THE LINK BETWEEN LOGLINEAR AND EXPONENTIAL RANDOM GRAPH MODELS FOR NETWORKS

Laura M. Koehly

Texas A&M University

Steven M. Goodreau

University of Washington

Martina Morris

University of Washington

ABSTRACT

Much progress has been made on the development of statistical methods for network analysis in the past ten years, building on the general class of exponential random graph (ERG) network models first introduced by Holland and Leinhardt (1981). Recent examples include "*p**" models (Wasserman and Pattison, 1996), and actor-oriented models (Snijders, 2001). For empirical application, ERG models currently require the equivalent of a network census: data on all dyads within the network. They can not be applied to sampled network data, the type increasingly collected in local (egocentric) network sample surveys. Conditional loglinear models have been adapted for analyzing such local network data (Marsden, 1981; Morris 1993). We show that these conditional loglinear models are related to the ERG model, though, somewhat surprisingly, not via the *a priori* blockmodels . Under certain conditions the two models are related via Bayes' rule. They do not yield equivalent predicted values except when fully saturated, but in practice, the differences are unlikely to be large. The alternate conditioning in the two models sheds light on the relationship between local and complete network data, and the role that models can play in bridging the gap between them.

1. INTRODUCTION

For many years, the methodology for model-based network analysis has developed along two distinct paths, one driven by a search for pragmatic approaches to network data collection, the other by developments in statistical theory for dependent data. The pragmatic approach assumed that for many populations of interest to social scientists, a network census – data on every node and every link in the network of interest – would be impossible to implement. These researchers sought methods that would enable the collection and analysis of sampled network data. This led to the development of the egocentric or local network survey design: a sample of the nodes (egos), with a name generator in the questionnaire to obtain a roster of their partners (alters), and name interpreters to collect information on these partners. No attempt is made to identify or enroll the alters. This approach makes local network data collection relatively cheap, easy to implement, and less intrusive than complete network data collection. Early examples include the Northern California Communities Study (Fischer 1982) and the core discussion partners network module used in the General Social Survey (Burt 1984).

Analytic strategies for local network data take one of two forms. The network information can be captured in node-specific summary measures, e.g. network size, heterogeneity, density, or mutuality, and the measures then treated as either response variables or covariates in a traditional linear model (e.g., Fischer 1982, Marsden 1987). Alternatively, the relational data (the tie itself) may be modeled explicitly as a function of the nodal attributes (Marsden 1981; Burt 1990). Typically, this type of analysis is conducted by forming a “mixing matrix” from the tie data – a contingency table of the ties that cross tabulates the attributes of the respondent (ego) by the attributes of their alter – and using a loglinear model (specifically, a generalized linear model with a log link and Poisson errors) to capture the degree of homophily in the matrix (Pagnini and Morgan 1990; Mare 1991; Morris 1991; Raymo and Xie 2000). The underlying approach grew naturally from earlier work on social mobility, in which the occupations of fathers and sons were compared (Goodman 1965). In both applications, a sampled dyad (father-son, ego-alter) is the unit of analysis, the diagonal of the matrix has a special meaning, and a member of the pair may contribute multiple dyads to the sample. The latter fact may induce some dependence among the observations, which can be handled by appropriate estimation methods (Yamaguchi 2003). In addition, it is necessary to have accurate estimates of the population subgroup sizes in order to distinguish between selection, activity level, and population composition effects (Morris 1993). Overall, this local network approach has proven quite practical, and data have been collected on a wide range of topics (e.g., Granovetter 1973; Laumann and Knoke 1987; Wellman and Wortley 1990; Massey 1990; Burt 1992; Morris 2004).

An important feature of this approach is that the observed data, and therefore the models, represent the presence of a tie, but not the absence. We therefore refer to these models as conditional loglinear models (CLLs).

The more theoretical line of statistical network analysis has sought to develop a comprehensive framework for investigating the structure of a network, assuming complete network data are available. This approach has given rise to a sequence of models for the probability of a tie, from the p_1 models first proposed in the late 1970s by Holland and Leinhardt, to the $p*$ models developed during the 1990s by Wasserman and Pattison. The key statistical advances underlying the development and application of these models has been the definition of the general class of exponential random graph models(ERG), and the development of estimation techniques that can be used with dependent data.

The earliest modeling began by representing modest forms of dependence between the network ties: reciprocity and transitivity. Holland and Leinhardt (1970, 1981) made impressive progress exploring the effects of these forms of dependence on network structure given the limited computational methods available at the time. Fienberg and Wasserman (1981) proposed additional terms to account for *a priori* blocks of network actors. *A priori* block models collapse actors into attribute classes in much the same way as the mixing matrices for local network data, but keep track of asymmetric and mutual ties as well as non-ties. Frank and Strauss (1986) took the next logical step, proposing the Markov random graph as a model for local dyad dependence, and identifying it explicitly as a special case of the general ERG (Besag 1977). The dependence is called Markovian because it extends only one step out from each dyad: two dyads are dependent if they share a node, and independent otherwise. In the last few years a wide range of more general forms of dyadic dependence have been explored with these models (Wasserman and Pattison 1996, Pattison and Wasserman 1999, Robins et al. 1999; Snijders 2001; Hoff, Raftery and Handcock 2003). The ERG class turns out to be very flexible, capable of representing such things as propensities for cycles, small world patterns and latent groups. They provide, for the first time, statistical models for generalized spatial and temporal dependence in networks.

Estimation of these models remains somewhat of a challenge. Following the work of Besag (1975, 1977), Strauss and Ikeda (1990) proposed using maximum pseudo-likelihood (MPL). This turned the problem into a simple logistic regression, making it possible to estimate these models using standard statistical software. Virtually all applications of ERG models to date have used MPL. While the MPL estimates are identical to the ML estimates when the dyads are independent, research suggest they may perform quite poorly under dyadic dependence

(Handcock 2003). True ML estimation requires Markov-Chain Monte Carlo methods (Geyer and Thompson 1992, Gilks et al. 1996). For various reasons, MCMC-based ML estimation continues to be difficult to implement (Handcock 2003), but progress is being made. Soon, the main limitation will be the availability of data, and that is a big change.

To date, these two statistical modeling frameworks -- for complete and local network data -- have been developed in isolation. Yet both are based on generalized linear models for exponential family distributions. It would be quite useful if the results from the two types of analysis were directly comparable. The relationships, however, are less direct than one might hope. In the limited case of saturated tie-independent models, the two models are equivalent, and their fitted values are related via Bayes' rule**.** For non-saturated tie-independent models (more often of interest to social scientists), they are not equivalent, but their fitted values are likely to be very similar in practice. The conditions under which equivalence holds, and the reasons for similarity when it does not, help to illuminate the similarities and differences between the two models. This understanding allows us to interpret the body of previous work on local network partnership data within the newer, more flexible ERG framework, bridging the gap between local and complete network data and making the first steps towards a coherent statistical framework for modeling networks.

In this paper, we explicate the relationships between ERG models and conditional loglinear mixing models for network data. We focus on the subclass of models that assume tie probabilities are independent, as this is the only subclass for which equivalence exists. The central findings are illustrated using data on a network of school friendships from the National Longitudinal Study of Adolescent Health (Add Health).

2. TERMINOLOGY AND NOTATION

Social network data include a set of social entities, generally referred to as *actors* or *nodes*, and a set of relational measurements, also known as *ties, links, arcs, lines, edges* or *partnerships,* that exist between pairs of those actors on some social relation. In the example we will be using below, actors are individual people and the relation is friendship. The number of actors in the network will be denoted by n; the fixed set of network actors will be represented by N, where $N = \{1,2,3,\ldots, n\}$. One generally measures some attribute variables on the actors such as sex, ethnic origin, religious affiliation, geographic location, or age. For simplicity, we shall assume for this paper that actors are coded according to a single nominal attribute that can take on K values; the results are easily generalizable to multiple and ordinal attributes. We define the sets C_k for $k = 1$ *to* K, whose elements are all those nodes possessing the k^{th} value of the attribute. (The ordering of attribute values is arbitrary for nominal attributes). The number of actors with attribute k is denoted n_k, so that $n = \sum_{k=1}^{K} n_k$.

