

The Urban Communes Data Set: a Gold Mine for Secondary Analysis

John Levi Martin, King-To Yeung, and Benjamin D. Zablocki¹
Rutgers University

INTRODUCTION

It seems a commonplace among social network researchers to assert that there are too many methods chasing too few data; perhaps accordingly, there has been a tradition of the systematic re-analysis of existing data. For one, classic data sets (the Sampson monastery data, the bank wiring room data, the Southern women) have been used to adjudicate between methods for the induction of structure (given other data such as subsequent leaving of the group in the Sampson case). But re-analysis has been an important theoretical, as well as methodological, tool: the Davis-Leinhardt (Davis 1979) collection of data sets was used to determine the tendencies of choice processes in informal social groups using the Holland-Leinhardt U|MAN triadic analysis (Holland and Leinhardt 1970; Holland and Leinhardt 1971; Holland and Leinhardt 1975), leading to sets of model with increasingly complex but interpretable principles of internal organization (e.g. Johnsen 1985). This pooled data set led to Davis and Leinhardt's (1972: 222f) rejection of Homans's (1951: 182ff) argument that influence would be intransitive when possible.

Of course, the theoretical conclusions drawn will depend upon the type of relation studied (perhaps Davis and Leinhardt would have indeed found intransitive influence if the relations they studied were relations of influence, as opposed to "liking") and the type of group. Children's play groups seem to have different organizational principles from adult task-oriented groups, and there are almost certainly shades of difference in between. For these reasons, there is a great need for further banks of data sets that can be used to test basic hypotheses about social networks, data sets which include multiple items and multiple groups. Thanks to the support of the National Science Foundation, we are able to make public what is probably the largest set of comparable network data on a wide variety of naturally occurring and non-task oriented adult groups in the world, the Urban Communes Data Set (UCDS) from 1974. Derived from a stratified national sampling frame, with research conducted in six cities, 60 different intentional communities were studied (and, as discussed below, are currently being studied), producing a dramatic wealth of network data. (Five groups declined to permit network data to be gathered during the first wave, although they relented in later waves.)

In addition to being naturally occurring *gemeinschaftliche* groups, these were relatively closed circles: the great majority of significant nonkinship relationships of our respondents were with fellow commune members. Consequently, these data are likely to be quite useful in examining social-psychological

¹This work was made possible by a grant by the National Science Foundation SES-99-06452; authors' names are listed alphabetically.

processes that may compare self-perception, esteem, and attitudes to group position, as well as other structural aspects of social networks. We go on briefly to describe the data set, the network data, and then how it may be obtained. Further information is available from our website, discussed below, and in Zablocki (1980).

THE DATA SET

Sampling

A commune was defined as a household in which five or more largely unrelated adults (either with children or of both sexes) voluntarily decided to live together with a collective identity so as to reach some ideological goal having to do with the achievement of community. In cases in which data collection began on a group that was later disqualified from analysis, we have retained that data for secondary analysts. Thus we have extensive network data on one group later determined to be basically involuntary (a rehabilitational commune in which entrance was a court-prescribed alternative to jail time).

To maximize geographical diversity, six large Standard Metropolitan Sampling Areas from different regions across the United States were chosen for analysis. Fieldworkers in each city first compiled a comprehensive census of communes within the SMSA. Communes were then selected on the basis of certain key variables such as ideological type, population size, number of children, type of neighborhood, and year founded. In the rare instances where access was flatly denied by a selected commune, the next highest group on the priority list was chosen for study.

There was, however, one significant deviation from this quota-based selection procedure. Early in the fieldwork, researchers located a number of nationwide "new religious movements", organized in the form of federations of communal households (whose members often moved among households in different states). Because of the obvious sociological and historical importance of these new religious federations, one NRM was selected for study. Representative households belonging to this NRM were included in the sample from each of the six cities. A group-level variable ("GURULAND") indicates whether a commune was part of this federation or not.

