Advance Access publication May 14, 2015

Political Analysis (2015) 23:400–414
doi:10.1093/pan/mpv012

Accounting for Right Censoring in Interdependent Duration Analysis

Jude C. Hays
University of Pittsburgh, 4600 Wesley W. Posvar Hall, Pittsburgh, PA 15260
e-mail: jch61@pitt.edu (corresponding author)

Emily U. Schilling
University of Iowa, 341 Schaeffer Hall, Iowa City, IA 52242
e-mail: emily-schilling@uiowa.edu

Frederick J. Boehmke
University of Iowa, 341 Schaeffer Hall, Iowa City, IA 52242
e-mail: frederick-boehmke@uiowa.edu

Edited by Jonathan Katz

Duration data are often subject to various forms of censoring that require adaptations of the likelihood function to properly capture the data generating process, but existing spatial duration models do not yet account for these potential issues. Here, we develop a method to estimate spatial-lag duration models when the outcome suffers from right censoring, the most common form of censoring. We adapt Wei and Tanner's (1991) imputation algorithm for censored (non-spatial) regression data to models of spatially interdependent durations. The algorithm treats the unobserved duration outcomes as censored data and iterates between multiple imputation of the incomplete, that is, right censored, values and estimation of the spatial duration model using these imputed values. We explore the performance of an estimator for log-normal durations in the face of varying degrees of right censoring via Monte Carlo and provide empirical examples of its estimation by analyzing spatial dependence in states' entry dates into World War I.

1 Introduction

The use of spatial econometrics to study both geographic and other forms of interdependence in politics has grown over the past two decades as new methods have been developed to analyze a variety of outcomes, including continuous, discrete, and duration outcomes. Despite this growth, there are still a number of methodological issues that have not been addressed sufficiently. Of particular interest in this article is the problem of right censoring in spatial duration models.

We focus on research in which the durations are observed in continuous time and interdependent across the units of analysis. There is a set period during which the units are at risk of experiencing the event. The covariates are fixed over this period, and some of the units do not experience the event. These conditions are not uncommon in political science research.

How do relative military capabilities affect the time it takes states to join expanding conflicts (Melin and Koch 2010)? Do a country's political institutions influence the time to devaluation in the midst of a currency crisis (Leblang 2003)? How do constituency factors determine the time until

Authors' note: We are grateful for funding provided by the University of Iowa Department of Political Science. Comments from Chris Zorn, Justin Grimmer, Matthew Blackwell, and participants at a University of Rochester seminar are gratefully acknowledged. Replication materials for this study are available from the *Political Analysis Dataverse* at http://dx.doi.org/10.7910/DVN/29603. Supplementary materials for this article are available on the *Political Analysis* Web site.

a representative takes a public stand on a pending piece of legislation (Box-Steffensmeier, Arnold, and Zorn 1997)? These are just a few of the many questions that involve interdependent and potentially right-censored durations.

Our goal is to develop a method to estimate continuous-time, spatial-lag duration models when the outcome suffers from right censoring. Spatial-lag models are characterized by their interdependent outcomes. Applied to duration analysis, these models make the time to an event for one unit a function of this duration for others. Right censoring occurs when we partially observe a duration: we know when an observation began and that it lasted until a least a given point in time, but we do not know exactly how long it survived (or would have survived) beyond this point. The combination of right censoring and interdependence significantly complicates duration analysis because the outcome for one unit depends on outcomes for others that are not fully observed and vice versa. In short, the methodological problem is that interdependence creates a single multivariate hazard function for the units in our samples, and when censoring occurs, we need to integrate this function in order to estimate the spatial-lag duration model.

To address this problem, we adapt Wei and Tanner's (1991) imputation algorithm for censored (non-spatial) regression data to models of spatially interdependent durations. The algorithm iterates between imputing the durations for right-censored observations and estimating the spatial-lag duration model given these imputed values. After a brief review of some existing methods for spatial duration analysis, we apply this algorithm to develop an estimator for log-normal durations with right censoring. The assumption of normality facilitates our extension of Wei and Tanner's method by allowing us to sample efficiently from the multivariate distribution of disturbances. We then explore the performance of the estimator in the face of varying degrees of right censoring and spatial interdependence via Monte Carlo followed by an empirical illustration that examines spatial dependence in states' entry dates into World War I (WWI). This offers a good example given the degree of interdependence in entry decisions and a significant amount of right censoring caused by countries that did not join before the conflict ended.

2 Spatial Duration Models

There are two prominent approaches to spatial duration analysis in the literature. The first is the spatial frailty approach originally developed in biostatistics by Banerjee, Carlin, and coauthors. The use of frailties has a long tradition in duration modeling. Frailties help account for the unobserved differences across units that make some more likely to experience the event of interest sooner than others, thereby playing a role similar to random effects. In spatial econometrics, spatially correlated frailties are associated with what is called the spatial error model.[1] Banerjee, Gelfand, and Carlin (2004) present both parametric and semi-parametric frameworks for continuous-time durations with conditionally autoregressive spatial frailties. The parametric variant of the spatial frailty model in Banerjee, Wall, and Carlin (2003) assumes a Weibull distribution for the baseline hazard, but a variety of parametric forms is feasible. Banerjee and Carlin (2003) estimate a semi-parametric version using a mixture of beta functions to flexibly model the integrated baseline hazard. More recently, in their geoadditive survival model, Hennerfeind, Brezger, and Fahrmeir (2006) extend this framework by generalizing the linear predictor and relaxing the proportional hazards assumption of the Cox model.

Importantly, these are not models of interdependence in the sense that duration outcomes for some units depend on those for others. Instead, these frailties typically are interpreted as reflecting omitted or unobservable covariates that cluster spatially. For example, Banerjee, Wall, and Carlin (2003) model infant mortality in Minnesota including covariate information on the child's sex, race, and birth weight as well as the mother's age and total number of previous births. Even with these covariates, the frailties exhibit spatial correlation, which the authors argue is due to omitted variables such as the overall quality of available health care and the mother's economic status among others. Similarly, Hennerfeind, Brezger, and Fahrmeir (2006) state that their geoadditive model is

[1]See LeSage and Pace (2009) for the connection between spatially correlated random effects and spatial error models.

"useful for detecting unobserved covariates which carry spatial information" (1073). Darmofal (2009) analyzes the timing of members of Congress's position taking on the North American Free Trade Agreement and finds evidence of spatial clustering in the frailties. The implication is that scholars "should consider the possibility that members from neighboring states share common unmeasured characteristics that impact the behavior of interest" (251).

The second approach extends discrete-time event history models to incorporate spatial dependence (Allison 1982; Beck, Katz, and Tucker 1998). Berry and Berry (1990), for instance, find that the probability an American state adopts a lottery at time t increases with the number of neighboring states that have adopted a lottery prior to time t. Gasiorowski (1995) shows that democratic breakdowns and transitions are less and more likely, respectively, as the regional proportion of democracies increases. Beardsley (2011) finds that states are more likely to experience armed conflict, either intrastate or interstate conflict, if their neighbors experienced armed conflict in the previous 2 years.

