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

Time lag spatial lag models

In our paper we describe a model with a temporally lagged spatial lag model as being a variant of an SLX model rather than a SAR model. This is because such a model does not have one of the defining characteristics of SAR models, endogenous regressors. As Lesage and Pace (2009: 192) write, such a model "relies on past period dependent variables and contains no simultaneous spatial interaction." In this section we provide some more details about models with temporally lagged spatial lag model terms and how the results from them are interpreted. To do this, we follow a set of notational and presentational conventions used in Chapter 4 of Elhorst (2014). In that chapter, Elhorst provides a taxonomy of a set of models that are dynamic in space and time.

Elhorst breaks the effects estimated by such models into four different categories: short-term direct effects, short-term indirect effects, long-term direct effects, and long-term indirect effects. The distinction between short-term and long-term effects is a common feature of time series models while, as we discuss in our paper, the distinction between direct and indirect effects is a common feature of spatial models.

For our purposes, we will consider a set of models beginning with one that is not dynamic in either time or space and then a select set of models that are dynamic in only one dimension before discussing the temporally lagged spatial lag model which is dynamic in both dimensions. Across all of these models we will assume a common panel data structure with observations that vary across units and over time. If we follow Elhorst's notational convention of using only temporal subscripts, a simple regression model without any spatial or temporal dynamics would be written as

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

which we call a "non-spatial static model."[1] Because this model has no temporal dynamics, all effects estimated from it will be short-term and because this model has no spatial dynamics, all effects from it will be direct. Thus for a one unit increase in a particular independent variable, $\mathbf{x}_k$, the effect will simply be an immediate (short-term direct) increase of β_k.

[1]Following this notational convention is convenient because it allows the temporal dimension for each term to be indentified by the subscripts and the spatial dimension of each term to be identified by the presence or absence of the connectivity matrix ($\mathbf{W}$). $\mathbf{y}_t$ and $\mathbf{X}_t$ contain the N observations for each unit at time t so that $\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\epsilon}_t$ expands into

If we add temporal dynamics to Equation 1 in the form of a lagged dependent variable, our model becomes

$$\mathbf{y}_t = \mathbf{y}_{t-1}\phi + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\epsilon}_t \tag{2}$$

which we will call a "non-spatial dynamic model." For a one unit increase in a particular independent variable, $\mathbf{x}_k$, there will now be both a short-term effect of β_k and a long-term effect of $\beta_k(1-\phi)^{-1}$. From a spatial perspective, both of these effects are direct because they are caused by changes in the value of the independent variable in one unit on the value of the dependent variable for that same unit.

If we add spatial dynamics to Equation 1 in the form of a spatially lagged independent variable, our model becomes

$$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}\mathbf{Z}_t\boldsymbol{\theta} + \boldsymbol{\epsilon}_t \tag{3}$$

which we will call a "temporally static SLX" model.[2] As the name implies, the effects from such a model will all be short-term. For a one unit increase in a particular independent variable, $\mathbf{x}_k$, there will be a short-term direct effect of β_k. This model, of course, also has indirect effects are come from the $\mathbf{W}\mathbf{Z}_t\boldsymbol{\theta}$. Thus, for instance, the effects of a global increase in a spatially-specified independent variable, $\mathbf{z}_k$, would be $\mathbf{W}\theta_k$.

If we add spatial dynamics to Equation 1 in the form of a spatially lagged dependent variable, our model becomes

$$\mathbf{y}_t = \rho\mathbf{W}\mathbf{y}_t + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\epsilon}_t \tag{4}$$

which we will call a "temporally static SAR" model. As with the temporally static SLX, the effects from such a model will all be short-term. The direct effect from a SAR model is $[(\mathbf{I}_N - \rho\mathbf{W})^{-1}\boldsymbol{\beta}_k\mathbf{I}_N]^{\bar{d}}$, where following Elhorst's notation, $\bar{d}$ is a calculation of the mean diagonal element of a matrix.[3] And the indirect effects from such a model are $[(\mathbf{I}_N - \rho\mathbf{W})^{-1}\beta_k\mathbf{I}_N]^{\overline{rsum}}$, where $\overline{rsum}$ is a calculation of the mean row sum of the non diagonal elements of a matrix.

Turning to the model of interest, we write the temporally lagged spatial lag model as

$$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}\mathbf{y}_{t-1}\theta + \boldsymbol{\epsilon}_t \tag{5}$$

using θ instead of ρ to emphasize that this model is essentially an SLX model. But, because this model is both temporally and spatially dynamic, it will have a combination of both short-term and long-term

---

$\begin{bmatrix}\mathbf{y}_1\\ \mathbf{y}_2\\ \vdots\\ \mathbf{y}_T\end{bmatrix} = \begin{bmatrix}\mathbf{X}_1\\ \mathbf{X}_2\\ \vdots\\ \mathbf{X}_T\end{bmatrix} + \begin{bmatrix}\boldsymbol{\epsilon}_1\\ \boldsymbol{\epsilon}_2\\ \vdots\\ \boldsymbol{\epsilon}_T\end{bmatrix}$ and each of the vectors and matrices in this expansion expands further into N items. E.g., each of the $\mathbf{y}_t = \begin{bmatrix}y_{1t}\\ y_{2t}\\ \vdots\\ y_{Nt}\end{bmatrix}$.

[2] As discussed in the paper, we specify the independent variables associated with spatial-X effects as $\mathbf{Z}$ to emphasize the point that the variables in $\mathbf{X}$ and $\mathbf{Z}$ do not have to be the same. Elhorst, and many others, specify both matrices identically as $\mathbf{X}$ or, as in Elhorst's Chapter 4 discussion of different combinations of models in time and space, $\mathbf{X}_t$.

