Multivar. Statist., pp. 211–227
E.–M. Tiit *et al.* (Eds)

METHODS FOR LONGITUDINAL SOCIAL NETWORK DATA: REVIEW AND MARKOV PROCESS MODELS

TOM A. B. SNIJDERS

Department of Statistics and Measurement Theory, University of Groningen, Grote Kruisstraat 2/1, 9712 TS Groningen, The Netherlands

ABSTRACT

Social network data pertain to social relations between individuals, or between other actors (countries, firms, etc.). The fact that each relation links two individuals, and each individual can be related with multiple others, results in complicated stochastic dependence structures. Longitudinal studies concerning the change and development of social networks in time are important for progress in sociology and related disciplines, but the literature does not contain enough satisfactory statistical models for longitudinal social network data. The paucity of satisfactory models is related to the difficulty of reconciling mathematical tractability with the expression of substantively interesting processes.

A review is given of models for longitudinal network data. Interesting developments are taking place in applying continuous-time Markov processes to model network change. Such models can incorporate sociologically important effects. They are in many aspects similar to computer simulation models that are common now in sociological theory, but contain a component of random change as an extra ingredient. Difficulties of mathematical tractability can be resolved by using stochastic approximation methods to obtain estimators and tests. This is done, however, at the expense of introducing some new problems: computer intense computations and questions about statistical efficiency. An example of this approach is given using a time series of personal preference relations in a closed group.

1. STATISTICAL MODELS FOR SOCIAL NETWORKS

Social networks, as studied in sociology and other social sciences, are the structures consisting of dyadic relations between individuals or other units (organisations, countries, etc.); a commonly used terms for these units is *actors*, stressing their active role in the constitution of the networks. This paper is concerned with so-called *entire networks*, where all relationships within a given set of n actors are considered. Such a network can be represented by an $n \times n$ matrix $X = (X_{ij})$, where X_{ij} is a number or a vector representing the relation directed from actor i to actor j $(i, j = 1, \ldots, n)$. Usually, self-relations are not

considered, so that the diagonal values X_{ii} are non-defined. An often used convention is $X_{ii} = 0$.

Traditionally, much attention has been paid to binary relations. In this case, X is a 0-1 matrix that is naturally seen as the adjacency matrix of a graph or directed graph. Pioneering work in the development of statistical methods for graphs was done by Frank (1971). The substantive meaning of the relationship can be, e.g., friendship, acquaintance, cooperation, opposition, etc.

Special problems that must be dealt with in the statistical analysis of social networks are constituted by the dependence structure of the elements of X:

- the dependence between the two halves X_{ij} and X_{ji} of the *dyad* (X_{ij}, X_{ji});
- the dependence between the elements of each row, corresponding to relations of the same *sender* i, and likewise among the elements of each column, referring to one *receiver* j;
- more complicated dependence structures, such as *transitivity*: "a friend of my friend is also my friend", and *group formation*.

These types of dependence have been modelled by stochastic dependence but also by using common parameters, e.g., row parameters and column parameters referring to sender and receiver characteristics. An important concept for the dyadic dependence is *symmetry* or *reciprocity*, indicating the tendency for X_{ij} and X_{ji} to be similar. In the most simple case of binary (0-1) data, this has led to the classification of dyads as *mutual*, $(x_{ij}, x_{ji}) = (1, 1)$, *asymmetric*, (0,1) or (1,0), and *null*, (0,0).

Another problem for the statistical analysis of social networks is

- the highly unbalanced nature of social network data: social structure does not have a crystalline regularity at all, so that there is usually not sufficient reason to assume simple balanced or symmetric structures.

In spite of the rather large amount of work that has been done on statistical and other data-analytic modelling of social network data, in many cases the currently available methods are not completely adequate to model the complexity of social relation patterns.

Longitudinal data is in principle much more informative about the studied phenomena in social networks than cross-sectional data; but they create additional problems for statistical analysis. A review of methods, including exploratory methods, for analysing change in networks was given by Frank (1991). The present paper is restricted to models that specify the joint probability distribution and discuss statistical procedures for time-series $X(t)$, $t \in \mathcal{T}$, of social networks for a constant set of actors.

2. SOME MODELS FOR LONGITUDINAL SOCIAL NETWORKS

2.1. *Discrete Markov chains*

It seems that the first statistical paper on longitudinal social networks was written by Katz and Proctor (1959). They propose a Markov chain model for binary relation networks, in which the dyads are assumed to be independent, each following a Markov chain for the 4 possible outcomes in the outcome space $\{0,1\}^2$, where the transition probabilities for the two asymmetric outcomes (0,1) and (1,0) are obviously required to be identical. This assumption of *dyad independence* has haunted the social network statistical literature, because the models become so complicated without it.

A recent paper that also proposes a Markov chain model is (Carley and Banks, 1995). They consider time series of directed graphs in which changes occur randomly, while the change rate parameter may change (e.g., decay) over time. The assumption of random changes is quite restrictive, and should be tested when these methods are applied.