Pairs of actors, whether or not they share a relational tie, are referred to as *dyads*. The value of the tie between two actors is denoted by X; for specific actors i,j the random variable is denoted X_{ij}. In the current discussion, we will assume that the tie relation is dichotomous, such that $X_{ij} = 1$ if actors i and j share a tie and $X_{ij} = 0$ if they do not.

Relations may be either *directed* or *nondirected*. The relation is *nondirected* if a tie is either present or absent between each actor pair ($X_{ij} = X_{ji}$ for all i,j pairs). A *directed* relation consists of measurements where the orientation of the ties between actors is meaningful. In this case X_{ij} need not equal X_{ji}. An example of a nondirected relationship would be “has sex with”; a directed relationship would be “sells drugs to”. With local network data there is often a directionality implied by the study design (separate from the relationship itself), such that respondents may be viewed as “sending” the relationship and their nominated partners as “receiving”. We will use an example based on directed data here, but the approach is easily extended to undirected relationships. Because of this generalizability we will use the word *tie* to describe a social relationship, since *arc* is generally restricted to directed relations, and *edge* or *line* to undirected relations, while *tie* remains more general (e.g. Wasserman and Faust 1994).

Figure 1 depicts three common forms of representation for network data, using a hypothetical directed network containing ten actors identified by location (urban/rural). The first representation is a *graph*, ***G***, consisting of a set of nodes joined by lines or arcs. The actors in *N* are the nodes in the graph. Relational ties are represented graphically by connecting two nodes with a directed line, $i \rightarrow j$, indicating that actor *i* initiates a relationship towards, or *chooses,* actor *j*. (Nondirected relationships are typically represented by a nondirected line, *i*—*j*.) The network can also be represented in a two-dimensional array called a *sociomatrix* or *adjacency matrix*, denoted by ***X*** with elements X_{ij}. If self-relations are disallowed, the main diagonal of the sociomatrix is ignored. For a nondirected relation one may assume that $X_{ji} = X_{ij}$ for all (*i*,*j*) pairs, or ignore the lower triangle, restricting analysis to those X_{ij} for which $i < j$.

The third form of representation begins by collapsing the sociomatrix into a *mixing matrix* or *contact matrix.* Rows and columns of the sociomatrix are aggregated within attribute classes, resulting in a smaller matrix in which cell entries t_{ab} indicate the total number of ties in a network among actor pairs with attributes *a* and *b*:

$$t_{ab1} = \sum_{i \in a} \sum_{j \in b} X_{ij} \qquad [1]$$

We refer to the two attribute dimensions as A and B, with $a,b = \{1 \dots k\}$ as attribute classes; since we assume a directed graph, dimensions A and B refer to the attribute classes of the sender (i) and receiver (j) of the relational tie, respectively. The third subscript Y represents tie value $y = \{0,1\}$, equaling 1 in the contact matrix since these are counts of ties present.[1] Information about the specific actors involved in the relationships is ignored in this matrix. As with the sociomatrix, the contact matrix is square for a directed relationship and triangular for a nondirected one. Square contact matrices may also be used for nondirected bipartite data, when the population can be divided into two classes with all partnerships between classes (e.g. a mixing matrix by race for heterosexual relationships, with the races for males and for females along the two margins).

The contact matrix ignores information about the absence of ties. This information can be represented in another matrix, which we will call the "non-contact" matrix, in which $y = 0$. The three dimensions imply three sets of marginals; we follow the standard notation representing the margins with a dot symbol in the relevant subscript. The total number of ties is thus represented by $t_{\bullet\bullet 1}$. The marginal table $t_{ab\bullet}$ represents the total number of dyads between two actors with a given attribute combination. In this marginal table A and B are always independent since $t_{ab\bullet}$ represents the number of possible a,b dyads and is simply the product of n_a and n_b for all a,b.[2] This constraint turns out to have important implications both for the patterns of mixing that occur in practice and for the model.

[1] Since much of the literature on loglinear models for partnership data examines only the contact matrix, this third subscript is often left implied given it always equals 1. We include it here for clarity in comparing with later models.

[2] This is true for all bipartite graphs (e.g. when modeling racial marriage patterns in heterosexuals), while for non-bipartite graphs it is only exactly true in the case where actors are allowed to share a tie with themselves. Otherwise, the number of homophilous dyads (those on the main diagonal of the contact matrix) in a group with n actors equals n^2-n rather than n^2. As n gets large, however, this difference becomes negligible. Since modeling as if on-diagonal relationships were allowed simplifies the analysis considerably and since its effects in large populations are small, we will do so throughout the paper.

In addition to the observed cell counts t_{aby}, we define the cell counts fit by a particular model as m_{aby} and the probability of an actor pair falling into a cell for a given model as π_{aby} ($=m_{aby}/m_{\bullet\bullet\bullet}$). The literature on conditional loglinear models is often concerned with the probability of a *tie* falling into a given cell (i.e. $m_{ab1}/m_{\bullet\bullet 1}$); we follow the tradition of this literature and denote this using two subscripts, π_{ab} (as distinct from $\pi_{ab1} = m_{ab1}/m_{\bullet\bullet\bullet}$).

3. MODELING THE GRAPH

The modeling approaches we compare are probabilistic, treating the X_{ij} ties as random variables with realizations x_{ij}. For dichotomous relations, the expected value of X_{ij} is thus equal to $P(X_{ij} = 1)$. A graph in which every potential partnership is independent and has an identical expected value is known as a *Bernoulli graph*. A graph obeys *conditional tie independence* (CTI) if its tie probabilities do not depend on one another given the attributes of the nodes; this model is sometimes referred to as an independence model in the network literature, dropping the "conditional" since complete independence models are rarely of interest. For directed relationships, *dyadic independence (*or more correctly, *conditional dyadic independence)* refers to a model in which tie probabilities are dependent on the value of the tie between the same two actors in the opposite direction, but not on other ties given the actor attributes[3]. Otherwise, ties are said to be *conditionally dependent*, analogously shortened to *dependent* in common usage. We will retain the longer but more accurate terms here for clarity. See Frank (1988) for a full discussion.

Nodes *i* and *j* are said to be *homogeneous* if they can be interchanged without affecting the probability of the graph. All nodes are homogeneous in a Bernoulli graph, while the

[3] Note that CTI and conditional dyadic independence are not identical; dyadic independence allows for a tie to depend on the state of the opposite tie between the same two actors, while CTI does not.

definition of CTI implies that nodes with the same attributes are homogeneous. Homogeneity constraints allow for a more parsimonious representation, but they represent substantive hypotheses that should be considered part of the model.

For the remainder of the paper we assume CTI. We will also assume that the order of the graph (i.e. the size of the population or the number of nodes) and its overall attribute composition are fixed, and we will leave these conditions out of our probability statements for simplicity. In the discussion, we will review the ability of different models to relax these assumptions.

3.1 *Conditional loglinear models for locally sampled networks*

In the CLL context each tie is a Bernoulli trial, the probability of which depends only on the attributes of the two actors involved. The cell counts t_{ab1} are the sum of these trials; since we have assumed a fixed population and attribute composition, these cell counts have a Poisson (if the total number of ties $t_{\bullet\bullet 1}$ is not fixed) or multinomial distribution (if it is).

The saturated CLL model can be expressed as:

$$\log \pi_{ab} = \lambda + \lambda_a^A + \lambda_b^B + \lambda_{ab}^{AB} \qquad [2]$$

where $\pi_{ab} = m_{ab1} / m_{\bullet\bullet 1}$. The first term represents a reference level for tie formation, the next two terms are main effects for the relative levels of tie formation for each group, and the last is an interaction effect for specific attribute pairings. Interaction effects can be used to saturate the model, or they can be constrained to index groups of cells. If the interaction effects are set to zero, one obtains the *marginal effects* model. The remaining parameter values are adjusted accordingly, and the odds ratios (the cross product for any four cells that form a rectangle) for the fitted cell probabilities must satisfy:

$$\frac{\pi_{a_1 b_1 1}\pi_{a_2 b_2 1}}{\pi_{a_1 b_2 1}\pi_{a_2 b_1 1}} = 1, \quad \forall a_1, a_2, b_1, b_2 \qquad [3]$$

This model fits the margins perfectly, but not necessarily the individual cell values.

