

# UC Irvine

## UC Irvine Previously Published Works

### Title

Galton's Problem as network autocorrelation

### Permalink

<https://escholarship.org/uc/item/3zq132jj>

### Journal

American Ethnologist, 11(4)

### ISSN

0094-0496

### Authors

DOW, MM  
BURTON, ML  
WHITE, DR  
[et al.](#)

### Publication Date

1984

### DOI

10.1525/ae.1984.11.4.02a00080

### 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

# Galton's Problem as network autocorrelation

**MALCOLM M. DOW**—Northwestern University

**MICHAEL L. BURTON**—University of California, Irvine

**DOUGLAS R. WHITE**—University of California, Irvine

**KARL P. REITZ**—University of California, Irvine

Ethnologists have long sought to establish and validate causal models of sociocultural phenomena by the use of correlational evidence across broad samples of societies. It is well known that correlation does not establish causation. However, evidence for causality is strengthened if a postulated correlation replicates in many different contexts, thus radically reducing the likelihood of hidden third factors, spurious correlations, or correlations due to chance composition of a sample.

In the use of correlation as a valid measure of the relationship between two variables, a necessary condition is that the sample units must be independent. Nonindependence can occur in a variety of ways, and its consequences can be quite severe. Even though a correlation exists between two variables, its estimate via a nonindependent sample can be grossly inflated or reduced. In cross-cultural research, independence of sample units is almost impossible to achieve, a fact that has made cross-cultural research open to a good deal of valid criticism. The eminent statistician Sir Francis Galton first pointed out the problem in his remarks following E. B. Tylor's (1889) presentation of his classic cross-cultural paper. The problem has since been known in anthropology as Galton's Problem (Naroll 1961, 1973; Schaeffer 1974).

The nonindependence of sample societies stems from the fact that humans have evolved

---

*Classical statistical inference procedures usually assume the independence of sample units. However, the assumption of independence is often unrealistic in cross-cultural research because societies in neighboring or historically related regions tend to be duplicates of one another across a wide variety of traits that are spread by historical fission, diffusion, or migration of peoples. A recent generalization of the usual regression model explicitly allows for networks of interdependencies among sample units as part of the model specification. Here, two new estimation procedures for this network autocorrelation model are compared to previously employed maximum likelihood procedures, and to the usual regression procedures which ignore interdependence. The results of comparisons based on simulated autocorrelation data and the reanalyses of two previously published empirical studies indicate that both of the procedures proposed here compare very favorably with the maximum likelihood approach, and both are vastly superior to the usual regression procedures when there is moderate to high autocorrelation (i.e., interdependence). [Galton's Problem, cultural diffusion, networks, cultural evolution, statistical methodology]*

---

Copyright © 1984 by the American Ethnological Society  
0094-0496/84/040754-17\$2.20/1

from a common evolutionary stock and from the diffusion of cultural traits among societies. Societies in neighboring or historically related regions tend to be duplicates of one another in terms of a wide variety of traits that are spread by historical fission, diffusion, or migration of peoples. The result is that neither the actual number of "independent" cases nor the effect of the interdependencies on trait correlations is generally known for any cross-cultural sample.

Galton's Problem applies with equal force to attempts to establish and validate multivariate causal models. Regression models are commonly used for this purpose. Regression coefficients may be under- or overestimated, depending on the extent and structure of non-independent cases. Here the problem is known as *autocorrelation*, where variables or error terms in the regression equation are correlated with the scores of related cases on the same variables or error terms. In technical terms, the presence of autocorrelation produces inefficient estimates of the regression coefficients: estimates will vary widely from the true coefficient. This problem is compounded by the fact that the variances of these coefficient estimates are systematically underestimated. As a result, the following may occur: (1) Where little or no correlation exists between two variables, presence of autocorrelation in the sample may result in an estimate of the regression coefficient not only substantially different from zero but also with an underestimate of its variance. This may lead the researcher to conclude that the correlation is significantly different than zero. (2) Where the relationship between two variables is indeed substantial and replicable across samples, the estimate of the same regression coefficient via two different autocorrelated samples may result in large differences between the two estimates. When this is coupled with an underestimated variance, a researcher may be led to conclude that the estimates are significantly different and that the relationship does not replicate.

Autocorrelation was first formulated as a problem with time-series data, for which events at a given point in time are generally not independent of the events at points just preceding in time. For cross-sectional data the comparable formulation was in terms of spatial autocorrelation, for which events at a given point in space are not independent of events at points nearby. In the past few years, several cross-cultural researchers have discussed the relevance of the spatial autocorrelation model to Galton's Problem (Loftin 1972; Naroll 1976; Pryor 1976; Simonton 1975; Wirsing 1974). That formulation requires measurement of interdependencies among societies in terms of spatial distance. The present authors (Dow, White, and Burton 1983; White, Burton, and Dow 1981) have generalized the spatial autocorrelation model to incorporate any kind of historical or diffusional relationship among societies. This network autocorrelation model can use the statistical solutions to the spatial autocorrelation model; only the measures of relationships among societies will differ.

Valid solutions to the spatial autocorrelation model have been developed only recently (Ord 1975; Cliff and Ord 1981; Doreian 1980, 1981; Bartels and Ketellapper 1979). In a previous paper (Dow, White, and Burton 1983) we review four models in which autocorrelation enters into the regression framework and suggest that one of them, the disturbances model, appears to be well suited for cross-cultural research. In this model autocorrelation appears in the error terms of the regression equation. This assumes that there are patches of related cases which fit the model better than expected by chance, other patches for which scores on the dependent variable are systematically overestimated, and remaining patches for which the dependent variable is underestimated. Such a model is consistent with the case in which the variables under study tend to travel (spread, diffuse) as a packet. The great advantage of this autocorrelated spatial disturbances model is that regression coefficients can be properly (efficiently) estimated even with the presence of autocorrelation, by the use of maximum likelihood (ML) procedures. Mathematical, empirical, and

simulation work (Dow, Burton, and White 1982) have all shown that these ML autocorrelation procedures yield markedly different and more accurate results, in the presence of autocorrelation, than ordinary least squares (OLS) regression, which assumes the absence of autocorrelation. The advantages of the ML procedure over OLS are of a major order even in small samples (Dow, Burton, and White 1982).

Autocorrelation solutions to Galton's Problem in the regression context depend on an explicit specification of the strength of relatedness between each pair of cases in the sample. A restricted and incomplete approach to such specifications, using one-dimensional arrays or "diffusion arcs," was introduced to the cross-cultural literature by Naroll (1961, 1973), Loftin (1972), and Simonton (1975). A more powerful two-dimensional or "diffusion proximity" specification was introduced by Wirsing (1974) and Pryor (1976). The weaknesses of the Loftin and Simonton arguments as autocorrelation solutions to Galton's Problem consist of underestimating the amount of diffusion by assuming that (1) diffusion operates only in one direction and (2) is never reciprocal. The weaknesses of the Wirsing and Pryor arguments are not in the specification of the interdependency matrix but in how this is incorporated into the estimation procedures for the regression model. Our network disturbances model builds in the proper estimation procedures.