2.2. *Loglinear models*

Sparked off by the work of Holland and Leinhardt (1981), many papers have appeared about loglinear models for categorical (binary or more complicated) networks. The basic model, called by Holland and Leinhardt the p_1 model, has probability function

$$K(\alpha,\beta,\rho)\exp\left(\sum_i \alpha_i x_{i+} + \sum_i \beta_i x_{+i} + \rho\sum_{ij} x_{ij}x_{ji}\right),$$

where an index replaced by a + indicates summation over that index. Some papers elaborating this approach are (Fienberg and Wasserman, 1981; Fienberg *et al.*, 1985; Wasserman and Weaver, 1985). Instead of using the incidental α_i and β_i parameters, these papers focus mostly on models where the senders and receivers are characterized by observed categorical actor-dependent variables. An extensive review of the literature on the p_1 and other models is given by Wasserman and Faust (1994).

The loglinear approach has been extended to longitudinal data by Wasserman (1987) for 2 time points, and by Wasserman and Iacobucci (1988) for multiple time points. The latter paper distinguishes between associative models, where quite complicated association patterns between the time-dependent values for each dyad (i,j) are modeled, and predictive models, that specify the conditional distribution of a dyad at each given time points, assuming its history is given. These models all assume independence between the dyads, however.

2.3. *Conditionally uniform models*

Another line of statistical models for social networks, mainly represented by directed graphs, is based on conditionally uniform models, i.e., models where the probability distribution is uniform, conditional on a certain set of statistics. Such a model reflects the notion that the conditioning statistics contain something that is relevant in the studied phenomena, and the rest is caused by random factors. Holland and Leinhardt (1975) initiated the study and application of this type of model, emphasizing the uniform model conditional on the dyad count, i.e., the numbers of mutual, asymmetric, and null dyads.

Conditionally uniform models become very complicated, however, when more informative conditioning statistics are considered. For example, conditioning on out-degrees and in-degrees is very relevant in sociological applications. When the relation can be interpreted as friendship or acquaintance, the in-degrees reflect the level of activity of the individual actors, and the out-degrees their popularity. These more complicated conditionally uniform models arise also when exact tests are required for the loglinear models. For example, to obtain an exact test for the hypothesis that the reciprocity parameter ρ in the p_1 model is equal to 0, one needs to condition on the sufficient statistics for the null hypothesis, which is just the vector of in-degrees and out-degrees; this leads to the uniform model, conditional on in- and out-degrees. This conditionally uniform model is difficult to handle statistically because these degrees can be regarded as incidental parameters, so that it is impossible to apply the usual asymptotic arguments. Algorithms for exact tests for this type of model were proposed by Snijders (1991).

Because of these complications, conditionally uniform models have been developed only for relatively simple conditioning statistics, and they are used more as a 'straw man' null hypothesis models than as models that pretend to represent reality. A conditional uniform model for longitudinal binary network data at 2 time points, conditional on the entire digraph at time 1, and on the numbers of newly formed and of disappeared relations for each actor, was developed by Snijders (1990).

2.4. *Continuous-time Markov chain models*

The most promising type of model is the continuous-time Markov model, where it is assumed that the social network, although observed only at discrete time points, follows a Markov process in continuous time. Holland and Leinhardt (1977a, 1977b); Wasserman (1977, 1979, 1980); Mayer (1984) proposed these models and developed estimation techniques for a number of special cases. Crucial for these estimation techniques is that either the dyads are assumed to be independent, so that a reduction by sufficiency leads to an outcome space with few elements, or other symmetrizing restrictions are made. An example of the latter is Wasserman's (1980) popularity model, where the rows of the adjacency matrix are assumed to be independent, and the change rate of

each X_{ij} depends only on the column sum (in-degree) X_{+j} which is an index for actor j's popularity, and where the assumption of stationarity is made. Leenders (1995) extended the possibilities for modeling longitudinal binary social network data as a continuous-time Markov processes, by dropping the assumption of stationarity and by allowing change rates for every dyad to be dependent on arbitrary covariates, provided that these remain constant between the moments of observation. In the following sections we present another elaboration of models that are also continuous-time Markov chains.

3. THE NEED FOR MODELS THAT REPRESENT SUBSTANTIVE THEORY

Most of the network models reviewed above are almost trivial from the point of view of sociological theory. The reason is the need to keep these models mathematically tractable in order to apply conventional statistical methods. For example, the assumption of independence between dyads excludes *a priori* almost all sociologically interesting interactions. Conditionally uniform models can serve at best as null models against which to test sociological theories. The continuous-time Markov models for networks offer more scope for expressing sociological theories, but they have not yet found much application because of the statistical difficulties associated with their practical use. The latter models can be seen as statistical models, not explicitly derived from a sociological theory, but in other aspects closely related to the models proposed below.

The remainder of this paper is devoted to statistical estimation of a class of models for evolution of the relation network and of individual behavior in a (non-changing) set of actors that are derived from sociological theories. A more extensive discussion of these models and associated methods for parameter estimation is given in (Snijders, 1995). The models treated are based on the assumption that each actor has his or her own goals which he/she tries to advance in accordance to his/her constraints and possibilities. The approach will be in the spirit of so-called methodological individualism: the driving force behind the network dynamics is constituted by the actors' actions, where each actor takes actions in order to further his own goals, which actions must be in the domain of his own behavior or of the directed relationships from him to others. More specifically, the models are in the rational choice tradition in sociology, as expounded, e.g., by Coleman (1990).

