PSYCHOMETRIKA—VOL. 66, NO. 2, 161–190
JUNE 2001

NETWORK MODELS FOR SOCIAL INFLUENCE PROCESSES

GARRY ROBINS

UNIVERSITY OF MELBOURNE

PHILIPPA PATTISON

UNIVERSITY OF MELBOURNE

PETER ELLIOTT

SWINBURNE UNIVERSITY OF TECHNOLOGY

This paper generalizes the p^* class of models for social network data to predict individual-level attributes from network ties. The p^* model for social networks permits the modeling of social relationships in terms of particular local relational or network configurations. In this paper we present methods for modeling attribute measures in terms of network ties, and so construct p^* models for the patterns of social influence within a network. Attribute variables are included in a directed dependence graph and the Hammersley-Clifford theorem is employed to derive probability models whose parameters can be estimated using maximum pseudo-likelihood. The models are compared to existing network effects models. They can be interpreted in terms of public or private social influence phenomena within groups. The models are illustrated by an empirical example involving a training course, with trainees' reactions to aspects of the course found to relate to those of their network partners.

Key words: social networks, social influence, p^* models, network effects models, attitudes, graphical models.

1. Introduction and Background

The formation and change of attitudes within groups has a long pedigree as a research question, even though it has appeared under different guises. Individuals within the same social system may tend to share certain attitudes, behaviors or beliefs. Research studies relating to, for instance, group norms (Sherif, 1936/1964), group or organizational culture (Schein, 1985), "collective cognitions" (Klimoski & Mohammed, 1994; Weick & Roberts, 1993), and social influence (Moscovici, 1985) attempt to describe similar processes of collective functioning, whereby cognition and social context interact. Such processes are observable in group settings, with several clear demonstrations since Sherif's (1936/1964) seminal work on group norms. Nevertheless, the interaction between social cognition and social context is of such complexity that there is considerable room for further modeling in this area (Pattison, 1994).

This article develops a new class of network models for social influence processes, by generalizing the p^* class of network models (Frank & Strauss, 1986; Pattison & Wasserman, 1999; Robins, Pattison & Wasserman, 1999; Strauss & Ikeda, 1990; Wasserman & Pattison, 1996) to incorporate individual attributes. An *attribute* is regarded as a variable measured at the level of the individual, as distinct from a network *tie*, which indicates a relationship between two indi-

This research was supported by grants from the Australian Research Council. The authors would like to acknowledge the help of Stanley Wasserman, Janice Langan-Fox and Larry Hubert, and would like to thank four anonymous reviewers for their helpful comments. Earlier versions of this article were presented at the North American Conference of the Psychometric Society, Lawrence, Kansas, June, 1999, and at the Australasian Mathematical Psychology Conference, Brisbane, Australia, December 1999.

Requests for reprints should be sent to Garry Robins, School of Behavioural Science, Department of Psychology, The University of Melbourne, Melbourne, Vic. 3010 AUSTRALIA. E-Mail: g.robins@psych.unimelb.edu.au

0033-3123/2001-2/1999-0705-A $00.75/0

viduals. This article follows Leenders (1997) in using *attribute* to refer not just to background variables that remain constant, such as sex and race, but also to changeable variables, including more psychologically based constructs such as measures of *attitude*.

Social influence occurs when an individual adapts his or her behavior, attitudes or beliefs to the behavior, attitudes or beliefs of others in the social system (Leenders, 1997). Influence does not necessarily require face-to-face interaction, but is based on information about other people. Social influence may arise when individuals affect others' behaviors, or when individuals imitate the behaviors of others, irrespective of the intention of the behavior's originator (Marsden & Friedkin, 1994). These processes have also been termed *contagion* (Leenders, 1997).

The models developed below describe the distribution of attributes across a network of relational ties. The ties are construed as the linkages between people that may enable the dissemination of influence. Whether a particular type of tie does indeed facilitate influence across a social structure is of course an empirical question. The models do not attempt to model explicit processes of interpersonal influences and how they might produce consensus in a group. Rather, they are intended to provide a means to investigate the extent to which shared opinions and attitudes can be explained by a pattern of social relations among individuals.

The models below essentially propose that social proximity (for instance, friendship) may lead to similar attitudes. It is worth noting other possible influence processes: for instance, Burt (1987) argued that people who occupy equivalent social positions tend to exhibit similarities. The network analytic literature has advanced various definitions of *equivalence* (see Wasserman & Faust, 1994), and the role of social position in influence structures has been explored in some depth by Noah Friedkin (see, for instance, Friedkin, 1993; Friedkin, 1998; Friedkin & Johnsen, 1997). Although the models below do not directly incorporate equivalence, certain aspects related to social position (such as popularity and expansiveness) are nevertheless modeled.

The use of actor attributes in network models is an established practice: Wasserman and Faust (1994), for instance, offer several applications. There has been a long-standing interest in two different processes that relate attribute and network variables. Firstly, individuals may change their relationships on the basis of the attributes of other individuals; that is, individual attributes may contribute to the formation or change of network ties. Secondly, network structure can affect individual characteristics, in that individuals may be influenced by others with whom they have network ties. Several researchers (e.g., Erickson, 1988; Leenders, 1997) have argued that these two processes are not mutually exclusive, but, rather, are intertwined. Nevertheless, as Erickson pointed out, it may be profitable to simplify the theoretical and analytical task by ignoring the effects of one process while examining the other. The second process—where network structure affects individual characteristics—is the focus of this article. The network is taken as given and the network ties are used to explain the distribution of attributes.

In many descriptive social influence theories, the conceptualization of social structure may lack complexity. For instance, social psychological theories typically rely on broad descriptions of social structure, such as ingroup/outgroup structures (e.g., Sherif, Harvey, White, Hood, & Sherif, 1961/1988) or majorities/minorities (Moscovici & Doise, 1994), although some recent work has recognized the relevance of more complex social topologies (Latane & L'Herrou, 1996).

More sophisticated modeling of social influence structures is relatively rare outside the network analytic literature. An early instance of the general approach adopted in this article—the use of spatially-based models to relate individual attributes to a network—is provided by Winsborough, Quarantelli and Yutzky (1963). Network researchers, such as Carley (1986, 1989), and Dunn & Ginsberg (1986), have also explicitly modeled the interplay between cognition and social structures, modeling that can be seen as relating to social influence processes. Perhaps the best-known mathematical model in this tradition is the network effects model (Doreian, 1982; Erbring & Young, 1979; Friedkin & Johnsen, 1990; Marsden & Friedkin, 1994):

$$\mathbf{x} = \alpha \mathbf{Y}\mathbf{x} + \mathbf{Z}\beta + \varepsilon \tag{1}$$

where $\mathbf{x}$ is a vector with an attitude of each person as entries, $\mathbf{Y}$ a matrix representing the relationships among the actors, $\mathbf{Z}$ a matrix of exogenous attribute variables that may be influential in shaping the attitudes, and ε a vector of residuals. Here α is a parameter for the effect of the network in transmitting attitudes, and β a vector of parameters for the effect of exogenous attributes on an individual's attitude. Friedkin and Johnsen (1997) summarize a discrete time version of this model and discuss the results emerging from different balances between the network and exogenous variables. Friedkin (1998) develops these models further, with an integration of social theory and mathematical modeling in the examination of social influence, and the incorporation of particular local structural information.

We view our own work as being in the tradition of these mathematical network models, but our starting point is the p^* class of models. The p^* family is a class of models for social networks with parameters reflecting a wide variety of possible structural features. The similarities between the models described here and the network effects tradition are clear. Importantly, both approaches explicitly model social influence in terms of dependencies arising from network ties. As discussed below, there are similarities in parameters and statistics relating to network effects in the models. As mentioned above, Friedkin's (1998) approach, as with ours, allows various structural features to be incorporated, although in his approach these effects enter the models less directly than in our modeling framework.

Some differences are also apparent. We use an auto-logistic framework, not the autoregressive approach of the network effects model, and the Hammersley-Clifford theorem underpins our models. As a result, potentially complex dependency structures can be used as a basis for model specification. The models described below include terms analogous to main effects that network effects models do not utilize, and also allow examination of a wider range of possibly important interactions. Network effects approaches, as in Friedkin (1998), on the other hand, allow a potentially large selection of other structural constructs to be included in the modeling process (as predictors of the "probabilities of interpersonal attachments"—Friedkin, p. 60; although his actual models, p. 57, do not use predictors that go beyond the structural effects discussed in this paper). Network effects approaches also make explicit use of various exogenous variables that are not present in the models below. In principle, these various innovations are aspects that might be included in further elaborations of the models presented here.

Section 2 of this paper presents an overview of p^* models. Section 3 provides the theoretical basis to generalize the models by incorporation of attribute variables through a directed dependence graph. For binary attributes, a variety of social influence models are presented in sections 4 and 5. Section 6 argues that these models allow the simultaneous investigation of individual, group-level and relational social influence effects. The theory for polytomous attribute models is presented in section 7. Empirical examples are provided in section 8.

2. p^* Models for Social Networks

2.1. Social Network Data Structures

For a set of n persons or *actors*, we represent a *relational tie* between persons i and j as a binary random variable Y_{ij} where $Y_{ij} = 1$ if person i considers person j as a partner under the relationship, and where $Y_{ij} = 0$, otherwise.

The matrix $\mathbf{Y} = [Y_{ij}]$ can also be regarded as corresponding to a random (directed) graph with the fixed node set $N = \{1, 2, \ldots, n\}$ and a (random) edge directed from node i to node j if $Y_{ij} = 1$. Let $\mathbf{y} = [y_{ij}]$ denote the matrix of realizations of the variable Y_{ij}, with $y_{ij} = 1$ if there is an observed tie from person i to person j, and $y_{ij} = 0$, otherwise.

Relational ties between some ordered pairs of actors may be regarded as impossible. Conventionally, in social network analysis actors are assumed not to have ties with themselves. In the models below, we follow this convention so that diagonal cells in the matrix $\mathbf{Y}$ (that is, variables of the form Y_{ii}) have entries that are not defined. In principle, however, the models can permit more general patterns of excluded cells. An ordered pair of actors between whom a relational tie is possible is termed a *couple*, and the set of all couples is denoted C.

2.2. p^* Network Models

The difficulty in modeling network data structures arises because the substantively interesting cases assume dependencies among the Y_{ij}. The simplest model of any interest assumes dyadic-independence: Y_{ij} is assumed independent of Y_{kl} for $\{i, j\} \neq \{k, l\}$. A loglinear dyadic independence model, the p_1 *model* (Holland & Leinhardt, 1981), is simple enough to be amenable to standard statistical techniques. More complicated dependencies, however, require different approaches.

In the area of spatial statistics it has long been recognised that dependencies among observations create special difficulties for the construction of probability models. Besag (1974), who was interested in stochastic models for spatial processes, used the *Hammersley-Clifford theorem* to formulate a conditional probability model for a finite system of spatially interacting random variables. Besag construed a lattice arrangement of *sites* to represent the spatial distribution associated with the random variables. A discrete-valued variable Z_i is associated with site i and a *neighborhood* relation among sites is used to specify conditional dependencies among variables. If two sites i and j are neighbors, then the corresponding variables are assumed to be conditionally dependent. If z_i is a realization of Z_i, then Z_i and Z_j are *conditionally independent* if

$$
\begin{aligned}
&P\big(Z_i = z_i, Z_j = z_j | \mathbf{Z}^{\#}_{S-\{i,j\}} = \mathbf{z}^{\#}_{S-\{i,j\}}\big) \\
&= P\big(Z_i = z_i | \mathbf{Z}^{\#}_{S-\{i,j\}} = \mathbf{z}^{\#}_{S-\{i,j\}}\big) P\big(Z_j = z_j | \mathbf{Z}^{\#}_{S-\{i,j\}} = \mathbf{z}^{\#}_{S-\{i,j\}}\big)
\end{aligned}
$$

and conditionally dependent otherwise. In this expression, S refers to the entire set of sites corresponding to the vector of all random variables $\mathbf{Z} = (Z_1, Z_2, \ldots, Z_s)$, $S - T$ refers to the set $S \cap \overline{T}$, and $Z^{\#}_{S-T}$ refers to the vector of variables from which $\mathbf{Z}_k$ has been deleted for all $k \in T$. (This notation with the suffix # is more fully explained in Appendix 1 where it is used to distinguish between different transformations of vectors.) Besag defined a *clique* as either a single site or a set of sites that were all neighbors of each other. In these terms, the Hammersley-Clifford theorem can be stated as follows.

Hammersley-Clifford theorem. For a system of s interacting variables $\mathbf{Z} = (Z_i)$, where each Z_i has a finite range of values and where the set of variables obey certain positivity conditions (see below), there is a set of functions Γ_A such that the probability distribution of $\mathbf{Z}$ will obey a relationship of the following form:

$$
\log\left[\frac{P(Z_i = z_i \big| Z^{\#}_{S-\{i\}} = z^{\#}_{S-\{i\}})}{P(Z_i = 0 \big| Z^{\#}_{S-\{i\}} = z^{\#}_{S-\{i\}})}\right] = z_i \sum_{T \subseteq S-\{i\}} \Gamma_{T\cup\{i\}}(\mathbf{z}^{\#}_{T\cup\{i\}}) \prod_{k\in T} z_k, \tag{2}
$$

where Γ_A is a function of $\mathbf{z}^{\#}_A$, and where, if the set of sites $T = \emptyset$, the term $\prod_{k\in T} z_k$ is taken to have the value 1. Moreover, $\Gamma_{T\cup\{i\}} = 0$ unless $T \cup \{i\}$ is a clique.