While the network disturbances model has great advantages for cross-cultural and causal-modeling research, its chief disadvantage to date has been the computational effort required by the ML estimation procedure. This is true in spite of Ord's (1975) useful simplification of the estimating routine. The procedure is formidable even with small samples, more so with large ones. And, unless the structure of the interdependencies in large samples assumes some particularly simple forms (Ord 1975:125), the ML computational routine cannot be employed.

The major purpose of this paper, then, is to examine two estimation procedures for the network autocorrelated disturbances model, both of which are considerably simpler computationally than the ML procedure. Although neither procedure generates estimates which are preferable to ML estimates on purely formal criteria, we compare the empirical performance of each procedure to the ML approach and to the usual OLS regression procedures using simulated autocorrelation data. The results of this simulation indicate that both methods compare very favorably with the ML procedure in terms of bias and efficiency of the regression coefficient estimates. Both methods also offer major improvements over the usual OLS regression when moderate to high autocorrelation is present. Also, both methods perform extremely well in reanalyses of two previously published autocorrelation examples.

In the following section we review briefly the matrix formalization of the network disturbances model and outline the three autocorrelation estimation methods.<sup>1</sup> We then present and discuss the results of the simulation. In the subsequent section we suggest significance testing procedures which can be used with these alternative routines. Finally, we present the results of our reexamination of two previously reported (White et al. 1981; Pryor 1976) empirical examples using all three autocorrelation methods and OLS.

## **autocorrelated disturbances regression model**

Although Galton's Problem was first raised in connection with the spatial and historical patterning of cultural traits, it is in a statistical sense actually a problem of model specification. A model is correctly specified when the equations in the model adequately correspond to the processes generating the observed variation in the data. Model specification is a matter of degree: the better the theoretical understanding of the underlying processes,

the closer the correspondence between the postulated equations and the empirical data, and the more precise the estimates of the model parameters.

From a cross-cultural perspective, cultural diffusion is one of the theoretically important underlying processes affecting the degree of variation in comparative data. The formal expression of this process in equational form is thus a necessary part of any correctly specified model used to analyze cross-cultural data. Before presenting the network autocorrelated disturbances model in equational form, however, we first examine the representation of cultural diffusion as a matrix of relationships among the sample units.

An important step in Pryor's (1976) attempt at a matrix formulation of Galton's Problem is the calculation of a "diffusion possibility index" between every pair of societies in a sample. Any factor thought to influence the probability of diffusion, such as distance and language similarity, can be used to calculate a "DP index" for each pair. As a purely hypothetical example, Pryor calculates the DP indices between pairs using the formula  $DP = 5 + 3L - 4D$ , where  $L$  is a 0,1 variable indicating language similarity (1) or not (0), and  $D$  is geographical distance measured in thousands of miles. Pryor computes a DP index for every pair drawn from six hypothetical societies ( $T, V, W, X, Y, Z$ ) and assembles the results into a "sample diffusion matrix," shown as the  $S$  matrix in Table 1.

Societies  $T$  and  $V$  in this  $S$  matrix would be most alike ( $DP = 6$ ), while  $T$  and  $Y$  have little similarity ( $DP = 1$ ), as measured by Pryor's hypothetical DP equation. It is entirely possible, of course, to calculate relationships using measures based only on distance or only on language similarity, rather than attempting to combine them in some ad hoc fashion. No matter which measures of diffusion or similarity are preferred, however, the interdependencies among sample units are still formalized as an  $S$  matrix.

The network autocorrelation disturbances model assumes that the  $N$  sample units are somehow differentially interrelated and that the interrelations can be operationalized as an  $N \times N$  "relational" matrix  $W$ . This matrix differs from Pryor's diffusion possibility matrix only in that each element is divided by its row sum; thus, the rows are rescaled to sum to 1. The  $W$  matrix in Table 1 is the diffusion possibility matrix  $S$  appropriately row-normalized to unity. Since a sample unit is not assumed to influence itself, the main diagonal of this matrix has all zeros. The entries in this matrix,  $w_{ij}$ , indicate the interaction probabilities between unit  $i$  and unit  $j$ . Note that now interactions may be reciprocal, but they need not be symmetric (i.e.,  $w_{ij}$  does not necessarily equal  $w_{ji}$ ).

This model is a generalization of the usual OLS (ordinary least squares) multiple regression model, which can be compactly stated in matrix form<sup>2</sup> as

$$Y = X\beta + \varepsilon \quad \varepsilon \sim \text{IND}(0, \sigma_\varepsilon^2 I) \tag{1}$$

Table 1. Matrix representation of sample unit interdependence.

		Pryor's (1976) sample diffusion possibility matrix					Row normalized relational matrix							
		T	V	W	X	Y	Z	T	V	W	X	Y	Z	
S =	T	—	6	5	3	1	2	T	0	.35	.29	.18	.06	.12
	V	6	—	4	4	2	1	V	.35	0	.24	.24	.12	.06
	W	5	4	—	2	2	3	W	.31	.25	0	.13	.13	.18
	X	3	4	2	—	4	5	X	.17	.22	.11	0	.22	.28
	Y	1	2	2	4	—	5	Y	.07	.14	.14	.29	0	.36
	Z	2	1	3	5	5	—	Z	.13	.06	.19	.31	.31	0

where

- Y is an  $N \times 1$  column vector of dependent variable observations;
- X is an  $N \times (K + 1)$  matrix of K independent variables plus an initial column of ones;
- $\beta$  is a  $(K + 1) \times 1$  column vector of regression coefficients;
- $\varepsilon$  is an  $N \times 1$  column vector of (independent) multivariate normal errors.

When the error term assumptions of this model are met, in particular the independence assumption, the ML estimation procedure and the usual OLS procedure yield equivalent results.<sup>3</sup> That is, estimates of the regression coefficients  $\beta$  and their standard errors are identical. However, given an interdependent set of observations, such as frequently occur in continuous area cross-cultural studies or time series studies, the error terms may not be independent of one another, and the above OLS model is incorrectly specified in this case.

