NONRESPONSE MODELS FOR SOCIAL NETWORK

STOCHASTIC PROCESSES

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the Graduate
School of The Ohio State University

By

Deborah J. Rumsey, B.S., M.S., M.S.

****

The Ohio State University

1993

Dissertation Committee:
Elizabeth A. Stasny
Robert Bartoszynski
Douglas E. Critchlow

Approved by

Elizabeth A. Stasny
Adviser

Department of Statistics
To My Family
ACKNOWLEDGMENTS

I would first like to express my gratitude to my adviser, Professor Elizabeth Stasny, for her continual support, inspiration, and friendship during the work on my dissertation, and throughout my time here at Ohio State. I would also like to thank Professors Robert Bartoszynski and Douglas Critchlow for their helpful comments and support, and Professor R.C. Srivastava, who alerted me to the subject of social networks in his stochastic processes course.

Other past professors have been exceptionally influential and inspiring to me during my years of graduate school, encouraging me at important points in my career, and I would like to acknowledge them: Professor Tony Barkauskas, University of Wisconsin-LaCrosse, and Professors Orlyn Edge, Kenneth Berk, and Mike Plantholdt at Illinois State University. Thank you for being such wonderful teachers and researchers, and for believing in me.

I would like to thank some very special friends at OSU for their encouragement, especially Lora Bohn, Joe Pultz, M. Steigerwald, Kit Kilin, Billie Dudley and Myrtle Pfouts. I also want to extend my appreciation to Mike and Gay O’Leary, for their continual support. My deepest expression of thanks goes to my family for their love and their faith in me every step of the way, with special thanks to my mother and father, for providing me with the strength and dedication to achieve this. Without you, this would not have been possible for me; my achievements are your achievements.

And to Kelly, thanks for always being there.
VITA

October 10, 1961  Born - Viroqua, Wisconsin

1984  B.S. in Mathematics, University of Wisconsin-LaCrosse, LaCrosse, Wisconsin

1987  M.S. in Mathematics, Illinois State University, Normal, Illinois

1990  M.S. in Statistics, Ohio State University, Columbus, Ohio

PUBLICATIONS


FIELDS OF STUDY

Major Field: Statistics
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ......................................................... iii  
VITA ................................................................. iv  
LIST OF TABLES .......................................................... viii  
LIST OF FIGURES .......................................................... ix

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II. Social Network Research; the Reciprocity, Popularity, and Expansiveness Models</td>
<td>7</td>
</tr>
<tr>
<td>2.1 Social Network Research</td>
<td>7</td>
</tr>
<tr>
<td>2.2 The Holland-Leinhardt General Model</td>
<td>15</td>
</tr>
<tr>
<td>2.2.1 The Holland-Leinhardt Modelling Framework</td>
<td>16</td>
</tr>
<tr>
<td>2.2.2 Embeddability, Identifiability, and Parameter Estimation</td>
<td>19</td>
</tr>
<tr>
<td>2.3 The Reciprocity Model</td>
<td>23</td>
</tr>
<tr>
<td>2.4 The Popularity and Expansiveness Models</td>
<td>27</td>
</tr>
<tr>
<td>III. Models for Link Nonresponse Under the Reciprocity Model</td>
<td>30</td>
</tr>
<tr>
<td>3.1 The Two-Stage Process for Nonresponse</td>
<td>30</td>
</tr>
<tr>
<td>3.1.1 The Observed Dyad Transition Matrix for Link Nonresponse Under the Reciprocity Model</td>
<td>31</td>
</tr>
</tbody>
</table>
3.1.2 Probabilities for Stages One and Two
Markov Chains ........................................ 37

3.2 Some Models for Link Nonresponse ...................... 42

3.2.1 Model 1: Random Link Nonresponse at
Time $t_2$ Only ........................................ 43

3.2.2 Model 2: Random Link Nonresponse at
Time $t_1$ Only ........................................ 50

3.2.3 Model 3: Nonrandom Link Nonresponse at
Time $t_2$ Only ........................................ 56

3.2.4 Model 4: Nonrandom Link Nonresponse at
Time $t_1$ Only ........................................ 61

3.2.5 Model 5: Random Link Nonresponse at
One or Both Time Periods .............................. 68

3.2.6 Model 6: Nonrandom Link Nonresponse at
One or Both Time Periods .............................. 76

IV. Models for Node Nonresponse Under the Popularity and
Expansiveness Models .................................... 85

4.1 The Two-Stage Process for Node Nonresponse .......... 85

4.1.1 The Observed Indegree Transition Matrix for Node
Nonresponse Under the Popularity Model ........... 88

4.2 Probabilities for Stages One and Two
Markov Chains ........................................... 96

4.3 Some Node Nonresponse Models under the Popularity
Model for the Data Markov Chain .................... 99

4.3.1 Model 1: Random Node Nonresponse at
Time $t_2$ Only (Popularity Model) ................... 101

4.3.2 Model 2: Random Node Nonresponse at
Time $t_1$ Only (Popularity Model) ................... 105

4.3.3 Model 3: Nonrandom Node Nonresponse at
Time $t_2$ Only (Popularity Model) ................... 110
4.3.4 Model 4: Nonrandom Node Nonresponse at Time $t_1$ Only (Popularity Model) .................. 112

4.3.5 Model 5: Random Node Nonresponse at One or Both Time Periods (Popularity Model) ...... 116

4.3.6 Model 6: Nonrandom Node Nonresponse at One or Both Time Periods (Popularity Model) ...... 122

4.4 Some Node Nonresponse Models under the Expansiveness Model for the Data Markov Chain .................. 128

4.4.1 Model 1: Random Node Nonresponse at Time $t_2$ Only (Expansiveness Model) ............... 130

4.4.2 Model 2: Random Node Nonresponse at Time $t_1$ Only (Expansiveness Model) ............... 131

V. Illustration ..................................................................................................................... 134

5.1 Nonresponse Data in the Social Network Setting .......... 134

5.2 Classroom Data with Generated Link Nonresponse ........ 136

5.3 Preliminary Findings on Node Nonresponse Models ...... 142

VI. Conclusions and Future Research ................................................................. 143

BIBLIOGRAPHY ............................................................................................................. 147
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Cell constraints under Model 1</td>
<td>45</td>
</tr>
<tr>
<td>2. Cell constraints under Model 2</td>
<td>51</td>
</tr>
<tr>
<td>3. Cell constraints under Model 3</td>
<td>70</td>
</tr>
<tr>
<td>4. The observed indegree transition matrix for node nonresponse</td>
<td>90</td>
</tr>
<tr>
<td>(g=4)</td>
<td></td>
</tr>
<tr>
<td>5. Blocks with non-zero cell counts with node nonresponse</td>
<td>92</td>
</tr>
<tr>
<td>at time $t_1$ only</td>
<td></td>
</tr>
<tr>
<td>6. Blocks with non-zero cell counts with node nonresponse</td>
<td>92</td>
</tr>
<tr>
<td>at time $t_2$ only</td>
<td></td>
</tr>
<tr>
<td>7. Blocks with non-zero cell counts with node nonresponse</td>
<td>95</td>
</tr>
<tr>
<td>at one or both time periods</td>
<td></td>
</tr>
<tr>
<td>8. Degrees of freedom under node nonresponse Model 1</td>
<td>102</td>
</tr>
<tr>
<td>9. Degrees of freedom under node nonresponse Model 2</td>
<td>106</td>
</tr>
<tr>
<td>10. Degrees of freedom under node nonresponse Model 3</td>
<td>111</td>
</tr>
<tr>
<td>11. Degrees of freedom under node nonresponse Model 4</td>
<td>113</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Digraph of a four-member social network</td>
<td>2</td>
</tr>
<tr>
<td>2. The Q matrix for the reciprocity model</td>
<td>24</td>
</tr>
<tr>
<td>3. The dyad transition matrix after Stage 1</td>
<td>25</td>
</tr>
<tr>
<td>4. The observed dyad transition matrix for link nonresponse</td>
<td>34</td>
</tr>
<tr>
<td>5. The observed dyad transition matrix in partitioned block form</td>
<td>35</td>
</tr>
<tr>
<td>6. Probabilities for elements in each ((k, \ell)) cell of the observed dyad transition matrix</td>
<td>41</td>
</tr>
<tr>
<td>7. Likelihood function for link nonresponse under Model 5 for link nonresponse at one or both time periods</td>
<td>71</td>
</tr>
<tr>
<td>8. Observed indegree transition matrix in partitioned block form ((g=4))</td>
<td>91</td>
</tr>
<tr>
<td>9. Illustration of a block combination configuration under Models 5 or 6</td>
<td>93</td>
</tr>
<tr>
<td>10. Illustration of the block combination configuration for given example</td>
<td>117</td>
</tr>
<tr>
<td>11. The observed outdegree transition matrix for (g=4)</td>
<td>129</td>
</tr>
<tr>
<td>12. The dyad transition matrix for classroom data after Stage 1</td>
<td>137</td>
</tr>
<tr>
<td>13. The generated transition matrix for classroom data under Model 1</td>
<td>138</td>
</tr>
<tr>
<td>14. Parameter estimates for classroom data under Model 1</td>
<td>138</td>
</tr>
<tr>
<td>15. Expected cell counts for classroom data under Model 1 (including link nonresponse)</td>
<td>139</td>
</tr>
</tbody>
</table>
16. Generated transition matrix for classroom data under Model 3...... 140
17. Parameter estimates for classroom data under Model 3............... 141
A social network is a group of individuals and a set of links between certain of those individuals based on some relationship(s) of interest. Social networks are used by scientists in various fields who study the structure of a group of individuals who are linked by one or more relations. The mathematical tool used to represent and study social networks is the digraph: a directed graph whose nodes or vertices represent the individuals in the group, and whose directed edges or arcs represent the existence of a particular relation in a certain direction. For example, suppose we have a group of 4 individuals: Mary, Sue, Bob, and Tom, and the relation we wish to study is friendship. Suppose we have the following situation at a particular point in time: Bob has friendships with Mary and Tom, Tom has a friendship with Sue, Mary is friends with Tom and Bob, and Sue is friends with no one in this group. Figure 1 shows the digraph representing this social network at this point in time:

The digraph for a social network may be summarized in an adjacency matrix $X$, whose elements have the following form:

$$X_{ij} = \begin{cases} 1, & \text{if } v_i \text{ relates to } v_j, \\ 0, & \text{otherwise} \end{cases} \quad (1.1)$$
If we let individuals Bob, Mary, Tom, and Sue be represented by vertices $v_1, v_2, v_3, v_4$, respectively, the adjacency matrix for this example would have the following form:

$$
X = \begin{bmatrix}
0 & 1 & 1 & 0 \\
1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0
\end{bmatrix}.
$$

(1.2)

A major contribution to the area of social network research was made in 1977 by Holland and Leinhardt (see 1977a) who propose a general modelling framework for the study of the evolution of a social network over time; through these models, we can study the long-term structures present in the social network, and measure the importance of certain structures to the process by which the social network structure changes from one time period to the next. These models involve modelling the social network as a stochastic process, where the adjacency matrix is modelled as a finite-state, continuous-time Markov-chain.

In the process of evolution of a social network over time, choices are made by each individual as to the formation, continuation, or dissolution of every relationship involving other members of the group. The modelling of these changes is based on certain structures in the network. Three particular models under the Holland-Leinhardt
framework are proposed by Wasserman (1977b), (1980); these we will study in detail:
1) the reciprocity model, 2) the popularity model, and 3) the expansiveness model.

We can illustrate the basic ideas of these three models with respect to the above example involving the classmates; for example, under the reciprocity model, Tom’s
decision as to whether or not to form a friendship with Sue is modelled to depend on
whether or not Sue considers Tom a friend of hers (i.e. the existence of a reciprocated
relationship). Under the popularity model, Tom’s decision regarding Sue would be
modelled to depend on how many other people in the group consider Sue to be a friend
of theirs (i.e. an indicator of Sue’s popularity is taken into account by Tom). Under the
expansiveness model, Tom’s decision regarding whether or not to form a friendship
with Sue would be modelled as a function of how many friends Tom already has (i.e.
Tom takes his level of expansiveness into account).

One element that has not been addressed regarding social network models is
finding a systematic way to handle missing data, particularly missing data due to
nonresponse. One can imagine that in the process of collecting data on a group of
individuals over one or more time periods, some missing data will occur due to
nonresponse; this has in fact occurred in the literature (see Taba (1955), and Sampson
(1968)).

Until now, the approach used in dealing with missing data due to nonresponse
has been to ignore data from any individual who does not respond at each and every
time period regarding the status of each relationship. But this approach assumes that
nonresponse occurs at random (see, for example, Little and Rubin (1987)) and that we
lose no information about the network by dropping this individual from the data set
apart from having a smaller sample size. This is not likely to be the case, as we can
imagine many situations, including the previous example, in which an individual may
not respond based on some relationship(s) in which he or she may be involved, or
based on some other structural property of the social network. Our models for nonresponse attempt to make use of the observed data to estimate what the states of the missing relationships would have been, had we been able to observe them. In doing this, we can provide more accurate estimates of the social network parameters.

We identify two types of nonresponse mechanisms in social networks: link nonresponse and node nonresponse. Link nonresponse involves information missing for one (or more) distinct relationship(s) in the network, i.e. information about one or more directed edges is missing. Node nonresponse occurs when one (or more) persons is completely missing and information on all links emanating from that person’s node is missing. We propose discrete-time Markov-chain models for link nonresponse and node nonresponse in the framework of the three models for the social networks as proposed by Wasserman (1980). We present both random (to serve as a benchmark) and nonrandom nonresponse models.

In each case, we model the nonrandom nonresponse as a function of factors involving the states of certain subgroups of the social network at one time period or another. These subgroups are determined by the model chosen for the social network; for example, in the reciprocity model, we study pairs of individuals, hence our nonresponse models under the reciprocity model will involve the states that the pairs were in at one of the time periods. Regarding our previous example with the classmates, Tom may be more likely to become a nonrespondent at the next time period regarding his relationship with Sue because at the present time period Tom and Sue are in an asymmetric relationship; in particular, Sue did not reciprocate Tom’s friendship. This would result in one case of link nonresponse. Under the popularity model, Sue may be likely to become a nonrespondent at the next time period by not participating in the survey at that time, since at the present time period she does not have any friends. This would result in node nonresponse.
Because an understanding of some standard social network models is required for readers of this dissertation, we will provide background information on social network models in Chapter 2; in particular, the work of Holland and Leinhardt (1977a), and Wasserman (1977b, 1980) will be discussed in detail. In Chapter 3, we present random and nonrandom nonresponse models for link nonresponse under the reciprocity model; in Chapter 4, we present random and nonrandom nonresponse models for node nonresponse models under the popularity and expansiveness models. In each case, we will present models which allow for missing data at time 2 but not at time 1, models which allow for missing data at time 1 but not at time 2, and finally models which allow for missing data at one or both time periods. (We will see that the third setting is more complicated in terms of the equations involved than the first two settings, because of the increase in the possibilities for the nonresponse.)

Along with our proposed nonresponse models, we develop estimates for the unobserved social network data, and for the nonresponse parameters. We present iterative procedures for those estimates which do not have a closed form, and we present examples where our models may be applied. In Chapter 4, we discuss our results and examples, and in Chapter 5 we offer suggestions for future research.

For topics discussed in this dissertation, it is not necessary that the reader know more than a few basic graph theory definitions; we will briefly review concepts which are used in later discussions. For more information on these and other topics in graph theory, we refer the reader to Harary (1965).

Formally, a digraph is a set of points, or nodes, \( v_1, \ldots, v_g \) (where \( g \) is the number of nodes), and a set of directed edges, where each edge connects two nodes in a directed manner (for example, nodes \( v_i \) and \( v_j \) can contain two directed edges between them; one edge from \( v_i \) to \( v_j \), and one from \( v_j \) to \( v_i \)). The indegree of node \( v_j \) is the total number of edges coming into node \( v_j \) from other nodes; the outdegree of node \( v_j \) is the
total number of edges leaving node \( v_i \) going out to other nodes. Unless otherwise stated, we assume the edges represent a binary situation (only presence or absence). Weighted graphs, which incorporate non-binary values placed on the edges to quantify the intensity of the relation, will not be considered here. For these discussions, we allow no self-loops (edges which begin and end at the same node) or multiple edges (more than one edge between the same pair of nodes).
CHAPTER II
SOCIAL NETWORK RESEARCH; THE RECIPROCITY,
POPULARITY, AND EXPANSIVENESS MODELS

2.1 Social Network Research

The area which we today call social network research began as early as the 1930's by individual scientists who were interested in studying relationships among people within a group. Basic topics of study included the process by which relationships occur, resulting structures appearing within the group, and changes that occur in this group structure over time. The data collected in these studies typically involved a researcher's observations, results of interviews of individual group members, or questionnaires. Although the analyses of the information collected in these early years of social network research was then limited to basic correlations and simple chi-square hypothesis tests, the questions raised, the data collected, and the conjectures proposed provide an important foundation upon which today's contributions are built. We find many places in current literature where the data from these early studies are reanalyzed, using more sophisticated and powerful techniques of mathematical modelling and statistical analysis, many times affirming those conjectures proposed so many years before. Therefore it is important as well as interesting to present an overview of this early research to get a glimpse of the impact it has had on modern day social network research.
The earliest organized studies of the social structure of groups were conducted in the late 1920's through early 1940's, mainly by sociologists and psychologists who had previously been taking a global approach, studying entire societies and cultures, and were now becoming more interested in the individual members of the group, and the dynamics that occur between the individuals, thereby affecting the group. For example, one such scientist, Sherif (1935, 1936), studied the relationship between the judgment of one individual regarding a specified stimulus, and the judgment of this individual (under the same stimulus) when he/she is placed in a small group working together.

Other studies conducted during this time involved such topics as the influence of a group on the individual, affects of social pressure, identification and measurement of conformity, and the emergence of leaders vs. followers. Another important object of study was the role of an individual within the group; in particular, role identification, role determination within the group, and the dynamics of changing roles over time. (Examples of this research are numerous, including Jackson (1944), Simpson (1938), and Newcomb (1943).) Particularly interesting role and hierarchy studies were conducted by Whyte (1943) in a study of street gangs, a study of families (Strodtbeck, (1954)), marriages (Motz, (1952)), workers (Jacobson, Charters, and Lieberman (1951)), and another study by Newcomb regarding a fraternity (1961). Hare (1962) provides a very detailed review paper giving a comprehensive overview of social network study up to the early 1960's.

By the 1950's, it was becoming apparent that some method of quantification of the concepts studied, such as friendship, esteem, and power, needed to be made to enable progress in a scientific direction, i.e. using mathematics and statistics. Bales (1950) proposed a method which was often used to code interactions that occurred within a group. Each unit of behavior by an individual was broken down and scored in a certain category of either task behavior (e.g. "Let's get to work"), or social-emotional
behaviors (e.g. "That's not funny"). Bion ([1948], for example) proposed another approach which integrated these two behavioral areas into a different scoring system. Through these systems, actions and reactions between individuals could be quantified, counted, summarized, and analyzed. Scientists began to ask questions regarding the changes that occur within the social structure of a group over time, although mathematical tools were not utilized yet.

Other areas of interest included the characteristics of an individual, and how these characteristics are related to factors such as popularity, influence, leadership, conformity, and success of the individual. Again, several studies were done in these areas, including one study involving 16 cabin groups of boys at camp (Lippitt, Polansky, and Rosen [1952]). This study suggested that those boys with high physical prowess, a high amount of social activity, and who were well liked, had the greatest amount of influence within the group, while intelligence and camp experience were not viewed as important.

Perception was also a very popular object of study in these early years, and continues to be important today. Group perception of an individual, the individual's perception of himself, and the relationship between these and the self-concept of the individual were among the areas of interest. (Examples include Cartwright [1952], Zander [1958], Manis [1955], and Taguriri and Kogan [1957], among many others; the reader is again referred to Hare [1962] for a comprehensive bibliography.)

The study of interpersonal choices became very important at this time. It was recognized that these were the very building blocks of a social network structure. With newly developed systems of quantification of the relationships, this area could move forward. Scientists were interested in how interpersonal choices are made, what influenced them, and how they change. A major contributor to this area was Jacob Moreno, a psychologist who began his study of interpersonal choices in the 1930's, and
whose ideas strongly influenced the area he named sociometry. He developed a system by which members of a group indicated their choices regarding other members of the group based on certain criterion; this process is called a sociometric test. (See Moreno (1953) for an example involving groups of delinquent women in a home.) Moreno also established a set of rules by which other scientists organized their data collection; this helped to provide some continuity to the area, which was important at this time. Moreno's work became very popular in the 1950's, and is still among the most cited in social network research.

One of Moreno's many important findings involving the study of social groups was that choices between individuals in the group result in certain configurations appearing more often than others, including the number of reciprocated choices between individuals, and the number of very popular individuals and isolates (very unpopular individuals) in the group (see Moreno and Jennings (1938)). These ideas have inspired many researchers to carry out census studies of certain structures of the group (for example, Davis (1968)) and to build mathematical models to measure the importance of factors such as reciprocity and popularity, and to predict the long-term structure of the group over time. Some of these researchers include Festinger (1955), Katz (1953), Katz and Proctor (1959), and Borgatta and Bales (1956).

The first use of mathematics in the area came in the 1950's. Frank Harary, known as the 'Father of Graph Theory', recognized that graph theory could become a natural mathematical companion to sociometry, offering many very important and powerful tools, including matrix manipulation, in an attempt to quantify, model, and analyze the theories presented. He thus introduced graph theory to the area of sociometry (see Harary and Ross (1954, 1957)). In 1953, Moreno began presenting the mathematical representation of his group data in matrix form (equivalent to the adjacency matrix of the corresponding graph). This representation, called a sociogram,
is still used today. Early studies influenced by graph theory include the study of cliques within a group, the degree of cohesiveness or 'connectedness' of a group, and the amount of balance within a group (in terms of the comparison of, say, positive with negative feelings toward members). See Borgatta, Couch and Bales (1954), Harary and Ross (1957), French (1956), Taba (1955), Newcomb (1961), and Sampson (1968) for examples of such studies. The last three references contain data sets which are commonly cited throughout social network literature, even today.

Early mathematical models for the study of group structure came in the 1940's and 1950's and were mainly deterministic in nature. These include the work of Hays and Bush (1954), Simon and Guetkow (1955 a,b), Cartwright and Harary (1956), Cohen (1956), and Hare (1961). Statistical tools used at this time had progressed to the use of factor analysis (see Cattell (1952)). The most sophisticated statistical models did not appear until the 1970's.

Of particular significance in the area of mathematical modelling of social groups was the contribution of Katz and Proctor (1959), who proposed the first stochastic model for sociometric data collected over time; their proposed model uses Markov-chain theory. The data collected for their application of this model is cited frequently in the literature. The study examined relationships between the 25 members of an eighth grade class occurring over the period of one term. (Stochastic modelling became increasingly popular in the 1970's, and is still very popular today; for other examples, see Mayer (1977), and Singer and Spilerman (1974a,b).) For a review of mathematical modelling of social network data in the 1960's and 1970's, see Holland and Leinhardt (1979) and Leinhardt (1977) for a summary of work prior to 1975; also see Burt (1980). More recent literature is reviewed by Rice and Richards (1985).

Work in the 1960's and early 1970's included identifying, counting, and analyzing certain structures in the data. Triads (groups of three) became an important
subgroup of study (Davis (1977), Holland and Leinhardt (1970, 1971, 1972, 1975a,b),
and Wasserman (1977a)), and were used to examine transitivity within a group (e.g.
suppose Bob likes Mary, and Mary likes Bill; for the triad to be transitive, Bob would
have to like Bill). Censuses were taken on the number of transitive triads within the
group, and this was compared to expected numbers of such triads under random
models, or other proposed models. Another structure of interest was the clique, or
cluster of strongly related individuals within the group (for a review of work in this
area, see Davis (1979)). With matrix theory and subsequent mathematical
developments, it was now possible to develop algorithms to identify and analyze cliques
within the group, and thus provide more solid evidence for some of the conjectures
postulated in the 1930's and 1940's.

By 1970, the sociometry literature contained as many as 900 studies of group
structure, as noted by Davis (1970). This led to the development of a data bank
containing sociometric data, an important source of later research (for example, Davis
and Leinhardt (1972)). The area was now identified as the study of social networks,
and had branched out to include other areas studying similar structures, such as politics,
physics, biology, education, chemistry, and agriculture. (See Wasserman (1978) for
examples of modelling in different areas related to social network research.) Hence, a
social network could now be used to represent such entities as groups of government
agencies, cities, molecules, or groups of animals.

The large volume of work produced in the area of social network theory during
this time is evidence of the great interest being generated in the area, and the many
different levels in which group structure can be studied. Two main branches of research
had evolved: macroanalysis, and microanalysis group structure. Macroanalysts studied
the group on a global level, looking for structural patterns, and comparing and analyzing
them (for example, White, Boorman, and Breiger (1976) study blockmodels), while
microanalysts studied the local aspects of the group, focusing on an individual, his/her characteristics, and the influence of the individual on the group structure (for example, Wasserman (1979) studies transition rates in reciprocity models).

One particular area in which mathematical models were developed in the 1970's (and in which research continues today) is the study of the evolution of the social network over time. Major contributions to this area have been made by Holland and Leinhardt (1977a,b), and Wasserman (1977b, 1980), among others. In 1977, Holland and Leinhardt introduced a general form of a stochastic model for a social network, in which the social network is assumed to be a continuous-time, finite state Markov chain. Their idea was to model the instantaneous transition rates as functions in terms of the structure of the social network. These models, known as Holland-Leinhardt models, proved to be a very important innovation in the area of social network research, in that they allow for the study of how a group approaches an equilibrium, as well as the prediction of what that equilibrium structure is (if it exists). Three particular models under the Holland-Leinhardt framework have been studied in detail by Wasserman (1977b, 1980); they include 1) the reciprocity model, 2) the popularity model, and 3) the expansiveness model. They are so called because the instantaneous transition rates in each model are functions of certain relevant structures in the group; for instance, under the reciprocity model, an individual tends to form relationships with people who will return the relationship. These three models provide an excellent basis for the study of the Holland-Leinhardt general model, and are the models upon which we build in this dissertation, allowing for link and node nonresponse in the social network data. Hence, we will review the basic definitions, concepts, and results from each of these models in Section 2.2.

In another major contribution to the area, Holland and Leinhardt (1981), and Fienberg and Wasserman (1981) described an exponential family of distributions for
social network data. This allowed for the use of log-linear models (and hence corresponding results for contingency table analysis) in the analysis of social network data. It is interesting to note that Holland and Leinhardt were inspired by the early theoretical predictions of Moreno (1934), regarding the increased frequency with which certain structures appeared within a social network, including reciprocated choices and very popular and very unpopular individuals. Sociogram simulation and model fitting using iterative algorithms appeared alongside these results, and have come to be research topics in their own right. (For example, see Fienberg, Meyer, and Wasserman (1981), Wasserman and Weaver (1985).) It should also be noted that the Markov-chain models presented by Holland and Leinhardt (1977a) prove to be a special case of those exponential family log-linear models presented by Holland and Leinhardt in 1981 (see Wasserman and Iocubucci (1988)).

A number of extensions and expansions of the above models have appeared in recent literature; many of which are attributed to Stanley Wasserman. These extensions include the analysis of valued relations (as opposed to the simple binary existence/nonexistence of relations previously studied; see Wasserman and Galaskiewicz (1984), and Wasserman and Iocubucci (1986), multiple relations (Fienberg, Meyer and Wasserman (1985)), conformity of two relations (such as self concept and likeability by others) (Wasserman (1987)), and the analysis of sequential social network data (collected over time) (Wasserman and Iocubucci (1988)). For additional references, see Freeman, White and Romney (1989). Many of these papers include analysis of the often cited monastery data set produced by Sampson (1968).

Most recently, Wasserman and his co-authors have proposed methods for canonical analysis of social network data (see Wasserman and Faust (1989), and Wasserman, Faust, and Galaskiewicz (1990)). Other current topics of interest include the study of bipartite groups (those involving two groups in which relations may only
occur between members of the same group) (Wasserman and Iocubucci (1991)). Other recent work includes the application of factor analysis, multidimensional scaling, cluster analysis, and Markov graphs (see for example Arabie, Carroll, and DeSarbo, (1987), Frank (1981), and Frank and Strauss (1983)).

Many of the questions which are examined in current social network research have stemmed from early theoretical predictions of scientists such as Moreno, Newcomb, Taba, and Sampson. With current new models, one can study these questions in more detail and depth; one can for instance measure the importance of certain factors within a group, decide whether or not certain factors conform to each other, identify clusters within the group, and study the structural changes in the social network over time. It is interesting to note that even though the area of social network research has grown tremendously, touching many different areas of research, and the techniques are much more sophisticated now than they were in the 1930's, many of the basic questions to which we seek answers are the same, and they very likely will be for some time.

2.2 The Holland-Leinhardt General Model

In 1977, Holland and Leinhardt proposed a general class of models to describe the process by which a social network evolves over time. Through these models, we can address questions regarding certain structural tendencies within a group; for example, do relationships tend to become reciprocated over time? How much time does the relationship between an arbitrary pair of individuals spend in a state of unbalance? Another important use of these models is for measuring factors such as reciprocity, popularity, and expansiveness of individuals and examining questions such as 'How much more likely is it that an individual will choose a popular individual than an
unpopular individual?' We can include covariates, such as gender, as part of the analysis. We can also use these models as benchmarks against which models with more complicated structures may be compared in a hypothesis test.

The nonresponse models proposed in this dissertation build on the underlying framework of the Holland-Leinhardt general model for the data, using three particular models, as studied by Wasserman (1977): 1) the reciprocity model, 2) the popularity model, and 3) the expansiveness model; hence, relevant definitions, concepts, and results important for discussions in this dissertation will be presented in the remaining subsections of this chapter.

### 2.2.1 The Holland-Leinhardt Modelling Framework

The adjacency matrix \( X = (X_{ij}(t)) \) for the social network is modelled as a finite-state, continuous-time Markov chain. Since there are no self-loops in the digraph, and all relationships are binary (0-1) for the cases we consider, the state space for \( X \) is the set of all \( g \times g \) binary matrices with zero diagonals (where \( g \) is the number of individuals in the social network). Therefore, there are \( 2^{g(g-1)} \) possible states for this process.

This general model is based on three main assumptions. The first assumption is involves a conditional independence of transitions. If \( V \) and \( W \) are realizations of the process for the adjacency matrix at times \( t \) and \( t+h \), respectively,

\[
P_{V|W}(t,t+h) = P\{X(t+h) = W | X(t) = V\} = \
\prod_{i,j} P\{X_{ij}(t+h) = W_{ij} | X(t) = V\} + o(h) \quad \text{as } h \to 0.
\] (2.1)
In the social network setting, this assumption means that two (or more) edges in the digraph for the social network cannot both change states in a short period of time (t, t+h).

The second assumption defines the transition intensity matrix \( Q = (q(X(t))) \) for the continuous-time process as a function of \( X(t) \), the structure of the social network at time t. Dependence on \( X(t) \) reflects the idea that choices made by individuals within the group are based on certain current structural elements of the entire group; this is the basic premise of the Holland-Leinhardt general model. Each entry in the adjacency matrix \( X(t) \) at time t reflects a relationship from one individual i to another individual j in the social network. It is through the formation and removal of these individual relationships that the structure of the social network changes, moving from one state to another over time. Therefore, the transition intensities for \( X(t) \) are defined in terms of these individual processes, called arc processes, with change intensities \( \lambda_{ij}(X(t),t) \), for i, j = 1, ..., g. Formally, assumption two is the following: Let

\[
P\left[X_{ij}(t + h) = 1 - v_{ij} \mid X(t) = V\right] = h\lambda_{ij}(V(t),t) + o(h) \quad \text{as } h \to 0. \quad (2.2)
\]

(Note that since \( w_{ij} \) can only take on the values 0 or 1, a change results in the value 1-\( w_{ij} \).) Specifically, the elements of the transition probability matrix are defined to be the following: the probability of a new arc forming from individual i to individual j is defined to be

\[
P\{X_{ij}(t+h) = 1 \mid X(t) = V, X_{ij}(t) = 0\} = h\lambda_{0ij}(X(t),t) + o(h) \quad \text{as } h \to 0 \quad (2.3)
\]

and the probability of an existing arc disappearing from individual i to individual j is defined to be the following:
\[ P\{X_{ij}(t+h) = 0 \mid X(t) = V, X_{ij}(t) = 1\} = h\lambda_{1ij}(X(t), t) + o(h) \text{ as } h \to 0. \quad (2.4) \]

\( \lambda_{0ij} \) and \( \lambda_{1ij} \) are non-negative functions of \( i, j, X, \) and \( t \). They are called individual choice intensities, or infinitesimal transition rates, for the continuous-time Markov chain. These rates depend on the structure of the social network at time \( t \), and possibly on certain characteristics of the individuals \( i \) and \( j \). From here, we can form \( Q = (q_{ij}(t)) \) the intensity matrix for \( X(t) \):

\[
q_{vw}(t) = \begin{cases} 
\lambda_{ij}(X(t), t), & \text{if } X(t) = V \text{ and } X(t + h) = W \text{ differ only by element } (i, j) \\
0, & \text{otherwise}
\end{cases}
\]

and for \( V = W \) we have

\[
q_{vw}(t) = -\sum_{V \neq W} q_{ij}(t) = -\sum_{i,j} \lambda_{ij}(X(t), t), \text{ as the diagonal elements of } Q.
\]  

\[(2.5)\]

The choice intensities \( \lambda_{ij}(X(t), t) \) can be expressed in terms of disjoint processes, one for new arc formation, \( \lambda_{0ij}(X(t), t) \), and one for arc removal, \( \lambda_{1ij}(X(t), t) \), which can then be modelled separately. The resulting representation is the following for \( i, j = 1, \ldots, g \):

\[
\lambda_{ij}(X(t), t) = (1 - x_{ij})\lambda_{0ij}(X(t), t) + x_{ij}\lambda_{1ij}(X(t), t). \quad (2.6)
\]
The third assumption of the Holland-Leinhardt general model disallows the unrealistic possibility for instantaneous states (states which occur for an instant only, and then change). Namely,

\[
\lim_{h \to 0} P\{X_{ij}(t + h) = X_{ij}(t) \mid X(t)\} = 1.
\]  (2.7)

Note that any specific functional form defined for \( \lambda_{ij}(X, t) \) in the general model results in a different model for a social network process, containing model parameters for factors such as reciprocity, popularity, and expansiveness (we have already mentioned models which study these three factors separately, the Reciprocity Model, the Popularity Model, and Expansiveness Model, and will discuss them in more detail in the Section 2.3).

The Holland-Leinhardt model is very general, departing from previous mathematical models in that it does not imply any specific form for the social network, such as balance (Cartwright and Harary (1956)).

Upon definition of the functional forms for the choice intensities, the object then becomes one of estimating the model parameters from the social network data. Since typical social network data is collected over discrete time points, the use of this data in a continuous-time Markov chain model presents interesting problems involving identifiability and embeddability, which we will discuss in the following subsection.

