

UC Irvine

UC Irvine Previously Published Works

Title

Multivariate Modeling with Interdependent Network Data

Permalink

<https://escholarship.org/uc/item/6pk977bw>

Journal

Cross-Cultural Research, 17(3-4)

ISSN

1069-3971

Authors

Dow, Malcolm M
White, Douglas R
Burton, Michael L

Publication Date

1982-08-01

DOI

10.1177/106939718201700304

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at

<https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Multivariate Modeling with Interdependent Network Data

Malcolm M. Dow and
Douglas R. White and
Michael L. Burton

In the recent comparative literature the problem of simultaneously modeling functional and diffusional effects is being penetrated from two directions. One approach emphasizes the similar problem which arises in regression-based time series analysis. A second approach focuses on the difficulties of constructing more realistic formal representations of sample unit interdependencies. Both approaches have yielded important and complementary, but distinct, insights. Here, we outline some recent methodological developments which synthesize both approaches into a comprehensive and unified analytical framework.

1.0 Introduction: Galton's Problem as Spatial Autocorrelation

The oldest fundamental criticism of cross-cultural research is that measures of trait interrelationships are problematic because observations are not independent from one case to the next. Because traits are often spread widely by the repeated historical fission and migration of peoples, neither the number of truly independent cases nor the exact nature of the interdependence among societies is generally known. The problem was first pointed out by the statistician Galton in his comments on Tylor's classic comparative paper delivered in 1889 (Tylor 1970), and it is well-known in anthropology as Galton's Problem (Naroll 1970; Schaefer 1974).

In the cross-cultural literature the notion of sample unit interdependence has recently become formalized in the concept of spatial autocorrelation (Loftin 1972; Simonton 1975; Naroll 1976). Unlike earlier approaches to the

Malcolm M. Dow is an Assistant Professor of Anthropology at Northwestern University. Douglas R. White is a Professor of Social Sciences at the University of California at Irvine. Michael L. Burton is an Associate Professor of Anthropology at the University of California at Irvine.

problem, the spatial autocorrelation approach does not require conceptualizing sample unit interdependence as the lack of independence. As Murdock and White (1969) show, the number of independent cases in even moderate to large cross-cultural samples may be so few that the usual statistical procedures are thus difficult or even impossible. Rather than attempting somehow to restore independence, the current emphasis is upon attempting to describe formally the existing interdependencies and incorporate these into the modeling and estimating procedures. From this perspective, then, the observations are seen as being "tied together, like bunches of grapes, not separate, like balls in an urn" (Stephan 1934).

The methodological problem of simultaneously modeling functional and diffusional effects is being penetrated from two directions. One approach has been to explore the effects on classical (parametric) statistical estimates and significance tests when the data are interdependent. A second approach has focused on the difficulties of constructing realistic formal representations of sample interdependence and associated significance tests. Thus far, both approaches have yielded important but complementary insights. That is, both approaches are still analytically and procedurally distinct. We briefly review each of these approaches before outlining some very recent methodological developments which synthesize them into a more comprehensive and unified analytical framework.

1.1 Time Series Approach to Spatial Autocorrelation

From this perspective the statistical problems found in the analysis of spatially interdependent data are essentially the same as the problems encountered in the analysis of serially interdependent time series data. Time series observations do not usually change precipitously from one observation to the next, but tend rather to move smoothly and regularly through time. Each observation thus tends to be like those close in time and to be less like those distant in time. When measures of association are computed using such serially dependent data, it is straightforward to show that they are unbiased but inefficient (Hibbs 1974; Johnston 1972). That is, although the empirical estimates are distributed around the true population parameter, they tend to be widely dispersed and hence unreliable.

The analogy between temporal and spatial series was first discussed in the cross-cultural literature by Loftin (1972). Loftin argued that since societies tend to resemble one another when they are geographically close, and tend to be less similar when they are geographically far apart, estimates based on spatially interdependent data should also be unbiased but unreliable. What

lends considerable bite to Loftin's analogy is his demonstration of exactly these effects in a previously published set of data (Ember 1971) testing the effects of Galton's Problem.

Even though parameter estimates based on serially dependent data are known to be unreliable, the analysis of any particular sample will produce estimates which appear to be considerably more reliable than is actually the case. As Loftin notes, this is because an interdependent data set will generally display less variation than a comparably sized sample of truly independent data. And, since the standard errors of estimates are functions of this variance, they will tend to be underestimated. The ratios of estimates to their standard errors will thus consistently produce inflated *t*- and *F*-values, leading to spurious attributions of significance to the various associations being tested. The situation is considerably more complicated when many competing causal hypotheses are being evaluated, since errors of inference may thus accumulate and seriously impair the model building process.

Simonton (1975) has proposed one way to tackle this problem of underestimated standard errors. Drawing upon a rather old suggestion in the time series literature, Simonton advocates the Orcutt-James (1948) procedure to test the significance of a correlation between two time series. Briefly, the idea here is to use measures of interdependence of each variable to find a smaller, more "efficient" sample size with which to compute standard errors. Since standard errors are inverse functions of sample size, reducing the sample *n* in this way will correct them upwards. Ratios of estimates to their standard errors will then decrease, lessening the possibility of spurious inference. However, this procedure cannot be generalized beyond the bivariate case; a separate reduced *n* must be computed for each pair of variables of interest. What this means in the multiple regression situation is not clear.

Despite the cumulative insights provided by the time series perspective into the statistical problems which can arise in analyzing interdependent data, there are inherent methodological limitations to this approach to modeling functional and diffusional effects. First,

The spatial autocorrelation method consists in somehow aligning societies in a worldwide sample into a single linear arrangement, such that neighbors in the alignment are the most closely related through geographical propinquity, linguistic relationship, and other known historical connection. Then for each variable (trait) being studied, we compute a spatial autocorrelation. That is to say, we compute the correlation between a society's score on that trait and its alignment neighbor's score on the same trait. (Naroll 1976:126)

That is, "spatial" autocorrelation is computed with respect to a one dimensional alignment of sample units. Naroll (1976:127) goes on to caution, however, that the uses of these spatial autocorrelations "depends entirely on the validity of the alignment as a measure of historical relationships." Loftin and Hill (1974) also caution that the validity of such correlations depends entirely on the accuracy of the alignment procedures, since the effects of alignment errors are to bias measures of autocorrelation towards the null of no autocorrelation. In short, "The problem is a fundamental one. Measurement is primarily in one dimension at a time, but interdependence may move simultaneously in many directions" (Loftin and Hill 1974:32).

A second and related problem concerns the sampling procedures required for unidimensional alignments. Obviously, the larger the sample the more likely it is that errors will occur in ordering the sample units. And, of course, the difficulties are considerably compounded when samples are taken from restricted geographical regions. Decisions on how to align a continuous area sample, particularly one of moderate to large size, may often involve many arbitrary choices, and methods other than spatial autocorrelation may be required in such cases. However, it is obviously preferable to have available a unified set of procedures which are applicable to both worldwide and continuous area studies.

1.2 Two Dimensional Spatial Autocorrelation

As we have just noted, the time series approach to modeling diffusion is severely limited by the necessary assumption that a sample of societies scattered about in two dimensions can be meaningfully reduced to a unidimensional series, where diffusion operates through the series in one direction only. Obviously this assumption is unrealistic. More realistic approaches to measuring spatial diffusion are discussed by Wirsing (1974a,b). Drawing upon some earlier geographical and statistical work in this area, Wirsing (1974b:201) presents three measures which allow diffusion to operate simultaneously in all directions. The most interesting autocorrelation coefficient from the present perspective is Geary's (1954) "contiguity ratio," which was developed to detect spatial patterning of measurements on a collection of nonoverlapping counties which subdivide a region. Underling the Geary statistic is the notion of a contiguity matrix which reflects the contiguity structure among n counties. This structure can be represented as an $n \times n$ matrix where the i, j th element equals 1 if county i and j are contiguous, and equals 0 otherwise. When the term "county" is

replaced by the term "society," Wirsing suggests that "the contiguity ratio can be used as a convenient measure of diffusion between societies with respect to certain variables or traits."

