PSYCHOMETRIKA—VOL. 61, NO. 3, 401–425
SEPTEMBER 1996

LOGIT MODELS AND LOGISTIC REGRESSIONS FOR SOCIAL NETWORKS: I. AN INTRODUCTION TO MARKOV GRAPHS AND p^*

STANLEY WASSERMAN

UNIVERSITY OF ILLINOIS

PHILIPPA PATTISON

UNIVERSITY OF MELBOURNE

Spanning nearly sixty years of research, statistical network analysis has passed through (at least) two generations of researchers and models. Beginning in the late 1930's, the first generation of research dealt with the distribution of various network statistics, under a variety of null models. The second generation, beginning in the 1970's and continuing into the 1980's, concerned models, usually for probabilities of relational ties among very small subsets of actors, in which various simple substantive tendencies were parameterized. Much of this research, most of which utilized log linear models, first appeared in applied statistics publications.

But recent developments in social network analysis promise to bring us into a third generation. The Markov random graphs of Frank and Strauss (1986) and especially the estimation strategy for these models developed by Strauss and Ikeda (1990; described in brief in Strauss, 1992), are very recent and promising contributions to this field. Here we describe a large class of models that can be used to investigate structure in social networks. These models include several generalizations of stochastic blockmodels, as well as models parameterizing global tendencies towards clustering and centralization, and individual differences in such tendencies. Approximate model fits are obtained using Strauss and Ikeda's (1990) estimation strategy.

In this paper we describe and extend these models and demonstrate how they can be used to address a variety of substantive questions about structure in social networks.

Key words: categorical data analysis, social network analysis, random graphs.

1. Introduction—The Evolution of Statistical Models for Social Networks

Statistical models have been used by researchers to study social networks for almost 60 years, beginning with the work of the early sociometricians in the late 1930's. The goal of these models was (and remains) the quantitative examination of the stochastic properties of social relations and the actors of a particular network. Nonstochastic models have certainly been considered by many, but such models do not have the nice properties of their statistical counterparts, such as goodness-of-fit statistics and the possibility of parametric significance tests of various network structural properties.

Recent developments in social network analysis have now brought us to a new generation of models, namely the Markov random graphs of Frank and Strauss (1986) and especially the estimation strategy for these models developed by Strauss and Ikeda (1990), described in brief in Strauss (1992). These models expand considerably the class of struc-

This research was supported by grants from the Australian Research Council and the National Science Foundation (#SBR93-10184). This paper was presented at the 1994 Annual Meeting of the Psychometric Society, Champaign, Illinois, June 1994. Special thanks go to Sarah Ardu for programming assistance, Laura Koehly and Garry Robins for help with this research, and to Shizuhiko Nishisato and three reviewers for their comments.
INTERNET *email* addresses: *pattison@psych.unimelb.edu.au* (PP); *stanwass@uiuc.edu* (SW). Affiliations: Department of Psychology, University of Melbourne (PP); Department of Psychology, Department of Statistics, and The Beckman Institute for Advanced Science and Technology, University of Illinois (SW).

Requests for reprints should be sent to Stanley Wasserman, University of Illinois, 603 East Daniel Street, Champaign, IL 61820.

0033-3123/96/0900-95265$00.75/0

tural models that can be investigated within the exponential family first proposed by Holland and Leinhardt (1981). Markov models, and their more general forms, which we label p^*, can be approximated by logistic regressions, thus giving the researcher easy access to a very large wealth of modeling tools.

The review and extension of this research is the goal of this and our companion paper (Pattison & Wasserman 1995).

We first give some notation which we will use throughout, and then turn to a brief recent history of statistical models for social networks. More thorough historical accounts can be found in chapters 13–16 of Wasserman and Faust (1994).

2. Some Notation

A *social network* is defined as a set of g social *actors* and a collection of r social *relations* that specify how these actors are relationally tied together. Of interest to us here will be networks with either $r = 1$ or 2 relations. Examples of social relations are "chooses as a friend" and "is a neighbor of", recorded for each pair of individuals in some set of actors.

We let $\mathcal{N}$ denote a set of actors: $\mathcal{N} = \{1, 2, \ldots, g\}$. A *dichotomous social relation*, $\mathcal{X}$, is a set of ordered pairs recording relational ties between pairs of actors. If the ordered pair (i, j) is in this set, then the first actor (i) in the pair has a relational tie to the second actor (j) in the pair; we write iXj, or more succinctly, $i \rightarrow j$.

A social relation can be either directed (i's tie to j may differ from j's tie to i) or nondirected (there is, at most, one nondirected tie connecting i and j), and can also be valued (the tie from i to j has a nondichotomous strength or value). The main statistical focus in the literature has been on models for single (or univariate), dichotomous, directed relations (represented as *directed graphs*). Since many of the models described here take on a different character when the relation under study is nondirected, we will (at times) describe the models and parameters which arise when the relation is not directed.

Any social relation can be represented by a $g \times g$ *sociomatrix*, $\mathbf{X}$, where the (i, j) entry in the matrix (which we denote by $X(i, j)$ or, sometimes, by X_{ij}) is the value of the tie from actor i to actor j on that relation. For a dichotomous relation, $\mathbf{X}$ is also the adjacency matrix of the directed graph representing the relation. For a valued relation, $X(i, j) = c$, the value of the tie from i to j. We typically assume that $c \in \{0, 1, 2, \ldots, C - 1\}$. For a dichotomous relation, $C = 2$, so that

$$X(i, j) = \begin{cases} 1 & \text{if } i \rightarrow j \\ 0 & \text{otherwise.} \end{cases}$$

Frequently of interest in network analysis are subsets of actors and all the ties that might exist among them. For example, for a dichotomous relation, a *dyad* is a pair of actors and all the ties between them, and can be in one of four states: null (no ties), asymmetric (one tie; two possibilities), and mutual (two ties).

In the general, multirelational case, relational ties are recorded for r relations. We assume that these relations, $\mathcal{X}_1, \mathcal{X}_2, \ldots, \mathcal{X}_r$, have associated sociomatrices $\mathbf{X}_1, \mathbf{X}_2, \ldots, \mathbf{X}_r$. And since these sociomatrices will be assumed to be random quantities, we will use lower-case bold-face characters (such as $\mathbf{x}$) to denote realizations of these random quantities. Our focus here is on the univariate situation ($r = 1$), while extensions to the multivariate situation can be found in the companion paper (Pattison & Wasserman, 1995).

Important network statistics for dichotomous directed relations include the *outdegree* and *indegree* of an actor, that is, the number of ties sent and received ($\Sigma_j X(i, j)$ and $\Sigma_i X(i, j)$, respectively); the number of ties, $L = \Sigma_{i,j} X(i, j)$, and the number of reciprocated ties (that is, such that $X(i, j) = X(j, i) = 1$; sometimes referred to as *mutual* dyads); and numbers of higher-level subgraphs, such as 2-out-stars (triples of nodes such that $i \rightarrow j$ and $i \rightarrow k$), 2-in-stars (triples of nodes such that $j \rightarrow i$ and $k \rightarrow i$), 2-mixed-stars (triples of nodes such that $i \rightarrow j$ and $k \rightarrow i$), cyclic triads (triples such that $i \rightarrow j$, $j \rightarrow k$, $k \rightarrow i$), and so on (there are many possibilities). We define

$$M = \sum_{i<j} X(i, j)X(j, i)$$

as the number of mutual dyads,

$$S = \sum_{i \neq j \neq k} X(i, j)X(i, k)$$

as the number of 2-out-stars,

$$T = \sum_{i \neq j \neq k} X(i, j)X(j, k)X(k, i)$$

as the number of cyclic triads, all for a single relation.

A *(directed) path* of length d from i to j is a sequence of nodes $\{i = i_1, i_2, \ldots, i_{d+1} = j\}$, beginning with i and ending with j, such that

$$X(i_1, i_2)X(i_2, i_3) \ldots X(i_d, i_{d+1}) = 1.$$

The shortest path between any pair of nodes is a *geodesic*, and the length of the geodesic(s) is the *distance* from i to j (denoted by d_{ij}). This geodesic distance can be defined a bit more formally as

$$d_{ij} = \text{minimum value of } k \text{ for which } X(i_0, i_1)X(i_1, i_2) \ldots X(i_{k-1}, i_k) = 1,$$

where $i_0 = i$ and $i_k = j$. The distance d_{ij} is undefined if there is no path from i to j. A graph is "disconnected" if some distances are undefined; the graph is then said to have more than one *component*. In a digraph, if at least one node has no path to another node, the digraph is not *strongly connected*.

If actors have been partitioned into a set of S *blocks* (or *positions*), we can define $\delta_{ij;rs}$ as an indicator variable, equal to 1 if actor i is in the rth block and actor j is in the sth.

There are many other simple graph statistics; some of these will be presented in later sections (see Tables 1 and 2 for nondirected relations and Tables 3 and 4 for directed relations).

We will let ϑ and ϖ be symbols to represent *logits*—log odds ratios, comparing the probability of one outcome of a random variable to the probability of another outcome, in a logarithm scale. The ϑ's will be logits based on Holland and Leinhardt's (1981) models, while the ϖ's will be logits based on the p^* models that we describe here.