2.2.2 Embeddability, Identifiability, and Parameter Estimation

It is reasonable to think of a social network as a constantly changing entity in which a change in relations could occur at any time; hence, it is appropriate to model it
as a continuous-time Markov chain. Our data, on the other hand, typically do not include a continuous historical record of the network; rather we have one or more distinct snapshots of the network taken at discrete time points (many times at evenly-spaced time intervals). The problem then becomes one of trying to estimate a continuous-time Markov chain with discrete-time data. This involves two objectives. First, to consider the embeddability problem: Can it be shown that there exists a class of continuous-time Markov chains from which this data might have arisen? Secondly, the identifiability problem: If indeed there exists such a class, is there a unique Markov chain which can be used to define our process? This problem had long been studied when the solution was finally discovered by Singer and Spilerman (1974b). The following discussion will explain the results as they apply to these social network models.

Suppose we have \( n \) observations of the continuous-time social network process for \( X \), namely, adjacency matrices of the social network taken at times \( t_1, \ldots, t_n \), denoted by \( X(t_1), \ldots, X(t_n) \). We pool the data across the \( n \) time periods to form an empirical, or observed transition matrix \( T \), where we let \( T_{ij}(t) \) be the number of observations of the social network process which were in state \( i \) at time \( t_1 \) and in state \( j \) at time \( t_2 \). From this, we obtain the estimated transition probability matrix \( \hat{P}(t_2 - t_1) = \hat{P}(t) \) (where \( \hat{p}_{ij} = (T_{ij}(t) / \sum_j T_{ij}(t)) \)) which contains the MLE's for the discrete-time transition probabilities. We check \( \hat{P} \) for embeddability and identifiability of a continuous-time Markov chain.

It is known that a continuous-time stationary Markov chain with intensity matrix \( Q \) is governed by the following set of differential equations (Karlin and Taylor (1975), p. 150):

\[
\frac{\partial P(t)}{\partial t} = QP(t), \quad \text{with} \ P(0) = I.
\] (2.8)
The solution to this system is

\[ P(t) = e^{Qt}, \quad t > 0, \quad \text{where} \quad e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} \]  

(2.9)

Now, given our empirical transition probability matrix \( \hat{P} \), we want to check for embeddability, i.e. if it can be represented as

\[ \hat{P}(t) = e^{\hat{Q}t} \quad \text{for some class of matrices} \quad \Theta, \quad \hat{Q} \in \Theta. \]  

(2.10)

If so, then we have the following:

\[ \hat{Q} = \frac{1}{t} \log \hat{P} \quad \forall \hat{Q} \in \Theta. \]  

(2.11)

Conditions for embeddability, as they apply to our social network models, are the following. Each of the \( r \) eigenvalues \( \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_r \) of \( \hat{P} \) must satisfy the condition

\[ \pi \left( \frac{1}{2} + \frac{1}{r} \right) \leq \arg(\log \varepsilon_i) \leq \pi \left( \frac{3}{2} - \frac{1}{r} \right). \]  

(2.12)

Singer and Spilerman point out that if all eigenvalues of \( \hat{P} \) are real, positive, and distinct, they automatically satisfy this condition, and additionally, \( \hat{Q} \) is unique.

It turns out that this is indeed typically the case. However, for additional conditions for embeddability and identifiability when the above does not hold, see Singer and Spilerman (1974, 1976).

Given the above conditions do hold, we can then find \( \hat{Q} \) in one of two ways:
1) \( \hat{Q} = \sum_{i=1}^{r} \log(e_i) \prod_{j \neq i} (\hat{P} - e_i1)/(e_i - e_j) \) \hspace{1cm} (2.13)

or

2) Diagonalize \( \hat{P} \) so that \( \hat{P} = H\Lambda H^{-1} \), where \( \Lambda = \text{diag}[e_1, \ldots, e_r] \) and \( H \) is the matrix of associated eigenvectors. We then have
\[ \hat{Q} = \log \hat{P} = \log[H\Lambda H^{-1}] = H[\log \Lambda]H^{-1}, \]

where
\[ \log \Lambda = \begin{bmatrix} \log e_1 & 0 & \cdots & 0 \\ 0 & \log e_2 & \cdots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \log e_r \end{bmatrix} \] \hspace{1cm} (2.14)

Now that we have \( \hat{Q} \), we can set \( \hat{Q} = Q \) (the theoretical intensity matrix), and solve for MLE's of the model parameters. This identifies the unique continuous-time Markov chain consistent with our data, hence one from which our data may have arisen.

If the data include more than two sociomatrices, the estimation process involves calculating the set of \( \hat{P} \) matrices. Two strategies have been proposed in proceeding from here: 1) use one of these matrices to find a \( \hat{Q} \) matrix, and the rest to test for time homogeneity in the model, or 2) repeat the above process for each \( \hat{P} \) and examine the likelihood function in areas around each solution to obtain 'pseudo-confidence regions' for the model parameters. For further discussion of these two strategies, see Singer and Spilerman (1974b, 1976), and Wasserman (1977b), respectively.

In this dissertation, we propose models to handle nonresponse in the social network data. Nonresponse in the setting of the Holland-Leinhardt general model results in an incomplete \( \hat{P} \) matrix. Our proposed models provide an estimate for the \( \hat{P} \) matrix, and from that point the general theory of the Holland-Leinhardt models can be applied to estimate model parameters such as reciprocity, etc. Our nonrandom
nonresponse models make use of the observed data in order to estimate the missing data, thus providing greater accuracy to estimates of the social network models parameters.

2.3 The Reciprocity Model

The reciprocity model involves a parameterization of the H-L model in which the decision by individual \( v_i \) to relate to individual \( v_j \) depends only on the presence or absence of the reciprocation of this relationship from \( v_j \) to \( v_i \). The model parameters measure the importance of reciprocity to relationship formation, continuation and dissolution. The functional forms for the change intensities under the reciprocity model are the following (as presented by Wasserman (1977)):

\[
\begin{align*}
\lambda_{0ij}(x,t) &= \lambda_0 + \mu_0 x_{ji} \quad \text{(new arc is formed)} \\
\lambda_{1ij}(x,t) &= \lambda_1 + \mu_1 x_{ji} \quad \text{(existing arc is terminated)}
\end{align*}
\]

Note that we have assumed time homogeneity for simplicity. \( \lambda_0 \) and \( \lambda_1 \) are measures of the overall rate of change for the group, while \( \mu_0 \) and \( \mu_1 \) measure the importance of the reciprocated arc. We see that the change intensities (hence the transition probabilities) of the choice process \( v_i \rightarrow v_j \) depend only on the presence or absence of a reciprocated arc. According to this model, the digraph process \( X(t) \) can be split into a set of \( \binom{g}{2} \) independent Markov process, one for each pair of nodes, providing a major simplification for all theory and applications involved. \( D_{ij}(t) = (X_{ij}(t), X_{ji}(t)) \) is defined as the dyad for the pair of nodes \( i \) and \( j \), where \( i < j = 1, ..., g \). Each dyad is a continuous-time Markov chain with four states:
Mutual \( D_{ij}(t) = (1,1) \)

Asymmetric \( D_{ij}(t) = (1,0) \) or \((0,1)\)

Null \( D_{ij}(t) = (0,0). \) \( (2.17) \)

Wasserman (1977) gives the infinitesimal transition matrix \( Q \), and the set of equilibrium probabilities, and estimates the parameters and/or certain functions of these parameters.

Application of this model to data in two different settings is exemplified. In setting (i), only one observed sociomatrix is available, and in setting (ii), two or more sociomatrices are available. We give the \( Q \) matrix for setting (i) under the reciprocity model below:

\[
Q = \begin{pmatrix}
(0,0) & (1,0) & (0,1) & (1,1) \\
(0,0) & -2\lambda_0 & \lambda_0 & \lambda_0 & 0 \\
(1,0) & \lambda_1 & -(-\lambda_0 + \lambda_1 + \mu_0) & 0 & \lambda_0 + \mu_0 \\
(0,1) & \lambda_1 & 0 & -(-\lambda_0 + \lambda_1 + \mu_0) & \lambda_0 + \mu_0 \\
(1,1) & 0 & \lambda_1 + \mu_1 & \lambda_1 + \mu_1 & -2(\lambda_1 + \mu_1)
\end{pmatrix}
\]

Figure 2. The \( Q \) matrix for the reciprocity model.

As an example of how these elements are derived, an arc changing from state \((1,0)\) to \((1,1)\) in a short period of time requires individual \( j \) to form a relationship with individual \( i \) in the presence of reciprocation of the relationship by individual \( i \) (hence \( x_{ji} = 1 \) in Equation \((2.15)\)). Thus, the \((2,4)\) element in \( Q \) is \( \lambda_0 + \mu_0 \).

Data from the social network collected over any two consecutive time periods \( t_1 \) and \( t_2 \) can be summarized in a 4x4 contingency table containing a total of \( \binom{g}{2} \) counts, one for each dyad in the social network. We will call this contingency table the observed dyad transition matrix; we give the form below.
<table>
<thead>
<tr>
<th>Time $t_1$</th>
<th>(0,0)</th>
<th>(1,0)</th>
<th>(0,1)</th>
<th>(1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0)</td>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
<td>$x_{13}$</td>
<td>$x_{14}$</td>
</tr>
<tr>
<td>(1,0)</td>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>$x_{23}$</td>
<td>$x_{24}$</td>
</tr>
<tr>
<td>(0,1)</td>
<td>$x_{31}$</td>
<td>$x_{32}$</td>
<td>$x_{33}$</td>
<td>$x_{34}$</td>
</tr>
<tr>
<td>(1,1)</td>
<td>$x_{41}$</td>
<td>$x_{42}$</td>
<td>$x_{43}$</td>
<td>$x_{44}$</td>
</tr>
</tbody>
</table>

Figure 3. The dyad transition matrix after Stage 1.

This contingency table has cell counts $x_{k\ell}$, where $k$ refers to the state of the dyad at time $t_1$ and $\ell$ refers to the state of the dyad at time $t_2$. Here $k$ or $\ell = 1, ..., 4$ indicates a dyad in a null, asymmetric (1,0), asymmetric (0,1), or mutual state, respectively. For example, $x_{32}$ is the number of dyads in the social network that were in state (0,1) at time $t_1$, and moved to state (1,0) at time $t_2$. (See Bishop, Fienberg, and Holland, (1975) for a discussion of the representation of Markov-chain data as contingency tables.)

The probabilities underlying the observed $x_{k\ell}$ may be modeled as follows. Let $\pi_k$ be the probability that a dyad is in state $k$ at time $t_1$ and $p_{k\ell}$ be the conditional probability that the dyad is in state $\ell$ at time $t_2$ given that it was in state $k$ at time $t_1$. The likelihood function in this case has the following form:

$$L \propto \prod_{i,j} (\pi_i p_{ij})^{x_{ij}}.$$  \hspace{1cm} (2.18)

From the observed transition matrix, one can obtain the empirical transition probability matrix, $\hat{P}$. Ordinarily, one would proceed by taking the derivative of $L$ with respect to the model parameters to find the MLE's; however, since we have $\hat{P} = e^{\hat{Q}t}$, which is an infinite power series of matrices (and recall the parameters of interest are in
the $Q$ matrix), these would be difficult to find. Wasserman proposes another strategy, involving solving this equation for $\hat{Q}$, setting this equal to $Q$, and solving for the parameters, as previously mentioned. He then conducts a grid search of the likelihood function in the area of the estimates to obtain 'pseudo-confidence regions' for the parameters.

In an interesting application, Wasserman (1980) fit the reciprocity model to classroom data collected by Katz and Proctor (1959). In this study, each of 25 members of an eighth-grade class was asked to list the three other students he/she would rather sit by (note we have a fixed-choice binary digraph). These surveys were taken four times during the school year (setting (ii)). The question of interest is the following: Does student i take into account whether student j wants to sit by him/her when making his/her choice? Katz and Proctor said yes. The interval estimates found by Wasserman for the four parameters did confirm the original conclusions: student i was much more likely to choose to sit by student j if student j had reciprocated the choice. Also, student i was much less likely to decide to remove student j from his/her list if student j had reciprocated that choice.

Wasserman (1980) also fits the reciprocity model to Newcomb's fraternity data (1961), and verified two of Newcomb's hypotheses: 1) mutuality between members of the social network increased over time, and 2) mutual relationships were more stable over time than non-mutual relationships.

The assumption of independence of the dyad processes under the reciprocity model is not always valid, but can be reasonably assumed in many cases; without this assumption, the mathematics becomes much more difficult. More work needs to be done in the area of building dependence structures into these models (see Wasserman (1980) for a discussion on this issue).
2.4 The Popularity and Expansiveness Models

Two additional models under the Holland-Leinhardt framework which are studied by Wasserman (1977, 1980) are the popularity model and the expansiveness model. These models are similar in that the choice process for each one involves more than one other node; in fact, information regarding all other nodes is included. They are also very similar in form, virtual complements of each other, as we will see.

The popularity model assumes that a choice by individual i regarding individual j depends (only) on the popularity of individual j. The popularity of individual j, in terms of the digraph, is measured by the indegree of node j, and is denoted $X_{\star j} = \sum_i X_{ij}(t)$. The higher the indegree, the more popular the individual is. (Note the maximum indegree for any individual is $g-1$.) The choice intensities for the popularity model are given by the following:

$$\lambda_{0ij}(X(t), t) = \lambda_0 + \pi_0 X_{\star j} \quad \text{ (new arc is formed)}$$  \hspace{1cm} (2.19)

$$\lambda_{1ij}(X(t), t) = \lambda_1 + \pi_1 X_{\star j} \quad \text{ (existing arc is removed).}$$ \hspace{1cm} (2.20)

Note that these choice intensities do not depend on the individual making the choice; rather, they depend on the other individual in the pair. Hence, each individual contemplating choosing individual j is involved in an identical choice process. This means the process for the entire social network is broken into $g$ independent processes, one for each individual. These processes are called column processes, $X_j$, where

$$X_j = (X_{1j}(t), X_{2j}(t), ..., X_{gj}(t))^\prime.$$ \hspace{1cm} (2.21)
(Data for the column process $X_i$ is found in the jth column of the adjacency matrix, where all incoming arcs into node j appear; hence the name column process.) Each column process has $2^{(g-1)}$ states, consisting of all binary gx1 vectors with element $x_{ij} = 0$. The number of states increases quickly with g, hence this model is more complicated than the dyad model. However, the reciprocity model uses more information regarding individual choices, and is more applicable in certain situations.

One simplification that can be made, as noted by Wasserman (1980), is to note that in terms of estimation and analysis, we only need to know the sum of the values of the column process, namely $X_{e,j} = \sum_i X_{ij}(t)$, and not the value of each particular $X_{ij}$, since all individuals are regarded as similar under this model. In other words, in determining whether to form a new relationship with an individual, only the popularity of that individual is taken into account, not the identities of persons currently 'liking' that person. This process has a much smaller state space, and greatly simplifies calculations and analyses. By this simplification, we see that the process for the social network under the popularity model is broken into g independent indegree processes, each being a Markov chain with g states, since each node has as its indegree a value between 0 and g-1. Data collected over two time periods are summarized in a gxg matrix containing counts of all nodes moving from indegree k at time $t_1$ to indegree $\ell$ at time $t_2$. This matrix is called the observed indegree transition matrix.

Among other results, Wasserman (1980) provides forms of the Q matrix for the popularity model in the case where $g = 4$. The same method for estimation applies to the popularity model, as for the reciprocity model, in terms of the $\hat{P}$ and $\hat{Q}$ matrices. He fits the model to Newcomb's fraternity data (1961). It turned out that popular individuals were more stable (remained popular) than unpopular individuals, who actually may lose popularity over time.
Under the expansiveness model, the choice by individual i regarding individual j depends on the expansiveness of individual i. Expansiveness refers to a measure of how outgoing the individual is who is making the choice, the idea being that a shy person is less likely to form new friendships than an outgoing, 'expansive' person. Therefore, the outdegree of individual i is the only thing taken into account. The choice functions under the expansiveness model are the following:

\[
\lambda_{0ij}(X(t),t) = \lambda_0 + \pi_0 X_{i\cdot} \quad \text{(new arc is formed)} \tag{2.22}
\]
\[
\lambda_{1ij}(X(t),t) = \lambda_1 + \pi_1 X_{i\cdot} \quad \text{(existing arc is removed).} \tag{2.23}
\]

Therefore, the process for the social network is broken into g independent outdegree processes, \(X_i = (X_{i1}(t), X_{i2}(t), ..., X_{ig}(t))\), where again, we only need to know the sum of the values of the column process, \(X_{i\cdot} = \sum_j X_{ij}\) for our analysis of the expansiveness model. Previously mentioned methods for estimation and analysis apply here as well. The data (outdegrees of each node at each time period) will be summarized in a \(g \times g\) outdegree transition matrix, similar to the indegree transition matrix under the popularity model.

These three models represent the underlying models for the data to which we add processes for nonresponse in the remaining chapters of this dissertation.
CHAPTER III
MODELS FOR LINK NONRESPONSE
UNDER THE RECIPROCITY MODEL

3.1 The Two-Stage Process for Link Nonresponse

In this chapter, we present six different link nonresponse models, assuming that the model for the social network data is the reciprocity model. As was described in Chapter 2, under the reciprocity model, the choice intensities for each individual regarding relations depend only on whether or not that relationship is reciprocated, and the stochastic process for the entire social network is broken into \( \binom{g}{2} \) independent dyad processes, one for each pair of individuals in the social network. Corresponding to this dyad process, we propose general models for link nonresponse by either member of a dyad at time \( t_1 \) only, time \( t_2 \) only, or both times. The dyads are assumed to be independent; hence we can have a number of occurrences of link nonresponse in the social network at one time. Three of our models are random nonresponse models, where nonresponse is modelled at time \( t_1 \), time \( t_2 \), and both time periods (respectively), and will be used as benchmarks. The other three models allow for nonrandom nonresponse, and the nonresponse depends on the state of the dyad at time \( t_1 \), or the time of the nonresponse, depending on the model.

As we recall from Chapter 1, nonresponse may occur in social network data in different ways and this will be reflected in our models. First, nonresponse can be on
the level of certain relations between specific individuals only. This type of nonresponse involves only isolated links in the digraph. We can imagine occurrences of missing links at different times between different individuals throughout, say, a 12-month study. We will label this scenario link nonresponse. (This is similar to item nonresponse in the survey-sampling terminology.) A second way in which nonresponse can occur in social network data is when an individual becomes a nonrespondent regarding all outgoing relationships with others in the social network. This means all outgoing links from that individual are missing. We will call this type of nonresponse node nonresponse; models for node nonresponse will be presented in Chapter 4. In this dissertation, we will consider link and node nonresponse models for data collected for two time periods: time t₁ and time t₂, although work on the models themselves (without the nonresponse) has been done for more than two time periods (see Wasserman (1980)).

3.1.1 The Observed Dyad Transition Matrix for Link Nonresponse Under the Reciprocity Model

The basic ideas used in our nonrandom link nonresponse models follow those from the area of survey sampling, and the work of Chen and Fienberg (1974, 1976) and Stasny (1983, 1987) on models for partially cross-classified data. Applying these ideas to the social network setting, we think of the observed social network data as being the end result of two stages. The first stage involves allocating the data to the 4x4 observed dyad transition matrix (as if all links were known). Then at the second stage, links may become missing (at t₁, t₂, or both time periods) due to link nonresponse; this means a dyad can lose part or all of its information for one or more time periods. This occurs according to several different scenarios, since each dyad
contains two pieces of information, and data for each dyad is collected for two time periods. In this section we will define terms which will be used throughout this dissertation, to describe these situations.

We use two Markov chains to model the two-stage process: one Markov chain models the data in stage one, and a second Markov chain models the link nonresponse for stage 2. We model the link nonresponse in terms of the observed dyad transition matrix (which may contain nonresponse data) using the model to allocate the link nonresponse data back into an estimate of the original 4x4 (unobserved) dyad transition matrix. This results in an estimated dyad transition matrix, no longer containing link nonresponse data. From there, we proceed with previous modelling techniques for the social network as described in Chapter 2. The ideas of Stasny (1983, 1987) regarding a two-stage Markov-chain model are modified to the social network setting and extended as the models become more complex.

Under the reciprocity model, we summarize the observed dyad data for the entire social network collected over two time periods in the observed dyad transition matrix. It contains counts of all dyads: those that have been completely reported, partially reported, and/or unreported for one or both time periods. We will define new terms which will help identify dyads in these different situations.

If a dyad is completely reported at time $t_1$, that is, if both individuals in the dyad respond at both time periods, we will say the dyad is "completely classified" at time $t_1$, or completely row classified. If one link is missing at time $t_1$, we will say the dyad is "partially classified" at time $t_1$, or partially row classified. If both links are missing at time $t_1$, we say the dyad is "unclassified" at time $t_1$, or is row unclassified. Similarly, for time $t_2$, a dyad can be completely classified at time $t_2$ (or column classified), partially classified at time $t_2$, (or partially column classified), or unclassified at time $t_2$ (column unclassified).
We consider the link nonresponse status of a dyad over two time periods when counting it in the observed dyad transition matrix. Following the terminology of Chen and Fienberg (1974, 1976), we say that a dyad is completely cross-classified if all information is known at both time periods, partially cross-classified if some (but not all) information is missing at one or both time periods, and completely unclassified if all information is missing for both time periods. Therefore a dyad which is completely row classified and completely column classified will be completely cross-classified in the observed dyad transition matrix. A dyad which is partially row classified or partially column classified will be partially cross-classified in the observed dyad transition matrix. And a dyad which is row unclassified and column unclassified will be completely unclassified in the observed dyad transition matrix.

As a result of dyads possibly containing missing data, the observed dyad transition matrix results in a 4x4 contingency table including the counts of all completely cross-classified data, and up to 15 supplementary row and column matrices including the counts of all partially cross-classified and completely unclassified dyads. This matrix is presented in Figure 4.

Note that in Figure 4, an 'M' in a row or column label in the observed dyad transition matrix identifies the element in the dyad where link nonresponse occurred (i.e. the link is missing) at that time period. For example, (M,1) indicates a dyad having one link missing from individual i to j, and the other link is present from individual j to i. The elements of the matrix are dyad counts; for example, $h_{21}$ represents the number of dyads that at time $t_1$ had the link from individual i to individual j present and the link from j to i missing, and at time $t_2$ the link from i to j was missing, and the link from j to i was absent. The number of dyads that were unclassified for both time periods is $z_{11}$, since both individuals were link nonrespondents at both times.
<table>
<thead>
<tr>
<th>Time $t_1$</th>
<th>(0,0)</th>
<th>(1,0)</th>
<th>(0,1)</th>
<th>(1,1)</th>
<th>Time $t_2$</th>
<th>(M,0)</th>
<th>(M,1)</th>
<th>(0,M)</th>
<th>(1,M)</th>
<th>(M,M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0)</td>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
<td>$x_{13}$</td>
<td>$x_{14}$</td>
<td>$a_{11}$</td>
<td>$a_{12}$</td>
<td>$e_{11}$</td>
<td>$e_{12}$</td>
<td>$o_{11}$</td>
<td></td>
</tr>
<tr>
<td>(1,0)</td>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>$x_{23}$</td>
<td>$x_{24}$</td>
<td>$a_{21}$</td>
<td>$a_{22}$</td>
<td>$e_{21}$</td>
<td>$e_{22}$</td>
<td>$o_{21}$</td>
<td></td>
</tr>
<tr>
<td>(0,1)</td>
<td>$x_{31}$</td>
<td>$x_{32}$</td>
<td>$x_{33}$</td>
<td>$x_{34}$</td>
<td>$a_{31}$</td>
<td>$a_{32}$</td>
<td>$e_{31}$</td>
<td>$e_{32}$</td>
<td>$o_{31}$</td>
<td></td>
</tr>
<tr>
<td>(1,1)</td>
<td>$x_{41}$</td>
<td>$x_{42}$</td>
<td>$x_{43}$</td>
<td>$x_{44}$</td>
<td>$a_{41}$</td>
<td>$a_{42}$</td>
<td>$e_{41}$</td>
<td>$e_{42}$</td>
<td>$o_{41}$</td>
<td></td>
</tr>
<tr>
<td>(M,0)</td>
<td>$b_{11}$</td>
<td>$b_{12}$</td>
<td>$b_{13}$</td>
<td>$b_{14}$</td>
<td>$d_{11}$</td>
<td>$d_{12}$</td>
<td>$g_{11}$</td>
<td>$g_{12}$</td>
<td>$u_{11}$</td>
<td></td>
</tr>
<tr>
<td>(M,1)</td>
<td>$b_{21}$</td>
<td>$b_{22}$</td>
<td>$b_{23}$</td>
<td>$b_{24}$</td>
<td>$d_{21}$</td>
<td>$d_{22}$</td>
<td>$g_{21}$</td>
<td>$g_{22}$</td>
<td>$u_{21}$</td>
<td></td>
</tr>
<tr>
<td>(0,M)</td>
<td>$f_{11}$</td>
<td>$f_{12}$</td>
<td>$f_{13}$</td>
<td>$f_{14}$</td>
<td>$h_{11}$</td>
<td>$h_{12}$</td>
<td>$t_{11}$</td>
<td>$t_{12}$</td>
<td>$w_{11}$</td>
<td></td>
</tr>
<tr>
<td>(1,M)</td>
<td>$f_{21}$</td>
<td>$f_{22}$</td>
<td>$f_{23}$</td>
<td>$f_{24}$</td>
<td>$h_{21}$</td>
<td>$h_{22}$</td>
<td>$t_{21}$</td>
<td>$t_{22}$</td>
<td>$w_{21}$</td>
<td></td>
</tr>
<tr>
<td>(M,M)</td>
<td>$s_{11}$</td>
<td>$s_{12}$</td>
<td>$s_{13}$</td>
<td>$s_{14}$</td>
<td>$v_{11}$</td>
<td>$v_{13}$</td>
<td>$y_{11}$</td>
<td>$y_{12}$</td>
<td>$z_{11}$</td>
<td></td>
</tr>
</tbody>
</table>

Figure 4. The observed dyad transition matrix for link nonresponse.
It is important that we differentiate a missing link from a link which is absent; a missing link is a link in which the status of the relationship is unknown, while in an absent link, the arc between the individuals is not present. It will be necessary to be able to distinguish between the two when collecting social network data.

It is useful to think of the observed dyad transition matrix as a partitioned matrix with the following form:

\[
\begin{bmatrix}
X_{4x4} & A_{4x2} & E_{4x2} & O_{4x1} \\
B_{2x4} & D_{2x2} & G_{2x2} & U_{2x1} \\
F_{2x4} & H_{2x2} & T_{2x2} & W_{2x1} \\
S_{1x4} & V_{1x2} & Y_{1x2} & Z_{1x1}
\end{bmatrix}
\]

Figure 5. Observed dyad transition matrix in partitioned block form.

Note here that the submatrix X is the 4x4 matrix containing the counts of all completely cross-classified dyads (those that are both row and column classified). (We also note that if missingness were ignored, X is the matrix that would have been analyzed.) Counts of dyads in which one individual was a nonrespondent at one time period only appear in blocks B and F (missing at time t₁ only), and blocks A and E (missing at time t₂ only). Counts of dyads in which one individual is a nonrespondent in one time period, and one (possibly the same one) is a nonrespondent in the other time period appear in blocks D, G, H, and T. Blocks O and S represent dyads where both individuals are respondents at one time period, and both individuals are nonrespondents at the other. Blocks V, Y, U, and W represent counts of dyads where both individuals are nonrespondents for one time period, and only one is a nonrespondent at the other time period. Block Z indicates the number of dyads that are unclassified, since both individuals are nonrespondents for both time periods.
In terms of the two-state process, the unobserved first stage involves all the dyads being completely cross-classified in the 4x4 block X according to their true states at times t₁ and t₂. Then, at the second stage, each member of the dyad (at either time t₁, t₂, or both) may become a link nonrespondent; this results in the dyad losing part or all information for one or both time periods, which moves it to one of the other 15 supplementary row or column block matrices in the observed dyad transition matrix.

We also note here that under certain circumstances in the models, we may not indeed have 16 individual matrices containing data, since some pairs of cells are indistinguishable under certain models (and this may occur across blocks), and since some models only allow for nonresponse at one time period. There will also be a certain amount of symmetry present in the probabilities for the data and nonresponse Markov chains in all three of the random nonresponse models, with some added constraints necessary for the nonrandom nonresponse models. For instance, if we model random link nonresponse at time t₂ only, then the (conditional) transition probability of moving from say, state (1,0) to state (M,0) is not really different from the transition probability of moving from state (1,0) to state (1,M); only the number of missing links is taken into account. On the other hand, if we model nonrandom link nonresponse at time t₂ only, with nonresponse at time t₂ dependent on the state of the dyad at time t₁, then we would want these two transition probabilities to be different; we present this idea in the context of our example.

Suppose at time t₁, Tom and Bob are in a (1,0) state (Tom likes Bob, but Bob does not reciprocate the friendship). Under a link nonresponse model where nonresponse occurs at one or both time periods, and the nonresponse depends on the state of the dyad at time t₁, we have two different situations: 1) Tom likes Bob, but 2) Bob doesn't like Tom. Under these circumstances, each individual will have a
different perspective on the situation in terms of responding or not responding. Perhaps Tom is more likely to become a nonrespondent during the next time period than Bob is, because Tom’s friendship is not being reciprocated; this means that from state (1,0) at time $t_1$, the probability of going to state (M,0) may be more likely than going to state (1,M). It seems reasonable that we do not assume these nonresponse transition probabilities will be the same, nor that the corresponding probabilities for the data Markov chain would be the same.

The question of number of distinct cells arises when one calculates the degrees of freedom for each model, and presents the greatest challenge in the models where link nonresponse occurs at one or both time periods. Some of these questions involve the matter of relabelling the nodes. We will discuss these issues further when we present the degrees of freedom for each model in turn.

### 3.1.2 Probabilities for Stages One and Two Markov chains

Applying the ideas of Stasny (1983, 1987), probabilities associated with each of the two stages of the process involve the initial and transition probabilities from two separate Markov chains: one Markov chain to model the dyad data (we will call it the data Markov chain), and another Markov chain to model the link nonresponse (we will call it the link nonresponse Markov chain). The probabilities associated with the data Markov chain are the following:

$$\pi_k = \text{the initial probability that a dyad is in state } k \text{ at time } t_1,$$

where

$$k = 1, ..., 4 \text{ represent states } N, A1, A2, M, \text{ respectively.}$$
\( p_{k, \ell} \) = the (conditional) transition probability of a dyad moving to state \( \ell \) at
time \( t_2 \) given that it was in state \( k \) at time \( t_1 \) (\( \ell = 1, \ldots, 4 \) and
\( k = 1, \ldots, 4 \)). \hfill (3.1)

The probabilities associated with the link nonresponse Markov chain follow,
beginning with estimates of the initial probabilities:

\[ \zeta_{rr}(k, \ell) = \text{the initial probability that both links in the dyad are known at time } t_1. \]

\[ \zeta_{r'r}(k, \ell) = \text{the initial probability that the first link in the dyad is missing and the second link is known at time } t_1. \]

\[ \zeta_{rr'}(k, \ell) = \text{the initial probability that the first link in the dyad is known and the second link is missing at time } t_1. \]

\[ \zeta_{r'r'}(k, \ell) = \text{the initial probability that both links in the dyad are missing at time } t_1. \) \hfill (3.2)

For any dyad in cell \((k, \ell)\) of the observed dyad transition matrix, there are sixteen (conditional) transition probabilities relating to link nonresponse, one for each combination of nonresponse/response states of a dyad over two time periods. We let

\[ \rho_{rr,rr}(k, \ell) = \text{the (conditional) transition probability of the dyad moving from both links known at time } t_1 \text{ to both links known at time } t_2. \]
\( \rho_{rr,rr}(k, \ell) = \) the (conditional) transition probability of the dyad moving from both links known at time \( t_1 \) to the first link missing and the second link known at time \( t_2 \).

\( \rho_{rr,rr'}(k, \ell) = \) the (conditional) transition probability of the dyad moving from both links known at time \( t_1 \) to the first link known and the second link missing at time \( t_2 \).

\( \rho_{rr,rr'}(k, \ell) = \) the (conditional) transition probability of the dyad moving from both links known at time \( t_1 \) to both links missing at time \( t_2 \). \hspace{1cm} (3.3)

The other 12 transition probabilities \( \rho_{rr,rr'}(k, \ell) \) through \( \rho_{rr',rr'}(k, \ell) \), are defined similarly; the first pair of subscripts indicates the state of missingness of the links in the dyad at time \( t_1 \) and the second pair of subscripts indicates the state of missingness of the links in the dyad at time \( t_2 \). Note that \( r \) indicates response for a link, while \( r' \) denotes nonresponse (i.e. missingness) for that link.

For notational ease, we represent the four states of missingness for the links in a dyad at either time period using the following notation:

\begin{align*}
1 &= r r = \text{both links known} \\
2 &= r' r = \text{first link missing, second link known} \\
3 &= r r' = \text{first link known, second link missing} \\
4 &= r' r' = \text{both links missing} \\
\end{align*} \hspace{1cm} (3.4)
So, for example, \( \rho_{\pi,\pi'}(k, \ell) \) will be denoted by \( \rho_{1,2}(k, \ell) \). (In general, we will use \( k \) and \( \ell \) as subscripts for the data Markov chain-parameters and \( i \) and \( j \) as subscripts for the nonresponse Markov-chain parameters.)

The following general constraints apply to the parameters in the most general form of the models for this two-stage process (for \( k, \ell = 1, \ldots, 4 \)):

1) \[ \sum_{k=1}^{4} \pi_k = 1 \]

2-5) \[ \sum_{\ell=1}^{4} \rho_{k\ell} = 1 \text{ for each } k=1, \ldots, 4 \]

6) \[ \sum_{i=1}^{4} \zeta_i(k, \ell) = 1 \]

7-10) \[ \sum_{j=1}^{4} \rho_{ij}(k, \ell) = 1 \text{ for } i = 1, \ldots, 4 \] \hspace{1cm} (3.5)

Note that this, the most general form of the model, cannot be fit given the observed data since it contains too many parameters. Observe, for the data Markov chain, there are 20 parameters with 5 constraints, and since there is a separate Markov chain, containing 20 parameters with 5 constraints for each cell \((k, \ell)\) of the observed dyad transition matrix, this means a final total of 255 free parameters; the maximum number of cells is \(9 \times 9 = 81\) with 1 constraint, leaving (a maximum of) 80 free cells. Clearly \( 255 >> 80 \), hence there would be too many parameters to be able to get a positive number of degrees of freedom for the general form of the model. In the next section, we propose constraints on the \( \zeta \) and \( \rho \) parameters that allow us to fit the models.

