Understanding the Role of Social Learning, Imitation and Social Norms in *Bacillus thuringiensis* Cotton Technology Adoption Decisions*

Annemie Maertens

Cornell University

Dec 14, 2008 - Preliminary and Incomplete, Not for Citation

Abstract

Recent studies in social sciences show that there exist important effects of social networks on technology adoption decisions. None of these studies have differentiated these social interaction effects along pathways. Using quasi-panel data, uniquely collected for this purpose in three villages in rural India covering the period 2001-2008, I show how one can distinguish between the influences of social learning, imitation and social norms on *Bacillus thuringiensis* cotton technology adoption decisions. Preliminary results indicate that social learning and imitation play a significant role in the decision to adopt Bt cotton technology.

*The data for this paper were gathered in India during the period August 2007 – November 2008, in collaboration with the International Crop Research Institute of the Semi-Arid Tropics (ICRISAT) in Patancheru, Andhra Pradesh, India. This research is funded through a combination of grants: a National Science Foundation Doctoral Dissertation Research Improvement Grant, the American Agricultural Economics Association McCorkle Fellowship, a Mario Einaudi Center for International Studies International Research Travel Grant, a Graduate School Research Travel Grant and an International Student and Scholar Office Grant. I am grateful towards the American Institute of Indian Studies for administrative support, the Department of Applied Economics and Management and Chris Barrett for additional financial support, and the Indian Statistical Institute in New Delhi for their hospitality. I'd also like to thank the enumerators and research assistants on the field: Shraavya Bhagavatula, Pramod Bangar, Sana Butool, V.D. Duche, Shital Duche, Anand Dhumale, Shilpa Indrakanti, Navika Harshe, Sapna Kale, Labhesh Lithikar, Nishita Medha, Ramesh Babu Para, Abhijit Patnaik, Dhere Madhav Parmeshwar, Gore Parmeshwar, Amidala Sidappa, Nandavaram Ramakrishna, K. Ramanareddy, P.D. Ranganath, and Arjun Waghmode. This draft has benefited considerably from my discussions with Cynthia Bantilan, Chris Barrett, Kaushik Basu, Larry Blume, Richard Bownas, Sommarat Chantarat, A.V. Chari, V.K. Chopde, William R. Coffman, Stephen Coate, Annelies Deuss, Ronald Herring, George Jakubson, Travis Lybbert, Hope Michelson, Felix Naschold, K.V. Raman, Chandrasekhara Rao, K.P.C. Rao, Mohan Rao, Nageswara Rao, Paulo Santos, Frank Shotkosky, Russell Toth, Thomas Walker, Kumar Acharya Ulu, Abhijit Patnaik, and K.C. Suri.

1 Introduction

When new agricultural technologies are introduced in developing countries, adoption often does not occur immediately. Instead farmers follow a complex pattern of gradual adoption, dis-adoption and sometimes non-adoption, resulting in an S-shaped adoption curve when plotting number of adopters against time (Feder et al. 1985, Besley and Case 1993, Sunding and Zilberman 2001). While prices, income and individual's attributes, such as risk aversion, are known determinants of adoption behavior, recent contributions in sociology and economics have pointed towards the importance of so-called social interaction effects, a general term encompassing both pecuniary and non-pecuniary effects of individuals on eachother, such as social learning, social norms and imitation.

Farmers learn about prices of inputs and outputs, use of inputs and expected yields given these inputs through experimentation, from the media, input dealers, and from eachother, the latter being referred to as social learning (Bandiera and Rasul 2006, Baird 2003, Conley and Udry 2001, Foster and Rosenzweig 1995, Moser and Barrett 2006, Munshi 2004). In addition, in conservative agricultural societies, a farmer's choice could be directly influenced by the choices of other farmers through the prevailing social norms; i.e., deviation from the standard agricultural practises entails a non-monetary cost, for instance, social exclusion (Appadurai 1989, Moser and Barrett 2006, Rogers 1965, Vasavi 1994). Finally, farmers can affect each other through one-on-one imitation, increasing or decreasing the demand for Bt cotton, respectively referred to, using Leibenstein's terminology (1950), bandwagon effect and snobeffect (Bandiera and Rasul 2006, Pomp and Burger 1995, Rogers 1965).

These studies face three main challenges. First, one needs to separate the social interaction effects from the correlated effects, i.e., the unobservables which seemingly coordinate the actions of the agents through similar constraints, for instance climatic and soil conditions. Second, one needs to differentiate between the various social interaction effects. This is not straightforward as these interactions will lead to the same reduced form effects in the data: correlated behavior. For instance, two farmers might be adopting a crop because they both learned from the same source about its profitability, or because one imitates the other, or because only if they coordinate and produce the same crop, will a trader find it worthwhile to come to the village and pick up the produce. However, differentiation along pathways is crucial from a policy perspective. The right strategy to stimulate increased and quicker uptake of promising technologies depends fundamentally on the structure of the technology adoption processes at the farm level; in other words, different constraints imply different policy measures. If a lack of information is the constraining factor, agricultural extension services might be a first best response. But which farmers does one then target: the well-educated, richer farmers, or the less-educated, poorer farmers? If farmers in a certain village seem to imitate the village leader, convincing the village leader of the benefits of the new technology might be sufficient. Third, one needs to find measures of the social relations of each individual in the sample, a difficult task due to the multi-dimensionality and sheer multitude of these relationships.

In this article, I build on the existing literature; using the adoption process of *Bacillus thuringiensis*, henceforth Bt, cotton in rural India as an example, I show how to distinguish between the influences of three different social interaction processes: social learning, imitation and social norms. The Bt technology was introduced in India in 2002. Studies have shown that adopting this technology can lead to a substantial reduction in pesticide use and sizable yield effects where pest damage by bollworms is not effectively controlled otherwise (Qaim 2003).

The remainder of this article is structured as follows. The next section gives some background information on the Bt cotton technology, the study area and the data collected. Section three outlines the theoretical model and the empirical identification strategy. Section four discusses the results and section five concludes.

2 Background

2.1 *Bacillus thuringiensis* cotton technology in India

India is one of the largest producers of cotton in the world. In 2008, India produced 5,443 thousand metric tons of cotton versus 2,985 thousand metric tons in the US. About one fourth of this production is exported. In terms of yield, India is at the tail-end of the distribution. The average cotton yield in 2008 in India is 579 kg/ha versus 951 kg/ha in the US and 1,325 kg/ha in China (see Table 1). However, yields in India have increased substantially since the early 90s (see Figure 2.1). The main cotton producing states in India are Gujarat, Maharashtra, Punjab and Andhra Pradesh, producing over 80 percent of the cotton in India, with an average yield of, respectively, 625 kh/ha, 253 kg/ha, 750 kg/ha and 381 kg/ha ginned cotton in 2006-07.[1]

Table 1.1: Basic cotton statistics of China, US and India

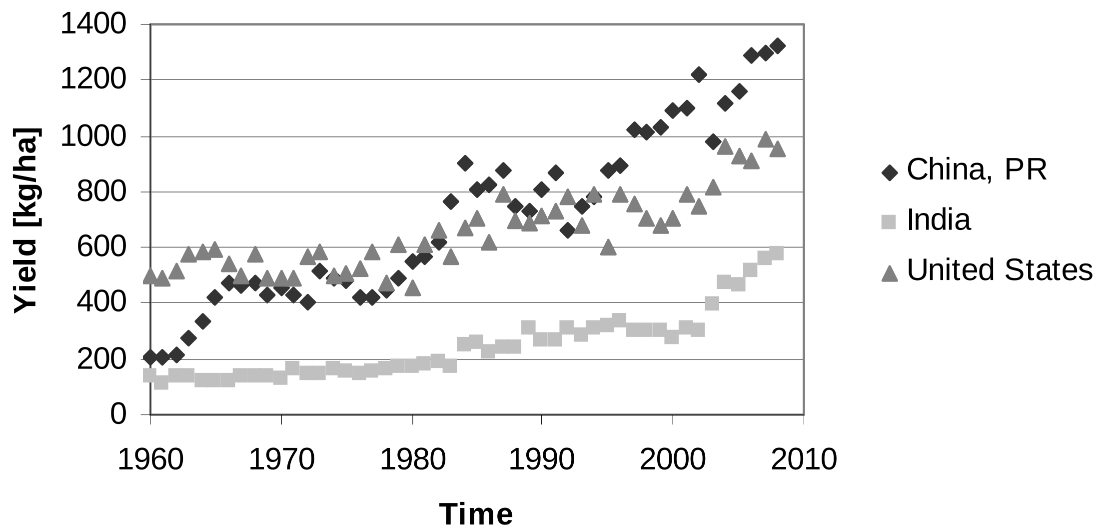
Country	Cotton Production*						Cotton Exports[1]	Cotton Yields[2]
	1980	1985	1990	1995	2000	2008	2008	2008
China	2,700	4,137	4,507	4,768	4,420	7,947	16	1,325
USA	2,421	2,924	3,376	3,897	3,742	2,985	2,830	951
India	1,322	1,964	1,989	2,885	2,380	5,443	1,328	579

Source: United States Department of Agriculture, PSD Online, Updated 10/10/2008.