This approach is widely used in political science research, and it works well when the moment of the event is not precisely observed, covariates are changing over the period units are at risk, and the interdependence in outcomes unfolds slowly. The type of event and length of time that units are at risk largely determine the appropriateness of discrete-time methods. For example, the onset of interstate wars, which are typically marked by formal declarations, is easier to pinpoint in time than civil wars, which can begin as spontaneous disorganized resistance. When units are at risk for decades or longer, even highly persistent covariates change considerably. Policy innovations that require legislative acts (e.g., lottery adoption) diffuse more slowly than those that require only executive action (e.g., currency devaluation).

Unfortunately, neither the spatial frailty nor the discrete-time event history approach is ideal when we have continuous interdependent durations with right censoring and fixed covariates over the period of risk. Consider the expansion of WWI, the illustration we present below. The period of risk lasted for less than 4.5 years (1567 days). Important covariates such as relative military capabilities change very little over this period, and any change that does occur is more a consequence than a cause of the conflict. Additionally, there is interdependence in the entry timing decisions of states, driven by rivalries and alliances, that occur over a matter of days rather than months or years. For research of this nature, we develop an estimation strategy for continuous-time spatial-lag models that account for right censoring. In this article, we consider the model for a log-normal duration process, but the approach can be generalized to other cases.

The data generating process for the spatial-lag duration model can be written in matrix notation as follows:

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \mathbf{X}\boldsymbol{\beta} + \mathbf{L}\mathbf{u}, \tag{1}$$

where $\mathbf{y}$ represents the natural log of the non-negative duration outcome and $\mathbf{W}$ specifies the connections between observations. The parameter ρ captures the degree of spatial dependence among the durations. In political science, frequently, spatial relationships are expected to be positive, leading to a positive ρ. The diagonal matrix $\mathbf{L}$ has $1/\lambda$ down the center diagonal, where λ represents the shape parameter for the error distribution and captures the existence and degree of duration dependence. From this structural form, we can derive the reduced form of the spatial model:

$$\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta} + (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{L}\mathbf{u}, \tag{2}$$

$$= \boldsymbol{\Gamma}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\Gamma}\boldsymbol{L}\boldsymbol{u}, \tag{3}$$

$$= \boldsymbol{\Gamma}\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{v}, \tag{4}$$

where $\boldsymbol{\Gamma} = (\mathbf{I} - \rho\mathbf{W})^{-1}$ and $\boldsymbol{v} = \boldsymbol{\Gamma}\boldsymbol{L}\boldsymbol{u}$ are the spatial multiplier matrix and reduced-form disturbances, respectively. The reduced-form equation makes evident the importance of including the spatial lag since the outcome for each unit depends on both the observed and unobserved components of other units through the spatial weights matrix. Ignoring this spatial dependence when it exists (i.e., $\rho \neq 0$) leads to biased coefficient estimates, as in the linear regression context (Franzese and Hays 2007; LeSage and Pace 2009).

The reduced-form equation also demonstrates the additional challenge posed by right censoring. Right censoring occurs when we do not observe the exact moment the event occurs, but know only that it occurred after a specific point in time. A standard duration model handles this easily by replacing the density of failure at a specific time with the survival function evaluated at the time of censoring, which captures the probability that the observation lasts longer than the censoring point. In the context of spatial durations, however, this becomes more complicated since the probability of surviving past a single point in time depends on the failure time of other observations and therefore the likelihood cannot be constructed via the marginal distributions of failure or censoring for each observation. Moreover, since the probability of failure for observed cases will often depend on the true failure time for right-censored cases, we lose critical information needed to model the failure time for all observations.

To explicitly introduce right censoring, assume we do not fully observe the duration, y_i, for some subset of observations but rather we observe only that it lasted at least until the censoring point. Thus, we divide observations into those cases that are censored, denoted with y_i^C, and those cases that we fully observe, denoted with y_i. In standard duration analysis, treating right-censored cases as fully observed when they are in fact right censored can lead to bias in the estimates of the parameters (Box-Steffensmeier and Jones 2004).

Given the lack of available alternatives, it is tempting to address right censoring in spatial duration models in one of two problematic ways. The first is to treat the time to censoring as an observed failure time. The second is to omit the censored observations from the analysis. Both of these approaches can lead to biased estimates. The first systematically shortens the duration for censored cases. The second leads to sample truncation, and unless the factors that produce the censoring are unrelated to those that affect when the event occurs, this will also tend to produce bias (Box-Steffensmeier and Jones 2004). Spatial dependence can exacerbate both sources of bias since even uncensored observations depend on the realization of the outcome variable in censored cases. We therefore require a method to recover the missing information about censored cases so that we can properly capture the dependence between observations, a task to which we turn in the next section.

3 An EM Approach for Imputation and Estimation

To develop an estimator that simultaneously addresses spatial interdependence and right censoring, we adapt Wei and Tanner's (1991) imputation method for (non-spatial) censored regression data to account for spatial interdependence. The logic of their approach lies in alternating between taking draws of the outcome variable for censored cases conditional on the censoring point and estimating the model using both the observed values for uncensored cases and the imputed values for censored ones. This approach allows estimation to proceed with computationally non-intensive regression models, which at one point would have been easier than doing maximum likelihood with a mix of censored and uncensored observations. We adapt it to spatial durations with censored data for the same reason since evaluating the likelihood for a high-dimensional multivariate distribution with a mix of observed and censored cases would prove challenging even today.

Extending the original approach requires accounting for the interdependence between observations as we iterate between the estimation and imputation stages. In order to do this, we need to sample from the multivariate distribution of reduced-form residuals. In the case of a multivariate normal distribution, we can use the Geweke–Hajivassiliou–Keane (GHK) sampler (Geweke 1989; Hajivassiliou and McFadden 1998; Keane 1994), which translates the correlated reduced-form errors into a linear combination of i.i.d. normal errors. This translation occurs via the Cholesky decomposition which, as an upper triangular matrix, allows us to iteratively solve for both the correlated reduced-form errors as well as their i.i.d. components. For each censored case, we obtain the minimum value of the i.i.d. normal error that ensures the realized value of the outcome variable will exceed the observed censoring point. We use this value to take a random draw for the i.i.d. error from a truncated (univariate) normal distribution and then substitute this draw into the solution for the next disturbance. Once this iterative process concludes, we calculate the dependent variable, reestimate the model, and repeat until the parameters stabilize.