Using the above notation, the general forms of the probabilities for dyads appearing in any of the cells of the observed dyad transition matrix are summarized in Figure 6. Note that, for example, there are 16 elements in the upper-left block, denoted in one general form. (For notational ease, we have dropped the dependence of \( \zeta \) and \( \rho \) parameters on \((k, \ell)\) in this figure.)
\[ \begin{array}{cccc}
\zeta_1 p_{1,1} \pi_k p_k \ell & \sum_{\ell=1}^2 \zeta_1 p_{1,2} \pi_k p_k \ell & \sum_{\ell=3}^4 \zeta_1 p_{1,2} \pi_k p_k \ell & \sum_{\ell=1}^2 \sum_{\ell=3}^4 \zeta_1 p_{1,3} \pi_k p_k \ell \\
\sum_{k=1}^4 \zeta_2 p_{2,1} \pi_k p_k \ell & \sum_{k=1, \ell=1}^2 \zeta_2 p_{2,2} \pi_k p_k \ell & \sum_{k=1, \ell=3}^4 \zeta_2 p_{2,2} \pi_k p_k \ell & \sum_{k=1}^2 \sum_{k=3}^4 \sum_{\ell=1, \ell=3}^2 \zeta_2 p_{2,3} \pi_k p_k, \ell \\
\sum_{k=1}^4 \zeta_3 p_{3,1} \pi_k p_k \ell & \sum_{k=1, \ell=1}^2 \zeta_3 p_{3,2} \pi_k p_k, \ell & \sum_{k=1, \ell=3}^4 \zeta_3 p_{3,2} \pi_k p_k, \ell & \sum_{k=1}^2 \sum_{k=3}^4 \sum_{\ell=1, \ell=3}^2 \zeta_3 p_{3,3} \pi_k p_k, \ell \\
\sum_{k=2,4} \zeta_4 p_{4,1} \pi_k p_k \ell & \sum_{k=1, \ell=1}^2 \sum_{k=1, \ell=3}^4 \zeta_4 p_{4,2} \pi_k p_k \ell & \sum_{k=1, \ell=3}^4 \sum_{k=1, \ell=3}^4 \zeta_4 p_{4,3} \pi_k p_k, \ell & \sum_{k=1}^2 \sum_{k=3}^4 \sum_{\ell=1, \ell=3}^2 \sum_{\ell=1, \ell=3}^2 \zeta_4 p_{4,4} \pi_k p_k, \ell \\
\end{array} \]

Figure 6. Probabilities for elements in each \((k, \ell)\) cell of the observed dyad transition matrix.
3.2 Some Models for Link Nonresponse

Under the reciprocity model for the data, we propose the following six Markov-chain models for link nonresponse:

**Model 1** Random link nonresponse at time $t_2$ only.

**Model 2** Random link nonresponse at time $t_1$ only.

**Model 3** Nonrandom link nonresponse at time $t_2$ only; probability of missingness depends on the reciprocity state of the dyad at time $t_1$, (N, A1, A2, or M).

**Model 4** Nonrandom link nonresponse at time $t_1$ only; probability of missingness depends on the reciprocity state of the dyad at time $t_1$, (N, A1, A2, or M).

**Model 5** Random link nonresponse at either or both time periods.

**Model 6** Nonrandom link nonresponse at either or both time periods; probability of missingness depends on the reciprocity state of the dyad at the time of the nonresponse (N, A1, A2, or M).

Of course, there are other variations of these models possible. For example, nonresponse at time $t_2$ with missingness depending on the state of the dyad at time $t_2$ (a modified version of Model 3 in terms of nonresponse). We present these six models as illustrations. Estimators for other models may be developed in the same manner as for the models described here.

We will consider each model in turn, finding degrees of freedom and forms for maximum likelihood estimates for the parameters (most of these will be iterative forms, since closed-form solutions for the maximum likelihood equations will be
In Chapter 5, we will present iterative procedures needed for obtaining certain parameter estimates, where we will illustrate fitting of the models to data. Since summations involving elements of transition probability matrices are fixed at 1 as we sum across rows, results for the link nonresponse models where missingness occurs at time $t_2$ only will be less complex in form; hence those models are presented first. We note that in social network settings, it makes more sense for the link nonresponse at time $t_1$ in Model 4 to depend on the state of the dyad at time $t_1$, the current time period, rather than the state of the dyad at time $t_2$, the next time period (the latter would be an ignorable nonresponse model and hence easier to fit). We propose, in Section 3.2.6, a model where nonresponse may occur at one or both time periods, and the nonresponse at each time period will depend on the state of the dyad at the time of the nonresponse. This model (Model 6) is the most realistic of those presented, and as one would expect, it leads to the most difficult forms for parameter estimates.

### 3.2.1 Model 1: Random Link Nonresponse at Time $t_2$ Only

As stated previously, the intention of presenting random nonresponse models is to use them as a benchmark to compare results with the nonrandom nonresponse models. Note that the parameters for the data Markov chain remain the same for the six models presented here, with the same five constraints mentioned previously. Thus, we focus our attention on the parameters for the link nonresponse Markov chain. The four link nonresponse parameters under Model 1 are the following: For each dyad in cell $(k, \ell)$, we have
\[ \zeta_i(k, \ell) = \begin{cases} 1, & \text{if } i = 1 \\ 0, & \text{otherwise} \end{cases} \]  
(3.6)

\[ \rho_{i,j}(k, \ell) = \begin{cases} \rho_{ij}, & \text{if } i = 1 \\ 0, & \text{otherwise} \end{cases} \]

(Note that we are not really counting \( \zeta \) as a parameter here, since it is a constant under Model 1; since Model 1 assumes no nonresponse at time \( t_1 \), \( \zeta \) must be 1. Also, note that throughout this dissertation, it is assumed that constraints 6-10 apply only to those values of \( \zeta \) and \( \rho \) not defined to be zero).

Note that the conditional transition probabilities for the link nonresponse model do not depend in any way on the actual (unobserved) structure of the dyads at either time period, since the nonresponse is assumed to be at random. It is as if we are randomly placing the dyads into three categories as to their response at time \( t_2 \): 1) those with no missing links, 2) those with one missing link, and 3) those with two missing links. For example, \( \rho_{14} \) is the conditional probability of a dyad being unclassified at time \( t_2 \) by having both links missing, given that both links were present at time \( t_1 \).

The likelihood function under Model 1 involves counts from blocks \( X, A, E, \) and \( O \) of the observed dyad transition matrix only, namely those blocks in which missing links occur only at time \( t_2 \). We calculate the degrees of freedom for Model 1 in the following way: There appear to be 36 cells in the observed dyad transition matrix which contain the observed data, with the constraint that the total number of counts in the cells is \( \binom{g}{2} \). In addition to this constraint, there are certain pairs of cells in the observed dyad transition matrix which are indistinguishable under Model 1 upon relabelling of the individuals in any dyad (there are 15 such pairs of cells).
Therefore, there are 16 constraints on the cells of the observed dyad transition matrix under Model 1. We present these cell constraints in the table below; cells with the same number are indistinguishable under Model 1.

<table>
<thead>
<tr>
<th></th>
<th>(0,0)</th>
<th>(1,0)</th>
<th>(0,1)</th>
<th>(1,1)</th>
<th>(M,0)</th>
<th>(M,1)</th>
<th>(0,M)</th>
<th>(1,M)</th>
<th>(M,M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0)</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>(1,0)</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>13</td>
</tr>
<tr>
<td>(0,1)</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>13</td>
</tr>
<tr>
<td>(1,1)</td>
<td>-</td>
<td>6</td>
<td>6</td>
<td>-</td>
<td>14</td>
<td>15</td>
<td>14</td>
<td>15</td>
<td>-</td>
</tr>
</tbody>
</table>

For example, constraint 7 reflects that the first individual in the dyad is a link nonrespondent at time $t_2$, while the other, a nonrespondent, maintains a null relationship. Given that the dyad was in a null state at time $t_1$, it is not important which individual is the link nonrespondent at time $t_2$. In this dissertation, the individuals within the dyad are indistinguishable except for the labelling of the individuals $i$ and $j$, where $i<j$; characteristics of the individual are not taken into account. However, under this set-up, individual characteristics may easily be added to the models; it is for this reason, as well as for ease in using Markov-chain models, that we do not make relabelling simplifications from the beginning.

We also have constraints on the 24 parameters of Model 1; 13 in number. Five of these have been mentioned previously, under the general form of the models: the initial probabilities, and conditional probabilities in each row must sum to 1. The remaining eight constraints apply because upon relabelling of the nodes in any particular dyad, we would expect the probabilities of the then relabelled states to remain the same. For example, suppose the dyad representing Tom and Bob is in state
(1,0) at time \( t_1 \), hence the initial probability would be \( \pi_2 \). (This means Tom likes Bob, but Bob does not like Tom.) If we relabel the two individuals, the state would then be (0,1), and the initial probability would be \( \pi_3 \). We expect these two probabilities to be the same under Model 1. Also, since this is a random link nonresponse model, we assume either individual is equally likely to be a link nonrespondent at time \( t_2 \), regardless of the state of the dyad at time \( t_1 \), or more precisely, the state of either link in the dyad at time \( t_1 \). (This is not true under nonrandom nonresponse models, as we will see later.) The 13 parameter constraints under Model 1 are listed below:

1-5) \[ \sum_k \pi_k = 1, \quad \sum_{\ell} p_{k\ell} = 1, \text{ for } k = 1, 2, \text{ and } 4, \quad \sum_j \rho_{1j} = 1 \]

6) \[ \pi_2 = \pi_3 \]

7) \[ \rho_2 = \rho_3 \]

8) \[ p_{12} = p_{13} \]

9) \[ p_{21} = p_{31} \]

10) \[ p_{22} = p_{33} \]

11) \[ p_{23} = p_{32} \]

12) \[ p_{24} = p_{34} \]

13) \[ p_{42} = p_{43} \] \hspace{1cm} (3.7)

(Note that constraints 9-12 imply \[ \sum_{\ell} p_{k\ell} = 1 \] for \( k = 3 \), hence we eliminate this latter constraint from the list.) Therefore, the degrees of freedom for Model 1 is (36 - 16) - (24 - 13) = 9.

The likelihood function under Model 1 is proportional to the following. For notational purposes, throughout this chapter, we note that unless otherwise stated products and sums over \( k \) and \( \ell \) are for \( k = 1, ..., 4 \) and \( \ell = 1, ..., 4 \), respectively.

\[
\prod_k \prod_{\ell} (\pi_k p_{k,\ell} \rho_{1j})^{x_{k\ell}} \times \prod_k \left[ \pi_k (p_{k,1} + p_{k,2}) \rho_{12} \right]^{x_{k1}} \times \prod_k \left[ \pi_k (p_{k,3} + p_{k,4}) \rho_{12} \right]^{x_{k2}} \\
\times \prod_k \left[ \pi_k (p_{k,1} + p_{k,3}) \rho_{13} \right]^{x_{k1}} \times \prod_k \left[ \pi_k (p_{k,2} + p_{k,4}) \rho_{13} \right]^{x_{k2}} \times \prod_k \left[ \pi_k \left( \sum_{\ell} p_{k,\ell} \right) \rho_{14} \right]^{q_{k1}}
\] \hspace{1cm} (3.8)
(Note that \( \zeta \) does not appear in this likelihood function because it is a constant under Model 1. Also, as a general note, we will always show the likelihood function in its most general form, without substitutions made for constraints such as \( \pi_2 = \pi_3 \); it will be easier to see how the likelihood function is formed if left in its most general state.)

Examining this likelihood function, we see that it can be split into the product of two factors, one involving the \( \pi \)'s and \( p \)'s, and the other involving the \( p \)'s; this is because under Model 1, nonresponse is random; this is a special case of ignorable nonresponse, as defined by Little and Rubin (1987). Under ignorable nonresponse, the data and nonresponse Markov chains are only dependent through quantities which are observable; we note that there is no dependence of the two Markov chains present in Model 1, however, since the nonresponse occurs at random. (Dependence will be seen in Models 3, 4, and 6, which are nonrandom link nonresponse models where the missingness depends on observed and/or unobserved quantities.) Under ignorable nonresponse, the likelihood function splits into two factors and we can find parameter estimates for the two Markov chains separately.

We maximize the likelihood function for Model 1 using Lagrange multipliers for the thirteen parameter constraints under this model. Maximum likelihood estimates (MLE's) for the probabilities for the data Markov chain are given below. For notational ease, we will denote summation over an index by a dot, and let \( r_{k*} \) denote the sum of all counts in row \( k \) of the observed dyad transition matrix

\[
(\text{e.g. } r_{1*} = x_{11} + x_{12} + x_{13} + x_{14} + a_{11} + \ldots + o_{11}).
\]

(3.9)

Also, we let \( r_{**} \) denote the grand total of all dyads (note under Model 1, \( r_{**} = \) the sum of rows 1 through 4 of the observed dyad transition matrix = \( \frac{g}{2} \), since we assume no nonresponse at time \( t_2 \).)
MLE's of the data Markov-chain parameters follow, starting with initial probabilities:

\[
\hat{\pi}_1 = \frac{r_{1*}}{r_{**}} \quad \hat{\pi}_2 = \frac{r_{2*} + r_{3*}}{2r_{**}} \quad \hat{\pi}_4 = \frac{r_{4*}}{r_{**}}.
\] (3.10)

Iterative forms for the MLE's of (conditional) transition probabilities follow. For each \(k = 1, \ldots, 4\), let

\[
c_{k,1} = x_{k,1} + a_{k,1} \left( \frac{p_{k,1}}{p_{k,1} + p_{k,2}} \right) + e_{k,1} \left( \frac{p_{k,1}}{p_{k,1} + p_{k,3}} \right) + o_{k,1} \left( \frac{p_{k,1}}{p_{k*}} \right)
\]

\[
c_{k,2} = x_{k,2} + a_{k,2} \left( \frac{p_{k,2}}{p_{k,1} + p_{k,2}} \right) + e_{k,2} \left( \frac{p_{k,2}}{p_{k,2} + p_{k,4}} \right) + o_{k,1} \left( \frac{p_{k,2}}{p_{k*}} \right)
\]

\[
c_{k,3} = x_{k,3} + a_{k,3} \left( \frac{p_{k,3}}{p_{k,3} + p_{k,4}} \right) + e_{k,1} \left( \frac{p_{k,3}}{p_{k,1} + p_{k,3}} \right) + o_{k,1} \left( \frac{p_{k,3}}{p_{k*}} \right)
\]

\[
c_{k,4} = x_{k,4} + a_{k,4} \left( \frac{p_{k,4}}{p_{k,3} + p_{k,4}} \right) + e_{k,2} \left( \frac{p_{k,4}}{p_{k,2} + p_{k,4}} \right) + o_{k,1} \left( \frac{p_{k,4}}{p_{k*}} \right).
\] (3.11)

For \(p_{k\ell}\)'s subject to relabelling constraints \(p_{k\ell} = p_{k'\ell'}\) we have

\[
\hat{p}_{k\ell} = \frac{c_{k\ell} + c_{k'\ell'}}{r_{k*} + r_{k'*}},
\] (3.12)

for example since \(p_{23} = p_{32}\), \(\hat{p}_{23} = (c_{23} + c_{32}) / (r_{2*} + r_{3*})\). For other \(p_{k\ell}\)'s, we have

\[
\hat{p}_{k\ell} = \frac{c_{k\ell}}{r_{k*}}.
\] (3.13)

We also wish to point out that in this dissertation, for notational simplicity, we refrain from including hat notation where parameters appear on the right-hand side of the
form of the estimate. Also, we will present all estimates in their most easily readable and understandable form; that is, without constraints included. For example, in the estimate for \( p_{12} \), we see both \( p_{12} \) and \( p_{13} \) on the right-hand side; we assume that it is known that \( p_{12} = p_{13} \), but we will not show the simplified version of this estimate with these substitutions made.

These estimates can be interpreted in the following way. \( \hat{\pi}_1 \) is the proportion of observed counts in row 1 of the dyad transition matrix; it estimates the initial probability of a dyad being in state (0,0) at time \( t_1 \). \( \hat{\pi}_2 \) estimates the (conditional) transition probability of a dyad moving from the null state at time \( t_1 \) to state (1,0) at time \( t_2 \). The estimate includes proportions of the counts in partially classified and unclassified columns where the proportions are determined by the estimated probability of being in state (1,0) at time \( t_2 \), given that a count is in a partially classified (block A column 1, or block E column 2) or unclassified (block O) cell.

Transition probabilities for the link nonresponse Markov chain have the following MLE's:

\[
\begin{align*}
\hat{p}_{11} &= \frac{x_{**}}{r_{**}} \\
\hat{p}_{12} &= \hat{p}_{13} = \frac{a_{**} + e_{**}}{2r_{**}} \\
\hat{p}_{14} &= \frac{o_{*1}}{r_{**}}
\end{align*}
\]  \hspace{1cm} (3.14)

These estimates can be interpreted in the following way. Given that each dyad was classified at time \( t_1 \), \( \hat{p}_{11} \) is the proportion of dyads which were completely cross-classified at both time periods; this estimates the conditional probability of a dyad having both links known at time \( t_2 \). \( \hat{p}_{12} = \hat{p}_{13} \) is the proportion of dyads which were partially cross-classified at time \( t_2 \); this estimates the (conditional) probability of a dyad having one link missing at time \( t_2 \). The (conditional) probability of a dyad
having both links missing at time $t_2$ is estimated by $\hat{\rho}_{14}$, the proportion of dyads which were unclassified at time $t_2$.

3.2.2 Model 2: Random Link Nonresponse at Time $t_1$ Only

This model is similar to Model 1, except that the nonresponse occurs at time $t_1$ rather than at time $t_2$. The four link nonresponse parameters under Model 2 follow. For each dyad in cell $(k, \ell)$ of the observed dyad transition matrix, we have:

$$\zeta_i(k, \ell) = \zeta_i, \text{ for } i = 1, \ldots, 4$$

(3.15)

$$\rho_{i,j}(k, \ell) = \begin{cases} 1, & \text{for } j = 1 \text{ and } i = 1, \ldots, 4 \\ 0, & \text{else} \end{cases}$$

(Note that we are not really counting $\rho$ as a parameter here, since it is a constant under Model 2.)

The initial link nonresponse probabilities do not depend on any specific structure of the dyads at either time period since, as in Model 1, nonresponse is assumed to be at random. It is as if we are randomly placing the dyads into three categories (depending on how many links are missing) at time $t_1$, similar to Model 1.

The likelihood function under Model 2 involves counts from cells in blocks X, B, F, and S of the observed dyad transition matrix, since missing data only occurs at time $t_1$. This means a dyad may lose partial or complete row classification at time $t_1$. There are again 36 cells with 16 constraints, as for Model 1, except that data appear in the first four columns of the observed dyad transition matrix, instead of the first four rows, as for Model 1. We have 24 parameters with 13 constraints, where the
parameter constraints differ from Model 1 only in that $\zeta$ appears in Model 2, and $\rho$
does not; hence the degrees of freedom for Model 2 is $(36 - 16) - (24 - 13) = 9$. (This is
the same as the number of degrees of freedom for Model 1.) The 15 pairs of
indistinguishable cells under Model 2 are presented below.

<table>
<thead>
<tr>
<th>Time $t_2$</th>
<th>(0,0)</th>
<th>(0,0)</th>
<th>(0,0)</th>
<th>(0,0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(0,0)</td>
<td></td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>(1,0)</td>
<td></td>
<td>2</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>(0,1)</td>
<td></td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>(1,1)</td>
<td></td>
<td></td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>(M,0)</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>(M,1)</td>
<td>11</td>
<td>12</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>(0,M)</td>
<td>7</td>
<td>9</td>
<td>8</td>
<td>10</td>
</tr>
<tr>
<td>(1,M)</td>
<td>11</td>
<td>13</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>(M,M)</td>
<td></td>
<td>15</td>
<td>15</td>
<td></td>
</tr>
</tbody>
</table>

The parameter constraints for Model 2 are given by the following:

1-5) Same as for Model 1  
6) $\sum \xi_i = 1$  
7-13) Same as for Model 1  
14) $\zeta_2 = \zeta_3$  
(3.16)

The likelihood function under Model 2 is therefore proportional to the following:
\[ \prod_k \prod_\ell (\pi_k p_k, \zeta_1)^{x_k, \ell} \times \prod_\ell \left[ (\pi_{1, \ell} + \pi_{2, \ell}) \zeta_2 \right]^{b_{1, \ell}} \times \prod_\ell \left[ (\pi_{3, \ell} + \pi_{4, \ell}) \zeta_2 \right]^{b_{2, \ell}} \times \prod_\ell \left[ (\pi_{1, \ell} + \pi_{3, \ell}) \zeta_3 \right]^{f_{1, \ell}} \times \prod_\ell \left[ (\pi_{2, \ell} + \pi_{4, \ell}) \zeta_3 \right]^{f_{2, \ell}} \times \prod_\ell \left[ \sum_k \pi_k p_k, \zeta_4 \right]^{g_{\ell, \ell}}. \]

(3.17)

(Note that \( p \) does not appear in this likelihood function, since it is a constant under Model 2.)

We see that because of the ignorable nonresponse, this likelihood function can be split into a product of two factors, one involving the \( \pi \)'s and the \( p \)'s, and the other involving the \( \zeta \)'s. Again, as in Model 1, because nonresponse is modelled as random, we have two independent Markov chains, and parameter estimates can be found for each one separately.

Maximizing the likelihood function for Model 2 using Lagrange multipliers for the thirteen parameter constraints, the following (iterative forms for the) MLE's for the Model 2 data Markov-chain parameters are obtained, beginning with the initial probabilities (here \( r_{**} \) is the sum of columns 1 through 4 of the observed dyad transition matrix = \( \begin{pmatrix} g_1 \\ g_2 \end{pmatrix} \)).

\[ \hat{\pi}_1 = \frac{1}{r_{**}} \sum_\ell \left( x_{1, \ell} + b_{1, \ell} \times \frac{\pi_{1, \ell}}{\pi_{1, \ell} + \pi_{2, \ell}} \right) + \left[ f_{1, \ell} \times \frac{\pi_{1, \ell}}{\pi_{1, \ell} + \pi_{3, \ell}} \right] + \left[ s_{1, \ell} \times \frac{\pi_{1, \ell}}{\sum_k \pi_k p_k, \ell} \right] \]

\[ \hat{\pi}_2 = \frac{1}{2r_{**}} \sum_\ell \left( x_{2, \ell} + b_{1, \ell} \times \frac{\pi_{2, \ell}}{\pi_{1, \ell} + \pi_{2, \ell}} \right) + \left[ f_{2, \ell} \times \frac{\pi_{2, \ell}}{\pi_{2, \ell} + \pi_{4, \ell}} \right] + \left[ s_{1, \ell} \times \frac{\pi_{2, \ell}}{\sum_k \pi_k p_k, \ell} \right] \]
\[
\hat{\pi}_4 = \frac{1}{r_{**}} \sum_\ell \left( x_{4,\ell} + \left[ b_{2,\ell} \times \frac{\pi_{4P4,\ell}}{\pi_{3P3,\ell} + \pi_{4P4,\ell}} \right] + \left[ f_{1,\ell} \times \frac{\pi_{3P3,\ell}}{\pi_{1P1,\ell} + \pi_{3P3,\ell}} \right] + \left[ s_{1,\ell} \times \frac{\pi_{3P3,\ell}}{\sum_k \pi_{kPk,\ell}} \right] \right).
\]

(3.18)

Iterative estimates of (conditional) transition probabilities: for each \( \ell = 1, \ldots, 4 \) we have the following. Let

\[
c_{1,\ell} = x_{1,\ell} + \left[ b_{1,\ell} \times \frac{\pi_{1P1,\ell}}{\pi_{1P1,\ell} + \pi_{2P2,\ell}} \right] + \left[ f_{1,\ell} \times \frac{\pi_{1P1,\ell}}{\pi_{1P1,\ell} + \pi_{3P3,\ell}} \right] + \left[ s_{1,\ell} \times \frac{\pi_{1P1,\ell}}{\sum_k \pi_{kPk,\ell}} \right]
\]

\[
c_{2,\ell} = x_{2,\ell} + \left[ b_{1,\ell} \times \frac{\pi_{2P2,\ell}}{\pi_{1P1,\ell} + \pi_{2P2,\ell}} \right] + \left[ f_{2,\ell} \times \frac{\pi_{2P2,\ell}}{\pi_{2P2,\ell} + \pi_{4P4,\ell}} \right] + \left[ s_{1,\ell} \times \frac{\pi_{2P2,\ell}}{\sum_k \pi_{kPk,\ell}} \right]
\]

\[
c_{3,\ell} = x_{4,\ell} + \left[ b_{2,\ell} \times \frac{\pi_{3P3,\ell}}{\pi_{3P3,\ell} + \pi_{4P4,\ell}} \right] + \left[ f_{1,\ell} \times \frac{\pi_{3P3,\ell}}{\pi_{1P1,\ell} + \pi_{3P3,\ell}} \right] + \left[ s_{1,\ell} \times \frac{\pi_{3P3,\ell}}{\sum_k \pi_{kPk,\ell}} \right]
\]

\[
c_{4,\ell} = x_{4,\ell} + \left[ b_{2,\ell} \times \frac{\pi_{4P4,\ell}}{\pi_{3P3,\ell} + \pi_{4P4,\ell}} \right] + \left[ f_{2,\ell} \times \frac{\pi_{4P4,\ell}}{\pi_{2P2,\ell} + \pi_{4P4,\ell}} \right] + \left[ s_{1,\ell} \times \frac{\pi_{4P4,\ell}}{\sum_k \pi_{kPk,\ell}} \right]
\]

(3.19)
and for \( k = 1, \ldots, 4 \) let \( d_k = \sum \lambda c_{k\ell} \). \hfill (3.20)

Then we have the following (iterative forms for) estimates for the \( p_{ij} \)'s:

\[
\hat{p}_{11} = \frac{c_{11}}{d_1}, \quad \hat{p}_{12} = \frac{c_{12} + c_{13}}{2d_1}, \quad \hat{p}_{14} = \frac{c_{14}}{d_1}
\]

\[
\hat{p}_{21} = \frac{c_{21} + c_{31}}{d_2 + d_3}, \quad \hat{p}_{22} = \frac{c_{22} + c_{32}}{d_2 + d_3}, \quad \hat{p}_{23} = \frac{c_{23} + c_{33}}{d_2 + d_3}, \quad \hat{p}_{24} = \frac{c_{24} + c_{34}}{d_2 + d_3}
\]

\[
\hat{p}_{41} = \frac{c_{41}}{d_4}, \quad \hat{p}_{42} = \frac{c_{42} + c_{43}}{2d_4}, \quad \hat{p}_{44} = \frac{c_{44}}{d_4}.
\] \hfill (3.21)

(Note that these forms reflect the constraints on the \( p_{ij} \)'s, as explained previously for Model 1; for example, since \( p_{12} = p_{13} \), in the (iterative) estimate it turns out that we average the counts in those corresponding cells.

These iterative forms are more complex than those for Model 1; in Model 1 the missing data is from time \( t_2 \), hence we sum across the row to include proportions of the missing data into our probability estimates, while in Model 2, the missing data is from time \( t_1 \), hence we sum down the columns to include proportions of the missing data in the probability estimates. Since the transition probabilities sum to one across rows in a probability transition matrix, while no such constraint exists for the columns, the iterative forms for the initial and transition probability estimates under Model 2 look more complicated than those under Model 1, even though the basic premise is the same for each model.

The iterative estimates for the data Markov-chain probabilities can therefore be interpreted in a way similar to those in Model 1; for example, \( \hat{p}_1 \) is the proportion of
observed counts in row 1 of the dyad transition matrix (i.e. the proportion of dyads in state (0,0) at time $t_1$). Since we model for missing data at time $t_1$, this estimate includes proportions of the partially classified and unclassified row counts from time $t_1$. The proportions are determined by the estimated probability of being in state (0,0) at time $t_1$, given that a count is in a partially classified (block B row 1, or block F row 1) or unclassified (block S) cell of the dyad transition matrix at time $t_1$. $\hat{p}_{12}$ estimates the conditional probability that a dyad moves from state (0,0) at time $t_1$ to the state (1,0) at time $t_2$. Using a Bayes' Rule argument, we see that $\hat{p}_{12}$ is the proportion of the estimate for $\hat{\pi}_1$ which is allotted to the state (1,0) at time $t_2$, since

$$p_{12} = \frac{\mathbb{P}[\text{dyad is in state (0,0) at time } t_1 \text{ and state (1,0) at time } t_2]}{\mathbb{P}[\text{dyad is in state (0,0) at time } t_1]}, \quad \text{where}$$

$$\mathbb{P}[\text{dyad is in state (0,0) at time } t_1] = \sum_k \mathbb{P}[\text{dyad is in (0,0) at time } t_1 \text{ and state } k \text{ at time } t_2]. \quad \text{(3.23)}$$

The initial probabilities for the Markov chain modelling the link nonresponse have the following MLE's:

$$\hat{\xi}_1 = \frac{x_{**}}{r_{**}},$$
$$\hat{\xi}_2 = \frac{b_{**} + f_{**}}{2r_{**}}, \quad \text{(3.24)}$$
$$\hat{\xi}_4 = \frac{s_{1**}}{r_{**}}.$$

We interpret these estimates in the following way: $\hat{\xi}_1$ is the proportion of dyads completely cross-classified at time $t_1$; it estimates the probability of a dyad
responding completely at time $t_1$. $2 \hat{\lambda}_2$ is the proportion of dyads partially classified at time $t_1$; it estimates the probability of a dyad responding partially at time $t_1$ (exactly one link known). $\hat{\lambda}_4$ is the proportion of dyads which were unclassified at time $t_1$; this estimates the probability that both links in the dyad are missing at time $t_1$. Note that since the link nonresponse is assumed to occur at random in Model 2, no other information is used about the missing dyads other than the number of links which were missing. In Model 3, we will incorporate information regarding the missing dyads into our nonrandom link nonresponse model.

3.2.3 Model 3: Nonrandom Link Nonresponse at Time $t_2$ Only

We assume in Model 3 that link nonresponse occurs at time $t_2$, as in Model 1, but that it does not occur at random; rather, the missingness of one or both links in the dyad depends on the state of the dyad at time $t_1$. In our example, suppose again that Bob did not respond regarding his relationship with Tom at month 2 (assuming Tom did respond regarding his relationship with Bob), this results in the dyad for Bob and Tom being in a state of partial classification at time $t_2$. Under this model, we will use the information about the status of the dyad for Tom and Bob at month 1 to help us estimate what Bob would have reported, if he had responded regarding his relationship with Tom.

We begin this section with a short discussion about nonrandom link nonresponse in the social network setting. It may be the case that if an individual is involved in a mutual or null relationship, that individual would be more likely to respond regarding the relationship than if the individual is involved in an asymmetric relationship. Two possible scenarios exist for an asymmetric relationship to be
present: 1) a friendship is not being reciprocated to the individual, or a friendship is not being reciprocated by that individual. In the case of our example, suppose Bob is the potential nonrespondent, and Bob and Tom are involved in an asymmetric relationship. This could be because Tom and Bob just had an argument, and Bob still likes Tom, but Tom no longer wants to reciprocate the friendship. Maybe Bob wouldn’t want to respond regarding this particular relationship at the next time period. This is an example of the first type of asymmetric relationship. As an example of the second type, suppose Tom likes Bob, but Bob is the one who does not want to reciprocate the friendship. This again relates an asymmetric relationship, but with a different tone. Hence it is important to differentiate between the position of individual i in a (1,0) dyad versus a (0,1) dyad when examining link nonresponse, especially in later models where we will discuss link nonresponse over two time periods.

The 16 link nonresponse parameters under Model 3 follow. For each dyad in cell \((k, \ell)\) of the observed dyad transition matrix, we have the following:

\[
\zeta_i(k, \ell) = \begin{cases} 
1, & \text{for } i = 1 \\
0, & \text{else}
\end{cases}
\]

\[
\rho_{i,j}(k, \ell) = \begin{cases} 
\rho_{i,j}(k), & \text{for } i = 1 \text{ and } j = 1, \ldots, 4 \\
0, & \text{else}
\end{cases}
\]

Note that there are only four (conditional) transition nonresponse parameters for each \(k\), \(\rho_{11}(k)\), \(\rho_{12}(k)\), \(\rho_{13}(k)\), and \(\rho_{14}(k)\), since nonresponse only occurs at time \(t_2\), not at time \(t_1\). (Also, we are not really counting \(\zeta\) as a parameter here since it is a constant under Model 3.)

Model 3 is a nonrandom link nonresponse model where nonresponse at time \(t_2\) depends on the state of that dyad at time \(t_1\). Hence, each of the four states of the dyad
transition matrix at time $t_1$ (N, A1, A2, or M), has its own Markov chain to model the link nonresponse at time $t_2$. Therefore under Model 3, we have one Markov chain modelling the data, and four Markov chains modelling the link nonresponse. As in Model 1, data will come from blocks X, A, E, and O of the observed dyad transition matrix.

The degrees of freedom for Model 3 is 4. As for Models 1 and 2, we have 36 cells with 16 constraints. Under Model 3, there are 36 parameters (20 for the data Markov chain, and 4 for each of the 4 Markov chains modelling the link nonresponse) with 20 constraints (11 for the data Markov chain and 9 for the nonresponse Markov chains). These 20 constraints are given by the following:

1-7) $\sum_k \pi_k = 1 \quad \sum_{\ell} p_{k\ell} = 1, \ i = 1, 2, \text{ and } 4 \quad \sum_j \rho_{ij}(k) = 1, \ k = 1, 2, \text{ and } 4$

8) $\pi_2 = \pi_3$

9-14) $p_{12} = p_{13} \quad p_{42} = p_{43}$
$p_{21} = p_{31} \quad p_{22} = p_{33} \quad p_{23} = p_{32} \quad p_{24} = p_{34}$

15-20) $\rho_{12}(1) = \rho_{13}(1) \quad \rho_{11}(2) = \rho_{11}(3) \quad \rho_{12}(2) = \rho_{13}(3)$
$\rho_{12}(3) = \rho_{13}(2) \quad \rho_{14}(2) = \rho_{14}(3) \quad \rho_{12}(4) = \rho_{13}(4)$. \hspace{1cm} (3.26)

(Note, for $k = 3$ we do not need $\sum_{\ell} p_{k\ell} = 1$, and for $k = 3$ we do not need $\sum_j \rho_{ij}(k) = 1$, because constraints 10-13 imply the former, and constraints 16-19 imply the latter.)

Many of these constraints again have to do with the fact that when we relabel the nodes in a dyad, we want the probabilities of that dyad appearing in a certain cell of the observed dyad transition matrix to be the same. For example, in constraint 19,
we have the probabilities of the following situations constrained to be equal: 1) (left-hand side) the dyad is in state (0,1) at time $t_1$, and the first individual in the dyad (individual i) is the nonrespondent at time $t_2$, and 2) (right-hand side) the dyad is in state (1,0) at time $t_1$, and the second individual (individual j) is the nonrespondent at time $t_2$. Note in both cases, it is the individual with the absent arc at time $t_1$ who is the nonrespondent at time $t_2$. It makes sense that these two scenarios should have the same probability, since they differ only in the labelling of the two nodes.

