

*This article deals with linear models for which data have been aggregated over well-defined geographic areas. Such data may be generated by spatial processes, and these may be represented in the form of spatial autocorrelation in the disturbance term or directly in the form of a spatial effect. This article details the derivation of Ord's (1975) MLE procedure for the spatial disturbances model and contrasts it with this MLE procedure for the spatial effects model. These alternative model specifications and estimation procedures are then illustrated by a variety of examples. These MLE procedures for the spatial models are also contrasted with conventional regression procedures (which ignored geographical space). If there is spatial autocorrelation present, an MLE procedure is preferable.*

## Linear Models with Spatially Distributed Data Spatial Disturbances or Spatial Effects?

PATRICK DOREIAN

*University of Pittsburgh*

### INTRODUCTION

When social processes are analyzed with conventional linear models and the data used for estimating these models have been aggregated for politically or administratively defined areas, for example, states, counties, or census tracts, the geography of the social process has been implicitly retained (Doreian, forthcoming). Empirical examples where linear relations have been estimated using data aggregated over well-defined geographical areas that comprise some region can be found in Inverarity

---

**AUTHOR'S NOTE:** *The estimation throughout the article is done in REPOMAT, a matrix algebra package designed by Norman P. Hummon. The programming for the spatial autocorrelation statistics was done by Philip Sidel of the Social Science Computing Center of the University of Pittsburgh. I am grateful to both for this programming. Comments made by one of the anonymous reviewers were particularly helpful and are gratefully acknowledged here.*

(1976), Ragin (1977), Chirot and Ragin (1975), Frisbie and Poston (1975), Matthews and Prothro (1963), Salamon and Van Evera (1973), Kernal (1973), Capecchi and Galli (1969), Mitchell (1969), and Doreian and Hummon (1976). Depending upon the substantive context, the combination of a data form that retains geographical space and the use of linear relations may lead to methodological problems that have to be considered. Essentially, the methodological problems hinge upon the issue of whether or not observations for a variable at one point of geographical space are interdependent with other observations for that variable at other points in geographical space. If there is such an interdependency, then the conventional methods for estimating linear equations become problematic and, in such instances, alternative procedures have to be sought. As the foregoing list of examples is far from exhaustive, and there are, in the literature, demonstrated instances of spatially distributed interdependencies (Ord, 1975; Doreian, forthcoming), these methodological problems are not hypothetical ones.

This article will address the following issues: (i) the representation of geographical space, (ii) determining whether or not spatial interdependencies exist, (ii) formulating alternative linear models that incorporate geographical space, and (iv) the estimation of these alternative models. The basic starting point for the entire discussion is the specification of the conventional linear population regression function;

$$Y = X\beta + \epsilon \quad [1]$$

where

$$E[\epsilon] = 0, E[\epsilon\epsilon'] = \sigma^2 I \quad [2]$$

with  $\epsilon$  being multivariate normal. Given a set of observations on  $Y$ , for example,  $y$ , and observations on the exogenous variables and the specifications of the above model, the prime empirical task is to estimate the vector of parameters,  $\beta$ , together with estimates of the standard errors of these estimates.

## *REPRESENTING GEOGRAPHICAL SPACE*

There are two broad representational strategies: measurements of distances between geographical locations within a region and partitioning the region into areas. Of course, these may be combined when distances between points (e.g., centroids or administrative centers) representing the areas are used. In this article, the focus is solely upon the partitioning strategy. Suppose a region,  $R$ , is partitioned into  $N$  mutually exclusive (and exhaustive) areas and that data exist for the  $N$  areas for all the variables of interest in a model. Doreian and Hummon (1976: 117-125) provide a general discussion of a matrix representation of geographical space that was extended by Doreian (forthcoming).

Consider the relation of adjacency. It is straightforward to determine if areas are adjacent to each other or not. Let  $S = [s_{ij}]$  be an  $(N \times N)$  matrix where  $s_{ij}$  is one if area  $i$  is adjacent to area  $j$ , and zero otherwise. Throughout, the  $s_{ii}$  are taken to be zero. The adjacency characteristics of  $R$  are completely specified in terms of  $S$ . The entries of  $S$  are either zero or one. However, a more general set of entries can be constructed where the entries can be viewed as weights. For example, let  $s_i$  be the row sum for the  $i^{\text{th}}$  row of  $S$ . Then a matrix,  $W = [w_{ij}]$ , can be constructed where  $w_{ij} = s_{ij}/s_i$ . The entries of  $W$  lie between zero and one (inclusive, although  $w_{ij} = 1$  is only possible for a pair of mutually adjacent, but otherwise disconnected, areas) and are proportions based upon the number of other areas adjacent to a specific area. This particular weighting scheme will be used throughout the article.<sup>1</sup> In the case of the Huk rebellion (Mitchell, 1969), considered below, control of an area by the rebels or the government forces has immediate consequences for adjacent areas. If military hardware and troops can be moved into an occupied area, then this area can be used to gain control of adjacent areas. There are many other social processes that can spread through geographical space where adjacency is a key spatial characteristic. In some cases, adjacency may be viewed as a special case of accessibility, in which case accessibility can then be used as the spatial characteristic. In the abstract, if a

relation can be defined over the areas of a region, this relation can be used to define a matrix of weights,<sup>2</sup> and, in general, the spatial structure of a region can be specified by some matrix,  $W$ , of weights. The elements of  $W$  are nonzero for adjacent pairs of areas and are zero otherwise. While it is emphasized that substantive concerns must dictate the choice of the spatial property represented, the following discussion assumes, simply, that some spatial property can be represented by a  $(N \times N)$  matrix  $W$ .

### *DETECTING SPATIAL INTERDEPENDENCIES*

Among the assumptions contained in [2] it is one to the effect that  $\epsilon$  is not autocorrelated;  $\epsilon_i$  and  $\epsilon_j$  are uncorrelated for  $i \neq j$ . In the time series context, (temporal) autocorrelation has been well studied where there are procedures for detecting autocorrelation and estimation strategies that take into account diagnosed autocorrelation (see, for example, Box and Jenkins, 1970; Hibbs, 1974). Of course,  $Y$  and the  $X$ s may also be autocorrelated. Spatial autocorrelation of either a variable or a disturbance term is the situation where the observations of variables or the values of the disturbance term for different areas are not independent. Cliff and Ord (1973) have dealt extensively with this problem. Determining whether spatial autocorrelation exists is a technical issue: Moran (1950) proposed a test statistic which was modified by Dacey and generalized by Cliff and Ord (1973: 12) to

$$I = (N/T) (y'Wy/y'y) \quad [3]$$

for a spatially distributed variable,  $y$ , where  $N$  is the number of areas partitioning the region and  $T$  is the sum of the weights of some appropriate weighting matrix. Cliff and Ord (1973: 13-15, 29-33) establish the distribution theory for  $I$  in order to test for spatial autocorrelation by treating  $(I - E[I])/(V[I])^{1/2}$  as a standardized normal deviate with  $E$  and  $V$  the expected value and variance operators, respectively. They extend this (1973: 87-97) to deal with residuals from a regression analysis. A brief account of

this effort is contained in Appendix A. Using their formulae, it is possible to test for spatial autocorrelation either in a variable of interest or in a residual obtained from a regression analysis. Where spatial autocorrelation is detected, it is necessary to deal with the problems its presence entails. There are several ways of doing this, and these are discussed in the following section.

