ANALYZING SOCIAL NETWORKS AS STOCHASTIC PROCESSES*

by

Stanley S. Wasserman

Technical Report No. 314

Department of Applied Statistics
School of Statistics
University of Minnesota
Saint Paul, MN 55108

March, 1978

Abstract

This paper present a new methodological for studying a social network of interpersonal relationships. The methods are based on a stochastic modelling framework that allows for the investigation of the changes that occur in a network over time. Specifically, we postulate that these changes can be modelled as a continuous time Markov chain. The transition rates for the chain are dependent upon a small set of parameters that measure the importance of various aspects of social structure on the probability of change.

We discuss the assumptions of the framework and describe two simple models that are applications of it. Several examples are presented and analyzed, and methods of parameter estimation are outlined. The models prove to be quite effective and allow us to better understand the evolution of a network.

I. Introduction: Modelling Change in Social Networks

Many new methodological tools have been proposed during the present decade for the analysis of social networks. These techniques either concentrate on the local or "micro" structure of a social group by examining subgraphs, particularly dyads and triads (see Holland and Leinhardt, 1970, 1975; Wasserman, 1977a) or attempt to represent a set of sociograms for a particular group by a substantively interesting collection of "blockmodels," a more global or "macro" paradigm (see White, Boorman, and Breiger, 1976; Arabie, Boorman, and Levitt, 1978).

Unfortunately, none of these techniques allow for the direct modelling of structural change in a social group. Holland and Leinhardt's methods utilize only one observation on a group, a single snapshot taken at a fixed time point in the evolution of the group, and thus completely ignore the issue of changing structure. White, et al., recognize the importance of studying structural change, and do discuss how a researcher can construct an array of blockmodels for a social group using observations on the group from several points in time. One can then study how the blockmodels for a group change over time, as the group approaches equilibrium. But they propose no model for this structural change, either stochastic or deterministic. Evolution is described simply in qualitative terms with no quantified statements concerning either the rates of movement of a group toward an equilibrium state or the internal or external forces governing how the present structure and environment of a group influence future structure. Indeed, they recognize that the primary analytic question still unanswered is how their hypothesized role structures evolve over time (Boorman and White, 1976, page 1442).

The basic premise of this paper is that a social group is a dynamic entity, with a gradually evolving structure that emerges over time. We shall tacitly assume that a group eventually reaches a statistical equilibrium, i.e. transitions between states continue to occur, but the probability that a group is in a specific state approaches a constant, limiting value. All of the models that are proposed in this paper are attempts to quantify how the present structure of a group "pushes" the group toward equilibrium. The existing methodologies are quite good for describing either local or global structure of a group already in equilibrium. This research hopefully will fill the existing methodological void and generate methods for studying the approach of a group toward its assumed equilibrium.

As an example of a small social group evolving over time, consider a sociometric study of a group, e.g. the college fraternity studied by Newcomb (1961) for fifteen weeks, during the second year of his two year investigation of how individuals become acquainted and the kinds of relationships between persons which emerge. During its existence, such a group undoubtedly passes through several stages of development at various rates. Some stages may be achieved quite rapidly, others rather slowly, and the duration of time spent in a given stage may be long or short. Sociomatrices collected longitudinally at different time points in the history of a group will reflect this structural development. As a simple example of change in structure, we examine Newcomb's seventeen member fraternity. Each member was asked to rank each of his fellow members on the basis of positive feeling. Consider the number of intransitive triads in the constructed binary digraph, for each of the fifteen weekly rankings made by the members (no rankings were done on week 9).

We compute τ_i, Holland and Leinhardt's standardized measure of group intransitivity for week i, $i = 0,1,\ldots, 15$, $i \neq 9$ (see Holland and Leinhardt, 1975, pages 35-37). Figure 1 shows how τ_i, a widely accepted measure of the social structure of a group, varies over time.

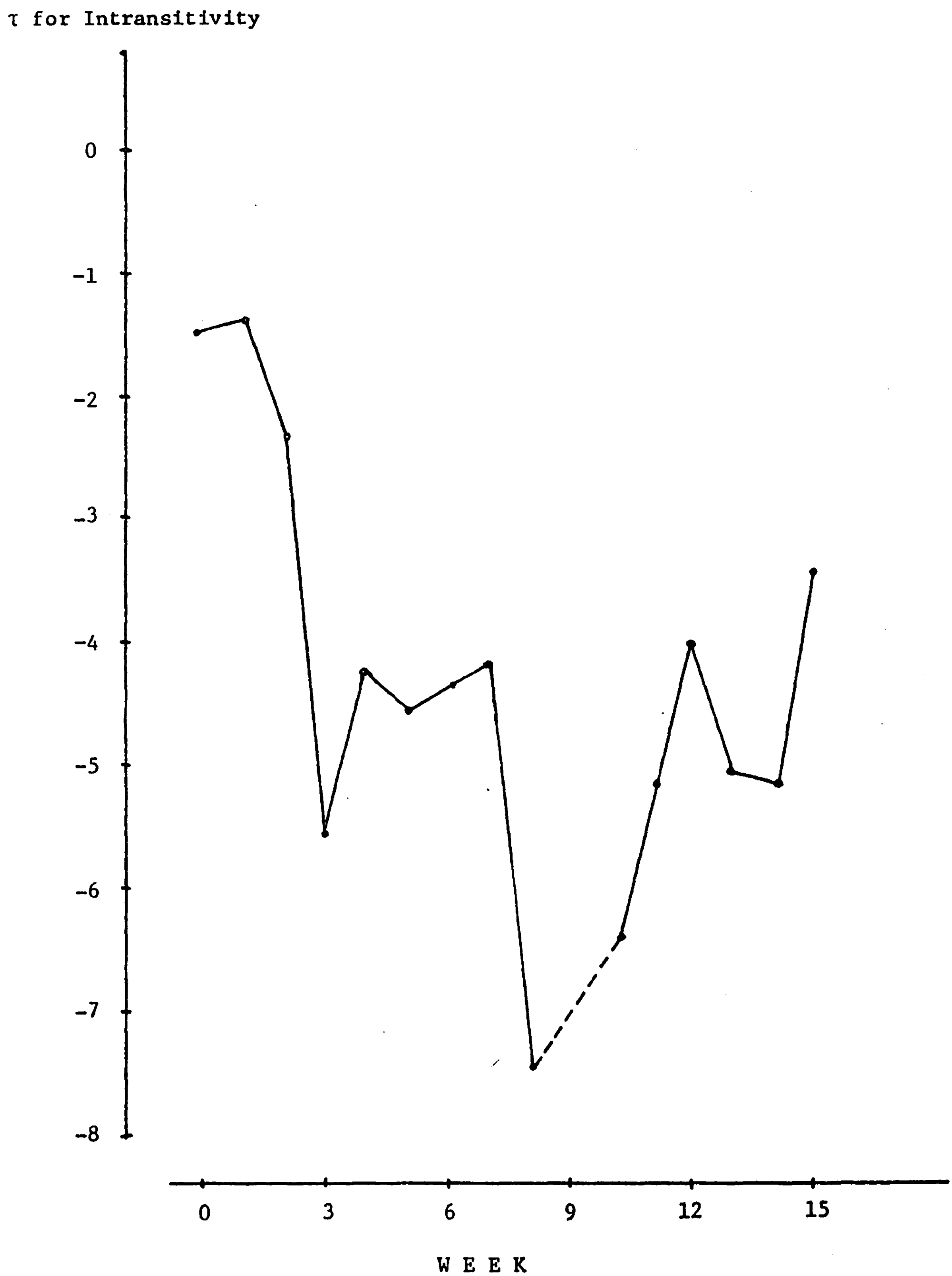
(Figure 1 about here)

Obviously, this fraternity is undergoing change. In the first two weeks of its development, intransitive triads do exist, but not as many as predicted by chance. After the second week, the group appears to exhibit a rather erratic equilibrium, except for the strange dip at week 8, with many fewer intransitivities than expected by change. A single observation taken on this group at week 8 would not at all be indicative of its average structure. Such conclusions can only be drawn from longitudinal data. We shall return to the analysis of this data set in a later section.

* *

In this paper several stochastic models of structural change are proposed which will constitute the foundation of a new methodology for the analysis of sociometric longitudinal data. We shall define social structure in terms of how various graph-theoretic properties of a group such as indegrees, outdegrees, and presence/absence of reciprocated arcs change over time. The stochastic models developed here have transition rates that are dependent upon the quantified social structure of a group. Holland and Leinhardt (1977b) elaborate further on this unique idea, characterizing how social structure can act as a generator of a network. The use of the proposed modelling frameworks to make precise sociological

Figure 1. Intransitivity Measure for Newcomb's Second Fraternity.



statements on group structure is perhaps the most important feature of these models.

In Section 2 of this paper we discuss two stochastic modelling frameworks for social networks, including a new paradigm based on a model for genetic nets of Kauffman (1969). We present two simple stochastic models, the reciprocity and popularity models, in Section 3, and fit these models to several data sets in Section 4. To conclude, we give some suggestions for future network modellers in Section 5. Other suggestions for network modelling are given in Wasserman (1978). Much of the material presented here is based on the author's unpublished doctoral dissertation (Wasserman, 1977b) and briefly discussed in Wasserman (1977c).

II. Postulating Modelling Frameworks for Networks

In this section we discuss two systems for modelling sociometric data, originating with Kauffman (1969) and Holland and Leinhardt (1977a). We use Kauffman's paradigm to explain with simple mathematical definitions and functions how a social network of interpersonal relationships can be viewed as a stochastic process. In this system, whenever a change occurs in a bond or link connecting two individual nodes, the probability that the change causes an arc to form or disappear depends on the current state of the process, the entire network, in a rather simple functional relationship.

