

HOMOGENEITY IN CONFIDING RELATIONS *

Peter V. MARSDEN

*Harvard University ***

Patterns of inbreeding (or homophily) and social distance in data from the 1985 General Social Survey on dyads discussing important matters are examined. Stratifying variables include age, education, race/ethnicity, religion, and sex. Discussion relations are most constrained by race/ethnicity, and least by sex and education. Inbreeding effects are present for all five stratifying variables, and account for virtually all structure in dyads classified by race/ethnicity and religion. Appreciable social distance biases in the formation of these strong ties are found for age and education, but not for other stratifying variables. The analysis illustrates the use of loglinear and log-multiplicative association models in the analysis of cross-classifications of attributes of alters and respondents.

1. Introduction

This paper reports analyses of the patterning of confiding relationships by demographic attributes, for a representative sample of Americans. It examines data from the 1985 General Social Survey (GSS) on pairs of persons who "discuss important matters" with one another. The analyses document both the strength of tendencies toward homophily or inbreeding, for persons sharing an attribute, and the extent of social distance differentiation, for persons having different attributes.

The cross-classification tables on which this paper focuses are of interest for several related reasons. At the most basic level, Blau (1977; Blau and Schwartz 1984) views the presence of tendencies toward

* This work was completed while I was on the faculty at the University of North Carolina at Chapel Hill. A previous version of this paper was presented at the Sixth Sunbelt Social Network Conference, Santa Barbara, CA, February, 1986. The work was supported by a grant from the Measurement Methods and Data Improvement program of the National Science Foundation (SES-8513327) and by the University Research Council of the University of North Carolina at Chapel Hill. I am indebted to Eric Leifer and Stanley Wasserman for helpful comments, and to Priscilla Preston for assistance in preparing the manuscript.

** Department of Sociology, Harvard University, William James Hall, Cambridge, MA 02138, U.S.A.

homophily as a cardinal criterion that an attribute is a meaningful dimension of social structure (see also Laumann 1966; Laumann and Pappi 1976: 51–54).

Moreover, the patterning of relationships by attributes has important implications for the stratification system of a society. Laumann and Senter (1976: 1307) suggest that “the amount of social stratification in a society can be defined usefully as the variable degree to which persons of approximately equal social status choose one another for intimate relationships such as marriage and friendship, beyond what would be expected on the basis of chance factors alone.” Under this conception, inbreeding tendencies reflect the degree of stratification or closure of associations. Systematic tendencies to form or avoid intimate social relationships can be used, furthermore, to infer dimensions generating social distance between “generalized social positions” defined by stratifying variables like occupation, religion or ethnicity (*e.g.* Laumann 1973; Marsden and Laumann 1978). The strength of tendencies to draw associates from proximate social positions is another indicator of the extent to which associations are stratified.

Social integration is the other side of stratification and closure. At the global level, Blau’s (1977) theory of structure is concerned with conditions promoting social integration defined as intergroup contact. Strong homophily tendencies imply less frequent intergroup relations, and thus impede integration. Fararo and Skvoretz (1987) demonstrate this point formally, discussing its implications for the extent and rate of diffusion.

Because homophily tendencies reflect the density of relationships among persons sharing an attribute, they are at once indicators of decreased global and increased local integration. As measures of local integration they index the capacity of a set of like persons to coordinate their actions on some matter of mutual concern (Laumann and Marsden 1979; Burt 1982). Notably, however, very strong homophily tendencies combined with a lack of ties to outsiders can compromise the effectiveness of collective action (Granovetter 1973).

The analyses reported here bear on these issues by providing documentation of a basic proposition of structural analysis: that strong ties tend to occur among persons of similar attributes (Burt, Marsden, and Rossi 1985). The next sections discuss data and methods. This is followed by results on the overall extent to which confiding relationships are structured by the attributes of age, education, race/ethnicity,

religious preference, and sex. More detailed analyses of the patterning are then presented; these reveal the inbreeding and social distance tendencies. The closing section gives comments and suggestions for further work.

2. The 1985 GSS network data

With a few exceptions, notably Blau and Schwartz's (1984) analyses of intermarriage based on U.S. Census data for Standard Metropolitan Statistical Areas, prior analyses of patterns of intimate association have been based on local samples such as the 1973 Detroit Area Study (Laumann 1973). On the one hand, such samples are useful because they permit the analyst to bring specific knowledge about an area, such as historical factors or local labor market conditions, to bear on interpretations. On the other hand, such samples cannot yield a description of associational patterns that generalizes beyond a locality.