[3] The components of the short-term direct effects for the SAR model are often separated into the pre-spatial direct effect, β_k, and then the spatially-filtered direct effect, $[(\mathbf{I}_N - \rho\mathbf{W})^{-1}\beta_k\mathbf{I}_N]^{\bar{d}}$.

effects as well as direct and indirect effects. In essence, the θ term in Equation 5 plays the role of an individual θ_k term inside $\boldsymbol{\theta}$ in Equation 3 with the addititonal complication that the temporal impact of this term works in a fashion along the lines of the ϕ term in Equation 2. Thus, the temporally lagged spatial lag model has unfiltered short-term direct effects, β_k, just like those of the non-spatial dynamic model and the temporally static SLX, but it has no short-term indirect effects. This is the case *precisely* because the spatially lagged component of the model, $\mathbf{W}\mathbf{y}_{t-1}\theta$, is also temporally lagged. To better understand how this term works, we can write Equation 5 back one time period as

$$\mathbf{y}_{t-1} = \mathbf{X}_{t-1}\boldsymbol{\beta} + \mathbf{W}\mathbf{y}_{t-2}\theta + \boldsymbol{\epsilon}_{t-1} \tag{6}$$

and then substitute the right-hand side of Equation 6 into Equation 5,

$$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}(\mathbf{X}_{t-1}\boldsymbol{\beta} + \mathbf{W}\mathbf{y}_{t-2}\theta + \boldsymbol{\epsilon}_{t-1})\theta + \boldsymbol{\epsilon}_t \tag{7}$$

which, if the data being modeled are temporally stationary, meaning $|\theta| < 1$, will lead to decreasing effects as we move more temporally distant from any change in lagged values of $\mathbf{y}_t$, $\mathbf{X}_t$, or $\boldsymbol{\epsilon}_t$. These longterm effects consist of own unit, or direct, effects of $[(\mathbf{I}_N - \theta\mathbf{W})^{-1}\beta_{1k}\mathbf{I}_N]^{\bar{d}}$ and indirect effects of $[(\mathbf{I}_N - \theta\mathbf{W})^{-1}\beta_{1k}\mathbf{I}_N]^{\overline{rsum}}$.

In Table 1 we provide the specifications of all 5 of the models that we discuss in this section and in Table 2 we provide a listing of the four different effects from each model. If we look across the entries for each model in Table 2, we can see that the temporally lagged spatial lag model is very different from the temporally static SAR. It does not have the defining characteristic of short-term endogenous effects. Instead, what it has is a combination of long-term direct effects and long-term indirect effects which combine elements of the non-spatial dynamic model and the temporally static SLX. And as we note in the paper, the SAR spatial multiplier matrix, $(\mathbf{I}_N - \rho\mathbf{W})^{-1}$ incorporates immediate feedback through terms like $\mathbf{W}^2$ and $\mathbf{W}^3$. In contrast, the temporal decay multiplier of the temporally lagged spatial lag model, $(\mathbf{I}_N - \theta\mathbf{W})^{-1}$, is simply a spatially weighted geometric lag (aka a Koyck lag) function which, as long as stationarity conditions are met, means that temporally more distant changes have smaller effects.

Table 1: Different specifications of temporal and spatially lagged models

Model Name	Specification
Non-spatial static model	$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\epsilon}_t$
Non-spatial dynamic model	$\mathbf{y}_t = \mathbf{y}_{t-1}\phi + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\epsilon}_t$
Temporally static SLX	$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}\mathbf{Z}_t\boldsymbol{\theta} + \boldsymbol{\epsilon}_t$
Temporally static SAR	$\mathbf{y}_t = \rho\mathbf{W}\mathbf{y}_t + \mathbf{X}_t\boldsymbol{\beta} + \boldsymbol{\epsilon}_t$
Temporally lagged spatial lag model	$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \mathbf{W}\mathbf{y}_{t-1}\theta + \boldsymbol{\epsilon}_t$

Notes:

Table 2: Effects from different specifications of temporal and spatially lagged models

Model Name	Short-term direct effects	Short-term indirect effects	Long-term direct effects	Long-term indirect effects
Non-spatial static model	β_k	none	none	none
Non-spatial dynamic model	β_k	none	$\beta_k(1-\phi)^{-1}$	none
Temporally static SLX	β_k	$\mathbf{W}\theta_k$	none	none
Temporally static SAR	$[(\mathbf{I}_N - \rho\mathbf{W})^{-1}\beta_k\mathbf{I}_N]^{\bar{d}}$	$[(\mathbf{I}_N - \rho\mathbf{W})^{-1}\beta_k\mathbf{I}_N]^{\overline{rsum}}$	none	none
Temporally lagged spatial lag model	β_k	none	$[(\mathbf{I}_N - \theta\mathbf{W})^{-1}\beta_k\mathbf{I}_N]^{\bar{d}}$	$[(\mathbf{I}_N - \theta\mathbf{W})^{-1}\beta_k\mathbf{I}_N]^{\overline{rsum}}$

Notes: In this table we follow some useful notational conventions used by Elhorst (2014) where $\bar{d}$ is a calculation of the mean diagonal element of a matrix and $\overline{rsum}$ is a calculation of the mean row sum of the non-diagonal elements of a matrix.

Additional Experiments

SAR Model Performance for SAR DGP

In this section we detail how the "correct" models perform for SAR and SLX data-generating processes, respectively. In both cases, the recovery rates of the coefficients are, as expected, close to 95%. We provide these results in Tables 3 and 4.

