A Comparison of the Small-Sample Properties of Several Estimators for Spatial-Lag Count Models *

Jude C. Hays (jchays@uiuc.edu)
Assistant Professor of Political Science, University of Illinois

Robert J. Franzese, Jr. (franzese@umich.edu)
Professor of Political Science, University of Michigan

23 July 2009

ABSTRACT: Political scientists frequently encounter and analyze spatially interdependent count data. Applications include counts of coups in African countries, of state participation in militarized interstate disputes, and of bills sponsored by members of Congress, to name just a few. The extant empirical models for spatially interdependent counts and their corresponding estimators are, unfortunately, dauntingly complex, computationally costly, or both. They also generally tend 1) to treat spatial dependence as nuisance, 2) to stress spatial-error or spatial-heterogeneity models over spatial-lag models, and 3) to treat all observed spatial association as arising by one undifferentiated source. Prominent examples include the Winsorized count model of Kaiser and Cressie (1997) and Griffith's spatially-filtered Poisson model (2002, 2003). Given the available options, the default approaches in most applied political-science research are to either to ignore spatial interdependence in count variables or to use spatially-lagged observed-counts as exogenous regressors, either of which leads to inconsistent estimates of causal relationships. We develop alternative nonlinear least-squares and method-of-moments estimators for the spatial-lag Poisson model that are consistent. We evaluate by Monte Carlo simulation the small sample performance of these relatively simple estimators against the naïve alternatives of current practice. Our results indicate substantial consistency improvements against minimal complexity and computational costs. We illustrate the model and estimators with an analysis of terrorist incidents around the world.

* Research supported in part by NSF grant #0318045. We thank Chris Achen, Jim Alt, Kenichi Ariga, Klaus Armingeon, Neal Beck, Jake Bowers, Jim Caporaso, Kerwin Charles, Bryce Corrigan, Tom Cusack, David Darmofal, Jakob de Haan, Scott De Marchi, John Dinardo, Zach Elkins, James Fowler, John Freeman, Fabrizio Gilardi, Kristian Gleditsch, Mark Hallerberg, John Jackson, Jonathan Katz, Mark Kayser, Achim Kemmerling, Gary King, Greg Koger; Hasan Kirmanoglu, Herbert Kitschelt, Jim Kuklinski, Don Lacombe, Chris Lawrence, David Lazer, Tse-Min Lin, Xiaobo Lu, Scott McClurg, Walter Mebane, Covadonga Meseguer, Mark Mizruchi, Michael Peress, Thomas Pluemper, David Prosperi, Dennis Quinn, Megan Reif, Frances Rosenbluth, Ken Scheve, John Scholz, Phil Schrodt, Ken Schultz, Jasjeet Sekhon, Chuck Shipan, Beth Simmons, Betsy Sinclair, David Siroky, John D. Stephens, Duane Swank, Wendy Tam-Cho, Vera Troeger, Craig Volden, Michael Ward, Greg Wawro, and Erik Wibbels – and the many others we have surely forgotten – for useful comments on this and/or other work in our broader project on spatial-econometric models in political science. All remaining errors are ours alone.

KEYWORDS: Interdependence, Spatial Econometrics, Spatial-Lag Models, Count Data, Poisson, Nonlinear Least-Squares, GMM Estimation

JEL CODES: C20, C21, C25.

I. Introduction:

Many phenomena that social scientists study are inherently, or by measurement, counts of events. Canonical political-science examples include counts of coups in African countries, of riots or acts of political violence within countries, of militarized interstate disputes (MIDs) or of states participating in MIDs, of terrorist attacks or of deaths or casualties by terrorist attacks, of bills sponsored by members of Congress, or of particular measures enacted in certain policy areas (the number of labor-rights provisions, for example). In all these political contexts, and indeed in most others and widely across the social sciences, substantively and theoretically, the numbers of events in some units depend on those of other units. Coups, riots, and violence at home are often spurred by example, fomentation, or otherwise from abroad; whether, which, and how many others join conflicts heavily condition states' entry to and involvement in MIDs; terrorists' choices of targets and natures and magnitudes of attacks in one area seem very likely to depend on what attacks they and others choose for other areas and what countermeasures are chosen in those and other areas; legislators' decisions to sponsor bills surely depend on their expectations or observations others' sponsorship activities; and governments' policy choices clearly depend on others' policies, whether by competition, learning, or by other mechanisms.

Indeed, as *Tobler's Law* (Tobler 1970) aptly sums: "Everything is related to everything else, but near things are more related than distant things." Furthermore, as Beck et al.'s (2006) pithy title reminds in corollary: "Space is More than Geography." That is, the substantive contents of the proximity in *Tobler's Law*, and so the pathways along which interdependence between units may operate, extend well beyond physical distance, contact, and contiguity (as some of the examples above attest). Long literatures in regional science, geography, and sociology elaborate from those disciplinary perspectives the multifarious mechanisms by which contagion may arise. Elkins & Simmons (2005) and Simmons et al. (2006), for instance, offer a list for international relations: coercion, competition, learning, and emulation.[1] In fact, as, e.g., Brueckner (2003) showed, strategic interdependence arises

[1] For fuller, closer match to prior traditions, add *cooperation* and *externality* to *competition*, combine *learning* and *emulation* as one, and add *relocation diffusion* (Haegerstrand 1967, 1970)— meaning the direct movement of some components of units i into other units j, such as by human migration or disease contagion. Note that aspects of these mechanisms may induce spatial association by *common-exposure* or *selection* effects, as opposed or in addition to by *interdependence* (see below). For example, *learning* from other units implies *contagion*, whereas *learning* from one's own experiences could implicate *common-exposure* sources of spatial

any time some unit(s)'s actions affect the marginal utility of other(s)'s actions. Given such externalities, *i*'s utility depends on both its policy and that of *j*.[2] In environmental policy, for instance, domestic welfare (or net political-economic benefits to policymakers) in each country will depend on the actions of others due to both environmental spillovers (e.g., of pollution) and economic ones (e.g., in regulatory costs). Optimizing behavior will yield best-response functions of *i*'s optimal policies as a function of *j*'s and *vice versa*. In this frame, positive externalities create free-rider incentives, which induce policies to move in opposite directions (i.e., as strategic substitutes), confer late-mover advantages, and make war-of-attrition (strategic delay or inaction) dynamics likely. Conversely, negative externalities create strategic complementarity, with policies moving in the same direction, yielding early-mover advantages and competitive races.[3] It's hard to imagine substantive contexts within *social* science, which is essentially defined by human strategic and nonstrategic interaction, where the actions, choices, or outcomes in some units did not depend on those in others.

The substantive/theoretical ubiquity and centrality of interdependence across political-science contexts notwithstanding, studies that accord it explicit attention are uncommon in general and almost nonexistent in event-count contexts.[4] Social-scientific interest in and applications of spatial modeling have burgeoned lately, however, due partly to advances in theory that imply interdependence and in methodology to address it; partly to global substantive developments that have raised perception of and attention to interconnectivity, at all levels, from micro/personal to macro/international; partly to advances in technology for obtaining and working with spatial data. Our aim in this paper is to aid the expansion of this obviously very welcome development—the dependence of outcomes in some units on those in others, *interdependence*, being so substantively and theoretically central across the political and other social sciences—to the study of political-science event-counts.

---

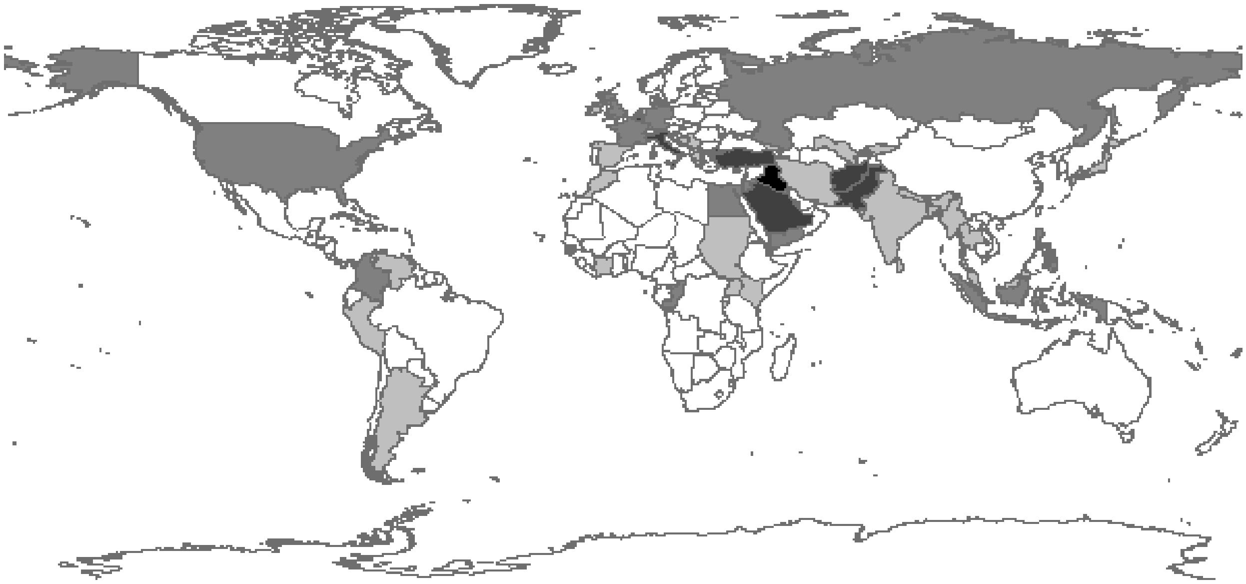
association insofar as units' experiences and lessons correlate spatially.

[2] In such microeconomic models, Manski (2000) shows externalities could arise from interactions, expectations, and/or preferences. Moreover, non-strategic interdependence could arise even without externalities. Examples and reviews of the booming microeconomic-theory literature with explicit models of interdependence include Akerlof (1997); Glaeser & Scheinkman (2003), Glaeser et al. (2003); Brock & Durlauf (2001).

[3] We eschew the terms *race to the bottom* (or *top*) and *convergence* because these competitive races need not foster convergence to top, bottom, or mean, and could spur divergence (see below and, for related further discussion of the observable regarding convergence, Plümper & Schneider 2007).