These stochastic models are so complicated that the development for them of statistical methods along classical lines (e.g., maximum likelihood estimators, minimum variance unbiased estimators, likelihood ratio tests) is difficult and seems to border on the impossible. On the other hand, computer simulation of these models is often quite feasible if only the models can be expressed in well-defined probability distributions. It will be shown below that computer simulation opens the possibility of (computer-intensive) methods of statistical inference, although these methods are not necessarily statistically efficient, or optimal in other statistical senses.

4. ELEMENTS OF STOCHASTIC ACTOR-ORIENTED NETWORK MODELS

The outcome space for our dynamic network model has the following basic components.

- The *set of actors*. In this paper a fixed and finite set of actors will be considered; addition of new actors, or exit from the network, will not be taken into consideration. The number of actors is denoted by n.
- The *network of relations* between the actors. All relations considered are directed relations (e.g., liking, esteem or power) because of the approach of methodological individualism. Relations commonly considered as undirected relations, such as friendship or cooperation, will be regarded as mutual directed relationships. Relations may be single, but it is more interesting to consider multiple and/or valued relations. Relationships between actors may (but need not) change in time, and may be determined by the social structure (e.g., hierarchy or kinship relations), but it is more interesting when the relationships can be purposely changed by the authors.

Time is considered in this model as a continuous parameter. The set of actors is fixed, but the relation network may change in time. The *state of the model* is the time-dependent value of the network. The probability distribution of the evolution of the model is assumed to be a continuous-time Markov process with finite outcome space. At random times (following a Poisson process of which the parameter is to be estimated), a random actor may make a change in the relations from him (or her) to the other actors. The transition 'mechanism' for this Markov process is described using the following ingredients.

- The *state of each actor* is a function of the state of the model on which depends his evaluation of this state in terms of his well-being. For example, in a model of friendship networks, the state of an actor could be the number of his friends, this concept is not necessary for the construction of the model, but it is often convenient.
- *Preference or utility functions* for each actor, defined as a function of the state of the actor. Utility functions can be the same for all actors, but in more complicated models they may differ between actors.
- The *actions* that an actor can take. Actions refer to changeable relations between the actors and other actors. For example, an action can be a friendship invitation to another actor, or the acceptance of another actor's invitation to a power contest. It is quite common that in the social structure there exist constraints to the actions that an author can take.
- The *choice* made by an actor to perform a certain action (or to refrain from doing so when the opportunity is offered) depends on the actor's

expectation of the utility of his state after the action. Ideally, the actor chooses the action that offers him the highest expected utility. Two limitations to the principle of utility maximization will be taken into account, however. First, the modeled utility functions will not be a perfect representation of the actors' utilities. Therefore, the utilities that propel the actors' choices may also contain random, i.e., unexplained, elements. (Random utility models are commonly used in econometric modeling (see, e.g. (Maddala, 1983; Pudney, 1989).) Second, an actor's future state may depend also on future actions of others, or on other things unknown to him; moreover, the actor's capacity for strategic foresight and general calculations is bounded. Therefore, instead of strictly maximizing his expected utility, each actor may use a *heuristic way* to approximate the expected utility associated to each of the alternatives for action available to him at a given moment. The choice of action by an actor will therefore be modeled as follows: for each alternative action available to the actor, the approximate, or perceived, expected utility is calculated; the actor does not necessarily choose the alternative with the highest perceived utility, but chooses among the alternatives with given probabilities which are a monotone increasing function of the perceived utility.

Analogous to linear regression modeling, these longitudinal models have to account for a degree of unexplained, or *random change*; the various theoretical effects introduced must push back this random aspect and explain the observed change as much as possible. In the process of model building, a sequence of increasingly complicated models can be fitted to the data, starting with a *null model* of random change.

5. ESTIMATION

A more extensive account of the estimation method proposed in this section, and of the example, is given in (Snijders, 1995).

The models considered in the remainder of this paper are Markov processes $(X(t))$ in continuous time of which the probability distribution is parametrized by a k-dimensional parameter θ. It is not assumed that the distribution of $X(t)$ is stationary. For a discrete set of time points $t = \tau_1, \ldots, \tau_M$, with $M \geqslant 2$, observations on $X(t)$ are available. For the type of models we have in mind, the likelihood function is too complicated to calculate, but Monte Carlo computer simulation of $X(t)$ is possible with the time parameter t starting at an arbitrary value τ_0 with a given outcome $x(\tau_0)$. In other words, a random drawing can be simulated from the conditional probability distribution of $X(t)_{t \geqslant \tau_0}$, given $X(\tau_0) = x(\tau_0)$. Because of the intractable likelihood function, estimation principles such as maximum likelihood are inapplicable. Therefore we propose an unconventional estimation method: the method of moments implemented

with Monte Carlo simulation. A related approach to estimation, also based on simulated expected values, was proposed by McFadden (1989) and Pakes and Pollard (1989). In this paper a different procedure is proposed, using stochastic approximation to solve the moment equations.