The 1985 GSS gathered data on interpersonal environments for a sample of 1534 persons drawn from the non-institutionalized population of the United States. These are the first survey network data representative of a national population. Considerations in gathering the data are discussed in Burt (1984); distributions of network size, density, heterogeneity, and kin-orientation are presented in Marsden (1987), for the sample as a whole and for subgroups defined by age, education, race/ethnicity, size of place, and sex; and the instrument used to gather the data is presented in Burt (1985).

Briefly, the network data center on those persons with whom a focal respondent "discussed important matters" within the six months prior to the GSS interview. Respondents were asked to name all such persons; they were then asked more specific questions about the first five persons named.¹ These "name descriptors" asked about each alter's age, education, race/ethnicity, religion, and sex; about the frequency, duration, intensity, and constituent role relations of the respondent's tie to each alter; and about the intensity of the ties between each pair of alters. The typical respondent named about three alters. These persons tend to have close relationships with one another

¹ Only 5.5 percent of the respondents named more than five alters (Marsden 1987), so few missing data result from this restriction on the follow-up questions.

Table 1
Respondent–alter dyads discussing important matters, classified by age, education, race/ethnicity, religion and sex

A. Age Respondent	Alter				
	< 30	30–39	40–49	50–59	60 or over
< 30	567	186	183	155	56
30–39	191	501	171	128	106
40–49	88	170	246	84	70
50–59	84	100	121	210	108
60 or over	34	127	138	212	387

B. Education Respondent	Alter							
	1–6	7–9	10–12	H.S. Grad.	Some College	Assoc. Degree	Bach. Degree	Grad. Degree
1–6 years	16	20	9	36	8	1	5	2
7–9 years	10	54	43	105	24	7	29	13
10–12 years	15	50	89	156	50	10	31	19
H.S. Grad	14	70	183	658	250	39	115	85
Some Coll.	6	28	76	276	225	44	137	93
Assoc. Deg.	2	2	13	55	38	22	45	35
Bach. Deg.	4	19	11	130	128	27	230	121
Grad. Deg.	1	3	9	50	30	12	92	149

C. Race/ethnicity Respondent	Alter			
	White	Black	Hispanic	Other
White	3806	29	30	20
Black	40	283	4	3
Hispanic	66	6	120	1
Other	21	5	3	34

D. Religious preference Respondent	Alter				
	Protestant	Catholic	Jewish	None	Other
Protestant	2129	305	22	83	30
Catholic	241	790	24	41	13
Jewish	13	7	68	5	1
None	92	66	12	131	14
Other	27	11	1	4	37

E. Sex Respondent	Alter	
	Male	Female
Male	1245	748
Female	970	1515

Table 2
Estimated design effects due to cluster sampling

	Age	Education	Race/ ethnicity	Religious preference	Sex
Design effect due to cluster-sampling of pairs within respondents	1.208	1.267	1.298	1.507	1.013
Overall design effect	1.812	1.901	1.947	2.261	1.520

(mean network density is 0.61), and some 55 percent of the typical interpersonal environment consists of persons bearing some kinship relation to the respondent (Marsden 1987).

Respondent-alter pairs are cross-classified by the age, education, race/ethnicity, religious preference, and sex of respondent (rows) and alter (columns) in Table 1. Standard demographic data for respondents were recoded into categories used for the name descriptors for alters, so the row and column classifications have identical categories.² Age was categorized using ten-year intervals. Because of occasional missing data on respondents or alters, table totals vary from 4167 (religion) to 4478 (sex).

2.1. Sampling design

The 1985 GSS used a multistage probability sampling design to the block or segment level; at the block level, quota sampling was used, with quotas based on age, employment status, and sex. The GSS documentation (Davis and Smith 1985: Appendix A) indicates that this sampling procedure means that the GSS sample generally has the efficiency of a simple random sample with two-thirds as many cases; so in general the design effect can be taken to be 1.5.

This analysis, however, is concerned with respondent-alter pairs. Pairs are cluster sampled within respondents; thus there may be a design effect beyond that due to the clustering involved in the selection of respondents. This would occur if individual respondents having a

² In the race/ethnicity classification, "Asian" and "Other" categories were combined into a single "Other" category, since cell frequencies are very small.

given attribute, who serve as clusters for the sampling of pairs, tend to cite alters who are more homogeneous than those cited by a typical respondent having that attribute. For example, there would be a loss of efficiency if some men disproportionately tend to cite men only, while other men cite women only; the pairs obtained by sampling them within respondents would then be more homogeneous than those obtained from some sampling frame of pairs.

A procedure proposed by Brier (1980) was adapted to estimate this additional design effect.³ Results, displayed in the first row of Table 2, show virtually no excess clustering of alters by sex, but appreciable effects of the sampling design for the other stratifying variables. The clustering of alters within respondents is especially apparent for religious preference.