[4] The rare exceptions include some of the literature on "the contagious coup hypothesis", which dates at least to Li & Thompson (1975), and [XXXX more review & citations here XXXX].

Figure 1: Map of Counts of Post-9/11 Terrorist Incidents



Empirically, the clustering or correlation of event counts on some dimension(s) of proximity, *spatial association*, is also obvious in a vast array of substantive contexts. For instance, the spatial clustering clearly manifests (in this case, along a geographic dimension) in Figure 1, which maps the dependent variable in our application, the count of post-9/11 terrorist events by country. However, outcomes may exhibit spatial association at least two distinct reasons. First, units may be responding similarly to similar exposure to similar exogenous internal/domestic or external/foreign stimuli (*common exposure*); second, units' responses may depend on others' responses (*contagion*). We may find terrorists' targeting and tactical choices, for example, to cluster geographically or along other dimensions of proximity, among states with strong ties to a particularly hated state, for instance, because states that are proximate on that dimension also experience similar exogenous domestic or foreign political-economic stimuli or because terrorist and/or counterterrorist policies toward/in one unit affect their policies toward/in other units nearby along this dimension.[5]

[5] A third possibility arises when the putative outcome variable, in this case acts or effects of terrorism, affects the variable along which clustering occurs (*selection*). Terrorist activities could shape alliances, for example, and we would very likely observe clustering of targeting and effects among allies. In Hays et al. (2009), we begin exploration of possibilities of modeling such network-behavior *coevolution*, i.e., *contagion* and *selection* simultaneously, and along with *common exposure*, in a linear spatial-autoregression context, with application to nation's active-labor-market policies. We do not raise that considerable further complication yet here in this event-count context. Likewise, Hays (2009) begins consideration of simultaneous within-unit and across-unit endogeneity (respectively, when $y_i \Leftrightarrow x_i$ and $y_i \Leftrightarrow y_j$) in the linear and binary spatial-autoregression context, but we do not explicitly extend that framework here to the event-count context yet.

We discussed elsewhere (Franzese & Hays 2003, 2004ab, 2005ab, 2006abc, 2007abcd, 2008abc, 2009) the severe empirical-methodological challenges in estimating *common exposure* and *interdependence* distinctly and well (a.k.a, *Galton's Problem*).[6] In this previous work, we explored specification, estimation, interpretation, and presentation of spatial- and spatiotemporal-lag linear-regression models, which reflect spatial interdependence directly and which therefore can distinguish *common exposure* from *contagion* as alternative substantive sources of observed spatial association. For such models, we gauged analytically and by simulation the biases of omitting interdependence or of including spatial lags to reflect it but ignoring the lags' simultaneity, finding the former a typically far graver concern, but that the latter also becomes appreciable as interdependence strengthens. We explained maximum-likelihood and moment estimators that redress the simultaneity, and our simulations showed them near-dominant as estimators. We explored various model-specification tests and the sensitivity of the estimators' performance to misspecification of the non-spatial component, of the pattern of spatial-connectivity, or of the strict assumptions of the consistent estimators. Finally, we showed how to calculate spatial and spatiotemporal dynamics and initial and equilibrium effects, along with their certainty estimates. We emphasized there and here that, regardless of how one's interests weigh among exogenous internal/domestic, external/foreign, or

---

[6] *Galton's* famous *Problem* is related to Manski's (also famous) *Reflection Problem* (1993, 1995, 2000), which in part is a formalization of Galton's profound comment revealing its full implications.

Galton originally raised the issue thus: "[F]ull information should be given as to the degree in which the customs of the tribes and races which are compared together are independent. It might be that some of the tribes had derived them from a common source, so that they were duplicate copies of the same original. ...It would give a useful idea of the distribution of the several customs and of their relative prevalence in the world, if a map were so marked by shadings and colour as to present a picture of their geographical ranges" (Darmofal 2009 [Galton 1889]). At http://en.wikipedia.org/wiki/Galton's_problem: "In [1888], Galton was present when Sir Edward Tylor presented a paper at the Royal Anthropological Institute. Tylor had compiled information on institutions of marriage and descent for 350 cultures and examined the correlations between these institutions and measures of societal complexity. Tylor interpreted his results as indications of a general evolutionary sequence, in which institutions change focus from maternal to paternal lines as societies grow more complex. Galton disagreed, noting that similarity between cultures could be due to borrowing, could be due to common descent, or could be due to evolutionary development; he maintained that without controlling for borrowing and common descent one cannot make valid inferences regarding evolutionary development. Galton's critique has become the eponymous Galton's Problem (Stocking 1968:175), as named by Raoul Naroll (1961, 1965 and Naroll & D'Andrade 1963) who proposed [some of] the first statistical solutions."

Manski's *Reflection Problem*, most specifically, refers to a series of proofs demonstrating, for instance, that identical exposure to identical-response-inducing shocks among a group cannot be distinguished (in a cross-section) from identical response to each other's outcomes.

context-conditional effects on one hand, or contagion/interdependence on the other, valid inferences regarding either possibility generally requires empirical models that specify and estimate both of them well because the two typically look much alike empirically and so the relative omission or inadequacy in the empirical model and estimates of the one will bias inferences in favor of the aspects of the other(s) most similar to it.[7]

This paper begins address in similar manner of Spatial-Count (or *S-Count*) models, as part of our broader such expansion to models with interdependence among qualitative or limited dependent variables (*S-QualDep*). As before, we stress substantively-theoretically guided (i.e., structural) specifications that can support counterfactual analyses of spatial or spatiotemporal responses in dependent-variable terms and that can distinguish the possible sources of spatial association. That is, we insist of our approach that it must address this difficult but crucial task of distinguishing and weighing the strengths of the alternative processes by which spatial association may arise. The approach must also afford estimation of the kinds of spatial and spatiotemporal effect-estimates of most substantive interest to political/social science, namely the responses across the units and over time to counter-factual shocks to explanatory variables in some units.[8] The kind of structural models we prefer and propose, first, take these two aims as essential; then, secondarily, simplicity is also desirable, as we wish to facilitate and foster widespread proper use of spatial-analytic methods among applied empirical researchers.

II. <u>Existing Estimators</u>:

In brief and in general, difficulties arise for *S-QualDep* models as the endogenous spatial and/or temporal lags of latent-outcomes, y_j^* and/or y^*_{t-1}, on their right-hand sides (RHS) are converted to observed quantities by computation of an integral (e.g., the cumulative

---

[7] All of this extends naturally to include the third possibility, *selection* (see note 5 and Hays et al. 2009).

[8] We would also note in passing here, although this is an argument for fuller development in another venue, that the presence of interdependence generally violates the crucial SUTVA assumption of matching methods or other "potential-outcomes" frameworks for causal inference, invalidating, or at least seriously complicating, such less structural/parametric approaches to empirical inference. As Rubin himself noted: "The two most common ways in which SUTVA can be violated appear to occur when (a) there are versions of each treatment varying in effectiveness or (b) there exists interference between units" (1990: 282). Conditions (a) and (b) are, in fact, the central issues of spatial econometrics and statistics, where their manifestations are called *spatial heterogeneity* and *spatial dependence*, respectively, and we have already argued extensively here and elsewhere that their presence is not merely "most common," but absolutely ubiquitous, across all of social science.

normal in probit). Since observations are spatially and/or temporally interdependent, log-likelihoods (or log-posteriors) of realized data (times priors) are single, non-separable logs of n-dimensional integrals rather than sums of logs of n unidimensional integrals as usual when assuming conditional independence of observations. The n-dimensional cumulative-normal of *S-Probit*, for instance, is exponentially more intense to compute than the non-spatial unidimensional ones, which prohibits direct optimization even in relatively small samples (see, e.g., Franzese & Hays 2009). Likewise, estimating spatial dynamics and effects, and their certainties, in terms of *QualDep* outcomes, rather than in parameter or latent-variable terms as is current practice, requires calculation of more n-dimensional integrals (see, e.g., Franzese & Hays 2009, Hays 2009). The large, current, and rapidly growing *S-QualDep* literature, with the notable exception of *S-Probit*,[9] orients mostly Bayesian, CAR (conditional autoregressive) rather than SAR (simultaneous autoregressive), and squarely within the spatial-statistical tradition. Our contribution is more toward the spatial-econometric analysis of *S-QualDep* models.[10] That is, we seek 1) estimators for the structural parameters of theoretically derived *S-QualDep* regression models, which will 2) allow estimates of the effects in terms of the outcomes of interest of counterfactual changes in explanatory variables, based on the estimated parameters of these models, and 3) that will also yield estimates of the uncertainty of these parameter and effect estimates. We seek estimates optimizing the usual desirable properties (unbiased, consistent, efficient, accurate standard-errors) in sample dimensions and research contexts typical of political/social science, of course, but we also prefer estimators be simple to implement, given that one of our core goals to spur and empower applied empirical researchers.

Besag (1974) was perhaps the first to offer an *S-Count* model with his introduction of the spatial-Poisson (*S-Poisson*) as one of his auto-models.[11] Besag's auto-models conquer the estimation challenges of *S-QualDep* by making assumptions that allow variables' joint

[9] Considerable development and applications of SAR models by Bayesian (MCMC by Metropolis-Hastings-within-Gibbs-sampling: LeSage 2000, LeSage & Pace 2009), frequentist (simulated-likelihood by recursive importance-sampling (RIS): Beron et al. 2003, Beron & Vijverberg 2004), and various GMM estimators (McMillan 1992) populate the *S-Probit* literature.

[10] We view the Bayesian-*versus*-frequentist aspect as more a practical than a philosophical issue, although sustaining both Bayesian and frequentist tools seems wisely prudent to us. Our aims depart more importantly from the literature in the other regards.

[11] Griffith & Haining (2006) give excellent summary of the history and recent developments in *S-Poisson*.

distributions to be expressed as products of their conditional distributions. That is, the key assumption is conditional independence; auto-models are conditional autoregressive (CAR), in other words, and, like all such, the required assumptions can be restrictive. They usually require symmetric spatial-interdependence structures (Anselin 2006), for starters. This may not be so constraining in many geographical contexts, since distance is symmetric, but it is manifestly implausible in many social-science applications: Belgium and the US affect each other equally, for instance? CAR models also usually require exclusively non-negative interdependence, although auto-Poisson requires non-positive (Besag 1974). Considering, e.g., Brueckner's (2003) demonstration that negative and positive externalities imply the former and the latter respectively, and that both positive and negative externalities occur commonly, either can obviously be highly confining.

