Binary interaction data, measuring the presence or absence of a relation between pairs of actors in a "dyadic interaction situation," are commonly gathered to study the social structure of the group of actors. Recent developments have made the statistical analysis of such data statistically easier and more substantively sophisticated. These developments allow researchers to simultaneously study several sociometric structural properties, such as reciprocity, differential popularity, and equivalence of actors. Building on this research, we review the stochastic models responsible for this breakthrough, and discuss methods for estimating expected values and model parameters. Throughout, we also highlight recent advances designed to incorporate nodal or actor attribute data into the relational data analysis. We conclude with an example illustrating these ideas based on conversational activities among actors in a group of eight people.

INTRODUCTION

Consider a collection or group of $g$ actors or individuals, and a sociometric relation, such as liking, helping, or communicating or simply interacting. Assume that we record

$$X_{ij} = 1 \quad \text{if actor } i \text{ "relates to" actor } j$$
$$= 0 \quad \text{otherwise}$$

for every ordered pair of actors. Clearly $X_{ij}$ need not equal $X_{ji}$. Such data, frequently called a univariate or a single relation social network or a realization of a directed graph, are common in small group studies designed to investigate the social structure of the group. Interest usually centers on the differential attractiveness, and to a lesser extent, the productivity or expansiveness of the actors. Also of primary importance is the level of reciprocity among the actors, and whether or not any tendencies toward cliquing or clustering are present.

This paper was presented at the 17th Annual Mathematical Psychology Meeting, Session on Measurement and Statistics, Society for Mathematical Psychology, August 21-23, 1984, Chicago, IL. This research was supported by a grant from the Research Board of the University of Illinois. We thank Phipps Arabie, Joseph Galaskiewicz, and two anonymous referees for comments on this paper. Send requests for reprints to Dr. Stanley Wasserman, Department of Psychology, and Department of Statistics, 603 East Daniel St., University of Illinois, Champaign, IL 61820.
A large number of statistical techniques currently exist for the analysis of a $g \times g$ array $X$ containing non-negative integers or continuous entries. In this more general case, $X_{ij}$ is usually interpreted as the strength of the relationship from actor $i$ to actor $j$. Kenny and LaVoie (1984) provide a review of many studies of this type. Some of these studies are analyzed as round robin designs (Warner, Kenny & Stoto, 1979; Wong, 1982, 1984). Others are variations of designs arising when actors are partitioned into subgroups based on attributes such as sex and race (Kraemer & Jacklin, 1979; Mendoza & Graziano, 1982), or, when symmetrized (forcing $X_{ij} = X_{ji}$), upper triangular contingency tables in which the same polytomy occurs on each margin (Larntz & Weisberg, 1976).

Statistical methodologists have just begun to address the binary version of this problem. Development of methods has lagged behind the substantive usage of this paradigm. Binary $X$ arrays are common in sociological studies, and arise frequently in the behavioral sciences when one observes a "clipped" or dichotomized quantity $X_{ij}$:

$$X_{ij} = 1 \quad \text{if} \quad Z_{ij} \geq t_{ij}$$

$$= 0 \quad \text{if} \quad Z_{ij} < t_{ij},$$

Here, $Z_{ij}$ is multivalued, maybe even continuous, and $t_{ij}$ is a threshold, above which we simply record 1 for $X_{ij}$; Or, such data can arise when one measures whether or not two actors (or "operational taxonomic units" in biological jargon) have an attribute (a single attribute) in common. There are many other situations yielding binary $X$ arrays.

A new approach to the statistical analysis of such data was initiated in 1977 by Holland and Leinhardt, and was described in detail in papers published by Holland and Leinhardt (1981) and Fienberg and Wasserman (1981). These authors introduced and generalized a model, or more precisely, an exponential family of distributions, termed $p_1$, for binary sociometric relations. In this paper, we discuss several important features of $p_1$: methods for calculating parameter estimates and their standard errors, and how to incorporate attribute variables into the analysis. We describe our approach to these problems, illustrating it with examples taken from Fienberg and Wasserman (1981). But first, we briefly review the theory surrounding $p_1$.

## 1. A Review of $p_1$

Consider a directed graph with $g$ nodes representing actors and a single binary sociometric relation such that at most one arc connects node $i$ to node $j$. We let $G$ denote the set of $g$ nodes, and describe the directed graph by means of a sociomatrix or adjacency matrix $X$ with elements $(X_{ij})$ defined earlier. The diagonal elements of $X$ are not of interest to us and are thus set to zero.
The models we are about to describe are best presented not using $X$, but in a
four-dimensional $g \times g \times 2 \times 2$ cross-classification $Y = (Y_{ijkl})$ with entries

$$Y_{ijkl} = 1 \quad \text{if} \quad (X_{ij}, X_{ji}) = (k, l)$$

$$= 0 \quad \text{otherwise}$$

for all $(i, j)$ and $k, l = 0, 1$. Note that the array $Y$ is symmetric in a special way,

$$Y_{ijkl} = Y_{jik}. \quad (1)$$

Define $\pi_{ijkl}$ as $P \{ (X_{ij}, X_{ji}) = (k, l) \}$ and let $\mu_{ijkl} = \log \pi_{ijkl}$. The $p_1$ exponential family of distributions postulates

$$\mu_{000} = \lambda_{ij}$$

$$\mu_{ij10} = \lambda_{ij} + \alpha_i + \beta_j + \theta$$

$$\mu_{ij01} = \lambda_{ij} + \alpha_i + \beta_i + \theta$$

$$\mu_{ij11} = \lambda_{ij} + \alpha_i + \alpha_j + \beta_i + \beta_j + 2\theta + \rho$$

subject to the constraints

$$\exp \{ \mu_{000} \} + \exp \{ \mu_{ij10} \} + \exp \{ \mu_{ij01} \} + \exp \{ \mu_{ij11} \} = 1 \quad (3)$$

for all dyads, and

$$\sum_{i=1}^{g} \alpha_i = \sum_{j=1}^{g} \beta_j = 0, \quad (4)$$

which are the standard, ANOVA-like constraints. Applications of this model are
given by Holland and Leinhardt (1981), Fienberg and Wasserman (1981), Fien-
berg, Meyer, and Wasserman (1985), and Rietz (1982).

The model postulates the admittedly strong assumption that the dyads are
statistically independent. This assumption of dyadic independence is open to attack; however, the generalization of $p_1$ from an “independent dyadic choice” model (as it is termed by some) to general dyadic dependence is quite difficult. One approach to solving this problem is to use triadic analyses to complement the models presented here (see Holland & Leinhardt, 1981). Another untired suggestion is to sample
dyads, thereby reducing or eliminating dependencies. Frank and Strauss (1983) use
the concept of a Markov graph (a special case of a Markov random field) to
introduce dependencies, but much more (hard) work needs to be done.