Table 3: Recovery Rates for the SAR Model Specification: SAR Data-Generating Process

	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
SAR model: $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \rho\mathbf{W}\mathbf{y} + \boldsymbol{\epsilon}$									
β	94.2	93.5	95.1	95.8	93.6	93.9	94.8	95.0	94.6
ρ	95.0	94.6	94.7	95.0	95.4	94.4	93.3	94.0	94.4

Second-Order Neighbor Model Performance for SAR DGP

One possibility that we explored in the paper addressed the particular functional form of the higher-order weights matrices. A specification mirroring the SAR by squaring $\mathbf{W}$ (seen in Equation 9 in the manuscript) produces non-zero values along the diagonal of the partial derivatives matrix, which means that there are feedback effects, and higher-order effects more generally. If deemed unnecessary by theory, the functional form of the weights matrix can be modified so that it expressly prohibit feedback effects. This specification would identify higher-orders of contiguity (see the second-order weights matrix, $\mathbf{W}_{\mathbf{2nd}}$ in Figure 1). The third set of experiments evaluates how well a model specified as

$$\mathbf{y} = \mathbf{x}\beta + \mathbf{W}\mathbf{x}\theta_1 + \mathbf{W}_{\mathbf{2nd}}\mathbf{x}\theta_2 + \boldsymbol{\epsilon} \tag{8}$$

deals with an SAR DGP where feedback effects are present. As we can see from the third section in Table 5, which is the same as Table 2 in the manuscript with this additional set of results, this change in the specification from $\mathbf{W}^2$ to $\mathbf{W}_{\mathbf{2nd}}$ results in a serious reduction in the recovery rates of the zero-order direct effect, which is now larger, on average, because it has to account for the spatial effects that would otherwise be modeled as feedback effects. Furthermore, the recovery rates for the second-order direct effects are, by construction, 0%.[4] We would only advocate this type of model specification if two conditions are met: first, the theory is quite clear about the impossibility of feedback effects, and second, these feedback effects are shown to be zero in robustness checks.

[4]Indeed, if we include second- and third-order contiguity weights matrices, we are able to model some of the higher order effects but still none of the feedback effects.

Table 4: Recovery Rates for the SLX Model Specification: SLX Data-Generating Process

	$\theta_1 = -0.8$; $\theta_2 = -0.8$	$\theta_1 = -0.8$; $\theta_2 = 0.01$	$\theta_1 = -0.8$; $\theta_2 = 0.4$	$\theta_1 = 0.4$; $\theta_2 = 0.4$	$\theta_1 = 0.4$; $\theta_2 = 0.01$	$\theta_1 = 0.4$; $\theta_2 = -0.8$	$\theta_1 = -0.2$; $\theta_2 = -0.2$	$\theta_1 = -0.2$; $\theta_2 = 0.01$	$\theta_1 = -0.2$; $\theta_2 = 0.4$
SLX model: $\mathbf{y} = \mathbf{x_1}\beta_1 + \mathbf{x_2}\beta_2 + \mathbf{W}\mathbf{x_1}\theta_1 + \mathbf{W}\mathbf{x_2}\theta_2 + \boldsymbol{\epsilon}$									
Direct									
β_1	95.0	96.1	95.5	94.7	94.7	95.0	94.7	95.3	94.4
β_2	93.7	94.9	95.5	95.7	95.0	94.3	95.5	95.0	95.0
Indirect									
θ_1	95.0	95.1	95.4	96.5	94.5	94.6	95.2	93.7	94.1
θ_2	95.3	94.9	94.5	95.0	95.1	94.7	94.6	94.8	95.2

Table 5: Recovery Rates for Various SLX Model Specifications: SAR Data-Generating Process