Between the marginal and saturated models lie a range of non-saturated interaction models, which involve grouping cells into categories representing layers of a given effect.[4] A simple one-parameter non-saturated interaction model is uniform homophily, which distinguishes between on- and off-diagonal cells. The parameter measures the general strength of assortative mixing, and this model can be tested for goodness-of-fit. One can also specify differential homophily factors for each diagonal cell (the “quasi-independence” model), linear or non-parametric distance off the diagonal for mixing by ordinal factors like age (the “diagonals parameter” model), and single-cell interaction terms (cf. Goodman 1984 and Morris 1991 for examples). Non-saturated interaction parameters can be thought of as a generalized marginal term, in the sense that the cells sharing a level of the interaction term will have their sum fit when the term is in the model. The interaction term levels I_{ab} can be represented as a *model matrix* (analogous to a design matrix from experimental studies), which helps to clarify their relationship to standard marginal models. For example, Table 1 contains the model matrix for a uniform homophily parameter with first level constraints in a four-value attribute, along with the model matrices for the marginal effects parameters. Together these would yield the model:

[4] With the exception of Goodman’s work, there is comparatively little statistical literature on non-saturated interaction models, despite widespread use of such models in the social sciences.

$$\log \pi_{ab} = \lambda + \lambda_a^A + \lambda_b^B + \lambda_{a,b}^{HOM} \qquad \begin{cases} \lambda_{a,b}^{HOM} = \lambda^{HOM}, & a = b \\ \lambda_{a,b}^{HOM} = 0, & a \neq b \end{cases} \qquad [4]$$

The two most common identification constraints are symmetric (or ANOVA) constraints, and first-level constraints. The latter set the first level or category effects for each variable and their interactions equal to zero, thus acting as a baseline for interpretation of the parameters associated with the remaining categories (Agresti 2002).

If the dependence induced by actors contributing multiple partnerships to the data is ignorable, the model can be fit using a generalized linear model with a log link and Poisson errors. Otherwise, the model can be fit using generalized estimating equations (GEE, Liang and Zeger 1986, Yamaguchi 2003). Under the assumption of tie-independence, parameter values for the fully saturated model can be stated as a function of the fitted cell probabilities:

$$\begin{aligned} \lambda &= \log(\pi_{111}) \\ \lambda_a^A &= \log(\pi_{a11} / \pi_{111}) \\ \lambda_b^B &= \log(\pi_{1b1} / \pi_{111}) \qquad [5] \\ \lambda_{ab}^{AB} &= \log\left(\frac{\pi_{ab1}\pi_{111}}{\pi_{a11}\pi_{1b1}} \right) \end{aligned}$$

A shorthand bracket notation is often used to identify the model terms: a single variable in brackets indicates that all the levels of that variable are included in model, two or more variables in brackets implies a full set of interaction terms for those variables, as well as all lower-order terms. The model in Equation [2] would be abbreviated as [AB], since this signifies a full set of AB interaction terms as well as marginal A terms and B terms.

To compare the CLLs to ERG models below, it is helpful to consider the unconditional form of this loglinear model (ULL). The ULL does not condition on the presence of a tie, instead it considers all dyads and treats the presence or absence of a tie as a third dimension with two levels (Y={0,1}). The saturated ULL [ABY] is represented as:

$$\log \pi_{aby} = \gamma + \gamma_a^A + \gamma_b^B + \gamma_y^Y + \gamma_{ab}^{AB} + \gamma_{ay}^{AY} + \gamma_{by}^{BY} + \gamma_{aby}^{ABY} \qquad [6]$$

The ULL class has marginal effects models and non-saturated interaction models analogous to the CLLs, as well as the same methods for fitting parameters. The ULLs are similar to the *a priori* block models introduced by Fienberg and Wasserman (1981), but they are not the same. We consider this relationship in more depth below.

3.2 *Random Graph Models for Complete Networks*

Exponential random graph (ERG) models with CTI reverse the conditioning of CLL, modeling the probability that actors share a tie given their attributes. ERG models use both the tie matrix and the non-tie matrix, treating the tie dimension as an outcome variable and modeling the log-odds that it is present. Population size and attribute composition are exogenously given in this model so the total number of dyads of each attribute combination ($t_{ab\bullet}$ for all *a*,*b*) is fixed.

The ERG model represents the probability function of the random graph ***G*,** defined by the sociomatrix ***X*,** as a linear combination of network statistics:

$$P(\mathbf{X} = \mathbf{x}) \;=\; c^{-1} \exp\left\{\theta' \mathbf{z}(\mathbf{x})\right\} \qquad [7]$$

(Besag 1974). The vector **z**(**x**) represents a set of network configurations, while the θ parameters represent the unknown weights of the linear function of network properties. The normalizing constant c is needed to ensure a proper probability distribution. Any dyad-based measure from the network may be included in **z**(**x**), although typically sums or sums of products of X_{ij} are used.[5]

This formulation is very general, and includes many network models proposed in the literature as special cases. The earliest of these, the p_1 models of Holland & Leinhardt (1981) included a θ parameter for overall partnership formation, actor-specific marginal parameters for each actor's expansiveness (number of ties they send) and attractiveness (number of ties they receive) as well as a parameter for mutuality (the tendency for a tie from actor i to j to be reciprocated as a tie from j to i). Some of the other models have already been noted above.

One of these deserves closer attention, because it is easy to confuse with the ULL model. Fienberg & Wasserman's (1981) *a priori* block model introduces exogenous attributes into the p_1 model by collapsing across individual indices into blocks defined by the attribute levels of each node. Within the cells defined by these blocks, the counts of the four paired tie values within dyads (0,0; 0,1; 1,0; 1,1) become the focus of analysis. Iacobucci (1994, pp. 605-674) provides a detailed review of the subsequent application of these models. In contrast to the ULLs above, *a priori* blockmodels preserve the counts of directed asymmetric and mutual ties within and between blocks. Because the counting strategy is different the *a priori* blockmodel parameters are not comparable to either the CLL or ULL model parameters. The method of estimation is similar, however, since the model form is loglinear.

[5] Examples include nodal degrees, the number of within-group ties (analogous to uniform homophily), or the number of transitive triads ($X_{ij} = X_{jk} = X_{ki} = 1$).

The only subclass of ERG models for which direct comparison can be made to the CLL specifications is CTI models. These models can have parameters for exogeneous attributes, but not for mutuality or any form of dyadic dependence. The saturated CTI model with homogeneity constraints is thus:

$$P\left(\mathrm{X}=\mathrm{x}\right)=c^{-1}\exp\left\{\theta z+\sum_{a=1}^{K}\theta_a^A z_a^A+\sum_{b=1}^{K}\theta_b^B z_b^B+\sum_{a=1}^{K}\sum_{b=1}^{K}\theta_{ab}^{AB}z_{ab}^{AB}\right\} \qquad [8]$$

where z = the total number of ties in the network, z_a^A = the number of ties initiated by actors in attribute class C_a, z_b^B = the number of ties received by actors in attribute class C_b, and z_{ab}^{AB} = the number of ties initiated by actors in C_a and received by actors in C_b . The θ, θ_a^A , θ_b^B and θ_{ab}^{AB} are the coefficient on each term. Reframed in logit form for an individual tie, Eq. [8] reduces to:

$$\operatorname{logit} P\left(X_{ij}=1\middle|i\in C_a, j\in C_b\right)=\theta+\theta_a^A+\theta_b^B+\theta_{ab}^{AB} \qquad [9]$$

Under CTI, unbiased estimates for the θ's can be obtained from logistic regression with the observed tie values as the outcome variable and the change in network statistics when that tie value is toggled (the “change statistic” δ_{ij}) as the predictors (Strauss and Ikeda 1990). This is a generalized linear model with a logit link function and binomial errors.