We note here that fitting $p_1$ to data via maximum likelihood is an iterative
process, and because of the sparseness of the $Y$ array, convergence can be quite slow. Strategies to speed up convergence are presented in Section 3.

Special cases of $p_1$ are obtained by setting various parameters of $p_1$ to zero. Tests
for appropriateness of various sets of the parameters (such as all the $\beta$’s are zero)
are described in detail in Fienberg and Wasserman (1981). As noted by these authors and others (e.g., Haberman, 1981), the standard large sample theory for situations such as this one (log linear models for categorical data—see Bishop, Fienberg, & Holland, 1975, Chap. 14) does not apply since the degrees of freedom for all tests (except the one setting only $p$ to zero) increase as the sample size ($g$) increases. Our approach to this problem, as described in Wasserman and Galaskiewicz (1984), is simply to compare the conditional likelihood ratio statistics to their appropriate degrees of freedom, and use them as summary statistics. We do not attach $p$-values to them. A large test statistic, relative to its degrees of freedom, implies the parameters in question are statistically important, and probably should not be dropped from the model. Last, we note that the Appendix of Fienberg et al. (1985) describes in detail degrees of freedom calculations.

2. Generalizations of $p_1$

One of the most important generalizations of this approach attempts to alleviate the inference problems inherent in using $p_1$. Again assuming simple binary interactions, we further postulate that prior to any statistical analysis, the $g$ actors have been partitioned into $K$ subgroups usually based on nodal/actor attributes. We call this subgrouping an a priori "stochastic blockmodel," to draw parallels between it and the a posteriori clusterings of sociometric data termed blockmodels (White, Boorman, & Breiger, 1976; Breiger, Boorman, & Arabie, 1975; Arabie, Boorman, & Levitt, 1978). A blockmodel is usually based on Lorrain and White's (1971) concept of structural equivalence. Two actors $i$ and $j$ exhibit this property if and only if for any individual $k$ in the group,

$$X_{ik} = X_{jk} \quad \text{and} \quad X_{ki} = X_{kj}.$$  

Fienberg and Wasserman (1981), and later Holland, Laskey, and Leinhardt (1983), introduced the analogous, but alternative, notion of stochastic equivalence, a more natural concept for statistical modeling, which concerns us here. Assuming that $p_1$ is appropriate, two actors $i$ and $j$ are called stochastically equivalent if and only if,

$$\alpha_i = \alpha_j \quad \text{and} \quad \beta_i = \beta_j.$$  

Notice that stochastic equivalence is weaker than structural equivalence. For two actors to be structurally equivalent, they must relate to all other actors in the group in exactly the same way, a highly unlikely event, whereas, two actors are stochastically equivalent only if their tendencies to initiate and receive interactions are equal (in a probabilistic sense). Two actors need not have identical rows and identical columns in the $X$ array to be stochastically equivalent. It is noteworthy that strict usage of structural equivalence is rarely adopted; rather, "lean fit" blockmodels are calculated based on a much weaker notion of structural equivalence: actors are blocked together if they have just a few interactions in common. No one knows how large or how small a "few" is. The lack of a "lean fit rule," a threshold
beyond which zeroblocks (blocks of mostly zeros) are considered oneblocks, makes blockmodel construction appear rather arbitrary. Use is made of the algorithm CONCOR (Breiger et al., 1975) to find partitions. The concreteness of the stochastic equivalence criterion is a definite advantage; furthermore, using it to construct subgroupings should prove it to be more flexible, since $\alpha_i$ and $\alpha_j$, and $\beta_i$ and $\beta_j$ need only be within a couple standard errors of each other for two actors to be blocked together.

Unfortunately, the details of a posteriori usage of our principle have yet to be specified. Many problems must be worked out, especially combinatorial difficulties which are common in partitioning algorithms (see, for example, Duran & Odell, 1974, Chap. 2). Holland, Laskey, and Leinhardt (1983) take a few tentative steps in this direction. This state of affairs is in sharp contrast to the a posteriori operationalization of structural equivalence using CONCOR, which is easy, quick computationally, and despite the arbitrariness of its implementation, substantively meaningful. For our purposes here, we shall use stochastic equivalence to obtain a priori partitions of the actors into subgroupings, based on exogenous information, and hope that in the near future, research will compare these two types of equivalences, in both the a priori and a posteriori settings.

To detail our approach, assume that prior to any relational data analysis, actors can be partitioned into $K$ subgroups, where, for a pair of actors $i$ and $j$ in subgroup $s$,

$$\alpha_i = \alpha_j = \alpha^{(s)}$$

$$\beta_i = \beta_j = \beta^{(s)}.$$  \hspace{1cm} (6)

Usually, this partition is based on one or more nodal attributes—for example, subgroup 1 may consist of white males, subgroup 2 of white females, etc. Of substantive interest is how likely it is that actors in one subgroup relate to actors in the same and other subgroups, and whether or not, given a set of nodal attributes, Eq. (6) is appropriate. Notice that this approach is a simplification of $p_1$, since stochastic equivalence reduces the numbers of $\alpha$ and $\beta$ parameters from $g$ to $K$. This use of categorical nodal characteristics, merging such actor attributes with relational data into a statistical model, is quite novel. Recently, Arabie (1981, 1984) and Frank and Wellman (1984) have also presented methods along these lines; Arabie using linear discriminant analysis of attribute variables to substantiate block model partitions, and Frank and Wellman, using categorical data methods similar to ours for the same purpose.

We label the $K$ subgroups defined by the categorical nodal attributes $G_1$, $G_2$, ..., $G_K$, where $G_s$ has $g_s$ actors. The most general $p_1$ subgroup models postulate that

$$\log P\{D_{ij} = (0, 0)\} = \lambda^{(rs)}$$

$$\log P\{D_{ij} = (1, 0)\} = \lambda^{(rs)} + \theta^{(rs)}$$

$$\log P\{D_{ij} = (0, 1)\} = \lambda^{(rs)} + \theta^{(sr)}$$

$$\log P\{D_{ij} = (1, 1)\} = \lambda^{(rs)} + \theta^{(rs)} + \theta^{(sr)} + \rho^{(rs)}$$  \hspace{1cm} (7)
for \( i \in G \), and \( j \in G \), subject to constraints (3). Note that the number of unknown parameters—\( K^2 \) choice effects \( \theta \), \((1/2)(K^2 + K)\) reciprocity effects \( \rho \), and \((1/2)(K^2 + K)\) values of \( \lambda \)—is now fixed, being determined by the external number of attribute categories. Chi-square likelihood ratio test statistics are easily calculated and correspond to valid tests—such as setting subsets of parameters equal to zero. Degrees of freedom for these tests are fixed, like \( K \). This fact alleviates the testing problems inherent with direct usage of \( p_1 \).