Recent advances to redress the strictly non-positive interdependence limitation include Kaiser and Cressie's (1997) suggestion of Winsorizing counts (restricting them to finite sets of integers), which affords limited positive interdependence within an auto-Poisson CAR framework. The Winsorized model, however, retains the computationally intensity and the problematic symmetry restriction of the auto-Poisson CAR (Griffith 2003). Griffith (2002, 2003) proposes a spatial-filtering approach to simultaneous autoregressive (SAR) Poisson in which each unit's mean is a function of the respective element of the eigenvector of the inverse spatial-covariance matrix. This seems promising, although eigenvector calculation heightens estimation demands, and this is especially so for the counterfactual effect-estimates with certainty-estimates that we seek. Moreover, by design, spatial filtering tends to deter or debar attempts to distinguish the alternative sources of spatial association. "Spatial filtering seeks to transform a variable containing spatial dependence into one free of it by partitioning the original georeferenced attribute variable into two synthetic variates: a spatial filter variate capturing latent spatial dependency that otherwise would remain in the response residuals and a nonspatial variate that is free of spatial dependence" (Griffith 2006: 166). The approach also assumes spatial association to arise from spatial heterogeneity or spatial-error processes. Thus, problematically from the perspective of our aims, spatial-filtering and CAR approaches generally tend 1) to treat spatial dependence as nuisance, 2) to stress spatial-error or spatial-heterogeneity models as opposed to spatial-lag models, and 3) to treat all observed spatial association as arising by one undifferentiated source, characteristically dependence in spatial econometrics, common shocks or, more generally, spatial heterogeneity, in spatial statistics, and selection in network analysis. The available estimators were simply not designed to address our aims, neither to distinguish alternative mechanisms by which spatial association arises nor to answer the counterfactual questions of our interest. Given that and their computational intensity, we propose instead a simpler pair of estimation strategies, Nonlinear Least-Squares (NLLS) and Method-of-Moments (MoM), for *S-Count* models—*S-Poisson* to start, with extensions to alternative count models to follow in later work—that are so-designed. Comparison of these simpler modeling and estimation strategies to the Winsorizing and spatial-filter strategies are also slated for future work; for now, Section IV offers comparison to the even simpler strategies of common current practice, ignoring and omitting interdependence or including spatial lags of observed counts in standard count models and so treating them as exogenous.

III. PROPOSED ESTIMATORS: Nonlinear Least-Squares (NLLS) and Generalized Method-of-Moments (GMM)

The starting point for empirical count models is the Poisson regression model, wherein realizations of the dependent variable, y_i=0,1,2,…, are observed counts of events taken to follow a Poisson distribution with probabilities:

$$\Pr(Y_i = y_i) = \left(e^{-\lambda_i}\lambda_i^{y_i}\right)\Big/\left(y_i!\right) \quad \text{, with } y_i \in \{0,1,2,\ldots,\infty\} \qquad (1),$$

where λ_i is the instantaneous rate at which the event occurs. A typical regression model parameterizes $\boldsymbol{\lambda}$, which is also the conditional mean of $\mathbf{y}$, as log-linear function of $\mathbf{X}$: $\ln(\boldsymbol{\lambda})=\mathbf{X}\boldsymbol{\beta}$, with $\mathbf{X}$ an $n\times k$ matrix of explanators and $\boldsymbol{\beta}$ a $k\times1$ vector of coefficients, like so:

$$\begin{aligned}\Pr(Y_i = y_i \mid \mathbf{x}_i) \equiv f(y_i \mid \mathbf{x}_i) = \left(e^{-\lambda_i}\lambda_i^{y_i}\right)\Big/\left(y_i!\right),\\ \text{with } E\left(y_i \mid \mathbf{x}_i\right) = \lambda_i = \exp\left(\mathbf{x}_i'\boldsymbol{\beta}\right)\end{aligned} \qquad (2).$$

The log-linearization guarantees $\lambda_i>0 \ \forall \ \mathbf{x}_i$ and $\boldsymbol{\beta}$. The standard approach is to estimate the model in (2) via maximum likelihood, with the log-likelihood implied by (1)-(2) being:

$$\ln L\left(\mathbf{y}\right) = \sum_i \left\{-\lambda_i + y_i\mathbf{x}_i'\boldsymbol{\beta} - \ln\left(y_i!\right)\right\} \qquad (3).$$

We are interested in more general models, and especially ones which allow the rates of event-occurrence to be dependent across observations. Since the Poisson distribution cannot easily accommodate this, our strategy is to abandon the explicitly Poisson form of (1) and the first line of (2), but retain the conditional-mean specification in the second line of (2). That conditional-mean equation implicitly defines a regression model like this:

$$y_i = \lambda_i + u_i = \exp(\mathbf{x}_i'\boldsymbol{\beta}) + u_i \tag{4},$$

where the usual sort of exogenous-regressors condition, E(u_i|$\mathbf{x}_i$)=0, assures unbiasedness. Disturbances in count data are inherently heteroschedastic because variances at high counts will exceed that at low. For one, counts are bounded below at, and so variation is truncated near, zero. I.e., typically, the mean and variance of count distributions are related. In the Poisson model, they are equal; the Binomial has mean np and variance np(1-p); etc.

We begin exploration of interdependent counts applying a model with additive errors:[12]

$$\begin{gathered}\mathbf{y} = \boldsymbol{\lambda} + \mathbf{u}, \text{ where } \boldsymbol{\lambda} \equiv E\left(\mathbf{y} \mid \mathbf{X}\right) = \exp\left[\left(I - \rho\mathbf{W}\right)^{-1}\mathbf{X}\boldsymbol{\beta}\right] \Leftrightarrow \\ \ln\boldsymbol{\lambda} = \left(I - \rho\mathbf{W}\right)^{-1}\mathbf{X}\boldsymbol{\beta} \Leftrightarrow \ln\boldsymbol{\lambda} = \rho\mathbf{W}\left(\ln\boldsymbol{\lambda}\right) + \mathbf{X}\boldsymbol{\beta}\end{gathered} \tag{5}.$$

$\mathbf{y}$ is an n×1 vector of observed counts; $\boldsymbol{\lambda}$ is the n×1 vector of their conditional means; $\mathbf{u}$ is an n×1 vector of additively separated stochastic-error terms, assumed *independently*, although heteroskedastically, distributed; $\mathbf{W}$ is an n×n spatial-weights or connectivity matrix with element $\{i,j\}$ reflecting the relative connectivity of unit j to i, and ρ the strength of interdependence by that pattern given in $\mathbf{W}$. The alternative expressions of the model in the second line of (5) help clarify its structure as a spatial log-linear regression, which naturally motivates the non-linear least-squares (NLLS) estimator to be introduced next. It also helps clarify why, as discussed in the next section, the better proxy for spatially lagged $\boldsymbol{\lambda}$ in a naïve (linear) least-squares estimator is the spatial lag of *logged* observed counts, rather than the currently common practice of spatially lagged observed counts.

We consider two estimators. First, define a parameter vector $\boldsymbol{\theta}$≡[$\boldsymbol{\beta}$ ρ]. The non-linear least-squares (NLLS) estimator for $\boldsymbol{\theta}$ simply minimizes the sum of squared errors $\mathbf{u}'\mathbf{u}$:

$$\underset{\boldsymbol{\theta}}{\text{Min}}\ \mathbf{u}'\mathbf{u} = \left[\mathbf{y} - \boldsymbol{\lambda}\right]'\left[\mathbf{y} - \boldsymbol{\lambda}\right] = \left[\mathbf{y} - e^{(I-\rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}}\right]'\left[\mathbf{y} - e^{(I-\rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}}\right] \tag{6}.$$

The least squares estimator is, as always:

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{NLLS} &= \left[\left(\frac{\partial E(\mathbf{y} \mid \mathbf{X})}{\partial\tilde{\boldsymbol{\theta}}}\right)'\left(\frac{\partial E(\mathbf{y} \mid \mathbf{X})}{\partial\tilde{\boldsymbol{\theta}}}\right)\right]^{-1}\left(\frac{\partial E(\mathbf{y} \mid \mathbf{X})}{\partial\tilde{\boldsymbol{\theta}}}\right)'\mathbf{y} \\ &= \left[\left(\frac{\partial\boldsymbol{\lambda}}{\partial\tilde{\boldsymbol{\theta}}}\right)'\left(\frac{\partial\boldsymbol{\lambda}}{\partial\tilde{\boldsymbol{\theta}}}\right)\right]^{-1}\left(\frac{\partial\boldsymbol{\lambda}}{\partial\tilde{\boldsymbol{\theta}}}\right)'\mathbf{y}\end{aligned} \tag{7}.$$

[12] The *S-Count* literature contains much discussion and debate regarding the appropriateness of additive or multiplicative error-forms. We prefer to start in the simpler additive form but extension to the multiplicative form, which will introduce some *S-QualDep* complications the additive form avoids, will be very important.

In the linear-regression case, $E(\mathbf{y}|\mathbf{X})=\mathbf{X\beta}$, so the gradient with respect to $\boldsymbol{\beta}$ is just $\mathbf{X}$, making the last line of (7) the very familiar $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$. In the nonlinear case, this last line involves $\boldsymbol{\theta}$ on both sides and cannot be reduced for closed-form solution for $\hat{\boldsymbol{\theta}}_{NLLS}$. Numeric optimization of (6) or calculation of (7) is quite possible though.

