Spatial Econometrics and Political Science

David Darmofal
Department of Political Science
University of South Carolina
350 Gambrell Hall
Columbia, SC 29208
Phone: (803) 777-5440
Email: darmofal@gwm.sc.edu

Abstract

Many theories in political science predict the spatial clustering of similar behaviors among neighboring units of observation. This spatial autocorrelation poses implications for both inference and modeling that are distinct from the more familiar serial dependence in time series analysis. In this paper, I examine how political scientists can diagnose and model the spatial dependence that is predicted by our theories. First, global and local measures of spatial autocorrelation are estimated to determine whether the data are spatially autocorrelated. If the data are spatially autocorrelated, the researcher attempts to model this dependence with a standard econometric specification and applies diagnostics to determine whether the covariates model the spatial autocorrelation. If the variables do not fully model the spatial dependence, the researcher estimates the spatial econometric model indicated by the diagnostics. Monte Carlo results and an empirical application to voting during the New Deal realignment highlight the importance of diagnosing and modeling spatial autocorrelation.

> "[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."

Sir Francis Galton at The Royal Anthropological Institute, 1888
The Journal of the Anthropological Institute
of Great Britain and Ireland 18: 270.

1 Introduction

Political science data are spatial data: the political behaviors, processes, and events we seek to understand occur at specific geographic locations. These locations, moreover, are often central to our understanding of these phenomena. Across a broad range of substantive concerns, from political communication to political conflict, democratization to dependency, policy diffusion to party mobilization, our theories posit that spatially proximate units are more likely to behave similarly than spatially distant units (Huckfeldt 1986; Vasquez 1995; Cardoso and Faletto 1979; Berry and Berry 1990). These theories, in short, predict positive spatial autocorrelation, the spatial clustering of similar behaviors among neighboring observations. Political science data are, in fact, particularly predisposed toward this positive spatial dependence.[1] Unlike more individuated concerns such as microeconomics, politics is, by nature, a collective concern. Shared political concerns combine with spatial proximity to promote familiarity. This familiarity in turn breeds both contempt and political conflict and interaction and political interdependence.

[1]Our interest in spatial econometrics is not on the full joint density, as spatial dependence would suggest, but instead on spatial autocorrelation as a moment of the joint density (Anselin and Bera 1998, 240). However, for ease of exposition, I follow common practice in spatial econometrics and use both terms to refer to spatial autocorrelation.

Until recently our ability to incorporate the spatial dimension of our theories in our models was quite limited, relying primarily on dummy variables to capture differences in behavior across geographically disparate units. Such an approach is suboptimal, as it is unable to address some of the central issues posed by spatially dependent data. Consider, for example, Sir Francis Galton's comment in the epigraph to this paper. Sir Galton's comment in response to Edward Tylor's presentation at The Royal Anthropological Institute in November 1888 clearly ranks among the most influential comments expressed at an academic presentation, remembered as it is more than a century later. Sir Galton's critique, which has since come to be known as Galton's problem, focuses on the critical substantive distinction between two alternative explanations for spatially dependent behavior.

On the one hand, spatial dependence may be produced by the diffusion of behavior between neighboring units. If so, the behavior is likely to be highly social in nature, and understanding the interactions between interdependent units is critical to understanding the behavior in question. Alternatively, neighboring units may share similar behaviors due simply to the units' independent adoptions of the behavior. If so, the spatial dependence observed in our data does not reflect a truly spatial process, but merely the geographic clustering of the sources of the behavior in question. Such dependence can be termed attributional dependence, as neighboring units have shared attributes that produce the clustering of behaviors. Clearly, determining which process is producing spatial dependence is critical to our understanding the behavior of interest.

A dummy variable approach is unable to distinguish between the two quite different explanations for spatial dependence. As proxies for our ignorance of the sources of spatial dependence, statistically significant parameters on dummy variables for geographic areas merely tell us that behaviors differ for units in these particular areas in contrast to the reference category. Such an approach cannot tell us whether the spatial dependence is consistent with diffusion or with the spatial clustering of the behavior's sources.

Recent advances in spatial econometrics, however, now allow us to address Galton's problem econometrically. While spatial econometric models come in a variety of forms, at their most basic level they share a common feature that distinguishes them from standard econometric models:

they explicitly model spatial autocorrelation.[2] Spatial econometric models allow us to address Galton's problem because each of the two alternative sources of spatial dependence posed by Galton presents its own distinct spatial econometric specification.

Spatial diffusion occurs because units' behavior is directly influenced by the behavior of "neighboring units." (As will be discussed, the definition of neighbors is generalizable, and need not imply contiguity). This diffusion effect corresponds to a positive and significant parameter on a spatially lagged dependent variable capturing the direct influence between neighbors.[3] Conversely, the geographic clustering of the sources of the behavior implies an alternative specification. Assuming that we are unable to model fully the sources of spatial dependence in the data generating process (DGP), these sources will produce spatial dependence in the error terms between neighboring locations. This spatial error dependence can be modeled via a spatially lagged error term.[4]

---

[2] The intellectual lineage of spatial econometric models can be traced to Isard's (1956) call for a regional science incorporating the spatial relationships between observed units. As a field, spatial econometrics includes a broad range of models in both frequentist and Bayesian perspectives. The most widely applied spatial econometric model is the cross-sectional model for continuous dependent variables. Given political scientists' extensive use of and familiarity with its non-spatial alternative, ordinary least squares, this paper focuses on this spatial econometric model. The discussion proceeds from a frequentist perspective. Scholars interested in a Bayesian perspective will also be interested in LeSage (1997).

[3] An artificial lag dependence may also be produced if there is a scale mismatch between the units of theoretical interest and the units that are available empirically, for example if the latter are nested within the former. This dependence of spatial autocorrelation estimates on the unit of analysis is known as the modifiable areal unit problem and is analogous to the issue of temporal aggregation in time series analysis (see, e.g., Freeman 1989). A positive and significant autoregressive parameter on a spatially lagged dependent variable is thus consistent with, but is not definitive with regard to the existence of a diffusion process.

[4] Spatial error dependence is also consistent with spatial clustering in measurement errors.

The substantive implications of properly modeling spatial dependence are intimately linked with methodological implications. Ignoring either form of spatial dependence in our models poses its own distinct implications for inference. For example, estimating an OLS model that ignores a diffusion effect in the DGP produces biased and inconsistent parameter estimates. Estimating an OLS model that ignores spatial clustering in the sources of the behavior produces inefficient parameter estimates, standard error estimates that are biased downward, and Type I errors.

Happily, the diagnosis and modeling of spatial dependence is a straightforward process that can be easily adapted by applied researchers. First, global and local measures of spatial autocorrelation are estimated to determine whether the data exhibit spatial autocorrelation. If the data do exhibit spatial autocorrelation, the researcher simply applies diagnostics to an OLS specification to determine whether the variables in the model sufficiently capture this spatial dependence. If the variables do not fully model the dependence, the diagnostics indicate whether the researcher should estimate a model with a spatially lagged dependent variable or a spatially lagged error term.[5]

The past decade has witnessed an increased use of spatial econometrics in a variety of disciplines, including economics, sociology, biostatistics, and urban planning. In contrast, despite the prominence of the spatial dimension in many of our theories, the application of spatial econometrics within political science remains quite limited.[6] By employing spatial econometrics, scholars in each of political science's empirical subfields can gain leverage on the spatial dependence that

[5]Rarely are models with joint lag and error dependence estimated. Largely this reflects the identification problem inherent in models with an identical spatial weights matrix for both the spatially lagged dependent variable and spatial autoregressive error dependence.

[6]Spatial econometric models have seen some limited use within political science. Perhaps the most visible application of spatial econometrics has occurred in international relations, where examples include Gleditsch and Ward (2000) and Gleditsch (2002). American politics applications include Busch and Reinhardt (2000) and Gimpel and Cho (2004). Comparative politics applications include Franzese and Hays (2003), O'Loughlin, Flint, and Anselin (1994), and Shin and Agnew (2002). Starr (1991), Starr and Most (1983), and Most and Starr (1980) also use a set of innovative approaches to examine the effects of spatial proximity on international conflict

is inherent in our theories.

The paper is structured as follows. I first examine how spatial dependence differs from the more familiar serial dependence in time series analysis and present Monte Carlo results on the effects of omitted spatial dependence on OLS estimates. Next, I examine the three sequential steps that researchers can adopt to diagnose and model the spatial dependence that is implied by our theories. Next, I demonstrate how these steps can be applied in practice through an empirical application to voting behavior during the New Deal realignment. The Monte Carlo and empirical results highlight the importance of diagnosing and modeling the spatial dependence that is predicted by our theories.

2 Modeling Spatial Autocorrelation

Formally, spatial autocorrelation implies a non-zero covariance between the values on a random variable for neighboring locations:

$$Cov(y_i, y_j) = E(y_i y_j) - E(y_i)E(y_j) \neq 0 \ \ \forall \ i \neq j, \tag{1}$$

where the i, j locations have a spatial interpretation (Anselin and Bera 1998, 241-242). The null hypothesis on a test of spatial autocorrelation is that the values on the random variable are distributed randomly in relation to space. That is, knowledge of units' spatial locations provides no leverage in predicting the units' values on the random variable. Positive spatial autocorrelation exists when neighboring units share similar values on the random variable, for example, when neighboring countries have similar economic policies or neighboring voters favor the same candidate. Negative spatial autocorrelation exists when neighboring units have dissimilar values on the random variable, for example, when neighboring countries have dissimilar economic policies or neighboring voters favor different candidates. Positive spatial autocorrelation will be more likely than negative spatial autocorrelation in most political science applications.