Wirsing's suggestions on measuring spatial autocorrelation, while clearly an improvement over previous unidimensional procedures, are nonetheless of limited usefulness in multivariate model building and testing. The tests he proposes focus on univariate or bivariate associations. As with earlier approaches, Wirsing considers the Galton Problem solved "if at least one trait can be shown to display no autocorrelation." However, correcting for the effects of multilateral dependencies in the multivariate regression situation depends on examining the residuals for autocorrelation. Hence, one cannot simply focus on the autocorrelation of single traits. The conjunction of autocorrelated residuals and even very limited autocorrelation of the independent variable(s) in a regression equation can result in seriously misleading inferences concerning the significance of the regression estimates (Martin 1974).¹

Pryor (1976) has also proposed several techniques to detect multilateral diffusion, each of which depends on an $n \times n$ "diffusion possibility matrix." The elements of this matrix, which are more general than the zero/one contiguity coefficients employed by Wirsing (1974), are simply a weighted average of any factors thought to influence diffusion between two societies, such as distance and common language, where the weights are chosen a priori. Pryor refers to these matrix elements as "diffusion possibility (DP) indices."

Given a suitably constructed diffusion possibility matrix, Pryor proposes a variety of ad hoc procedures which may "suggest" the possibility of diffusion with respect to traits of interest. Basically, the tests compare the average of the DP indices between all pairs of societies both of which have the trait to the average of the DP indices of those pairs of societies only one of which has the trait. This clustering test can be carried out with respect to the dependent or independent variables of interest, or on a combination of independent variables weighted using the appropriate regression weights. A similar procedure can be carried out using the residuals from a regression equation. In each case, however, the "test" of diffusion reduces to a simple comparison of mean values.²

1.3 Network Autocorrelation

The problems encountered when applying classical statistical methods to spatially interdependent data have increasingly come to the attention of geographers, economists, and statisticians over the last few years. One ma-

major line of this research has been the development of multivariate regression models for analyzing two-dimensional series data. These models all incorporate measures of the interconnectedness among societies. Basic analytical and, especially, computational difficulties in estimating the parameters for these complex models have recently been overcome, and several powerful and flexible multivariate models are now well within computational reach.

Several advantages are apparent for these models over previous attempts to deal with interdependent data sets. First, any network of interrelationships which can be reasonably hypothesized can be easily formalized as a connectivity or relational matrix, and this matrix incorporated into the estimation procedures, provided that significant autocorrelation is found with respect to the network. Each society in a cross-cultural sample may be linked to many others without regard to direction, and any kind of linkage or flow between the units employed in constructing the connectivity matrix. Different hypotheses concerning interdependencies can be investigated by testing various network matrices, or several matrices may be combined into a single matrix. Second, the estimation procedures provide consistent (and hence asymptotically unbiased and efficient) estimates. Thus, increasing sample sizes will result in more "precise" estimates. Third, complex patterns of relationships among sets of variables may be modeled using several regression equations. More elaborate causal schemes can now be investigated.

Before discussing these regression models, we briefly examine the formal representation of a network of interdependencies among sample units as a connectivity matrix. We point out the need for a hypothesis concerning the underlying structure of interrelationships, and we examine the nature of the estimation problems associated with the usual regression procedures. We briefly outline maximum likelihood procedures which overcome these difficulties. Only the general nature of these procedures is discussed. Estimating equations and significance testing procedures will be discussed in a later paper.

2.0 Nature of Spatial and Network Autocorrelation

Simply stated, *spatial autocorrelation* implies that what happens at one location in space is in some way related to what happens at nearby locations. For example, if a large continuous geographical area is subdivided into a set of smaller nonoverlapping areas, then we could reasonably hypothesize that the values on a variable at one location would be systematically related to values on the variable at contiguous or nearby

areas. More generally, *network autocorrelation* implies that the attributes of each node of the network can be predicted in part from a knowledge of the attributes at related nodes.

The underlying structure of interdependencies among a set of units can be formally expressed in terms of a connectivity matrix. Any structural arrangement connecting n sample units can be represented as an n by n matrix as shown in Figure 1.

Several features of the above matrix formulation should be noted. First, by convention a unit is not connected to itself, and so a connectivity matrix will have zeros along the main diagonal. Second, although no restrictions need be placed on the weights c_{ij} , it will be more convenient in the analysis that follows if each row is scaled to sum to one. That is, each element of the C matrix is divided by its row sum to a new stochastic matrix W . One result of this scaling is that in general $w_{ij} \neq w_{ji}$, since i and j will not usually be connected to the same number of other units. Third, this matrix formulation permits a very rich expression of the idea of interdependence. Any unit may be connected to any other unit, not simply those which are actually contiguous, spatially or otherwise. Fourth, any kind of linkage or flow between units can be represented by the appropriate element of the W matrix. Thus both direct and indirect feedback effects between units can be modeled in this manner.

Selection of the weights is of major importance, since spurious results may arise if the hypothesized weights do not correspond to any real process (Cliff and Ord 1973). Commonly in geographical research, the weights (c_{ij}) are estimated according to some notion of space-friction constraints on the possibility of effects from one unit to another. The simplest function in this case is an exponentially decaying distance function such as D_{ij}^{-a} , where D_{ij} is the distance from location i to j and a is a suitable exponent chosen a priori. Notice, however, that if the units are a collection of subregions, then some decision must be made as to the exact location of the points within the regions which are to be used in measuring distance. Actually, the situation is considerably more complex than simply computing distances, since consideration of size and shape of the various regions may also be of importance in realistically describing the interactions among a collection of regions. Cliff and Ord (1973) attempt to overcome this problem by computing weights based on both distance and proportion of boundary in common. Gatrell (1979) has constructed a measure of interaction among Swedish towns based on geographical distance and number of telephones. Bodson and Peeters (1975) employ distance and minimum public transport time when constructing an "accessibility function" to generate interaction coefficients among 44 Belgian districts.

Figure 1. Matrix Representation of Network Interdependencies

$$C = \begin{bmatrix}
 & 1 & 2 & \bullet & \bullet & i & j & \bullet & \bullet & n \\
 1 & 0 & C_{12} & \bullet & & & & & \bullet & C_{1n} \\
 2 & C_{21} & 0 & & & & & & & \bullet \\
 \bullet & \bullet & & 0 & & & & & & \\
 \bullet & \bullet & & & 0 & & & & & \\
 i & & & & & \bullet & & & & \\
 j & & & & & & \bullet & C_{ij} & & \\
 \bullet & & & & & C_{ji} & 0 & & & \\
 \bullet & & & & & & & \bullet & & \bullet \\
 \bullet & \bullet & & & & & & & \bullet & C_{n-1n} \\
 n & C_{n1} & \bullet & & & & & \bullet & C_{nn-1} & 0
 \end{bmatrix}$$

Clearly, the form of the weighting matrix will depend on the problem at hand and the available data. In cross-cultural research, many possibilities exist for constructing weights. Apart from the spatial functions which are obviously applicable to a diffusion process, other pertinent network effects which have commonly been hypothesized are language similarity, whether or not two societies belong to the same state, and trading relationships.

A commonly hypothesized structure for time series observations is that each is related to the preceding observation. This simple model can be written as $Y_i = \rho Y_{i-1} + u_i$, where the u_i are normally distributed random error items. Likewise, the Linked Pair method of measuring spatial autocorrelation hypothesizes that each society in an alignment is related only to the succeeding society in the alignment. This hypothesis about the underlying structure of interdependencies can be more formally expressed using a connectivity matrix as suggested above. The appropriate matrix in this case is shown in Figure 2.