The second row of Table 2 gives an overall design effect, obtained by multiplying the effect estimated here by 1.5. This reflects the joint effects of clustering in the sampling of the GSS respondents within geographic areas and of clustering alters within respondents. The chi-square statistics obtained in the analyses reported here have been divided by the appropriate overall factor to adjust for reductions in sample efficiency.

3. Association models

The next section briefly summarizes the strength of the association between the attributes of respondents and alters using simple measures of association (*e.g.* Reynolds 1984; Agresti 1984). Most of the analyses, however, rely on loglinear and log-multiplicative models for cross-classified data (*e.g.* Goodman 1984). The association models used here to

³ Brier's (1980) procedure applies to a within-cluster sampling design in which an observation may fall in any cell of the table, while the sampling of alters within respondents is conditional on a respondent's attribute; this scheme resembles a within-cluster product-multinomial design. I estimated separate design factors by Brier's method for each category of a respondent attribute variable (*e.g.* one for men, one for women) and report a weighted average of these as the design effect in the first row of Table 2. Note that all analyses condition on the marginal totals for attributes of the respondent, as is appropriate for a product-multinomial design (Fienberg 1980).

study patterns of homophily and social distance are special cases of the following log-multiplicative model for cell frequencies F_{ij}

$$\log F_{ij} = \alpha_i + \beta_j + \phi\mu_i\nu_j + \delta_k. \quad (1)$$

This is Goodman's (1979, 1984) row and column effects model (II), with additional parameters δ_k referring to "regions" containing subsets of cells in a table. Parameters α_i and β_j are the usual main effects related to the differential size of row (respondent) and column (alter) marginal totals. Interaction effects between rows and columns are represented by the terms $\phi\mu_i\nu_j$ and δ_k .⁴

Model (1) is log-multiplicative instead of loglinear by virtue of the presence of the third term, which asserts that interactions between rows and columns are due to the product of ϕ , an intrinsic index of association (Goodman 1985) with scores μ_i for rows and ν_j for columns. In the general case, the category scores μ_i and ν_j are taken as unknown parameters to be estimated.⁵ In the present context of a square cross-classification in which the categories of the row and column variables are the same, the scores may be restricted such that $\mu_i = \nu_i$; the model is then termed a "homogeneous row and column effects model." It asserts that tendencies to form relationships between social groups are governed by their proximity (as measured by the product $\mu_i\nu_j$) on a single dimension of social distance. The extent to which that dimension is taken into account in the formation of inter-

⁴ The model of Equation (1) can be generalized to include multiple sets of category scores, per Goodman (1985). On parallels between this approach and the canonical correlation approach to the analysis of cross-classifications, see Goodman (1981); for parallels with correspondence analysis, see Goodman (1986).

⁵ The location and variability of the scores μ_i and ν_j are restricted as follows for purposes of ensuring identification:

$$\sum_i \mu_i P_i = \sum_j \nu_j P_j = 0; \quad \text{and}$$

$$\sum_i \mu_i^2 P_i = \sum_j \nu_j^2 P_j = 1.$$

These restrictions involve no loss of generality; see details in Goodman (1979, 1984, 1985). The usual side constraints on main effects α_i and β_j are also imposed:

$$\sum_i \alpha_i = \sum_j \beta_j = 0.$$

category contacts is given by the magnitude of the index of association ϕ .⁶ Note that the product of scores $\mu_i \nu_j$ will be largest for social groups with extreme scores, and thus that this model incorporates certain “edge effects” (Verbrugge 1977) on self-selection tendencies.

The parameters δ_k in model (1) are special parameters that apply to particular regions of the cross-classification, and they may be used to represent additional components of the interaction of rows and columns. Most commonly, the δ_k are used to refer to individual diagonal cells, and thus give inbreeding effects over and above those implied by the social distance model discussed above. The use of parameters δ_k for individual cells is equivalent to “blinking out” those cells (Goodman 1968) or setting their expected frequencies equal to those observed. Regions of the table may, however, be defined quite generally; see Goodman (1972), Hauser (1978, 1979), or Marsden (1981).

Numerous special cases of the above model are of interest here. When no parameters δ_k are introduced, one obtains a unidimensional social distance model. Category scores μ_i and ν_i may be equated, in which case row and column effects are homogeneous, or they may be different, in which case row and column effects are heterogeneous. If category scores can be hypothesized to be known *a priori*, then the loglinear “uniform association” model of Goodman (1979, 1984; see also Agresti 1984) is obtained. Hypotheses about the intervals separating category scores may be tested by comparing the adequacy of the uniform association model and the row and column effects model.

