



Statistical Analysis of Multiple Sociometric Relations

Stephen E. Fienberg; Michael M. Meyer; Stanley S. Wasserman

Journal of the American Statistical Association, Vol. 80, No. 389. (Mar., 1985), pp. 51-67.

Stable URL:

<http://links.jstor.org/sici?sici=0162-1459%28198503%2980%3A389%3C51%3ASAOMSR%3E2.0.CO%3B2-%23>

Journal of the American Statistical Association is currently published by American Statistical Association.

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/astata.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact support@jstor.org.

Statistical Analysis of Multiple Sociometric Relations

STEPHEN E. FIENBERG, MICHAEL M. MEYER, and STANLEY S. WASSERMAN*

Loglinear models are adapted for the analysis of multivariate social networks, a set of sociometric relations among a group of actors. Models that focus on the similarities and differences between the relations and models that concentrate on individual actors are discussed. This approach allows for the partitioning of the actors into blocks or subgroups. Some ideas for combining these models are described, and the various models and computational methods are applied to the analysis of data for a corporate interlock network of the 25 largest organizations in Minneapolis/St. Paul and for a classic network of 18 monks in a cloister. The computational techniques all involve variations on the standard iterative proportional-fitting algorithm used extensively in the analysis of multidimensional contingency tables.

KEY WORDS: Loglinear model; Directed graph; Social network; Sociometric data; Iterative proportional fitting; GLIM model.

1. INTRODUCTION

Sociometric relations are typically defined for a set of social actors. A social network is a construct describing these actors and the various relations that exist among them. As used in the social sciences, actors have been individuals, organizations, cities, or even nation states; relations have ranged from kinship or friendship to transfers of scarce resources to corporate board-of-directors interlocks.

Moreno (1934) was the first social scientist to study individual networks in a systematic manner, and he was apparently the first network researcher to use mathematics. Much of his terminology, including such phrases as "sociogram," "sociomatrix," and "sociometric test," is still in use today. Festinger (1949) and Katz (1947, 1953, 1955) developed Moreno's ideas, focusing on matrix representations of sociometric data, the popularity of actors, mutuality of relationships in social groups, and even the representation of interpersonal relations as a stochastic process (see Katz and Proctor 1959). Formal graph theory, as reviewed here in Section 2, was introduced to social network research by Cartwright and Harary (1958), in an attempt to quantify the social psychological theories of Heider (1958).

Since these pioneering efforts, sociologists, social psychologists, and social anthropologists have repeatedly used the social-network paradigm. Davis and Leinhardt (see Davis 1970) scanned the "sociometry" literature and found nearly 900 examples of social networks from diverse small groups. Since 1970, social network analysis has grown rapidly in popularity. Leinhardt (1977) presented a collection of 24 previously published papers that provide an historical perspective on social network analysis, and a collection of papers in a volume edited by Holland and Leinhardt (1979) summarized the state of the art as of about 1975. Burt (1980) discussed more recent sociological developments, Wasserman (1978) reviewed alternative mathematical models for small group behavior, and Frank (1981) summarized some of the statistical theory on random graphs. Almost none of this research on the analysis of social networks has appeared in statistical journals, with the exception of some of the work by Katz and the paper by Wasserman (1980). There are just a few papers with substantial statistical content.

In a landmark statistical paper for network analysis, Holland and Leinhardt (1981) proposed an exponential family of probability distributions for the analysis of a single sociometric relation. Fienberg and Wasserman (1981a) discussed simple computational procedures for fitting these models and proposed some extensions to model networks in which the actors fall into natural subgroups. These distributions include parameters that relate characteristics of individual actors (e.g., popularity) to differential rates for entering into or severing sociometric relations. In Fienberg et al. (1981), we described a related class of models for multiple relations, extending Holland and Leinhardt's family to more than one relation by focusing on the associations among the relations rather than on influences of individual actors. Here we bring these two types of analyses together and present some "combined models" for the analysis of multivariate directed graphs. These models incorporate actor and subgroup parameters, and quantities to measure the degree of interrelatedness of the different relations.

Methods to study a multivariate-directed graph, which focus solely on the relations and ignore individual social actors, are forms of *macroanalysis*. Data for such analyses consist of aggregate counts of the different structural patterns, which occur within the network. The methods for studying local structure in a network by using the *triad census* (Holland and Leinhardt 1975; Wasserman 1977) can be labeled macroanalytic. Alternatively, we could study the attributes of the actors and how these attributes affect the existing ties between them. Such a study is a *microanalysis*, and it promises a more fine-grained investigation.

Both the macroanalysis and microanalysis approaches have

* Stephen E. Fienberg is Professor of Statistics and Social Science, Carnegie-Mellon University, Pittsburgh, PA 15213; Michael M. Meyer is Assistant Professor, Department of Statistics, University of Wisconsin, Madison, WI 53706; and Stanley S. Wasserman is Associate Professor, Departments of Psychology and Mathematics, University of Illinois, Champaign, IL 61820. This research was supported by National Science Foundation Grant SES80-08573 to the University of Minnesota and by Office of Naval Research Contract N00014-80-C-0637 to Carnegie-Mellon University. The authors thank David Hinkley, Paul Holland, Samuel Leinhardt, and Peter Marsden for helpful discussions, which are reflected in various ways in this article. Special thanks are due to Joseph Galaskiewicz for posing several of the questions that led to the work reported here, and to Jack Litzau for computational assistance. The referees and associate editors provided helpful comments on earlier drafts of this material.

substantive value. A macroanalysis of a group centers on the global structure of its relations, asking questions such as the following: Which relation exhibits the strongest "reciprocity" or is most likely to have symmetric flows? Are there any "multiplex" patterns, flows of different relations in the same direction? Are there any patterns of "exchange," in which a flow in one direction for one relation is reciprocated by a flow in the opposite direction for a different relation? Are there any higher order interactions, involving three or more flows for two or more relations?

A microanalysis of a group is a local study, turning attention to the level at which data are actually gathered. Most microanalyses have been limited to groups with data on just a single relation. The primary concern of such studies is the individual group member: Which actors have the most prestige or popularity? Which actors are involved in many relations, which in few? Do actors enter into mutual, symmetric relationships at different rates? Such questions, while concerned with individual actor effects, are often answered by examining dyadic or triadic relationships.

As an example, we consider the now classic study of 18 monks in an isolated American monastery, conducted by Sampson (1969) and partially analyzed by Holland and Leinhardt (1981), Breiger (1981), Reitz (1982), and many others. Sampson studies four types of relations: Affect, Esteem, Influence, and Sanction. Actors were asked to give three positive choices [e.g., which three brothers do you like best (positive affect)] and three negative choices [e.g., which three brothers are you most antagonistic towards (negative affect)] for each of the four types. In this way, data were gathered on eight relations: (a) Like and (b) Antagonism (Affect), (c) Esteem and (d) Disesteem, (e) Influence and (f) Negative Influence, and (g) Praise and (h) Blame (Sanction).

We define

$$x_{ijr} = 1, \quad \text{if actor } i \text{ chooses actor } j \text{ on relation } r \\ = 0, \quad \text{otherwise,}$$

$$i, j = 1, 2, \dots, 18, \quad r = 1, 2, \dots, 8,$$

and arrange these data into eight binary sociomatrices, $\mathcal{X} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_8\}$, each of dimensions 18×18 . Versions of these arrays are given in Figure 1, where the rows and columns have been permuted to reflect constructed subgroupings of the actors. Since the 1s, 2s and 3s in Figure 1 refer to order of choices, we set all nonzero entries equal to 1 to obtain binary sociomatrices.

Holland and Leinhardt (1981) used the Like relation to illustrate their new methods. Other researchers (White et al. 1976; Breiger et al. 1975) have studied all the relations, but in nonstatistical attempts to aggregate the 18 monks in other substantively meaningful ways. In later sections of this article, we analyze a version of the network, which aggregates over positive and negative effects, searching for both macro- and micro-models that provide good statistical descriptions of the relationships among the actors.

Most sociometric research, both empirical and mathematical, is preoccupied with overly simplistic descriptions of group structure. This is very apparent in Burt's (1980) review. The

goal of this article is to build upon the ideas of Holland and Leinhardt to develop models for the simultaneous macro- and micro-analysis of multiple relational networks. These models aid in the formulation and testing of theories concerning group dynamics. In the next section we review Holland and Leinhardt's model and our extensions of it, and in Section 3 we illustrate these ideas in an analysis of a 1976 corporate interlock network from the Twin Cities (Minneapolis/St. Paul). We emphasize the many substantive findings that can be obtained from this form of statistical modeling. In Section 4 we present several models for the analysis of data from multivariate directed graphs, and we conclude by demonstrating these ideas on Sampson's network.

2. BACKGROUND: MODELS FOR SINGLE RELATIONAL DATA

A directed graph, or digraph, consists of a set of g nodes and sets of directed arcs or "choices" connecting pairs of nodes. Digraphs are natural mathematical representations of social networks, where the nodes represent individuals, organizations, or other social actors and the arcs represent relations: directed attitudes, feelings such as friendship, or transfers of goods or information. A digraph is frequently summarized by a set of $g \times g$ sociomatrices $\{\mathbf{X}_r\}$, one for each of the R defined relations. The g diagonal terms of each sociomatrix, X_{iir} , are defined to be zero.

First consider a digraph with a single relation, $R = 1$. The row total, X_{i+} , is referred to as the out-degree of node i , and the corresponding column total, X_{+i} , as the in-degree of node i . A matrix \mathbf{x} can be thought of as the realization of a matrix of random variables, \mathbf{X} , where we assume that the $\binom{g}{2}$ pairs or dyads,

$$D_{ij} = (X_{ij}, X_{ji}), \quad i < j,$$

are independent bivariate random variables, with $2^2 = 4$ possible realizations when we study dyads rather than individual actors:

$$D_{ij} = (1, 1): \text{mutual} \\ = (1, 0) \text{ or } (0, 1): \text{asymmetric} \\ = (0, 0): \text{null.}$$

Note that the two asymmetric realizations cannot be distinguished unless the actors are labeled.

A multivariate-directed graph, or multigraph, is described by a collection of random sociomatrices $\mathcal{X} = \{\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_R\}$. We assume that the $\binom{g}{2}$ dyads,

$$\mathbf{D}_{ij} = \begin{pmatrix} X_{ij1}, X_{ji1} \\ X_{ij2}, X_{ji2} \\ \vdots \\ X_{ijR}, X_{jiR} \end{pmatrix}, \quad i < j,$$

are independent $2R$ -variate random variables with 2^{2R} possible realizations.