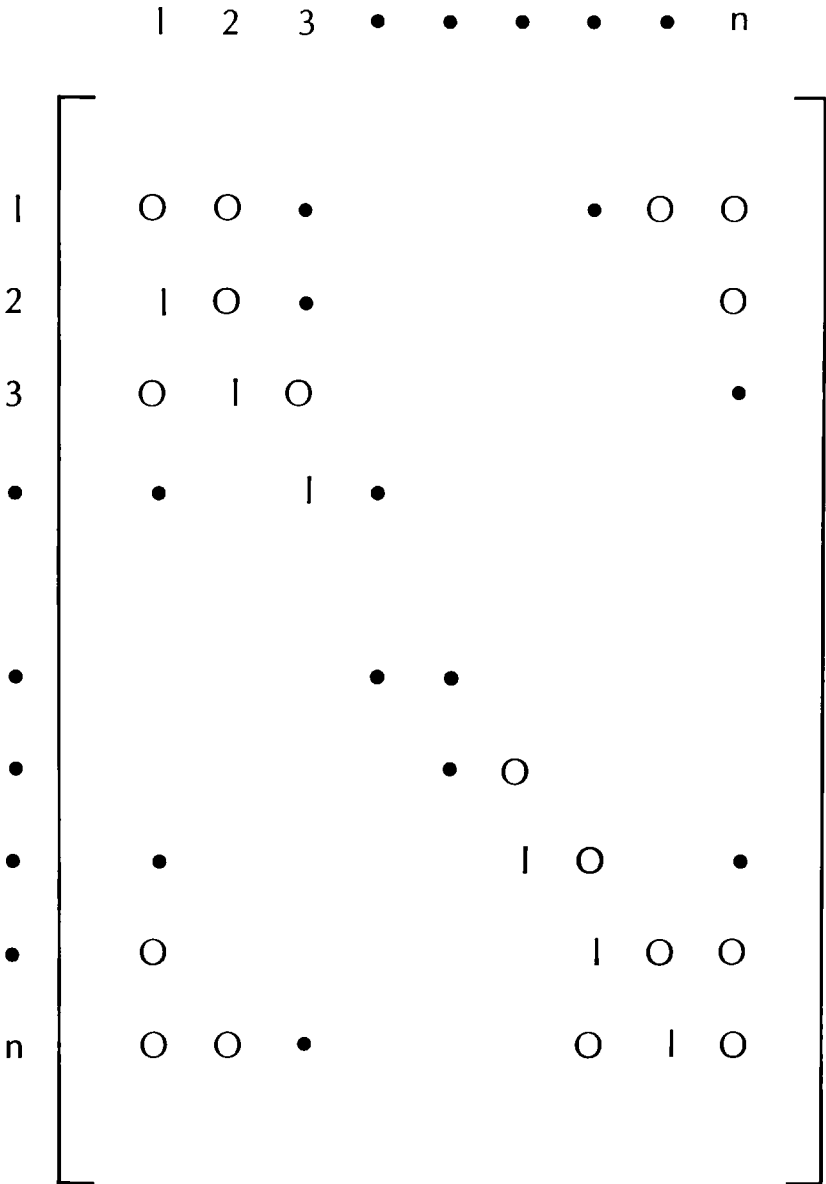
Using this W matrix, and writing Y as an $n \times 1$ column vector of trait scores, and u as an $n \times 1$ column vector of random error terms, the above simple time series equation can be represented as

$$Y = \rho WY + u \quad (2.1)$$

In this equation, WY represents a column vector of trait scores which are "lagged" one period with respect to the Y scores. Hence the simple regression coefficient is a measure of the flow of variable Y with respect to the structure represented by the W matrix. That is, ρ measures the autocorrelation of Y with respect to W .

For a time series the matrix W is certainly a reasonable hypothesis, since the essential asymmetry of time is reflected in the structure of weights in W . Observation i is affected by observation $i-1$, and there is no reciprocal effect, which intuitively corresponds to the notion that current events cannot influence past events. This natural temporal asymmetry results in a rather special form for the W matrix: all of the elements above the main diagonal are zero. So long as W has this upper triangular structure and the error terms meet the usual assumptions, ordinary least squares (OLS) regression procedures will generate consistent and relatively efficient parameter estimates (Malinvaud 1970). However, this fact depends heavily on the W matrix having the above triangular structure, since this corresponds to the assumption that the errors are not correlated with the independent (WY) variable. In time series this is not an unreasonable assumption, but in the two dimensional spatial case, where societies are related to one another in a multilateral fashion, the assumption that the errors are uncorrelated with the WY variable while being a function of the Y variable is clearly

Figure 2. First-Order Lag Time Series Matrix



untenable. In this case, the OLS regression procedures produce inconsistent estimates, that is, the estimates do not converge to the true population parameters no matter how much the sample size is increased.

The failure of the usual regression procedures to deal with multilateral spatial or network interdependencies raises some new estimation problems. In the following section we review briefly the sources of bias and inconsistency in the OLS estimating procedures, and discuss the methods recently proposed which overcome these limitations and which allow consistent (asymptotically unbiased and efficient) estimates to be found. Although the discussion here focuses on the regression model, it should be noted that partial correlation procedures are incorrect for the same reasons that OLS regression is incorrect. And, although correlations between Y and WY will be unbiased, the distributions of such statistics are quite different from the usual distributions (Cliff and Ord 1973).

3.0 Maximum Likelihood Estimation of Network Autocorrelation

Within the regression framework, then, the network autocorrelation problem is twofold: formulation of a plausible network structure W , and estimation of the autocorrelation coefficient. While the difficulties of constructing a W matrix will depend on the problem at hand, estimating a network autocorrelation coefficient belongs to a more general set of problems.

In the usual application of OLS regression analysis, coefficients are estimated by the method of moments. That is, the means, variances, and covariances of all of the variables are calculated, and these are then used to find estimates of the population parameters of the regression model. An alternative procedure which does not employ sample moment is the method of maximum likelihood. The basic idea behind this latter procedure is to try to locate estimates of population parameters which are most likely to have generated the observed sample values.

For the normal regression model with independent errors, maximum likelihood (ML) estimation and ordinary least squares (OLS) are equivalent procedures. For example, in the time series case, where W has an upper triangular structure as previously mentioned, the procedures produce the same parameter estimates and variances. However, the two methods are importantly different in the case where there are multilateral dependencies among the sample units and the W connectivity matrix has a more general form. In this latter situation, ML procedures are superior to OLS, since they yield asymptotically consistent estimates, whereas OLS estimates are inconsistent (Hepple 1976). To see this important and crucial difference in

the properties of these two estimation procedures we consider again the simple regression model. We begin our discussion with the following univariate case in which there is one autocorrelated variable Y ; later we will consider the more general case in which there is a dependent variable and one or more independent variables.

$$Y = \rho WY + v \quad (3.1)$$

Again, Y is an $n \times 1$ column vector of scores on variable Y , W is an $n \times n$ connectivity matrix, and the errors, v , are an $n \times 1$ column vector of normally distributed random errors with mean zero, no autocorrelation or heteroscedasticity (i.e., $v \sim N(0, \sigma_v^2 I)$).³

This basic equation can be rewritten as follows:

$$\begin{aligned} v &= Y - \rho WY \\ &= (I - \rho W)Y \\ &= AY \text{ where } A = (I - \rho W) \text{ and } I \text{ is} \end{aligned}$$

the $n \times n$ identity matrix.

The usual OLS regression problem is to minimize the sum of squares of the error terms. This is equivalent to maximizing the following likelihood function:

$$L(v) = \left[\frac{\times}{2\pi\sigma^2} \right]^{\frac{N}{2}} \exp \left[\frac{-v'v}{2\sigma^2} \right] \quad (3.2)$$

since this function is maximized when the negative exponent, the sum of squared error terms, is minimized. However, since the v are unobserved it is preferable to make a change of variable in the likelihood function to the observed Y s. After the change of variable, the new likelihood function is

$$L(Y) = \left[\frac{1}{2\pi\sigma^2} \right]^{\frac{N}{2}} \exp \left[\frac{-(AY)'(AY)}{2\sigma^2} \right] |\text{Det}(A)| \quad (3.3)$$

where $\text{Det}(A)$ is the Jacobian of the linear transformation of the Y 's to the v 's ($v = AY$).⁴ Analytically, it is more convenient to deal with the logarithm of this function, which introduces no new difficulties, since the maximum of the above function will remain the same under such a monotonic transformation. The log-likelihood function is thus

$$\ln(Y) = -\frac{N}{2} \ln(2\pi\sigma^2) - \frac{1}{2\sigma^2} (AY)'(AY) + \ln|\text{Det}(A)| \quad (3.4)$$

The fundamental difference between OLS and ML procedures concerns the last term of this equation, $\text{Det}(\mathbf{A})$. In the time series case, and where a one dimensional alignment of societies is analyzed using, for example, the Linked Pair statistic, the matrix $\mathbf{A} = \mathbf{I} - \rho\mathbf{W}$ has a rather special structure: it has all ones on the main diagonal apart from the first term and all zeros within the upper triangular portion of the matrix. Now, the determinant of such a matrix is simply the product of the main diagonal elements (Hepple 1976), and so the determinant of matrix \mathbf{A} , i.e., $\text{Det}(\mathbf{A})$, is easily computed. In time series analysis the first diagonal element is usually set equal to $(1 - \rho^2)^{1/2}$ which is thus the value of $\text{Det}(\mathbf{A})$. However, since this term remains constant when the sample size increases, its magnitude becomes swamped by the other terms of the likelihood function, and so it is asymptotically negligible. Hence minimizing the sum of squared errors and maximizing the above log-likelihood function are equivalent procedures in this situation. However, there is no such equivalence when \mathbf{W} has a more general structure, that is, when \mathbf{W} is no longer upper triangular, since the determinant is no longer simply the product of its main diagonal elements. Not only does the determinant of \mathbf{A} not collapse to a simple quantity, but it is neither negligible nor invariant to sample size (Hepple 1976). Whittle (1954) has shown that the expectation of the bias by ignoring this term is nonzero, and that the OLS regression coefficient is thus asymptotically biased and inconsistent.

