Tamil Nadu Socio-Economic Mobility Survey: An Overview

Yale Economic Growth Center & Center for Microfinance, IFMR

November 13, 2014
Preface

The purpose of the Tamil Nadu Socioeconomic Mobility Survey (TNSMS) is to have an interdisciplinary study that looks at the economic and social mobility of households, individuals and communities over a period of medium-term and long-term. The strategy behind the study is to allow all those interested in understanding development to investigate relationships between seemingly disparate factors that determine economic and social transformations. To do so, we have implemented a large-scale panel survey in the Indian state of Tamil Nadu that will extend for at least 15 years. Data collection for the baseline survey round has been completed in 2010.

TNSMS collects data at the individual level in health, education, employment, cognitive skills, tests scores for people under 14 years of age, involvement in social and political organizations and migration. This is supplemented by data at the household level on agricultural patterns, social networks and wealth. The data also includes information on community facilities, infrastructure, access to public services and participation in organizations. The 2010 TNSMS covers a sample of 9171 households from 30 of the 32 districts across urban and rural Tamil Nadu, India.

The Yale EGC-CMF Tamil Nadu Socioeconomic Mobility Survey is a collaborative effort between the Yale Economic Growth Center (ECG) and the Center for Micro Finance (CMF), IFMR and funded principally by the EGC and ICICI Foundation for Inclusive Growth.

This document describes the various features of the data collected by the TNSMS. One can also find a description of the sampling strategy the response rates and field operations implemented for the purpose of data collection.
Acknowledgements

The Tamil Nadu Socioeconomic Mobility Survey is a collaborative effort between Yale Economic Growth Center and the Center for Micro Finance, IFMR and funded principally by the EGC and ICICI Foundation for Inclusive Growth. The field work, questionnaire development, public use file creation are the combined efforts of the Principal Investigator; Prof. Mark Rosenzweig at Yale University; Research Manager Amy Jensen Mowl at CMF, IFMR and Research Associate Vaishnavi Pratap at CMF, IFMR. This project would not have been possible but for the efforts of Prof. Mark Rosenzweig at Yale University and the Executive Director of IFMR LEAD, Sharon Buteau and President of IFMR C.V. Krishnan.
Contents

0.1 INTRODUCTION ........................................ 4
    0.1.1 Scope ........................................ 5
    0.1.2 Inter-disciplinary Approach .................... 6
0.2 Census Questionnaire ................................. 7
    0.2.1 Sampling Design and Sample Size ............... 7
    0.2.2 Scope of the survey ............................ 7
0.3 Household Questionnaire ............................. 7
    0.3.1 Sampling Design and Sample Size ............... 7
    0.3.2 Scope of the survey ............................ 8
0.4 Individual Questionnaire ............................ 9
    0.4.1 Sampling Design and Sample Size ............... 9
    0.4.2 Scope of the survey ............................ 9
0.5 Married women Questionnaire ........................ 10
    0.5.1 Sampling Design and Sample Size ............... 10
    0.5.2 Scope of the survey ............................ 10
0.6 Community Questionnaire ............................ 10
    0.6.1 Sampling Design and Sample Size ............... 10
    0.6.2 Scope of the survey ............................ 11
0.7 Data Cleaning and Publication ...................... 12
0.8 Tables .............................................. 12
0.1 INTRODUCTION

Economic development is inherently a dynamic but often slow process characterized by changes in institutions, the natural environment and by the mobility of individuals. Advances in the scientific understanding of both the multi-level causes and consequences of development have been hampered by the lack of comprehensive data tracking individuals, their environment and institutions over the long term. Our goal is to provide a new laboratory for carrying out a wide range of potential studies of the medium- and long-term changes, or lack of changes, that take place during the process of development. We are particularly interested in providing a framework for the investigation of the large set of issues that have remained outside the scope of scientific analysis because of the short time frame and narrow methodological focus of most existing surveys in developing countries. By committing to a very long term, comprehensive study that obtains information on individuals and the built and natural environment in which they reside through their life courses, we can provide the tools required for a transformation of the study of development.