Besag's (1974) proof of this theorem used the positivity conditions $P(\mathbf{Z} = \mathbf{0}) \neq 0$, and that, if $P(Z_i = z_i) > 0$ for all i, then $P(Z_1 = z_1, Z_2 = z_2, \ldots, Z_s = z_s) > 0$. (Robins et al., 1999, showed that the second condition can be relaxed.) Besag showed that any set of Γ-functions resulted in a coherent probability description of the system. If the system is binary, so that $z_i = 0$ or 1, the Γ-functions become constants, the parameters of the model.

In what follows below, we often use the term ω_i to specify the conditional probability expression

$$\log\left[\frac{P\left(Z_i = z_i \middle| Z^{\#}_{S-\{i\}} = z^{\#}_{S-\{i\}}\right)}{P\left(Z_i = 0 \middle| Z^{\#}_{S-\{i\}} = z^{\#}_{S-\{i\}}\right)}\right].$$

Equation (2) is the *conditional form* of the Hammersley-Clifford model. It is straightforward to derive an equivalent *joint form*, which expresses a probability distribution of spatially arranged variables:

$$P(\mathbf{Z} = \mathbf{z}) = \frac{1}{c}\exp\sum_{T\subseteq S}\Gamma_T(\mathbf{z}^{\#}_T)\prod_{k\in T} z_k, \tag{3}$$

where

$$c = \sum_{\mathbf{Z}}\exp\sum_{T\subseteq S}\Gamma_T(\mathbf{z}^{\#}_T)\prod_{k\in T} z_k$$

normalizes the probability distribution.

Frank and Strauss (1986) were the first to apply this probability description to social networks. In place of a spatial array, they envisaged a more abstract *dependence graph.* Translating Besag's system, a vertex in the dependence graph represented a site and an edge represented a neighborhood relationship between the two respective sites. For social network data, our notion of a *couple* is the equivalent of *site*, and each vertex in the dependence graph represents a couple (i, j) corresponding to a network variable, Y_{ij}. The presence of an edge in the dependence graph between couples (i, j) and (s, t) signifies a conditional dependence between Y_{ij} and Y_{st}. For a single dichotomous network, the Hammersley-Clifford theorem then specifies that

$$p^*(\mathbf{y}) = P(\mathbf{Y} = \mathbf{y}) = \frac{1}{c}\exp\left(\sum_{T\subseteq C}\gamma_T\prod_{(s,t)\in T} y_{st}\right) \tag{4}$$

where

i. T is a subset of C, the set of all couples;
ii. γ_T is a parameter corresponding to T and is nonzero only if T is a clique (that is, T comprises a single couple, or (i, j) and (s, t) are neighbors for all pairs of couples in T); and
iii. $c = \sum_{\mathbf{Y}}\exp(\sum_{T\subseteq C}\gamma_T\prod_{(s,t)\in T} y_{st})$.

This model was introduced by Frank and Strauss (1986) and more fully explicated in three papers that dealt with a single dichotomous relation (Wasserman & Pattison, 1996), networks of multiple relations (Pattison & Wasserman, 1999) and valued relations (Robins et al., 1999).

Maximum likelihood estimation of parameters is computationally intractable for network models with complex dependence graphs (e.g., those that are connected). Strauss and Ikeda (1990) proposed the use of *maximum pseudo-likelihood* estimation as an approximate technique, a suggestion adopted by Wasserman and Pattison (1996). For a binary network, the pseudo-likelihood function is given by:

$$PL(\Gamma) = \prod_{(i,j)\in C} P\left(Y_{ij} = 1|\mathbf{Y}^{\#}_{C-(i,j)}\right)^{y_{ij}} P\left(Y_{ij} = 0|\mathbf{Y}^{\#}_{C-(i,j)}\right)^{(1-y_{ij})}$$

where Γ is a vector of all parameters γ_T in (4). The *maximum pseudo-likelihood estimator* is the value of Γ that maximises the pseudo-likelihood function. Strauss and Ikeda showed that pseudo-likelihood estimation can be conducted using standard logistic regression procedures (see Wasserman & Pattison for a description of how this is done).

As discussed by Robins, Elliot, and Pattison (2001), the use of pseudo-likelihood estimation may be an interim phase: The development of Monte Carlo techniques (Besag & Clifford, 1989; Crouch & Wasserman, 1988; Geyer & Thompson, 1992), may provide alternative estimation procedures. Other possible alternative estimation procedures recently used in network studies include the method of moments (e.g., Van de Bunt, Van Duijn, & Snijders, 1999).

3. Generalizing p^* for Social Influence Models

A useful generalization of the p^* class of models is obtained by including attribute variables in the dependence graph. We assume polytomous attribute variables. Moreover, to model social influence processes, whereby relational ties shape attributes, a form of directionality needs to be incorporated into the dependence graph. The directionality is required to represent the hypothesis that relational ties may affect attributes, but that attributes do not shape relational ties. In contrast, the dependence graph used by Frank and Strauss (1986) represented nondirected dependencies among variables.

3.1. Graphical Modeling

The graphical modeling literature (Cox & Wermuth, 1996; Edwards, 1995; Lauritzen, 1996; Whittaker, 1990; see also Prendergast et al., 1996) provides insights into the representation of directed dependencies. As with the Frank and Strauss (1986) dependence graph, graphical models have variables represented as vertices, with the absence of an edge between vertices signifying that the two variables are conditionally independent.

Directed edges in a graphical model relate explanatory to response variables (Cox & Wermuth, 1996). As is standard in graph theory, an arrow in the graph represents a directed edge. So, a directed edge from a to b in the graph can be used to represent the situation where Z_b is assumed to be a variable in response to explanatory variable Z_a. We adopt the terminology whereby Z_a is referred to as a *parent* of Z_b and Z_b as a *child* of Z_a (Lauritzen & Spiegelhalter, 1988). We utilize such a directed graph to represent dependencies in models that examine the distribution of one set of variables (the *child* block of variables), given the values of another set of variables (the *parent* variables). In our particular case, we wish to examine the distribution of attributes, given a set of observed network ties. The dependencies are directed in the sense that, if Z_a is a *parent* of Z_b, then the functional form of a probability expression for Z_b, conditional on all other variables, depends on the value of Z_a (see Appendix 1). On the other hand, we make no claims about the form of a conditional probability expression for Z_a.

Graphs including both directed and undirected edges can represent a coherent probability structure when the vertex set satisfies a particular partial ordering. The vertices are able to be partitioned into *blocks*, with nondirected edges within a block, and with only directed edges from one block to another, such that all arrows are pointed in the one direction. A graph satisfying this condition is termed a *chain graph* (Wermuth & Lauritzen, 1990). In this paper, we use a two-block chain graph, which can be defined as a graph containing two sets of vertices—which may be termed *parent* and *child* vertices, respectively—with the only edges between the two sets being directed from parent vertices to child vertices, these being the only directed edges in the graph. Nondirected edges may occur within blocks.

In general, an undirected graph can be derived from a chain graph with equivalent conditional independence properties (in graphical modeling, these are referred to as *Markov properties*—see Lauritzen, 1996, or Whittaker, 1990, for a summary of the various results). The undirected graph derived from a directed dependence graph is often referred to as a *moral graph* (Lauritzen & Spiegelhalter, 1988), because it involves introducing edges between parents of the same child (the so-called *marrying of the parents*). For the two-block chain graphs of this article, we define a moral graph as the undirected graph with the same vertex set, but with edges between two vertices a and b in the moral graph if they are connected by an edge or an arrow in the original graph, or if they are both parents of the same child. (This is a particular case of a more general definition; see Lauritzen, p. 7)

A *clique* in a (moral) graph is a single vertex, or a subset of vertices with each pair connected by an edge. A *maximal clique* is a clique that is not a subset of any other clique.

3.2. *A Variant of the Hammersley-Clifford Theorem*

For social influence models we concentrate on the two-block chain graph, with attribute variables $\mathbf{X} = (X_i)$, $i \in \mathsf{N}$ in the child block (response variables) and network variables $\mathbf{Y}$ in the parent block (explanatory variables). (Although in what follows we do not consider exogenous variables as in the network effects model of (1), exogenous variables could be simply incorporated as additional parent variables.) The interest is in a conditional probability description $P(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y})$, modeling the probabilities of observing particular attributes as a function of the network ties.

Theorem. Given a block of parent network variables $\mathbf{Y}$ and assuming appropriate positivity conditions, a conditional probability description for a block of attributes $\mathbf{X} = (X_i)$, $i \in N$, of a set of actors is given by:

$$\omega_i = \log\left[\frac{P(X_i = x_i|\mathbf{X}^{\#}_{J-\{i\}} = \mathbf{x}^{\#}_{J-\{i\}}, \mathbf{Y} = \mathbf{y})}{P(X_i = 0|\mathbf{X}^{\#}_{J-\{i\}} = \mathbf{x}^{\#}_{J-\{i\}}, \mathbf{Y} = \mathbf{y})}\right]$$
$$= x_i \sum_{J\in\zeta(i)} \sum_{R\subseteq J-\{i\}} \sum_{Q\subseteq \mathrm{pa}(R)} \Gamma_{R\cup Q\cup\{i\}}(\mathbf{x}^{\#}_{R\cup\{i\}}, \mathbf{y}^{\#}_{Q}) \prod_{k\in R} x_k \prod_{(s,t)\in Q} y_{st} \tag{5}$$

Here, $\zeta(i)$ is the set of maximal cliques relating to the attribute variable of actor i (that is, if we treat the child block as a dependence graph in its own right, ignoring the parent block, then $\zeta(i)$ is the set of maximal connected subgraphs that include i). In the expansion, J represents the subsets of actors that are involved in the maximal cliques, $\zeta(i)$, and R represents the actors in J, excluding i. As well, $\mathrm{pa}(R)$ is the set of network couples on which parent variables are defined for the set of attribute variables indexed by R. If Q or R are empty, then the relevant product is taken to have the value 1, and $\Gamma_{\emptyset} = 0$. Moreover, the expansion in (5) can be represented by a moral graph wherein each directed edge of the chain graph is replaced by a nondirected edge, and an edge is inserted between two parent variables of the same child. Γ_T is nonzero only if its index set T is a clique in the moral graph.

The proof of this theorem is given in Appendix 1.

Equation (5) is the conditional form. As for (3), there is a joint form that expresses a probability distribution for vectors of attributes, conditional on an observed network:

$$P(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y}) = \frac{1}{c}\exp \sum_{R\subseteq\zeta} \sum_{Q\subseteq \mathrm{pa}(R)} \Gamma_{R\cup Q}(\mathbf{x}^{\#}_{R}, \mathbf{y}^{\#}_{Q}) \prod_{k\in R} x_k \prod_{(s,t)\in Q} y_{st}, \tag{6}$$

where ζ is the set of maximal cliques relating to the attribute variables. Again, $\Gamma_{R\cup Q}$ is nonzero only if $R \cup Q$ is a clique in the moral graph.

When attributes are dichotomous, $x_i = 1$ in (5), and the Γ-terms become constants (labeled γ) that constitute the parameters of the model. We then have a simplified version of (5):

$$\omega_i = \log\left[\frac{P(X_i = 1|\mathbf{X}^{\#}_{J-\{i\}} = \mathbf{x}^{\#}_{J-\{i\}}, \mathbf{Y} = \mathbf{y})}{P(X_i = 0|\mathbf{X}^{\#}_{J-\{i\}} = \mathbf{x}^{\#}_{J-\{i\}}, \mathbf{Y} = \mathbf{y})}\right]$$
$$= \sum_{J\in\zeta(i)} \sum_{R\subseteq J-\{i\}} \sum_{Q\subseteq \mathrm{pa}(R)} \gamma_{R\cup Q\cup\{i\}} \prod_{k\in R} x_k \prod_{(s,t)\in Q} y_{st}. \tag{7}$$

In the case of binary attributes, suppose that $X_i = 1$ signifies that i "possesses" the attribute and that $X_i = 0$ signifies the opposite. Then, the parameter $\gamma_{R\cup Q\cup\{i\}}$ is associated with the statistic

$$\prod_{k\in R} x_k \prod_{(s,t)\in Q} y_{st}.$$

If the parameter is positive, the odds of i possessing the attribute is enhanced as long as the actors in R also have the attribute and as long as the network ties are in place on the couples in Q. Social influence arises because i's attribute is affected by the attributes of the actors in R, who may have social relations with i through the network ties in Q. In sections 4 and 5, we give examples of models based on specific dependency structures that define certain classes of R and Q.

Our strategy for model development is to hypothesize a dependence structure represented by a chain dependence graph. This can then be expressed in terms of the expansion of (5), or of (7) for binary attributes. We then derive simpler models by restricting the number of vertices we consider in $R \cup Q$. This last step is akin to concentrating on main effects and lower order interaction terms, by setting higher-order interaction terms to zero.

3.3. *Sufficient Statistics and Homogeneity Constraints*

Frank and Strauss (1986) assumed a *Markov* condition for conditional dependence among network variables. In a *Markov directed graph*, possible ties are assumed to be conditionally dependent whenever they have an actor in common: that is, the variables Y_{ij} and Y_{st} are conditionally dependent if and only if $\{i, j\} \cap \{s, t\} \neq \emptyset$. By assuming that these are the only dependencies, Frank and Strauss (1986) showed that sufficient statistics for the model are confined to indicators of certain network configurations: *ties*, *reciprocal ties*, *in-stars*, *out-stars*, *mixed-stars*, and all possible triadic configurations.