Clearly, then, the usual regression procedures are invalid where there are multilateral dependencies among the sample units. However, the ML estimates which maximize the log-likelihood function have the desirable properties of consistency (Ord 1975; Hepple 1976). Also, since the ML estimates are asymptotically normally distributed, variances and covariances can be computed and inferential procedures applied.

The major computational difficulty in obtaining the estimate ρ which maximizes the likelihood function above involves evaluating $\text{Det}(\mathbf{A})$ term. Since ρ is found by a direct search procedure, this determinant would have to be evaluated at each iteration, a computational burden. Recently, Ord (1975) has proposed a new procedure which brings the maximum likelihood solution well within computational reach. Ord shows that the heavy computational step of evaluating the determinant of \mathbf{A} at each iteration can be reduced to the problem of computing the eigenvalues of the \mathbf{W} matrix one time only, and then using these eigenvalues in a simple formula at each iteration. That is, given the eigenvalues of matrix \mathbf{W} , $\lambda_1, \lambda_2, \dots, \lambda_n$, then

$$\text{Det}(\mathbf{A}) = \text{Det}(\mathbf{I} - \rho\mathbf{W}) = \prod_{i=1}^n (1 - \rho\lambda_i) \quad (3.5)$$

The computational savings involved in using this simple product function in place of the complete evaluation of $\text{Det}(\mathbf{A})$ are considerable.⁵

Using Ord's simple identity, and after condensing the log-likelihood function by eliminating σ^2 , the ML estimate of ρ is the ρ which minimizes the following expression:

$$-\frac{2}{N} \sum_{i=1} \ln(1 - \rho \lambda_i) + \ln(\mathbf{Y}'\mathbf{Y} - 2\mathbf{Y}'\mathbf{W}\mathbf{Y} + \rho^2(\mathbf{W}\mathbf{Y})'(\mathbf{W}\mathbf{Y}))$$

Once a suitable ρ is found its significance may be ascertained using the asymptotic variance-co-variance matrix of σ^2 and ρ (Ord 1975: 124).

Ord (1975) notes that it is useful though not necessary to have the matrix \mathbf{W} row scaled to unity, which implies that $\rho < 1$. This latter restriction enhances the interpretability of ρ in the more complex models discussed below.

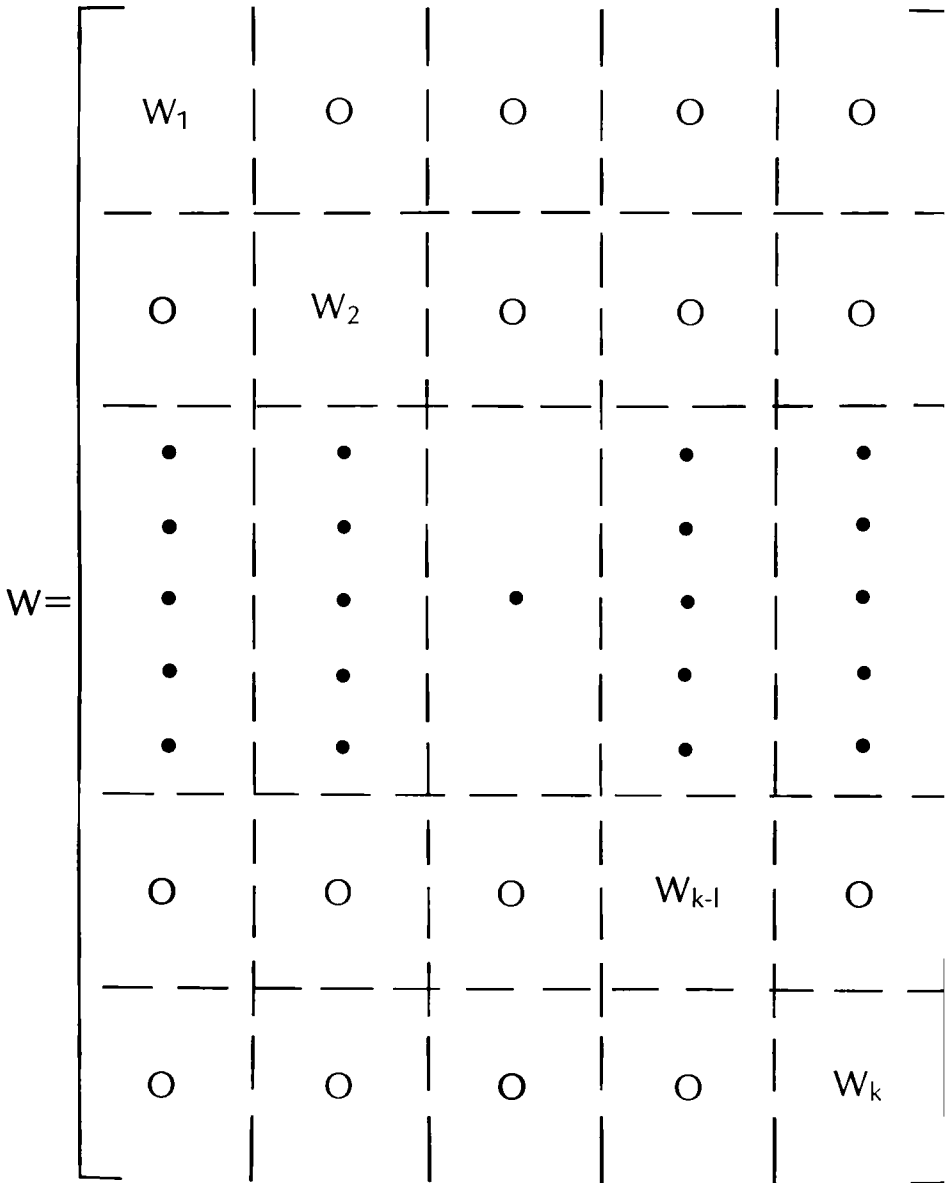
One point to note with respect to calculating the ML estimates of spatial or network autocorrelation is that the largest real valued nonsymmetric connectivity matrix \mathbf{W} for which eigenvalues can be computed given current computational facilities is around 70×70 . However, considerably larger samples can be handled using this approach if the \mathbf{W} matrices can be forced into block structure. That is, if by elimination of a few elements of \mathbf{W} it can be arranged in the structure shown in Figure 3, then the eigenvalues of the entire \mathbf{W} matrix can be obtained easily from the eigenvalues of each sub-matrix. Careful blocking of the larger matrix should require few elements to be deleted, and this will not affect the asymptotic properties of the ML coefficients. Other simplifications are also possible (Ord 1975).

4.0 Some Related Regression Models for Network Autocorrelated Data

The methodological focus so far has been on the problem of finding a statistically satisfactory measure of autocorrelation with respect to an hypothesized multilateral structure of interdependencies among the data units. However, the univariate situation discussed above is obviously of limited utility in model building and testing. In most usual research settings, the investigator is interested in the pattern of relationships among a set of variables. Fortunately, the ML approach is straightforwardly extended to the multiple regression situation.

The models that have been developed allow for network autocorrelation in basically two ways: either by formulating these effects as an explicit variable and including it as an independent variable in the equation to be

Figure 3. Block Diagonal Network Matrix



estimated, or as a change in the assumptions concerning the structure of the error terms. It is also possible to incorporate both kinds of autocorrelation into a single model.

4.1 Single Network Models

i) Network Disturbances Model. In regression analysis the assumption of independence applies not to the variables but to the error terms. That is, the usual specification of the regression model $Y = X\beta + \varepsilon$, $\varepsilon \sim N(0, \sigma_\varepsilon^2 I)$, requires no assumptions concerning independence with respect to the measured variables. However, the error term assumptions are crucial. These assumptions require the errors to be (1) normally distributed with zero mean, (2) homoscedastic (i.e., equal variances), and (3) uncorrelated with each other (i.e., no autocorrelation). While there is some evidence to show that the regression model is robust with respect to violations of normality and homoscedasticity (Macdonald 1976; Bohrnstedt and Carter 1971), violation of the nonautocorrelation assumption can have disastrous consequences. Serious errors of inference may result if the errors are autocorrelated and this is not corrected, since again the t- and F- statistics will generally be inflated (Hibbs 1974; Martin 1974).