Notes: [1] in thousand metric tons; [2] in kg/ha (ginned) cotton

Main losses in cotton production in India are due to its predominant cultivation under rainfed conditions and its susceptibility to 166 species of insects, pests and diseases. Today, around 50 percent of pesticides used in India are used on cotton (ISAAA 2005). The major pests affecting cotton are: jassids, aphids, white fly, and bollworms. The cotton bollworm complex comprises the American bollworm (*Helicoverpa armigera*), pink bollworm (*Pectinophora gossypiella*), spiny bollworm (*Earias insulana*) and spotted bollworm (*Earias vittella*) (ISAAA 2005, Asia-Pacific Consortium of Agricultural Biotechnology 2006).

[1]Source: INDIASTAT. Selected State-wise Estimated Yield of Cotton in India in 2006-07. Note the all-India average is 421 kg/ha.



Evolution of the cotton yield in China, US and India Source: same as Table 1

As a response to the bollworm pest problems, Monsanto developed the Bt GM (Genetically Modified) technology during the 1980s. In collaboration with the Maharashtra Hybrid Seed Company (Mahyco), the technology was then introduced into several of Mahyco's hybrid breeding lines during the 1990s. In 2002, the Genetic Engineering Approval Committee (GEAC) approved the commercial release of three Bt cotton cultivars of Mahyco. As of August 2008, 225 cotton cultivars with one of the Bt constructs have been approved by GEAC. These Bt cultivars contain a gene (*cry gene*) sourced from the soil bacterium *Bacillus thuringiensis* in their DNA sequence.[2] This gene produces a protein that is toxic to several insects of the Lepidoptera order; amongst others, the American bollworm, the spiny bollworm, the spotted bollworm, and to a lesser extent, the pink bollworm. The Bt gene does not effectively control against all bollworms and provides no protection against other pests and diseases. Also, when a Bt gene is inserted in the DNA of a plant it only affects its pest resistance. It does not affect its duration, drought resistance, fiber length, expected yield etc. These properties are determined by the cultivar in which the gene was inserted. However, only few studies compare two "isogenic" cultivars (i.e., having the same genes),

[2]This protein, when entering the gut of the insect in the larvae phase, meets a receptor protein, it binds with it, and punctures the wall of the intestine, which leads to paralysis and eventually death of the insect. This receptor protein is only found in insects of the Lepidoptera order.

the Bt cultivar versus the non-Bt cultivar. The effect on pesticide use, yield and hence profits from switching from any non-Bt to any Bt cultivar can therefore be decomposed in two effects: the Bt effect and the germplasm effect. The Bt technology can lead to a substantial reduction in pesticide use and sizable yield effects where pest damage by bollworms is not effectively controlled otherwise.

2.2 The data collected

The three villages selected for this study are part of the Village Level Studies (VLS) program of the International Crop Research Institute of the Semi-Arid Tropics (ICRISAT). In this program, ICRISAT followed up 300 randomly selected households from six villages during the period 1975-1985 on a three-weekly basis. This dataset, known as the first generation VLS, contains detailed household (income, wealth, consumption, and labor) and plot level (input/output) data.[3] In 2001, ICRISAT restarted the panel, revisting the first generation VLS households and their split-offs, in addition to x newly added households to make the sample representative for each village in terms of land-holding size.[4] During the period August 2007-November 2008, I collected an additional round of data among the three out of six villages: Aurepalle in the Mahbubnagar district in Andhra Pradesh and Kanzara and Kinkhed in the Akola district in Maharashtra. I covered 245 VLS farmers, 20 progressive farmers and 3 village leaders using a combination of experiments and questionnaires, including both quantitative and qualitative, objective and subjective, closed and open-end questions. Among the VLS farmers, I conducted a household questionnaire and plot-level questionnaires for each plot. Among the progressive farmers, I carried out a progressive farmer questionnaire, containing a larger recall-section than the household questionnaire. Finally, I completed a village questionnaire, including information on climate and village infrastructure, with the assistance of the village pradhan, three knowledgeable people in each village, the Mandal/Tehsil Revenue Office[5] and the District Collector's Office.

[3] For an overview of the goals, methods and outcomes of the first generation VLS see Singh et al. (1985) and Walker and Ryan (1990).

[4] For an overview of the goals, methods and outcomes of the second generation VLS up to 2005 see Bantilan et al. (2006) and Rao and Charyulu (2007)

[5] Mandal refers to the third-level administrative area in Andhra Pradesh, below state and district. The equivalent in other states is tehsil.

The household questionnaire includes sections on household composition, landholding (including soil characteristics), agricultural machinery and income, and a recall section on cotton adoption, production and marketing.[6] It also contains a section on self-perception regarding risk-aversion, time-preferences and ability, perceived health and environmental hazards associated with Bt cotton and social networks. In addition, I elicited risk aversion, beliefs regarding yield distributions and social networks within the village using an experimental set-up. The risk game, based on Lybbert and Just (2007), consists of four hypothetical farming seasons; for each season I elicited the farmer's maximim willingness-to-pay for a bag of cotton seed that gives a particular yield distribution. The results of this game provide a proxy for risk aversion. The yield distribution game, based on Lybert et al. (2007), asks the farmer to construct four yield density function of two Bt cultivars and two non-Bt cultivars of their choice conditional on the plot characteristics of their own plots. The plot-level questionnaire includes questions on the per-plot agricultural inputs used and outputs produced, included prices.

I measure social networks in three different ways. The first way is to ask the farmer how many farmers he knew in each year since 2001-02 in different social groups (total, village, relatives) that adopted Bt cotton in that year and what the experience of these farmers was with Bt cotton. This method has the advantage of capturing all information and imitation links of the farmer, but, as we know little about these contacts, provides little information on the nature of learning and imitation. Secondly, I use a random matching within sample experiment based on Conley and Udry (2001) and Santos and Barrett (2007). Through this method, I elicit the characteristics of the relationship between two randomly drawn respondents, both members of the VLS sample. Each VLS respondent is matched up with six randomly drawn VLS respondents and four fixed progressive farmers. The questions asked regarding this relationship include: how long have you known person X?, how frequently do you talk to person X?, how risk averse are you compared to person X?, how profitable are you on your farm compared to person X?, and a set of questions on the knowledge of the X's farming activities in terms of inputs and outputs. As we know the characteristics of person X drawn, this method can shed more light on the nature of learning and imitation, but

[6] This section overlaps with the data collected in the 2001-2007 ICRISAT-VLS.

might - under certain conditions - incorrectly represent the population of information and imitation contacts of the respondent. Thirdly, I asked the respondent who he would go to for information if they had problems with their cotton crop, including some characteristics of person X's farming activities and the relationship between the respondent and person X. This method provides a proxy of all the strong information links of the respondent. To control for information from institutional sources, I added a section on the information obtained since 2001-02 from contacts with extension agents, NGOs, input dealers and ICRISAT.

The questionnaire includes one direct question regarding social norms ("when you adopted Bt cotton, did you experience any resistance from fellow farmers, relatives or others?"). As such, identification of the social norms effect will be solely based on the implications of theoretical model. As some of the resistance in agricultural societies find its roots in a collective fear of the unknown, I collected data on the farmer's views and the perceived views of different social groups of the effects of Bt cotton on animal health, human health and the environment.