### LINEAR SPATIAL EQUATIONS

There are two alternatives to equations 1 and 2 whereby geographical space can be incorporated into the specification of the equation to be estimated.

The first alternative is the "spatial disturbances model," and is based directly upon the notion of spatial autocorrelation. The specification of the spatial disturbances model is given in equations 4, and 5, and 6:

$$Y = X\beta + \epsilon \quad [4]$$

$$\epsilon = \rho W\epsilon + \nu \quad [5]$$

$$E[\nu] = 0, E[\nu\nu'] = \sigma^2 I \quad [6]$$

with  $\nu$  being multivariate normal. The empirical task is to estimate  $\rho$ ,  $\beta$ ,  $\sigma^2$ , and the standard errors of these estimates. Here, the spatial autocorrelation is dealt with primarily as a technical problem: While there is a spatial process specified in [5], it is of secondary interest to the appropriate estimation of [4].

The second alternative is labeled a "spatial effects model" and is based on an argument that the values of  $Y$  are systematically related to values of  $Y$  in adjacent areas (Ord, 1975; Mitchell, 1969; Dorean, 1981). Here, spatial autocorrelation is dealt with both substantively and as a technical problem. If there is a well-defined spatial process for the endogenous variable, it can be included in the model directly. On the basis of a substantively meaningful specification of  $W$ , the following specification is used:

$$Y = \rho WY + X\beta + \epsilon \quad [7]$$

$$E[\epsilon] = 0, E[\epsilon\epsilon'] = \sigma^2 I \quad [8]$$

where  $\epsilon$  is multivariate normal. The parameter,  $\rho$ , is a spatial parameter. If it is significantly different from zero, a spatial process can be said to be operating, otherwise, there is no spatial process and the specification of equation 1 suffices.<sup>3</sup> For the spatial effects model also, the empirical task is to estimate  $\rho$ ,  $\beta$ ,  $\sigma^2$ , and the standard errors of these estimates.

Thus, one modeling choice facing the researcher is one among the conventional regression specification, the spatial disturbances model, and the spatial effects model. The choice is facilitated in two ways. The procedures for detaching spatial autocorrelation help decide whether or not the conventional regression specification is appropriate. If no spatial autocorrelation is present, then such a specification is warranted, but if spatial autocorrelation is present, then one of the other two specifications should be pursued.<sup>4</sup> The choice between the spatial disturbances model and the spatial effects model is primarily a substantive one.

### *MAXIMUM LIKELIHOOD PROCEDURES FOR SPATIAL MODELS*

Ord (1975) has provided MLE methods for both the spatial disturbances and the spatial effect models. Given the terseness of his presentation, Doreian (forthcoming) has provided derivations of the statistical properties of, and sociological examples of, the spatial effects model. This will be summarized after the following similar treatment of the spatial disturbances model.

As the  $\nu$  are multivariate normal, the joint likelihood function of the  $\nu$  is given by

$$L(\nu) = (1/\omega 2\pi)^{N/2} \exp[-(1/2\omega)\nu'\nu]$$

where  $\omega = \sigma^2$  to simplify notation. However, the  $\nu$  are not observed. The relation between  $\epsilon$  and  $\nu$  is given by  $\nu = (I - \rho W)\epsilon = A\epsilon$  where  $A = I - \rho W$ . The joint likelihood function of the  $\epsilon$  is given by

$$L(\epsilon) = |A| (1/\omega 2\pi)^{N/2} \exp[-1/2\omega(A\epsilon)'(A\epsilon)]$$

where  $|A|$  is the jacobian of the variable change from the  $\nu$  to the  $\epsilon$ , and the corresponding log-likelihood function is

$$l(\epsilon) = \text{const} - (N/2)\ln\omega - 1/2\omega [\epsilon'A'A\epsilon] + \ln|A| \quad [9]$$

Finally, as the  $Y$  are observed, rather than the  $\epsilon$ , the log-likelihood function (given  $Y = y$ ) to be maximized is

$$l(y) = \text{const} - (N/2)\ln\omega - 1/2\omega[y'A'Ay - 2\beta'X'A'Ay + \beta'X'A'AX\beta] + \ln|A| \quad [10]$$

From [10],

$$\partial l / \partial \beta = -1/2\omega[-2X'A'Ay + 2X'A'AX\beta] \quad [11]$$

Setting [11] to zero gives

$$\hat{\beta} = [X'A'AX]^{-1}X'A'Ay \quad [12]$$

Thus, if  $\rho$  is known,  $\hat{\beta}$  is obtained readily from [12], which amounts to a regression of  $Ay$  on  $AX$ . Minimizing [10] or [9] with respect to  $\omega$  is also straightforward:

$$\partial l / \partial \omega = -N/2\omega + 1/2\omega^2[\epsilon'A'A\epsilon] \quad [13]$$

Setting [13] to zero gives<sup>5</sup>

$$\hat{\omega} = \hat{\sigma}^2 = 1/N [\epsilon'A'A\epsilon] \quad [14]$$

With  $\hat{\beta}$  estimated, equation 14 provides  $\hat{\sigma}^2$ , but both depend on  $\rho$ . This parameter can be estimated by a direct search procedure.<sup>6</sup>

The search procedure is immensely simplified by Ord's observation that the determinant of  $A$  is given by (Ord, 1975: 121)

$$|A| = \prod_{i=1}^N (1 - \rho\lambda_i) \quad [15]$$

where  $\lambda_i$  are the eigenvalues of  $W$ . Substituting [14] back into [9] indicates that

$$l(y; \rho, \hat{\omega}, \hat{\beta}) = \text{const} - N/2 \ln \hat{\omega} + \ln |A|$$

From this, it is clear that  $\hat{\rho}$  is the value of  $\rho$  which minimizes

$$\ln \hat{\omega} - (2/N) \sum_{i=1}^N \ln(1 - \rho \lambda_i) \quad [16]$$

From [14], see Note 5,

$$N\hat{\omega} = y'A'Ay - y'A'AX\beta - \beta'X'A'Ay + \beta'X'A'AX\beta$$

Substituting into this equation for  $\hat{\beta}$  from [12] gives, upon simplification,

$$\begin{aligned} N\hat{\omega} &= y'A'Ay - y'A'AX(X'A'AX)^{-1}X'A'Ay \\ &= y'A'[I - AX((AX)'AX)^{-1}(AX)']Ay \\ &= y'A'PAy \end{aligned}$$

where  $P = I - (AX)((AX)'AX)^{-1}(AX)'$ . It follows, from substituting this into [16], that  $\hat{\rho}$  is the value of  $\rho$  that minimizes

$$\ln(y'A'PAy) - 2/N \sum_{i=1}^N \ln(1 - \rho \lambda_i) \quad [17]$$

The direct search is made on the values of [17] for the values of  $\rho$  in permitted range ( $\rho \leq 1/\lambda_{\max}$ ). With  $\hat{\rho}$  found by the direct search procedure, equations 12 and 14 can then be used to estimate  $\beta$  and  $\omega$ , respectively.