The variance-covariance matrix of $\hat{\boldsymbol{\theta}}_{NLLS}$ also follows the standard least-squares form:

$$Var\left(\hat{\boldsymbol{\theta}}_{NLLS}\right)=\left[\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)'\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)\right]^{-1}\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)'\boldsymbol{\Omega}\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)\left[\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)'\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)\right]^{-1}\text{, where }\boldsymbol{\Omega}=Var(\mathbf{u}) \qquad (8).$$

In the homoskedastic and independent $\boldsymbol{\Omega}=\sigma^2\mathbf{I}$ case, we could pull the scalar σ^2 out of the middle and this would reduce to an even more familiar $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$–like form. As already noted, count data will be heteroskedastic, so $\boldsymbol{\Omega}$ diagonal with diagonal elements σ_i^2 is more appropriate. Paralleling what is true for linear least-squares, NLLS would be BUE for this model with homoskedasticty; with the surely heteroskedastic structure, it is unbiased but inefficient, with estimated variance-covariances consistent but neither efficient nor generally unbiased. A generalized (weighted) NLLS estimator would be BUE if we knew $\boldsymbol{\Omega}$:

$$\hat{\boldsymbol{\theta}}_{GNLLS}=\left[\left(\frac{\partial\boldsymbol{\lambda}}{\partial\tilde{\boldsymbol{\theta}}}\right)'\boldsymbol{\Omega}^{-1}\left(\frac{\partial\boldsymbol{\lambda}}{\partial\tilde{\boldsymbol{\theta}}}\right)\right]^{-1}\left(\frac{\partial\boldsymbol{\lambda}}{\partial\tilde{\boldsymbol{\theta}}}\right)'\boldsymbol{\Omega}^{-1}\mathbf{y}\text{ , where }\boldsymbol{\Omega}\equiv Var(\mathbf{u}) \qquad (9).$$

$$V\left(\hat{\boldsymbol{\theta}}_{GNLLS}\right)=\left[\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)'\boldsymbol{\Omega}^{-1}\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)\right]^{-1} \qquad (10).$$

As usual, if we do not know, but can consistently estimate (NLLS would do), the feasible generalized (weighted) NLLS estimator would be asymptotically BUE:[13]

$$\hat{\boldsymbol{\theta}}_{FGNLLS}=\left[\left(\frac{\partial\boldsymbol{\lambda}}{\partial\tilde{\boldsymbol{\theta}}}\right)'\hat{\boldsymbol{\Omega}}^{-1}\left(\frac{\partial\boldsymbol{\lambda}}{\partial\tilde{\boldsymbol{\theta}}}\right)\right]^{-1}\left(\frac{\partial\boldsymbol{\lambda}}{\partial\tilde{\boldsymbol{\theta}}}\right)'\hat{\boldsymbol{\Omega}}^{-1}\mathbf{y}\text{ , where }\hat{\boldsymbol{\Omega}}\equiv\widehat{\mathrm{var}}(\hat{\mathbf{u}}) \qquad (11).$$

$$\widehat{\mathrm{var}}\left(\hat{\boldsymbol{\theta}}_{FGNLLS}\right)=\left[\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)'\hat{\boldsymbol{\Omega}}^{-1}\left(\frac{\partial\boldsymbol{\lambda}}{\partial\hat{\boldsymbol{\theta}}}\right)\right]^{-1} \qquad (12).$$

The gradients of $\boldsymbol{\lambda}=E(\mathbf{y}|\mathbf{X})$ with respect to the parameters, $\boldsymbol{\theta}$, used in the NLLS estimator defined in (7) and (12), are given by horizontal concatenation of these gradients:

$$\frac{\partial\boldsymbol{\lambda}}{\partial\boldsymbol{\beta}}=\mathrm{vec}_k\left\{\frac{\partial\boldsymbol{\lambda}}{\partial\beta_k}\right\}=\mathrm{vec}_k\left\{\boldsymbol{\lambda}\odot\left[(\mathbf{I}-\rho\mathbf{W})^{-1}\mathbf{x}_k\right]\right\}=\left[\boldsymbol{\lambda}\odot(\mathbf{I}-\rho\mathbf{W})^{-1}\right]\mathbf{X} \qquad (13),$$

[13] In this version, unfortunately, we erroneously mix NLLS coefficient estimator, (7), and FGNLLS variance-covariance estimator, (12), in the Monte Carlos. Thus, our coefficient estimates are consistent but inefficient and our standard errors inaccurate (apparently from the results, not too badly so and not asymptotically).

$$\frac{\partial \boldsymbol{\lambda}}{\partial \rho} = \boldsymbol{\lambda} \odot \left[(\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{W} (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \boldsymbol{\beta} \right] \quad (14),$$

where the $\odot$ between $\boldsymbol{\lambda}$ and the bracketed vectors in (13) and (14) indicates element-by-element (i.e., Hadamard) multiplication, with the Hadamard product operating on each of the k columns in the $(n\times1)\odot(n\times k)$ expression ending (13). The matrix-functions $vec_k\{\mathbf{a}_k\}$ and $diag\{\mathbf{b}\}$ are, respectively, the $n\times k$ horizontal concatenation of the k $n\times1$ vectors $\mathbf{a}_k$ and the $k\times k$ diagonal matrix formed with the $k\times1$ vector $\mathbf{b}$ on the diagonal and zeros elsewhere. Horizontal concatenation of the $n\times k$ matrix in (13) and the $n\times1$ vector in (14) then yields the aforementioned $n\times(k+1)$ matrix of gradients:

$$\frac{\partial \boldsymbol{\lambda}}{\partial \hat{\boldsymbol{\theta}}} = \left[\left\{ \boldsymbol{\lambda} \odot (\mathbf{I} - \rho \mathbf{W})^{-1} \right\} \mathbf{X} \qquad \boldsymbol{\lambda} \odot \left\{ (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{W} (\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \boldsymbol{\beta} \right\} \right] \quad (15).$$

A closely related, nearly as simple generalized methods-of-moments GMM estimator is:

$$E\left[\mathbf{g}_i (y_i, \mathbf{z}_i, \boldsymbol{\theta}) \right] = \mathbf{0}, \text{ where } \mathbf{z}_i = \left[\mathbf{x}_i \quad [\mathbf{W}\mathbf{X}]_i \right] \quad (16),$$

with $\mathbf{g}_i$ an $r\times1$ vector (of moment conditions). The vector $\mathbf{z}_i$ contains exogenous variables, including spatial instruments. When $r=k$, where k is the number of parameters to estimate, then the population moment/orthogonality conditions are

$$\sum_{i=1}^{n} \mathbf{g}_i (y_i, \mathbf{z}_i, \boldsymbol{\theta}) = \mathbf{0} \quad (17).$$

More specifically, we exploit the orthogonality of the variables in $\mathbf{z}$ are with $\mathbf{u}$ to derive the sample conditions:

$$\sum_{i=1}^{n} \left(y_i - \exp\left[(\mathbf{I} - \rho \mathbf{W})^{-1} \mathbf{X} \boldsymbol{\beta} \right]_i \right) \mathbf{z}_i = \mathbf{0} \quad (18).$$

In this "just-identified" case when $r=k$, a unique solution to (18) exists, yielding the GMM estimator $\hat{\boldsymbol{\theta}}_{GMM}$, which is asymptotically normally distributed and root-n consistent for $\boldsymbol{\theta}$:

$$\sqrt{n} \left(\hat{\boldsymbol{\theta}}_{GMM} - \boldsymbol{\theta} \right) \xrightarrow{d} \mathbf{N}\left[\mathbf{0}, \mathbf{V} \right] \quad (19).$$

A robust sandwich estimator for the variance-covariance matrix $\mathbf{V}$ in (19) is then:

$$\widehat{\mathrm{var}} \left(\hat{\boldsymbol{\theta}}_{GMM} \right) = \hat{\mathbf{A}}^{-1} \hat{\mathbf{B}} \left(\hat{\mathbf{A}}' \right)^{-1} \quad (20),$$

$$\text{where } \hat{\mathbf{A}} = \frac{1}{n} \sum_{i=1}^{n} \left. \frac{\partial \mathbf{g}_i (y_i, \mathbf{z}_i, \boldsymbol{\theta})}{\partial \hat{\boldsymbol{\theta}}'} \right|_{\hat{\boldsymbol{\theta}}_{GMM}} \quad (21),$$

$$\text{and } \hat{\mathbf{B}} = \frac{1}{n} \sum_{i=1}^{n} \left\{ \left[\mathbf{g}_i (y_i, \mathbf{z}_i, \hat{\boldsymbol{\theta}}_{GMM}) \right] \left[\mathbf{g}_i (y_i, \mathbf{z}_i, \hat{\boldsymbol{\theta}}_{GMM}) \right]' \right\} \quad (22).$$

Note that the standard error for $\hat{\theta}_i$ is calculated as $\sqrt{\hat{\mathbf{V}}_{ii} / n}$.

This moment estimator in the $r{=}k$ case corresponds to the just-identified case of indirect least-squares in the linear-regression context. When $r{>}k$, the parameters are overidentified. A unique solution is obtained by minimizing the quadratic form:

$$\hat{\boldsymbol{\theta}}_{GMM} = \text{Argmin}_{\tilde{\boldsymbol{\theta}}}\left[\mathbf{g}_i(y_i,\mathbf{z}_i,\tilde{\boldsymbol{\theta}})\right]'\mathbf{Q}\left[\mathbf{g}_i(y_i,\mathbf{z}_i,\tilde{\boldsymbol{\theta}})\right] \qquad (23),$$

where $\mathbf{Q}$ is an $r{\times}r$ symmetric positive-definite matrix of weights that tune the minimization. Note for intuition-building analogy that, if $\mathbf{Q}$ is the identity, the minimized quadratic sum is unweighted, so the estimator corresponds to two-stage least-squares (2SLS) in the linear-regression context. With $\mathbf{Q}$ equal the (estimated) inverse of the variance-covariance matrix for $\mathbf{y}|\mathbf{X}$, the estimator corresponds to the linear-regression GMM, which is just a weighted 2SLS. Indeed, the analogy *FGLS is to OLS as GMM is to 2SLS* helps build much useful intuition. The variance-covariance estimate for this overidentified case is

$$\text{var}\left(\hat{\boldsymbol{\theta}}_{GMM}\right) = \left(\hat{\mathbf{A}}\mathbf{Q}\hat{\mathbf{A}}\right)^{-1}\left(\hat{\mathbf{A}}\mathbf{Q}\hat{\mathbf{B}}\mathbf{Q}\hat{\mathbf{A}}\right)\left(\hat{\mathbf{A}}\mathbf{Q}\hat{\mathbf{A}}\right)^{-1} \qquad (24)$$