What implications does spatial autocorrelation pose for inference? At first glance, spatial

---

and democratization.

autocorrelation exhibits a surface similarity with the more familiar temporal dependence in time series analysis. Both are instances of dependent data. And in both the cross-sectional spatial case and the longitudinal time series case, the dependence can be modeled either via a lagged dependent variable or via the error term.

Spatial dependence, however, is not the cross-sectional analog of serial dependence. The critical distinction between the two forms of dependence arises from the dimensionality of the dependence. In the time series case, dependence is unidimensional: the past influences the present. Cross-sectional spatial dependence, in contrast, is typically conceived of as multidimensional. In the diffusion case, neighbors influence the behavior of their neighbors *and vice versa*. In the case of attributional dependence, errors for neighboring observations exhibit simultaneous dependence.[7] The simultaneous, multidimensional nature of spatial dependence leads to implications for inference that are distinct from the time series case. This also significantly complicates estimation of spatial econometric models incorporating this spatial dependence.

2.1 Spatial vs. Temporal Autocorrelation

To understand the differing implications of spatial and temporal autocorrelation, it is helpful to consider Anselin's (1988, 34) general spatial model for cross-sectional data:

$$y = \rho \mathbf{W} y + \varepsilon$$

$$\varepsilon = \lambda \mathbf{W} \varepsilon + \xi, \quad (2)$$

[7]This multidimensional perspective is the predominant perspective in spatial econometrics, as it is consistent both with the simultaneous, interactive nature of social contact and the simultaneous generation of errors in cross-sectional data. In contrast, a conditional perspective assumes that spatial dependence is unidirectional, with a unit's value on the variable conditioned on its neighbors' values, but not vice versa. In the conditional perspective, the process generating the neighbors' values on the variable is unexplained by the model (see Anselin 2003, 157). In practice, the number of dimensions is typically fixed at two in multidimensional models of spatial dependence.

where y is an N by 1 vector of observations on the dependent variable, $\mathbf{W}y$ is a spatially lagged dependent variable with spatial weights matrix $\mathbf{W}$, ρ is the spatial autoregressive parameter for the spatially lagged dependent variable, ε is an N by 1 vector of error terms, $\mathbf{W}\varepsilon$ is a spatially lagged error term with spatial weights matrix $\mathbf{W}$, λ is the spatial autoregressive parameter for the spatially lagged error term, and $\xi \sim N(0,\Omega)$, where $\Omega_{ii} = h_i(\mathbf{z}\alpha)$. When $\alpha = 0$, $h = \sigma^2$, and the errors are homoskedastic. The spatial lag model consistent with a diffusion process in the DGP results from setting λ equal to zero. The spatial error model consistent with attributional dependence results from setting ρ equal to zero.

Consider, first, the spatial lag model, where $\lambda = 0$. This bears some similarity to a time series model with a lagged dependent variable:

$$y_t = \rho y_{t-1} + \varepsilon_t, \tag{3}$$

where the lag in (3) is temporal rather than spatial, as it is in (2). In the time series case, OLS is a biased but consistent estimator of ρ in the absence of serial correlation and other misspecification errors. Thus, although the OLS estimator should not be relied on in small samples, it can still be employed for asymptotic inference.

In contrast to the time series case, OLS estimates of the autoregressive parameter ρ in a spatial lag model will be biased and inconsistent, regardless of whether the errors exhibit dependence. The distinction between the spatial and temporal cases exists because of the multidimensional nature of spatial dependence. Consider first, the bias of the OLS estimator of the spatial autoregressive parameter ρ. The expected value of the OLS estimator, $\hat{\rho}$, is:

$$\begin{aligned} E(\hat{\rho}) &= (y'\mathbf{W}'\mathbf{W}y)^{-1}y'\mathbf{W}'(\rho\mathbf{W}y + \varepsilon) \\ &= \rho + (y'\mathbf{W}'\mathbf{W}y)^{-1}y'\mathbf{W}'\varepsilon. \end{aligned} \tag{4}$$

As in the time series case, the expected value of the estimator will not equal the true value of ρ. In the spatial case, the sources of bias are twofold. First, as Anselin (1988, 78) notes, just as in the time series case, the complex nature of the matrix inverse induces a non-zero correlation with the error term. Unique to the spatial case, however, the expected value of $y'\mathbf{W}'\varepsilon$ is also non-zero whenever $\rho \neq 0$ due to the multidimensional nature of spatial, as opposed to temporal, dependence.

The consistency of the OLS estimator in the time series case exists only because y_{t-1} is uncorrelated with ε_t when there is no serial correlation in the errors. This does not hold in the multidimensional spatial case. We can reformulate the spatial lag model in (2) as:

$$y = (\mathbf{I} - \rho\mathbf{W})^{-1}\varepsilon, \tag{5}$$

where $(\mathbf{I} - \rho\mathbf{W})$ is the Leontief inverse, which acts as a spatial multiplier, linking the spatially lagged dependent variable to the errors at all locations. As Anselin and Bera (1998, 246-47) show, in contrast to the unidimensional time-series case, in the multidimensional spatial case the matrix inverse is a full matrix, producing an infinite series, $(\mathbf{I} + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \ldots)\varepsilon$. As a result, $\mathbf{W}y_i$ is correlated not only with ε_i, but also with the errors at all other locations.

Because of the simultaneous nature of spatial dependence, the OLS estimator $\hat{\rho}$ is inconsistent, regardless of whether there is dependence in the error term or not. This can be seen via the probability limit:

$$\text{plim } N^{-1}(y'\mathbf{W}'\varepsilon) = \text{plim } N^{-1}\varepsilon'\mathbf{W}(\mathbf{I} - \rho\mathbf{W})^{-1}\varepsilon, \tag{6}$$

where the error term takes a quadratic form and only when $\rho = 0$ does the probability limit equal zero (Anselin 1988, 58).

As Anselin (1988) shows, if a diffusion process exists in the DGP and a spatially lagged dependent variable is omitted from the model altogether, the result is biased and inconsistent parameter estimates for the covariates in the model, reflecting omitted variable bias. Estimation of the spatial lag model incorporating the spatially lagged dependent variable must proceed via either maximum likelihood estimation or an instrumental variables specification.

The second principal spatial model is a spatial error model where now the dependence pertains to the error term, rather than to a spatially lagged dependent variable. The spatial error model bears a resemblance to a time series model with serially correlated errors. The implications of spatial error dependence are similar, though not identical, to those of serial correlation in time series. As in the time series case, OLS parameter estimates remain unbiased, but are no longer efficient. In the presence of spatial error dependence, standard error estimates will be biased downward, producing Type I errors (Anselin 1988). The loss of information implicit in this spatial error dependence must be accounted for in estimation in order to produce unbiased standard error estimates. Similar to the case of a spatially lagged dependent variable, the simultaneous error dependence produces a non-zero covariance between the error terms at all locations, via the spatial multiplier:

$$E[\varepsilon\varepsilon'] = \sigma^2(\mathbf{I} - \lambda\mathbf{W})^{-1}(\mathbf{I} - \lambda\mathbf{W}')^{-1}, \tag{7}$$

with this covariance declining as the order of contiguity increases (Anselin 1988).

In the unidimensional serial correlation case, iterative FGLS estimators such as the Cochrane-Orcutt and Durbin estimators may be applied, producing consistent estimates of the autoregressive parameter for the serial dependence. These approaches are not applicable in the case of multidimensional spatial dependence. A spatial analog of the Cochrane-Orcutt estimator does not produce consistent estimates of the autoregressive parameter, λ (Anselin 1988, 110). And as Kelejian and Prucha (1997, 108) demonstrate, λ is unidentified in a spatial analog of the Durbin two-step method. As a result, estimation of the autoregressive spatial error parameter, λ, proceeds via maximum likelihood estimation.

3 Defining Neighbors Via a Spatial Weights Matrix

The first step in the process of diagnosing and modeling spatial autocorrelation is determining whether one's data exhibit such autocorrelation in the absence of covariates. If the researcher finds no evidence of spatial dependence in the data in this step, she can proceed by estimating a standard econometric model. If, however, diagnostics indicate the presence of spatial autocorrelation, she will next wish to model this dependence with covariates.

When diagnosing this spatial autocorrelation in cross-sectional data, constraints must be imposed on the covariances between observations, since the parameters of the complete covariance matrix are unidentified (Anselin 2002, 256). There are $N(N-1)$ potential spatial correlations in a cross-sectional data set. Clearly, there is insufficient information in cross-sectional data to estimate each of these separate covariances. Instead, the spatial autocorrelation must be parameterized in a limited number of parameters. In most applied work, a single autoregressive parameter is estimated to capture this spatial dependence.

Typically, spatial data in political science take the form of lattice data: observations on polygons such as counties, states, or nations.[8] When employing lattice data, the researcher constrains the dependence via a spatial weights matrix, which defines the j neighbors of unit i for whom spatial autocorrelation with i is permissable. The standard spatial weights matrix is an n x n matrix, $\mathbf{W}$, in which all of the j neighbors of i have non-zero values, ($\mathbf{w}_{ij} \neq 0$), the k non-neighbors of i have zero values, ($\mathbf{w}_{ik} = 0$), and i is not a neighbor of itself ($\mathbf{w}_{ii} = 0$). Generally, spatial weights matrices are row-standardized so that the sum of the weights for each observation equals 1. As a result, the spatial influence from neighbors is a weighted average of this influence across the j neighbors.

