

Introduction

These data were collected in the framework of Dr. Annemie Maertens' PhD dissertation during the period August 2007 – July 2009. The dissertation was undertaken from Cornell University, but executed in India in collaboration with the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). The project was sponsored through an NSF Doctoral Dissertation Improvement Grant (Grant No. 0649330).

The main goal of this project was to collect and analyze household survey data in the Indian states of Andhra Pradesh and Maharashtra in order to gain a better understanding of the role of social networks and identity in economic decision-making. The first panel of this research studied the role of social learning and social pressures in *Bacillus thuringiensis* (Bt) cotton adoption using data from three villages (Aurepalle, Kanzara and Kinkhed).

The data collection consisted of five phases: (1) qualitative round (to determine the topic of the two panels); (2) trial round (to field test the questionnaires); (3) training round (to train the enumerators); (4) quantitative collection round (to collect the household level, village level data); (4') data entry of (4) ; (5) data validation round (to collect additional data to correct the missing variables and inconsistencies uncovered in (4')). The first phase was completed by the beginning of November 2007, the second by the end of December 2007, the third by the end of January 2008 and the fourth by the end of July 2008 and fifth by the end of July 2009.

The data collected are available free of charge. I would like to request that those intending to use the data would register by emailing am445@cornell.edu, stating their contact details and research topic. Any questions and comments can be addressed to am445@cornell.edu.

The villages selected for this study are part of the Village Level Studies (VLS) program ICRISAT. In this program, ICRISAT followed 300 randomly selected households from six villages during the period 1975-1985 every three weeks. In 2001, ICRISAT restarted the panel, revisiting 185 of the first generation VLS households and their split-offs, in addition to 261 newly added households. This data collection is currently ongoing:

To obtain the 1975-1985 and 2001-2006 data:

<http://www.icrisat.org/gt-mpi/knowledgeBase/Databases/vls.asp>

To obtain the 2001-2006 data, see also:

<http://www.economics.ox.ac.uk/members/stefan.dercon/icrisat/ICRISAT/index.html>

Published papers resulting from these data

Maertens. Forthcoming. 'Who Cares What Others Think (or Do)? Social Learning and Social Pressures in Cotton Farming in India'. *American Journal of Agricultural Economics*.

Maertens, A., AV Chari and D.R. Just. 2014. 'Why Farmers Sometimes Love Risks: Evidence from India.' *Economic Development and Cultural Change*, 62(2): 239-274.

Maertens, A. and C.B. Barrett. 2013. 'Empirical Methods for Identifying Social Network Effects on Technology Adoption.' *American Journal of Agricultural Economics Papers and Proceedings*, 95(2):353-359.

Structure of the sample

In the first panel of the project, I covered 246 ICRISAT-VLS farmers and 20 non-ICRISAT-VLS farmers. Among the VLS farmers, I conducted a household questionnaire, a plot-level questionnaire for each plot and a recall module if the household had not been part of the ICRISAT-VLS sample since 2001 (these recall data are available on request). Among the additional farmers, whom I labelled progressive farmers, I carried out a (progressive farmer) questionnaire and a supplementary network questionnaire. The latter dataset is available on request. I completed a village questionnaire with the assistance of the village head, three knowledgeable people in each village, the Mandal/Tehsil Revenue Office and the District Collector's Office and collected daily and monthly rainfall data from the local meteorological offices.

Overview of the documents

Codebook

Datasets following the structure in the Table below

ICRISAT-VLS Data		Progressive Farmer Data	Village Data
Input-Output Questionnaire	Household Questionnaire		
Section VIA/B/C/D	Section I	PFSectionI	VSectionI
	Section IIA/B	PFSection IIA/B	VSectionII
	Section III	PFSection IIIA/B	VSectionIIIA/B/C
	Section IVA/B	PFSection IV	Rainfall Data
	Section V	PFSection V	
	Section VII	PFSection VI	
	Section VIII	PFSection VII	
	Section IXA/B/C/D	PFSection VIII	
	Section X	PFSection IX	
	Section XI		

With the assistance of Oxford University, I constructed the following file to link the households over time. This file was constructed using individual level information from ICRISAT/Oxford and the information collected in the recall module. The linking “variable” is the most important agricultural decision maker of the household as identified in Section XIII and Section IXA.