	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
Simple SLX model: $\mathbf{y} = \mathbf{x}\beta + \mathbf{W}\mathbf{x}\theta + \boldsymbol{\epsilon}$									
Direct: Total	94.9	94.4	94.8	93.7	95.5	93.5	95.0	94.0	88.9
0 Order	77.9	91.9	95.1	93.5	95.5	93.4	94.8	90.7	81.0
2nd Order	–	–	–	–	–	–	–	–	–
3rd Order	–	–	–	–	–	–	–	–	–
Indirect: Total	15.1	49.9	85.2	93.7	95.2	93.5	40.3	0	0
1st Order	91.4	93.8	95.5	93.1	95.2	94.5	94.4	93.8	93.1
2nd Order	–	–	–	–	–	–	–	–	–
3rd Order	–	–	–	–	–	–	–	–	–
SLX model with a squared term: $\mathbf{y} = \mathbf{x}\beta + \mathbf{W}\mathbf{x}\theta_1 + \mathbf{W}^2\mathbf{x}\theta_2 + \boldsymbol{\epsilon}$									
Direct: Total	95.0	94.0	95.1	93.8	95.4	94.0	95.2	94.4	91.0
0 Order	94.2	95.1	95.2	94.2	93.6	94.8	94.8	94.9	95.8
2nd Order	93.2	93.9	94.7	94.6	94.0	94.7	93.6	93.4	93.8
3rd Order	–	–	–	–	–	–	–	–	–
Indirect: Total	93.0	94.7	94.8	94.7	93.9	94.5	93.3	83.7	2.6
1st Order	93.0	94.9	95.9	94.0	94.9	94.9	94.4	93.2	92.6
2nd Order	93.2	93.9	94.7	94.6	94.0	94.7	93.6	93.4	93.8
3rd Order	–	–	–	–	–	–	–	–	–
SLX model with a second-order term: $\mathbf{y} = \mathbf{x}\beta + \mathbf{W}\mathbf{x}\theta_1 + \mathbf{W}_{\mathbf{2nd}}\mathbf{x}\theta_2 + \boldsymbol{\epsilon}$									
Direct: Total	95.5	93.6	95.0	93.9	96.3	93.3	95.7	93.8	91.0
0 Order	76.1	89.9	95.1	93.7	96.3	93.8	95.3	89.5	77.9
2nd Order	0	0	0	0	0	0	0	0	0
3rd Order	–	–	–	–	–	–	–	–	–
Indirect: Total	95.3	94.0	95.1	95.8	95.0	95.4	94.5	90.6	36.0
1st Order	96.5	95.4	95.9	94.2	95.2	95.6	95.0	91.7	88.4
2nd Order	95.0	94.5	94.7	95.8	94.5	95.3	94.1	94.6	93.6
3rd Order	–	–	–	–	–	–	–	–	–
SLX model with squared and cubed terms: $\mathbf{y} = \mathbf{x}\beta + \mathbf{W}\mathbf{x}\theta_1 + \mathbf{W}^2\mathbf{x}\theta_2 + \mathbf{W}^3\mathbf{x}\theta_3 + \boldsymbol{\epsilon}$									
Direct: Total	95.0	94.2	94.6	94.3	96.1	94.4	94.5	93.7	93.4
0 Order	94.2	95.1	95.3	94.1	93.9	94.9	94.3	94.9	95.9
2nd Order	93.9	94.0	94.8	95.2	94.2	95.2	93.8	93.2	93.3
3rd Order	94.2	93.6	94.7	93.5	95.0	94.4	95.0	93.9	93.4
Indirect: Total	95.5	95.1	95.6	94.4	95.2	94.5	93.9	92.6	87.0
1st Order	94.3	94.7	95.1	93.9	94.9	93.7	93.9	95.2	94.4
2nd Order	93.9	94.0	94.8	95.2	94.2	95.2	93.8	93.2	93.3
3rd Order	94.2	93.6	94.7	93.5	95.0	94.4	95.0	93.9	93.4

Spatial Error Model (SEM)

A popular alternative to the SAR and SLX models is the spatial error model (SEM).[5] Instead of the spatial dependence arising in the outcomes (as in y_i influencing y_j, and vice versa) or in the observables (as in x_i influencing y_j), the SEM models spatial dependence in the unobservables, or errors:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \lambda \mathbf{W} \mu + \boldsymbol{\epsilon} \tag{9}$$

where the overall error is decomposed into ϵ, "a spatially uncorrelated error term that satisfied the normal regression assumption, and [μ], which is a term indicating the spatial component of the error term" (Ward and Gleditsch 2008: 65-66). If λ is 0, then there is no spatial dependence in the errors and an OLS can be safely estimated; if λ is not 0, then "we have a pattern of spatial dependence between the errors for connected observations" (Ward and Gleditsch 2008: 66). This poses no complications for generating quantities of interest, as "the differences in the independent variables in i do not have effects on outcomes in observations connected to i" (Ward and Gleditsch 2008: 67). Essentially, this means that variables will only have a direct effect (i.e., x_i on y_i) and no indirect effects (such as x_i on y_j), feedback or otherwise.[6]

We echo Beck, Gleditsch and Beardsley's (2006: 30) conclusion that the SEM is not appropriate in most political science applications. This is because a variable can have an impact on neighboring observations if omitted (and thus part of the error term), but not if it is included. Consider the example of economic growth in an SEM model:

> [...] remember that the "errors" are just the variables that we either chose not to measure, or could not measure. In particular, they are errors from the perspective of the analyst, not the perspective of policy makers in the country. Thus, if Germany grew more quickly because of some variable not included in the specification, that growth would affect all other countries. But if Germany grew more quickly because it had a left government, and if that variable were included in the specification, then this extra German growth would have no impact on the growth in other countries (30).

It remains to be seen how SLX and SAR models perform when the true data-generating process features spatial dependence in the unobservables, which is typically consistent with an SEM model. In other words, if one is uncertain about the true spatial process at work and diagnostics are uncertain, how dangerous is it to begin with a SLX or SAR model if there is spatial dependence in the unobservables? LeSage and Pace (2009: 157-159) point out that the result is "unbiased but inefficient coefficient estimates" and "inference regarding dispersion of the explanatory variables based on the asymptotic variance-covariance matrix for the SAR model will be misleading, since error dependence is ignored when constructing the variance-covariance matrix".

To explore this possibility we simulate data based on the following equation (Darmofal 2015: 102):

[5] The description of the SEM draws heavily from Ward and Gleditsch (2008: 65-67).

[6] Darmofal (2015: 107-108) succinctly states that "because the spatial multiplier in a spatial error model pertains only to the errors, substantive covariates do not vary in their equilibrium effects based on the spatial locations of the observations in a spatial error formulation."

$$\mathbf{y} = \mathbf{x}\beta + \boldsymbol{\epsilon}, \quad \epsilon = \lambda \mathbf{W} \epsilon + \xi \tag{10}$$

with matrix $\mathbf{X}$ containing a single variable drawn from a uniform distribution, $x \in [-10, 10]$, and where $\beta = 1$. $\mathbf{W}$ is an $N \times N$ symmetric row-standardized contiguity weights matrix, where each element below the diagonal is randomly drawn from a Bernoulli distribution. We simulated 1000 data sets at each of nine different scenarios defined by the strength of the spatial error dependence, $\lambda \in \{-0.8, 0.8\}$.[7] As with the experiments presented in the manuscript, we focus on the recovery rates for direct and indirect effects of x since those reflect both the coefficients and their uncertainty. Table 5 shows the results of the experiments.