This model can also be abbreviated as [AB], indicating that the right-hand side of the equation contains a similar set of terms as in the saturated CLL model. The left-hand side of the equation is different, however. A marginal effects model in this context also involves setting the AB interaction terms to 0. This is commonly referred to as a model of independence for A and

B, but it is not the same as the independence model for the CLL. While the logit is now an additive function of row and column effects alone, *A* and *B* are not independent conditional on Y. The model instead implies:

$$\frac{\pi_{a_1 b_1 1}\pi_{a_2 b_2 1}}{\pi_{a_1 b_2 1}\pi_{a_2 b_1 1}} = \frac{\pi_{a_1 b_1 0}\pi_{a_2 b_2 0}}{\pi_{a_1 b_2 0}\pi_{a_2 b_1 0}} \qquad [10]$$

We draw out the implications below.

3.3 *Linking ERG models and conditional loglinear models*

Conditional loglinear models predict $P(i \in C_a, j \in C_b \mid X_{ij} = 1)$, while ERG models with CTI predict $P(X_{ij} = 1 \mid i \in C_a, j \in C_b)$. These are related by Bayes' rule:

$$P(i \in C_a, j \in C_b \mid X_{ij} = 1) = \frac{P(X_{ij} = 1 \mid i \in C_a, j \in C_b)\, P(i \in C_a, j \in C_b)}{P(X_{ij} = 1)} \qquad [11]$$

The two conditional probabilities are linked by the two marginal probabilities for ties and attributes: P(X_{ij} = 1) is the fraction of all dyads in the network that have a tie, and $P(i \in C_a, j \in C_b)$ is the joint distribution of nodal attributes for all dyads. Bayes' rule thus provides a simple explicit expression for transforming the predicted conditional probabilities from one model to that of the other.[6]

Since both CLLs and ERG models are generalized linear models and Bayes' rule provides a link between them, it would be natural to expect that models in one class would have an equivalent representation in the other, in the sense of a model yielding the same fitted cell probabilities. However, due to the nature of the conditioning in each model, the only equivalent models are fully saturated models. Non-saturated models from each class that appear comparable in terms of predictors in fact yield different outcomes. Intuitively, this is because non-saturated ERG models use information from the non-tie layer to fit values in the tie layer, while the CLL ignores information in the non-tie layer.

The ULL provides an explicit bridge for comparing CTI models in the ERGM and CLL frameworks. All ERG models with CTI correspond to a 3-way ULL that contains the following terms and no other (Agresti 2002, p. 332)

- a full set of AB interaction terms;
- a Y marginal term;
- every term in the ERG model;
- every term in the ERG model crossed by Y.

[6] The distinction in conditioning is similar to those observed in a series of papers by Robins, Pattison, and Elliott (2001a, 2001b) that distinguish between social selection and social influence. A social selection model assumes that individuals select partners based on attributes, analogous to the process underlying the CTI-based ERGMs. Social influence models assume that network structure can influence individual characteristics such as beliefs. The probability of the attribute is conditional on the tie, as with the CLL. A major difference lies in the fact that most of the literature using CLLs for mixing matrices, although modeling the probability of the attribute conditional on the tie, do not actually purport a causal link from tie to attribute but rather leave the direction of causality unspecified.

The AB interaction terms in the ULL ensure that the cells in the $t_{ab\bullet}$ marginal matrix are fit exactly. These establish the population size, marginal attribute composition, and the numbers of dyads (not ties) among attribute groups. Any equivalent ULL must have this [AB] term in the model, because population size and composition are exogenous to the ERG model.

Each CLL corresponds to a ULL containing the following terms:

- a Y marginal term;
- every term in the CLL model;
- every term in the CLL model crossed by Y.

Here the *Y* term sets the number of ties, so that $m_{\bullet\bullet 1}$ can be fit exactly to $m_{\bullet\bullet 1}$ in the ULL model. Each of the terms crossed by *Y* allows the CLL terms to be represented in the ULL tie layer independently of the non-tie layer. That independence means that the fitted cell probabilities in the ULL ($m_{ab1}/m_{\bullet\bullet\bullet}$) can be divided by the probability of a tie ($m_{\bullet\bullet 1}/m_{\bullet\bullet\bullet}$) fit by the Y term to yield the CLL cell probabilities ($m_{ab1}/m_{\bullet\bullet 1}$).

Table 2 lays out the set of equivalencies among ERG models, ULLs, and CLLs. We use the symbol U_{AB} as a general symbol representing any set of non-saturated interaction terms between *a* and *b*. Note that non-saturated ERG models have no corresponding CLL. This follows deductively from the two sets of equivalence rules above.[7]

To make these differences clearer, we highlight the most familiar model -- marginal effects -- in the two frameworks. Table 2 makes explicit how these differ. While this is the model that we commonly think of as implying that A and B are independent, independence

[7] Imagine that there exists some ERG model with an equivalent CLL. The ULL that is equivalent to this ERGM must contain an [AB] interaction term. If the ULL contains [AB] then its CLL equivalent must also contain [AB]. If the CLL contains [AB] then the ULL must contain [ABY], which means it is fully saturated.

clearly means different things in the two models. For CLL, it means that A and B are independent conditional on Y. For ERG models, independence means "no 3-way association"; all three variables are pairwise dependent, but each pair is conditionally independent given the third. This does not mean that A and B are independent in either layer of Y; instead, the pattern of dependence is the same in each layer. The difference is also evident when comparing the fitting constraints, Eq. [3] for the CLL, and Eq. [10] for the ERG model.

The model of "no 3-way association" in the loglinear setting is one of the more difficult to interpret in practice, yet it corresponds to the basic marginal effects ERG model. There is no simpler definition in the ERG context because of the implicit constraint that A and B are independent in the marginal matrix of all dyads, $m_{ab\bullet}$. Since the tie and non-tie matrices must sum to this marginal matrix, the cell values in one layer determine the other when the totals are fixed. A and B can only exhibit conditional independence in both layers under a narrow range of conditions: if either the sender or the receiver attribute subgroups are homogeneous with respect to tie formation. (When both groups are, or one group is and ties are undirected, this degenerates into the Bernoulli model). In essence, the additional implicit constraint $m_{ab\bullet}$ creates an inverse form of Simpson's paradox; two attributes are independent in the marginal table, but when stratified by a third variable (here, tie value), they are not independent in each stratum.

In practice, however, the fitted values from the two marginal effects models are likely to be similar. Social networks for populations of reasonable size are generally quite sparse, because the number of ties in a population generally scales roughly with the population size, while the number of dyads varies with the square of population size. If almost all dyads have $Y = 0$, then we can assume that in any sparse dataset $t_{ab0} \cong t_{ab\bullet}$ for all a,b. Thus

$$\frac{t_{a_1 b_1 0} t_{a_2 b_2 0}}{t_{a_1 b_2 0} t_{a_2 b_1 0}} \approx \frac{t_{a_1 b_1 \square} t_{a_2 b_2 \square}}{t_{a_1 b_2 \square} t_{a_2 b_1 \square}} \approx 1. \qquad [12]$$

This means the right hand side of Eq. [10] is approximately equal to 1 for sparse matrices, reducing it to Eq. [3], implying that the two models will yield approximately equal results. Bayes' rule can be used to transform the fitted cell probabilities from one model to the other to determine the magnitude of the difference. In our experience, sparse matrices of at least a few hundred people yield marginal models in which cell counts differ by at most a tenth of a partnership. For non-sparse matrices from small settings such as an office or classroom, the differences will be small as long as the sizes and activity levels of the different attribute classes are roughly equal. This latter set of conditions is seen in the example below.

4. EXAMPLE: THE ADD HEALTH STUDY

We use the friendship nomination data from the first wave of the National Longitudinal Study of Adolescent Health (Add Health) to demonstrate the results above. Add Health is a nationally representative study of students in grades 7 through 12, and the first wave was conducted in 1994-1995. The study was school-based, and students were provided with a roster of all students in the school and asked to select up to five close male friends and five close female friends. Complete details of this and subsequent waves of the study can be found in Resnick et al. (1997) and Udry and Bearman (1998) and at *http://www.cpc.unc.edu/projects/addhealth*.

We use friendship data from one school comprising 71 students divided into six grades (with 15, 13, 16, 10, 13 and 4 students in grades 7 through 12, respectively). The ties are directional since it is possible person A could name B as a friend without B nominating A. The limit on nominations means that the data are not complete, but we will assume for convenience that a lack of nomination in these data implies the absence of a friendship.