For both digraphs and multigraphs, the assumption that the dyads are independent random variables is crucial to the independent-dyadic-choice models described here. There is some

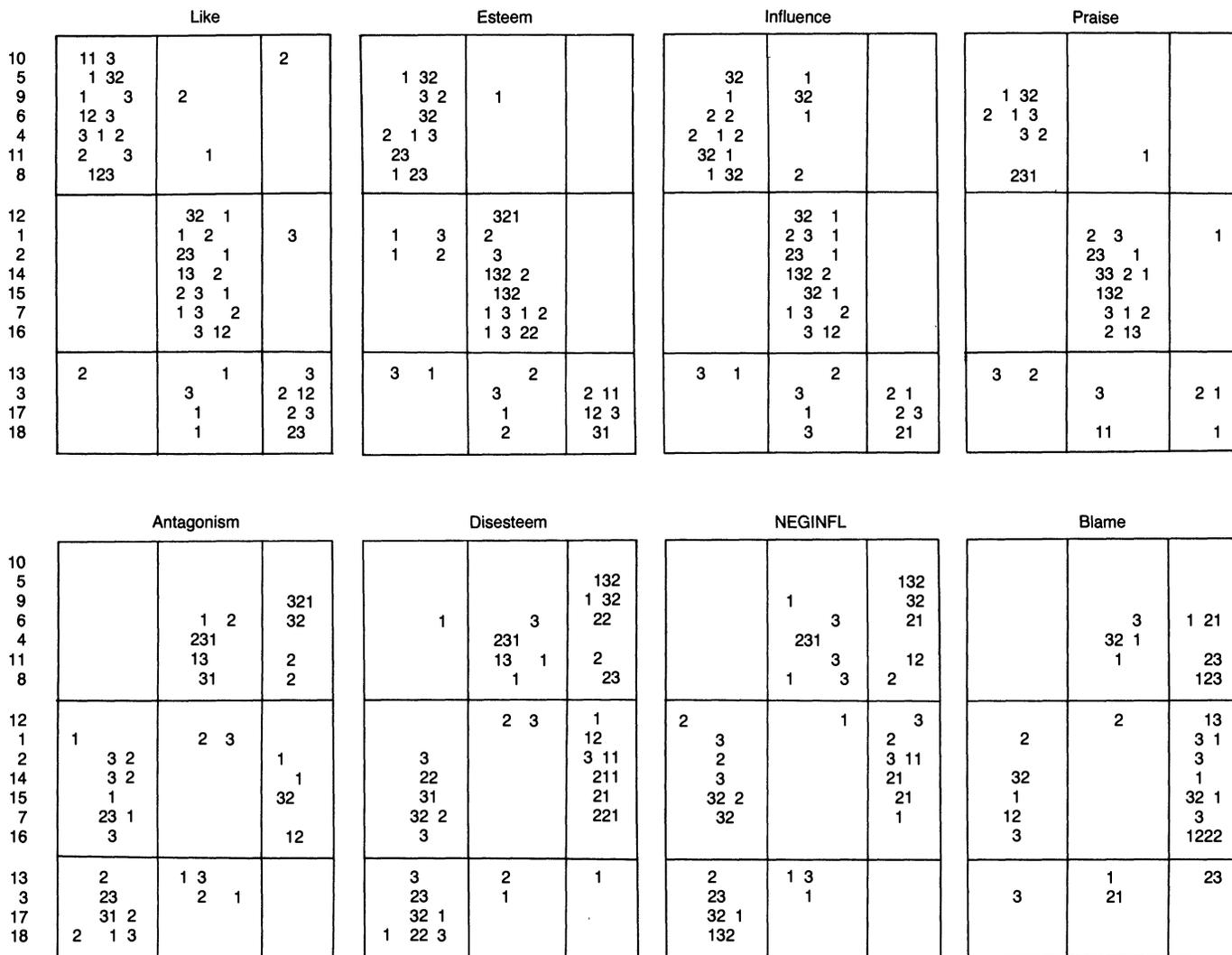


Figure 1. Sampson's (1969) Data.

weak evidence from social psychology (see Heider 1958) that individuals consider the current state of all $g - 1$ dyads they are involved in when deciding whether or not to initiate new choices or break old ones. Nonetheless, there is no conclusive evidence of this empirical tendency. In networks involving actors other than people, it is not at all clear whether a dyadic relationship between a pair of actors is independent of all other dyads. What is true, however, is that a relaxation of this assumption is very difficult to implement statistically. Frank and Strauss (1983) have taken a few steps in this direction using theory from Markov random fields, but the technical breakthrough that would allow us to consider the probability of specific dyad states conditional on many other dyads has yet to be achieved. For now, we choose to view these models as local approximations, at least an order of magnitude superior to independent-choice models that incorporate no dyadic structure at all.

Holland and Leinhardt (1981) introduced a class of models, labeled p_i , to model microbehavior in a social group, on which only one relation has been defined. We now describe these models and explain how their analysis can be accomplished

by using standard computational approaches to the analysis of loglinear models for categorical data. We then outline some extensions of these models that allow for grouping of individual actors. Further details can be found in Fienberg and Wasserman (1981a). In Section 4 we extend this approach to the analysis of multigraph data.

Consider a network of g nodes and a single relation, and represent the sociomatrix \mathbf{X} as a four-dimensional $g \times g \times 2 \times 2$ cross-classification $\mathbf{Y} = (Y_{ijkh})$, where the subscripts i and j refer to the two actors in a dyad, and k and h refer to the dyad state:

$$Y_{ijkh} = 1, \text{ if } D_{ij} = (X_{ij}, X_{ji}) = (k, h) \\ = 0, \text{ otherwise.} \tag{2.1}$$

For example, $Y_{ij11} = 1$ if D_{ij} is a mutual dyad. Note that the 2×2 tables $\mathbf{Y}_{ij}(i \neq j)$ contain one 1 and three 0s. Furthermore, $Y_{ijkh} = Y_{jihk}$, and the marginal totals of these 2×2 tables correspond to indicator variables for X_{ij} and X_{ji} . Because each margin is either (0, 1) or (1, 0), the interior of the table is completely determined by its marginal totals.

We denote a realization of \mathbf{Y} by $\mathbf{y} = (y_{ijkh})$ and let π_{ijkh} be the probability of the observation (k, h) for the dyad (i, j) , where

$$\sum_{k,h} \pi_{ijkh} = 1, \tag{2.2}$$

and we define $\mu_{ijkh} = \log \pi_{ijkh}$. The Holland–Leinhardt p_1 class of models is as follows:

$$\begin{aligned} \mu_{ij00} &= \lambda_{ij}, \\ \mu_{ij10} &= \lambda_{ij} + \alpha_i + \beta_j + \theta, \\ \mu_{ij01} &= \lambda_{ij} + \alpha_j + \beta_i + \theta, \\ \mu_{ij11} &= \lambda_{ij} + \alpha_i + \alpha_j + \beta_i + \beta_j + 2\theta + \rho_{ij}, \end{aligned} \tag{2.3}$$

where

$$\sum_{i=1}^g \alpha_i = \sum_{j=1}^g \beta_j = 0, \tag{2.4}$$

and $\rho_{ij} \equiv \rho$.

The substantive meanings associated with these parameters are described at length in Holland and Leinhardt (1981) and Reitz (1982). In short, the $\{\alpha_i\}$ are often interpreted as effects measuring individual expansiveness or productivity, the $\{\beta_j\}$ are interpreted as individual popularity or attractiveness effects, and θ is a parameter interpreted as measuring overall choice volume (as represented by the presence of directed arcs). The parameter ρ reflects the rate at which the pairs of actors engage in mutual, reciprocated relationships and is often described as a measure of mutuality or reciprocity. Throughout this article, we consider various generalizations and extensions of the p_1 model, and we restrict the use of θ to choice-like parameters and the use of ρ to reciprocity- or mutuality-like parameters.

The sufficient statistics for the parameters of p_1 are easily expressed as margins of \mathbf{y} :

$$\begin{aligned} \frac{1}{2}y_{++11} &= M, && \text{number of mutuals,} \\ y_{i+1+} &= x_{i+}, && \text{out-degree of node } i, \\ y_{+j1+} &= x_{+j}, && \text{in-degree of node } j, \\ y_{++1+} &= x_{++}, && \text{total number of choices.} \end{aligned} \tag{2.5}$$

Through the use of the full \mathbf{y} array and its redundancies, one can show that fitting p_1 to the \mathbf{x} array is equivalent to fitting the “no three- or four-factor” interaction loglinear model to \mathbf{y} . A proof of this equivalence was given in Meyer (1981). Thus we can fit p_1 to data by using the standard iterative proportional fitting procedure (IPFP) applied to \mathbf{y} . Furthermore, the special cases of p_1 , listed in table 1 of Holland and Leinhardt (1981), all have equivalent loglinear models for \mathbf{y} , and thus they can also be fit using the standard IPFP. The equivalent models are given in table 2 of Fienberg and Wasserman (1981a). To compute likelihood ratio goodness-of-fit statistics we divide the usual statistic computed on \mathbf{y} by 2 to adjust for the duplication (i.e., $y_{ijkh} = y_{jihk}$).

An important generalization of p_1 starts with Equations (2.3) with constraints (2.4) and further postulates that

$$\rho_{ij} = \rho + \rho_i + \rho_j, \quad i < j, \tag{2.6}$$

where the $\{\rho_i\}$ are normalized to sum to zero. The effect of reciprocity now depends additively on the individual actors in a dyad, and the $\{\rho_i\}$ measure the rates at which actors are likely to enter into mutual, symmetric relationships. The allowance for individual mutuality parameters is substantively important in many circumstances, particularly those in which some, but not all, actors can use symmetric relationships to their advantage. This model provides one goodness-of-fit test for p_1 (see Fienberg and Wasserman 1981b), since it contains p_1 as a special case, when $\rho_1 = \rho_2 = \dots = \rho_g = 0$. Note, however, that this model still assumes independence of dyads so it can not be used to detect departures from this crucial assumption. Furthermore, the distribution of the log likelihood ratio test statistic comparing p_1 to (2.6) may not necessarily be asymptotically chi-squared with $g - 1$ degrees of freedom.

The natural asymptotic setting for this type of network problem is one in which the size of the network is allowed to increase, that is, $g \rightarrow \infty$. Here the sample size N is the number of dyads, $\binom{g}{2}$, that is, $N = O(g^2)$. We are interested in the asymptotic distribution of a likelihood ratio test statistic with $df = O(N^{1/2})$, and standard asymptotic theory is not applicable here. Haberman (1981) discussed aspects of this problem. One way to sidestep this unsolved theoretical problem is to recast our models so that the number of parameters does not necessarily grow with the sample size.

We now describe a variant on p_1 for single relational sociometric data that assumes that the g actors have been partitioned into K subgroups based on external, extrarelational information. Questions asked in such situations include: How likely is it that actors in one subgroup relate to actors in other subgroups, and How structurally similar are actors in a given subgroup? We label the subgroups G_1, G_2, \dots, G_K , where the partition of actors is mutually exclusive and exhaustive, and assume that subgroup G_k contains g_k actors, such that $g_1 + g_2 + \dots + g_K = g$. The approach we propose should be contrasted with data-dependent grouping techniques, such as those used in cluster analysis. For example, White et al. (1976) (also see Breiger 1981) aggregated the 18 monks from Sampson’s cloister into 3 “blocks” or subgroups, containing $g_1 = 7, g_2 = 7, g_3 = 4$ actors. This aggregation is reflected in Figure 1, where the rows and columns of the \mathbf{X} matrices have been rearranged so that the first seven rows and columns refer to actors in G_1 , and so forth. Breiger et al. (1975) constructed a slightly different partition. We note that these partitions, called “blockmodels,” were accomplished by data-dependent grouping algorithms that attempt to collect together all actors that are “structurally equivalent” (see Lorrain and White 1971).

We modify Equations (2.3) by introducing inter- and intra-subgroup choice and reciprocity parameters:

$$\begin{aligned} \mu_{ij00} &= \lambda^{(ij)} \\ \mu_{ij10} &= \lambda^{(ij)} + \theta^{(rs)} \\ \mu_{ij01} &= \lambda^{(ij)} + \theta^{(sr)} \\ \mu_{ij11} &= \lambda^{(ij)} + \theta^{(rs)} + \theta^{(sr)} + \rho^{(rs)}, \end{aligned} \tag{2.7}$$

$i \in G, \text{ and } j \in G_s.$

The parameters $\{\theta^{(rs)}\}$ are choice effects, and the $\{\rho^{(rs)}\}$ are

reciprocity effects. The parameters $\{\lambda^{(ij)}\}$ are included to insure that the y_{ijkh} sum to 1 for each dyad. One special case of the subgroup model (2.7) sets $\rho^{(rs)} = 0$ for all r and s . Holland and Leinhardt (1981) noted that if we further define

$$\pi^{(rs)} = P\{X_{ij} = 1 \mid i \in G_r \text{ and } j \in G_s\}, \quad (2.8)$$

then, in this special case,

$$\theta^{(rs)} = \log\left(\frac{\pi^{(rs)}}{1 - \pi^{(rs)}}\right) = \text{logit}(\pi^{(rs)}). \quad (2.9)$$

A second special case of (2.7) is also a special case of p_1 in which we have a simple additive model for $\theta^{(rs)}$. All actors in subgroup G_r have a common α , $\alpha^{(r)}$, and a common β , $\beta^{(r)}$. We set

$$\begin{aligned} \theta^{(rs)} &= \theta + \alpha^{(r)} + \beta^{(s)} \\ \rho^{(rs)} &= \rho. \end{aligned} \quad (2.10)$$

This model is equivalent to p_1 if $K = g$, and it is a simplification, in the sense that we reduce the number of α 's (and β 's) from $g - 1$ to $K - 1$.