To be precise, the algorithm starts by estimating a spatial regression model that treats the censoring point as the observed event time. Using these results, we calculate the reduced-form residuals. The data generating process follows from equation (4). The reduced-form error $\boldsymbol{\upsilon}$ represents a linear combination of i.i.d. errors. Using these reduced-form residuals, we then calculate the Cholesky decomposition, $\mathbf{A}^{-1}$, of the expected covariance matrix of these errors:

$$E[\hat{\boldsymbol{\upsilon}}\hat{\boldsymbol{\upsilon}}'] = E[(\boldsymbol{\Gamma}\boldsymbol{L}\hat{\mathbf{u}})(\boldsymbol{\Gamma}\boldsymbol{L}\hat{\mathbf{u}})'], \tag{5}$$

$$= (\boldsymbol{\Gamma}\boldsymbol{L})E[\hat{\mathbf{u}}\hat{\mathbf{u}}'](\boldsymbol{\Gamma}\boldsymbol{L})', \tag{6}$$

$$= (\boldsymbol{\Gamma}\boldsymbol{L})(\boldsymbol{I})(\boldsymbol{\Gamma}\boldsymbol{L})', \tag{7}$$

$$= (\boldsymbol{\Gamma}\boldsymbol{L})(\boldsymbol{\Gamma}\boldsymbol{L})'. \tag{8}$$

For a log-normal distribution, $\boldsymbol{L}$ captures the standard deviation as a departure from the standard normal error $\mathbf{u}$ and has σ_u down its diagonal. We can therefore simplify this for the log-normal case to $\boldsymbol{\sigma}_u^2\Gamma\Gamma'$.

We can then write the reduced-form errors as a linear combination of i.i.d. normal errors: $\hat{\boldsymbol{\upsilon}} = \mathbf{A}^{-1}\boldsymbol{\eta}$. Since $\mathbf{A}^{-1}$ is upper triangular, we can iteratively solve for the corresponding value of η one observation at a time starting with the last observation. Assume without loss of generality that we have solved for the kth observation. When we solve for the $(k-1)$th observation, if the observation is uncensored this results in the calculated values. If the $(k-1)$th observation is censored, we can calculate the minimum necessary value for the duration to exceed the censoring point given the calculated errors for the previous observations. The value of censoring point is calculated in the same way that we solve for the uncensored η_k. Once the censoring point is determined for k–1, we take a random draw from the censored standard normal distribution to obtain a value that is greater than the censoring point to impute η_{k-1} (see our online appendix for an example).

In order to create the imputed reduced-form spatial errors, we multiply the $\boldsymbol{\eta}$ by $\mathbf{A}^{-1}$. Using these imputed values for the errors, an imputed $\mathbf{y}$ is calculated from the combination of $\hat{\mathbf{y}}$ from the previous spatial-lag model and the imputed spatial errors. This method returns the same value of y for each of the uncensored cases and imputed outcomes for the censored cases. The imputed $\mathbf{y}$ is then run in a spatial-lag model, and the parameter vector is saved.

To account for estimation uncertainty in our procedure, we multiply impute M draws of the errors for censored cases. After each step of the EM process, we take the average of the M parameter vectors as our current estimate to begin the next step. The algorithm continues until the results converge, which we assess by whether the log-likelihood changes by less than a small amount, for example, 0.0001.[2] Once it converges, we calculate the final estimates by combining each of the multiply imputed data sets using Rubin's (2009) formula:

$$\overline{\boldsymbol{\theta}} = \frac{1}{M}\sum_{m=1}^{M}\hat{\boldsymbol{\theta}}_m, \tag{9}$$

$$\mathrm{Var}(\overline{\boldsymbol{\theta}}) = \frac{1}{M}\sum_{m=1}^{M}\mathrm{Var}(\hat{\boldsymbol{\theta}}_m) + \frac{M+1}{M}\left(\frac{1}{M-1}\sum_{m=1}^{M}(\hat{\boldsymbol{\theta}}_m - \overline{\boldsymbol{\theta}})^2\right). \tag{10}$$

[2]Tanner and Wong (1987) discuss the conditions under which the algorithm converges.

https://doi.org/10.1093/pan/mpv012 Published online by Cambridge University Press

4 Monte Carlo

In order to evaluate our approach, we conduct a series of Monte Carlo simulations.[3] These allow us to study the EM algorithm's properties, both statistically and computationally, and to compare the estimates that it produces to those that one would obtain from either ignoring the censoring or ignoring the spatial interdependence as well as to those from the original uncensored data as a benchmark comparison. To make these comparisons across a wide range of circumstances, we vary both the amount of spatial interdependence and censoring in the data.

Our data generating process proceeds as follows. We start with one hundred units spread out evenly across a ten-by-ten grid. We construct a spatial dependence matrix, **W**, based on queen contiguity. We then generate one hundred i.i.d. observations of a single independent variable, **X**, according to the standard normal distribution. For simplicity's sake, we set $\boldsymbol{L} = \boldsymbol{I}$, resulting in the following data generating process:

$$\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1}(-1 - 1 \times \mathbf{X}) + (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{u}. \quad (11)$$

With one hundred observations, our approach to generating **W** results in about around 7% connectivity between units. As is common, we row standardized the spatial weights matrix by dividing each element by the sum of the elements in its row, which produces a matrix in which all of the elements represent proportions and each row sums to 1. Since most spatial relationships in political science are expected to be positive, we run simulations with ρ varying from 0 to 0.75 by increments of 0.25. We hold the independent variables constant across all of the simulations.

We introduce censoring through a common censoring point for all observations. This mimics what occurs when some observations have not failed by the end of the study, such as when all units have not adopted a policy in an event history analysis or when the event of interest becomes infeasible at a certain point in time, as in our application to the timing of countries' entry into WWI. Based on the distribution of the dependent variable that results from our data generating process, we selected censoring points that vary from −1 to 0.5 by increments of 0.5. The amount of censoring that occurs ranges from 50% to nearly 0% depending on the degree of censoring and the amount of spatial correlation. (See Supplementary Fig. A.1 in our online appendix.)

For each combination of values of spatial interdependence and censoring points, we generate draws of **u** from a standard normal distribution, calculate the value of **y**, then apply our censoring rule so that $y_i^c = min\{y_i, C\}$. We estimate four models: a naïve spatial model that treats all realizations of y_i^c as uncensored, a naïve log-normal duration model that accounts for the censoring but ignores the spatial dependence, our EM spatial duration model with imputation of censored values, and a log-normal spatial duration model using the uncensored value y_i. The latter serves as a best-case scenario against which to compare our EM approach since spatial models often show some degree of bias in parameter estimates with relatively small samples. We repeat this five hundred times for each combination of the parameters.

For our EM estimator we set the maximum number of iterations for each draw to one hundred and use fifteen imputations for each step of the EM process. With high degrees of censoring, our proposed estimator occasionally fails to produce estimates. This happens about a third of the time when the degree of censoring is high and the spatial correlation is low, circumstances which make it difficult for our estimation procedure to draw strength across observations. (See Supplementary Fig. A.2 in our online appendix.) In exploring individual draws, we found this to be an issue usually resolved by changing the seed and rerunning the estimator, so it should not pose a significant problem for most applications. Standard errors are calculated according to equation (10).

Figure 1 presents the average bias in the estimates for the intercept, slope coefficient, and spatial dependence parameters, with the first two having true values of −1 and the latter varying from 0 to 0.75. We report detailed results for all parameters in Supplementary Tables A.1–A.3 of the online appendix. Four patterns emerge quite clearly. First, the benchmark estimates evidence a potential,

[3]Replication materials for this study are available from the *Political Analysis Dataverse* (Hays, Schilling, and Boehmke 2015).