I add these data to six rounds of the ICRISAT-VLS data (2001-07), containing household composition, landholding, per-plot agricultural inputs used and output produced, machinery, income and wealth data, to form a quasi-panel dataset, encompassing seven cropping years, from 2001-02 to 2007-08.

Table 2.1 gives the basic statistics of the three villages. Aurepalle, with 925 households individuals is the largest village of the three. It is situated in the drought-prone Telangana region of Andhra Pradesh. The soils in this region, alfisols, are generally sufficient for good crop production, but as they are acidic in nature, require fertilizers and tend to be prone to erosion. Kanzara and Kinkhed, with respectively 319 and 189 households, are located in the somewhat less dought-prone Akola district of West Maharashtra. The soils in this region, vertisols, with their high clay content and poor physical condition are very susceptible to erosion, but as their chemical properties are excellent, can give high yield if properly managed.The average education level of the respondent (i.e., the main decision-maker with regard to agriculture) is low, especially in Aurepalle (2.21 years). In Aurepalle, Kanzara and Kinkhed, respectively, 60%, 85% and 83% of the sampled households have farmed cotton in the last seven years and of these, respectively, 76%, 47% and 10% have adopted Bt cotton at any point in time in the last seven years.

Table 2.1: Introducing the three study villages

	Aurepalle	Kanzara	Kinkhed
Number of households in village	925	319	189
Number of households in sample	127	63	55
Soils[1]	Alfisol	Vertisols	Vertisol
Average rainfall (mm/year)[2]	542	1140	1052
Distance to nearest town (km)	10;12	9	12
Phone service in the village (year establishment)	1978;2003	2002	1993
% of households that farms cotton[3]	60	85	83
% of cotton farmers that adopts Bt cotton[3]	76	47	10
Average education level of respondent (in years)	2.21	6.5	6.4
Average number of household members	4.22	4.87	4.50
Average yearly income (Rs)[4]	43,543	53,720	38,087

Notes: [1]Source: Walker and Ryan (1990) using the USDA Soil Taxonomy system; [2]Period: 2005-2007; [3]period:2001-2008; [4] in 2004-2005. The two figures for Aurepalle refer to the main village of Aurepalle and the sub-village Nallavaripalli respectively

3 Model and identification strategy

3.1 Theoretical model

Consider a representative farmer, denoted with subscript i. Subscript i is implied unless otherwise noted. Appendix B gives a list of the notation used in the model. The farmer's per-period well-being u is a function of the farmer's consumption, denoted c, and a non-material satisfaction term, denoted s. [7] I assume that this (differentiable) utility function takes the following shape:

$$u_t \equiv u(c_t, s_t) \tag{3.1}$$

[7]Note that I abstract from intra-household issues and the labor-leisure decision.

$$\frac{\partial u}{\partial c_t} > 0 \text{ and } \frac{\partial^2 u}{\partial c_t^2} < 0 \text{ and } \frac{\partial u}{\partial s_t} > 0 \tag{3.2}$$

The satisfaction term is included to capture social norms and imitation effects and is defined as in 3.3, with $A_{i,bt,t}$ and $A_{j,bt,t}$, respectively, farmer's i and farmer j's acreage under Bt cotton at time t, N_i the set of farmers to whom farmer i is connected, α_{ij} a parameter capturing the importance of individual j for farmer i in determining the social norm, $\mathbf{A}_{i,bt,t}$ a 1 by $|N_i|$ vector with each element of the vector equal to $A_{i,bt,t}$ and $\mathbf{A}_{j,bt,t}$ a 1 by $|N_i|$ vector with the $|N_i|$ elements equal to $A_{j,bt,t}$ for all farmers j farmer i is connected to.[8]

$$s_{it} \equiv s\left[\left|A_{i,bt,t} - \sum_{\forall j \in N_i} \alpha_{ij} A_{j,bt,t}\right|, |\mathbf{A}_{i,bt,t} - \mathbf{A}_{j,bt,t}|\right] \tag{3.3}$$

The first term in 3.3, denoted by Δ_{it}, is the absolute value of the deviation of farmer's i behavior from the social norm. This social norm is represented by the weighted sum of the actions of the other farmers within i's network, with $\sum_{j \in N_i} \alpha_{ij} = 1$. The second term in 3.3, a 1 by $|N_i|$ vector, captures imitation effects. Each element of this vector, denoted by Δ_{ijt}, represents the deviation of farmer's i behavior from farmer j's behavior. I assume that deviation from the social norm entails a loss in satisfaction.

$$\frac{\partial u}{\partial \Delta_{it}} < 0 \tag{3.4}$$

Turning to the production side of the model, the farmer possesses four kinds of capital: labor (L), land (A), unproductive wealth (W) and knowledge (K). Labor and land are assumed to be fixed over time. [9] The stochastic production function of a crop-cultivar, denoted $k \in \{1, 2, ..K\}$, is defined as in 3.5, with L_{kt} and A_{kt}, respectively, labor and land allocated to crop-cultivar k in period t, $\mathbf{x}_{kt}$ a vector of variable inputs to crop-cultivar in period t and ϵ_k capturing unexpected shocks caused by weather fluctuations, diseases and pests etc. For tractability, I abstract from the learning process about the production function and include

[8] Note that I abstract from the various Bt-cotton crop cultivars available.

[9] Note that I did not include a human capital state variable. This also implies that I abstract from the health implications of pesticide use. Even though the health implications (for the farmer) of incorrect pesticide use are known to be severe in developing countries (Antle and Pingali 1994, Pingali et al. 1994), the data does not allow me to look at these issues.

a crop-cultivar specific knowledge term K_{kt} in the production function:

$$y_{kt} \equiv y_k \left(A_{kt}, L_{kt}, x_{kt}, K_{kt}, \epsilon_k\right) \tag{3.5}$$

$$\epsilon_k \sim N(0, \sigma_k) \tag{3.6}$$

Knowledge can now be defined as a vector of crop-cultivar specific knowledge:

$$K = (K_1, K_2, ...K_K) \tag{3.7}$$

Immediately after period t begins, crops are harvested. Farmer i observes y_{ikt} of all crop-cultivars on his fields and gets to know y_{jkt} of the fields of the farmers he is connected to. He updates his knowledge to K_t and his wealth to W_t and makes his decisions: consumption, c_t, net-borrowing, d_t, labor, L_{kt+1}, land, $A_{kt+1}$and $\mathbf{x}_{kt+1}$variable inputs for each crop-cultivar.

Assuming a time-separable utility function and discrete time steps, $\tau \in \{t, t+1, ...\}$, where τ is the 'integration dummy', the farmer maximizes at each time period t the discounted flow of instantaneous utility over an infinite horizon[10]:

$$V(W_t, K_t) = \max_{\left\{c_t, d_t, \{L_{kt+1}, A_{kt+1}, x_{kt+1}\}_{\forall k}\right\}_{\forall t}} E\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} u(c_\tau, s_\tau)\right] \tag{3.8}$$

In 3.8, V(•) is the value function of the constrained maximization problem at time t, $\delta \in (0, 1)$ is the discount rate, summarizing preferences over time, and E is the expected value operator given knowledge K. The farmer maximizes the objective function subject to multiple

[10]Note that I have abstracted from the fact that some variable input decisions are made throughout the growing season, for instance, pesticides. However, as far as the decision to adopt Bt cotton versus non-Bt cotton is concerned, and the acreage under Bt cultivation, what matters is the expected pesticide use, not the actual pesticide use.

constraints:

$$\sum_{\forall k} L_{kt+1} = L \tag{3.9}$$

$$\sum_{\forall k} A_{kt+1} = A \tag{3.10}$$

$$c_t + \sum_{\forall k} \mathbf{p}_i.\mathbf{x}_{kt+1} - d_t + r.d_{t-1} \leq W_t \tag{3.11}$$

In 3.11, the budget constraint, p_c is set equal to zero, $\mathbf{p}_i$ denotes the vector of variable input prices and r denotes the gross interest rate. 3.9 reflects the labor constraint and 3.10 reflects the land resource constraint. The farmer's choices are also subject to the following laws of motion:

$$W_{t+1} = \sum_{\forall k} p_k.y_{k,t+1} + \left[W_t - \left[c_t + \sum_{\forall k} \mathbf{p}_i \mathbf{x}_{kt+1} - d_t + r.d_{t-1}\right]\right] \tag{3.12}$$

$$K_{it+1} = K\left(K_{it}, \{y_{ikt}\}_{\forall k}, \{y_{jkt}\}_{\forall k, \forall j \in N_i}\right) \tag{3.13}$$