For details about these and other generalizations and specializations of p_1 , and for comments on fitting these subgroup models to single relational data, see Fienberg and Wasserman (1981a). In Section 4 we give a multivariate generalization of this model.

3. ANALYSIS OF A SINGLE RELATION IN A CORPORATE NETWORK

To illustrate these models, and to present some additional methods, we consider 1976 data on a network of the 25 largest publicly owned corporations headquartered in the Twin Cities of Minneapolis and St. Paul. A firm is included in the network if it is among *Fortune* magazine's 500 largest industrials, 50 largest commercial banks, 50 largest life-insurance companies, 50 largest financial companies, 50 largest retailers, 50 largest transportation companies, and 50 largest utilities. These companies are listed in Table 1, along with their ranks and location.

A preliminary analysis and thorough discussion of this network (shown in Figure 2) was given by Galaskiewicz and Wasserman (1981). An arc (or a corporate interlock) exists from firm i to firm j if an officer of firm j is on the corporate board of directors of firm i . An interesting feature of this network is the exclusion of dyadic interactions in which the two firms of the dyad have the same Standard Industrial Code. These "competitive" dyads have been excluded because of Securities and Exchange Commission anti-trust regulations that prevent interlocks between firms in the same industry. There are 27 of these structurally-zero dyads.

A variety of models was fitted to two versions of this network. One version included all 25 firms, and the other included only 20 firms, excluding four firms that do not interact with the others (have zero in-degrees and out-degrees)—American Hoist and Derrick, IDS, Gamble-Skogmo, and North Central Airlines—and a firm, Land O'Lakes, that is a cooperative and hence not strictly publicly owned. The calculation of degrees of freedom (df), discussed in the Appendix, is tricky because of the structural zeros and the zero in-degrees and out-degrees. In general we follow an approach similar to that suggested by

Table 1. Twin Cities Corporate-Network Actors

Actor	Fortune Rank (1976)	City
<i>Manufacturers</i>		
Minnesota Mining & Manufacturing (3M)	56	St. Paul
Honeywell	67	Minneapolis
General Mills	84	Minneapolis
Control Data	170	Minneapolis
Pillsbury	173	Minneapolis
Land O'Lakes	180	Minneapolis
International Multifoods	233	Minneapolis
Bemis	318	Minneapolis
Peavy	361	Minneapolis
Heorner-Waldorf	382	St. Paul
American Hoist and Derrick	434	St. Paul
Economics Laboratory	500	St. Paul
<i>Commercial Banks</i>		
Northwest Bankcorporation	18	Minneapolis
First Bank System	20	Minneapolis
<i>Life Insurance Companies</i>		
Minnesota Mutual Life Insurance	41	St. Paul
Northwestern National Life Insurance	42	Minneapolis
<i>Diversified Financial Companies</i>		
St. Paul Companies	20	St. Paul
Investors Diversified Services (IDS)	28	Minneapolis
<i>Retailing Companies</i>		
Dayton Hudson	20	Minneapolis
Gamble-Skogmo	22	Minneapolis
<i>Transportation Companies</i>		
Burlington Northern R.R.	10	St. Paul
Northwest Orient Airlines	18	St. Paul
North Central Airlines	48	Minneapolis
Soo Line R.R.	49	Minneapolis
<i>Utilities</i>		
Northern States Power	28	Minneapolis

Bishop et al. (1975, pp. 115–116). In Table 2, we report likelihood ratio (G^2) statistics and degrees of freedom for just two models. We note that goodness-of-fit statistics such as G^2 do not have the usual asymptotic chi-squared distributions in this setting (e.g., see Holland and Leinhardt 1981 and Haberman 1981). Thus we choose to use them here in a rough and somewhat informal manner. Since G^2/df is approximately equal to 1 for $g = 25$ and $g = 20$, the very simple model with a single parameter appears to provide a reasonable description of both versions of the data. This implies that the actors in neither version exhibit differential productivity or attractiveness and that there is no tendency toward reciprocity. If we accept this simple model as being appropriate, we are taking the elements in X to be independent identically distributed Bernoulli random variables with $p = P\{X_{ij} = 1\}$ and log odds ratio $\theta = \log(p/(1 - p))$. Maximum likelihood estimates (MLE's) of θ are -2.49 ($g = 25$) and -2.01 ($g = 20$). This yields $\hat{p} = .0906$ ($g = 25$) and $\hat{p} = .1553$ ($g = 20$). If we now compare p_1 to this simple choice model, we have $\Delta G^2 = G_2^2 - G_1^2 = 138.97$ with $\Delta \text{df} = \text{df}_2 - \text{df}_1 = 353$ for the $g = 25$ case, and $\Delta G^2 = 93.57$ with $\Delta \text{df} = 165$ for $g = 20$. In either case $\Delta G^2/\Delta \text{df}$ is much less than 1. Nonetheless, the inapplicability of standard asymptotic theory does not allow us to conclude that the extra parameters of the p_1 model are unnecessary.

How then can we get a handle on the appropriateness of dyadic models such as p_1 ? Our approach is to focus on natural partitions of the firms based on extrarelational information, and we compare the fit of p_1 -like models for nested partitions.

		Large Mpls								Large St. Paul			Small Mpls					Small St. Paul								
		2	3	4	5	6	7	13	14	19	20	1	17	21	22	8	9	16	18	23	24	25	10	11	12	15
Large Mpls	2	X	M	X	1					M	X	1	M		X											
	3	M	X		X	X	X	M		1	X		M		X	X									X	
	4	X		X											1											
	5		X		X		X									X	1									
	6		X		X																					
	7		X		X		X			1						X										
	13	1	M		1			X	X					1		1	1	M							1	
14	1						X	X	M		1	M	1													
19	M			1					M	X	X															
20									X	X						X	X								X	
Large St Paul	1	X	X								X				X								X		X	
	17											X	1	1												
	21		M						M				X							X	X					
	22	M									1		X						X							
Small Mpls	8	X	X								X				X								X		X	
	9	X			X		X			X					X										X	
	16		1					M		X						X									X	
	18												X				X									
	23												X		X			X								
	24						1		1			X	X		1	1			X	X						
25						1		1											X							
Small St Paul	10										X				X								X			
	11	X									X				X								X	X		
	12										X				X								1	X	X	
	15									X		1	1	1		X									X	

Figure 2. Twin Cities Corporate Network Organized by Size and Location: (X) Structural Zero; (M) Mutual Relationship; (1) Asymmetric Relationship. The network contains 8 mutual and 32 asymmetric choices. $X_{++} = 48$.

Suppose we have two possible mutually exclusive and exhaustive partitions of a set of g actors, $\mathcal{G} = \{G_1, G_2, \dots, G_K\}$ and $\mathcal{H} = \{H_1, H_2, \dots, H_L\}$, such that $K < L$, and the G 's are unions of the H 's. For example, let $g = 6$, and define $G_1 = \{1, 2, 3\}$, $G_2 = \{4, 5, 6\}$, $H_1 = \{1, 2\}$, $H_2 = \{3\}$, and $H_3 = \{4, 5, 6\}$; then $G_1 = H_1 \cup H_2$ and $G_2 = H_3$. Thus \mathcal{G} is an aggregation of \mathcal{H} .

We consider whether or not to further aggregate the actors into K subgroups, assuming that the actors are already partitioned into L subgroups; that is, can we combine some of the L existing subgroups to form K larger ones? Note that if $L = g$, then we ask whether or not we should do any aggregation at all. We test

$$H_0 : p_1 \text{ applied to } K \text{ subgroups is appropriate}$$

versus

$$H_A : p_1 \text{ applied to } L \text{ subgroups is appropriate.}$$

The version of p_1 applied to subgroups is given by Equations (2.7) and (2.10). In terms of the model parameters, there are $L - 1$ each of the $\alpha^{(i)}$ and $\beta^{(i)}$ effects under H_A and $K - 1$ each under H_0 . The α 's and β 's for the subgroups that are aggregated under H_0 are equated. Since H_0 is a special case of H_A , if we assume that the model under H_A is correct, then the conditional likelihood ratio statistic $G^2(H_0 | H_A) = G^2(H_0) - G^2(H_A)$, with $g(g - 1) - 2K - [g(g - 1) - 2L] = 2(L - K)$ degrees of freedom can be used to test H_0 versus H_A . If $L = g$, then the test statistic has $2(g - K)$ degrees of freedom. Again we note that these df and the likelihood ratio statistics should be used with caution. It is best just to interpret the conditional likelihood ratio statistics, which have $2(L - K)$ df. Here, because of the aggregation, it is more reasonable to compare these statistics to the corresponding chi-square reference distributions than it is in the original setting where we

are concerned with the fit of p_1 and its variants with no aggregation.

For the 1976 Twin Cities corporate network, we focus on three partitions using the information in Table 1: $\mathcal{G}_1 = \{G_{11} = \text{Minneapolis firms}, G_{21} = \text{St. Paul firms}\}$; $\mathcal{G}_2 = \{G_{12} = \text{large firms}, G_{22} = \text{small firms}\}$; and $\mathcal{H} = \{H_1 = \text{large Minneapolis firms}, H_2 = \text{large St. Paul firms}, H_3 = \text{small Minneapolis firms}, H_4 = \text{small St. Paul firms}\}$. "Size" of a firm is determined by the *Fortune* ratings: "large" firms rank among the larger 250 manufacturing firms, or the 25 larger banks, life insurance companies, and so forth. Note that both \mathcal{G}_1 and \mathcal{G}_2 are aggregations of \mathcal{H} . The $25 \times 25 \times 2 \times 2$ y array, aggregated to a $4 \times 4 \times 2 \times 2$ array to reflect the \mathcal{H} partition, is given as Table 3 in the form of 16×2 tables. In Table 3, for each of the 12 intergroup tables, the (1, 1) entry represents the number of null relationships, the (2, 2) entry represents the number of mutual relationships, and the (1, 2) and (2, 1) entries represent the asymmetric relationships. For the four intragroup tables the counts of null and mutual relationships in the (1, 1) and (2, 2) cells are doubled and the count of asymmetric relationships is duplicated and given in both the (1, 2) and (2, 1) cells. Thus when Table 3 is viewed as an 8×8 table, the entries on the diagonal are doubled and those below the diagonal are duplicated in the symmetrically placed cells above the diagonal.

Table 2. Goodness of Fit of p_1 Model for Corporate Network Data

Model	$g = 25$		$g = 20$	
	G^2	df	G^2	df
$p_1(\theta, \rho, \{\alpha_i\}, \{\beta_j\})$	186.69	192	182.89	176
$p_1(\theta, \rho = \alpha_i = \beta_j = 0)$	324.66	545	276.46	341

Table 3. 1976 Twin Cities Corporate Network Relations Aggregated Into Four Subgroups Based on Location and Size

	Large				Small				
	Minneapolis		St. Paul		Minneapolis		St. Paul		
Large									
Minneapolis	54	7	31	0	53	5	35	2	H_1
	7	8	4	3	4	1	1	0	
St. Paul	31	4	6	3	25	0	11	3	H_2
	0	3	3	0	0	0	0	0	
Small									
Minneapolis	53	4	25	0	38	2	25	0	H_3
	5	1	0	0	2	0	0	0	
St. Paul	35	1	11	0	25	0	8	1	H_4
	2	0	3	0	0	0	1	0	
	H_1		H_2		H_3		H_4		

NOTE: See the description of the table in the text.