Optimal GMM uses the weights matrix $\mathbf{Q}$ that maximizes the estimator's efficiency. This is done in two steps. Typically, in the first step, $\mathbf{Q}$ is set to $\mathbf{I}$, and $\boldsymbol{\theta}$ is estimated using (23), i.e., by the equivalent of 2SLS, and $\mathbf{B}$ is estimated using (22). The optimal GMM estimates for $\boldsymbol{\theta}$ then minimizes:

$$\hat{\boldsymbol{\theta}}^{opt}_{GMM} = \text{Argmin}_{\tilde{\boldsymbol{\theta}}}\left[\mathbf{g}_i(y_i,\mathbf{z}_i,\tilde{\boldsymbol{\theta}})\right]'\hat{\mathbf{B}}^{-1}\left[\mathbf{g}_i(y_i,\mathbf{z}_i,\tilde{\boldsymbol{\theta}})\right] \qquad (25).$$

Restating all this by analogy may aid the intuition. The first stage is estimated consistently by unweighted instrumentation to provide a consistent estimate of the variance-covariance of $\mathbf{y}|\mathbf{X}$. Weighting by the inverse of that consistently estimated variance-covariance matrix then yields asymptotically efficient estimates of the parameters. The variance-covariance estimates for these estimated parameters, finally, are given by:

$$\hat{\boldsymbol{\Sigma}}^{opt}_{GMM} = \widehat{\text{var}}\left(\hat{\boldsymbol{\theta}}^{opt}_{GMM}\right) = \left(\hat{\mathbf{A}}'\hat{\mathbf{B}}^{-1}\hat{\mathbf{A}}\right)^{-1} \qquad (26).$$

IV. Comparing Estimators: Monte Carlo Analyses of Naïve and Consistent Estimators

We explore the small-sample properties of the two consistent estimators described above (NLLS and GMM) as well as the two naïve estimators of current common practice. We call these two estimators of current common practice, i.e., the one omitting spatial dependence altogether and the one that includes spatial lags but assumes them exogenous, the *non-spatial* (NS) and the *spatially-lagged observed-counts* (SLOC) estimators, respectively. The former is just a standard maximum-likelihood estimation of a Poisson count-regression model; the latter also applies standard maximum-likelihood estimation of a Poisson count-regression model, but including spatially weighted averages of observed counts among the regressors, i.e., in the conditional mean specification. We evaluate two versions of the SLOC estimator. One uses $\mathbf{Wy}$ for this spatial lag, the other uses $\mathbf{W}(\ln(\mathbf{y}+\mathbf{1}))$. The former is what extant literature typically applies, but the latter as easily implements something closer to the log-linear specification of the model's conditional mean, $\mathrm{E}(\mathbf{y}|\mathbf{X})$, and so should perform better as imperfect but expedient alternative to consistent alternatives, our NLLS or GMM or the spatial-filtering or Winsorized alternatives suggested elsewhere.

Our data-generating process (DGP) is the *S-Poisson* model described above. For $\mathbf{W}$, we use a row-standardized binary-contiguity matrix for the 48 contiguous U.S. states. We set β to 0.6, and consider both moderate (Table X) and high spatial interdependence (Table X), ρ=0.3 and ρ=0.5 respectively, for a smaller and a larger sample-size, n={96,240}. To create the 96×96 and 240×240 $\mathbf{W}$'s we Kronecker product the original 48×48 $\mathbf{W}$ with 2×2 and 5×5 identity matrices, $\mathbf{I}_2$ and $\mathbf{I}_5$. This could reflect, for example, multiple observations (2 or 5) of event counts in a panel (though an oddly temporally independent one) of the 48 states.

The NS estimator will of course suffer omitted-variable bias. In typical omitted-variable fashion, the estimated direct effect of x_i on the conditional mean of y_i will include indirect effects from spatial feedback. With reinforcing (e.g., all-positive) interdependence this will imply an inflationary bias in $\hat{\beta}_{NS}$, as we have shown before in other model-contexts (e.g., Franzese & Hays 2007c, 2008b). The estimators employing SLOC are biased too; however, in this DGP, unlike in the spatial-lag models we have considered previously (e.g., Franzese & Hays 2004a, 2006b, 2007cd, 2008ab), the spatially lagged observed counts do not covary with the residuals. They are exogenous, so these estimates do not, as they did before, suffer simultaneity biases, which are typically (i.e., in the reinforcing feedback case) inflationary. Rather, these estimators suffer measurement-error-induced attenuation-bias. Simply: as seen most clearly from the conditional-mean specification in second line of (2), interdependence in this model is in the conditional mean, not in the actual counts, and $\mathbf{Wy}$ is poor proxy for $\mathbf{W}[E(\mathbf{y}|\mathbf{X})]$, i.e., for $\mathbf{W}[\ln(\boldsymbol{\lambda})]$.We propose that spatial lag $\mathbf{W}[\ln(\mathbf{y}+\mathbf{1})]$ provides better proxy. We will call the SLOC estimator using $\mathbf{Wy}$, SLOC_1, and that using $\mathbf{W}[\ln(\mathbf{y}+\mathbf{1})]$, SLOC_2.

Table 1: Estimator Comparison for S-*Poisson* Model, Moderate Interdependence

Monte Carlo Results (500 Trials) for $\mathbf{y}=\boldsymbol{\lambda}+\mathbf{u}$, where $\boldsymbol{\lambda}=\exp[(\mathbf{I}-\rho\mathbf{W})^{-1}\beta\mathbf{x}]$, with $\beta=.6$, $\rho=.3$, and $n=\{96,240\}$

Parameter	***Result***	***NS***	***SLOC$_1$***	***SLOC$_2$***	***NLLS***	***GMM***
		***n*=96 / 240**	***n*=96 / 240**	***n*=96 / 240**	***n*=96 / 240**	***n*=96 / 240**
β=.6	mean	.794 / .794	.664 / .666	.615 / .613	.597 / .600	.596 / .599
	s.d.	.016 / .010	.044 / .026	.052 / .032	.056 / .034	.054 / .032
	rmse	.194 / .194	.078 / .071	.054 / .035	.056 / .034	.054 / .032
	mean s.e.	.016 / .010	.038 / .024	.047 / .030	.052 / .033	.052 / .033
	overconfidence	.990 / .935	1.157 / 1.081	1.101 / 1.109	1.077 / 1.025	1.035 / .965
ρ=.3	mean	*n/a*	.054 / .054	.260 / .263	.302 / .299	.304 / .300
	s.d.	*n/a*	.016 / .010	.066 / .042	.075 / .046	.072 / .044
	rmse	*n/a*	.246 / .246	.077 / .056	.075 / .046	.072 / .044
	mean s.e.	*n/a*	.014 / .009	.063 / .039	.069 / .044	.068 / .043
	overconfidence	*n/a*	1.136 / 1.080	1.051 / 1.065	1.087 / 1.054	1.059 / 1.009

Table 1 reports the results of 500 trials using the five estimators (NS, SLOC$_1$, SLOC$_2$, NLLS, GMM) for the moderate interdependence (ρ=.3) case. The NS estimator exhibits an upward bias as expected, approximately 30% of the true value. It gives all of the effect $\partial\lambda_i/\partial x_i$, both direct and indirect, to β. The SLOC$_1$ estimator also overestimates β, though by less than the non-spatial estimator. It divides the total effect between β and ρ, but because of the measurement error (and misspecification), it gives too much to β and too little to ρ. Alternatively, we could say the measurement error (*cum* misspecification) causes underestimation of ρ, which induces overestimation of β. The magnitudes of these biases are huge and appreciable, respectively: about -.25 or -80% for $\hat{\rho}_{SLOC_1}$ and +.065 or +11% for $\hat{\beta}_{SLOC_1}$. The implied post-feedback total effects, which may be proxied by $\hat{\beta}/(1-\hat{\rho})$,[14] are poorly estimated too: about .70 *versus* a correctly approximate .85. The standard-error estimates are also, unsurprisingly, overconfident. Specifically, standard deviations of the sampling distributions for $\hat{\beta}_{SLOC_1}$ and $\hat{\rho}_{SLOC_1}$ are respectively 15.7% and 13.6% larger than the average of the estimated standard errors in the small sample and 8.1% and 8.0% larger respectively in the larger sample. The SLOC$_2$ estimator that uses observed *log* counts does

[14] Properly, these will depend on $\mathbf{W}$ also, but the proxy serves well enough for qualitative comparison.

much better. In the case of ρ, the parameter which SLOC_1 estimates especially poorly, the bias drops to about -.04 or -13.3% and the root-mean-squared-errors drop from .246 for SLOC_1 to .077 for SLOC_2 in the smaller sample and from .246 to .056 in the larger. The post-feedback total effects are much better estimated also: at about .83 in either sample. The standard errors from the SLOC_2 estimator are more accurate too: overconfidence ranging from 5% to 11% across parameters and sample sizes. Meanwhile, the consistent estimators perform a bit better still. Even in the smaller sample, both parameter estimates are essentially unbiased: a worst case of about -1.3% for GMM. GMM, conversely, performs ever-so-slightly better in RMSE terms. The approximate total-effect estimates are spot on the approximately .85 truth in all cases. Standard-error estimates are reasonably accurate for both as well, although overconfidence does range from -3.5% to a noticeable +8.7%.

Table 2: Estimator Comparison for S-Poisson Model, Strong Interdependence

Monte Carlo Results (500 Trials) for $\mathbf{y}=\boldsymbol{\lambda}+\mathbf{u}$, where $\boldsymbol{\lambda}=\exp[(\mathbf{I}\text{-}\rho\mathbf{W})^{-1}\beta\mathbf{x}]$, with β=.6, ρ=.5, and n={96,240}

Parameter	***Result***	***NS***	***SLOC$_1$***	***SLOC$_2$***	***NLLS***	***GMM***
		***n*=96 / 240**	***n*=96 / 240**	***n*=96 / 240**	***n*=96 / 240**	***n*=96 / 240**
β=.6	mean	1.052 / 1.053	.762 / .764	.610 / .611	.598 / .598	.598 / .599
	s.d.	.011 / .007	.036 / .023	.041 / .027	.042 / .027	.043 / .028
	rmse	.453 / .453	.166 / .166	.042 / .029	.042 / .027	.043 / .029
	mean s.e.	.011 / .007	.029 / .018	.036 / .022	.041 / .026	.041 / .027
	overconfidence	.970 / .963	1.262 / 1.301	1.154 / 1.205	1.035 / 1.06	1.028 / 1.03
ρ=.5	mean	*n/a*	.059 / .059	.478 / .477	.501 / .501	.501 / .501
	s.d.	*n/a*	.006 / .004	.037 / .024	.038 / .025	.039 / .025
	rmse	*n/a*	.441 / .441	.043 / .033	.038 / .025	.039 / .025
	mean s.e.	*n/a*	.005 / .003	.035 / .022	.037 / .023	.036 / .024
	overconfidence	*n/a*	1.2455 / 1.288	1.051 / 1.095	1.03 / 1.057	1.059 / 1.056

Table 2 reports the Monte Carlo results for the strong-interdependence (ρ=.5) case. The same patterns as found under moderate interdependence are mostly repeated, although the performances of the NS and SLOC estimators both deteriorate, as seen by notable increases in their root-mean-squared-errors. The SLOC_2 estimator continues to overestimate β and underestimate ρ, but the inaccuracies remain small (about +1.75% and -4.5%, respectively) and even smaller for the implied approximate total effects (1.1684 *versus* 1.2, or -.0316 or

-2.6%). In fact, the $SLOC_2$ estimator's performance improves as interdependence-strength increases. This is because the measurement error for $\ln(\mathbf{y}+\mathbf{1})$ *versus* $\ln(\boldsymbol{\lambda})$ declines with the strength of interdependence, given our particular DGP: with positive ρ and $\mathbf{y}$, we expect fewer observed zeros on average in the generated data as the λ_i will be larger. The NLLS and GMM estimators continue to perform best, with NLLS now slightly outperforming GMM under the stronger interdependence, although the differences remain small.