Figure 2 provides a graph of the data, while Table 3 shows the corresponding contact and non-contact matrices. We begin by fitting a CLL and an ERG model with main effects only and the ULL that corresponds to each. A quick glance at Figure 2 makes it clear that there is a strong preference for students of all grades to nominate friends in their own grade; we thus also run the CLL and ERG models for main effects with uniform homophily. In each case, first-level identification constraints were used, and both were fit using the *glm* macro in R (Ihaka and Gentleman 1996).

The parameter estimates for the marginal effects models are shown in Table 4. These allow for a comparison of the CLL and ERG model to their respective ULL parameterizations. In the case of the ERG model, the ULL parameter values in the first column of Table 4 are those that fit the $t_{ab\bullet}$ cells exactly; there are no corresponding parameters in the ERG model since the $t_{ab\bullet}$ values were conditioned upon in the construction of the model. The ULL parameters in the second column (those that represent the patterns in the $Y = 1$ layer) have values equal to the ERG model parameters. Note that all 25 ([a-1]*[b-1s]) of the AB interaction terms in the ULL list are very close to 0 and are described by summary statistics rather than enumerated; they are modeling the log of the ratio of odds ratios between the tie and non-tie matrix, and those ratios are all very close to 1 for the reason explained in the previous section (see Eq. [12]). Table 5 emphasizes their negligible effect on both fitted cell probabilities and values by providing summary counts of the order of magnitude in difference invoked by their inclusion or exclusion in model fit.

For the CLL marginal effects model, the corresponding ULL model does not have any [AB] interaction parameters. There are exactly twice as many parameters in the ULL parameterization as in the CLL model, since the ULL model is fitting both layers; the first column of the ULL values fits the appropriate independence model in the non-tie layer, while the second column then fits independence in the tie layer. With the first-level parameterization, each CLL parameter equals the sum of the two parameters in the ULL in the same row in Table 4.

Note the strong similarity between the ULL parameters for the two models, despite the fact that they are not identical. We can also see the similarity between the models by applying Bayes' rule to calculate the fitted probabilities of having a nodal attribute composition given the presence of a tie. Take the example of a 7th grade sender and 8th grade receiver; under the marginal effects ERG model, Eq. [11] would yield:

$$P\left(i \in C_7, j \in C_8 \mid X_{ij} = 1\right) = \frac{\dfrac{\exp\left(-2.895 - 0.102\right)}{1 + \exp\left(-2.895 - 0.102\right)} * \left(\dfrac{5 + 190}{305 + 4736}\right)}{\left(\dfrac{305}{305 + 4736}\right)} = 0.0304 \qquad [13]$$

The marginal effects CLL model gives this value directly as:

$$P\left(i \in C_7, j \in C_8 \mid X_{ij} = 1\right) = \frac{\exp\left(2.468 - 0.240\right)}{305} = 0.0304 \qquad [14]$$

Other combinations are compared in Table 6; despite being different models, the fitted cell probabilities obtained from the two are generally equal down to the fourth decimal place.

Figure 2 made clear the strong tendency for ties to be homophilous by grade. A likelihood ratio test confirms that adding a single uniform homophily parameter significantly improves the model fit in either modeling framework. Using the uniform homophily parameters accentuates the differenced between the fitted cell probabilities of the two frameworks, although the differences are still on the order of the third decimal place. Thus, even for relatively small, dense social networks, the practical differences between the two modeling frameworks are not very large.

5. DISCUSSION

Conditional loglinear models and ERG models are both generalized linear models based on the exponential family that can be used to represent attribute mixing in networks, but they condition on different aspects of the data. CLLs condition on the tie being present, and model the patterns in partner selection, while equivalent CTI ERG models condition on the attribute composition of the population, and model the distribution of ties and non-ties. This helps to clarify how one might choose between these models (and data collection strategies) based on context. For large populations in which people form ties with only a small fraction of possible partners, it seems reasonable to assume that the non-ties (or at least the great majority of them) are not explicitly chosen. In this case, the CLLs would be a reasonable approach for analysis. In small settings such as schools or offices or isolated populations, the patterns of non-ties (do not collaborate, do not get along, can not marry) may be as intentionally chosen as the ties. This is also the situation in which complete data are more feasible to collect, and here ERG models may be a better choice. Whatever the relative theoretical merits of each model, however, the similarity of the fitted values in practice suggests that there is little to be gained by selecting the model based on the form of conditioning. That leaves one free to choose on other grounds.

One of the other grounds to consider is the flexibility of the modeling framework. In this paper, attention has been limited to the restricted set of “comparable” models, so that the similarities and differences between the frameworks can be clearly identified. The set of comparable models share two assumptions: fixed population attribute composition and dyadic (tie) independence. In practical applications these assumptions may be a severe handicap. For example, when network models are used in a dynamic context, such as modeling the transmission of HIV through a network of partnerships, it is often desirable to allow for population composition changes (for example, to allow for group specific infection and mortality rates). The CLLs make it relatively easy to relax fixed attribute composition, and replace it with the much weaker assumption that population sizes and preferences are separable. This is because the CLL mixing parameters are specified in terms of odds ratios, which allows the margins to change independently of selection patterns, making it straightforward to integrate exogenously changing population sizes into a dynamically changing mixing matrix (Morris, 1991). Currently, it is not clear how this would be accomplished in the ERG modeling framework.

On the other hand, it is often important to be able to relax the dyadic independence assumption to be able to model such things as dependence among dyads. A good example (in the same HIV transmission context) is the rule of serial monogamy in sexual partnerships, which imposes a very strong form of dyadic dependence: the probability of a link between two nodes is zero if either node is already linked to another. Here ERG models have the clear advantage, as they can model forms of dyadic dependence explicitly (Wasserman and Pattison, 1996; Pattison and Wasserman, 1999). Because non-ties cannot be modeled explicitly in the CLL, there is no way to represent this form of dependence.

The ability to represent dyadic dependence is the reason that ERG models have attracted so much interest in contemporary network analysis. Dependence among social ties has always been the theoretical heart of network analysis – from balance theory and cognitive networks to kinship structure and role algebras. While most applications of ERG models have focused on locally connected subsets of the graph (Markov graphs), ERG models can incorporate a much wider range of interdependence, including global network properties such as connectivity, centrality and distance. Using Markov Chain Monte Carlo algorithms for estimation, ERG models also place the problem of inference for conditional dependence on a firm statistical footing. This has important practical implications, as it enables one to test whether dyadic independence, or limited forms of dependence, provide a reasonable fit to the data in particular contexts. When the answer to this question is yes, network structure may still be present (as in attribute mixing), but the data requirements for estimating such structural parameters are much simpler.

The difference between these two statistical frameworks for modeling networks currently poses a difficult choice for network analysts: flexibility to model population changes while constrained to dyadic independence, vs. flexibility to model dependence constrained by static population composition. This would be a sorry state of affairs if it were not such an improvement over the recent past. Statistical models for networks have been long in coming, and the choice now afforded, even if not ideal, is a temporary inconvenience. In short order, the constraints of the current frameworks will be eclipsed by developments in both network sampling and modeling. Much of the impetus for current work is coming from the need for network models in applied settings. Especially in the field of HIV/AIDS prevention, the value of a network perspective is tremendous, providing a unique and powerful perspective on the dynamics of transmission and the opportunities for prevention. The spillover effects of resources devoted to this aim are making it possible for basic network methods to develop at a remarkable pace. These include not only the local and complete data approaches already discussed, but various "partial network" sampling schemes, including various forms of snowball samples, link tracing, and adaptive sampling (Frank and Snijders 1994, Thompson and Frank 2000); such methods result in complex modeling issues but also hold great promise.

The theoretical perspective afforded by network analysis has been regarded by some as the heart of a true social theory – because it makes the relation, and the positions defined by relations, the unit of analysis. While we are still some years from having a viable statistical framework that can replace the linear regression model, it is now a matter of time. The links with the growing literature in evolutionary game theory and the general study of agent-based interacting dynamic systems (Bowles, 2003; Gintis, 2000; Macy, 2002; Majeski, 1999; Padgett, 1993; Parker, 2003; Watts, 2003), suggest that the social sciences are converging in a remarkable way to the empirical study of social relations and interaction.