Given an interdependent set of observations and an appropriate W relational matrix, the OLS model can be respecified as a network autocorrelated disturbances model:

$$Y = X\beta + \varepsilon \tag{2}$$

$$\varepsilon = \rho W\varepsilon + v \quad v \sim \text{IND}(0, \sigma_v^2 I) \tag{3}$$

This model states that the disturbances  $\varepsilon$  from the previous OLS regression model are not independent. Rather, each disturbance is a weighted average of the disturbances at related units, the weights being the nonzero  $w_{ij}$  in the appropriate row of W, times a scalar  $\rho$  which is analogous to a correlation coefficient, plus an  $N \times 1$  column vector of random components v. Rho ( $\rho$ ) expresses the extent of the error dependencies (i.e., the overall autocorrelation in the system of variables under study). Note that when  $\rho = 0$  (i.e., there is no autocorrelation present with respect to W), this model is identical to the OLS model in equation 1. When W has all zero elements, corresponding to complete independence of each sample unit from all the others, this model is again identical to equation 1. Thus, the network disturbances model is a true generalization of the usual OLS regression model.

Some insight into the nature of the network disturbances model can be obtained by algebraically combining equations 2 and 3. Solving for  $\varepsilon$  in terms of v in equation 3 we get

$$\varepsilon - \rho W\varepsilon = v \tag{4}$$

then  $(I - \rho W)\varepsilon = v$  (5)

and  $\varepsilon = (I - \rho W)^{-1}v$  (6)

where I is the  $N \times N$  identity matrix. Substituting back into equation 2 we get

$$Y = X\beta + (I - \rho W)^{-1}v. \tag{7}$$

Premultiplying this equation by  $(I - \rho W)$  gives

$$(I - \rho W)Y = (I - \rho W)X\beta + v \tag{8}$$

or  $Y - \rho WY = (X - \rho WX)\beta + v$  (9)

This latter equation is more simply written as

$$Y^* = X^* \beta + v \quad v \sim \text{IND}(0, \sigma_v^2 I) \tag{10}$$

where  $Y^* = Y - \rho WY$

and  $X^* = \begin{cases} X_1^* & = & X_1 & - & \rho W X_1 \\ \vdots & & \vdots & & \vdots \\ X_{K+1}^* & = & X_{K+1} & - & \rho W X_{K+1} \end{cases}$  (11)

Note that after this common linear transformation of all variables the error terms in equation 10 now meet OLS assumptions; hence, unbiased and relatively efficient estimates of the regression coefficients  $\beta$  and their standard errors can be obtained using the usual OLS procedures on the transformed data.

Within the regression framework, then, the network autocorrelation problem is twofold:

construction of a plausible relational matrix  $W$  and estimation of the autocorrelation parameter  $\rho$ . Selection of the weights  $w_{ij}$  for the  $W$  matrix is of major importance, since spurious results may be obtained if the hypothesized matrix does not correspond to any real process (Cliff and Ord 1973). In geographical research, interaction effects are often estimated according to notions 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  $D_{ij}^{-\alpha}$ , where  $D_{ij}$  is the distance from location  $i$  to location  $j$  and  $\alpha$  is a suitable exponent chosen a priori. Cliff and Ord (1973) compute weights based on both distance and proportion of boundary in common between the 28 counties of Eire. Gatrell (1979) has constructed a measure of interaction among Swedish towns based on distance and number of telephones. Bodson and Peeters (1975) employ distance and minimum public transportation time when constructing an "accessibility function" to generate the interaction weights among 44 Belgian *arrondissements*.

Clearly, the form of the weighting matrix will depend on the problem at hand and on the available data. In cross-cultural research, many possibilities exist for computing 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 measure of language similarity has been proposed and employed in previous cross-cultural examples (White et al. 1981; Dow in press). Language similarity and geographical distance measures are used individually to construct the two  $W$  matrices employed in the analyses reported below.<sup>4</sup>

If there is autocorrelation, the network disturbances model has desirable statistical properties that the OLS model lacks. First, although the OLS estimates of regression coefficients are unbiased, they are more highly dispersed around the true population parameter than are the ML estimates. Second, the OLS estimates of the sampling variances of the regression coefficients will generally underestimate the true variances. Hence, with respect to estimates based on a single sample, underestimating the variances of the regression coefficients will lead to spurious attributions of significance to particular independent variables. By contrast, in replication studies across several interdependent samples, the investigator would tend to conclude that a valid model fails to replicate because of large differences in estimates of the same parameter due to their unreliability (i.e., wide dispersion around the true parameter). Thus, both single and multiple sample replication studies are biased toward finding differences where none exist (type 1 error) if OLS estimation is used with interdependent samples (Dow, Burton, and White 1982). This is a crucial issue in cross-cultural research whereby functional relationships are expected, if valid, to replicate across the major geographical/cultural regions of the world.

### three estimation procedures

**maximum likelihood (ML)** In the usual application of the regression model, estimates of the coefficients and their standard errors can be obtained by the method of moments. That is, means, variances, and covariances of all variables are calculated and can then be used to find estimates of the population parameters. An alternative procedure which does not employ sample moments is maximum likelihood (ML) estimation. The basic idea behind this procedure is to try to locate estimates which are "most likely" to have generated the observed sample values. Details on the estimation theory for the ML network disturbances model and other ML autocorrelation models are given in Ord (1975), Doreian (1980), Cliff

and Ord (1981). Since we shall compare two simpler solutions (IGLS and IRR; see below) to the ML solution, we present here an outline of the ML solution.

The basic problem is to locate estimates  $\hat{\rho}, \hat{\beta}$  and  $\hat{\sigma}_\epsilon^2$  for the network disturbances model, plus the standard errors of these estimates. The major computational problem is to find the ML estimate of  $\hat{\rho}$  which permits the simple variable transformations in equation 11. The appropriate ML estimate is the value  $\hat{\rho}$  which minimizes the following expression (Doreian 1980)<sup>5</sup>:

$$\ln(Y'A'PAY) - \frac{2}{N} \sum_{i=1}^N \ln(1 - \rho\lambda_i) \quad (12)$$

where

$$\begin{aligned} A &= I - \rho W \quad (I \text{ is the } N \times N \text{ identity matrix}) \\ P &= I - [(AX)(AX)'AX]^{-1}(AX)' \\ \lambda_i &= i^{\text{th}} \text{ eigenvalue of } W. \end{aligned}$$

For the results reported below, expression 12 was minimized by direct search of the interval  $(-1, 1)$ .<sup>6</sup> As the sample  $N$  increases, repeated evaluation of expression 12 becomes computationally burdensome. Note that the summation term requires the  $N$  eigenvalues of the  $W$  matrix. As mentioned above, unless  $W$  assumes some particularly simple form, its eigenvalues are virtually impossible to compute for a sample of  $N$  over about 90.<sup>7</sup>

With  $\hat{\rho}$  obtained from the above procedures the following parameters can be estimated:

$$\hat{\beta} = [X'A'AX]^{-1} X'A'AY \quad (13)$$

which amounts to a generalized least squares (GLS) regression of  $AY$  on  $AX$ , and

$$\hat{\sigma}_\epsilon^2 = \frac{1}{N} [\hat{\epsilon}' \Lambda \hat{\epsilon}] \quad (14)$$