When data are obtained from a sample with known interdependencies, it is possible to incorporate this structure into the error term assumptions. Specifically, given any hypothesized network matrix W the usual regression model can be respecified as the following simultaneous equations:

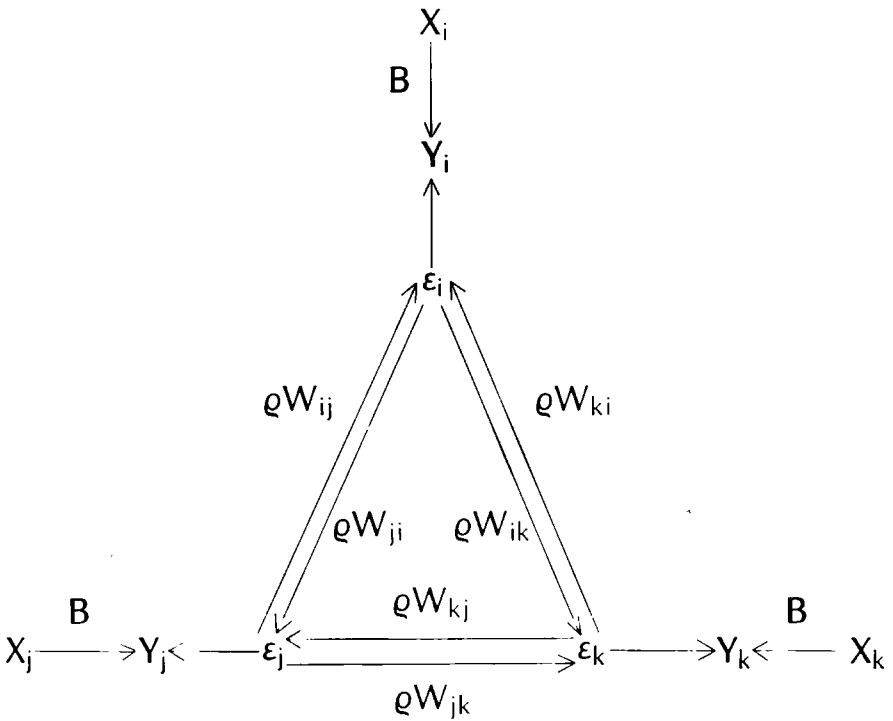
$$Y = X\beta + \varepsilon \quad (4.1)$$

$$\varepsilon = \rho W\varepsilon + \nu \quad \nu \sim N(0, \sigma_\nu^2 I) \quad (4.2)$$

Here, the error terms are hypothesized as being autocorrelated with respect to W . The X matrix contains K independent variables plus an initial column of ones which results in the intercept term appearing as the first element of the $K+1$ column vector of coefficients, β . Note that when $\rho = 0$, this is simply the OLS model, which shows that the above model is an appropriate generalization of OLS. Likewise, if the W matrix has all zeroes, corresponding to complete independence of each unit from all of the others, this model is again identical to OLS.

The nonrecursive nature of the relationship among the variables implied by equations 4.1 and 4.2 is more immediately grasped by inspecting the corresponding structural diagram. For clarity and simplicity, we consider only three interrelated network units and one independent variable. Exten-

Figure 4. Structural Diagram for the Network Disturbances Model



sion to n nodes and several independent variables is straightforward. Thus, given three related nodes i, j , and k , the relationships implied by the above simultaneous equations are shown in Figure 4.

Figure 4 illustrates quite clearly that embedding the network matrix W in the error term equation avoids the question of how the hypothesized network might interact with and affect the measured X, Y variables. In an exploratory stage of research it may not be possible to specify precisely such interactions. As the nature of various network structures and processes become better understood, it should then be possible to make more detailed specifications. Hence, concern at this stage of analysis would probably focus on obtaining efficient estimates of the β s, and perhaps on gaining some understanding of various networks of interest. Also, it should be noted that if in the exploratory stage of analysis there are several networks of interest, then row scaling the W matrices to unity will allow comparisons of the feedback parameters ρ by eliminating the differences due to the different scales used to construct the W matrices.

Some further insights into this model are obtained if the equations are transformed and combined in the following manner:

From 4.2 we get $\epsilon - \rho W\epsilon = \nu$

Thus $(I - \rho W)\epsilon = \nu$

and $\epsilon = (I - \rho W)^{-1} \nu$ (4.3)

Substituting this latter expression into A.1 above gives

$$y = X\beta + (I - \rho W)^{-1} \nu$$

and, after premultiplying each side of this equation by $(I - \rho W)$ we get⁵

$$(I - \rho W)Y = (I - \rho W)X\beta + \nu$$

or,

$$Y - \rho WY = (X - \rho WX)\beta + \nu$$
 (4.4)

Equation 4.4 implies that for a known ρ , if we subtract out of each variable score that portion due to the influence of contiguous or related units, i.e., if we form new variables $Y^* = Y - \rho WY$ and $X^* = X - \rho WX$, then OLS applied to the regression equation,

$$Y^* = X^*\beta + \nu$$
 (4.5)

will yield unbiased and efficient estimates of β . This is so because the errors are now by definition not autocorrelated. In general, ρ will not be known a priori. In his initial formulation of the disturbances model Ord (1975) proposed an iterative procedure which begins by estimating ρ using the OLS residuals from equation 4.1 as input to the ML procedure described in section 3.0 above. This estimate is then used to form new variables Y^* and X^* as just described. The residuals from OLS applied to these new variables are then used to find a second estimate of ρ . The procedure iterates until successive values of ρ converge to within some prespecified decimal place.

Although Ord's procedure is simply an estimating routine devised to locate consistent parameter estimates, it does suggest that the disturbances model may be appropriate if one suspects that the underlying network somehow interacts with the measured variables. At each iteration a common transformation is applied to all of the variables even though they may be differentially affected by the network structure. A common transformation is all that is required to produce well-behaved errors and thus permit efficient estimates to be found.⁶

It is possible that the disturbances model will suggest significant autocorrelation among the regression residuals when there is "really" no autocorrelation present. This can be caused by either of two types of model misspecification: (1) nonlinear relationships between the dependent and independent variables which are not included in the model, and (2) the

omission of important explanatory variables. The first problem is easily handled by employing simple nonlinear variable transformations or by adding interaction terms. However, if these procedures do not remove the autocorrelation among the residuals, then additional explanatory variables may have to be introduced. In time series analysis, it is common to introduce a "lagged" dependent variable as a new independent variable in this situation. We now consider the more general case where network effects are introduced into the model as an independent variable.

ii) Network Effects Model. This model is formally specified as:

$$Y = \rho WY + X\beta + \varepsilon \quad \varepsilon \sim N(0, \sigma^2 I) \quad (4.6)$$

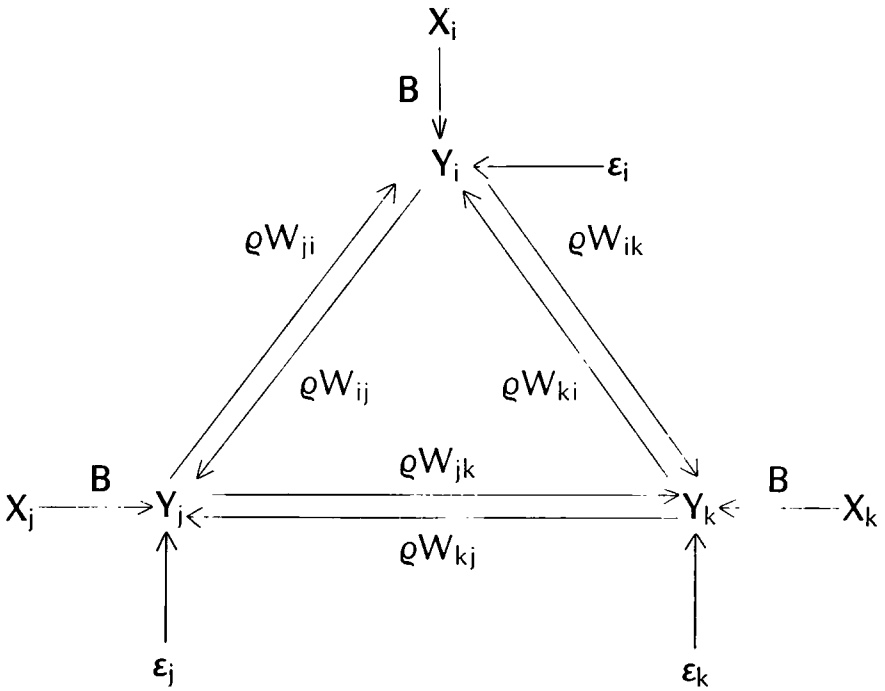
For this model, then, only the dependent variable is assumed to be auto-correlated with respect to the W network. The independent X variables are unaffected by the network, and the errors are also unaffected. WY is not the same as the other independent variables here, since it is by definition correlated with the error terms. OLS regression procedures are invalid in this situation and cannot be used to estimate the model parameters, since violation of the assumption of independence between "lagged" dependent variables and errors throws OLS off the beam and produces biased and inconsistent estimates (Johnston 1972)⁷. Maximum likelihood procedures are available (Ord 1975) to estimate the parameters of this model which are not affected by the interdependency among WY and the error terms.