We then present the framework of Holland and Leinhardt, discussing the mathematical assumptions for the infinitesimal transition rates and the characteristics of the state space of the process. The reciprocity and popularity models described in the later sections of this paper utilize this framework. First, we give a few necessary mathematical definitions and explanations.

We use continuous time stochastic processes to model change in social networks. A group's structural development obviously progresses continuously through time, and a discretization of this progression may force unnecessary and unrealistic assumptions upon the process of change. In addition, we choose to model this structural development of a group as a finite state Markov chain. One can always argue that in the social sciences, Markov models are gross oversimplifications of reality. It is very unlikely that past individual group behavior has no direct influence on future structure. However, our goal is not the construction of elaborate, perhaps non-Markovian, models incorporating parameters for all the factors of social structure that modify interpersonal relationships; rather, we shall postulate <u>simple</u> models so that we can make use of the wealth of information on the analysis of continuous time Markov chains. By doing so, we can concentrate on the interpretation and evaluation of the models, and will not be bogged down by the mathematics that usually accompany non-Markovian models in continuous time with complex state spaces.

We deal with the simplest type of interpersonal relation, the existence or nonexistence of a choice between two individuals. Our models place probability functions on these binary arcs that specify how likely it is that a specific arc appears or disappears in the next interval of time. Moreover, the length of time that an arc remains unchanged is a random variable, exponentially distributed by the Markov assumption (this rather stringent assumption is easily relaxed by working with semi-Markov models, in which waiting times have arbitrary distributions).

We let $\underset{\sim}{X}(t)$ be a binary matrix valued stochastic process, with elements $(X_{ij}(t))$, where

$$X_{ij}(t) = \begin{cases} 1, \text{ if } i \text{ chooses } j \text{ at time } t \\ 0, \text{ otherwise.} \end{cases} \tag{1}$$

Occasionally we will write $i \rightarrow j$ if i chooses j. We assume that the group under investigation has g members, so that $\underset{\sim}{X}(t)$ is a $(g \times g)$ binary matrix, with main diagonal fixed at zero by convention. As mentioned previously, t is a continuous parameter. The matrices $\underset{\sim}{x}, \underset{\sim}{w}, \underset{\sim}{y}, \underset{\sim}{z}, \ldots$ are single states or realizations of the continuous time process. There are $2^{g(g-1)}$ possible realizations of the process, so that $\mathcal{S}$, the state space of $\underset{\sim}{X}(t)$, is quite large, but still finite.

II.1 Interpreting a Social Network as a Stochastic Process

We have informally postulated that a network is an evolving entity, governed by a complex social structure that influences when, where, and how changes in individual arcs occur. How can we best view this process of structural change in the simplest possible manner? Can we construct a dynamic system that mimics the forces causing change and the changes themselves in a network? Kauffman's (1969) work on the modelling of random nets of binary genes provides a useful and simplistic framework for network evolution. We shall apply it to this situation not as a suggested method for predicting future structure but as a way of mathematically comprehending how changes in interpersonal relations could arise.

Consider a specific arc in a social network. At some time t, this arc, $X_{ij}(t)$, which we shall write as x, is either "on", indicating that a relationship exists between individual i and individual j at this point in time, or "off", indicating no relationship. By our Markov assumption, this binary variable x will remain in its 1 or 0 state for an exponentially distributed length of time. Its average waiting time, the reciprocal of the parameter of this exponential distribution, is a complex function of the <u>entire</u> network. This implies that the length of time for a relationship to emerge in a newly formed group, devoid of any ties, may be much longer than in a group that has been in existence for quite a while with many interpersonal relationships present. In addition, the average waiting times will, in general, not be constant across the set of binary arcs.

Suppose that during every instant of time $(t,t+h)$, the binary x evaluates the current structure of the network and decides whether to change or not. Thus, $y = X_{ij}(t+h)$ may or may not be equal to $x = X_{ij}(t)$. Presumably, an individual in a social group is constantly evaluating his or her position in the group, and continually making decisions on forming new ties or severing old ones. We can model this process of evaluation and decision as follows:

1) x receives binary inputs from n other arcs, $z_1, z_2, \ldots, z_n$. These z's are merely elements of $\underset{\sim}{X}(t)$ and are the set of relationships in the network that influence individual i's choice or non-choice of individual j.

2) x also examines the length of time τ since it last underwent change.

3) x's decision to change or not is governed by the binary output function f_{ij}; in fact, $y = f_{ij}(x, \tau, z_1, \ldots, z_n)$.

4) The model can allow at most one arc to change during the interval $(t,t+h)$, following the Holland-Leinhardt (1977a) framework, or can relax this restriction as in Mayer's (1977) "party" framework.

This representation of a relationship as a binary variable, whose value at the end of a very small interval of time is determined by a binary output function of binary inputs and time since last change, is the key conceptualization of this paradigm. It conveys a clear understanding of the process of structural change in a network of a group of individuals. Its primary utility is to give the user of the methodology proposed here new insights into the role of stochastic processes as modelling tools for social networks.

II.2 The Basic Modelling Framework

The Holland-Leinhardt (1977a) modelling framework postulates two assumptions that define a very general family of models. The first is that $\underset{\sim}{X}(t)$ is a Markov chain:

Assumption 1. Markov

$\underset{\sim}{X}(t)$ is a standard Markov chain with finite state space S, and probability transition function $P_{\underset{\sim}{x}\underset{\sim}{y}}(t,h)$ defined as

$$P_{\underset{\sim}{x}\underset{\sim}{y}}(t,h) = P\left\{\underset{\sim}{X}(t+h) = \underset{\sim}{y} \mid \underset{\sim}{X}(t) = \underset{\sim}{x}\right\}. \quad (2)$$

A standard Markov chain has a probability transition matrix equal to the identity matrix when the interval $(t,t+h)$ shrinks to a width of zero. By this first assumption, the future behavior of the group is completely determined by the present state of the process, $\underset{\sim}{X}(t)$; i.e. the past is irrelevant.

The second assumption of the framework states that for a small interval of time, the changes in the relationships between individuals in the group are statistically independent:

Assumption 2. Conditional Change Independence

$$P_{\underset{\sim}{x}\underset{\sim}{y}}(t,h) = \prod_{i,j} P\left\{X_{ij}(t+h) = y_{ij} \mid \underset{\sim}{X}(t) = \underset{\sim}{x}\right\} + o(h) \quad \text{as } h \to 0. \quad (3)$$

The probability transition matrix, for small h, can be factored into the probabilities of change for all arcs. Statistically, this means that the changes in arcs are independent, just as the factorization of a multivariate density function into a product of univariate densities implies that the random variables in question are independent. Consequently, the probability that any two arcs act in collusion and change simultaneously, is essentially zero.

Assumption 2 greatly simplifies the mathematics needed to fit this framework to data. Mayer (1977) has objected to this assumption, and has developed a new framework that allows for many simultaneous arc changes, occurring as "party" events. Unfortunately, his models are difficult to analyze mathematically, and consequently, he has only studied very simple models generated by his framework.

In the small interval of time (t,t+h), a relationship can change in only two ways: 1) if there is no bond from individual i to individual j at time t, such a bond may be present at time t+h; and 2) a bond linking individual i to individual j at time t, may not be present at time (t+h). In mathematical terms, a "0" may become a "1", and vice versa. Hence, we can further define the elements of the probability transition matrix, defined in (2) and (3), as

$$\begin{aligned} P\left\{X_{ij}(t+h) = 1 \mid \underset{\sim}{X}(t) = \underset{\sim}{x},\ X_{ij}(t) = 0\right\} &= h\lambda_{0ij}(\underset{\sim}{x},t) + o(h) \\ P\left\{X_{ij}(t+h) = 0 \mid \underset{\sim}{X}(t) = \underset{\sim}{x},\ X_{ij}(t) = 1\right\} &= h\lambda_{1ij}(\underset{\sim}{x},t) + o(h). \end{aligned} \tag{4}$$

These functions, λ_{0ij} and λ_{1ij}, are the infinitesimal transition rates for the continuous time, Markov chain $\underset{\sim}{X}(t)$. Note that they depend on both $\underset{\sim}{x}$, the current state of the process, and the time t. The unique feature of this framework is that it allows these functions to depend on properties of the nodes and on various features of the group structure at time t, modelled as graph-theoretic functions. Holland and Leinhardt (1977a, 1977b) and Wasserman (1977b) give examples of various λ_{0ij} and λ_{1ij}.

A researcher may develop a stochastic model for a social network simply by specifying a functional form for these infinitesimal transition rates (4). Two examples are given in the next section. This framework allows him or her to define "social structure" by a set of graph-theoretic quantities, such as indegree of each node or the number of intransitive triads that node i is involved in at time t, and to combine these to form the transition rates of the process. Thus, we can model how various aspects of group structure, suitably quantified, affect the future relationships in a group evolving through time as a stochastic process.

III. Two Simple Stochastic Models

In this section we present two stochastic models for social networks, each incorporating assumptions 1 and 2 of the modelling framework. The first model is for reciprocity of friendship, where the tendency over time for a choice of individual j by individual i, $i \to j$, depends only on whether or not $j \to i$. Analytically, this model yields a state space with only 4 states and is quite easy to work with. The reciprocity model has been studied in detail by Wasserman (1977d).

We also discuss a model for popularity in which the transition rates for a change in the relationship from individual i to individual j depends on how popular individual j is, as measured by the indegree of node j. Both of these models are defined simply by assuming the Holland-Leinhardt assumptions are operating and by specifying the infinitesimal transition rates (4). In Section IV, we discuss the estimation of the unknown model parameters.