3. A Brief History

The models that are considered in this paper express each relational tie as a stochastic function of actor or network structural properties. An important example of such a model is Holland and Leinhardt's (1977, 1981) p_1 model. The p_1 model includes parameters for tie density, the propensity for reciprocity of ties, and individuals' tendencies to express and receive ties. This model and its many generalizations make the strong assumption of dyadic independence, an assumption that has been criticized in the literature (see chapter 15 of Wasserman & Faust, 1994). Consequently, developments which relax the assumption are of considerable importance. Frank and Strauss' (1986) research on Markov random graphs was the first such generalization, and has great potential. Markov graphs permit dependencies among any ties that share a node (for example, X_{ij} and X_{ik}, or X_{ij} and X_{jk}).

The important work of Strauss and Ikeda (1990) has made these models computationally available. Strauss and Ikeda investigated a *pseudolikelihood* estimation procedure, a generalization of maximum likelihood, that uses an approximate likelihood function which does not assume dyadic independence (see comments in Iacobucci & Wasserman 1990; and chapters 15 and 16 of Wasserman & Faust 1994). Strauss and Ikeda derived a pseudolikelihood as a function of each data point (x_{ij}), conditional on the rest of the data. Any interdependencies in the data can be directly modeled by this statistical conditioning, so no assumptions need to be made that the data points are all independent.

Strauss and Ikeda compared the performance of standard maximum likelihood (ML) estimates to their maximum pseudo-likelihood (MP) estimates both in a simulation study, and by analyzing the "like" relation measured on the monks in the monastery studied by Sampson (1968). In the simulations, they looked at the performance of the estimates in five replicated networks containing fifteen, twenty, or thirty actors. They found that MP and ML estimates performed equally well, as evaluated by a root mean squared error measure. The estimates had greater standard errors for the networks with fewer actors, but this would be true of any procedure—better precision usually occurs with larger data sets.

Under all the conditions for which both ML and MP estimates could be estimated, the two performed similarly. One of the main advantages in the use of MP estimation is that there are conditions under which MP estimates exist, but ML estimates do not. In addition, the MP approach further expands the applicability of p_1 because it can be used to fit models that do not assume dyadic independence, such as those described by Frank and Strauss (1986).

Strauss and Ikeda's comparisons also address the issue of how well the maximum likelihood estimation of p_1 parameters performs even under conditions where the assumption of dyadic independence is known to be violated. The fact that the ML estimates are as good as MP estimates is good news. One can proceed to use the relatively simple methods without much concern that violation of the assumption of dyadic independence will greatly affect the results. The biggest advantage (to us) of the MP estimation technique is its ease of use—as we note below, logistic regression computational procedures can be used to fit these models, and hence, to give very good approximations of fitted p_1-type models.

Details of such models, which we will refer to here by p^*, follow.

4. p_1, p^*, and Logit Models

Before describing the models, p_1 and p^*, we note that we will illustrate the methods described here on an example taken from Vickers and Chan (1981) and Vickers (1981). They obtained network data from 29 students in Grade 7 in a school in Melbourne, Victoria in Australia. They asked the students to nominate their classmates on a number of relations, including the following:

- Who do you get on with in the class?
- Who are your best friends in the class?
- Who would you rather not be friends with?
- Who would you prefer to work with? and
- Who would you rather not work with?

Each of these questions gives rise to a dichotomous relation, with measurements recorded in a 29 $\times$ 29 sociomatrix. We focus our attention here on the first relation ("Getting on with"), and the fourth ("Work with"). Both of these relations are directed. We label the two, $\mathcal{X}_g$, and $\mathcal{X}_w$ (for Get on With, and Work with). The two matrices for these relations are shown in Tables 11 and 12 found in the Appendix.

We note that Actors 1 through 12 are boys, while Actors 13 through 29 are girls. We will use these data to illustrate the models described below.

4.1. p_1

We quickly review p_1, referring the reader to more complete treatments (Fienberg & Wasserman, 1981; Holland & Leinhardt, 1981; Reitz, 1982; Wasserman, 1987; Wasserman & Faust, 1994; Wasserman & Galaskiewicz, 1984) if needed.

Consider the dyad $D_{ij} = (X_{ij}, X_{ji})$, and the four possible states that this random quantity can be in:

- $(X_{ij}, X_{ji}) = (0, 0)$—Null dyad,
- $(X_{ij}, X_{ji}) = (1, 0)$—Asymmetric dyad $(i \to j)$,
- $(X_{ij}, X_{ji}) = (0, 1)$—Asymmetric dyad $(j \to i)$,
- $(X_{ij}, X_{ji}) = (1, 1)$—Mutual dyad.

We consider the substantive tendencies which might cause a particular dyad to be in one of these states (for example, mutual dyads might arise if both actors were quite popular). Such tendencies are then incorporated as additive effects into a model for the logarithms of the probabilities of the four dyadic states.

We first create a new matrix, **Y**, from **X**, which aids the computations for the model. We define the entries of this $g \times g \times 2 \times 2$ matrix as follows:

$$Y_{ijkl} = \begin{cases} 1 & \text{if the dyad } D_{ij} \text{ takes on the values } (X_{ij} = k, X_{ji} = l) \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

Thus, the four entries in **Y** which are associated with the pair of actors i and j (the 2×2 submatrix associated with the ith level of variable 1 and the jth level of variable 2) code the state of the dyad D_{ij}. To specify p_1, we postulate the following log linear equations:

$$\log P(Y_{ij00} = 1) = \lambda_{ij}$$

$$\log P(Y_{ij10} = 1) = \lambda_{ij} + \theta + \alpha_i + \beta_j$$

$$\log P(Y_{ij01} = 1) = \lambda_{ij} + \theta + \alpha_j + \beta_i$$

$$\log P(Y_{ij11} = 1) = \lambda_{ij} + 2\theta + \alpha_i + \alpha_j + \beta_j + \beta_i + \rho. \tag{2}$$

The parameters of this model are explained at length in Fienberg and Wasserman (1981), Wasserman (1987), or chapter 15 of Wasserman and Faust (1994). In brief, the α's reflect substantive tendencies toward expansiveness (they are "subject" effects), the β's, attractiveness or popularity ("partner" effects), θ is an overall choice parameter, while ρ models tendencies toward reciprocation or mutuality.

If the relation in question is not directed, then model (2) has a simpler form. For nondirected relations, there are only two dyad states, and only two types of parameters (see Wasserman & Faust, 1994).

Perhaps a better way to present this model is to relax the restriction to dichotomous relations. We can assume that the relation is *valued*, and that every tie has a strength, with a score measured on an integer scale from 0 to $C - 1$. If we define k as the value of the tie from i to j and l as the value of the tie from j to i, then we can generalize **Y** so that its elements $Y_{ijkl} = 1$, when $X_{ij} = k$ and $X_{ji} = l$. The following model statement (from

Wasserman & Iacobucci, 1986; see also Anderson & Wasserman, 1995) generalizes the four statements of p_1:

$$\log P(Y_{ijkl} = 1) = \lambda_{ij} + \theta_k + \theta_l + \alpha_{i(k)} + \alpha_{j(l)} + \beta_{j(k)} + \beta_{i(l)} + \rho_{kl}. \tag{3}$$

Regardless of the nature of the relation, dyads are assumed to be statistically independent, so that the joint probability distribution, and hence the likelihood function, is simply a product of dyadic probabilities. Models such as p_1 are referred to as *dyadic independence* models (for obvious reasons).

Estimation (after placing necessary and logical constraints on model parameters), fitting, and testing are described in any of the above references. Simplifications arise if relations are ordinal, or if actors can be assumed to fall naturally into subgroups (so that actor-level parameters can be equated for all actors within a particular subgroup). Fitting is not particularly easy, and the asymptotic theory for testing is questionable (at best). These two problems (along with the desire for models that do allow dyads to be dependent) motivate the research into more complicated models that we now present. These models have the useful feature that they include dyadic independence models such as p_1 as a special case.

4.2. p^*—*Theory, Special Cases, Estimation*

In order to specify the exponential family of models p^*, which contains Markov random graphs, as well as p_1 as a special case (in a sense that we mention below), we need a bit more notation. From $\mathbf{X}$, the sociomatrix for a single, dichotomous, directed relation, we define three new relations, with sociomatrices easily constructed from $\mathbf{X}$. First, we define $\mathbf{X}_{ij}^+$ as the sociomatrix for the relation formed from $\mathcal{X}$ where the tie from i to j is forced to be present: $\mathbf{X}_{ij}^+ = \{X_{kl}, \text{with } X_{ij} = 1\}$. Next, we define $\mathbf{X}_{ij}^-$ as the sociomatrix for the relation formed from $\mathcal{X}$ where the tie from i to j is forced to be absent (or to be at level 0): $\mathbf{X}_{ij}^- = \{X_{kl}, \text{with } X_{ij} = 0\}$. Lastly, we define $\mathbf{X}_{ij}^c$ as the *complement* relation for the tie from i to j: $\mathbf{X}_{ij}^c = \{X_{kl}, \text{with } (k, l) \neq (i, j)\}$. The complement relation has no relational tie coded from i to j—one can view this single variable as missing. All told, these three new relations are needed to define p^*.

The original specification of these models was just for a single, dichotomous relation. Most of the early work focused on nondirected relations. Generalizations to valued relations, and to more than one relation, were mentioned in passing (in concluding remarks) by Frank and Strauss (1986, sec. 6) and by Strauss and Ikeda (1990, sec. 5). We address all of these generalizations in this paper and in our companion paper, Pattison and Wasserman (1995).