The guiding principles of the studies are flexibility and a long-run commitment. We aim to provide a research infrastructure that will encompass a wide range of innovative studies of the process of economic development, many of which cannot be predicted in advance. Some examples of the types of research that will become possible with these data are:

- the long-term impacts on individuals of health shocks, childhood nutritional status, and schooling
- local institutional transformation and economic development
- the interactions between ecological transformation and patterns of growth
- the dynamics of household formation and restructuring
- randomized evaluation of development interventions with long-term follow-up
- patterns of long-term occupational or expenditure mobility

The main objective of the Tamil Nadu Socioeconomic Mobility Survey is to provide a scientific framework for a wide range of potential studies of the medium- and long-term changes that are taking place during the process
of development. Our strategy is to permit the investigation of unexpected connections between the multiple transformations that occur during the process of economic development. To do so, we have implemented a large-scale panel survey in the Indian state of Tamil Nadu that will extend for at least 15 years. Data collection for the baseline survey round has been completed in 2010.

The Yale EGC-CMF Tamil Nadu Socioeconomic Mobility Survey is a collaboration between the Yale Economic Growth Center and the Center for Micro Finance, IFMR and funded principally by the EGC and ICICI Foundation for Inclusive Growth. The Economic Growth Center at Yale is leading this research program, but will rely on strong long-term institutional collaboration. In India, the Centre for Microfinance Research at the Institute for Financial Management and Research, an economic research and policy think tank, has partnered with the EGC on the Tamil Nadu Socioeconomic Mobility survey. In addition to collaborating in development of the survey instruments and sampling procedures, the CMF is responsible for operational management of the survey.

0.1.1 Scope

Our strategy is to permit the investigation of unexpected connections between the multiple transformations that occur during the process of economic development. To do so, the project team has planned large scale panel surveys in three developing countries to continue for at least 15 years.

We will conduct surveys once every three years. The interval length is chosen to conserve on costs, and because three years is sufficient to begin to observe significant changes beyond the short-run fluctuations that have been the focus of many other studies. In India, the survey begins with a random sample of 10,000 households in 400 communities in rural and urban areas of the state of Tamil Nadu. Every three years following the initial survey, each individual in the original 10,000 households would be followed for re-interviews. Individuals and households will be followed irrespective of whether they move out of their original community, or where they move to. This aspect of the study design is essential to capture the important, but heretofore understudied, spatial component of economic mobility as we follow individuals over their life courses.

A very broad range of data will be required to maximize the usefulness
of this research infrastructure. Each round of the survey will consist of three key elements: a census of all inhabitants of the community; comprehensive household-level surveys; and community inventories.

0.1.2 Inter-disciplinary Approach

The goal of this project is to shed new light on long-run processes of economic development. Development economics provides the intellectual framework that guides the overall design of project. Progress on understanding economic development, however, relies on insights from a broad range of disciplines. The broad scope of this project provides a natural setting for collaboration across disciplines.

A central goal of the project is to understand the pathways through which local social and political institutions influence patterns of economic development, while taking into account the possibility that the process of economic development itself may shape the evolution of these institutions. There is an extremely broad range of such institutions which we will examine in the context of this long-term study. Three specific examples include: farmers security of land tenure; processes of dispute resolution within villages; and the role of caste and kin groups in occupational and residential choices.

A parallel goal of the project is to examine the dynamic relationships between economic change and broader measures of human well-being. We will investigate long-term interactions between economic development and individual health, nutrition and education.

The project will clarify our understanding of the long-run connections between environmental and economic change. For example, we will monitor how land resource management and farming systems evolve with economic growth. Similarly, we can study how local environmental threats (indoor air pollution, agricultural chemicals, water-borne diseases) are related to individual health and to production and consumption choices.