III.1 Reciprocity Model

Suppose that

$$\begin{aligned} \lambda_{0ij}(\underset{\sim}{x},t) &= \lambda_0 + \mu_0 x_{ji} \\ \lambda_{1ij}(\underset{\sim}{x},t) &= \lambda_1 + \mu_1 x_{ji} \; ; \end{aligned} \qquad (5)$$

that is, the probability that a choice $i \to j$ is made, or that a choice $i \to j$ is withdrawn in the interval of time $(t,t+h)$ depends only on the presence or absence of the reciprocated choice $j \to i$. Note that we have made a further simplification by assuming that λ_{0ij} and λ_{1ij} do not depend on the time t; hence, the transition rates are time homogeneous, and the network stochastic process is stationary in time.

The transition rates (5) depend on x_{ji} by a linear equation, with coefficients λ_0 and μ_0, or λ_1 and μ_1. The parameters λ_0 and λ_1 are measures of the overall rate of change of the group. The parameters μ_0 and μ_1 measure the "importance" of a reciprocated arc, but as we shall discuss in the next section, there are better measures of the importance of x_{ji}. It is likely that $\lambda_0 > \lambda_1$, since the number of choices in a group increases rather than decreases over time. In addition, the presence of a choice $j \to i$ should increase the tendency for a choice $i \to j$ to form; hence, $\lambda_0 + \mu_0 \geq \lambda_0$, or $\mu_0 \geq 0$. Similarly, the presence of a choice $j \to i$ should decrease the tendency for a choice $i \to j$ to disappear. Then $\lambda_1 + \mu_1 \leq \lambda_1$, or $\mu_1 \leq 0$, but the restriction $\lambda_1 + \mu_1 \geq 0$ (since $\lambda_1 + \mu_1$ is a probability) forces $\mu_1 \geq -\lambda_1$.

To summarize, we expect the four parameters in question to obey the following inequalities:

$$\begin{aligned} \lambda_0 &\geq \lambda_1 \\ \mu_0 &\geq 0 \\ -\lambda_1 \leq \mu_1 &\leq 0 \ . \end{aligned} \tag{6}$$

There are g(g-1) relationships in a social network of g individuals, and g(g-1) pairs of transition rates (5) for the reciprocity model. But for all these pairs, the rates for the relationship between i and j depend only on the relationship between i and j. Hence, if we define

$$D_{ij}(t) = (X_{ij}(t), X_{ji}(t)) \tag{7}$$

as the dyad for the pair of nodes (i,j), each dyad is independent of all other dyads. $D_{ij}(t)$ is, of course, a stochastic process, in fact a Markov chain, with a state space $\mathcal{D}$ containing 4 states: a mutual state $D_{ij}(t) = (1,1)$, two asymmetric states $D_{ij}(t) = (1,0)$ or $(0,1)$, and a null state

$D_{ij}(t) = (0,0)$. By the parameterization of the reciprocity model, the entire $\underset{\sim}{X}(t)$ digraph process can be represented as a set of $\binom{g}{2}$ independent dyad processes. Each is a continuous time, 4 state Markov chain with identical infinitesimal generator, $\underset{\sim}{Q}$, the matrix of infinitesimal transition rates between states, given in Table 1.

(Table 1 about here)

Because of the simplicity of the reciprocity model, we can use data on the transitions of pairs to estimate the parameters of the model, λ_0, λ_1, μ_0, μ_1. For a given group observed at one or more points in time, we have a "sample" of $\binom{g}{2}$ independent observations on the dyad process. For example, Katz and Proctor (1959) study pair transitions for an 8th grade classroom of 25 boys and girls. Their original analysis is based on the assumption that each dyad follows a Markov model in discrete time. In the next section we use their data and estimate the reciprocity model parameters.

The moments, $E\{X_{ij}(t)\}$ and $E\{X_{ij}(t)X_{ji}(t)\}$, and the probability transition matrix for the dyad process are very complicated expressions. These quantities depend on the eigenvalues and eigenvectors of the infinitesimal generator shown in Table 1. They are given, along with computational details, in Wasserman (1977b). Fortunately, the steady state, equilibrium probabilities of the process, which are given below, are simple to compute, and easy to comprehend.

Table 1

Infinitesimal generator for the dyad process

		$D_{ij}(t+h)$ STATE (0,0)	(1,0)	(0,1)	(1,1)
$D_{ij}(t)$ STATE	(0,0)	$-2\lambda_0$	λ_0	λ_0	0
	(1,0)	λ_1	$-(\lambda_0+\lambda_1+\mu_0)$	0	$\lambda_0+\mu_0$
	(0,1)	λ_1	0	$-(\lambda_0+\lambda_1+\mu_0)$	$\lambda_0+\mu_0$
	(1,1)	0	$\lambda_1+\mu_1$	$\lambda_1+\mu_1$	$-2(\lambda_1+\mu_1)$

Let

$$\begin{aligned}
\pi_M(t) &= P\{D_{ij}(t) = (1,1)\} && \text{(Mutual)} \\
\pi_{A1}(t) &= P\{D_{ij}(t) = (1,0)\} && \text{(Asymmetric)} \\
\pi_{A2}(t) &= P\{D_{ij}(t) = (0,1)\} && \text{(Asymmetric)} \\
\pi_N(t) &= P\{D_{ij}(t) = (0,0)\} && \text{(Null)}
\end{aligned} \tag{8}$$

be the elements of $\underset{\sim}{\pi}(t)$, the vector of probabilities of the 4 state dyad process. If we let $t \to \infty$, $\underset{\sim}{\pi}(\infty)$ is the vector of equilibrium probabilities, which are

$$\begin{aligned}
\pi_M(\infty) &= \frac{\lambda_0(\lambda_0+\mu_0)}{(\lambda_0+\lambda_1)(\lambda_1+\mu_1) + \lambda_0(\lambda_0+\mu_0+\lambda_1+\mu_1)} \\
\pi_{A1}(\infty) = \pi_{A2}(\infty) &= \frac{\lambda_0(\lambda_1+\mu_1)}{(\lambda_0+\lambda_1)(\lambda_1+\mu_1) + \lambda_0(\lambda_0+\mu_0+\lambda_1+\mu_1)} \\
\pi_N(\infty) &= \frac{\lambda_1(\lambda_1+\mu_1)}{(\lambda_0+\lambda_1)(\lambda_1+\mu_1) + \lambda_0(\lambda_0+\mu_0+\lambda_1+\mu_1)} \, .
\end{aligned} \tag{9}$$

These probabilities can be found by first showing that the dyad process is reversible, that is the distribution of $D_{ij}(t-\tau)$ and $D_{ij}(t+\tau)$ are equal for all t and $\tau < t$. Equilibrium probabilities are then simply found from the set of reversibility equations (again, see Wasserman, 1977b, for more details).

III.2 Popularity Model

The popularity model has transition rates

$$\begin{aligned} \lambda_{0ij}(\underset{\sim}{x},t) &= \lambda_0 + \pi_0 x_{+j} \\ \lambda_{1ij}(\underset{\sim}{x},t) &= \lambda_1 + \pi_1 x_{+j} \quad . \end{aligned} \tag{10}$$

Individual i's choice or non-choice of individual j depends only on the number of individuals who choose j, x_{+j}, the indegree of j. As with the reciprocity model, the transition rates are time homogeneous, making the network process stationary in time. There are also g(g-1) pairs of rates with the popularity model, but for a fixed j, the (g-1) pairs for various i are identical. That is, each individual in the group has a unique pair of transition rates that are constant for all choices to be made of him or her. Thus, there are only g distinct pairs of rates. Also note that the choices made of individual j do not depend on the number of choices received by the other group members.

The parameters λ_0 and λ_1 are again measures of the overall rate of change for the group. The popularity parameters π_0 and π_1 measure the importance of the "popularity" of individual j on the relationships from group members to individual j. Since a large indegree should increase the tendency for individual j to be chosen by individual i, π_0 is undoubtedly positive and unrestricted in magnitude. Now consider π_1. A large indegree should decrease the chance of choices disappearing. Hence $\pi_1 < 0$. However, π_1 is restricted in size. Since $(\lambda_1 + \pi_1 x_{+j})$ must be positive, we can conclude that the condition

$$\lambda_1 + (g-1)\pi_1 > 0$$

must hold; consequently,

$$-\lambda_1/(g-1) < \pi_1 < 0 \ .$$

Let

$$\underset{\sim}{X}_{.j}(t) = (X_{1j}(t),\ X_{2j}(t),\ldots,\ X_{gj}(t))' \qquad (11)$$

be the j^{th} column process of the digraph process $\underset{\sim}{X}(t)$. Each of the g column processes has 2^{g-1} states, consisting of all possible zero-one vectors of length g, with one entry, $X_{jj}(t)$, fixed at zero. By the parameterization (10) of the popularity model, the stochastic processes $\{\underset{\sim}{X}_{.j}(t)\}$ are independent and identically distributed. The entire digraph process can be represented as g independent column processes consisting of the columns of $\underset{\sim}{X}(t)$, each a continuous time, 2^{g-1} state Markov chain. The infinitesimal generator for the column process, with $g = 4$, is given in Table 2. For simplicity, we have written $\underset{\sim}{X}_{.j}(t)$ as a vector of length 3, ignoring the single zero entry in the vector.

(Table 2 about here)

The solid lines in the generator shown in Table 2 group together those realizations of the column process with equal number of choices. Notice the block-diagonal appearance of the generator. This block-diagonal feature of the generator is apparent for all g. One can easily imagine the structure of the generator for general g: $\lambda_0 + (k-1)\pi_0$ immediately above the diagonal, and $\lambda_1 + (k+1)\pi_1$ immediately below the diagonal, in blocks, where k is the number of ones in the vectors contained in the blocks. The remaining off-diagonal elements in the matrix are zero, demonstrating the "birth-and-death" nature of the column process.