4.2.1. *Logit Models*

We now describe the general, log linear form of p^*, which we will then present in its logit formulation. We begin with log linear models of the form

$$\Pr(\mathbf{X} = \mathbf{x}) = \frac{\exp\{\boldsymbol{\theta}'\mathbf{z}(\mathbf{x})\}}{\kappa(\boldsymbol{\theta})}, \tag{4}$$

where $\boldsymbol{\theta}$ is a vector of model parameters and $\mathbf{z}(\mathbf{x})$ is a vector of network statistics. These models are of exponential family form, in which the probability function depends on an exponential function of a linear combination of network statistics. Such models arise frequently, not only for studies of social networks, but also in spatial modeling, statistical mechanics, and even in test theory (Strauss, 1992). We refer to models of the form (4) with the label p^*. One might need constraints on the elements of $\boldsymbol{\theta}$ to insure a set of uniquely-determined parameters (as we illustrate later with our examples).

In model (4) for social networks, the $\boldsymbol{\theta}$ parameters are the weights of the linear combination, and are usually unknown, and hence, must be estimated. The function κ, in the denominator of model (4), is a constant that insures that the probability distribution is indeed proper, summing to one over the sample space of the random variable $\mathbf{X}$—all possible directed graphs. Examples of various statistics $\mathbf{z}$ are numerous, and (for a directed relation) include the number of mutual dyads, M, the number of ties, L, and the outdegree of the ith actor, x_{i+}. A very wide range of such statistics can be found in Tables 1, 2, 3, and 4.

A very simple example of model (4), which is quite similar to the simplest Bernoulli graph distribution as well as to a special case of p_1, is

$$\Pr(\mathbf{X} = \mathbf{x}) = \frac{\exp\{\theta L\}}{\kappa(\theta)}, \tag{5}$$

with a single parameter θ and which depends on only the number of relational ties.

The problem with distributions of this form is the normalizing constant. In order for probabilities to be computed, one must be able to calculate κ, which is just too difficult for most networks. This prevents easy maximum likelihood estimation of the model parameters (see Frank & Strauss, 1986) except in special circumstances. But, there are tricks, as Fienberg and Wasserman (1981) discovered for p_1, and as we describe below, for p^*.

As described by Strauss and Ikeda (1990), one can turn this loglinear model into a logit model, using the dichotomous nature of the random variable X_{ij}. We first condition on the complement of X_{ij}, and consider just the probability that the tie from i to j is present:

$$\begin{aligned}\Pr(X_{ij} = 1 | \mathbf{X}_{ij}^c) &= \frac{\Pr(\mathbf{X} = \mathbf{x}_{ij}^+)}{\Pr(\mathbf{X} = \mathbf{x}_{ij}^+) + \Pr(\mathbf{X} = \mathbf{x}_{ij}^-)} \\ &= \frac{\exp\{\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}_{ij}^+)\}}{\exp\{\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}_{ij}^+)\} + \exp\{\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}_{ij}^-)\}}\end{aligned} \tag{6}$$

which has the advantage of *not* depending on the normalizing constant. We next consider the odds ratio of the presence of a tie from i to j to its absence, which simplifies model (6):

$$\begin{aligned}\frac{\Pr(X_{ij} = 1 | \mathbf{X}_{ij}^c)}{\Pr(X_{ij} = 0 | \mathbf{X}_{ij}^c)} &= \frac{\exp\{\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}_{ij}^+)\}}{\exp\{\boldsymbol{\theta}'\mathbf{z}(\mathbf{x}_{ij}^-)\}} \\ &= \exp\{\boldsymbol{\theta}'[\mathbf{z}(\mathbf{x}_{ij}^+) - \mathbf{z}(\mathbf{x}_{ij}^-)]\}.\end{aligned} \tag{7}$$

From this, the log odds ratio, or *logit*, model has the rather simple expression:

$$\varpi_{ij} = \log\left\{\frac{\Pr(X_{ij} = 1 | \mathbf{X}_{ij}^c)}{\Pr(X_{ij} = 0 | \mathbf{X}_{ij}^c)}\right\} = \boldsymbol{\theta}'[\mathbf{z}(\mathbf{x}_{ij}^+) - \mathbf{z}(\mathbf{x}_{ij}^-)]. \tag{8}$$

If we define $\boldsymbol{\delta}(x_{ij}) = [\mathbf{z}(\mathbf{x}_{ij}^+) - \mathbf{z}(\mathbf{x}_{ij}^-)]$, then the logit model (8) simplifies succinctly to $\varpi_{ij} = \boldsymbol{\theta}'\boldsymbol{\delta}(x_{ij})$. The expression $\boldsymbol{\delta}(x_{ij})$ is the vector of network statistics that arises when the variable X_{ij} changes from 1 to 0. This version of the model, in which a log odds ratio is equated to a linear function of the components of $\boldsymbol{\delta}(x_{ij})$, will be referred to as the *logit* p^* model for a single, dichotomous relation.

As one can see, to specify a logit p^* model, one chooses a priori a collection of network statistics that is supposed to affect the log odds of a tie being present to absent.

The model itself depends on the hypothesized structural features of the network. For each hypothesized structural feature (such as transitivity), there is a corresponding network statistic (such as, T, the total number of transitive triads) and a corresponding "explanatory variable" in the logit model; the explanatory variable is the change in the network statistic when the tie from node i to node j changes from being present to absent.

The model is easy to construct when the relation is dichotomous, so that logits are simple and well-defined. When the relation is valued, one must be careful about which logits to model—there will be $C - 1$ logits for a dichotomous relation that takes on integer values from 0 to $C - 1$. We discuss this at length in Pattison and Wasserman (1995).

4.2.2. *Dependence Graphs*

In the original specification of model p^*, Frank and Strauss (1986) viewed it as a generalization of model p_1 and all its relatives, designed to relax the restriction of the earlier models to independent dyads. Frank and Strauss first presented p^* as a model incorporating a complicated dependence structure for nondirected relations, and then generalized this to directed relations. Even though Frank and Strauss (1986) presented the general specification of p^*, they concentrated (as can be seen by the title of that paper) on the special case of p^* incorporating a Markov assumption. This special case is based on the theory for Markov random fields, applied to graphs (and hence, labeled *Markov random graphs*). As Frank and Strauss desired, one no longer had to assume that dyads were statistically independent.

Markov graphs are generalizations of Markov random fields designed for spatial interaction models (Kindermann & Snell, 1980; Ripley, 1981; Speed, 1978; Strauss, 1977), and are based on the work of Ising (1925) for models of rectangular arrays of binary variables, or *lattices* (Wasserman, 1978). Any two sites on a lattice are *neighbors* if they are sufficiently close. One then postulates an Ising model for the entire set of lattice variables which has the same exponential family form as model (4) with parameters $\boldsymbol{\theta}$.

There are a variety of ways of viewing p^*. One important aspect or view of these models is the dependence structure they assume for the lines of the graph or arcs of the digraph. A *dependence graph*, in the context of a social network, indicates which relational ties (or subsets of relational ties) are conditionally independent. The dependence structure of a random directed graph is simply a graph whose nodes are all possible relational ties in the original relation and whose ties specify which ties in the relation are *conditionally dependent*, given the remaining relational ties. Two ties are conditionally dependent if the conditional probability that the ties both are present, given the other ties in the network, is not equal to the product of their marginal conditional probabilities. The dependence graph has lines connecting all pairs of conditionally dependent ties. There are many ways of specifying conditional dependence between a pair of ties, which lead to a variety of distinct dependence graphs.

For example, the dependence digraph arising from p_1 has lines connecting each relational tie to its dyadic pair (and no others). Other examples are described below.

To fully appreciate this view of p^*, one needs to consider the structure of the dependence graph and how a particular structure is reflected in p^* model parameters. We do this formally, using the Hammersley-Clifford theorem (discussed by Besag, 1974) which establishes that a random directed graph has a probability that depends only on the complete subgraphs of the dependence graph. (A *complete subgraph* of a graph is a subset of nodes in which every pair of nodes in the subset is linked by a line. Thus, a complete subgraph in the dependence graph corresponds to a set of ties in the original random directed graph, every pair of which is conditionally dependent, given the rest of the graph.) The theorem also establishes that sufficient statistics for a loglinear model for the random directed graph are of the form $\Pi_{(i,j)\in A} X_{ij}$, where A is a complete subgraph of the dependence graph. The product here is computed across a set of edges in the original directed graph that are pairwise mutually conditionally dependent. The induced subgraph in the original directed graph corresponding to this set of edges is termed a *sufficient subgraph* for the loglinear model. (The subgraph induced by A is simply the collection of all edges in A and the nodes in the original graph to which they are incident.) This important result implies that one need only worry about complete subgraphs of the dependence graph.

Frank and Strauss described several important examples of specific dependence structures.

1. The first is the case of a Bernoulli directed graph, in which all edges are conditionally independent. The dependence graph of a Bernoulli graph comprises a collection of isolated nodes, each corresponding to an edge in the Bernoulli graph.
2. A second is the case of dyad independence in which X_{ij} is conditionally dependent only on X_{ji}, given the rest of the graph (Strauss & Ikeda, 1990). In this case, the complete subgraphs of the dependence graph are of the form $\{X_{ij}\}$ or $\{X_{ij}, X_{ji}\}$.
3. A third type of dependence structure is one in which edges that have an actor in common are conditionally dependent (Frank & Strauss, 1986). From this, Frank and Strauss define a *Markov directed graph* as a random directed graph for which arcs in the dependence graph connect pairs of possible ties if and only if they have an actor in common, such as X_{ij} and X_{ik}, X_{ij} and X_{ki}, X_{ij} and X_{jk}, or X_{ij} and X_{kj}. It immediately follows that in a Markov graph, the complete subgraphs of the dependence graph are simply *triads* and *k-stars*. In a Markov digraph, the complete subgraphs of the dependence graph are mutual dyads, triads, and stars of order 3 or more. Frank and Strauss give the complete subgraphs (with 4 or fewer nodes) of the dependence graph of a Markov digraph.