5.1. *Method of moments*

The method of moments proposed here is based on the conditional distributions of $X(\tau_{m+1})$ given $X(\tau_m)$, because we concentrate on the development dynamics of the network. We consider situations where observations at $M = 2$ time points are available; or where no stationarity assumption is made, so that separate estimations ar performed for each time interval $[\tau_{m-1}, \tau_m]$. Parameter θ can be estimated by a conditional moment estimator based on a k-dimensional statistic of the form

$$S\big(X(\tau_{m-1}), X(\tau_m)\big). \tag{1}$$

The function S must be chosen such that its expectation, conditional on $X(\tau_{m-1})$, is a coordinate-wise increasing function of θ (although this monotonicity is not always easy to prove; we often may have to rely on the intuitive plausibility of this monotonicity). For given data $x(\tau_{m-1})$, $x(\tau_m)$ the estimate $\hat{\theta}$ is defined to be the solution of

$$\begin{aligned} &\mathrm{E}_{\hat{\theta}}\big\{S(X(\tau_{m-1}), X(\tau_m)) \mid X(\tau_{m-1}) = x(\tau_{m-1})\big\} \\ &\quad = S(x(\tau_{m-1}), x(\tau_m)). \end{aligned} \tag{2}$$

The delta method can be used to derive an approximate covariance matrix for $\hat{\theta}$. Denote

$$\Sigma_\theta = \mathrm{Cov}_\theta\big\{S(X(\tau_{m-1}), X(\tau_m)) \mid X(\tau_{m-1}) = x(\tau_{m-1})\big\}, \tag{3}$$

$$D_\theta = \frac{\partial}{\partial\theta}\,\mathrm{E}_\theta\big\{S(X(\tau_{m-1}), X(\tau_m)) \mid X(\tau_{m-1}) = x(\tau_{m-1})\big\}; \tag{4}$$

then the approximate covariance matrix of $\hat{\theta}$ is

$$\mathrm{Cov}\,(\hat{\theta}) \approx {D_\theta}^{-1}\Sigma_\theta {D'_\theta}^{-1}. \tag{5}$$

5.2. *Stochastic approximation*

We are in a situation where we wish to solve the equation (2), while we cannot evaluate the left-hand side explicitly, but we do have a possibility to generate random variables with the desired distribution. Stochastic approximation methods, in particular adaptive variants of the Robbins–Monro procedure, can be used to obtain approximate solutions. For an introduction to stochastic approximation and the Robbins–Monro procedure, and results about the consistency of adaptive Robbins–Monro procedures, we refer to (Ruppert, 1991).

The proposed procedure is represented here in abbreviated notation as a recursive procedure to find the value of θ that solves

$$\mathrm{E}_\theta Z = 0 \tag{6}$$

for some random variable Z with probability distribution $F(\theta)$, where Z and θ are k-dimensional. The random variable Z represents the statistic (1) and its distribution is induced by the conditional distribution of $X(\tau_m)$, given $X(\tau_{m-1}) = x(\tau_{m-1})$.

The basic recursion formula for the Robbins–Monro (1951) procedure with step-size $1/N$ is

$$\theta_{N+1} = \theta_N - \frac{1}{N} D_N^{-1} Z_N(\theta_N), \tag{7}$$

where $Z_N(\theta)$ is a random variable with expected value $\mathrm{E}_\theta Z$. These steps are made for $N = 1, 2, \ldots$ up to a suitable total number $N_{\max}$. The optimal value of D_N is the derivative matrix $(\partial \mathrm{E}_\theta Z / \partial \theta)$. In adaptive Robbins–Monro procedures (Venter, 1967), this matrix is estimated during the approximation process as an average of difference quotients. For the stable estimation of these difference quotients, it is essential to use common random numbers (see also (Ruppert, 1991, Section 4.3)). The common random numbers technique is a well-known technique in Monte Carlo simulation, used to reduce the variance of differences between generated random variables. This technique operates by generating two or more random variables using the same stream of random numbers, obtained by employing the same initialisation of the random number generator.

In order to estimate D_θ and Σ_θ, we employ the following procedure. Since the parameters θ_j, $j = 1, \ldots, k$, may have different 'natural scales', positive numbers r_j are used that indicate these scales; they should be in the order of magnitude of the standard errors of the estimates of θ_j, i.e., the square roots of the diagonal elements of (5). These can be determined from earlier experience, or by trial and error. Define e_j as the scaled j-th unit vector ($e_{jj} = r_j$, $e_{jh} = 0$ for $h \neq j$) and let c_N denote a non-increasing sequence of small positive numbers. Generate random variables

$$\begin{aligned} Z_{N0} &\sim F(\hat{\theta}_N), \\ Z_{Nj} &\sim F(\hat{\theta}_N + c_N e_j) \qquad (j = 1, \ldots, k) \end{aligned} \tag{8}$$

using common random numbers. In order to obtain sufficient stability for the resulting process (7), it is advisable that the derivative matrices D_N be estimated with a precision that is not too low, even in the first steps of the process. For this reason, n_0 simulations are carried out before making any steps of the type (7), just to have a sufficiently stable value for D_N. Specifically,

values for (8) are simulated for $N = 1 - n_0, \ldots, 0$, with $\hat{\theta}_N = \hat{\theta}_1$, the initial value. The derivatives $\partial \mathrm{E}_\theta Z_i / \partial \theta_j$ are estimated by

$$D_{Nij} = \frac{1}{N + n_0} \sum_{n=1-n_0}^{N} \frac{Z_{nji} - Z_{n0i}}{c_n r_j}, \tag{9}$$

and the recursion process (7) is carried out for $N \geqslant 1$. We used $n_0 = 10$ and $N_{\max} = 200$ or 400.