The popularity model is analytically identical to a model for expansiveness with parameterization

$\underset{\sim}{X}_{.j}(t+h)$

$_j(t+h)$

	(0,0,0)	(0,0,1)	(0,1,0)	(1,0,0)	(0,1,1)	(1,0,1)	(1,1,0)	(1,1,1)
(0,0,0)	$-3\lambda_0$	λ_0	λ_0	λ_0	0	0	0	0
(0,0,1)	$\lambda_1+\pi_1$	$-2(\lambda_0+\pi_0)$ $-(\lambda_1+\pi_1)$	0	0	$\lambda_0+\pi_0$	$\lambda_0+\pi_0$	0	0
(0,1,0)	$\lambda_1+\pi_1$	0	$-2(\lambda_0+\pi_0)$ $-(\lambda_1+\pi_1)$	0	$\lambda_0+\pi_0$	0	$\lambda_0+\pi_0$	0
(1,0,0)	$\lambda_1+\pi_1$	0	0	$-2(\lambda_0+\pi_0)$ $-(\lambda_1+\pi_1)$	0	$\lambda_0+\pi_0$	$\lambda_0+\pi_0$	0
(0,1,1)	0	$\lambda_1+2\pi_1$	$\lambda_1+2\pi_1$	0	$-(\lambda_0+2\pi_0)$ $-2(\lambda_1+2\pi_1)$	0	0	$\lambda_0+2\pi_0$
(1,0,1)	0	$\lambda_1+2\pi_1$	0	$\lambda_1+2\pi_1$	0	$-(\lambda_0+2\pi_0)$ $-2(\lambda_1+2\pi_1)$	0	$\lambda_0+2\pi_0$
(1,1,0)	0	0	$\lambda_1+2\pi_1$	$\lambda_1+2\pi_1$	0	0	$-(\lambda_0+2\pi_0)$ $-2(\lambda_1+2\pi_1)$	$\lambda_0+2\pi_0$
(1,1,1)	0	0	0	0	$\lambda_1+3\pi_1$	$\lambda_1+3\pi_1$	$\lambda_1+3\pi_1$	$-3(\lambda_1+3\pi_1)$

Table 2. Infinitesimal Generator for Popularity Model and Column Process, g=4.

$$\lambda_{0ij}(\underset{\sim}{x},t) = \lambda_0 + \varepsilon_0 x_{i+}$$
$$\lambda_{1ij}(\underset{\sim}{x},t) = \lambda_1 + \varepsilon_1 x_{i+} \quad . \tag{12}$$

The expansiveness model reduces $\underset{\sim}{X}(t)$ to a set of g independent and identically distributed <u>row processes</u> consisting of the rows of $\underset{\sim}{X}(t)$. These row processes are mathematically equivalent to the column processes, with ε_0 and ε_1, the parameters measuring the importance of the willingness to make many choices, replacing π_0 and π_1 in all formulas.

As with the reciprocity model, computations of moments and probability transition matrix for the column (or row) process are quite difficult. However, by examining the indegree process, derived from the column process we can easily compute the equilibrium probabilities of $\underset{\sim}{X}_{\cdot j}(t)$.

Define

$$I_j(t) = \sum_i X_{ij}(t) \tag{13}$$

as the <u>indegree process</u>, the sum of the number of ones in the j^{th} column process at time t. $I_j(t)$ is a continuous time birth-and-death process, since by assumption an indegree can only increase by 1, decrease by 1, or remain the same in a small interval of time. Consequently, we know that the equilibrium probabilities of the indegree process are

$$P\{I_j(\infty) = k\} = \frac{\gamma_k}{\Sigma\gamma_k}, \; k = 0, 1, \ldots, (g-1),$$

where (14)

$$\gamma_k = \binom{g-1}{k} \prod_{j=0}^{k-1} \frac{(\lambda_0 + j\pi_0)}{(\lambda_1 + (j+1)\pi_1)}$$

(see Karlin and Taylor, 1975, page 137).

Conditional on k, the number of ones in a column vector, every one of the $\binom{g-1}{k}$ vectors with k ones and (g-1-k) zeros are equally likely. The probability mass function of the column process is just a multiple of the probability mass function of the indegree process; specifically,

$$P\{\underset{\sim}{X}_{.j}(t) = \underset{\sim}{x}_{.j}\} = \frac{1}{\binom{g-1}{k}} P\{I_j(t) = k\} \quad . \tag{15}$$

Letting $t \to \infty$ in equation (15) yields

$$\begin{aligned} P\{\underset{\sim}{X}_{.j}(\infty) = \underset{\sim}{x}_{.j}\} &= \frac{\gamma_k}{\binom{g-1}{k} \Sigma\gamma_k} \\ &= \prod_{j=0}^{k-1} \frac{(\lambda_0 + j\pi_0)}{(\lambda_1 + (j+1)\pi_1)} \end{aligned} \tag{16}$$

for $k = 0,1,\ldots,(g-1)$, where $\underset{\sim}{x}_{.j}$ has k ones and (g-1-k) zeros, as the equilibrium probabilities of $\underset{\sim}{X}_{.j}$. The column process can be analyzed by means of the simple indegree process $I_j(t)$.

IV. Fitting Stochastic Models to Social Networks

It would be very misleading if the reader were left with the impression that the purpose of this research is simply to propose stochastic models for social networks. This is far from the truth. The worth of these models lies not in their theoretical elegance, but in how well they fit the data. Consequently, the data analysis problems that arise in the evaluation of these models is at least as important as their mathematical analysis.

Wasserman (1977b) shows that complete evaluation of stochastic models for networks is very difficult, primarily because one must estimate parameters from a continuous time model with only discrete data. This problem arises because an investigator can only observe a social network at a few discrete points in time. To adequately analyze a continuous time model, one needs a continuous record of the history of the group. We comment first on data collection schemes in this section.

An important aspect of this new methodology is to devise "good" estimation procedures for the parameters of the proposed models when one does not have a continuous record of the group. A start in this direction has been made by Wasserman (1977d). We will discuss parameter estimation for the reciprocity model and the popularity model in this section.

IV.1 Data Collection

The primary assumption in this research is that a social network is an evolving entity. Consequently, if we are to study how a social network changes over time, we need to observe and gather data on the group at several points in time. Moreover, at the very least as shown by Singer and Spilerman (1976), data should be collected longitudinally with the time between repeats and the number of repeats systematically varied.

Suppose that we have observations on one of the processes derived from the $\underset{\sim}{X}(t)$ digraph process mentioned in the previous section. For example, if we consider the dyad process, one observation on the group gives us $\binom{g}{2}$ independent observations on the dyad process, $D_{12}(t)$, $D_{13}(t), \ldots,$ $D_{1g}(t)$, $D_{23}(t), \ldots,$ $D_{g-1,g}(t)$, each a 4-state continuous time Markov chain. Further suppose that this set of observations on the derived process is observed at times $t = t_0, t_1, \ldots, t_n$, where the t_j are distinct positive numbers.

If $n = 0$, we have a single set of observations on the derived process and are not able to form a transition matrix since, a calculation that requires observations taken at two points in time. Estimation using this single set of observations is based solely on the probability distribution of the derived process, or if the group is near steady-state, on the equilibrium probabilities, such as those given in equation (9) for the dyad process.

If $n \to \infty$, we essentially have a continuous record of the process, in the sense that $t_{j+1} - t_j \to 0$. Usually t_∞ is a finite time, say τ. A continuous record is the complete history of the process over the interval (t_0,τ), including all the transitions from state to state (the discrete skeleton) and all the waiting times in each state. Billingsley (1961) discusses estimation of the parameters of a Markov chain assuming a continuous record of the process. This is an ideal situation. Unfortunately, it is quite rare for an investigator to have such complete data on the evolution of a group. The estimation problems that arise when one does not have a continuous record are quite complex, and many are still unsolved. Perhaps future network investigators will make an effort to gather continuous records, thus easing the mathematical burdens for analysis.

If n is finite and nonzero, we can define a set of empirical probability transition matrices (see Anderson and Goodman, 1957, or Wasserman, 1977b). For moderate n, there is a substantial number of estimated transition matrices, $\binom{n}{2}$ to be exact, that can be used to determine whether the data adhere to the assumptions of a continuous time Markov chain. This is the most common longitudinal data collection scheme. The group is observed at several points in time, perhaps equally spaced, perhaps not. We first discuss how one "fits" the reciprocity model to data consisting of only one observation on the group, and second, to data consisting of two or more observations.

IV.2 Estimation of Reciprocity Model Parameters

Suppose that we have a single observation of some group, and we form a sociomatrix $\underset{\sim}{x}$. Throughout this section, we shall assume that the reciprocity model is operating, with transition rates (5). By the assumptions of the model, the matrix $\underset{\sim}{x}$ consists of $\binom{g}{2}$ independent dyads $D_{ij}(t)$, each a continuous time Markov chain. Since the labeling of the nodes in a directed graph is arbitrary, we cannot distinguish a (1,0) asymmetric dyad from a (0,1) asymmetric dyad -- we only know that the dyad in question is neither mutual or null. Consequently, the information in $\underset{\sim}{x}$ can be summarized by three sufficient statistics:

$$
\begin{aligned}
&1)\quad M(t) = \sum_{i>j} x_{ij}x_{ji} = \text{number of mutuals} \\
&2)\quad A(t) = \sum_{i>j} [(1-x_{ij})x_{ji} + x_{ij}(1-x_{ji})] = \text{number of asymmetrics} \qquad (17) \\
&3)\quad N(t) = \sum_{i>j} (1-x_{ij})(1-x_{ji}) = \text{number of nulls.}
\end{aligned}
$$

The three statistics obey the restriction

$$M(t) + A(t) + N(t) = \binom{g}{2} \quad . \qquad (18)$$

We estimate the four unknown parameters of the reciprocity model

$(\lambda_0, \lambda_1, \mu_0, \mu_1)$ by maximizing the likelihood function of the parameters given the single sociomatrix $\underset{\sim}{x}$. For computational ease we use the steady state values of π_M, π_A, π_N, the probabilities of the states of the dyad process, given in equation (9). Wasserman (1977b) gives a detailed examination of the likelihood function, and hence, we shall only briefly comment on some of its features.