Two clear patterns emerge from Table 6. The first pattern is that the SAR model does a poor job of recovering the first-order indirect effect (which is actually 0) in the SEM DGP for models of λ that are lower than lower than 0 or larger than 0.2. In those case, the SAR model finds false evidence that x_i influences y_j through the outcomes (i.e., by influencing y_i). Since the SAR model finds "phantom" first-order indirect effects that do not actually exist, it also does a poor job in recovering the true average total effects (not shown). The second pattern is that the various specifications of the SLX model—ranging from a simple model with only first-order indirect effects to one with first- through third-order effects—recovers estimates of the direct and indirect effects that are consistently close to 95% for all values of λ. The SLX is more flexible in this respect because the θs are effectively 0, which rules out higher-order effects. For these two reasons, it appears as though the SLX model is more robust to incorrectly specifying the spatial dependence when it actually occurs in the unobservables. If one is only concerned about generating meaningful inferences for the explanatory variables, then this type of misspecification is not problematic; of course, if one is interested in modeling the actual spatial process, then the SLX would be unable to show that the spatial dependence actually exists in the unobservables.

Application

In this section we explore substantive effects from the defense spending application in the manuscript, as well as some models to show how deftly SLX handles conditional spatial dependence.

As shown in the SAR model of Table 7, both civil and interstate wars influence defense burdens. Of course, the coefficients themselves are only the estimated direct effects of those covariates on defense burden. To understand the total effect of the covariates, it is important to utilize the partial derivatives approach. In the manuscript we showed that the total impact of the covariates on the defense burden depend on the coefficient (β), the size of the change in x (in the case of civil and interstate wars, the difference between a value of 0 and 1), the global spatial autocorrelation coefficient (ρ), and each state's distribution of neighbors ($\mathbf{W}$). Each state potentially has a different configuration of neighbors, which leads to different indirect effects for each.

To simplify matters, we examine the average direct, indirect and total effects in Table 8. From this table we can see that the estimated average total effect of a civil war at time $t-1$ in a neighboring state on the defense burden of the focal state at time t is a reduction in military spending as a percentage of GDP of -0.62%. The effect of a civil war in a state results in a reduction in that state's defense burden,

[7] It is worth noting that values of λ close to the absolute value of 1 are exceedingly rare in practice.

Table 6: Recovery Rates for SAR and Various SLX Model Specifications: SEM Data-Generating Process

	-0.8	-0.6	-0.4	-0.2	0	0.2	0.4	0.6	0.8
SAR model: $\mathbf{y} = \mathbf{x}\beta + \rho\mathbf{W}\mathbf{y} + \boldsymbol{\epsilon}$									
Direct: Total	94.6	94.3	94.5	93.4	95.3	93.8	95.1	94.7	93.4
0 Order	94.7	94.4	94.3	93.8	95.6	93.6	95.3	94.7	93.5
2nd Order	98.0	99.4	99.8	99.8	100	99.9	99.9	99.6	99.0
3rd Order	100	100	100	100	100	100	100	100	100
Indirect: Total	59.9	72.1	83.8	90.1	93.9	97.7	97.3	96.2	93.1
1st Order	66.9	77.6	87.3	91.2	94.0	94.3	90.0	82.6	73.1
2nd Order	98.0	99.4	99.8	99.8	100	99.9	99.9	99.6	99
3rd Order	100	100	100	100	100	100	100	100	100
Simple SLX model: $\mathbf{y} = \mathbf{x}\beta + \mathbf{W}\mathbf{x}\theta + \boldsymbol{\epsilon}$									
Direct: Total	94.9	94.8	94.9	93.8	95.5	93.5	95.3	94.9	93.9
0 Order	94.9	94.8	94.9	93.8	95.5	93.5	95.3	94.9	93.9
2nd Order	–	–	–	–	–	–	–	–	–
3rd Order	–	–	–	–	–	–	–	–	–
Indirect: Total	94.0	95.5	95.8	93.1	95.2	94.5	94.8	94.3	95.6
1st Order	94.0	95.5	95.8	93.1	95.2	94.5	94.8	94.3	95.6
2nd Order	–	–	–	–	–	–	–	–	–
3rd Order	–	–	–	–	–	–	–	–	–
SLX model with a squared term: $\mathbf{y} = \mathbf{x}\beta + \mathbf{W}\mathbf{x}\theta_1 + \mathbf{W}^2\mathbf{x}\theta_2 + \boldsymbol{\epsilon}$									
Direct: Total	95.0	94.1	95.0	93.8	95.4	94.0	95.2	95.1	94.0
0 Order	94.2	95.1	95.2	94.2	93.6	94.7	94.7	95.0	95.3
2nd Order	94.4	94.0	94.9	94.6	94.0	94.7	93.8	93.9	94.8
3rd Order	–	–	–	–	–	–	–	–	–
Indirect: Total	96.2	95.1	95.0	94.7	93.9	94.4	93.6	92.2	93.3
1st Order	94.8	95.5	96.0	94.0	94.9	94.9	94.8	93.9	95.6
2nd Order	94.4	94.0	94.9	94.6	94.0	94.7	93.8	93.9	94.8
3rd Order	–	–	–	–	–	–	–	–	–
SLX model with squared and cubed terms: $\mathbf{y} = \mathbf{x}\beta + \mathbf{W}\mathbf{x}\theta_1 + \mathbf{W}^2\mathbf{x}\theta_2 + \mathbf{W}^3\mathbf{x}\theta_3 + \boldsymbol{\epsilon}$									
Direct: Total	94.8	94.2	94.6	94.3	96.1	94.5	94.8	94.5	95.2
0 Order	94.2	95.0	95.3	94.1	93.9	94.8	94.3	94.9	95.4
2nd Order	94.8	94.0	95.0	95.2	94.2	95.3	93.9	93.3	94.3
3rd Order	94.6	93.8	94.8	93.6	95.0	94.6	95.3	94.3	94.4
Indirect: Total	95.8	95.5	95.8	94.5	95.2	94.5	94.0	93.2	93.4
1st Order	94.7	94.8	95.1	93.9	94.9	93.7	93.9	95.2	94.3
2nd Order	94.8	94.0	95.0	95.2	94.2	95.3	93.9	93.3	94.3
3rd Order	94.6	93.8	94.8	93.6	95.0	94.6	95.3	94.3	94.4

