

Microfinance Data Documentation

Siying Cao *

April 21, 2017

1 Introduction to Data Collection

This is a uniquely rich data set on network structure and microfinance participation in 43 rural villages in Karnataka, a state in southern India. The network data were collected in anticipation of the introduction of services in these villages by a microfinance institution, Bharatha Swamukti Samsthe (BSS), and were drawn from detailed surveys of households covering a wide range of interactions. BSS then entered these villages and provided the authors with data on MF take-up at regular intervals. In 2006, 6 months before BSSs entry into any village, the authors conducted a baseline survey in all 75 villages (in their initial sample).¹ The survey consisted of a village questionnaire, a full census that collected data on all households in the villages, and a detailed follow-up survey fielded to a subsample of individuals. For publications and working papers that draw on this data set, check out Esther Duflo’s website at <https://economics.mit.edu/faculty/eduflo/social>²

2 Data Sets Description

In the data repository, you will find folders “Network Data” and “Demographics and Outcomes”. The “Network Data” folder contains adjacency matrices for each of the 75 villages surveyed.³ About half of households received detailed survey in which individuals were asked to list the names of people with whom they shared a certain relationship.⁴ For each variable, an **individual** matrix and a **household** matrix were constructed. A relationship between households exists if any household member indicated a relationship with members from the other household. 12 dimensions

*Please send any comments to siyingc-at-uchicago-dot-edu

¹Prior to BSSs entry, these villages had almost no exposure to MF institutions, and limited access to any type of formal credit.

²Note the Download Data Link at the bottom of the page is not working. However, it is preserved in Diffusion of Microfinance data coming with their paper.

³The 75 villages are numbered 1-77, villages 13 and 22 are missing

⁴Households were randomly sampled and stratified by religion and geographic sub-region

of network data were collected, including names of those who they borrow money from, give advice to, help with a decision, borrow kerosene or rice from, lend kerosene or rice to, lend money to, obtain medical advice from, engage socially with, are related to, go to temple with, invite to ones home, and visit in anothers home. The folder also includes the ALL network which is a union and an AND network which is the intersection, done both at the individual and household level. Matlab file `adjacencymatrix` stores the ALL network for each village in a single cell. The networks are undirected. The matrices in csv file format lack row or column headers, but in the folder “adjacency matrix keys”, there are files that link each row to a particular individual or household.

“Demographics and Outcomes” contains 2 files. `Household_characteristics.dta` has demographic information about a households home (e.g. roofing material, type of latrine, quality of access to electric power) and a dummy for whether anyone in the household became a microfinance client, as indicated by leader. `Individual_characteristics.dta` has individual demographic information. The individual questionnaire was administered to a sample of about 46% of all households per village, and asked for information including age, caste, education, language, native home, etc.

Finally the folder “Village Level Leader Centrality” contains a single dta file that records various measures of leader centrality (eigenvector centrality, communication centrality, closeness centrality, etc.) and **village-level** information about the average number of people in a household, fraction of leaders, average savings as well as the microfinance take-up rate in each village.

3 Research Guide

This data set provides a great opportunity to study network structure and information diffusion in the development context. Banerjee et al. (2013) collected and used this data to examine the spread of information about microfinance. They estimate a structural model of diffusion and participation to tease out the two channels through which individuals/households take up MF. Using model parameter estimates, they construct a new measure of communication centrality, called diffusion centrality that could predict how widely and quickly others would be likely to make use of the information the leaders (the first-informed individuals) passed on. The unique multi-dimensional network measures may also be an exciting venue to explore. An example is Jackson et al. (2012).

References

- [1] Abhijit Banerjee, Arun G. Chandrasekhar, Esther Duflo, and Matthew O. Jackson. The Diffusion of Microfinance. *Science*, 341(6144):1236498, July 2013.
- [2] Matthew O. Jackson, Tomas Rodriguez-Barraquer, and Xu Tan. Social Capital and Social Quilts: Network Patterns of Favor Exchange. *The American Economic Review*, 102(5):1857–1897, 2012.