Clearly, the definition of neighbors is a critical decision in the modeling of spatial autocorrelation. If one misspecifies the spatial dependence in the DGP by treating non-neighbors as neighbors, or vice versa, subsequent spatial autocorrelation estimates will be biased. Closely related is the form and extent of spatial dependence between neighbors. Do all of the j neighbors of i exert the same influence on i? Or is this influence greater for neighbors closer to i? In defining neighbors and the form of spatial dependence between these neighbors, the constraints on potential spatial dependence incorporated in the weights matrix should reflect theoretical expectations. Commonly employed neighbor definitions include contiguity, k-nearest neighbor,

[8]The other two principal forms of spatial data are geostatistical data (sample data drawn from a continuous spatial surface) and point pattern data (geo-coded event data). Geostatistical data are unlikely to be employed in political science applications since, unlike sample data in the natural sciences, most political science samples are not drawn from a continuous spatial field. Although point pattern data have natural applicability within political science in spatial survival models, such models are beyond the scope of this paper and I focus instead on the lattice data perspective that is likely to be of greatest use within political science.

distance-decay, and non-spatial definitions.

The simplest definition of neighbors is the first-order contiguity case. Here, there are three possibilities. A rook contiguity definition defines units sharing a common edge with unit i as neighbors of i. A bishop contiguity definition defines units sharing a common vertex with i as neighbors of i. A queen contiguity definition combines the rook and bishop definitions as any unit sharing either a common edge or vertex with i is defined as a neighbor of i. Thus, in the queen contiguity definition, all polygons contiguous to i are neighbors of i (Anselin 1988, 18).

Alternatively, one may wish to relax the strict contiguity neighbor definition, but still retain a nearness conception of spatial influence. A k-nearest neighbor definition retains the nearness conception while not assuming that there is any substantive importance to the Euclidean distance between units. In a k-nearest neighbor definition, all units among the k nearest neighbors of unit i are treated as neighbors of i, while the $k + 1, \ldots, k + n$ units are treated as non-neighbors. Clearly the value k should be theoretically informed.

Often, the researcher will wish to posit that the spatial autocorrelation between units will decline as the distance between units increases. This conception of spatial dependence is consistent with Tobler's (1970, 236) first law of geography, in which "everything is related to everything else, but near things are more related than distant things." This implies a distance-decay function, as in a gravity or entropy model. The distance-decay function may also incorporate a size measure, such as population size, GDP, length of border, or other factors posited to influence autocorrelation independent of the distance between observations.

The size measure highlights an important consideration in modeling autocorrelation. The lattice data approach to dependence between observations is sufficiently generalizable that we need not incorporate a spatial component in this dependence at all. We can, for example, posit a social network effect, in which dependence is not a function of the spatial distance between units, but instead, is a function of these units' membership in a common social network, no matter how spatially dispersed. Alternatively, Beck, Gleditsch, and Beardsley (2006) model dependence as a function of trade flows between countries.

4 Monte Carlo Analysis

With this definition of weights matrices in hand, we can now employ Monte Carlo experiments to examine the performance of the OLS estimator when spatial dependence is present, but as in most applied research, this dependence is ignored in the model specification.[9] Recall from Section 2.1 the expected performance of the OLS estimator if spatial dependence in the data is ignored. If spatial autocorrelation consistent with a diffusion process exists in the data generating process and a spatially lagged dependent variable is omitted from the OLS model, this will produce biased and inconsistent OLS parameter estimates for covariates in the model.[10] If, alternatively, the spatial autocorrelation pertains only to the errors, OLS will remain an unbiased estimator, but it will no longer be efficient. Standard error estimates will be biased downward, producing Type I errors.

In the Monte Carlos, the DGP for the case of spatial lag dependence takes the form:

$$y = \rho \mathbf{W} y + \beta_0 + \beta_1 \mathbf{x} + \varepsilon, \tag{8}$$

where $\varepsilon \sim N(0, \sigma^2 \mathbf{I})$. The DGP for the case of spatial error dependence takes the form:

$$y = \beta_0 + \beta_1 \mathbf{x} + \varepsilon$$

$$\varepsilon = \lambda \mathbf{W} y \varepsilon + \xi, \tag{9}$$

where $\xi \sim N(0, \sigma^2 \mathbf{I})$. In both cases, the independent variable, $\mathbf{x}$, is normally distributed with

[9]The Monte Carlos are based on modified versions of the R code for omitted spatial error dependence in Anselin (2005).

[10]Of course, the OLS estimate of the spatial autoregressive parameter for the lagged dependent variable will be biased and inconsistent if this term is included in the OLS model; estimation should instead proceed via maximum likelihood estimation or an instrumental variables specification. However, given that a diffusion process is often implied by our theories and yet is often ignored in practice, it is critical to examine the performance of the OLS estimator for non-spatial covariates when this diffusion process is ignored.

a mean of 0 and a standard deviation of 3. I set $\beta_0 = \beta_1 = 1$. The OLS estimates for each experiment reflect a standard OLS specification that ignores the spatial lag or error dependence.

I examine the bias of the OLS estimates of β_1 when spatial lag and error dependence in the DGP are omitted from the OLS specification. I also examine the OLS estimates of the standard error of β_1 when spatial error dependence is omitted from the OLS specification. I examine the performance of OLS varying both the number of observations (and the corresponding spatial weights matrices) and the degree of spatial autocorrelation, as reflected in the autoregressive parameters ρ and λ. For each set of experiments, the observations are arrayed in regular square lattices. Monte Carlos are performed for four different sizes of square lattice structures: a 5 by 5 lattice ($n = 25$), a 10 by 10 lattice ($n = 100$), a 20 by 20 lattice ($n = 400$), and a 30 by 30 lattice ($n = 900$). In each case, a queen contiguity definition of neighbors is employed. The performance of OLS is examined for ten values of both ρ and λ: -.9, -.7, -.5, -.3, -.1, .1, .3, .5, .7, .9. For each combination of lattice size and ρ or λ value, 1000 replications are performed.

4.1 Monte Carlo Results for Omitted Spatial Lag Dependence

Table 1 reports the bias of the OLS estimator of β_1 when spatial dependence consistent with a diffusion process exists in the DGP but is omitted from the OLS specification. As can be seen, OLS performs well at low levels of both positive and negative spatial autocorrelation. The bias of the OLS estimator, however, increases appreciably as $|\rho|$ increases to .5 and beyond. Moreover, there is an asymmetric effect, as bias is markedly more problematic at high levels of positive spatial autocorrelation. With an n of 25, the OLS estimator overstates the true value of β_1 by 15% when $\rho = .7$ and by 35% when $\rho = .9$. These results are consistent with the expectations from Section 2.1.

4.2 Monte Carlo Results for Omitted Spatial Error Dependence

In contrast to the Monte Carlo results with omitted lag dependence, the Monte Carlo experiments with omitted spatial error dependence show no appreciable bias in the OLS estimator. Employing proportional measures of bias, as in Table 1, there is no bias to two decimal places across the four different lattice sizes and the ten values of λ. Consistent with expectations from

Section 2.1, OLS remains an unbiased estimator of slope parameters even at high levels of positive and negative spatial error dependence.

As expected, OLS standard errors are unduly optimistic in the presence of spatial error dependence. Table 2 reports the ratio of the OLS standard errors to the true standard errors. At low levels of spatial error dependence, OLS standard errors remain unbiased. However, as both positive and negative error dependence increase in size, the standard errors reported by OLS increasingly understate the true standard errors. Similar to the case of omitted spatial lag dependence, the problem is particularly acute for high levels of positive spatial error autocorrelation. For example, with an n of 900 and a λ of .9, the OLS standard error is only .57 the size of the true standard error. In the presence of both positive and negative spatial error dependence, inference based on OLS is likely to lead to Type I errors.

5 Global Measures of Spatial Autocorrelation

The Monte Carlo results highlight the importance of diagnosing the spatial autocorrelation that is implied by many of our theories in political science. This diagnosis of spatial autocorrelation may proceed at either the global or local levels. Once one has defined "neighboring" observations via the spatial weights matrix, the researcher can employ tests for global spatial autocorrelation to examine whether the data as a whole exhibit spatial autocorrelation (against a null of spatial randomness) as well as the strength and direction (positive or negative) of any spatial autocorrelation. Using the same weights matrix, the researcher can also employ tests for local spatial autocorrelation (again, against a null of spatial randomness) to identify the particular observations that are autocorrelated with neighboring observations on the random variable and also determine the strength and the direction of this spatial autocorrelation. Typically, researchers first employ global tests of spatial autocorrelation and subsequently employ local tests to decompose the global result.

The most commonly applied global spatial autocorrelation measure is Moran's I. For Moran's I, value (dis)similarity between neighboring observations is measured as deviations from the mean on the variable of interest. The global Moran's I is thus:

$$I = \frac{N}{S}\frac{\Sigma_i \Sigma_j \mathbf{w}_{ij}(y_i - \overline{y})(y_j - \overline{y})}{\Sigma_i (y_i - \overline{y})^2}, \tag{10}$$

where N is the number of observations, S is the sum of the weights, $\mathbf{w}_{ij}$ is the (i,j)th element in the spatial weights matrix $\mathbf{W}$, y_i and y_j are the values on the random variable at locations i and j, and $\overline{y}$ is the mean on y.[11]

A positive global Moran's I that differs significantly from the expected value under the null indicates positive spatial autocorrelation – the clustering of similar values on the random variable among neighboring observations. A negative global Moran's I that differs significantly from the expected value under the null indicates negative spatial autocorrelation – the clustering of dissimilar values on the random variable among neighboring observations.