Frank and Strauss' use of Markov (di)graphs greatly simplifies the possible dependence structures that can arise, and leads naturally to a class of parametric models. Specifically, loglinear models for graphs with Markov dependence can be seen to depend only on the complete set of triads and k-stars, and not on tetrads or other more complicated subgraphs (we note that such subgraph structures are defined in chapter 4 of Wasserman & Faust, 1994). Further simplifications of such models arise with assumptions of homogeneity of model parameters, and assumptions that some parameters are zero.

Occasionally, it may make sense to propose more complex dependence structures. Suppose, for instance, that there is a *coloring* on the nodes of a graph or directed graph (that is, each node in the graph is assigned exactly one color from a specified set of colors). The coloring may specify the status of network actors on some attribute such as family membership; in this case, the coloring would require as many colors as there are families (and we assume, in this simple example, that each actor is a member of exactly one family). We define a *block random graph* to be a p^* random graph in which two possible ties (i, j) and (k, l) are conditionally dependent if and only if nodes i, j, k and l all have the same color; in other words, ties are conditionally dependent if they are incident to nodes of the same color. The corresponding dependence graph consists of a set of disconnected components, one for each color in the color set and each component is complete. The complete subgraphs of the dependence graph are thus the collection of these components and all of their subgraphs. Sufficient statistics for the corresponding log linear model parameters are products of observed ties, where the products are taken over subsets of ties linking individuals of the same color. A block random graph would allow one to model interdependencies among relational ties among family members.

4.2.3. Some Models

To generate a specific member of the logit p^* family, one needs to specify the vector of network statistics $\mathbf{z}(\mathbf{x})$ and to determine the need for any constraints on the parameters $\boldsymbol{\theta}$.

As shown by Frank and Strauss (1986), one Markov graph model for a nondirected relation depends on the number of cyclic triads and the numbers of k-stars in the graph representation of the relation. Such a model needs k-stars as high as $k = (g - 1)$, and as low as $k = 1$ (the number of ties), so there are exactly $(g - 1) + 1 = g$ components in $\mathbf{z}(\mathbf{x})$. One need not work with a more complicated model, assuming of course, that one wants homogeneity (parameters not depending on the individual actors).

A simpler Markov graph model used by Frank and Strauss (1986) is referred to as the *triad model* and depends only on the number of relational ties, L, the number of 2-stars, S, and the number of cyclic triads, T. A special case of this model is one-dimensional, and depends only on S, the number of 2-stars, and is referred to as the *clustering model*. This latter model is the only model that Frank and Strauss estimate with maximum likelihood. For a directed relation, the model becomes more complicated, since there are different types of stars (for example, there are "in-stars" of order 2, for which both ties end at actor i, as well as "out-stars" of order 2, for which both ties originate at actor i; and of course, there are many different types of triad counts).

Another special Markov directed graph model mentioned by Strauss and Ikeda (1990), is the model:

$$\varpi_{ij} = \log\left\{\frac{\Pr\,(X_{ij} = 1|\mathbf{X}_{ij}^c)}{\Pr\,(X_{ij} = 0|\mathbf{X}_{ij}^c)}\right\} = \theta + \rho(x_{ji}) + \alpha_i + \beta_j. \tag{9}$$

This special case of p^* is similar to p_1, and so we label it p_1^*. There is also a version of this model for nondirected relations, similar to nondirected p_1 (see chapter 15 of Wasserman & Faust, 1994), and one can easily add blockmodel-type parameters, as do Wasserman and Galaskiewicz (1984) and Wang and Wong (1987) to p_1. Nondirected p_1^* is equivalent to nondirected p_1, since the relational ties are actually completely independent (in this special instance).

We can put model (9) into the standard p^* model form. First define the vector $\boldsymbol{\theta}$ as the $(2g + 2)$-dimensional parameter vector with components

$$\boldsymbol{\theta} = (\theta, \alpha_1, \ldots, \alpha_g, \beta_1, \ldots, \beta_g, \rho)',$$

where one would constrain the α's and the β's in some way (perhaps both sets of g parameters would sum to zero). With this choice of $\boldsymbol{\theta}$, the vector of graph statistics is

$$\mathbf{z}(\mathbf{x}) = (L, x_{1+}, \ldots, x_{g+}, x_{+1}, \ldots, x_{+g}, M)',$$

so that the vector of changes in $\mathbf{z}(\mathbf{x})$ that arises when x_{ij} changes from a 1 to a 0 is

$$\boldsymbol{\delta}(x_{ij}) = (1, 0, \ldots, 0, 1, 0, \ldots, 0, 0, \ldots, 0, 1, 0, \ldots, 0, x_{ji})',$$

where the 1's fall in the first, $(i + 1)$st, and $(g + 1 + j)$th entries (corresponding to the θ, α_i, and β_j components of $\boldsymbol{\theta}$).

Thus, it appears that p_1 arises as a logit p^* model. The components of $\mathbf{z}(\mathbf{x})$ for p_1^* are the set of indegrees, the set of outdegrees, M, and L, which are also the sufficient statistics for the parameters of p_1; however, and most importantly, this special case p_1^* does not assume dyadic independence (as we discuss at bit more at length, later in this section). One way to view p_1^* is as an alternative method for estimation of the parameters of p_1.

One can take p_1^* and generalize it in ways not possible with the dyadic independence model p_1. Such models incorporate graph statistics into $\mathbf{z}(\mathbf{x})$ that depend on higher-order graph properties; specifically, counts of the frequencies of the sufficient subgraphs. Adding counts of k-stars is one possibility—Frank and Strauss' triad model is an example. There are many other possibilities, which we discuss in brief below. One can postulate rather complicated "explanatory graph statistics" as components of $\mathbf{z}(\mathbf{x})$.

4.2.4. *Structural Parameters for Networks*

The class of logit models for directed graphs just described allows very considerable breadth in the formulation of candidate models for a particular relation. As a result, it is important to begin to describe what considerations might influence the choice of a model and to explicate some possible relationships among models. In this section, we review briefly some major themes of structural analysis for networks, and show how such themes might be reflected in certain types of loglinear models. Our account is not intended to be comprehensive but rather to illustrate how a wide range of questions about network structure can be formulated and explored within this framework.

Structural themes. Wasserman and Faust (1994) review descriptive structural analyses of single network relations that have been prominent in the network literature. These analyses have dealt with at least three broad themes:

1. The degree of clustering of nodes in a network as well as the presence of cohesive subsets;
2. Patterns of connectivity and reachability in the network, and especially the distribution of centrality and prestige across the actors; and
3. The similarity among network positions and the degree to which network structure can be summarized in a blockmodel.

In several of these themes, it has been argued that questions about network structure should be assessed only after allowing for a variety of constraints that may be regarded as lower-level features of the data.

In an early paper, for instance, Holland and Leinhardt (1973, 1979) observed that restrictions on the outdegree of each node imposed constraints on possible structural patterns that could be observed at a more global level. Similarly, Holland and Leinhardt (1975), Wasserman (1987), and Snijders (1991) have argued that multirelational questions and questions dealing with triads and other higher-order structures are best assessed only after taking account of the distribution of various nodal or dyadic network properties. Thus, we will consider models that assess the degree to which a network is clustered, or centralized, in the presence of certain lower-level properties of the network.

A basic distinction applying to the parameters of logit models for structural features of a network is that between homogeneous and nonhomogeneous effects. An effect is defined to be *homogeneous* if it is assumed to be equal for all pairs of nodes i and j; for instance, the reciprocity effect in the p_1^* model is assumed to be homogeneous, whereas the expansiveness and attractiveness effects are not. (Note that Frank and Strauss define a model to be homogeneous if all isomorphic graphs have the same probability. Thus, a model is homogeneous in the sense of Frank and Strauss if all of its parameters are homogeneous in the sense defined here.)