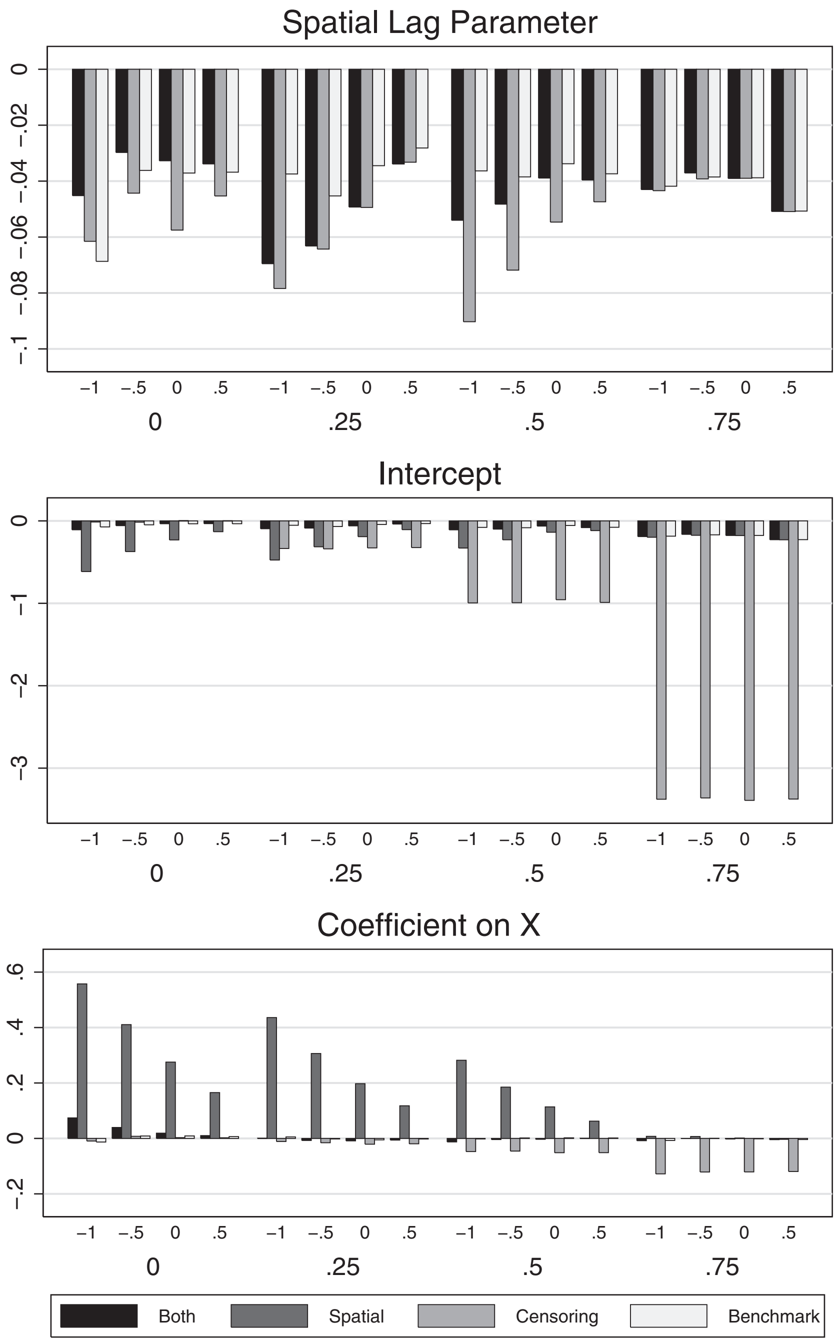

Fig. 1 Average bias from EM approach, naïve duration model, naïve spatial model, and uncensored benchmark spatial duration model, varying the amount of spatial correlation and the amount of censoring. *Notes:* Results represent the average deviation from the true parameter value across five hundred simulations, excluding cases for which the EM algorithm did not converge within one hundred iterations. EM algorithm performed with fifteen imputations. Duration coefficients estimated in time to failure format.

though slight, negative bias, especially for the intercept and spatial dependence parameters. This is consistent with other simulations of spatial estimators and should be kept in mind when evaluating the performance of the other estimators since it represents the best-case scenario in which no censoring occurs.

Second, the naïve spatial estimator suggests bias for all three parameters, with the deviations from the true values most severe for the coefficient and intercept. These deviations get smaller as the spatial correlation increases or the censoring parameter decreases. Since this estimator ignores censoring, it makes sense that it does worse as censoring increases. The magnitude of the apparent bias can be quite severe, with average estimates frequently ranging from 80% to 45% of the true value of -1. Whenever there is positive spatial dependence, there is also an attenuation bias in the average estimates for ρ that is decreasing in the strength of interdependence, ranging from a high of 60% to a low of 15%. To sum, because it ignores the censoring problem, the naïve spatial estimator understates both the effect of the covariate and the strength of interdependence.

Third, the naïve duration model exhibits the opposite pattern. It does quite well with little interdependence, which we expect, since when $\rho = 0$ our data correspond exactly to a standard duration model. When the spatial correlation reaches 0.25, however, the intercept begins to show some apparent bias, which quickly becomes much worse: the average bias is over 300% of the magnitude of the true value when $\rho = 0.75$. The slope coefficient exhibits a similar pattern, though on a much smaller scale, with deviations up to 10% of the true effect. In other words, our results show that the naïve duration model produces coefficient estimates that overstate on average the direct effect of the covariate.

Fourth, the EM estimator produces average estimates for all three parameters near their true values. The results shown here exhibit some potential bias in estimating the intercept and slope parameters for $\rho = 0$ and small values of C, which makes sense since in those circumstances one has much less information about the outcomes for the censored cases given that their values are independent of the outcomes in uncensored cases.[4] We also see some slight deviation in the estimate of the spatial parameter relative to the benchmark case with moderate spatial dependence, but not with large or zero dependence. When spatial correlation is very low, a simple duration estimator may be preferred, and when it is high, a simple spatial estimator may suffice, but our estimator performs about as well as the alternatives even in these situations and clearly outperforms both in all other cases.

Given the greater complexity of our estimator, including the multiple imputation component, we also want to consider the relative precision of its estimates. We start by evaluating the accuracy of the reported standard errors for our approach and then move to a mean-squared error comparison across alternatives. Figure 2 provides a comparison of the average standard errors across the five hundred draws to the standard deviation of the sampling distribution for the estimated coefficient on **X**. Generally speaking, our method provides reasonably accurate estimates of uncertainty, in line with Wei and Tanner (1991). For the other parameters, see our online appendix.

Moving to a comparison across estimators, Fig. 3 plots the square root of the sum of the variance and the squared bias of each parameter for all four estimators. The plots show clear evidence that even with its greater complexity our EM estimator generally equals or outperforms both of the naïve ones and by a wide margin for the spatial estimator with spatial interdependence less than 0.75 and for the duration estimator with spatial dependence equal to or greater than 0.5. The results for the spatial dependence parameter are generally comparable across models.