There are several special cases of this general subgroup model (7). Setting \( \rho^{(rs)} = 0 \) for all subgroups yields a logit model (see Sect. 3 of Fienberg et al., 1985). A \( p_1 \) for subgroups specifies that

\[
\rho^{(rs)} = \rho, \quad \text{for all } r \text{ and } s
\]

and

\[
\tilde{\theta}^{(rs)} = \theta + \alpha^{(r)} + \beta^{(s)}
\]

(see Fienberg & Wasserman, 1981). Fienberg et al. (1985) point out that when \( K < g \), a comparison of likelihood ratio statistics for \( p_1 \) and for this model indicates whether the actors are stochastically equivalent. We note that Wasserman and Galaskiewicz (1984) present other \( p_1 \) subgroup models, all based on simplifications of the parameters in Eqs. (7). For example, we could allow

\[
\tilde{\theta}^{(rs)} = \theta + \alpha^{(r)} + \beta^{(s)} + (\alpha\beta)^{(rs)}
\]

and/or that

\[
\rho^{(rs)} = \rho + \rho^{(r)} + \rho^{(s)}.
\]

The log likelihood function for the parameters in model (7) (see Fienberg & Wasserman, 1981, pp. 182–183) shows that the minimal sufficient statistics for all of the parameters in Eqs. (7)–(11) are lower dimensional margins of the four-dimensional array \( w \), which has entries

\[
W_{rst} = \sum_{i \in G} \sum_{j \in G} y_{ijkl}.
\]

The \( w \)'s are the counts of the four different dyad types within and between subgroups, computed simply by pooling the \( 2 \times 2 \) submatrices of the \( y \) array. Model fitting and parameter estimation details are given in the next two sections.

### 3. Model Fitting Strategies

As we have stated, \( p_1 \) itself is simply a special case of a subgroup model, obtained by letting \( K \), the number of subgroups, equal \( g \). Consequently, in this section, when presenting the methods for fitting these models, we illustrate with a subgroup
model, operating directly on the $K \times K \times 2 \times 2$ array $W$. To apply these results to $p_1$, assume the $p_1$ subgroup model, replace $W$ by $Y$, and let $K = g$ in the following equations.

All of the models described in the previous section are log linear, and, as Fienberg and Wasserman (1981) point out, maximum likelihood estimates of the $\{\pi_{ijkl}\}$ are found by calculating a $\hat{\pi}$ array with sufficient statistics equal to their expected values under the model in question. One method of calculating $\hat{\pi}$ is based on Darroch and Ratcliff's (1972) results that allow one to construct an iterative algorithm to solve these maximum likelihood equations. This is the approach taken by Holland and Leinhardt (1981) and Fienberg and Wasserman (1981); using an iterative proportional fitting procedure (IPFP). Another method, also discussed here, a general Newton–Raphson (NR) algorithm, is described in detail in Haberman (1978). We note that an interesting realization by Fienberg and Wasserman (see "Comments" following Holland & Leinhardt, 1981) and an important breakthrough by Meyer (1981, 1982) allowed us to simplify substantially the original model fitting guidelines of Holland and Leinhardt.

Details for specific models (i.e., the minimal sufficient statistics and which margins to fit) are given by Fienberg and Wasserman (1981) and Wasserman and Galaskiewicz (1984). Furthermore, to compute likelihood ratio goodness-of-fit statistics, we simply divide the usual statistic computed on $y$ (after calculating $\hat{\pi}$ from $w$) by 2 to adjust for the duplication of cells in $y$ (Eq. (1)). The most practical outcome of Meyer's theoretical results is that we can use standard IPFP or NR computer programs for modeling contingency tables to fit our models to data. The remainder of this section is a summary of these programs, examples, and suggested guidelines based on our experience, for using them.

We studied both a direct and an indirect approach to computing maximum likelihood estimates $\hat{\pi}$. The indirect approach uses IPFP to calculate this array. The estimated $u$-terms (log linear model parameters in Bishop et al., 1975, notation) for the models fit to the four-dimensional arrays are then simple functions of the $\{\hat{\pi}_{ijkl}\}$. The direct approach focuses specifically on the estimated $u$-terms. It sets up an iterative series of approximations to find $u$-terms that maximize the log likelihood function. We report here on the following "canned" software programs designed to fit log linear models to categorical data using any of these approaches:

(i) SPSS-X LOGLINEAR.
(ii) CTAB.
(iii) SPSS-X (Release 2) HILOGLINEAR.
(iv) BMDP BMDP4F.

The first one is direct, and the last three are indirect. All but the second are widely available. CTAB is a simple IPFP FORTRAN program based on Haberman's (1972) algorithm, and is described in Haberman (1978), Fienberg (1980), and Kennedy (1983). Programs very similar to CTAB are also known by other names, such as ECTA.
The LOGLINEAR procedure in SPSS-X release 1.1, run on an IBM 4341 model M2 with OS/MVT operating system (SPSS, Inc., 1983), uses the NR algorithm to produce maximum likelihood estimates of $u$-terms. We investigated three indirect model-fitting programs, all of which use IPFP. The first, CTAB (see Wasserman, 1982), is an interactive program that uses IPFP to calculate maximum likelihood estimates of expected cell counts. SPSS-X release 2, a recent (1984) update of SPSS-X, contains HILOGLINEAR, an IPFP for multidimensional categorical data arrays. Documentation is available from SPSS, Inc. HILOG (Hierarchical Loglinear procedure) is designed for model selection, as it uses backward elimination from the saturated model to choose the simplest subset model that fits the data. Lastly, the IPFP BMDP4F, from the 1983 release of BMDP (Dixon, 1983) run on a CDC CYBER 175, is the most flexible of the three IPFP routines discussed here, but is not interactive. Considerably more detail about these programs, and how to utilize them in this context, is available from the authors.

We used these four “canned” packages to study data analyzed by Fienberg and Wasserman (1981). We fit $p_1$ to the information flows existing among a collection of 16 business organizations, and $p_1$ subgroup models to the flows of support among a full set of 73 organizations categorized by two binary attribute variables: (a) whether each organization is owned by people in the community (local) or people outside the community (extralocal); and (b) whether each organization has public or private ownership. Thus, we have a $16 \times 16$ array $x$ (given in Table 1) which we transform into a $16 \times 16 \times 2 \times 2$ array $y$, and a $2^2 \times 2^2 \times 2 \times 2$ array $w$ (given in Table 2) based on a partition of actors into $K = 2^2 = 4$ subgroups. These tables, especially the first, are of interest because their analysis is not without problems.

The reader may want to first become acquainted with the Fienberg and Wasserman (1981) analysis of these tables.