It is of course desirable to perform the usual inferential procedures on  $\hat{\rho}$  and  $\hat{\beta}$ . Consistent estimates of their standard errors are obtained from

$$V(\hat{\sigma}_\epsilon^2, \hat{\rho}, \hat{\beta}) = \sigma^4 \begin{bmatrix} N/2 & \sigma^2 \text{tr}(B) & 0' \\ \sigma^2 \text{tr}(B) & \sigma^4 [\text{tr}(B'B) - \alpha] & 0' \\ 0 & 0 & \sigma^2 X'A'AX \end{bmatrix}^{-1} \quad (15)$$

where

$$\begin{aligned} \alpha &= - \sum_{i=1}^N \lambda_i^2 / (1 - \rho\lambda_i)^2 \\ B &= WA^{-1} \\ 0' &= 1 \times (K + 1) \text{ row vector of zeros} \\ 0 &= (K + 1) \times 1 \text{ column vector of zeros.} \end{aligned}$$

That is, with the appropriate estimates of  $\hat{\sigma}_\epsilon^2$ ,  $\hat{\rho}$ , and  $\hat{\beta}$  inserted into equation 15, and the total expression evaluated, the standard error of  $\hat{\rho}$  is given by the square root of the (2,2) element of  $V$ , and the standard errors of the  $\hat{\beta}_j$ 's by the square roots of the remaining main diagonal entries.

**Iterative generalized least squares (IGLS)** The iterative generalized least squares procedure is based in part on suggestions made by Ord (1975) and Bodson and Peeters (1975). For the ML approach, the criterion used to decide on the best estimate  $\hat{\rho}$  was that the estimate minimize equation 12, which is equivalent to maximizing the total log-likelihood function. In the IGLS procedure, we sidestep the difficulties associated with finding the eigenvalues of  $W$  and the repeated evaluation of equation 12. We iteratively search the interval  $-1 < \rho < 1$  as before; now, however, we (1) insert the search values of  $\hat{\rho}$  directly into equation 11 and perform the variable transformations; (2) apply OLS to the resulting equation 10 and obtain estimated residuals  $\hat{\epsilon}$ ; (3) retain as the "best" estimate of  $\rho$  the  $\hat{\rho}$  which results in the minimum sum of squares of the residuals from equation 10; (4) retain as our



estimates of  $\beta_j$  the  $\hat{\beta}_j$ 's corresponding to this  $\hat{\rho}$ ; (5) retain the associated standard errors of the  $\hat{\beta}_j$ .

As with the above direct-search ML estimation, it is possible that the search procedure may locate a local rather than a global minimum  $p$  within the search interval. A fairly fine-grained search would make this unlikely, however. The simulation results reported below clearly suggest that this procedure will estimate a global minimum. It is also possible that no unique minimum sum of squares of errors may be found. This possibility did not occur in any of the 200 simulation direct searches reported below; nor did it arise with any of the empirical examples.

**Iterative residual regression (IRR)** The iterative residual regression procedure originated in Durbin's (1960) approach to time series analysis and is discussed in Bodson and Peeters (1975) for the spatial autocorrelation case. For this procedure the steps are: (1) estimate  $Y = Y\beta + \varepsilon$  using OLS and obtain the residuals  $\hat{\varepsilon} = Y - X\hat{\beta}$ ; (2) estimate  $\hat{\rho}$  by regressing  $\hat{\varepsilon}$  on  $W\hat{\varepsilon}$  using OLS (i.e., on  $\hat{\varepsilon} = \rho W\hat{\varepsilon} + v$ ); (3) use this  $\hat{\rho}$  to transform variables as in equation 11; (4) estimate equation 10 using OLS and obtain a new residual vector  $\hat{v}$ ; (5) use these residuals to obtain another estimate  $\hat{\rho}$  as in step 2; (6) terminate the procedure when successive estimates of  $\hat{\rho}$  are within some prespecified absolute value<sup>8</sup>; (7) retain the last value of  $\hat{\rho}$  and associated  $\hat{\beta}_j$  and their standard errors.

Whether this estimation procedure converges to some minimum is unknown for multi-lateral interdependencies. For the 200 simulations of this model, the maximum limit we set of 10 iterations was never reached. For the simulation data analyzed here, convergence was very rapid, averaging less than four iterations. Again, the close similarity of the results of this method to the results of the other two procedures suggests that it does indeed converge to a global minimum.

### generation of data and relational matrix $W$

The procedures employed here to generate the simulation data are similar to those reported in Dow, Burton, and White (1982). The following model was used to generate the data:

$$Y = O + 1 \bullet X + \varepsilon \tag{16}$$

$$X = \lambda WX + \zeta \tag{17}$$

$$\varepsilon = \rho W\varepsilon + v \tag{18}$$

where  $\zeta \sim \text{IND}(50,64I)$  and  $v \sim \text{IND}(0,81I)$ .

Since our main interest is in the relative performance of each of the three procedures with respect to the bias and efficiency of the  $\beta$  estimate, the intercept was set equal to zero. First, we generated a random column vector  $\zeta$  with mean and variance as in expression 18. An autocorrelated independent  $X$  variable was then obtained using

$$X = (I - \lambda W)^{-1} \zeta \tag{19}$$

The extent and direction of OLS estimation error is a function of the degree of autocorrelation of the independent variable (i.e.,  $\lambda$ ) and of the errors (i.e.,  $\rho$ ); therefore, we set  $\lambda = .4$ . For each independent  $X$  vector we generated an autocorrelated error vector by drawing a random vector  $v$  and transforming it as before:

$$\varepsilon = (I - \rho W)^{-1} v \tag{20}$$

Given the  $X$  and  $\varepsilon$  vectors, we then constructed the dependent  $Y$  variable using equation 15. That is, we simply add the  $X$  and the  $\varepsilon$  vectors. These steps were repeated 50 times for values of  $\rho = .23, .43, .63, .83$  and a fixed sample size of 35. Only positive values of  $\lambda$  and  $\rho$

were used, since negative autocorrelation is less interesting theoretically and less likely empirically. Also, the expense of computing using the ML program precluded examination of all possible autocorrelation values.

The 35 × 35 relational matrix *W* used in generating the data via equations 19 and 20 was obtained from the language family relationships of 35 Sub-Saharan African societies from the Standard Cross-Cultural Sample (Murdock and White 1969). Briefly, linguistic similarity is measured on the genetic tree of languages as an inverse function of the number of nodes along the path between the two languages. Further details on the construction of this and other cross-cultural relational matrices are provided in White et al. (1981).