The improvements of NLLS and GMM over $SLOC_2$, though, are not dramatic in either the moderate- or the strong-interdependence case and, in fact, the advantages of the former seem may decline with the strength of interdependence. However, we strongly suspect that the consideration of multiplicative-error formulations of *S-Count* models (see note 12) to come in future work will show that formulation magnifies these advantages of the simple consistent estimators NLLS and GMM over the even simpler naïve expedient $SLOC_2$ because in the multiplicative formulation, unlike in the additive one, the SLOC estimators would also suffer simultaneity biases. That remains conjecture for now, however.

V. <u>Empirical Application</u>: Counts of Terrorist Events around the World

For an empirical application, we use data from Plümper & Neumayer (2008), which in turn derive from the Mickolus et al. (2003) ITERATE dataset, to assemble counts of post-9/11 terrorist events in 162 countries of the world.[15] We created a strict territorial-contiguity weights-matrix for these countries using the Correlates of War dataset (COW: Stinnet et al. 2002). There are 26 isolates among these 162 countries, and we have data on all the covariates for 139 of them. To begin, we calculate Moran's I, both local and global versions, to consider the evidence of generic clustering (spatial clustering along contiguity lines). In the 162-country sample, we find I=0.096. In this size sample, with this $\mathbf{W}$, global Moran's I has an expected value and standard deviation under the null of independence of -0.006 and 0.035, respectively. Using the asymptotic-normal approximation, the resulting "t-statistic" of 2.924 would have two-tailed significance of 0.003. I.e., we very easily reject the null of no spatial association along contiguity lines. The local Moran's I's reveal that these results are mainly driven by relations among event-counts within the Middle East and in other familiar hot-spots. Countries with significantly positive local association, I_i, (at the

[15] Important recent studies on the theoretical and empirical determinants of terrorist events include Blomberg et al. 2004; Braithwaite & Li 2007; Bird, Blomberg, & Hess 2008; Neumayer & Plümper 2008.

.05 level by a normal-approximation analogous to that for global I), indicating pockets of positively correlated event-counts in the local neighborhood (hot-or cold-spots near other hot- or cold-spots), are Afghanistan, Jordan, Kuwait, Pakistan, Saudi Arabia, and Turkey. Only Syria has significantly negative local association, I_i, although others on the negative side, indicating cold- near hot-spots or hot- near cold-spots, include countries in Western Europe and North Africa, Canada and the U.S., and Russia and the Baltics, for instance. Iraq, Afghanistan, and Pakistan suffered the largest numbers of events, with 203, 42, and 39 respectively. Slightly more than half the sample, 85 countries, experienced no incidents.

Substantively, our core interest is whether terrorist activity exhibits a geographically-based interdependence around the world. Do exogenous increases in the expected number of terrorist events in one country affect the expected number of such events in its neighbors, for instance? Evidence of spatial clustering alone does not answer this question. Important political, social, and economic determinants of terrorism—e.g., regime-type and economic development—cluster spatially, so neighbors are exposed to correlated stimuli that could explain the geographical patterns that we just noted. To understand better the role geography plays in terrorist activity, and in particular to evaluate whether terrorist events truly exhibit spatial interdependence and not merely association, we must specify a spatial-lag count-model of terrorist events that addresses *Galton's Problem*, i.e., that controls well for these alternative *common-exposure* explanations of event counts. Nearly all previous studies recognize some role for geography, but few use models that appropriately specify these alternatives with spatial lags and controls for possibly clustered exogenous internal and external conditions, and even fewer use estimators appropriate for such specifications.

Our empirical models include six explanatory variables derived from theory. In one strong contending theory, Blomberg et al. (2004) argue that terrorism stems from feelings of deprivation or relative deprivation and a desire to alter the *status quo* combined with the absence of democratic mechanisms to promote change. By this logic, terrorists will target elites within poor non-democratic countries or wealthy nations abroad. Accordingly, we include several economic variables—real GDP per capita, trade openness, and growth in real GDP per capita—to capture these propositions,[16] with all three variables taken from the Penn World Table (PWT: Heston et al. 2006), and the country's democracy score from

[16] In future refinements, we should interact the (relative-)deprivation measures with the democracy score to reflect better the conditional proposition of Blomberg et al.'s (2004) argument.

Polity IV (Marshall et al. 2008). To these four, we add a dummy variable, labeled *Conflict*, indicating whether the country has been involved in either an interstate or civil conflict this decade. Conflict data for post-9/11 era are not yet available in COW and so are taken from Lee (2009). We also control country size by including the country's population (PWT).

Finally, we move in our empirical application to a negative-binomial model-specification for the likelihood-based estimators, NS, $SLOC_1$, and $SLOC_2$, because preliminary estimates of Poisson specifications overwhelmingly indicated overdispersion (positive dependence of events within the counts). The negative-binomial distribution relaxes the equality of mean and variance *cum* zero-dependence of events within counts implied by the Poisson. The S-GMM and S-NLLS estimators, on the other hand, make no specific assumptions about the probability-functional form for the count process. Their assumptions include only a specific conditional-mean function and the non-covariance of that function with the stochastic term. Unlike Poisson models, then, S-GMM and S-NLLS do not impose equality of the mean and variance of the underlying process. Accordingly, we label the models estimated by S-GMM and S-NLLS in Table 3 generically *Spatial Count* rather than *Negative Binomial* models.

The last four columns of Table 3 reports four sets of model-estimates, one each for the non-spatial model and NS estimator, the naïve-spatial $SLOC_1$ (spatially lagged observed-count) model and estimator, the naïve-spatial $SLOC_2$ (spatially lagged log observed-count-plus-one) model and estimator, and the spatial-lag model and S-GMM estimator.[17]

[17] We have not yet programmed the S-NLLS estimator, although it will not be difficult and we have no reason to expect its estimates to differ appreciably from those of the S-GMM estimator.

Table 3: Spatial Count Models of Terrorist Incidents (Post-9/11)

	Non-Spatial Model	**Naïve-Spatial Model 1**	**Naïve-Spatial Model 2**	**Spatial Model**
Population	2.110	2.550	3.042	-.230
	(3.046)	(2.922)	(3.029)	(.938)
Real GDP per capita	.052**	.044**	.026	.021**
	(.023)	(.022)	(.022)	(.010)
Trade Openness	-.281	-.277	-.435	-.396
	(.361)	(.329)	(.330)	(.276)
Growth in Real GDP per capita	-.001	.005	.013	-.029
	(.023)	(.021)	(.022)	(.057)
Democracy Score	-.067	-.023	-.019	-.016
	(.054)	(.054)	(.052)	(.061)
Conflict	1.967***	2.072***	1.551***	1.220**
	(.514)	(.492)	(.491)	(.567)
Spatial Lag		.054**	.786**	.893**
		(.022)	(.234)	(.440)
Observations	*139*	*139*	*139*	*139*
Model	*Negative Binomial*	*Negative Binomial*	*Negative Binomial*	*Spatial Count*
Estimator	*NS*	$SLOC_1$	$SLOC_2$	*S-GMM*
Log-Likelihood	*-282.186*	*-277.510*	*-276.434*	*n/a*

Notes: Standard errors in parentheses. The spatial lags are generated with a row-normalized binary-contiguity weighting matrix, using shared territorial borders as the criterion. ***, **, * significant at 1%, 5%, 10%.

Across the models and estimators, we find the key internal explanatory factors to be a country's real GDP *per capita* and its participation in an interstate or civil conflict. The spatial model-estimates all also indicate significant positive interdependence, now better distinguished from mere association insofar as the rest of those models sufficiently control exogenous internal and external conditions that may correlate spatially. Comparing the model-estimates, we notice first that, as we would expect given the results of our Monte Carlo simulations, the non-spatial model-estimates accord generally greater direct role to the internal explanatory factors, as seen in the generally higher absolute-value coefficients, than do the properly spatial model-estimates. Next, we note that, again as our simulations would lead us to expect, the naïve-spatial $SLOC_1$ (spatially lagged observed-count) model and estimator suggests dramatically weaker interdependence in terrorist event-counts than either the better-proxied naïve-spatial $SLOC_2$ (spatially lagged log observed-count-plus-one) model and estimator or the better-specified and S-GMM consistently estimated spatial-lag model. The much smaller and very likely underestimated of interdependence relative to the latter two model-estimates induces, in turn, generally larger and very likely overestimation of the roles of the non-spatial regressors' roles. Indeed, the role of *conflict* seems perhaps marginally greater by the SLOC_1 model-estimates than it would seem from the non-spatial model-estimates. With the better-proxied naïve-spatial SLOC_2 model-estimates, however, we see greater log-likelihood, much greater positive interdependence, and generally smaller direct role for domestic conditions, especially for economic development and *conflict*, which have emerged as the two key factors. Indeed, economic development is not statistically significant at conventional levels by these model-estimates. Finally, and once more as our simulations would lead us to expect, the better-specified and consistent S-GMM model-estimates suggest stronger still interdependence in terrorist event-counts, and lesser still direct roles for domestic factors, especially *conflict* again.