Nonhomogeneous effects can be further distinguished according to whether they are assumed to vary freely for all pairs of nodes (i, j) or to be subject to equality constraints within classes of the set of possible edges of the network. For instance, a stochastic blockmodel (first defined by Fienberg & Wasserman 1981; and elaborated upon by Wasserman & Galaskiewicz 1984; and Wang & Wong 1987) specifies that density effects for node pairs (i, j) and (k, l) are equal whenever nodes i and k are in the same class, and j

Table 1: Homogeneous parameters and graph statistics for logit models for nondirected relations

Label	Parameter	Graph statistic z(x)
Homogeneous effects		
Choice	θ	$L = X_{++}$
Clusterability	τ	$T = \sum_{i,j,k} X_{ij}X_{jk}X_{ki}$
2-stars	σ	$S = \sum_{i,j,k} X_{ij}X_{ik}$
$\vdots$	$\vdots$	$\vdots$
k-stars	υ_k	$U_k = \sum_{i,j_1,j_2,\ldots,j_k} X_{ij_1}X_{ij_2}\cdots X_{ij_k}$
$\vdots$	$\vdots$	$\vdots$
k-paths	π_k	$P_k = \sum_{i,j_1,j_2,\ldots,j_k} X_{ij_1}X_{j_1j_2}\cdots X_{j_{k-1}j_k}$
$\vdots$	$\vdots$	$\vdots$
Connectivity	ν	N = minimum number of edges whose removal disconnects the graph (edge connectivity)
Indices of prominence homogeneity (centralization)		
	ϕ_1	C_{G_1} = sum of lengths of geodesics (d_{ij})
	ϕ_2	C_{G_2} = variance of geodesic lengths (d_{ij})
	ϕ_3	C_E = maximum geodesic length (eccentricity)
	ϕ_4	C_C = variance of geodesic length sums (d_{i+}) (closeness centralization)
	ϕ_5	C_D = variance of X_{i+}'s (degree centralization)
	ϕ_6	C_B = variance of numbers of geodesics containing i (betweeness centralization)
Association with fixed **Y**	Γ	$G = \sum_{i,j} X_{ij}Y_{ij}$

Table 2: Actor-level parameters and graph statistics for logit models for nondirected relations

Label	Parameter	Graph statistic z(x)
Subgroup–specific/Block–level effects		
Block effects	ζ_{rs}	$B_{rs} = \sum_{i,j} X_{ij}\delta_{ij;rs}$
Node–specific/actor–level effects		
Differential choice	γ_i	X_{i+} = degree (degree centrality)
Differential closeness	$\phi_4(i)$	$d_{i.}$ = average geodesic distance (closeness centrality)
Differential connectedness	ν_i	N = minimum number of edges incident to i whose removal disconnects the graph
Differential betweeness	$\phi_6(i)$	B_i = number of geodesics containing i

and l are in the same class, of some a priori partition of the node set into classes or blocks. More generally, any partition of the set of edges may be used to define equality constraints for an effect in a logit model.

Nondirected relations. Tables 1 and 2 set out a collection of possible parameters for logit models for nondirected network relations. The parameters are distinguished in the table according to the homogeneity of their associated effects; in addition, they may be differentiated by the type of structural property that they are intended to represent and the nature of the conditional dependence among network relations that is presumed to underlie them.

For instance, the parameters listed in Tables 1 and 2 reflect structural features such as: the degree to which a network is dense, and the degree to which denseness varies according to some prespecified relation Y or some blockmodel $\mathcal{B}$; the degree to which the network displays clustering; the degree to which nodes in the network have degree of at least k, and the degree to which there is variation in nodal degrees; the degree of connectivity in the network and the level and variability of connectedness of node pairs; and the degree to which nodes in the network differ in their connectedness, closeness, and centrality in the network.

Some of these parameters may be specified in logit models assuming edge independence (for instance, θ, ϕ_5, ζ_{rs}, and γ_i); some in logit models for Markov random graphs (such as τ, σ, and v_k); and some implicitly assume more complex conditional dependencies (such as π_k, ν, ν_i, ϕ_1, ϕ_2, ϕ_3, ϕ_4, ϕ_6, $\phi_4(i)$, and $\phi_6(i)$). For example, the k-path parameter π_k allows dependencies among all potential edges lying on paths of length k, in contrast to Markov random graphs which may be conceptualized as permitting dependencies only among edges on paths of length 2.

Note that it is not always easy to specify the pattern of assumed dependencies: the connectivity parameter ν, for instance, is associated with a pattern of dependencies among edges that might best be regarded as having arbitrary complexity.

The choice of an hypothesized logit model for an observed network relation depends on several considerations, which we now outline.

First, of course, the effects of primary interest (such as clustering, centralization, and so on) are likely to be dictated by theoretical concerns. Indeed, a large number of descriptive studies provide some preliminary support to the claim that clustering, centralization, positional similarity, and so on, are often both theoretically and empirically important network features.

Second, assumptions about conditional dependence among possible edges of the network relation may be recognized by including parameters corresponding to sufficient subgraphs in a model. Thus, if Markov random graphs are assumed, then models might contain parameters for density, clustering and k-stars (for $k = 2, \ldots, g - 1$), although, we might also fit models in which some of these parameters are set equal to zero and others set to be equal to each other across sets of possible edges.

Third, as indicated earlier, it may sometimes be useful to control for lower-order effects by including their associated parameters in the logit model. For example, one could examine the degree to which a network exhibits clustering after variations among nodes in "expansiveness" have been taken into account. This might be a particularly useful strategy where the lower-order effects have been constrained by methodological restrictions (for instance, where the maximum number of nodes mentioned in response to a relational question is limited by design).

Fourth, such effects can be hypothesized to be homogeneous (Table 1) or nonhomogeneous (Table 2) and, in the latter case, be subject to equality constraints or not. Such generality allows one to focus on a sequence of nested models and a comparison of the fit of these nested models. It may be noted here that we are allowing any effect of the model to be subject to equality constraints, not just density, as in stochastic blockmodels (see chapter 16 of Wasserman & Faust, 1994).

For instance, a plausible model for a relation may specify the presence of two subgroups having denser within-group than between-group ties, and a tendency for a centralized structure within each group, but the degree of centralization might be hypothesized to be greater in one group than another. (Thus, in this two-block case, we might propose a logit model that includes four density parameters for within- and between-block relations, and two centralization parameters, one for each subgroup.) Such parameters are listed in Table 2.

Further, the equality constraints need not be restricted to having the form of a blockmodel; instead, any partition on the set of possible edges may define the equality constraints. An example arises in the case where information is available on a second network relation and where the presence of a relation of this second kind may be regarded as "affecting" the first relation. For instance, for the data presented in Tables 1 through 3, we might propose that the probability of a Work tie between nodes i and j varies according to whether i and j are linked by a "Get on With" tie. Such a model has two density parameters, one for node pairs (i, j) that are linked by a "Get on With" tie, and one for node pairs not linked in this way. Such parameters are listed at the end of Table 1.

Fifth, we need to consider possible dependencies among effects and insure that the logit model is identified. Consider, for instance, the nondirected version of the p_1^* model, for which the vector $\mathbf{z}(\mathbf{x})$ of explanatory statistics consists of the total number of ties L, and the degree X_{i+} of each node i. Since $L = \underline{\Sigma X_{i+}}$, there is a linear dependence among the explanatory variables; hence, we need to impose an additional constraint (such as requiring that the parameters associated with the degrees sum to zero, $\Sigma\ \gamma_i = 0$), or that one of the parameters is fixed at zero ($\gamma_1 = 0$, for example). Some other notable dependencies among parameters in Table 2 are that block density effects sum to the overall density effect (and that any differential block effects sum to the overall corresponding effect), that star effects are dependent on differential choice, overall density, and so forth. Caution is clearly necessary here.

Directed relations. Some parameters of interest for logit models for directed relations are set out in Table 3 (homogeneous effects and effects subject to equality constraints) and Table 4 (node-specific effects). As the tables make clear, there are a large number of effects of possible interest when modeling a directed graph, including centralization, clusterability, prestige, connectivity, transitivity, reciprocity, and block effects. Many of the listed node-specific effects are derived from indices of centralization and prestige, reflecting the prominence of these notions in the network literature (see chapter 5 of Wasserman & Faust, 1994; see also Faust & Wasserman, 1992; Koehly & Wasserman, 1994).

Several of the effects also reflect the frequency of various forms of triads in a directed relation. For instance, the cyclicity parameter τ_C reflects the frequency of cyclic triads in the relation. One possible approach, similar in spirit to that described by Holland and Leinhardt (1975, 1978), is to define parameters corresponding to the 16 distinct types of triad (the triad isomorphism classes) in a directed relation (see also chapter 14 of Wasserman & Faust, 1994). One can then assess the fit of a number of models of interest, especially the class described by Johnsen (1985, 1986). More complete and complex association parameters are also possible, such as those focussing on conformity, exchange and complementarity (see Wasserman, 1987).

Table 3: Homogeneous parameters and graph statistics for logit models for directed relations

Label	Parameter	Graph statistic z(x)
Homogeneous effects		
Choice	θ	$L = X_{++}$
Reciprocity	ρ	$M = \sum X_{ij}X_{ji}$
Transitivity	τ_T	$T_T = \sum_{i,j,k} X_{ij}X_{jk}X_{ik}$
Intransitivity	τ_I	$T_I = \sum_{i,j,k} X_{ij}X_{jk}(1 - X_{ik})$
Cyclicity	τ_C	$T_C = \sum_{i,j,k} X_{ij}X_{jk}X_{ki}$
2-in-stars	σ_1	$S_1 = \sum_{i,j,k} X_{ji}X_{ki}$
2-out-stars	σ_2	$S_2 = \sum_{i,j,k} X_{ij}X_{ik}$
2-mixed-stars	σ_3	$S_3 = \sum_{i,j,k} X_{ji}X_{ik}$
⋮	⋮	⋮
k-in-stars	υ_{kI}	$U_{kI} = \sum_{i,j_1,j_2,\ldots,j_k} X_{j_1i}X_{j_2i}\cdots X_{j_ki}$
k-out-stars	υ_{kO}	$U_{kO} = \sum_{i,j_1,j_2,\ldots,j_k} X_{ij_1}X_{ij_2}\cdots X_{ij_k}$
⋮	⋮	⋮
3-paths	π_3	$P_3 = \sum_{i,j,k,l} X_{ij}X_{jk}X_{kl}$
⋮	⋮	⋮
k-paths	π_k	$P_k = \sum_{i,j_1,j_2,\ldots,j_k} X_{ij_1}X_{j_1j_2}\cdots X_{j_{k-1}j_k}$
Connectivity	ν	N = minimum number of edges whose removal disconnects the digraph (edge connectivity)
Indices of prominence homogeneity (centralization/prestige)		
	ϕ_1	C_{G_1} = sum of lengths of geodesics (d_{ij})
	ϕ_2	C_{G_2} = variance of geodesic lengths (d_{ij})
	ϕ_3	C_E = maximum geodesic length (eccentricity)
	ϕ_4	C_C = variance of geodesic length sums (d_{i+}) (closeness centralization)
	ϕ_5	C_D = variance of X_{i+}'s (degree centralization)
	ϕ_6	C_B = variance of numbers of geodesics containing i (betweeness centralization)
	ϕ_7	P_P = variance of geodesic length sums (d_{+i}) (proximity group prestige)
	ϕ_8	P_D = variance of X_{+i}'s (degree group prestige)
Association with fixed **Y**	Γ	$G = \sum_{i,j} X_{ij}Y_{ij}$
Block effects	ζ_{rs}	$B_{rs} = \sum_{i,j} X_{ij}\delta_{ij;rs}$

Interpretation of parameters. Parameters corresponding to each of the effects listed in these tables may be interpreted in terms of their contribution to the "likelihood" of occurrence of networks with the relevant feature of interest. For instance, a large positive value of a parameter corresponding to an index of clusterability means that clusterable networks are more likely to occur (*ceteris paribus*). Further, this tendency can be assessed statistically, at least approximately, by comparing the fit of two models, one with the parameter and one without.