A *reciprocal* tie occurs between i and j when $y_{ij} = y_{ji} = 1$. A *star* has a number of ties directed towards and away from a particular node. We refer to an (s, t, r)-star when $s + r$ ties are directed to a node, $t + r$ ties directed away from a node, and r of these incoming and outgoing ties are reciprocated (in other words, the actor represented by the node has s incoming ties, t outgoing ties and r reciprocated ties). The *order* of an (s, t, r) star is said to the $s + t + 2r$. A reciprocal tie can then be considered as a (0, 0, 1) star of order 2. A *k-in-star* is a (k, 0, 0) star, whereas a *k-out-star* is a (0, k, 0) star. A *k-mixed-star* is of the form (s, t, 0) where $s, t \neq 0$ and $s + t = k$. Figure 1 depicts these configurations for stars of *order* 2.

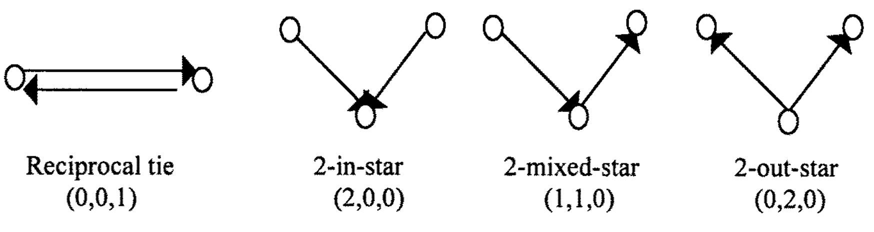

FIGURE 1.
Stars of order 2.

To make the models identifiable, homogeneity constraints of some sort are required. Following the homogeneity strategy originally introduced by Frank and Strauss (1986), Pattison and Wasserman (1999) discussed a general strategy of assuming that parameters corresponding to certain isomorphic configurations of array entries of $\mathbf{Y}$ are equal. For Markov directed graphs, sufficient statistics then become counts of various stars and triads. In the models below, we take a similar approach to homogeneity. The attributes can be construed as colorings on the nodes of the graph of the network, so that we impose homogeneity across parameters when they refer to isomorphic configurations with the isomorphism preserving both edges and colorings. The isomorphism classes represent various types of local configurations in the network. An example of isomorphic configurations for an important model discussed below is depicted in Figure (4). The model includes simple star and reciprocal configurations with colored nodes.

We note that the model defined by (7) is *autologistic*. Equation (7) represents a set of equations (one for each observation i). The variables appear as "explanatory" variables for certain equations in the set (i.e., X_j appears on the RHS in the equation for ω_i) and as "response" variables in other equations (i.e., X_j appears on the LHS in the equation for ω_j). If there are conditional dependencies among attribute variables, certain parameters appear in more than one equation and therefore need to be equated as there is only one parameter for each clique in (5) and (6).

The equating of parameters has important implications that are quite interpretable. For instance, consider the parameter γ_T for the index set $\{i, j, (i, j)\}$. This corresponds to the statistic $X_i X_j Y_{ij}$ in (6). This parameter relates to the effect of a tie between actor i and actor $j(Y_{ij})$ on the attributes of each of the two actors (X_i and X_j). Assuming binary attributes, as in (7), for ω_i, γ_T arises when $R = \{j\}$ and $Q = \{(i, j)\}$, and the estimator for the parameter is $Y_{ij}X_j$. In this case, the parameter relates to the effect on an actor (i) possessing the attribute, when the actor's choice of network partners (i.e., all j such that $Y_{ij} = 1$) also possess the attribute ($X_j = 1$). Because this effect involves choices by the actor we term it an *out-network* effect. On the other hand in the version of Equation (7) for ω_j, γ_T arises when $R = \{i\}$ and $Q = \{(i, j)\}$, and the estimator for the parameter is $Y_{ij}X_i$. In this case, the parameter relates to the effect on an actor (j) possessing the attribute, when those who choose the actor as a network partner (i.e., all i such that $Y_{ij} = 1$) also possess the attribute ($X_i = 1$). Because this effect involves choices of the actor by others, we term it an *in-network* effect. Even though these two cases involve out- and in-network effects, respectively, the parameters are identical.

There may be theoretical reasons to distinguish between those who select or nominate a person as a network partner, and those whom that person selects, as representing possible different generators of influence. For instance, a person may be influenced by those who give him or her attention (*in-network influence*), or a person may be influenced by those he or she respects (*out-network influence*). It is also possible that influence is transmitted by reciprocal ties (*reciprocal network influence*). Although the distinction between in-network and out-network influence may have theoretical value, models without a temporal component cannot distinguish between them. This makes good sense: supposing actor i nominates actor j as a network partner, it is impossible to determine whether similar attitudes of i and j arise through the out-network or through the in-network. It is precisely such parameters that are equated through the requirements of (5), (6) and (7).

4. Binary-Attribute Influence Models: Independent Attributes

The framework described above allows a method to generate and fit models arising from particular dependency assumptions. In social behavioral research, appropriate dependency structures are not self-evident and the challenge for effective model development is to specify realistic dependency structures. While there are some theoretical leads, it is impossible at this stage of knowledge to defend a single set of dependencies as most plausible. Our approach is to construct models under several different sets of assumptions in order to understand the relationship of model form to dependence assumptions. By developing a hierarchical set of nested models, we are also able to compare simpler with more complex models. We can then assess the contribution provided by a more complex dependency structure, and so gain understanding about social behavior when a more complex dependency assumption improves explanatory power.

With this approach in mind, we start with simple dependency structures. We assume that attributes are binary and that the attributes are conditionally independent of each other.

For independent attributes, $\zeta(i) = \{i\}$ and (7) becomes:

$$\omega_i = \sum_{Q \subseteq \mathrm{pa}(i)} \gamma_{\{i\} \cup Q} \prod_{(s,t) \in Q} y_{st}$$

and if homogeneity across actors is assumed so that the γ-terms are not dependent on i, the result is

$$\omega_i = \sum_{Q \subseteq \mathrm{pa}(i)} \gamma_Q \prod_{(s,t) \in Q} y_{st}. \tag{8}$$

One of the simplest descriptions of influence would be to assume that an individual's attribute is conditionally dependent only on the individual's expressed ties. In the directed dependence graph, Y_{ij} is then a parent variable to X_i for all j. The directed dependence graph and moral dependence graph for this case are presented in Figure 2.

The network variables here all relate to outgoing ties. The only possible isomorphic network configurations are out-stars of various orders. With homogeneity imposed across out-stars, (8) becomes

$$\omega_i = \theta + \sigma_{0,1,0} \sum_{k \neq i} y_{ik} + \sigma_{0,2,0} \sum_{k \neq i} \sum_{\substack{j < k \\ j \neq i}} y_{ik} y_{ij} + \dots$$

where θ is an intercept term and a $\sigma_{r,s,t}$ parameter refers to the effect of an (r, s, t)-star. The parameter $\sigma_{0,1,0}$ relates to the number of network partners selected by actor i (*outdegree*) and can be interpreted as representing the effect of an actor making a large number of network choices (an *expansive* actor). If $\sigma_{0,1,0}$ is positive, then actors who are expansive are more likely to possess the attribute, compared to actors who are not expansive. As each of the parameters relates to out-stars of increasing order, this model is referred to as an *out-star independent attribute* model. The formula is equivalent to an arbitrary polynomial of degree $n - 1$ in the out-degree of actor i and, therefore, an arbitrary function of the out-degree.

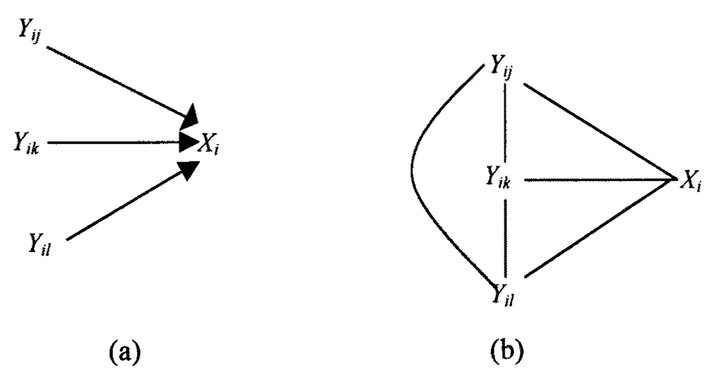

FIGURE 2.
Representation of (a) directed dependence graph and (b) moral dependence graph for out-star model. Note: Only nodes connected to X_i are represented.

It is simple to formulate a counterpart *in-star model*, with actors assumed to be influenced by those who report ties to them and parameters interpreted in terms of the number of choices made by others of the focal actor (*in-degree*), which is related to *popularity*. Again this is an arbitrary polynomial of degree $n - 1$, but this time of the in-degree.

It is also not complicated to combine these models, hypothesizing that an individual's attribute is conditionally dependent on both ingoing and outgoing ties. The moral graph is similar to that represented in Figure 2, except that every network variable involving i is included as a parent variable of X_i, that is, $\text{pa}(i) = \{Y_{ij}, Y_{ji} : j \in N, j \neq i\}$. This implies that every possible network star involving i is represented by a clique in the moral graph. Imposing homogeneity constraints across stars of the same order and type, and—purely for the sake of illustration—limiting the order of configurations to two, the following model results:

$$\begin{aligned}\omega_i = \theta &+ \sigma_{0,0,1} \sum_{k \neq i} y_{ki} y_{ik} + \sigma_{1,0,0} \sum_{k \neq i} y_{ki} + \sigma_{0,1,0} \sum_{k \neq i} y_{ik} \\ &+ \sigma_{2,0,0} \sum_{k \neq i} \sum_{\substack{j<k \\ j \neq i}} y_{ki} y_{ji} + \sigma_{0,2,0} \sum_{k \neq i} \sum_{\substack{j<k \\ j \neq i}} y_{ik} y_{ij} + \sigma_{1,1,0} \sum_{k \neq i} \sum_{j \neq i} y_{ki} y_{ij}\end{aligned} \tag{9}$$

where, as before, θ is the intercept term and the σ parameters refer to various reciprocity, in-star, out-star and mixed-star effects. What we have done here is to simplify the out-degree and in-degree polynomials to an arbitrary quadratic of the in- and out-degrees.

It is worthwhile commenting on the interpretation to be applied to out- and in-star parameters of order higher than 1. Robins (1998) showed that in univariate p^* models higher order out- and in-star parameters provided a refinement of expansiveness and popularity interpretations based solely on the lowest order out-star and in-star parameters. The effect of the higher order parameters is to vary the strength of the expansiveness and popularity effects across different ranges of outdegrees and indegrees, respectively. Robins (1998) concluded that, as a first approximation, it is often convenient to ignore these refinements and set to zero higher order out-star and in-star parameters, unless the interest is specifically directed toward the detail of expansiveness or popularity. For hierarchical models, it follows that parameters pertaining to mixed-stars other than (1, 1, 0)-mixed stars should also be set to zero: for instance, a higher order (r, 1, 0) mixed-star with $r > 1$ contains an (r, 0, 0) in-star as a lower order configuration, and with lower order parameters set to zero it makes sense to set the higher order parameter to zero as well. As a simplifying first approximation, this approach is followed below. What in effect we are doing here is to investigate models in which the major effects in the polynomial expressions concerning in- and out-degrees are assumed to be linear.

5. Dependent Binary Attribute Models

A more interesting, and arguably more realistic, notion of social influence allows influence effects to arise not just from the ties that an individual has with others but also from the attributes of those others. Such models relax the assumption of independent attributes. A possible alternative is that an individual's attribute is conditionally dependent on the attributes of every other individual. In other words, all X_i are conditionally dependent on each other in the dependence graph, so that the subgraph of the dependence graph with attribute variables as vertices is *complete* (i.e., there are edges between all vertices).

Unfortunately, any model with a complete dependence graph among the response variables is not identifiable and does not meet the positivity conditions of the Hammersley-Clifford theorem (Robins, 1998). Several alternative approaches are discussed by Robins, including selective homogeneity constraints arising from *partial conditional independence* structures, the subject of ongoing work (Pattison & Robins, 2000). For the purposes of this article, we confine ourselves to an alternative item-response model.

5.1. Modeling Responses to Individual Items

This approach requires that the attribute variables are measured by responses to a set of items intended to measure the one shared construct. Instead of modeling an overall attribute variable based on the total score of responses, the proposal is to model responses to individual items. This type of model is particularly suited to investigating attitudes among a group of people measured through multi-item scales.

The following simple Markov-related assumptions are made for a dependence graph. Item responses by an actor i are assumed to be conditionally dependent on network variables in which i is one of the actors in the network couple. Item-responses are conditionally dependent on each other if they are responses by the same actor, or if they are responses to the same item.

Denote i's response to the s-th item as $X_i^{(s)}$. Let $\mathbf{X}$ be the matrix with $X_i^{(s)}$ as the entry for the cell (i, s). The dependency assumptions can be represented in the directed dependence graph and moral graph depicted in Figure 3. Note that there is no longer a complete dependence graph among the response variables, because there is no dependency between $X_j^{(s)}$ and $X_i^{(t)}$ for distinct (j, s) and (i, t).