Although both this interpretation of the global Moran's I and the form of the measure suggest a correlation coefficient, Moran's I differs from a correlation coefficient in two key respects. The expected value of Moran's I under the null is not zero, but instead is $\frac{-1}{N-1}$, and thus is a function of the number of observations (Anselin 1992, 133). As a consequence, the expected value of Moran's I under the null is negative, though it approaches zero asymptotically. Moreover, unlike a correlation coefficient, Moran's I is not bounded at $\pm$ 1. Instead, the bounds are a function of the data and will generally be narrower than $\pm$ 1 (Cliff and Ord 1981, 21).

Inference on Moran's I takes either of two approaches (see Anselin 1992, 134). If the random variable is normally distributed with a constant variance, then Moran's I is asymptotically normally distributed under the null. Inference then proceeds by comparing the observed z-value to its probability given the normal distribution. Alternatively, one can apply a randomization approach in which the observed values on the random variable are randomly permuted across all locations. Note that this randomization with regard to spatial location corresponds to the null

[11] Geary's c is a second, less commonly applied measure of global spatial autocorrelation. Because Geary's c defines value (dis)similarity as the squared difference in values between neighboring observations, it gives greater weight to extreme values than does Moran's I (Cliff and Ord 1981, 14-15).

of spatial randomness. A Moran's I is then calculated for each permutation to form an empirical reference distribution and the observed Moran's I is compared to the reference distribution to determine pseudo-statistical significance.

6 Local Measures of Spatial Autocorrelation

Often we will wish to move beyond global measures of spatial autocorrelation to identify the specific units that exhibit spatial autocorrelation with their neighbors. Local indicators of spatial association (LISA statistics) are employed to disaggregate the spatial dependence diagnosed with the global measures of spatial autocorrelation and identify local spatial autocorrelation. Anselin (1995, 94) defines a LISA statistic as any statistic satisfying the following two conditions: the LISA for each observation measures the extent of significant spatial clustering of similar values around the observation, and the sum of LISAs for all observations is proportional to a corresponding global indicator of spatial association.[12]

The principal LISA statistic is the local Moran's I:

$$I_i = \frac{\Sigma_j \mathbf{w}_{ij}(y_i - \overline{y})(y_j - \overline{y})}{(y_i - \overline{y})^2}, \tag{11}$$

where the notation is as in its global analog in (10). The interpretation of values of the local Moran's I is analogous to its global counterpart.[13]

In addition to the identification of local spatial clustering, the correspondence between LISA statistics and global spatial autocorrelation measures carries significant additional advantage in

[12]Although Anselin defines a LISA in terms of the clustering of similar values, his definition is unnecessarily restrictive. LISA statistics, like their global analogs, identify both positive and negative spatial autocorrelation. Alternative measures of local spatial autocorrelation have also been developed that are not proportional to a global analog. Most prominent among these are the G_i and G_i^* statistics developed by Getis and Ord (Getis and Ord 1992, Ord and Getis 1995).

[13]The local Geary's c is an alternative LISA statistic. Like its global analog, it defines value (dis)similarity as the squared differences in values between neighboring observations.

decomposing the global measures. By estimating LISA statistics, the researcher can identify which observations are consistent with the global pattern of positive or negative spatial autocorrelation and which observations run counter to this global pattern. High leverage observations can also be identified. Often this follows a two-sigma rule: observations with LISAs that are more than two standard deviations from the mean can be examined to determine whether they are unduly influencing the global measure (Anselin 1995, 97).

Inference on the local Moran proceeds in the same manner as for its global analog. The researcher can rely on asymptotics and assume a normal distribution or employ randomization. In contrast to the global approach, however, only as many observations as are in each observation's neighborhood set need be resampled from the permuted values in the randomization approach.

7 Spatial Heterogeneity

Once the researcher has diagnosed univariate spatial autocorrelation via global and local measures, the next step is to attempt to model this autocorrelation with covariates. In this step, the researcher estimates a standard econometric model and applies diagnostics to determine whether the covariates fully model the spatial dependence. If the diagnostics indicate that the covariates do not fully model this dependence, these diagnostics indicate whether a spatial lag or a spatial error model is in order. When estimating the standard econometric model, it is critical to determine whether the spatial autocorrelation diagnosed in the previous step is the product of a second form of spatial effect, spatial heterogeneity in parameters. If such behavioral heterogeneity has produced any of the spatial autocorrelation that was diagnosed, this heterogeneity should be included in the model specification. In the next section, I examine diagnostics for spatial autocorrelation that can be applied to OLS models, including OLS models that incorporate behavioral heterogeneity. First, however, I examine how the researcher can model this behavioral heterogeneity.

The intuition of how spatial heterogeneity in parameters may produce univariate spatial autocorrelation is straightforward. If a covariate differs in its effects on the political phenomenon of interest across units, and if the effects are similar among neighboring units, this may produce similar values on the dependent variable. Often this heterogeneity will be of a discrete form, where a parameter or parameters are homogeneous within spatial subsets of the data and heterogeneous across these subsets, reflecting distinct spatial regimes in the data. The LISA statistics estimated in the previous step are useful in suggesting potential spatial regimes in the data.[14]

This concept of discrete spatial heterogeneity in parameters can be found in many political science theories. Consider, for example, studies of American voting behavior where behavioral parameters are hypothesized to operate differently in the South vs. the non-South. Or consider the comparative politics concepts of the global North and South where relationships such as those between financial weakness and capital account liberalization or between globalization and welfare state spending differ in discrete spatial subsets of the data (Brooks 2004; Rudra 2002). We can model this spatial heterogeneity in parameters via spatial switching regressions.

7.1 Spatial Switching Regressions

A spatial switching regression, or spatial regimes model, applies spatial Chow tests to diagnose structural instability in parameters across regimes. The spatial regimes model takes the form:

$$\begin{bmatrix} y_i \\ y_j \end{bmatrix} = \begin{bmatrix} \mathbf{X}_i & 0 \\ 0 & \mathbf{X}_j \end{bmatrix} \begin{bmatrix} \beta_i \\ \beta_j \end{bmatrix} + \begin{bmatrix} \varepsilon_i \\ \varepsilon_j \end{bmatrix}, \tag{12}$$

[14]Alternatively, one may posit continuous spatial heterogeneity in parameters, where the parameter values exhibit a continuous spatial drift as one moves across spatial dimensions. Such continuous parameter heterogeneity may be modeled via a spatial expansion model (where the behavioral parameters are modeled as a function of the x,y coordinates of the observations), or a geographically weighted regression (in which spatially proximate units are given greater weight in the calculation of the spatially varying parameters than more spatially distant units) (Fotheringham, Charlton, and Brunsdon 1998).

where i,j index discrete spatial subsets of the data. The spatial Chow test then proceeds as a test of the null that $\beta_i = \beta_j$ via an F test:

$$\mathbf{C} = [(\mathbf{e}_R^{'}\mathbf{e}_R - \mathbf{e}_U^{'}\mathbf{e}_U)/K][\mathbf{e}_U^{'}\mathbf{e}_U/(N-2K)] \sim F_{K,N-2K}, \tag{13}$$

where $\mathbf{e}_R$ and $\mathbf{e}_U$ are the OLS residuals from a restricted model (with the equality restriction imposed) and from an unrestricted model, N is the number of observations, and K is the number of regressors (Anselin 1990, 192).

If the switching regression models fully the sources of univariate spatial autocorrelation, a spatial econometric specification is not required. Spatial dependence may, however, persist even after any behavioral heterogeneity has been modeled. If so, the researcher will wish to model both the spatial heterogeneity and the spatial dependence.[15] In the next section, I examine several diagnostics that can be applied to OLS models, including OLS spatial regimes models, to determine whether spatial autocorrelation persists in the presence of covariates.

8 Diagnostics for Spatial Dependence in OLS Models

8.1 The Moran's *I* and Kelejian-Robinson Diagnostics for Spatial Error Dependence in an OLS Model

Given the wide familiarity with the Moran's I as a diagnostic for univariate spatial autocorrelation, it is not surprising that the test has been extended to the diagnosis of spatial dependence in the presence of covariates. The Moran's I test for spatial error dependence in OLS regression

[15]The researcher will also wish to model spatial heterogeneity and spatial dependence jointly because the latter can lead to overrejection of the null of parameter stability in spatial Chow tests on OLS models (Anselin 1990). Thus, if the researcher finds parameter instability in a spatial Chow test for an OLS regimes model and diagnostics also indicate continued spatial lag or error dependence, the researcher should next estimate the spatial model indicated by the diagnostics to determine whether the parameter instability persists in a spatial econometric or instrumental variables model that explicitly models the spatial dependence.

residuals takes the form:

$$I = \frac{N}{S}\frac{\mathbf{e}'\mathbf{W}\mathbf{e}}{\mathbf{e}'\mathbf{e}}, \tag{14}$$

where N is the number of observations, S is the sum of the weights, $\mathbf{e}$ are the residuals from an OLS regression, and $\mathbf{W}$ is the spatial weights matrix. As can be seen from the form of (14), Moran's I is the spatial analog of the Durbin-Watson test for serial correlation in residuals. Although Greene (2003, 192) recommends such a test as a diagnostic for cross-sectional autocorrelation, the Moran's I is actually a highly unreliable diagnostic for spatial error dependence in OLS residuals. Anselin and Rey (1991) find that the Moran's I diagnostic for error dependence underrejects the null in the presence of non-normal errors and overrejects the null in the presence of heteroskedasticity. Moreover, the Moran's I diagnostic also has power against lag dependence, rejecting the null when lag rather than error dependence is present. As a result, it should not be relied on for guidance for the proper spatial model.