Perhaps the most interesting models here are those that begin with the graph statistics of p_1^*, and add counts of k-stars across all actors, or only for those actors in the same blockmodel class. Latter models with such parameters are referred to as *Markov block-*

Table 4: Actor-level parameters and graph statistics for logit models for directed relations

Label	Parameter	Graph statistic z(x)
Node-specific/actor-level effects		
Differential expansiveness	α_i	X_{i+} = outdegree (degree centrality)
Differential attractiveness	β_i	X_{+i} = indegree (degree prestige)
Differential closeness	$\phi_4(i)$	$d_{i.}$ = average geodesic distance from (closeness centrality)
Differential closeness	$\phi_7(i)$	$d_{.i}$ = average geodesic distance to (proximity prestige)
Differential betweeness	$\phi_6(i)$	B_i = number of geodesics containing i
Differential connectedness	ν_i	N = minimum number of edges incident to i whose removal disconnects the digraph

models by Strauss and Ikeda. All in all, there is a real wealth of models here, far more than is possible within the p_1 framework.

We note that many other effects of possible interest may be used in this general framework; just some of the possibilities are mentioned in the tables. The particular effects of interest (and their interpretation) depend on the content of the relation **X** and the questions of interest. Many of the widely used indices of structure in a network are closely related to the graph statistics listed in the table (see Wasserman & Faust, 1994). These indices include those for centralization, clusterability, prestige, connectivity, transitivity, and reciprocity.

4.2.5. *Estimation*

Estimation of the parameters of p_1, or of any model that assumes dyadic independence, is not particularly difficult. The likelihood function, being the product of the probabilities for each dyad, is easy to write down. Such is not the case for dyadic dependence models, such as p^*. As noted earlier, the likelihood function for the parameters $\boldsymbol{\theta}$ of p^* depends on the complicated normalizing constant $\kappa(\boldsymbol{\theta})$, which makes maximum likelihood estimation difficult. Indeed, work with this paradigm was limited until Strauss and Ikeda (1990) realized that new, approximate estimation techniques could be used to estimate $\boldsymbol{\theta}$ and the probabilities of relational ties.

Frank and Strauss (1986) discussed inference for Markov graphs, but their discussion was limited. They stated:

> Standard likelihood techniques for the Markov models are not immediately applicable because of the complicated functional dependence of the normalizing constant on the parameters. If the number of [actors] is less than, say, six, it is feasible to evaluate [the normalizing constant, κ] by direct enumeration; otherwise, fitting the model to data can be a difficult problem. (p. 836)

They gave a few possible approaches, but the only promising approach is via logistic regressions. This idea was described in brief by Frank and Strauss, but it was not until Strauss and Ikeda (1990) that the theory behind this approximation was discussed in detail. It then became possible to fit these models, albeit approximately.

The likelihood function for the general form of p^*, model (4), is

$$L(\boldsymbol{\theta}) = \frac{\exp\{\boldsymbol{\theta}'\mathbf{z}(\mathbf{x})\}}{\kappa(\boldsymbol{\theta})},$$

where the dependence on the normalizing constant can easily be seen. An approximate estimation approach, proposed by Strauss (1986) and Strauss and Ikeda (1990), utilizes tools made popular in models for rectangular lattices; specifically, we define the *pseudo-likelihood function* to be

$$PL(\boldsymbol{\theta}) = \prod_{i \neq j} \Pr(X_{ij} = 1 | \mathbf{X}_{ij}^c)^{x_{ij}} \Pr(X_{ij} = 0 | \mathbf{X}_{ij}^c)^{(1 - x_{ij})} \quad (10)$$

and a *maximum pseudolikelihood estimator* to be the value of $\boldsymbol{\theta}$ that maximizes (10). MP estimators are much easier to calculate that ML estimators. MP estimators differ from ML estimators for all but the simplest models (those for which the conditional probabilities are indeed independent of the complement relation). Basically, this "pseudo-" approach assumes conditional independence of the relational ties.

The theorem given below, taken from Strauss and Ikeda (1990), who also gave a proof, gives the very important result that estimation of $\boldsymbol{\theta}$ can be accomplished via logistic regression using any standard logistic regression model-fitting routine.

Theorem. Consider a given logit p^*, as specified in (8). Maximizing the pseudolikelihood given in (10) is equivalent to maximizing the likelihood function for the fit of logistic regression to the model (8) for independent observations $\{x_{ij}\}$. Such logistic regressions can be fit using iteratively reweighted Gauss-Newton computational techniques, as implemented by any logistic regression model package.

The proof of the theorem uses the fact that the derivatives of the pseudolikelihood, set equal to zero, are identical to those obtained from a logistic regression, with the relational variables as data values. Thus, fitting p^* can be done by using the logit p^* form and assuming that the relational variables are actually statistically independent. The idea for this theorem was first suggested by Frank and Strauss (1986) for estimation of the parameters in the triad model.

We have used the statistical software package (SPSS) to fit p^* models, while Strauss and Ikeda used BMDP. There is nothing special about the choice of a package. One treats the $g(g - 1)$ observed binary relational quantities as the measurements on the logit response variable, and then codes a set of explanatory variables, corresponding to the variables specified by $\mathbf{z}(\mathbf{x})$ in the logit p^* formulation (defined here as $\boldsymbol{\delta}(x_{ij})$, the change in the $\mathbf{z}(\mathbf{x})$ vector of network statistics that arises when the variable x_{ij} changes from 1 to 0). An example of such a vector was given earlier, for p_1^*. We illustrate this with several examples shortly. We have written a little C program that takes a sociomatrix $\mathbf{x}$ and produces the vector of measurements on the response variable and the matrix of explanatory variables (both of which have $g(g - 1)$ rows for a directed relation, and $g(g - 1)/2$ for a nondirected relation), which can then be fed to a statistical package.

It is important to note that one does not always have to use iteratively reweighted least squares techniques to fit logistic regressions. For some models, the observed data can be rearranged into a multidimensional contingency table in such a way that the response variable and the explanatory variables are margins of the table. In such cases, as described by Fienberg (1980, chap. 6) and Agresti (1990, chap. 5), the logistic regressions (actually, the logit models) are equivalent to loglinear models containing both the interaction of all the explanatory variables simultaneously, and selected interactions of the explanatory

Table 5: Logit models with homogeneous parameters

	Model	Number of Parameters	Likelihood ratio Statistic	Sum of Absolute Residuals
1	Choice	1	1114.8	400.6
2	Mutuality + Choice	2	977.8	334.6
3	Transitivity + Choice + Mutuality	3	809.9	268.5
4	Cyclicity + Choice + Mutuality	3	976.2	333.8
5	2–Out–Stars + Choice + Mutuality	3	858.7	287.1
6	2–In–Stars + Choice + Mutuality	3	949.7	323.2
7	2–Mixed–Stars + Choice + Mutuality	3	970.3	331.1
8	Degree–centralization + Choice + Mutuality	3	962.8	328.2
9	Degree–prestige + Choice + Mutuality	3	976.8	334.1
10	Transitivity and Cyclicity + Choice + Mutuality	4	725.1	233.5
11	All 2–Stars + Choice + Mutuality	5	786.6	256.8
12	All 2–Stars + Transitivity and Cyclicity + Choice + Mutuality	7	689.3	220.1
13	All 2–Stars, Transitivity and Cyclicity, Centralization and Prestige + Choice + Mutuality	9	681.6	217.5

variables with the response. Thus, parameters can be estimated a bit more easily with iterative proportional fitting. Such is the case with p_1, as well as with p_1^*.

5. Example—"Get on with" Relation

We now consider the fit of a variety of p^* models, for the "Get on with" relation for the Year 7 students shown in Table 11. We also use the "Work with" relation (to study the association between these two relations), shown in Table 12. There are $g = 29$ actors, which can be partitioned *a priori* into boys and girls. This categorization based on gender will be used in our analyses (in order to get *block* effects).

We first note that the example studied here is for illustrative purposes only. We have fitted a large number of models to this relation simply to demonstrate the flexibility of this approach. One would *not* do this in practice. Model fitting should not be done with no substantive forethought.

The set of models and some fit statistics (likelihood ratio statistics and the sum of absolute residuals) are reported in Tables 5 through 10. The tables give fits of some logit models to the "Get on with" relation.