Let $\omega_i^{(s)}$ refer to the conditional logit for i's response to the s-th item. Then, with some simplifying assumptions described immediately below, the following model results:

$$\begin{aligned}\omega_i^{(s)} &= \theta + \lambda \sum_{k\neq i} x_k^{(s)} + \nu \sum_{t\neq s} x_i^{(t)} + \sigma_{0,1,0} \sum_{k\neq i} y_{ik} + \sigma_{1,0,0} \sum_{k\neq i} y_{ki} + \sigma_{0,0,1} \sum_{k\neq i} y_{ik}y_{ki} \\ &\quad + \gamma_{0,1,0} \sum_{t\neq s} x_i^{(t)} \sum_{k\neq i} y_{ik} + \gamma_{1,0,0} \sum_{t\neq s} x_i^{(t)} \sum_{k\neq i} y_{ki} + \gamma_{0,0,1} \sum_{t\neq s} x_i^{(t)} \sum_{k\neq i} y_{ik}y_{ki} \\ &\quad + \eta \sum_{k\neq i} x_k^{(s)}(y_{ik} + y_{ki}) + \eta_R \sum_{k\neq i} x_k^{(s)} y_{ik}y_{ki}. \end{aligned} \tag{10}$$

In this model, interactions among item-responses are assumed to be zero, and interactions among network variables are restricted to reciprocal ties. Here, the λ parameter relates to responses by other actors to the same item, the ν parameter relates to responses by the focal actor to other items, the γ parameters concern an interaction of the focal person's responses with network variables, and the η parameters deal with interactions between other actors' responses to the same item and network variables.

Homogeneity constraints have again been applied across isomorphic configurations of the same order and of the same type. The isomorphism classes of configurations are depicted in

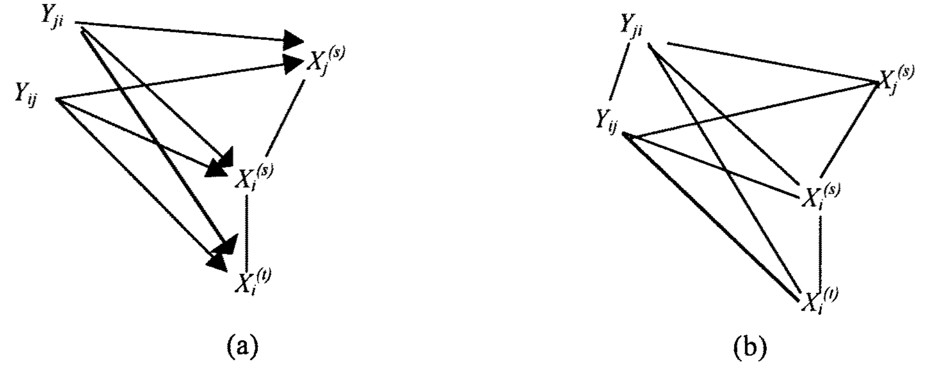

FIGURE 3.
Representation of (a) directed dependence graph and (b) moral graph for item-response model. Note: Only nodes connected to $X_i^{(s)}$ are represented.

Figure 4. Constraints relating to the λ and ν parameters assume, respectively, that actors influence each others' responses to the same degree, irrespective of the item (that is, no particular item is more likely to be the subject of influence than any other item), and that no particular actor is more predisposed to be influenced than any other. (The latter constraint assumes that there is a homogeneous predisposition to respond to items in a particular way—such as, to agree to items—and can be relaxed if there are sufficient items to allow individual parameters to be modeled.) Analogous constraints are applied to the γ parameters.

5.2. Relations with IRT and Network Effects Models

It is worthwhile commenting on similarities and differences between this approach and standard Item Response Theory (IRT) models. IRT models contain individual item parameters. Our homogeneity assumptions above imply that items are modeled as interchangeable. This is primarily a matter of simplicity at this point, and there are no impediments to a more constrained homogeneity assumption that would allow different effects for different items. (In the small-n examples below, there is simply not enough data to fit such a model.) A model for individual item parameters would have the same form as (10) but with separate parameters (except for the σ parameters) for different items.

IRT models, however, typically assume local independence once individual item and person parameters are incorporated in the model. The models that we are presenting here do not make that assumption. For instance, in Figure 3, an approach analogous to IRT would be to incorporate in the parent block an additional indicator variable for each item and for each person, each item indicator variable being a parent of all responses to that particular item, and each person indicator variable a parent of all responses by that person. Local independence then would imply that responses by different individuals to the same item are explained by the item and person variables, that is, that there would be no connection between $X_i^{(s)}$ and $X_j^{(s)}$ in Figure 3, given that they are both responses to the same item. An IRT-type model based on this type of dependency assumption would then have the following form (ignoring interactions between network variables and item indicator variables):

$$\omega_i^{(s)} = \theta_i + \beta_s + \sigma_{0,1,0} \sum_{k \neq i} y_{ik} + \sigma_{1,0,0} \sum_{k \neq i} y_{ki} + \sigma_{0,0,1} \sum_{k \neq i} y_{ik} y_{ki}, \tag{11}$$

where θ_i is a parameter relating to each person while β_s relates to each item.

But in so doing, terms such as $x_k^{(s)} y_{ik}$ are lost to the resulting model. These are precisely the terms that relate networks and attributes, not just in the models developed here but also in the Network Effects models. Of course, there are many circumstances when a local independence assumption is appropriate. The models here, however, deal with situations when local independence may not be observed.

It is also worth comparing this model with network effects models. Some differences are apparent in that our framework is auto-logistic, rather than auto-regressive. In addition, the network effects model in (1) applies to a single attribute, rather than a scale of items as is the case in (10) (although conceptually the generalization of (1) would not be difficult). Nevertheless, as noted above, both models contain parameters relating to the interaction between network and attribute variables, parameters from which inferences can be made about social influence effects arising from the network. An auto-logistic version of the network effects model, then, would include the η parameters but other parameters in (10) have no counterparts in (1). Such a model in effect would assume certain interactions without assessing main effects, and would also set some other interactions to zero.

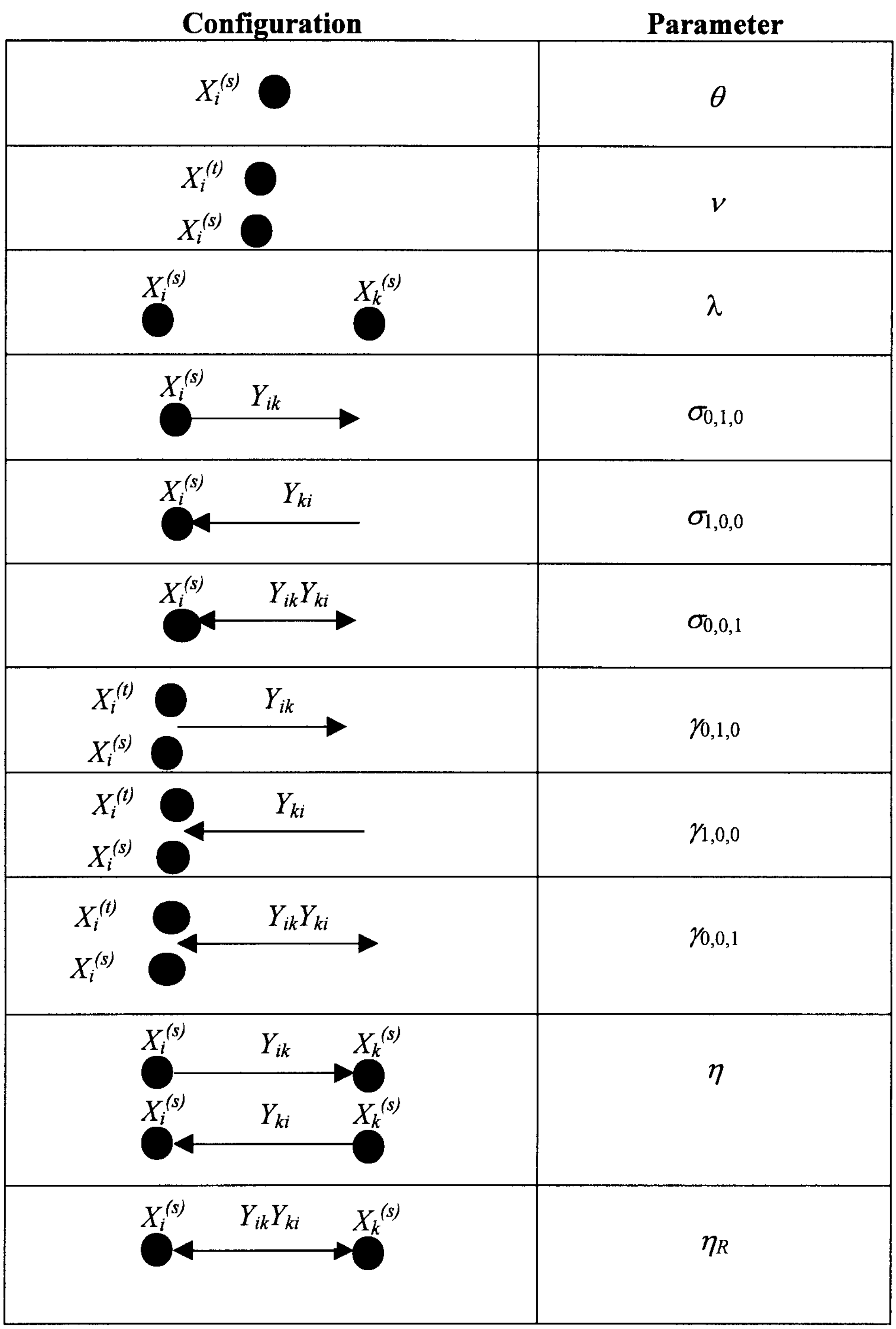

FIGURE 4.
Classes of isomorphic configurations. Note: Single headed arrows represent ties; double headed arrows represent reciprocal ties; dots represent a positive response to an item by an individual.

6. Interpretation: Models for Group Processes

For model (10), there are several possible approaches to the interpretation of a parameter such as λ. In more traditional IRT terms, such a parameter could be seen as a consequence of the distribution of item difficulties. In the development of network effects models, however, similar parameters have been discussed in terms of a more substantive interpretation relating specifically to social influence phenomena. In (10), λ relates to the responses of all other persons to item s, irrespective of relationships within the group. Erbring and Young (1979) discussed such a parameter as a "contextual" effect. They were concerned that in the absence of any interaction among individuals, such an effect represented "social telepathy" and that to invoke constructs such as "common fate" or "group norms" as an interpretation was to ignore other mediating variables that might be better modeled directly. They preferred to drop the parameter from their models, but their arguments were applied to a closely supervised examination, where no public responses are available. Our data sets, however, relate to groups of people who discuss issues both privately, as well as publicly in group meetings. Here the possibility of public responses shaping others' attitudes is obviously possible, irrespective of network ties, without any need to invoke social telepathy, common fate or group norms.

In this article, we construe the parameter as a general attitudinal level effect, one that could possibly be interpreted as representing "public" effects. In this case λ might encapsulate two different influence processes: first, the "public influence" process, whereby the group as a whole exerts influence on individuals and, secondly, influence transmitted through networks other than $\mathbf{Y}$. (Although we are attracted to and use this type of interpretation in what follows, we accept that the interpretation of the parameter is likely to depend strongly on the circumstances in which the data is collected. More work is required both in model development and in empirical situations to identify the various components that might contribute to the effects associated with λ.)

In a similar way, the ν parameter relates to i's response to items other than s. This effect describes a propensity for people in the group to respond in a particular way. If we accept these interpretations of the λ and ν parameters, together they pertain to effects associated with others in the group without consideration of the pattern of network ties. We term these types of effects *collective effects*. (We are assuming throughout this discussion that higher order effects are absent or negligible.)

The σ parameters pertain to network ties (incoming and outgoing ties), regardless of the item-responses of other actors. The $\sigma_{1,0,0}$ and $\sigma_{0,1,0}$ parameters relate to i's levels of popularity and expansiveness in both directed and reciprocated networks. In a very rudimentary sense, these parameters relate to the *social position* of i in the network. The γ parameters concern the interaction of i's responses to other items with social position statistics. The effects pertaining to the σ and γ parameters are termed *social position effects*.

The η parameters, on the other hand, concern the interaction of network ties and others' item-responses. A substantial η implies that an important predictor of i's response is the item-responses among the network partners of i. A resultant inference is that social influence is transmitted through the network. These social influence parameters are termed *network effects*.

If these interpretations are accepted, then (10) contains three different elements: collective effects, social position effects, and network effects. These are utilised to predict individual attributes. The model, then, contains several features of the modeling of group process that are usefully combined into the one model:

i. variables relating to individuals in the group ($X_i^{(s)}$);
ii. parameters relating to the "collective functioning" of the group (λ and ν);
iii. parameters relating to an individual's social position (the σ and γ parameters);
iv. parameters that specify the spread of influence through the network (η).

We do not claim (10) as a particularly sophisticated model of group structure and process. Nevertheless, it does begin to address a number of different features that are theoretically interesting in group process (see, for instance, Markovsky & Chaffee, 1995; Markovsky & Lawler, 1994). An aspect that is not addressed in the model is the emergence of group structure over time (see Arrow & McGrath, 1995; Stokman & Zeggelink, 1996), a feature that is particularly apposite in influence processes (Friedkin & Johnsen, 1990).

Desirably, models of group structure and process should allow some interpretation of the nature of individual group members; the nature of the group "as a whole"; interactions among members; process within the group; and temporal change. To incorporate these features in the one model is not straightforward. Equation (10) is a simple example that may allow some interpretation in the first four of these features through the variables and parameters listed above in the items i through iv. Temporal change is discussed in the conclusion.

7. Attitudes: Polytomous Attribute Measures

The above discussion has been restricted to binary item-responses. This section derives polytomous models for an attitude questionnaire with three response categories for each item: agreement, neutral, disagreement. Following the approach of Robins et al. (1999), we develop a trichotomous logit model, with the neutral response as the baseline category.