Kelejian and Robinson propose an alternative, non-parametric diagnostic for error dependence that does not require normality in errors. The Kelejian-Robinson diagnostic takes the form:

$$KR = \frac{\hat{\gamma}'\mathbf{Z}'\mathbf{Z}\hat{\gamma}}{\hat{\sigma}^4} \sim \chi^2_K, \tag{15}$$

where $\hat{\gamma}$ is the parameter vector from an auxiliary regression of the cross-products of residuals and the cross-products of covariates in the matrix, $\mathbf{Z}$, in which the cross-products are all paired observations for which there is a non-zero correlation, and K is the number of regressors in the observation matrix, $\mathbf{Z}$ (Anselin 1992, 179). The term, $\hat{\sigma}^4$, is any consistent estimator of σ^4. Although the diagnostic has the advantage of not assuming normality in the errors, it has the disadvantage of being limited to the case of spatial error dependence.

8.2 Lagrange Multiplier and Robust Lagrange Multiplier Diagnostics for Spatial Lag and Spatial Error Dependence in an OLS Model

Given the poor performance of the Moran's I diagnostic and the limitation of the Kelejian-Robinson diagnostic, the researcher typically will wish to apply alternative diagnostics for spatial dependence. Here, the principal diagnostics are four Lagrange Multiplier diagnostics. Following

Anselin and Rey's (1991, 119) notation, the one-directional Lagrange Multiplier diagnostic for spatial lag dependence takes the form:

$$LM_{Lag} = [N\mathbf{e}'\mathbf{W}y/\mathbf{e}'\mathbf{e}]^2[N(\mathbf{WX}\hat{\beta})'\mathbf{M}(\mathbf{WX}\hat{\beta})/\mathbf{e}'\mathbf{e} + \text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)]^{-1}, \quad (16)$$

where N is the number of observations, $\mathbf{e}$ are the OLS residuals, $\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, $\hat{\beta}$ is the OLS estimate of β, tr is the matrix trace operator, and $\mathbf{W}$ is the spatial weights matrix for the spatially lagged dependent variable. The one-directional Lagrange Multiplier diagnostic for spatial error dependence takes the form:

$$LM_{Error} = [N\mathbf{e}'\mathbf{We}/\mathbf{e}'\mathbf{e}]^2[\text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)]^{-1}. \quad (17)$$

The two tests are one-directional in that they do not consider the presence of the alternative form of spatial dependence. As a consequence, they are not robust when the alternative form of dependence is present. The Lagrange Multiplier diagnostic for spatial lag dependence will reject the null even if $\rho = 0$, if spatial error dependence is present. Similarly, the Lagrange Multiplier diagnostic for spatial error dependence will reject the null even if $\lambda = 0$, if spatial lag dependence is present (Anselin and Bera 1998, 274). As a result, researchers should also apply Lagrange Multiplier diagnostics that are robust to the alternative form of dependence.

When the alternative form of spatial dependence is present, the simple unidirectional LM tests converge to a noncentral χ^2 distribution featuring an additional noncentrality parameter. Recognizing this, Bera and Yoon (1993) developed modified Lagrange Multiplier tests that account for the noncentrality parameter, and are, as a consequence, robust to misspecification of the form of spatial dependence. Anselin, Bera, Florax, and Yoon (1996) subsequently extended Bera and Yoon's modified LM tests to the diagnosis of spatial lag and spatial error dependence in OLS specifications. In essence, the spatial lag dependence is estimated in the diagnostic for lag dependence by accounting for any spatial error dependence that may exist. Likewise, spatial error dependence is estimated in the diagnostic for error dependence by accounting for any spatial lag dependence that may exist. The robust Lagrange Multiplier diagnostic for spatial lag dependence developed by Anselin, Bera, Florax, and Yoon (1996, 83) as an extension of the Bera and Yoon modified LM test takes the form:

$$LM^*_{Lag} = \frac{(\mathbf{e}'\mathbf{W}y/s^2 - \mathbf{e}'\mathbf{W}\mathbf{e}/s^2)^2}{N\tilde{J}_{\rho\cdot\beta} - t}, \tag{18}$$

where $s^2 = \frac{\mathbf{e}'\mathbf{e}}{N}$, and $N\tilde{J}_{\rho\cdot\beta} = [t + (\mathbf{WX}\beta)'\mathbf{M}(\mathbf{WX}\beta)/s^2]$, with $\mathbf{M} = \mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$, and $t = \text{tr}(\mathbf{W}'\mathbf{W} + \mathbf{W}^2)$. The robust Lagrange Multiplier diagnostic for spatial error dependence in an OLS model takes the form:

$$LM^*_{Error} = [\mathbf{e}'\mathbf{W}\mathbf{e}/s^2 - t(N\tilde{J}_{\rho\cdot\beta})^{-1}(\mathbf{e}'\mathbf{W}y/s^2)]^2/[t - t^2(N\tilde{J}_{\rho\cdot\beta})^{-1}]. \tag{19}$$

Although the four Lagrange Multiplier diagnostics assume normally distributed errors, Anselin, Bera, Florax, and Yoon (1996, 93) find that the robust LM diagnostics are robust to non-normality and Anselin and Rey (1991, 124) find that the one-directional LM diagnostics are robust to non-normality in large samples. Anselin, Bera, Florax, and Yoon also find that the robust LM diagnostics are less prone to Type I errors than their non-robust counterparts in the presence of the alternative form of spatial dependence. As Anselin and Bera (1998, 277) note, however, the robustification of the Lagrange Multiplier diagnostics does not come without a price. The robust LM diagnostics have reduced power in comparison to the unidirectional LM diagnostics in the absence of the alternative form of spatial dependence. As a consequence, they are more prone to Type II errors than their non-robust counterparts.

The researcher is thus advised to estimate both the one-directional and robust Lagrange Multiplier diagnostics. Given the propensity of the one-directional diagnostics toward Type I errors, an inability to reject the nulls on either of the one-directional tests indicates that the OLS specification is sufficient. If the null is rejected on one or both of the one-directional diagnostics, the researcher should consult the robust diagnostics. The researcher should estimate the spatial model indicated by the larger robust LM diagnostic.

9 Modeling Spatial Dependence

If the OLS diagnostics indicate the existence of spatial lag dependence, the researcher has two options. She can estimate a mixed regressive, spatial autoregressive (spatial lag) model via maximum likelihood estimation. Alternatively, she can estimate an instrumental variables specification incorporating instruments for the spatially lagged dependent variable. If the OLS diagnostics indicate the existence of spatial error dependence, the researcher may choose to estimate a more fully specified OLS model to model the spatial dependence or may choose to estimate a maximum likelihood model incorporating the spatial dependence in the errors. I examine spatial lag and spatial error models in turn next.

9.1 Maximum Likelihood Spatial Lag Estimation

The mixed regressive, spatial autoregressive model, or spatial lag model, extends the pure spatial autoregressive model in (2) to include a set of covariates and associated parameters:

$$y = \rho \mathbf{W} y + \mathbf{X}\beta + \varepsilon, \tag{20}$$

where $\mathbf{X}$ is an N by K matrix of observations on the covariates, β is a K by 1 vector of parameters, and the remaining notation is as in (2). (Both spatial lag and spatial error models can easily be modified to include spatial regimes). As discussed in Section 2, if spatial lag dependence exists in the DGP, OLS estimates of the spatial autoregressive parameter, ρ, will be biased and inconsistent, regardless of whether spatial dependence exists in the error term or not. As a result, the researcher may wish to estimate ρ via maximum likelihood estimation. Here, the log-likelihood function takes the form:

$$L_{Lag} = \sum_i \ln(1 - \rho\omega_i) - \frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma^2) - \frac{(y - \rho\mathbf{W}y - \mathbf{X}\beta)'(y - \rho\mathbf{W}y - \mathbf{X}\beta)}{2\sigma^2}, \tag{21}$$

where ω_i are the eigenvalues of the spatial weights matrix (Anselin and Bera 1998, 255).

9.2 Instrumental Variables Spatial Lag Estimation

Maximum likelihood estimation of the spatial lag model assumes that the errors are normally distributed. If diagnostics such as the Jarque-Bera indicate that the errors are not normally distributed, the researcher should instead estimate an instrumental variables model. Here, the researcher will employ instruments for the spatially lagged dependent variable that are asymptotically uncorrelated with the error term. Assuming that the proper instruments can be found, the instrumental variables estimator will be consistent (as the instruments will be uncorrelated with the error), but will not be the most efficient estimator. The relative efficiency will depend upon the choice of the instruments (Anselin 1988, 84; Anselin and Bera 1998, 259). Anselin (1988, 85) suggests two potential instruments for the spatially lagged dependent variable. The researcher may employ spatially lagged versions of the covariates in the model as instruments for the spatially lagged dependent variable. Alternatively, the researcher may use spatially lagged predicted values of the dependent variable.