Intriguingly, and in these regards not well foreshadowed by our simulations, S-GMM seems to offer notable efficiency gains over, or at least to allocate its coefficient-estimation efficiency across regressors quite differently (and, at least in this case, more fortuitously) than, the well-proxied SLOC_2 model-estimates. In particular, the coefficients on population, real GDP *per capita*, and trade openness are smaller but also much more precisely estimated. Population, for which the other model-estimates had reported sizable but highly imprecise coefficient estimates, is now seen with much greater precision to have much smaller coefficient. The estimated direct role of real GDP *per capita*, though smaller still than that from the SLOC_2 model-estimates, is much more precisely estimated and once again comfortably significant. Likewise, the direct role of trade openness is a bit smaller by these estimates than those of SLOC_2, though larger than those of NS or SLOC_1, but it approaches significance in the S-GMM model-estimates only. On the other hand, standard errors for growth, democracy, conflict, and the spatial lag are a bit larger by the S-GMM than by the other model-estimates, although the last two retain quite comfortable significance, with larger spatial-lag and smaller conflict coefficient as already noted.

As we have emphasized previously, interpretation of effects in spatially dynamic models, as in any model beyond the strictly linear-additive, cannot stop with noticing coefficient point-estimates. Effects, always and everywhere, are derivatives or differences such as $\partial E(y) / \partial x$ or $\Delta E(y) / \Delta x$. In spatial contexts, in fact, we will generally have interest in vectors of derivatives, or gradients, (a) $\nabla_{x_i} E(\mathbf{y} \mid \mathbf{X})$ or (b) $\nabla_{\varepsilon_i} E(\mathbf{y} \mid \mathbf{X})$, or even matrices of derivatives, (c) $\nabla_{\mathbf{x}} E(\mathbf{y} \mid \mathbf{X})$ or (d) $\nabla_{\boldsymbol{\varepsilon}} E(\mathbf{y} \mid \mathbf{X})$.[18] In order, these expressions indicate (a) the estimated responses across the set of n units (including i) to a marginal change in regressor x in unit i, (b) the estimated responses across the set of n units to a marginal change in the expected outcome (in this case, event-count) in unit i, (c) the set of estimated responses across all n units to some set of marginal changes in regressor x in some set of units j, row-by-row, and (d) the set of estimated responses across all n units to some set of marginal changes in the expected outcome in some set of units j, row-by-row.[19,20] For instance, we may be interested in how some subset or the entire set of countries would respond to some hypothetical shock to the wealth of some country, some subset of countries, or all countries. Table 4 illustrates with an example of the estimated effects on expected terrorist event-counts in selected countries of a hypothetical conflict in Iran.

Table 4: Estimated Responses in Expected Terrorist Event Counts across Selected Countries from a Hypothetical *Conflict* in Iran

	25th Percentile	**50th Percentile**	**75th Percentile**
Iran, Islamic Rep.	4.89	6.55	8.7
Iraq	1.77	5.75	18.02
Pakistan	1.29	2.47	4.06
Kuwait	0.21	0.91	5.03
Syrian Arab Republic	0.25	0.89	4.15
Israel	0.05	0.44	4.36
Bangladesh	0.05	0.42	3.19
Afghanistan	0.14	0.39	1.29
Nepal	0.07	0.35	1.58
Azerbaijan	0.19	0.33	0.58

To calculate these spatially distributed responses to hypothetical shocks of substantive interest, note first that S-GMM estimates are consistent and asymptotically normal. Thus,

[18] We may sometimes wish instead or additionally some scalar summary of these estimated vectors or matrices of responses. LeSage & Pace (2009) offer some good suggestions for such scalar summary effect-measures.

[19] We have noted elsewhere how it is convenient, in thinking about the responses across units i to hypothetical shocks to the outcome in units j, to conceive of ε as having a stochastic part and fixed component that can be shocked thusly. With additively separated ε, as in our specification in this draft, effects (b) and (d) are zero.

[20] In the spatiotemporal context, moreover, we would have interest in these sorts of effects in the short run, in the long run, and dynamically between, as explained in our previous work.

we can use the estimated parameters and their estimated variance-covariance matrix to draw from the asymptotic sampling distribution, a multivariate-normal with these means and variance-covariance, to bootstrap confidence intervals for the effects of covariates on terrorist event counts. The effect we want to calculate is the expected (or estimated) response across all n countries of a hypothetical shift (from zero to one) in one country's, i's, Iran's, value for one variable, $x_{i,k}$, Iran's *conflict* indicator:

$$\frac{dE(\mathbf{y} \mid \mathbf{X})}{dx_{i,k}} = \frac{d\boldsymbol{\lambda}}{dx_{i,k}} = \frac{d\left\{e^{(\mathbf{I}-\rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}}\right\}}{dx_{i,k}} \qquad (27),$$
$$= e^{(\mathbf{I}-\rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}} \odot \left[\left(\mathbf{I}-\rho\mathbf{W}\right)^{-1}\right]_i \beta_k = \boldsymbol{\lambda} \odot \mathbf{m}_i \beta_k$$

where $\mathbf{m}_i$ is the i^{th} column of the spatial multiplier, $\mathbf{M}\equiv(\mathbf{I}\text{-}\rho\mathbf{W})^{-1}$, and $\odot$ again is Hadamard or element-by-element multiplication (see, e.g., Anselin 2003, and Franzese & Hays 2006a, 2008b, for more on spatial multipliers). The expression in (27) is an $n\times1$ vector of estimated responses across all units to this hypothetical shock. Notice that, since the first term in this product involves **x**-values for all observations, i.e., all of **X**, in the term $(\mathbf{I}\text{-}\rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}$, the effect of a change in some $x_{i,k}$, i.e., in some x in some unit, depends on the values of <u>all</u> the x in <u>all</u> the units. In nonlinear models, it is common that the effect of one variable depends on the values of all the variables, since all the values affect where in the nonlinear function $E(\mathbf{y}|\mathbf{x})$, i.e., at what value of $E(\mathbf{y}|\mathbf{x})$, lies when we evaluate a counterfactual change (of some specific size) in x. In spatial contexts, we have $E(\mathbf{y}|\mathbf{X})$, i.e., the expected outcomes in each unit depend on those in all the units, i.e., at what value in this nonlinear function we are considering some dx depends on all **X**. This dimensional extension in the expression of counterfactuals arises in all spatial nonlinear models.[21]

Table 4 illustrates one informative way to present the estimates of such substantive counterfactuals. It reports the median and interquartile range of the estimated responses in terrorist event-counts around the world from a counterfactual conflict in Iran. Specifically, step-by-step: first, set Iran's *conflict* score in the dataset to one, keeping all other variables for Iran and all variables for all other units at their sampled values, and calculate $E(\mathbf{y}|\mathbf{X})$ for a draw of $\tilde{\boldsymbol{\theta}} \equiv \left\{\tilde{\rho},\tilde{\boldsymbol{\beta}}\right\}$ from $\tilde{\boldsymbol{\theta}} \sim \mathbf{N}(\hat{\boldsymbol{\theta}}^{opt}_{GMM},\hat{\boldsymbol{\Sigma}}^{opt}_{GMM})$. Again keeping *ceteris paribus*, set Iran's *conflict* to zero and calculate $E(\mathbf{y}|\mathbf{X})$ for the same draw of $\tilde{\boldsymbol{\theta}}$. The differences are the $n\times1$ vector of estimated responses across all countries to this hypothetical $d\mathbf{X}$ for this draw of $\tilde{\boldsymbol{\theta}}$.

[21] Franzese & Hays (2009), e.g., explain and discuss the analogous issue in S-Probit, for instance.

Repeating, say, 1000 times yields a parametric bootstrap of the sampling distribution for this substantive counterfactual effects-estimate.[22] Table 4 presents those the median and interquartile range of those estimates[23] for the countries with the ten largest median estimated responses the hypothetical *conflict* for Iran, using the first 1,000 draws with $\tilde{\rho}$ within the range of nonsingularity/invertibility of $\mathbf{I}-\tilde{\rho}\mathbf{W}$, which is $-1/|\omega_{min}| < \tilde{\rho} < 1/\omega_{max}$, where ω_{min} and ω_{max} are the minimum and maximum eigenvalues of $\mathbf{W}$. Logically, the largest median effect is for Iran itself, +6.55 expected terrorist-events; sensibly, the next largest is for Iraq, at +5.75, interestingly nearly as large. The estimated median response to a counterfactual *conflict* for Iran is also sizable for another first-order neighbor, Pakistan at +2.5, but the estimated responses fade rapidly from there. The median estimated response across all countries, for instance, is a paltry +.000224, or essentially zero.

VI. Conclusion:

Using a log-linearization of standard count models, like the Poisson or Negative Binomial, incorporating spatial autoregression into the systematic component of the model proves rather easy and straightforward. Given that additive separability of the stochastic component in this log-linearized form, as we show, least-squares based estimation, *at least once properly specified to scale*, has good bias properties, and there is not very much efficiency enhancement to be had from generalized-method-of-moments (weighted) instrumentation. Accordingly, at least insofar as this log-linearization and separability seems appropriate substantively and theoretically to the empirical application, applied researchers can make good estimates of their spatial-autoregressive count models simply by including spatial lags of *log of one plus observed counts*. In this substantive application, some interdependence in terrorist-event counts does appear to manifest, although additional exploration would seem to suggest that this is through exogenous spatial spillovers, rather than by contagion in the outcomes.

---

[22] Alternatively, similarly bootstrapping (27) directly would yield the average estimated derivative at some $\mathbf{X}$.

[23] A parametric bootstrap in a dynamic model may occasionally draw $\tilde{\boldsymbol{\theta}}$ yielding explosive, i.e., nonstationary, responses. If even one draw does so, then variance-based measures like confidence intervals will be explosive. For this reason, practitioners often report percentile ranges instead.

References

Akerlof, G.A. 1997. "Social Distance and Social Decisions." *Econometrica* 65(5):1005-27.

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

Beron, K.J., Murdoch, J.C., Vijverberg, W.P.M. 2003. "Why Cooperate? Public Goods, Economic Power, and the Montreal Protocol," *Review of Economics and Statistics* 85(2):286-97.