The nature of the feedback process implied by equation 4.6 is again easily grasped from inspection of a corresponding structural diagram. Again, we consider only three related network nodes and a single independent variable. Figure 5 shows the interrelationships among the variables specified by the network effects model.

It is possible to introduce the WY variable into this model simply to "control" for its effects and thus obtain efficient regression estimates. However, Figure 5 strongly suggests that the feedback process among the sample units with respect to the dependent variable may also be of considerable interest. That is, one may also be interested in the effects of the network (WY) on the dependent variable (Y) while holding constant individual unit attributes (X s). And, again, row scaling W to unity has the advantage of allowing comparisons of different network schemes on the dependent variable.

While Figure 5 appears to suggest direct feedback effects it should be recalled that the W matrices need not express any kind of contiguity, spatial or otherwise. In a cross-cultural sample, for example, two societies could be

Figure 5. Structural Diagram for the Network Effects Model



classified as similar on the basis of subsistence type and yet have no known relationship with one another. Interpretation of the feedback parameter q in this model is thus dependent on the network problem at hand.

4.2 Multiple Network Models

Previously we mentioned how rather complex hypotheses concerning the structure of a connectivity matrix W could be combined to produce a single W . Thus, Cliff and Ord (1973) combined both distance and proportion of common boundary of counties to produce a single connectivity matrix. In some situations, however, it may be more desirable to separate the effects of distinct underlying processes which are presumed to be operating simultaneously with respect to the variables of interest. In cross-cultural studies, for example, it may be of interest to examine the effects of both geographical and linguistic effects within the same model. Or, we may be interested not only in the effects of immediately contiguous societies, but also the effects of more distant, not necessarily contiguous, societies, and so on.

i) Multiple Network Disturbances. In this situation the general autoregressive model is:

$$Y = X\beta + \varepsilon \quad (4.7)$$

$$\varepsilon = \rho_1 W_1 \varepsilon + \rho_2 W_2 \varepsilon + \nu \quad \nu \sim N(0, \sigma_\nu^2 I) \quad (4.8)$$

These equations may be rearranged and combined as before to give

$$Y - \rho_1 W_1 Y - \rho_2 W_2 Y = (X - \rho_1 W_1 X - \rho_2 W_2 X)\beta + \nu \quad (4.9)$$

Again, a common transformation results in a well-behaved error term in this equation, and again OLS can be applied to the transformed variable to yield satisfactory estimates of the population parameters. In the case that the W_1 and W_2 matrices are systematically related, as they are, for example, when they represent contiguous societies and those at further lags, the maximum likelihood estimating equations can again be simplified using a slight extension of Ord's (1975) previously discussed procedure. In the more general case, though, Ord's results do not apply and more complex estimation procedures are required (Brandsma and Ketellapper 1979). Although we are not concerned here with estimating equations and the various computational problems involved in obtaining valid and efficient estimates, it is of interest to note that both negative and positive autocorrelation processes may simultaneously occur, and that this event presents no particular estimation problems. However, the substantive conclusion here is that the occurrence of a negative spatial or other network autocorrelation coefficient should not be regarded as a solution to the problem of cultural diffusion in and of itself, as Naroll (1976) has suggested. Rather, the occurrence of a negative coefficient should be interpreted as an indication that other processes may also be operative, and that a fuller understanding of the diffusion process requires critical examination of negative and positive autocorrelation processes.

The structural diagram for this simultaneous equation model is essentially the same as for the single disturbance model shown in Figure 1, where we now replace each ρW_{ij} with $\rho_1 W_{1ij} + \rho_2 W_{2ij}$.

For this model, row scaling both W_1 and W_2 to unity again permits a direct comparison of the ρ_1 and ρ_2 estimates by removing the confounding effects of differences in measurement scales used to construct W_1 and W_2 .

ii) Joint Network Effects—Network Disturbances Model. This model combines two previous models into the following two simultaneous equations: (Doreian 1980b)

$$Y = \rho_1 W_1 Y + X\beta + \varepsilon \quad (4.10)$$

$$\varepsilon = \rho_2 W_2 \varepsilon + \nu \quad \nu \sim N(0, \sigma_\nu^2 I) \quad (4.11)$$

As with the multiple disturbances model, the W_1 and W_2 network matrices can either bear some intrinsic relationship to one another (contiguous, lag one), or they may represent quite distinct types of networks (language, distance). In the case of the multiple disturbances, however, it makes no difference which matrix is labeled W_1 and which W_2 . That is, both matrices are treated "equally" in the estimation procedure. This is not the case with the above joint network effects-network disturbances model. Only in the case that the two matrices are identical (as they may be) will it not matter how they are labeled. Otherwise, the coefficient associated with a particular network matrix will be different depending on whether it is entered into the estimating procedures as the W_1 and W_2 matrix. In general, then, the joint model requires some additional understanding of the impact of network structures and processes on the dependent variable of interest.

The structural diagram corresponding to this rather complex joint effects-disturbances model is shown in Figure 6. (For clarity, we omit most of the feedback terms from the diagram.)

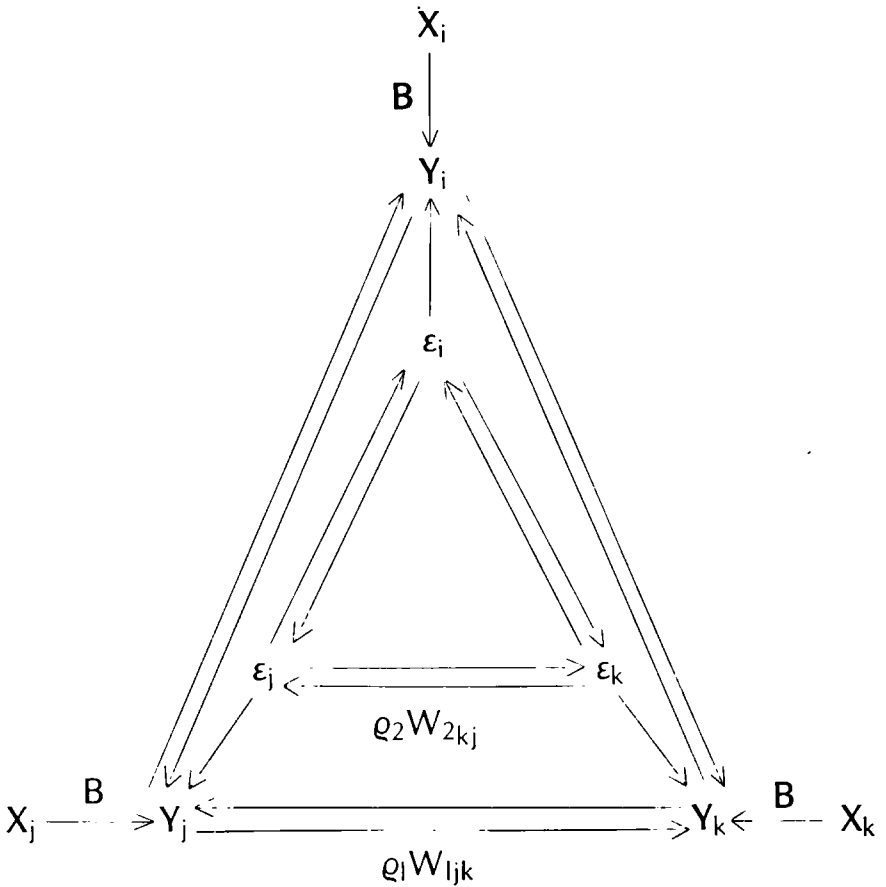
Interpretation of the ρ_1 feedback parameter again must be made with reference to the nature of the network relationship operationalized by the W_1 matrix.

The above four models clearly do not exhaust the possibilities for extending the regression model to deal with the network autocorrelation problems. But, though it is easy to see how to state more complex models, it is considerably more difficult to identify substantive issues which are sufficiently well developed to require even the two multiple network models discussed here.

4.3 *The Choice of Model*

The effects model assumes that only the dependent variable is autocorrelated, whereas the disturbances model assumes that the same autocorrelation process operates on all of the variables. The effects model is clearly the appropriate model for the diffusion of a single variable by contagion. A good example of this would be the adoption of innovations (Rogers 1971). In models of adoption of innovations it is usually assumed that the innovations diffuse through social networks, so that position in the network has an important effect on an individual's adoption of the innovation. This is a good example of network autocorrelation where the variable to be explained (adoption of the innovation) is autocorrelated. It is often safe to assume that other variables which affect the propensity to innovate, such as social status (Cancian 1979) or ethnicity, are subject to much less auto-

Figure 6. Structural Diagram for the Joint Network Effects-Network Disturbances Model



correlation than the dependent variable, so that they can be safely assumed for all practical purposes not to be autocorrelated. They would occur in a regression equation as non-autocorrelated independent variables.