In 3.12 $\mathbf{p}_k$ denotes the price of crop-cultivar k . Note from 3.11 that large borrowings can negative liquid wealth. As such, either per-period constraints or a transversality condition needs to be imposed on the farmer in order to avoid a endless cycle of debt accumulation. Opting for the latter, I impose:

$$\lim_{t \longrightarrow \infty} \left[\delta^t \frac{\partial u(c_t)}{\partial c_t} W_t\right] = 0 \tag{3.14}$$

Finally, the control variables in this dynamic optimization problem are subject to a non-negativity constraint and the state variables to a set of initial conditions:

$$c_t, L_{kt+1}, A_{kt+1}, \mathbf{x}_{kt+1}, d_t \geq 0 \quad \forall k \tag{3.15}$$

$$W(0) = W_0 \text{ and } K(0) = K_0 \tag{3.16}$$

Under certain conditions, the sequential problem as defined by 3.8 - 3.16 is equivalent to the simpler two-period problem, i.e., their value functions and solutions are identical. Appendix A derives the First Order Conditions and Envelop Conditions of this two-period problem. Let's now focus on the production decisions regarding Bt cotton. Assume an interior solution, i.e., $L_{bt,t} > 0$, $A_{bt,t} > 0$ and $\mathbf{x}_{bt,t} > 0$. Combining the First Order Conditions and the Envelope Conditions results in the following three Euler Equations:

$$L_{bt,t} : \delta.E\left[\frac{\partial u_{t+1}}{\partial c_{t+1}}.p_{bt}.\frac{\partial y_{bt,t+1}}{\partial L_{bt,t+1}} + \delta.\frac{\partial u_{t+2}}{\partial c_{t+2}}.\left[p_{bt}.\frac{\partial y_{bt,t+2}}{\partial K_{bt,t+2}}\right]\frac{\partial K_{t+1}}{\partial y_{bt,t+1}}.\frac{\partial y_{bt,t+1}}{\partial L_{bt,t+1}}\right] = \lambda_L \tag{3.17}$$

$$A_{bt,t} : \left[\frac{\partial u_t}{\partial s_t}\frac{\partial s_t}{\partial \Delta_{it}} + \left\{\frac{\partial s_t}{\partial \Delta_{ijt}}\right\}_{\forall_j}\right] + \delta.E\left[\begin{array}{c}\frac{\partial u_{t+1}}{\partial c_{t+1}}.p_{bt}.\frac{\partial y_{bt,t+1}}{\partial A_{bt,t+1}} + \\ \delta.\frac{\partial u_{t+2}}{\partial c_{t+2}}.\left[p_{bt}.\frac{\partial y_{bt,t+2}}{\partial K_{bt,t+2}}\right]\frac{\partial K_{t+1}}{\partial y_{bt,t+1}}.\frac{\partial y_{bt,t+1}}{\partial A_{bt,t+1}}\end{array}\right] = \lambda_A \tag{3.18}$$

$$\mathbf{x}_{bt,t} : \left[\frac{\partial u_t}{\partial c_t}. - p_i\right] + \delta E\left[\begin{array}{c}\frac{\partial u_{t+1}}{\partial c_{t+1}}.p_{bt}.\frac{\partial y_{bt,t+1}}{\partial \mathbf{x}_{bt,t+1}} + \\ \delta.\frac{\partial u_{t+2}}{\partial c_{t+2}}.\left[p_{bt}.\frac{\partial y_{bt,t+2}}{\partial K_{bt,t+2}}\right]\frac{\partial K_{t+1}}{\partial y_{bt,t+1}}.\frac{\partial y_{bt,t+1}}{\partial \mathbf{x}_{bt,t+1}}\end{array}\right] = 0 \tag{3.19}$$

In 3.17 and 3.18 λ_L and λ_A, respectively, denote the Lagrange multiplier on the labor and land constraint. Following 3.17 the farmer choses his labor allocation such that the expected marginal utility of allocating one unit of labor to each one of the crops is equal for all crops and equal to the shadow price of labor. The expected marginal utility of allocating one more unit of labor to Bt cotton is the sum of two terms: the additional utility due to additional production and hence consumption in the next period and the additional utility due to increased knowledge on the Bt cotton production function which will effect production and consumption in the next-next period. In the case of land, social norms and imitation enter the picture. The first term in 3.18 reflects the loss or gain in utility due to synchronising one's actions with the established social norm and/or with specific individuals. In case of the variable inputs, the farmer additionaly trades-off losses in current consumption with gains in future consumption.

To conclude this section, consider the binary Bt cotton adoption decision. Denote the optimal solution at time period t conditional on Bt cotton adoption as $\left(c_t^*, d_t^*, \left\{L_{kt+1}^*, A_{kt+1}^*, x_{kt+1}^*\right\}_{\forall k}\right)$ and the corresponding satisfaction and state variables as $s_t^*, W_{t+1}^*, K_{t+1}^*$ and the optimal solution conditional on non-adoption as $\left(c_t^{**}, d_t^{**}, \left\{L_{kt+1}^{**}, A_{kt+1}^{**}, x_{kt+1}^{**}\right\}_{\forall k}\right)$ and the corresponding satisfaction and state variables as $s_t^{**}, W_{t+1}^{**}, K_{t+1}^{**}$. When making the discrete adoption decision at period t, the farmer compares the following two value functions:

$$V_{bt}(W_t, K_t) = [u(c_t^*, s_t^*)] + \delta E\left[V(W_{t+1}^*, K_{t+1}^*)\right] \tag{3.20}$$

$$V_{non-bt}(W_t, K_t) = [u(c_t^{**}, s_t^{**})] + \delta E\left[V(W_{t+1}^{**}, K_{t+1}^{**})\right] \tag{3.21}$$

3.2 Econometric identification strategy

There are several ways in which one can approach estimation of the theoretical model outlined above. One could specify functional forms for 3.1, 3.3 and 3.5, derive the policy functions of interest ($L_{bt}(\cdot), A_{bt}(\cdot), x_{bt}(\cdot)$ and the binary Bt adoption decision rule) and estimate these using GLS or maximum likelihood. On the other end of the spectrum, one can represent the reduced form policy function using a flexible functional forms such as a trans-log specification, or approximate the reduced form policy function using a second-order Taylor Expansion. The reduced form policy functions for L_{bt}, A_{bt} and $\mathbf{x}_{bt}$ are:

$$L_{i,bt,t} = L(\{p_{k,t}\}_{\forall k}, \mathbf{p}_{x,t}, \sum_{\forall j \in N_i} \alpha_{ij} A_{j,bt,t}, \{A_{j,bt,t}\}_{\forall j \in N_i}, W_{it}, K_{it}, L_i, A_i) \tag{3.22}$$

$$A_{i,bt,t} = A(\{p_{k,t}\}_{\forall k}, \mathbf{p}_{x,t}, \sum_{\forall j \in N_i} \alpha_{ij} A_{j,bt,t}, \{A_{j,bt,t}\}_{\forall j \in N_i}, W_{it}, K_{it}, L_i, A_i) \tag{3.23}$$

$$\mathbf{x}_{i,bt,t} = x(\{p_{k,t}\}_{\forall k}, \mathbf{p}_{x,t}, \sum_{\forall j \in N_i} \alpha_{ij} A_{j,bt,t}, \{A_{j,bt,t}\}_{\forall j \in N_i}, W_{it}, K_{it}, L_i, A_i) \tag{3.24}$$

With knowledge at time t determined by (see 3.13):

$$K_{it} = K\left(K_{it-1}, \{y_{ikt}\}_{\forall k}, \{y_{jkt}\}_{\forall k, \forall j \in N_i}\right) \tag{3.25}$$

Similarly, the binary adoption decision has the following reduced form:

$$V_{bt} - V_{non-bt} = V(\{p_{k,t}\}_{\forall k}, \mathbf{p}_{x,t}, \sum_{\forall j \in N_i} \alpha_{ij} A_{j,bt,t}, \{A_{j,bt,t}\}_{\forall j \in N_i}, W_{it}, K_{it}, L_i, A_i) \tag{3.26}$$