Intelligent investigation of these broad questions requires collaboration with experts in many fields, including public health, education, environmental sciences, political science, demography, psychology and sociology.
0.2 Census Questionnaire

0.2.1 Sampling Design and Sample Size

The sampling design involved having equal representation from rural and urban areas. In the case of rural areas, 200 villages were selected from a total list of Census of India villages (2001) in Tamil Nadu, India. In the case of urban areas, the study utilizes the National Sample Surveys Urban Frame Survey (UFS) block divisions. The urban selected areas consists of 200 randomly selected UFS blocks from the total number of UFS blocks as listed by the UFS in Tamil Nadu, India. The urban sample covers small, medium and large towns including slum communities. The questionnaire was administered to the knowledgable member of the household.

0.2.2 Scope of the survey

The census of all inhabitants of the community, based on house-to-house interviews, provides the sample frame for the next stage, and also permits measurement of the broad characteristics of inhabitants of community and their linkages with each other.

The rural household census captures household size, and demographic and economic information about the household and its members (occupation, caste, age, education, landowning details, and income from all sources), and location, including GPS coordinates. The urban household listing captures household location and size, as well as demographic and economic information for the head of the household.

0.3 Household Questionnaire

0.3.1 Sampling Design and Sample Size

In each of the urban and rural areas selected for the survey approximately 25 randomly selected households were selected for the household questionnaire. The survey was administered to the most knowledgable adults in the household, which in most cases was the head of the household. A comprehensive household-level survey, based on face-to-face interviews, captures detailed socioeconomic and health data for each sample household and its members. The survey has been successfully completed for all persons from 9,171 house-
holds located in 200 rural villages and 200 sampling units in urban areas of Tamil Nadu, India.

0.3.2 Scope of the survey

The questionnaire has been administered to the most knowledgeable person(s). The modules for the household survey cover aspects that are covered in a regular household sampling survey such as household roster, durable goods, household assets, living condition and consumption expenditure. But it departs from the regular household surveys in collecting other features such as information on non co-residents, cash transfers from government schemes, participation in social welfare schemes and social networks and information on variety of transactions with partners. The survey also a very details component on employment concerning agricultural practices, farm labour, cropping patterns, farm assets, livestock, land tenure, irrigation and business transactions. For those not working in agriculture there is a comprehensive section on non-farm enterprises and self-employment that covers questions concerning financial transactions, number of employees and partners, labour contracts and type of business. While the features that this survey module shares with other standard surveys such as the National Sample Survey (NSS) makes the data credible and comparable to other surveys. At the same time the additional features allow the data to be distinguished from standard household surveys, making it a unique data set that allows one to explore relationships that are not a trivial exercise using other data sets. The modules include:

- Household roster
- Information on non co-resident family members, spouses, and extended family
- Cash transfers
- Participation in social welfare schemes and government benefits
- Agriculture, including cropping patterns, inputs/outputs, assets, livestock, land tenure and transactions, irrigation, farm assets, use of extension services and technology transfer
- Non-farm enterprises and self-employment
- Transfers, financial assets, insurance, borrowing, lending, and gifts
- Durable goods and household assets
0.4 Individual Questionnaire

0.4.1 Sampling Design and Sample Size

There were approximately 25 households that were selected from the census to be administered the household as well as the individual survey. From the 9,171 households that were covered the individual surveys were administered to 32,713 adult individuals between the ages of 15 to 60 years across urban and rural areas. The cognitive tests were administered to children below the age of 14 years.

0.4.2 Scope of the survey

The individual module is a combination of features from the standard individual questionnaires administered by the National Sample Survey Organisation (NSSO) and additional features that allow researchers to study unexplored and relevant relationships. The standard sections include questions on the respondent’s education history, health, earnings, type of employment, utilization of public services and membership in different social organisations. The unusual features include cognitive tests, achievement and a specialised cognitive tests for children under 14 years of age, activities of daily living (ADL), anthropometric measurements, membership in political organisations and detailed migration histories. The modules include:

- Housing conditions
- Consumption expenditure
- Social networks and information on transactions partners for all transactions in the data

- Education history and basic cognitive testing for all members, as well as achievement tests for children under age 14

- Health status and utilization of health services, including anthropometric measurements, ADLs, as well as specialized modules for young children

- Migration history

- Occupation, earnings, and characteristics of employment
• Political and social activities and membership

0.5 Married women Questionnaire

0.5.1 Sampling Design and Sample Size

The questionnaire was administered to ever-married women between the age of 15 and 60 years and in each of the randomly selected 25 households approximately. The survey is successfully administered to 13,478 ever-married women from the sample households in both rural and urban areas.