While the likelihood is a function of four parameters, it contains only 2 "pieces of information", M(t) and N(t), since A(t) is determined by knowledge of g in equation (18). Consequently, we can only estimate 2 functions of the parameters. These two functions are the "change" ratios:

$$\theta_1 = (\lambda_0+\mu_0)/(\lambda_1+\mu_1)$$
$$\theta_2 = \lambda_0/\lambda_1 \tag{19}$$

The first ratio

$$\theta_1 = \frac{\lambda_0+\mu_0}{\lambda_1+\mu_1} \approx \frac{P\{X_{ij}(t+h) = 1 \mid X_{ij}(t) = 0, X_{ji}(t) = 1\}}{P\{X_{ij}(t+h) = 0 \mid X_{ij}(t) = 1, X_{ji}(t) = 1\}} \tag{20}$$

is the ratio of the probabilities of change, in a small unit of time, given the <u>presence</u> of a reciprocated choice. For example, if $\theta_1 = 6$, then in the presence of $j \to i$, a non-choice from i to j is 6 times more likely to change into a choice than a choice to a non-choice, in a small interval of time. Similarly,

$$\theta_2 = \frac{\lambda_0}{\lambda_1} \approx \frac{P\{X_{ij}(t+h) = 1 \mid X_{ij}(t) = 0, X_{ji}(t) = 0\}}{P\{X_{ij}(t+h) = 0 \mid X_{ij}(t) = 1, X_{ji}(t) = 0\}} \tag{21}$$

has an equivalent interpretation, but in the <u>absence</u> of a reciprocated choice. We expect the ratio θ_1 to be larger than the ratio θ_2 since reciprocated arcs should increase the chance of $i \to j$ relative to the chance of $i \not\to j$.

The maximum likelihood estimates of θ_1 and θ_2, found by Wasserman (1977b) are:

$$\hat{\theta}_1 = \frac{2M(t)}{A(t)}$$
$$\hat{\theta}_2 = \frac{A(t)}{2N(t)} \quad . \tag{22}$$

At the end of this section, we discuss an example, evaluating $\hat{\theta}_1$ and $\hat{\theta}_2$ for the group under study.

There are two other ratios that are more interesting than the change ratios θ_1 and θ_2. These ratios are

$$\kappa_1 = (\lambda_0+\mu_0)/\lambda_0$$
$$\kappa_2 = (\lambda_1+\mu_1)/\lambda_1 \tag{23}$$

and directly measure the importance of reciprocated arcs. For example, with

$$\kappa_1 \approx \frac{P\{X_{ij}(t+h) = 1 \mid X_{ij}(t) = 0,\ X_{ji}(t) = 1\}}{P\{X_{ij}(t+h) = 1 \mid X_{ij}(t) = 0,\ X_{ji}(t) = 0\}} \tag{24}$$

the numerator and denominator differ only in the presence or absence of the choice $j \rightarrow i$. A value for 3 for κ_1 indicates that a change in a small interval of time from a non-choice to a choice is 3 times more likely in the presence of $j \rightarrow i$ than in the absence of $j \rightarrow i$. The ratio κ_2 in a similar way measures the effect of a reciprocated arc on the change from a choice to a non-choice. We suspect that $\kappa_1 > 1$ and $\kappa_2 < 1$, but unfortunately, these "reciprocity" ratios cannot be estimated via maximum likelihood with only one observation on the digraph.

The ratio of θ_1 to θ_2 or κ_1 to κ_2 is also an interesting quantity and more informative than considering θ_1 and θ_2 separately. We have

$$\frac{\theta_1}{\theta_2} = \frac{\kappa_1}{\kappa_2} \approx$$

$$\frac{P\{X_{ij}(t+h) = 1 | D_{ij}(t) = (0,1)\}}{P\{X_{ij}(t+h) = 0 | D_{ij}(t) = (1,1)\}} \bigg/ \frac{P\{X_{ij}(t+h) = 1 | D_{ij}(t) = (0,0)\}}{P\{X_{ij}(t+h) = 0 | D_{ij}(t) = (1,0)\}} =$$

$$\frac{P\{X_{ij}(t) = 0,\ X_{ij}(t+h) = 1 | X_{ji}(t) = 1\}}{P\{X_{ij}(t) = 1,\ X_{ij}(t+h) = 0 | X_{ji}(t) = 1\}} \bigg/ \frac{P\{X_{ij}(t) = 0,\ X_{ij}(t+h) = 1 | X_{ji}(t) = 0\}}{P\{X_{ij}(t) = 1,\ X_{ij}(t+h) = 0 | X_{ji}(t) = 0\}} \qquad (25)$$

an odds ratio. The ratio gives the increase in the odds of a new choice $i \to j$ coming into existence during the interval $(t, t+h)$ due to the presence of $j \to i$. With no reciprocity effect, $\log(\theta_1/\theta_2) = 0$; if $\log(\theta_1/\theta_2) > 0$, a positive reciprocity effect is present implying that reciprocated choices increase the odds of new choices.

We now consider the data analyzed by Katz and Proctor (1959) mentioned in section 3. The dyad censuses, the statistics $M(t)$, $A(t)$, $N(t)$, are given in Table 3. The data are remarkably constant, which may indicate that the group has reached an equilibrium. Of course, data on the transitions of pairs are necessary to completely verify this. We calculate $\hat{\theta}_1$, $\hat{\theta}_2$, and $\log(\hat{\theta}_1/\hat{\theta}_2)$ for each of the four time points in Table 4.

(Tables 3 and 4 about here)

The four estimates of θ_1 differ from the four estimates of θ_2. The ratio of probabilities is about 6 or 7 times greater in the presence of $j \to i$ than in its absence. The log odds ratios also show a positive reciprocity effect. Also note that all the θ_1 and θ_2 estimates are less than unity. This implies that a change from a choice to a non-choice is more probable than a change from a non-choice to a choice. This rather

Dyad Census	t_1 September	t_2 November	t_3 January	t_4 May
M(t)	15	13	14	16
A(t)	45	46	47	43
N(t)	240	241	239	241
$\binom{g}{2}$	300	300	300	300

Table 3. Dyad Censuses for 8th Grade Classroom from Katz and Proctor (1959).

Estimate	t_1 September	t_2 November	t_3 January	t_4 May
$\hat{\theta}_1$	0.677	0.565	0.596	0.744
$\hat{\theta}_2$	0.094	0.095	0.098	0.089
$\log(\hat{\theta}_1/\hat{\theta}_2)$	0.851	0.774	0.784	0.922

Table 4. Change and log odds ratios for Katz and Proctor Data

strange result is due to the large number of null relations in the group.

A study of the change ratios, θ_1 and θ_2, and the log odds ratio, $\log(\theta_1/\theta_2)$, over many groups would be quite informative. Unfortunately, with only one observation on the group we cannot estimate the four parameters of the reciprocity model or even the reciprocity ratios κ_1 and κ_2.

We now consider the situation in which we have more than a single observation on the group, again assuming that the reciprocity model is operating. We first assume that we have two observations on the group, taken at times t_1 and $t_2 > t_1$, and sociomatrices $\underset{\sim}{X}(t_1)$ and $\underset{\sim}{X}(t_2)$. We examine each of the $\binom{g}{2}$ dyad pairs $(X_{ij}(t_1), X_{ji}(t_1))$ and $(X_{ij}(t_2), X_{ji}(t_2))$, and determine the state of D_{ij} at times t_1 and t_2. These dyad transitions can be arranged in a 4 x 4 table, with rows corresponding to the dyad state at time t_1, and columns to the t_2 state. As an example, consider Table 5, taken from Katz and Proctor (1959), giving the transitions from September to November of the dyadic relations in the eight grade classroom under investigation. The row and column margins of the table are identical to the dyad censuses for these two months, shown in Table 3.

(Table 5 about here)

We let $\underset{\sim}{T}$ denote such a matrix of transitions, with entries $(t_{k\ell})$, where the subscripts k and ℓ are defined as follows:

$$\begin{aligned} k,\ell &= 1 = \text{Null } (0,0) \\ k,\ell &= 2 = \text{Asymmetric } (1,0) \\ k,\ell &= 3 = \text{Asymmetric } (0,1) \\ k,\ell &= 4 = \text{Mutual } (1,1). \end{aligned} \qquad (26)$$

		NOVEMBER			
		Null	Asymmetric	Mutual	
SEPTEMBER	Null	217	22	1	240
	Asymmetric	21	Same direction 17 Reversed 2	5	45
	Mutual	3	5	7	15
		241	46	13	300

Table 5. Dyad Transitions from September to November for Katz and Proctor data.

The entry t_{14} is the number of nulls that become mutuals, $t_{23} + t_{32}$, the number of asymmetric dyads that reverse direction, etc. The elements of $\underset{\sim}{T}$ are sufficient statistics for the model parameters.

The likelihood function of the four unknown parameters given $\underset{\sim}{T}$ depends on the conditional probabilities of the dyad states at time t_2 given the states at t_1, and the probabilities of each state at the first observation. The conditional probabilities are the elements of the probability transition matrix of the dyad process, and as mentioned in section III, are very complicated, non-linear expressions, being sums of exponential functions. The probabilities of each state at the first observation are the $\underset{\sim}{\pi}(t_1)$ probabilities, defined in equation (8). Because these probabilities are so complicated, we will not give the likelihood function here, referring the reader to Wasserman (1977b) for details.