The software programs that we used did not perform as desired with the 16-actor information-flow relation. The indirect methods “converged” to the default convergence criterion of 0.01 in approximately 4000 iterations. However, the $u$-terms had not truly converged. Throughout successive groups of iterations, the $u$-term estimates changed—often in the first decimal place. Even after an additional 2000 iterations, the situation was no better. The direct method also failed to converge. SPSS-X's LOGLINEAR procedure terminated execution after several hours of computing time (on a CYBER 175) because of underflow and overflow errors.

The problems we encountered probably resulted from both the size and the nature of $16 \times 16$ $X$ array. Sparse contingency tables require a large number of iterations when IPFP routines must fit many margins. The Newton–Raphson algorithm estimates $u$-terms directly; here, there are 319 (unique) $u$-terms to be

---

1 We thank Ken R. Smith of the University of Utah and J. Clyde Mitchell of Nuffield College, Oxford, for questioning Fienberg and Wasserman's analysis of the data in Table 1. As we point out here, these authors did not let $p_1$ fitted values converge, preferring to stop at 350 iterations of an IPFP program similar to CTAB. In some respects, this paper serves as an answer to the questions they raised.
TABLE 1

Sociomatrix of Business Organizations Based on Information Forms from Fienberg and Wasserman (1981) and Associated Statistics

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
<th>16</th>
</tr>
</thead>
<tbody>
<tr>
<td>Out-degree</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Farm Equipment Co.</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2. Clothing Mfg. Co.</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3. Farm Supply Co.</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>4. Mechanical Co.</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5. Electrical Equip. Co.</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6. Metal Products Co.</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7. Music Equipment Co.</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>8. 1st Towertown Bank</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>9. Towertown Savings and Loan</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>10. Bank of Towertown</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11. 2nd Towertown Bank</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>12. Brinkman Law Firm</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>13. Cater Law Firm</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>14. Knapp Law Firm</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>15. Towertown News</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>16. WTWR Radio</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Note. (120) = 120 dyads; of these, 43 are mutuals, 6 are asymmetric, and 71 are null.

estimated. Therefore, the algorithm requires the inversion of a 319 × 319 matrix at each iteration. Not only is this procedure time consuming, but it is subject to precision errors, especially when matrix entries have extreme (in this case very small) values.

The subgroup support-flow relation fared much better with regard to the u-term estimation. CTAB converged to a tolerance of 10^{-10} after just 30 iterations, and computed u-terms that had also converged. Similarly, BMDP4F converged in a small number of iterations, produced u-terms identical to CTAB, and computed standard errors as well.

The direct method was also successful. SPSS-X's LOGLINEAR procedure produced u-terms identical to CTAB and BMDP4F. LOGLINEAR also produced standard errors of u-terms, and reached convergence in a very reasonable amount of computing time.

The subgroup relation probably fared well because the w array is small and less sparse. Furthermore, the w array contains reasonably large counts, not just zeros and ones. Iterative proportional fitting routines manipulate less sparse contingency tables quite well. The Newton–Raphson process only needed to estimate 31 parameters in this problem, hence it manipulated 31 × 31 matrices rather than 319 × 319 matrices.
PARAMETER ESTIMATION

TABLE 2
w-Arrays for Intra- and Intersubgroup Choices among the Four Subgroups
(from Fienberg & Wasserman, 1981)

<table>
<thead>
<tr>
<th></th>
<th>$G_1$</th>
<th>$G_2$</th>
<th>$G_3$</th>
<th>$G_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(a) Information</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G_1$</td>
<td>404</td>
<td>44</td>
<td>422</td>
<td>20</td>
</tr>
<tr>
<td>$G_2$</td>
<td>44</td>
<td>158</td>
<td>30</td>
<td>100</td>
</tr>
<tr>
<td>$G_3$</td>
<td>382</td>
<td>24</td>
<td>356</td>
<td>29</td>
</tr>
<tr>
<td>$G_4$</td>
<td>24</td>
<td>32</td>
<td>20</td>
<td>57</td>
</tr>
<tr>
<td>(b) Money</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G_1$</td>
<td>424</td>
<td>81</td>
<td>433</td>
<td>97</td>
</tr>
<tr>
<td>$G_2$</td>
<td>81</td>
<td>64</td>
<td>24</td>
<td>18</td>
</tr>
<tr>
<td>$G_3$</td>
<td>454</td>
<td>4</td>
<td>435</td>
<td>0</td>
</tr>
<tr>
<td>$G_4$</td>
<td>4</td>
<td>0</td>
<td>27</td>
<td>0</td>
</tr>
<tr>
<td>(c) Support</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$G_1$</td>
<td>484</td>
<td>56</td>
<td>463</td>
<td>44</td>
</tr>
<tr>
<td>$G_2$</td>
<td>56</td>
<td>54</td>
<td>34</td>
<td>31</td>
</tr>
<tr>
<td>$G_3$</td>
<td>402</td>
<td>26</td>
<td>361</td>
<td>6</td>
</tr>
<tr>
<td>$G_4$</td>
<td>26</td>
<td>8</td>
<td>60</td>
<td>35</td>
</tr>
</tbody>
</table>

4. Estimation of Parameters and Standard Errors

We now turn to the primary purpose of the paper. Assume that one has computed the estimated $p_{ij}$ probabilities of dyadic interaction. In this section we describe how to use these probability estimates, contained in the $\hat{p}$ array, to calculate estimates of model parameters and the associated estimated asymptotic standard errors. We show how one can manipulate the $\hat{p}$'s to obtain parameter estimates and how these parameter estimates can be simply derived as functions of
the estimated \( u \)-terms from the log linear model fit to either \( y \) or \( w \). Further, because of this latter simple relationship, the estimated asymptotic standard deviations (EASD) of model parameters are also functions of the EASD's of the \( u \)-terms.

Fienberg and Wasserman (1981, pp. 165–166) present one method for calculating \( p_1 \) parameter estimates. Specifically, they show that since

\[
\rho = \log\left(\frac{\hat{\pi}_{11} \hat{\pi}_{00}}{\hat{\pi}_{10} \hat{\pi}_{01}}\right)
\]

is a constant for all dyads, we can average the estimates of \( \rho \) over all dyads to obtain the MLE

\[
\hat{\rho} = \frac{1}{\binom{2}{2}} \sum_{i<j} \log\left(\frac{\hat{\pi}_{ij0} \hat{\pi}_{0j1}}{\hat{\pi}_{i0j} \hat{\pi}_{01j}}\right).
\]

Using the same general strategy, they suggest computing

\[
\hat{\alpha}_i - \hat{\alpha}_{i'} = \log\left(\frac{\hat{\pi}_{ij0} / \hat{\pi}_{i00}}{\hat{\pi}_{ij0} / \hat{\pi}_{i00}}\right) \quad (i \neq i')
\]

over many values of \( j \) and then averaging; similarly,

\[
\hat{\beta}_j - \hat{\beta}_{j'} = \log\left(\frac{\hat{\pi}_{ij1} / \hat{\pi}_{i00}}{\hat{\pi}_{ij1} / \hat{\pi}_{i00}}\right) \quad (j \neq j')
\]

can be computed over many values of \( i \) and then averaged. Reitz (1982) gives more details of this method.