Overall, then, we take these results as providing solid evidence in favor of our EM spatial duration model for censored data. With even modest levels of censoring, it appears to outperform a naïve spatial duration model that ignores the censoring in root mean-squared error terms. The results also indicate that it provides estimates that do not deviate much from the benchmark model with fully observed data. With extremely high rates of censoring, some apparent bias does emerge, but given the size of the standard deviations it may not be meaningful.

It is worth noting that the basic results of these Monte Carlo experiments hold even when we reduce the sample size significantly. Our imputation approach continues to outperform the

[4]We confirmed this by increasing the amount of censoring but do not report the results given the increasingly low rate at which we obtained estimates.

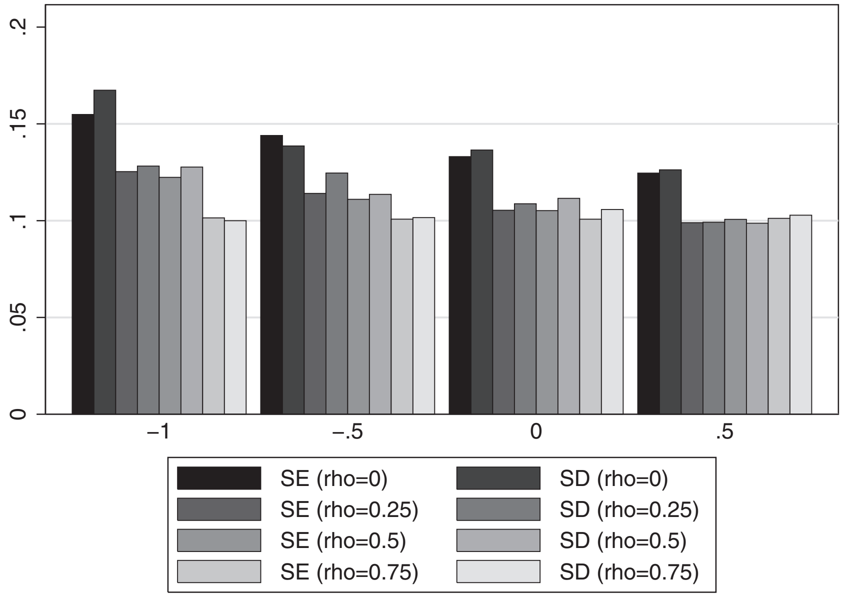

Fig. 2 Comparison of standard deviations and standard errors, varying the amount of spatial correlation and the amount of censoring. *Notes:* Results represent the average across five hundred simulations, excluding cases for which the EM algorithm did not converge within one hundred iterations. EM algorithm performed with fifteen imputations. Standard errors represent the average value across the five hundred iterations, while the standard deviation comes from the sampling distribution of the parameter estimates.

alternatives when we cut the sample size in half, using forty-nine units arranged on a seven-by-seven grid.[5] This is roughly the sample size in our illustration, which we present in the next section.

5 Illustration: The Diffusion of WWI

There is a large literature in international relations on the diffusion or contagion of war (Most and Starr 1980; Levy 1982; Siverson and Starr 1991; Gartner and Siverson 1996; Kadera 1998; Melin and Koch 2010; Radil, Flint, and Chi 2013). According to one line of thought, the spread of war is analogous to the spread of infectious disease. War is theorized to diffuse through geographical proximity, rivalries, and military alliances among other mechanisms. Duration models provide a natural framework for empirically evaluating theories of conflict diffusion. Given a particular level of conflict "exposure," the question is how long it will take before a country succumbs to the scourge of war. Right censoring presents a significant methodological challenge to duration analyses of war diffusion, however. Wars end before all the potential joiners have entered the conflict. An armistice is like a vaccine or the end of a clinical trial.

We model the WWI entry timing decisions of states using a spatial-lag model of interdependent durations. WWI is an excellent case for studying the diffusion of war. It began as a localized conflict that over the course of 4 years expanded to include half of the independent states in the international system. Had the war continued, undoubtedly, more states would have been drawn into in the conflict.[6]

[5]We do not report the results of these small sample experiments, but they are available upon request.

[6]Our model assumes that all of the countries in the sample were at risk of joining the war. We are comfortable with this assumption. It is critical to recognize the difference between choosing a side in the conflict and direct military involvement. Failure in this context (i.e., joining the war) just means giving up one's neutrality. All of the participants issued declarations of war, but not all of them fought. Many Latin American countries joined the war, for instance, but Brazil was the only country with significant direct military involvement. We believe all of the politically independent countries in the world at the time were at risk of abandoning their neutrality. Nevertheless, for some applications, there may be units that are not at risk and, for these cases, it would be useful to develop a split-population version of the spatial-lag duration model.