If, on the other hand, parameter ϕ is constrained to be zero, one obtains loglinear association models based on regions only, per Goodman (1972). When a single parameter δ contrasts cells on and off the main diagonal, the “constant inbreeding” model of Marsden (1981) is obtained. When different parameters δ_k apply to each of the diagonal cells, one obtains the usual model of quasi-independence (Goodman 1968), which in the present setting can be termed a model of “differential inbreeding”, since it permits homophily tendencies to vary across categories.

The results presented below of analyses of the classifications in Table 1 give maximum likelihood estimates. They were obtained using

⁶ Specifically, ϕ is the estimated log-odds ratio for a 2-by-2 table cross-classifying two respondent groups with scores μ_i that differ by 1.0, and two alter groups with scores ν_j that differ by 1.0.

either ANOAS (Clogg and Shockey 1982) or the LOGLINEAR routine implemented in SPSSX (SPSS, Inc. 1983).

4. Overall associations between attributes

To begin, I examined the overall degree of association between respondent and alter attributes. Table 3 displays values of two measures of association calculated for the classifications in Table 1. The statistic λ gives the proportional reduction in error obtained by predicting conditional rather than marginal modes, while γ takes the ordering of categories into account, giving the proportional excess of "concordant" over "discordant" pairs of cases (Reynolds 1984; Agresti 1984).

Table 3 shows that the race/ethnic classifications of respondents and alters are most highly associated. Education and sex have the lowest values of λ , but the corresponding values of γ are nonetheless of moderate magnitude. Age and religion have intermediate values of λ ; the value of γ for age, however, reveals a strong association once the ordering of categories is taken into account.

These statistics are suggestive of the overall magnitude of association for the classifications in Table 1, but they do not reveal anything about the structure or pattern of the association. They do not show whether subgroups differ in their homophily tendencies, whether they differentiate between outgroups in terms of social distance, or how strong any inbreeding and social distance tendencies are. These issues are addressed by the association models discussed in the next section.

Table 3
Overall association of respondent and alter attributes

Attribute	λ (symmetric)	γ
Age	0.348	0.747
Education	0.067	0.453
Race/ethnicity	0.594	- ^a
Religious preference	0.379	- ^a
Sex	0.183	0.444

^a γ not reported for unordered attributes having more than two categories.

Table 4
 Association models for classifications in Table 1^a

Model	G^2	df
A. Age		
Independence	835.86	16
Constant inbreeding	178.86	15
Differential inbreeding (quasi-independence)	113.95	11
Uniform association	264.82	15
Differential inbreeding, uniform association	42.67	10
Homogeneous row and column effects	217.60	12
Differential inbreeding, homogeneous row and column effects	21.50	7
Heterogeneous row and column effects	209.83	9
Differential inbreeding, heterogeneous row and column effects	16.25	4
Quasi-symmetry	16.87	6
B. Education		
Independence	636.02	49
Constant inbreeding	358.43	48
Differential inbreeding (quasi-independence)	285.81	41
Uniform association	125.80	48
Differential inbreeding, uniform association	58.81	40
Homogeneous row and column effects	118.16	42
Differential inbreeding, homogeneous row and column effects	48.91	34
Heterogeneous row and column effects	110.76	36
Differential inbreeding, heterogeneous row and column effects	42.93	28
Quasi-symmetry	24.39	21
C. Race/ethnicity		
Independence	1314.73	9
Constant inbreeding	13.99	8
Differential inbreeding (quasi-independence)	3.10	5
Homogeneous row and column effects	242.25	6
Differential inbreeding, homogeneous row and column effects	1.69	3
Heterogeneous row and column effects	187.11	4
Differential inbreeding, heterogeneous row and column effects	0.06	1
Quasi-symmetry	1.69	3
D. Religious preference		
Independence	1067.28	16
Constant inbreeding	66.04	15
Differential inbreeding (quasi-independence)	9.07	11
Homogeneous row and column effects	312.55	12

Table 4 (continued)

Model	G^2	df
Differential inbreeding, homogeneous row and column effects	5.15	7
Heterogeneous row and column effects	310.14	9
Differential inbreeding, heterogeneous row and column effects	2.00	4
Quasi-symmetry	2.79	6
E. Sex		
Independence	161.46	1

^a Chi-square statistics reported here have been divided by the appropriate overall design effect from Table 2.

5. Inbreeding and social distance patterns

The major results of analyses of the classifications in Table 1 appear in Tables 4 and 5. Table 4 gives likelihood ratio test statistics (G^2) for loglinear and log-multiplicative models applied to the classifications, while Table 5 presents parameter estimates for selected models.