An attribute model derived for binary data, such as (10), can be generalized to a trichotomous version with the same dependence structure. The imposition of homogeneity demands a more complicated parameterization to account for three possible responses to each item. Robins et al. (1999) showed that conversion of a trichotomous data structure to a three-way binary (person by item by response) array permits estimation of the model. Certain cells that represent situations with zero probability (that is, cells that imply a variable can have two distinct values simultaneously) need to be excluded from the model. The response dimension of this array can be decomposed into two levels—an a and a d level corresponding to the *agreement* and *disagreement*. We create two dummy variables, $a_i^{(s)}$ and $d_i^{(s)}$:

$$\begin{aligned}
&\text{if neutral response, } a_i^{(s)} = d_i^{(s)} = 0;\\
&\text{if agreement, } a_i^{(s)} = 1 \quad \text{and} \quad d_i^{(s)} = 0;\\
&\text{if disagreement, } a_i^{(s)} = 0 \quad \text{and} \quad d_i^{(s)} = 1.
\end{aligned}$$

As can be seen from this coding, for estimation purposes we omit cells in the three-way array that imply that $a_i^{(s)} = d_i^{(s)} = 1$.

Denote agreement and disagreement conditional logits, respectively, as:

$$\omega_i^{(s)}(a) = \log\left[P\left(a_i^{(s)} = 1|\mathbf{x}^{\#}_{J-(i,s,a)}, \mathbf{y}, d_i^{(s)} \neq 1\right) \Big/ P\left(a_i^{(s)} = 0|\mathbf{x}^{\#}_{J-(i,s,a)}, \mathbf{y}, d_i^{(s)} \neq 1\right)\right],$$

and

$$\omega_i^{(s)}(d) = \log\left[P(d_i^{(s)} = 1|\mathbf{x}^{\#}_{J-(i,s,d)}, \mathbf{y}, a_i^{(s)} \neq 1) \Big/ P\left(d_i^{(s)} = 0|\mathbf{x}^{\#}_{J-(i,s,d)}, \mathbf{y}, a_i^{(s)} \neq 1\right)\right],$$

where $\mathbf{x}$ is the observed three-way array of binary variables, indexed by the set of cells with nonzero probabilities, J, where $(i, s, a) \in J$ refers to the cell for the variable recording person i's possible agreement to item s, and with $\mathbf{y}$ the set of observed network ties.

The resulting model expands (10) in that each parameter has several counterparts in the trichotomous case. For instance, a parameter involved in predicting the conditional log odds of i's agreement to an item is distinct from a parameter relating to the prediction of i's disagreement. Moreover, the prediction of i's agreement may be based on the agreements of others, or the disagreement of others, so that different parameters result. For example, the η parameter in

(10) has three counterparts in a trichotomous model: η^{aa} (the chances of agreement to an item when network partners agree to that item); η^{dd} (the chances of disagreement to an item when network partners also disagree); and η^{ad} (the chances of having the opposite opinion from network partner.)

The agreement logit is as follows:

$$\begin{aligned}\omega_i^{(s)}(a) = {}& \theta^a + \lambda^{aa}\sum_{k\neq i} a_k^{(s)} + \lambda^{ad}\sum_{k\neq i} d_k^{(s)} + \nu^{aa}\sum_{t\neq s} a_i^{(t)} + \nu^{ad}\sum_{t\neq s} d_i^{(t)} + \sigma_{0,1,0}^{a}\sum_{k\neq i} y_{ik}\\ &+ \sigma_{1,0,0}^{a}\sum_{k\neq i} y_{ki} + \sigma_{0,0,1}^{a}\sum_{k\neq i} y_{ik}y_{ki} + \gamma_{0,1,0}^{aa}\sum_{t\neq s} a_i^{(t)}\sum_{k\neq i} y_{ik} + \gamma_{0,1,0}^{ad}\sum_{t\neq s} d_i^{(t)}\sum_{k\neq i} y_{ik}\\ &+ \gamma_{1,0,0}^{aa}\sum_{t\neq s} a_i^{(t)}\sum_{k\neq i} y_{ki} + \gamma_{1,0,0}^{ad}\sum_{t\neq s} a_i^{(t)}\sum_{k\neq i} y_{ki} + \gamma_{0,0,1}^{aa}\sum_{t\neq s} a_i^{(t)}\sum_{k\neq i} y_{ik}y_{ki}\\ &+ \gamma_{0,0,1}^{ad}\sum_{t\neq s} d_i^{(t)}\sum_{k\neq i} y_{ik}y_{ki} + \eta^{aa}\sum_{k\neq i} a_k^{(s)}(y_{ik}+y_{ki}) + \eta^{ad}\sum_{k\neq i} d_k^{(s)}(y_{ik}+y_{ki})\\ &+ \eta_R^{aa}\sum_{k\neq i} a_k^{(s)} y_{ik}y_{ki} + \eta_R^{ad}\sum_{k\neq i} d_k^{(s)} y_{ik}y_{ki}. \end{aligned} \tag{12}$$

There is an analogous formulation for the disagreement logit. Homogeneous effects that occur "across categories" (such as the network effects association between agreement and disagreement, i.e., η^{ad}) are represented in both logits and require the equating of parameters. When we consider parameters from both agreement and disagreement logits we have the following equalities: $\lambda^{ad} = \lambda^{da}$, $\gamma_{0,1,0}^{ad} = \gamma_{0,1,0}^{da}$, $\gamma_{1,0,0}^{ad} = \gamma_{1,0,0}^{da}$, $\gamma_{0,0,1}^{ad} = \gamma_{0,0,1}^{da}$, $\eta^{ad} = \eta^{da}$, and $\eta_R^{ad} = \eta_R^{da}$. So, the disagreement logit is

$$\begin{aligned}\omega_i^{(s)}(d) = {}& \theta^d + \lambda^{ad}\sum_{k\neq i} a_k^{(s)} + \lambda^{dd}\sum_{k\neq i} d_k^{(s)} + \nu^{ad}\sum_{t\neq s} a_i^{(t)} + \nu^{dd}\sum_{t\neq s} d_i^{(t)}\\ &+ \sigma_{0,1,0}^{d}\sum_{k\neq i} y_{ik} + \sigma_{1,0,0}^{d}\sum_{k\neq i} y_{ki} + \sigma_{0,0,1}^{d}\sum_{k\neq i} y_{ik}y_{ki}\\ &+ \gamma_{0,1,0}^{ad}\sum_{t\neq s} a_i^{(t)}\sum_{k\neq i} y_{ik} + \gamma_{0,1,0}^{dd}\sum_{t\neq s} d_i^{(t)}\sum_{k\neq i} y_{ik} + \gamma_{1,0,0}^{ad}\sum_{t\neq s} a_i^{(t)}\sum_{k\neq i} y_{ki}\\ &+ \gamma_{1,0,0}^{dd}\sum_{t\neq s} a_i^{(t)}\sum_{k\neq i} y_{ki} + \gamma_{0,0,1}^{ad}\sum_{t\neq s} a_i^{(t)}\sum_{k\neq i} y_{ik}y_{ki} + \gamma_{0,0,1}^{dd}\sum_{t\neq s} d_i^{(t)}\sum_{k\neq i} y_{ik}y_{ki}\\ &+ \eta^{ad}\sum_{k\neq i} a_k^{(s)}(y_{ik}+y_{ki}) + \eta^{dd}\sum_{k\neq i} d_k^{(s)}(y_{ik}+y_{ki})\\ &+ \eta_R^{ad}\sum_{k\neq i} a_k^{(s)} y_{ik}y_{ki} + \eta_R^{dd}\sum_{k\neq i} d_k^{(s)} y_{ik}y_{ki}. \end{aligned} \tag{13}$$

7.1. *Fitting the Model: Model Selection Process*

The principal model that we fit includes all the parameters in (12) and (13). In considering the fit of the model, we are guided by the pseudo-likelihood deviance and the mean absolute residual, as in Wasserman and Pattison (1996), but we also compare the principal model to a simpler model that does not include any network terms, that is, a model with only the λ and ν parameters. As these parameters relate to the group as a whole, we term this model the *Collective Effects Model*. To infer that the network transmits social influence, a model with network effects parameters has to improve on the Collective Effects model.

For both the collective effects and the network models, we reduce the models by removing unimportant parameters. We use a backward selection approach, removing parameters that do not make a contribution of 4 to the pseudo-likelihood deviance. (The analogous maximum likelihood p-value would be about 0.05.) In the reduced network model, we retain all parameters that are retained in the reduced collective effects model to ensure a nested relationship between the two models, and we also assess effects involving interactions with discarded "collective effects" terms. We check our backward selection of parameters by forward selection. (Occasionally, with smaller data-sets, we find that collinearity or other features of the data can lead to estimation problems when all parameters are entered in the network model. In these cases, forward selection is the preferred approach.) Any network-related parameter retained in the model is then interpreted. As an exploratory aid to interpretation, we have also fitted submodels involving only the network of reciprocal ties, to investigate the effect, if any, of reciprocal influence.

8. Empirical Example: Attitudes on a Training Course

An example of a trichotomous model based on the logits in (12) and (13) is presented by fitting the model to data from a training course conducted by a major Australian government business enterprise. The training course was designed to improve staff participation in the workplace and had a heavy emphasis on group work, either in collective discussions or in smaller subgroups. The group comprised six males and eight females. After four days of training, participants were asked to complete two sociometric questions, listing those with whom they had most communicated during the course (*Course Interaction*); and with whom they had most interacted socially outside training (*Social Interaction*).

Participants were also asked to respond to twenty statements relating to teamwork within the group (Appendix 2), indicating their agreement or disagreement with each statement on a five-point Likert scale, with the midpoint signifying "undecided". (The two points on the scale for agreement and for disagreement were subsequently collapsed into agree and disagree categories.) Relevant items were scored in reverse, so that all agreements were in the direction of a positive statement about the group's performance. In what follows, then, the term "agreement" signifies satisfaction about the functioning of the group, whereas "disagreement" indicates a view that the group was not functioning as fully or as cohesively as it might have.

Table 1 presents the levels of agreement by course participants to each of the items. A number of items were quite uncontroversial and, overall, trainees responded positively about their group, with some 77% of responses expressing agreement. Some items, however, suggested dissension from the notion of a smoothly functioning or completely consensual group, with a total of 12% of responses undecided and 11% disagreements. The imposition of homogeneity

TABLE 1.
Matrix of responses to 20 training items on a 5-point Likert scale ($n = 14$)

1	2	3	2	5	1	2	1	2	1	1	3	2	1	2	2	1	4	1	2
2	2	2	2	2	2	3	2	2	1	2	2	2	2	2	2	2	2	2	2
2	2	2	2	2	2	3	4	2	2	2	2	4	2	2	2	2	4	2	3
2	2	2	2	2	2	2	2	2	1	2	1	2	2	2	2	1	2	2	1
2	2	2	3	2	1	2	2	2	2	3	2	2	2	2	3	2	3	2	2
2	2	3	3	3	2	2	4	2	2	3	2	3	2	2	2	2	2	2	2
2	2	3	2	2	2	3	2	2	2	2	3	3	2	3	2	2	3	2	2
1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	1	2
2	2	2	4	2	2	2	4	3	2	3	2	4	4	4	4	2	4	2	3
1	1	4	1	2	1	2	4	4	1	1	2	5	2	4	4	1	4	2	4
1	1	3	2	2	1	2	5	4	2	2	3	4	3	4	4	1	4	2	3
1	1	1	1	2	1	3	2	2	1	3	1	1	2	2	2	1	2	1	2
2	2	2	2	5	2	2	3	2	2	2	2	2	2	2	2	1	2	2	1
2	2	3	3	4	2	2	2	3	2	2	2	3	2	2	4	2	4	2	2

TABLE 2.
Collective effects model

Parameter	Pseudo-likelihood estimate	Standard error (approx.)
θ^a	−3.00	1.43
θ^d	−3.28	0.91
λ^{aa}	0.26	0.09
λ^{dd}	0.63	0.20
ν^{aa}	0.16	0.06
ν^{dd}	0.52	0.14

across items has to be seen as a first approximation, and with a larger data set there might be value in investigating the effects of relaxing this homogeneity assumption.

For each of the two networks, maximum pseudo-likelihood baseline logits were fitted. The two networks were modeled separately. It would be possible to develop a multi-network model by incorporating multivariate p^* models (Pattison & Wasserman, 1999) into (12) and (13). The number of parameters, however, would escalate accordingly, and given the small number of actors in this example, estimation problems result.

8.1. *Collective Effects Models*

The collective effects model is used as a basis to compare later network models to establish the presence of network effects, which have to improve the fit of the collective effects model to be regarded as meaningful. The collective effects model had a pseudo-likelihood deviance of 185.8 with 6 parameters retained and a mean absolute residual of 0.246. Parameter estimates are provided in Table 2. The Table includes standard errors as calculated by the logistic regression procedure as a guide only. Given the pseudo-likelihood estimation procedures, these have to be seen as approximate at best.

Brief inspection of the estimates reveals that, as might be expected, there are effects reflecting a tendency for actors to agree when others also agree ($\hat{\lambda}^{aa}$) and to disagree when others disagree ($\hat{\lambda}^{dd}$). There are also effects for actors having propensities generally to agree and to disagree ($\hat{\nu}^{aa}$ and $\hat{\nu}^{dd}$). The absence of substantial $\hat{\lambda}^{ad}$ and $\hat{\nu}^{ad}$ effects suggests that, overall, tendencies to take a contrary view are not strong.