There is another issue to address regarding the parameters under Model 3. Since this is a nonrandom nonresponse model, we also note that $r$ depends on $k$, the state of the dyad at time $t_1$. Hence, we maintain the differences between certain pairs of (conditional) transition probabilities, such as $\rho_{11}(1)$ and $\rho_{11}(2)$, and $\rho_{12}(2)$ and $\rho_{12}(3)$. Note in particular, the situation represented by this latter set of $\rho$'s. $\rho_{12}(2)$ means the dyad is in state (0,1) at time $t_1$ and the second individual (the one with a link present at time $t_1$) is the nonrespondent at time $t_2$. $\rho_{12}(3)$ means the dyad is in state (1,0) at time $t_1$ and the second individual (the one with no link present at time $t_1$) is the nonrespondent at time $t_2$. Under this nonrandom nonresponse model, we want these two probabilities to remain different, since they involve different situations for the individual members of the dyad, as we eluded to previously.

The likelihood function under Model 3 is proportional to the following:

$$
\prod_k \prod_{t} \left[ \pi_k p_{k,t} \rho_{11}(k) \right]^{a_{k1}} \times \prod_k \left[ \pi_k \rho_{12}(k) \sum_{\ell=1,2} p_{k,\ell} \right]^{a_{k1}} \times \prod_k \left[ \pi_k \rho_{12}(k) \sum_{\ell=3,4} p_{k,\ell} \right]^{a_{k2}}
\times \prod_k \left[ \pi_k \rho_{13}(k) \sum_{\ell=1,3} p_{k,\ell} \right]^{e_{k1}} \times \prod_k \left[ \pi_k \rho_{13}(k) \sum_{\ell=2,4} p_{k,\ell} \right]^{e_{k1}} \times \prod_k \left[ \pi_k \rho_{14}(k) \sum_{\ell=1}^4 p_{k,\ell} \right]^{o_{k1}}.
$$

(3.27)
This likelihood function is of the same form as that of Model 1 in the sense that link nonresponse occurs at time \( t_2 \) for both models; it differs from the likelihood function for Model 1 in that the \( \rho \)'s depend on \( k \).

We know that the Markov chains for the data and the link nonresponse are dependent, since the conditional probability of a dyad losing one or both links at time \( t_2 \) depends on the state of the dyad at time \( t_1 \). However, the nonresponse is ignorable because it does not depend on something that was unobserved because of the nonresponse; hence the likelihood function will split into two parts as discussed previously, and estimates for the parameters from the data Markov chain can be found separately from those of the link nonresponse Markov chain.

Iterative estimates for probabilities for the Markov chain generating the data under Model 3 will be the same as those under Model 1 (see (3.10)-(3.13)). Iterative estimates for probabilities for the four Markov chains modelling the link nonresponse follow.

\[
\hat{\rho}_{11}(1) = \frac{x_{1*}}{r_{1*}}, \quad \hat{\rho}_{12}(1) = \frac{a_{1*} + e_{1*}}{2r_{1*}}, \quad \hat{\rho}_{14}(1) = \frac{o_{11}}{r_{1*}} \\
\hat{\rho}_{11}(2) = \frac{x_{2*} + x_{3*}}{r_{2*} + r_{3*}}, \quad \hat{\rho}_{12}(2) = \frac{a_{2*} + e_{3*}}{r_{2*} + r_{3*}}, \quad \hat{\rho}_{13}(2) = \frac{a_{3*} + e_{2*}}{r_{2*} + r_{3*}}, \quad \hat{\rho}_{14}(2) = \frac{o_{21} + o_{31}}{r_{2*} + r_{3*}} \\
\hat{\rho}_{11}(4) = \frac{x_{4*}}{r_{4*}}, \quad \hat{\rho}_{12}(4) = \frac{a_{4*} + e_{4*}}{r_{4*}}, \quad \hat{\rho}_{14}(4) = \frac{o_{41}}{r_{4*}}. \tag{3.28}
\]

We can interpret these iterative estimates in the following way. \( \hat{\rho}_{11}(k) \) is the proportion of dyads from state \( k \) at time \( t_1 \) (\( N, A1, A2, \) or \( M \)) that were completely cross-classified at both time periods; this estimates the probability that a dyad from state \( k \) at time \( t_1 \) will have both links known at both time periods. Note that each
estimate includes counts across row k of the dyad transition matrix only, due to the dependence of the link nonresponse on the state of the dyad at time \( t_1 \). \( \hat{\rho}_{12}(k) = \hat{\rho}_{13}(k) \) is the proportion of dyads from state k at time \( t_1 \) which were partially cross-classified at time \( t_2 \); this estimates the probability that a dyad will have exactly one missing link at time \( t_2 \). The probability of a dyad having both links missing at time \( t_2 \) is estimated by \( \hat{\rho}_{14}(k) \), the proportion of dyads from state k at time \( t_1 \) which were incompletely cross-classified at time \( t_2 \).

3.2.4 Model 4: Nonrandom Link Nonresponse at Time \( t_1 \) Only

Model 4 combines features of Models 2 and 3, and this can be seen in the parameters and their maximum likelihood estimates (and iterative forms). Under Model 4, link nonresponse occurs at time \( t_1 \) only, and the nonresponse is assumed to be nonrandom. As in Model 3, the link nonresponse depends on the state of the dyad at time \( t_1 \). As stated previously, it seems more appropriate in the social network setting to assume that link nonresponse at time \( t_1 \) depends on the observations at time \( t_1 \) rather than on time \( t_2 \). As in our example, if Bob did not respond regarding his relationship with Tom at time \( t_1 \), it seems more realistic to think that he is taking the structure of the dyad at that time period into account, rather than somehow thinking ahead to time \( t_2 \). This will mean nonresponse is nonignorable, since it depends on something which we did not observe, namely the actual state of the dyad at time \( t_1 \). (Bob observed it, but did not report the status of his link.) Therefore, the likelihood function will not split into a product of two factors as before, and the estimates for the Markov chains modelling the data and the link nonresponse are interlaced.
Under Model 4, the missingness depends on the state of the dyad at time $t_1$. The link nonresponse parameters for Model 4 are similar to those under Model 2, except that they depend on $k$, the state of the dyad at time $t_1$. These parameters are the following (for each dyad in cell $(k, \ell)$ of the observed dyad transition matrix, $k, \ell=1, ..., 4$):

$$ \zeta_i(k, \ell) = \zeta_i(k), \text{ for } i = 1, ..., 4 $$

$$ \rho_{i,j}(k, \ell) = \begin{cases} 1, & \text{for } j = 1 \text{ and } i = 1, ..., 4 \\ 0, & \text{else} \end{cases} \quad (3.29) $$

(Note there are sixteen in number; four $\zeta$ parameters for each $k$, with the $\rho$'s simply constant.)

We have 4 Markov chains modelling the link nonresponse, as we did in Model 3; one for each possible state of a dyad at time $t_1$. However, since the nonresponse occurs at different time periods in Model 3 and Model 4, but for both models the nonresponse depends on the state of the dyad at time $t_1$, Models 3 and 4 are not symmetric to each other as Models 1 and 2 were. (We note that if one wished to model the nonresponse at time $t_2$ in Model 3 to depend on the state of the dyad at time $t_2$, one could do that in a similar manner to our Model 4; in this case the degrees of freedom would be 4 also.) We wish to present one model with ignorable nonresponse at one time period and one with nonignorable nonresponse at one time period, to demonstrate the differences in the two situations, and leave it to the individual investigator to chose which model is the most reasonable for the given situation.

Data under Model 4 involves counts from cells in blocks $X, B, F, \text{ and } S$ of the observed dyad transition matrix, the same as under Model 2. The degrees of freedom under Model 4 is 4, as in Model 3. There are 36 parameters with 20 constraints as in
Model 3; 11 for the data Markov chain and 9 for the nonresponse Markov chain.

However, since nonresponse occurs at time $t_1$ instead of time $t_2$, the $\rho$ parameters (and constraints) will be replaced by $\zeta$ parameters (and constraints). The nine constraints on the $\zeta$'s are the following:

\[ \sum_i \zeta_i(k) = 1, \text{ for } k = 1, 2, \text{ and } 4 \]
\[ \zeta_1(2) = \zeta_1(3), \quad \zeta_2(1) = \zeta_3(1), \quad \zeta_2(2) = \zeta_3(3), \quad \zeta_2(3) = \zeta_3(2), \quad \zeta_2(4) = \zeta_3(4), \quad \zeta_4(2) = \zeta_4(3). \quad (3.30) \]

The likelihood function under Model 4 has the following form:

\[
\prod_k \prod_t \left( \pi_k p_t \zeta_1(k) \right)^{x_{k,t}} \times \prod_t \left[ \sum_{k=1,2} \left( \pi_k p_t \zeta_2(k) \right)^{r_{1,t}} \times \prod_t \left[ \sum_{k=3,4} \left( \pi_k p_t \zeta_2(k) \right)^{r_{2,t}} \right] \right] \times \prod_t \left[ \sum_{k=1,3} \left( \pi_k p_t \zeta_3(k) \right)^{r_{3,t}} \right] \times \prod_t \left[ \sum_{k=2,4} \left( \pi_k p_t \zeta_3(k) \right)^{r_{4,t}} \right] \right]. \quad (3.31)
\]

Since the nonresponse is nonignorable, this likelihood function will not split into two factors for the data Markov chain and the link nonresponse Markov chains. This means the maximum likelihood parameter estimates for these two processes must be obtained simultaneously.

The (iterative forms for the) maximum likelihood parameter estimates under Model 4 for the data Markov chain are similar to those found for Model 2, except that the initial link nonresponse parameters $\zeta$ appear in the Model 4 estimates of the $\pi_i$ and $\beta_j$ (they were not present in the $\pi$ and $\rho$ estimates under Model 2). Because of nonrandom nonresponse under Model 4, the $\zeta$ parameters now depend on $k$, (the rows), and hence will not cancel out as they did in the Model 2 estimates (which
contain column sums). The (iterative) parameter estimates for the data Markov chain follow (here $r_{**}$ is the sum of rows 1 through 4 of the observed dyad transition matrix).

\[
\hat{\pi}_1 = \frac{1}{r_{**}} \sum_t \left[ x_{1,t} + \frac{b_{1,t} \pi_{11} \zeta_{22}(1)}{\sum_{k=1,2} \pi_{11} \zeta_{22}(k)} + \frac{f_{1,t} \pi_{11} \zeta_{33}(1)}{\sum_{k=1,3} \pi_{11} \zeta_{33}(k)} + \frac{s_{1,t} \pi_{11} \zeta_{44}(1)}{\sum_k \pi_{11} \zeta_{44}(k)} \right]
\]

\[
\hat{\pi}_2 = \frac{1}{2r_{**}} \sum_t \left[ x_{2,t} + \frac{b_{2,t} \pi_{22} \zeta_{22}(2)}{\sum_{k=1,2} \pi_{22} \zeta_{22}(k)} + \frac{f_{2,t} \pi_{22} \zeta_{33}(2)}{\sum_{k=1,3} \pi_{22} \zeta_{33}(k)} + \frac{s_{1,t} \pi_{22} \zeta_{44}(2)}{\sum_k \pi_{22} \zeta_{44}(k)} \right]
\]

\[
\hat{\pi}_3 = \frac{1}{r_{**}} \sum_t \left[ x_{3,t} + \frac{b_{3,t} \pi_{33} \zeta_{22}(3)}{\sum_{k=3,4} \pi_{33} \zeta_{22}(k)} + \frac{f_{1,t} \pi_{33} \zeta_{33}(3)}{\sum_{k=1,3} \pi_{33} \zeta_{33}(k)} + \frac{s_{1,t} \pi_{33} \zeta_{44}(3)}{\sum_k \pi_{33} \zeta_{44}(k)} \right]
\]

\[
\hat{\pi}_4 = \frac{1}{r_{**}} \sum_t \left[ x_{4,t} + \frac{b_{4,t} \pi_{44} \zeta_{22}(4)}{\sum_{k=3,4} \pi_{44} \zeta_{22}(k)} + \frac{f_{2,t} \pi_{44} \zeta_{33}(4)}{\sum_{k=2,4} \pi_{44} \zeta_{33}(k)} + \frac{s_{1,t} \pi_{44} \zeta_{44}(4)}{\sum_k \pi_{44} \zeta_{44}(k)} \right].
\] (3.32)

Iterative estimates of the (conditional) transition probabilities follow. These estimates are again similar to those for the Model 2 conditional transition probabilities, except that their formulas contain nonresponse parameters, which do not disappear in the situation where nonresponse is nonignorable. Let

\[
c_{1,t} = x_{1,t} + \frac{b_{1,t} \pi_{11} \zeta_{22}(1)}{\sum_{k=1,2} \pi_{11} \zeta_{22}(k)} + \frac{f_{1,t} \pi_{11} \zeta_{33}(1)}{\sum_{k=1,3} \pi_{11} \zeta_{33}(k)} + \frac{s_{1,t} \pi_{11} \zeta_{44}(1)}{\sum_k \pi_{11} \zeta_{44}(k)}
\]
\[ c_{2,t} = x_{2,t} + \frac{b_{1,t} \pi_{2,t} \xi_{2}(2)}{\sum_{k=1,2} \pi_{2,t} \xi_{2}(k)} + \frac{f_{2,t} \pi_{2,t} \xi_{3}(2)}{\sum_{k=2,4} \pi_{k,t} \xi_{3}(k)} + \frac{s_{1,t} \pi_{2,t} \xi_{4}(2)}{\sum_{k} \pi_{k,t} \xi_{4}(k)} \]

\[ c_{3,t} = x_{3,t} + \frac{b_{2,t} \pi_{3,t} \xi_{2}(3)}{\sum_{k=3,4} \pi_{3,t} \xi_{2}(k)} + \frac{f_{1,t} \pi_{3,t} \xi_{3}(3)}{\sum_{k=1,3} \pi_{k,t} \xi_{3}(k)} + \frac{s_{1,t} \pi_{3,t} \xi_{4}(3)}{\sum_{k} \pi_{k,t} \xi_{4}(k)} \]

\[ c_{4,t} = x_{4,t} + \frac{b_{3,t} \pi_{4,t} \xi_{2}(4)}{\sum_{k=3,4} \pi_{4,t} \xi_{2}(k)} + \frac{f_{2,t} \pi_{4,t} \xi_{3}(4)}{\sum_{k=2,4} \pi_{k,t} \xi_{3}(k)} + \frac{s_{1,t} \pi_{4,t} \xi_{4}(4)}{\sum_{k} \pi_{k,t} \xi_{4}(k)} \]

with \( d_{k} = \sum_{t} c_{k,t} \) for \( k = 1, \ldots, 4 \).

Then the (iterative forms for) the MLE's of the (conditional) transition probabilities are the following:

\[ \hat{p}_{11} = \frac{c_{11}}{d_{1}}, \quad \hat{p}_{12} = \frac{c_{12} + c_{13}}{2d_{1}}, \quad \hat{p}_{14} = \frac{c_{14}}{d_{1}}, \]

\[ \hat{p}_{21} = \frac{c_{21} + c_{31}}{d_{2} + d_{3}}, \quad \hat{p}_{22} = \frac{c_{22} + c_{32}}{d_{2} + d_{3}}, \quad \hat{p}_{23} = \frac{c_{23} + c_{33}}{d_{2} + d_{3}}, \quad \hat{p}_{24} = \frac{c_{24} + c_{34}}{d_{2} + d_{3}}, \]

\[ \hat{p}_{41} = \frac{c_{41}}{d_{4}}, \quad \hat{p}_{42} = \frac{c_{42} + c_{43}}{2d_{4}}, \quad \hat{p}_{44} = \frac{c_{44}}{d_{4}}. \]

(3.35)

Interpretations for these iterative forms are analogous to those for Model 2, using Bayes' Rule arguments with conditional probabilities. We therefore move on to the parameter estimates for the Markov chain modeling the link nonresponse, where we find the results are very different from what we have previously seen in Models 1-3. This is due to the nonignorable nonresponse present under Model 4. Under this
model, the nonresponse is based on the state of the dyad at time $t_1$, and we do not have this information, since this is the very information that is missing at time $t_1$. This means that the parameter estimates for the data Markov chain will be present in the formulas for the nonresponse Markov chain parameter estimates. The (iterative forms for the) MLE's link nonresponse Markov chain parameters under Model 4 are given below, beginning with the case when $k=1$.

$$
\hat{\xi}_1(1) = \frac{x_{1*}}{\sum_{l} \left[ \frac{b_{1,l} \pi_{1} p_{1,l} \xi_{2}(1)}{\sum_{k=1}^{2} \pi_{k} p_{k,l} \xi_{2}(k)} + \frac{f_{1,l} \pi_{1} p_{1,l} \xi_{3}(1)}{\sum_{k=1}^{3} \pi_{k} p_{k,l} \xi_{3}(k)} + \frac{s_{1,l} \pi_{1} p_{1,l} \xi_{4}(1)}{\sum_{k} \pi_{k} p_{k,l} \xi_{4}(k)} \right]} = \frac{x_{1*}}{(\hat{\pi}_1)(r_{**})}
$$

$$
\hat{\xi}_2(1) = \sum_{l} \left[ \frac{b_{1,l} \pi_{1} p_{1,l} \xi_{2}(1)}{\sum_{k=1}^{2} \pi_{k} p_{k,l} \xi_{2}(k)} + \frac{f_{1,l} \pi_{1} p_{1,l} \xi_{3}(1)}{\sum_{k=1}^{3} \pi_{k} p_{k,l} \xi_{3}(k)} \right] \left/ \left(2\hat{\pi}_1)(r_{**}) \right/ \right.
$$

$$
\hat{\xi}_4(1) = \sum_{l} \frac{s_{1,l} \pi_{1} p_{1,l} \xi_{4}(1)}{\sum_{k} \pi_{k} p_{k,l} \xi_{4}(k)} \left/ \left(\hat{\pi}_1)(r_{**}) \right/ \right.
$$

(3.36)

where $\hat{\pi}_1$ is given by equation set (3.32).

Similarly, if we let $\hat{\pi}_2$ and $\hat{\pi}_3$ be given by equation set (3.32), we have the following (iterative) estimates for the case when $k=2$:

$$
\hat{\xi}_1(2) = \frac{x_{2*} + x_{3*}}{(\hat{\pi}_2 + \hat{\pi}_3)(r_{**})}
$$

$$
\hat{\xi}_2(2) = \sum_{l} \left[ \frac{b_{1,l} \pi_{2} p_{2,l} \xi_{2}(2)}{\sum_{k=1}^{2} \pi_{k} p_{k,l} \xi_{2}(k)} + \frac{f_{1,l} \pi_{3} p_{3,l} \xi_{3}(3)}{\sum_{k=1}^{3} \pi_{k} p_{k,l} \xi_{3}(k)} \right] \left/ \left(\hat{\pi}_2 + \hat{\pi}_3)(r_{**}) \right/ \right.
$$
\[ \hat{\zeta}_3(2) = \sum \left[ \frac{b_{2,\ell} \pi_3 p_3 \xi_2(3) + f_{2,\ell} \pi_2 p_2 \xi_3(2)}{\sum_{k=1,2} \pi_k p_k \xi_2(k)} + \frac{f_{2,\ell} \pi_2 p_2 \xi_3(2)}{\sum_{k=2,4} \pi_k p_k \xi_3(k)} \right] / (\hat{\pi}_2 + \hat{\pi}_3)(r_{**}) \]

\[ \hat{\zeta}_4(2) = \sum \left[ \frac{s_{1,\ell} \pi_2 p_2 \xi_4(2) + s_{1,\ell} \pi_3 p_3 \xi_4(3)}{\sum_k \pi_k p_k \xi_4(k)} + \frac{s_{1,\ell} \pi_3 p_3 \xi_4(3)}{\sum_k \pi_k p_k \xi_3(k)} \right] / (\hat{\pi}_2 + \hat{\pi}_3)(r_{**}) \]  

(3.37)

and for \( k = 4 \) we have the following (where \( \hat{\pi}_4 \) is given by equation set (3.32)):

\[ \hat{\zeta}_1(4) = \frac{x_{4**}}{(\hat{\pi}_4)(r_{**})} \]

\[ \hat{\zeta}_2(4) = \sum \left[ \frac{b_{2,\ell} \pi_4 p_4 \xi_2(4) + f_{2,\ell} \pi_3 p_3 \xi_3(4)}{\sum_{k=3,4} \pi_k p_k \xi_2(k)} \right] / (2\hat{\pi}_4)(r_{**}) \]

\[ \hat{\zeta}_4(4) = \sum \left[ \frac{s_{1,\ell} \pi_4 p_4 \xi_4(4)}{\sum_k \pi_k p_k \xi_4(k)} \right] / (\hat{\pi}_4)(r_{**}) \]  

(3.38)

We again interpret these (iterative) estimates using a Bayes' Rule argument for conditional probabilities; for example, knowing that the dyad was in state 1 at time \( t_1 \), estimates for the individual probabilities of a dyad responding completely, partially, or not at all at time \( t_1 \) are proportions of those counts that were in state 1 at time \( t_1 \) which were completely classified, partially classified, or unclassified at time \( t_1 \), respectively. The proportions are determined by the estimated probabilities of being in state 1 at time \( t_1 \) given that the dyad provided information regarding both links, only one link, or
none at all. Therefore, we have a conditional probability within a conditional probability.

### 3.2.5 Model 5: Random Link Nonresponse at One or Both Time Periods

The last two models we propose in Chapter 3 are perhaps the most realistic, yet the most complex in form. Models 5 and 6 allow for link nonresponse occurring during one time period (t₁ or t₂), or both time periods (t₁ and t₂). For instance, in our example, Bob may be a nonrespondent for both of the months we are studying, Sue may be a nonrespondent for only one of those months, while Mary may be a nonrespondent for the other month. All of these situations can be handled under both Models 5 and 6. Model 5 assumes the link nonresponse occurs at random, while Model 6 assumes the link nonresponse at either time period depends on the state of the dyad at the time of the nonresponse.

As we will see, assuming link nonresponse may occur at either time period or both time periods means we are observing counts from all of the cells of the dyad transition matrix as shown in Table 1. All formulas for parameter estimates under Models 5 and 6 will hence be more complex, since they include terms from all blocks of this matrix. We also need to examine the underlying structures more carefully to determine the indistinguishable cell constraints and degrees of freedom in these cases.

Under Model 5, we have the following 20 nonresponse parameters:

\[
\begin{align*}
\zeta_i(k, \ell) &= \zeta_i, \ i = 1, ..., 4, \\
\rho_{ij}(k, \ell) &= \rho_{ij}, \ i, j = 1, ..., 4.
\end{align*}
\]
The likelihood function under Model 5 involves counts from all blocks of the observed dyad transition matrix. It has a complicated form, but it is likely that many of the cell counts will be zero, simplifying this function. Note that under Model 5 these nonresponse probabilities do not depend on any structure of the dyads, since the link nonresponse is assumed to be at random. But since link nonresponse is modelled for two time periods, we have many more parameters than we did in the previous models.

Specifically, we have 81 cell counts with 37 constraints (including 36 pairs of indistinguishable cells, and the overall constraint for the total count); this leaves 45 free cells. There are 40 parameters with 22 constraints (leaving 18 free parameters). Hence we have \((81-37) - (40-22) = 26\) degrees of freedom. This is the largest number of degrees of freedom for any of our models. This is due to the large number of cells in the now 9x9 observed dyad transition matrix, and the fact that nonresponse is random, hence Model 5 includes no additional parameters over previous models with only 36 cells. An adjustment to this degrees of freedom may be necessary if certain configurations result in cell counts of zero, leaving some parameters which cannot be estimated. We need to examine the cell counts to check that the chi-square conditions hold for assessing the fit of the model. The constraints on the cells and parameters are identical in form to those under previous random nonresponse models (Models 1 and 2), applied to two time periods. The basic idea is that cells and parameters that are equal, up to relabelling of the nodes in a dyad, should be indistinguishable. The 36 pairs of indistinguishable cells in the 9x9 observed dyad transition matrix are given in Table 3; cells with the same number are considered indistinguishable.

The set of 22 parameter constraints is analogous to those from previous models. We have the usual 2 constraints on the \(\pi\)'s, and the usual 9 constraints on the \(p\)'s, as seen in all previous models. We also have 2 constraints on the \(\zeta\)'s (similar to
those for the \( \pi \)'s) and 9 constraints on the \( \rho \)'s, as seen in Model 3, which are similar to those for the \( \rho \)'s. These constraints are all due to relabelling, as previously discussed. The likelihood function under Model 5 is given in Figure 7. (For notational ease, we have dropped the dependence of \( \zeta \) and \( \rho \) parameters on \((k, \ell)\) in Figure 7.)

<table>
<thead>
<tr>
<th>( t_1 )</th>
<th>((0,0))</th>
<th>((1,0))</th>
<th>((0,1))</th>
<th>( t_2 )</th>
<th>((M,0))</th>
<th>((M,1))</th>
<th>((0,M))</th>
<th>((1,M))</th>
<th>((M,M))</th>
</tr>
</thead>
<tbody>
<tr>
<td>((0,0))</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td>t2</td>
<td>7</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>-</td>
</tr>
<tr>
<td>((1,0))</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>9</td>
<td>10</td>
<td>11</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>((0,1))</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>5</td>
<td>11</td>
<td>12</td>
<td>9</td>
<td>10</td>
<td>15</td>
</tr>
<tr>
<td>((1,1))</td>
<td>-</td>
<td>6</td>
<td>6</td>
<td>t2</td>
<td>13</td>
<td>14</td>
<td>13</td>
<td>14</td>
<td>-</td>
</tr>
<tr>
<td>((M,0))</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
<td>25</td>
<td>26</td>
<td>27</td>
<td>28</td>
<td>35</td>
</tr>
<tr>
<td>((M,1))</td>
<td>20</td>
<td>21</td>
<td>22</td>
<td>23</td>
<td>29</td>
<td>30</td>
<td>31</td>
<td>32</td>
<td>36</td>
</tr>
<tr>
<td>((0,M))</td>
<td>16</td>
<td>18</td>
<td>17</td>
<td>19</td>
<td>27</td>
<td>28</td>
<td>25</td>
<td>26</td>
<td>35</td>
</tr>
<tr>
<td>((1,M))</td>
<td>20</td>
<td>22</td>
<td>21</td>
<td>23</td>
<td>31</td>
<td>32</td>
<td>29</td>
<td>30</td>
<td>36</td>
</tr>
<tr>
<td>((M,M))</td>
<td>-</td>
<td>24</td>
<td>24</td>
<td>-</td>
<td>33</td>
<td>34</td>
<td>33</td>
<td>34</td>
<td>-</td>
</tr>
</tbody>
</table>

Under Model 5, nonresponse is ignorable, hence, this likelihood function will split into two factors, one for each Markov chain. The (iterative) MLE's are complex in form, as expected; for the data Markov chain, we will give the (iterative) MLE's for one initial probability, \( \pi_1 \), and four corresponding conditional transition probabilities, \( p_{11}, p_{12}, p_{13}, \) and \( p_{14} \). Estimates for all other parameters for the data Markov chain are of similar form to those shown, following patterns we have seen developed in previous sections. Here \( r_{xx} \) is the sum of all 9 rows of the observed dyad transition matrix.
Figure 7. Likelihood function under Model 5 for link nonresponse at one or both time periods.
\[ \hat{\pi}_1 = \frac{1}{r_{\ast \ast}} \left\{ x_{1 \ast} + a_{1 \ast} + c_{1 \ast} + o_{1 \ast} + \sum_{t} b_{1t} \left( \frac{\sum_{k=1,2} \pi_{k1t}}{\sum_{k=1,2} \pi_{kkt}} \right) + d_{11} \left( \frac{\sum_{t=1,2} \pi_{1t}}{\sum_{k=1,2} \sum_{t=1,2} \pi_{kt}} \right) \right\} + \]

\[ d_{12} \left( \frac{\sum_{t=3,4} \sum_{k=1,2} \pi_{k1t}}{\sum_{k=1,2} \sum_{t=3,4} \pi_{kkt}} \right) + g_{11} \left( \frac{\sum_{t=1,3} \pi_{1t}}{\sum_{k=1,2} \sum_{t=1,3} \pi_{kt}} \right) + g_{12} \left( \frac{\sum_{t=2,4} \pi_{1t}}{\sum_{k=1,2} \sum_{t=2,4} \pi_{kt}} \right) + \]

\[ u_{11} \left( \frac{\sum_{t} \pi_{1t}}{\sum_{k=1,2} \sum_{t} \pi_{kt}} \right) + f_{1t} \left( \frac{\pi_{1t}}{\sum_{k=1,3} \pi_{kt}} \right) + h_{11} \left( \frac{\sum_{t=1,2} \pi_{1t}}{\sum_{k=1,3} \sum_{t=1,2} \pi_{kt}} \right) + \]

\[ h_{12} \left( \frac{\sum_{t=3,4} \pi_{1t}}{\sum_{k=1,3} \sum_{t=3,4} \pi_{kt}} \right) + t_{11} \left( \frac{\sum_{t=1,3} \pi_{1t}}{\sum_{k=1,3} \sum_{t=1,3} \pi_{kt}} \right) + t_{12} \left( \frac{\sum_{t=2,4} \pi_{1t}}{\sum_{k=1,3} \sum_{t=2,4} \pi_{kt}} \right) + \]

\[ w_{11} \left( \frac{\sum_{t} \pi_{1t}}{\sum_{k=1,3} \sum_{t} \pi_{kt}} \right) + s_{1t} \left( \frac{\pi_{1t}}{\sum_{k=1,3} \pi_{kt}} \right) + v_{11} \left( \frac{\sum_{t=1,2} \pi_{1t}}{\sum_{k=1,3} \sum_{t=1,2} \pi_{kt}} \right) + \]

\[ v_{12} \left( \frac{\sum_{t=3,4} \pi_{1t}}{\sum_{k=1,3} \sum_{t=3,4} \pi_{kt}} \right) + y_{11} \left( \frac{\sum_{t=1,3} \pi_{1t}}{\sum_{k=1,3} \sum_{t=1,3} \pi_{kt}} \right) + y_{12} \left( \frac{\sum_{t=2,4} \pi_{1t}}{\sum_{k=1,3} \sum_{t=2,4} \pi_{kt}} \right) + z_{11} \left( \frac{\sum_{t} \pi_{1t}}{\sum_{k=1,3} \sum_{t} \pi_{kt}} \right) \]

(3.40)

Notice that since link nonresponse is modelled for both time periods, there are many ways in which a dyad in, say, state (0,0) could appear in the observed dyad transition matrix. We now involve many more factors in each estimate because of the many (16) different combinations of response/nonresponse available for both links in a dyad over two time periods. (Iterative forms for the) MLE's for \( p_{11}, ..., p_{14} \) follow.
Let \( c_{11} = x_{11} + a_{11} \frac{P_{11}}{\sum_{\ell=1,2} P_{1\ell}} + e_{11} \frac{P_{11}}{\sum_{\ell=1,3} P_{2\ell}} + o_{11} \frac{P_{11}}{\sum_{\ell} P_{1\ell}} + b_{11} \frac{\pi_1 P_{11}}{\sum_{k=1,2} \pi_k P_{k1}} + 
abla_{11} \frac{\pi_1 P_{11}}{\sum_{k=1,2} \sum_{\ell=1,2} \pi_k P_{k\ell}} + g_{11} \frac{\pi_1 P_{11}}{\sum_{k=1,2} \sum_{\ell=1,3} \pi_k P_{k\ell}} + u_{11} \frac{\pi_1 P_{11}}{\sum_{k=1,2} \sum_{\ell} \pi_k P_{k\ell}} +
abla_{11} \frac{\pi_1 P_{11}}{\sum_{k=1,3} \sum_{\ell=1,2} \pi_k P_{k\ell}} + t_{11} \frac{\pi_1 P_{11}}{\sum_{k=1,3} \sum_{\ell=1,3} \pi_k P_{k\ell}} +
abla_{11} \frac{\pi_1 P_{11}}{\sum_{k=1,3} \sum_{\ell=1,2} \pi_k P_{k\ell}} + v_{11} \frac{\pi_1 P_{11}}{\sum_{k=1,3} \sum_{\ell=1,2} \pi_k P_{k\ell}} + \nabla_{11} \frac{\pi_1 P_{11}}{\sum_{k=1} \sum_{\ell=1,3} \pi_k P_{k\ell}} + w_{11} \frac{\pi_1 P_{11}}{\sum_{k=1,3} \sum_{\ell=1,3} \pi_k P_{k\ell}} + z_{11} \frac{\pi_1 P_{11}}{\sum_{k=1} \sum_{\ell=1,3} \pi_k P_{k\ell}} + (3.41)

\[
c_{12} = x_{12} + a_{12} \frac{P_{12}}{\sum_{\ell=1,2} P_{1\ell}} + e_{12} \frac{P_{12}}{\sum_{\ell=2,4} P_{1\ell}} + o_{11} \frac{P_{12}}{\sum_{\ell} P_{2\ell}} + b_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,2} \pi_k P_{k2}} + d_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,2} \sum_{\ell=1,2} \pi_k P_{k\ell}} + g_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,2} \sum_{\ell=2,4} \pi_k P_{k\ell}} + u_{11} \frac{\pi_1 P_{12}}{\sum_{k=1,2} \sum_{\ell} \pi_k P_{k\ell}} +
abla_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,3} \sum_{\ell=1,2} \pi_k P_{k\ell}} + f_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,3} \sum_{\ell=1,2} \pi_k P_{k\ell}} + t_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,3} \sum_{\ell=1,3} \pi_k P_{k\ell}} +
abla_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,3} \sum_{\ell=1,2} \pi_k P_{k\ell}} + s_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,3} \sum_{\ell=1,2} \pi_k P_{k\ell}} + v_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,3} \sum_{\ell=1,2} \pi_k P_{k\ell}} + \nabla_{12} \frac{\pi_1 P_{12}}{\sum_{k=1} \sum_{\ell=1,3} \pi_k P_{k\ell}} + w_{12} \frac{\pi_1 P_{12}}{\sum_{k=1,3} \sum_{\ell=1,3} \pi_k P_{k\ell}} + z_{12} \frac{\pi_1 P_{12}}{\sum_{k=1} \sum_{\ell=1,3} \pi_k P_{k\ell}} + \nabla_{12} \frac{\pi_1 P_{12}}{\sum_{k=1} \sum_{\ell=1,3} \pi_k P_{k\ell}} + (3.42)
\]
\[ c_{13} = x_{13} + a_{12} \frac{p_{13}}{\sum_{\ell=3,4} p_{1\ell}} + e_{11} \frac{p_{13}}{\sum_{\ell=1,3} p_{1\ell}} + o_{11} \frac{p_{13}}{\sum_{\ell} p_{1\ell}} + b_{13} \frac{\pi_1 p_{13}}{\sum_{k=1,2} \sum_{\ell} \pi_k p_{k\ell}} + \]

\[ d_{12} = \sum_{k=1,2} \sum_{\ell=3,4} \frac{\pi_1 p_{13}}{\pi_k p_{k\ell}} + g_{12} \sum_{k=1,3} \sum_{\ell=2,4} \frac{\pi_1 p_{13}}{\pi_k p_{k\ell}} + u_{11} \sum_{k=1,2} \sum_{\ell} \frac{\pi_1 p_{13}}{\pi_k p_{k\ell}} + \]

\[ f_{13} = \sum_{k=1,3} \frac{\pi_1 p_{13}}{\pi_k p_{k\ell}} + h_{12} \sum_{k=1,3} \sum_{\ell=2,4} \frac{\pi_1 p_{13}}{\pi_k p_{k\ell}} + t_{11} \sum_{k=1,3} \sum_{\ell=1,3} \frac{\pi_1 p_{13}}{\pi_k p_{k\ell}} + \]

\[ w_{11} = \sum_{k=1,3} \sum_{\ell} \frac{\pi_1 p_{13}}{\pi_k p_{k\ell}} + s_{13} \sum_{k=1} \frac{\pi_1 p_{13}}{\pi_k p_{k3}} + v_{12} \sum_{k=1,3} \sum_{\ell=3,4} \frac{\pi_1 p_{12}}{\pi_k p_{k\ell}} + \]

\[ y_{11} = \sum_{k=1,3} \sum_{\ell=1,3} \frac{\pi_1 p_{13}}{\pi_k p_{k\ell}} + z_{11} \sum_{k} \frac{\pi_1 p_{13}}{\pi_k p_{k\ell}} \]

\[ (3.43) \]

\[ c_{14} = x_{14} + a_{12} \frac{p_{14}}{\sum_{\ell=3,4} p_{1\ell}} + e_{12} \frac{p_{14}}{\sum_{\ell=2,4} p_{1\ell}} + o_{11} \frac{p_{14}}{\sum_{\ell} p_{1\ell}} + b_{14} \frac{\pi_1 p_{14}}{\sum_{k=1,2} \sum_{\ell} \pi_k p_{k\ell}} + \]

\[ d_{12} = \sum_{k=1,2} \sum_{\ell=3,4} \frac{\pi_1 p_{14}}{\pi_k p_{k\ell}} + g_{12} \sum_{k=1,3} \sum_{\ell=2,4} \frac{\pi_1 p_{14}}{\pi_k p_{k\ell}} + u_{11} \sum_{k=1,2} \sum_{\ell} \frac{\pi_1 p_{14}}{\pi_k p_{k\ell}} + \]

\[ f_{14} = \sum_{k=1,3} \frac{\pi_1 p_{14}}{\pi_k p_{k3}} + h_{12} \sum_{k=1,3} \sum_{\ell=2,4} \frac{\pi_1 p_{14}}{\pi_k p_{k\ell}} + t_{12} \sum_{k=1,3} \sum_{\ell=1,3} \frac{\pi_1 p_{14}}{\pi_k p_{k\ell}} + \]

\[ w_{11} = \sum_{k=1,3} \sum_{\ell} \frac{\pi_1 p_{14}}{\pi_k p_{k\ell}} + s_{14} \sum_{k=1} \frac{\pi_1 p_{14}}{\pi_k p_{k4}} + v_{12} \sum_{k=1,3} \sum_{\ell=3,4} \frac{\pi_1 p_{12}}{\pi_k p_{k\ell}} + \]

\[ y_{12} = \sum_{k=1,3} \sum_{\ell=2,4} \frac{\pi_1 p_{14}}{\pi_k p_{k\ell}} + z_{11} \sum_{k} \frac{\pi_1 p_{14}}{\pi_k p_{k\ell}} \]

\[ (3.44) \]

Then we have the following (iterative) estimates for the (conditional) probabilities for \( k = 1 \) in the data Markov chain (where \( \hat{\pi}_1 \) is given by the equation (3.40)):

\[ \hat{p}_{11} = \frac{c_{11}}{(\hat{\pi}_1)(r_{**})} \quad \hat{p}_{12} = \frac{c_{12} + c_{13}}{(2\hat{\pi}_1)(r_{**})} \quad \hat{p}_{14} = \frac{c_{14}}{(\hat{\pi}_1)(r_{**})} \]

\[ (3.45) \]
Our usual Baye’s rule arguments apply here, and we see that the numerators of these estimates, $c_{11}$, $c_{12} + c_{13}$, and $c_{14}$ split the counts of those dyads in state 1 at time $t_1$ (estimated by $(\hat{\pi}_1)(r_{**})$) proportionate to those which moved to states 1, 2 (or 3) and 4, respectively, at time $t_2$. Note some terms contain sums across rows ($k$), while others contain sums added across columns ($\ell$), and some contain both. Again, this is because of the many ways that missingness can occur under Model 5; missingness only at time $t_1$ means we add over columns to obtain estimates, and missingness at time $t_2$ only means we add across rows to obtain estimates, so missingness over both time periods would mean summing across both rows and columns to obtain the estimates.