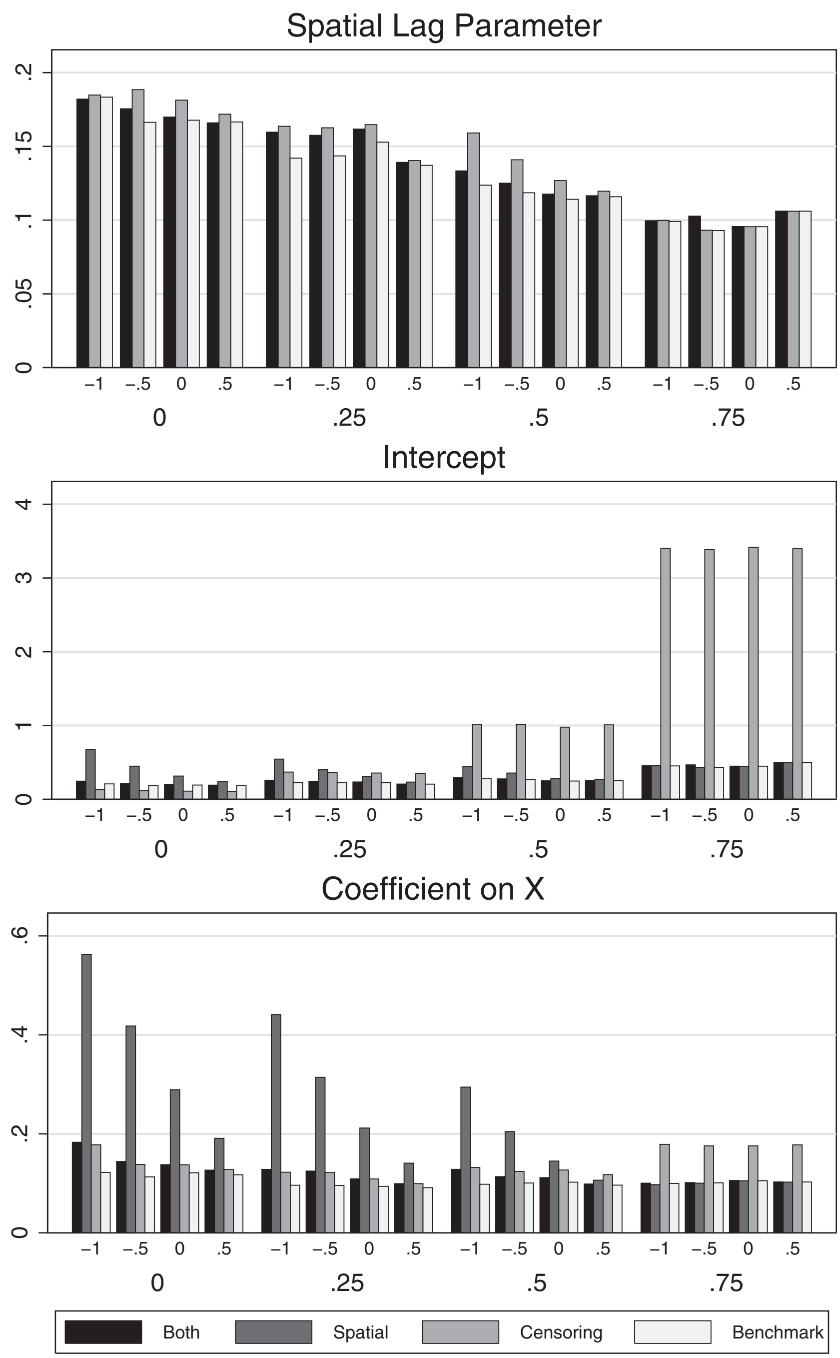

Fig. 3 Comparison of root mean standard errors from EM approach, naïve duration model, and uncensored benchmark spatial duration model, varying the amount of spatial correlation and the amount of censoring. *Notes:* Results represent the average across five hundred simulations, excluding cases for which the EM algorithm did not converge within one hundred iterations. EM algorithm performed with fifteen imputations. $\mathrm{RMSE}^2(\hat{\boldsymbol{\theta}}) = (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0)^2 + \mathrm{Var}(\hat{\boldsymbol{\theta}})$.

One could treat the entry timing decisions of these states as independent and driven purely by domestic and international structural factors such as regime type, trade exposure, and relative military capabilities, but this approach is unsatisfactory. Ultimately, each state's decision about when to enter the war was heavily influenced by the entry timing decisions of others, and any empirical analysis should take this interdependence into account. We incorporate three forms of interdependence into our models: geographical distance, rivalry, and defensive alliances. These sources of interdependence suggest that states will be influenced by the participation and entry timing decisions of their neighbors, rivals, and allies.

To capture the role of geography in the spread of war, we use an inverse distance spatial weights matrix, which is standard in the spatial econometrics literature. With this matrix, every state's entry timing decision is influenced by every other state's decision, but the interdependence between geographically proximate states is much stronger than it is for distant ones. Of course, the geographical notion of distance can be extended to social, political, and economic contexts (Beck, Gleditsch, and Beardsley 2006), and this is how we approach our rivalry and alliance weights matrices. Every state's entry timing decision is influenced by every other state's decision, but the interdependence between rivals and allies is much stronger than for states not connected by a rivalry or alliance. To be more specific about how we generate our matrices, consider the case of international rivalry. If state i has no rivals, then its weights for $i \neq j$ are

$$w_{ij} = \frac{1}{n-1}. \tag{12}$$

If state i has both rivals and non-rivals, then its weights for non-rivals $i \neq j$ are

$$w_{ij} = \frac{(1-r)}{n-n_r-1}, \tag{13}$$

where r is a parameter on the interval $[n_r/(n-1), 1]$ that determines the relative influence of rivals and non-rivals and n_r is i's total number of rivals. For i's rivals, the weights are

$$w_{ij} = \frac{r}{n_r}. \tag{14}$$

With this function for w_{ij}, everyone influences everyone else equally is a special case where $r = n_r/(n-1)$. This produces a weights matrix for which all of the off-diagonal elements are $1/(n-1)$. At the other extreme, $r = 1$, the function produces rows in the weights matrix where *only* rivals have influence. We report the results for $r = 1$ below, but our findings are robust to the choice of r.[7]

We identify all of the defensive alliances that were in force at the onset of the war using Gibler (2009), being careful to exclude alliances that were formed during the conflict. In total, there were sixteen defensive alliances connecting eleven of the countries in our sample. We code all of the rivalries, both proto-rivalries (short-term) and enduring ones, that existed at the onset of the war using Klein, Goertz, and Diehl (2006). Again, we do not include rivalries that emerged during the conflict. At the onset of the war, there were thirty-one active rivalries involving twenty-six of the states in our sample.

Following Radil, Flint, and Chi (2013), our dependent variable is the number of days before entering WWI. Of the forty-four sample countries, half eventually joined the conflict. All the spatial weights matrices are row standardized. We include three important covariates in the analysis. National capabilities are the (rescaled) COW CINC index scores (Singer, Bremer, and Stuckey 1972); democracy is the polity measure of regime type (Marshall and Jaggers 2002); and trade is the value of total trade in current US dollars (Barbieri 2002). We estimate the three types of models evaluated in our Monte Carlo: a non-spatial duration model, a naïve spatial model that treats the time of the censoring as an observed failure time, and our multiple imputation model. Based on the Monte Carlo results, we expect the estimates from non-spatial duration model to overstate the effects of the covariates on war joining. This is particularly true for variables such as national capabilities that cluster among states that are linked by the mechanisms or "vectors" through which

[7]Given the functional-form assumptions, one could estimate the parameter r. In this case, the unobserved weight is a function of an observed relationship (rivalry or alliance) with an estimable parameter. This would be roughly analogous to replacing inverse distance in a weights matrix with a powered exponential function of distance (e.g., see Ripley 2004).

Table 1 Comparison of log-normal duration models of the timing of entry into WWI

		Distance		*Alliance*		*Rivalry*	
Spatial lag	*None*	*Naive*	*Imp.*	*Naive*	*Imp.*	*Naive*	*Imp.*
Constant	9.211*** (1.366)	6.724*** (1.196)	7.898** (2.309)	6.018*** (1.031)	6.707*** (1.492)	4.581*** (1.147)	4.586** (1.876)
Capabilities	−0.315** (0.133)	−0.180** (0.750)	−0.279** (0.116)	−0.141** (0.743)	−0.189* (0.107)	−0.142** (0.682)	−0.303* (0.184)
Democracy	0.017 (0.087)	0.001 (0.047)	−0.007 (0.082)	0.013 (0.046)	0.024 (0.078)	−0.000 (0.043)	0.061 (0.116)
Trade	−0.208 (0.281)	−0.186 (0.159)	−0.250 (0.270)	−0.262* (0.152)	−0.389 (0.252)	−0.138 (0.143)	−0.090 (0.358)
σ^2	3.070 (0.500)	3.523*** (0.752)	7.897** (2.309)	3.243*** (0.694)	6.442*** (1.786)	2.832*** (0.614)	12.944** (4.972)
ρ		0.120 (0.139)	0.214 (0.163)	0.313** (0.155)	0.471*** (0.158)	0.447*** (0.142)	0.675*** (0.118)

Notes. $N = 44$. Standard errors in parentheses. ***$P < 0.01$, **$P < 0.05$, *$P < 0.1$.

conflict diffuses. For example, national capabilities in our sample are more than twice as high among the states that are connected by alliance networks. Additionally, because the naïve spatial model fails to account for right censoring, we expect its estimates to understate both the strength of interdependence as well as the direct effects (i.e., those prior to any spatial or network feedback) of the covariates on war joining.