Given the specification of the spatial disturbances model, it is necessary to assess whether  $\rho$  is truly nonzero and to perform the usual inference procedures on  $\beta$ . The asymptotic variance-covariance matrix,  $V$ , of the parameter estimates, is given by

$$V^{-1} = -E \left[ \frac{\partial^2 \ell}{\partial \theta_r \partial \theta_s} \right] \quad [18]$$



where  $\theta_r$  and  $\theta_s$  denotes pairs of parameters being estimated (Kendall and Stuart, 1967: 55). The derivation of the asymptotic variance-covariance matrix,  $V$ , is given in Appendix B. With the definitions

$$\alpha = - \sum_{i=1}^N \lambda_i^2 / (1 - \rho \lambda_i)^2$$

and

$$B = WA^{-1}$$

$$V(\hat{\omega}, \hat{\rho}, \hat{\beta}) = \omega^2 \begin{bmatrix} N/2 & \omega \text{tr}(B) & 0' \\ \omega \text{tr}(B) & \omega^2 (\text{tr}(B'B) - \alpha) & 0' \\ 0 & 0 & \omega X'A'AX \end{bmatrix}^{-1} \quad [19]$$

where  $0$  is a column vector of zeros. Notice if either (i)  $\rho = 0$  or (ii)  $W = 0$ , the null matrix, then  $A = I$  and the specialized nonspatial outcomes are reached: The spatial disturbances model is a proper generalization of the conventional regression model.

The case for the spatial effects model can now be briefly summarized (for more detail, see Doreian, forthcoming). From equation 7,  $\epsilon = Ay - X\beta$  and the joint likelihood function for the  $y_i$  is given by

$$L(y) = |A| (1/2\pi\omega)^{N/2} \exp [1/2\omega[Ay - X\beta]' [Ay - X\beta]] \quad [20]$$

The log-likelihood function is given by

$$l(y) = \text{const} - (N/2) \ln \omega - 1/2\omega[y'A'Ay - 2\beta'X'Ay + \beta'X'X\beta] + 1/n |A| \quad [21]$$

which has to be minimized with respect to  $\rho$ ,  $\omega$ , and  $\beta$ . Minimizing  $l(y)$  with respect to  $\beta$  and  $\omega$  gives

$$\hat{\beta} = (X'X)^{-1}X'Ay \quad [22]$$

and

$$\hat{\omega} = (1/N)y'A'Ay \quad [23]$$

as the estimating equations for  $\beta$  and  $\omega$ , respectively. With these "estimates" substituted back into [21],

$$l(y; \rho, \hat{\omega}, \hat{\beta}) = \text{const} - N/2 \ln \hat{\omega} + \ln |A|$$

(exactly as before). It is straightforward to show that  $\rho$  is the value of  $\rho$  that minimizes

$$- (2/N) \sum_{i=1}^N \ln(1 - \rho\lambda_i) + \ln(y'My - 2\rho y'MWy + \rho^2(Wy)'MWy) \quad [24]$$

where  $M = I - X(X'X)^{-1}X'$  is an idempotent and symmetric matrix, and this minimization is also done via a direct search procedure. The variance-covariance matrix for these estimators is

$$V(\hat{\omega}, \hat{\rho}, \hat{\beta}) = \omega^2 \begin{bmatrix} N/2 & \omega \text{tr}(B) & 0 \\ \omega \text{tr}(B) & \omega^2 \text{tr}(B'B) + \omega \beta' X' B' B X \beta - \alpha \omega^2 \omega X' B X \beta \\ 0 & \omega X' B X \beta & \omega X' X \end{bmatrix}^{-1} \quad [25]$$

where, as before,

$$B = WA^{-1}$$

$$\alpha = - \sum_{i=1}^N \lambda_i^2 / (1 - \rho\lambda_i)^2$$

and 0 is a column vector of zeros.

Thus, for all three alternatives of the linear model specification, there are maximum-likelihood estimation equations and the formulae for the variance-covariance matrices for these estimates. We turn now to consider some of their empirical applications.

## *EMPIRICAL EXAMPLES*

### *A MODEL OF INSURGENCY*

The first example concerns the Huk insurgency in the Philippines studied by Mitchell (1969) and Dorean and Hummon (1976). The list of exogenous variables involved is: the percentage of the population speaking the Pampangan dialect, *P*; farmers as a percentage of the population, *FMP*; owners as a percentage of the population, *OWN*; percentage of cultivated land given to sugar production, *SGR*; presence of mountains (a dummy), *MNT*; and presence of swamps, *SWP* (another dummy variable). The dependent variable of interest is the level of Huk control operationalized as the percentage of barrios in a municipality under the control of the Huks. Huk control was located, for the most part, in those areas of central Luzon that were Pampangan in the sense that this ethnic group was dominant in those areas and that the majority of the population spoke that particular dialect. According to Mitchell (1969), there has been a historical cleavage between Pampangans and other ethnic groups. Further, this cleavage persisted, as there was a strong element of mistrust between the Pampangans and other neighboring ethnic groups.

It has been argued (see, for example, Mitchell, 1969), that agrarian political insurgency movements are often associated with certain kinds of economic conditions. Accordingly, certain economic variables were included in the list of exogenous variables that are concerned with land tenure and mode of production. If there is merit in the argument claiming that the desire for land by the peasantry is a motivating force in rebellion, Mitchell observes, then, other things being equal, it will be expected that barrios with few farmers owning their own land be under Huk control. Farmers as a percentage of the population and farm owners as a percentage of farmers are the land tenure variables used in the model. The production variable is the percentage of cultivated land given over to the production of sugar, as this particular agricultural product was important for the economy of the local area. Further, there was a preponder-

ance of cane estates where there were many farm laborers and migrant workers who, it could be argued, would develop sympathy for an agrarian movement. In addition, some purely geographical variables have been included in the light of the argument that guerrilla activity is facilitated by geographical terrain that is difficult for government forces to police. To this end, a dummy variable representing mountainous terrain and a dummy variable representing swamp land were included in the model. Finally, Mitchell specified that the cultural variable, the percentage of the population speaking the Pampangan dialect, is used multiplicatively with the remaining exogenous variables.

In addition to specifying a relation between the level of Huk control and various cultural, economic, and geographical variables, it is possible to argue that the level of Huk control in one area is interdependent with the level of Huk control in adjacent areas (see Mitchell, 1969). Since this is the case, a model that has a spatial effect built into it, namely equation 7, would be an appropriate specification where the matrix,  $W$ , is the matrix of weights constructed from the adjacency matrix. If there is no such spatial interdependency as far as the level of Huk control is concerned, then the straightforward conventional population regression function, equation 1, would be specified. The former is the spatial effects model, and the latter is the nonspatial model.

Of course, a preliminary step is to examine whether or not the dependent variable of interest is spatially autocorrelated. When the spatial autocorrelation statistic,  $I$ , was calculated for these data, it was found to have a value of 0.76, and when the standard normal deviate was constructed from this, according to the formula in Appendix A, its value was found to be 7.81; this indicates that the level of Huk control is a spatially autocorrelated variable. Moreover, a regression of Huk control on the (multiplicative) exogenous variables specified in the model does not remove this spatial autocorrelation. Seemingly, then, the specification of the spatial effects equation is appropriate.