The estimates for the nonresponse Markov chain follow, beginning with the initial probabilities:

$$
\hat{\pi}_1 = \frac{\sum_{i=1}^4 r_{i*}}{r_{**}}, \quad \hat{\pi}_2 = \frac{\sum_{i=5}^8 r_{i*}}{r_{**}}, \quad \hat{\pi}_4 = \frac{r_{9*}}{r_{**}}.
$$

(3.46)

The estimates for the (conditional) transition probabilities are the following:

$$
\hat{\rho}_{11} = \frac{x_{**}}{\sum_{i=1}^4 r_{i*}}, \quad \hat{\rho}_{12} = \frac{a_{**} + e_{**}}{2\sum_{i=1}^4 r_{i*}}, \quad \hat{\rho}_{14} = \frac{c_{11}}{\sum_{i=1}^4 r_{i*}},
$$

$$
\hat{\rho}_{21} = \frac{b_{**} + f_{**}}{\sum_{i=5}^8 r_{i*}}, \quad \hat{\rho}_{22} = \frac{d_{**} + t_{**}}{\sum_{i=5}^8 r_{i*}}, \quad \hat{\rho}_{23} = \frac{g_{**} + h_{**}}{\sum_{i=5}^8 r_{i*}}, \quad \hat{\rho}_{24} = \frac{u_{**} + w_{**}}{\sum_{i=5}^8 r_{i*}}.
$$

$$
\hat{\rho}_{41} = \frac{s_{1*}}{r_{9*}}, \quad \hat{\rho}_{42} = \frac{v_{1*} + y_{1*}}{2r_{9*}}, \quad \hat{\rho}_{44} = \frac{z_{11}}{r_{9*}}.
$$

(3.47)
We interpret these parameters in a similar manner to that in previous models; as before, we see the estimates reflect the constraints due to relabelling.

3.2.6 Model 6: Nonrandom Link Nonresponse at One or Both Time Periods

This model is the most complex of all models presented; it represents the situation where nonresponse is modelled for one or two time periods, and the nonresponse is assumed to be nonrandom. Under Model 6, the nonresponse at time $t_1$ or $t_2$ depends on the state of the dyad at the time of the nonresponse. For example, in our group of four students, considering the dyad between Tom and Bob, we may have the following situation. Suppose Bob and Tom were in a (0,1) relationship (respectively) at time $t_1$, and a (1,0) relationship at time $t_2$. Since Bob is not reciprocating the friendship at time $t_1$, it may be that Tom is more likely to be the nonrespondent at time $t_1$, based on the status of the dyad at time $t_1$. Similarly, at time $t_2$, things have switched; now Tom is the one not reciprocating the friendship. Perhaps Tom is more likely to be the nonrespondent at time $t_2$. The point is, each person in the dyad at each time period examines the status of the dyad, and at that time decides whether or not to be a link nonrespondent, based on that status.

As in Model 5, the data may appear in any of the 16 blocks of the observed dyad transition matrix. Model 6 has 100 parameters; we have $4 \pi$ parameters, $16 \rho$ parameters, $16 \zeta$ parameters, and $64 \rho$ parameters. There are 16 nonresponse Markov chains, one for each combination of dyad states over both time periods. Each nonresponse Markov chain has $4 \zeta$ parameters and $16 \rho$ parameters, but the parameters are not unique for each nonresponse Markov chain; for example, a dyad in state 1 at time $t_1$ and state 2 at time $t_2$ will have the same $\zeta$ parameters as the
nonresponse Markov chain for a dyad in state 1 at time $t_1$ and state 4 at time $t_2$. The 80 parameters included in the 16 nonresponse Markov chains are given below (for any dyad in cell $(k, \ell)$ of the observed dyad transition matrix).

$$
\zeta_i(k, \ell) = \zeta_i(k), \text{ for } i = 1, \ldots, 4 \quad \rho_j(k, \ell) = \rho_j(\ell), \text{ for } i, j = 1, \ldots, 4.
$$

(3.48)

Since the nonresponse depends on unobservable quantities at each time period, it will be nonignorable; this means the likelihood function will not split into 2 separate factors, and the parameters for the data Markov chain and the nonresponse Markov chains must be estimated simultaneously. The likelihood function for Model 6 is the same as that for Model 5, except that the $\zeta$ parameters depend on $k$, and the $\rho$ parameters depend on $\ell$; therefore we will not present it here. (We refer the reader to Figure 7 for the likelihood function under Model 5.)

The degrees of freedom for Model 6 is 2. There are 81 cells with 37 constraints, as for Model 5; this means there are 44 free cells. There are 100 parameters with 58 constraints, leaving a total of 42 free parameters. The usual 11 constraints on the $\pi$ and $p$ parameters also apply under Model 6. The remaining 47 constraints are on the parameters in the 16 nonresponse Markov chains, including the 9 constraints on the $\zeta$ parameters as seen previously in Model 4 (see (3.30)). The remaining 38 constraints on the $\rho$ parameters are unique to Model 6, and are presented below.

For $i = 1, 4$ and $k = 1, 4$:

$$
\zeta_i(2) = \zeta_i(3), \quad \zeta_2(k) = \zeta_3(k)
$$

$$
\rho_{12}(k) = \rho_{13}(k), \quad \rho_{11}(2) = \rho_{11}(3), \quad \rho_{12}(2) = \rho_{13}(3),
$$

$$
\rho_{12}(3) = \rho_{13}(2), \quad \rho_{14}(2) = \rho_{14}(3)
$$

(3.49)
For $i = 2, 3$ and $k = 1, ..., 4$:
\[
\rho_{21}(k) = \rho_{31}(k), \quad \rho_{23}(k) = \rho_{32}(k), \\
\rho_{22}(k) = \rho_{33}(k), \quad \rho_{23}(k) = \rho_{32}(k).
\]

Also,
\[
\sum_j \rho_{1j}(\ell) = 1, \text{ for } \ell = 1, 2, 4 \\
\sum_j \rho_{2j}(\ell) = 1, \text{ for } \ell = 1, ..., 4 \\
\sum_j \rho_{4j}(\ell) = 1, \text{ for } \ell = 1, 2, 4. \tag{3.50}
\]

Therefore, the degrees of freedom is $(81 - 37) - (100 - 58) = 44 - 42 = 2$ under Model 6. As explained for Model 3, which is also a nonrandom nonresponse model, we place constraints on the parameters so that upon relabelling of the individuals in a dyad, we still retain differences such as the status at $t_1$ of a time $t_2$ link nonrespondent. For example, if one link is missing within the Tom-Bob dyad, was it Tom or Bob who was the link nonrespondent? We need to know this for the next time period, because if again a link is missing, there is a difference between Tom being a nonrespondent twice, and Tom being a nonrespondent once and Bob being a nonrespondent once.

Note that we have many more parameters in Model 6 than we had in any of the other models, because of the nonrandom nonresponse at two time periods; this explains the large drop in the degrees of freedom from Model 5 to Model 6.

We again maximize the likelihood function using LaGrange multipliers to obtain MLE’s of the 42 free parameters under Model 6. Again, due to the complexity of the form of these (iterative) estimates, and the fact that we know patterns develop with these (iterative) estimates, as we have seen in previous models, we present a subset of the MLE’s here without loss of generality. We begin with (iterative) estimates for the data Markov chain in the case when $k=1$. 
\[
\hat{\pi}_i = \frac{1}{r^{**}_i} \left\{ x_{i*} + a_{i*} + e_{i*} + o_{i*} + \sum_{\ell} b_{i\ell} \left( \frac{\pi_{i\ell} \zeta_{\ell 3} \rho_{32}(\ell)}{\sum_{k=1,2} \pi_k P_{k,\ell 3} \zeta_{\ell 2}(k) \rho_{32}(\ell)} \right) \right\} + \\
\frac{d_{11} \left( \sum_{\ell=1,2} \sum_{k=1,2} \pi_{k\ell} \zeta_{\ell 2}(k) \rho_{22}(\ell) \right)}{\sum_{k=1,2} \sum_{\ell=1,2} \pi_{k\ell} \zeta_{\ell 2}(k) \rho_{22}(\ell)} + d_{12} \left( \sum_{\ell=3,4} \sum_{k=1,2} \pi_{k\ell} \zeta_{\ell 2}(k) \rho_{22}(\ell) \right) + \\
g_{11} \left( \sum_{\ell=1,3} \sum_{k=1,2} \pi_{k\ell} \zeta_{\ell 2}(k) \rho_{23}(\ell) \right) + g_{12} \left( \sum_{\ell=2,4} \sum_{k=1,2} \pi_{k\ell} \zeta_{\ell 2}(k) \rho_{23}(\ell) \right) + \\
u_{11} \left( \sum_{\ell=1,2} \sum_{k=1,2} \pi_{k\ell} \zeta_{\ell 3}(k) \rho_{34}(\ell) \right) + f_{1\ell} \left( \frac{\pi_{i\ell} \zeta_{\ell 3}(k)}{\sum_{k=1,3} \pi_k P_{k,\ell 3} \zeta_{\ell 3}(k)} \right) + \\
h_{11} \left( \sum_{\ell=1,2} \sum_{k=1,3} \pi_{k\ell} \zeta_{\ell 3}(k) \rho_{32}(\ell) \right) + h_{12} \left( \sum_{\ell=3,4} \sum_{k=1,3} \pi_{k\ell} \zeta_{\ell 3}(k) \rho_{32}(\ell) \right) + \\
t_{11} \left( \sum_{\ell=1,3} \sum_{k=1,2} \pi_{k\ell} \zeta_{\ell 3}(k) \rho_{33}(\ell) \right) + t_{12} \left( \sum_{\ell=2,4} \sum_{k=1,2} \pi_{k\ell} \zeta_{\ell 3}(k) \rho_{33}(\ell) \right) + \\
w_{11} \left( \sum_{\ell=1,3} \sum_{k=1,2} \pi_{k\ell} \zeta_{\ell 3}(k) \rho_{34}(\ell) \right) + s_{1\ell} \left( \frac{\pi_{i\ell} \zeta_{\ell 4}(k) \rho_{41}(\ell)}{\sum_{k=1,3} \pi_k P_{k,\ell 4} \zeta_{\ell 4}(k) \rho_{41}(\ell)} \right) + \\
y_{11} \left( \sum_{\ell=1,3} \sum_{k=1,3} \pi_{k\ell} \zeta_{\ell 3}(k) \rho_{43}(\ell) \right) + y_{12} \left( \sum_{\ell=2,4} \sum_{k=1,3} \pi_{k\ell} \zeta_{\ell 3}(k) \rho_{43}(\ell) \right) + \\
z_{11} \left( \sum_{k=1,2} \sum_{\ell=1,3} \pi_{k\ell} \zeta_{\ell 4}(k) \rho_{44}(\ell) \right) \right\}.
\]

(3.51)

Note that because of the nonrandom nonresponse at one or both time periods, we have
\(\zeta\) and \(\rho\) parameters appearing in this (iterative) estimates (as well as (iterative)
estimates for \(\pi_2\) through \(\pi_4\), which are not shown here). (Iterative forms for the)MLE's for \(p_{11}, \ldots, p_{14}\) follow.
Let $c_{11} =$

\[ x_{11} + a_{11} \sum_{\ell=1,2} \frac{p_{11}p_{12}(1)}{\sum_{\ell=1,2} p_{12}(\ell)} + e_{11} \sum_{\ell=1,3} \frac{p_{11}p_{13}(1)}{\sum_{\ell=1,3} p_{13}(\ell)} + o_{11} \sum_{\ell=1,4} \frac{p_{11}p_{14}(1)}{\sum_{\ell=1,4} p_{14}(\ell)} + \]

\[ b_{11} \sum_{k=1,2} \frac{p_{11}p_{12}(1)p_{22}(1)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + d_{11} \sum_{k=1,2} \frac{p_{11}p_{12}(1)p_{22}(1)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + g_{11} \sum_{k=1,2} \frac{p_{11}p_{12}(1)p_{22}(1)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + \]

\[ u_{11} \sum_{k=1,2} \frac{p_{11}p_{12}(1)p_{22}(1)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + f_{11} \sum_{k=1,3} \frac{p_{11}p_{13}(1)p_{33}(1)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + h_{11} \sum_{k=1,3} \frac{p_{11}p_{13}(1)p_{33}(1)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + \]

\[ t_{11} \sum_{k=1,3} \frac{p_{11}p_{13}(1)p_{33}(1)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + w_{11} \sum_{k=1,3} \frac{p_{11}p_{13}(1)p_{33}(1)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + s_{11} \sum_{k=1,3} \frac{p_{11}p_{13}(1)p_{33}(1)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + \]

\[ y_{11} \sum_{k} \frac{p_{11}p_{13}(1)p_{33}(1)}{\sum_{k} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + z_{11} \sum_{k} \frac{p_{11}p_{13}(1)p_{33}(1)}{\sum_{k} p_{k}p_{r}p_{s} (k)p_{33}(\ell)}. \] (3.52)

Let $c_{12} =$

\[ x_{12} + a_{12} \sum_{\ell=1,2} \frac{p_{12}p_{12}(2)}{\sum_{\ell=1,2} p_{12}(\ell)} + e_{12} \sum_{\ell=1,3} \frac{p_{12}p_{13}(2)}{\sum_{\ell=1,3} p_{13}(\ell)} + o_{12} \sum_{\ell=1,4} \frac{p_{12}p_{14}(2)}{\sum_{\ell=1,4} p_{14}(\ell)} + \]

\[ b_{12} \sum_{k=1,2} \frac{p_{12}p_{12}(2)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + d_{12} \sum_{k=1,2} \frac{p_{12}p_{12}(2)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + g_{12} \sum_{k=1,2} \frac{p_{12}p_{12}(2)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + \]

\[ u_{12} \sum_{k=1,2} \frac{p_{12}p_{12}(2)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + f_{12} \sum_{k=1,3} \frac{p_{12}p_{13}(2)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + h_{12} \sum_{k=1,3} \frac{p_{12}p_{13}(2)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + \]

\[ t_{12} \sum_{k=1,3} \frac{p_{12}p_{13}(2)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + w_{12} \sum_{k=1,3} \frac{p_{12}p_{13}(2)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + s_{12} \sum_{k=1,3} \frac{p_{12}p_{13}(2)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + \]

\[ v_{12} \sum_{k} \frac{p_{12}p_{13}(2)}{\sum_{k} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + y_{12} \sum_{k} \frac{p_{12}p_{13}(2)}{\sum_{k} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + z_{12} \sum_{k} \frac{p_{12}p_{13}(2)}{\sum_{k} p_{k}p_{r}p_{s} (k)p_{33}(\ell)}. \] (3.53)

Let $c_{13} =$

\[ x_{13} + a_{13} \sum_{\ell=1,2} \frac{p_{13}p_{13}(3)}{\sum_{\ell=1,2} p_{13}(\ell)} + e_{13} \sum_{\ell=1,3} \frac{p_{13}p_{13}(3)}{\sum_{\ell=1,3} p_{13}(\ell)} + o_{13} \sum_{\ell=1,4} \frac{p_{13}p_{14}(3)}{\sum_{\ell=1,4} p_{14}(\ell)} + \]

\[ b_{13} \sum_{k=1,2} \frac{p_{13}p_{13}(3)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + d_{13} \sum_{k=1,2} \frac{p_{13}p_{13}(3)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + g_{13} \sum_{k=1,2} \frac{p_{13}p_{13}(3)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + \]

\[ u_{13} \sum_{k=1,2} \frac{p_{13}p_{13}(3)}{\sum_{k=1,2} p_{k}p_{r}p_{s} (k)p_{22}(\ell)} + f_{13} \sum_{k=1,3} \frac{p_{13}p_{13}(3)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + h_{13} \sum_{k=1,3} \frac{p_{13}p_{13}(3)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + \]

\[ t_{13} \sum_{k=1,3} \frac{p_{13}p_{13}(3)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + w_{13} \sum_{k=1,3} \frac{p_{13}p_{13}(3)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + s_{13} \sum_{k=1,3} \frac{p_{13}p_{13}(3)}{\sum_{k=1,3} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + \]

\[ v_{13} \sum_{k} \frac{p_{13}p_{13}(3)}{\sum_{k} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + y_{13} \sum_{k} \frac{p_{13}p_{13}(3)}{\sum_{k} p_{k}p_{r}p_{s} (k)p_{33}(\ell)} + z_{13} \sum_{k} \frac{p_{13}p_{13}(3)}{\sum_{k} p_{k}p_{r}p_{s} (k)p_{33}(\ell)}. \]
\[ u_{11} \sum_{k=1,2} \sum_{\ell} \pi_k \pi_{k3} \zeta_2 (k) \rho_{24}(\ell) + f_{13} \sum_{k=1,3} \pi_k \pi_{k3}(k) + h_{12} \sum_{k=1,3} \sum_{\ell=3} \pi_k \pi_{k3} \zeta_3 (k) \rho_{32}(\ell) + t_{11} \sum_{k=1,3} \sum_{\ell=1,3} \pi_k \pi_{k3} \zeta_3 (k) \rho_{33}(\ell) + w_{11} \sum_{k=1,3} \sum_{\ell=1,3} \pi_k \pi_{k3} \zeta_4 (k) \rho_{34}(\ell) + s_{13} \sum_{k=1,3} \pi_k \pi_{k3} \zeta_4 (k) + v_{12} \sum_{k=5,4} \sum_{\ell=3} \pi_k \pi_{k5} \zeta_4 (k) \rho_{42}(\ell) + y_{11} \sum_{k=5,4} \sum_{\ell=1,3} \pi_k \pi_{k5} \zeta_4 (k) \rho_{43}(\ell) + z_{11} \sum_{k=5,4} \sum_{\ell=1,3} \pi_k \pi_{k5} \zeta_4 (k) \rho_{44}(\ell). \]

(3.54)

Let \( c_{14} = \)

\[ x_{14} + a_{12} \sum_{\ell=3,4} \pi_{14} \rho_{12}(\ell) + c_{12} \sum_{\ell=2,4} \pi_{14} \rho_{13}(\ell) + o_{11} \sum_{\ell} \pi_{14} \rho_{14}(\ell) + b_{14} \sum_{k=1,2} \sum_{\ell} \pi_k \pi_{k5} \zeta_2(k) \rho_{24}(\ell) + d_{12} \sum_{k=1,3} \sum_{\ell=3,4} \pi_k \pi_{k5} \zeta_2(k) \rho_{42}(\ell) + e_{12} \sum_{k=1,3} \sum_{\ell=3,4} \pi_k \pi_{k5} \zeta_4(k) \rho_{42}(\ell) + f_{14} \sum_{k=1,3} \sum_{\ell=1,3} \pi_k \pi_{k5} \zeta_4(k) \rho_{42}(\ell) + g_{12} \sum_{k=1,3} \sum_{\ell=3,4} \pi_k \pi_{k5} \zeta_4(k) \rho_{42}(\ell) + h_{12} \sum_{k=1,3} \sum_{\ell=3,4} \pi_k \pi_{k5} \zeta_4(k) \rho_{42}(\ell) + \]

(3.55)

Then we have the following (iterative) estimates for the (conditional) transition probabilities for \( k=1 \) in the data Markov chain (where \( \hat{\pi}_1 \) is given by equation (3.51)): 

\[ \hat{p}_{11} = \frac{c_{11}}{(\hat{\pi}_1)(r_{11})} \quad \hat{p}_{12} = \frac{c_{12} + c_{13}}{(2\hat{\pi}_1)(r_{11})} \quad \hat{p}_{14} = \frac{c_{14}}{(\hat{\pi}_1)(r_{11})}. \]

(3.56)

Our interpretations follow those for Model 5 estimates for the data Markov chain parameters, except for the presence of the \( \zeta \) and \( \rho \) parameters in the (iterative) estimates due to nonignorable nonresponse possibly present at both time periods.
(Iterative) MLE's for selected parameters of the nonresponse Markov chains under Model 6 are presented below, beginning with the $\zeta(k)$ parameters, when $k=1$.

\[
\zeta_1(l) = \left( x_{1*} + a_{1*} + e_{1*} + o_{1*} \right) \left( r_{1*} \hat{\pi}_1 \right)
\]

\[
\zeta_2(l) = \left\{ \sum_{\ell} b_{1\ell} \left( \frac{\pi_1 p_{1\ell} \zeta_2(1)}{\sum_{k=1,2} \pi_k p_k \zeta_2(k)} \right) + d_{11} \left( \frac{\sum_{\ell=1,2} \pi_1 p_{1\ell} \zeta_2(1) p_{22}(\ell)}{\sum_{k=1,2} \sum_{\ell=1,2} \pi_k p_k \zeta_2(k) p_{22}(\ell)} \right) + \right.
\]

\[
g_{11} \left( \frac{\sum_{\ell=1,3} \pi_1 p_{1\ell} \zeta_2(1) p_{23}(\ell)}{\sum_{k=1,2,\ell=1,3} \pi_k p_k \zeta_2(k) p_{23}(\ell)} \right) + g_{12} \left( \frac{\sum_{\ell=2,4} \pi_1 p_{1\ell} \zeta_2(1) p_{23}(\ell)}{\sum_{k=1,2,\ell=2,4} \pi_k p_k \zeta_2(k) p_{23}(\ell)} \right) + \right.
\]

\[
u_{11} \left( \frac{\sum_{\ell=1,2} \pi_1 p_{1\ell} \zeta_2(1) p_{24}(\ell)}{\sum_{k=1,2} \sum_{\ell=1,2} \pi_k p_k \zeta_2(k) p_{24}(\ell)} \right) + v_{12} \left( \frac{\pi_1 p_{13}(1)}{\sum_{k=1,3} \pi_k \zeta_3(k)} \right) + \right.
\]

\[
u_{11} \left( \frac{\sum_{\ell=1,2} \pi_1 p_{1\ell} \zeta_3(1) p_{32}(\ell)}{\sum_{k=1,3,\ell=1,2} \pi_k p_k \zeta_3(k) p_{32}(\ell)} \right) + h_{12} \left( \frac{\sum_{\ell=3,4} \pi_1 p_{1\ell} \zeta_3(1) p_{32}(\ell)}{\sum_{k=1,3,\ell=3,4} \pi_k p_k \zeta_3(k) p_{32}(\ell)} \right) + \right.
\]

\[
u_{11} \left( \frac{\sum_{\ell=1,3} \pi_1 p_{1\ell} \zeta_3(1) p_{33}(\ell)}{\sum_{k=1,3,\ell=1,3} \pi_k p_k \zeta_3(k) p_{33}(\ell)} \right) + t_{12} \left( \frac{\sum_{\ell=2,4} \pi_1 p_{1\ell} \zeta_3(1) p_{33}(\ell)}{\sum_{k=1,3,\ell=2,4} \pi_k p_k \zeta_3(k) p_{33}(\ell)} \right) + \right.
\]

\[
u_{11} \left( \frac{\sum_{\ell=1,3} \pi_1 p_{1\ell} \zeta_3(1) p_{34}(\ell)}{\sum_{k=1,3,\ell=1,3} \pi_k p_k \zeta_3(k) p_{34}(\ell)} \right) \right\} / \left( r_{1*} \hat{\pi}_1 \right)
\]

\[
\zeta_4(l) = \left\{ \sum_{\ell} s_{1\ell} \left( \frac{\pi_1 p_{1\ell} \zeta_4(1) p_{41}(\ell)}{\sum_{k=1,3} \pi_k p_k \zeta_4(k) p_{41}(\ell)} \right) + y_{11} \left( \frac{\sum_{\ell=1,3} \pi_1 p_{1\ell} \zeta_4(1) p_{43}(\ell)}{\sum_{k=1,3} \pi_k p_k \zeta_4(k) p_{43}(\ell)} \right) \right\} / \left( r_{1*} \hat{\pi}_1 \right)
\]

\[
y_{12} \left( \frac{\sum_{\ell=2,4} \pi_1 p_{1\ell} \zeta_4(1) p_{43}(\ell)}{\sum_{k=2,4} \pi_k p_k \zeta_4(k) p_{43}(\ell)} \right) + z_{11} \left( \frac{\sum_{\ell} \pi_1 p_{1\ell} \zeta_4(1) p_{44}(\ell)}{\sum_{k} \pi_k p_k \zeta_4(k) p_{44}(\ell)} \right) \right\}
\]

\[
(3.57)
\]

\[
(3.58)
\]

\[
(3.59)
\]
We now present the (iterative) MLE's for the \( \rho \) parameters when \( i=1 \) and \( \ell=1 \).

Let 
\[
\hat{d} = x_{1*} + \sum_{k} a_{k1} \left( \frac{p_{k1}}{\sum_{\ell=1,2} p_{k\ell} \rho_{12}(\ell)} \right) + \sum_{k} e_{k1} \left( \frac{p_{k1}}{\sum_{\ell=1,3} p_{k\ell} \rho_{13}(\ell)} \right) + \sum_{k} o_{k1} \left( \frac{p_{k1}}{\sum_{\ell} p_{k\ell} \rho_{14}(\ell)} \right)
\]

(3.60) Then we have the following estimates:

\[
\hat{\rho}_{11}(1) = \frac{x_{1*}}{\hat{d}}
\]

\[
\hat{\rho}_{12}(1) = \frac{\sum_{k} a_{k1} \left( \frac{p_{k1}}{\sum_{\ell=1,2} p_{k\ell} \rho_{12}(\ell)} \right) + \sum_{k} e_{k1} \left( \frac{p_{k1}}{\sum_{\ell=1,3} p_{k\ell} \rho_{13}(\ell)} \right)}{\hat{d}}
\]

\[
\hat{\rho}_{14}(1) = \frac{\sum_{k} o_{k1} \left( \frac{p_{k1}}{\sum_{\ell} p_{k\ell} \rho_{14}(\ell)} \right)}{\hat{d}}.
\]

(3.61)

The greatest difference between Model 5 (the random nonresponse model) and Model 6 (the nonrandom nonresponse model) lies within the estimates for the nonresponse Markov chain parameters. Since under Model 6, nonresponse at either time period depends on the state of the dyad at the time of the nonresponse, which is unobservable, the parameter estimates will be more complicated since we must estimate parameters based on a characteristic that was not observed. This means that in estimating \( \xi_{2}(1) \), for example, we cannot know precisely how many dyads with one link missing at time \( t_{1} \) were in state \((0,0)\) at time \( t_{1} \). Therefore, the estimate includes
the estimated proportion of those dyads in rows 5-8 of the observed dyad transition matrix which were indeed in state (0,0) at time \( t_1 \), given they were in \((M,0) = \{(0,0) \text{ or } (1,0)\}\) or \((0,M) = \{(0,0) \text{ or } (0,1)\}\) at time \( t_1 \). Again, estimated proportions include the estimated conditional probabilities of such events, as we have seen in previous models. (Note that under Model 5, the nonresponse is random, and is estimated by the proportion of dyads that appear in rows 5-8 of the observed dyad transition matrix.)

Similarly, with the (conditional) transitional probabilities, for example \( p_{22}(1) \), it is not known how many dyads that were in state 1 at time \( t_2 \) moved from a state of partial response (one link missing) to a state of partial nonresponse (not necessarily the same). Again, a proportioned estimate is made, based on the conditional probability of a dyad being in state \((0,0)\) at time \( t_2 \), given that it was in state \((M,0)\), or \((0,M)\) at time \( t_2 \). This situation has not appeared previously, because nonrandom nonresponse at time \( t_2 \) in previous models depended on the state of the dyad at time \( t_1 \), hence was ignorable nonresponse. (Note, denominator summations included sums over rows in previous such models, and not columns as they are in these model estimates.)

Iterative procedures for these estimates, and adjusting for sparseness of cells upon fitting these models will be discussed in Chapter 5 with the applications.
CHAPTER IV
MODELS FOR NODE NONRESPONSE UNDER THE
POPULARITY AND EXPANSIVENESS MODELS

4.1 The Two-Stage Process for Node Nonresponse

As mentioned in Chapter 1, we identify two types of nonresponse in the social network setting: link nonresponse (as discussed in Chapter 3), and node nonresponse. In this chapter, we propose random and nonrandom models for node nonresponse in the setting where the data is modelled with the popularity or expansiveness models as studied by Wasserman (1987).

Recall, node nonresponse occurs in a social network when an individual does not respond regarding any relationships with other group members. Therefore, in the digraph, all links emanating from that individual's node are missing. In our example, for instance, Sue may decide to become a nonrespondent at time $t_2$; therefore the links from Sue to Tom, Bob, and Mary will be missing at time $t_2$. Note that node nonresponse is a generalization of link nonresponse with the constraint that any missing link belongs to a set of g-1 missing links, all coming from one of the node nonrespondents.

In this chapter, we initially present models for nonresponse where the model for the data is the popularity model, under which individual i takes the popularity of individual j into account when deciding whether or not to form a relationship with
him/her. In the nonrandom nonresponse models that we propose, we allow node nonresponse to depend on the popularity of the individual making the decision. Similar nonresponse models will be presented where the model for the data (and the nonresponse) is the expansiveness model; however, these models will differ from those of the popularity model because under the both expansiveness model for the data, and the corresponding node nonresponse model, the expansiveness of individual i (the individual making the decision) is taken into account, not the expansiveness of individual j. These latter models will be discussed only very briefly for two reasons. First, they are direct extensions of the nonresponse models under the popularity model, since expansiveness model regards links in the reverse order of the popularity model. Also, the popularity model seems to appear more often in applications than the expansiveness model.

As was described in Chapter 2, under the popularity model, the choice intensity for each individual i regarding a relationship with individual j depends only on the popularity of individual j. The stochastic process for the social network is thus broken into g independent column processes, one for each individual in the social network. Corresponding to this column process, we formulate general models for node nonresponse by any individual in the social network at time t_1 only, t_2 only, or both times. We will propose six models in this chapter; random and nonrandom nonresponse at time t_1 only, random and nonrandom nonresponse at time t_2 only, and random and nonrandom nonresponse at both time periods.

The basic ideas of link nonresponse apply in the case of node nonresponse as well, except for two major differences. First, we are operating under the popularity model for the data, instead of the reciprocity model, hence, the data (including nonresponse) will be summarized in the gχg observed indegree transition matrix as described in Chapter 2 with supplementary row and column matrices (as opposed to
the observed dyad transition matrix under the reciprocity model). The counts appearing in this matrix are counts of individual nodes, not dyads. This means we can discuss nonresponse probabilities for each individual, not just for a pair of individuals as for the dyad processes in Chapter 3.