9.3 Maximum Likelihood Spatial Error Estimation

Alternatively, diagnostics may indicate the existence of spatial error dependence. The regressive model with autoregressive error dependence takes the form:

$$y = \mathbf{X}\beta + \varepsilon$$

$$\varepsilon = \lambda \mathbf{W} \varepsilon + \xi, \tag{22}$$

where the notation is as in (2) and (20). As discussed in Section 2, OLS estimates of the autoregressive parameter, λ, will not be consistent. As a result, if the researcher cannot model the spatial error dependence with covariates in a more fully specified OLS model, she may wish to estimate a maximum likelihood model incorporating the dependence in the errors.[16] Here, the

[16] Alternatively, one can employ a generalized moments estimator to account for the spatial error dependence (see Kelejian and Prucha 1999).

log-likelihood takes the form:

$$L_{Error} = \sum_i \ln(1 - \lambda\omega_i) - \frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln(\sigma^2) - \frac{(y - \mathbf{X}\beta)'(\mathbf{I} - \lambda\mathbf{W}_2)'(y - \mathbf{X}\beta)}{2\sigma^2}. \qquad (23)$$

10 Spatial Autocorrelation and Voting Behavior During the New Deal Realignment

I employ data on voting behavior during the New Deal realignment to demonstrate how spatial dependence can be diagnosed and modeled by applied researchers. Theory predicts spatial dependence in voting behavior during such critical realignments. Realignment theory emphasizes emerging issue concerns as the impetus for changes in voting behavior (Sundquist 1983). Voters in neighboring electorates are likely to share similar policy concerns; as a consequence, we should expect similar shifts in voting behavior in neighboring electorates in response to these shared concerns. Such shifts, moreover, may reflect a diffusion process; for example, shared social networks that transcend local electorates may promote shared changes in voting behavior in response to these emergent policy concerns. Indeed, in an analysis that does not employ spatial econometrics, Nardulli (1995) finds that neighboring counties exhibit similar changes in voting behavior during realignments.

I examine changes in county-level support for the Democratic Party between the 1928 and 1932 presidential elections.[17] Because the county is the lowest unit of analysis for which we have voting data on all local electorates in the United States during this period, a county-level analysis provides particular utility in examining spatial dependence in voting behavior. Change in the Democratic vote between the Al Smith and Franklin D. Roosevelt candidacies is measured as a proportion of the eligible electorate $\left(\frac{\text{Democratic Vote}_{1932}}{\text{Eligible Electorate}_{1932}} - \frac{\text{Democratic Vote}_{1928}}{\text{Eligible Electorate}_{1928}}\right)$. This reflects the argument by Andersen (1979) that the mobilization of non-voters, rather than the conversion of active partisans, may have played a critical role in producing the increased support for the Democrats during this period.

[17]The focus is thus on immediate changes in voting behavior as the New Deal realignment was taking place, rather than on enduring changes in voting behavior.

10.1 Diagnosing Spatial Dependence in Voting Behavior

The first step in assessing spatial dependence in changes in Democratic support between the 1928 and 1932 elections is diagnosing this dependence in the absence of covariates via global and local measures of spatial autocorrelation. If the spatial autocorrelation measures indicate no spatial dependence, we can proceed with estimating a standard OLS specification. I estimate spatial dependence in the data using the global and local Moran's I's.

The global Moran's I for the changes in Democratic support is .67 ($p < .01$, employing the randomization approach with a queen contiguity neighbor definition).[18] The global Moran's I thus diagnoses strong positive global spatial autocorrelation. At the local level, more than 40 percent of the counties exhibit significant positive spatial autocorrelation with their neighbors as estimated by the local Morans (also employing the randomization approach and a queen contiguity neighbor definition). The strong global spatial dependence reflects a common local pattern and is not the product of a few high outlying cases.

The local spatial autocorrelation can be seen visually in Figure 1, which plots the counties' local spatial autocorrelation by the direction of the spatial autocorrelation and by the counties' relationship to the mean county-level change in Democratic support between 1928 and 1932. Counties with insignificant local Morans are plotted in white. Positively autocorrelated counties that shared above average shifts toward the Democrats with their neighbors are plotted in dark blue and labeled, for ease of exposition, as high-high cases. Positively autocorrelated counties that shared below average shifts toward the Democrats with their neighbors are plotted in light blue and are labeled as low-low cases. Negatively autocorrelated counties with stronger shifts toward the Democrats than their neighbors are plotted in dark green and labeled as high-low cases. Negatively autocorrelated counties with weaker shifts toward the Democrats than their neighbors are plotted in light green and labeled as low-high cases.[19]

---

[18]The global and local Moran's I, OLS, and instrumental variables estimates in this paper were computed using SpaceStat™ v.1.91.

[19]The average county-level change in Democratic support between 1928 and 1932 was .14 (an average increase of 14 percentage points in Democratic support among the eligible electorate).

The local Moran map suggests discrete spatial regimes in the data. There was a strong shift toward the Democrats between 1928 and 1932 in the Plains, where the Dust Bowl had hit between the two elections (an average increase in Democratic support of 24 percentage points in Plains counties shaded in dark blue). In contrast, much of the South evidenced much smaller movement toward the Democrats, unsurprising given the Democratic dominance that already existed in the South (an average increase of only 4 percentage points in Southern counties shaded in light blue). Given these discrete spatial subsets of the data, a spatial regimes model in which the behavioral parameters are allowed to vary in the Plains and the South is indicated.

10.2 Modeling Spatial Dependence and Spatial Heterogeneity in Voting Behavior

Given the findings of significant positive global and local spatial autocorrelation and the indication of potential behavioral heterogeneity, the next step is to attempt to model this spatial dependence and spatial heterogeneity with covariates. I model changes in county-level support for the Democrats between 1928 and 1932 as a function of six county-level covariates suggested by the literature on voting behavior during the New Deal realignment. $\textit{Republican Vote}_{1928}$ measures the Republican vote in the county in the 1928 presidential election (again, as a proportion of the eligible electorate) and accounts for changes from Republican support to Democratic support between the 1928 and 1932 elections. $\textit{Non-Voting}_{1928}$ measures the non-voting population in the county in the 1928 presidential election as a proportion of the eligible electorate and reflects Andersen's thesis that the increase in Democratic support resulted from the mobilization of previous non-voters. *Foreign-Born* is the proportion of foreign-born individuals in the county in 1932 and reflects Andersen's (1979) argument that immigrants played a pivotal role in producing the increase in Democratic support. *Rural* is a dichotomous variable coded 1 if the county was

---

The mean in counties with positively autocorrelated above average change (plotted in dark blue) was .23. The mean in counties with positively autocorrelated below average change (plotted in light blue) was .04. The mean in counties plotted in dark green was .17 while the mean in counties plotted in light green was .11. The mean in counties with insignificant LISAs (plotted in white) was .14.

a rural county and 0 otherwise. This covariate examines Brown's (1991) argument that the surge in Democratic support was disproportionately a rural phenomenon, rather than the urban phenomenon suggested by Andersen and others.

I also include two interaction terms to assess the effects of interactions between electoral and demographic features of local electorates. The interaction term $\textit{Non-Voting}_{1928} * \textit{Foreign-Born}$ examines whether shifts toward the Democrats were promoted by the joint presence in a county of two groups likely to have weak prior attachments to the political parties, non-voters and immigrants. The interaction term $\textit{Republican Vote}_{1928} * \textit{Foreign-Born}$ examines whether shifts toward the Democrats were promoted or impeded by the joint presence of immigrants and prior Republican supporters. Such prior Republican support may have discouraged movement toward the Democrats or encouraged it through the interaction of conversion and mobilization.

The estimates of the OLS spatial regimes model are listed in column 4 of Table 3. In both the Plains and the South, shifts to the Democrats in 1932 were stronger where Republicans had run stronger in 1928, where there were large pools of non-voters in 1928, and in rural counties. Counterintuitively, however, the OLS estimates indicate that the joint presence of pools of non-voters and immigrants impeded shifts toward the Democrats in the Plains in 1932. Thus, precisely in those counties where we would expect particularly strong movement toward the Democrats due to the presence of two groups with weak attachments to the parties, the OLS estimates indicate that shifts to the Democrats were instead impeded.[20]

The diagnostics for spatial dependence in the presence of the covariates in the OLS model are reported in Table 4. As can be seen, each of the six test statistics is statistically significant; the covariates in the OLS model do not fully model the spatial dependence diagnosed in the data. Given the limitations of the Moran's I and Kelejian-Robinson diagnostics for spatial dependence in OLS models, the choice of the proper spatial model should be based on the Lagrange Multiplier

[20]The OLS spatial Chow tests, not reported in Table 3 due to space limitations, indicate spatial heterogeneity. The OLS Chow tests are significant for the intercept, $\textit{Republican Vote}_{1928}$, $\textit{Republican Vote}_{1928} * \textit{Foreign-Born}$, and $\textit{Non-Voting}_{1928} * \textit{Foreign-Born}$ ($p < .05$). The Chow test on the full model also indicates behavioral heterogeneity ($p < .01$).

diagnostics. The robust LM test statistic for spatial lag dependence is larger than the statistic for spatial error dependence. As a result, a spatial lag model is indicated.