An estimate of the covariance matrix Σ_θ can also be obtained from the random variables generated in the recursion process. If $\hat{\theta}_N$ is close to its limiting value $\hat{\theta}$, Z_{N0} generated according to (8) will have approximately the covariance matrix Σ_θ. The expected value $\mathrm{E}Z_{N0} - \mathrm{E}_\theta Z$ is approximately $D_\theta(\hat{\theta}_N - \theta)$. As a consequence, Σ_θ can be estimated by

$$\hat{\Sigma}_\theta = \frac{1}{N} \sum_{n=1}^{N} H_{Nn} H'_{Nn}, \tag{10}$$

where

$$H_{Nn} = \left(Z_{n0} - \bar{Z}_{(N)0} - D_N(\hat{\theta}_N - \bar{\theta}_{(N)})\right),$$

$$\bar{Z}_{(N)0} = \frac{1}{N} \sum_{n=1}^{N} Z_{n0}, \qquad \bar{\theta}_{(N)} = \frac{1}{N} \sum_{n=1}^{N} \hat{\theta}_n.$$

The resulting estimator for the covariance matrix (5) of $\hat{\theta}$ is

$$\widehat{\mathrm{Cov}}(\hat{\theta}) = D_N^{-1} \hat{\Sigma}_\theta {D'_N}^{-1}. \tag{11}$$

All nice properties of adaptive Robbins–Monro procedures that have been mathematically proved, are asymptotic for $N_{\max} \to \infty$. A good starting value for the recursions is important; from a poor starting value, it will take a very large number of recursion steps (7) to reach the solution of (6). An interactive procedure is used, where the drift in the values for θ_N is monitored, and where the value of N can be reset to a lower value if the drift is still appreciable after a relatively large number of steps.

6. NEWCOMB'S FRATERNITY

The book by Newcomb (1961), reports on an extensive longitudinal study of two groups of students living together in a student fraternity house. In this section a longitudinal model is proposed to express some of the theoretical mechanisms that, according to Newcomb's analysis, govern the development of the friendship network in these groups. This model is intended to give an example of the way of modeling proposed in the preceding section, and as a reconstruction of a part of Newcomb's theory.

6.1. *A model incorporating reciprocity and balance*

The set of actors is the set of $n = 17$ men living in the house in year II; the data used are those reported in the UCINET program (Borgatti, *et al.*, 1992). The relational data are given, for each moment where they are available, by sociometric rankings made by each man of the 16 others. We shall interpret this relation as liking. The ranking matrices are available for 15 almost consecutive weeks. (Data for week 9 are missing.) The ranking of actor j by actor i is denoted r_{ij}, where the value 1 indicates highest preference. The vector

$$r_{i*} = (r_{ij})_{j=1,\ldots,n;\ j\neq i}$$

thus is the permutation of the numbers 1 to $n-1 = 16$ indicating the preference ordering of actor i. The entire preference matrix is denoted by $\mathbf{r}$. The diagonal of this matrix is meaningless, and will be conventionally defined as 0. The weeks are indicated by the time parameter $t = 1, \ldots, 16$ ($t \neq 9$). Ranks r_{ij} or matrices $\mathbf{r}$ referring to a specific time point t are denoted $r_{ij}(t)$ or $\mathbf{r}(t)$, respectively.

The state of the actor, mentioned in Section 4, can be taken to consist of the complete preference matrix. (For this simple model, the concepts of the actor's state is not separately needed; it is mentioned here only for the sake of formal completeness.)

The preference function is the crucial part of the model, and must express some principal parts of the sociological theory developed and used by Newcomb. It will be convenient to work with a *tension function* rather than a utility function; a tension function is a function which the actors wish to minimize, and it may be considered equivalent to a constant minus the utility function. When starting with a bounded utility function, the tension function can be defined as the maximum of the utility function minus its present value. (Hoede (1990) and Zeggelink (1994) use tension functions in a similar way.) The principal effects proposed by Newcomb are *reciprocity* and *balance*, and we shall assume that each actor i wishes to minimize a tension function $p_i(\mathbf{r})$ which is the weighted sum of a reciprocity and a balance component. The reciprocity effect means that the actor prefers that others like him to the same degree as he likes them. The corresponding component of the tension function is defined as

$$p_i^{(1)}(\mathbf{r}) = \sum_{j=1,\ j\neq i}^{n} (r_{ij} - r_{ji})^2. \tag{12}$$