The hierarchy in Table 4 shows the three aggregations and gives the associated likelihood and conditional likelihood ratio statistics for testing the significance of aggregations. Note that $G^2(\mathcal{K} | p_1) = 401.80 - 186.69 = 215.11$ is less than the corresponding difference in df, 346, so that one might argue that aggregating the 25 actors into 4 subgroups is strongly supported by a formal test of goodness of fit. Formal inferences here are nonetheless complicated as the standard chi-squared reference distribution is not appropriate for this conditional test. The statistic $G^2(\mathcal{K})$ is clearly small relative to the nominal df, however, and simplicity of the \mathcal{K} aggregation is so desirable that we find the \mathcal{K} version of p_1 to be a very attractive model. The statistics $G^2(\mathcal{G}_1 | \mathcal{K}) = 59.82$ and $G^2(\mathcal{G}_2 | \mathcal{K}) = 124.01$, each with 4 df, yield p -values less than 10^{-4} , so further aggregation is not advisable.

The foregoing analysis allows us to make a more formal statistical assessment about the appropriateness of the p_1 model as opposed to the simple choice model. The two 4-df comparisons involving nested partitions, of \mathcal{G}_1 within \mathcal{K} and of \mathcal{G}_2 within \mathcal{K} , lend strong support for the conclusion that there are differences in the values of choice parameters between Minneapolis and St. Paul firms, and between large and small firms. If such differences exist, then we can conclude that the simple choice model is not appropriate and that a dyadic model of the p_1 form is a distinct improvement. We are still left with the question of whether aggregation by size and location within the p_1 structure is justified, but formal theory does not help us answer this question.

There is one substantial advantage in using aggregated versions of these models. Besides the ease with which the maximum likelihood cell estimates can be computed (we need only a $K \times K \times 2 \times 2$ table, where K is usually quite a bit smaller than g), the standard χ^2 distributions are more appropriate as reference distributions for the resulting test statistics. This is because the number of parameters ($2K$ with the p_1 -subgroup model) is fixed and does not increase in the limit, as $g \rightarrow \infty$. The problems that arise in testing when using models with parameters for each actor (see Haberman 1981) are fortunately attenuated when actors are aggregated. Of course, in the corporate interlock data this form of alternative asymptotics is

only, at best, an approximation, as the actors were selected from the *Fortune* 500 or 50 lists, and this selection criterion automatically requires that $g \leq 800$. But with $K = 4$ and $g = 25$, the approximation seems to be a reasonable one.

In the following section we generalize this approach to the case of multiple relations.

4. MODELS FOR MULTIPLE-RELATION DATA

We now turn our attention to networks of actors on which several relations are defined. We discuss three types of models: (a) models with neither actor nor group parameters, (b) models with only group parameters, and (c) models with both actor and group parameters. The first type is a family of models for the macroanalysis of the multiple relations that ignores any differences between actors. These models were briefly described in Fienberg et al. (1981) and were used implicitly by Galaskiewicz and Marsden (1978) to study resource flows between organizations in a midwestern community. Models for multiple-relation data were first suggested by Davis (1968).

The most useful models for multiple relations are those that include parameters to reflect different choice tendencies of the actors, particularly when they have been partitioned into groups. If each group is a singleton, then we have a different set of parameters for each actor; however, in practice this is likely to be a very large number. Thus the assumption of a specific partition, chosen as a consequence of extrarelatational information, allows us to parsimoniously limit the number of parameters and (as is the case with single relational data) use standard χ^2 asymptotic distributions for testing.

The last type of model is a generalization of the family of models for multiple relational data sets in which the actors have been partitioned into mutually exclusive and exhaustive groups. The assumption that all actors in a specific group relate to actors in other groups and to other actors in the same groups in identical ways may not always be the case. There may be subtle individual differences among the actors in a subgroup. Thus, the third type of models allows us to add individual actor parameters to study these differences to the second type of models with just group parameters.

We conclude this article by illustrating the fitting and the interpretation of these models in the context of Sampson's network of 18 monks. In this example, we have four positive relations and four negative relations, and three subgroups, empirically determined by the use of clustering algorithms. Although the subgroups were formed using the data to be analyzed in our illustration, we proceed as if the subgroups were preformed.

Table 4. Goodness of Fit of Aggregated Versions of p_1 Model for Corporate Network Data

Aggregation	G^2	df	ΔG^2	Δdf
p_1 —no aggregation; 25 actors	186.69	192		
\mathcal{K} —aggregation by size and location; $L = 4$	401.80	538		
\mathcal{G}_1 —aggregation by location; $K_1 = 2$	461.62	542	59.82	4
\mathcal{G}_2 —aggregation by size; $K_2 = 2$	525.63	542	124.01	4

4.1 Models for the Macroanalysis of Multiple Relations

In order to model the macro-aspects of multigraphs, we need to develop a notation for the 2^{2R} possible realizations of the $\{D_{ij}\}$ and a representation for the table of summary counts of these realizations obtained by adding across dyads. Since these models assume no individual actor differences, the sufficient statistics for the model parameters are margins of this 2^{2R} table.

We consider first the case $R = 2$. Here there are two relations between each pair of actors. For the purposes of this discussion we label the relations as *like* and *admire*. We can summarize the pairwise relationships in a 2^4 table of counts with the four factors as follows:

1. Actor one likes actor two (yes and no).
2. Actor one is liked by actor two (yes and no).
3. Actor one admires actor two (yes and no).
4. Actor one is admired by actor two (yes and no).

Consider the following pair of actors. Actor A likes actor B but does not admire him/her. Actor B admires Actor A and likes him/her. When this pair of actors is tabulated, they would add a count of one to the (yes, yes, yes, no) cell. If we reverse the role of actors, then the same pair will also add a count of one to the (yes, yes, no, yes) cell. We refer to this representation of the data as the *w*-array and denote the entries as $\{w_{ii'jj'}\}$, where the subscripts are paired—for example, i and i' for the relation *like*, and each subscript takes the values yes and no. Obviously all of the relationships are double counted in this representation. The double counting results in the four diagonal cells (with $i = i'$ and $j = j'$) being doubled and the 12 off-diagonal cells being six symmetric pairs of exact duplicates.

Another way of looking at the data is by classifying the relationship. For each of the two relations, *like* and *admire*, there are three *types* of relationship:

1. Mutual (denoted by M)—for example, where each actor likes the other
2. Asymmetric (denoted by A)—for example, where one actor likes the other but the relationship is not reciprocated
3. Null (denoted by N)—for example, where neither actor likes the other.

When we put the two relations, *like* and *admire*, together, there are seven possible types of combined relationships: (i) NN , (ii) AN or NA , (iii) MN or NM , (iv) AA , (v) AA , (vi) AM or MA , and (vii) MM . The only pairs that are not straightforward are (iv) AA and (v) AA , which represent the two different ways that asymmetric relationships can occur. The pair AA represents a dyad in which one actor both likes and admires the other actor but neither feeling is reciprocated. The other pair AA represents the case in which the only relationships are that one actor likes the other and actor two admires the first—that is, there is a form of reciprocity in which the *admire* relation is exchanged with the *like* relation. We can now represent the basic data in a nonduplicated form using this new notation and perspective. Let $\{Z_{ab}\}$, for $a, b = M, A(\bar{A}), N$, denote unduplicated counts for the 10 different pairs. They can be arrayed in table form as in Table 5.

The easiest way to visualize these situations is through directed graphs, as given in Figure 3. The arrows indicate relationships. Double-headed arrows represent reciprocated relationships, whereas the single-headed arrows are one-way flows. Each of the seven pictures corresponds to a distinct case. For example, consider case (vi). The picture shows a situation in which relation 1 is asymmetric and relation 2 is mutual.

For each of the 10 relationships illustrated in the seven parts of Figure 3 we can envision a parameter or effect corresponding to the heightened or diminished presence of that relationship across the network. Unidirectional relationships involve pure “choice” effects and are denoted by θ 's, whereas simple bidirectional or exchange relationships involving pure “reciprocity” are denoted by ρ 's. Some relationships have embedded within them both choice and reciprocity and we denote corresponding effects by $\rho\theta$'s to indicate the combination. For example, the full mutuality relationship can be thought of either as mutuality (reciprocity) on two relationships or as reciprocity of a multiplex choice relationship.

Some of the aspects of our models can be demonstrated with the $R = 2$ case, but most of the interesting problems occur with $R > 2$. With this in mind we turn now to an $R = 4$ example and a more general discussion of models combining different “effects.”

Figure 4 contains summaries of Sampson's data, shown in Figure 1, in the form of two 2^8 *w*-arrays of counts for pairs of monks, one for the four positive relations and the other for the four negative relations. For each we have $R = 4$, with entries $\{w_{ii'jj'kk'hh'}\}$, where the subscripts appear in pairs, one pair for each relation, and take the values Y and N . Within each table, each pair is counted twice, once from the perspective of each member, yielding a total count of $2 \times \binom{8}{2} = 306$. For example, the upper left-hand entry in Figure 4(a), which is 180, is a doubled count of the 90 pairs of monks with all null relations. An all-null relationship is perceived the same by each participant, hence the doubled count. The next two counts in the first row are both 6. These duplicates occur because they correspond to perceptions from opposite ends of an asymmetric relationship. One participant views it as a (Y, N) relationship, and the other views it as a (N, Y) relationship. Similar duplication occurs for all asymmetric relationships; that is, relationships that are perceived differently by the two participants.

Classifying the data by relationship type, we obtain the layout in Table 6, which eliminates the cells that occur twice. In general, a 2^{2R} *w*-array contains $2^{R-1}(2^R + 1)$ unique cells. Among these are 2^R cells whose counts are duplicated, that is, occur twice in *w*. If we eliminate the doubling and duplication in the eight-dimensional *w*-arrays given in Figure 4, we get two arrangements of 136 cells, whose counts correctly total

Table 5. Nonredundant Arrangement of Cells for Two Relations

Relation 1	Relation 2		
	M	A	N
M	Z_{MM}	Z_{MA}	Z_{MN}
A	Z_{AM}	Z_{AA}	Z_{AN}
N	Z_{NM}	Z_{NA}	Z_{NN}

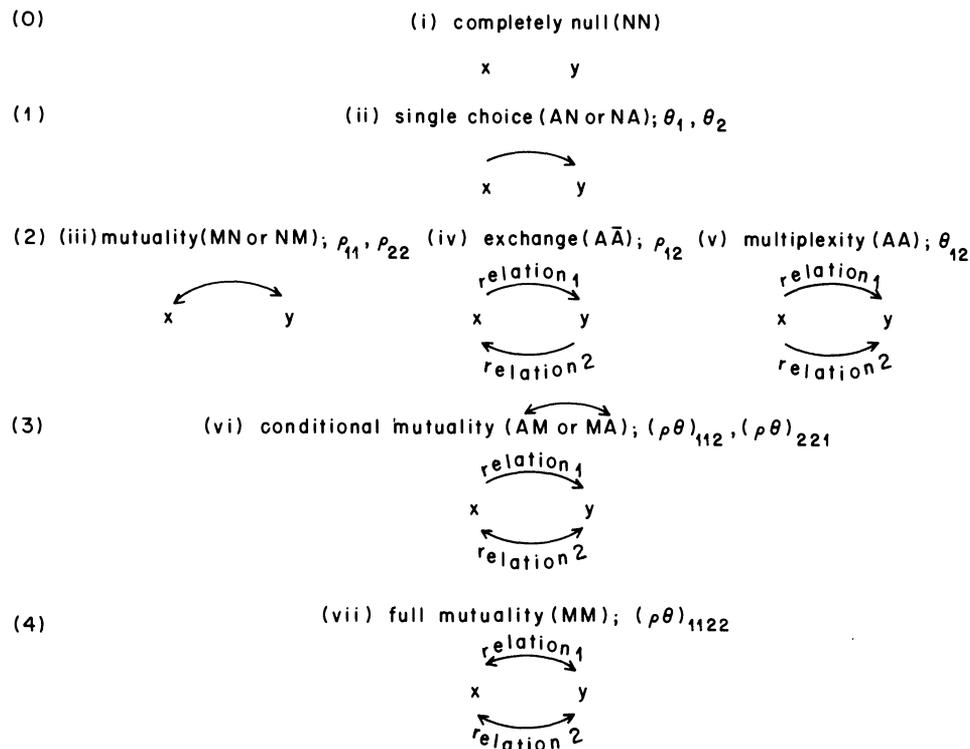


Figure 3. Patterns of Choices for Dyads With $R = 2$ Relations. Quantities to the right of the dyadic pattern names are parameters associated with that type of pattern. Numbers on the left are the numbers of choices in the patterns.