Table 1 shows the estimated equations for the nonspatial model, the spatial effects model, and also, for purposes of comparison, the spatial disturbances model. Each model has the

TABLE 1  
Estimation Outcomes for the Three Specifications of the Linear Model for the Huk Rebellion Example

<u>Non-Spatial Model</u>	
(OLS)	
$\hat{Y} = 1.15 + 3.79 P*FMP - 1.91 P*OWN + 0.46 P*SGR + 38.38 P*MNT + 17.17 P*SWP$	
(2.94) (10.94)	(0.44) (0.16) (7.02) (7.94)
$R^2 = .73$	
<u>Spatial Effects Model</u>	
$\hat{Y} = -0.88 + 0.47 WY + 2.27 P*FMP - 1.07 P*OWN + 0.18 P*SGR + 30.46 P*MNT + 12.43 P*SWP$	
(2.48) (0.11) (0.81) (0.38) (0.14) (5.89) (6.52)	
FIT = .80 $\hat{\sigma}^2 = 126.4$	
(23.9)	
<u>Spatial Disturbances Model</u>	
$\hat{Y} = 3.83 + 2.31 P*FMP - 1.11 P*OWN + 0.39 P*SGR + 28.38 P*MNT + 12.62 P*SWP$	
(4.43) (0.81) (0.43) (0.18) (6.62) (7.48)	
FIT = .71 $\hat{\rho} = 0.58$ $\hat{\sigma}^2 = 148.9$	
(0.13) (28.67)	

NOTE: Numbers in parentheses are standard errors.

coefficients of interest estimated together with the estimated standard errors for these coefficients and an indication of the quality of fit of the model.<sup>7</sup> In order to assess whether or not individual coefficients are significant, a two-tailed test has been employed with a significance level arbitrarily set at 0.05.<sup>8</sup> Several things are immediately apparent from a comparison of the first two panels of Table 1. First, all coefficient estimates for the nonspatial model are greater in magnitude than the corresponding estimated coefficients for the spatial effects model. Second, all estimates of the standard errors of the coefficients are higher for the nonspatial model than for the spatial effects model. This suggests serious deficiencies in the nonspatial model in the form of upward biases (in magnitude) in the coefficient estimates and in the estimates of the standard errors of these coefficient estimates. In this particular case, the two biases appear to be partially offsetting as far as inferential purposes are concerned, but nevertheless, they do point to serious deficiencies in the nonspatial model. When the two estimated forms are compared with inferential purposes in mind, it is clear that they lead to different inferences concerning those exogenous variables that are deemed important for predicting the level of insurgent control. Specifically, both the sugar cane production variable and the presence of swamp land (both defined interactively with the Pampangan dialect variable) are deemed to be significant for the nonspatial model, whereas they are not significant for the spatial effects model. To the extent that there is a spatial effect operative, the use of the nonspatial model also leads to serious inferential problems. On this basis, it appears that the use of a nonspatial model is inappropriate for spatially distributed data where spatial autocorrelation is present, and there are good arguments for using a spatial effects model.

In this particular example, it is clear that a spatial effects model is warranted. However, it is also instructive to examine the outcome of a specification of a spatial disturbances model. If we examine the coefficient estimates, we again see that the nonspatial model has coefficient estimates that are inflated relative to the coefficient estimates for a model that incorporates a spatial

disturbance term (see panel 3 of Table 1). When the estimates of the standard errors for the coefficient estimates are examined, there is not the same clear pattern that was observed for the spatial effects specification. Four of the standard error estimates are lower for the spatial disturbances model; two are higher than those for the nonspatial model. When inferential concerns are addressed, there is one difference that emerges from a comparison of the nonspatial model with the model that has a spatially autocorrelated disturbance term. The variable depicting the presence of swamp land is deemed to be significant under the nonspatial model whereas it is not under the spatial disturbances model. When the spatial effects model and the spatial disturbances model are compared, there is a single inferential difference concerning the production variable of sugar cane production. However, the choice between a spatial effects model and a spatial disturbances model is not one that is readily settled by examining the estimation outcomes from the two specifications. Rather, the choice concerning the way in which spatial autocorrelation is to be dealt with is made on substantive grounds. In this particular context, the spatial effects model seems more compelling than the spatial disturbances model. However, the comparison between the two in this context is instructive insofar as it shows that the two different strategies for dealing with spatial autocorrelation can also lead to different inferences concerning those exogenous variables that are deemed important for predicting the levels of the endogenous variable of interest.

#### *DETERMINANTS OF PRESIDENTIAL ELECTORAL SUPPORT*

The next set of examples is taken from a project exploring at the macro level the determinants of presidential voting behavior in the parishes of Louisiana.<sup>9</sup> Here, the dependent variable of interest is the percentage of the voting electorate who support particular Democratic presidential candidates.<sup>10</sup>

Howard and Grenier (1976) have, by means of factorial ecology, delineated several dimensions that are important charac-

teristics of Louisiana. Three particularly salient dimensions are: percentage black, B; percentage Catholic, C; and percentage urban, U. Each of these variables, in the aggregate, can be seen as predictors of electoral turnout and partisan electoral behavior in presidential elections. The overwhelming majority of the electorate is, and has been, registered with the Democratic Party. Until 1944, voting for the Democrats was seen as voting for the status quo and, in particular, white supremacy (Howard, 1971). With the adoption, in the national Democratic platform, of a mild, embryonic civil-rights plank, there was a good deal of defection from the Democratic ranks and support for the States' Rights candidacy of Thurmond. It is reasonable to ask whether the concentration of blacks in particular parishes (counties) is a predicting factor of the support (or lack of support) for particular candidates. The distinction has often been noted between Northern Louisiana and Southern Louisiana where one of the characteristics is given by the variable, percentage Catholic. The Southern parishes are characterized by a higher proportion of the Cajun population, whereas the Northern parishes have a preponderance of people with a Protestant affiliation. Again, if religion is an important characteristic differentiating the parishes, it is reasonable to ask whether or not this also is a predicting characteristic for presidential voting support. Throughout the period from the 1930s through to the 1970s, Louisiana underwent a period of transformation from a predominantly agricultural state to one characterized by higher levels of industrialization and urbanization. Again, it is reasonable to ask whether, in the face of this long-term trend in the changing composition of the population, percentage urban is a predictor of partisan political behavior. Another exogenous variable considered in these examples is a measure of black political equality which attempts to operationalize the extent to which blacks are enfranchised in relation to their numbers in a particular parish. There are further predictors that can be included, but the examples in this article will focus only on these four exogenous variables.

Given that 1948 saw the high water mark of support for the Democrat party, at least among white voters; that some presi-



dential elections were characterized by the presence of States' Rights candidates (for example, Thurmond in 1948 and Wallace in 1968); that religion became salient in the Kennedy candidacy; that there was, in 1965, the Civil Rights Voting Act and a consequent jump in black enfranchisement; it is unreasonable to expect that the parameterization of a model linking voting support for particular candidates to the exogenous variable would be fixed through time. Thus, it is not appropriate to use any of the pooled cross section and time-series procedures that would assume homogeneity of parameterization throughout the period for which cross sections were being pooled: It is much more reasonable to examine the elections individually and try to chart the parameter changes in the estimates through time. If this course of action is adopted, it becomes important that the specification of the models and the estimation procedures used lead to parameter estimates that are as precise as possible. In the light of the results for the Huk rebellion, it is clear that the nonspatial model is likely to produce parameter estimates that are inflated and also estimates of the standard errors of these parameter estimates that are inflated. It cannot be assured that the amount of inflation in these estimates would be fixed from period to period, and any comparison that we might make is compromised by imprecision in the estimating procedure.