Two examples of studies where the effects model seems the appropriate model are Doreian's studies of the Huk rebellion and of voting in a Southern state (Doreian 1980a,b). In both cases, we can assume rather high contagion effects. An important feature of both of these situations is that the dependent variable, participation in the rebellion in one case, or voting Democratic in the other, fluctuate rather rapidly, whereas the independent

variables, which pertain to socioeconomic features such as urbanization and land ownership, tend to change slowly over time. Much of the short-run fluctuation, then, can be modeled simply by autocorrelation in the dependent variable.

The disturbance model seems to us to be the appropriate model for much of holistic cross-societal research. Here it is often realistic to assume that entire systems of traits diffuse together. This would especially be the case with migration or fissioning of entire societies, as, for example, with the Athapaskan, Bantu, Malayo-Polynesian, Indo-European, and Nilotic migrations. Because the disturbances model provides an appropriate model for this kind of diffusion, which has been of considerable interest to anthropologists, we think it will be the best model for much cross-cultural research. We should add that there are other situations than the migrations of entire societies, where several traits diffuse together. In East Africa, for example, societies from different language families share such traits as age set systems, female circumcision, removal of the incisor teeth of children, the presence of subcastes of blacksmiths and of hunters who are also ritual circumcision specialists, and female husbands. Examples of such diffusion of trait clusters can be found in most regions of the world.

The mixed model is intriguing because it provides for a combination of joint diffusion of most traits with a different form of diffusion of the dependent variable. It seems likely to us that many situations will be found where this is the appropriate model. We do not yet have a computational algorithm for this model, so do not have the direct experience with it that we have with the disturbances and effects models. In the example which follows, all tests were done with both the disturbances model and the effects model. In all cases, the effects model produced unsatisfactory results, with high autocorrelation coefficients and negligible regression coefficients. The disturbances model, on the other hand, showed interpretable evidence of autocorrelation, but also produced meaningful regression coefficients that were similar in sign but different in magnitude to regression coefficients from OLS analysis of the same data.

5.0 An Empirical Example

Space limitations do not allow us to include an elaborate empirical example here. To give a feel for the workings of the autocorrelation analysis we present part of the results from a study of variation in the sexual division of labor in African agriculture (Burton, White and Dow 1981). The sample consists of the 34 societies of the standard cross-cultural sample (Murdock and White 1969) that are in Africa and also have agriculture.

The dependent variable in this study is the degree of female participation in agriculture. This is computed as the sum of the sexual division of labor codes for three tasks: soil preparation, crop tending, and harvesting. The variable is coded from 3 to 15; higher values correspond to greater female participation in agriculture. We use two variables to predict female participation in agriculture. From reasoning about ecological circumstances of food production, we have predicted that there will be greater female participation in root crop agriculture than in cereal crop agriculture. This variable is labeled C, for crop type. From an argument about the organization of work, we have predicted that there will be less female agricultural participation in societies that have slavery. This variable is labeled S.

The ordinary least squares equation shows significant coefficients for both variables:

$$\begin{array}{rcccc}
 T = 24.4 & -4.76C & -2.64S & R^2 = .47 & (5.1) \\
 (2.7) & (1.22) & (.73) & & \\
 & p < .001 & p < .005 & &
 \end{array}$$

In this study we used two different W matrices. The first, W_L , is based on genetic relations among languages. The second, W_D , is based on an exponentially decreasing function of geographical distances. In equation 5.2 below we see that the distance W_D matrix provides clear evidence for spatial autocorrelation. The regression coefficients, are still significant, but of slightly different magnitude than before, and the measure of goodness of fit,⁸ R^2 , has increased over the OLS measure, as it should, since there is significant autocorrelation:

$$\begin{array}{rcccc}
 T^* = 21.8 & -4.44C^* & -1.78S^* & R^2 = .61 & \rho = .49 \\
 (2.8) & (1.33) & (.66) & & (.20) \\
 & p < .001 & p < .005 & & p < .05
 \end{array}
 \quad (5.2)$$

Using the language-based W_L matrix, equation 5.3 shows significant autocorrelation and an even stronger measure of fit. The regression coefficients are still significant, and the standard errors of the estimates have decreased, which is what we would expect when there is autocorrelation.

$$\begin{array}{rcccc}
 T^* = 19.2 & -3.19C^* & -1.86S^* & R^2 = .67 & \rho = .69 \\
 (2.9) & (1.07) & (.61) & & (.18) \\
 & p < .005 & p < .005 & & p < .001
 \end{array}
 \quad (5.3)$$

One may now ask whether these two matrices are not just somewhat different measures of the same diffusion process. To test that possibility, we reasoned that a plausible source of language-related diffusion could be the well-known historical migration of the Bantu peoples. We constructed a dichotomous variable, B, that takes the values of 1 for Bantu societies and 0 for all others. When we entered this into the OLS regression equation it showed a significant coefficient for Bantuness:

$$\begin{array}{ccccccc}
 T = 16.4 & -4.06C & -1.79S & +3.41B & R^2 = .63 & & \\
 (2.3) & (1.04) & (.66) & (.95) & & & (5.4) \\
 p < .001 & & p < .005 & p < .001 & & &
 \end{array}$$

Finally, we added the Bantu variable to the spatial autocorrelation analysis.

$$\begin{array}{ccccccc}
 T = 16.4 & +4.06C^* & -1.79S^* & +3.41B^* & R^2 = .67 & \rho = .19 & (5.5) \\
 (3.1) & (.98) & (.61) & (.89) & & (.26) & \\
 p < .001 & & p < .005 & p < .001 & & n.s. &
 \end{array}$$

The spatial autocorrelation term is no longer significant. Once we have taken account of the spatial clustering of Bantu societies, there is no longer spatial autocorrelation in the system of variables which is to say that there is no other significant spatial autocorrelation than that which can be accounted for by the Bantu migration. This is a desirable end result for an autocorrelational analysis because we are able to provide a meaningful description of the process that produced the observed autocorrelation.

For the above example, comparison of the OLS and ML procedures would not lead us to draw any conflicting inferences concerning the importance of particular independent variables: all of the regression coefficients were highly significant using both procedures. However, the observed variations in the magnitudes of the coefficients and their standard errors clearly indicates that such conflicting inferences could easily arise (see Doreian 1980a; Hepple 1976).

Conclusion

The above examination of the spatial autocorrelation approach in the comparative literature has led to some interesting conclusions. First, the currently proposed multivariate techniques fail to deal adequately with the problem of sample unit interdependence. Artificially constructed one-dimensional alignments are a fundamental source of measurement error, and

in failing to model any real underlying process they introduce the possibility of spurious results. Second, the various correlational and regression procedures proposed to model the associations among traits are not analytically generalizable to the more realistic situation of several variables and multilateral dependencies among sample units. Moreover, if the usual multivariate correlation or regression procedures are applied without modification to interdependent data, they yield inconsistent or inefficient estimates.

While the notion of multilateral dependencies is attractive theoretically, it introduces new analytical and computational difficulties. Recent results have brought some of these more complex models well within computational reach, however. Four multivariate regression models which explicitly hypothesize spatial or other networks of relationships among the sample units were presented and discussed here. The emphasis in this type of modeling shifts from attempting somehow to restore independence to describing realistically the underlying structure of interdependencies in a connectivity matrix, and then incorporating this structure into the estimating procedures. Although these models are certainly complex conceptually and computationally, they constitute a more powerful formal characterization of Galton's Problem than previous approaches.

REFERENCE NOTES

1. In a footnote to his next Galton Problem paper Wirsing (1975) observes that the Geary statistic may be used to test for autocorrelation among regression residuals. However, he offers no further suggestions as to how such two dimensional autocorrelation might be incorporated into the regression model.
While it is true that the Geary contiguity ratio is a valid test for autocorrelation among regression residuals, it is not the most powerful test available. A test initially proposed by Moran, modified by Dacey and generalized by Cliff and Ord (1973), was shown by Cliff and Ord and independently by Hepple (1974) to be statistically more powerful. Computational details for this I statistic are provided by Cliff and Ord (1973), although Hepple (1976) has provided a matrix formulation which considerably simplifies the algebra.
2. Pryor's claim that his procedures have "greater statistical power" than previous methods, is clearly an overstatement. Since the true degrees of freedom are unknown it is obviously impossible to make any determination of the type II error rates, hence it is impossible to construct a power function for any of his tests. Given that Pryor nowhere refers to type II error rates, it seems likely that he does not actually intend to claim greater "statistical" power for his tests. Aside from drawing attention to the need for more realistic measures of effects between societies (i.e., Diffusion Possibility Indices), it is doubtful that any of Pryor's procedures are statistically more powerful than the clustering tests based on zero/one diffusion effects discussed by Wirsing (1974b).
3. For simplicity, we assume that the Y variable is centered about its mean. Thus the intercept term may be ignored.
4. See Dhrymes (1970:10) for a discussion of Jacobians.
5. We assume $(I - \rho W)$ is invertible. This will almost always be the case in practice.