Table 7: Non-Spatial OLS, SAR and SLX Models of Neighborhood Effects on Defense Burdens

	OLS Model 1	**SAR Model 2**	**SLX Model 3**
Spatial Estimates (ρ and θ)			
ρ		0.07***	
		(0.004)	
Contiguity×Civil War$_{t-1}$			-0.11*
			(0.07)
Contiguity×Interstate War$_{t-1}$			0.18***
			(0.07)
Ally×Interstate War$_{t-1}$			-0.006
			(0.03)
Contiguity×Defense Burden$_{t-1}$			0.006***
			(0.002)
Defense Pact×Defense Burden$_{t-1}$			0.003***
			(0.001)
Non-Spatial Estimates (β)			
Civil War$_{t-1}$	-0.47***	-0.45***	-0.39***
	(0.17)	(0.16)	(0.15)
Interstate War$_{t-1}$	0.46***	0.45***	0.37**
	(0.16)	(0.16)	(0.16)
Total Population (Logged)$_{t-1}$	0.03	0.03	0.02
	(0.02)	(0.02)	(0.02)
Alliance with US	-0.21**	-0.22**	-0.15*
	(0.10)	(0.09)	(0.08)
ΔUS Defense Burden	-0.04	-0.03	-0.05
	(0.07)	(0.07)	(0.08)
ΔUSSR/Russia Defense Burden	0.006	-0.01	-0.05***
	(0.02)	(0.02)	(0.02)
US Ally×Δ US Defense Burden	0.14	0.11	0.21
	(0.12)	(0.11)	(0.13)
US Ally×Δ USSR Defense Burden	0.005	0.01	0.02
	(0.03)	(0.03)	(0.02)
Annual Trend	-0.007***	-0.006***	-0.004**
	(0.002)	(0.002)	(0.002)
1992	1.13***	0.74***	-1.26***
	(0.28)	(0.27)	(0.21)
Defense Burden$_{t-1}$	-0.12***	-0.12***	-0.14***
	(0.006)	(0.01)	(0.006)
ΔDefense Burden$_{t-1}$	-0.14***	-0.13***	-0.14***
	(0.01)	(0.02)	(0.01)
Constant	0.46**	0.41*	0.22
	(0.22)	(0.21)	(0.17)
N	6,328	6,328	7,266

Note: Models include regional fixed effects. The SAR model excludes isolates.
* p-value < 0.1; ** p-value < 0.05; *** p-value < 0.01

on average, of -0.46%. Note that this effect is slightly larger than the coefficient for *civil war*$_{t-1}$; the difference is the result of feedback effects, or the effects of civil war in state i influencing its neighbor j, which feeds back to affect state i. The average indirect effect—or, the average effect of a civil war in the focal state on other states is -0.17%. These effects demonstrate that civil wars can meaningfully impact states' defense burdens, and over a quarter of the overall effect spills over into neighboring states. The effects of *interstate war*$_{t-1}$ are similar in magnitude and distribution between direct and indirect effects, except for being positive. On average, experiencing an interstate war at time $t-1$ increases that state's defense burden by 0.46%, and spills over to increase the defense burdens of neighboring states by 0.17%.

As we demonstrated with our Monte Carlo experiments, the consequences of model choices can range from understating the overall effects, to making the opposite inference regarding indirect effects. Recall that in the case of civil and interstate wars, the SLX variables were consistent with the pattern of positive spatial dependence in the SAR model (consistent with the positive ρ). When we compare the various effects of the SLX to the SAR model for both of these variables (see Table 2), we see that the SLX produces average total and indirect effects that are much larger (almost two and three times larger in the case of *interstate war*$_{t-1}$), and smaller direct effects. Since the SLX model does not force the spatial autocorrelation to be represented by one parameter, the indirect effects are free to vary in size based on the particular covariate. In the case of *defense burden*$_{t-1}$, the average total effects are much smaller in the SLX model because the coefficients for the two SLX variables are signed in the opposite direction from the *defense burden*$_{t-1}$ variable. In the SAR model, contrary to expectations, increasing one's defense burden by 1% is estimated to decrease contiguous neighbors' defense burdens by -0.04%; in the SLX model, this same change is estimated to increase contiguous neighbors' defense burdens by 0.02%. The latter estimation technique is more flexible and provides more realistic inferences.