While it is possible to estimate a model linking candidate support to the exogenous predictors for each of the elections, for example, from 1936-1976, the number of these empirical examples would be excessive for this article; therefore, only a few examples are considered here. At issue, however, is the precision of the estimates and also the inferences concerning which of the exogenous variables are predictors of presidential candidate support for particular elections.

As before, we are faced with a choice among three models: the conventional regression formulation that ignores space (the nonspatial model), the spatial effects model, and the spatial disturbances model. If either the intended variable of interest or the disturbance term are spatially autocorrelated, then the nonspatial model is not appropriate.<sup>11</sup> However, in contrast to the Huk

example, it is not immediately clear which of the two spatial models is more appropriate. In the case of predicting support for particular presidential candidates in elections in the state of Louisiana, there is room for debate whether there is a spatial effects process operative or whether the spatial autocorrelation ought to be dealt with as a technical problem through the specification of the spatial disturbances model. There are few, if any, attempts on the part of the leaders of one parish to impose an electoral outcome on adjacent parishes, but it is still possible that there is a spatial process operative.<sup>12</sup> On (weak) grounds, I would specify the spatial effects model in preference to the spatial disturbances model. In the following examples, all three specifications will be presented in order that their estimated outcomes may be compared.

The first election considered for purposes of illustration is the 1948 election, the results of which, for each of the possible specifications, are shown in Table 2. When the nonspatial model and the spatial effects model are compared, there are similar patterns of the Huk example as far as the magnitude of the estimates is concerned. All of the estimated coefficients for the nonspatial model are inflated relative to the coefficients for the spatial effects model—as are the estimates of the standard errors of the coefficient estimates. If elections were being compared across years with respect to the parameter estimates, this inflation could be important and would be grounds for preferring the spatial-effects model. As far as inference is concerned, the two specifications would lead the researcher to the same inference concerning the exogenous variables that are predictive of Democratic presidential support in 1948. It appears that the percentage black in a parish is negatively related to the percentage supporting the Democratic candidate<sup>13</sup>, whereas the level of black political equality is positively related to Democratic support. Neither the variables of percentage Catholic nor percentage urban have predictive utility for this election. When the spatial effects model and spatial disturbances models are compared, there is one sharp difference concerning the inference that is made about the exogenous variables that are predictors of the support for the

TABLE 2  
Estimation Outcomes for the Three Specifications of the Linear Model for  
Democratic Presidential Support in Louisiana in 1948

<u>Non-Spatial Model</u>	$\hat{Y} = 44.42 - 0.29 B - 0.07 C - 0.07 U + 0.40 BPE$ $(5.13) (0.10) (0.04) (0.08) (0.21)$			
	$R^2 = 0.18$			
<u>Spatial Effects Model</u>	$\hat{Y} = 20.05 + 0.61 WY - 0.18 B - 0.057 C - 0.05 U + 0.37 BPE$ $(5.91) (0.11) (0.08) (0.034) (0.07) (0.16)$			
	$FIT = 0.46 \quad \hat{\sigma}^2 = 97.27$ $(17.71)$			
<u>Spatial Disturbances Model</u>	$\hat{Y} = 41.47 - 0.17 B - 0.09 C - 0.06 U + 0.43 BPE$ $(5.73) (0.10) (0.06) (0.06) (0.16)$			
	$\hat{\rho} = 0.66 \quad FIT = 0.16 \quad \hat{\sigma}^2 = 93.84$ $(0.10) (17.23)$			

NOTE: Numbers in parentheses are standard errors.

Democratic presidential candidate in 1948. According to the spatial effects model, percentage black is an important predictor, whereas percentage white is not an important predictor of democratic presidential support under the spatial disturbances model. This substantive difference can be used as further evidence to support the preferences for the spatial effects specification as opposed to the spatial disturbances specification.<sup>14</sup> Given that this was the 1948 election, that the electorate was overwhelmingly white, and that at the national level the democratic party platform had a mild civil rights plank in it and was seen as betraying the white supremacy ideology, it is reasonable to predict that a sizable black community in a parish would act as a trigger to mobilize a white vote against the Democratic candidates and in favor of the States' Rights (white supremacist) candidacy of Thurmond.

This particular difference emphasizes that the choice between the spatial effects model and the spatial disturbances model is likely to entail more than the simple choice regarding the way in which the problem of spatial autocorrelation is handled. That is, while the arguments for or against either of the two spatial representations should rest primarily on substantive arguments, it has to be recognized that the models will not, in their estimated form, always be the same with respect to the inclusion of exogenous variables. The particular choice of the spatial representation does have consequence for the results stemming from the estimation of either spatial model.

The next example is furnished by the 1952 election, the results of which are shown in Table 3. There are no important differences with respect to the magnitude of the estimated coefficients; nor is there a great difference between the estimates of the standard errors of the estimated coefficients. The one point at which a difference emerges concerns whether the spatial effects model or the spatial disturbances model is the better approach in dealing with spatial autocorrelation. On the basis of testing for autocorrelation, both the dependent variable of interest and the residual left after ordinary least squares regression show spatial autocorrelation. When the spatial effects model is estimated, it is clear that the spatial coefficient for that model is not significant,

TABLE 3  
Estimation Outcomes for the Three Specifications of the Linear Model for  
Democratic Presidential Candidate Support in 1952

<u>Non-Spatial Model</u>	
$\hat{Y} = 55.39 - 0.06 B - 0.002 C - 0.12 U + 0.18 BPE$	
(5.18) (0.10) (0.03) (0.06) (0.06)	
$R^2 = 0.23$	
<u>Spatial Effects Model</u>	
$\hat{Y} = 41.62 + 0.26 WY - 0.06 B - .007 C - .12 U + 0.17 BPE$	
(9.73) (0.15) (0.09) (0.03) (0.05) (0.06)	
FIT = 0.27	$\hat{\sigma}^2 = 91.15$
	(16.21)
<u>Spatial Disturbances Model</u>	
$\hat{Y} = 56.24 - 0.11 B - .008 C - .10 U + .20 BPE$	
(5.10) (0.09) (0.04) (0.05) (0.06)	
$\hat{\rho} = 0.37$	FIT = 0.22
(0.15)	$\hat{\sigma}^2 = 87.33$
	(15.65)

NOTE: Numbers in parentheses are standard errors.

whereas the estimated spatial coefficient for the spatial disturbances model is significant. This outcome poses a dilemma for the researcher who opted, on theoretical grounds, for the spatial effects model. It would appear that, from the estimation of such a model, the incorporation of the spatial effects term is not warranted, and this would point the researcher back to the nonspatial model. On the other hand, when the spatial autocorrelation is treated as a technical problem, the spatial disturbances model seems an adequate way of dealing with this particular problem, and the spatial disturbances model appears preferable. One resolution of this dilemma would be to fit a model that has both a spatial effects term and a spatial disturbances term simultaneously. This is an agenda item pointed to by this example, and work on developing the estimation procedure for such a combined spatial effects and spatial disturbances model is currently under way.