## simulation results

Our previous study of the ML disturbances model (Dow, Burton, and White 1982) showed that there were no significant gains over OLS for low levels of error autocorrelation (i.e.,  $\rho < .5$ ). Hence, we did not employ our ML program on the simulated data where  $\rho = .23$  or  $.43$ . We compare the performance of the three autocorrelation procedures to each other and to OLS with respect to bias and efficiency of the regression coefficient  $\beta$ , and then compare the three autocorrelation procedures with respect to bias in the estimate of the autocorrelation parameter  $\rho$ .

**bias in the  $\beta$  estimates:  $E(\beta - \hat{\beta})$**  Table 2 gives the average  $\beta$  over 50 replications at each level of  $\rho$ . Since the true  $\beta = 1$ , the average bias introduced by each procedure corresponds to the values reported in Table 2, minus 1. Clearly, there is very little bias using any of the autocorrelation methods, as we expected from the previous discussion. For autocorrelation at  $.43$  and above, OLS displays the highest bias.

Table 2. Average of 50 estimates of  $\beta$  at each level of  $\rho$  (true  $\beta = 1$ ).

$\rho$	OLS	IGLS	IRR	ML
.23	1.050	1.057	1.059	—
.43	1.044	1.007	1.013	—
.63	1.070	1.008	1.009	1.012
.83	1.135	1.011	1.010	1.006

**relative efficiency of  $\beta$  estimates** As mentioned previously, the problem with OLS analyses of autocorrelated data is that the regression coefficient estimates, though unbiased, are inefficient. That is, the variance of the estimates is large relative to that obtainable through autocorrelation procedures. It is usual to compare estimators using a measure which combines both bias and efficiency, the mean square error (MSE), since the overall performance of an estimator depends on both quantities. Using expectation notation, MSE is defined as

$$E(\beta - \hat{\beta})^2 = (\text{bias})^2 + \text{variance} \quad (21)$$

The MSE is the expected value (i.e., average) of the squared differences of the estimates  $\hat{\beta}$  from the true population parameter  $\beta$ .

Table 3 reports the ratios of MSE's of  $\beta$  for all procedures. If the MSE ratio equals 1, then the procedures being compared are equally efficient; if the ratio is less than 1, then the procedure used to calculate the numerator is more efficient (smaller MSE) than the procedure used to calculate the denominator; and if the ratio is more than 1, the opposite is true.

Table 3. Ratios of MSE ( $\beta$ ) for all estimation procedures.

$\rho$	<u>OLS</u> IGLS	<u>OLS</u> IRR	<u>IGLS</u> IRR	<u>IGLS</u> ML	<u>IRR</u> ML
.23	.751	.595	.792	—	—
.43	1.021	.946	.926	—	—
.63	2.372	2.384	1.005	1.041	1.038
.83	7.394	7.396	1.000	.989	.986

The first two columns of Table 3 indicate that at  $\rho = .23$ , OLS is more efficient than either IGLS or IRR, and that there is little to distinguish between the OLS and either of these two procedures at  $\rho = .43$ . However, at the higher levels of  $\rho$  both IGLS and IRR show huge gains in efficiency over OLS, anywhere from about 100 to 700 percent. The third column indicates that there is virtually no difference in the performance of IGLS or IRR at these higher levels of  $\rho$ . The last two columns also indicate, rather surprisingly, that for the data analyzed here, at least, there is little to distinguish between IGLS and IRR and the more formally correct ML procedure. The former two are marginally more efficient at moderately high levels of autocorrelation ( $\rho = .63$ ), while ML is marginally more efficient at high autocorrelation levels ( $\rho = .83$ ).

**bias in  $\rho$  estimates** Table 4 shows that at low levels of autocorrelation ( $\rho = .23$ ) neither IGLS nor IRR yield satisfactory estimates of  $\rho$ . At the next level of autocorrelation ( $\rho = .43$ ) both procedures tend to underestimate the true  $\rho$ , while at higher levels they both tend to overestimate  $\rho$  by about 15 percent. ML provides better estimates of  $\rho$  than either IGLS or IRR at these higher levels of autocorrelation. However, this difference appears to have little or no impact on overall efficiency of the regression estimates, as shown in Table 3.

Table 4. Average of 50 estimates of  $\rho$ .

$\rho$	IGLS	IRR	ML
.23	-.045	-.222	—
.43	.376	.315	—
.63	.769	.759	.600
.83	.940	.928	.867

### significance testing procedures for autocorrelation parameter $\rho$

In the above discussion of the ML computational procedures, expression 14 was given for obtaining the standard error of  $\rho$ . Since  $\rho/\sqrt{\text{var}(\rho)}$  is asymptotically normally distributed, the likelihood of significant autocorrelation is easily evaluated. The exact distribution of the  $\hat{\rho}$  estimate generated by IGLS or IRR is not known.

One approach to discerning whether or not a statistically significant network process is operative is to employ Cliff and Ord's (1973) I-statistic. This statistic can be applied to the residuals obtained from an initial OLS regression to assess whether significant autocorrelation is occurring with respect to a given relational matrix  $W$ .<sup>9</sup> Hepple (1976) has provided a similar statistic for testing OLS residuals which is considerably simpler to compute than the I-statistic. Given the OLS vector of residuals  $\hat{\epsilon}$ , Hepple's s-statistic can be converted to a standard normal deviate as follows:

$$Z = \frac{s - E(s)}{\sqrt{\text{var}(s)}} \quad (22)$$

where

$$s = \varepsilon' W \varepsilon / \hat{\varepsilon}' \varepsilon \quad (23)$$

$$E(s) = \text{tr}(G'W) / (N-K) \quad (24)$$

$$\text{Var}(s) = 2 \left[ \frac{(N-K)\text{tr}(G'W)^2 - (\text{tr}(G'W))^2}{(N-K)^2(N-K+2)} \right] \quad (25)$$

and

$$G = I - [X(X'X)^{-1}X'] \text{ (I is an } N \times N \text{ identity matrix).}$$

Hence, if a significant  $s$  were detected, either IGLS or IRR could be used to obtain an estimate of the degree of autocorrelation present. The  $Z$  score obtained from this statistic is computed and reported in the empirical examples discussed below.

An alternative approach to testing the  $\hat{\rho}$  generated by either IGLS or IRR for significance would be to accept the  $\hat{\rho}$  they yield as a good approximation to the ML  $\hat{\rho}$  and insert the IGLS or IRR value into a likelihood ratio test (Hoel 1971; Brandsma and Ketellapper 1979). However, the simulation results reported here suggest that this would only offer a very approximate test of significance, so we do not pursue the possibility at this time.

## empirical reanalyses

Pryor (1976) presents three tests for the detection of diffusion of cultural traits. As noted above, his method is based on the idea of a "diffusion possibility matrix" which he suggests could be constructed from the physical distances between societies or on language similarities. As an example he presents data from 60 societies rated on seven variables. He hypothesizes that "the presence of a significant amount of market exchange of labor is directly proportional to the level of economic development of the society and also to whether it relies heavily on herding as a source of subsistence goods" (1976:740). He also hypothesizes that "the presence of gambling in a society is directly proportional to the presence of a commercial money and to the presence of considerable socioeconomic differences and is inversely related to whether or not the society is a nomadic herding society" (1976:740). Using distance as a criterion for the possibility of diffusion, Pryor suggests via his three tests that diffusion of the measured traits is likely. Having found some evidence of diffusion, his options at the time he wrote the paper were severely limited. As we shall show with our more current methodology, Pryor did remarkably well. He notes that gambling was prevalent particularly in North America, so he includes whether a society was a North American society as a (0,1) dummy variable and recomputes his regression equation.

Using our methods, we reanalyzed the data as presented by Pryor. Instead of a diffusion possibility matrix, we computed a network relational matrix based on physical distances, as in our previous example. Using that matrix and our two network autocorrelation procedures, we recomputed the regression coefficients and present them here, along with Pryor's estimates (Tables 5 and 6). The standard deviations of the regression coefficients are reported in parentheses below the coefficients.

In the model for the presence of gambling (Table 5), we find significant autocorrelation ( $z = 6.853$ ). Our estimate of  $\rho$  is large ( $\rho = .795$ ), but our regression coefficients are substantially the same as reported by Pryor in his first analysis, with one important exception. Pryor (1976:741) reports a value of .167 (.114) as the coefficient and standard deviation for the presence of socioeconomic inequality. From these results alone he would have had to conclude that socioeconomic inequality was not a causal factor in the presence of gambling. Using the same variables along with the matrix of distance similarities, we find a

Table 5. Presence of gambling.

Coeff. name	Pryor's 1st results	Pryor's 2nd results	IRR	IGLS	IRR with NA
z	—	—	6.853	6.853	1.263
p	—	—	.795	.800	.292
Intrcpt	.123	-.281	-.341 (.337)	-.347 (.346)	-.369 (.175)
MON	.470 (.193)	.415 (.126)	.465 (.129)	.466 (.129)	.422 (.122)
SOCED	.167 (.114)	.237 (.075)	.321 (.067)	.322 (.067)	.274 (.068)
NMHRD	-.692 (.298)	-.430 (.196)	-.638 (.213)	-.638 (.213)	-.479 (.194)
NA	—	1.33 (.151)	—	—	1.291 (.198)
R <sup>2</sup>	.221	.676	.198	.197	.656

Notes: MON is the presence of commercial money.  
 SOCED is the presence of socioeconomic inequality.  
 NMHRD is the presence of nomadic herding.  
 NA society is in North America.  
 ( ) is the standard error.

coefficient of .321 (.067), which is double Pryor's coefficient, with half the standard error. This means that a relationship which remains undetected using standard techniques is shown to be highly significant when the autocorrelation is taken into account.

When Pryor included North America as a dummy variable, the coefficient for socioeconomic inequality increased and, more importantly, was significant. By use of this dummy variable Pryor found a relationship not seen in his first analysis. At the same time  $R^2$  increases from .221 to .676, which indicates that a great deal of the variation in the occurrence of gambling can be explained by the presence of the society in North America. When we include North America in our regression equation, there is no longer significant autocorrelation ( $z = 1.263$ ) and our regression coefficients coincide with Pryor's. This indicates that the autocorrelation which we report in our first analysis is to be found mainly in North America.

Again applying the techniques to the model for labor market exchange (Table 6), the test for autocorrelation gives a value of  $z = -.492$ , which is insignificant. Although computation of the regression estimates under a model of autocorrelated disturbances is inappropriate in the case where autocorrelation is insignificant, we have presented those computations to illustrate that estimates of the main parameters of interest remain essentially the same as in standard OLS procedures. First note that the results of the two procedures

Table 6. Labor market exchange.

Coeff. name	Pryor's results	IRR	IGLS
z	—	-.492	-.492
p	—	-.198	-.200
Intrcpt	-.103	-.095 (.091)	-.094 (.091)
ECDEV	.017 (.003)	.017 (.003)	.017 (.003)
HERD	.551 (.138)	.563 (.133)	.564 (.133)
R <sup>2</sup>	.452	.437	.437

Notes: z is the normally distributed measure of the presence of autocorrelated disturbances.  
 p is the coefficient of autocorrelation.  
 ECDEV is the level of economic development.  
 HERD means the most important source of food comes from herding.  
 ( ) is the standard error.

(IRR and IGLS) are almost identical. This indicates that the process probably has converged to a global minimum. Although Pryor reports evidence of diffusion, we find that autocorrelation is statistically insignificant and that his OLS coefficients are valid and correct. He does report that clustering occurred only among the dependent variables, which is true and is an indication that perhaps the disturbances model is not correct and that the "effects" model for autocorrelation may be more appropriate.<sup>10</sup> It should also be emphasized that since there is insignificant autocorrelation, the negative values for  $\rho$  are meaningless.

These results verify and illustrate our simulation findings. First, there is little difference between the results of IGLS and IRR; however, since they are entirely different procedures, the fact that they give identical results is an almost sure check that they have converged to a global minimum. Second, we have illustrated a case in which significant autocorrelation did not exist and yet the results of the two procedures give results identical to OLS. Finally, and of greatest importance to cross-cultural researchers, is the fact that a relationship which could have been completely missed using standard techniques has been detected using either IGLS or IRR. If it had not been for Pryor's astute observations, he might have been forced to abandon a relationship of substantial importance in his theorizing.

In an earlier paper (White et al. 1981) we propose and explore a model of the causes of the sexual division of labor in African agriculture. One of the relationships that we tested is total female agricultural participation as a function of crop type (C: 1 = root, 2 = cereal) and slavery (S: 1 = absent, 2 = incipient, 3 = present). The total-female-involvement-in-agriculture dependent variable, labeled  $T$ , is obtained as the sum of three five-point scales indicating female contribution to harvesting, soil preparation, and crop tending, and ranges from a low of 3 to a high of 15.

The sample used in our previous study consists of the 31 African societies from the Standard Cross-Cultural Sample (Murdock and White 1969) that have some degree of agriculture and for which we have data on crop type and slavery. To measure interdependence among these societies we employed both geographical distance and linguistic relatedness criteria. That is, we constructed two  $W$  connectivity matrices, one based on an exponentially decreasing function of distance and the other on a tree of historical relations among all 31 languages.<sup>11</sup> (Details on the construction of these matrices and the variables used, as well as the theoretical rationale for the model, are available in White et al. 1981.)

Table 7 shows the results of our reanalyses. The OLS and ML results for both  $W$  matrices are taken from White et al. (1981). We include here the parameter estimates obtained using the above-described IGLS and IRR procedures.<sup>12</sup> For the language  $W$  matrix, all three network disturbances procedures show very similar magnitudes for the regression coefficients and their standard deviations. These estimates are quite different from the OLS estimates, although there are no reversals in inferring significance of the independent variables: each regression coefficient is still more than twice its standard deviation. The disturbances parameter estimate  $\hat{\rho}$  is significant for all three procedures, although the estimate is higher for IGLS and IRR. Both of these latter procedures show  $R^2$ 's almost identical to OLS and ML.<sup>13</sup> For the distance  $W$  matrix, the estimates for all three autocorrelation procedures are very similar. The  $\hat{\rho}$ 's are again all significant, though less so than with the language matrix. The  $R^2$ 's are very slightly lower for IGLS and IRR than for OLS or ML.

## conclusions

The formal representation of the network autocorrelated disturbances model in matrix terms, as in equations 2 and 3, offers a quite natural characterization of Galton's Problem.  $W$  relational matrices permit rich expression to the idea of cultural diffusion, and

Table 7. Results of all estimation procedures for the total-female-participation-in-agriculture example.

<i>OLS</i>	$T = 24.4$ (2.7)	$-4.76C$ (1.22)	$-2.64S$ (.73)	$R^2 = .47$	
<i>Language W matrix</i>					
<i>ML</i>	$T^* = 19.2$ (2.9)	$-3.19C^*$ (1.07)	$-1.86S^*$ (.61)	$R^2 = .45$	$\hat{\rho} = .69$ (.18)
<i>IGLS</i>	$T^* = 18.5$ (3.5)	$-3.12C^*$ (1.09)	$-1.71S^*$ (.62)	$R^2 = .45$	$p < .001$ $\hat{\rho} = .81$ $Z = 3.67$
<i>IRR</i>	$T^* = 18.6$ (3.5)	$-3.12C^*$ (1.09)	$-1.71S^*$ (.63)	$R^2 = .45$	$p < .001$ $\hat{\rho} = .81$ $Z = 3.67$ $p < .001$
<i>Distance W matrix</i>					
<i>ML</i>	$T^* = 21.8$ (2.8)	$-4.44C^*$ (1.33)	$-1.78S^*$ (.66)	$R^2 = .45$	$\hat{\rho} = .49$ (.20)
<i>IGLS</i>	$T^* = 21.6$ (3.4)	$-4.43C^*$ (1.45)	$-1.47S^*$ (.67)	$R^2 = .45$	$p < .05$ $\hat{\rho} = .76$ $Z = 2.56$
<i>IRR</i>	$T^* = 21.6$ (3.3)	$-4.43C^*$ (1.45)	$-1.48S^*$ (.67)	$R^2 = .43$	$p < .02$ $\hat{\rho} = .76$ $Z = 2.56$ $p < .02$

hypotheses concerning diffusional effects may be investigated by constructing  $W$  matrices based on various theoretical considerations. The  $W$  matrices based on measures of language similarity and geographical distance used in the cross-cultural examples reported here, for example, correspond quite closely to theoretical notions concerning the mechanisms underlying cultural diffusion processes.

Since the network disturbance model is a true generalization of the OLS multiple regression model, the same kinds of complex hypotheses testing and data analyses that are possible with the latter are possible with the former. Even where observations are interdependent, hypotheses concerning the effects of an independent variable on a dependent variable while holding constant one or more additional independent variables can be tested. More complex analyses of direct and indirect effects, such as in path analysis or structural equation modeling, are now possible using cross-cultural survey data. Replication of such complex models across major world geographical/cultural regions, where the most important regional autocorrelation effects are specified within the model, is an important step toward formulating models of great generality.

The results of the simulation study clearly show that all three autocorrelation procedures completely dominate OLS regression whenever autocorrelation effects are moderate to high ( $\rho \geq .5$ ). Thus, one of the three autocorrelation procedures is clearly preferable to OLS at these levels of autocorrelation. For the simulation data analyzed here, there is surprisingly little to distinguish between the more formally correct ML autocorrelation procedure and the two newly proposed procedures in terms of bias and efficiency of the regression coefficient estimates. From the standpoint of computational simplicity, however, either IGLS or IRR is clearly preferable to ML.

Neither our simulation results nor our empirical reanalyses allow us to draw any firm conclusions concerning the relative merits of IRR and IGLS. It is possible that for larger sample sizes one of the two procedures will come to dominate the other in terms of the MSE of the regression coefficients or the  $\rho$ . For estimating a more general model with two or more relational matrices embedded in the error structure, residual regression seems to be the more natural approach to estimation.

Our confidence in both of the new and computationally simpler autocorrelation procedures as accurate and reliable solutions to Galton's Problem is considerably strengthened by the results of our empirical reanalyses. In the reanalyses of our own previous example, both IGLS and IRR generate estimates that are trivially different from our previously reported ML results. Also, in the reanalysis of Pryor's examples both procedures uncover a substantively important relationship that does not appear using conventional OLS routines, and which may have been missed originally but for Pryor's astute observations. While it is obvious that no technique will replace the investigator's understanding and grasp of a problem, the task of theory building and testing using interdependent cross-cultural data may be made substantially easier if appropriate techniques, such as IGLS or IRR, are employed.

## notes

*Acknowledgments.* This research was supported by a grant from the National Science Foundation to Michael Burton and Douglas White.

<sup>1</sup> The general matrix formulation of Galton's Problem presented and discussed in this paper is an extension of the usual multiple regression analysis. Hence, we assume that regression analysis is appropriate to the problem at hand.

<sup>2</sup> Although the matrix formulation employed in this paper introduces some difficulties in that it requires some familiarity with elementary matrix operations, it leads to the most natural expression of the more general network autocorrelated disturbances model. To follow the matrix algebra formulation of OLS and the network autocorrelation disturbances model presented here requires some knowledge of addition, subtraction, and multiplication of vectors and matrices, and of matrix inverses. An introduction to these operations in the context of the multiple regression model can be found in Hilton (1976); more advanced treatments are given in Johnston (1972) and Pedhazur (1982).

<sup>3</sup>  $\varepsilon \sim \text{IND}(0, \sigma_\varepsilon^2 I)$  is a compressed statement of several assumptions made about the error terms in the OLS regression model. This expression says that the errors are independent, normally distributed (IND) with zero mean and no autocorrelation or heteroscedasticity ( $0, \sigma_\varepsilon^2 I$ ).

<sup>4</sup> For discussion and examples of the  $\alpha$ -coefficient and a language similarity measure suitable for cross-cultural research, see White et al. (1981).

<sup>5</sup> The  $\ln$  terms in expression 12 are natural (i.e., base  $e$ ) logarithms.

<sup>6</sup> A coarse first search of the interval was conducted using steps of .1 to locate a minimum. A second search of the  $\pm .1$  interval around this minimum in steps of .01 was then carried out to locate the final minimum  $\hat{\beta}$ .