6. Doreian (1980a) has developed full ML procedures for this model which simultaneously solve for ρ and β . However, the relative merits of Doreian's and Ord's procedures are not known. We mention both only to indicate that each appears to carry different implications concerning autocorrelation in the measured variables, though both are basically concerned with locating valid and efficient regression estimates.
7. A "lag one" matrix W may be constructed by setting $w_{ij} = 1$ if there exists a society k such that i and k are contiguous and k and j are contiguous, otherwise $w_{ij} = 0$. Generalization of this notion to higher order "lags" is immediate.
8. The R^2 's reported here are simply the squared correlations between the dependent y variables and the predicted \hat{y} variables. Because of interdependence, these R^2 's cannot strictly be given the usual variance explained interpretation. They do give an approximate indication of the degree of fit of the model, however.

REFERENCES

- Bodson, P. and D. Peeters
1975 Estimation of the Coefficients of a Linear Regression in the Presence of Spatial Autocorrelation. An application to a Belgian Labor-Demand Function. *Environment and Planning* 7:455-472.
- Bohrnstedt, George and Thomas Carter
1971 Robustness in Regression Analysis. *In* *Sociological Methodology*. H. Costner, ed. pp. 118-146. San Francisco: Jossey-Bass.
- Brandsma, Andre S., and R. H. Ketellapper
1979 A Biparametric Approach to Spatial Autocorrelation. *Environment and Planning* 11:51-58.
- Burton, Michael L., Douglas R. White, and Malcolm M. Dow
1981 Sexual Division of Labor in African Agriculture: A Network Autocorrelation Analysis. *American Anthropologist*, 83, in press.
- Cancian, Frank
1979 *The Innovator's Situation*. Stanford: Stanford University Press.
- Cliff, Andrew D. and J. Keith Ord
1973 *Spatial Autocorrelation*. London: Pion.
1975 Model Building and the Analysis of Spatial Pattern in Human Geography. *Journal of the Royal Statistical Society. Series A*: 297-328.
- Dhrymes, Phoebus
1970 *Econometrics*. New York: Harper and Row.
- Doreian, Patrick
1980a Linear Models with Spatially Distributed Data: Spatial Disturbances versus Spatial Effects? *Sociological Methods and Research* 9:29-60.
1980b A Maximum Likelihood Procedure for a Linear Model with a Spatial Effect and a Spatial Disturbance Term. Manuscript, Department of Sociology, University of Pittsburgh.
1981 Estimating Linear Models with Spatially Distributed Data. *In* S. Leinhardt, ed. *Sociological Methodology*. San Francisco: Jossey-Bass.
- Dow, Malcolm M.
1979 Structure and Process in Cultural-Ecological Complexes: A Comparative Canonical-Redundancy and Causal Analysis. Ann Arbor: Xerox University Microfilms, #7919299
Ph.D. Dissertation, School of Social Sciences, University of California, Irvine.
- Dow, Malcolm M., Douglas R. White, and David E. Hansen
1979 Regression Analysis and Network Autocorrelation. Manuscript, School of Social Sciences, University of California at Irvine.
- Driver, Harold E.
1966 Geographical-Historical versus Psycho-Functional Explanations of Kin Avoidances. *Current Anthropology* 7:131-160.

- Duncan, Otis D.
1975 *Structural Equation Models*. New York: Academic Press.
- Ember, Melvin
1971 An Empirical Test of Galton's Problem. *Ethnology* 10:98-106.
- Gatrell, A.C.
1979 Autocorrelation in Spaces. *Environment and Planning* 11:507-516.
- Geary, Robert C.
1954 The Contiguity Ratio and Statistical Mapping. *The Incorporated Statistician* 5: 115-141.
- Heath, Dwight B.
1958 Sexual Division of Labor and Cross-Cultural Research. *Social Forces* 37:77-79.
- Hepple, Lawrence W.
1974 *Econometric Estimation and Model-Building with Spatial Series*. Ph.D. Dissertation, University of Cambridge, England.
1976 A Maximum-Likelihood Model for Econometric Estimation with Spatial Series. *In* *Theory and Practice in Regional Science*. I. Masser, ed. pp. 90-104. London: Pion.
- Hibbs, Douglas
1974 Problems of Statistical Estimation and Causal Inference in Dynamic Time Series Models. *In* *Sociological Methodology*. H. Costner, ed. pp. 252-308. San Francisco: Jossey-Bass.
- Johnston, Jack
1972 *Econometric Methods*. New York: McGraw Hill.
- Loftin, Colin
1972 Galton's Problem as Spatial Autocorrelation: Comments on Ember's Empirical Test. *Ethnology* 11:425-435.
1975 Partial Correlation as an Adjustment Procedure for Galton's Problem. *Behavior Science Research* 10:131-42.
- Loftin, Colin and Robert Hill
1974 A Comparison of Alignment Procedures for Tests of Galton's Problem. *In* *Studies in Cultural Diffusion*. J. Scafer, ed. pp. 23-61. New Haven: Human Relations Area Files.
- MacDonald, Kenneth
1976 Causal Modeling in Politics and Sociology. *Quality and Quantity* 10:189-208.
- Malinvaud, Edmond
1970 *Statistical Methods of Econometrics*. Amsterdam: North-Holland.
- Martin, Ronald L.
1974 On Spatial Dependence, Bias and the Use of First Spatial Differences in Regression Analysis. *Area* 6:185-194.
- Murdock, George P. and Douglas R. White
1969 Standard Cross-Cultural Sample. *Ethnology* 8:329-369.
- Naroll, Raoul
1964 A Fifth Solution to Galton's Problem. *American Anthropologist* 66:863-67.
1970 Galton's Problem. *In* *Handbook of Method in Cultural Anthropology*. R. Naroll and R. Cohen, eds. Garden City, N.Y.: Natural History Press. Reprinted 1973, New York: Columbia University Press.
1976 Galton's Problem and HRAFLIB. *Behavior Science Research* 11:123-146.
- Orcutt, G.H., and S.F. James
1948 Testing the Significance of Correlations Between Time Series. *Biometrika* 35:397-418.
- Ord, Keith
1975 Estimation Methods for Models of Spatial Autocorrelation. *Journal of the American Statistical Association* 70:120-126.

- Pryor, Frederick
1976 The Diffusion Possibility Method: A More General and Simpler Solution to Galton's Problem. *American Ethnologist* 3:731-749.
- Rogers, Everett
1971 *Communication of Innovations: A Cross-Cultural Approach*. New York: Free Press.
- Ross, Marc and Elizabeth Homer
1976 Galton's Problem in Cross-National Research. *World Politics* 1:1-28.
- Schaefer, James, ed.
1974 *Studies in Cultural Diffusion: Galton's Problem*. New Haven: Human Relations Area Files, HRAFlex Books.
- Simonton, Dean Keith
1975 Galton's Problem, Autocorrelation, and Diffusion Coefficients. *Behavior Science Research* 4:239-249.
- Stephan, Frederick F.
1934 Sampling Errors and the Interpretation of Social Data Ordered in Time and Space. *Journal of the American Statistical Association* 29:165-166.
- Taylor, Edward B.
1970 On a Method of Investigating the Development of Institutions. *In Readings in Cross-Cultural Methodology*. F. Moore, ed. New Haven: HRAF Press.
- Whittle, Peter
1954 On Stationary Processes in the Plane. *Biometrika* 41:434-449.
- Wirsing, Rolf
1974a *Matrix Methods: New Ways of Measuring Diffusion*. Ph.D. Dissertation, Department of Anthropology, SUNY at Buffalo.
1974b *Measuring Diffusion: The Geary Method and the Dacey Method*. *In Studies in Cultural Diffusion: Galton's Problem*. James M. Schaefer, ed. pp. 197-220. New Haven: HRAF Press.
1975 Second Order Partials as a Means to Control for Diffusion. *Behavior Science Research* 2:143-159.