The attributes of respondents and alters are significantly associated in all five classifications, as the poor fit of the independence model to each (Table 4) shows. The principal object of this section is to analyze the form of these associations, in particular the degree to which they are attributable to inbreeding or homophily patterns involving self-selection, as distinct from social distance patterns involving differentiation of outgroups along graded dimensions.⁷

5.1. Race/ethnicity

The results in Table 4 show that the strong tendency for discussion ties to be patterned by race/ethnicity is almost entirely accounted for by

⁷ The primary focus is on the form of interaction between an attribute of a respondent and the corresponding attribute of an alter. The marginal distributions of the classifications in Table 1 are also of some interest; and they are quite similar. Indices of dissimilarity between row and column percentages range from a minimum of 1.1, for race/ethnicity, to 8.3, for age. Thus few "prestige" effects (Laumann 1966) are evident in these citations. Note that the marginal distributions in Table 1 do not yield accurate accounts of the distributions of respondents in the GSS. This is because respondents appear in Table 1 in proportion to the number of alters they cite; and there are clear differences in network size by age, education, and race/ethnicity (Marsden 1987; Marsden and Hurlbert 1987).

Table 5
Estimated parameters for selected models

A. Age, model of homogeneous row/column effects and differential inbreeding		
Age group	Inbreeding effect	Row/column score
< 30	1.148	-0.328
30-39	0.896	-0.224
40-49	0.693	-0.208
50-59	0.919	-0.126
60+	-6.992	0.885
ϕ		7.146
B. Education, model of homogeneous row/column effects and differential inbreeding		
Education level	Inbreeding effect	Row/column score
1-6 years	1.579	-0.494
7-9 years	0.778	-0.344
10-12 years	0.146	-0.345
High school graduate	0.329	-0.113
Some college	0.332	0.115
Associate's degree	0.814	0.260
Bachelor's degree	0.261	0.435
Grad./prof. degree	0.988	0.485
ϕ		5.136
C. Race/ethnicity, models of constant and differential inbreeding		
Race/ethnic group	Constant inbreeding effect	Differential inbreeding effect
White	3.099	2.562
Black	3.099	4.199
Hispanic	3.099	2.920
Other	3.099	3.153
D. Religious preference, model of differential inbreeding		
Religious group	Inbreeding effect	
Protestant	1.611	
Catholic	1.565	
Jewish	4.222	
None	1.931	
Other	3.117	
E. Sex, constant inbreeding (saturated model)		
Sex	Inbreeding effect	
Male	0.478	
Female	0.478	

self-selection or inbreeding tendencies. The model of independence, which asserts no association between the race/ethnicity of respondents and that of alters, has $G^2 = 1314.73$ (panel C, Table 4). This is reduced by nearly 99 percent, to 13.99, by positing a constant level of inbreeding across the four race/ethnic groups, using only one degree of freedom. This model has an associated p -level of 0.082. Permitting the inbreeding tendencies to vary across groups further reduces G^2 to 3.10 (5 df, $p = 0.685$), a very good fit.

Adding social distance parameters—row and column scores—to the differential inbreeding model reduces G^2 by an insignificant amount, but adding parameters for differential inbreeding to a model containing row and column effects reduces G^2 dramatically. Thus the patterning of discussion ties by race/ethnicity is entirely attributable to inbreeding tendencies; ties with race/ethnic groups other than that to which a respondent belongs are formed randomly, rather than in relation to a dimension of social distance.⁸

The parameter estimates in Table 5 show how the strength of inbreeding tendencies varies across race/ethnic groups. Inbreeding is very high for all four groups, but highest for blacks ($\hat{\delta}_2 = 4.199$) and lowest for whites ($\hat{\delta}_1 = 2.562$). These inbreeding effects are very strong, and the fact that the differential inbreeding model fits the data better than the constant inbreeding model means that differences in the magnitude of these effects are significant. Of equal note here is the fact that, for these four major race/ethnic groups, no social distance dimensions need be posited over and above the inbreeding effects in order to explain adequately the patterning of discussion ties by race/ethnicity.

5.2. Religion

The results of the analysis of the table cross-classifying respondent-alter pairs by religious preference are quite similar to those for the race/ethnicity table. Virtually all of the association in this cross-classification can be accounted for by differential inbreeding effects, as shown in Table 4. Adding a single, constant inbreeding parameter to

⁸ This conclusion, of course, holds only for the four-category classification of race/ethnicity used here, and may not apply when race/ethnic groups are distinguished in greater detail (e.g. Laumann 1973).

the independence model reduces G^2 by 94 percent, to 66.04 (15 df). Permitting inbreeding tendencies to vary across religious groups reduces G^2 by an additional 56.97, using 4 degrees of freedom. The overall G^2 for the model of differential inbreeding, 9.07 on 11 df, has an associated p -level of 0.642.