There are several problems with this approach. The first is what to do when some of the \( \hat{\pi}_{ij0} \)'s are zero. The second is exactly how many values of \( \hat{\alpha}_i - \hat{\alpha}_{i'} \) and \( \hat{\beta}_j - \hat{\beta}_{j'} \) to compute. We were basically displeased with the arbitrariness of the method. Consequently, we looked for a more logically defensible method.

The best, although rather time consuming, method that we found utilizes the fact that the logarithms of the estimated probabilities are linear functions of the model parameters. We have, assuming \( p_1 \),

\[
\log\left(\frac{\hat{\pi}_{ij0} / \hat{\pi}_{i00}}{\hat{\pi}_{i0j} / \hat{\pi}_{i00}}\right) = \hat{\theta} + \hat{\alpha}_i + \hat{\beta}_j
\]

\[
\log\left(\frac{\hat{\pi}_{ij0} / \hat{\pi}_{000}}{\hat{\pi}_{i0j} / \hat{\pi}_{00j}}\right) = \hat{\theta} + \hat{\alpha}_j + \hat{\beta}_i
\]

\[
\log\left(\frac{\hat{\pi}_{ij1} / \hat{\pi}_{i00}}{\hat{\pi}_{i0j} / \hat{\pi}_{i00}}\right) = 2\hat{\theta} + \hat{\alpha}_i + \hat{\alpha}_j + \hat{\beta}_i + \hat{\beta}_j + \hat{\rho}.
\]

Consequently, the \( 2g + 2 \) model parameters\(^2\) can be found by least squares using a "response" vector of length \( 3(\binom{2}{2}) \)—three entries for each dyad as given in Eq. (15) which we call \( \tilde{x} \)—and a vector of parameters\(^3\) of length \( 2g \). One must construct a

\(^2\) Remember the \( \{A_1\} \) are present simply to ensure that the estimated probabilities in each \( 2 \times 2 \) table sum to unity; hence they are rarely of interest.

\(^3\) Also remember that \( \hat{\theta}_x = -\sum_{i=1}^{2g} \hat{\alpha}_i \) (similarly for \( \hat{\beta}_x \)), so we really have just \( 2g \) free parameters.
A parameter estimation technique involves using a matrix $X$ containing 0's, 1's, -1's, and 2's using effect coding as in ANOVA. Details are given in the Appendix. It is easy to verify that the estimated parameters equal $(X'X)^{-1} X'y$ and are the unique MLE's. Note that to implement this method we need only $\hat{\theta}$ and the design matrix $X$. Further, this method can be used for any model (including those for subgroups) simply by constructing $X$ appropriately.

Another approach can be implemented when the canned program, used to calculate $\hat{\theta}$, also calculates estimated $u$-terms. Again assuming $p_1$, we fit the no three-factor interaction model to $y$ so that

$$
\begin{align*}
\mu_{ijkl} &= \log \pi_{ijkl} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{4(l)} + u_{12(\bar{i})} \\
+ u_{13(\bar{k})} + u_{14(\bar{i})} + u_{23(\bar{j})} + u_{24(\bar{j})} + u_{34(\bar{k})} \\
(\text{for } k, l = 0, 1)
\end{align*}
$$

is obtained. Equating this model to the standard $p_1$ Eqs. (2), yields

**Theorem.** Assuming we fit $p_1$ to a single relational data set for $g$ actors using a no three-factor interaction model for the translated $y$ array, then

$$
\begin{align*}
\alpha_i &= 2\hat{u}_{13(i1)} = 2\hat{u}_{24(i1)} \\
\beta_j &= 2\hat{u}_{23(j1)} = 2\hat{u}_{14(j1)} \\
\theta &= 2\hat{u}_{3(i1)} + 2\hat{u}_{34(10)} = 2\hat{u}_{4(11)} - 2\hat{u}_{34(11)} \\
\hat{\rho} &= 4\hat{u}_{34(11)}
\end{align*}
$$

where the estimated $u$-terms arise from the model (16).

**Proof.** From Eq. (16), note that

$$
\begin{align*}
\mu_{ij10} - \lambda_{i00} &= u_{3(i1)} - u_{3(10)} + u_{34(10)} - u_{34(00)} \\
+ u_{13(i1)} - u_{13(10)} + u_{23(j1)} - u_{23(j0)}
\end{align*}
$$

so that, from Eq. (13),

$$
\begin{align*}
(\mu_{ij10} - \mu_{ij00}) - (\mu_{ij'10} - \mu_{ij'00}) &= \alpha_i - \alpha_{i'} \\
= u_{13(i1)} - u_{13(i'1)} - u_{13(i0)} + u_{13(i'0)} \\
= 2u_{13(i1)} - 2u_{13(i'1)}
\end{align*}
$$

since $u_{13(i0)} + u_{13(i1)} = 0$ for all $i$. Clearly,

$$
\alpha_i = 2u_{13(i1)} = 2u_{24(i1)} \quad \text{(because of symmetries in $y$)}
$$

is a solution to the above equations; furthermore, the $\{\alpha_i\}$ sum to zero as is required. All of the above calculations also hold for the estimated $\mu$'s and $u$'s. Similarly,

$$
\beta_j = 2\hat{u}_{23(j1)} = 2\hat{u}_{14(j1)}.
$$
And it is easy to verify, using Eq. (12), that
\[ \hat{\rho} = \hat{\mu}_{000} + \hat{\mu}_{111} - \hat{\mu}_{101} - \hat{\mu}_{011} = 4\hat{u}_{34(11)}. \]

Last, from Fienberg and Wasserman (1981), we have
\[ \hat{\theta} = \log \hat{\pi}_{j10} - \log \hat{\pi}_{j00} - \hat{\alpha}_i - \hat{\beta}_j \]
for all \((i, j)\), so that
\[
\hat{\theta} = \hat{u}_{3(1)} - \hat{u}_{3(0)} + \hat{u}_{34(10)} - \hat{u}_{34(00)} + \hat{u}_{13(i)} - \hat{u}_{13(i0)} \\
+ \hat{u}_{23(j1)} - \hat{u}_{23(j0)} - 2\hat{u}_{13(i1)} - 2\hat{u}_{14(j1)} \\
= \hat{u}_{3(1)} - \hat{u}_{3(0)} + \hat{u}_{34(10)} - \hat{u}_{34(00)} \\
= 2\hat{u}_{3(1)} + 2\hat{u}_{34(10)} = 2\hat{u}_{4(1)} - 2\hat{u}_{34(11)}. \quad \text{Q.E.D.}
\]

Note that this theorem provides a check on whether the estimated probabilities have actually converged. If \( \hat{u}_{13(i1)} \neq \hat{u}_{24(i1)} \), or \( \hat{u}_{23(j1)} \neq \hat{u}_{14(j1)} \), an IPFP or NR algorithm needs to run more iterations.