Now suppose we have more than 2 observations of the group. We denote the sociomatrices by $\underset{\sim}{X}(t_1)$, $\underset{\sim}{X}(t_2)$, ..., $\underset{\sim}{X}(t_n)$, where n = number of observations. We initially form (n-1) matrices $\underset{\sim}{T}_m$, m = 1, 2, ..., (n-1), where the elements of $\underset{\sim}{T}_m = (t_{k\ell m})$ give the transitions of the dyads at time t_m to new states at time t_{m+1}. The dyad process is stationary so that the probability transition matrix for the dyads depends only on $t_{m+1} - t_m$. The likelihood function is

$$L(\underset{\sim}{\theta}|\underset{\sim}{X}(t_1) = \underset{\sim}{x}_1, \underset{\sim}{X}(t_2) = \underset{\sim}{x}_2, \ldots, \underset{\sim}{X}(t_n) = \underset{\sim}{x}_n) =$$
$$[\pi_M(t_1)^{M(t_1)}\pi_A(t_1)^{A(t_1)}\pi_N(t_1)^{N(t_1)}]. \prod_{m=1}^{n-1} \prod_{k,\ell} p_{kl}(t_{m+1} - t_m)^{t_{k\ell m}} \tag{27}$$

where $\underset{\sim}{\theta} = (\lambda_0, \lambda_1, \mu_0, \mu_1)'$ and

$$p_{k\ell}(t_{m+1} - t_m) = P\{D_{ij}(t_{m+1}) \text{ in state } \ell | D_{ij}(t_m) \text{ in state } k\} \tag{28}$$

are elements of the probability transition matrix. When we sample at regular intervals, so that $t_2-t_1 = t_3-t_2 = \ldots = t_n-t_{n-1} = \tau$, the likelihood (27) reduces to

$$L(\underset{\sim}{\theta}|\underset{\sim}{x}_1, \underset{\sim}{x}_2,\ldots, \underset{\sim}{x}_n) =$$

$$\pi_M(t_1)^{M(t_1)}\pi_A(t_1)^{A(t_1)}\pi_N(t_1)^{N(t_1)} \cdot \prod_{k,\ell} p_{kl}(\tau)^{\sum_m t_{k\ell m}} \tag{29}$$

With this equidistant sampling, we can pool all the dyad transitions across time points, to form a "super-matrix" of transitions $\underset{\sim}{T} = \sum_m \underset{\sim}{T}_m$, giving us a sample size of $(n-1)\binom{g}{2}$ dyads.

To estimate the four parameters of the reciprocity model, we can differentiate the likelihood function (27) to obtain a system of four equations in four unknowns. Unfortunately, because of the complexity of the derivatives, direct solution is virtually impossible. We must rely on either a graphical exploration of the likelihood function in five dimensions, or an approximate solution obtained via a Newton-Raphson-type iterative algorithm to find the maximum likelihood parameter estimates. We will illustrate the former.

We first compute $\hat{\underset{\sim}{P}}(t_i-t_j)$, an empirical probability transition matrix for the dyad process. If we have $n > 2$ observations on the group, there are $\binom{n}{2}$ $\hat{\underset{\sim}{P}}$ matrices, each of which can be studied further. We then utilize recent research of Singer and Spilerman (1974, 1976) on the embeddability of empirical probability transition matrices as transition matrices for continuous time Markov chains. We compute estimates of the infinitesimal generator for the process

$$\hat{\underset{\sim}{Q}} = \frac{1}{t_i - t_j} \hat{\underset{\sim}{P}}(t_i - t_j), \quad i > j, \tag{30}$$

for as many of the $\hat{\underset{\sim}{P}}$'s as possible. Since the unknown parameter $\underset{\sim}{\theta}$ is a one-to-one function of the elements of the infinitesimal generator $\underset{\sim}{Q}$ given in Table 1, we can obtain a set of parameter estimates $\{\hat{\underset{\sim}{\theta}}\}$ that are hopefully near the true maximum of the likelihood. We evaluate L at each of the $\hat{\underset{\sim}{\theta}}$'s and explore the likelihood in the vicinity of the $\hat{\underset{\sim}{\theta}}$ that gives the largest value of L, $\hat{\underset{\sim}{\theta}}^*$, by using finer and finer grids with center at $\hat{\underset{\sim}{\theta}}^*$.

An example will help illustrate. Katz and Proctor sample their eighth grade classroom in September, November, January, and May. September, November, and January are roughly equidistant in time, as are September, January, and May. We shall use the September-November and September-January transition matrices to estimate $\underset{\sim}{\theta}$. These matrices are given in Table 6, where we have lumped together the asymmetric states for simplicity.

(Table 6 about here)

Estimates of the theoretical infinitesimal generator of the dyad process, $\hat{\underset{\sim}{Q}}$, are given in Table 7, along with the theoretical generator $\underset{\sim}{Q}$. We let t_0 = September, t_1 = November, t_2 = January, and define $t_1 - t_0 = 1$ time unit, equivalent to 3 months. Because of the length of the first half of the school year, September-January is a period of 4½ months, thus $t_2 - t_1 = 1.5$ time units. These time intervals, 1 and 1.5, were used as divisors for the calculation of the $\hat{\underset{\sim}{Q}}$ matrices in Table 7. Except

$\hat{\underset{\sim}{P}}$ (November - September)

NOVEMBER

SEPTEMBER	Null	Asymmetric	Mutual
Null	0.467	0.333	0.200
Asymmetric	0.111	0.422	0.467
Mutual	0.004	0.092	0.904

$\hat{\underset{\sim}{P}}$ (January - September)

JANUARY

SEPTEMBER	Null	Asymmetric	Mutual
Null	0.400	0.200	0.400
Asymmetric	0.133	0.333	0.534
Mutual	0.008	0.121	0.871

Table 6. Empirical Transition Matrices from Katz and Proctor data.

for the first row of $\hat{\underset{\sim}{Q}}$, the two estimated generators are quite similar. The difference in the first rows is probably due to the large number of nulls that changed to mutuals from September to January, as seen in Table 6.

(Table 7 about here)

Setting the calculated elements of $\hat{\underset{\sim}{Q}}$ equal to their theoretical values yields the "rough" parameter estimates

$$\begin{aligned} \hat{\lambda}_0 &= 0.076 \\ \hat{\lambda}_1 &= 0.773 \\ \hat{\mu}_0 &= 0.202 \\ \hat{\mu}_1 &= -0.353 \end{aligned} \tag{31}$$

for September-November, and

$$\begin{aligned} \hat{\lambda}_0 &= 0.077 \\ \hat{\lambda}_1 &= 0.642 \\ \hat{\mu}_0 &= 0.203 \\ \hat{\mu}_1 &= -0.387 \end{aligned} \tag{32}$$

for September-January, remarkably close.

The change and reciprocity ratios for the group calculated from the September-November estimates (31) are:

$$\begin{aligned} \hat{\theta}_1 &= 0.098 \\ \hat{\theta}_2 &= 0.662 \end{aligned} \tag{33}$$

and

$$\begin{aligned} \hat{\kappa}_1 &= 3.658 \\ \hat{\kappa}_2 &= 0.543 \ . \end{aligned} \tag{34}$$

The estimates (33) are virtually identical to those given in Table 4 based

$\hat{\underset{\sim}{Q}}$ September – November

	Null	Asymmetric	Mutual
Null	-0.867	0.814	0.053
Asymmetric	0.278	-1.051	0.773
Mutual	-0.011	0.158	-0.147

$\hat{\underset{\sim}{Q}}$ September – January

	Null	Asymmetric	Mutual
Null	-0.673	0.347	0.326
Asymmetric	0.280	-0.922	0.642
Mutual	-0.017	0.163	-0.146

$\underset{\sim}{Q}$ Infinitesimal Generator

	Null	Asymmetric	Mutual
Null	$-2(\lambda_1+\mu_1)$	$2(\lambda_1+\mu_1)$	0
Asymmetric	$(\lambda_0+\mu_0)$	$-(\lambda_0+\lambda_1+\mu_0)$	λ_1
Mutual	0	$2\lambda_0$	$-2\lambda_0$

Table 7. $\underset{\sim}{Q}$ matrices for Katz and Proctor Data

on single observations. Obviously, the group is near equilibrium since the estimates in Table 4, computed with the steady-state probabilities, are so similar to the estimates (33), computed with no steady-state assumptions. The reciprocity ratios (34) are quite important. A non-choice is nearly 4 times more likely to change into a choice in the presence of a reciprocated arc, and a choice is only half as likely to disappear in the presence of a reciprocated arc, than in its absence. These two reciprocity ratios summarize the effect of a reciprocated choice very nicely.

We now use the point $\hat{\underset{\sim}{\theta}} = (0.076, 0.073, 0.202, -0.353)'$ and explore the likelihood function (27) in the neighborhood of this point in 5-dimensional space. We shall try to determine whether L, or equivalently $\log L$, has a true maximum in the vicinity of $\hat{\underset{\sim}{\theta}}$.