As was the case for race/ethnicity, adding homogeneous row and column effects to the model of differential inbreeding reduces G^2 by an insignificant amount (3.92 on 4 df). Adding the inbreeding effects to a model already including homogeneous social distance effects, on the other hand, improves the fit of the model dramatically; G^2 drops by 307.40 using 5 df (see Table 4).

Parameter estimates for the differential inbreeding model, presented in panel D of Table 5, show that inbreeding tendencies are strongest for Jewish people, and weakest for Catholics and Protestants; persons of other or no religion have intermediate inbreeding tendencies. These inbreeding tendencies are again quite strong, though on average not of the magnitude found for the classification of pairs by race/ethnicity. Once again, the good fit of the model of differential inbreeding implies that no distinctions are drawn among outgroups by respondents; apparent differential social distances between religious groups can be understood in terms of their different tendencies to make *any* outgroup citations (see Marsden 1981).⁹

5.3. Sex

Since sex has only two categories, all models except the independence model are equivalent to the saturated model, which accounts for cell frequencies entirely. I have parameterized the saturated model as a model of constant inbreeding here. It shows tendencies to cite same-sex persons as alters; these tendencies are substantially weaker than the inbreeding tendencies for race/ethnicity and religion that have been presented above. This is due to the many kin alters cited in the GSS network data; inbreeding tendencies by sex are much stronger when calculated for nonkin respondent–alter pairs only ($\hat{\delta} = 1.124$). The

⁹ Again, this conclusion applies only to the five category classification of religious preference available for the GSS network data; it may not hold when more fine-grained distinctions are drawn among Protestant denominations, for example (Laumann 1973; Marsden and Laumann 1978).

inbreeding tendencies are negligible ($\hat{\delta} = -0.073$) for kin pairs; the kin alters cited include many spouses and cross-sex siblings, parents, children, and other relations.

5.4. Age

Inspection of Table 4 shows that the interaction between the ages of respondent and alter has a quite different pattern than the inbreeding configurations observed for race/ethnicity, religion, and sex. The differential inbreeding model reduces the lack of fit of the independence model by 86 percent, to 113.95 (11 df). This latter value, however, remains highly significant. Introduction of homogeneous row and column effects further reduces G^2 to 21.50 (7 df). The bulk of this reduction (71.28) is attributable to a single dimension of social distance along which the age categories are equally spaced; but a significant further reduction (21.17, 3 df) is obtained when the age categories are permitted to have unequally spaced scores on the distance dimension. The patterning of citations by age reflects both homophily and social distance effects. Each of these effects is highly significant, net of the other.

Note that the model of homogeneous row and column effects and differential inbreeding is associated with a very low ($p = 0.003$) descriptive level of significance. I will return later to the issue of what might account for this lack of fit. At this point, though, it is of interest to inspect the parameters of the homogeneous row and column effects and differential inbreeding model for age, which appear in panel A of Table 5.

The estimated scores for age categories on the distance continuum show that there is a great deal of social distance between older (60 and above) persons, and all younger people; the scores for the remaining age categories, while ordered as one would anticipate, are comparatively close to each other, and far from the score for the 60 and above group. The index of association ϕ has a value of 7.146 here, so this distance dimension plays a quite important part in structuring the citations by age.

The inbreeding effects here are effects on inbreeding *net* of those implied by a group's position on the distance continuum. For the four youngest age groups inbreeding effects are positive, though of lower magnitude than those found for the race/ethnicity or religion classifi-

cations. The inbreeding parameter for the 60 and over age group is, however, negative and very large ($\hat{\delta}_5 = -6.992$). What this means is that persons in this group make many fewer inbreeding citations than would be expected based on their extreme position on the social distance continuum; recall that this position implies a very strong inbreeding effect for this group ($5.597 = 7.146 (0.885^2)$), and the large value of $\hat{\delta}_5$ serves mainly to counteract that effect; nonetheless there does seem to be a net "outbreeding" tendency for the older age category within this model ($5.597 - 6.992 = -1.395$), indicating that persons 60 and over, unlike those in the other age groups, have a disproportionate likelihood of making outgroup citations. It seems likely that this reflects cross-generational kin ties.

As remarked earlier, the model of unidimensional social distance with differential inbreeding does not account for all association between the ages of respondents and alters. It does account for more than 97 percent of the association; but the remaining G^2 value (21.50 on 7 df) is highly significant. Two possibilities for understanding the remaining lack of fit were considered. One is a heterogeneous row and column effects model, which permits the category scores for respondents and alters to differ. Use of this model together with differential inbreeding effects reduces G^2 by only 5.25 on 3 df. A second is the model of quasi-symmetry, which can be viewed as a multidimensional distance model (see Goodman 1985); it subsumes all possible homogeneous row and column effects (involving multiple sets of scores) on the association. Its fit is also inadequate ($G^2 = 16.87$, 6 df).