Taking the second-order Taylor expansion around the optimal decision of 3.22, one obtains the following econometric specification for the amount of labor put into the production of Bt cotton, $L_{bt,t}$, with $\gamma_L \sim N(0, \sigma_{\gamma_L})$. Note that the derivation of the econometric specification of $A_{bt,t}$ and $\mathbf{x}_{bt,t}$ are analogous.

$$
\begin{aligned}
L_{i,bt,t} \;=\; & \left\{\frac{\partial L_{bt}}{\partial p_k}.p_{k,t}\right\}_{\forall k} + \left\{\frac{\partial^2 L_{bt}}{\partial p_{kt}^2}.p_{k,t}^2\right\}_{\forall k} + \frac{\partial L_{bt}}{\partial \mathbf{p}_x}.\mathbf{p}_{x,t} + \frac{\partial^2 L_{bt}}{\partial \mathbf{p}_x^2}.\mathbf{p}_{x,t}^2 + \\
& \frac{\partial L_{bt}}{\partial \Delta_i}.\sum_{\forall j\in N_i} \alpha_{ij} A_{j,bt,t} + \frac{\partial^2 L_{bt}}{\partial \Delta_i^2}.\left[\sum_{\forall j\in N_i} \alpha_{ij} A_{j,bt,t}\right]^2 + \left\{\frac{\partial L}{\partial \Delta_{ij}}.A_{j,bt,t}\right\}_{\forall j\in N_i} \\
& + \left\{\frac{\partial^2 L}{\partial \Delta_{ij}^2}.A_{j,bt,t}^2\right\}_{\forall j\in N_i} + \frac{\partial L_{bt}}{\partial W}.W_{it} + \frac{\partial^2 L_{bt}}{\partial W^2}.W_{it}^2 + \frac{\partial L_{bt}}{\partial K}.K_{it} + \\
& \frac{\partial^2 L_{bt}}{\partial K^2}.K_{it}^2 + \frac{\partial L_{bt}}{\partial L}.L_i + \frac{\partial^2 L_{bt}}{\partial L^2}.L_i^2 + \frac{\partial L_{bt}}{\partial A}.A_i + \frac{\partial^2 L_{bt}}{\partial A^2}.A_i^2 + \gamma_{A,it}
\end{aligned}
\tag{3.27}
$$

Following 3.26 one can derive the following probit specification for the binary adoption decision, with γ_{bt} ˜$N(0, \sigma_{\gamma_{bt}})$ and $X = (\{p_{k,t}\}_{\forall k}, \mathbf{p}_{x,t}, \sum_{\forall j\in N_i} \alpha_{ij} A_{j,bt,t}, \{A_{j,bt,t}\}_{\forall j\in N_i}, W_{it}, K_{it}, L_i, A_i)$.

$$
P(bt_{it}=1) = P(V_{bt} - V_{non-bt} \geqslant 0) = P(X'.\theta + \gamma_{bt} \geqslant 0) = \Phi(X'.\theta) \tag{3.28}
$$

Note that 3.27 (and similarly, the econometric specification of $A_{bt,t}$ and $\mathbf{x}_{bt,t}$) and 3.28 assume that there are no individual differences in the specification of the utility function 3.1, the social norms and imitation function 3.3 and the production functions 3.5. These assumptions are not necessarily correct. Especially the common utility function is problematic as it assumes no individual differences with regard to risk preferences. As preferences with regard to risk are known to influence the adoption decision, I include a risk-coefficient (R_{r_i}) in 3.27 (and similarly, the econometric specification of $A_{bt,t}$ and $\mathbf{x}_{bt,t}$) and 3.28. This risk coefficient is calculated from the results of the risk experiment (see Appendix C).

The data provides information on $\{p_{k,t}\}_{\forall k}$, $\mathbf{p}_{x,t}$, W_{it}, L_i and A_i. But what about the terms $\sum_{\forall j\in N_i} \alpha_{ij} A_{j,bt,t}$, $\{A_{j,bt,t}\}_{\forall j\in N_i}$ and K_{it}? Starting with the latter, from 3.25 one can see that knowledge at time t is determined by knowledge at time t-1 and the realizations of yields given the set of inputs of farmer i and his information contacts at time t. Knowledge at time t-1 is determined by knowledge at time t-2 and the realizations of yields given the set of inputs of farmer i and his information contacts at time t-1, etc. Given the finite time horizon of the panel, one can rewrite 3.25, with $t \in \{2,3,4,5,6,7\}$ corresponding to $\{2002-03, 2003-04, 2004-05, 2005-06, 2006-07, 2007-08\}$.

$$K_{it} = K\left(\left\{y_{ik1}, y_{ik2}, y_{ik3}, ..., y_{ikt-1}; \{y_{jk1}, y_{jk2}, ..., y_{jkt-1}\}_{\forall j \in N_i}\right\}_{\forall k}\right) \tag{3.29}$$

To transform 3.29 into an econometric specification, two questions need to be solved. First, which other farmers j does one select? Second, how does one map $(y_{jkt}, A_{kt}, L_{kt}, x_{kt})$ onto K_k? Note that the latter involves the question of aggregation accross experiences of different farmers.

Ideally, following the theoretical model, one would like to include all the information contacts of each farmer. This is exactly what the first social network question does, asking the number of farmers farmer i knew in each year since 2001-02 that adopted Bt cotton. Assuming a linear K-mapping (with a_1 and a_2, respectively, the coefficients on own experience and others' experience), no updating of knowledge regarding the non-Bt crops, one can replace the K_{it} term in 3.27 (and similarly, the econometric specification of $A_{bt,t}$ and $\mathbf{x}_{bt,t}$) and 3.28 with, where I is an indicator function taking the value of one if i (or j) adopt Bt cotton and 0 if i (or j) do not adopt Bt cotton:

$$K_{bt} = a_1. \sum_{\tau=1}^{\tau=t-1} I_{i,bt,\tau} + \sum_{\tau=1}^{\tau=t-1} \sum_{\forall j} I_{j,bt,\tau} \tag{3.30}$$

Admittingly, this linear mapping is a rather crude method of capturing the increase in knowledge from own and others' experience. Ideally, one would like to take into account certain aspects of j's production (input and output) and the relationship between farmer i and his contact j. Due to time constraints asking about these aspects is practically not feasible. [11]

As such, one needs to think about sampling the set of contacts of each farmer. Two of the most popular techniques are respondent-driven snow-ball sampling and taking the "network of a sample". The first technique is useful when one is interested in properties of the network

[11] In addition, farmers often incorrectly report their peer's behavior. This might or might not be a problem depending on what drives the farmer's action: the actual behavior of j or the pereceived behavior of j (see Hogset and Barrett 2008).

itself, but as it clearly results in a non-representative sample of the households, it is not a useful technique for the economic analysis of the effects of social networks on something else (Scott 1991). The second technique, taking the "network of a sample", artificially truncates the network, and clearly is not representative for the “network of the population” and as such will result in biased estimates of the micro-economic behavior (Santos and Barrett 2007). The reason behind this bias is a positive covariance between the behavior of the contacts sampled and the behavior of the contacts that aren’t sampled and as such are part of the error term (see Appendix E). As such, I propose to use a third method, the random matching within sample method. Depending on the structure of underlying network, this method has the potential to provide unbiased estimates of the social learning effects. Note that in both "network of a sample" and "random matching within sample" one needs to control for information coming from outside of the village as this might create spurious correlation. The advantage of using this method is now one can take certain aspects of j’s production and the relationship into account. In case of 3.28 it’s knowledge about the profitability of Bt cotton that will drive the binary adoption decision, or:

$$K_{bt} = a_1. \sum_{\tau=1}^{\tau=t-1} I_{i,bt,\tau} + a_2. \sum_{\tau=1}^{\tau=t-1} \sum_{\forall j \text{ linked to i in } \mathrm{S}_i} I_{j,bt,\tau} + a_3 \sum_{\tau=1}^{\tau=t-1} INF_\tau \tag{3.31}$$