Author Note

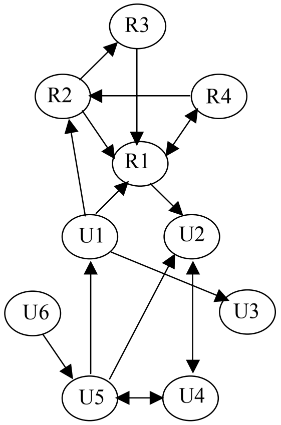
Laura M. Koehly, Department of Psychology, Texas A&M University; Steven M. Goodreau, Center for AIDS and STD and Center for Statistics and the Social Sciences, University of Washington; Martina Morris, Department of Sociology, Center for Statistics and the Social Sciences, University of Washington. The first two authors contributed equally to the intellectual content of this manuscript.

This research was supported by NIH grants HD34957 and DA12831. Portions of this paper were presented at the 2000 IUSSP Conference on Partnership Networks and the Spread of HIV and Other Infections, Chiang Mai, Thailand; the 2000 Annual Meeting of the International Network for Social Network Analysis, Vancouver, BC; and the 2001 annual meeting of the methodology section of the American Sociological Association, Minneapolis, MN.

Correspondence concerning this article should be addressed to Laura M. Koehly, Department of Psychology, Texas A&M University, College Station, TX 77843-4235. e-mail: koehlyl@tamu.edu.

FIGURE 1: Representations of network data

Graph



Sociomatrix

	R1	R2	R3	R4	U1	U2	U3	U4	U5	U6
R1	-	0	0	1	0	1	0	0	0	0
R2	1	-	1	0	0	0	0	0	0	0
R3	1	0	-	0	0	0	0	0	0	0
R4	1	1	0	-	0	0	0	0	0	0
U1	1	1	0	0	-	0	1	0	0	0
U2	0	0	0	0	0	-	0	1	0	0
U3	0	0	0	0	0	0	-	0	0	0
U4	0	0	0	0	0	1	0	-	1	0
U5	0	0	0	0	1	1	0	1	-	0
U6	0	0	0	0	0	0	0	0	1	-

Contact matrix and non-contact matrix

Y=1 (tie)

	R	U	$t_{i\bullet 1}$
R	6	1	7
U	2	8	10
$t_{\bullet j1}$	8	9	17

Y=0 (no tie)

	R	U	$t_{i\bullet 0}$
R	6	23	29
U	22	22	44
$t_{\bullet j0}$	28	45	73

FIGURE 2: Add Health Friendship Data, by grade

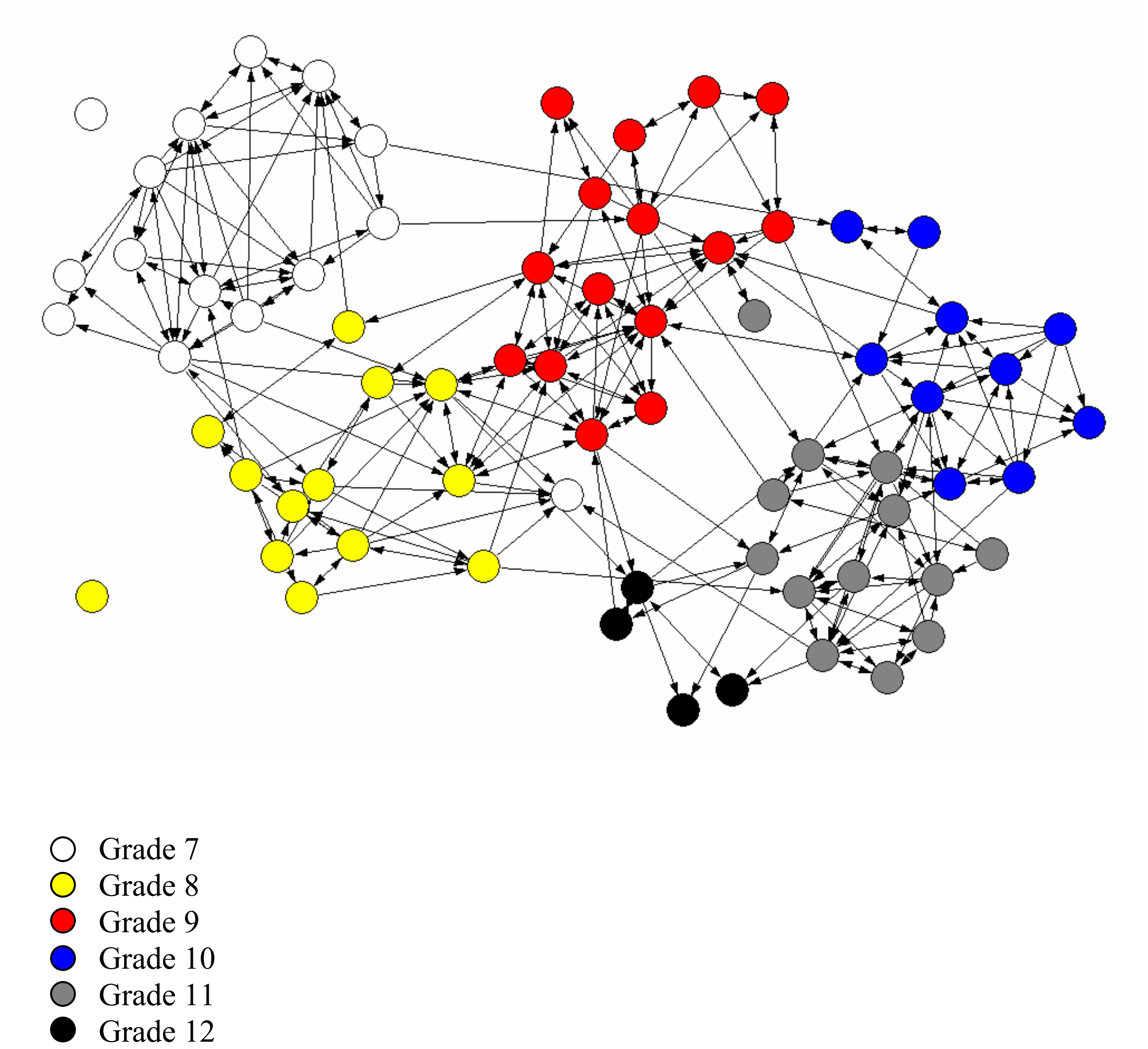


TABLE 1: Model matrices, 4x4 table

Model matrix for uniform homophily

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Implicit model matrices for marginal effects with first-level constraints

$$\begin{array}{ccc} A=2 & A=3 & A=4 \\ \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \end{array}$$

$$\begin{array}{ccc} B=2 & B=3 & B=4 \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{array}$$

TABLE 2: Corresponding models

Name	**Cond. loglinear model**	**Uncond. loglinear model**	**ERG model**
	cell count in layer Y=1 is a function of:	*cell count is a function of:*	*logit(Y) is a function of:*
Saturated	[AB]	[ABY]	[AB]
Bernoulli graph		[AB][Y]	[-]
Marginal Effects (ERG model) (i.e. no 3-way interaction)		[AB] [AY] [BY]	[A][B]
Marginal Effects (CLL) (i.e. independence of A and B conditional on Y)	[A][B]	[AY][BY]	
Non-saturated interaction (ERG model)		[AB] [AY] [BY] [U_{AB} Y]	[A][B] and [U_{AB}]
Non-saturated interaction (CLL)	[A][B] and [U_{AB}]	[AY] [BY] [U_{AB} Y]	

Notation follows Fienberg (1977) and many others. [X] refers to terms for each value of variable X. [XY] refers to a full set of interaction terms for X by Y, as well as terms for each level of X alone and of Y alone. U_{XY} ("U" for "unsaturated") indicates that some but not all of the set of interaction terms are included in the model (e.g. uniform homophily). Any interaction term implies that all lower order terms are included as well. The model pairs enclosed in each square make clear the lack of equivalence between ERG models and CLLs. In each case, the ULL that corresponds to the ERG model contains an [AB] interaction term that is missing from the ULL corresponding to the CLL. Although all of the other terms are identical between the two models, the presence or absence of the [AB] terms change the values and interpretations of the others.