We first compute $\log L$ for a large grid to determine whether the function has a global maximum. We set up a grid using the values

$\lambda_0 = 0.10, 0.25, 0.50, 0.75, 1.00$

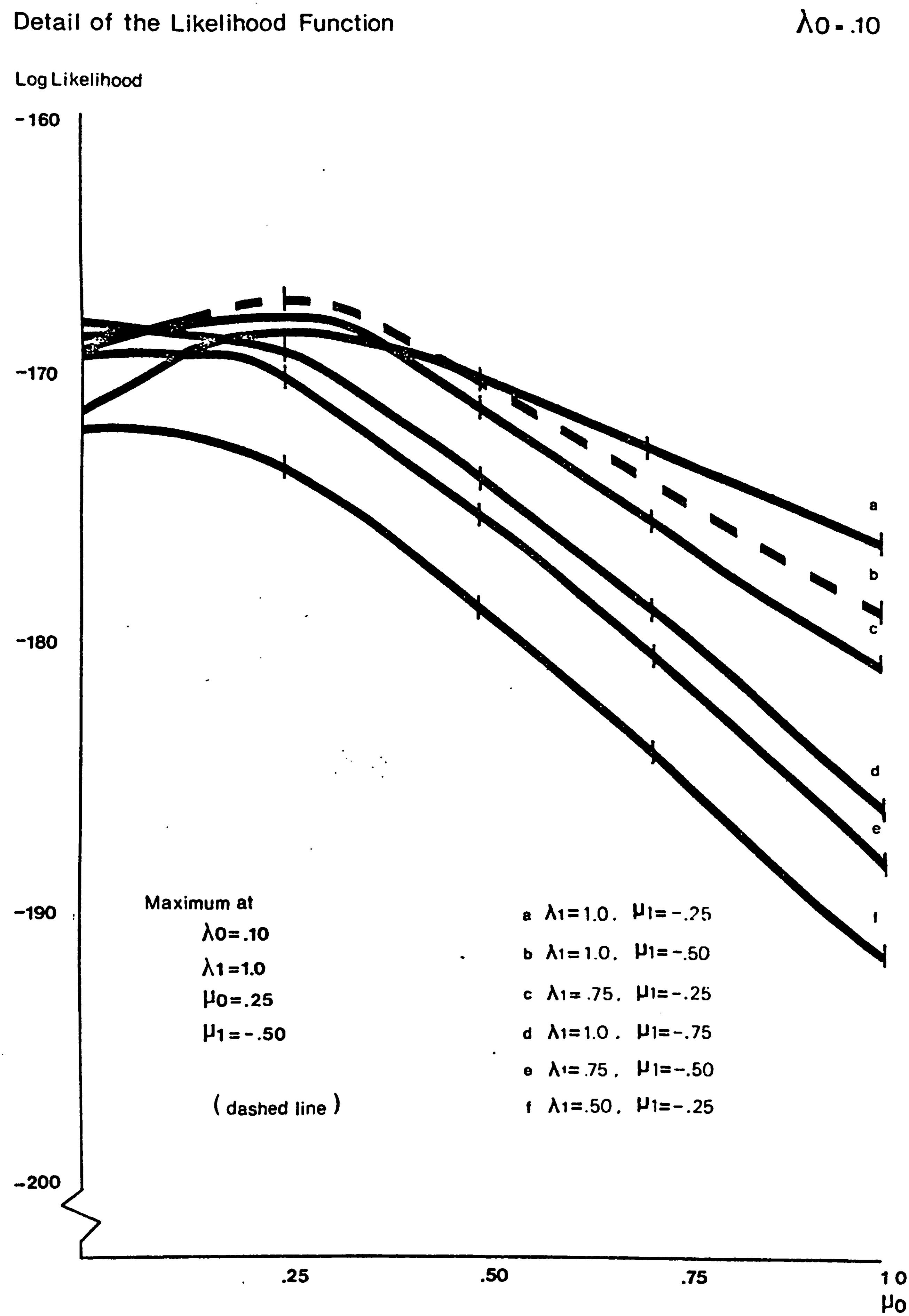
$\mu_0 = 0, 0.25, 0.50, 0.75, 1.00$

$(\lambda_1,\mu_1) = (0.50,-0.25), (0.75,-0.25), (1.0,-0.25),$
$(0.75,-0.50), (1.0,-0.50), (1.0,-0.75)$.

Remember that $\lambda_1 + \mu_1 > 0$, so that these two parameters cannot vary freely. After our computations, we see that the function is monotonically decreasing in λ_0, and increases for small μ_0 when λ_0 is small and then monotonically declines for larger μ_0. The function also appears to show a general increase for increasing λ_1 and μ_1. We suspect that $\log L$ has a unique maximum for very small $\lambda_0 < 0.10$, μ_0 near 0.25, λ_1 near

Figure 2

Detail of the Likelihood Function



0.75, and μ_1 near -0.25. This range of values corresponds very closely to $\hat{\underset{\sim}{\theta}}$. Figure 2 gives a detailed view of the surface, with λ_0 fixed at 0.10. The likelihood appears to be rather flat in the vicinity of $\hat{\underset{\sim}{\theta}}$.

(Figure 2 about here)

We make a further exploration of L with a much smaller grid centered at (0.0875, 0.30, 0.85, -0.45). The function has a maximum of -164.4 at (0.075, 0.80, 0.15, -0.40), very near $\hat{\underset{\sim}{\theta}}$. Because of the flatness of log L near this maximum, a continuum of points are all possible maxima. If we examine the range of points with log L > -164.6, we see that the parameters may vary as:

$$\begin{aligned} 0.05 &\le \lambda_0 \le 0.10 \\ 0.70 &\le \lambda_1 \le 1.00 \\ 0.15 &\le \mu_0 \le 0.25 \\ -0.50 &\le \mu_1 \le -0.30 \ . \end{aligned} \tag{35}$$

We can consider the ranges (35) as a "pseudo-confidence" region.

IV.3 Estimation of Popularity Model Parameters

Recall the popularity model, discussed previously, with transition rates dependent upon the indegree of the chosen individual. The model has four parameters, λ_0 and λ_1, measures of the overall rate of change, and π_0 and π_1, measures of the effect of popularity on choices made. With this set of transition rates, given in equation (10), the set of column processes $\{X_{\cdot j}(t)\}$, equation (11), are independent and identically distributed as continuous time, 2^{g-1} state Markov chains. Moreover, the set of birth-and-death indegree processes $\{I_j(t)\}$, equation (13), carry all of the analytic information in the column processes necessary to estimate the model parameters.

As an example, again consider the fraternity studied by Newcomb (1961), and Nordlie (1958). The indegree process for each of the 17 fraternity members is observed for fifteen consecutive weeks, except for week 9. We form indegree censuses for each week, counting the number of fraternity members with indegree 0, indegree 1, etc. The indegree censuses for the fraternity are given in Table 8.

(Table 8 about here)

The data are difficult to comprehend in this tabular form. Figure 3 shows a graphical simplification of these data, box-and-whisker plots (Tukey (1977) or Leinhardt and Wasserman (1978)) for the indegree processes at each week. The box demarcates the quarters of each batch of 17 numbers, while the line drawn through the box marks the median. Whiskers are drawn down to the minimum and up to the maximum. One can imagine dashed lines connecting the quarters and the medians to further simplify our study of the changes in $\{I_j(t)\}$ over time.

(Figure 3 about here)

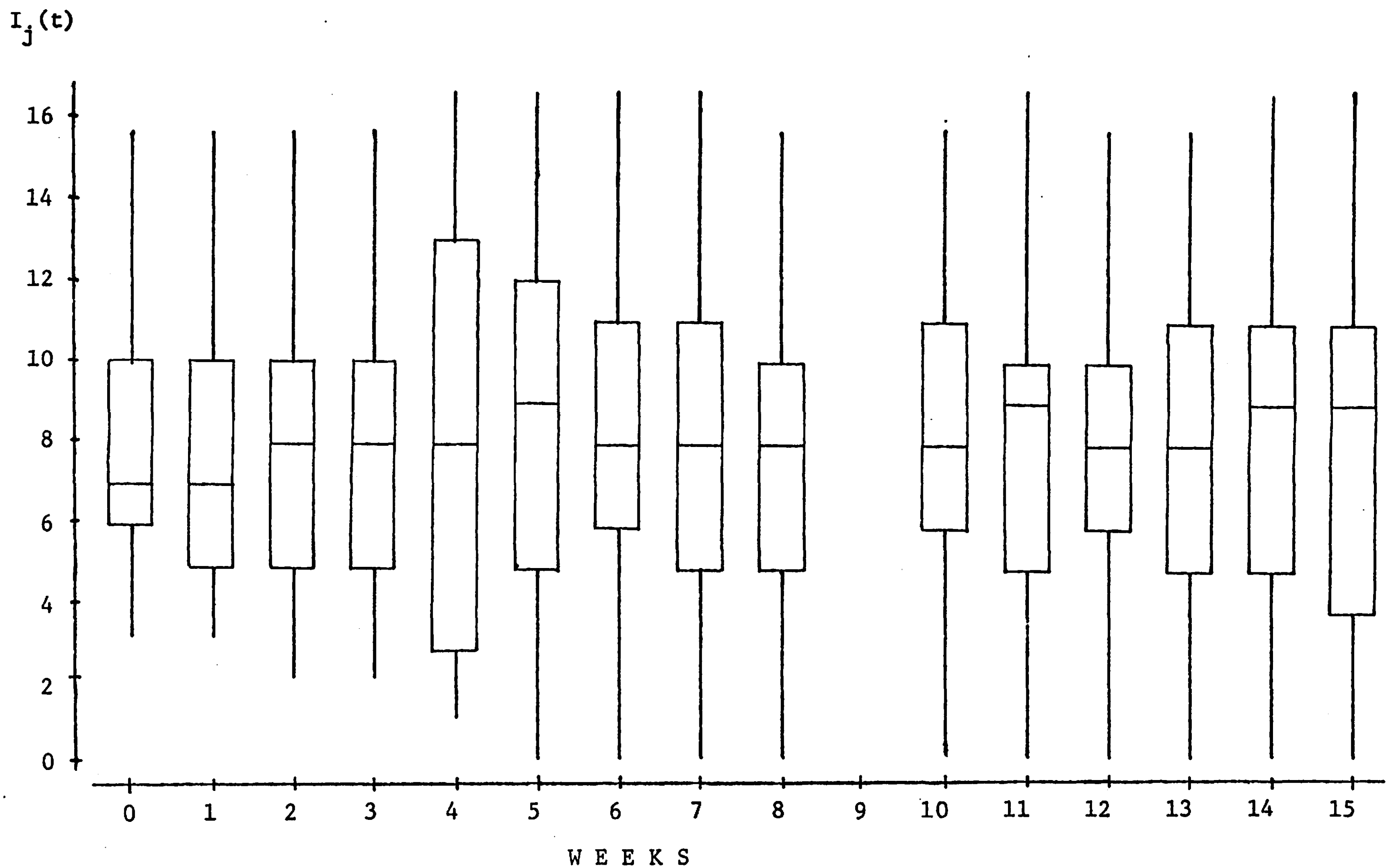
These data show good constancy through time. A change occurs in the $\{I_j(t)\}$ at week 4, increasing the dispersion, but the $\{I_j(t)\}$ settle back into the pattern of the earlier weeks at week 6, although with an increased range. This constancy probably insures that estimates of the popularity model parameters based on single observations of the group will be reasonably stable over time.