153. In Table 6, we give one possible arrangement of these 136 cells in a form resembling a four-dimensional $3 \times 3 \times 3 \times 3$ cross classification, in which some of the 81 cells, involving joint classification by two or more asymmetric relations, are further subclassified. We denote the counts in Table 6 by $\mathbf{z} = \{z_{abcd}; a, b, c, d = M, A, \bar{A}, N\}$, where $M =$ Mutual, $A =$ Asymmetric, $\bar{A} = \bar{A}$ symmetric, and $N =$ Null. If asymmetric links in different relations go in the same direction, they are all labeled by A . If some go in one direction and some in the other, they are labeled by two different subscripts, A and \bar{A} , with identical subscripts for those relations whose asymmetric arcs go in the same direction. We arbitrarily assign the subscript A to the first asymmetric relationship, going from left to right. Each cell contains a symbol for the corresponding entry, below which is the actual count corresponding to an entry or entries in Figure 4(a).

To understand the structure and notation in Table 6, we consider some examples. First, the $z_{NNNN} = 90$ entry of Table 6 is exactly one-half of the count of 180 in the (1, 1) entry of Figure 4(a). In general, cells in Figure 4(a) corresponding to only mutual or null relationships on all four dimensions contain counts that are double those in the corresponding cells of Table 6. Next the (2, 2) entry in Figure 4(a) (i.e., the count 1) corresponds to the Praise = Null (no, no), Esteem = Null (no, no), Influence = Asymmetric (no, yes), and Like = Asymmetric (no, yes). This is the cell in Table 6 labeled z_{ANAN} . Similarly the $z_{\bar{A}AMM}$ entry ($= 1$) is the Like = Asymmetric (yes, no), Esteem = Asymmetric (no, yes) (with the flow going the other way), Praise = Mutual (yes, yes), and Influence = Mutual (yes, yes) cell and corresponds to the (16, 7) and (16, 10) entries in Figure 4(a). Thus, the key to the correspondence between the cells in the tables involving at least

one asymmetric relationship is that the factors in Figure 4 need to be read in pairs.

There are certainly advantages, from both the modeling and the conceptual standpoints, in rearranging these tabled counts in several ways. First of all, we can eliminate redundancies and duplicated cells to arrive at a table (such as Table 6) containing only the unique counts. Clearly this aids our modeling task. Second, there are many ways to structure a multiway table. Some approaches will allow us to better understand the flows of information or feelings between actors. We give one example in Figure 4, where the dependencies among the relations are highlighted.

We wish to model p_{abcd} , the probability that a randomly selected dyad would be assigned to cell (a, b, c, d) of Table 6, where

$$\sum_{\text{all cells}} p_{abcd} = 1. \tag{4.1}$$

We define

$$\begin{aligned} \xi_{abcd} &= \log p_{abcd}, \text{ if all of } a, b, c, \text{ and } d \text{ are equal to} \\ &\quad \text{either } M \text{ or } N \\ &= \log(p_{abcd}/2), \text{ if one or more of } a, b, c, \text{ or } d \\ &\quad \text{equals } A, \end{aligned} \tag{4.2}$$

and we develop a class of linear models for the $\{\xi_{abcd}\}$, which yields an affine translation of a class of loglinear models for the $\{p_{abcd}\}$. The reasons for this approach were discussed by Fienberg et al. (1981); primarily, we introduce the factor of $\frac{1}{2}$ for A cells to make our models consistent with the univariate model of Holland and Leinhardt (1981).

The models for the $\{\xi_{abcd}\}$ are linear in sets of parameters

Praise Influence		Esteem		No				Yes						
		Like		No		Yes		No		Yes				
		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes			
No	No	No	180	6	6	2	6	2	1	6	1	2		
	Yes	No	4	1						3	8	1		
Yes	No	No	3	1			1	13						
	Yes	No	2	1			2	1						
No	No	No	3							1	1			
	Yes	No	2							2	13			1
Yes	No	No					1			1				
	Yes	No												

(a)

Blame Neg=Inf		Disesteem		No				Yes						
		Antagonism		No		Yes		No		Yes				
		No	Yes	No	Yes	No	Yes	No	Yes	No	Yes			
No	No	No	164	3	3		3	7		3	7			
	Yes	No	7							7	5	1		4
Yes	No	No	7	1			6	1		1				
	Yes	No									2			
No	No	No	7							6	1			
	Yes	No	2							1	7			
Yes	No	No					2							
	Yes	No					2							
Yes	No	No									2			
	Yes	No											1	4

(b)

Figure 4. Sampson's Cloister Data Aggregated Over Actors Displayed as 2⁸ Tables for (a) Positive Relations and (b) Negative Relations.

that reflect the various distinct types of dyadic patterns. In Fienberg et al. (1981) we discussed the $R = 3$ case. Figure 3 and the associated notation show the possible distinct relationships and the parameters involved in the $R = 2$ case. For each of the 10 possible dyadic patterns there is a corresponding parameter. For example, case (vi) has a parameter $(\rho\theta)_{122}$ for the type AM (the 1 represents the asymmetric arc for relation 1 and the 22 represents the mutual arcs for relation 2), and $(\rho\theta)_{112}$ for MA . The parameters have a hierarchical structure. For example, if $(\rho\theta)_{122}$ is present, then this implies that $\theta, \theta_1, \theta_2, \theta_{12}, \rho_{12}$, and ρ_{22} are all present.

As mentioned before, the $\{\xi_{ab}\}$ are linear functions of these hierarchically structured parameters. Thus suppose we consider the possibility of effects associated with only levels 0, 1, and 2 of Figure 3. Then the model for the translated log-probabilities is

$$\begin{aligned}
 \xi_{NN} &= \mu \text{ (for normalization),} \\
 \xi_{AN} &= \mu + \theta_1, & \xi_{NA} &= \mu + \theta_2, \\
 \xi_{MN} &= \mu + 2\theta_1 + \rho_1, & \xi_{NM} &= \mu + 2\theta_2 + \rho_2, \\
 \xi_{A\bar{A}} &= \mu + \theta_1 + \theta_2 + \rho_{12}, & \xi_{AA} &= \mu + \theta_1 + \theta_2 + \theta_{12}, \\
 \xi_{MA} &= \mu + 2\theta_1 + \theta_2 + \theta_{12} & \xi_{AM} &= \mu + \theta_1 + 2\theta_2 + \theta_{12} \\
 &+ \rho_{11} + \rho_{12}, & &+ \rho_{22} + \rho_{12}, \\
 \xi_{MM} &= \mu + 2\theta_1 + 2\theta_2 + 2\theta_{12} + \rho_{11} + \rho_{22} + 2\rho_{12}. \quad (4.3)
 \end{aligned}$$

For further explanations and illustrations, see Fienberg et al. (1981). When we extend this form of modeling to the $R = 4$ case, there are 22 distinct types of relationships and a total of 81 possible parameters.

The parameters in this family of models are GLIM-like in structure (see Nelder and Wedderburn 1972). A parameter is included in the model if and only if the corresponding effect (such as choice, conditional multiplexity, etc.) is present. In general, we reiterate that the models are also hierarchical: If we set some parameters equal to zero, all related higher-order terms are also zero.

To fit these models to multivariate networks, we apply the general results for fitting loglinear models given in Haberman (1974) or Appendix II of Fienberg (1980). The minimal sufficient statistics (MSS's) are linear combinations of the elements of the z array, with coefficients of 0, 1, or 2. The fitted values of these elements are found by solving the likelihood equations, which set the MSS's equal to their estimated expected values. We could use a version of generalized iterative proportional fitting due to Darroch and Ratcliff (1972). This method allows us to work with the z -table but requires that we explicitly specify (i.e., write out) each of the sufficient statistics. This is not only tedious but also error prone. Even worse, the generalized iterative proportional fitting procedure can be slow to converge and is computationally inefficient.

An alternative method of fitting these models works directly

Table 6. Nonredundant Arrangement of Cells for Positive Relations From Sampson's Data

Like	Praise											
	M, Influence			A, Influence			N, Influence					
	M	A	N	M	A	N	M	A	N			
M, Esteem												
M	Z _{MMMM} 0	Z _{MMAM} 0	Z _{MMNM} 0	Z _{MMMA} 0	Z _{MMAA} 0	Z _{MMĀĀ} 0	Z _{MMNA} 0	Z _{MMM̄N} 0	Z _{MMAN} 0	Z _{MMNN} 0		
A	Z _{MAMM} 0	Z _{MAAM} 0	Z _{MAĀM} 0	Z _{MANM} 0	Z _{MAMA} 0	Z _{MAAA} 0	Z _{MAĀĀ} 0	Z _{MANA} 1	Z _{MAMN} 0	Z _{MAAN} 1	Z _{MAĀN} 0	Z _{MANN} 0
N	Z _{MNMM} 0	Z _{MNAM} 0	Z _{MNNM} 0	Z _{MNMA} 0	Z _{MNAA} 0	Z _{MNĀĀ} 0	Z _{MNNA} 0	Z _{MNM̄N} 0	Z _{MNAN} 0	Z _{MNNN} 1		
A, Esteem												
M	Z _{AMMM} 0	Z _{AMAM} 0	Z _{AMĀM} 0	Z _{AMNM} 0	Z _{AMMA} 0	Z _{AMAA} 1	Z _{AMĀĀ} 0	Z _{AMNA} 0	Z _{AMMN} 0	Z _{AMAN} 0	Z _{AMĀN} 0	Z _{AMNN} 0
A	Z _{AAMM} 0	Z _{AAAM} 0	Z _{AAĀM} 0	Z _{AANM} 0	Z _{AAMA} 1	Z _{AAAA} 13	Z _{AAĀĀ} 0	Z _{AANA} 0	Z _{AAMN} 0	Z _{AAAN} 8	Z _{AAĀN} 0	Z _{AANN} 2
N	Z _{ANMM} 0	Z _{ANAM} 0	Z _{ANĀM} 0	Z _{ANNM} 0	Z _{ĀĀAMA} 0	Z _{ĀĀAA} 0	Z _{ĀĀĀĀ} 0	Z _{ĀĀANA} 0	Z _{ĀĀMN} 0	Z _{ĀĀAN} 0	Z _{ĀĀĀN} 0	Z _{ĀĀNN} 1
					Z _{ANMA} 0	Z _{ANAA} 1	Z _{ANĀĀ} 0	Z _{ANNA} 1	Z _{ANMN} 0	Z _{ANAN} 1	Z _{ANĀN} 0	Z _{ANNN} 6
					Z _{ANMĀ} 0	Z _{ANĀĀ} 0	Z _{ANĀĀ} 0	Z _{ANNĀ} 0				
N, Esteem												
M	Z _{NMMM} 0	Z _{NMAM} 0	Z _{NMNM} 0	Z _{NMMA} 0	Z _{NMAA} 0	Z _{NMĀĀ} 0	Z _{NMNA} 0	Z _{NMM̄N} 0	Z _{NMAN} 0	Z _{NMNN} 0		
A	Z _{NAMM} 0	Z _{NAAM} 0	Z _{NAĀM} 0	Z _{NANM} 1	Z _{NAMA} 0	Z _{NAAA} 2	Z _{NĀĀĀ} 0	Z _{NANA} 1	Z _{NAMN} 0	Z _{NAAN} 3	Z _{NAĀN} 0	Z _{NANN} 6
N	Z _{NNMM} 0	Z _{NNAM} 0	Z _{NNNM} 0	Z _{NNMA} 0	Z _{NNAA} 2	Z _{NNĀĀ} 0	Z _{NNNA} 3	Z _{NNM̄N} 0	Z _{NNAN} 4	Z _{NNNN} 90		

NOTE: Corresponds to Figure 4(a).

with the w-table and uses the standard IPFP. It relies on a “trick” that is outlined in Fienberg et al. (1981) and fully explained in Meyer (1982). The following results form the basis of the technique.

