." *American Politics Research* 32 (5): 521–45.

Gujarati, Damodar N. 2003. *Basic Econometrics*, 4th ed. New York: The McGraw-Hill Companies.

Hollyer, James R., B. Peter Rosendorff, and James Raymond Vreeland. 2011. "Democracy and Transparency." *The Journal of Politics* 73 (4): 1191–205.

Houle, Christian, Mark A. Kayser, and Jun Xiang. 2016. "Diffusion or Confusion? Clustered Shocks and the Conditional Diffusion of Democracy." *International Organization* 70 (4): 687–726.

Jacob, Suraj, John A. Scherpereel, and Melinda Adams. 2014. "Gender Norms and Women's Political Representation: A Global Analysis of Cabinets, 1979–2009." *Governance* 27 (2): 321–45.

Jordana, Jacint, David Levi-Faur, and Xavier Fernández i Marín. 2011. "The Global Diffusion of Regulatory Agencies: Channels of Transfer and Stages of Diffusion." *Comparative Political Studies* 44 (10): 1343–69.

Karch, Andrew, Sean C. Nicholson-Crotty, Neal D. Woods, and Ann O'M. Bowman. 2016. "Policy Diffusion and the Pro-Innovation Bias." *Political Research Quarterly* 69 (1): 83–95.

Keele, Luke, and Nathan J. Kelly. 2006. "Dynamic Models for Dynamic Theories: The Ins and Outs of Lagged Dependent Variables." *Political Analysis* 14 (2): 186–205.

Krieckhaus, Jonathan. 2006. "Democracy and Economic Growth: How Regional Context Influences Regime Effects." *British Journal of Political Science* 36 (2): 317–40.

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

Linos, Katerina. 2011. "Diffusion Through Democracy." *American Journal of Political Science* 55 (3): 678–95.

Lipsmeyer, Christine S., and Ling Zhu. 2011. "Immigration, Globalization, and Unemployment Benefits in Developed EU States." *American Journal of Political Science* 55 (3): 647–64.

Lopez-Cariboni, Santiago, and Xun Cao. 2015. "Import Competition and Policy Diffusion." *Politics & Society* 43 (4): 471–502.

Maggetti, Martino, and Fabrizio Gilardi. 2016. "Problems (and Solutions) in the Measurement of Policy Diffusion Mechanisms." *Journal of Public Policy* 36 (1): 87–107.

Makse, Todd, and Craig Volden. 2011. "The Role of Policy Attributes in the Diffusion of Innovations." *The Journal of Politics* 73 (1): 108–24.

Mallinson, Daniel J. 2019. "Who Are Your Neighbors? The Role of Ideology and Decline of Geographic Proximity in the Diffusion of Policy Innovations." *Policy Studies Journal*. https://doi.org/10.1111/psj.12351

Miller, Michael K. 2016. "Democracy by Example? Why Democracy Spreads When the Worlds Democracies Prosper." *Comparative Politics* 49 (1): 83–116.

Mitchell, Joshua L. 2018. "Does Policy Diffusion Need Space? Spatializing the Dynamics of Policy Diffusion." *Policy Studies Journal* 46 (2): 424–51.

Neumayer, Eric, and Thomas Plümper. 2010. "Making Spatial Analysis Operational: Commands for Generating Spatial Effect Variables in Monadic and Dyadic Data." *Stata Journal* 10 (4): 585–605.

———. 2016. "W." *Political Science Research and Methods* 4 (1): 175–93.

Plümper, Thomas, Vera E. Troeger, and Hannes Winner. 2009. "Why Is There No Race to the Bottom in Capital Taxation?" *International Studies Quarterly* 53 (3): 761–86.

R Core Team. 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Rogers, Everett M. 2004. *Diffusion of Innovations*. New York: The Free Press.

Shipan, Charles R., and Craig Volden. 2006. "Bottom-Up Federalism: The Diffusion of Antismoking Policies from U.S. Cities to States." *American Journal of Political Science* 50 (4): 825–43.

———. 2008. "The Mechanisms of Policy Diffusion." *American Journal of Political Science* 52 (4): 840–57.

———. 2012. "Policy Diffusion: Seven Lessons for Scholars and Practitioners." *Public Administration Review* 72 (6): 788–96.

Siverson, Randolph M., and Harvey Starr. 1990. "Opportunity, Willingness, and the Diffusion of War." *American Political Science Review* 84 (1): 47–67.

Sugiyama, Natasha Borges. 2008. "Theories of Policy Diffusion: Social Sector Reform in Brazil." *Comparative Political Studies* 41 (2): 193–216.

Sylvester, Steven M., and Donald P. Haider-Markel. 2016. "Buzz Kill: State Adoption of DUI Interlock Laws, 2005–11." *Policy Studies Journal* 44 (4): 491–509.

Volden, Craig. 2006. "States as Policy Laboratories: Emulating Success in the Children's Health Insurance Program." *American Journal of Political Science* 50 (2): 294–312.

Walker, Jack L. 1969. "The Diffusion of Innovations Among the American States." *American Political Science Review* 63 (3): 880–99.

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

Weyland, Kurt. 2005. "Theories of Policy Diffusion Lessons from Latin American Pension Reform." *World Politics* 57 (2): 262–95.

Whitaker, Eric A., Mitchel N. Herian, Christopher W. Larimer, and Michael Lang. 2012. "The Determinants of Policy Introduction and Bill Adoption: Examining Minimum Wage Increases in the American States, 1997–2006." *Policy Studies Journal* 40 (4): 626–49.

Whitten, Guy D., Laron K. Williams, and Cameron Wimpy. Forthcoming. "Interpretation: The Final Spatial Frontier." *Political Science Research and Methods*.

Williams, Laron K. 2015. "It's All Relative: Spatial Positioning of Parties and Ideological Shifts." *European Journal of Political Research* 54 (1): 141–59.

———. 2016. "Long-Term Effects in Models with Temporal Dependence." *Political Analysis* 24 (2): 243–62.

Williams, Laron K., and Guy D. Whitten. 2012. "But Wait, There's More! Maximizing Substantive Inferences from TSCS Models." *The Journal of Politics* 74 (3): 685–93.

Wimpy, Cameron, Laron K. Williams, and Guy D. Whitten. Forthcoming. "X Marks the Spot: Unlocking the Treasure of Spatial-X Models." *The Journal of Politics*.

Supporting Information

Additional supporting information may be found in the online version of this article at the publisher's web site: