Title
Sexual Division of Labor in African Agriculture: A Network Autocorrelation Analysis

Journal
American Anthropologist, 83(4)

ISSN
0002-7294

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

Publication Date
1981-12-01

DOI
10.1525/aa.1981.83.4.02a00040

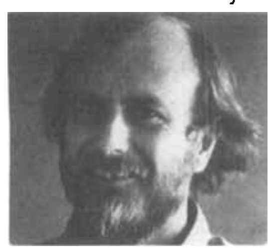
Sexual Division of Labor in African Agriculture: A Network Autocorrelation Analysis

DOUGLAS R. WHITE
MICHAEL L. BURTON*
University of California, Irvine

MALCOLM M. DOW
Northwestern University

A model of causes and consequences of sexual division of labor in agriculture is tested using a sample of African societies. Crop type and the presence or absence of slavery are shown to be effective predictors of the degree of female contribution to agricultural subsistence, and the degree of polygyny is shown to be affected by female agricultural contribution and the form of residence. Autocorrelation effects are found and are shown to be a consequence of Bantu societies having higher female participation in agriculture than would otherwise be expected. This effect is an example of one of the kinds of phenomena that anthropologists have referred to as Galton's problem. [sexual division of labor, cultural ecology, Galton's Problem, Africa]

IN ANTHROPOLOGICAL THEORY, variations in the sexual division of labor have often been seen as causes of variations in residence patterns, marriage practices, beliefs about gender, socialization patterns, and many other aspects of human behavior and belief. For example, it has been hypothesized that matrilocality (Driver 1956), polygyny (Gillin 1948; Heath 1958; Goody 1977) and altruistic behavior by children (Whiting and Whiting 1975) are consequences of high female contributions to subsistence activity. Schlegel and Barry (1980) find a correlation between high female contribution to subsistence and the presence of adolescent initiation ceremonies for either sex. Given the broad range of these hypotheses, it is important to understand the causes of variability in



MICHAEL BURTON is Associate Professor of Anthropology in the School of Social Sciences at the University of California, Irvine, where he has taught since 1969. He received a Ph.D. degree from Stanford University in 1968. His research interests include cognitive anthropology, the organization of work, sex roles, and cross-cultural methodology. He did fieldwork in Kenya during 1973-74 with the Kikuyu and Maasai, and has published on the sexual division of labor, occupational cognition, the Maasai age set system, and quantitative methodology.

DOUGLAS R. WHITE is Professor of Anthropology and Comparative Culture at the University of California, Irvine. He was co-principal investigator with G. P. Murdock for the Cross-Cultural Cumulative Coding Center at the University of Pittsburgh, where he taught from 1967 to 1976. He received his Ph.D. from the University of Minnesota in 1967. His publications on the theory of social structure range from Mexican and North American Indian comparative and case studies to previous cross-cultural studies on the division of labor, and to more general mathematical models and methods. He is currently publishing a network theory of role, status, and social structure using various case studies.

MALCOLM M. DOW received his B.A. (Mathematics) and Ph.D. (Anthropology) from the University of California, Irvine. He held an NIH Postdoctoral Fellowship in Epidemiology at UCLA in 1979-80 and is now Assistant Professor of Anthropology at Northwestern University. His current research interests include mathematical anthropology, cultural ecology, sociocultural epidemiology, demographic anthropology, and social networks.

0002-7294/81/040824-26$3.10/1

the sexual division of labor. The present paper examines possible causes of variation in the sexual division of labor in African agriculture, and tests a hypothesis about one consequence of that variation, relating increased female subsistence contributions to increases in the degree of polygyny. The method of analysis involves a new solution to Galton's problem, network autocorrelation analysis, that incorporates regional interactions among societies due to diffusion or to historical interconnections, as well as functional relationships. The functional relationships are expected, if valid, to replicate from one region to another, while the historical and diffusional effects are not. The present paper is the first of several regional studies. The sample for this study consists of the 36 African societies from the Standard Cross-Cultural Sample (Murdock and White 1969) that have some degree of agriculture.

In Africa, most cultivation of crops is by horticultural methods, with the plow being present only in regions of direct contact with the societies of the Circum-Mediterranean area. We include the entire African continent in our sample, unlike Murdock and White (1969), who include much of Northern Africa in the Circum-Mediterranean culture area, hence we do have some societies in the present sample that practice intensive agriculture with plows. We shall use the term "agriculture" in its most extended sense to refer to any crop cultivation, including both horticulture and intensive agriculture.

As Lancaster (1976) and others have noted, female contribution to agriculture is higher in Africa than in most regions of the world. This fact makes Africa an interesting region within which to examine variation in the division of labor, since the African continent includes societies with the highest rates of female participation in subsistence of all world societies. Africa is also interesting because of the absence of the plow and of prescriptive monogamy from most of the continent. These two variables are highly linked in much of the Old World (Burton and Reitz 1981) and their absence from most of Africa makes it possible to test models without having to be concerned with the confounding influence of those two variables.

Although rates of female participation in African agriculture are high, on the average, it would be a mistake to assume that African cultivation is a unitary phenomenon and that there is no variation to explain. There is a great diversity of cultivation practices within Africa, and also a considerable variation in the division of labor in agriculture.

POSSIBLE CAUSES OF THE SEXUAL DIVISION OF LABOR

A previous study of the sexual division of labor emphasized constraints on the allocation of tasks (Burton, Brudner, and White 1977; White, Burton, and Brudner 1977), and found strong cross-societal regularities in task allocations across production sequences. The regularities take the form of implicational sequences. In these sequences if women do a particular task within a given society, then we can infer that women also do each of some other set of tasks. For example, if women do soil preparation, then they also do planting, crop tending, and harvesting. We found several such sequences for which the regularities pertained across the standard cross-cultural sample of 186 societies, with very few exceptions. The model adduced to explain those relationships was based on universal constraints on women's productive activity, due to childbirth and nursing. The implications produce a partial ordering of tasks within any given production sequence from tasks that are most easily performed by women to tasks that in some or all societies are more difficult for women to do, given the constraints. The latter tasks require long absences from home, or would be particularly dangerous to women or infants, and often focus on the acquisition of raw materials, rather than on secondary processing. Hence, men in preindustrial societies tend to dominate the beginning stages of production sequences, and women's activities tend to be focused in the later stages of the production sequences, in secondary processing activities.

Although the constraint model has a very good fit to the data, it does not account for all of the observed variance in sexual division of labor. Rather, it accounts for the constraints on that variation. Take, for example, the relationship between soil preparation and crop tending. Soil preparation is much more likely to be done by men than is crop tending, and the constraint model says that there will not be societies where women do the soil preparation while men do the crop tending. There are, however, three possibilities that are not excluded by the constraint model, with many empirical examples of each. These are: (1) women do both tasks; (2) men do both tasks; and (3) men do soil preparation while women do crop tending. Nor does the constraint model exclude the possibility of both tasks being shared equally. Hence, although the constraint model limits the number of possibilities, there is still considerable variance to explain.

Societal complexity has often been mentioned as a cause of variation in the sexual division of labor. Murdock and Provost (1973) discuss several examples of the effect of societal complexity upon the sexual division of labor, most notably upon housebuilding and several crafts. These tasks become occupational specialties in complex societies with the consequence that they are done by men rather than by women. The recent literature on women's work has produced many examples of a similar process, in which economic development appears to cause women to lose their economic position with respect to such tasks as farming and marketing (Romalis 1979). With respect to agriculture, several scholars have noted the strong relationship between the presence of the plow and male involvement in agriculture. Boserup (1965, 1970) proposes population pressure as a primary cause of changes from long fallow to short fallow agriculture, and notes the statistical association between short fallow agriculture, predominantly male contributions to subsistence, and monogamy. Her argument concerning population pressure has been controversial (Bronson 1972; Barlett 1976; Cowgill 1975) but her description of the consequences of agricultural intensification has been generally accepted. She does not, however, propose actual mechanisms that would explain the negative association between intensive agriculture and the use of predominantly male agricultural labor, and Burton and Reitz (1981) provide evidence that the observed relationship between the plow and high male contributions to subsistence is an artifact of Galton's problem. The two traits appear to have diffused together throughout much of Europe and Asia with no causal relationship pertaining between them.

The present paper will focus on the organization of production, with emphasis on two aspects, the material constraints of the ecological system, and social relations of production. Material constraints pertain to the ways in which the material circumstances of production interact with constraints upon women's activities. For example, with respect to hunting, it has been noted that women often participate in hunts where large groups of people surround animals and drive them over cliffs, but that they do not participate in long-distance hunting expeditions. In the former case, the kill is done relatively near to the base camp, and extensive travel is not required. In the latter case, extensive travel under dangerous conditions may be required. Similar reasoning may apply with respect to the circumstances of agricultural production, which may vary greatly across agricultural systems. Under the rubric of agricultural production we must include not only production in the field, but also secondary processing of the food in preparation for storage or cooking. Martin and Voorhies (1975) take the latter approach in their explanation for the fact that women participate more in root crop agriculture than in cereal crop agriculture. They say that cereal crops require more secondary processing (winnowing, grinding, etc.) than do root crops, and that the high involvement of women in the secondary processing of cereal crops prevents a high involvement in production in the field.

Relations of production include the presence or absence of wage labor markets, of conscription and corvee labor, or slavery, as well as considerations of the economics of the family. Given the constraints on women's activity discussed above, the division of labor is best seen as the consequence of decisions about optimization of labor input to productive activity. Women's allocation of labor to a given activity, then, will depend upon (1) the constraints placed upon this allocation by other activities in which women are engaged; (2) the constraints of the specific task; (3) the relative payoff to women of the particular task as opposed to other tasks; and (4) the relative availability of other categories of people (men, children, or slaves) to do the task. If men are away for extensive periods of time due to wage labor, warfare, long-distance hunting or fishing trips, or transhumant pastoralism, then women's participation in agriculture may increase. Conversely, if slaves are available, women's participation in agriculture may decrease.

In the present paper we will consider one variable pertaining to material circumstances and one variable pertaining to relations of production. In the first case, we will examine the contrast between societies mainly dependent upon root crops and societies mainly dependent upon cereal crops. With respect to relations of production, we shall examine the effect of slavery upon the sexual division of labor.

We have mentioned Martin and Voorhies's assertion that cereal crops require more secondary processing than root crops, thereby reducing the participation of women in primary production. Another important difference between root crops and cereal crops can be seen in the temporal organization of harvesting. Root crops can typically be harvested at a more leisurely pace than cereal crops because the roots can be left underground, where they are protected from the elements, from plant diseases, and from pests. Because of the time pressure involved, cereal crop harvesting is more likely than root crop harvesting to require the mobilization of a high percentage of the able people in the community, often involving most adults of both sexes and many of the children, whereas root crop harvesting is more likely to be done by women alone. This difference in time pressure appears to be exaggerated, at least in Africa, by the spatial distributions of the two kinds of crops. Cereal crops tend to be grown in areas that have sharp alternations of rainy seasons and dry seasons, such as the Sahel region, and the East African Highlands, whereas root crops tend to be grown in tropical lowland areas that have less seasonal variation in climate. This may be due to the fact that grasses place excessive nutrient demands upon tropical rain forest environments due to their quick growth (Meggers 1971; Geertz 1963), or because it is harder to store cereal grains than roots in humid climates, or because cereals do not fare well under extremely heavy rainfall. The tropical lowland societies tend to use slash-and-burn techniques, with intercropping, in which crops are planted and harvested at staggered intervals. By contrast with more intensive systems, slash-and-burn agriculture requires high labor inputs to crop tending, especially weeding, relative to the inputs in tasks such as soil preparation, and the crop tending is often done by women. Agricultural regimes that require medium labor inputs with no periods of peak demand are more compatible with the child-rearing constraints on women's activity than are regimes which require sudden mobilization of large numbers of people. Tropical root crop systems fit this description very well. Furthermore, in many such systems men are involved in hunting, warfare, or both, and are less likely than women to be available to participate in cultivation. Hence, the relative benefits to family units from female participation in agriculture are higher than in more intensive systems, where hunting is no longer possible, and warfare has been taken over by the state. As Netting (1968) notes, extensive agricultural systems in Africa are characterized by predominantly female participation, whereas in more intensive systems men tend to participate more in agriculture, as a result of the increased labor demands of intensive agriculture.

As a test of these assertions about the ecological differences between root crop agriculture and cereal crop agriculture, we have computed contingency tables on the entire standard cross-cultural sample (Murdock and White 1969) showing the relationship

TABLE I. RELATIONSHIP BETWEEN CROP TYPE AND INTENSITY OF CULTIVATION.

	Roots	Cereals
Horticulture	10	1
Shifting cultivation	18	34
Intensive cultivation	5	50

of crop type to agricultural intensity, the length of the dry season, and the presence of domesticated animals. In Table I we see that root crops are concentrated in horticultural and shifting agricultural systems, whereas cereal crops are found predominantly in intensive systems (gamma = .79).

In Table II we see that root crops are not grown in environments with long dry seasons (gamma = .84) and in Table III we see a strong relationship between cereal crops and the relative importance of animal husbandry (gamma = .79). Differences in crop type, then, are dramatically related to variations on a set of important ecological variables.

Boserup (1970) notes that female participation in production may decrease in stratified societies, where lower-status people (slaves, serfs, landless laborers) are available to do the agricultural labor. In those circumstances, high-status women will not participate in agricultural labor. The overall effect will be to reduce the total amount of female participation in agriculture. We test this hypothesis using slavery as a measure of social stratification because it is the single stratification variable that pertains specifically to the organization of economic production. Hence it should be the most accurate measure of the presence of stratified relations of production. Other stratification variables, such as the presence of superordinate levels of political administration, may have only a very tenuous relationship, if any, to actual relations of production.

In addition to testing for the effects of crop type and slavery on the sexual division of labor, we will consider the Boserup (1965, 1970) hypothesis that the sexual division of labor is affected mainly by population density.

Our hypotheses have antecedents in the literature on African horticulture. Strong parallels can be seen as early as Baumann (1928). Baumann was a member of the German *kulturkreis* or "Zones of Culture" school which emphasized evolutionary accounts combined with both functional and diffusionist hypotheses. His evolutionary arguments reflect such 19th-century concerns as the supposed transition from matriarchy to patriarchy. He hypothesizes that agriculture was invented by women as a consequence of their gathering activities, and that "plough culture banishes the woman from the actual cultivation of the fields to a little garden in front of the house, where, as at the beginning of her agricultural activity, she grows a few vegetables and roots" (1928:291). Putting aside his projective evolutionary reconstructions, and his generally colonial and deprecatory attitudes towards native women, there are remarkable parallels between his functional and diffusionist observations and the general elements of our model. We regard these parallels as a source for convergent validation of our model, in that his observations are based on a statistical analysis of ethnographic reports from more than 200 societies in Sub-Saharan Africa.

One parallel is that Baumann views climatic and environmental factors as closely associated with differences in the participation of the sexes in agriculture. He observes

TABLE II. RELATIONSHIP BETWEEN CROP TYPE AND LENGTH OF DRY SEASON.

	Roots	Cereals
1-2 months	7	6
3-5 months	8	24
6-8 months	1	25
9-12 months	0	11

TABLE III. RELATIONSHIP BETWEEN CROP TYPE AND IMPORTANCE OF ANIMAL HUSBANDRY.

Animal husbandry	Roots	Cereals
0-15%	14	16
16-35%	3	42
36-55%	1	7
over 55%	0	7

that the northern boundary of women's predominance in agriculture is constrained by the limits of the tropical forest region, and that the boundary of female agriculture also tends to coincide with that between root crops and cereal grains.

Baumann's view of the economies of labor is also similar to our own. He emphasizes that land clearing is more difficult in the forest than in the savannah, and that males often perform clearing in the forest zones in spite of the predominance of female labor, whereas soil preparation is more difficult in the savannah than in the forest. He notes the higher male participation in agriculture in the savannah region of the Sudan than in the West African and Congo forest regions, and more generally, that women are much more likely to participate in root cultivation than in cereal cultivation. He clearly discusses the year-round nature of crop tending:

> "In the hot zone the destruction of weeds is the farmer's most important task (Kärger 1889:44)." It is just from these tropical regions, such as the Congo and Melanesia, that the accounts of the never-ending, exceedingly heavy work of the women in the fields and of the men's extreme idleness, are too numerous and weighty to be disregarded. Women's work in hoe culture is in fact continuous, and extends over the whole agricultural year; the men's share, namely the clearing, lasts a limited time and forms no part of the actual process of hoe culture (1928:294).

Baumann also notes an association between domesticated animals, predominance of male labor in husbandry, slavery, and the use of slaves in agriculture.

POSSIBLE CONSEQUENCES OF THE SEXUAL DIVISION OF LABOR

Residence

There are two main theories of the determinants of postmarital residence. The first is that the sexual division of labor in subsistence activities may favor cooperative male groups, cooperative female groups, mixed cooperative groups, or independent homesteads, favoring in turn virilocality, uxorilocality, bilocality, or neolocality. There is considerable variability across regions in the empirical relationship between residence and the sexual division of labor in subsistence activities. We find a worldwide Pearson's correlation of + .21 between patrilocal residence and male participation in agriculture. For the Americas this correlation is even stronger, being + .29. Driver (1956) found an equally strong relationship between residence and contribution to subsistence within North America. For the African continent, however, the same correlation is nearly zero (− .08). The hypothesized relationship between sexual division of labor and residence does not seem to hold for Africa. The second main theory of residential determinants argues that defensive activities such as warfare and feuding are primary determinants of residential groups. Divale (1974) argues that groups which have migrated into a region where they are surrounded by hostile groups (e.g., in the past 500 years) are likely to develop matrilocal residence as a means of distributing descent-related males across local communities, thus serving to integrate the entire society for defensive purposes. Ember and Ember (1971) argue that local warfare or feuding stresses solidarity of male warriors and thus favors patrilocal groups, while distant warfare stresses female community solidarity and thus matrilocality.

Obviously, both subsistence and defense are fundamental aspects of the division of labor, and residential arrangements seem to be a very likely consequence of the overall division of labor, including both subsistence and warfare. It is not simply the contribution of the sexes that constitutes the determinant, if any, of residence patterns, but the relative importance and structure of the task in terms of spatial and community organization.

Polygyny and Marriage Systems

Murdock (1949:36), Heath (1958) and others have argued that the degree of female contribution to subsistence is a predictor of the degree of polygyny. This seems to be particularly likely in agricultural societies, as opposed to pastoral societies which may have high polygyny but low female contribution to subsistence. In the latter type of society, wives are economically dependent, but produce the personnel (children) through which cooperative work groups and marriage alliances are maintained.

Polygyny is clearly affected by factors other than female workload, including possibly warfare and adult sex ratios (Ember 1974). Nonsororal polygyny is also constrained by matrilocality (Murdock 1949:31). We would expect lower rates of nonsororal polygyny in matrilocal societies than in patrilocal societies. Most of the polygynous societies in our African sample practice nonsororal polygyny; only two are coded as practicing sororal polygyny.

CAUSES AND CONSEQUENCES OF THE DIVISION OF LABOR: A MULTIVARIATE MODEL

The preceding theoretical considerations can be combined into a multivariate model of the causes and consequences of the sexual division of labor in agriculture. The present model will be specific to Africa. Variables which may be important in explaining worldwide variations in sexual division of labor, such as climate and the presence of the plow, will not be included in the African model. These are factors which are relative constants for the region: tropical climate and the absence of the plow, for example, characterize most of the preindustrial African societies. Nor will we include bilateral inheritance, which Goody (under the rubric "diverging devolution") considers important to explaining monogamy (1977). Neither bilateral inheritance nor monogamy occurs with high frequency in Africa. Our multivariate model is an example of the kind of comparative regional study advocated by Jorgensen (1979), for which the goal is to replicate across regions models formulated within specific cultural areas.

The model for Africa is found in Figure 1. We see cereal crop agriculture and slavery as leading to decreased female participation in agriculture. Following Heath (1958) and others, we predict that high female participation in agriculture will lead to higher amounts of polygyny. Patrilocal residence facilitates nonsororal polygyny, and will have an effect on polygyny that is independent from the effect of the agricultural division of labor. Finally, as we noted above, there is a zero correlation within Africa between female participation in agriculture and marital residence, a relationship which could be different in a study of some other region, such as North America.

AUTOCORRELATION: ECOLOGICAL, DIFFUSIONAL, AND GENETIC EFFECTS

The oldest fundamental criticism of cross-cultural studies is that measures of trait interrelationships are problematic because observations are not independent from one case to the next. If traits spread through diffusion or multiply through the historical fission of societies, one does not know how many independent instances one is observing. Measures of association based on the variance accounted for may be seriously inflated if the ob-

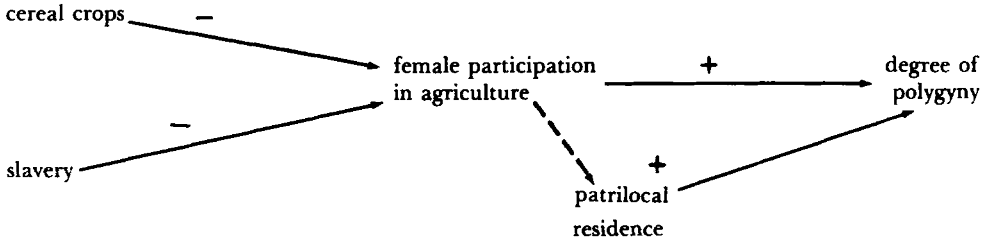


Fig. 1. Model of relationships among sexual division of labor, societal complexity, and marriage.

served cross-societal variance is lowered by diffusional effects. This objection has come to be known as Galton's problem (Naroll 1970a; Schaefer 1974).

Previous attempts to justify causal or functional inferences from cross-cultural data have taken one of three approaches. Murdock (1949:117) argues that phenomena which exhibit the greatest incidence of parallel independent invention are the most suitable for comparative analysis. This is a valid observation, but one that limits comparative research to a small and perhaps even nonexistent set of problems, where spatial clustering of traits as evidence of nonindependence is absent. Another strategy for testing purely functional hypotheses while discounting historical ones has been to draw a sample of genuinely independent societies, or screen the sample to eliminate nonindependent occurrences. As Murdock and White (1969) show, this can result in miniscule samples (e.g., $N = 15$) such that multivariate analyses are difficult or impossible, and there is still a question of whether historical influences have truly been removed. Other scholars, such as Driver (1956) and Naroll (1970b), have sought to remove the restriction on both problem area and sample construction by simultaneously testing both diffusional-historical hypotheses and functional hypotheses, and attempting to separate these different types of effects. One of the strategies developed by Naroll has been to draw a sample, define a relationship between societies which might be important for diffusion (e.g., spatial propinquity) or genetic relationship (e.g., language affinity), measure the degree of spatial or linguistic clustering as a test of the historical hypotheses, and then try to correct the intertrait correlations for the effects of historical connectedness, leaving as a residual an estimate of the true "functional" correlation.

Cross-cultural anthropologists have recently recognized that Galton's problem can be seen as an example of spatial autocorrelation (Loftin 1972, 1975; Loftin and Hill 1974; Naroll 1976; Simonton 1975). Autocorrelation refers to the statistical relatedness of units that are connected to or proximal to each other. The original econometric formulation of autocorrelation was in terms of temporal autocorrelation, in which observations of a variable for a given time period are highly correlated to observations of the same variable for the time period immediately preceding or following. In spatial autocorrelation, observations at some point in space are correlated to observations at nearby points. Statistical techniques for temporal and spatial autocorrelation analysis have been developed by geographers and economists (Cliff and Ord 1973, 1975; Hibbs 1974; Ord 1975). Given known patterns of spatial and/or temporal distribution of a variable, a particular society's score on this variable can be predicted from its position in the spatial or temporal series, as a function of the characteristics of related units. Subtracting a society's predicted score due to autocorrelation from its actual score gives a set of first differences which are not subject to Galton's objection.

Rather than testing causal predictions where a societal variable Y is a function of X, as in the classic regression model,

$$Y = A + \Sigma B_i X_i + U$$

where A and each B_i are constants, each X_i is an independent variable and U is a random error term, a model which incorporates autocorrelation might assume that scores of both Y and X_i can be partially predicted on the basis of autocorrelation. Call these predictions $P(Y)$ and $P(X_i)$. A causal model which incorporates autocorrelation then, might view Y as a function of autocorrelation and of the first differences $(X - P(X))$:

$$Y = P(Y) + A + \Sigma(X_i - P(X_i))\, B_i + V$$

where V is a truly independent (nonautocorrelated) random error term. This model is not subject to Galton's objection. It says that there is autocorrelation, predicts the scores of Y and X_i from autocorrelation, subtracts these scores to get the first differences Y^* and Y_i^*, and predicts that these differences are related:

$$Y^* = A + \Sigma X_i^* B_i$$

In this model, the "functional" component of the relation between variables X_i and Y is that between the first differences, X_i^* and Y^*. The "historical" component is that between X_i and $P(X_i)$ or Y and $P(Y)$.

In this paper we will use the regression framework for spatial autocorrelation analysis, but shall extend it to a more general formulation of autocorrelation, in which societies are seen as occupying positions in networks of relations among societies, and in which autocorrelation exists among societies that are close in the networks. Examples of such networks could include trade or market networks, networks of political relationship, or networks of historical linguistic relationships. Spatial autocorrelation analysis is a special case of network autocorrelation analysis in which the relation between societies is measured by spatial distance. The present paper will focus on two measures of relatedness among societies: spatial distance and the historical relatedness of languages.

At this point we must speak to a common objection to the use of the regression analysis in cross-cultural data, which pertains to the level of measurement of much cross-cultural data, stating that regression analysis should only be used with interval scale data. This objection overlooks the high degree of robustness of regression analysis in situations that deviate from the exact formulation of the model. It also understates the amount of interval or ordinal data in cross-cultural codes.[1]

Network autocorrelation analysis provides an estimation procedure for computing regression coefficients that gives unbiased and efficient estimates of the effects of the explanatory variables as well as the effects of autocorrelation, which are not subject to Galton's objection. These estimates can be compared across regions in replication experiments, and the present study is the first in a series of studies in which we hope to find replicable relationships across major world regions. We now turn to some of the details of the estimation procedure.

Autocorrelation is said to exist in the regression model when estimates of the parameters (B's) which minimize the sum of the squared error terms for each observation (ordinary least squares or OLS) produce error terms which are not independent. The error in one observation is correlated with errors in related observations. In the presence of autocorrelation, ordinary least-squares (OLS) procedures still produce unbiased estimates of the parameters, but these estimates are not the most efficient. This means that OLS procedures for computing variances of estimated coefficients are incorrect, and standard tests of significance and correlation are invalid. The consequences of violating the nonautocorrelation assumption are serious: the t-ratios for individual coefficients are generally inflated, leading to spurious attributions of significance to independent variables. Also, since the estimate of the error variance is generally downwardly biased, t-, F-, and R^2- statistics tend to be exaggerated (Hibbs 1974; Dow, White, and Hansen 1979).

In the usual application of regression analysis the sample units are assumed to be independently drawn from the population of interest. This independence is not required by the regression model. Sampling in this manner by no means guarantees that the assumption of independent errors (i.e., nonautocorrelation of errors) is met. Frequently, even the assumption that the cases have been randomly selected is unrealistic. For example, if economic variables are sampled at fixed time intervals, it is highly unlikely that cases and error terms are independent. The observations are clearly ordered in time and since economic variables usually do not change precipitously, each observation is related to other observations close in time. A similar situation is encountered when the sample units are selected from a two-dimensional "spatial series." For example, if a large geographic area is subdivided into regions and data are collected from each region, a plausible alternate hypothesis for the error terms is that each is a simple linear function of the error terms of "nearby" or continguous regions plus a random component. To model this hypothesis requires a proximity matrix which reflects the structure of spatial interdependencies. The proximity matrix is easily constructed as a $(N \times N)$ matrix where the entries measure the relative proximities of the societies. Each element of the proximity matrix is then divided by its row sum to form a new stochastic matrix W. The assumption of autocorrelated errors can be stated in compaçt matrix notation as

$$U = \rho WU + V$$

where V is a truly random term, and p is the "autocorrelation coefficient," constrained to values between -1 and $+1$ so long as the W matrix is stochastic. If p is found to be significantly nonzero, beyond sampling variability, the ordinary-least-squares regression coefficients are, as stated above, unbiased but are not minimum variance. Modified procedures must be employed to restore efficient or "best" estimation.

Whether or not observations are sampled randomly in time and space, autocorrelation may be present in the error terms of a specific regression model, with respect to a "network" matrix, W. Two basic models which hypothesize network interdependence among the sample units are available (Cliff and Ord 1973, 1975; Ord 1975; Doreian 1980, 1981; Dow, White and Burton 1981). Of these, we use the *First-Order Network Autoregressive Model,* formally stated as

$$Y = A + XB + U$$
$$U = \rho WU + V$$

where X is a matrix of independent variables. Here $XB = X_iB_i$. Combining these equations and rearranging terms:

$$Y - \rho WY = A + (X - \rho WX)B + V$$

If new variables are created by subtracting out of the original variables the "systematic" part due only to the position of the sample unit within a hypothesized "network" W, i.e.,

$$Y^* = Y - \rho WY$$
$$X^* = X - \rho WX$$

then the regression equation

$$Y^* = A + X^*B + V$$

will provide unbiased and minimum variance ("best") estimates of the regression coefficients for the population (Martin 1974). This is so because the disturbances (V) of this latter equation are by definition not autocorrelated with respect to the network of relations represented by the W matrix.

For cross-cultural research, we think that the autoregressive or autocorrelated disturbances model will be applicable to sociocultural phenomena, since for any of the models we have proposed, diffusional or linguistic relatedness processes should affect both the independent and dependent variables. We do not have any hypotheses in which we would expect autocorrelation only in the dependent variable and not in the independent variables. If we did, we should use the other main autoregression model, the effects model (Doreian 1980).[2]

COMPUTING THE STOCHASTIC INTERACTION MATRIX, *W*

The *W* matrix is derived from a matrix of proximities among societies. These proximities can be measured from spatial relations, social contact such as trade, or historical connections such as common language. To obtain the *W* matrix from the proximity matrix, the diagonals are set to zero, row sums are computed, and each entry of the proximity matrix is divided by its row sum. W_{ij} then measures the proportion of the total proximity of item i to all other societies that is accounted for by its proximity to society j.

One measure of spatial proximity is the actual geographical distance between societies. This is an appropriate measure for a situation in which no single society occupies a large portion of the geographical plane. If there were a single society, such as China, that occupied a very large area, then it could become difficult to decide which point within the area occupied by the large society to use as the referent point for the distance measurement. Geographical barriers, such as oceans, also can make direct measurement inappropriate. In such cases, geographical distance should be measured by traversing known or likely migration routes, rather than by using the shortest physical distance.

The present paper uses a modified measure of spatial proximity, an exponential decreasing function of actual geographical distance. In preliminary investigations we found that the straight-line distance model gave too much weight to societies that were far removed from each other. By so doing, it obscured the effects of diffusion, which tends to operate over relatively short distances. An exponential decreasing function decreases the weight given to distant societies relative to the weight given to close societies. The actual formula used for the spatial proximity from society i to society j (S_{ij}) is

$$S_{ij} = e^{-(D_{ij}/120)}$$

where D_{ij} = distance between i and j in miles. This formula gives about one-third the weight to a neighbor 300 miles away that it gives to a neighbor 150 miles away. The empirical effect is to assign most of the weight in the *W* matrix to the closest neighbors to any given society within a 200-300 mile radius. We use spatial distance as a proxy for social contact in the form of trade, intermarriage or warfare, and we think most historical contact has been among pairs of societies within this size of radius.

Proximity measures can also be constructed that are based on shared social process or common history. Two kinds of measures of prior social contact are particularly appropriate to cross-cultural studies. The first kind of measure pertains to common linguistic history; the second to common political history, such as incorporation into empires. The present paper will use a linguistic measure.

A straightforward way to measure linguistic history is to measure proximity on the genetic tree of languages as an inverse function of the number of nodes along the path between the two languages. In Figure 2, for example, there is one node between languages A and B, whereas languages A and C are separated by three nodes and languages A and F are separated by five nodes. These distances can easily be transformed into proximities, using an appropriate linear transformation; for example, subtracting each distance from the maximum distance on the tree. Here the proximity measure between A and B would be two, the proximity measure between the C and D would be one and the proximity of A to E, F, G, or H would be zero.

To measure genetic relations among the 43 African languages, we constructed a tree of historical relations among all 43 languages using work by Greenberg (1966) and Guthrie (1967). We then used this simple measure of proximity on the tree to construct the *W* matrix for linguistic relations.[3]

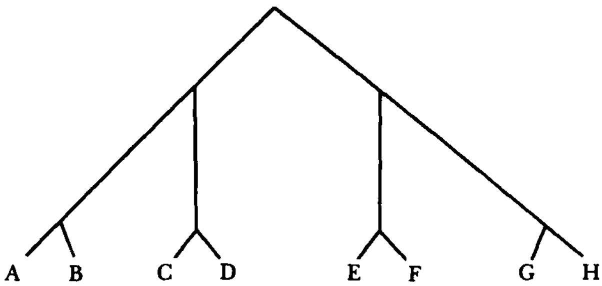


Fig. 2. A hypothetical tree structure.

TESTS OF THE MODEL OF SEXUAL DIVISION OF LABOR

Murdock and Provost's article (1973a) on sexual division of labor provides the codes for the present study. These authors coded the 186 societies of the standard cross-cultural sample (Murdock and White 1969) for the sexual division of labor in 50 activities. Each activity is coded a on five-point scale. For the purpose of regression analysis, we have assigned numerical values to the Murdock and Provost codes as follows:

1: Entirely male *(M)*
2: Predominantly male *(N)*
3: Divided equally between the sexes *(E)*
4: Predominantly female *(G)*
5: Entirely female *(F)*

To test our model of causes of the division of labor in agriculture, we use variables on crop types (Murdock and Morrow 1970) and slavery (Murdock 1967). The *Ethnographic Atlas* has a four-category code for slavery: absence or near absence; incipient; reported but type not identified; and hereditary and socially significant. We merged the last two categories to produce a three-category variable. To measure female contribution to agricultural activity we use the sum of three variables: harvesting, soil preparation, and crop tending. This produces a scale which ranges from 3 (100% male participation in all three activities) to 15 (100% female participation in all three activities). We report results below using this composite score, but all relationships were also tested separately for each of the three agricultural tasks, yielding identical patterns in all cases.

We first test the relationship of total female agricultural participation, labeled by T, to the two independent variables, using ordinary-least-squares regression. Letting S stand for slavery (1 = absent, 2 = incipient, 3 = present) and C stand for crop type (1 = root, 2 = cereal), the ordinary-least-squares equation shows a strong relationship:[4]

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

Using the language W matrix, the autocorrelation analysis indicates significant linguistic autocorrelation and a stronger relationship; the coefficients for crop type and slavery have decreased, but are still significant.

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

Using the exponential decreasing distance W, the autocorrelation analysis also indicates significant spatial autocorrelation:

$$T^* = \underset{(2.8)}{21.8} - \underset{\substack{(1.33)\\ p < .001}}{4.44C^*} - \underset{\substack{(.66)\\ p < .005}}{1.78S^*} \qquad R^2 = .61 \qquad \rho = \underset{\substack{(.20)\\ p < .05}}{.49}$$

Figure 3 plots residuals from the OLS equation for the relationship of female participation in agriculture, to crop type and slavery. We see from this figure that virtually all societies south of 5 degrees north latitude have positive residuals, and that virtually all societies north of 7 degrees north latitude have negative residuals. Southern African societies have more female agricultural participation than would be predicted by the OLS regression, and Northern African societies have less female agricultural participation than would be predicted by the equation. West African societies close to the coast in Southern Nigeria, Ghana, and the Ivory Coast have positive residuals, whereas West African societies further north in the Sahel region have negative residuals.

Because of the strength of the linguistic autocorrelation term, and because Bantu societies are concentrated in Southern Africa (and their historical predecessor came from the coastal region of West Africa where are found the related Niger-Congo languages) we hypothesized that the autocorrelation pattern in Figure 3 could be due to the rapid expansion of Bantu societies within the last 2,000 years. Bantu societies appear to have originally practiced tropical root-crop agriculture, with a high female contribution to the agricultural labor. Introducing a linguistic autocorrelation term lowers the regression coefficients between the sexual division of labor, crop type, and slavery, suggesting that part of that relationship may have been due to presence in the sample of Bantu societies. As a preliminary test of these hypotheses, we computed a contingency table showing relationships among crop type, slavery, the sexual division of labor in agriculture, and membership in the Bantu language group (Table IV).

Until recently, the approach taken to analyzing this table would have been to examine separately the bivariate relationships among the four variables. There are six such relationships—slavery to crop type *(SC)*, slavery to Bantu language *(SB)*, slavery to female contribution *(SF)*, crop type to Bantu language *(CB)*, crop type to female contribution *(CF)*, and Bantu language to female contribution *(BF)*. Taking this approach, one would conclude that all six relationships were valid. Such a conclusion would be the invalid result of analyzing a multivariate system at the bivariate level. The best approach to multivariate analysis of contingency tables is log-linear analysis (Bishop, Fienberg, and Holland 1975). This approach has proven useful in the analysis of cross-cultural data with regional effects (Burton and Reitz 1981). The strategy for log-linear analysis of complex systems of three or more variables depends upon the hierarchical nesting of models. One begins with a complete model, which in the present case would be the model containing all six bivariate relationships. Then terms are deleted, iteratively. At each stage, the term that is deleted is the one which makes the smallest contribution to the goodness of fit measure, G^2, provided that its contribution is not statistically significant. This process is the reverse of the stepwise procedure of linear regression, in which terms are added, provided their contributions are statistically significant.

The deletion record for the current analysis is shown in Table V. In the first row we see the complete model. G^2 is a measure that is arithmetically close to χ^2. The p level in column three is an estimate of the probability that the sample data could have been generated by the model, so a high p level shows a good fit of the model to the data. In the first row we see that the p level for the complete model is .995. There is no question that this model provides an adequate fit to the data; however, we must examine whether the model is redundant. The term *SB*, for the relationship between language and slavery, has

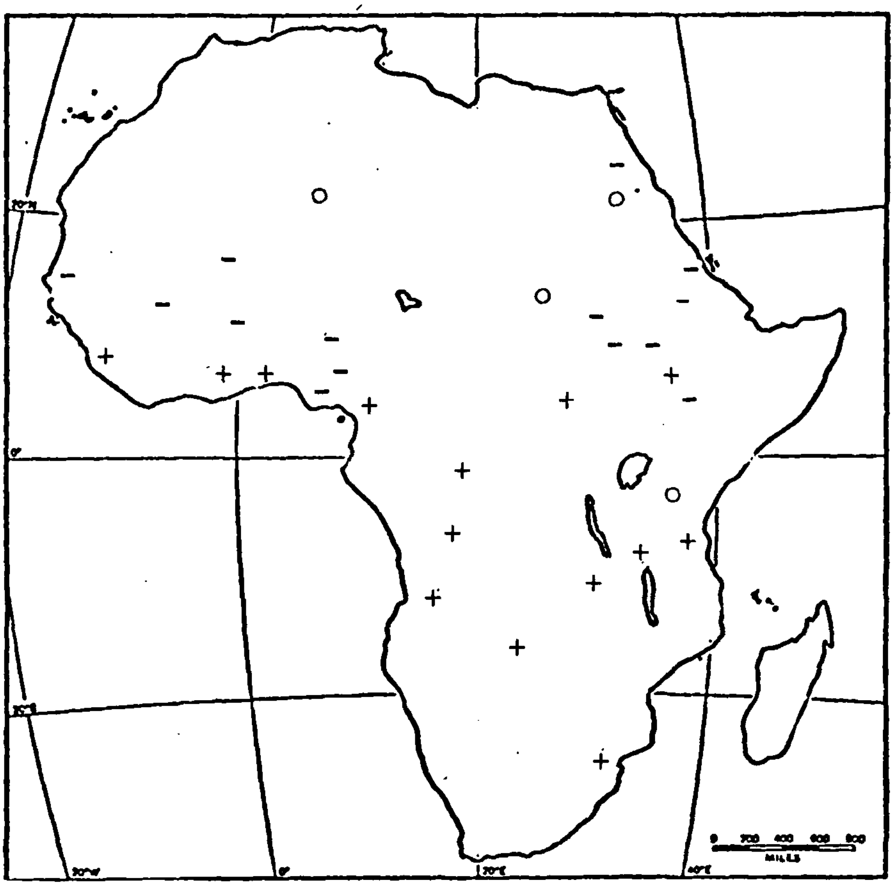


Fig. 3. Residuals from regression predicting sexual division of labor in agriculture.

+ More female participation in agriculture than predicted by OLS regression
o Same amount of female participation in agriculture as predicted by OLS regression
– Less female participation in agriculture than predicted by OLS regression

the lowest contribution to G^2 – .12. We see in row two that deletion of the term *SB* results in a model with a very good fit to the data. The last column is an estimate of the probability that a difference in G^2 as large as .12 could have occurred by chance. It is apparent that it could have, so the term *SB* is deleted, meaning that we conclude that there is no independent relationship between slavery and membership in the Bantu language family. In row three we see that the term *BC*, for relationship between language and crop type, contributes only .53 to the G^2 measure, and that this contribution is not significant; and once *BC* is deleted, we see in row four that the term *SC*, for the relationship between slavery and crop type, also makes a nonsignificant contribution to G^2. This leaves a model in which the female contribution to agriculture has independent relationships to language, crop type, and slavery. Further deletion of any one of those relationships will result in a significant change in G^2, as is shown in the last four rows of the table.

Both log-linear analysis and regression analysis are multivariate methods, and the log-linear analysis provides a useful replication of regression analysis, which may be helpful for those readers who still have qualms about the use of regression analysis with cross-

TABLE IV. CROSS-TABULATION: CROP TYPE, DIVISION OF LABOR IN AGRICULTURE, BANTU LANGUAGES, AND SLAVERY.

Female Contribution to Agriculture

Slavery	Crop	Language	Low: 1-6	Medium: 7-11	High: 12-15
None	Root	Non-Bantu	0	0	0
		Bantu	0	0	0
	Cereal	Non-Bantu	0	1	0
		Bantu	0	0	2
Incipient	Root	Non-Bantu	0	0	1
		Bantu	0	0	2
	Cereal	Non-Bantu	1	1	1
		Bantu	0	1	1
High	Root	Non-Bantu	0	2	0
		Bantu	0	0	1
	Cereal	Non-Bantu	9	5	0
		Bantu	0	1	2

$N = 31$

cultural data. In the OLS analysis we see the same effects for language family, crop type, and slavery as were found with the log-linear analysis:

$$T = \underset{(2.3)}{16.4} \; -\underset{\substack{(1.04)\\p < .001}}{4.06C} \; -\underset{\substack{(.66)\\p < .05}}{1.79S} \; +\underset{\substack{(.95)\\p < .001}}{3.41B} \qquad R^2 = .63$$

The language autocorrelation analysis now shows a reduced amount of linguistic autocorrelation, with a significant effect for Bantu language membership:

$$T^* = \underset{(3.0)}{15.6} \; -\underset{\substack{(.95)\\p < .001}}{3.25C^*} \; -\underset{\substack{(.57)\\p < .05}}{1.58S^*} \; +\underset{\substack{(.99)\\p < .01}}{2.52B^*} \qquad R^2 = .72 \qquad \rho = \underset{\substack{(.24)\\p < .05}}{.59}$$

The spatial autocorrelation analysis now shows no spatial autocorrelation, leading to the conclusion that the observed spatial autocorrelation can be accounted for by the spatial clustering of the societies with Bantu languages:

$$T^* = \underset{(3.1)}{16.4} \; -\underset{\substack{(.98)\\p < .001}}{4.06C^*} \; -\underset{\substack{(.61)\\p < .005}}{1.79S^*} \; +\underset{\substack{(.89)\\p < .001}}{3.41B^*} \qquad R^2 = .67 \qquad \rho = \underset{\substack{(.26)\\\text{n.s.}}}{.19}$$

An alternative theory of the causes of sexual division of labor in agriculture is found in Boserup (1970), who holds that male-dominated agriculture is a consequence of a shift from long-fallow (extensive) to short-fallow (intensive) agriculture, which is in turn a consequence of increases in population density. We have done a test of this alternative hypothesis, using the two variables, agricultural intensity and population density. Results of that test show very little relationship between the female contribution to agriculture labor and either population density or agriculture intensity. Using ordinary least squares, regressions of the female contribution to agriculture labor on either variable, or on both, produces an R^2 value of nearly zero, with no significant relationships. The autocorrelation analysis shows evidence for both linguistic and spatial autocorrelation, and again

TABLE V. SUMMARY OF LOG-LINEAR ANALYSIS OF TABLE I.[5]

Model	Degrees of freedom	G^2	p	Difference in G^2 from preceding model	Degrees of freedom for difference	p level for deletion
FB, BC, FS, SC, SB, FC	17	6.68	.995	—	—	—
FB, BC, FS, SC, FC	19	6.80	.995	.12	2	.943
FB, FS, SC, FC	20	7.21	.995	.41	1	.469
FB, FC, FS	22	10.01	.985	2.81	2	.246
FB, FS	23	16.50	.830	6.57	1	.012
FB, FC	25	19.42	.780	9.41	3	.025
FS, FC	23	28.56	.200	18.55	2	.0001

shows no functional relationships. Letting D = population density and A = agricultural intensity we have for the linguistic autocorrelation analysis:

$$T^* = \underset{\substack{(3.2)\\ \text{n.s.}}}{10.48} - \underset{\substack{(.36)\\ \text{n.s.}}}{.49D^*} - \underset{(.69)}{.01A^*} \qquad R^2 = .50 \qquad \rho = \underset{\substack{(.14)\\ p < .001}}{.79}$$

and for the spatial autocorrelation analysis:

$$T^* = \underset{\substack{(2.99)\\ \text{n.s.}}}{12.72} - \underset{\substack{(.39)\\ \text{n.s.}}}{.47D^*} - \underset{(.70)}{.36A^*} \qquad R^2 = .47 \qquad \rho = \underset{\substack{(.14)\\ p < .001}}{.69}$$

We can safely conclude that the observed variations in the division of agricultural labor across Africa are not due to variations in agricultural intensity or population density. Hence, although one of our predictor variables, crop type, does appear to be related to agricultural intensity, we can conclude that it is the material circumstances of production of the crops themselves that should be used as explanatory variables, not agricultural intensity per se.

PREDICTIONS OF THE DEGREE OF POLYGYNY

Our model predicts that high female participation in agriculture will lead to increased degree of polygyny, all other things being equal, and that patri-virilocal residence will act to facilitate polygyny. To test this part of the model we use a variable pertaining to type of marriage taken from the CCCCC data on settlement patterns and community organization (Murdock and Wilson 1972). The four categories with numerical codes are:

1: Polyandry
2: Monogamy
3: Limited polygyny (less than 20% of all marriages)
4: Intensive polygyny (more than 20% of all marriages)

In the African sample there are no cases of polyandry, and only two cases of monogamy—the Amhara and the Tuareg. Of the remaining 41 societies, 15 have limited polygyny and 26 have intensive polygyny. From a map of the distribution of polygyny across Africa (Figure 4), it is very obvious that the variable is spatially clustered. Intensive polygyny is clustered in West and West Central Africa, limited polygyny in Northeast and East Central Africa. A secondary cluster of intensive polygyny is found in East Africa. Monogamy is found only in the northern half of the continent, and the two monogamous societies are proximal to societies with limited polygyny. The question here is whether there are also significant relationships between polygyny and the independent variables, and whether there is significant spatial clustering of polgyny other than that which could be accounted for by the spatial clustering of the independent variables that we have already noted. The disturbances autocorrelation model is addressed to precisely this kind of question. We test the model using the variables:

T = Total female contribution to harvesting, soil preparation and tending
P = Degree of polygyny
R = Marital residence, coded as follows: 1 = uxorilocal; 2 = avunculocal; and 3 = virilocal

The sample includes 34 agricultural societies for which there is information on all three variables. The ordinary-least-squares equation shows a relatively strong relationship in the predicted direction:

$$\begin{array}{llll} P = 1.60 & +.093T & +.42R & R^2 = .40 \\ \quad (.48) & (.023) & (.14) & \\ & p < .001 & p < .005 & \end{array}$$

This equation is easily interpreted. A hypothetical matrilocal society in which men did all of the agricultural labor would be predicted to have a polygyny value of 2.3, and would hence be predicted to be monogamous. At the other extreme, a society in which women did all of the agricultural labor, and in which there was patrilocal residence, would be predicted to have a polygyny value of 4.3, and would, therefore, be predicted to have intensive polygyny. The equation makes sense in terms of our predictions about the effects of residence and of female agricultural labor on the degree of polygyny.

A test of the same model using the linguistic autocorrelation matrix shows no evidence for linguistic autocorrelation:

$$\begin{array}{lllll} P^* = 1.60 & +.092T^* & +.42R^* & R^2 = .43 & \rho = .49 \\ \quad (.43) & (.022) & (.13) & & (.29) \\ & p < .001 & p < .005 & & \text{n.s.} \end{array}$$

Hence, although the sexual division of labor in agriculture exhibits strong linguistic autocorrelation, the relationship between that sexual division of labor and the degree of polygyny does not exhibit linguistic autocorrelation. We can interpret this to mean that the division of labor diffused as part of the Bantu migration, but that form of marriage has a separate causal relationship to the division of labor.

Spatial autocorrelation analysis has different results; there is clear evidence for spatial autocorrelation in the distribution of polygyny:

$$\begin{array}{lllll} P^* = 1.77 & +.06T^* & +.43R^* & R^2 = .51 & \rho = .45 \\ \quad (.44) & (.02) & (.13) & & (.21) \\ & p < .005 & p < .005 & & p < .05 \end{array}$$

We think this spatial autocorrelation has a clear interpretation. Figure 5 is a plot of the residuals from the ordinary-least-squares regression, predicting degree of polygyny from female contribution to agriculture and from residence. A plus mark means the society has more polygyny than is predicted by the regression; a minus means it has less polygyny than is predicted by the regression. If the predicted amount of polygyny is within .1 of the actual amount of polygyny, a zero is indicated. We see from this map that there is a band of societies in West and Central Africa, just north of the equator, that has more polygyny than predicted. Another group of societies in North and Northeast Africa has less polygyny than predicted, while societies in Central and South Africa have either less polygyny than predicted or the same amount as predicted. These clusterings appear to result from diffusion. Much of the northern half of the African continent has been influenced by two major monotheistic religions (Islam and Coptic Christianity) that prohibit or limit polygyny. At the same time, there appears to be a diffusion effect in West Africa. Many of the societies in West Africa have relatively low female participation in agriculture yet still have intensive polygyny. Hence, polygyny levels in West Africa are higher than would be predicted by the ordinary-least-squares equation alone.

To test the hypothesis that there is a diffusion effect on the polygyny variable due to the presence of monotheistic religions that prohibit or limit the practice of polygyny, we examine a variable from the Ethnographic Atlas pertaining to the presence of high gods. This variable has four categories: high gods absent; high gods present but not active in human affairs; high gods present and active in human affairs, but not supportive of

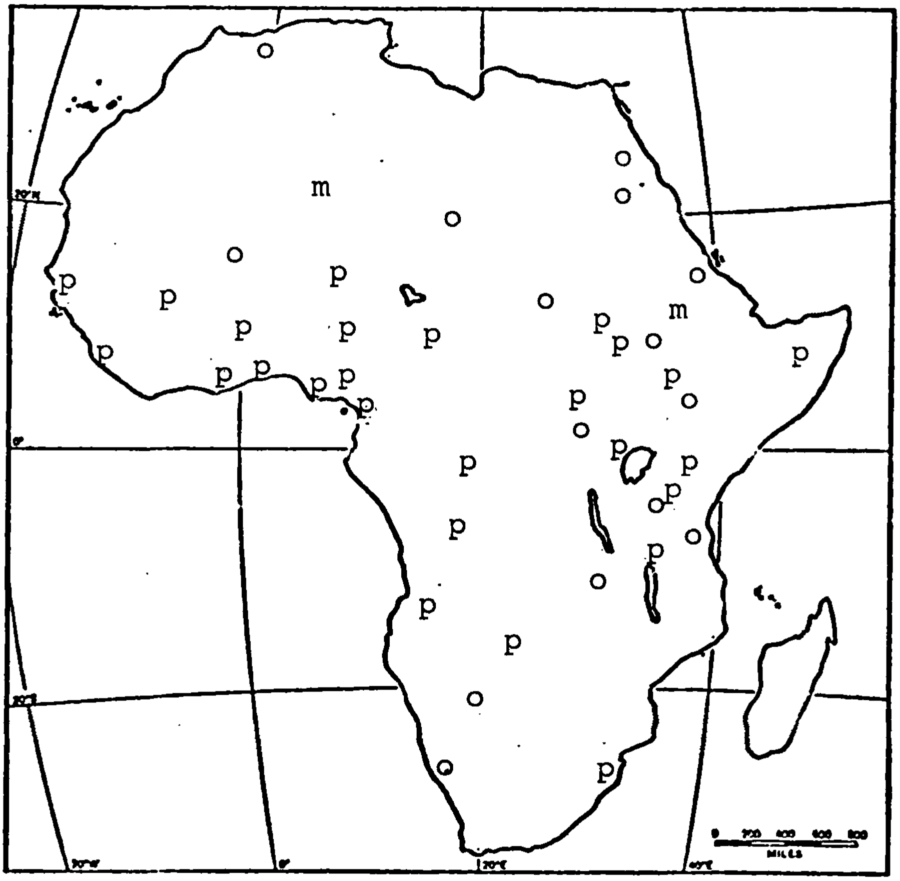


Fig. 4. Distribution of polygyny in Africa.

m Monogamy
o Limited polygyny
p Intensive polygyny

morality; and high gods present, active, and specifically supportive of human morality. Eleven of the African societies in our sample fall into this latter category, and it is in this category that we would expect to find the effect of monotheism upon marriage practices. To test the effect of this variable on polygyny we have computed a three-way contingency table among the female contribution to agriculture, the degree of polygyny, and the presence of a high god supporting human morality (Table VI).

We can see from this table the effect on polygyny of belief in a moralistic high god. Two of the 11 societies with a high god are monogamous, and of the remainder, only two out of nine practice extensive polygyny. By contrast, among societies that do not have such a high god, 17 out of 21 practice extensive polygyny. For African societies with a moralistic high god, limited polygyny is the norm, whereas for societies without such a high god, extensive polygyny is clearly the norm. The relationship between polygyny and female contribution to subsistence can also be seen quite clearly in this table. Surprisingly, there is also a striking relationship between the high god variable and the female contribution to subsistence: none of the societies with a moralistic high god has a high female contribution to agricultural subsistence. Whether this effect is due specifically to the ef-

TABLE VI. RELATIONSHIPS AMONG HIGH GOD, FEMALE CONTRIBUTION TO SUBSISTENCE, AND POLYGYNY.

	No high god, or high god that does not support morality	High god that supports morality
Female contribution to subsistence		
3-8		
Monogamy	0	2
Limited polygyny	3	7
Extensive polygyny	5	2
9-15		
Monogamy	0	0
Limited polygyny	1	0
Extensive polygyny	12	0

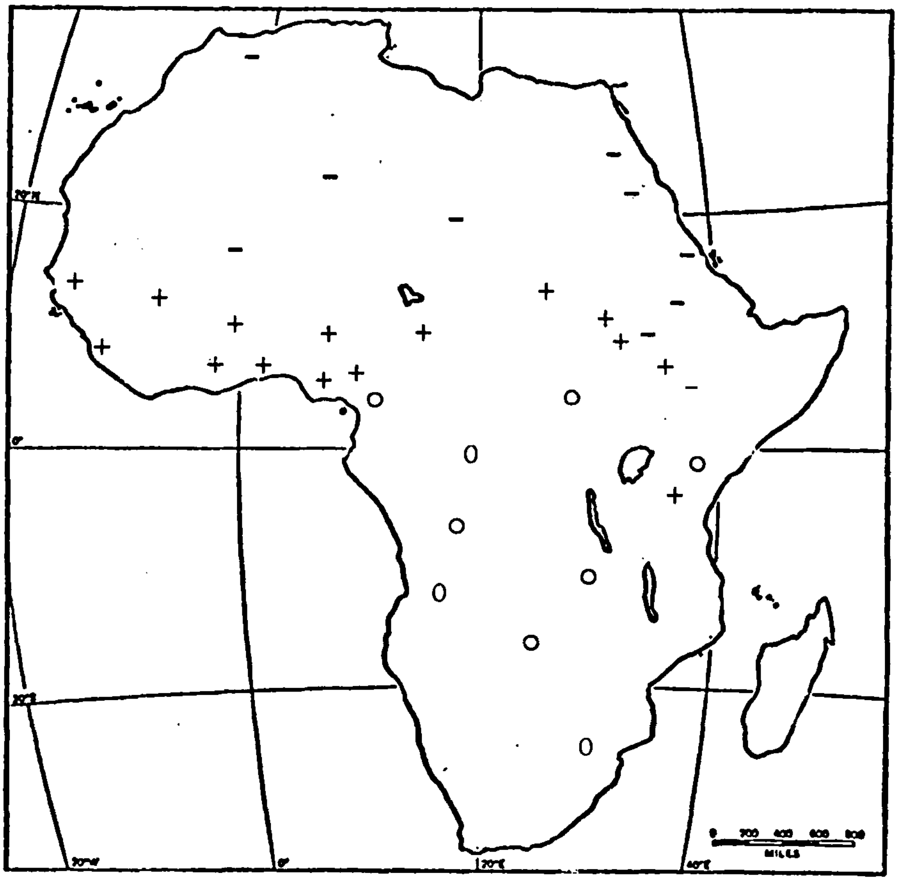


Fig. 5. Residuals from regression predicting polygyny.

+ More polygyny than predicted by the regression
o Same amount of polygyny as predicted
– Less polygyny than predicted

fects of religion upon the division of labor, through institutions such as female seclusion, or whether it is due to the variables that were discussed earlier in this paper remains to be investigated in future research.[6]

OBSERVATIONS ON GALTON'S PROBLEM

Galton's problem is not simply a problem of spatial or historical patterning, it is a problem of the applicability of particular statistical models based on particular assumptions. Correlational and regression models assume the absence of autocorrelation of error terms. The existence of autocorrelation can, however, be specified as part of the model. In this case we are not concerned about the lack of independence in the sampling units since the statistical model specifically incorporates interdependencies. If the model is correctly specified, then the prediction errors should behave like white noise, and will not exhibit autocorrelation. In this context, Galton's problem is one of specification. If the specification is correct, an examination of error terms, as in Figures 3 and 5, should show the errors to be nonautocorrelated. Figures 3 and 5, however, show the patterning of residuals for models which were incorrectly specified. Figure 3 represents the errors from a regression model for female participation in agriculture with no spatial or linguistic effects (interdependencies), but the error terms are obviously spatially autocorrelated. The autocorrelated-disturbances model includes spatial or linguistic interdependencies, and removes those autocorrelations dependent upon the specification of those interdependencies. If properly specified, the autoregressive model efficiently estimates not only the regression coefficients of the independent variables, but their variances, making possible the comparison of different regional replications of the model. Only if the slopes of the regression coefficients (i.e., magnitudes) replicate from one region to another (as determined by significance tests which make use of consistent estimates of their variances) can we regard a model as properly specified on a worldwide basis. The determination of such variances is of crucial importance for the testing of theory on a worldwide basis.

Attention to comparison of slopes of independent variables in the autoregression and OLS models, with and without incorporation of autocorrelation effects, shows the close

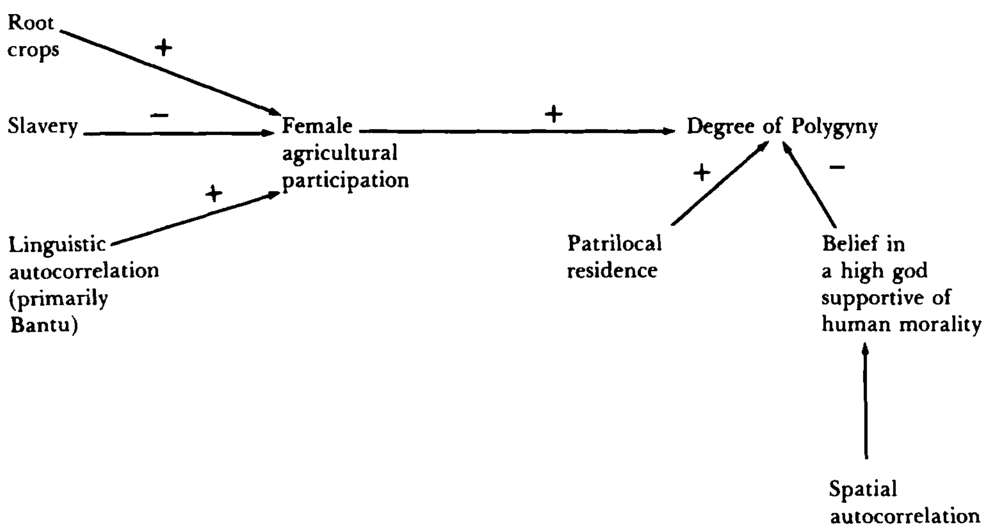


Fig. 6. Final Model of System of Variables Pertaining to the Agricultural Division of Labor.

correspondence between autocorrelation and the problem of ecological correlation. In the latter case the aggregation of observations of different types leads to regression slopes which differ in the aggregate from those found within each type. In the autocorrelation case it is not aggregation but interaction which leads to a distortion of the true slopes of the independent variables when autocorrelated disturbances are ignored.

SUMMARY AND CONCLUSIONS

Autocorrelation analysis of relationships among crop type, slavery, residence, polygyny, and female participation in agriculture have validated the model that was formulated in Figure 1. We have found evidence for two processes of autocorrelation within the system of variables pertaining to the agricultural division of labor. First, the agricultural division of labor is linguistically autocorrelated; a phenomenon that we attribute to the rapid and relatively recent Bantu expansion. Second, polygyny is spatially autocorrelated, a fact that we attribute to the presence of monotheistic religions in North and Northeast Africa. The final model looks very much like the original model, with the addition of autocorrelation terms (Figure 6), and of the variable for belief in a high god.

This study demonstrates the usefulness for cross-cultural research of the multiple regression approach to autocorrelation. The present paper is a comparative regional study. Functional relationships are examined within a region in which certain variables, such as the plow and climate, display less variation than in a worldwide sample. The comparative regional study allows us to examine genetic relations among societies and other autocorrelation processes, such as trade, diffusion, and conquest, in more detail than is possible with a worldwide sample. Within the African sample, two observations about autocorrelation emerge from the present study that we think are particularly important. The first is that autocorrelation does exist, and that the multiple regression approach allows us to separate the effects of autocorrelation from the functional relationships. Autocorrelation changes the strength of relationships, but does not obscure them, so that all relationships that are observed with the autocorrelation analysis can also be observed using more traditional analyses, such as ordinary-least-squares regression, or contingency table analysis. Thus, the autocorrelation analysis allows us to be more precise about the exact form of relationships, but does not change our understanding of the direction of relationships. Whether this last generalization will hold up in further replications of the same study, or in different studies, remains to be seen.

The second important generalization is that the observed autocorrelation effects are due to historical processes that are already well known to African ethnologists. Thus, we can have some confidence that anthropologists have been on the right track. The historical processes that they have focused upon are the very same processes which appear in our autocorrelation analysis. No new, previously unnoticed, historical relationships appear in the present analysis. The strength of the present analysis is that it allows much more specific statements about the ways in which the historical processes have affected particular relationships. Ethnologists have long known that Bantu societies had a high female participation in subsistence. The present study says that this fact is partly the consequence of the historical connectedness of Bantu societies. Bantu societies have more female participation in agriculture than can be accounted for on the basis of other facts about their economic and ecological circumstances. Similarly, ethnologists have long known that there is a high incidence of polygyny in West Africa; and the present study suggests that this is in part the consequence of some regional phenomenon, such as diffusion.

The strength of the autoregressive approach to a regional study, such as the present one, is that it makes possible unbiased and consistent estimates of the relationships among the variables, whereas a model that did not take autocorrelation into account would produce unbiased but inconsistent estimates. For example, in the present study the OLS model overestimated the strength of the relationships among polygyny, female agricultural participation, crop type, and slavery. Obtaining accurate estimates through the autocorrelation analysis will allow for the next step in the analysis, to compare results across regions. The final goal, then, will be to put together models from the several world regions, each one taking into consideration the most important within-region autocorrelation effects, to produce models of greater generality.

NOTES

* The two co-principal investigators share primary and equal responsibility for this paper.

Acknowledgments. The network autocorrelation techniques used in this paper were programmed in APL on the UCI Sigma-7 computer by David Hansen, using algorithms developed by Patrick Doreian (1980, 1981). Support for this project was provided by an NSF grant to Douglas R. White, by the School of Social Sciences of UCI, by a grant from the Committee on Research of the Academic State of UCI, and by a grant from the Committee on Instructional Development of UCI. We are grateful to Russ Bernard, Frank Cancian, Patrick Doreian, Linton Freeman, Ellen Greenberger, Robert LeVine, David Strauss, Beatrice Whiting, and John Whiting for helpful comments on this paper.

[1] It has long been common practice to include nominally measured independent variables in regression analysis. In econometrics, these are called dummy variables, and they usually serve to indicate policy changes in time series analysis. It is also possible in regression analysis to use nominal variables as dependent variables, provided their distribution is not highly skewed (Johnston 1972:183).

[2] We have replicated several of the equations from this paper using the effects model. We find that the effects model tends to yield a high autocorrelation coefficient, no significant functional relationships, and a lower R^2 value than does the disturbances model. We believe this is because the effects model treats autocorrelation as if there were no autocorrelation in the independent variable. By so doing, it inflates the estimate of autocorrelation in the dependent variable by throwing all autocorrelation onto the dependent variable only. The reduced R^2 is evidence that the effects model is misspecified. A later paper will discuss the differences between the two models in detail.

[3] There are four major phyla of African languages—Click, Afroasiatic, Nilo-Saharan, and Congo-Kordofanian. They differ in the number of levels in their historical reconstruction. This may be an artifact of the number of extant languages in each phylulm. Congo-Kordofanian, which includes the Bantu languages and is found in all parts of the African continent, has six levels, whereas the Click phylum, found only in parts of Southern and Eastern Africa, has only two levels. Our method of constructing the W matrix normalizes for these differences in numbers of levels. A language receives a positive value in the W matrix only for languages in the same phylum. All other entries are zero. Each value, W_{ij}, is computed by dividing by the row sum of the matrix of proximities among languages. This has the effect of normalizing for differences in the height of the four tree structures. The method of constructing language W matrices has another consequence—the W matrix can be completely partitioned into four separate W matrices. This would provide a very useful computational shortcut if we were using a very large sample, in that we could compute the eigenvalues for each of the submatrices separately, then combine them. For all practical purposes it is difficult to compute eigenvalues for a matrix larger than 70×70, but by combining smaller submatrices in this way we could work with a considerably larger sample size.