Result 1. For the class of affine translations of hierarchical loglinear models described before, each set of MSS's is equivalent to a set of marginal totals for the 2^{2R} table (i.e., the w-table) with doubled and duplicated counts.

Result 2. For each affine translation of a loglinear model for the z-table, there is a corresponding loglinear model for the w-table, with equivalent estimated expected values, once we take account of the duplication and doubling.

To calculate the fitted values for a z-table model one must first fit the corresponding w-table model. The fitted values for the w-table are then transformed back into a z-table structure by removing the duplications and halving the doubled cells. The remarkable fact is that the fitted w-table is the image of some fitted z-table and thus the back transformation is possible. In fact this is the basis of the whole transformation technique.

Provided the model fit to the w-table enforces all of the sufficient statistic constraints and provided the resulting fitted values satisfy the symmetry requirements to be a w-table, it is possible to carry out this transformation. Such a try-it-and-see approach certainly works in our situation; however, one can actually prove that the results will be correct for the models we have considered (see Meyer 1982 for details). While this approach does yield fitted values, it does not give an easy method of calculating degrees of freedom. For this we have to rely on first principles and some hand calculation.

4.2 Models for Both Microanalysis and Macroanalysis: Actor and Group Effects

We now consider models for multiple relations that allow the actors in the network to engage in relations at possibly different rates and include both actor and group effects. To review, we suppose that the R sociometric relations defined for a group of g actors are binary, and the presence/absence of directed links between actors is recorded in the form of R sociomatrices. As before, we concentrate on the dyadic rela-

tionships between the $\binom{g}{2}$ pairs of actors i and j , represented by the $2R$ -variate \mathbf{D}_{ij} , with realization \mathbf{d}_{ij} .

Primarily to limit the number of parameters, we now assume that the actors have been partitioned into K mutually exclusive and exhaustive subgroups, G_1, G_2, \dots, G_K . This allows the models to generate easily interpretable likelihood ratio statistics. In practice, it can be very useful to also allow for the inherent differences in the actors. If there are single actors that behave contrary to the group as a whole (or to the collection of subgroups), then they can be placed into their own singleton subgroups. Thus, their individual differences can still be modeled directly.

In this section we outline models that can include both actor and group effects even when all actors are placed into one of the K subgroups. These models contain all of the previous models as special cases. The R sociomatrices are used to construct a table of pseudocounts of size $g \times g \times (2 \times 2)^R$. From this multivariate version of the \mathbf{y} -array, we can aggregate the $(2 \times 2)^R$ tables to form a $K \times K \times (2 \times 2)^R$ table, whose entries are the frequencies of the different dyadic relationship patterns between actors of a group partitioned into K subgroups. As in the earlier cases, it is most convenient to work with the full $g \times g \times (2 \times 2)^R$ data table but to describe models in terms of the unduplicated data array. This approach also grants us a considerable degree of flexibility in fitting the models. For many of the models it is possible to consider collapsed or aggregated versions of the data that would result in smaller data tables. We believe, however, that the unification that is introduced by always considering the full data table outweighs the occasional advantage of having a smaller table.

We will begin our discussion by concentrating on the choice parameters in an $R = 2$ relation network. Consider the general model

$$\log P(\mathbf{D}_{ij} = \mathbf{d}_{ij}) = \lambda^{(ij)} + \sum_{r=1}^R \theta_r^{(ij)} x_{ijr} + \sum_{r=1}^R \theta_r^{(ji)} x_{jir}, \quad \text{for all } i > j, \quad (4.4)$$

where x_{ijr} equals 1 if there is an arc linking individual i to j on relation r . The model in expression (4.4) includes different choice parameters for each pair of individuals. The parameters $\lambda^{(ij)}$ are normalizing constants and are required so as to meet the marginal constraints of the problem.

Initially, we focus our attention on just the first relation—that is, $r = 1$ —and give some special cases of model (4.4). If we wished to consider a model that asserted that the response depended only on the chooser, we would allow $\theta_1^{(ij)} = \theta + \theta_1^{(i,\cdot)}$. Similarly, dependence only on the chosen actor would lead to $\theta_1^{(ij)} = \theta + \theta_1^{(\cdot,j)}$. Obviously we could allow chooser and chosen actor effects (but excluding the interaction) by specifying $\theta_1^{(ij)} = \theta + \theta_1^{(i,\cdot)} + \theta_1^{(\cdot,j)}$. Another version of this model would be to suppose that individual actors assert influence only through the groups to which they belong. We need to introduce a set of group-specific parameters, and we shall use the usual parameters with superscripts in square braces. For example, $\theta^{[1,\cdot]}$ would indicate a choice parameter from group 1, $\theta^{[\cdot,3]}$ a choice parameter from group 3, and $\theta^{[1,3]}$ a choice parameter that depended on both groups. To ease the specification of models, let γ be the function that maps indi-

vidual indexes into the appropriate group indexes. Now to specify a model with group effects we could write $\theta_1^{(ij)} = \theta_1 + \theta_1^{[\gamma(i),\cdot]}$ for the choosing group $\gamma(i)$, $\theta_1^{(ij)} = \theta_1 + \theta_1^{[\cdot,\gamma(j)]}$ for the chosen group $\gamma(j)$, $\theta_1^{(ij)} = \theta_1 + \theta_1^{[\gamma(i),\cdot]} + \theta_1^{[\cdot,\gamma(j)]}$ for both, or even $\theta_1^{(ij)} = \theta_1 + \theta_1^{[\gamma(i),\cdot]} + \theta_1^{[\cdot,\gamma(j)]} + \theta_1^{[\gamma(i),\gamma(j)]}$ to indicate group-by-group choice interactions. If we aggregate over groups, then these models are just “actor” models in which the actors are the groups. In summary, we have just described four basic classes of models—those that involve parameters $(\theta^{(i,j)})$ for each pair of actors, and those with individual parameters $(\theta^{(i,\cdot)})$ and $(\theta^{(\cdot,j)})$ for actors and the corresponding notions, $(\theta^{[d,\cdot]})$ and $(\theta^{[\cdot,e]})$, for groups, where d and e index over the K subgroups. Thus some possible choice-only models for one relation are

- $\log P(D_{ij} = d_{ij})$
- = (1) θ , constant
 - (2) $\theta + \theta^{(i,\cdot)}$, chooser
 - (3) $\theta + \theta^{(\cdot,j)}$, chosen
 - (4) $\theta + \theta^{(i,\cdot)} + \theta^{(\cdot,j)}$, chooser and chosen
 - (5) $\theta + \theta^{(i,\cdot)} + \theta^{(\cdot,j)} + \theta^{(i,j)}$, interaction
 - (6) $\theta + \theta^{[\gamma(i),\cdot]}$, group chooser
 - (7) $\theta + \theta^{[\cdot,\gamma(j)]}$, group chosen
 - (8) $\theta + \theta^{[\gamma(i),\cdot]} + \theta^{[\cdot,\gamma(j)]}$, group chooser and chosen
 - (9) $\theta + \theta^{[\gamma(i),\cdot]} + \theta^{[\cdot,\gamma(j)]} + \theta^{[\gamma(i),\gamma(j)]}$, group interaction.

It is possible to mix and match among these models. For example, the model $\theta + \theta^{(i,\cdot)} + \theta^{(\cdot,j)} + \theta^{[\gamma(i),\gamma(j)]}$ allows individual actor parameters and a group interaction. As soon as we contemplate such models, we need to note that there is a partial hierarchy to the models just displayed, which we represent in Figure 5. The diagram indicates that any parameter at level i implies all of those parameters at levels less than i . The modeling strategy that we outlined can be used for other types of flows (e.g., mutual, exchange, and multiplex). For multiple relations (i.e., $R \geq 2$), we need to model not only choice parameters (i.e., θ 's), but also mutuality or reciprocation parameters (i.e., ρ 's), and the more complex combinations of these (i.e., $\rho\theta$'s) in various forms. For mutuality parameters, one cannot distinguish chooser from chosen, and as a consequence we write $\rho_{11}^{(ij)} = \rho_{11}^{(i,\cdot)} = \rho_{11}^{(\cdot,j)}$. Thus for $R = 2$ each type of parameter illustrated in Figure 5 can be the basis of a

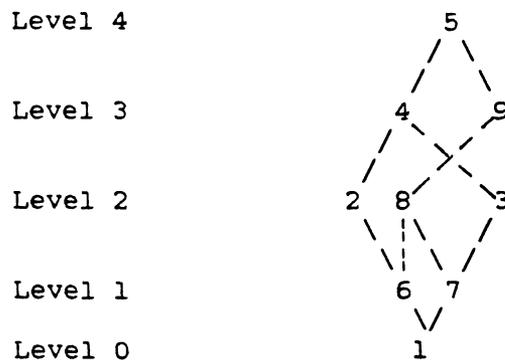


Figure 5. Hierarchy Displaying Levels of Parameters in Selected Log-Linear Models.

modeling strategy similar to that outlined here for simple choices. In these cases we need to be concerned about the hierarchical structure linking different parameter types as well as the hierarchical structure within parameter types.

4.3 Fitting Models

If we restrict ourselves to actor parameters, then we can fit the models described before using the IPFP to adjust simple margins of the symmetric $g \times g \times (2 \times 2)^R$ data array. When models with group parameters are included it is still possible to use the IPFP, but now a more general notion of margin is needed.

Let us consider two relations and the model $\log P(\mathbf{D}_{ij} = \mathbf{d}_{ij}) = \lambda^{(ij)} + \theta_1 + \theta_1^{(i,\cdot)}$, i.e., a choice parameter on only the first relationship. The sufficient statistics for this model in the symmetric data array are

$$[12] \sum_{khmn} x_{ijkhmn} \text{ for all } i, j,$$

$$[13] \sum_{jhmn} x_{ijkhmn} \text{ for all } i, k,$$

$$[24] \sum_{ikmn} x_{ijkhmn} \text{ for all } j, h.$$

Now consider the model $\lambda^{(ij)} + \theta_1 + \theta_1^{(j(\cdot),\cdot)}$. For this model the sufficient summary is

$$\sum_{khmn} x_{ijkhmn} \text{ for all } i, j,$$

$$\sum_{i \in G_d} \sum_{jhmn} x_{ijkhmn} \text{ for } d = 1, \dots, G \text{ and for all } k,$$

$$\sum_{j \in G_d} \sum_{ikmn} x_{ijkhmn} \text{ for } d = 1, \dots, G \text{ and for all } h.$$

We extend the usual square-bracket notation to this situation. Recall that [12] indicates that for each value of i and j we should sum over all other dimensions in the table. We shall use the notation $[g(1) g(2)]$ to indicate that for each \mathcal{G}_d and \mathcal{G}_e we should sum over all entries in the table. A simple example should help to explain the notation.

Consider the following 3×3 table:

	1	2	3
1	y_{11}	y_{12}	y_{13}
2	y_{21}	y_{22}	y_{23}
3	y_{31}	y_{32}	y_{33}

Let $\mathcal{G}_1 = \{1, 2\}$ and $\mathcal{G}_2 = \{3\}$. Then the [1] margin is the triple (y_{1+}, y_{2+}, y_{3+}) , the $[g(1)]$ margin is the pair $(y_{1+} + y_{2+}, y_{3+})$, and the $[g(1) g(2)]$ margin is the table

$y_{11} + y_{12} + y_{21} + y_{22}$	$y_{13} + y_{23}$
$y_{31} + y_{32}$	y_{33}

In effect we have collapsed over the groups. It is an easy application of the IPFP to fit models that use this generalized notion of margin. Most statistical packages (e.g., BMDP or IMSL) that contain an IPFP routine cannot be easily cajoled into fitting such models without some tinkering to allow for the generalized definition of margin that we have presented.