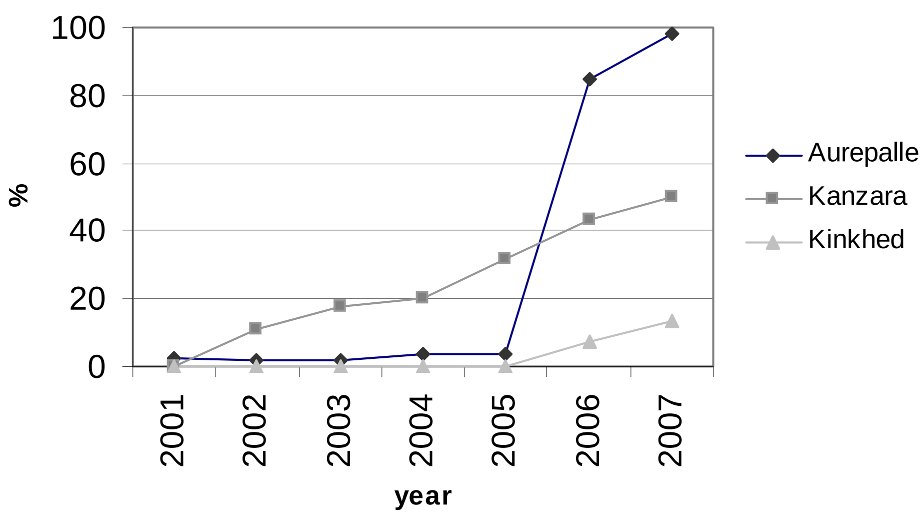
With INF denoting a masure of the information coming from outside of the village. And "j linked to i in S_i" referring to "respondent i draws or is given the card of j in the random matching within sample game, knows j, thinks that j is a cotton farmer and knows whether j cultivates Bt cotton and the yield of j".

With regard to the social norm term, $\sum_{\forall j \in N_i} \alpha_{ij} A_{j,bt,t}$, I give an equal importance to all cotton farmers in the village,i.e., $\alpha_{ij} = \alpha = \frac{1}{\text{number cotton farmers in the village}}$ In the imitation term, $\{A_{j,bt,t}\}_{\forall j \in N_i}$, only the four most "influential" progressive farmers of each village are included as j’s. These are aggregated into one term, each given an equal weight.

4 Results

Figure 4 plots the number of Bt cotton farmers as percentage of the total number of cotton farmers versus time. The adoption process is markedly different in the three villages. While Kanzara displays a smooth adoption process, Aurepalle shows a sudden jump in 2005 and Kinkhed a reluctant take-off in the same year. In Andhra Pradesh, the first Mahyco Bt cotton cultivars were approved by the GEAC in 2004. This might be one of the causes of the delayed adoption in Aurepalle. The current adoption rate in Aurepalle, Kanzara and Kinkhed, respectively, stand at 98%, 50% and 14%.

Tables 4.1 summarizes the properties of the distribution of the average (perceived) yield of Bt and non-Bt cotton cultivars of the yield distribution game. These reflect the current beliefs about Bt versus non-Bt cotton in the three villages. One can see that the average of the average (perceived) yield is higher in Bt cotton versus non-Bt cotton. This difference is however smaller in Kinkhed compared to Kanzara and Aurepalle.

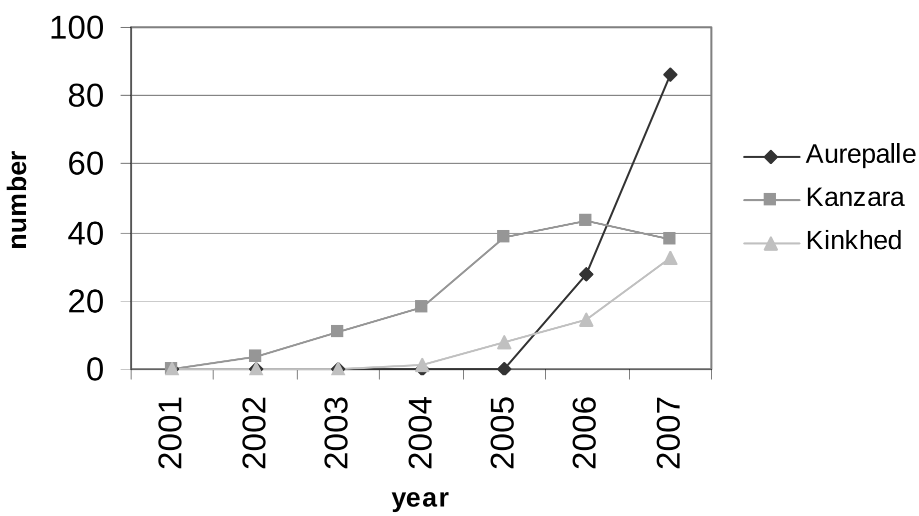


Adoption curve of Bt cotton

Table 4.1: Properties of the distribution of the average (perceived) yield of cotton

	Bt Cotton				Non-Bt Cotton			
	Average	Deviation	Minimum	Maximum	Average	Deviation	Minimum	Maximum
Aurepalle	7.6	2.0	3.7	13.1	4.8	1.1	3.1	8.5
Kanzara	6.2	2.4	1.4	11.9	3.9	1.3	1.3	6.8
Kinkhed	5.5	1.3	3.7	12.1	3.7	0.7	2.1	6.9

Figure 4 shows the average number of farmers known that adopt Bt cotton in each year in the three villages. Disregarding the 2007 data point, Figure 3 resembles Figure 2 in shape for all three villages. If social learning is present, the number of farmers known in the past years should affect the adoption in the current year; if social norms and imitation are present, the number of farmers known in the current year should affect the adoption in the current year.



Average number of farmers known that adopt Bt cotton

Concerns about the health and environmental implications of this new technology might be one of factors driving social norms. In Aurepalle, Kanzara and Kinkhed, respectively, 13%, 51% and 55% of the respondents "agreed" or "strongly agreed" with (at least) one of the following statements: "Bt cotton is hazardous for animal health: they might get sick or die when they eat the fodder", "Bt cotton is hazardous for human health: if you touch it too much, you might get sick" and "Bt cotton is hazardous for the environment: it damages crops and soils"

To conclude the descriptive statistics, Table 4.2 presents some selected results of the random matching within sample experiment. Recall that each respondent draws 6 name cards of VLS respondents and is given a set of 4 fixed cards with name of progressive farmers. Denote the farmer on the card by x. The first column indicates the number of times that the respondent knew x as a percentage of all the cards drawn. One can see that in a small village like Kinkhed, literally everyone knows everyone. The second column gives, conditional on knowing x, the number of times that the respondent thought that x was a farmer, again as a percentage. Similarly, the third column indicates, conditional on x thought to be a farmer, the number of times that the respondent thought that x was a cotton farmer. The fourth column gives, conditional on x thought to be a cotton farmer, the number of times the respondent knew both whether x was cultivating Bt versus non-Bt cotton and the yield obtained by x. Once again, in Kinkhed, farmers not only knew each other by name, but also by cultivar and yield.

Table 4.2: Results of the random matching within sample experiment

	Know x?	Does x farm?	Does x farm cotton?	Is there a "link"?
Aurepalle	88%	68%	57%	33%
Kanzara	99%	81%	70%	36%
Kinkhed	100%	83%	89%	60%

Table 4.3A. presents the results of the probit regressions using. 3.30 to define the learning term and a second order term for both learning and social norms, omitting the imitation term. As the experience term perfectly predicts adoption, i.e., once one adopts Bt cotton, one does not dis-adopt, this variable was omitted. Any interaction term of the learning variable with the average experience of the other farmers is insignificant (results not reported).[12] The first

[12] Note that including an interaction term using an ordinal variable measuring experience of a group of other farmers ("very positive" up to "very negative") poses an aggregation problem across year. I included only last year's experience interacted with five experience dummies.

two columns present the results of the regression one without the year*aurepalle interaction terms. Both the probit coefficient and the marginal effects at the average are presented. The second two columns present the results of regression two including the year*aurepalle interaction terms. Note that having one acre of land (i.e., of cultivable land) and one additional able, adult household member increases the probability of adoption significantly. The price of cotton is insignificant, as are the female and male wages. Somewhat surprisingly the price of the Bt cotton seed is significant only at the 10% level in the first regression (and has the "wrong" sign) and not significant in the second regression. This could be explained by the way this price is measured, i.e., as the official price in each state, and as such might not always reflect the price the farmer is facing. As the VLS has changed the way in which income is measured in 2005, and income is not comparable across years, I opted not to include this variable. [13] The first and second order learning term are significant in both regressions at the 1% level. According to the second regression, knowing ten more people that have adopted Bt cotton in the past increases the probability of adopting with 2%. Note that the sign of the second-order learning term indicates a concave learning effect. The first-order social norm term is significant in both regression, but the second-order term is only significant in the second regression. The coefficients in both regressions point at a concave social norm effect. Note that the size of the social norm effect is substantial, having one more acre/farmer under Bt cotton increases the probability of adoption with 11% in the first regression and 40% in the second regression. None of the year*aurepalle interaction effects included are significant in the second regression. The insignificance of the risk dummy in both regression one and two is probably due to the rather crude measurement of risk aversion.[14]

[13] Before 2005 income was measured through a direct question to the respondent who is asked to recall his/her income one year back. From 2005 onwards, income is measured through a set of modules as the difference between production valued at market prices and expenditures valued at market prices including family labor on a three weekly basis of the different production activities.