The balance effect means that the actor prefers that others to whom he is close view 'the world' in the same way as he views it himself. The group of other persons in the house is considered as a significant part of the world that determines an important part of the balance effect. Accordingly, the balance effect is understood more restrictively as the actor's preference that his friends in the fraternity house have the same preference order for the various other

persons in the house as he has himself. For defining the balance component, we use a non-increasing function $\phi(k)$ defined for $k = 1, \ldots, n-1$ which measures the closeness to i of the actor whom he accords rank $r_{ij} = k$. Assuming, somewhat arbitrarily, that the opinions of actor i's 5 closest friends in the house are especially important to him, this function is defined as

$$\phi(k) = \begin{cases} (6-k)/5 & \text{for } k = 1, \ldots, 5; \\ 0 & \text{for } k > 5. \end{cases} \tag{13}$$

The difference between two actors' views of their housemates is measured by the sum of squared differences of rankings,

$$\sum_{h=1,\ h\neq i,j}^{n} (r_{ih} - r_{jh})^2.$$

The balance component of the tension function is defined as

$$p_i^{(2)}(\mathbf{r}) = \sum_{j=1,\ j\neq i}^{n} \phi(r_{ij}) \sum_{h=1,\ h\neq i,j}^{n} (r_{ih} - r_{jh})^2. \tag{14}$$

The entire tension function is now

$$p_i(\mathbf{r}) = \alpha_1 p_i^{(1)}(\mathbf{r}) + \alpha_2 p_i^{(2)}(\mathbf{r}). \tag{15}$$

The parameters α_1 and α_2 indicate the importance of balance and reciprocity, respectively. We assume that the actor has no strategic foresight, but is completely myopic: he chooses (probabilistically) between the various possible actions on the basis of the value that his tension function will have immediately after the action has been taken.

We now come to the actions that can be taken by the actor. These actions are changes in his preference ordering. It is assumed that the changes in the actors' preferences occur slowly, more or less continuously, and that each actor is immediately aware of the changes in the others' preferences. These frequent but small changes are modeled as follows.

The time unit is a week, but time is regarded as a continuous parameter within weeks. Each actor has opportunities for action, i.e., for changing his preference order, at random time points in the week. These opportunities arise independently for the different actors, and follow, for each actor, a Poisson process with common intensity parameter λ. This means that, at every point on the continuous time axis and for every actor, the time that the actor has to wait until his next opportunity for action has a negative exponential distribution with mean $1/\lambda$; for each actor, the number of opportunities for action per week is a random variable having a Poisson distribution with expected value λ. The actions that each actor may take on these moments are *interchanges of*

preferences: when actor i has an opportunity for action and if he has adjacent preferences for j and h, e.g., $r_{ij} = k$, $r_{ih} = k+1$ for some number k, $1 \leqslant k \leqslant n-1$, then he may interchange these preferences leading to $r_{ij} = k+1$, $r_{ih} = k$. When the actor has an opportunity for action, he may also leave his preference order the same. Including this last possibility, the total number of possible actions therefore is n. A sequence of many small changes of this kind can change any ordering into any other ordering.

Given the preference orders of all the actors, the n alternative actions available to actor i at a given moment can lead to n different values for the preference matrix $\mathbf{r}$; indicate these values by $\mathbf{r}^{(1)}$ to $\mathbf{r}^{(n)}$. Their tension values for actor i are $p_i(\mathbf{r}^{(1)})$ to $p_i(\mathbf{r}^{(n)})$. One of these values, associated with the 'no change' alternative, is equal to his present tension. It is assumed that the probability of taking an action is a logistic function of the tension values:

$$P\left\{\text{change to } \mathbf{r}^{(k)}\right\} = \frac{\exp\left(-p_i\left(\mathbf{r}^{(k)}\right)\right)}{\sum_{l=1}^{n} \exp\left(-p_i\left(\mathbf{r}^{(l)}\right)\right)}. \tag{16}$$

This function can be motivated by a random utility argument as in (Maddala, 1983). This functional dependence of choice probabilities on utilities is often called the *multinomial logit model.* Note that the exponential function used in (16) has a unit (and therefore non-variable) scale parameter; this accords with the use of the two free scale parameters α_1 and α_2 in the definition of the tension function p_i.

The three parameters in this model are λ, α_1, and α_2. A higher value of λ leads to more rapid change. A higher value of α_1 will tend to push r_{ij} toward r_{ji}. A higher value of α_2 will have two effects: friendship choices (lower values of r_{ij}) of i are encouraged to those j who have similar preferences as i; and changes of i's preferences are encouraged into the direction of the preferences of his closer friends.

The probabilistic model for friendship development in the fraternity is now complete. Mathematically speaking, it is a continuous-time Markov chain for the discrete stochastic matrix $\mathbf{r}$. Special sub-models are:

- $\alpha_1 = \alpha_2 = 0$: purely random change;
- $\alpha_2 = 0$: change on the basis of reciprocity only.

The parameter λ cannot be set to 0, because that would imply that no change occurs at all. It is possible to consider the model where $\alpha_1 = 0$, $\alpha_2 > 0$, where changes occurs on the basis of balance only while reciprocity plays no role. This sub-model seems, however, very implausible theoretically, so we will not pay attention to this possibility.