<sup>7</sup> Associated with any  $N \times N$  square relational matrix  $W$  are  $N$  numbers, not necessarily distinct, called its eigenvalues. Special computational routines are required to calculate the eigenvalues for all but extremely small  $W$  matrices. An extended discussion of eigenvalues, their properties, and algorithms to calculate them is given in Green (1976). Because of the limitations of our current computer facility, the maximum sample size our ML autocorrelation procedures can handle is about 60.

<sup>8</sup> We stopped when successive values of  $\hat{\beta}$  were within an absolute value of .01.

<sup>9</sup> The  $I$ -statistic can also be used to test for autocorrelation of any individual variable in the equation. Details on the computation of the  $I$ -statistic and its standard error for significance testing are given in Cliff and Ord (1973) and Doreian (1980). For an interesting biological example of the use of the  $I$ -statistic to investigate diffusion of human alleles, throughout Europe, with the spread of farming during the Neolithic, see Sokal and Menozzi (1982).

<sup>10</sup> The "effects" model hypothesizes that only the dependent variable is autocorrelated with respect to a given  $W$  matrix. This model is specified as follows:

$$Y = \rho WY + X\beta + \varepsilon \quad \varepsilon \sim \text{NID}(0, \sigma_\varepsilon^2 I)$$

See Doreian (1980) and Erbring and Young (1979) for discussion of this model.

<sup>11</sup> In this paper we estimate the network disturbances model separately for the language matrix and for the distance matrix. The network disturbances model can be generalized to include both matrices simultaneously as follows:

$$Y = X\beta + \varepsilon \\ \varepsilon = \rho_1 W_1 \varepsilon + \rho_2 W_2 \varepsilon + v \quad v \sim \text{IND}(0, \sigma_v^2 I)$$

For an empirical example of this more general model applied to cross-cultural data, see Dow (1984).



<sup>12</sup> All of the computations reported in this paper for the IRR and IGLS procedures were carried out using BASIC language programs written by Reitz and Dow. Listings of these programs are available from the authors upon request.

<sup>13</sup> The  $R^2$  reported in our previous paper (White et al. 1981) are incorrect for the ML procedures, using both the language and distances  $W$  matrices. Instead of the .67 and .61 reported, the correct  $R^2$  are .45 and .45, respectively, as shown here. Again, the  $R^2$  reported here are simply the squared Pearson product-moment correlation between the original dependent variable scores and the predicted scores (i.e.,  $r^2_{y\hat{y}}$ ).

## references cited

Bartels, Cornelius, and Ronald Ketellapper, eds.

1979 *Exploratory and Explanatory Statistical Analysis of Spatial Data*. The Hague: Martinus Nijhoff.

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.

Brandtsma, Andries, and Ronald Ketellapper

1979 A Biparametric Approach to Spatial Autocorrelation. *Environment and Planning* 11:51-58.

Cliff, Andrew, and Keith Ord

1973 *Spatial Autocorrelation*. London: Pion.

1981 *Spatial Processes: Models and Applications*. London: Pion.

Doreian, Patrick

1980 Linear Models with Spatially Distributed Data: Spatial Disturbances or Spatial Effects? *Sociological Methods and Research* 9:29-60.

1981 Estimating Linear Models with Spatially Distributed Data. *In Sociological Methodology*

1981. S. Lienhardt, ed. pp. 359-388. San Francisco: Jossey-Bass.

Dow, Malcolm M.

in press A Biparametric Approach to Network Autocorrelation: Galton's Problem. *Sociological Methods and Research*.

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

1982 Network Autocorrelation: A Simulation Study of a Foundational Problem in Regression and Survey Research. *Social Networks* 4:169-200.

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

1983 Multivariate Modeling with Interdependent Network Data. *Behavior Science Research* 17: 216-245.

Durbin, J.

1960 Estimation of Parameters in Time-Series Regression Models. *Journal of the Royal Statistical Society* 22B:139-153.

Erbring, Lutz, and Alice Young

1979 Individuals and Social Structure: Contextual Effects as Endogenous Feedback. *Sociological Methods and Research* 7:396-430.

Gatrell, Antony

1979 Autocorrelation in Spaces. *Environment and Planning* 11:507-516.

Green, Paul E.

1976 *Mathematical Tools for Applied Multivariate Analysis*. New York: Academic Press.

Hepple, Leslie

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.

Hilton, Gordon

1976 *Intermediate Politometrics*. New York: Columbia University Press.

Hoel, Paul

1971 *Introduction to Mathematical Statistics*. New York: Wiley.

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.

Murdock, George P., and Douglas R. White

1969 Standard Cross-Cultural Sample. *Ethnology* 8:329-369.

Naroll, Raoul

1961 Two Solutions to Galton's Problem. *Philosophy of Science* 28:15-39.

1973 Galton's Problem. *In Handbook of Methods in Cultural Anthropology*. R. Naroll and R. Cohen, eds. pp. 974-989. New York: Columbia University Press.

- 1976 Galton's Problem and HRAFLIB. *Behavior Science Research* 2:123-148.
- Ord, Keith  
 1975 Estimation Methods for Models of Spatial Interaction. *Journal of the American Statistical Association* 70:120-126.
- Pedhazur, Elazar J.  
 1982 Multiple Regression in Behavioral Research. New York: Holt, Rinehart & Winston.
- Pryor, Frederic  
 1976 The Diffusion Possibility Method: A More General and Simpler Solution to Galton's Problem. *American Ethnologist* 4:731-749.
- Sokal, Robert, and Paolo Menozzi  
 1982 Spatial Autocorrelations of HLA Frequencies in Europe Support Demic Diffusion of Early Farmers. *The American Naturalist* 119:1-17.
- Schaefer, James, ed.  
 1974 Studies in Cultural Diffusion: Galton's Problem. New Haven, CT: HRAFlex Books.
- Simonton, Dean Keith  
 1975 Galton's Problem, Autocorrelation, and Diffusion Coefficients. *Behavior Science Research* 10:239-248.
- Tylor, Edward B.  
 1889 On a Method of Investigating the Development of Institutions Applied to the Laws of Marriage and Descent. *Journal of the Royal Anthropological Institute* 18:272ff.
- White, Douglas R., Michael L. Burton, and Malcolm M. Dow  
 1981 Sexual Division of Labor in African Agriculture: A Network Autocorrelation Analysis. *American Anthropologist* 83:824-849.
- Wirsing, Rolf  
 1974 Measuring Diffusion: The Geary Method and the Dacey Method. In *Studies in Cultural Diffusion: Galton's Problem*. James M. Schaefer, ed. pp. 198-219. New Haven, CT: HRAF Press.

Submitted 30 November 1982

Revised version received 19 August 1983

Accepted 13 October 1983