As far as inference is concerned, each of the alternative specifications does point to the same substantive outcomes. Given that the spatial term is not significant for the spatial effects model and that  $\hat{\rho}$  for the spatial disturbances model is not very large, this is not surprising. All of the specifications agree that percentage black and percentage Catholic in a parish are not predictors of the level of support for the Democratic presidential candidate, whereas percent urban and the measure of black political equality are predictor of that support. As far as a comparison with the preceding elections is concerned, it would appear that for the first Eisenhower election percentage black is no longer a predictor of Democratic candidate support whereas percentage urban now is (with a negative coefficient). While black political equality remains a predictor of democratic presidential candidates support, the absolute values of the coefficient have declined during the period from the 1948 election to the 1952 election.

When the 1956 election was examined, the only difference that emerged was, again, between the spatial effects model and the spatial disturbances model. Here, the spatial coefficient for the spatial effects model was significant whereas the estimate of the

spatial parameter for the spatial disturbances model was not. This would suggest that, in this instance, the spatial effects approach, both on substantive and methodological grounds, was the better way of dealing with spatial autocorrelation. The final example included is that of the 1960 election, in which no differences appear at all between the various specifications of a linear model. The results of the estimation of these alternative specifications are shown in Table 4. In comparison to the previous two elections, we now have percentage Catholic as a major predictor of Democratic presidential support. Given that Kennedy was the Democratic candidate, was Catholic, and, moreover, that religion became an issue in the 1960 presidential election, it is not surprising that the parishes with high percentages of Catholics in them gave solid support to the Catholic presidential candidate. Both percentage urban and percentage black political equality remain as predictors of the level of Democratic presidential candidate support. The magnitude of the coefficient for the percentage urban area remains, as before, negative, whereas the value of the coefficient for black political equality is somewhat higher than in the two preceding elections in the 1950s. Given that both the spatial effects and the spatial disturbances are clearly present, this would-be model could be improved by the incorporation of spatial effects and spatial disturbances simultaneously.

### *DISCUSSION*

Given that a linear relationship between an endogenous variable and a set of exogenous variables is specified and that either the endogenous variable of interest or a disturbance term is spatially autocorrelated, some form of a spatial model should be considered. The nonspatial model estimated by conventional regression procedures is not a reliable representation and should be avoided when there is a spatial phenomenon to be analyzed. Whether the researcher chooses the spatial effects model, the spatial disturbances model, or a combination of the two is a theoretical decision. Further, as the examples have indicated,

**TABLE 4**  
**Estimation Outcomes for the Three Specifications of the Linear Model for**  
**Democratic Presidential Candidate Support in 1960**

Non-Spatial Model

$$\hat{Y} = 21.03 + 0.01 B + 0.30 C - 0.11 U + 0.39 BPE$$

(4.40) (0.08) (0.04) (0.04) (0.06)

$$R^2 = .88$$

Spatial Effects Model

$$\hat{Y} = 13.78 + 0.31 WY - .004 B + 0.22 C - 0.10 U + 0.29 BPE$$

(4.67) (0.09) (0.07) (0.05) (0.04) (0.06)

$$FIT = .90 \quad \hat{\sigma}^2 = 49.78$$

(8.83)

Spatial Disturbances Model

$$\hat{Y} = 26.93 - 0.11 B + 0.37 C - 0.07 U + 0.25 BPE$$

(4.50) (0.07) (0.06) (0.03) (0.06)

$$\hat{\rho} = 0.68 \quad FIT = 0.86 \quad \hat{\sigma}^2 = 39.91$$

(0.10) (7.34)

NOTE: Numbers in parentheses are standard errors.



inferences concerning the specific roles of exogenous variables may vary according to the spatial model chosen.

There was no clear-cut pattern concerning the numerical differences between the estimates for the spatial disturbances and the spatial effects models. The extent to which the two types of spatial phenomena, and their representation, can be confused can be explored via Monte Carlo simulations: Data can be generated under either regime for variations in  $\rho$  and  $\sigma^2$  and estimated according to each specification to see the extent to which the models can be genuinely distinguished in an empirical context. This is a topic for further investigation. Another avenue of investigation raised in the article is the examination of a specification that includes both a spatial effect and spatial autocorrelation. The estimation of such a model is likely to be quite complex, and it is currently being pursued.

It is clear that for linear models employing spatially distributed data, attention must be paid to the spatial characteristics of the phenomena being studied. Various strategies have been proposed for doing this together with maximum likelihood methods that should prove of increasing value in the analysis of geographically distributed social phenomena.

## APPENDIX A

### TESTING FOR SPATIAL AUTOCORRELATION

$W$  is an appropriately defined matrix of spatial weights. Define

$$\begin{aligned} w_{i.} &= \sum_j w_{ij} & w_{.j} &= \sum_i w_{ij} \\ S_1 &= 1/2 \sum_{i \neq j} \sum (w_{ij} + w_{ji})^2 \\ S_2 &= \sum_i (w_{i.} + w_{.i})^2 \end{aligned}$$

The test statistic,  $I$ , defined by Cliff and Ord (1973) for testing the spatial autocorrelation of a variable,  $y$ , is

$$I = (N/T) (y'Wy/y'y) \quad [A.1]$$

where  $N$  is the number of areas and

$$T = \sum_{i \neq j} \sum w_{ij}$$

For this situation,  $E[I] = -1/N$  and  $V[I] = (N^2S_1 - NS_2 + 3T^2)/N(N-1)T^2$ , and the standardized normal deviate can be constructed.

Suppose a residual,  $\hat{\epsilon}$ , has been returned from a (nonspatial) regression analysis. In order to test for spatial autocorrelation of  $\hat{\epsilon}$ , the following test statistic is defined:

$$I = (N/T) (\hat{\epsilon}'W\hat{\epsilon}/\hat{\epsilon}'\hat{\epsilon}) \quad [A.2]$$

Cliff and Ord (1973) derive expression for  $E[I]$  and  $V[I]$  in order to construct a standardized normal deviate. Define

$$D = [d_{ij}] = X(X'X)^{-1}X'$$

Then,

$$E[I] = -(T + N \sum_{i \neq j} \sum w_{ij} d_{ij}) / (N - K)T \quad [A.3]$$

where there are  $K$  exogenous variables (including the column of 1's for the intercept). The expression for

$$V[I] = \frac{N}{(N-K)T^2} \left[ \left[ \frac{N^2S_1 - NS_2 + 3T^2}{N^2} \right] + \frac{1}{N} \sum_i \sum_j (w_{i.} + w_{.j}) \right. \\ \left. (w_{j.} + w_{.j}) d_{ij} + \right. \\ \left. 2(\sum_{i \neq j} \sum w_{ij} d_{ij})^2 - [\sum_{i \neq j} \sum \sum (w_{ik} + w_{ki}) (w_{jk} + w_{kj}) d_{ij} + \sum_{i \neq j} \sum (w_{ij} + w_{ji})^2 d_{ii}] + \right. \\ \left. \frac{1}{N} \sum_{i \neq j} \sum \sum (w_{ij} + w_{ji}) (w_{ik} + w_{ki}) (d_{ii} d_{jk} - d_{ij} d_{ik}) \right] \frac{1}{(N-K)^2} \quad [A.4]$$