INDEGREE CENSUS

# of individuals with indegree of \ WEEKS	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	0	0	0	0	1	1	1	1		1	1	1	1	1	1
1	0	0	0	0	2	1	1	0	1		1	1	1	0	2	1
2	0	0	1	1	1	2	0	2	0		2	0	0	1	0	1
3	1	1	0	2	2	0	1	0	1		0	2	1	1	0	1
4	1	2	1	1	0	0	1	1	1		0	0	1	1	0	1
5	1	2	4	1	0	1	0	1	1		0	1	0	1	2	0
6	5	2	0	0	2	2	2	1	0		1	0	1	0	1	1
7	2	2	1	3	1	1	2	1	2		1	2	3	1	1	1
8	1	1	3	1	1	0	1	2	1		3	0	1	3	1	1
9	1	2	2	3	1	1	1	1	1		2	4	2	2	2	2
10	1	1	1	1	2	2	0	2	4		1	2	2	1	1	0
11	0	1	1	1	1	0	4	1	1		1	1	0	3	2	3
12	1	1	1	0	0	3	1	1	0		0	0	1	0	2	2
13	2	0	1	1	2	2	1	2	2		2	1	1	0	1	0
14	0	0	0	1	1	0	0	0	0		1	1	0	0	0	1
15	1	2	1	1	0	0	0	0	1		1	0	2	2	0	0
16	0	0	0	0	1	1	1	1	0		0	1	0	0	1	1

Table 8. Indegree Censuses for Newcomb's (1961) fraternity.

Figure 3. Indegree Processes for Newcomb's Fraternity.



Suppose we observe the group at one point in time, say t_k, recording data from one of the 16 weekly observations. The $\{I_j(t_k),\ j = 1,2,\ldots,17\}$ can be used to estimate $\underset{\sim}{\phi} = (\lambda_0,\lambda_1,\pi_0,\pi_1)'$ via maximum lieklihood. As with estimation of the reciprocity model parameters, we assume that the group is near equilibrium, and use the steady-state probabilities, γ_k, defined in equation (14), in the likelihood function. If we define the indegree census as

$$C_k(t) = \#\text{ of the } I_j(t) \text{ equal to } k\ , \quad k = 0,1,\ldots,(g-1) \tag{36}$$

then the likelihood function of $\underset{\sim}{\phi}$ is simply

$$L(\underset{\sim}{\phi}|\underset{\sim}{X}(t) = \underset{\sim}{x}) = \prod_{k=0}^{(g-1)} \gamma_k^{\;C_k(t)} \ . \tag{37}$$

Unlike the likelihood $L(\underset{\sim}{\phi}|\underset{\sim}{X}(t) = \underset{\sim}{x})$, we can estimate $\underset{\sim}{\phi}$ directly, if $(g-1) > 4$, but unfortunately, differentiation of (37) and solution of the resulting system of 4 equations in 4 unknowns is difficult, and must be done approximately using computer algorithms.

If we observe the group at two (or more) points in time, we can record indegree transitions, i.e. the number of individuals with indegree i at time t_k and indegree j at time t_ℓ, $t_\ell > t_k$. Probability transition matrices for the indegree process can be quite large, $g \times g$ in size, making their analysis rather difficult. We prefer to work directly with the indegree censuses; fortunately, we can estimate all four parameters even with a single observation on the group.

V. Considerations for Prospective Stochastic Networkers

We social network researchers are entering the stochastic process game at a rather late date. Our colleagues interested in social mobility have been playing the game for nearly twenty-five years, since the classic work of Blumen, Kogan, and McCarthy (1955). Hopefully, we can learn from mobility studies and not duplicate the mistakes made by mobility researchers over the years.

The framework discussed in this paper is unique to the social network discipline because of its allowances for parameterizations. One can define how the social structure of a group influences future structure in many ways. By defining this "operative" social structure using graph-theoretic quantities, one can combine these quantities to form the transition rates of the process. Unfortunately, all but a few models, such as the reciprocity and popularity model discussed here, are quite difficult to manipulate mathematically. The specified transition rates may not allow a decomposition of the network into a set of independent and identical processes (see Mayer, 1977, section X). We can still learn about these models by Monte Carlo simulation of networks evolving as stochastic processes with specified transition rates.

Perhaps the central issue in this research still to be addressed is model evaluation -- how do we determine which model, chosen from some set of models, provides the best fit for the data. Since all of our models are time-homogeneous Markov, we must first verify that the observed realization of the process has fixed transition rates, and that the Markov assumption is valid. Here, some of the ideas of Singer and Spilerman (1976, 1977) can be applied. Once verified, we can ask which of the models under consideration provides the best description of the data.

In addition to these mathematical concerns, we need to answer some sociological questions such as why model parameters might differ across groups and how to theoretically determine functional forms for the λ_{0ij} and λ_{1ij}. The questions can only be answered through empirical sociological research. Hopefully, the ideas and methods presented in this paper will motivate sociologists to consider networks as evolving entities and to study the implications of this assumption.

VI. Acknowledgements

This paper is based on invited presentations given at the Mathematical Social Science Board Advanced Research Symposium on Stochastic Process Models of Social Structure, December 1977, Carnegie-Mellon University, and the International Network for Social Network Analysis Colloquium on New Directions in Structural Analysis, March 1978, University of Toronto. I am indebted to Paul Holland and Samuel Leinhardt for their continued interest in my research and to Thomas Mayer for his diligent reading of my dissertation and for his comments that are reflected in this paper.

VII. References

Anderson, T.W. and L.A. Goodman (1957). "Statistical inference about Markov chains." Annals of Mathematical Statistics, 28, 89-110.

Arabie, P., S.A. Boorman, and P.R. Levitt (1978). "Constructing blockmodels: How and why." Journal of Mathematical Psychology, to appear.

Billingsley, P. (1961). Statistical Inferences for Markov Processes. Chicago: The University of Chicago Press.

Blumen, I., M. Kogan, and P.J. McCarthy (1955). The Industrial Mobility of Labor as a Probability Process. Ithaca, N.Y.: Cornell University Press.

Boorman, S.A. and H.C. White (1976). "Social structure from multiple networks. II. Role structures." American Journal of Sociology, 81, 1384-1446.

Holland, P.W. and S. Leinhardt (1970). "A method for detecting structure in sociometric data." American Journal of Sociology, 70, 492-513.

Holland, P.W. and S. Leinhardt (1975). "Local structure in social networks" in Sociological Methodology 1976, edited by D.R. Heise, San Francisco: Jossey-Bass, Inc., 1-45.

Holland, P.W. and S. Leinhardt (1977a). "A dynamic model for social networks." Journal of Mathematical Sociology, 5, 5-20.

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

Karlin, S. and H.M. Taylor (1975). A First Course in Stochastic Processes, Second Edition. New York: Academic Press.

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

Kauffman, S.A. (1969). "Metabolic stability and epigenesis in randomly contructed genetic nets." Journal of Theoretical Biology, 22, 437-467.

Leinhardt, S. and S.S. Wasserman (1978). "Exploratory data analysis: An introduction to selected methods" in Sociological Methodology 1979, edited by K. Schuessler. San Francisco: Jossey-Bass, in press.

Mayer, T.F. (1977). "Parties and networks: Stochastic models for relationship networks." Paper presented at the MSSB Advanced Research Symposium on Stochastic Process Models for Social Structure, 1-2 December, 1977, Carnegie-Mellon University, Pittsburgh, Pennsylvania.

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

Nordlie, P.G. (1958). "A longitudinal study of interpersonal attraction in a natural group setting," Ph.D. dissertation, Department of Social Psychology, University of Michigan.

Singer, B. and S. Spilerman (1974). "Social mobility models for heterogeneous populations." Sociological Methodology 1973-1974, edited by H.L. Costner. San Francisco: Jossey-Bass.

Singer, B. and S. Spilerman (1976). "Representation of social processes by Markov models." American Journal of Sociology, 82, 1-54.

Singer, B. and S. Spilerman (1977). "The representation of multi-wave panel data by non-stationary continuous-time Markov chains." Paper presented at the MSSB Advanced Research Symposium on Stochastic Process Models for Social Structure, 1-2 December, 1977, Carnegie-Mellon University, Pittsburgh, Pennsylvania.

Tukey, J.W. (1977). Exploratory Data Analysis. Reading, Mass.: Addison-Wesley.

Wasserman, S.S. (1977a). "Random directed graph distributions and the triad census in social networks." Journal of Mathematical Sociology, 5, 61-86.

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

Wasserman, S.S. (1977c). "A survey of mathematical models for graphs." Proceedings of the American Statistical Association Social Statistics Section, 860-864.

Wasserman, S.S. (1977d). "A stochastic model for directed graphs with transition rates determined by reciprocity." Technical Report No. 305, University of Minnesota, School of Statistics.

Wasserman, S.S. (1978). "Models for binary directed graphs and their applications." Advances in Applied Probability, 10, to appear.

White, H.C., S.A. Boorman, and R.L. Breiger (1976). "Social structure from multiple networks. I. Blockmodels of roles and positions." American Journal of Sociology, 81, 730-780.