The second major departure from the reciprocity model is that since we model for node nonresponse rather than link nonresponse, the consequences of the nonresponse regarding other members of the social network are different; while under link nonresponse only the other individual in that particular dyad is affected by a nonrespondent, under node nonresponse, all of the other g-1 nodes are affected. As we will see in the next subsection, node nonresponse will reduce the set of possible cells in which the data will appear in the observed indegree transition matrix to those among two or four block subsets. (Recall that in the link nonresponse models under the reciprocity model for the data, the number of blocks where observed data appeared was also reduced to a subset of blocks, but a larger subset than in this case.)

In modelling the node nonresponse, we again begin with the ideas of Chen and Fienberg (1974), and Stasny (1987) on models for partially cross-classified data, extending the two-stage model for nonresponse. The data will include indegrees for each node in the social network at time $t_1$ and time $t_2$. Again, we think of this data as being the end result of a two-stage process. The first (unobserved) stage involves allocating the data to the $g \times g$ block of the observed indegree transition matrix, as if all indegrees were known. At the second stage, one or more nodes may become nonrespondents at time $t_1$, $t_2$, or both time periods (subject to the constraints of the models). This means all other nodes lose the information from the incoming link involving a nonresponding node. This results in the observed indegree transition matrix with supplementary rows and columns for the partially classified nodes, and completely unclassified nodes. (The same definitions regarding complete and partial
classification in the observed indegree transition matrix apply here to nodes as they did to dyads in the observed dyad transition matrix.)

We again use two Markov chains in modelling the two-stage process: one Markov chain models the data in stage one, and a second Markov chain models the node nonresponse for stage 2, using the model to allocate the nonresponse data back into an estimated version of the gxe indegree transition matrix. From here, the techniques of Wasserman (1980) apply, as described in Chapter 2, for the popularity model (and expansiveness model). (Note that under the expansiveness model, the data would be summarized in an observe outdegree transition matrix, and similar ideas would apply.)

4.1.1 The Observed Indegree Transition Matrix for Node Nonresponse Under the Popularity Model

Under the popularity model, we summarize the observed data for the entire social network over two time periods in the observed indegree transition matrix. It will contain counts of those nodes whose indegree is known, partially known, or completely unknown, at time $t_1$, time $t_2$, or both (depending on the model).

In this setting, the set of possible states for the nodes with partial information follows a certain pattern. This is because once it is known (at Stage 2) how many nodes have not responded (say $M_1$ node nonrespondents at time $t_1$, and $M_2$ node nonrespondents at time $t_2$), we know some information regarding the actual (unobserved) indegree of each node. The actual indegree of any node which responded at time $t_1$ could be at most $M_1$ more than what was observed at time $t_1$, while the actual indegree of any node not responding at time $t_1$ could be at most $M_1-1$ more than what was observed at time $t_1$. Similarly, the observed indegree of a
responding node at time $t_2$ could increase by at most $M_2$, while the observed indegree of a node not responding at time $t_2$ could increase by at most $M_2 - 1$. For example, in the case where $g = 4$ with $M_1 = 1$ and $M_2 = 1$, responding nodes have possible indegree sets $(0,1), (1,2), \text{or} (2,3)$ at $t_1$ and $t_2$ while the nonresponding node(s) have possible indegrees of 0, 1, 2, or 3 at both times. Here, the notation $(0,1)$ means that 0 is the observed indegree for the responding node, and with another possible incoming edge from the one missing node, the indegree could increase to 1; therefore, the possible actual indegrees for that node would be 0 or 1.

Note that as the number of individuals, $g$, in the social network increases, the size of the observed indegree transition matrix ($g \times g$) increases, as do the number of factors in the likelihood function and parameter estimates (note this was not the case in the dyad models of Chapter 3). This means sparse data will be an issue under the popularity model for the data, as opposed to the reciprocity model, where the size of the observed dyad transition matrix is fixed, and as $g$ increases, the number of edges appearing in the cell counts increases.

For notational simplicity, for all further discussions, we will present models in the particular case of four individuals, as in our classroom example; hence $g = 4$. In the general case however, some simplifications do occur because of the block patterns in which the data appear, and the structural zeroes that result in the other blocks. There is also the topic of scarcity of counts in certain cells; we will discuss these issues later.

The observed indegree transition matrix (including node nonresponse data) in the case where $g = 4$ is shown in Table 4. In general, the row and column labels refer to the set of possible indegrees for a node at either time period, given that the first element in the set is the observed indegree at that time period. For example, an indegree counted in row 9 of this matrix has observed indegree 1, and this could
<table>
<thead>
<tr>
<th></th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>(0,1)</th>
<th>(1,2)</th>
<th>(2,3)</th>
<th>(0,1,2)</th>
<th>(1,2,3)</th>
<th>(0,1,2,3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$x_{00}$</td>
<td>$x_{01}$</td>
<td>$x_{02}$</td>
<td>$x_{03}$</td>
<td>$a_{00}$</td>
<td>$a_{01}$</td>
<td>$a_{02}$</td>
<td>$e_{00}$</td>
<td>$e_{01}$</td>
<td>$o_{00}$</td>
</tr>
<tr>
<td>1</td>
<td>$x_{10}$</td>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
<td>$x_{13}$</td>
<td>$a_{10}$</td>
<td>$a_{11}$</td>
<td>$a_{12}$</td>
<td>$e_{10}$</td>
<td>$e_{11}$</td>
<td>$o_{10}$</td>
</tr>
<tr>
<td>2</td>
<td>$x_{20}$</td>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>$x_{23}$</td>
<td>$a_{20}$</td>
<td>$a_{21}$</td>
<td>$a_{22}$</td>
<td>$e_{20}$</td>
<td>$e_{21}$</td>
<td>$o_{20}$</td>
</tr>
<tr>
<td>3</td>
<td>$x_{30}$</td>
<td>$x_{31}$</td>
<td>$x_{32}$</td>
<td>$x_{33}$</td>
<td>$a_{30}$</td>
<td>$a_{31}$</td>
<td>$a_{32}$</td>
<td>$e_{30}$</td>
<td>$e_{31}$</td>
<td>$o_{30}$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$b_{00}$</td>
<td>$b_{01}$</td>
<td>$b_{02}$</td>
<td>$b_{03}$</td>
<td>$d_{00}$</td>
<td>$d_{01}$</td>
<td>$d_{02}$</td>
<td>$g_{00}$</td>
<td>$g_{01}$</td>
<td>$u_{00}$</td>
</tr>
<tr>
<td></td>
<td>$b_{10}$</td>
<td>$b_{11}$</td>
<td>$b_{12}$</td>
<td>$b_{13}$</td>
<td>$d_{10}$</td>
<td>$d_{11}$</td>
<td>$d_{12}$</td>
<td>$g_{10}$</td>
<td>$g_{11}$</td>
<td>$u_{10}$</td>
</tr>
<tr>
<td></td>
<td>$b_{20}$</td>
<td>$b_{21}$</td>
<td>$b_{22}$</td>
<td>$b_{23}$</td>
<td>$d_{20}$</td>
<td>$d_{21}$</td>
<td>$d_{22}$</td>
<td>$g_{20}$</td>
<td>$g_{21}$</td>
<td>$u_{20}$</td>
</tr>
<tr>
<td></td>
<td>$f_{00}$</td>
<td>$f_{01}$</td>
<td>$f_{02}$</td>
<td>$f_{03}$</td>
<td>$h_{00}$</td>
<td>$h_{01}$</td>
<td>$h_{02}$</td>
<td>$t_{00}$</td>
<td>$t_{01}$</td>
<td>$w_{00}$</td>
</tr>
<tr>
<td></td>
<td>$f_{10}$</td>
<td>$f_{11}$</td>
<td>$f_{12}$</td>
<td>$f_{13}$</td>
<td>$h_{10}$</td>
<td>$h_{11}$</td>
<td>$h_{12}$</td>
<td>$t_{10}$</td>
<td>$t_{11}$</td>
<td>$w_{10}$</td>
</tr>
<tr>
<td></td>
<td>$s_{00}$</td>
<td>$s_{01}$</td>
<td>$s_{02}$</td>
<td>$s_{03}$</td>
<td>$v_{00}$</td>
<td>$v_{01}$</td>
<td>$v_{02}$</td>
<td>$y_{00}$</td>
<td>$y_{01}$</td>
<td>$z_{00}$</td>
</tr>
</tbody>
</table>

Table 4. The observed indegree transition matrix for node nonresponse ($g=4$).
increase by as much as two, meaning two of the remaining nodes are missing at time $t_1$; hence this node has indegree in the set $(1,2,3)$ at time $t_1$.

Again, it is useful to think of the observed indegree transition matrix as a partitioned block matrix with the following form (for $g = 4$ individuals)

$$
\begin{bmatrix}
X_{4 \times 4} & A_{4 \times 3} & E_{4 \times 2} & O_{4 \times 1} \\
B_{3 \times 4} & D_{3 \times 3} & G_{3 \times 2} & U_{3 \times 1} \\
F_{2 \times 4} & H_{2 \times 3} & T_{2 \times 2} & W_{2 \times 1} \\
S_{1 \times 4} & V_{1 \times 3} & Y_{1 \times 2} & Z_{1 \times 1}
\end{bmatrix}
$$

Figure 8. Observed indegree transition matrix in partitioned block form ($g = 4$).

The submatrix $X$ contains counts of all completely cross-classified nodes. Note this will contain no counts unless no nodes are missing for at least one time period. In fact, after Stage 2, when it is known how many nodes have not responded at each time period (say $M_1$ and $M_2$), all nonstructural-zero counts will fall in a one, two, or four-submatrix block of the above matrix, as determined by the values of $M_1$ and $M_2$. For example, suppose 2 nodes are missing at time $t_1$ only (hence none are missing at time $t_2$). The responding nodes each have exactly 2 missing incoming links at time $t_1$, and none missing at time $t_2$; this places them in block $F$. The nonresponding nodes have exactly one missing incoming link at time $t_1$, putting them in block $B$. All other blocks would have counts of zero in each cell.

In general, if the node nonresponse occurs at time $t_1$ only, then the data will appear in one of the following one or two-block combinations:
Table 5. Blocks with non-zero cell counts with node nonresponse at $t_1$ only.

<table>
<thead>
<tr>
<th>Blocks</th>
<th>Number of Missing Nodes at $t_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>0</td>
</tr>
<tr>
<td>X, B</td>
<td>1</td>
</tr>
<tr>
<td>B, F</td>
<td>2</td>
</tr>
<tr>
<td>F, S</td>
<td>3</td>
</tr>
<tr>
<td>S</td>
<td>4</td>
</tr>
</tbody>
</table>

When there are two blocks with non-zero cell counts, the nonresponding nodes will appear in the first block of the two-block sequence, and the responding nodes will appear in the second block (since they lose more information). Note for $g = 4$, the highest number of missing nodes is four, and the data fall in block S (note this situation is highly unlikely to occur, and would leave us no information with which to make an estimate in the nonrandom nonresponse case).

Similarly, if the node nonresponse occurs at time $t_2$ only, data will only appear in one of the following one or two-block combinations:

Table 6. Blocks with non-zero cell counts with node nonresponse at $t_2$ only.

<table>
<thead>
<tr>
<th>Blocks</th>
<th>Number of Nodes Missing at Time $t_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>0</td>
</tr>
<tr>
<td>X, A</td>
<td>1</td>
</tr>
<tr>
<td>A, E</td>
<td>2</td>
</tr>
<tr>
<td>E, O</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>4</td>
</tr>
</tbody>
</table>
there are two blocks with non-zero cells, the nonresponding nodes will again appear in the first block of the two-block sequence, while the responding nodes will appear in the second block.

If we model node nonresponse at one or both time periods, data will generally appear in a four-block section of the observed indegree transition matrix, corresponding to the union of the individual one or two-block submatrices for each of the separate time periods, resulting in a one, two, or four-block submatrix. For example, if we model node nonresponse at both time periods and $M_1=2$ while $M_2=1$, then the data will appear in at most a 4-block submatrix containing the blocks B, D, F, and H in the following way:

- B contains counts of nodes missing at both times $t_1$ and $t_2$
- D contains counts of nodes missing at time $t_1$ only
- F contains counts of nodes missing at time $t_2$ only
- H contains counts of nodes not missing at either time period

If we place these blocks in the order in which they appear in the observed indegree transition matrix, we have the following ($r$ denotes node response, $r'$ denotes node nonresponse)

\[
\begin{array}{c|cc}
  & t_2 \\
\hline
  t_1 & r' & r \\
  t_1' & B & D \\
  r & F & H \\
\end{array}
\]

Figure 9. Illustration of a block combination configuration under Models 5 or 6.
This pattern for responding/nonresponding nodes holds for any four-block combination in the observed indegree transition matrix.

In general, for node nonresponse allowed at both time periods, the data will appear in one of the one, two, or four-block combinations as listed in Table 7.

Since our nonrandom node nonresponse models allow the missingness of a node to depend on the popularity (indegree) of that person, we still use the observed indegree transition matrix to obtain our information; no additional information from the adjacency matrix of the social network is needed. The subset of blocks in which the data appear identifies the number of missing nodes at each time period, and the cell location in a particular block of the submatrix allows us to identify the indegree of nodes in any state of response/nonresponse. We therefore have in the observed indegree transition matrix all the information we need to estimate the parameters for the models proposed in this chapter. In Chapter 6, we mention models where this is not necessarily the case.
Table 7. Blocks with non-zero cell counts with node nonresponse at one or both time periods.

<table>
<thead>
<tr>
<th>Blocks</th>
<th>Number of Missing Nodes at Time $t_1$</th>
<th>Number of Missing Nodes at Time $t_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>X, A</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>A, E</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>E, O</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td>O</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>X, B</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>B, F</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>F, S</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>S</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>X, A, B, D</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B, D, F, H</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>A, E, D, G</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>F, H, S, V</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>E, O, G, U</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>D, G, H, T</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>H, T, V, Y</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>G, U, T, W</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>T, W, Y, Z</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>S, V</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>O, U</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>V, Y</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>U, W</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Y, Z</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>W, Z</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Z</td>
<td>4</td>
<td>4</td>
</tr>
</tbody>
</table>
4.2 Probabilities for Stages One and Two Markov Chains

Applying the ideas of Stasny (1987), probabilities associated with each of the two stages of the process involve the initial and transition probabilities from two separate Markov chains: one Markov chain to model the indegree for each node (we will call it the data Markov chain), and another Markov chain to model the node nonresponse (we will call it the node nonresponse Markov chain). Note that since the indegree of any node is between 0 and g-1, we enumerate the states of the data Markov chain accordingly. The probabilities associated with the data Markov chain are given by the following (again, we consider g = 4, but any value of g would be handled in a similar manner without loss of generality).

\[ \pi_k = \text{the initial probability that a node has indegree } k \text{ at time } t_1, \text{ where } k = 0, ..., 3 \text{ represent states for the indegree of that node.} \]

\[ p_{k,\ell} = \text{the (conditional) transition probability of a node having indegree } \ell \text{ at time } t_2 \text{ given that it had indegree } k \text{ at time } t_1 \text{ (} \ell = 0, ..., 3 \text{ and } k = 0, ..., 3). \]

\[ \zeta_r(k, \ell) = \text{the initial probability that the individual is a node respondent at time } t_1. \]

\[ (4.1) \]

The probabilities associated with the node nonresponse Markov chain follow, beginning with estimates of the initial probabilities for an individual in the (k, \ell) cell of the (gxg unobserved) indegree transition matrix after Stage 1:
\[ \zeta_r(k, \ell) = \text{the initial probability of the individual is a node nonrespondent at time } t_1. \] (4.2)

For any node in cell \((k, \ell)\) of the observed indegree transition matrix, there are four (conditional) transition probabilities relating to node nonresponse (in contrast to sixteen for the dyad model, which required information regarding a pair of individuals). We present these transition probabilities below.

\[ \rho_{r,r}(k, \ell) = \text{the (conditional) transition probability of the individual moving from node response at time } t_1 \text{ to node response at time } t_2. \]

\[ \rho_{r'r}(k, \ell) = \text{the (conditional) transition probability of the individual moving from node nonresponse at time } t_1 \text{ to node response at time } t_2. \]

\[ \rho_{r'r'}(k, \ell) = \text{the (conditional) transition probability of the individual moving from node nonresponse at time } t_1 \text{ to node nonresponse at time } t_2. \] (4.3)

For notational ease, in the node nonresponse Markov chain we represent the two states of missingness for the nodes at either time period numerically using the following notation:
1 = r = respondent (indegree partially known at best if missing nodes exist)
2 = r' = nonrespondent (indegree known if no other missing nodes exist, partially known otherwise).

(4.4)

So, for example, \( \rho_{r'r}(k, \ell) \) will be denoted by \( \rho_{2,1}(k, \ell) \).

The following general constraints apply to the parameters in the most general form of the models for this two-stage process (for \( k, \ell = 0, ..., 3 \)):

1) \( \sum_{i=0}^{3} \pi_i = 1 \)

2-5) \( \sum_{j=0}^{3} p_{ij} = 1 \) for each \( i = 0, ..., 3 \)

6) \( \sum_{i=1}^{2} \xi_i(k, \ell) = 1 \)

7-10) \( \sum_{j=1}^{2} \rho_{ij}(k, \ell) = 1 \) for \( i = 1, 2 \)

(4.5)

Note that under the popularity model, the data are the observed indegrees of each individual node (0, 1, ..., g-1); there is no relabelling symmetry present, as there was in the dyad model. Therefore, we will not need to examine possible constraints to handle relabelling, as we did in Chapter 3.

We cannot fit this most general form of the model given the observed data for any \( g \), since it contains too many parameters, as we will illustrate in the following discussion. For the data Markov chain, the maximum number of cells in which the data appear would be the four-block combination in the upper left-hand corner of the observed indegree transition matrix (X, A, B, D); in general, this submatrix would contain

\( g^2 + 2g(g-1) + (g-1)^2 = 4g^2 - 4g + 1 \)

(4.6)
cells, with one constraint, hence there are \(4g^2 - 4g = 4g(g-1)\) free cells.

Regarding the parameters, the data Markov chain contains \(g + g^2\) parameters
(20 in the case where \(g = 4\)), with \(1 + g\) constraints (5 here). This leaves \((g + 1)(g - 1)\)
free parameters. Each \((k, \ell)\) cell of the data Markov chain has its own node
nonresponse Markov chain with six parameters and three constraints, as previously
mentioned. Therefore, there are \(g^2(6-3) = 3g^2\) free node nonresponse parameters.
Therefore, the total number of free parameters is

\[
(g + 1)(g - 1) + 3g^2 = 4g^2 - 1. \tag{4.7}
\]

Hence, the maximum number of degrees of freedom possible for this most
general form of the model would be \(4g(g-1) - (4g^2 - 1) = 1 - 4g < 0\) for any \(g > 0\).
Hence, the most general form of these models cannot be fit. Note that the degrees of
freedom may be altered because there are only \(g\) counts allocated to the \(4g^2 - 4g + 1\)
cells and some probabilities may not be able to be estimated; these issues need to be
addressed on an individual case basis. In the next section we propose constraints on
the \(\zeta\) and \(\rho\) parameters that allow us to fit certain models.

4.3 Some Node Nonresponse Models Under the Popularity Model

for the Data Markov Chain

Under the popularity model for the data, we propose the following six Markov-
chain models for link nonresponse:
Model 1  Random node nonresponse at time $t_2$ only.
Model 2  Random node nonresponse at time $t_1$ only.
Model 3  Nonrandom node nonresponse at time $t_2$ only; probability of missingness depends on the indegree state of the node at time $t_1$.
Model 4  Nonrandom node nonresponse at time $t_1$ only; probability of missingness depends on the indegree state of the node at time $t_1$.
Model 5  Random node nonresponse at either or both time periods.
Model 6  Nonrandom node nonresponse at either or both time periods; probability of missingness depends on the indegree state of the node at the time of the nonresponse.

We present these six models in this chapter. Estimators for other similar models may be developed in the same manner as for the models described here.

We will consider each model in turn, finding degrees of freedom and forms for parameter estimates. In Chapter 5, we will present iterative procedures needed for obtaining certain parameter estimates, where we present applications of the models. As with the models for link nonresponse, Model 6 is the most realistic of those presented, and as one would expect, it leads to the most difficult forms for parameter estimates.
4.3.1 Model 1: Random Node Nonresponse at Time \( t_2 \) Only (Popularity Model)

The data Markov chain parameters remain the same as stated in (4.5), with the \( (g + 1 = 5) \) constraints as mentioned. Therefore, we will focus on the node nonresponse Markov chain. (Note, unless otherwise stated in Chapter 4, \( k \) and \( \ell = 0, 1, 2, 3 \) for all discussions, since \( g = 4 \).) The four node nonresponse parameters under Model 1 are the following:

\[
\zeta_i(k, \ell) = \begin{cases} 
1, & \text{for } i = 1 \\
0, & \text{for } i = 2 
\end{cases} 
\tag{4.8}
\]

\[
\rho_{ij}(k, \ell) = \begin{cases} 
\rho_{ij}, & \text{for } i = 1, \text{ and } j = 1, 2 \\
0, & \text{otherwise} 
\end{cases} 
\]

Again, note \( \zeta \) is not really considered a parameter here; since node nonresponse is not allowed at \( t_1 \), \( \zeta \) remains constant under Model 1. As discussed in Chapter 3, since this is a random nonresponse model, \( \rho \) does not depend on the social network data in any way, it is assumed that the nodes are randomly placed into a response or nonresponse category at Stage 2, according to some probability \( \rho \).

Data under Model 1 appear in one of the one or two-block combinations as seen in Table 6. In the likelihood function, we use a set of collectively exhaustive indicator functions (actually indicator constants, given the data) to denote the blocks in which the data appear. For example,

\[
I_{X,A} = \begin{cases} 
1, & \text{if data appear in blocks } X \text{ and } A \\
0, & \text{otherwise} 
\end{cases} \tag{4.9}
\]
Degrees of freedom depend on the number of structural zeroes which appear in the data; hence, on the blocks in which the data appear. For a given set of observed data, possible degrees of freedom under Model 1 are given below. Note that for all of Chapter 4, the constraints general constraints (4.5) apply to each model respectively, regarding summation of relevant parameters to one. No additional constraints are required for the node nonresponse models. Therefore we will not present them for each model separately. Degrees of freedom are calculated in the same manner as for Chapter 3 models.

Table 8. Degrees of freedom under node nonresponse Model 1.

<table>
<thead>
<tr>
<th>Number of Missing Nodes</th>
<th>Blocks</th>
<th>Number of Free Cells</th>
<th>Number of Parameters</th>
<th>Degrees of Freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>X</td>
<td>16-1=15</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>X,A</td>
<td>28-1=27</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>A,E</td>
<td>20-1=19</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>E,O</td>
<td>12-1=11</td>
<td>16</td>
<td>&lt;0</td>
</tr>
<tr>
<td>4</td>
<td>O</td>
<td>4-1=3</td>
<td>16</td>
<td>&lt;0</td>
</tr>
</tbody>
</table>

Note that the saturated model (where the degrees of freedom = 0) fits the model exactly; in this situation, we do not count the nonresponse parameters since there is no nonresponse data with which to estimate them (also note no nonresponse parameter estimates are needed). In the case where we have many missing nodes, relative to the number of nodes in the social network (in general, when \( M_1 \) and/or \( M_2 \) is large relative to \( g \)), the model cannot be fit. Since it is reasonable to assume that in most situations we will not have a high percentage of missing nodes for any particular time
period, we are not distressed by these results. For most cases, we will be able to fit the model given the observed data. Note that in the link nonresponse models of Chapter 3, we do not have the degrees of freedom depend on the number of nodes; however, we do need to adjust the degrees of freedom for sparseness of data in the cells, as we also have to do here. We will discuss this issue in Chapter 5, when we apply the models to actual data.

The likelihood function under Model 1 is proportional to the following (again, note that $g = 4$).

\[
\prod_{k=0}^{3} \prod_{\ell=0}^{3} \left[ \pi_k (p_{k,\ell} + (p_{k,\ell+1} + p_{k,\ell+2}) (p_{11} + p_{12} I_{AX}) \right] \times \prod_{k=0}^{3} \prod_{\ell=0}^{2} \left[ \pi_k (p_{k,\ell} + p_{k,\ell+1} + p_{k,\ell+2}) (p_{11} + p_{12} I_{AX}) \right]
\times \prod_{k=0}^{3} \left[ \pi_k (p_{k,0} + p_{k,1} + p_{k,2} + p_{k,3}) (p_{11} + p_{12} I_{EO}) \right]^{e_{u}}
\times \prod_{k=0}^{3} \left[ \pi_k (p_{k,0} + p_{k,1} + p_{k,2} + p_{k,3}) (p_{11} + p_{12} I_{EO}) \right]^{o_{u}}
\]

(4.10)

Note that given the data, the indicator functions are simply constant, 0 or 1. Note also that $\zeta$ does not appear in the likelihood function since it is a constant under Model 1. Since nonresponse is ignorable here, the likelihood function can be split into two factors, as discussed previously. We again maximize this function using Lagrange multipliers for the constraints. The following MLE's are obtained, beginning with those for the initial probabilities for the data Markov chain. (Again, the $r_{**}$ of Chapter 3 applies here, except that here $r_{**} = g$.)

\[
\hat{\pi}_k = \frac{r_{k**}}{r_{**}}, \text{ for } k = 0, ..., 4
\]

(4.11)

The (iterative) MLE's for the (conditional) transition probabilities follow.
\[
\hat{p}_{k,0} = \frac{1}{r_{k*}} \left\{ x_{k0} + a_{k0} \left( \frac{p_{k0}}{\sum_{\ell=0}^{\infty} p_{k\ell}} \right) + e_{k0} \left( \frac{p_{k0}}{\sum_{\ell=0}^{\infty} p_{k\ell}} \right) + o_{k0} \left( \frac{p_{k0}}{\sum_{\ell=0}^{\infty} p_{k\ell}} \right) \right\}
\]

\[
\hat{p}_{k,1} = \frac{1}{r_{k*}} \left\{ x_{k1} + a_{k0} \left( \frac{p_{k1}}{\sum_{\ell=0}^{\infty} p_{k\ell}} \right) + a_{k1} \left( \frac{p_{k1}}{\sum_{\ell=1}^{\infty} p_{k\ell}} \right) + e_{k0} \left( \frac{p_{k1}}{\sum_{\ell=0}^{\infty} p_{k\ell}} \right) + e_{k1} \left( \frac{p_{k1}}{\sum_{\ell=1}^{\infty} p_{k\ell}} \right) + o_{k1} \left( \frac{p_{k1}}{\sum_{\ell=0}^{\infty} p_{k\ell}} \right) \right\}
\]

\[
\hat{p}_{k,2} = \frac{1}{r_{k*}} \left\{ x_{k2} + a_{k1} \left( \frac{p_{k2}}{\sum_{\ell=1}^{\infty} p_{k\ell}} \right) + a_{k2} \left( \frac{p_{k2}}{\sum_{\ell=2}^{\infty} p_{k\ell}} \right) + e_{k0} \left( \frac{p_{k2}}{\sum_{\ell=0}^{\infty} p_{k\ell}} \right) + e_{k1} \left( \frac{p_{k2}}{\sum_{\ell=1}^{\infty} p_{k\ell}} \right) + o_{k0} \left( \frac{p_{k2}}{\sum_{\ell=0}^{\infty} p_{k\ell}} \right) \right\}
\]

\[
\hat{p}_{k,3} = \frac{1}{r_{k*}} \left\{ x_{k3} + a_{k2} \left( \frac{p_{k3}}{\sum_{\ell=2}^{\infty} p_{k\ell}} \right) + e_{k2} \left( \frac{p_{k3}}{\sum_{\ell=2}^{\infty} p_{k\ell}} \right) + o_{k0} \left( \frac{p_{k3}}{\sum_{\ell=0}^{\infty} p_{k\ell}} \right) \right\}.
\]

(4.12)

As in Chapter 3, for notational simplicity, we omit the hat notation on the right side of these iterative forms. The interpretation for these parameters is similar to that for the corresponding data Markov chain parameters under Model 1 of Chapter 3. For the initial probabilities \(\pi_k\), the estimate is simply the proportion of counts that appear in
the kth row at time \( t_1 \). For the conditional transition probabilities, \( p_{k\ell} \), we again use a Bayes' Rule argument; each estimate includes proportions of counts in partially classified and unclassified columns where the proportions are determined by the estimated probability of being in state \( \ell \) at time \( t_2 \), given that a count is in a partially classified or unclassified cell containing \( \ell \) as a member of the indegree set.

The (conditional) transition probabilities for the node nonresponse Markov chain have the following MLE's:

\[
\hat{\rho}_{12} = \frac{1}{r_{**}} \left( \sum_{k=0}^{3} \sum_{\ell=0}^{3} x_{k\ell} I_{XA} + \sum_{k=0}^{3} \sum_{\ell=0}^{2} a_{k\ell} I_{AE} + \sum_{k=0}^{3} \sum_{\ell=0}^{1} e_{k\ell} I_{EO} + \sum_{k=0}^{3} o_{k0} I_{O} \right)
\]

\[
\hat{\rho}_{11} = \frac{1}{r_{**}} \left( \sum_{k=0}^{3} \sum_{\ell=0}^{3} x_{k\ell} I_{X} + \sum_{k=0}^{3} \sum_{\ell=0}^{2} a_{k\ell} I_{XA} + \sum_{k=0}^{3} \sum_{\ell=0}^{1} e_{k\ell} I_{AE} + \sum_{k=0}^{3} o_{k0} I_{EO} \right).
\]  

(4.13)

We interpret these estimates in the following way. Given that each node was classified at time \( t_1 \), \( \rho_{11} \) is estimated by the proportion of nodes which were completely cross-classified. Notice this depends on which blocks the data appear in; for example, a node appearing in block \( X \) is a responding node at time \( t_2 \) (hence completely cross-classified) if the data appear only in block \( X \), but it is a nonresponding node at time \( t_2 \) (hence partially cross-classified) if the data appear in blocks \( X \) and \( A \).

### 4.3.2 Model 2: Random Node Nonresponse at Time \( t_1 \) Only (Popularity Model)

This model is similar to Model 1, except that the node nonresponse occurs at time \( t_1 \) rather than at time \( t_2 \). The two node nonresponse parameters under Model 2
follow. For each node in cell \((k, \ell)\) of the observed indegree transition matrix we have the following:

\[
\zeta_i(k, \ell) = \zeta_i, \text{ for } i = 1, 2
\]

\[
\rho_{ij}(k, \ell) = \begin{cases} 
1, & \text{for } j = 1 \text{ and } i = 1, 2 \\
0, & \text{else}
\end{cases}
\]  

(4.14)

(Again, we are not really counting \(\rho\) as a parameter here, since it is a constant under Model 2.) Also, note that the initial node nonresponse probabilities do not depend on any specific structure of the social network at either time period.

Data under Model 2 appear in one of the one or two-block combinations as shown in Table 9. Degrees of freedom, as for Model 1, depend on the number of structural zeroes which appear in the data. For a given set of observed data, possible degrees of freedom under Model 2 are the following:

<table>
<thead>
<tr>
<th>Number of Missing Nodes</th>
<th>Number of Blocks</th>
<th>Number of Free Cells</th>
<th>Number of Free Parameters</th>
<th>Degrees of Freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>X</td>
<td>16-1=15</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>X,B</td>
<td>28-1=27</td>
<td>16</td>
<td>11</td>
</tr>
<tr>
<td>2</td>
<td>B,F</td>
<td>20-1=19</td>
<td>16</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>F,S</td>
<td>12-1=11</td>
<td>16</td>
<td>&lt;0</td>
</tr>
<tr>
<td>4</td>
<td>S</td>
<td>4-1=3</td>
<td>16</td>
<td>&lt;0</td>
</tr>
</tbody>
</table>

Table 9. Degrees of freedom under node nonresponse Model 2.
(Note these are similar to the degrees of freedom possibilities under Model 1, as expected.) Again, the saturated model fits the data perfectly, and for three or more missing nodes, the model cannot be fit.