## APPENDIX B

DERIVATION OF THE VARIANCE-COVARIANCE  
MATRIX OF THE MLE ESTIMATORS

In order to obtain  $V$ , it is necessary to obtain the second partial derivatives of  $\ell(y)$  with respect to the parameters being estimated. First, some preliminary remarks: (i)  $\partial\epsilon/\partial\beta = X$ ; (ii) for general quadratics  $\partial(h'\beta)/\partial\beta = h$ ; and  $\text{tr}(g'Sh) = \text{tr}(Shg')$  for any matrix  $S$  and (conformable) vectors  $g$  and  $h$ .  
From 9

$$\partial\ell/\partial\rho = \partial/\partial\rho \ln |A| - 1/2\omega \partial/\partial\rho \epsilon' A' A \epsilon \quad [B.1]$$

This can be taken in two steps:

(i) from the form of  $|A|$  given in 15,

$$\partial/\partial\rho \ln |A| = \partial/\partial\rho \sum_{i=1}^N \ln(1 - \rho\lambda_i) = - \sum_{i=1}^N \lambda_i / (1 - \rho\lambda_i)$$

(ii)

$$\begin{aligned} \partial/\partial\rho \epsilon' A' A \epsilon &= \partial/\partial\rho (\epsilon' \epsilon - 2\rho\epsilon' W + \rho^2 \epsilon' W' W \epsilon) \\ &= -2\epsilon' W \epsilon + 2\rho\epsilon' W' W \epsilon \\ &= -2\epsilon' [I - \rho W'] W \epsilon \\ &= -2\nu' W \epsilon \end{aligned}$$

These two results are substituted into [B.1] to give

$$\partial\ell/\partial\rho = - \sum \lambda_i / (1 - \rho\lambda_i) + 1/\omega \nu' W \epsilon \quad [B.2]$$

Equations 11, 13, and B.2 are the basis for obtaining the required second partial derivatives.

From 11

$$\partial^2\ell/\partial\beta^2 = - (1/\omega) X' A' A X \quad [B.3]$$

From 13;

$$\begin{aligned} (a) \quad \partial^2\ell/\partial\beta^2 &= N/2\omega^2 - 2/2\omega^3 \epsilon' A' A \epsilon \\ &= 1/2\omega^2 [N - 2/\omega \epsilon' A' A \epsilon] \\ &= 1/2\omega^2 [N - 2/\omega N \omega] \quad (\text{at the minimum}) \\ &= - N/2\omega^2 \quad [B.4] \end{aligned}$$

$$\begin{aligned}
 \text{(b)} \quad \partial^2 \ell / \partial \omega \partial \rho &= 1/2 \omega^2 \partial / \partial \rho [\epsilon' A' A \epsilon] \\
 &= -1/2 \omega^2 2 \nu' W \epsilon \quad (\text{from ii above}) \\
 &= -1/\omega^2 \nu' W \epsilon \quad [\text{B.5}]
 \end{aligned}$$

$$\begin{aligned}
 \text{(c)} \quad \partial^2 \ell / \partial \omega \partial \beta &= 1/2 \omega^2 \partial / \partial \beta [\epsilon' A' A \epsilon] \\
 &= (1/2 \omega^2) \partial / \partial \beta [(y - X\beta)' A' A (y - X\beta)] \\
 &= (1/2 \omega^2) [-2X' A' A y + 2X' A' A X \beta] \\
 &= 0 \text{ at the minimum} \quad [\text{B.6}]
 \end{aligned}$$

From B.2;

$$\begin{aligned}
 \text{(a)} \quad \partial^2 \ell / \partial \rho^2 &= \partial^2 / \partial \rho^2 (\ln |A|) + 1/\omega \partial / \partial \rho (\nu' W \epsilon) \\
 &= \partial / \partial \rho \left( -\sum_{i=1}^N \lambda_i / (1 - \rho \lambda_i) \right) 1/\omega \partial / \partial \rho (\epsilon' W \epsilon - \rho \epsilon' W' W \epsilon) \\
 &= -\sum_{i=1}^N \lambda_i^2 / (1 - \rho \lambda_i)^2 - (1/\omega) \epsilon' W' W \epsilon \\
 &= \alpha - (1/\omega) \epsilon' W' W \epsilon \quad [\text{B.7}]
 \end{aligned}$$

where  $\alpha = -\sum_{i=1}^N \lambda_i^2 / (1 - \rho \lambda_i)^2$

$$\begin{aligned}
 \text{(b)} \quad \partial^2 \ell / \partial \rho \partial \beta &= (1/\omega) \partial / \partial \beta (\nu' W \epsilon) \\
 &= (1/\omega) \partial / \partial \beta (\nu' W (y - X\beta)) \\
 &= (-1/\omega) \nu' W X \quad [\text{B.8}]
 \end{aligned}$$

We now take the expected values of these second derivatives:

From B.4

$$E[\partial^2 \ell / \partial \omega^2] = -N/2 \omega^2$$

From B.7

$$\begin{aligned}
 E[\partial^2 \ell / \partial \rho^2] &= \alpha - (1/\omega) E[\epsilon' W' W \epsilon] \\
 &= \alpha - (1/\omega) E[\nu' A^{-1} W' W A^{-1} \nu] \\
 &= \alpha - (1/\omega) E[\text{tr} \nu' B' B \nu] \text{ with } B = W A^{-1} \\
 &= \alpha - (1/\omega) E[\text{tr} B' B \nu \nu'] \\
 &= \alpha - (1/\omega) \text{tr}(B' B) \sigma^2 I \\
 &= \alpha - \text{tr}(B' B)
 \end{aligned}$$

From B.3

$$E[\partial^2 \ell / \partial \beta^2] = -(1/\omega) X' A' A X$$

From B.5

$$\begin{aligned} E[\partial^2 \ell / \partial \omega \partial \rho] &= -(1/\omega^2) E[\nu' W \epsilon] \\ &= -(1/\omega^2) E[\nu' W A^{-1} \nu] \\ &= -1/\omega \operatorname{tr} B \end{aligned}$$

From B.6

$$E[\partial^2 \ell / \partial \omega \partial \beta] = 0, \text{ as a column vector}$$

From B.8

$$E[\partial^2 \ell / \partial \rho \partial \beta] = -1/\omega E[\nu' W X] = 0, \text{ as a column vector}$$

Substituting these expected values into [18] and noting the negative sign gives the variance-covariance matrix of the coefficient estimates:

$$V(\hat{\omega}, \hat{\rho}, \hat{\beta}) = \omega^2 \begin{bmatrix} N/2 & \omega \operatorname{tr}(B) & 0' \\ \omega \operatorname{tr}(B) & \omega^2 \operatorname{tr}(B'B) - \omega^2 \alpha & 0' \\ 0 & 0 & \omega X' A' A X \end{bmatrix}^{-1} \quad [\text{B.9 and 19}]$$

## NOTES

1. Other weighting schemes can be considered (see Mitchell 1969; Doreian, 1981). For example, if  $b_{ij}$  is the length of the common border shared by areas  $i$  and  $j$  and if  $b_{ii}$  denotes the total perimeter of area  $i$ , then a matrix of weights can be defined by  $w_{ij} = b_{ij}/b_{ii}$  for  $i \neq j$  together with  $w_{ii} = 0$ .