Waves

The data collection occurred in three large periods, separated by intervals of around 12 years. Within the first of these periods, the first wave of data collection, which we are currently making public, was conducted in 1974. A second wave took place in 1975 of groups that were still in existence. A smaller wave took place in 1976, only collecting certain data among a minority of the remaining groups. At these times, new members were incorporated into the sampling frame. In the second period, 1984-1987, two waves of data (the first a limited pilot study) were collected on these same respondents, although the vast majority were no longer in communal groups. These focused on the consequences of unconventional behaviors and the possible changes in value orientations as persons left the "sacred canopy" of the group. In the most recent period, a pilot study has been completed and an additional wave of data from former members is currently being collected; this focuses on the relationship between the retention of network ties and the continued commitment to the beliefs, values, and behaviors associated with group life. This most recent wave of data will be made public as soon as it is all collected and cleaned. We hope that there will be sufficient interest in the longitudinal aspects of the data set to warrant cleaning and making public the intermediate waves, though they lack the closed network aspect of the earliest waves.

Data

The UCDS itself is a multiwave, multimethod study. While questionnaire techniques predominated, close to 1,500 hours of person-to-person, open-ended interviews were recorded on tape and on interview schedules, and these formed an indispensable part of the data library, although these are not being made public because of the difficulty of preserving anonymity of the respondents.

At the group level, fieldworkers in cooperation with members of the communes themselves were asked to complete fifty-four closed-ended items (such as dates of founding, type of authority, principles of decision making, economic organization, method of child rearing, type and intensity of ideology, and marital and sexual arrangements).

At the individual level, four or five members of each commune participated in a long interview (from two to four hours) covering each's early life, current life, present goals, and participation in the ongoing activities of the commune. This schedule is a combination of open- and closed-ended items. Other members were given a shorter version containing twenty of the most important closed-ended questions from the longer version. The short form comprises the minimal data set desired for each commune member. All members were also given an attitude questionnaire consisting of ninety-nine Likert-type items drawn, for the most part, from standard scales used on national surveys for the assessment of alienation, preference, self-esteem and self-perception. The most important aspects of the individual level data are made public in a single data file. A special survey of 32% of members who left in the first year of data collection was also conducted. Finally, and most importantly for our purposes, is the network data, which we go on to describe in greater detail, after some remarks about the probable degree of validity of these data.

Validity

The face validity of the data collected is high primarily because of the positive dispositions of almost all of the respondents toward the research project and because the first two waves were collected not by questionnaire but through face-to-face interviews. This validity has been demonstrated in two ways: (1) high consistency in reporting factual background information at widely separated time periods; and (2) high consistency, in the first wave, between interview data and data gathered through participant-observation.

The extent of cooperation and the apparent quality of the data collected were, on the whole, quite high. All the communes participated in the gathering of the systematic commune-level information. At the individual level there was a more selective but still quite high rate of response. For the first wave, we received personal background data from 81 percent of the possible urban respondents ($N = 667$), more than half of whom participated in autobiographical interviews. Attitude data were collected from 60 percent of the sample, as these questions were less interesting to many of the participants than questions focusing on their own lives, the group, or their relations with other members.

Relationship data were obtained from 70 percent (80% if we exclude the groups which refused on principle to allow for the distribution of these questionnaires). Ethnographic and interview cross-checks revealed a uniformly high quality to the responses, as the questions were generally judged to be relevant by the commune members and hence answered with care. Finally, the network data has been coded three times by independent coders to eliminate errors. As a consequence, we believe the overall quality of the resulting data to be very high.

THE NETWORK DATA

Network data were collected by a special instrument that had two parts. In the first, respondents were asked to volunteer the names of persons that they thought had particular characteristics. Characteristics asked include "supportive," "decisive," "influential," "loving," "interested in fixing up the house," "interested in kids," "strong," "sexy," "dominant," "charismatic," "intuitive," "holy," "passive," "dependent," and "narcissistic." Cases in which the respondent indicated that "everyone" possessed some characteristic were coded differently from cases in which the respondent chose specific alters. Respondents were also asked to report who cleaned the bathroom, cleaned the kitchen, did the laundry, went shopping, put a child to bed, called a meeting together, confronted someone at a meeting, made a statement of policy to the group, dealt with outside authorities, or mentioned that the house was messy.