Thus the remaining association between the ages of respondents and alters involves some asymmetry, and this asymmetry cannot be described by the use of different unidimensional sets of scores for the age categories of respondents and alters. Examination of residuals from the model of quasi-symmetry reveals a tendency, net of all symmetric effects on the association, for respondents to cite alters in the next youngest age category, while underselecting those in the next oldest age category. This bias could be modeled by adding diagonal parameters per Goodman (1972, 1985) to the models used here.

The asymmetry and multidimensionality in the age classification is significant at conventional levels, but it should not be overemphasized; these features are associated with less than 3 percent of the association in the table. The bulk of the patterning of citations by age can be understood in terms of inbreeding tendencies for the four younger age

groups together with a dimension of social distance separating the oldest people from the younger groups.

5.5. Education

Many results of the analysis of the table classifying pairs by the education levels of respondents and alters are like those from the analysis by age. The homogeneous row and column effects, differential inbreeding model accounts for more than 92 percent of the patterning of citations by education, but a marginally significant additional component of association remains, since this model has a G^2 value of 48.91 with 34 df ($p = 0.047$, panel B, Table 4). Both the inbreeding and the social distance effects contribute significantly, net of the other, to structuring the association.¹⁰ Heterogeneous row and column effects do not account for the remaining association, but introduction of all possible symmetric effects does reduce G^2 to insignificance; the model of quasi-symmetry has $p = 0.275$.

Parameter estimates for the homogeneous row and column effects, differential inbreeding model are given in panel B of Table 5. These show an ordering of the row-column scores by education level that one would anticipate, and few departures from equal spacing of the categories. The estimated scores show that there is no social distance between the education categories “7–9 years” and “10–12 years”, and that persons holding a Bachelor’s degree may be somewhat closer to those holding Graduate or Professional degrees than they are to those with Associate’s degrees. This social distance dimension is important in structuring citation patterns, as indicated by the high index of association, $\hat{\phi} = 5.136$.

All eight education categories have inbreeding levels higher than those implied by the social distance component of the model. “Edge effects” are suggested by the fact that the two highest net inbreeding parameters are associated with the two extreme education categories, but the effects do not grow monotonically smaller as one moves from either end toward the middle categories of education.

¹⁰ There is, however, only weak evidence of unequal spacing of the education categories (conditional $G^2 = 9.9$ on 6 df).

6. Summary and discussion

The GSS data on respondent–alter dyads who discuss important matters confirm the proposition that persons involved in strong social relations tend to have similar attributes. They reflect the extreme racial/ethnic segregation of American society, while revealing comparatively higher levels of intergroup contact for sex and education. Confiding relations are also appreciably patterned by religious preference and age.

The form of homogeneity differs for the attributes studied here. For the nominal ones—race/ethnicity and religious preference—its primary basis involves ingroup–outgroup distinctions only. These vary in magnitude across race/ethnic and religious groups. For age and education, both inbreeding and social distance biases exert strong effects on the structuring of respondent–alter citations. Sex homogeneity, of course, is trivially accounted for by a model positing inbreeding. These differences in the form of homogeneity correspond to Blau's (1977) distinction between nominal and graduated parameters of social structure (Marsden 1981).

These findings are first steps that describe homogeneity patterns. At least two extensions should be explored. One direction is to attempt to understand how homogeneity patterns arise, and why they vary in magnitude across subgroups. A principle useful for this task is Blau's (1977) emphasis on opportunities for contact; see also Feld's (1981) discussion of foci for the organization of social ties. Due to the homogeneity of foci (families, neighborhoods, etc.) inbreeding effects and social distance differentiation should be much reduced when adjusted for opportunities for contact.

A second extension involves study of the way in which inbreeding and social distance tendencies vary with tie strength, in light of the proposition that the similarity of those involved in relations should rise with strength (Granovetter 1973). The analyses here concern strong ties only; some related results which bear on this point, using the same general modeling approach taken here, are presented in Marsden (1986).

References

- Agresti, Alan
1984 *Analysis of Ordinal Categorical Data*. New York: Wiley.