2. There are, potentially, a very large number of weighting schemes and spatial characteristics that can be selected for representation. Some will be more compelling and fruitful than others. The plethora of choices has led some researchers (for example, Arora & Brown, 1977) to abandon this approach to the specification of geographical space. Such an abandonment is premature (see Doreian, forthcoming).

3. If  $y$  is not spatially autocorrelated, the spatial-effects model would not be explored further. By the same argument, if  $\epsilon$  is obtained from estimating [1] by conventional regression procedures and is found not to be spatially autocorrelated, then the spatial disturbances model would not be explored further. In both cases, the model expression in [1] and [2] would suffice.

4. The problem of dealing simultaneously with spatial effects and spatial disturbances is much more complex and is not tackled in this paper.

5. Of course,  $\epsilon$  is not observed, but, as  $\hat{\epsilon} = Y - X\beta$ , equation 14 is equivalent to

$$\hat{\omega} = 1/N[y'A'Ay - 2\beta'X'A'Ay + \beta'X'A'AX\beta]$$

which is obtained from [10] rather than [9].

6. Ord (1975: 122) suggests an iterative procedure for obtaining  $\hat{\rho}$  which fared badly both for real data and artificially generated data.

7. This measure,  $FIT = r_{yy}^2$ , is used to give some indication of the fit of the model. It should not be interpreted as a measure of variance explained. A better measure is being pursued.

8. In cases where more than one coefficient is (separately) nonsignificant, the models were reestimated with variables dropped singly. The resulting estimated model was always the same as that obtained from dropping nonsignificant variables multiply. The analog of the partial joint F-test for these spatial models has not yet been developed.

9. The author, in collaboration with Charles Grenier, is pursuing this analysis of Louisiana political dynamics in which geographic space is included. The final models are unlikely to include the specific equations contained herein. These equations are being used to illustrate the strategies and problems of estimating models with spatial distributed data.

10. In general, models can be formulated to account for the support for the Democratic candidate, the Republican candidate, and States' Rights candidates. In this article, our attention is confined only to some equations that predict the level of support for the Democratic candidate.

11. The dependent variables considered here are all spatially autocorrelated as indicated by the value of  $I$ , and regression of them on their corresponding exogenous variables does not completely remove this spatial autocorrelation.

12. It could be further argued that a spatial effects process might be more likely in an era before mass television and radio communication and would fade through time as the mass media became a more uniform source of information across geographic regions.

13. Note, however, that we cannot make statements concerning individual voting behavior. One interpretation consistent with the negative coefficient of percentage black is that, when the size of the black population was higher, the white vote was mobilized against the Democratic candidate, since, at the national level, the Democratic platform mildly disavowed white supremacist sentiment, and this was seen as a betrayal. Given the virtual exclusion of blacks from the electorate, this is not an unreasonable interpretation, but the results do not directly support it. By way of contrast; for the 1968 and 1972 elections, the coefficient for this variable was strongly positive. With both blacks and whites in the electorate, such a simple interpretation as above is no longer possible, although there is evidence that the black vote was mobilized for Humphrey (against Wallace, particularly) in 1968 and for McGovern in 1972.

14. This is said with some caution, however, as I have earlier remarked that, in general, the choice between the two spatial models is not based on the exogenous variables. In this

instance there is a good deal of evidence that B should be in the model regardless of the specifics of its formulation. Clearly, the more we know of a process, the more we would be able to do this, but then, the inferential issues being discussed would be less important. Conversely, the less we know of a process, the more important is the correct representation of the spatial properties.

## REFERENCES

- ARORA, S. S. and M. BROWN, (1977) "Alternative approaches to spatial autocorrelation: an improvement over current practice." *Int. Regional Sci. Rev.* 2 (Fall): 67-78.
- BOX, G.E.P. and G. M. JENKINS (1970) *Time Series Analysis: Forecasting and Control*. San Francisco: Holden-Day.
- CAPECCHI, V. and G. GALLI (1969) "Determinants of voting behavior in Italy: a linear causal model of analysis," pp. 235-283 in M. Dogan and S. Rokkan (eds.) *Social Ecology*. Cambridge, MA: MIT Press.
- CHIROT, D. and C. RAGIN (1975) "The market, tradition and peasant rebellion: the case of Romania in 1907." *Amer. Soc. Rev.* 40 (August): 428-444.
- CLIFF, A. D. and J. K. ORD (1973) *Spatial Autocorrelation*. London: Pion.
- DOREIAN, P. (forthcoming) "On the estimation of linear models with spatially distributed data." *Soc. Methodology*.
- and N. HUMMON (1976) *Modeling Social Processes*. New York: Elsevier North-Holland.
- FRISBIE, W. P. and D. L. POSTON (1975) "Components of sustenance organization and non-metropolitan change: a human ecological investigation." *Amer. Soc. Rev.* 40 (December): 773-784.
- HIBBS, D. (1974) "Problems of statistical estimation and causal inference in time series regression models," pp. 252-308 in H. L. Costner (ed.) *Sociological Methodology*. San Francisco: Jossey-Bass.
- HOWARD, P. (1971) *Political Tendencies in Louisiana*. Baton Rouge: Louisiana State Univ. Press.
- and C. GRENIER (1978) "A longitudinal analysis of political change: a proposed methodological approach." *SOECO Papers on Social Ecology*, Ljubljana.
- INVERARITY, J. M. (1976) "Populism and lynching in Louisiana 1889-1896: A test of Erikson's theory of the relationship between boundary crises and repressive justice." *Amer. Soc. Rev.* 41 (April): 262-280.
- KENDALL, M. G. and A. STUART (1967) *The Advanced Theory of Statistics*. London: Hafner.
- KERNALL, S. "Comment: a re-evaluation of black voting in Mississippi." *Amer. Pol. Sci. Rev.* 67 (December): 1307-1318.
- MATTHEWS, D. R. and J. W. PROTHRO (1963) "Social and economic factors and negro voter registration in the South." *Amer. Pol. Sci. Rev.* 57 (March): 24-44.
- MITCHELL, E. J. (1969) "Some econometrics of the Huk rebellion." *Amer. Pol. Sci. Rev.* 63: 1129-1171.
- MORAN, P.A.P. (1950) "A test for serial independence of residuals." *Biometrika* 3: 178-181.

- ORD, K. (1975) "Estimation methods for models of spatial interaction." J. of the Amer. Stat. Assn. 70 (March): 120-126.
- RAGIN, C. (1977) "Class, status, and 'reactive ethnic cleavages': the social bases of political regionalism." Amer. Soc. Rev. 42 (June): 438-450.
- SALAMON, L. M. and S. VAN EVERA (1973) "Fear, apathy, and discrimination: a test of three explanations of political participation." Amer. Pol. Sci. Rev. 67 (December): 1288-1306.

*Patrick Doreian is an Associate Professor in the Department of Sociology at the University of Pittsburgh. He is coauthor with Norman Hummon of Modeling Social Processes and continues to work on the dynamic modeling of social phenomena. The current article is a part of a project on spatial analysis in which the political dynamics of Louisiana are being studied. He is also currently working on procedures for analyzing social network data.*