[14] A farmer is considered risk averse if his willingness to pay for the first yield distribution is higher than or equal to his willingness to pay for the second yield distribution. Recall that the first distribution second-order stochastically dominates the second distribution.

Table 4.3A: Results of probit regression one and two(without experience interaction terms)

	N=1074		N=917	
	Coefficient	dF/dx	Coefficient	dF/dx
Land	0.0593307***	0.007846	0.0574067***	0.00895
Member	0.1630798***	0.021566	0.1682035***	0.02622
Pcotton	0.01161	0.0015352	0.0199229	0.00311
Pmale	0.01153	0.0015251	-0.0074698	-0.00116
Pfemale	0.02833	0.003747	0.0310347	0.00484
Pbtseed	0.000077*	0.0000102	-0.0001367	-2.10E-05
Learning	0.008454***	0.001118	0.013103***	0.00204
Learning2	-0.0000164***	-2.17E-06	-0.0000228***	-3.60E-06
Norm	0.8638165**	0.1142327	2.606657***	0.40628
Norm2	-0.14415	-0.019062	-1.006277**	-0.15684
Riskdummy	-0.00114	-0.0001509	0.0525934	0.00845
2004*aur			0.4689471	0.09342
2005*aur			2.5716	0.77412
2006*aur			1.602766	0.46405
2007*aur			0.0230077	0.00362
Constant	-5.02087		-4.146105	

Notes: *** significant at 1% level, ** significant at 5% level, * significant at 10% level

Table 4.3 B continues using the same specification for the learning term but including the imitation term. One can see that the learning term remains significant in regression three and four, and it's size is similar. The social norm terms became insignificant through inclusion of the imitation term in both regression three and four. The imitation term and it's square are significant in both regression three and four. It's coefficient is large; one acre/PF more under Bt cotton increases the probability of adoption with, respectively, 7% and 11%. In regression four, two of the year*aurepalle interaction terms become significant. This might point at an incorrect measurement of the price of Bt cotton seed in Aurepalle in those years.

Table 4.3B: Results of probit regression three and four (without experience interaction terms)

	N=1074		**N=917**	
	Coefficient	dF/dx	Coefficient	dF/dx
Land	0.059404***	0.0067928	0.0579811***	0.0087744
Member	0.1573156***	0.0179888	0.1625935***	0.0246056
Pcotton	-0.0306528	-0.0035051	0.0071266	0.0010785
Pmale	0.0584956***	0.0066889	0.0213299	0.0032279
Pfemale	-0.0201931	-0.0023091	-0.0538998	-0.008157
Pbtseed	0.0004956**	0.0000567	0.000074	0.0000112
Learning	0.0085215***	0.0009744	0.0115225***	0.0017437
Learning2	-0.0000162***	-1.85E-06	-0.0000201***	-3.04E-06
Norm	-1.187079	-0.1357407	-0.9748837	-1.48E-01
Norm2	0.5442277*	0.0622316	0.2765595	0.0418523
Imitation	0.6230185***	0.0712413	0.7744032***	0.1171919
Imitationsquare	-0.0468424***	-0.0053564	-0.0474653**	-0.007183
Riskdummy	-0.0696148	-0.0076748	0.0318526	0.0048856
2004*aur			1.216314*	0.3225766
2005*aur			1.378482	0.3782205
2006*aur			3.077141***	0.8705021
2007*aur			3.409177	0.9114619
Constant	-6.823411		-4.349147	

Notes: *** significant at 1% level, ** significant at 5% level, * significant at 10% level

Table 4.4 presents the results of the probit regression, using 3.31 to define the learning term and a second-order for learning, social norms and imitation. Interestingly, the male wage is significant in regression three and positive. The first and second order term of learning is significant both regression five and six. Knowing both cultivar and yield of one additional farmer in the past increases the probability of adoption with respectively 1.2% and 1.6%. Knowing both cultivar and yield of ten more additional farmers would therefore increase the probability of adoption with 12% and 16%, respectively. Note the difference in size compared to the regressions in Tables 4.3 A and 4.3B. The second order norm term is significant in the fifth regression, but has a positive sign. Again, the imitation terms are significant and strong, with a size similar to the ones reported in Table 4.3B. The year*aurepalle interaction dummies are significant for three years.

Table 4.4: Results of probit regression five and six

	N=1074		N=917	
	Coefficient	dF/dx	Coefficient	dF/dx
Land	0.0687622***	0.0073253	0.0721596***	0.0102359
Member	0.1515644***	0.0161462	0.151955***	0.0215549
Pcotton	-0.0238671	-0.0025426	0.0032866	0.0004662
Pmale	0.0699416***	0.0074509	0.0442642	0.0062789
Pfemale	-0.029196	-0.0031103	-0.0629037	-0.008923
Pbtseed	0.0005281*	0.0000563	0.0001811	0.0000257
Learning	0.1136583***	0.0121081	0.1112562***	0.0157818
Learning2	-0.0049745**	-0.0005299	-0.0040113*	-0.000569
Norm	-1.471534	-0.156763	-1.405001	-0.1993
Norm2	0.6583627**	0.0701356	0.5070752	0.071929
Imitation	0.6298538***	0.0670985	0.8692402***	0.1233024
Imitation2	-0.0435762***	-0.0046422	-0.0541322**	-0.007679
Riskdummy	0.0083608	0.0008947	0.1121323	0.0167032
2004*aur			1.437179**	0.3916529
2005*aur			1.623972***	0.4565078
2006*aur			2.321791**	0.693331
2007*aur			2.703037	0.794781
Constant	-7.534237		-5.780368	

Notes: *** significant at 1% level, ** significant at 5% level, * significant at 10% level

5 Conclusion

Preliminary results indicate at strong social learning and imitation effects, independent of the way in which social learning is measured. The social norms effect seem to depend on the inclusion of an imitation term, pointing towards the fact that the social norms might be mainly driven by the adoption behavior of the progressive farmers. In several cases, the aurepalle* year fixed effects are significant. This might be due the incorrect measurement of the price of Bt cotton seed in Andhra Pradesh.

The measurement of several variables can be improved. The prices of Bt cotton seeds should reflect the prices that farmers are facing in the market in each year and a variable income should be included. Similarly, risk aversion should be measured using a continuous variable rather than a dummy variable. In addition, both the α's in the social norm term, the weights of the individual farmers in the social learning terms as well as the identity of the farmers to be included in the imitation term can be further investigated.

The empirical analysis presented so far can be extended towards the input decisions regarding Bt cotton: acreage, labor and pesticides. In addition, the econometric model used could be refined taking into account the "time-effects" of the seemingly irreversible adoption decision.

Finally, the theoretical model presented leads itself to a more structural estimation of the social norms and imitation function (3.3) and production functions.

6 References

Antle, John M. and Prabhu L. Pingali. 1994. "Pesticides, Productivity, Farmer Health: A Philippine Case Study." American Journal of Agricultural Economics, 76:3, pp. 418-30.

Appadurai, Arjun 1989. "Transformations in the Culture of Agriculture," in Contemporary Indian Tradition: Voices on Culture, Nature and the Challenge of Change. Carla M. Borden ed. Washington and London: The Smithsonian Institution Press, pp. 173-186.