We report the results in Table 1, where some familiar patterns emerge. Before discussing these results, however, a word of caution is in order. In the Monte Carlo experiments, we were able to compare the models because we knew the true data generating process. We could say that, on average, the naïve duration model overstated the effect of the covariate while the naïve spatial model understated it. With the application, we do not know the data generating process. Moreover, model comparison across the three approaches is complicated by the fact that either the likelihoods are not directly comparable (naïve duration versus EM imputation) or, when the likelihoods are comparable, the models treat the value of the outcome variable differently for censored cases (naïve spatial versus EM imputation). Consequently, we cannot compare likelihood-based measures of model fit such as the AIC or SBC. Instead, we interpret the results of our application in light of the Monte Carlo experiments, recognizing that we cannot be certain that the interpretation is correct.

With this in mind, several things about the results are worth noting. First, among the covariates, national capabilities is the only one that has a robust statistically significant effect on the war-participation timing of states. Military power is associated with early entry into the war. Second, among the spatial lags, only the rivalry and alliance lags are statistically significant. We do not find that geography matters for the participation timing decisions of states. This may seem surprising at first; however, the simple fact that this was a *world* war rather than a localized conflict means that geography was less significant than would otherwise be the case. For WWI, military power, rivalry, and alliances were more important determinants of participation than a state's geographical location.[8]

When we compare the estimators, we find very similar patterns to those in the Monte Carlo experiments. Focusing on national capabilities, we see that the estimated coefficient from the non-spatial duration model is much larger than the estimates from the other models.[9] The difference is particularly stark when we use the alliance weights matrix (national capabilities and alliance membership are strongly correlated) and cannot be explained by the fact that the spatial model decomposes the total effect of national capabilities into a direct effect, given by the estimated coefficient, and an indirect effect through spatial feedback, which is identified from the diagonal elements of the spatial multiplier matrix. If we calculate the spatial multipliers from our imputation model, $(\mathbf{I} - \rho\mathbf{W})^{-1}$, we find that for the defensive alliance weights matrix the average multiplier is 1.05 and the maximum is 1.29. Inflating the direct effect estimates by these multipliers gives $1.05 \times -0.189 = -0.199$ and $1.29 \times -0.189 = -0.244$. The estimated effect from the non-spatial model, −0.315, is

[8]This is true whether we use an inverse distance or contiguity spatial weights matrix, once we control for national capabilities and trade. The contiguity spatial lag is significant in a model without these two covariates.

[9]Again, we need to be careful interpreting these differences across the three approaches, which could be due to chance rather than the systematic tendencies uncovered in the Monte Carlo experiments.

29% larger than the maximum effect from the alliance spatial model and 58% larger than the average effect. Given the natural log scale, these effects are given approximately in percentage change. Consider the case of Germany, which has a CINC index score of 16 and spatial multiplier of 1.12. Our imputation model predicts that a one-unit decrease in the relative power capabilities of Germany would have increased its logged survival time by $1.12 \times 0.189 = 0.212$, approximately 21.2%, which compares to a 31.5% increase in the expected number of days before Germany joins the war from the non-spatial model.

Additionally, relative to the imputation model, the estimates from the naïve spatial model imply a smaller direct effect of national capabilities on survival times as well as weaker interdependence. This suggests much smaller spillover effects from one country to another. The strength of spillovers is determined by the off-diagonal elements of the spatial multiplier matrix. Element (i,j) of the matrix gives the spillover from unit i to unit j. Using the imputation model, the estimated spillover from Germany to Italy, for example, is 0.257, which implies that a one-unit decrease in Germany's relative military capabilities leads to an approximately 5% increase in the expected number of days before Italy joins the war ($0.257 \times 0.189 = 0.049$). Using the naïve spatial model, the same spillover effect is a 2% increase ($0.138 \times 0.141 = 0.019$).[10]

To sum, we believe that our imputation approach to interdependent duration analysis leads to a more accurate understanding of the diffusion of WWI. If the Monte Carlo experiments provide a useful guide, when compared with the non-spatial and naïve spatial approaches, ours produces more accurate estimates of the total effects of covariates—the estimates from the non-spatial duration model seem to overstate the effect of national capabilities, for example, while the estimates from the naïve spatial model understate this effect—as well as, in the case of the naïve spatial model, a better decomposition into direct effects and interdependence-driven multipliers.

6 Conclusion

In this article, we develop an approach to estimate interdependent duration models when the outcome suffers from right censoring. Interdependent duration outcomes abound in all fields of political science and include the time it takes states to enter wars or alliances; the diffusion of policies across countries or subnational units; and the timing of candidate entry decisions into election contests or issue position taking by elected representatives. Simply put, politics and strategic behavior frequently generate duration interdependence across actors.

One problem for studying interdependent durations in politics is that right censoring prevents us from fully observing the consequences of this interdependence. Adapting Wei and Tanner's imputation algorithm for censored (non-spatial) data to models of spatially interdependent durations provides a method for estimating a spatial duration model with right censoring. Our Monte Carlo analysis shows that this approach performs well and is almost always preferable to simply ignoring the censoring problem or the spatial interdependence. We see a number of ways in which future work could profitably extend our methods.

First, it would be helpful to develop analogous estimators that allow for other forms of duration dependence. Our focus on log-normal duration processes is driven by the relative ease of drawing from a multivariate normal distribution via importance sampling and the GHK simulator. We believe working with other distributions, such as the Weibull, is possible, but it requires evaluating the integrated (interdependent) baseline hazard function. In the case of the Weibull, the baseline hazard function is a sum of extreme value distributions. One could either sample efficiently from this distribution using the results in Nadarajah (2008) or by using a more "brute force" approach, that while computationally more costly, should still be feasible. More generally, we would like to offer a semi-parametric approach. This could follow along the lines of Wei and Tanner, who sample from a Kaplan–Meier estimate of the distribution of residuals. However, this hinges critically on

[10]Note that confidence intervals for any of these effects can be computed easily using parametric bootstrap techniques. To do so, take random draws from the sampling distribution for ρ and β, calculate the implied multipliers, and record the desired percentiles of the resulting distribution. The 90% confidence interval for the spillover from Germany to Italy, for example, is [0.0003, 0.1485].

the assumption that these residuals are i.i.d. Alternatively, one could use a mixture of beta distributions as in Banerjee and Carlin (2003) to produce very flexible forms of duration dependence (see also Royston and Parmar 2002).