The models in Table 5 are homogeneous models, specifying that tie probabilities depend only on $\mathbf{z}(\mathbf{x})$ statistics such as the total number of ties, the number of mutual ties, the number of transitive or cyclic triads, and the numbers of each of the three types of 2-stars which occur for a directed relation (in-stars, out-stars and mixed-stars). These models include the directed versions of Frank and Strauss' triad and clustering models. Table 3, discussed earlier, displays a list of some possible parameters.

The models in Table 5 all contain an overall choice parameter (θ), and, in view of the

Table 6: Parameter estimates for Models 13 and 30

Model	Effect	Parameter estimate
13	Choice	-1.18
	Mutuality	1.98
	Transitivity	0.26
	Cyclicity	-0.20
	2-in-stars	-0.01
	2-out-stars	-0.15
	2-mixed-stars	-0.08
	Degree-centralization	1.29
	Degree-prestige	-0.49
30	Mutuality	1.33
	Transitivity	0.13
	Choice (boy-boy)	-2.22
	Choice (boy-girl)	-2.95
	Choice (girl-boy)	-4.35
	Choice (girl-girl)	-3.19

Table 7: Logit models in which within-block effects differ from between-block effects (blocks are based on gender)

	Model	Number of Parameters	Likelihood ratio Statistic	Sum of Absolute Residuals
14	Choice + Mutuality + Transitivity + Mutuality-Within-Blocks	4	798.8	263.9
15	Choice + Mutuality + Transitivity + Choice-Within-Blocks	4	791.6	262.6
16	Choice + Mutuality + Transitivity + Transitivity-Within-Blocks	4	801.2	265.5

substantially better fit of Model 2 compared to Model 1, all except Model 2 contain a reciprocity parameter (ρ). Models 4 through 9 add, one at a time, various additional parameters to Model 2, and show, amongst other things, that adding a transitivity parameter leads to a relatively large improvement in fit. Model 10 adds a cyclicity parameter to Model 3 and establishes that the cyclicity effect is substantially stronger in the presence of a transitivity effect. Models 11 and 12 add parameters for the three types of 2-stars to Models 3 and 10, respectively; Model 13 adds the degree-centralization and degree-prestige parameters to Model 12. The addition of 2-stars leads to a modest improvement in fit, while the degree-prestige and degree-centralization parameters lead to no appreciable increase in fit.

The parameter estimates for Model 13 are listed in Table 6. One can see that there is certainly a tendency for relational ties that increase mutuality, degree-centralization, and transitivity to increase the log odds, and hence to be more likely to be present. Ties that increase the other statistics are less likely to be present.

Table 8: p_1^* and Generalizations

	Model	Number of Parameters	Likelihood ratio Statistic	Sum of Absolute Residuals
17	p_1^*	58	683.7	218.5
18	p_1^* + Choice–within–blocks	59	644.7	204.5
19	p_1^* + Transitivity	59	661.0	210.5
20	p_1^* + Cyclicity	60	675.9	216.1
21	p_1^* + Degree–centralization	59	683.6	218.3
22	p_1^* + Degree–prestige	59	668.0	210.1
23	p_1^* + Transitivity + Choice–within–blocks	60	639.5	202.9
24	p_1^* + Transitivity + Cyclicity	60	661.0	210.5
25	p_1^* + Transitivity + Cyclicity + Degree–prestige + Degree–centralization	62	644.7	202.2
26	$p_1^* + G$	59	602.5	189.0
27	p_1^* + Transitivity + G	60	595.3	186.8
28	p_1^* + Transitivity + Cyclicity + Degree–centralization + Degree–prestige + G + Choice–within–blocks	64	572.7	177.0

It is apparent from the fit of these models that in order to adequately summarize the structural tendencies in these data, one would probably want to include at least choice, mutuality, and transitivity parameters in a logit model for the data, and possibly cyclicity, 2-out-star, and 2-in-star parameters as well.

Table 7 summarizes the fit of some models that permit differential within-block and between-block effects for choice, mutuality, and transitivity. Since blocks are based on gender in this example, these models therefore permit different choice, mutuality and transitivity effects for within-gender (boy-boy and girl-girl) and between-gender (boy-girl

Table 9: Logit models with differential block-density effects (generalizations of Wang and Wong blockmodels)

	Model	Number of Parameters	Likelihood ratio Statistic	Sum of Absolute Residuals
29	Mutuality + Differential choice	5	895.7	302.0
30	Mutuality + Transitivity + Differential choice	6	753.0	247.0
31	Mutuality + Transitivity + Differential choice + Cyclicity + All 2–stars + Degree–centralization + Degree–prestige	12	675.7	215.5
32	Choice + Mutuality + G	3	780.1	254.5
33	Choice + Mutuality + G + Transitivity	4	702.9	226.7
34	Mutuality + Differential choice + G	6	755.1	245.8
35	Differential choice + Mutuality + Transitivity + G	7	684.5	220.1
36	Mutuality + G + Transitivity + Degree–centralization + Cyclicity + Degree–prestige + All 2–stars + Differential choice	13	626.4	197.9

Table 10: General differential block effects (based on gender)

	Model	Number of Parameters	Likelihood ratio Statistic	Sum of Absolute Residuals
37	Choice + Differential mutuality	5	935.5	315.2
38	Choice + Mutuality + Differential degree–prestige	4	967.8	330.6
39	Choice + Mutuality + Differential degree–centralization	4	948.1	322.4
40	Differential mutuality + Differential choice	8	887.0	299.3
41	Differential mutuality + Differential choice + Transitivity	9	744.6	244.2
42	Choice + Mutuality + Differential transitivity	6	748.7	244.6
43	Mutuality + Differential transitivity + Differential choice	9	739.8	243.1
44	Differential choice + Differential mutuality + Differential transitivity	12	729.0	240.0
45	Differential mutuality + Differential choice + G	9	735.9	239.1
46	Mutuality + Differential choice + Differential within–block transitivity	7	831.0	275.6
47	Differential choice + Differential mutuality + G + All 2–stars + Degree–prestige + Transitivity	16	610.1	193.4
48	Differential choice + Mutuality + Transitivity + Differential within–block transitivity	8	751.0	245.7

and girl-boy) ties. Comparing Models 14 through 16 with Model 3, we see a modest tendency for stronger within-block choice effects, but only small tendencies for stronger within-block reciprocity and transitivity.

The models summarized in Table 8 include p_1^* and various generalizations. Model 18 adds a single within-block choice parameter to Model 17, with a reasonable increase in fit. This model is a logit formulation of a stochastic blockmodel in which block effects are all equal and occur only on the diagonal of the blockmodel. Models 19 through 25 augment p_1^* with one or more other structural parameters, but Model 18 appears to provide one of the more useful improvements in fit over p_1^*. Models 26 through 28 illustrate how different density effects can be fitted in the presence and absence of another relation (the "Work With" relation). The parameter γ corresponds to a graph statistic (G) that is a count of the number of pairs of nodes joined by both "Work With" and "Get on With" ties, and so reflects the greater likelihood of a "Get on With" tie in the presence of a "Work With" tie.

We also considered models with differential choice effects according to the gender of the sender and recipient of a tie (see Table 9). Model 29 specifies a mutuality effect and different choice parameters for each of these four types of tie (boy-boy, boy-girl, girl-boy, girl-girl). It is a better fit than Model 2, suggesting that the likelihood of ties varies among and between blocks. Model 30 adds a transitivity parameter to Model 2, while Model 31 adds various other homogeneous effects as well. Models 32 through 36 add a parameter for association with the "Work With" relation. It may be noted that Model 30 appears to provide a better fit to the data than Model 15; thus, the model that constrains within-gender effects to be equal (that is, boy-boy and girl-girl) and between-gender effects to be equal (that is, boy-girl and girl-boy) does not provide as good a fit to the data as the model

Table 11: Vickers and Chan data — Get on with relation

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1	0	0	0	0	0	1	0	1	0	0	1	1	0	1	0	1	1	0	1	0	1	1	0	0	0	1	1	0	0
2	1	0	1	1	1	1	1	1	0	1	1	0	0	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	0
3	0	0	0	1	0	0	1	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	1	1	0	1	1	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	1	1	0	0	0	1	0	1	0	1	1	0	0	1	1	1	0	0	1	1	1	1	1	0	0	0	1	1	1
6	1	1	1	1	1	0	1	1	0	1	1	1	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	0	1
7	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
9	0	1	1	1	0	0	1	1	0	0	1	1	0	1	0	1	1	1	1	0	0	0	1	1	1	1	0	1	1
10	0	1	0	0	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	0	0	0	1	1	0	1	0	0	1	0	0	1	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1
14	0	1	0	0	1	0	0	0	0	0	1	0	0	0	1	1	1	0	1	1	1	1	1	1	0	1	1	0	0
15	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1
16	1	1	0	0	1	0	0	0	0	1	1	1	0	1	1	0	1	0	1	1	1	1	1	1	0	1	0	0	0
17	1	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	0	0	1	0	0	0	0	1	0	1	0	1	1
18	0	0	0	0	0	0	0	1	1	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	0	1	1	1	1	1	0	1	1	0	0
20	0	1	0	0	1	0	0	0	0	0	1	0	1	1	0	1	0	0	1	0	1	1	1	1	1	1	1	0	0
21	1	1	0	0	1	0	0	0	0	0	1	0	0	1	0	1	0	0	1	1	0	1	1	0	0	0	0	0	0
22	0	1	0	0	1	0	0	0	0	0	1	0	0	1	1	1	0	0	1	1	1	0	1	0	0	1	1	0	1
23	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	1	0	0	1	1	1	0	1
24	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1
25	0	0	0	0	0	0	0	1	1	0	1	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	0	1	1	0	0	0	1	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	1	1	1	1	0	0	0	1	0	1	1
28	0	1	0	1	1	0	0	1	1	0	0	1	1	0	1	0	1	0	0	1	1	0	1	1	0	1	1	0	1
29	0	1	0	1	0	0	0	1	1	0	0	0	1	0	0	0	1	0	0	0	0	0	1	1	0	1	1	1	0

fitting four separate choice effects. The parameter estimates for Model 30 are shown in Table 6 and suggest that boy-girl ties are actually much more likely than girl-boy ties.