Another advantage in the flexibility of the SLX is the ability to properly model conditional patterns of spatial dependence. In the first SLX model (Model 2 in Table 7), we demonstrated that the spatial effects of *defense burden*$_{t-1}$ were conditioned by patterns of neighbors via contiguity and defense pact. To explore this further, let us examine how the flexibility of SLX models allows us to easily estimate region-specific SLX variables.[8] Due to security agreements, colonial histories, regional organizations, and other characteristics (see **?**, 429–430 for a summary), covariates that might spillover in one region are contained in another. In Models 4 and 5 (Table 9) we estimate separate region-specific parameters for both variables (*civil war*$_{t-1}$ and *interstate war*$_{t-1}$). By doing so, we can show how the effects of wars depend on the regions in which they occur.

The results in Table 9 (and the effects depicted in Table 2) demonstrate that the non-conditional SLX model (Model 3) clouded a great deal of region-specific heterogeneity in the spatial patterns. The resulting average indirect effect of *civil war*$_{t-1}$ in Model 3 was slightly negative (-0.36). This value represents a rough average of the indirect effects across regions, and obscures the fact that in one of the regions (Africa) civil wars in one's region actually increases states' defense burdens. This is also the case in *interstate war*$_{t-1}$, as both Europe and Asia/Oceania respond in the opposite manner as the other regions to interstate wars in the region. These inferences—while relatively easy to derive in the SLX setting—would be prohibitively difficult, if not impossible, with an SAR model.

One set of circumstances where the SAR model is generally more appropriate than the SLX is in the case of higher-order effects beyond the first-order. If these effects are expected to occur simultaneously,

[8] For this example, we use the Correlates of War's regional classification (see Stinnett et al. 2002 for a description).

Table 8: Average Direct, Indirect and Total Effects of Covariates on Defense Burden

Variables	**W Specification**	**Avg. Effects**	**SAR**	**SLX**			
			M2	M3	M4	M5	M6
Civil War$_{t-1}$	Contiguity	Direct	-0.46	-0.39			
		Indirect	-0.17	-0.36			
		Total	-0.62	-0.74			
Civil War$_{t-1}$	Region (Europe)	Indirect			-4.63		
	Region (Middle East)	Indirect			-6.32		
	Region (Africa)	Indirect			2.04		
	Region (Asia)	Indirect			-1.37		
	Region (Americas)	Indirect			-1.54		
Interstate War$_{t-1}$	Contiguity	Direct	0.46				
		Indirect	0.17				
		Total	0.63				
	Contiguity + Alliance	Direct		0.37			
		Indirect		0.51			
		Total		0.88			
	Contiguity ($\mathbf{W}$)	Indirect					0.51
	Contiguity ($\mathbf{W}^2$)	Direct					0.05
		Indirect					0.19
		Total					0.24
	Contiguity ($\mathbf{W}^3$)	Direct					0.01
		Indirect					0.11
		Total					0.12
	Contiguity ($\sum_{n=1}^{3} \mathbf{W^n}$)	Direct					0.30
		Indirect					0.80
		Total					1.10
Interstate War$_{t-1}$	Region (Europe)	Indirect				-1.43	
	Region (Middle East)	Indirect				2.16	
	Region (Africa)	Indirect				0.21	
	Region (Asia)	Indirect				-0.84	
	Region (Americas)	Indirect				0.44	
Defense Burden$_{t-1}$	Contiguity	Direct	-0.12				
		Indirect	-0.04				
		Total	-0.16				
	Contiguity + Defense	Direct		-0.14			
		Indirect		0.05			
		Total		-0.09			

Note: SAR Model 1 uses a binary, un-row-standardized contiguity weights matrix.

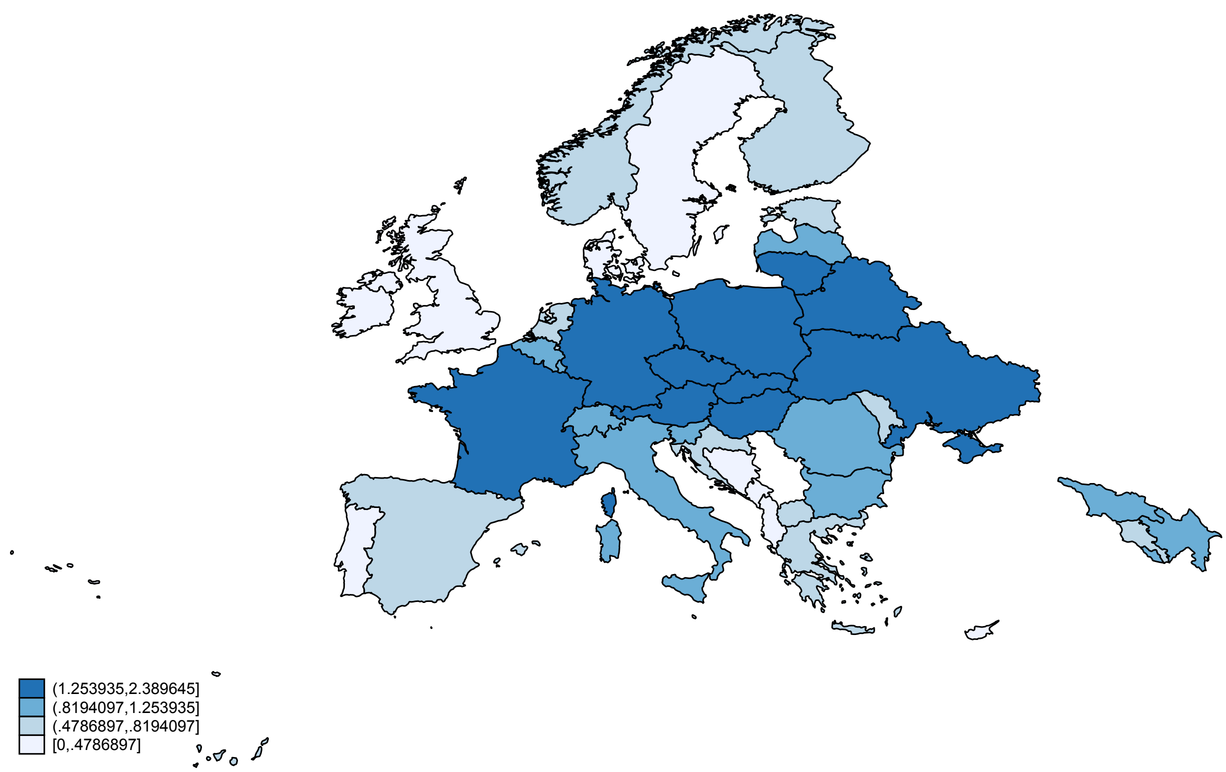
Table 9: Conditional SLX Models of Neighborhood Effects on Defense Burdens