Beron, K.J., Vijverberg, W.P.M. 2004. "Probit in a Spatial Context: A Monte Carlo Analysis," in L. Anselin, R.J.G.M. Florax, & S.J. Rey, eds., *Advances in Spatial Econometrics: Methodology, Tools and Applications*. Berlin: Springer-Verlag.

Besag, J. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems." *Journal of the Royal Statistical Society B*, 36:192-225

Blomberg, S.B., Hess, G.D., Weerapana, A. 2004. "An Economic Model of Terrorism," *Conflict Management and Peace Science* 21(1):17-28.

Brock, W.A., Durlauf, S.N. 2001. "Discrete Choice with Social Interactions." *Review of Economic Studies* 68(2):235-60.

Brueckner, J. K. 2003. "Strategic Interaction Among Governments: An Overview of Empirical Studies." *International Regional Science Review* 26(2):175-88.

Correlates of War Project. *Direct Contiguity Data, 1816-2006*, v. 3.1. http://correlatesofwar.org.

Darmofal, D. 2009. "Bayesian Spatial Survival Models for Political Event Processes," *American Journal of Political Science* 53(1):241-57.

Elkins, Z., Simmons, B. 2005. "On Waves, Clusters, and Diffusion: A Conceptual Framework," *Annals of the American Academy of Political and Social Science* 598(1):33-51.

Franzese, R., Hays, J. 2003. "Modeling Spatial Relationships in International and Comparative Political Economy: An Application to Globalization and Capital Taxation in Developed Democracies," presented at the Annual Meetings of the Midwest Political Science Association.

Franzese, R., Hays, J. 2004a. "Empirical Modeling Strategies for Spatial Interdependence: Omitted-Variable vs. Simultaneity Biases," presented at the 21st Summer Meeting of the Society for Political Methodology.

Franzese, R., Hays, J. 2004b. "Modeling International Diffusion: Inferential Benefits and Methodological Challenges, with an Application to International Tax Competition," *Wissenschaftszentrum-Berlin* SP II 2004 – 12.

Franzese, R., Hays, J. 2005a. "Modeling Spatial Interdependence in Comparative and International Political Economy with an Application to Capital Taxation," presented at the annual Meeting of the Midwest Political Science Association.

Franzese, R., Hays, J. 2005b. "Spatial Econometric Modeling, with Application to Employment Spillovers and Active-Labor-Market Policies in the European Union," presented at Groningen University, workshop on 'Partisan Politics, Political Autonomy, and Policy Harmonization across Europe'.

Franzese, R., Hays, J. 2006a. "Calculating and Presenting Spatio-Temporal Patterns of Diffusion and Interdependence," presented at the General Conference of the European Consortium of Political Research, Pisa.

Franzese, R., Hays, J. 2006b. "Spatiotemporal Models for Political-Science Panel and Time-Series-Cross-Section Data," presented at the 23rd Annual Summer Meetings of the Society for Political Methodology.

Franzese, R., Hays, J. 2006c. "Strategic Interaction among EU Governments in Active-Labor-Market Policymaking: Subsidiarity and Policy Coordination under the European Employment Strategy," *European Union Politics* 7(2):167-89.

Franzese, R., Hays, J. 2007a. "Empirical Models of International Capital-Tax Competition," in *International Taxation Handbook*, G. Gregoriou, C. Read, eds., Elsevier Press, pp. 43-72.

Franzese, R., Hays, J. 2007b. "Interdependence in Comparative & International Political Economy, with Applications to Economic Integration and Strategic Fiscal-Policy Interdependence," presented at Paris 13 (*Université Paris*), Axe 5: PSE.

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

Franzese, R., Hays, J. 2007d. *Spatial Econometric Models of Interdependence*. Book prospectus.

Franzese, R., Hays, J. 2008a. "Empirical Modeling of Spatial Interdependence in Time-Series Cross-Sections," in S. Pickel, G. Pickel, H-J. Lauth, D. Jahn, eds., *Neuere Entwicklungen und Anwendungen auf dem Gebiet der Methoden der vergleichenden Politikwissenschaft, Band II. Wiesbaden: Westdeutscher Verlag*, forthcoming.

Franzese, R., Hays, J. 2008b. "Empirical Models of Spatial Interdependence," J. Box-Steffensmeier, H. Brady, D. Collier, eds., *Oxford Handbook of Political Methodology*, Oxford UP.

Franzese, R., Hays, J. 2008c. "Interdependence in Comparative Politics: Substance, Theory, Empirics, Substance," *Comparative Political Studies* 41(4/5):742-80.

Franzese, R., Hays, J. 2009. "The Spatial Probit Model of Interdependent Binary Outcomes: Estimation, Interpretation, and Presentation," presented at The Annual Meetings of the Public Choice Society.

Galton, F. 1889. "Discussion: 'On a method of investigating the development of institutions applied to laws of marriage and descent,' E. Tylor," *Journal of the Anthropological Institute* 18:270.

Glaeser, E.L., Scheinkman, J. 2003. "Non-market Interactions," *Advances in Economics and Econometrics: Theory and Applications, Eighth World Congress* 1:339-69.

Glaeser, E.L., Scheinkman, J., Sacerdote, B.L. 2003. "The Social Multiplier." *Journal of the European Economic Association* 1(2-3):345-53.

Griffith, D. 2002. "A Spatial Filtering Specification for the Auto-Poisson Model," *Statistics and Probability Letters* 58:245-51.

Griffith, D. 2003. *Spatial Autocorrelation and Spatial Filtering: Gaining Understanding through Theory and Scientific Visualization*. Berlin, Germany: Springer-Verlag.

Griffith, D. 2006. "Assessing Spatial Dependence in Count Data: Winsorized and Spatial Filter

Specification Alternatives to the Auto-Poisson Model," *Geographical Analysis* 38(2):160-79.

Griffith, D., Haining, R. 2006. "Beyond Mule Kicks: The Poisson Distribution in Geographical Analysis," *Geographical Analysis* 38(2):123-39.

Haegerstrand, T. 1967. *Innovation Diffusion as a Spatial Process*. (translation: A. Pred). Chicago: University of Chicago Press.

Haegerstrand, T. 1970. "What About People in Regional Science?" *Papers in Regional Science* 24(6):7-21.

Hays, J.C. 2009. "Bucking the System: Using Simulation Methods to Estimate and Analyze Systems of Equations with Qualitative and Limited Dependent Variables." Paper given at the 2nd Annual St. Louis Area Methods Meeting (SLAMM), Washington University in St. Louis (https://netfiles.uiuc.edu/jchays/www/Hays_SLAMM09_Revised.pdf).

Hays, J.C., Kachi, A. 2009. "Full Information Maximum Likelihood Estimation of Interdependent Weibull Duration Models with an Application to Government Formation and Survival." Available at SSRN: http://ssrn.com/abstract=1372937.

Hays, J., Kachi, A., Franzese, R. 2009. "An m-STAR Model of Dynamic, Endogenous Interdependence – a.k.a. Network-Behavior Coevolution – in the Social Sciences," presented at the 3rd World Congress of the Spatial Econometrics Association.

Heston, A., Summers, R., Aten, B. 2006. *Penn World Table Version 6.2*, Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, September 2006.

Kaiser, M.S., Cressie, N. 1997. "Modeling Poisson Variables with Positive Spatial Dependence," *Statistics & Probability Letters* 35(4):423-32.

Lee, R. 2009. "The History Guy: Wars and Conflicts of the 21st Century," http://www.historyguy.com/21st_century_wars.html.

LeSage, J.P. 2000. "Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models," *Geographical Analysis* 32(1):19-35.

LeSage, J.P., Pace, R.K. 2009. *Introduction to Spatial Econometrics*. Boca Raton: CRC Press.

Li, R., Thompson, W. 1975. "The 'Coup Contagion' Hypothesis," *Journal of Conflict Resolution* 19(1):63–88.

McMillen, D.P. 1992. "Probit with Spatial Autocorrelation," *Journal of Regional Science* 32:335-48.

Manski, C. F. 1993. "Identification of Endogenous Social Effects: The Reflexion Problem," *Review of Economic Studies* 60:531-42.

Manski, C. F. 1995. *Identification Problems in the Social Sciences*. Cambridge: Harvard University Press.

Manski, C.F. 2000. "Economic Analysis of Social Interactions." *Journal of Economic Perspectives* 14(3): 115-36..

Marshall, M.G., Jaggers, K., Gurr, T.R. 2008. *Polity IV Project: Political Regime Characteristics and Transitions, 1800-2008*. George Mason University, Colorado State University, and University of Maryland.

Mickolus, E.F., Sandler, T., Murdock, J.M., Flemming, P.A. 2003. INTERNATIONAL TERRORISM: ATTRIBUTES OF TERRORIST EVENTS, 1968-1977 [ITERATE 2] [Computer file]. ICPSR07947-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor], 1982. doi:10.3886/ICPSR07947.

Naroll, R. 1961. "Two Solutions to Galton's Problem," *Philosophy of Science* 28(1):15-39.

Naroll, R., D'Andrade, R.G. 1963. "Two Further Solutions to Galton's Problem," *American Anthropologist* 65(5):1053-67.

Naroll, R. 1965. "Galton's Problem: The Logic of Cross-Cultural Analysis," *Social Research* 32:428-51.

Neumayer, E., Plümper, T. 2008. "International Terrorism and the Clash of Civilizations: Was Huntington Right After All?" *British Journal of Political Science*.

Pluemper, T., Schneider, C. 2006. "The Computation of Convergence, or: How to Chase a Black Cat in a Dark Room," Available at SSRN: http://ssrn.com/abstract=955868.

Simmons, B., Dobbin, F., Garrett, G. 2006. "Introduction: The International Diffusion of Liberalism," *International Organization* 60(4):781–810.

Stinnett, D.M., Tir, J., Schafer, P., Diehl, P.F., Gochman, C. 2002. "The Correlates of War Project Direct Contiguity Data, Version 3." *Conflict Management and Peace Science* 19(2):58-66.

Stocking, G.W. 1968. "Edward Burnett Tylor," in D.L. Sills, ed., *International Encyclopedia of the Social Sciences*, v. 16. New York: McMillan Company: pp. 170-177.

Tobler, W.R. 1970. "A Computer Model Simulation of Urban Growth in the Detroit Region," *Economic Geography* 46(2):234-40.