A theorem such as the one above can be proved for any special case or generalization of \( p_1 \). In other words, for any relative of \( p_1 \), the parameters of the model are simple estimates of the \( u \)-terms belonging to the model's equivalent log linear model fit to the four-dimensional \( y \) or \( w \) array. Some of these models, their equivalent log linear models, and the relationships between the model parameters and \( u \)-terms are given in Table 3.

To calculate the estimated asymptotic standard deviations (EASDs) we use the fact that all model parameter estimates are simple linear functions of the estimated \( u \)-terms of the log linear models fit to \( y \) or \( w \). So, for example, for \( p_1 \):
\[
\text{EASD}(\hat{\alpha}_i) = \text{EASD}(2\hat{u}_{13(i1)}) = \sqrt{2} \text{EASD}(\hat{u}_{13(i1)}), \quad \text{for all } i; \\
\text{EASD}(\hat{\beta}_j) = \text{EASD}(2\hat{u}_{23(j1)}) = \sqrt{2} \text{EASD}(\hat{u}_{23(j1)}), \quad \text{for all } j; \\
\text{EASD}(\hat{\rho}) = \text{EASD}(4\hat{u}_{34(11)}) = 2 \text{EASD}(\hat{u}_{34(11)}).
\]

The EASDs of the estimated \( u \)-terms are linear combinations of the fitted cell means; for theory see Haberman (1978, chap. 4). As we noted in the previous study, several canned programs provide them as output.

CTAB, BMDP, and SPSS-X converged quickly when fitting the \( p_1 \) model to the four-group support relation. The converged \( u \)-term estimates from the three procedures agreed to the third decimal place (CTAB and BMDP only print \( u \)-terms to three decimal places). The \( u \)-terms needed to compute \( p_1 \) parameter estimates using the method described above in the theorem are in Table 4, followed by computations of the parameters. These computed parameter estimates agreed to three decimal places with least squares estimates obtained from CTAB fitted values (see Appendix).
### TABLE 3
Models, Model Parameters, and Equivalent u-Terms

#### I. Models for ungrouped actors

<table>
<thead>
<tr>
<th>Model</th>
<th>Log linear model for y</th>
<th>Parameters</th>
<th>Equivalent u-terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) ( p_1 )</td>
<td>(ii) Equal popularity</td>
<td>( \theta )</td>
<td>( 2u_{3(1)} + 2u_{34(10)} )</td>
</tr>
<tr>
<td></td>
<td>( { x_i } )</td>
<td>( { \beta_j } )</td>
<td>( 2u_{3(3)} )</td>
</tr>
<tr>
<td></td>
<td>( { \rho } )</td>
<td>( 2u_{34(11)} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[12][13][14][23][24][34]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ii) ( p_1 )</td>
<td>(iii) Differential reciprocity</td>
<td>( \theta )</td>
<td>( 2u_{3(1)} + 2u_{34(10)} )</td>
</tr>
<tr>
<td></td>
<td>( { x_i } )</td>
<td>( { \beta_j } )</td>
<td>( 2u_{3(3)} )</td>
</tr>
<tr>
<td></td>
<td>( { \rho } )</td>
<td>( 2u_{34(11)} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[12][13][24][34]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iii) ( p_1 )</td>
<td>( { x_i } )</td>
<td>( { \beta_j } )</td>
<td>( 2u_{3(3)} )</td>
</tr>
<tr>
<td></td>
<td>( { \rho } )</td>
<td>( 2u_{34(11)} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[12][134][234]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### II. Models for grouped actors

<table>
<thead>
<tr>
<th>Model</th>
<th>Log linear model for ( w )</th>
<th>Parameters</th>
<th>Equivalent u-terms ((i \in G_i, j \in G_j))</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) ( p_1 ) for subgroups</td>
<td>(ii) Interacting subgroups</td>
<td>( \theta )</td>
<td>( 2u_{3(1)} + 2u_{34(10)} )</td>
</tr>
<tr>
<td></td>
<td>( { x_i } )</td>
<td>( { \beta_j } )</td>
<td>( 2u_{3(3)} )</td>
</tr>
<tr>
<td></td>
<td>( { \rho } )</td>
<td>( 2u_{34(11)} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[12][13][14][23][24][34]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ii) ( p_1 )</td>
<td>(iii) Interacting subgroups/ differential reciprocity</td>
<td>( \theta )</td>
<td>( 2u_{3(1)} + 2u_{34(10)} )</td>
</tr>
<tr>
<td></td>
<td>( { x_i } )</td>
<td>( { \beta_j } )</td>
<td>( 2u_{3(3)} )</td>
</tr>
<tr>
<td></td>
<td>( { \rho } )</td>
<td>( 2u_{34(11)} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[123][134][34]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(iii) ( p_1 )</td>
<td>(iv) Saturated</td>
<td>( \theta )</td>
<td>( 2u_{3(1)} + 2u_{34(10)} )</td>
</tr>
<tr>
<td></td>
<td>( { x_i } )</td>
<td>( { \beta_j } )</td>
<td>( 2u_{3(3)} )</td>
</tr>
<tr>
<td></td>
<td>( { \rho } )</td>
<td>( 2u_{34(11)} )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>[1234]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Similarly, \( p_j \) parameter standard errors are computed as functions of \( u \)-term estimated standard errors. The latter are produced as output by SPSS-X and BMDP. SPSS-X computes standard errors of all unique \( u \)-terms. In our four-group case, for example, \( \hat{u}_{13(11)} + \hat{u}_{13(21)} + \hat{u}_{13(31)} + \hat{u}_{13(41)} = 0 \), so SPSS-X computes standard errors of \( \hat{u}_{13(11)} \), \( \hat{u}_{13(21)} \), and \( \hat{u}_{13(31)} \) only. The standard error of \( \hat{u}_{13(41)} \) may be computed using the relationship

\[
\hat{\sigma} = 2\hat{u}_{34(11)} = 4\hat{u}_{34(11)} = 4(0.648) = 2.592
\]

\[
\hat{\beta} = 2\hat{u}_{4(1)} - 2\hat{u}_{34(11)} = 2(-0.567) - 2(0.648) = -2.430
\]

SPSS-X produces the variance-covariance matrix of \( u \)-term estimates, from which the needed covariances may be read. BMDP prints \( u \)-terms and ratios of \( u \)-terms to their standard errors. Standard errors, then, may be found by dividing. However, more accurate values are obtained using the variance-covariance matrix of \( u \)-term estimates, as with SPSS-X. Standard errors computed by BMDP and SPSS-X for this problem agreed to three significant digits; the results are in Table 5.