TABLE 3: Add Health: Reported friendships and imputed non-friendships by grade of nominator and nominee for one school

Friendships

		Grade of nominee						
		7	**8**	**9**	**10**	**11**	**12**	
	7	52	5	*	*	*	*	59
	8	8	33	9	*	*	*	52
Grade	**9**	*	10	70	*	4	*	86
Of	**10**	*	*	3	30	10	*	43
Nominator	**11**	*	*	*	7	43	4	57
	12	*	*	*	*	*	5	8
		61	48	86	39	60	11	305

Non-friendships

		Grade of nominee						
		7	**8**	**9**	**10**	**11**	**12**	
	7	173	190	239	149	195	60	1006
	8	187	136	199	130	168	51	871
Grade	**9**	240	198	186	159	204	63	1050
Of	**10**	150	130	157	70	120	40	667
Nominator	**11**	194	169	206	123	126	48	866
	12	60	52	63	40	50	11	276
		1004	875	1050	671	863	273	4736

* cell contains fewer than 3 observations

TABLE 4: Parameter values for Add Health, marginal effects models with corresponding ULL models

ERG marginal effects model

ULL				ERG model	
γ	5.362	$\gamma^{Y}_{y=1}$	-2.895	θ	-2.895
$\gamma^{A}_{a=8}$	-0.144	$\gamma^{AY}_{a=8,y=1}$	0.018	$\theta^{A}_{a=8}$	0.018
$\gamma^{A}_{a=9}$	0.044	$\gamma^{AY}_{a=9,y=1}$	0.335	$\theta^{A}_{a=9}$	0.335
$\gamma^{A}_{a=10}$	-0.411	$\gamma^{AY}_{a=10,y=1}$	0.095	$\theta^{A}_{a=10}$	0.095
$\gamma^{A}_{a=11}$	-0.150	$\gamma^{AY}_{a=11,y=1}$	0.116	$\theta^{A}_{a=11}$	0.116
$\gamma^{A}_{a=12}$	-1.295	$\gamma^{AY}_{a=12,y=1}$	-0.706	$\theta^{A}_{a=12}$	-0.706
$\gamma^{B}_{b=8}$	-0.138	$\gamma^{BY}_{b=8,y=1}$	-0.102	$\theta^{B}_{b=8}$	-0.102
$\gamma^{B}_{b=9}$	0.046	$\gamma^{BY}_{b=9,y=1}$	0.299	$\theta^{B}_{b=9}$	0.299
$\gamma^{B}_{b=10}$	-0.403	$\gamma^{BY}_{b=10,y=1}$	-0.044	$\theta^{B}_{b=10}$	-0.044
$\gamma^{B}_{b=11}$	-0.151	$\gamma^{BY}_{b=11,y=1}$	0.135	$\theta^{B}_{b=11}$	0.135
$\gamma^{B}_{b=12}$	-1.304	$\gamma^{BY}_{b=12,y=1}$	-0.411	$\theta^{B}_{b=12}$	-0.411

γ^{AB}_{ab} interaction terms (n=25):

mean = 0.000
max. = 0.009
min. = -0.009
std. dev.= 0.0036

CLL marginal effects model

ULL				CLL	
γ	5.363	$\gamma^{Y}_{y=1}$	-2.894	λ	2.468
$\gamma^{A}_{a=8}$	-0.144	$\gamma^{AY}_{a=8,y=1}$	0.018	$\lambda^{A}_{a=8}$	-0.126
$\gamma^{A}_{a=9}$	0.043	$\gamma^{AY}_{a=9,y=1}$	0.334	$\lambda^{A}_{a=9}$	0.377
$\gamma^{A}_{a=10}$	-0.411	$\gamma^{AY}_{a=10,y=1}$	0.095	$\lambda^{A}_{a=10}$	-0.316
$\gamma^{A}_{a=11}$	-0.150	$\gamma^{AY}_{a=11,y=1}$	0.115	$\lambda^{A}_{a=11}$	-0.034
$\gamma^{A}_{a=12}$	-1.293	$\gamma^{AY}_{a=12,y=1}$	-0.704	$\lambda^{A}_{a=12}$	-1.998
$\gamma^{B}_{b=8}$	-0.138	$\gamma^{BY}_{b=8,y=1}$	-0.102	$\lambda^{B}_{b=8}$	-0.240
$\gamma^{B}_{b=9}$	0.045	$\gamma^{BY}_{b=9,y=1}$	0.299	$\lambda^{B}_{b=9}$	0.343
$\gamma^{B}_{b=10}$	-0.403	$\gamma^{BY}_{b=10,y=1}$	-0.044	$\lambda^{B}_{b=10}$	-0.447
$\gamma^{B}_{b=11}$	-0.151	$\gamma^{BY}_{b=11,y=1}$	0.135	$\lambda^{B}_{b=11}$	-0.017
$\gamma^{B}_{b=12}$	-1.302	$\gamma^{BY}_{b=12,y=1}$	-0.411	$\lambda^{B}_{b=12}$	-1.713

TABLE 5: Order of magnitude differences in cell counts for ULL marginal effects model fits with and without AB interaction parameters:

order of mag.	Y=1 layer	Y=0 layer
10^{-1} - 10^{0}	**1**	**13**
10^{-2} - 10^{-1}	**6**	**11**
10^{-3} - 10^{-2}	**15**	**1**
10^{-4} - 10^{-3}	**3**	-
Identical	**11**	**11**
total # cells	**36**	**36**

(Note the first-level contraints parameterization always results in 11 values being identical in each layer for this model).

TABLE 6: Order of magnitude differences in the Y=1 layer for ERGM vs. CLL model fits:

order of mag.	marginal effects model		uniform homophily model	
	cell prob.	cell counts	cell prob.	cell counts
10^{-1} - 10^{0}	-	-	-	**7**
10^{-2} - 10^{-1}	-	**10**	-	**26**
10^{-3} - 10^{-2}	-	**22**	-	**2**
10^{-4} - 10^{-3}	**2**	**4**	**19**	**1**
10^{-5} - 10^{-4}	**22**	-	**16**	-
10^{-6} - 10^{-5}	**12**	-	**1**	-
total # cells	**36**	**36**	**36**	**36**

REFERENCES

Agresti, A. 2002. *Categorical Data Analysis*. New York: Wiley-Interscience.

Besag, J. E. 1974. "Spatial Interaction and the Statistical Analysis of Lattice Systems." *Journal of the Royal Statistical Society Series B* 36:192-236.

---. 1975. “Statistical Analysis of Non-lattice Data.” *The Statistician*. 24: 179-195.

---. 1977. “Some Methods of Statistical Analysis for Spatial Data.” *Bulletin of the International Statistical Association*. 47:77-92.

Bowles, S., J. K. Choi, and A. Hopfensitz. 2003. “The co-evolution of individual behaviors and social institutions.” *Journal Of Theoretical Biology* 223: 135-147.

Burt, R. S. 1983. “Network Data from Archival Records.” In Burt, R. S., and Minor, M. J. (Eds.), *Applied Network Analysis*, pp. 158-174. Beverly Hills: Sage.

---. 1984. “Network Items and the General Social Survey.” *Social Networks* 6: 293-340.

---. 1990. "Kinds of relations in American discussion networks." In Calhoun, C., M. W. Meyer, and W. R. Scott (Eds.), *Structures of power and constraint: papers in honor of Peter M. Blau.* Cambridge: Cambridge University Press.

---. 1992. *Structural holes : the social structure of competition.* Cambridge: Harvard University Press.

Buve, A., M. Carael, R. J. Hayes, B. Auvert, B. Ferry, N. J. Robinson, S. Anagonou, L. Kanhonou, M. Laourou, S. Abega, E. Akam, L. Zekeng, J. Chege, M. Kahindo, N. Rutenberg, F. Kaona, R. Musonda, T. Sukwa, L. Morison, H. A. Weiss and M. Laga. 2001. "Multicentre Study on Factors Determining Differences in Rate of Spread of Hiv in Sub-Saharan Africa: Methods and Prevalence of HIV Infection." *Aids* 15 Suppl 4:S5-14.