Second, the respondent was asked to fill out a page-long relationship census for each other person in the group. First the respondent was asked to describe her relationship with the person in question in her own words; then a number of closed choice questions followed. The respondent (ego) was asked to judge how much time the two spent together ("just by yourselves"), to report whether or not she knew alter's father's job (informal discussions with a number of social psychologists indicated that this was a good way of identifying people who had spent some time talking with each other about themselves), which of the two had more power in the relationship, and whether alter ever acted like a mother, father, sister, brother, son or daughter in their relationship (multiple choices were allowed and coded).

Then, each respondent was asked to describe the tenor of her relationship with alter by answering whether each of a set of descriptions of the relationship was true, false, or sometimes true: these were "work together," "mind children together," "sleep together," "loving," "hateful," "parental," "exciting," "awkward," "tense," "jealous," "unreciprocated in some way," "exploitative," "improving," "sexual," to indicate whether the relation was more important to the respondent than to alter, or whether (a separate item) the relation was more important to alter than to the respondent. Finally, the respondent was asked to indicate five significant persons in her life, who could be outside the commune. If any commune member was listed, this was noted in the data through an indicator variable.

The possibility for detailed analysis of multiplex relations is obvious given the many relational items asked; even more exciting is the possibility of embedding such analysis in contextual comparisons given the group-level data, or tying network position or effects to individual level variables.

GETTING THE DATA

The respondents were promised confidentiality and that their responses would only be used for purposes of social scientific research. Accordingly, we are taking the strictest precautions to ensure that anonymity is preserved. This means that much of the contextual data that could conceivably be used to identify groups cannot be made public; we also are restricting access to those with a departmental affiliation who can present a clear research and/or educational objective. We also ask that researchers reciprocate by making public any measures or indices that prove helpful, and supplying programs when possible to replicate analyses.

To learn more about the data, you may go to our web-site.² You can also follow links from the Rutgers sociology department (sociology.rutgers.edu); it is listed under "Links to Other Sites" under "Rutgers Sites of Special Interest." From our home page you may learn more about the data set. If you want to download the data, send an e-mail message to us (follow the "contact us" link) explaining your project, and we will return a user-name and a password with which you can reach the password-protected data page.

Analysis of pooled dyadic data from separate networks poses a problem for conventional network software; accordingly, a user-friendly program for Dyadic Analyses of Multiple Networks, previously described in these pages (Martin 1999) is also on the web-site. This program will allow the user not only to analyze the data using QAP models, but create files for UCINET and KRACKPLOT analysis. We believe that the Urban Commune Data Set will provide network researchers with a vast array of opportunities for comparative network research, as one can make structural analyses that compare different groups, the same groups across time, or different relations within the same groups. This may give us a better understanding of the ways in which network properties, frequently assumed to be universal, may depend on group properties hitherto unexplored.

² <http://sociology.rutgers.edu/UCDS/UCDS.htm>

REFERENCES

- Davis, J.A. 1979. The Davis/Holland/Leinhardt Studies: An Overview. In *Perspectives on Social Network Research*, P. W. Holland and S. Leinhardt, eds., 51-62. New York: Academic Press.
- Davis, J.A. and S. Leinhardt. 1972. The Structure of Positive Interpersonal Relations in Small Groups. In *Sociological Theories In Progress Volume 2*, J. Berger, M. Zelditch, Jr., and B. Anderson (eds.), 218-251. Boston: Houghton Mifflin
- Holland, P.W. and S. Leinhardt. 1970. A Method for Detecting Structure in Sociometric Data. *American Journal of Sociology* 70: 492-513.
- Holland, P.W. and S. Leinhardt. 1971. Transitivity in Structural Models of Small Groups. *Comparative Group Studies* 2:107-124.
- Holland, P.W. and S. Leinhardt. 1975. "Local Structure in Social Networks," *Sociological Methodology* 1976: 1-45.
- Homans, G.C. 1950. *The Human Group*. New York: Harcourt, Brace and Company.
- Johnsen, E.C. 1985. Network Macrostructure Models for the Davis-Leinhardt Set of Empirical Sociomatrices. *Social Networks* 7: 203-224.
- Martin, J.L. 1999. A General Permutation-Based QAP Analysis Approach for Dyadic Data from Multiple Groups. *Connections* 22(Spring).
- Zablocki, B. 1980. *Alienation and Charisma: A Study of Contemporary American Communes*. New York: The Free Press.