- Blau, Peter M.
1977 *Inequality and Heterogeneity*. New York: Free Press.
- Blau, Peter M. and Joseph E. Schwartz
1984 *Cross-Cutting Social Circles: Testing a Macrosociological Theory of Intergroup Relations*. New York: Academic Press.
- Brier, Stephen S.
1980 "Analysis of contingency tables under cluster sampling." *Biometrika* 67: 591–596.
- Burt, Ronald S.
1982 "A note on cooptation and definitions of constraint," in: Peter V. Marsden and Nan Lin (eds.) *Social Structure and Network Analysis*, pp. 219–233. Beverly Hills: Sage.
1984 "Network items and the General Social Survey." *Social Networks* 6: 293–339.
1985 "General Social Survey network items." *Connections* 8: 119–123.
- Burt, Ronald S., Peter V. Marsden and Peter H. Rossi
1985 "A research agenda for survey network data." Manuscript, Center for the Social Sciences, Columbia University.
- Clogg, Clifford C. and James W. Shockey
1982 "ANOAS: A computer program for the analysis of association in two-way cross-classifications having ordered categories." Mimeographed. Population Issues Research Center, Pennsylvania State University.
- Davis, James A. and Tom W. Smith
1985 *General Social Surveys, 1972–1985: Cumulative Codebook*. Chicago: National Opinion Research Center.
- Fararo, Thomas J. and John Skvoretz
1987 "Unification research programs: Integrating two structural theories." *American Journal of Sociology* 92: 1183–1209.
- Feld, Scott L.
1981 "The focused organization of social ties." *American Journal of Sociology* 86: 1015–1035.
- Fienberg, Stephen E.
1980 *The Analysis of Cross-Classified Categorical Data*. Second Edition. Cambridge, MA: MIT Press.
- Goodman, Leo A.
1968 "The analysis of cross-classified data: Independence, quasi-independence, and interactions in contingency tables with or without missing entries." *Journal of the American Statistical Association* 63: 1091–1131.
1972 "Some multiplicative models for the analysis of cross-classified data," in: L. LeCam, J. Neyman and E.L. Scott (eds.) *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 649–696. Berkeley, CA: University of California Press.
1979 "Simple models for the analysis of association in cross-classifications having ordered categories." *Journal of the American Statistical Association* 74: 537–552.
1981 "Association models and canonical correlation in the analysis of cross-classifications having ordered categories." *Journal of the American Statistical Association* 76: 320–334.
1984 *The Analysis of Cross-Classified Data Having Ordered Categories*. Cambridge, MA: Harvard University Press.
1985 "The analysis of cross-classified data having ordered and/or unordered categories: Association models, correlation models, and asymmetry models for contingency tables with or without missing entries." *Annals of Statistics* 13: 10–69.
1986 "New methods for analyzing the intrinsic character of qualitative variables using cross-classified data." Manuscript, Departments of Sociology and Statistics, University of Chicago.

- Granovetter, Mark S.
1973 "The strength of weak ties." *American Journal of Sociology* 78: 1360–1380.
- Hauser, Robert M.
1978 "A structural model of the mobility table." *Social Forces* 56: 919–953.
1979 "Some exploratory methods for modeling mobility tables and other cross-classified data," in: K.F. Schuessler (ed.) *Sociological Methodology 1980*, pp. 413–458. San Francisco: Jossey-Bass.
- Laumann, Edward O.
1966 *Prestige and Association in an Urban Community*. Indianapolis: Bobbs-Merrill.
1973 *Bonds of Pluralism: The Form and Substance of Urban Social Networks*. New York: Wiley Interscience.
- Laumann, Edward O. and Peter V. Marsden
1979 "The analysis of oppositional structures in political elites: Identifying collective actors." *American Sociological Review* 44: 713–732.
- Laumann, Edward O. and Franz U. Pappi
1976 *Networks of Collective Action: A Perspective on Community Influence Systems*. New York: Academic Press.
- Laumann, Edward O. and Richard Senter
1976 "Subjective social distance, occupational stratification, and forms of status and class consciousness: A cross-national replication and extension." *American Journal of Sociology* 81: 1301–1338.
- Marsden, Peter V.
1981 "Models and methods for characterizing the structural parameters of groups." *Social Networks* 3: 1–27.
1986 "Heterogeneity and tie strength: An analysis of second-order association." Paper presented at the Annual Meetings of the Southern Sociological Society, session on Categorical Data Analysis, New Orleans, LA, April.
1987 "Core discussion networks of Americans." *American Sociological Review* 52: 122–131.
- Marsden, Peter V. and Jeanne S. Hurlbert
1987 "Small networks and selectivity bias in the analysis of survey network data." *Social Networks* 9: 333–349.
- Marsden, Peter V. and Edward O. Laumann
1978 "The social structure of religious groups: A replication and methodological critique," in: S. Shye (ed.) *Theory Construction and Data Analysis in the Behavioral Sciences*, pp. 81–111. San Francisco: Jossey-Bass.
- Reynolds, H.T.
1984 *Analysis of Nominal Data*. Second Edition. Beverly Hills: Sage.
- SPSS, Incorporated
1983 *SPSSX User's Guide*. New York: McGraw-Hill.
- Verbrugge, Lois M.
1977 "The structure of adult friendship choices." *Social Forces* 56: 576–597.