Fienberg, S. E. 1977. *The Analysis of Cross-Classified Categorical Data*. Cambridge, Mass.: MIT Press.

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

Fischer, C. S. 1982. *To Dwell among Friends: Personal Networks in Town and City*. Chicago: University of Chicago Press.

Frank, O. 1988. “Random Sampling and Social Networks: A Survey of Various Approaches.” *Mathematiques, Informatique, et Sciences Humaines* 26: 19-33.

Frank, O. and D. Strauss. 1986. "Markov Graphs." *Journal of the American Statistical Association* 81:832-42.

Frank, O.and T. Snijders. 1994. "Estimating the size of hidden populations using snowball sampling." *Journal of Official Statistics* 10:53-67.

Geyer, C. J. and E. A. Thompson. 1992. "Constrained Monte Carlo Maximum Likelihood for Dependent Data." *Journal of the Royal Statistical Society Series B* 54:657-99.

Gilks, W. R., S. Richardson and D. J. Spiegelhalter. 1996. *Markov Chain Monte Carlo in Practice*. London: Chapman & Hall.

Gintis, H. 2000. *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton: Princeton University Press.

Goodman , L. A. 1965. "On the Statistical Analysis of Mobility Tables." *The American Journal of Sociology* 70:564-585.

Granovetter, M. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1360-80.

Handcock, M. 2003. "Statistical models for social networks: Inference and degeneracy." In Breiger, R., K. Carley, and P. Pattison (Eds.) *Dynamic Social Network Modeling and Analysis,* pp. 229-240. Washington, DC: National Academy Press.

Hoff, P. D., A. E. Raftery, and M. S. Handcock. 2002. "Latent space approaches to social network analysis." *Journal of the American Statistical Association* 97:1090-1098.

Holland, P. W. and S. Leinhardt. 1970. “A Method for Detecting Structure in Sociometric Data.” *American Journal of Sociology* 72:492-513.

---. 1976. "Local Structure in Social Networks." *Sociological Methodology* 7:1-45.

---. 1981. “An Exponential Family of Probability Distributions for Directed Graphs (with discussion). *Journal of the American Statistical Association* 76: 33-65.

Iacobucci, D. 1994. “Statistical Analysis of Single Relational Networks”. In Wasserman, S. and K. Faust. *Social Network Analysis: Methods and Applications* (pp. 605-674). New York: Cambridge University Press.

Ihaka, R. and R. Gentleman. 1996. "R: A Langauge for Data Analysis and Graphics." *Journal of Computational and Graphical Statistics* 5:299-314.

Laumann, E. O., J. H. Gagnon, S. Michaels, R. T. Michael and J. S. Coleman. 1989. "Monitoring the AIDS Epidemic in the United States: A Network Approach." *Science* 244:1186-9.

Laumann, E. O. and D. Knoke. 1987. *The organizational state: social change in national policy domains*. Madison: University of Wisconsin Press.

Liang, K.-Y. and S. L. Zeger. 1986. “Longitudinal Data Analysis Using Generalized Linear Models.” *Biometrika* 73:13-22.

Macy M.W. and Willer R. 2002. From factors to actors: Computational sociology and agent-based modeling. *Annual Review of Sociology* 28: 143-166.

Majeski S, Sylvan D. 1999. “How foreign policy recommendations are put together: A computational model with empirical applications.” *International Interactions* 25 (4): 301-332.

Mare, R. D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review* 56:15-32.

Marsden, P. V. 1981. "Models and methods for characterizing the structural parameters of groups." *Social Networks* 3:1-27.

---. 1987. "Core Discussion Networks of Americans" *American Sociological Review* 52:122-131.

---. 1988. “Homogeneity in Confiding Relations.” *Social Networks* 10: 57-76.

Massey, D. S. 1990. "The Social and Economic Origins of Immigration." *Annals AAPSS* 510: 60-72.

Morris, M. 1991. "A Loglinear Modeling Framework for Selective Mixing." *Mathematical Biosciences* 107:349-77.

---. *In press*. *Network Epidemiology: A Handbook For Survey Design and Data Collection.* Oxford: Oxford University Press.

Morris, M. and L. Dean. 1994. "Effect of Sexual Behavior Change on Long-Term Human Immunodeficiency Virus Prevalence among Homosexual Men." *American Journal of Epidemiology* 140:217-32.

Padgett J.F., Ansell C.K. 1993. “Robust Action and the Rise of the Medici, 1400-1434 *American Journal of Sociology* 98 (6): 1259-1319.

Pagnini, D. L. and S. P. Morgan. 1990. "Intermarriage and Social Distance Among U.S. Immigrants at the Turn of the Century." *The American Journal of Sociology* 96:405-432.

Parker D.C., Manson S.M., Janssen M.A., Hoffmann M.J., Deadman P. Multi-agent systems for the simulation of land-use and land-cover change: A review. 2003. *Annals Of The Association Of American Geographers* 93 (2): 314-337.

Pattison, P. E., and G. L. Robins. 2002. Neighbourhood-based models for social networks. *Sociological Methodology* 32:301-337.

Pattison, P. and S. Wasserman. 1999. "Logit Models and Logistic Regressions for Social Networks: II. Multivariate Relations." *British Journal of Mathematical & Statistical Psychology* 52:169-93.

Raymo, J. M. and Y. Xie. 2000. "Temporal and Regional Variation in the Strength of Educational Homogamy." *American Sociological Review* 65:773-781.

Resnick, M. D., P. S. Bearman, R. W. Blum, K. E. Bauman, K. M. Harris, J. Jones, J. Tabor, T. Beuhring, R. E. Sieving, M. Shew, M. Ireland, L. H. Bearinger and J. R. Udry. 1997. "Protecting Adolescents from Harm. Findings from the National Longitudinal Study on Adolescent Health." *Journal of the American Medical Association* 278:823-32.

Rindfuss, R., A. Jampaklay, B. Entwisle, Y. Sawangdee, K. Faust, P. Prasartkul *In press*. "The Collection and Analysis of Social Network Data in Nang Rong, Thailand." in M. Morris (ed.) *Network Epidemiology: A Handbook For Survey Design and Data Collection.* Oxford: Oxford University Press.

Robins, G., P. Elliott and P. Pattison. 2001a. "Network Models for Social Selection Processes." *Social Networks* 23: 1-30.

Robins, G., P. Pattison and P. Elliott. 2001b. "Network Models for Social Influence Processes." *Psychometrika* .

Robins, G., P. Pattison and S. Wasserman. 1999. "Logit Models and Logistic Regressions for Social Networks: III. Valued Relations." *Psychometrika* 64:371-94.

Snijders, T. A. B. 2001. "The Statistical Evaluation of Social Network Dynamics." *Sociological Methodology* 31:361-395.

Strauss, D. and M. Ikeda. 1990. "Pseudolikelihood Estimation for Social Networks." *Journal of the American Statistical Society* 85:204-12.

Thompson, S.K. and O. Frank. 2000. "Model-based estimation with link-tracing sampling designs." *Survey Methodology* 26:87-98.

Udry, J. R. and P. S. Bearman. 1998. "New Methods for New Research on Adolescent Sexual Behavior." In *New Perspectives on Adolescent Risk Behavior*, edited by R. Jessor. Cambridge ; New York: Cambridge University Press.

Wasserman, S. and K. Faust. 1994. *Social Network Analysis.* Cambridge: Cambridge University Press.

Wasserman, S. and P. Pattison. 1996. "Logit Models and Logistic Regressions for Social Networks: I. An Introduction to Markov Graphs and P*." *Psychometrika* 60:401-25.

Watts, D.J. 2003. *Six Degrees: The Science of a Connected Age.* New York: W.W. Norton.

Wellman, B. and S. Wortley. 1990. " Different Strokes from Different Folks: Community Ties and Social Support." *The American Journal of Sociology* 96: 558-588.

Yamaguchi, K. 2003. "A Liang-Zeger Method for Modeling Dyadic Interdependence in the Analysis of Social Networks." *Sociological Methodology* 33:343.