It is relatively simple, however, to extend an IPFP algorithm to handle these generalized margins. To do so, one needs to modify the margin-handling aspects of the algorithm (i.e., the formulation of the steps in the iteration corresponding to the MSS's) so as to recognize an arbitrary partition rather than just the natural margins. It is also necessary to provide an easy method for specifying generalized margins. The easiest method of doing this is to allow the user to specify a table of integers, which indicates which cells belong to which levels of a generalized margin. In the preceding example, the $[g(1) g(2)]$ margin consists of four levels, and the table of integers we would use to specify this margin would be

1	1	2
1	1	2
3	3	4

We now use the notation to show which models correspond to certain parameterizations for $R = 2$. Table 7 lists some of the choice models and a small selection of other possible models. In the table, i and j index the actors, and d and e index the groups. There are many possible models with many possible combinations of population, group, and individual parameters. We list only the highest-order (in a hierarchical interaction sense) parameters in the model. Thus, if $\theta^{[d,e]}$ is listed, $\theta^{[d,\cdot]}$, $\theta^{[\cdot,e]}$, and θ are also included, and if $\rho_{12}^{[d,e]}$ is listed, $\theta_1^{[d,e]}$, $\theta_2^{[d,e]}$, $\rho_{12}^{[d,\cdot]}$, and $\rho_{12}^{[\cdot,e]}$ are also included in addition to their implied lower-order terms.

4.4 An Example: Sampson's Data

To demonstrate the ubiquity and apparent complexity of these social network models, we have taken a somewhat unusual (and bold) approach to the analysis of Sampson's network of 18 monks, described in Section 1. We view the four positive attributes (like, esteem, influence, and praise) as realizations of a single positive-effect process, and the four negative relations (antagonism, disesteem, negative influence, and blame) in a similar manner. There is substantial justification for this pooling. White et al. (1976) found that when the 18 actors are aggregated into three blocks, the concrete social structure of this network is much the same across the four pairs of positive/negative relations: "A top-esteemed block (consisting of 7 actors) unambivalently positive toward itself, in conflict with . . . a second, more ambivalent block (also of 7 actors) to which is attached a block of losers (of size 4)." We label these blocks or subgroups as

$$\mathcal{G}_1 = \{1, 2, \dots, 7\}, \quad \mathcal{G}_2 = \{8, 9, \dots, 14\},$$

$$\mathcal{G}_3 = \{15, 16, 17, 18\}.$$

Therefore we aggregate over both sets of relations by summing the four sociomatrices for the positive relations, and the four negative relations, to obtain one positive and one negative relation matrix. These arrays, given in Figure 6, have entries indicating the number of times actor i chooses actor j , either on the positive or negative choices. Each of the four positive relations is paired with a negative relation, and thus the arrays in Figure 6 can be thought of as marginal totals for the aggregated table we actually wish to analyze.

Table 7. A Selection of Possible Models for $R = 2$

Highest Order Parameters in the Model	Margins of x to Be Fitted	Newly Introduced Parameter(s)
Choice		
θ_1	[12] [3] [4]	Simple choice
$\theta_1^{(i)}$	[12] [13] [24]	Chooser choice
$\theta_1^{(j)}$	[12] [14] [23]	Chosen choice
$\theta_1^{(i)}, \theta_1^{(j)}$	[12] [13] [14] [23] [24]	Chooser and chosen choices
$\theta_1^{(g)}$	[12] [g(1) 3] [g(2) 4]	Chooser group effects
$\theta_1^{(e)}$	[12] [g(1) 4] [g(2) 3]	Chosen group effects
$\theta_1^{(g,e)}$	[12] [g(1) g(2) 3] [g(1) g(2) 4]	Chooser and chosen group interaction
Mutuality		
$\rho_{11}^{(i)}$	[12] [134] [234]	Chooser and chosen mutuality, relation 1
$\rho_{22}^{(j)}$	[12] [g(1) 56] [g(2) 56]	Chooser and chosen group mutuality, relation 2
$\rho_{11}^{(g,e)}$	[12] [g(1) g(2) 34]	Chooser and chosen group mutuality interaction, relation 1
Multiplex		
$\theta_{12}^{(i)}$	[12] [135] [246]	Chooser multiplex, relations 1 and 2
$\theta_{12}^{(e)}$	[12] [g(1) 46] [g(2) 35]	Chosen group multiplex, relations 1 and 2
$\theta_{12}^{(g,e)}$	[12] [g(1) g(2) 35] [g(1) g(2) 46]	Chooser and chosen group multiplex interaction, relations 1 and 2
Exchange		
$\rho_{12}^{(i)}$	[12] [136] [245]	Chooser and chosen exchange, relations 1 and 2
$\rho_{12}^{(g), \rho_{12}^{(e)}}$	[12] [g(1) 45] [g(2) 36] [g(1) 36] [g(2) 45]	Chooser and chosen group exchange, relations 1 and 2
$\rho_{12}^{(g,e)}$	[12] [g(1) g(2) 36] [g(1) g(2) 45]	Chooser and chosen group exchange interaction, relations 1 and 2

The techniques we have used to analyze 0-1 sociomatrices are directly applicable here. The data are transformed into a six-dimensional contingency table, by treating the four pairs of positive/negative relations as replications, and we then fit log-linear models to this table. If we do not use grouping effects, then all we do is fit standard models that have certain margins of the table as sufficient statistics. With grouping effects, we still fit standard log-linear models, but the sufficient statistics are now the generalized version of margin that we outlined before. Furthermore, with multiple observations on each actor, the asymptotic basis for the goodness-of-fit statistics stands on firmer ground. In our analysis we have examined the $18 \times 18 \times (2 \times 2) \times (2 \times 2)$ (corresponding to actor

\times actor \times positive \times negative) version of this table and have used the three groups given before.

A priori, some choices are unlikely to be reciprocated across relations. We should find a simple-choice or group-choice model to be an adequate summarization of the flows of attitudes, both positive and negative, across and between these three, substantively different, subgroups. A summary of some of the models that we fit to this network is given in Table 8. Note that the models are hierarchical, with each model containing all of the terms of the models preceding it in the table.

The difference in goodness of fit between models 2 and 3 (which in a sense is a measure of the impact of the grouping effect) is statistically significant at any reasonable level of

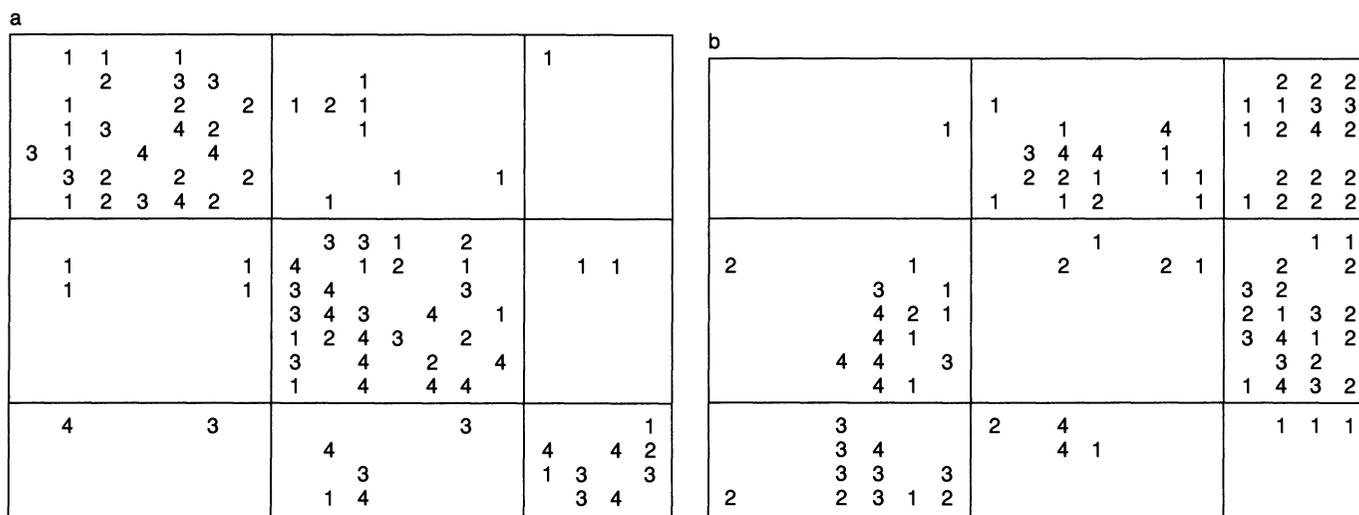


Figure 6. Sampson's Cloister Data Aggregated Over Relations: (a) Aggregated Positive; (b) Aggregated Negative.

Table 8. Summary of Fit of Several Models on Sampson's Data

Model	Highest Order Parameters	Margins	df	G ²	X ²
1	$\lambda^{(g)}$	[12]	2,295	2,835	6,252
2	θ_1, θ_2	[12] [3] [4] [5] [6]	2,293	1,453	3,392
3	$\theta_1^{(g,1)}, \theta_1^{(g,e)}$ $\theta_2^{(g,1)}, \theta_2^{(g,e)}$	[12] [g(1) 3] [g(1) 4] [g(1) 5] [g(1) 6] [g(2) 3] [g(2) 4] [g(2) 5] [g(2) 6]	2,285	1,395	3,350
4	$\rho_{12}^{(g,1)}, \rho_{12}^{(g,e)}$	[12] [g(1) 36] [g(2) 36] [g(1) 45] [g(2) 45]	2,279	1,368	3,088
5	$\rho_{12}^{(g,e)}$	[12] [g(1) g(2) 36] [g(1) g(2) 45]	<1,800	1,180	2,158

significance and is typical of the improvement resulting from the addition of simple grouping parameters. Similarly the difference in G² values for models 4 and 5 is also large, but it is less than the difference in degrees of freedom. The small number of degrees of freedom for models is caused by a large number of fitted zeros. Indeed, any model that includes even an overall multiplex (θ_{12}) parameter induces at least 2,142 fitted zeros out of the $18 \times 18 \times 4 \times 4 = 5,184$ cells in the table, and the goodness-of-fit statistics are not dramatically improved by the inclusion of these parameters. We note the very large differences between the G² and X² values in Table 8, which go in the opposite direction from that suggested by the argument given in Larntz (1978). The only explanation we can offer is the presence of the large proportion of observed zero cells. The degrees of freedom in Table 8 are calculated as follows. There are 2,592 observations. Model 1 has 153 (= 9 × 17) parameters and 144 fitted zeros, which result in 2,295 degrees of freedom. For each of the subsequent entries in the table the reduction on degrees of freedom is just the number of added parameters, except for model 5, where there are a substantial number of additional fitted zeros.

It appears that model 4, which includes different reciprocity effects for each group, provides a reasonable description of the data. This model is what we would expect to find because the groups we used were determined by a block-model algorithm, whose objective is to produce groups with a high degree of reciprocity. Hence our final model is not surprising.

A more thorough analysis of the data for this network should include a detailed study of the similarities of the four pairs of positive/negative relations and should experiment with other, more refined partitions of the actors, as suggested by Breiger et al. (1975). We have just touched the surface of a rather large, and certainly rich, set of longitudinal data. We have studied the monastery structure only at the midpoint of a 12-month period, during which a crisis over theology occurred, and the group split up.

5. CONCLUSION

In this article, we have considered a variety of loglinear models for micro- and macro-analysis of binary social-network data, and we have demonstrated how these models can be treated in a unified manner. The models we have considered describe important aspects of the data, and we have had the good fortune to be able to take advantage of relatively easy estimation methods for model fitting.

Unfortunately, large data sets and corresponding large models

are almost inevitable with the type of sociometric relational problems that we have described here. Our modeling has been consciously and unconsciously influenced by what it is possible for us to compute. The models with separate group effects seem to be at the limits of the computational methodology that we have presented. Other models that could be considered interesting (e.g., additional relationships between the groups, akin to ordered category models for contingency tables) have not been mentioned. This is not because we find them uninteresting, but rather because the computational prospect of fitting such models is daunting.