10.3 Instrumental Variables Spatial Lag Model

The Jarque-Bera diagnostic rejects the null of normality in the OLS errors ($p < .01$). As a consequence, an instrumental variables specification should be estimated. For instruments, I employ lagged versions of the six covariates in the model. I again model behavioral heterogeneity in parameters via a switching regression specification. Because the Breusch-Pagan diagnostic indicates heteroskedasticity ($p < .01$), I estimate the instrumental variables model with a group-wise heteroskedasticity variable in which the error variances are allowed to vary across the two spatial regimes in the data. The estimates from the instrumental variables spatial regime model are reported in column 6 of Table 3.

As can be seen from the estimate for the spatial autoregressive parameter, ρ, there is significant spatial lag dependence in the presence of covariates. Accounting for this spatial dependence produces some noticeable differences in the sizes of the effects of the covariates in the model. Most strikingly, the counterintuitive negative effect of the $\textit{Non-Voting}_{1928} * \textit{Foreign-Born}$ interaction disappears in the instrumental variables specification, failing to reach statistical significance. Thus, had we ignored the spatial dependence in the data and estimated a standard OLS model, as we typically do, we would have concluded that the joint presence of two groups with weak prior attachments to the political parties, non-voters and immigrants, impeded shifts toward the Democratic party. By modeling the spatial dependence in changes in support for the Democrats between 1928 and 1932, we can see that this negative effect is merely a methodological artifact from ignoring the spatial dependence in the OLS specification.

Modeling the spatial dependence through an instrumental variables specification also affects our understanding of the $\textit{Republican Vote}_{1928} * \textit{Foreign-Born}$ interaction. The OLS model indicates no effect of such an interaction on changes in Democratic support. The instrumental variables model, in contrast, indicates that there is an additional gain for the Democrats in the joint presence of populations particularly likely to experience conversion and mobilization during this period. Although it is impossible with aggregate data to determine the contributions of mobilization and conversion to the upsurge in Democratic support, the instrumental variables estimates are consistent with the presence of both forms of realignment dynamics. In contrast to the OLS estimates, the instrumental variables estimates give no reason to believe that past partisanship or habitual non-voting impeded the shift to the Democrats in the Plains.[21]

11 Conclusion

Since the time of Galton, political scientists have been aware that spatial clustering in political phenomena may be produced by two quite distinct processes: diffusion processes, or, alternatively, the independent occurrences of the phenomena among neighboring units. In recent decades, scholars have been aware of the methodological problems posed by this spatial dependence. Until recently, however, scholars lacked a methodological approach that could simultaneously speak to Galton's problem and address the hurdles to inference posed by spatially dependent data. Spatial econometrics now provides scholars a rigorous method for addressing the spatial dimension that is central to many of our theories of political behaviors, processes, and events.

This paper has sought to highlight both the importance and the ease of modeling the spatial dependence in our data. As the Monte Carlo and empirical results demonstrate, spatially dependent data pose significant challenges to inference that are not easily addressed by techniques developed to handle the more familiar serial dependence in time series analysis. Instead, the researcher working with spatially dependent data will often wish to employ spatial econometrics to model this spatial autocorrelation. The diagnosis and modeling of this spatial autocorrelation, moreover, is a simple three-step process.

The application of spatial econometrics is aided by two parallel developments that are occurring in the social sciences. First, the past decade has witnessed an explosion in the availability of geo-coded data in the social sciences. Today, institutions such as the Census Bureau and

[21] There is again evidence of spatial heterogeneity in the instrumental variables estimates. As can be seen in Table 3, the instrumental variables Chow tests are significant for the intercept, $\textit{Republican Vote}_{1928}$, $\textit{Republican Vote}_{1928} * \textit{Foreign-Born}$, and $\textit{Non-Voting}_{1928} * \textit{Foreign-Born}$. The Chow test on the full model also indicates behavioral heterogeneity ($p < .01$).

state governments provide free downloads of geo-coded data and corresponding ESRI shape files (which significantly ease the creation of spatial weights matrices). Additionally, spatial econometric estimators are increasingly being incorporated in statistical software. This includes packages dedicated to spatial econometrics, such as SpaceStat and GeoDa. It also includes general purpose packages with which political scientists are already familiar, such as R, Stata, and WinBUGS.

As researchers, we routinely test for violations of modeling assumptions such as non-normality or heteroskedasticity. But while time series analysts routinely test for dependence in their data, testing for spatial dependence in cross-sectional political science research is rare. Happily, researchers can now employ easily applied methods and tools to diagnose and model the spatial dependence that is predicted by our theories. In doing so, researchers are likely to draw renewed attention to the interactions and interdependence that are at the core of many of the political phenomena we examine.

References

Andersen, Kristi. 1979. *The Creation of a Democratic Majority, 1928-1936*. Chicago: University of Chicago Press.

Anselin, Luc. 1988. *Spatial Econometrics: Methods and Models*. Kluwer: Dordrecht.

Anselin, Luc. 1990. "Spatial Dependence and Spatial Structural Instability in Applied Regression Analysis." *Journal of Regional Science* 30(2): 185-207.

Anselin, Luc. 1992. "SpaceStat Tutorial: A Workbook for Using SpaceStat in the Analysis of Spatial Data." Typescript. University of Illinois at Urbana-Champaign.

Anselin, Luc. 1995. "Local Indicators of Spatial Association – LISA." *Geographical Analysis* 27(2): 93-115.

Anselin, Luc. 2002. "Under the Hood: Issues in the Specification and Interpretation of Spatial Regression Models." *Agricultural Economics* 27: 247-267.

Anselin, Luc. 2005. "Exploring Spatial Data with GeoDa™: A Workbook." Typescript. University of Illinois at Urbana-Champaign and Center for Spatially Integrated Social Science.

Anselin, Luc, and Anil K. Bera. 1998. "Spatial Dependence in Linear Regression Models with an Introduction to Spatial Econometrics." In *Handbook of Applied Economic Statistics*, eds. Aman Ullah and David E.A. Giles. New York: Marcel Dekker.

Anselin, Luc, Anil K. Bera, Raymond Florax, and Mann J. Yoon. 1996. "Simple Diagnostic Tests for Spatial Dependence." *Regional Science and Urban Economics* 26: 77-104.

Anselin, Luc, and Serge Rey. 1991. "Properties of Tests for Spatial Dependence in Linear Regression Models." *Geographical Analysis* 23(2): 112-131.

Beck, Nathaniel, Kristian Gleditsch, and Kyle Beardsley. 2006. "Space is More than Geography: Using Spatial Econometrics in the Study of Political Economy." Forthcoming. *International Studies Quarterly*

Bera, Anil K., and Mann J. Yoon. 1993. "Specification testing with Locally Misspecified Alternatives." *Econometric Theory* 9: 649-658.

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

Brooks, Sarah M. 2004. "Explaining Capital Account Liberalization in Latin America: A Transitional Cost Approach." *World Politics* 56(3): 389-430.

Brown, Courtney. 1991. *Ballots of Tumult: A Portrait of Volatility in American Voting*. Ann Arbor: University of Michigan Press.

Busch, Marc L., and Eric Reinhardt. 2000. "Geography, International Trade, and Political Mobilization in U.S. Industries." *American Journal of Political Science* 44(4): 703-719.

Cardoso, Fernando Henrique, and Enzo Faletto. 1979. *Dependency and Development in Latin America*. Berkeley, CA: University of California Press.

Cliff, A.D., and J.K. Ord. 1981. *Spatial Processes: Models and Applications*. London: Pion.

Fotheringham, A.S., M.E. Charlton, and C. Brunsdon. 1998. "Geographically Weighted Regression: A Natural Evolution of the Expansion Method for Spatial Data Analysis." *Environment and Planning A* 30: 1905-1927.

Franzese, Robert J., Jr., and Jude C. Hays. 2003. "Modeling Spatial Relationships in International and Comparative Political Economy: An Application to Globalization and Capital Account Taxation in Developed Democracies." Typescript. University of Michigan.

Freeman, John R. 1989. "Systematic Sampling, Temporal Aggregation and the Study of Political Relationships." *Political Analysis* 1: 61-98.

Getis, Arthur, and J.K. Ord. 1992. "The Analysis of Spatial Association by Use of Distance Statistics." *Geographical Analysis* 24(3): 189-206.

Gimpel, James G., and Wendy K. Tam Cho. 2004. "The Persistence of White Ethnicity in New England Politics." *Political Geography* 23: 987-1008.

Gleditsch, Kristian Skrede. *All International Politics is Local: The Diffusion of Conflict, Integration, and Democratization.* Ann Arbor, MI: University of Michigan Press.

Gleditsch, Kristian Skrede, and Michael D. Ward. 2000. "Peace and War in Time and Space: The Role of Democratization." *International Studies Quarterly* 43: 1-29.

Greene, William H. 2003. *Econometric Analysis*, 5th Ed. Upper Saddle River, NJ: Prentice Hall.

Huckfeldt, Robert. 1986. *Politics in Context: Assimilation and Conflict in Urban Neighborhoods.* New York: Agathon Press.

Isard, Walter. 1956. *Location and Space-Economy: A General Theory Relating to Industrial Location, Market Areas, Land Use, Trade, and Urban Structure.* New York: Wiley.

Kelejian, Harry H., and Ingmar R. Prucha. 1997. "Estimation of Spatial Regression Models with Autoregressive Errors by Two-Stage Least Squares Procedures: A Serious Problem." *International Regional Science Review* 20(1 & 2): 103-111.