8.2. *Course Interaction Network*

Parameter estimates for the model for the course interaction network (with insubstantial parameters removed according to the criteria above) are presented in Table 3. The model has a pseudo-likelihood deviance of 196.6 with 7 parameters, an improvement over the collective

TABLE 3.
Course interaction model

Parameter	Pseudo-likelihood estimate	Standard error (approx.)
θ^a	−3.00	1.43
θ^d	−4.60	1.28
λ^{aa}	0.26	0.09
λ^{dd}	0.40	0.25
ν^{aa}	0.16	0.06
ν^{dd}	0.70	0.19
η^{dd}	0.98	0.35

effects model of 10.8 for one parameter. The mean absolute residual of the model is 0.234, compared with the mean absolute residual of 0.246 for the collective effects model. (A trichotomous version of the IRT-type model in (11), but with homogeneity imposed across persons, and with parameters removed according to the above criteria, results in a similar deviance level of 197 but with 11 parameters. The model, however, shows clear overfitting with several extreme parameter estimates, so that more detailed comparison is pointless.)

The effects in the agreement logit are the same as for the collective effects model, but there is an additional effect in the disagreement logit, η^{dd}, which corresponds to the levels of disagreement by network partners. A person is more likely to disagree with an item if network partners also disagree with that item. This effect also showed up in a submodel based only on the network of reciprocal ties (η_R^{dd}) but the reciprocal parameter dropped out in the presence of η^{dd}. The full network is presented in Figure 5 and the reciprocal network in Figure 6.

Residual analysis is revealing here. We present this analysis not to overemphasize results arising from a single case study, but to illustrate the type of interpretations that can be made under this approach. By comparing residuals from both models across each of the twenty items, we can determine where the course interaction model improves the collective effects model. For most items, residuals are not greatly changed, but there is a substantial improvement for the fifth item, "The training group has been dominated by a few individuals". (Recall that the item is scored in reverse, so disagreement here signifies endorsement of the item.) For the disagreement logit, for items other than the fifth item the absolute mean differences between residuals from the two models was less than 0.13, and for the majority of items less than 0.1. For the fifth item, however, the mean difference between residuals from the two models was 0.37. (As the only

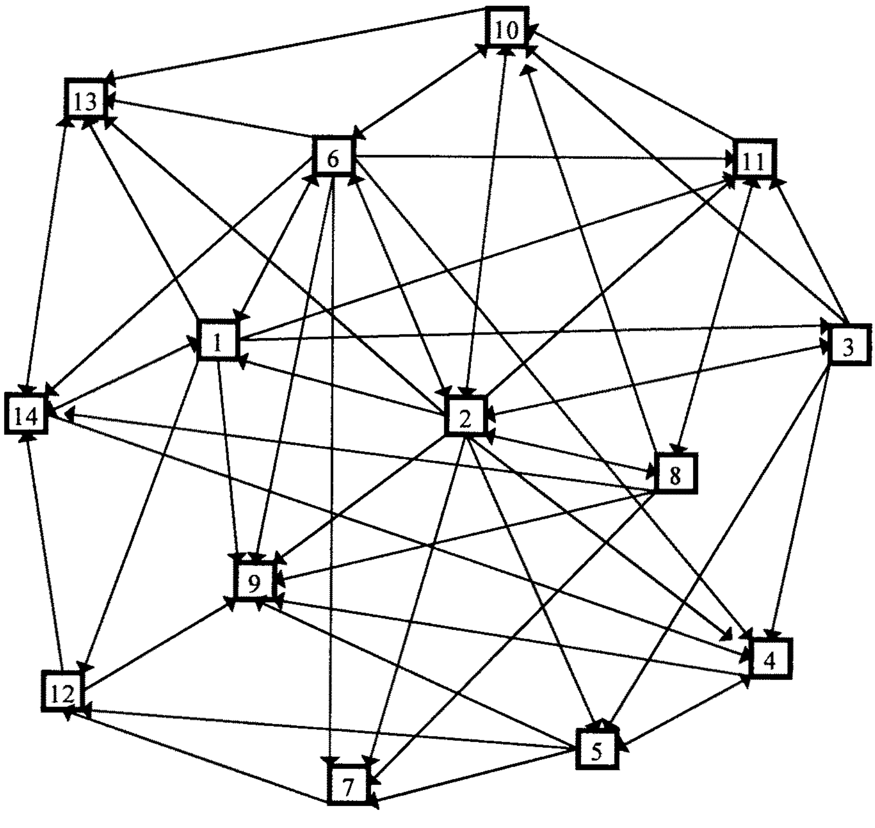

FIGURE 5.
Course interaction network.

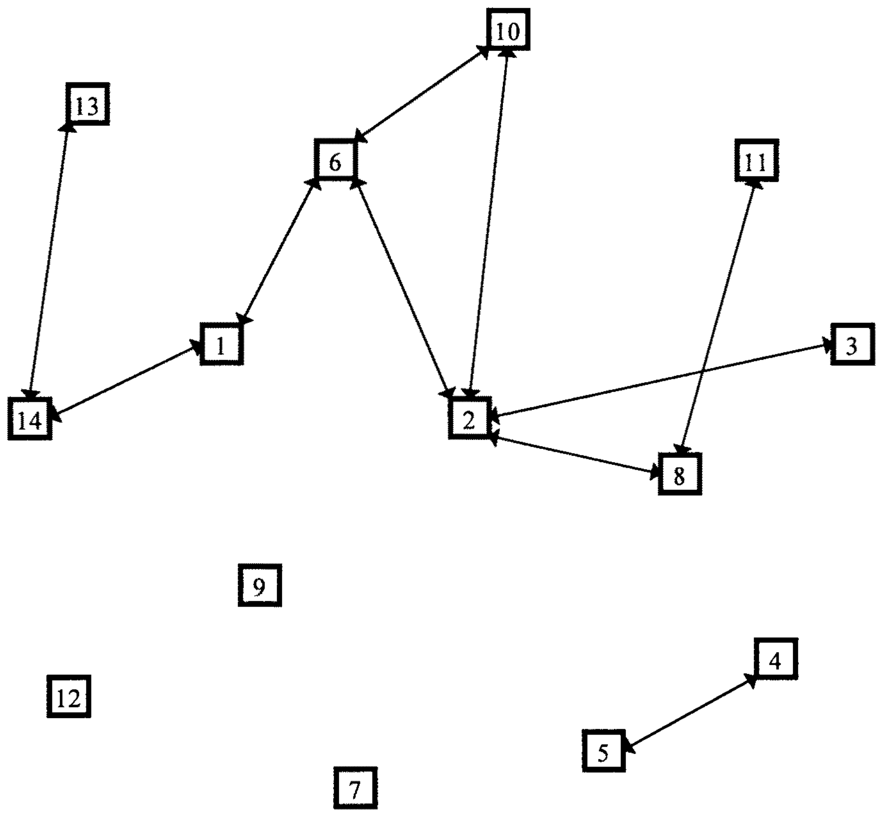

FIGURE 6.
Reciprocal course interaction network.

additional parameter in the network model relates to the disagreement logit, there is no change in residuals for the agreement logit.) Three individuals (1, 13, 14) disagreed with the fifth item, and one (6) was undecided.

Ignoring the directionality of ties in the directed network, we can see that these four individuals make up a clique in the network. (Of course, the directionality of ties is irrelevant here because the model does not differentiate in-network and out-network influence.) In fact, these four individuals participate in the majority of cliques in the network and, in most of these cliques, they make up at least half the membership. The ties among the four, then, are an important feature of the network structure. Because of the simpler nature of the reciprocal network it is easier to examine possible influence structures among the four actors. The four are a connected subgroup in the reciprocal network (Figure 6).

The position of actor 6 in the reciprocal network is interesting. This individual provides the only connection between the "dissenting subgroup" of actors 1, 13 and 14 and a second connected subgroup of actors (2, 3, 8, 10 and 11) who did not believe the course was dominated. In technical terms, actor 6 is a *cutpoint* in the graph of the mutual network (e.g., Wasserman & Faust, 1994), meaning that if actor 6 were not present, the graph would break down into further separated components. In substantive terms, actor 6 provides the only reciprocal ties that could carry influence between the two subgroups. Actor 6 is also the only individual who was undecided about whether the course was dominated. It is tempting to speculate that the position of actor 6, connecting two groups with opposing attitudes on this issue, could have shaped his or her neutrality; and that the two groups are insulated from each other on this issue by the neutrality of their only connection, actor 6. It is unclear whether the neutrality of an individual in such

TABLE 4.
Social interaction model

Parameter	Pseudo-likelihood estimate	Standard error (approx.)
θ^a	−3.00	1.43
θ^d	−3.48	1.01
λ^{aa}	0.26	0.09
λ^{dd}	0.34	0.23
ν^{aa}	0.16	0.06
ν^{dd}	0.49	0.15
η^{dd}	0.91	0.36

a position might be an effort to preserve different options for action (Leifer, 1988), perhaps to maximise the powerful position of bridging two otherwise unconnected subgroups (Burt, 1992), or—in contrast—arises because different subgroups are pressuring that individual to adopt contradictory positions, preventing the individual sustaining both ties and strongly-held attitudes (Krackhardt, 1997).

It is worthwhile summarizing what this model together with the residual analysis has shown. Although the model is homogeneous across both individuals and items, the residual analysis has indicated four individuals who form a clique in the network and who share a common concern about one particular aspect of the course's functioning. Despite the approximations in pseudo-likelihood estimation, the modeling approach developed here provides some evidence that network effects do indeed arise from network processes, and that this can be interpreted as an instance of social influence.

Moreover, it may be claimed that this network effect is primarily a private exchange between individuals, given that variables from the collective model are included. The network effect seems to be over and above the public effects incorporated in the collective model. The inference might be that the perception of a few dominating the course, rather than being openly discussed, was shared privately among some individuals through their interaction ties.

8.3. Social Interaction Network

Parameter estimates for the social interaction model are presented in Table 4. The model has a pseudo-likelihood deviance of 194.9 with 7 parameters and a mean absolute residual of 0.236.

Here the parameter estimate $\hat{\eta}^{dd}$, as with the course interaction network, suggests that an actor is more likely to disagree with an item if network partners also disagree with that item. The social interaction network is depicted in Figure 7.

Residual analysis here reveals that the model improves residuals, not so much for particular items, but for some individuals. The disagreement predictions for actors 1, 3 and 11, especially, are improved for a number of items. It is interesting to note that these actors form a clique with actor 10 who has the highest levels of disagreement. This might suggest that when those four actors were together socially (either as the one subgroup, or dyadically), their disagreement with the group's functioning could have been a topic they discussed.

9. General Conclusions

The course and social interaction models presented above describe how certain influence processes work in this training group from the perspectives of two different networks. This simple empirical example, however, is suggestive of some possible hypotheses for further exploration with more data. For instance, various networks may be involved in the transmission of influence in different ways, over and above the collective effects model. Influence might occur not only through public knowledge but also through private dyadic interactions.

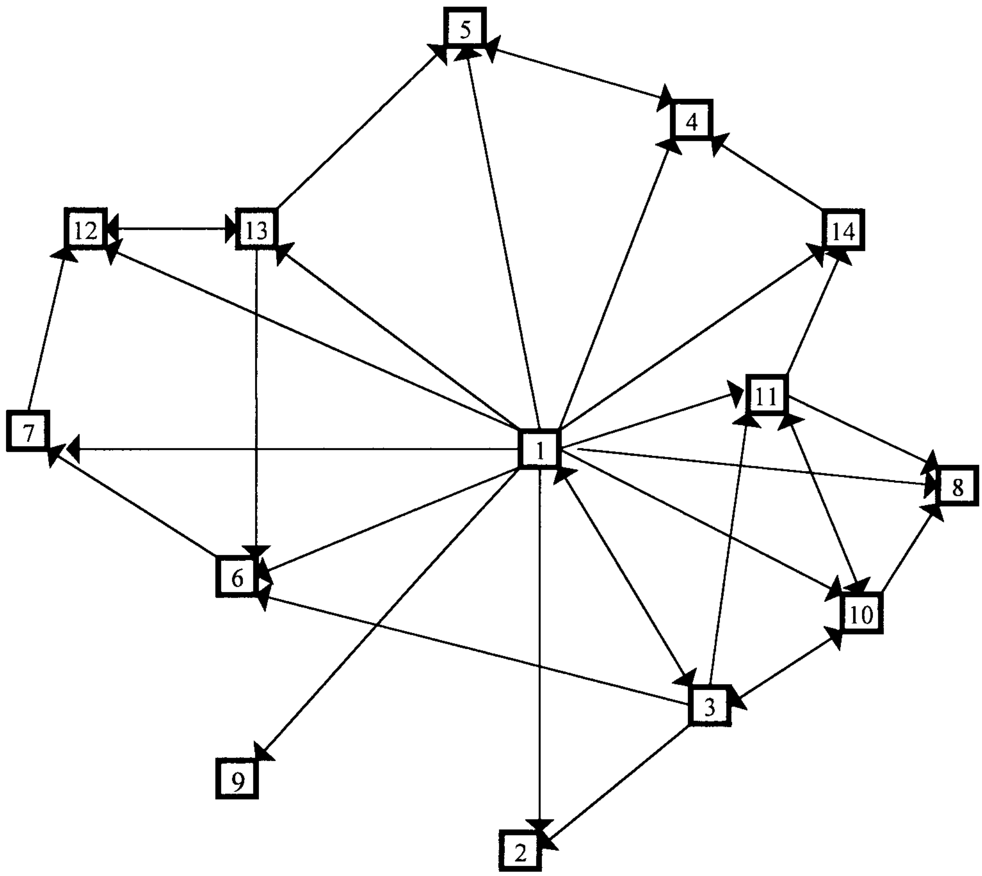

FIGURE 7.
Social interaction network.