There are two basic theoretical problems with the models presented that we have highlighted at various points in this article. First, there is the assumption of dyadic independence, which is fundamental to models we have considered. To date no one has suggested a feasible approach that would allow for a check on this assumption. The Markov graph models of Frank and Strauss (1983) offer one possible route to relaxing this assumption, but further work needs to be done before this possibility can be realized. The second problem is the lack of a suitable asymptotic framework that could provide reference distributions for inference purposes, especially for assessing goodness of fit. For further details, see the discussion of this problem in Fienberg and Wasserman (1981b) and Haberman (1981). We commend these problems to the attention of theoretically-oriented readers.

We believe that we have indicated how more general models could be formulated, and we have presented some of the techniques that are appropriate for fitting the models to actual data. Further advances in methodology in this area are likely to be as dependent upon advances in numerical algorithms or computer hardware as they will be on new statistical ideas.

APPENDIX

In the next several paragraphs, we comment on the calculations of degrees of freedom (df) for the models used in the analysis of the corporate interlock network data of Section 3. First we note that a network of *g* actors and a single relation between the actors is reorganized into a $g \times g \times 2 \times 2$ contingency table containing $\binom{g}{2}$ dyads. Each dyad contains two pieces of information, the choice/nonchoice of actor *j* by actor *i* and the choice/nonchoice of actor *i* by actor *j* so that we begin with $2 \times \binom{g}{2} = g(g - 1)$ df for modeling.

For a network with no structural zeros (except self-choices $\{x_{ii}\}$, which are always fixed at zero), Holland and Leinhardt's p_1 model contains ($g - 1$) expansiveness parameters,

α 's, $(g - 1)$ popularity parameters, β 's, a grand mean, θ , and a reciprocity parameter, ρ . Thus the df associated with p_1 is $g(g - 1) - (g - 1) - (g - 1) - 1 - 1 = g(g - 3)$. Special cases and generalizations of p_1 have df obtained by adding or subtracting to $g(g - 3)$ the number of independent parameters either set to zero or added to p_1 [as with model (2.6), which has $g(g - 3) - (g - 1) = g(g - 4) + 1$ df]. This rule also applies to subgroup models, such as (2.10).

The situation with p_1 is slightly more complicated when the sociomatrix contains structural zeros and/or has entire rows or columns of zeros. In the first case, we identify the number of such structural zero dyads and reduce the df available for modeling by that number. For example, the 25-actor corporate-interlock network contains 27 "outlawed" dyads, so we begin with $2[\binom{25}{2} - 27] = 546$ df, rather than 600.

If the original sociomatrix contains no zero in-degrees or out-degrees and if the inclusion of any structural zeros in the analysis has also not forced any rows or columns to have zero sums, then we proceed with df calculations as outlined before, using the corrected df available for modeling. To consider the situation for p_1 in which the sociomatrix contains zero in-degrees or out-degrees, the simplest way to proceed is to classify the actors based on whether they enter into any relationships with the other actors. The four types into which all g actors can be placed are as follows:

1. Type A—actors with nonzero in-degrees and out-degrees
2. Type B—actors with zero out-degrees and nonzero in-degrees
3. Type C—actors with zero in-degrees and nonzero out-degrees
4. Type D—actors with zero in-degrees and out-degrees.

We then consider a dyad and the type of actors in the dyad. If both actors are of Type A, we have two df for the dyad as usual. If one actor is of Type A and the other is of Type B or C, we have just one df for the dyad and the same for one actor of Type B and one of Type C. For the other five pairs of types, we have zero df for the dyad. With regards to our example, we have 11 actors in A, 5 in B, 5 in C, and 4 in D. We need only consider the following dyads:

1. Within A— $\binom{11}{2}$ dyads, but 4 are structural zeros = 51 dyads
2. Between A and B— 11×5 dyads, but 8 are zeros = 47 dyads
3. Between A and C— 11×5 dyads, but 3 are zeros = 52 dyads
4. Between B and C— 5×5 dyads, but 2 are zeros = 23 dyads.

Within A, each dyad contributes two df, and the other three pairs of types contribute just one, which gives $2 \times 51 - 47 - 52 - 23 = 224$ dyads available for modeling, or 448 df. Note that if we only fit θ and/or ρ , the in-degrees and out-degrees are not sufficient statistics, so we keep the full complement of $2 \times [\binom{g}{2} - \text{number of structural zero dyads}]$ df, and the preceding reasoning must be altered if we fit the $\{\alpha_i\}$ but not $\{\beta_j\}$ parameters, or vice versa.

Define m_α and m_β as the numbers of α 's and β 's that can be estimated—that is, the numbers of parameters $\neq -\infty$. We then

subtract $m_\alpha + m_\beta + 2$ from df for modeling to arrive at the df for p_1 . For our example, $m_\alpha = 15$ and $m_\beta = 15$, so we have $224 - 15 - 15 - 2 = 192$ df. A general rule is as follows: Let n_i be the number of actors of each type, $i = A, B, C, D$; then with no structural zero dyads,

$$\text{df for } p_1 = \binom{n_A}{2} \times 2 + n_A n_B + n_A n_C + n_B n_C - m_\alpha - m_\beta - 2. \quad (\text{A.1})$$

If there are structural zero dyads, let $s(i, j)$ be the number of structural zero dyads between actors of types i and j . Then,

$$\text{df for } p_1 = [\binom{n_A}{2} - s(A, A)] \times 2 + n_A n_B - s(A, B) + n_A n_C - s(A, C) + n_B n_C - s(B, C) - m_\alpha - m_\beta - 2. \quad (\text{A.2})$$

Note that for model (2.6), m_ρ , the number of ρ 's that can be estimated, is equal to n_A . Thus we need to subtract n_A from (A.1) or (A.2) to arrive at the df for this generalization of p_1 .

[Received October 1982. Revised July 1984.]

REFERENCES

- Bishop, Y. M. M., Fienberg, S. E., and Holland, P. W. (1975), *Discrete Multivariate Analysis: Theory and Practice*, Cambridge, MA: MIT Press.
- Breiger, R. L. (1981), Comment on "An Exponential Family of Profitability Densities for Directed Graphs," by P. W. Holland and S. Leinhardt, *Journal of the American Statistical Association*, 76, 51-54.
- Breiger, R. L., Boorman, S. A., and Arabia, P. (1975), "An Algorithm for Clustering Relational Data With Application to Social Network Analysis and Comparison With Multidimensional Sealing," *Journal of Mathematical Psychology*, 12, 328-383.
- Burt, R. S. (1980), "Models of Network Structure," *Annual Review of Sociology*, 6, 79-141.
- Cartwright, D., and Harary, F. (1958), "Structural Balance: A Generalization of Heider's Theory," *Psychology Review*, 63, 277-293.
- Daroch, J. N., and Ratcliff, D. (1972), "Generalized Iterative Scaling of Loglinear Models," *The Annals of Mathematical Statistics*, 43, 1470-1480.
- Davis, J. A. (1968), "Statistical Analysis of Pair Relationships: Symmetry, Subjective Consistency, and Reciprocity," *Sociology*, 31, 102-119.
- (1970), "Clustering and Hierarchy in Interpersonal Relations: Testing Two Graph Theoretical Models on 742 Sociograms," *American Sociological Review*, 35, 843-852.
- Festinger, L. (1949), "The Analysis of Sociograms Using Matrix Algebra," *Human Relations*, 2, 153-158.
- Fienberg, S. E. (1980), *The Analysis of Cross-classified Categorical Data* (2nd ed.), Cambridge, MA: MIT Press.
- Fienberg, S. E., Meyer, M. M., and Wasserman, S. (1981), "Analyzing Data From Multivariate Directed Graphs: An Application to Social Networks," in *Interpreting Multivariate Data*, ed. V. Barnett, London: John Wiley, 289-306.
- Fienberg, S. E., and Wasserman, S. (1981a), "Categorical Data Analysis of Single Sociometric Relations," in *Sociological Methodology 1981*, ed. S. Leinhardt, San Francisco: Jossey-Bass, 156-192.
- (1981b), Comment on "An Exponential Family of Probability Densities for Directed Graphs," by P. W. Holland and S. Leinhardt, *Journal of the American Statistical Association*, 76, 54-57.
- Frank, O. (1981), "A Survey of Statistical Methods for Graph Analysis," in *Sociological Methodology*, ed. S. Leinhardt, San Francisco: Jossey-Bass, 110-155.
- Frank, O., and Strauss, D. (1983), "Markov Graphs," unpublished manuscript.
- Galaskiewicz, J., and Marsden, P. V. (1978), "Interorganizational Resource Networks: Formal Patterns of Overlap," *Social Science Research*, 7, 89-107.
- Galaskiewicz, J., and Wasserman, S. (1981), "A Dynamic Study of Change in a Regional Corporate Network," *American Sociological Review*, 46, 475-484.
- Haberman, S. J. (1974), *The Analysis of Frequency Data*, Chicago: University of Chicago Press.
- (1981), Comment on "An Exponential Family of Probability Densities

- for Directed Graphs," by P. W. Holland and S. Leinhardt, *Journal of the American Statistical Association*, 76, 60–62.
- Heider, F. (1958), *The Psychology of Interpersonal Relations*, New York: John Wiley.
- Holland, P. W., and Leinhardt, S. (1975), "Local Structure in Social Networks," in *Sociological Methodology 1976*, ed. D. R. Heise, San Francisco: Jossey-Bass, 1–54.
- (eds.) (1979), *Perspectives in Social Network Research*, New York: Academic Press.
- (1981), "An Exponential Family of Probability Densities for Directed Graphs," *Journal of the American Statistical Association*, 76, 33–51.
- Katz, L. (1947), "On the Matrix Analysis of Sociometric Data," *Sociometry*, 10, 233–241.
- (1953), "A New Status Index Derived From Sociometric Analysis," *Psychometrika*, 18, 39–43.
- (1955), "Measurement of the Tendency Toward Reciprocation of Choice," *Sociometry and the Science of Man*, 18, 659–665.
- Katz, L., and Proctor, C. H. (1959), "The Concept of Configuration of Interpersonal Relations in a Group as a Time-Dependent Stochastic Process," *Psychometrika*, 24, 317–327.
- Lametz, K. (1978), "Small-Sample Comparisons of Exact Levels of Chi-Squared Goodness-of-Fit Statistics," *Journal of the American Statistical Association*, 73, 253–263.
- Leinhardt, S. (ed.) (1977), *Social Networks: A Developing Paradigm*, New York: Academic Press.
- Lorrain, F. P., and White, H. C. (1971), "Structural Equivalence of Individuals in Social Networks," *Journal of Mathematical Sociology*, 1, 49–80.
- Meyer, M. M. (1981), *Application and Generalizations of the Iterative Proportional Fitting Procedure*, unpublished Ph.D. thesis, University of Minnesota, School of Statistics.
- (1982), "Transforming Contingency Tables," *Annals of Statistics*, 10, 1172–1181.
- Moreno, J. L. (1934), *Who Shall Survive?* Washington, DC: Nervous and Mental Disease Publishing.
- Nelder, J. A., and Wedderburn, R. W. (1972), "Generalized Linear Models," *Journal of the Royal Statistical Society, Ser. A*, 135, 370–384.
- Reitz, K. P. (1982), "Using Loglinear Analysis With Network Data: Another Look at Sampson's Monastery," *Social Networks*, 4, 243–251.
- Sampson, S. F. (1969), *Crisis in a Cloister*, unpublished Ph.D. thesis, Cornell University, Dept. of Sociology.
- Wasserman, S. (1977), "Random Directed Graph Distributions and the Triad Census in Social Networks," *Journal of Mathematical Sociology*, 5, 61–86.
- (1978), "Models for Binary Directed Graphs and Their Applications," *Advances in Applied Probability*, 10, 803–818.
- (1980), "Analyzing Social Networks as Stochastic Processes," *Journal of the American Statistical Association*, 75, 280–294.
- White, H. C., Boorman, S. A., and Breiger, R. L. (1976), "Social Structure From Multiple Networks I: Blockmodels of Roles and Positions," *American Journal of Sociology*, 81, 730–780.