Finally once parameter estimates and standard errors have been calculated, we may test parameters for significance. Ratios of parameters to their standard errors are presented in Table 6. These ratios are distributed asymptotically as standard normal random variables and so we use tabulated percentage points to test significance.
### TABLE 5

Standard Errors of Parameters for the Four-Group Support Relation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>EASD</th>
<th>EASD</th>
<th>EASD</th>
<th>EASD</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{\alpha}_{13(11)}$</td>
<td>0.0390</td>
<td>EASD($\hat{\alpha}_{23(11)}$)</td>
<td>0.0377</td>
<td>EASD($\hat{\alpha}_{34(11)}$)</td>
</tr>
<tr>
<td>$\hat{\alpha}_{13(21)}$</td>
<td>0.0419</td>
<td>EASD($\hat{\alpha}_{23(21)}$)</td>
<td>0.0433</td>
<td>EASD($\hat{\alpha}_{41(11)}$)</td>
</tr>
<tr>
<td>$\hat{\alpha}_{34(31)}$</td>
<td>0.0412</td>
<td>EASD($\hat{\alpha}_{23(31)}$)</td>
<td>0.0363</td>
<td>EASD($\hat{\alpha}_{41(11)}$)</td>
</tr>
<tr>
<td>$\hat{\alpha}_{13(41)}$</td>
<td>0.0731</td>
<td>EASD($\hat{\alpha}_{23(41)}$)</td>
<td>0.0665</td>
<td></td>
</tr>
</tbody>
</table>

### TABLE 6

Ratios of $p_1$ Parameter Estimates to Their EASDs for the Four-Group Support Relation

| Parameter | Ratio |  | Ratio |  | Ratio |  |
|-----------|-------|  |       |  |       |  |
| $\hat{\alpha}_1$ | $\frac{0.252}{0.0552}$ | 4.565 | $\frac{-0.266}{0.0533}$ | $-4.991$ | $\frac{0.214}{0.0593}$ | 3.609 | $\frac{-0.742}{0.0612}$ | $-12.124$ |
| $\hat{\alpha}_2$ | $\frac{-0.092}{0.0583}$ | -1.578 | $\frac{0.510}{0.0513}$ | 9.942 | $\frac{-0.372}{0.1034}$ | -3.598 | $\frac{0.498}{0.0926}$ | 5.378 |
| $\hat{\alpha}_3$ | $\frac{0.252}{0.0552}$ | 4.565 | $\frac{-0.266}{0.0533}$ | $-4.991$ | $\frac{0.214}{0.0593}$ | 3.609 | $\frac{-0.742}{0.0612}$ | $-12.124$ |
| $\hat{\alpha}_4$ | $\frac{-0.092}{0.0583}$ | -1.578 | $\frac{0.510}{0.0513}$ | 9.942 | $\frac{-0.372}{0.1034}$ | -3.598 | $\frac{0.498}{0.0926}$ | 5.378 |
| $\hat{\beta}_1$ | $\frac{2.592}{0.0476}$ | 54.454 | $\frac{-2.430}{0.0820}$ | -29.634 | $\frac{0.214}{0.0593}$ | 3.609 | $\frac{-0.742}{0.0612}$ | $-12.124$ |
| $\hat{\beta}_2$ | $\frac{-0.092}{0.0583}$ | -1.578 | $\frac{0.510}{0.0513}$ | 9.942 | $\frac{-0.372}{0.1034}$ | -3.598 | $\frac{0.498}{0.0926}$ | 5.378 |
5. An Example

Warner (1978) conducted a study involving four male and four female participants, focusing on the proportion of time spent speaking by each actor in the \(^{28}\) possible dyads. These data were analyzed by Warner et al. (1979) and Wong (1982) using methods for round robin designs. Each of the 28 dyads conversed on three separate occasions for about 12 to 15 min each time. The raw data, \(x_{ijk}^*\) = proportion of time spent speaking by actor \(i\) to actor \(j\) on occasion \(k\), are given in Table 7 of Warner et al. (1979) and Table 1 of Wong (1982). We massaged this \(x^*\) array to yield

\[
x_{ij} = 1 \quad \text{if} \quad \frac{\sum_k x_{ijk}^*}{k} > 50% \quad\text{otherwise,}
\]

using a constant threshold of \(t_{ij} = 50\%\). Thus, a 1 in our \(x\) array indicates that actor \(i\) dominated the conversation with actor \(j\). This dichotomized array is given in Table 7. In the table, actors 1, 2, 5, and 6 are male, and actors 3, 4, 7, and 8 are female. We note that of the 28 dyads, 4 are mutuals, indicating that both actors in the pair were vociferous, and 6 are nulls, indicating a large amount of “dead air” in their conversations.

An initial analysis of these data revealed that the four females behaved in much the same way, while the males were quite heterogeneous. This can be seen in Fig. 1 where we plot \(\hat{\beta}_i\) versus \(\hat{\alpha}_i\), for each of the eight actors, from a \(\rho_1\) analysis. Clearly, the females cluster nicely, unlike the males. We note that \(\rho\) is quite small and that both the \(\{\hat{\alpha}_i\}\) and \(\{\hat{\beta}_i\}\) are statistically large. Note that these findings differ from those of Warner et al. (1979), who do not find large “partner” or \(\beta\) effects, and Wong (1982), who finds large differences among both the males and females.

We next grouped together the four females to generate subgroup analyses with

<table>
<thead>
<tr>
<th>Actor</th>
<th>1*</th>
<th>2*</th>
<th>3</th>
<th>4</th>
<th>5*</th>
<th>6*</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1*</td>
<td></td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2*</td>
<td>0</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>1</td>
<td></td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5*</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>6*</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Note. * = male actor.
five subgroups (1 for the females and 1 for each male). The goodness-of-fit statistics for the models we considered are given in Table 8. Included in the table are statistics for a second subgrouping in which all females and all males are aggregated into two sex-based subgroups. A glance at the $G^2$'s for all actor models and for the five-subgroup models indicates that the models with the $\theta$, $\{x_i\}$, and $\{\beta_j\}$ parameters are most appropriate.

With this in mind, we can determine which of the aggregations is correct. We test

$H_0$: five subgroups—females aggregated, males unaggregated

versus

$H_A$: eight subgroups—females and males unaggregated.