In the example, the two networks demonstrated social influence effects through the transmission of disagreement and neutrality. The type of influence that is transmitted through networks may be different from that arising from public knowledge. For instance, in the training group, non-agreement may have been less publicly acceptable and so was discussed more privately through network structures. Moreover, individuals may choose different networks for the transmission of influence. Some local features of networks, but not others, may have significance in transmission of influence. For instance, some cliques may be important to the shaping of particular attitudes, whereas others may not be.

Of course, the issue of what constitutes a properly specified model remains. The proper specification of dependencies among particular structural or individual variables is crucial to the efficacy of the models we have developed. In this paper, we have not taken a strong position on what are the appropriate variables for inclusion in social influence models, nor are we suggesting that the type of dependencies we utilize as illustrations are necessary or optimal. Our aim is simply to set up a particular modeling framework that we believe will prove useful in answering such questions, in conjunction with more empirical and theoretical work relating specifically to social influence phenomena. The issue of what is an appropriate dependency structure for social processes in general (including social influence processes) is inevitably a vexed one that will perhaps always be open to question and further investigation. In ongoing work, we are investigating the effects of generalized dependency structures other than the Markov random graph assumption, considering both more constrained and more generalized dependencies (Pattison & Robins, 2000).

We wish to avoid any confusion between the models we have developed and standard logistic regression. As noted above, these models are auto-logistic, not logistic. The principal and major difference is that in the p^* framework there is no clear differentiation between dependent and independent variables. In the case of the models of this paper (as distinct from previous work on p^*, such as Wasserman & Pattison, 1996), at least some of the variables—those in the parent block of the two block chain graph—are clear counterparts of independent variables. Nevertheless, as pointed out above, the variables in the child block enter into the models in the form of both independent and dependent variables, at least when there are dependencies within the child block (which is the case for the more interesting models). Moreover, the Hammersley-Clifford theorem requires the equating of certain parameters, a process without any counterpart in standard logistic regression. We know from our own experience in developing these models that lack of attention to the equating of parameters can result in models that are not justified by the theorem, and that the equating of particular parameters is not an immediately intuitive step.

The general approach of this article can also be adapted to model processes other than social influence. The variant of the Hammersley-Clifford theorem presented in Appendix 1 can be adapted with attributes as parent variables and network ties as child variables. This enables the modelling of what Leenders (1997) referred to as *selection models*, where relationships emerge from the similarities and dissimilarities among the attributes of individuals. Models along these lines are explicated in Robins, Elliott and Pattison (2001) and are the subject of further investigation in an applied empirical context by Elliott (2000).

The next step is to develop models that allow the simultaneous modelling of social influence and selection processes, for, as discussed above, the two processes are likely to be intertwined in many circumstances (Leenders, 1997). This will require a further adaptation of the Hammersley-Clifford theorem and is a matter for ongoing work. The formulation of these models will involve the consideration of temporal network models, including models that can differentiate between in- and out-network effects discussed in this article.

Appendix 1: The Hammersley-Clifford Theorem and Directed Dependence Graphs

The directed network dependence graph for the models of this article has two blocks with network variables in a "parent" block and attribute variables in a "child" block. Following the proof of Besag (1974), this appendix proves a version of the Hammersley-Clifford theorem appropriate to a general two-block chain graph, where the aim is the prediction of the child variables. Let $\mathbf{X} = (X_1, X_2, \ldots, X_s)$ denote the child variables and $\mathbf{Y} = (Y_1, Y_2, \ldots, Y_t)$ the parent variables. Let $S = \{1, 2, \ldots, s\}$ and $T = \{s+1, s+2, \ldots, s+t\}$. Define the complete set of variables as $\mathbf{Z} = (Z_k | k \in S \cup T)$, where $Z_k = X_k$ for $k \in S$ and $Z_k = Y_{k-s}$ for $k \in T$.

For the proof, the notation used by Robins et al. (1999) is helpful. Firstly, the notation provides a simple means to express the setting to zero of a number of variables. Let $\mathbf{w}$ be an arbitrary vector $(w_1, w_2, \ldots, w_n)$ with $J = \{1, 2, \ldots, n\}$. Denote by $\mathbf{w}_A$, for $A \subseteq J$, the vector $\mathbf{w}$ with the entries indexed by $J - A$ set to zero. For instance, $\mathbf{w}_{\{i\}} = (0, \ldots, 0, w_i, 0, \ldots, 0)$ and $\mathbf{w}_{J-\{i\}} = (w_1, w_2, \ldots, w_{i-1}, 0, w_{i+1}, \ldots, w_n)$.

Secondly, the notation can be used to represent a subset of variables. Let $|A| = m$. Denote by $\mathbf{w}_A^\#$ the ordered m-tuple derived from $\mathbf{w}$ by excluding the entries indexed by $J - A$ but retaining the natural ordering induced by J, so that $\mathbf{w}_A^\# = (w_{a_1}, w_{a_2}, \ldots, w_{a_m})$ where $\{a_1, a_2, \ldots, a_m\} = A$ and $a_j < a_j + 1$. For instance $\mathbf{w}_{\{i\}}^\# = (w_i)$ and $\mathbf{w}_{J-\{i\}}^\# = (w_1, w_2, \ldots, w_{i-1}, w_{i+1}, \ldots, w_n)$. As Robins et al. explain, more generally the notation can be applied to higher-way arrays.

Mutual conditional dependencies occur among the $\mathbf{X}$ variables and directed dependencies from the $\mathbf{Y}$ to the $\mathbf{X}$ variables. We follow Besag (1974) in defining a mutual conditional dependence (*neighborhood*) as occurring among child variables X_i and X_j if

$$P\left(X_i = x_i | \mathbf{Z}^{\#}_{S\cup T-\{i\}} = \mathbf{z}^{\#}_{S\cup T-\{i\}}\right)$$

can be expressed in a functional form dependent on x_j and if

$$P\left(X_j = x_j | \mathbf{Z}^{\#}_{S\cup T-\{i\}} = \mathbf{z}^{\#}_{S\cup T-\{i\}}\right)$$

can be expressed in a functional form dependent on x_i. For a directed dependence, where Y_k is a *parent* of site X_i we mean that the functional form of

$$P\left(X_i = x_i | \mathbf{Y}, \mathbf{X}^{\#}_{S-\{i\}}\right)$$

is dependent on y_k but not so dependent if Y_k is not a parent of X_i. As the models of interest are based on the conditional probability $P(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y})$, we are not concerned about conditional probability expressions for $\mathbf{Y}$.

Because

$$\frac{P(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y})}{P(\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}|\mathbf{Y} = \mathbf{y})} = P\left(X_i = x_i|\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}, \mathbf{Y} = \mathbf{y}\right),$$

then

$$\frac{P(\mathbf{X} = \mathbf{x}_{S-\{i\}}|\mathbf{Y} = \mathbf{y})}{P(\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}|\mathbf{Y} = \mathbf{y})} = P\left(X_i = 0|\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}, \mathbf{Y} = \mathbf{y}\right)$$

when $x_i = 0$, so that

$$\frac{P(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y})}{P(\mathbf{X} = \mathbf{x}_{S-\{i\}}|\mathbf{Y} = \mathbf{y})} = \frac{P(X_i = x_i|\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}, \mathbf{Y} = \mathbf{y})}{P(X_i = 0|\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}, \mathbf{Y} = \mathbf{y})}. \tag{A1}$$

Define

$$Q(\mathbf{z}) = \log\left[P(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y})/P(\mathbf{X} = \mathbf{0}|\mathbf{Y} = \mathbf{y})\right].$$

From (A1), for $i \in S$,

$$\begin{aligned} Q(\mathbf{z}) - Q(\mathbf{z}_{S\cup T-\{i\}}) &= \log\left[\frac{P(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y})}{P(\mathbf{X} = \mathbf{x}_{S-\{i\}}|\mathbf{Y} = \mathbf{y})}\right] \\ &= \log\left[\frac{P(X_i = x_i|\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}, \mathbf{Y} = \mathbf{y})}{P(X_i = 0|\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}, \mathbf{Y} = \mathbf{y})}\right]. \end{aligned} \tag{A2}$$

Given the RHS of this equation and the definitions of neighborhood and parent associations among the variables, the functional form of $Q(\mathbf{z}) - Q(\mathbf{z}_{S\cup T-\{i\}})$ can only include x_i itself, x_j where X_j is a neighbor of X_i, or y_t where Y_t is a parent of X_i.

Now a series of functions of the z_k are introduced, indexed by various subsets of $S \cup T$, and defined recursively. First define $\Gamma_{\{k\}}(z_k)$ by $z_k\Gamma_{\{k\}}(z_k) = Q(\mathbf{z}_{\{k\}})$. Define a Γ-function for a subset of $S \cup T$ with only two elements, say $\{i, j\}$:

$$z_i z_j \Gamma_{\{i,j\}}(z_i, z_j) = Q(\mathbf{z}_{\{i,j\}}) - z_i\Gamma_{\{i\}}(z_i) - z_j\Gamma_{\{j\}}(z_j).$$

Note that $\Gamma_{\{i,j\}}(z_i, z_j) = \Gamma_{\{i,j\}}(\mathbf{z}^{\#}_{\{i,j\}})$. In general, a Γ-function is recursively defined for a subset A of $S \cup T$ in terms of the Q-expression for the relevant variables and the Γ-functions for the subsets of A:

$$\Gamma_A(\mathbf{z}^{\#}_A)\prod_{i\in A} z_i = Q(\mathbf{z}_A) - \sum_{M\subset A}\Gamma_M(\mathbf{z}^{\#}_A)\prod_{i\in M} z_i.$$

With these definitions in place, there is an expansion of Q

$$Q(\mathbf{z}) = \sum_{M \subseteq S \cup T} \Gamma_M(\mathbf{z}_M^{\#}) \prod_{i \in M} z_i, \tag{A3}$$

where $\Gamma_\emptyset$ is defined as zero.

It then follows that

$$\begin{aligned} Q(\mathbf{z}) - Q(\mathbf{z}_{S\cup T-\{i\}}) &= \sum_{M \subseteq S \cup T} \Gamma_M(\mathbf{z}_M^{\#}) \prod_{k \in M} z_k - \sum_{M \subseteq S \cup T-\{i\}} \Gamma_M(\mathbf{z}_M^{\#}) \prod_{k \in M} z_k \\ &= \sum_{M \subseteq S \cup T-\{i\}} \Gamma_{M\cup\{i\}}(\mathbf{z}_{M\cup\{i\}}^{\#}) \prod_{k \in M\cup\{i\}} z_k \\ &= z_i \sum_{M \subseteq S \cup T-\{i\}} \Gamma_{M\cup\{i\}}(\mathbf{z}_{M\cup\{i\}}^{\#}) \prod_{k \in M} z_k, \end{aligned} \tag{A4}$$

where if $M = \emptyset$ then $\prod_{k\in M} z_k$ is defined as one.

This equation holds for all possible observed values $\mathbf{z}$. In particular, take a set of variables $\mathbf{Z}_R^{\#}$ where $i \in R$ and assume that all other observed values are zero. Then because $z_k = 0$ unless $k \in R$ we have from (A2) and (A4):

$$\begin{aligned} Q(\mathbf{z}_R) - Q(\mathbf{z}_{S\cup T-\{i\}}) &= \log\left[\frac{P(X_i = x_i|\mathbf{Z}_{R-\{i\}}^{\#} = \mathbf{z}_{R-\{i\}}^{\#}, \mathbf{Z}_{S\cup T-R}^{\#} = \mathbf{0})}{P(X_i = 0|\mathbf{Z}_{R-\{i\}}^{\#} = \mathbf{z}_{R-\{i\}}^{\#}, \mathbf{Z}_{S\cup T-R}^{\#} = \mathbf{0})}\right] \\ &= x_i \sum_{M \subset R-\{i\}} \Gamma_{M\cup\{i\}}(\mathbf{z}_{M\cup\{i\}}^{\#}) \prod_{k\in M} z_k \end{aligned} \tag{A5}$$

Now suppose that the variables indexed by $R - \{i\}$ are all child variables but none is a neighbor of X_i. In this case $Q(\mathbf{z}) - Q(\mathbf{z}_{S\cup T-\{i\}})$ does not depend on any of the values from that set of variables. Then it follows from (A5) that $\Gamma_{M\cup\{i\}}(\mathbf{z}_{M\cup\{i\}}^{\#}) = 0\ \forall M \subseteq R - \{i\}$. If, on the other hand, R does index a set of neighbors of X_i, because i has been chosen arbitrarily there will be an equation of form similar to (A5) for all elements of R. It follows that, for nonzero $\Gamma_{M\cup\{i\}}$, the variables indexed by elements of M are all neighbors of each other, that is, they form a clique.

Suppose that R indexes only parent variables, but that none is a parent of Y_i. Then, again, $Q(\mathbf{z}) - Q(\mathbf{z}_{S\cup T-\{i\}})$ does not depend on any of the values from that set of variables, so it follows that

$$\Gamma_{M\cup\{i\}}(\mathbf{z}_{M\cup\{i\}}^{\#}) = 0\ \forall M \subseteq R - \{i\}.$$

So for nonzero $\Gamma_{M\cup\{i\}}$, where M indexes only parent variables, these are all parents of Y_i.