0.5.2 Scope of the survey

The purpose of the questionnaire is to obtain detailed information on women’s health, fertility, pregnancy care, educational background, time use and child care. This particular questionnaire is a unique exercise in collecting information that is otherwise available in disparate surveys such as the National Family Health Survey (NHFS) and the National Sample Survey (NSS). Having both these features combined into one survey allows one to study relationships that are otherwise not possible. The modules include:

• Marital history
• Fertility history and detailed birth history
• Time use survey for woman as well as for children younger than 14 years
• Child labor
• Contraception and reproductive health

0.6 Community Questionnaire

0.6.1 Sampling Design and Sample Size

A community inventory, based on key informant interviews, documents a broad range of natural and institutional features of the community. The
Community survey was administered to local community leaders, village elders, self-help group leaders, anganwadi teachers and others who may be holding positions in public offices in the survey area. The purpose of collecting information on community level data is to better understand the impact of community factors have on households and individuals. This is not a common feature among most standard households surveys, allowing researchers to explore relationships between formal institutions and community organisations and household or individual status. The community data has been administered from 195 rural areas and 200 urban survey areas.

0.6.2 Scope of the survey

The surveys were first administered to local community leaders - village leaders in rural areas and the appropriate government officials in urban areas. The responses of these leaders is then supplemented using information and data from other sources which include individuals who hold a position in public offices. The questions in the survey cover access to and utilisation of public health and education institutions, availability of infrastructure, presence of development programmes, etc. The survey departs from the regular community surveys by collecting information on macro-factors like wages, prices, settlement patterns and land tenure systems, etc. The modules include:

- Infrastructure
- Formal and informal financial institutions
- Presence of various development programs
- Political organizations and local governance
- Settlement patterns and land tenure systems
- Agriculture and livestock production
- Business structure and employment
- Local prices and wage rates
- Health and education services and facilities
0.7 Data Cleaning and Publication

The data entry process was implemented using the data entry software with various checks and balances. It is the policy of CMF, IFMR and Yale EGC to have the data double entered and reconciled for errors. The process was deemed complete only after the error rates were found to be below the acceptable level of 0.05% after reconciliation. During this process it was ensured that personal identification information is stored separately and entered independently as per the Human Subjects Committee guidelines of Yale University.

The public use files have been made keeping in mind the requirements of researchers for selecting, merging and working on data using their own analysis files. Hence apart from the basic cleaning of the data we have refrained from over-cleaning or adding any form sample weights to the data. We feel these are decisions that are best left to researches. This should allow researchers the flexibility of utilizing the data to best suit their requirements.

0.8 Tables

Table 1: Sample Size Targets and Completed Questionnaires

<table>
<thead>
<tr>
<th>Sample Size</th>
<th>Community Questionnaire</th>
<th>Household Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Target</td>
<td>Actual</td>
</tr>
<tr>
<td>Rural</td>
<td>200</td>
<td>195</td>
</tr>
<tr>
<td>Urban</td>
<td>200</td>
<td>200</td>
</tr>
</tbody>
</table>

Table 2: Matched Individual and Household Questionnaires

<table>
<thead>
<tr>
<th></th>
<th>Matched Questionnaires</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>RURAL Individual Questionnaires</td>
<td>18,104</td>
<td>100%</td>
</tr>
<tr>
<td>Ever-Married Questionnaires</td>
<td>9733</td>
<td>100%</td>
</tr>
<tr>
<td>URBAN Individual Questionnaires</td>
<td>14,609</td>
<td>98%</td>
</tr>
<tr>
<td>Ever-Married Questionnaires</td>
<td>3747</td>
<td>97.8%</td>
</tr>
</tbody>
</table>