The likelihood function under Model 2 is proportional to the following (with \( g = 4 \)):

\[
\prod_{k=0}^{3} \prod_{\ell=0}^{3} \left[ \pi_k p_k,\ell \left( x_{2I_{XB}} + \zeta_1 I_{X} \right) \right]^{x_{k\ell}} \times \prod_{k=0}^{3} \prod_{\ell=0}^{3} \left[ \left( \pi_k p_k,\ell + \pi_{k+1} p_{k+1,\ell} \right) \left( x_{2I_{BF}} + \zeta_1 I_{XB} \right) \right]^{b_{k\ell}}
\times \prod_{k=0}^{3} \prod_{\ell=0}^{3} \left[ \left( \pi_k p_k,\ell + \pi_{k+1} p_{k+1,\ell} + \pi_{k+2} p_{k+2,\ell} \right) \left( x_{2I_{FS}} + \zeta_1 I_{BF} \right) \right]^{f_{k\ell}}
\times \prod_{\ell=0}^{3} \left( \left( \pi_0 p_{0,\ell} + \pi_1 p_{1,\ell} + \pi_2 p_{2,\ell} + \pi_3 p_{3,\ell} \right) \left( x_{2I_{S}} + \zeta_1 I_{FS} \right) \right)^{s_{k\ell}}.
\]

(4.15)

Again, as in Model 1, the nonresponse is ignorable, and we can estimate the parameters for the data and node nonresponse Markov chains separately. The following (iterative) MLE's for the data Markov chain are obtained, beginning with those for the initial probabilities:

\[
\hat{\pi}_0 = \frac{1}{r_{**}} \left\{ x_{0*} + \sum_{\ell=0}^{3} \frac{\pi_0 p_{0,\ell}}{\sum_{k=0}^{3} \pi_k p_k,\ell} + \sum_{\ell=0}^{3} \frac{f_{0,\ell}}{2} \frac{\pi_0 p_{0,\ell}}{\sum_{k=0}^{3} \pi_k p_k,\ell} + \sum_{\ell=0}^{3} \frac{s_{0,\ell}}{3} \frac{\pi_0 p_{0,\ell}}{\sum_{k=0}^{3} \pi_k p_k,\ell} \right\}
\]

\[
\hat{\pi}_1 = \frac{1}{r_{**}} \left\{ x_{1*} + \sum_{\ell=0}^{3} \frac{\pi_1 p_{1,\ell}}{\sum_{k=0}^{3} \pi_k p_k,\ell} + \sum_{\ell=0}^{3} \frac{b_{1,\ell}}{2} \frac{\pi_1 p_{1,\ell}}{\sum_{k=0}^{3} \pi_k p_k,\ell} + \sum_{\ell=0}^{3} \frac{f_{1,\ell}}{2} \frac{\pi_1 p_{1,\ell}}{\sum_{k=0}^{3} \pi_k p_k,\ell} + \sum_{\ell=0}^{3} \frac{3}{\sum_{k=0}^{3} \pi_k p_k,\ell} \frac{\pi_1 p_{1,\ell}}{3} \right\}
\]
\[ \hat{\pi}_2 = \frac{1}{r_{**}} \left\{ x_{2*} + \frac{3}{\sum_{k=1}^{2} \pi_k p_{k,\ell}} \pi_{22,\ell} + \frac{3}{\sum_{k=2}^{2} \pi_k p_{k,\ell}} \pi_{22,\ell} + \frac{3}{\sum_{k=0}^{2} \pi_k p_{k,\ell}} \pi_{22,\ell} + \frac{3}{\sum_{k=0}^{2} \pi_k p_{k,\ell}} \pi_{22,\ell} \right\} \]

\[ \hat{\pi}_3 = \frac{1}{r_{**}} \left\{ x_{3*} + \frac{3}{\sum_{k=2}^{3} \pi_k p_{k,\ell}} \pi_{33,\ell} + \frac{3}{\sum_{k=1}^{3} \pi_k p_{k,\ell}} \pi_{33,\ell} + \frac{3}{\sum_{k=0}^{3} \pi_k p_{k,\ell}} \pi_{33,\ell} \right\} \] (4.16)

The (iterative) MLE's for the (conditional) transition probabilities follow:

\[ \hat{p}_{0,\ell} = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0\ell} + b_{0\ell} \frac{1}{\sum_{k=0}^{1} \pi_k p_{k,\ell}} \pi_{00,\ell} + \frac{1}{\sum_{k=0}^{1} \pi_k p_{k,\ell}} \pi_{00,\ell} + \frac{1}{\sum_{k=0}^{1} \pi_k p_{k,\ell}} \pi_{00,\ell} \right\} \]

\[ \hat{p}_{1,\ell} = \frac{1}{(\hat{\pi}_1)(r_{**})} \left\{ x_{1\ell} + b_{1\ell} \frac{1}{\sum_{k=0}^{3} \pi_k p_{k,\ell}} \pi_{11,\ell} + \frac{1}{\sum_{k=1}^{3} \pi_k p_{k,\ell}} \pi_{11,\ell} + \frac{1}{\sum_{k=0}^{3} \pi_k p_{k,\ell}} \pi_{11,\ell} \right\} \]
\[ \hat{p}_{2\ell} = \frac{1}{(\tilde{r}_2)(r_{**})} \left\{ x_{2*} + b_{1\ell} \frac{\pi_{22p_{2\ell}}}{\sum_{k=1}^{2} \pi_{k2p_{k,\ell}}} + b_{2\ell} \frac{\pi_{22p_{2\ell}}}{\sum_{k=2}^{3} \pi_{k2p_{k,\ell}}} + f_{0\ell} \frac{\pi_{22p_{2\ell}}}{\sum_{k=0}^{3} \pi_{k2p_{k,\ell}}} \right\} \]

\[ + \left( s_{0\ell} \frac{\pi_{22p_{2\ell}}}{\sum_{k=0}^{3} \pi_{k2p_{k,\ell}}} \right) \]

\[ p_{3\ell} = \frac{1}{(\tilde{r}_3)(r_{**})} \left\{ x_{3*} + b_{2\ell} \frac{\pi_{33p_{3\ell}}}{\sum_{k=2}^{3} \pi_{k3p_{k,\ell}}} + f_{1\ell} \frac{\pi_{33p_{3\ell}}}{\sum_{k=1}^{3} \pi_{k3p_{k,\ell}}} + s_{0\ell} \frac{\pi_{33p_{3\ell}}}{\sum_{k=0}^{3} \pi_{k3p_{k,\ell}}} \right\}. \quad (4.17) \]

The interpretation for these estimates is similar to those for Model 1, based on conditional probabilities. Again, these iterative forms are more complex than those for Model 1. As we saw in Chapter 3, because we have missing data at time 1, we sum down the columns to include proportions of the missing data in our estimates, and we do not have a constraint on the column totals as we do on the row totals.

The initial probabilities for the Markov chain modelling the node nonresponse have the following MLE's:

\[ \hat{\zeta}_1 = \frac{1}{r_{**}} \left( \sum_{k=0}^{3} \sum_{\ell=0}^{3} x_{k\ell}I_{X} + \sum_{k=0}^{3} \sum_{\ell=0}^{2} b_{k\ell}I_{XB} + \sum_{k=0}^{3} \sum_{\ell=0}^{1} f_{k\ell}I_{BF} + \sum_{k=0}^{3} s_{k0}I_{FS} \right) \]

\[ = \frac{1}{r_{**}} \left( x_{**}I_{X} + b_{**}I_{XB} + f_{**}I_{BF} + s_{**}I_{FS} \right) \]

\[ \hat{\zeta}_2 = \frac{1}{r_{**}} \left( x_{**}I_{XB} + a_{**}I_{BF} + c_{**}I_{FS} + s_{**}I_{S} \right). \quad (4.18) \]
The interpretation of these estimates follows that of Model 1; the estimates are the proportions of completely (and partially) cross-classified nodes at time $t_1$. No other information is used, since node nonresponse is assumed to be random under this model.

### 4.3.3 Model 3: Nonrandom Node Nonresponse at Time $t_2$ Only

(Popularity Model)

Under this nonrandom node nonresponse model, node nonresponse is not assumed to be random; rather, it depends on the popularity of the individual (node) at time $t_2$, reflected by the indegree of the corresponding node at time $t_2$. In our example of Chapter 1, since Sue was chosen by no one, she is unpopular at time $t_1$; hence, she may be more likely than the others to be a node nonrespondent at time $t_1$. We try to capture this difference with this model. The node nonresponse under Model 3 depends on an observable quantity, hence is ignorable.

The eight node nonresponse Markov chain parameters under Model 3 are the following (for each node in cell $(k, \ell)$ of the observed indegree transition matrix):

\[
\zeta_i(k, \ell) = \begin{cases} 
1, & \text{for } i = 1 \\
0, & \text{else}
\end{cases}
\]

\[
\rho_{ij}(k, \ell) = \begin{cases} 
\rho_{ij}(k), & \text{for } i = 1 \text{ and } j = 1, 2 \\
0, & \text{else}
\end{cases}.
\]

(4.19)

Note, each initial state $k$ of the data Markov chain has a separate node nonresponse Markov chain; hence, we have four node nonresponse Markov chains in this case.
The blocks in which the data will appear, and the possible resulting degrees of freedom are given below.

<table>
<thead>
<tr>
<th>Number of Missing Nodes</th>
<th>Number of Blocks</th>
<th>Number of Free Cells</th>
<th>Number of Parameters</th>
<th>Degrees of Freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>X</td>
<td>16-1=15</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>X,A</td>
<td>28-1=27</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>A,E</td>
<td>20-1=19</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>E,O</td>
<td>12-1=11</td>
<td>19</td>
<td>&lt; 0</td>
</tr>
<tr>
<td>4</td>
<td>O</td>
<td>4-1=3</td>
<td>19</td>
<td>&lt; 0</td>
</tr>
</tbody>
</table>

Note, we have a saturated model (which fits the data exactly) in the cases where there are zero or two missing nodes; in the case where we have more than two missing nodes, we cannot fit the model.

The likelihood function under node nonresponse Model 3 is the same as that for Model 1 except for the dependence of $\rho$ on $k$; hence we will not show it here. We refer the reader to (4.10) for the likelihood function under Model 1. Also, the MLE's for the $\pi$ and $p$ parameters are the same as those under Model 1 (see (4.11)-(4.12)). Therefore, we present the MLE's for the node nonresponse Markov-chain parameters under Model 3.

\[
\hat{\rho}_{1|1}(k) = \frac{1}{r_{k^{*}}} (x_{k^{*}}I_X + a_{k^{*}}I_{XA} + e_{k^{*}}I_{AE} + o_{k^{*}}I_{EO})
\]
We interpret these estimates in the same manner as for Model 3 of Chapter 3. Recall that when the data are in blocks \( X \) and \( A \), for example, there is one nonrespondent at time \( t_2 \), the node in block \( X \) is the node nonrespondent; it loses no information and are hence is completely cross-classified. Those nodes in block \( A \) are the respondents at time \( t_2 \); they lose one piece of information from the node nonrespondent, and are hence partially cross-classified. The estimates are found by taking proportions across each row only, since we are using the information that the node was in row \( k \) at time \( t_1 \) to reflect the popularity of the individual.

### 4.3.4 Model 4: Nonrandom Node Nonresponse at Time \( t_2 \) Only

**Popularity Model**

Model 4 combines features of Models 2 and 3, and this is reflected in the estimates. Under this model, node nonresponse occurs at time \( t_2 \) only, and is assumed to depend on the state of the indegree of the node at time \( t_2 \), the time of the nonresponse. This means that the response is nonignorable, and the estimates for the data and node nonresponse Markov chains will have to be made simultaneously. As stated in Chapter 3, this setting where the nonrandom nonresponse depends on the structure of the social network at the time of the nonresponse seems to be more appropriate; however, we present the two settings for illustrations and comparison of ignorable and nonignorable response models.

The eight node nonresponse Markov-chain parameters under Model 4 are the following (for each node in cell \((k, \ell)\) of the observed indegree transition matrix):

\[
\hat{\rho}_{12}(k) = \frac{1}{r_k}(x_k \cdot I_{X_k} + a_k \cdot I_{A_k} + e_k \cdot I_{E_k} + o_k \cdot I_{O_k}).
\]
\[ \zeta_{i}(k, \ell) = \begin{cases} \zeta_{i}(k), & \text{for } i = 1, 2 \\ 0, & \text{else} \end{cases} \]

\[ \rho_{ij}(k, \ell) = \begin{cases} 1, & \text{for } i = 1 \text{ and } j = 1, 2 \\ 0, & \text{else} \end{cases} \]

(4.21)

We have four Markov chains modelling the node nonresponse, as we did in Model 3; one for each possible indegree of a node at time \( t_1 \). The blocks in which the data will appear, and the possible resulting degrees of freedom for Model 4 are given below.

<table>
<thead>
<tr>
<th>Number of Missing Nodes</th>
<th>Number of Blocks</th>
<th>Number of Free Cells</th>
<th>Number of Parameters</th>
<th>Degrees of Freedom</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>X</td>
<td>16-1=15</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>X,B</td>
<td>28-1=27</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>B,F</td>
<td>20-1=19</td>
<td>19</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>F,S</td>
<td>12-1=11</td>
<td>19</td>
<td>&lt; 0</td>
</tr>
<tr>
<td>4</td>
<td>S</td>
<td>4-1=3</td>
<td>19</td>
<td>&lt; 0</td>
</tr>
</tbody>
</table>

(Note the possible degrees of freedom under Model 4 are the same as those under Model 3, as expected, since the number of cells and parameters match in the two models.) The likelihood function under Model 4 resembles that of Model 2, except for dependence of the \( \zeta \) parameters on \( k \); therefore we will not present it here (see (4.15) for the likelihood function under Model 2).
The (iterative forms for the) MLE's under Model 4 for the data Markov chain are similar to those for Model 2, except that the initial node nonresponse parameters \( \zeta \) appear in the Model 4 estimates of the \( \pi \) and \( p \) parameters (they were not present in the corresponding estimates under Model 2). Because of nonrandom nonresponse under Model 4, the \( \zeta \) parameters now depend on \( k \), (the row classification) and hence will not cancel out as they did in the Model 2 estimates (which contain column sums). Hence, we present only those (iterative forms for the) MLE's for parameters in the case where \( k = 0 \) and \( \ell = 0, ..., 3 \) without loss of generality. We begin with the (iterative) MLE's for the data Markov chain parameters:

\[
\hat{\pi}_0 = \frac{1}{r_{**}} \left\{ x_{0*} + \sum_{\ell=0}^{3} b_{0,\ell} \frac{\pi_0 p_{0,\ell} (\zeta_1(0)I_{BF} + \zeta_2(0)I_{XB})}{\sum_{k=0}^{1} \pi_k p_{k,\ell} (\zeta_1(k)I_{BF} + \zeta_2(k)I_{XB})} + \sum_{\ell=0}^{3} s_{0,\ell} \frac{\pi_0 p_{0,\ell} (\zeta_1(0)I_S + \zeta_2(0)I_{FS})}{\sum_{k=0}^{3} \pi_k p_{k,\ell} (\zeta_1(k)I_S + \zeta_2(k)I_{FS})} \right\}
\]

(4.22)

\[
\hat{p}_{0,\ell} = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0,\ell} + b_{0,\ell} \frac{\pi_0 p_{0,\ell} (\zeta_1(0)I_{BF} + \zeta_2(0)I_{XB})}{\sum_{k=0}^{1} \pi_k p_{k,\ell} (\zeta_1(k)I_{BF} + \zeta_2(k)I_{XB})} + \sum_{\ell=0}^{3} s_{0,\ell} \frac{\pi_0 p_{0,\ell} (\zeta_1(0)I_S + \zeta_2(0)I_{FS})}{\sum_{k=0}^{3} \pi_k p_{k,\ell} (\zeta_1(k)I_S + \zeta_2(k)I_{FS})} \right\}
\]

(4.23)

The interpretation for these estimates is similar to those for Model 2, based on the Bayes' Rule argument with conditional probabilities. We therefore move on to the
parameter estimates for the node nonresponse Markov-chain initial probabilities, where the results will be different from previous models, due to the nonignorable nonresponse under Model 4. We present these selected (iterative) forms below in the case where \( k=0 \); others follow in similar form.

\[
\hat{\zeta}_1(0) = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0*}I_X + \sum_{\ell=0}^3 b_{0\ell} \frac{\pi_0P_{0,\ell}\zeta_1(0)I_{XB}}{\sum_{k=0}^3 \pi_kP_{k,\ell}(\zeta_2(k)I_{BF} + \zeta_1(k)I_{XB})} + \sum_{\ell=0}^3 \frac{\pi_0P_{0,\ell}\zeta_1(0)I_{BF}}{\sum_{k=0}^3 \pi_kP_{k,\ell}(\zeta_2(k)I_{FS} + \zeta_1(k)I_{FS})} \right\}
\]

\[
\hat{\zeta}_2(0) = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0*}I_{XB} + \sum_{\ell=0}^3 b_{0\ell} \frac{\pi_0P_{0,\ell}\zeta_2(0)I_{BF}}{\sum_{k=0}^3 \pi_kP_{k,\ell}(\zeta_2(k)I_{BF} + \zeta_1(k)I_{XB})} + \sum_{\ell=0}^3 \frac{\pi_0P_{0,\ell}\zeta_2(0)I_{FS}}{\sum_{k=0}^3 \pi_kP_{k,\ell}(\zeta_2(k)I_{FS} + \zeta_1(k)I_{FS})} \right\}.
\]

We interpret these parameter estimates again using a Bayes' Rule argument for conditional probabilities; here, given that the node had indegree 0 at time \( t_1 \), estimates for the probability of a node responding at time \( t_1 \) are given by the estimated proportions of those counts which were partially classified at time \( t_1 \) (assuming \( M_1>0 \)) that were actually in state 0 at time \( t_1 \). Notice again, in this setting, a responding node at time \( t_2 \) is partially classified at time \( t_2 \) if there is at least one node missing, since it
loses information from the missing node(s). A nonresponding node may be completely classified at time $t_2$, if there are no other missing nodes at that time, or partially classified at time $t_2$, if there are other missing nodes (since it loses information regarding its indegree). If more than one node is missing, all nodes are partially classified at time $t_2$, yet the missing nodes lose the least amount of information, becoming 'less partially classified'.

4.3.5 Model 5: Random Node Nonresponse at One or Both Time Periods

(Popularity Model)

As seen with the link nonresponse models of Chapter 3, the last two models we present in this chapter are perhaps the most realistic, and the most complex in form. Each model allows for node nonresponse to occur at one or both time periods; Model 5 assumes the node nonresponse occurs at random while Model 6 assumes the node nonresponse depends on the indegree of the node (i.e. the popularity of the individual) at the time of the nonresponse.

Under both Models 5 and 6, the data appear in any of the one, two, or four-block combinations as seen in Table 7. Due to the unwieldiness of the resulting likelihood function, we present each of these models in a special case only; to obtain estimates for other cases, similar procedures can be followed, parallel to what has been previously presented in this chapter. Therefore, for all discussion involving Models 5 and 6, we will assume $M_1=1$ and $M_2=1$ ($g = 4$ individuals throughout); that is, there is one missing node at each time period (not necessarily the same one). Note this means there will be no indicator functions in the likelihood function or in the estimates.

Given this special case, under Model 5 there are 26 parameters with 8 constraints, leaving 18 free parameters; the usual 20 parameters from the data Markov
chain, and 6 from the node nonresponse Markov chain. These latter parameters are the following:

\[
\zeta_i(k, \ell) = \begin{cases} 
\zeta_i, & \text{for } i = 1, 2 \\
0, & \text{else}
\end{cases} 
\]

\[
\rho_{ij}(k, \ell) = \begin{cases} 
\rho_{ij}, & \text{for } i = 1, 2 \text{ and } j = 1, 2 \\
0, & \text{else}
\end{cases}.
\]

Data will appear in blocks X, A, B, and D in the form of Figure 10,

\[\begin{array}{c|cc}
  & t_2 \\
  & \hline 
  t_1 & r' & r \\
 r' & X & A \\
r & B & D 
\end{array}\]

Figure 10. Illustration of the block combination configuration for given example.

where \(r\) denotes response and \(r'\) denotes nonresponse at each time period. Notice that the node appearing in block X did not respond at either time period (as opposed to Chapter 3, where respondents at both time periods appeared in the X submatrix). In this setting, there are 49 cells, with one constraint, leaving 48 free cells. Therefore, the degrees of freedom for this particular situation under Model 5 is \((49 - 1) - (26 - 8) = 30\); but note that there are only four counts so the degrees of freedom will almost surely have to be adjusted. The likelihood function under Model 5 for this setting is proportional to the following:
\[
\prod_{k=0}^{3} \prod_{\ell=0}^{3} \left[ \pi_k p_{k,\ell} s_2 p_{22} \right]^{x_{2w}} \times \prod_{k=0}^{3} \prod_{\ell=0}^{2} \left[ \pi_k \sum_{v=0}^{1} (p_{k,\ell+v} s_2 p_{21}) \right]^{a_{2w}} \\
\times \prod_{k=0}^{2} \prod_{\ell=0}^{3} \left[ \sum_{u=0}^{1} (\pi_{k+u} p_{k+u,\ell} s_{12} p_{12}) \right]^{b_{2w}} \times \prod_{k=0}^{2} \prod_{\ell=0}^{2} \left[ (\sum_{u=0}^{1} \sum_{v=0}^{1} \pi_{k+u} p_{k+u,\ell+v} s_{11} p_{11}) \right]^{c_{2w}}.
\]

(4.26)

Under Model 5, nonresponse is ignorable, hence the estimates for the data and node nonresponse Markov chains can be made separately. We present the (iterative) MLE's for the data Markov-chain initial probabilities below.

\[
\hat{\pi}_0 = \frac{1}{r_{**}} \left\{ x_{0*} + a_{0*} + \sum_{\ell=0}^{3} b_{0\ell} \frac{\pi_{0} p_{0,\ell}}{\sum_{k=0}^{3} \pi_k p_{k,\ell}} + \sum_{\ell=0}^{2} d_{0\ell} \frac{\pi_{0} (p_{0,\ell} + p_{0,\ell+1})}{\sum_{k=0}^{1} \pi_k (p_{k,\ell} + p_{k,\ell+1})} \right\}
\]

\[
\hat{\pi}_1 = \frac{1}{r_{**}} \left\{ x_{1*} + a_{1*} + \sum_{\ell=0}^{3} b_{1\ell} \frac{\pi_{1} p_{1,\ell}}{\sum_{k=0}^{3} \pi_k p_{k,\ell}} + \sum_{\ell=0}^{2} d_{1\ell} \frac{\pi_{1} (p_{1,\ell} + p_{1,\ell+1})}{\sum_{k=1}^{2} \pi_k (p_{k,\ell} + p_{k,\ell+1})} \right\}
\]

\[
\hat{\pi}_2 = \frac{1}{r_{**}} \left\{ x_{2*} + a_{2*} + \sum_{\ell=0}^{3} b_{2\ell} \frac{\pi_{2} p_{2,\ell}}{\sum_{k=1}^{3} \pi_k p_{k,\ell}} + \sum_{\ell=0}^{2} d_{2\ell} \frac{\pi_{2} (p_{2,\ell} + p_{2,\ell+1})}{\sum_{k=2}^{3} \pi_k (p_{k,\ell} + p_{k,\ell+1})} \right\}
\]
\[ \hat{\pi}_3 = \frac{1}{r_{**}} \left\{ x_{3*} + a_{3*} + \sum_{\ell=0}^{3} b_{2\ell} \frac{\pi_{3,\ell}}{\sum_{k=2}^{3} \pi_{k,k,\ell}} + \sum_{\ell=0}^{2} d_{2\ell} \frac{\pi_{3}(p_{3,\ell} + p_{3,\ell+1})}{\sum_{k=2}^{3} \pi_{k,k,\ell}(p_{k,\ell} + p_{k,\ell+1})} \right\}. \] (4.27)

The (iterative) MLE's for the (conditional) transition probabilities are given below for \(k = 0\) and 1; estimates for \(k = 2\) are similar to those for \(k = 1\), and estimates for \(k = 3\) are similar to those for \(k = 0\); this is because of the symmetry present in the way the indegrees appear in the observed indegree transition matrix. First, for \(k = 0\) we have:

\[ \hat{p}_{0,0} = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0,0} a_{00} \frac{p_{0,0}}{\sum_{\ell=0}^{1} p_{0,\ell}} + b_{00} \frac{\pi_{0,0,0}}{\sum_{k=0}^{1} \pi_{k,k,0}} + d_{00} \frac{\pi_{0,0,0}}{\sum_{k=0}^{1} \sum_{\ell=0}^{1} \pi_{k,k,\ell}} \right\} \]

\[ \hat{p}_{0,1} = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0,1} a_{00} \frac{p_{0,1}}{\sum_{\ell=0}^{1} p_{0,\ell}} + a_{01} \frac{p_{0,1}}{\sum_{k=0}^{1} \pi_{k,k,1}} + b_{01} \frac{\pi_{0,0,1}}{\sum_{k=0}^{1} \sum_{\ell=0}^{1} \pi_{k,k,\ell}} \right\} \]

\[ \hat{p}_{0,2} = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0,2} a_{01} \frac{p_{0,2}}{\sum_{\ell=1}^{2} p_{0,\ell}} + a_{02} \frac{p_{0,2}}{\sum_{k=0}^{2} \pi_{k,k,2}} + b_{02} \frac{\pi_{0,0,2}}{\sum_{k=0}^{2} \sum_{\ell=1}^{2} \pi_{k,k,\ell}} \right\} \]

\[ \hat{p}_{0,3} = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0,3} a_{02} \frac{p_{0,3}}{\sum_{\ell=2}^{3} p_{0,\ell}} + a_{03} \frac{p_{0,3}}{\sum_{k=0}^{3} \pi_{k,k,3}} + b_{03} \frac{\pi_{0,0,3}}{\sum_{k=0}^{3} \sum_{\ell=2}^{3} \pi_{k,k,\ell}} \right\} \]
\[
\hat{p}_{0,3} = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{03} + a_{02} \frac{p_{0,3}}{3 \sum_{\ell=2} p_{0,\ell}} + b_{03} \frac{\pi_0 p_{0,3}}{\sum_{k=0} \pi_k P_{k,3}} + d_{02} \frac{\pi_0 p_{0,3}}{\sum_{k=0} \sum_{\ell=2} \pi_k P_{k,\ell}} \right\} \tag{4.28}
\]

Recall, as in the last chapter, \(\hat{\pi}_k\) refers to the iterative estimate for \(\pi_k\), given by (4.27).

Next, for \(k = 1\) we have the following (iterative) MLE's:

\[
\hat{p}_{1,0} = \frac{1}{(\hat{\pi}_1)(r_{**})} \left\{ x_{10} + a_{10} \frac{p_{1,0}}{\sum_{\ell=0} P_{1,\ell}} + b_{00} \frac{\pi_1 p_{1,0}}{\sum_{k=0} \pi_k P_{k,0}} + b_{10} \frac{\pi_1 p_{1,0}}{\sum_{k=1} \pi_k P_{k,0}} + \right\}
\]

\[
d_{00} \frac{\pi_1 p_{1,0}}{\sum_{k=0} \sum_{\ell=0} \pi_k P_{k,\ell}} + d_{10} \frac{\pi_1 p_{1,0}}{\sum_{k=1} \sum_{\ell=0} \pi_k P_{k,\ell}} \right\}
\]

\[
\hat{p}_{1,1} = \frac{1}{(\hat{\pi}_1)(r_{**})} \left\{ x_{11} + a_{10} \frac{p_{1,1}}{\sum_{\ell=0} P_{1,\ell}} + a_{11} \frac{p_{1,1}}{\sum_{\ell=1} P_{1,\ell}} + b_{01} \frac{\pi_1 p_{1,1}}{\sum_{k=0} \pi_k P_{k,0}} + b_{11} \frac{\pi_1 p_{1,1}}{\sum_{k=1} \pi_k P_{k,0}} + \right\}
\]

\[
d_{00} \frac{\pi_1 p_{1,1}}{\sum_{k=0} \sum_{\ell=0} \pi_k P_{k,\ell}} + d_{01} \frac{\pi_1 p_{1,1}}{\sum_{k=0} \sum_{\ell=1} \pi_k P_{k,\ell}} + d_{10} \frac{\pi_1 p_{1,1}}{\sum_{k=1} \sum_{\ell=0} \pi_k P_{k,\ell}} + d_{11} \frac{\pi_1 p_{1,1}}{\sum_{k=1} \sum_{\ell=1} \pi_k P_{k,\ell}} \right\}
\]

\[
\hat{p}_{1,2} = \frac{1}{(\hat{\pi}_1)(r_{**})} \left\{ x_{12} + a_{11} \frac{p_{1,2}}{\sum_{\ell=1} P_{1,\ell}} + a_{12} \frac{p_{1,2}}{\sum_{\ell=2} P_{1,\ell}} + b_{02} \frac{\pi_1 p_{1,2}}{\sum_{k=0} \pi_k P_{k,0}} + b_{12} \frac{\pi_1 p_{1,2}}{\sum_{k=1} \pi_k P_{k,0}} + \right\}
\]

\[
d_{01} \frac{\pi_1 p_{1,2}}{\sum_{k=0} \sum_{\ell=1} \pi_k P_{k,\ell}} + d_{02} \frac{\pi_1 p_{1,2}}{\sum_{k=0} \sum_{\ell=2} \pi_k P_{k,\ell}} + d_{11} \frac{\pi_1 p_{1,2}}{\sum_{k=1} \sum_{\ell=1} \pi_k P_{k,\ell}} + d_{12} \frac{\pi_1 p_{1,2}}{\sum_{k=1} \sum_{\ell=2} \pi_k P_{k,\ell}} \right\}
\]
\[ \hat{p}_{1,3} = \frac{1}{(\hat{\pi}_1)(r_{**})} \left\{ x_{13} + a_{12} \frac{p_{1,3}}{\sum_{\ell=0}^3 p_{1,\ell}} + b_{03} \frac{\pi_1 p_{1,3}}{\sum_{k=0}^3 \pi_k p_{k,0}} \right\} + \frac{d_{02} \frac{\pi_1 p_{1,3}}{\sum_{k=0}^3 \pi_k p_{k,2}} + d_{12} \frac{\pi_1 p_{1,3}}{\sum_{k=1}^3 \sum_{\ell=2}^3 \pi_k p_{k,\ell}}}{\sum_{k=0}^3 \sum_{\ell=2}^3 \pi_k p_{k,\ell}} \right\}. \]  

(4.29)

Our usual Bayes' Rule arguments apply here; for example, when \( k=1 \) we see that the estimates for \( p_{10}, \ldots, p_{13} \) split the counts of those nodes with indegree 1 at time \( t_1 \) (estimated by \( (\hat{\pi}_1)(r_{**}) \)) proportionate to those that moved to indegrees 0, 1, 2, or 3, respectively, at time \( t_2 \). Note some terms contain sums across rows (\( k \)), while others contain sums across columns (\( \ell \)), and some contain both. Again, this is because of the many possible ways missingness can occur under Model 5.

The initial probabilities for the Markov chain modelling the node nonresponse have the following MLE's:

\[ \hat{\zeta}_1 = \frac{(b_{**} + d_{**})}{r_{**}} \]  

(4.30)

\[ \hat{\zeta}_2 = \frac{(x_{**} + a_{**})}{r_{**}} \]

and the (conditional) transition probabilities have the following MLE's:

\[ \hat{\rho}_{11} = \frac{d_{**}}{(\hat{\zeta}_1)(r_{**})} \quad \hat{\rho}_{21} = \frac{a_{**}}{(\hat{\zeta}_2)(r_{**})} \]

(4.31)

\[ \hat{\rho}_{12} = \frac{b_{**}}{(\hat{\zeta}_1)(r_{**})} \quad \hat{\rho}_{22} = \frac{x_{**}}{(\hat{\zeta}_2)(r_{**})}. \]
These estimates are interpreted in the same manner as explained previously. Note that cells from each of the blocks appear in the corresponding estimate, as expected, given the form of the data (see Figure 10).

4.3.6 Model 6: Nonrandom Node Nonresponse at One or Both Time Periods
(Popularity Model)

This model is the most complex of all node nonresponse models presented; it represents the situation where node nonresponse is modelled for one or two time periods, and the nonresponse is assumed to be nonrandom, depending on the indegree of the node at the time of the nonresponse.

Again, because of the cumbersome form that the likelihood function and resulting estimates would take on if we presented the general form of this model, we will present it in a special case, where $M_1=1$ and $M_2=1$; $g = 4$ as usual. Data appear in the four-block combination X, A, B, D. There are again 49 cells with 1 constraint, as seen in Model 5. The number of parameters under Model 6 is 44, with 17 constraints, leaving 27 free parameters, and thus $(49 - 1) - (44 - 17) = 21$; again there are actually only four counts, so the degrees of freedom will be less than this in practice. We have the usual 20 parameters with 5 constraints from the data Markov chain, and the remaining 24 parameters come from 16 node nonresponse Markov chains, one for each combination of indegree states over the two time periods. These 24 nonresponse parameters are given below.
\[ \zeta_i(k, \ell) = \begin{cases} \zeta_i(k), & \text{for } i = 1, 2 \\ 0, & \text{else} \end{cases} \] (4.32)

\[ \rho_{ij}(k, \ell) = \begin{cases} \rho_{ij}(\ell), & \text{for } i = 1, 2 \text{ and } j = 1, 2 \\ 0, & \text{else} \end{cases} \]

Note that the initial node nonresponse probabilities depend on \( k \), the indegree of the node at time \( t_1 \), while the (conditional) transition probabilities depend on \( \ell \), the indegree of the node at time \( t_2 \). In each case, the node nonresponse probability depends on the indegree of the node (i.e. the popularity of the individual) at the time of the nonresponse. Since the nonresponse depends on unobservable quantities at each time period, it will be nonignorable. The likelihood function for Model 6 is the same as that for Model 5, except for this dependence of \( \zeta \) and \( \rho \) on \( k \) and \( \ell \), respectively; the usual constraints apply. The likelihood function under Model 6 is given below.

\[
\prod_{k=0}^{3} \prod_{\ell=0}^{3} \left[ \pi_k p_{k, \ell} \zeta_2(k) \rho_{22}(\ell) \right] x_{k \ell} \times \prod_{k=0}^{3} \prod_{\ell=0}^{2} \left[ \pi_k \sum_{v=0}^{1} (p_{k, \ell+v} \xi_2(k) \rho_{21}(\ell + v)) \right]^{-a_{k \ell}} \\
\times \prod_{k=0}^{2} \prod_{\ell=0}^{3} \left[ \sum_{u=0}^{1} (\pi_{k+u} p_{k+u, \ell} \xi_1(k + u) \rho_{12}(\ell)) \right]^{b_{k \ell}} \\
\times \prod_{k=0}^{2} \prod_{\ell=0}^{2} \left[ \sum_{u=0}^{1} \sum_{v=0}^{1} \pi_{k+u} p_{k+u, \ell+v} \zeta_1(k + u) \rho_{11}(\ell + v) \right]^{c_{k \ell}}.
\]

The (iterative) MLE's of the data Markov chain parameters will be the same as those found under Model 5, except for the dependence of \( \zeta \) on \( k \) and the dependence of \( \rho \) on \( \ell \). (Recall, since we have \( M_1 = 1 \), and \( M_2 = 1 \), there will be no indicator
functions in the estimates. We therefore will only present the MLE’s for the $\pi$ and $p$
parameters in the case where $k = 0$, without loss of generality.

\[
\hat{\pi}_0 = \frac{1}{r_{**}} \left\{ x_{0*} + a_{0*} + \sum_{\ell=0}^{3} b_{0,\ell} \frac{\pi_{0,\ell} \zeta_{S_1}(1)}{\sum_{k=0}^{1} \pi_{k,\ell} \zeta_{S_1}(k)} + \sum_{\ell=0}^{2} d_{0,\ell} \frac{\pi_{0,\ell} \zeta_{S_1}(1)}{\sum_{k=0}^{1} \pi_{k,\ell} \zeta_{S_1}(k)} \right\}
\]

(4.34)

\[
\hat{p}_{0,0} = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0,0} + a_{0,0} \frac{p_{0,0} \varphi_{21}(0)}{\sum_{\ell=0}^{1} p_{0,\ell} \varphi_{21}(\ell)} + b_{0,0} \frac{\pi_{0,0} \zeta_{S_1}(0) \rho_{12}(0)}{\sum_{k=0}^{1} \pi_{k,0} \zeta_{S_1}(k) \rho_{12}(0)} + \frac{\pi_{0,0} \zeta_{S_1}(0) \rho_{11}(0)}{\sum_{k=0}^{1} \sum_{\ell=0}^{1} \pi_{k,\ell} \zeta_{S_1}(k) \rho_{11}(\ell)} \right\}
\]

\[
\hat{p}_{0,1} = \frac{1}{(\hat{\pi}_0)(r_{**})} \left\{ x_{0,1} + a_{0,0} \frac{p_{0,1} \varphi_{21}(1)}{\sum_{\ell=0}^{1} p_{0,\ell} \varphi_{21}(\ell)} + a_{0,1} \frac{p_{0,1} \varphi_{21}(1)}{\sum_{\ell=1}^{1} \sum_{\ell=0}^{1} p_{0,\ell} \varphi_{21}(\ell)} + b_{0,1} \frac{\pi_{0,1} \zeta_{S_1}(0) \rho_{12}(1)}{\sum_{k=0}^{1} \pi_{k,1} \zeta_{S_1}(k) \rho_{12}(1)} + \frac{\pi_{0,1} \zeta_{S_1}(0) \rho_{11}(1)}{\sum_{k=0}^{1} \sum_{\ell=0}^{1} \pi_{k,\ell} \zeta_{S_1}(k) \rho_{11}(\ell)} \right\}
\]

\[
+ d_{0,0} \frac{\pi_{0,0} \zeta_{S_1}(0) \rho_{11}(1)}{\sum_{k=0}^{1} \sum_{\ell=0}^{1} \pi_{k,\ell} \zeta_{S_1}(k) \rho_{11}(\ell)} + d_{0,1} \frac{\pi_{0,1} \zeta_{S_1}(0) \rho_{11}(1)}{\sum_{k=0}^{1} \sum_{\ell=0}^{1} \pi_{k,\ell} \zeta_{S_1}(k) \rho_{11}(\ell)} \right\}
\]
\[
\hat{p}_{0,2} = \frac{1}{(\hat{\pi}_0)(\tau_{**})} \left\{ x_{02} + a_{01} \frac{p_{0,2}r_{21}(2)}{\sum_{\ell=1} p_{0,\ell}r_{21}(\ell)} + a_{02} \frac{p_{0,2}r_{21}(2)}{\sum_{\ell=2} p_{0,\ell}r_{21}(\ell)} + b_{02} \frac{\pi_0p_{0,2}\zeta_1(0)r_{12}(2)}{\sum_{k=0} \pi_kp_{k,2}\zeta_1(k)r_{12}(2)} \right. \\
+ d_{01} \frac{\pi_0p_{0,2}\zeta_1(0)r_{11}(2)}{\sum_{k=0} \sum_{\ell=1} \pi_kp_{k,\ell}\zeta_1(k)r_{11}(\ell)} + d_{02} \frac{\pi_0p_{0,2}\zeta_1(0)r_{11}(2)}{\sum_{k=0} \sum_{\ell=2} \pi_kp_{k,\ell}\zeta_1(k)r_{11}(\ell)} \right\}
\]

\[
\hat{p}_{0,3} = \frac{1}{(\hat{\pi}_0)(\tau_{**})} \left\{ x_{03} + a_{02} \frac{p_{0,3}r_{21}(3)}{\sum_{\ell=2} p_{0,\ell}r_{21}(\ell)} + b_{03} \frac{\pi_0p_{0,3}\zeta_1(0)r_{12}(3)}{\sum_{k=0} \pi_kp_{k,3}\zeta_1(k)r_{12}(3)} \right. \\
+ d_{02} \frac{\pi_0p_{0,3}\zeta_1(0)r_{11}(3)}{\sum_{k=0} \sum_{\ell=2} \pi_kp_{k,\ell}\zeta_1(k)r_{11}(\ell)} \right\}
\]

(4.35)

Note that because of the nonignorable nonresponse at one or both time periods, we have \( \zeta \) and \( \rho \) parameters appearing in the iterative estimates of the \( \pi \)'s and p's. Our interpretations follow those made previously, with proportions based on estimated probabilities of a node being in each particular state at time \( t_2 \), given (here with \( k = 0 \)) that the indegree was 0 at time \( t_2 \). Under Model 6, the nonresponse estimates will be different from those under Model 5 since they also use the information regarding the indegree of each node at times \( t_1 \) and \( t_2 \). We follow with (iterative) MLE's for selected parameters for the node nonresponse Markov chains under Model 6; in particular, those for which \( k = 1 \) and \( \ell = 1 \).