[4] We have used unstandardized regression coefficients throughout this paper. The number under the coefficient is the t statistic, used to test the significance of the coefficient. Unstandardized coefficients are suitable for comparative research in which the goal is to compare analyses across data sets. Coefficients which have been standardized relative to a particular data set cannot be compared across data sets (Dow 1979). We have listed a measure of goodness of fit—R^2—for both the OLS

regression equations and the autocorrelation regression equations. In both cases, this measure is the square of the correlation between observed and predicted values. In the autocorrelation case, however, R^2 does not have the same distribution as the R^2 from OLS, so one cannot use the same significance tests on it, nor should it be interpreted as a proportion of variance explained.

[5] The table has structural zeroes, that is, cells that are zero because of marginal totals that are zero. Tables with structural zeroes require adjustments in the measures of degrees of freedom, as described in Bishop, Fienberg, and Holland (1975). The degrees of freedom numbers in the table have been adjusted to compensate for zero cells.

[6] We have performed a log-linear analysis upon the data of Table VII, and the analysis suggests that the best model for these data is one in which there are independent relationships of polygyny to the female subsistence contribution, polygyny to the moralistic high god, and female subsistence contribution to the moralistic high god.

REFERENCES CITED

Barlett, Peggy F.
1976 Labor Efficiency and the Mechanism of Agricultural Evolution. Journal of Anthropological Research 32:124-140.

Baumann, Hermann
1928 The Division of Work According to Sex in African Hoe Culture. Africa (London) 1: 289-319.

Bishop, Yvonne M. M., Stephen E. Fienberg, and Paul W. Holland
1975 Discrete Multivariate Analysis: Theory and Practice. Cambridge: MIT Press.

Boserup, Ester
1965 The Conditions of Agricultural Growth: The Economics of Agrarian Change Under Population Pressure. Chicago: Aldine.
1970 Women's Role in Economic Development. London: Allen & Unwin.

Bronson, Bennet
1972 Farm Labor and the Evolution of Food Production. *In* Population Growth: Anthropological Implications. Brian Spooner, ed. pp. 190-218. Cambridge: MIT Press.

Burton, Michael L., Lilyan A. Brudner, and Douglas R. White
1977 A Model of the Sexual Division of Labor. American Ethnologist 4:227-251.

Burton, Michael L., and K. Reitz
1981 The Plow, Female Contribution to Agricultural Subsistence, and Polygygy: A Log-Linear Analysis, Behavior Science Research 16:4.

Cliff, A. D., and J. K. Ord
1973 Spatial Autocorrelation. London: Pion.
1975 The Comparison of Means when Samples Consist of Spatially Autocorrelated Observations. Environment and Planning 7:725-734.

Cowgill, George L.
1975 On Causes and Consequences of Ancient and Modern Population Changes. American Anthropologist 77:505-525.

Divale, William Tulio
1974 Migration, External Warfare, and Matrilocal Residence. Behavior Science Research 9:75-133.

Doreian, Patrick
1980 Linear Models with Spatially Disturbed Data: Spatial Disturbances or Spatial Effects? Sociological Methods and Research 9:29-60.
1981 Estimating Linear Models with Spatially Distributed Data. Sociological Methodology 1981:359-388. S. Leinhardt, ed. San Francisco: Jossey-Bass.

Dow, Malcolm McLaren
1979 Structure and Process in Cultural-Ecological Complexes: A Comparative Canonical-Redundancy and Causal Analysis. Ph.D. Dissertation, School of Social Sciences, University of California, Irvine.

Dow, Malcolm M., Douglas R. White, and Michael L. Burton
1981 Multivariate Modeling With Interdependent Network Data. Behavior Science Research. In Press.
Dow, Malcolm, Douglas R. White, and David Hansen
1979 Regression Analysis and Network Autocorrelation. Manuscript. School of Social Sciences, University of California, Irvine.
Driver, Harold E.
1956 An Integration of Functional, Evolutionary, and Historical Theory by Means of Correlations. Indiana University Publications in Anthropology and Linguistics Memoir 12:1-35. Suppl. to International Journal of American Linguistics 22:1.
Ember, Melvin
1974 Warfare, Sex Ratio, and Polygyny. Ethnology 13:197-206.
Ember, Melvin, and Carol R. Ember
1971 The Conditions Favoring Matrilocal versus Patrilocal Residence. American Anthropologist 73:571-94.
Geertz, Clifford
1963 Agricultural Involution: the Process of Ecological Change in Indonesia. Berkeley and Los Angeles: University of California Press.
Gillin, John P.
1948 The Ways of Men, an Introduction to Anthropology. New York: Appleton-Century.
Goody, Jack
1977 Production and Reproduction: A Comparative Study of the Domestic Domain. Cambridge: Cambridge University Press.
Greenberg, Joseph
1966 The Languages of Africa. Indiana University Research Center in Anthropology, Folklore, and Linguistics, Publication #25. Bloomington: Indiana University.
Guthrie, Malcolm
1967 The Classification of the Bantu Languages. London: Dawsons.
Heath, Dwight B.
1958 Sexual Division of Labor and Cross-Cultural Research. Social Forces 37:77-79.
Hibbs, D.
1974 Problems of Statistical Estimation and Causal Inference in Time Series Regression Models. Sociological Methodology 1973-74. pp. 252-308. San Francisco: Jossey-Bass.
Johnston, J.
1972 Econometric Methods. New York: McGraw-Hill.
Jorgensen, Joseph G.
1979 Cross-Cultural Comparisons. Annual Review of Anthropology 8:309-331.
Kärger, Karl
1889 Brasilianische Wirtschaftsbilder: Erlebnisse und Forschungen. Berlin: Gergonne.
Lancaster, C. S.
1976 Women, Horticulture, and Society in Sub-Saharan Africa. American Anthropologist 78:539-564.
Loftin, Colin
1972 Galton's Problem as Spatial Autocorrelation: Comments on Ember's Empirical Test. Ethnology 9:425-435.
1975 Partial Correlation as an Adjustment Procedure for Galton's Problem. Behavior Science Research 10:131-141.
Loftin, Colin, and Robert N. Hill
1974 A Comparison of Alignment Procedures for Tests of Galton's Problem. *In* Studies in Cultural Diffusion: Galton's Problem. James M. Schaefer, ed. pp. 25-61. New Haven: Human Relations in Files Press.
Martin, M. Kay, and Barbara Voorhies
1975 Female of the Species. New York & London: Columbia University Press.
Martin, Ronald L.
1974 On Spatial Dependence, Bias, and the Use of First Spatial Differences in Regression Analysis. Area 6:185-194.

Meggers, Betty J.
1971 Amazonia: Man and Culture in a Counterfeit Paradise. Chicago: Aldine.
Murdock, George Peter
1949 Social Structure. New York: Macmillan.
1967 Ethnographic Atlas. Ethnology 9:122-225.
Murdock, George P., and Diana O. Morrow
1970 Subsistence Economy and Supportive Practices: Cross-Cultural Codes 1. Ethnology 9: 302-330.
Murdock, George P., and Catherina Provost
1973 Factors in the Division of Labor by Sex: A Cross-Cultural Analysis. Ethnology 12: 203-225.
Murdock, George P. and Douglas R. White
1969 Standard Cross-Cultural Sample. Ethnology 8:329-369.
Murdock, George P., and Suzanne F. Wilson
1972 Settlement Patterns and Community Organization: Cross-Cultural Codes 3. Ethnology 11:254-305.
Naroll, Raoul
1970a Galton's Problem. *In* A Handbook of Method in Cultural Anthropology. Raoul Naroll and Ronald Cohen, eds. pp. 974-989. New York: Columbia University Press.
1970b Cross-Cultural Sampling. *In* A Handbook of Method in Cultural Anthropology. Raoul Naroll and Ronald Cohen, eds. pp. 889-926. New York: Columbia University Press.
1976 Galton's Problem and HRAFLIB. Behavior Science Research 11:123-147.
Netting, Robert McC.
1968 Hill Farmers of Nigeria: Cultural Ecology of the Kofyar of the Jos Plateau. Seattle and London: University of Washington Press.
Ord, J. K.
1975 Estimation Methods for Models of Spatial Interaction. Journal of the American Statistical Association 70:120-126.
Romalis, Shelly
1979 Sexual Politics and Technological Change in Colonial Africa. *In* Sex Roles in Changing Cultures. Ann McElrody and Carolyn Matthiasson, eds. Department of Anthropology. SUNY Buffalo.
Schaefer, James M.
1974 Studies in Cultural Diffusion: Galton's Problem. New Haven: Human Relations Area Files Press.
Schlegel, Alice, and Herbert Barry, III
1980 The Evolutionary Significance of Adolescent Initiation Ceremonies. American Ethnologist 7:696-715.
Simonton, Dean Keith
1975 Galton's Problem, Autocorrelation, and Diffusion Coefficients. Behavior Science Research 10:239-248.
White, Douglas R., Michael L. Burton, and Lilyan A. Brudner
1977 Entailment Theory and Method: A Cross-Cultural Analysis of the Sexual Division of Labor. Behavior Science Research 12:1-24.
Whiting, Beatrice B., and John W. M. Whiting
1975 Children of Six Cultures: A Psycho-Cultural Analysis. Cambridge: Harvard University Press.

Submitted 18 March 1980
Revised manuscript submitted 24 April 1981
Accepted 15 May 1981