Second, models for interdependent durations can also be extended to accommodate other interesting features of duration data. For example, our estimator can be extended to account for possible competing risks since units that fail by one risk are treated as right censored for other risks (Schilling 2014). Researchers using duration data must also be concerned with whether their data meet the proportional hazards assumption, which says that the relative effect of a variable on the hazard must stay constant over time. Often this does not hold. When it does not, the standard fix involves interacting the offending variable with a function of time, usually the natural logarithm (Box-Steffensmeier and Jones 2004). This fix means moving to a duration model that accommodates time-varying covariates, which our estimator does not. Developing an extension that accounts for non-proportional hazards, or time-varying covariates in general, would not be trivial since it involves a move from a single duration outcome per unit to one observation per unit per period, which creates more ways to account for spatial interdependence.

Finally, we can view this as part of a broader class of problems with missing data in spatial regression models. Here, we only lose part of the information about the value of the dependent variable, but in many empirical applications we lose the entire value of the dependent variable whether due directly to its missingness or to listwise deletion from missing covariates. With independent observations, the consequences and fixes for regression-type models are well known (King et al. 2001), but the consequences of such missingness in spatially interdependent data are not, and our results suggest that modifications to existing solutions may be needed.

Conflict of interest statement. None declared.

References

Allison, P. D. 1982. Discrete-time methods for the analysis of event histories. *Sociological Methodology* 13(1):61–98.

Banerjee, S., and B. P. Carlin. 2003. Semiparametric spatio-temporal frailty modeling. *Environmetrics* 14(5):523–35.

Banerjee, S., A. E. Gelfand, and B. P. Carlin. 2004. *Hierarchical modeling and analysis for spatial data*. Boca Raton, FL: CRC Press.

Banerjee, S., M. M. Wall, and B. P. Carlin. 2003. Frailty modeling for spatially correlated survival data, with application to infant mortality in Minnesota. *Biostatistics* 4(1):123–42.

Barbieri, K. 2002. *The liberal illusion: Does trade promote peace?* Ann Arbor, MI: University of Michigan Press.

Beardsley, K. 2011. Peacekeeping and the contagion of armed conflict. *Journal of Politics* 73(4):1051–64.

Beck, N., K. S. Gleditsch, and K. Beardsley. 2006. Space is more than geography: Using spatial econometrics in the study of political economy. *International Studies Quarterly* 50(1):27–44.

Beck, N., J. N. Katz, and R. Tucker. 1998. Taking time seriously: Time-series-cross-section analysis with a binary dependent variable. *American Journal of Political Science* 42:1260–88.

Berry, F. S., and W. D. Berry. 1990. State lottery adoptions as policy innovations: An event history analysis. *American Political Science Review* 84(2):395–415.

Box-Steffensmeier, J. M., L. W. Arnold, and C. J. Zorn. 1997. The strategic timing of position taking in Congress: A study of the North American free trade agreement. *American Political Science Review* 91(2):324–38.

Box-Steffensmeier, J. M., and B. S. Jones. 2004. *Event history modeling: A guide for social scientists*. New York: Cambridge University Press.

Darmofal, D. 2009. Bayesian spatial survival models for political event processes. *American Journal of Political Science* 53(1):241–57.

Franzese, Robert J., and J. 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.

Gartner, S. S., and R. M. Siverson. 1996. War expansion and war outcome. *Journal of Conflict Resolution* 40(1):4–15.

Gasiorowski, M. J. 1995. Economic crisis and political regime change: An event history analysis. *American Political Science Review* 89(04):882–97.

Geweke, J. 1989. Bayesian inference in econometric models using Monte Carlo integration. *Econometrica* 57(6):1317–39.

Gibler, D. M. 2009. *International military alliances, 1648–2008*. CQ Press.

Hajivassiliou, V. A., and D. L. McFadden. 1998. The method of simulated scores for the estimation of LDV models. *Econometrica* 66(4):863–96.

Hays, J. C., E. U. Schilling, and F. J. Boehmke. 2015. *Replication data for: Accounting for right censoring in interdependent duration analysis*. http://dx.doi.org/10.7910/DVN/29603 Dataverse [Distributor] V1 [Version].

Hennerfeind, A., A. Brezger, and L. Fahrmeir. 2006. Geoadditive survival models. *Journal of the American Statistical Association* 101(475):1065–75.

Kadera, K. M. 1998. Transmission, barriers, and constraints a dynamic model of the spread of war. *Journal of Conflict Resolution* 42(3):367–87.

Keane, M. P. 1994. A computationally practical simulation estimator for panel data. *Econometrica* 62(1):95–116.

King, G., J. Honaker, A. Joseph, and K. Scheve. 2001. Analyzing incomplete political science data: An alternative algorithm for multiple imputation. *American Political Science Review* 95(01):49–69.

Klein, J. P., G. Goertz, and P. F. Diehl. 2006. The new rivalry dataset: Procedures and patterns. *Journal of Peace Research* 43(3):331–48.

Leblang, D. 2003. To devalue or to defend? The political economy of exchange rate policy. *International Studies Quarterly* 47(4):533–60.

LeSage, J., and R. K. Pace. 2009. *Introduction to spatial econometrics*. Boca Raton, FL: CRC Press.

Levy, J. S. 1982. The contagion of great power war behavior, 1495–1975. *American Journal of Political Science* 26(3):562–84.

Marshall, M. G., and K. Jaggers. 2002. Polity IV project: Political regime characteristics and transitions, 1800–2002. http://systemicpeace.org/polity/polity4.htm (accessed July 11, 2013).

Melin, M. M., and M. T. Koch. 2010. Jumping into the fray: Alliances, power, institutions, and the timing of conflict expansion. *International Interactions* 36(1):1–27.

Most, B. A., and H. Starr. 1980. Diffusion, reinforcement, geopolitics, and the spread of war. *American Political Science Review* 74(4):932–46.

Nadarajah, S. 2008. Exact distribution of the linear combination of p Gumbel random variables. *International Journal of Computer Mathematics* 85(9):1355–62.

Radil, S. M., C. Flint, and S.-H. Chi. 2013. A relational geography of war: Actor-context interaction and the spread of World War I. *Annals of the Association of American Geographers* 103(6):1468–84.

Ripley, B. D. 2004. *Spatial statistics*. Hoboken, NJ: John Wiley & Sons.

Royston, P., and M. K. Parmar. 2002. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in Medicine* 21(15):2175–97.

Rubin, D. B. 2009. *Multiple imputation for nonresponse in surveys*, vol. 307. Hoboken, NJ: Wiley.com.

Schilling, E. 2014. Timing of congressional position taking: An estimator for a spatial duration model with competing risks. Presented at the Visions in Methodology Conference, 2014. http://visionsinmethodology.org/conferences/2014-conference/ (accessed July 11, 2013).

Singer, J. D., S. Bremer, and J. Stuckey. 1972. Capability distribution, uncertainty, and major power war, 1820–1965. In *Peace, war, and numbers*, ed. B. M. Russett, 19–48. Beverly Hills, CA: Sage Publications.

Siverson, R. M., and H. Starr. 1991. *The diffusion of war: A study of opportunity and willingness*. University of Michigan Press.

Tanner, M. A., and W. H. Wong. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* 82(398):528–40.

Wei, G. C., and M. A. Tanner. 1991. Applications of multiple imputation to the analysis of censored regression data. *Biometrics* 47(4):1297–309.