The initial probabilities for the node nonresponse Markov chain (iterative) MLE's are given by the following:
\[ \hat{\xi}_1(0) = \frac{1}{(\hat{\pi}_0(r_{**}))} \left\{ \sum_{\ell=0}^{3} b_{0\ell} \sum_{k=0}^{1} \pi_0 p_{0,\ell} \xi_1(1) \right\} + \sum_{\ell=0}^{2} d_{0\ell} \sum_{k=0}^{1} \pi_0 (p_{0,\ell} + p_{0,\ell+1}) \xi_1(1) \left\} \right. \\
\hat{\xi}_2(0) = \frac{x_{0**} + a_{0**}}{(\hat{\pi}_0(r_{**}))}. \] (4.36)

Note that in forming these estimates, the iterative estimate \( \hat{\pi}_0 \) is split into two parts above; one for nonresponding node probabilities, and the other for responding node probabilities.

The (iterative forms for the) MLE's for the transition probabilities (again, where \( \ell = 0 \)) are given by the following:

Let \( d = \left\{ x_{**} + \sum_{k=0}^{3} a_{k0} \frac{p_{k0} \rho_{21}(0)}{\sum_{v=0}^{1} p_{k,v} \rho_{21}(v)} \right\} + b_{**} + \left. \sum_{k=0}^{2} d_{k0} \frac{\sum_{u=0}^{1} \pi_{k+u} p_{k+u,0} \xi_1(k+u) \rho_{11}(0)}{\sum_{u=0}^{1} \sum_{v=0}^{1} \pi_{k+u} p_{k+u,v} \xi_1(k+u) \rho_{11}(v)} \right\}. \) (4.37)

Then we have the following (iterative) estimates.

\[ \hat{\rho}_{11}(0) = \left( \sum_{k=0}^{2} \frac{\sum_{u=0}^{1} \pi_{k+u} p_{k+u,0} \xi_1(k+u) \rho_{11}(0)}{\sum_{u=0}^{1} \sum_{v=0}^{1} \pi_{k+u} p_{k+u,v} \xi_1(k+u) \rho_{11}(v)} \right) / d \]
\[ \hat{\rho}_{12}(0) = \frac{b_{*0}}{d} \]

\[ \hat{\rho}_{21}(0) = \sum_{k=0}^{3} a_{k0} \frac{p_{k,0}\rho_{21}(0)}{\sum_{\nu=0}^{1} p_{k,\nu}\rho_{21}(\nu)} / d \]

\[ \hat{\rho}_{22}(0) = x_{*0}/d . \]  

Note that the estimates for \( \rho_{12} \) and \( \rho_{22} \) are based on observed counts only; this is because the node in either situation is a nonrespondent at time \( t_2 \) (recall \( M_1 = 1 \) and \( M_2 = 1 \) here), and it does not lose any information from the other nodes regarding its indegree. The nonresponse is based on an observable quantity (indegree at time \( t_2 \)) in this case, which makes it ignorable (for that node only). All nodes responding at time \( t_2 \), however, do lose information at time \( t_2 \) from the missing node, hence the nonresponse is based on an unobservable quantity, which means it is nonignorable. The estimates for \( \rho_{11} \) and \( \rho_{21} \), therefore, are based on estimated proportions, as seen previously. In situations where there is more than one missing node at either time period, all nodes lose some information at the time of the nonresponse. In general, the nonresponding nodes lose the least amount of information.

In Chapter 5, we will discuss application of these models to data sets.
4.4 Some Node Nonresponse Models under the Expansiveness Model for the Data Markov Chain

Under the expansiveness model for the data, the choice intensities are based on the expansiveness of the individual making the decision, reflected by the outdegree of the corresponding node at that time. Note that when an individual becomes a node nonrespondent in this case, the outdegree is completely unknown since all outgoing edges are missing due to the node nonresponse; this means there will be no partial information for an individual at the time when he/she is missing. Also, in contrast to node nonresponse under the popularity model where all other individuals are affected by a node nonrespondent, under the expansiveness model only the missing node is affected by nonresponse. These differences make this situation easier to handle in terms of the nonresponse models, yet they bring limitations to the model fitting, as we will see when discussing degrees of freedom, for example.

Under the expansiveness model, we summarize the observed data in the observed outdegree transition matrix. This matrix will be of size $g \times g$, and the possible outdegree for each node will range from 0 through $g-1$, with only one supplementary row and column where we summarize missing outdegrees at times $t_2$, and $t_1$, respectively. We present the observed outdegree transition matrix in Figure 11 below, and as before, we present all further models and discussions in the case where $g = 4$ for clarity, without loss of generality. (Note that in practice, one would probably be working with more than 4 individuals in any given situation.)
(Here $M$ indicates a missing node; hence a node with outdegree completely unknown.)

There are four different combinations in which the data will appear in this matrix, depending on the number of missing nodes and the time period(s) in which missingness occurs: data appear in block $X$ if there is no missing data at either time period, blocks $X$ and $A$ if missingness occurs time $t_2$ only, blocks $X$ and $B$ if missingness occurs at time $t_1$ only, and blocks $X$, $A$, $B$, and $D$ if missingness occurs at both time periods.

There are fewer states in the observed outdegree transition matrix under the expansiveness model than in the observed indegree transition matrix under the popularity model, since there is only one place where missing data can fall; parameter estimation is clearly different in two models. If we define the parameters for nonresponse the same as we have previously done with the same six models as presented previously, we have the following situations in the case where $g = 4$.

The random nonresponse models for one time period only have 3 degrees of freedom each, except in the case where there is no missing data; in this case the saturated model fits the data exactly. The nonrandom nonresponse models for one time period only have 0 degrees of freedom; hence we can fit the data exactly. (This

<table>
<thead>
<tr>
<th>Time $t_1$</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>$x_{00}$</td>
<td>$x_{01}$</td>
<td>$x_{02}$</td>
<td>$x_{03}$</td>
<td>$a_{00}$</td>
</tr>
<tr>
<td>1</td>
<td>$x_{10}$</td>
<td>$x_{11}$</td>
<td>$x_{12}$</td>
<td>$x_{13}$</td>
<td>$a_{10}$</td>
</tr>
<tr>
<td>2</td>
<td>$x_{20}$</td>
<td>$x_{21}$</td>
<td>$x_{22}$</td>
<td>$x_{23}$</td>
<td>$a_{20}$</td>
</tr>
<tr>
<td>3</td>
<td>$x_{30}$</td>
<td>$x_{31}$</td>
<td>$x_{32}$</td>
<td>$x_{33}$</td>
<td>$a_{30}$</td>
</tr>
<tr>
<td>M</td>
<td>$b_{00}$</td>
<td>$b_{01}$</td>
<td>$b_{02}$</td>
<td>$b_{03}$</td>
<td>$d_{00}$</td>
</tr>
</tbody>
</table>

Figure 11. The observed outdegree transition matrix for $g = 4$. 
is still useful for examination of the nonresponse probabilities, to see if we are picking up some nonrandomness in the data.) For nonresponse at both time periods, the random model fits exactly, and the nonrandom model cannot be fit. (This shows a disadvantage to using the expansiveness model.)

We present estimates for Models 1 and 3; results for other models follow in a similar manner, as we have seen in the preceding sections.

4.4.1 Model 1: Random Node Nonresponse at Time \( t_2 \) Only

(Expansiveness Model)

We present these results for the case when \( g = 4 \). The usual 20 parameters for the data Markov chain apply here as well, except that the states are the outdegree of each node, rather than the indegree. (The possible number of states is still \( g \).) Under Model 1, the four node nonresponse parameters are the following (for any node in cell \((k, \ell)\) of the outdegree transition matrix at Stage 1, \(k, \ell = 0, \ldots, 3\)):

\[
\zeta_i(k, \ell) = \begin{cases} 
1, & \text{for } i = 1 \\
0, & \text{for } i = 2
\end{cases}
\]

\[
\rho_{ij}(k, \ell) = \begin{cases} 
\rho_{ij}, & \text{for } i = 1, \text{ and } j = 1,2 \\
0, & \text{otherwise}
\end{cases} \quad (4.39)
\]

Data under Model 1 will appear in block \( X \) (in the case where there is no missing data), or blocks \( X \) and \( A \). (Note that responding nodes appear in block \( X \), while nonrespondents appear in block \( A \); this is in contrast to the node nonresponse models in the previous sections.) The likelihood function is proportional to the following:
\[
\prod_{k=0}^{3} \prod_{\ell=0}^{3} \left( \pi_{k}\pi_{k,\ell}\bar{p}_{11} \right)^{x_{k\ell}} \times \prod_{k=0}^{3} \left( \sum_{\ell=0}^{3} \pi_{k}\pi_{k,\ell}\bar{p}_{12} \right)^{a_{k\ell}}.
\]

(4.40)

The following forms for MLE's for the data Markov chain are obtained:

\[
\hat{\pi}_{k} = \frac{x_{k*} + a_{k0}}{r_{k*}} = \frac{r_{k*}}{r_{**}}. 
\]

(4.41)

\[
\hat{p}_{k\ell} = \frac{1}{r_{k*}} \left( \frac{x_{k\ell} + a_{k0}}{\sum_{\ell=0}^{3} p_{k\ell}} \right) = \frac{x_{k\ell} + a_{k0}p_{k\ell}}{r_{k*}}. 
\]

(4.42)

The following MLE's for the node nonresponse Markov chain are obtained:

\[
\hat{\rho}_{11} = \frac{x_{**}}{r_{**}} 
\]

\[
\hat{\rho}_{12} = \frac{a_{0*}}{r_{**}}. 
\]

(4.43)

The usual interpretations apply to these parameter estimates.

4.4.2 Model 3: Nonrandom Node Nonresponse at Time \(t_2\) Only

(Expansiveness Model)

Under Model 3, the node nonresponse at time \(t_2\) is assumed to be nonrandom, depending on the expansiveness of the person at time \(t_2\), the state of the outdegree at
the time of the nonresponse (analogous to previous versions of Model 4 in Chapters 3 and 4). Under Model 3, the sixteen node nonresponse parameters are the following (for any node in cell \((k, \ell)\) of the outdegree transition matrix at Stage 1):

\[
\zeta_i(k, \ell) = \begin{cases} 
1, & \text{for } i = 1 \\
0, & \text{for } i = 2 
\end{cases}
\]

\[
\rho_{ij}(k, \ell) = \begin{cases} 
\rho_{ij}(\ell), & \text{for } i = 1, \text{ and } j = 1, 2 \\
0, & \text{otherwise} 
\end{cases} 
\quad (4.44)
\]

Data under Model 3 will appear in block \(X\), or \(X\) and \(A\), as for Model 2. The likelihood function is proportional to the following:

\[
\prod_{k=0}^{3} \prod_{\ell=0}^{3} \left( \pi_k p_{k, \ell} \rho_{11}(\ell) \right)^{x_{k\ell}} \times \prod_{k=0}^{3} \left( \sum_{\ell=0}^{3} \pi_k p_{k, \ell} \rho_{12}(\ell) \right)^{a_{k\ell}}.
\]

\[
(4.45)
\]

The following forms for MLE's for the data Markov chain are obtained:

\[
\hat{\pi}_k = \frac{x_{k\bullet} + a_{k0}}{r_{\bullet\bullet}} = \frac{r_{k\bullet}}{r_{\bullet\bullet}} 
\]

\[
(4.46)
\]

\[
\hat{p}_{k\ell} = \frac{1}{r_{k\bullet}} \left( \frac{x_{k\ell} + a_{k0}}{\frac{3}{3} \sum_{\ell=0}^{3} p_{k\ell} \rho_{12}(\ell)} \right).
\]

\[
(4.47)
\]

The following forms for MLE's for the node nonresponse Markov chain are obtained:
\[ \hat{\rho}_{11}(\ell) = \frac{x_{*\ell}}{d} \]

\[ \hat{\rho}_{12}(\ell) = \sum_{k=0}^{3} a_{k0} \frac{\pi_k p_{k\ell} p_{12}(\ell)}{\sum_{\ell=0}^{3} \pi_k p_{k\ell} p_{12}(\ell)} \]

where \( d = x_{*\ell} + \sum_{k=0}^{3} a_{k0} \frac{\pi_k p_{k\ell} p_{12}(\ell)}{3 \sum_{\ell=0}^{3} \pi_k p_{k\ell} p_{12}(\ell)} \). \hspace{3cm} (4.48)

The usual interpretations apply to these parameter estimates; note that since the nonresponse is at time \( t_2 \), and depends on the state of the outdegree at that time, the summations move down columns of the observed outdegree transition matrix.
CHAPTER V
ILLUSTRATION

5.1 Nonresponse Data in the Social Network Setting

In this chapter, we apply selected nonresponse models to social network data containing missing values which have been generated according to different nonresponse rates. Since it is typical practice for researchers in the area of social networks to discard information regarding individuals who respond partially or do not respond (see Taba (1955) and Sampson (1968), for example) we were unable to find actual data that had been collected and reported in the form needed to apply these models. It is important that in the future, nonresponse be taken into account when collecting data. This can be accomplished in a number of ways.

Most importantly, the way in which the members of the social network are asked to respond to the questions may be modified. For example, in Sampson's monastery data (1968), it is noted that certain members of the order voluntarily left at the eleventh month, because of conflicts within the group caused by expulsions of certain other members. The social networks are shown for the time period before those members left, but for the following time period, the available data shows only the relationships between the remaining group members. It would have been interesting to see what the remaining members had felt about those who left during that time period, and use that information as a data set with four node nonrespondents. It is noted that one of the members of the
group left at a different time period because of extreme loneliness and the isolation that he felt. Perhaps this would have been a good data set to apply the nonrandom node nonresponse models from Chapter 4, if the information were available. Sampson noted that in addition to conducting surveys to collect information about the relationships between the monks, he also conducted interviews with the individuals; when survey information was not provided by an individual regarding relationships with the others, the information collected during the interviews was used as a best guess as to the actual status of the relationship at that time. This could lead to data inaccuracies due to interviewer bias.

Sampson's monastery data is an example of a fixed-choice survey where, for example, each monk was asked to list the three monks he liked best, instead of being asked to provide a list of all individuals he considered highly. This imposes constraints on the responses of the individuals, as noted by Holland (1977), and hinders our understanding of the social structure of the group. Holland proposes that surveys be conducted using a roster format, where each individual is given a list of all group members, and is asked to make a choice regarding each individual relationship at that time. The proposed choices for answers were dislike and like, and the individual is instructed to not circle either if he/she did not feel strongly. We propose that another possible response be added to this list, so that we have the following possible responses: like, dislike, and no opinion. If all three are left blank, we would count it as nonresponse. This would again separate the absence of a relationship from the possible presence of a relationship, but where the individual has chosen not to inform us of it.

As another example, in Taba's (1955) classroom data, each student is asked to list the classmates he/she would most like to sit next to. Again, it may be more informational to give the student a list of all individuals in the group, and give possible responses to the question, such as: want to sit next to, do not wish to sit next to, no
opinion, do not wish to respond. This would differentiate the absence of the relationship from actual nonresponse. In this data set, any individuals who responded only partially were dropped from the data, and the results were not reported, although in the description of the experiment, and the analysis, it is noted that certain individuals were very isolated from the group, etc., which may have warranted their nonresponse. Had there been nonresponse on record, we could have applied our models to it.

The models proposed in this dissertation will be useful in situations where nonresponse has been documented and is clearly separated from the absence of relationships. In the future, as techniques for data collection improve, and researchers are made aware of the significance of nonresponse, these models will be more applicable.

5.2 Classroom Data with Generated Link Nonresponse

To provide an illustration of the nonresponse models presented in this dissertation, we start with the data of Taba (1955) in which each student in a classroom of 25 eighth graders was asked to list the members he/she would like to sit by. The data was collected at different time points during the academic year. Wasserman (1980) concludes that the relationships recorded in this data set are fairly stable, and fits the reciprocity model to it. We take one of the social networks with data collected in September and January as the two time points, and generate link nonresponse according to different nonresponse probabilities. We then apply two models from Chapter 3, and compare our results. It is important to note that work involving model fitting is still in the exploratory phase. The amount of nonresponse occurring in the social network setting has not been estimated, and until data is collected from which we can get some idea of nonresponse rates, the nonresponse probabilities we use must be speculative.
The dyad transition matrix for the data collected on this classroom of 25 students over the two time periods (September and January) is given in the following:

<table>
<thead>
<tr>
<th></th>
<th>January</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0,0)</td>
<td>(1,0)</td>
<td>(0,1)</td>
<td>(1,1)</td>
</tr>
<tr>
<td>September</td>
<td>217</td>
<td>9</td>
<td>13</td>
<td>1</td>
</tr>
<tr>
<td>(0,0)</td>
<td>9</td>
<td>9</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(1,0)</td>
<td>12</td>
<td>1</td>
<td>8</td>
<td>3</td>
</tr>
<tr>
<td>(0,1)</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>7</td>
</tr>
</tbody>
</table>

Figure 12. Dyad transition matrix for classroom data at Stage 1.

(Note that with 25 individuals, there are 300 total dyads in the social network.)

We move to Stage 2 of the two-stage process by generating missing links according to two different sets of nonresponse probabilities, one for each of link nonresponse Models 1 and 3. The iterative (maximum likelihood) estimates were computed using Fortran programs, where iterations were performed until the difference between consecutive estimates was less than 0.0005. Estimates converged quickly, for a range of starting values. For Models 1 and 3, we present the nonresponse probabilities used to generate the link nonresponse, and the resulting dyad transition matrix observed at the end of Stage 2.

Recall, for Model 1, we assume the nonresponse occurs randomly at time \( t_2 \). The nonresponse probabilities chosen for link nonresponse generation are the following:

\[
\rho_1 = 0.9025, \quad \rho_2 = 0.0475, \quad \rho_3 = 0.0475, \quad \rho_4 = 0.0025. \quad (5.1)
\]

The nonresponse probabilities used to generate the link nonresponse under Model 1 assume independence of the individual in the dyad, and 0.05 as the probability of an individual becoming a link nonrespondent at time \( t_2 \). Using a random number
generator for 300 uniform (0,1) random variables, we simulated that 29 dyads lost
partial information, and one lost all information at Stage 2. The resulting observed dyad
transition matrix is given in Figure 13.

<table>
<thead>
<tr>
<th></th>
<th>(0,0)</th>
<th>(1,0)</th>
<th>(0,1)</th>
<th>(1,1)</th>
<th>(M,0)</th>
<th>(M,1)</th>
<th>(0,M)</th>
<th>(1,M)</th>
<th>(M,M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>September</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0,0)</td>
<td>192</td>
<td>9</td>
<td>13</td>
<td>1</td>
<td>14</td>
<td>0</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(1,0)</td>
<td>9</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(0,1)</td>
<td>11</td>
<td>1</td>
<td>7</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(1,1)</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 13. Generated dyad transition matrix for classroom data under Model 1.

Upon fitting Model 1 to this data, we obtained the following estimates for the
parameters in the data and link nonresponse Markov chains, as shown in Figure 14.

<table>
<thead>
<tr>
<th>( \hat{P}_{11} )</th>
<th>( \hat{P}_{12} )</th>
<th>( \hat{P}_{13} )</th>
<th>( \hat{P}_{14} )</th>
<th>( \hat{P}_i )</th>
<th>( \hat{f}_i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>.8988</td>
<td>.0485</td>
<td>.0485</td>
<td>.0042</td>
<td>.9000</td>
<td>.8000</td>
</tr>
<tr>
<td>.4674</td>
<td>.3619</td>
<td>.0455</td>
<td>.1253</td>
<td>.0483</td>
<td>.0750</td>
</tr>
<tr>
<td>.4674</td>
<td>.0455</td>
<td>.3619</td>
<td>.1253</td>
<td>.0483</td>
<td>.0750</td>
</tr>
<tr>
<td>.2000</td>
<td>.1404</td>
<td>.1404</td>
<td>.5191</td>
<td>.0033</td>
<td>.0500</td>
</tr>
</tbody>
</table>

Figure 14. Parameter estimates for classroom data under Model 1.

The expected cell counts under Model 1 for this generated data are given by
Figure 15. Note the estimates follow the constraints for Model 1 regarding the
relabelling symmetry within members of the dyad. The estimates for the nonresponse
parameters are close to those chosen for nonresponse generation. The \( \chi^2 \) and \( G^2 \) values
under this model are 21.6191 and 19.5801, respectively. We compare these values to a Chi-Square distribution with 9 degrees of freedom. These values are admittedly high, suggesting a poor fit of the Model. However, over half of the contribution to these values is made by one cell: the dyad in row 2 column 9 of the observed dyad transition matrix. This count corresponds to the rare event that a dyad in state (1,0) at $t_1$ loses both links at time $t_2$. Only one dyad of the 300 in the entire network was generated with both links missing, and according to the nonresponse probabilities, we would expect that dyad to be in row 1, where 80% of the dyads appear, yet the dyad appears in a cell where the expected cell count is low (given that there are very few values in that row at time $t_1$). This indicates a strong sensitivity of these models to situations involving extremely small expected counts in which dyads do appear at the end of Stage 2. It should be noted that if the rare event hadn’t occurred, the $X^2$ and $G^2$ values would be close to 9, indicating a satisfactory fit of Model 1.

<table>
<thead>
<tr>
<th></th>
<th>January</th>
<th></th>
<th>September</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0,0)</td>
<td>(0,1)</td>
<td>(1,0)</td>
</tr>
<tr>
<td>(0,0)</td>
<td>194</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>(1,0)</td>
<td>9</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>(0,1)</td>
<td>9</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>(1,1)</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Figure 15. Expected cell counts for classroom data under Model 1 (including link nonresponse).

Under Model 3, the link nonresponse at time $t_2$ is assumed be nonrandom, depending on the state of the dyad at time $t_1$. We generated link nonresponse under Model 3 with the following chosen nonresponse probabilities:
\[
\begin{align*}
\rho_{11}(1) &= .81 & \rho_{12}(1) &= .09 & \rho_{13}(1) &= .09 & \rho_{14}(1) &= .01 \\
\rho_{11}(2) &= .72 & \rho_{12}(2) &= .18 & \rho_{13}(2) &= .08 & \rho_{14}(2) &= .02 \\
\rho_{11}(3) &= .72 & \rho_{12}(3) &= .08 & \rho_{13}(3) &= .18 & \rho_{14}(3) &= .02 \\
\rho_{11}(4) &= .9025 & \rho_{12}(4) &= .0475 & \rho_{13}(4) &= .0475 & \rho_{14}(4) &= .0025
\end{align*}
\]

(5.2)

We assumed here that an individual in an asymmetric dyad (at time \(t_1\)) has a higher chance of being a link nonrespondent (at time \(t_2\)) than an individual in a mutual or null state. Also, within the asymmetric dyad, the individual whose friendship is not being reciprocated is more likely to be the link nonrespondent. The generated dyad transition matrix for the classroom data under Model 3 is given in Figure 16.

\[
\begin{array}{|c|cccc|cccc|}
\hline
& (0,0) & (1,0) & (0,1) & (1,1) & (M,0) & (M,1) & (0,M) & (1,M) & (M,M) \\
(0,0) & 182 & 6 & 9 & 1 & 13 & 3 & 23 & 2 & 1 \\
(1,0) & 7 & 8 & 0 & 2 & 2 & 0 & 1 & 0 & 1 \\
(0,1) & 9 & 0 & 5 & 3 & 1 & 1 & 4 & 1 & 0 \\
(1,1) & 3 & 3 & 2 & 7 & 0 & 0 & 0 & 0 & 0 \\
\hline
\end{array}
\]

Figure 16. Generated dyad transition matrix for classroom data under Model 3.

Upon fitting Model 3 to this data, we obtained the following estimates for the parameters in the data and link nonresponse Markov chains, as shown in Figure 17.

The fit of this and the remaining models to this data have similar results, showing sensitivity to rare events. However, the probabilities seem reasonable; and the nonresponse probability estimates could be used in their own right to examine the occurrence of nonresponse in the social network. Note that this data set contains relatively few counts in rows 2-4 of the observed dyad transition matrix; 72% of the
data, in fact, is in cell (1,1). Dyads with missing information at time $t_2$ that were in an asymmetric or mutual state at time $t_1$ have potentially greater influence over those in the null state at time $t_1$, since expected cell counts will be low in those rows.

<table>
<thead>
<tr>
<th>$\hat{p}_{i1}$</th>
<th>$\hat{p}_{i2}$</th>
<th>$\hat{p}_{i3}$</th>
<th>$\hat{p}_{i4}$</th>
<th>$\hat{x}_i$</th>
<th>$\hat{p}_{1j}(1)$</th>
<th>$\hat{p}_{2j}(2)$</th>
<th>$\hat{p}_{3j}(3)$</th>
<th>$\hat{p}_{4j}(4)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>.9052</td>
<td>.0439</td>
<td>.0439</td>
<td>.0071</td>
<td>.8000</td>
<td>.8112</td>
<td>.6032</td>
<td>.6032</td>
<td>.7917</td>
</tr>
<tr>
<td>.4864</td>
<td>.3709</td>
<td>.0001</td>
<td>.1426</td>
<td>.0750</td>
<td>.0904</td>
<td>.1746</td>
<td>.1111</td>
<td>.0833</td>
</tr>
<tr>
<td>.4864</td>
<td>.0001</td>
<td>.3709</td>
<td>.1426</td>
<td>.0750</td>
<td>.0904</td>
<td>.1111</td>
<td>.1746</td>
<td>.0833</td>
</tr>
<tr>
<td>.2000</td>
<td>.1667</td>
<td>.1667</td>
<td>.4667</td>
<td>.0500</td>
<td>.0042</td>
<td>.0476</td>
<td>.0476</td>
<td>.0000</td>
</tr>
</tbody>
</table>

Figure 17. Parameter estimates for classroom data under Model 3.

Note in general that as $g$ increases, the number of dyads increases quickly; a group of 35 individuals will have 595 dyads, and presumably more nonresponse. This will mean higher expected cell counts, and rare events will have less of an effect on $X^2$ or $G^2$. More work needs to be done to assess the performance of these models under different situations, with different types of data sets and different nonresponse probabilities. These results are for illustrative purposes only, and reflect the fact that this is the first attempt to study nonresponse in the social network setting.

We also note that pooling data across more than two time periods may alleviate some of the sensitivity problem we are currently experiencing with these models, since this would also give increased cell counts, and provide a more stable setting in which to handle the nonresponse data.
5.3 Preliminary Findings on Node Nonresponse Models

Since the node nonresponse models for popularity and expansiveness present a complex situation involving degrees of freedom, we see that more work needs to be done before these models can be applied to data. The sparseness of cell counts makes this a very delicate situation; recall, there will be only $g$ counts in the observed indegree transition matrix. Wasserman (1990) works with a subset of the observed indegree transition matrix where most of the non-zero counts occur, disregarding the rest of the matrix. For example, if there are 25 individuals in the group, there is a small occurrence of an individual having more than say, 15 friendships. This form of collapsing cells is an approach that could be explored in our setting as well.

One way to approach this problem is to pool data across many different time periods, as described by Wasserman (1977). This adds information, as long as the time-homogeniety assumptions hold. For example, with Newcomb's fraternity data (1961), twelve social networks are pooled together to obtain the estimated probability transition matrix, and the estimated intensity matrix for this data. We may consider doing the same thing here with the node nonresponse model. This does not present a problem if the number of missing dyads is the same for each time period, since all data would appear in the same one, two, or four-block submatrix of the observed indegree transition matrix at the end of Stage 2. If, however, there are different numbers of missing individuals at different time periods, the data will appear in different submatrices of the observed indegree transition matrix, and the number of parameters to be estimated would increase greatly, possibly nullifying the increased number of non-zero cells. More work needs to be done in this area to determine when pooling would be reasonable and beneficial, in terms of degrees of freedom and estimability of parameters.
CHAPTER VI
CONCLUSIONS AND FUTURE RESEARCH

In this dissertation, we investigated nonresponse in social network data. In Chapter 2, we saw the many directions the area of social network research has taken, increasing in sophistication over the years, yet still exploring many of the same questions that social network researchers asked as early as the 1920's. We saw a general modelling framework, proposed by Holland and Leinhardt (1977a), which models the evolution of a social network over time. Our contribution to this area lies in handling nonresponse data in the social network setting; in particular, we propose models for nonresponse where data is modelled according to one of three particular models in the Holland-Leinhardt framework: 1) the reciprocity model, 2) the popularity model, and 3) the expansiveness model.

We have identified two types of nonresponse in social network data: link nonresponse and node nonresponse, and situations where each may occur. We proposed six models for each type of nonresponse: for link nonresponse under the reciprocity model for the data and node nonresponse for data under both the popularity model and expansiveness model. We modelled for both random and nonrandom nonresponse, occurring at one time period only, and both time periods. We presented models for both ignorable and nonignorable nonresponse. Using the ideas presented in Chapters 3 and 4 of this dissertation, an investigator has a means to handle nonresponse under a host of different situations.
We know upon investigation of the area that nonresponse data is virtually not available at this time, since it is typical practice to discard any data containing missing information; however, we believe that nonresponse is important in the social network setting, and that nonresponse data can help us learn more about the structure of the social network, and the dynamics which cause it to change over time. We saw instances where the nonresponse data, had it been available, could have provided more information, and we presented suggestions for improvement of data collection techniques in the social network setting. We are confident that upon the availability of such data, our proposed models will prove to be quite useful; for estimation of the actual (unobserved) data transition matrix under one of the Holland-Leinhardt models, or for investigation of the nonresponse parameters in their own right.

We identified some of the problems that occur at this yet preliminary stage of model fitting of data with generated random and nonrandom link nonresponse; namely, the sensitivity of the models when some expected cell counts are very small. In terms of the link nonresponse models, the next step would be to try pooling data across different time points, to add information and perhaps decrease the sensitivity of the model by increasing the expected cell counts (recall the link nonresponse models proposed have four states to the data Markov chain only, and there is no dependence on g, as there is with the node nonresponse models). More simulations are needed to identify the situations under which the models may be used with confidence.

With the node nonresponse models, the next step would be to examine the effect of pooling data, or collapsing cells with little or no data, and the consequences in terms of degrees of freedom, since in this situation the degrees of freedom is related to the placement of the counts in the block structure described in Chapter 4. Again, more simulations need to be done to examine under which conditions the node nonresponse models may be used. It is known that if different numbers of missing nodes are present
at different time periods, data will appear in more than one of the one, two, or four-block combinations, and this leads to an increase in the number of parameters which may offset what the increased cell counts could accomplish. However, if it is reasonable to expect consistently low numbers of nonrespondents at each time period (say 0 or 1), then pooling may be beneficial.

We must also try different rates of nonresponse with both the link and node nonresponse models; recall, since there is a shortage of nonresponse data at this time, the probabilities chosen to generate the link nonresponse in Chapter 5 were only speculative.

Other important elements of future research in this area include variance estimation using the observed information matrix, and exploration of other iterative procedures, such as the EM algorithm, for estimation of the parameters for which there is no closed form (Dempster, Laird, and Rubin (1977)). It would also be very interesting to expand these models so they include terms that would measure the importance of nonresponse, as the Holland-Leinhardt model parameters measure the importance of reciprocity and expansiveness, for example, and to also work on nonresponse models for the log-linear framework introduced by Holland and Leinhardt (1980). We can also investigate the estimation of the Q matrix in the Holland-Leinhardt (1977a) framework using the estimated $\hat{P}$ matrix found via our models.

We can also examine models for link nonresponse under the popularity model for the data, and models for node nonresponse under the reciprocity model for the data; the latter case may alleviate some of the problems encountered in Chapter 5 regarding dependence on $g$ (in terms of the number of parameters and sparseness of cell counts). One could also investigate models which allow for both link and node nonresponse simultaneously, since it is reasonable to assume this could occur in a social network.
Nonresponse rates in the social network setting have yet to be examined; indeed, nonresponse data in this area at this time is very hard to find, because the techniques which have been used up to this point have ignored the possible importance and effect of nonresponse data. It is important that data collection methods be modified to include the possibility of nonresponse. As data becomes available, we can assess the occurrence of nonresponse in social networks, and move forward in the modelling area. It is our hope that the ideas presented in this dissertation motivate that progress.