6.2. *Statistics for moment estimation*

In order to apply the method of moments, we need to choose statistics that capture the effects of the three parameters in the model. The effects of the parameters were indicated above: λ determines the rate of change, α_1 reciprocity,

and α_2 balance. A statistic that is relevant for the amount of change from time t, with matrix $\mathbf{r}(t)$, to time $t+1$, with matrix $\mathbf{r}(t+1)$, is the normed sum of squared differences

$$\text{Dist}(t, t+1) = \frac{6}{n^2(n-1)(n-2)} \sum_{1 \leqslant i,j \leqslant n;\ i \neq j} \bigl(r_{ij}(t+1) - r_{ij}(t)\bigr)^2. \qquad (17)$$

The norming factor follows from asymptotic considerations under the null model where $\alpha_1 = \alpha_2 = 0$. Statistics that are relevant for the parameters α_1 and α_2 are the totals for reciprocity and balance over the set of all actors:

$$\text{Rec}(t+1) = \frac{2}{n(n-1)} \sum_{1 \leqslant i < j \leqslant n} \bigl(r_{ij}(t+1) - r_{ji}(t+1)\bigr)^2, \qquad (18)$$

$$\text{Bal}(t+1) = \frac{1}{c} \sum_{\substack{1 \leqslant i,j \leqslant n \\ i \neq j}} \phi(r_{ij}(t+1)) \sum_{h \neq i,j} \bigl(r_{ih}(t+1) - r_{jh}(t+1)\bigr)^2, \qquad (19)$$

where c is a norming constant,

$$c = n(n-2) \sum_{k=1}^{n-1} \phi(k) .$$

The statistic that has the role of (1) in the estimation process, is defined as

$$S(\mathbf{r}(t+1), \mathbf{r}(t)) = \bigl(\text{Dist}(t, t+1),\ \text{Rec}(t+1),\ \text{Bal}(t+1)\bigr). \qquad (20)$$

The statistics $\text{Rec}(t+1)$ and $\text{Bal}(t+1)$ depend only on $\mathbf{r}(t+1)$, not on $\mathbf{r}(t)$. However, their probability distribution as it is used in the simulations for the method of moments does depend on $\mathbf{r}(t)$, as this is the starting value for the simulations in the time interval $(t, t+1)$.

6.3. *Parameter estimates*

It is not *a priori* clear whether the parameters λ, α_1 and α_2 may be assumed to be constant during the whole observation period. Therefore, results are presented of estimation of parameters separately for all weeks. The estimation has been performed for all transitions from week t to week $t+1$: $(t = 1, \ldots, 15)$, where weeks 8 and 9 have been taken together because data for the beginning of week 9 are missing. The outcomes of statistics (20) are presented in Table 1.

The parameter estimates are presented in Table 2.

The estimates of λ quickly decrease from $\hat{\lambda} = 192$ for the first period to values around 40 for the later periods. This means that larger changes in preferences occurred in the beginning (the persons living in the house did not know each other before moving in) than later. The estimates for α_1 and α_2

Table 1.
Outcomes of distance, reciprocity and balance statistics

Period t	Dist(t, $t+1$)	Rec($t+1$)	Bal($t+1$)
1	0.3538	26.43	29.89
2	0.1934	24.34	28.83
3	0.1500	25.88	26.57
4	0.1597	27.23	22.11
5	0.1199	29.69	19.91
6	0.0872	30.69	21.25
7	0.0810	26.68	20.16
8–9	0.0960	27.53	20.39
10	0.1067	28.28	18.02
11	0.1123	30.22	20.64
12	0.1062	29.51	20.59
13	0.0787	31.03	21.49
14	0.0948	29.84	18.58
15	0.1012	30.94	17.39

Table 2.
Reciprocity with Balance Model: Robbins–Monro Moment Estimates

Period t	$\hat{\lambda}$	S.E.($\hat{\lambda}$)	$\hat{\alpha}_1$	S.E.($\hat{\alpha}_1$)	$\hat{\alpha}_2$	S.E.($\hat{\alpha}_2$)
1	192.1	32.1	0.0071	0.0015	0.0035	0.0008
2	93.8	15.7	0.0104	0.0042	0.0041	0.0023
3	67.0	11.4	0.0058	0.0023	0.0046	0.0009
4	72.6	14.4	0.0058	0.0024	0.0070	0.0011
5	53.0	9.9	0.0039	0.0037	0.0073	0.0013
6	37.0	4.7	0.0034	0.0053	0.0042	0.0017
7	31.5	3.6	0.0247	0.0062	0.0084	0.0016
8–9	42.5	5.9	0.0096	0.0035	0.0061	0.0019
10	45.0	7.6	0.0076	0.0060	0.0088	0.0027
11	50.3	6.9	0.0035	0.0032	0.0035	0.0016
12	46.7	6.7	0.0086	0.0029	0.0057	0.0011
13	32.9	4.1	0.0028	0.0055	0.0044	0.0018
14	38.7	10.5	0.0081	0.0044	0.0092	0.0041
15	44.3	5.7	0.0056	0.0083	0.0075	0.0054

are quite variable; period 7 stands out with a high value for $\hat{\alpha}_1$. The average estimates are 0.0076 for α_1 and 0.0060 for α_2.