Kelejian, Harry H., and Ingmar R. Prucha. 1999. "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model." *International Economic Review* 40(2): 509-33.

LeSage, James P. 1997. "Bayesian Estimation of Spatial Autoregressive Models." *International Regional Science Review* 20: 113-129.

Most, Benjamin A., and Harvey Starr. 1980. "Diffusion, Reinforcement, Geopolitics, and the Spread of War." *American Political Science Review* 74(4): 932-46.

Nardulli, Peter F. 1995. "The Concept of a Critical Realignment, Electoral Behavior, and Political Change." *American Political Science Review* 89(1): 10-22.

O'Loughlin, John, Colin Flint, and Luc Anselin. 1994. "The Geography of the Nazi Vote: Context, Confession and Class in the Reichstag Election of 1930." *Annals of the Association of American Geographers* 84: 351-380.

Ord, J.K., and Arthur Getis. 1995. "Local Spatial Autocorrelation Statistics: Distributional Issues and an Application." *Geographical Analysis* 27(4): 286-306.

Rudra, Nita. 2002. "Globalization and the Decline of the Welfare State in Less-Developed Countries." *International Organization* 56(2): 411-445.

Shin, Michael, and John Agnew. 2002. "The Geography of Party Replacement in Italy, 1987-1996." *Political Geography* 21: 221-242.

Starr, Harvey. 1991. "Democratic Dominoes: Diffusion Approaches to the Spread of Democracy in the International System." *Journal of Conflict Resolution* 35(2): 356-381.

Starr, Harvey, and Benjamin A. Most. 1983. "Contagion and Border Effects on Contemporary African Conflict." *Comparative Political Studies* 16(February): 92-117.

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

Tylor, Edward B. 1889. "On a Method of Investigating the Development of Institutions; Applied to Laws of Marriage and Descent." *The Journal of the Anthropological Institute of Great Britain and Ireland* 18: 245-272.

Vasquez, John A. 1995. "Why Do Neighbors Fight? Proximity, Interaction, or Territoriality." *Journal of Peace Research* 32(3): 277-293.

Table 1: Bias of the OLS Estimator with Omitted Spatially Lagged Dependent Variable

	ρ									
N	-.9	-.7	-.5	-.3	-.1	.1	.3	.5	.7	.9
25	.09	.05	.02	.01	.00	.01	.03	.07	.15	.35
100	.12	.07	.04	.01	.00	.00	.01	.04	.11	.29
400	.11	.06	.04	.02	.00	.00	.01	.04	.09	.25
900	.08	.04	.02	.00	.00	.00	.02	.06	.14	.36

Table 2: Ratio of OLS Standard Error to True Standard Error with Spatial Error Dependence

	λ									
N	-.9	-.7	-.5	-.3	-.1	.1	.3	.5	.7	.9
25	.87	.89	.95	.97	1.03	1.01	.98	.96	.85	.71
100	.85	.93	.98	.96	1.00	.98	.96	.97	.86	.64
400	.86	.93	.95	1.02	1.00	.97	.98	.92	.85	.59
900	.87	.92	.93	1.02	.98	1.00	.96	.92	.83	.57

Figure 1: Change in Democratic Vote, 1928-1932

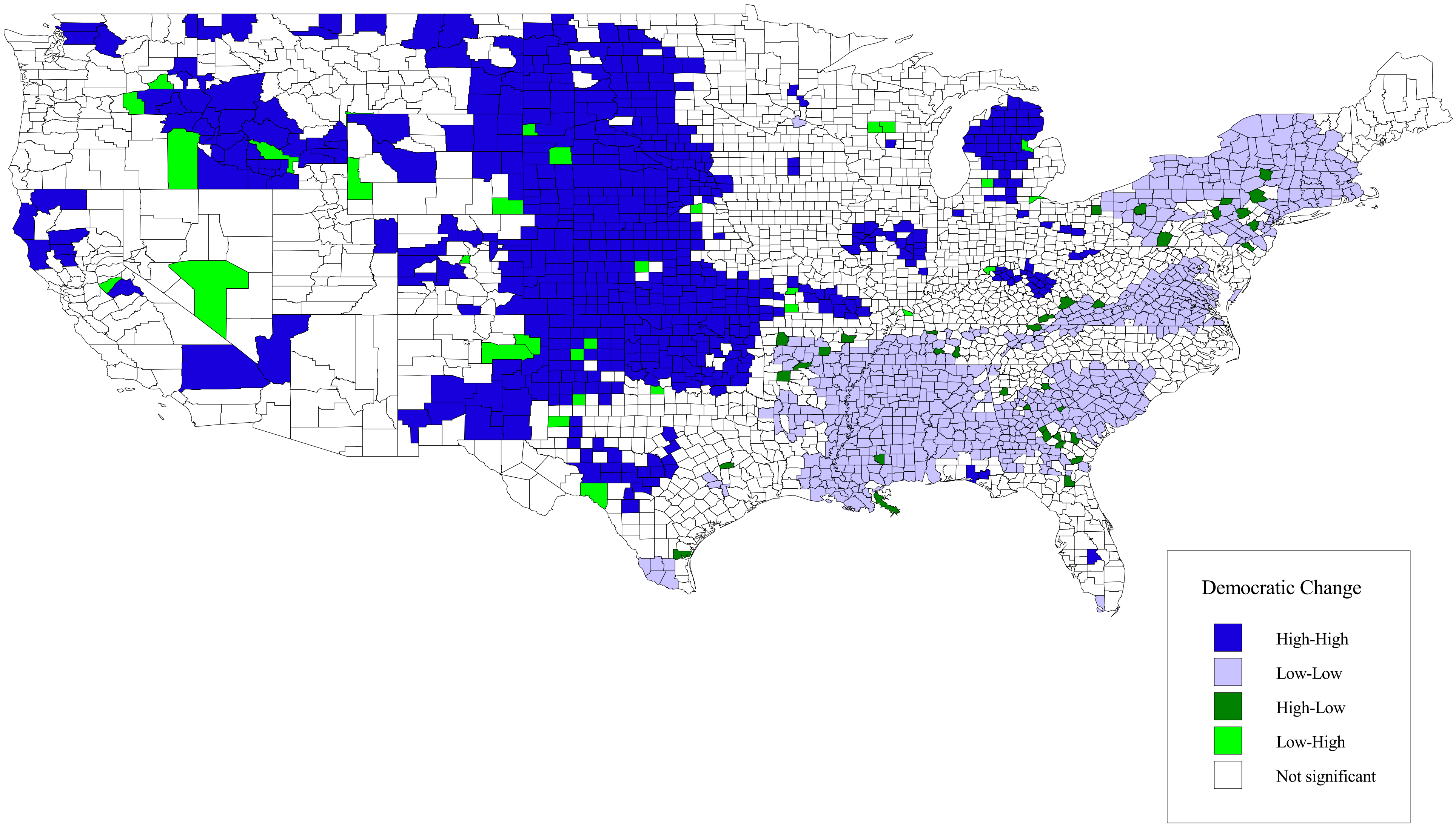

Table 3: OLS and IV Spatial Lag Estimates for Change in Democratic Vote, 1928-1932

Section	Covariate	Estimator	Estimates	Estimator	Estimates	IV Chow
Plains	Intercept	OLS	-.10**	IV Lag	-.10**	11.30**
			(.03)		(.02)	
South	Intercept	OLS	-.24**	IV Lag	-.22**	
			(.03)		(.03)	
Plains	Republican Vote_{1928}	OLS	.32**	IV Lag	.20**	30.33**
			(.05)		(.04)	
South	Republican Vote_{1928}	OLS	.69**	IV Lag	.50**	
			(.05)		(.04)	
Plains	Non-Voting_{1928}	OLS	.33**	IV Lag	.22**	.13
			(.06)		(.04)	
South	Non-Voting_{1928}	OLS	.27**	IV Lag	.24**	
			(.03)		(.03)	
Plains	Foreign-Born	OLS	.60	IV Lag	.02	.00
			(.32)		(.25)	
South	Foreign-Born	OLS	.57	IV Lag	.04	
			(.81)		(.70)	
Plains	Rep.Vote_{1928}*Foreign-Born	OLS	.82	IV Lag	.90*	7.15**
			(.53)		(.40)	
South	Rep.Vote_{1928}*Foreign-Born	OLS	-5.19**	IV Lag	-2.39*	
			(1.32)		(1.17)	
Plains	Non-Voting_{1928}*Foreign-Born	OLS	-2.25**	IV Lag	-.84	5.75*
			(.68)		(.52)	
South	Non-Voting_{1928}*Foreign-Born	OLS	3.46**	IV Lag	1.58	
			(.96)		(.84)	
Plains	Rural	OLS	.04**	IV Lag	.03**	.72
			(.01)		(.01)	
South	Rural	OLS	.05**	IV Lag	.02**	
			(.01)		(.01)	
	ρ			IV Lag	.55**	
					(.05)	
	Model Chow					50.78**

$N = 1651$

* $p < .05$, ** $p < .01$

Table 4: Diagnostics for Spatial Dependence in OLS Model

Moran's I (error)	26.24*
Kelejian-Robinson (error)	625.40*
Lagrange Multiplier (error)	660.40*
Robust Lagrange Multiplier (error)	35.15*
Lagrange Multiplier (lag)	691.30*
Robust Lagrange Multiplier (lag)	66.05*

* $p < .01$