The test statistic is $G^2(H_0 \mid H_A) = G^2(\theta, \{x_i\}, \{\beta_j\} \mid H_0) - G^2(\theta, \{x_i\}, \{\beta_j\} \mid H_A) = 37.96 - 26.04 = 11.92$ with $47 - 41 = 6$ degrees of freedom, a moderately large test statistic. We cannot reject the null hypothesis, implying that the females do act as a cohesive group. We next ask whether or not the males communicate with each other and the females identically, by testing

$H_0'$: two subgroups—females aggregated, males aggregated

versus

$H_A'$: five subgroups—females aggregated, males unaggregated.
TABLE 8
Statistics for Subgroup Analyses of Warner (1978) Communication Data

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameters</th>
<th>( G^2 ) All actors</th>
<th>( G^2 ) Five subgroups</th>
<th>( G^2 ) Two subgroups</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p_1 )</td>
<td>all</td>
<td>25.86 (40)</td>
<td>37.64 (46)</td>
<td>69.61 (52)</td>
</tr>
<tr>
<td>( p_1 - \rho )</td>
<td>{( \alpha_i ), ( \beta_j ), ( \theta )}</td>
<td>26.04 (41)</td>
<td>37.96 (47)</td>
<td>70.81 (53)</td>
</tr>
<tr>
<td>( p_1 - {\alpha_i } - \rho )</td>
<td>{( \beta_j ), ( \theta )}</td>
<td>61.74 (48)</td>
<td>65.76 (51)</td>
<td>72.69 (54)</td>
</tr>
<tr>
<td>( p_1 - {\beta_j } - \rho )</td>
<td>{( \alpha_i ), ( \theta )}</td>
<td>53.10 (48)</td>
<td>56.36 (51)</td>
<td>74.74 (54)</td>
</tr>
</tbody>
</table>

Note. All actors = eight subgroups; five subgroups = females aggregated; two subgroups = females and males aggregated; degrees of freedom are given following the \( G^2 \)'s in parentheses.

The test statistic is \( G^2(H_0 | H_A) = G^2(\theta, \{\alpha_i\}, \{\beta_j\} | H_0) - G^2(\theta, \{\alpha_i\}, \{\beta_j\} | H_A) = 70.81 - 37.96 = 32.85 \) with \( 53 - 47 = 6 \) degrees of freedom, a very large statistic. Clearly, our initial analysis was correct—while the females are quite similar, undifferentiating among the males would be a mistake.

The sending and receiving parameter estimates for this best fitting model are

\[
\hat{\alpha}_F = -7.75 \quad \hat{\beta}_F = 7.17 \\
\hat{\alpha}_{M1} = 22.08 \quad \hat{\beta}_{M1} = -5.70 \\
\hat{\alpha}_{M2} = -9.14 \quad \hat{\beta}_{M2} = -13.55 \\
\hat{\alpha}_{M3} = 3.29 \quad \hat{\beta}_{M3} = 6.81 \\
\hat{\alpha}_{M4} = -8.52 \quad \hat{\beta}_{M4} = 5.26
\]

We see that the females are unlikely to talk but are likely to be spoken to. Male 1 appears quite talkative, and males 2 and 4 less so, while male 2 also is unlikely to be spoken to. Examination of the standard errors indicates that the following groupings are justified:

speaking

\[
\begin{array}{cccc}
\text{Male} & \text{Male} & \text{Females} & \text{Male} \\
2 & 4 & 3 & 1 \\
\end{array}
\]

and spoken to

\[
\begin{array}{cccc}
\text{Male} & \text{Male} & \text{Male} & \text{Females} \\
2 & 1 & 4 & 3 \\
\end{array}
\]
Assuming that $p_1$ is the appropriate model, then Eq. (15) defines a system of $3^2$ linear equations, one triple for each dyad $i, j (i > j)$. We can represent the system in matrix algebra as

$$z = X\gamma.$$  \hspace{1cm} (A1)

The $z$ vector's $3^2$ entries are the left side of Eq. (15) for each dyad. The parameter vector $\gamma$ consists of the $2g$ distinct parameters to be estimated. Since we let

$$\alpha_g = -\sum_{i=1}^{g-1} \alpha_i \quad \text{and} \quad \beta_g = -\sum_{j=1}^{g-1} \beta_j,$$

we use unique parameters $\theta, \alpha_1, \alpha_2, \ldots, \alpha_{g-1}, \beta_1, \beta_2, \ldots, \beta_{g-1}, \rho$. The design matrix $X$ contains the coefficients of these parameters in the linear system.

To be exact, the design matrix consists of $2g$ columns, each corresponding to a parameter, $\theta, \alpha_1, \alpha_2, \ldots, \alpha_{g-1}, \beta_1, \beta_2, \ldots, \beta_{g-1}, \rho$, respectively. Each row of $X$ represents one of the $3^2$ equations of the system. For example, the first row of $X$ represents the equation

$$\log(\hat{p}_{1210}/\hat{p}_{1200}) = \hat{\theta} + \hat{\alpha}_1 + \hat{\beta}_2 \quad (i = 1, j = 2)$$

so there are 1's in the columns corresponding to $\theta, \alpha_1, \text{and} \beta_2$, and 0's elsewhere. Remaining rows are defined in a similar manner, unless $\alpha_g$ or $\beta_g$ are present in the equation. In this case, we replace $\alpha_g$ by $-\alpha_1 - \alpha_2 - \cdots - \alpha_{g-1}$ and $\beta_g$ by $-\beta_1 - \beta_2 - \cdots - \beta_{g-1}$, and express the row of $X$ in terms of the new coefficients. For example, the equation

$$\log(\hat{p}_{g310}/\hat{p}_{g300}) = \hat{\theta} + \hat{\alpha}_g + \hat{\beta}_3 \quad (i = g, j = 3)$$

becomes

$$\log(\hat{p}_{g310}/\hat{p}_{g300}) = \hat{\theta} - \hat{\alpha}_1 - \hat{\alpha}_2 - \cdots - \hat{\alpha}_{g-1} + \hat{\beta}_3$$

and so the corresponding row has 1's in the $\theta$ and $\beta_3$ columns; -1's in the $\alpha_1, \alpha_2, \ldots, \alpha_{g-1}$ columns; and 0's elsewhere.

The least squares solution of the system $\hat{z} = X\hat{\gamma}$ yields maximum likelihood parameter estimates of the $p_1$ parameters. We used readily available software subroutines to produce parameter estimates from fitted values presented by various log linear routines. Specifically, we used an IMSL (IMSL, Inc., 1982) subroutine that finds high accuracy least squares solutions of (A1) via a sweep operation algorithm. Computing time was moderate, and the parameter estimates agreed with direct estimates (those found as functions of $\nu$-terms). Note, that the above approach is easily modified for different $\gamma$ vectors from different $p_1$-type models.
REFERENCES


WASSERMAN, S. (1982). CTAB user’s guide for the University of Illinois at Urbana-Champaign. Unpublished manuscript.


RECEIVED: July 21, 1984