	Model 4	Model 5	Model 6
Spatial Estimates (θ)			
Contiguity$\times$Interstate War$_{t-1}$			0.16
			(0.12)
Contiguity$^2\times$Interstate War$_{t-1}$			0.02
			(0.05)
Contiguity$^3\times$Interstate War$_{t-1}$			0.002
			(0.01)
Region$\times$Civil War$_{t-1}$	-0.16	-0.008	
	(0.10)	(0.03)	
Region (Middle East)$\times$Civil War$_{t-1}$	-0.06		
	(0.13)		
Region (Africa)$\times$Civil War$_{t-1}$	0.23**		
	(0.11)		
Region (Asia)$\times$Civil War$_{t-1}$	0.11		
	(0.12)		
Region (Americas)$\times$Civil War$_{t-1}$	0.21*		
	(0.12)		
Region$\times$Interstate War$_{t-1}$	-0.001	-0.05	
	(0.02)	(0.05)	
Region (Middle East)$\times$Interstate War$_{t-1}$		0.12*	
		(0.07)	
Region (Africa)$\times$Interstate War$_{t-1}$		0.06	
		(0.05)	
Region (Asia)$\times$Interstate War$_{t-1}$		0.02	
		(0.06)	
Region (Americas)$\times$Interstate War$_{t-1}$		0.06	
		(0.10)	
Non-Spatial Estimates (β)			
Civil War$_{t-1}$	-0.41***	-0.39***	-0.42***
	(0.15)	(0.15)	(0.15)
Interstate War$_{t-1}$	0.44***	0.44***	0.24
	(0.15)	(0.15)	(0.21)
Total Population (Logged)$_{t-1}$	0.04**	0.04**	0.03*
	(0.02)	(0.02)	0.02)
Annual Trend	-0.003*	-0.004**	-0.004***
	(0.002)	(0.002)	0.002)
1992	-1.00***	-1.04***	-1.05***
	(0.20)	(0.20)	(0.20)
Defense Burden$_{t-1}$	-0.12***	-0.12***	-0.12***
	(0.006)	(0.006)	(0.006)
ΔDefense Burden$_{t-1}$	-0.15***	-0.15***	-0.15***
	(0.01)	(0.01)	(0.01)
Constant	0.22	0.24	0.23
	(0.17)	(0.17)	(0.17)
N	7,266	7,266	7,266

Note: Models include regional fixed effects.

then it is generally correct to estimate an SAR. If, however, the higher-order effects are based on spatial clustering in the observables, then the SLX can be modified to estimate higher-order effects. In Model 6 (Table 9) we add two higher-order effects to represent the possibility that changes in defense burdens by second- and third-order contiguous neighbors might have an effect on the focal state. Adding the two higher-order effects SLX variables has the benefit of allowing for feedback effects (as shown in the direct effects for second- and third-order contiguity in Table 2), but has the downside of increasing multicollinearity (both higher-order effects have variance inflation factors above 11). In this case, all three SLX variables are positive, and it is clear that the positive effects decline considerably at each additional order of contiguity. F-tests suggest that we cannot reject the null hypothesis that the two higher-order coefficients are jointly equal to 0, which means that the specification with first-order contiguity is sufficient.

Figure 1: Total Effects of $Interstate\ War_{t-1}$ across European States in 2007 (Model 6)



Note: The values represent the total effects—including direct and indirect effects—for each European state (i.e., the total in each row of the partial derivatives matrix), given an interstate war in every European state in 2007. Russia, which is omitted for graphical purposes, has the largest total effects of any European state (2.74).

While the F-tests suggest that the specification with the first-order contiguity is sufficient,[9] it is instructive to graphically explore how the **W** (i.e., the distribution of contiguous neighbors in this model) influences the size of the indirect effects. In Figure 1 we depict the average total effects of $interstate\ war_{t-1}$ for all European countries in 2007 (except for Russia for illustration purposes). There is a great deal of variation in the size of the effects, and that variation is largely consistent with the historical record detailing countries' responses to interstate wars. For example, the countries that have the largest total indirect

[9]F-tests suggest that we cannot reject the null hypothesis that the two higher-order coefficients are jointly equal to 0.

effects are also those with the most contiguous neighbors (Russia, Germany, Poland, Ukraine, and Austria).[10] On the other hand, countries like Malta, Cyprus, United Kingdom and Ireland have the smallest total effects.

[10] This is a function of our choice to not row-standardize the weights matrix.