Suppose M indexes a union of parent and child variables. In the simplest case, suppose the variables are y_k and x_l. As the functional form of $Q(\mathbf{z}) - Q(\mathbf{z}_{S\cup T-\{i\}})$ can only include variables that are neighbors of X_i or parents of Y_i, if Y_k is not a parent of X_i or if X_l is not a neighbor of X_i, then $\Gamma_{\{i,l,s+k\}} = 0$. Suppose that Y_k is a parent of X_i, that X_l is a neighbor of X_i, but that Y_k is not a parent of X_l. Then, as $Q(\mathbf{z}) - Q(\mathbf{z}_{S\cup T-\{l\}})$ can only include neighbors or parents of X_l, so $\Gamma_{\{i,l,s+k\}}$, which is a term in the expansion of both $Q(\mathbf{z}) - Q(\mathbf{z}_{S\cup T-\{i\}})$ and $Q(\mathbf{z}) - Q(\mathbf{z}_{S\cup T-\{l\}})$, must be zero. Accordingly, $\Gamma_{\{i,l,s+k\}}$ is nonzero only when two conditions hold: Y_k is a parent of both X_i and X_l; and X_i and X_l are neighbors of each other. Analogous arguments show that $\Gamma_{A\cup B}$, for $A \subseteq S$ and $B \subseteq T$, is nonzero if and only if the variables indexed by A (child variables) form a clique of neighbors, and if and only if each of the variables indexed by B (parent variables) is a parent of every variable indexed by A.

In other words, the only nonzero Γ-terms in the expansion (A5) are those pertaining to variables that comprise a clique in the moral graph of the chain graph. This moral graph is formed from the chain graph by drawing an edge between any two parent variables of the same child variable ("marrying the parents") and by replacing all arrows by undirected edges.

Note that this definition of the moral graph differs slightly from the graphical modeling version for a chain graph given by Lauritzen (1996, p. 7). In the Lauritzen version, parents of subsets of connected children are married whereas here only parents of the same child are married. Our definition and the graphical modeling version concur in the special case of a *directed acyclic graph* (a chain graph with all chain components consisting of one vertex, Lauritzen, p. 7.) Our version of a moral graph applies to a two block chain graph when the aim is to model a conditional factorization $P(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y})$, rather than a more general recursive factorization (see Lauritzen). If, for instance, it was intended also to model $P(\mathbf{Y} = \mathbf{y})$, then the Lauritzen version would have to be used. As it is, the additional edges in the Lauritzen moral graph never enter into a model for $P(\mathbf{X} = \mathbf{x}|\mathbf{Y} = \mathbf{y})$ because they never form part of cliques that include any of the $\mathbf{X}$ variables.

Denote conditional logits by

$$\omega_i = \log\left[\frac{P(X_i = x_i|\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}, \mathbf{Y} = \mathbf{y})}{P(X_i = 0|\mathbf{X}^{\#}_{S-\{i\}} = \mathbf{x}^{\#}_{S-\{i\}}, \mathbf{Y} = \mathbf{y})}\right]$$

Then from (A2) and (A5) the Hammersley-Clifford set of equations is:

$$\omega_i = x_i \sum_{M\in\zeta(i)} \sum_{R\subseteq M-\{i\}} \sum_{Q\subseteq \mathrm{pa}(R)} \Gamma_{R\cup Q\cup\{i\}}(\mathbf{x}^{\#}_{R\cup(i)}, \mathbf{y}^{\#}_{Q}) \prod_{k\in R} x_k \prod_{j\in Q} y_j \tag{A6}$$

where $\zeta(i)$ denotes the set of maximal cliques of neighbors that include X_i and $\mathrm{pa}(R)$ denotes the set of parent variables of all child variables indexed by R. When $R = \emptyset$ or $Q = \emptyset$, the products $\prod_{k\in R} x_k$ and $\prod_{j\in Q} y_j$ are defined as equal to 1.

Appendix 2: Teamwork Questionnaire Items

1. The training group has consistently worked together in a cooperative and understanding manner.
2. I feel an atmosphere of honesty and trust exists within this group.
3. Some members of the group seemed reluctant or unable to participate fully or freely.
4. The training group did not complete all the tasks of the course adequately.
5. The training group has been dominated by a few individuals.
6. The facilitators really got everyone involved in the group exercises.
7. The training group has been able to develop consensus on the central issues of the course, even if there have been some strongly held opinions.
8. I was cautious about expressing my full opinions in this group.
9. Communication within the training group has not been as open and effective as it might have been.
10. I feel there is a very friendly environment within this group.
11. The training group has been highly motivated to tackle the course exercises.
12. There were such strong disagreements on some issues that the training group has become fragmented or split.
13. I felt the training group could have handled the tasks and exercises of the course more smoothly and efficiently.
14. The training group has given plenty of feedback—including praise and encouragement—to members during the course.
15. I participated in the course exercises as they came up but without feeling fully involved in the group itself.
16. The training group could have done with clearer direction.
17. Every member of the group has been listened to and has had their opinions heard.

18. On some important issues, my own opinions were different from most others in the group.
19. The training group has approached the course in an intelligent and responsible way.
20. There could have been better team spirit within the training group.

Statements 3, 4, 5, 7, 8, 9, 12, 13, 15, 16, 18 and 20 were scored in reverse, so that all agreements were in the direction of a positive statement about the group's performance.

References

Arrow, H., & McGrath, J.E. (1995). Membership dynamics in groups at work: A theoretical framework. In L.L.Cummings & B.M.Staw (Eds.), *Research in organizational behavior, Vol. 17* (pp. 373–411). Greenwich, CT: JAI Press.

Besag, J.E. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society, Series B, 36*, 96–127.

Besag, J.E., & Clifford, P. (1989). Generalized Monte Carlo significance tests. *Biometrika, 76*, 633–642.

Burt, R.S. (1987). Social contagion and innovation: Cohesion versus structural equivalence. *American Journal of Sociology, 92*, 205–211.

Burt, R.S. (1992). *Structural holes: The social structure of competition.* Cambridge, MA: Harvard University Press.

Carley, K. (1986). An approach for relating social structure to cognitive structure. *Journal of Mathematical Sociology, 12*, 137–189.

Carley, K. (1989). The value of cognitive foundations for dynamic social theory. *Journal of Mathematical Sociology, 14*, 171–208.

Cox, D.R., & Wermuth, N. (1996). *Multivariate dependencies—Models, analysis and interpretation.* London: Chapman & Hall.

Crouch, B., & Wasserman, S. (1988, May). Fitting p^*: Monte Carlo maximum likelihood estimation. Paper presented at International Conference on Social Networks, Barcelona, Spain.

Doreian, P. (1982). Maximum likelihood methods for linear models. *Sociological Methods & Research, 10*, 243–269.

Dunn, W.N., & Ginsberg, A. (1986). A sociocognitive network approach to organizational analysis. *Human Relations, 40*, 955–976.

Edwards, D. (1995). *Introduction to graphical modelling.* New York, NY: Springer-Verlag.

Elliott, P. (2000). *Statistical models for the relationship between personal attributes and social ties: Applications to adolescent female peer groups.* Unpublished doctoral dissertation, University of Melbourne, Department of Psychology.

Erbring, L., & Young, A.A. (1979). Individuals and social structure: Contextual effects as endogenous feedback. *Sociological Methods & Research, 7*, 396–430.

Erickson, B.H. (1988). The relational basis of attitudes. In B. Wellman & S.D. Berkowitz (Eds.), *Social structures: A network approach* (pp. 99–121). Cambridge, UK: Cambridge University Press.

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

Friedkin, N.E. (1993). Structural bases of interpersonal influence in groups: A longitudinal case study. *American Sociological Review, 58*, 861–872.

Friedkin, N.E. (1998). *A structural theory of social influence.* New York, NY: Cambridge University Press.

Friedkin, N.E., & Johnsen, E.C. (1990). Social influence and opinions. *Journal of Mathematical Sociology, 15*, 193–205.

Friedkin, N.E., & Johnsen, E.C. (1997). Social positions in influence networks. *Social Networks, 19*, 210–222.

Geyer, C.J., & Thompson, E.A. (1992). Constrained Monte Carlo maximum likelihood for dependent data. *Journal of the Royal Statistical Society, Series B, 54*, 657-699.

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

Krackhardt, D. (1997, February). *Identity, control and Simmelian ties.* Paper presented at The White Tie Event in honour of Harrison White, San Diego, CA.

Klimoski, R., & Mohammed, S. (1994).Team mental model: Construct or metaphor? *Journal of Management, 20*, 403–437.

Latane, B., & L'Herrou, T. (1996). Spatial clustering in the conformity game: Dynamic social impact in electronic groups. *Journal of Personality and Social Psychology, 70*, 1218–1230.

Lauritzen, S.L. (1996). *Graphical models.* Oxford, UK: Oxford University Press.

Lauritzen, S.L., & Spiegelhalter, D.J. (1988). Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society, Series B, 50*, 157–224.

Leenders, R.Th.A.J. (1997). Longitudinal behavior of network structure and actor attributes: Modeling interdependence of contagion and selection. In P. Doreian & F.N. Stokman (Eds.), *Evolution of social networks* (pp. 165-184). Amsterdam: Gordon & Breach.

Leifer, E. (1988). Interaction preludes to role-setting: Exploratory local action. *American Sociological Review, 53*, 865–878.

Markovsky, B., & Chaffee, M. (1995). Social identity and solidarity. In B. Markovsky, K. Heimer, & J. O'Brien (Eds.), *Advance in group processes, Vol. 12* (pp. 249–270). Greenwich, CT: JAI Press.

Markovsky, B., & Lawler, E.J. (1994). A new theory of group solidarity. In B.Markovsky, K.Heimer, & J. O'Brien (Eds.), *Advance in group processes, Vol. 11* (pp. 113-137). Greenwich, CT: JAI Press.

Marsden, P.V., & Friedkin, N.E. (1994). Network studies of social influence. In S. Wasserman & J. Galaskiewicz (Eds.), *Advances in social network analysis* (pp. 3–25). Thousand Oaks, CA: Sage.

Moscovici, S. (1985). Social influence and conformity. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology: Volume II Special fields and applications.* (pp. 347–412). New York, NY: Random House.
Moscovici, S. & Doise, W. (1994). *Conflict and consensus: A general theory of collective decisions* (W.D. Halls, Trans.). London: Sage.
Pattison, P. (1994). Social cognition in context. In S. Wasserman & J. Galaskiewicz (Eds.), *Advances in social network analysis* (pp. 79–109). Thousand Oaks, CA: Sage.
Pattison, P., & Robins, G.L. (2000, April). *Neighbourhood-based models for social networks: Beyond Markovian neighbourhoods.* Paper presented at International Sunbelt Social Network Conference.
Pattison, P. & Wasserman, S. (1999). Logit models and logistic regressions for social networks: II. Multivariate relations. *British Journal of Mathematical and Statistical Psychology, 52,* 169–193.
Prendergast, J.F., Gange, S.J., Newton, M.A., Lindstrom, M.J., Palta, M., & Fisher, M.R. (1996). A survey of methods for analyzing clustered binary response data. *International Statistical Review, 64,* 89–118.
Robins, G.L. (1998). *Personal attributes in interpersonal contexts: Statistical models for individual characteristics and social relationships.* Unpublished doctoral dissertation, University of Melbourne, Department of Psychology.
Robins, G.L., Elliott, P., & Pattison, P. (2001). *Network models for social selection processes. Social Networks, 23,* 1–30.
Robins, G.L., Pattison, P., & Wasserman, S. (1999). Logit models and logistic regressions for social networks: III. Valued relations. *Psychometrika, 64,* 371–394.
Schein, E.H. (1985). *Organizational culture and leadership.* San Francisco, CA: Jossey-Bass.
Sherif, M. (1964). *The psychology of group norms.* New York, NY: Harper & Row. (Original work published 1936)
Sherif, M., Harvey, O.J., White, B.J., Hood, W.R., & Sherif, C.W. (1988). *Intergroup conflict and cooperation: The robber's cave experiment.* Norman, OK: University of Oklahoma. (Original work published 1961)
Stokman, F.N., & Zeggelink, E.P.H. (1996). Self organizing friendship networks. In W.B.G. Liebrand & D.M. Messick (Eds.), *Frontiers in social dilemmas research* (pp. 385-418). Berlin, Heidelberg: Springer-Verlag.
Strauss, D., & Ikeda, M. (1990). Pseudolikelihood estimation for social networks. *Journal of the American Statistical Association, 85,* 204–212.
Van de Bunt, G.G., Van Duijn, M.A.J., & Snijders, T.A.B. (1999). Friendship networks through time: An actor-oriented dynamic statistical network model. *Computational & Mathematical Organization Theory, 5,* 167–192.
Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* Cambridge, UK: Cambridge University Press.
Wasserman, S., & Pattison, P. (1996). Logit models and logistic regressions for social networks: I. An introduction to Markov graphs and p^*. *Psychometrika, 61,* 401–425.
Weick, K.E., & Roberts, K.H. (1993). Collective mind in organizations: Heedful interrelating on flight decks. *Administrative Science Quarterly, 38,* 357–381.
Wermuth, N., & Lauritzen, S.L. (1990). On substantive research hypotheses, conditional independence graphs and graphical chain models. *Journal of the Royal Statistical Society, Series B, 52,* 21–50.
Whittaker, J. (1990). *Graphical models in applied multivariate statistics.* Chichester, UK: John Wiley & Sons.
Winsborough, H.H., Quarantelli, E.L., & Yutzky, D. (1963). The similarity of connected observations. *American Sociological Review, 28,* 977–983.

Manuscript received 19 JAN 1999
Final version received 3 JAN 2000