Finally, the models presented in Table 10 illustrate how differential effects can be investigated for other structural parameters, including reciprocity, degree-centralization, degree-prestige, and transitivity.

6. Concluding Comments

The previous section, implementing the models described in this paper, illustrates the enormous number of models that one can fit to a social relation. The parameters in these

Table 12: Vickers and Chan data — Work with relation

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29
1	0	1	0	0	1	1	0	0	0	0	1	1	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	1	1	0	1	0	1	0	0	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0
3	0	0	0	1	0	0	1	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	1	0	0	0	1	0	1	0	0	1	0	0	0	1	1	0	0	1	0	1	0	0	0	0	0	0	0	0
6	1	1	1	1	1	0	0	1	0	0	1	0	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1
7	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
8	1	1	1	1	1	1	1	0	0	1	1	1	0	1	0	1	1	0	1	1	1	1	1	1	0	0	0	0	0
9	0	0	1	0	0	0	0	1	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	0	0	0
10	0	1	0	0	1	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
11	1	1	1	1	1	1	1	1	0	1	0	1	0	1	0	1	0	0	0	1	1	1	1	0	0	0	0	0	0
12	1	1	1	0	1	1	0	0	0	0	1	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	1	1	1	1	1	1	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	1	1	1	1	1	0	1	1	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	1	0	0	1	1	1	0	0	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	1	1	1	1	1	0	0	1	1	1
20	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	0	1	1	1	0	0	0	1	0	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	0	1	1	0	0	0	0	0	0
22	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	0	1	0	0	1	1	0	0
23	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	1	1	1	1	0	1	0	0	0	0	0
24	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	1	0	0	0	0	0	0	1	1	0	1
25	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	1	0	0
27	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
28	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	1	1	0	1	1	0	1
29	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	1	1	0

models, in a very novel way, quantify the effects of structural tendencies on the probability of ties being present or absent. This parameterization of structural effects is a unique feature of this methodology. One could quickly become lost in these tables however; substantive theory should always guide such fitting.

Clearly, much of what we do in this paper relies on an approximate estimation strategy, since standard maximum likelihood estimation is quite hard for large networks. Rather than working with a likelihood function, we work with a pseudo-likelihood function, which we then maximize (and thus, obtain the equivalence to logistic regression).

Maximum likelihood estimation appears to be very difficult in this setting, just as it is in much of modern applied statistics (for example, with covariance structure models and test theory). Further investigation of the quality of maximum pseudolikelihood estimates is certainly called for, and will be the topic of future research.

Comparisons of p_1-type models and the ML estimates of their parameters with logit p^* models and their approximate, MP parameter estimates have been made by Strauss and Ikeda (1990). We have also investigated how much is "lost" in MP estimation in very small networks (Walker, 1995). The bottom line from this initial research (which is quite reassuring) is that approximate MP estimates are quite close to their exact, ML counterparts. We can proceed to postulate p^*-type models, and fit them approximately, and probably not lose too much in the process (over exact estimation).

These preliminary results will certainly be augmented in the future, particularly since the advent of computationally-intensive ML estimation techniques (such as the Gibbs sampler and Markov chain Monte Carlo ideas) should make ML estimation more feasible.

In spite of this approximate estimation approach, the models and estimation strategy proposed here have substantial benefits. This approach has tremendous flexibility to express plausible and interesting structural assumptions, coupled with ease in model fitting.

There is more to be done to generalize these models to other types of relations. Pattison and Wasserman (in press) describe some of these extensions to valued and bivariate relations.

References

Agresti, A. (1990). *Categorical data analysis*. New York: John Wiley and Sons.

Anderson, C. J., & Wasserman, S. (1995). Logmultilinear models for valued social relations. *Sociological Methods & Research, 24*, 96–127.

Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B, 36*, 192–236.

Faust, K., & Wasserman, S. (1992). Centrality and prestige: A review and synthesis. *Journal of Quantitative Anthropology, 4*, 23–78.

Fienberg, S. E. (1980). *The analysis of cross-classified, categorical data* (2nd ed.). Cambridge, MA: The MIT Press.

Fienberg, S. E., & Wasserman, S. (1981). Categorical data analysis of single sociometric relations. In S. Leinhardt (Ed.), *Sociological methodology 1981* (pp. 156–192). San Francisco: Jossey-Bass.

Frank, O., & Strauss, D. (1986). Markov random graphs. *Journal of the American Statistical Association, 81*, 832–842.

Holland, P. W., & Leinhardt, S. (1973). The structural implications of measurement error in sociometry. *Journal of Mathematical Sociology, 3*, 85–111.

Holland, P. W., & Leinhardt, S. (1975). The statistical analysis of local structure in social networks. In D. R. Heise (Ed.), *Sociological methodology 1976* (pp. 1–45). San Francisco: Jossey-Bass.

Holland, P. W., & Leinhardt, S. (1977). Notes on the statistical analysis of social network data. Unpublished manuscript.

Holland, P. W., & Leinhardt, S. (1978). An omnibus test for social structure using triads. *Sociological Methods & Research, 7*, 227–256.

Holland, P. W., & Leinhardt, S. (1979). Structural sociometry. In P. W. Holland & S. Leinhardt (Eds.), *Perspectives on social network research* (pp. 63–83). New York: Academic Press.

Holland, P. W., & Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *Journal of the American Statistical Association, 76*, 33–65. (with discussion)

Iacobucci, D., & Wasserman, S. (1990). Social networks with two sets of actors. *Psychometrika, 55*, 707–720.

Ising, E. (1925). Beitrag zur theorie des ferramagnetismus. *Zeitschrift fur Physik, 31*, 253–258.

Johnsen, E. C. (1985). Network macrostructure models for the Davis-Leinhardt set of empirical sociomatrices. *Social Networks, 7*, 203–224.

Johnsen, E. C. (1986). Structure and process: Agreement models for friendship formation. *Social Networks, 8*, 257–306.

Kindermann, R. P., & Snell, J. L. (1980). On the relation between Markov random fields and social networks. *Journal of Mathematical Sociology, 8*, 1–13.

Koehly, L., & Wasserman, S. (1994). Classification of actors in a social network based on stochastic centrality and prestige. Manuscript submitted for publication.

Pattison, P., & Wasserman, S. (in press). Logit models and logistic regressions for social networks: II. Extensions and generalizations to valued and bivariate relations. *Journal of Quantitative Anthropology.*

Reitz, K. P. (1982). Using log linear analysis with network data: Another look at Sampson's monastery. *Social Networks, 4,* 243–256.

Ripley, B. (1981). *Spatial statistics.* New York: Wiley.

Sampson, S. F. (1968). *A novitiate in a period of change: An experimental and case study of relationships.* Unpublished doctoral dissertation, Department of Sociology, Cornell University, Ithaca, NY.

Snijders, T. A. B. (1991). Enumeration and simulation methods for 0-1 matrices with given marginals. *Psychometrika, 56,* 397–417.

Speed, T. P. (1978). Relations between models for spatial data, contingency tables, and Markov fields on graphs. *Supplement Advances in Applied Probability, 10,* 111–122.

Strauss, D. (1977). Clustering on colored lattices. *Journal of Applied Probability, 14,* 135–143.

Strauss, D. (1986). On a general class of models for interaction. *SIAM Review, 28,* 513–527.

Strauss, D. (1992). The many faces of logistic regression. *The American Statistician, 46,* 321–327.

Strauss, D., & Ikeda, M. (1990). Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association, 85,* 204–212.

Vickers, M. (1981). *Relational analysis: An applied evaluation.* Unpublished Masters of Science thesis, Department of Psychology, University of Melbourne, Australia.

Vickers, M., & Chan, S. (1981). *Representing classroom social structure.* Melbourne: Victoria Institute of Secondary Education.

Walker, M. E. (1995). *Statistical models for social support networks: Application of exponential models to undirected graphs with dyadic dependencies.* Unpublished doctoral dissertation, University of Illinois, Department of Psychology.

Wang, Y. J., & Wong, G. Y. (1987). Stochastic blockmodels for directed graphs. *Journal of the American Statistical Association, 82,* 8–19.

Wasserman, S. (1978). Models for binary directed graphs and their applications. *Advances in Applied Probability, 10,* 803–818.

Wasserman, S. (1987). Conformity of two sociometric relations. *Psychometrika, 52,* 3–18.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* Cambridge, England: Cambridge University Press.

Wasserman, S., & Galaskiewicz, J. (1984). Some generalizations of p_1: External constraints, interactions, and non-binary relations. *Social Networks, 6,* 177–192.

Wasserman, S., & Iacobucci, D. (1986). Statistical analysis of discrete relational data. *British Journal of Mathematical and Statistical Psychology, 39,* 41–64.

Manuscript received 1/20/95
Final version received 6/6/95