Taken individually, many of the parameter estimates for α_1 and α_2 exceed twice their standard error, which suggests that they are significantly larger than 0. To obtain tests for the significance of the α_1 and α_2 parameters in which the 14 periods are combined, the average parameter estimates for α_1 and α_2 were divided by their standard errors. The resulting t-values were 6.2 for α_1 and 9.5 for α_2, indicating strongly significant reciprocity and balance effects.

REFERENCES

Borgatti, S., Everett, M. G. and Freeman, L. C. (1992). *UCINET IV Version 1.0 Reference Manual*. Analytic Technologies, Columbia.

Carley, K. and Banks, D. (1995). Models for fixed-node networks: Model fitting and model testing. *Social Networks* (to appear).

Coleman, J. S. (1990). *Foundations of Social Theory*. Harvard University Press, Cambridge, London.

Fienberg, S. and Wasserman, S. (1981). Categorical data analysis of directed graphs: model for a single generator. In: *Sociological Methodology 1981*, S. Leinhardt (Ed.), Jossey-Bass, San Francisco.

Fienberg, S. E., Meyer, M. M. and Wasserman, S. (1985). Statistical analysis of multiple sociometric relations. *JASA* **80**, 51–67.

Frank, O. (1971). *Statistical Inference in Graphs*. FOA Repro, Stockholm.

Frank, O. (1991). Statistical analysis of change in networks. *Statistica Neerlandica* **45**, 283–293.

Hoede, C. (1990). Social atoms: kinetics. In: *Social Networks Through Time*, pp. 45–63, J. Weesie and H. D. Flap (Eds), ISOR, Utrecht.

Holland, P. and Leinhardt, S. (1975). Local structure in social networks. In: *Sociological Methodology 1976*, D. R. Heise (Ed.), Jossey-Bass, San Francisco.

Holland, P. and Leinhardt, S. (1977a). A dynamic model for social networks. *J. Math. Sociology* **5**, 5–20.

Holland, P. and Leinhardt, S. (1977b). Social structure as a network process. *Zeitschrift für Soziologie* **6**, 386–402.

Holland, P. and Leinhardt, S. (1981). An exponential family of probability distributions for directed graphs. *JASA* **76**, 33–50.

Katz, L. and Proctor, C. H. (1959). The concept of configuration of interpersonal relations in a group as a time-dependent stochastic process. *Psychometrika* **24**, 317–327.

Leenders, R. Th. A. J. (1995). Models for network dynamics: a Markovian framework. *J. Math. Sociology*, to appear.

Maddala, G. S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge.

Mayer, T. (1984). Parties and networks: stochastic models for relationship networks. *J. Math. Sociology* **10**, 51–103.

McFadden, D. (1989). A method of simulated moments for estimation of discrete response models without numerical integration. *Econometrica* **57**, 995–1026.

Newcomb, T. (1961). *The Acquaintance Process*. Holt, Rinehart & Winston, New York.

Pakes, A. and Pollard, D. (1991). The asymptotic distribution of simulation experiments. *Econometrica* **57**, 1027–1057.

Pudney, S. (1989). *Modelling Individual Choice*. Basil Blackwell, Oxford.

Robbins, H. and Monro, S. (1951). A stochastic approximation method. *Ann. Math. Statist.* **22**, 400–407.

Ruppert, D. (1991). Stochastic approximation. In: *Handbook of Sequential Analysis*, B. K. Gosh and P. K. Sen (Eds), Marcel Dekker, New York.

Snijders, T. A. B. (1990). Testing for change in a digraph at two time points. *Social Networks* **12**, 359–373.

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

Snijders, T. A. B. (1995). Stochastic actor-oriented dynamic network analysis. *J. Math. Sociology*, to appear.

Venter, J. H. (1967). An extension of the Robbins–Monro procedure. *Ann. Math. Statist.* **38**, 181–190.

Wasserman, S. (1977). Stochastic models for directed graphs. Ph.D. dissertation. University of Harvard, Dept. of Statistics.

Wasserman, S. (1979). A stochastic model for directed graphs with transtion rates determined by reciprocity. In: *Sociological Methodology 1980*, K. F. Schuessler (Ed.), Jossey-Bass, San Francisco.

Wasserman, S. (1980). Analyzing social networks as stochastic processes. *JASA* **75**, 280–294.

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

Wasserman, S. and Faust, K. (1994). *Social Network Analysis*: *Methods and Applications*. Cambridge University Press, New York, Cambridge.

Wasserman, S. and Iacobucci, D. (1988). Sequential social network data. *Psychometrika* **53**, 261–282.

Wasserman, W. and Weaver, S. (1985). Statistical analysis of binary relational data: parameter estimation. *J. Math. Psychology* **29**, 406–427.

Zeggelink, E. P. H. (1994). Dynamics of structure: an individual oriented approach. *Social Networks* **16**, 295–333.