
0. Summary

Date: September 16, 2013

This project contains data from a survey of social networks in 75 villages in rural southern Karnataka, a state in India. A census of households was conducted, and a subset of individuals was asked detailed questions about the relationships they had with others in the village. This information was used to create network graphs for each village, included here. Basic information for all households and all surveyed individuals is also provided.

Previously, these 75 villages had been identified by Bharatha Swamukti Samsthe (a microfinance institution) as places where it planned to start operations. Six months after the survey, BSS began expanding to these villages and in two years eventually spread to 43 of them. Information about which households participated in BSS's microfinance program was matched to the household census.

There are four sections in this package:

1. "Matlab Replication" contains data and mfiles needed to simulate the model. Data includes matrices describing the relationship between individuals and households in each village, demographic information for households and individuals, and information on which households participated in BSS's program.
2. "Stata Replication" contains data and do files needed to run the reduced-form analysis of leader centrality measures.
3. "Data" contains individual adjacency matrices for each of the surveyed villages, demographic and outcome data, and keys to link the former to the latter.
4. "Survey Instruments" has copies of the surveys that were administered.

The following are some notes on the sections.

1. Matlab replication

The "Matlab Replication" folder contains the following files:

- 1.1 "adjacency matrix": This file contains the adjacency data for each of the 75 villages.
- 1.2 "hhcovariates ##": These files contain data on number of rooms in a household, number of beds, whether the household has private/government/no electricity, whether the household has own/common/no latrine, number of rooms per capita and number of beds per capita.
- 1.3 "HHhasAleader ##": These files contain data on whether a household member is a village leader.
- 1.4 "Omega_rel ##", "Omega_abs ##"
- 1.5 "mf ##" : These files contain microfinance participation data.

2. Stata Replication

In the "Stata Replication" folder, there are two files that contain information on leaders and villages.

2.1 "cross_sectional.dta" contains leader centrality data (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, etc. as well as the microfinance take-up rate in each village.

2.2 "panel.dta" contains the empirical and simulated microfinance take up rate across villages over time. Leader centrality and village-level information is also included.

3. Data

The "Data" folder contains two subfolders: "Network Data" and "Demographics and Outcomes." In the "Network Data" folder, there are adjacency matrices for each of the 75 villages surveyed. The 75 villages are numbered 1-77 (villages 13 and 22 are missing.) About half of households received detailed surveys in which individuals were asked to list the names of people with whom they shared a certain relationship. Households were randomly sampled and stratified by religion and geographic sub-region.

For each variable, an individual matrix and a household matrix were constructed. A relationship between households exists if any household members indicated a relationship with members from the other household. These questions were asked in the individual survey.

Individuals were asked 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 one's home
- visit in another's home.

We also include the ALL network which is a union and an AND network which is the intersection. This is done both at the individual and household levels.

The networks are undirected (each matrix is symmetric). Each row or column of a matrix represents one individual. The matrices lack row or column headers, but in the folder "adjacency matrix keys" there are files that link each row (or the equivalent column) to a particular individual or household.

For instance, using the file "adj_templecompany_vilno_1.csv," we discover that the individual in the fifth row goes to the same temple as the individual in the eighth row. Using "key_vilno_1.csv," we can determine that in the individual in the fifth row has an ID of 100201, and the individual in the eighth row has an ID of 100204. We can then look up demographic information about these individuals using the dataset in the "Demographics and Outcomes" folder by matching on the ID variable.

To make this process easier, the datasets in "Demographics and Outcomes" already contain a variable called "adjmatrix_key" which lists that observation's corresponding row number in the adjacency matrix.

3.1 List of household network files.

```
adj_borrowmoney_HH_vilno_##  
adj_giveadvice_HH_vilno_##  
adj_helpdecision_HH_vilno_##  
adj_keroricecome_HH_vilno_##  
adj_keroricego_HH_vilno_##  
adj_lendmoney_HH_vilno_##  
adj_medic_HH_vilno_##  
adj_nonrel_HH_vilno_##  
adj_rel_HH_vilno_##  
adj_templecompany_HH_vilno_##  
adj_visitcome_HH_vilno_##  
adj_visitgo_HH_vilno_##  
  
adj_allVillageRelationships_HH_vilno_##  
adj_andRelationships_HH_vilno_##
```

3.2 List of individual network files

```
adj_borrowmoney_vilno_##  
adj_giveadvice_vilno_##  
adj_helpdecision_vilno_##  
adj_keroricecome_vilno_##  
adj_keroricego_vilno_##  
adj_lendmoney_vilno_##  
adj_medic_vilno_##  
adj_nonrel_vilno_##  
adj_rel_vilno_##  
adj_templecompany_vilno_##  
adj_visitcome_vilno_##  
adj_visitgo_vilno_##  
  
adj_allVillageRelationships_vilno_##  
adj_andRelationships_vilno_##
```

3.3 List of key files

```
key_HH_vilno_##  
key_vilno_##
```

In the "Demographics and Outcomes" folder, there are two files that describe other information about the individuals and households surveyed.

3.4 "Household_characteristics.dta" has demographic information about a household's home (roof type, number of rooms, latrine type, etc.) and a dummy that indicates whether anyone in the household became a microfinance client. The household survey was conducted for every household.

3.5 "Individual_characteristics.dta" has individual demographic information (age, caste, religion, language, occupation, etc.) This more detailed survey was conducted among a little under half of households, and also asked for social network information, which is represented in the previous tables.

4. Survey instruments

The "Survey Instruments" folder contains the surveys used in the study. Most questions contained in this data were obtained using "household-final.doc" and "individual-final.doc."