Asia-Pacific Consortium of Agricultural Biotechnology. 2006. Bt Cotton in India - A Status Report, New Delhi.

Baird, Sarah. 2003. "Modeling Technology Adoption Decisions: An Analysis of High Yielding Variety Seeds in India During the Green Revolution." Working Paper. Department of Agricultural and Resource Economics, University of California at Berkeley.

Bandiera, Oriana and Imran Rasul. 2006. "Social Networks and Technology Adoption in Notherns Mozambique." The Economic Journal, 116:514, pp. 869-902.

Bantilan, M.C.S., P. Anand Babu, G.V. Anupama, H. Deepthi, and R. Padmaja. 2006. "Dryland Agriculture: Dynamic Challenges and Priorities." Research Bulletin No. 20, GT-IMPI, ICRISAT.

Barrett, Christopher B. 2005a. "Smallholder Identities and Social Networks The Challenge of Improving Productivity and Welfare," in The Social Economic of Poverty: On Identities, Communities, Groups and Networks. Christopher B. Barrett ed. London and New York: Routledge, pp. 214-43.

Barrett, Christopher B. ed. 2005b. The Social Economics of Poverty. On Identities, Communities, Groups and Networks. London and New York: Routledge.

Basu, Kaushik. 2000. Prelude to Political Economy: A Study of the Social and Political Foundations of Economics. Oxford University Press.

Bertrand, Marianne and Sendhil Mullainathan. 2001. "Do People Mean What They Say? Implications for Subjective Survey Data." MIT Economics Working Paper No. 01-04.

Besley, Timothy and Anne Case. 1993. "Modeling Technology Adoption in Developing Countries." The American Economic Review, 83:2, pp. 396-402.

Binswanger, Hans P. 1980. "Attitudes toward Risk: Experimental Measurement in Rural India." American Journal of Agricultural Economics, 62:3, pp. 395-407.

Chamley, Christophe P. 2004. Rational Herds, Economic Models of Social Learning. Cambridge University Press.

Conley, Timothy G and Christopher Udry. 2003. "Learning about a new technology: Pineapple in Ghana." Working Paper, Department of Economics, Yale University.

Crost, Benjamin, Bhavani Shankar, Richard Bennett, and Stephen Morse. 2007. "Bias from Farmer Self-Selection in Genetically Modified Crop Productivity Estimates: Evidence from Indian Data." Journal of Agricultural Economics, 58:1, pp. 24-36.

David, G. Shourie and Y.V.S.T. Sai. 2002. "Bt Cotton: Farmer's Reactions." Economic and Political Weekly, November 16, 2002.

Dev, S. Mahendra and N. Chandrasekhara Rao. 2006. Socio-Economic Impact Assessment of Transgenic Cotton in Andhra Pradesh. Working Paper. Centre for Economic and Social Studies, Hyderabad.

Doherty, V.S. 1982. "A Guide to the Study of Social and Economic Groups and Stratification in ICRISAT's Indian Village Level Studies." GT-IMPI, ICRISAT.

Engle-Warnick, Jim, Javier Escobal, and Sonia Laszlo. 2006. "Risk Preference, Ambiguity Aversion and Technology Choice: Experimental and Survey Evidence from Peru." Presented at NEUCD 2006 at Cornell University.

Feder, Gechon and Sara Savastano. 2006. "The Role of Opinion Leaders in the Diffusion of New Knowledge: The Case of Integrated Pest Management." World Development, 34:7, pp. 1287-300.

Feder, Gechon, Richard Just, and David Zilberman. 1985. "Adoption of Agricultural Innovations in Developing Countries: A Survey." Economic Development and Cultural Change, 32:2, pp. 255-97.

Floyd, C.N., A.H. Harding, D.P. Paddle, D.P Rasali, K.D. Subedi, and P.P. Subedi. 1999. "The Adoption and Associated Impact of Technologies in the Western Hills of Nepal." Agricultural Research and Extension Network. Network Paper No. 90.

Foster, Andrew D. and Mark R. Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." Journal of Political Economy, 103:6, pp. 1176-209.

Griliches, Zvi. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." Mimeo.

Hubbell, Bryan J., Michele C. Marra, and Gerald A. Carlson. 2000. "Estimating the Demand for a New Technology: Bt Cotton and Insecticide Policies." American Journal of Agricultural Economics, 82:1, pp. 118-32.

Hogset, Heidi and Christopher B. Barrett. 2008. "Social Learning, Social Influence and Projection Bias: A caution on inferences based on proxy–reporting of peer behavior." Working Paper Molde University College and Cornell University.

ISAAA, International Service for the Acquisition of Agri-Biotech Applications 2005. "The Story of Bt Cotton in India - Video."

Jackson, Matthew O. and Brian W. Rogers. 2007. "Meeting Strangers and Friends of Friends: How Random Are Social Networks." American Economic Review, 97:3, pp. 890-915.

Just, David R. and Travis J. Lybbert. 2007. "Risk Averters that Love Risk? Average versus Marginal Risk Aversion. ." Working Paper, AEM, Cornell University and ARE, UC Davis.

Krishnan, Pramila and Emanuela Sciubba. 2006. "Links and Architecture in Village Networks." Birkbeck Working Paper in Economics and Finance, University of London.

Leibenstein, H. 1950. "Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand." The Quarterly Journal of Economics, 64:2, pp. 183-207.

Liu, Elaine. 2008. "Time to Change What we Sow: Risk Preferences and Technology Adoption Decisions of Cotton Farmers in China." Job-Market Paper. Department of

Economics. Princeton University

Lybbert, Travis J. 2006. "Indian Farmer's Valuation of Yield Distributions: Will poor farmers value 'pro-poor' seeds?" Food Policy, 31:5, pp. 415-41.

Lybbert, Travis J. and David R. Just. 2007. "Is Risk Aversion Really Correlated with Wealth? How Estimated Probabilities Introduce Spurious Correlation "American Journal of Agricultural Economics, 89:4, pp. 839-1224.

Lybbert, Travis J., Christopher B. Barrett, John G. Mccpeak, and Winnie K. Luseno. 2007. "Bayesian Herders: Updating of Rainfall Beliefs in Response to External Forecasts." World Development 35:3, pp. 480–97.

Moscardi, Edgardo and Alain de Janvry. 1977. "Attitudes toward Risk among Peasants: An Econometric Approach." American Journal of Agricultural Economics, 59:4, pp. 710-16.

Moser, Christine M and Christopher B. Barrett. 2006. "The Complex Dynamics of Smallholder Technology Adoption: The Case of SRI in Madagascar." Agricultural Economics, 35:3, pp. 373–388.

Naik, Gopal, Matin Qaim, Arjunan Subramnian, and David Zilberman. 2005. "Bt Cotton Controversy: Some Paradoxes Explained." Economic and Political Weekly, April 9 2005, pp. 1514-17.

Narayanamoorthy, A and S S Kalamkar. 2006. "Is Bt Cotton Cultivation Economically Viable for India Farmers? An Empirical Analysis." Economic and Political Weekly, June 30, 2006, pp. 2716-24.

Pingali, Prabhu L., Cynthia B. Marquez, and Florencia G. Pelis. 1994. "Pesticides and Philippine Rice Farmers Health: A Medical and Economic Analysis." American Journal of Agricultural Economics, 76:3, pp. 587-92.

Pomp, Marc and Kees Burger. 1995. "Innovation and Imitation: Adoption of Cocoa by Indonesian Smallholders." World Development 23:3, pp. 423-31.

Qaim, Matin and David Zilberman. 2003. "Yield Effects of Genetically Modified Crops in Developing Countries." Science, 299: 7 February 2003, pp. 900-902.

Qaim, Matin, Arjunan Subramanian, Gopal Naik, and David Zilberman. 2006. "Adoption of Bt Cotton and Impact Variability: Insights from India." Review of Agricultural Economics, 28:1, pp. 48-58.

Qaim, Matin. 2003. "Bt Cotton in India: Field Trial Results and Economic Projections." World Development, 31:12, pp. 2115–27.

Rao, K.P.C, 2008. "Documentation of the Second Generation Village Level Studies 2001-2004". GT-IMPI, ICRISAT.

Rao, K.P.C. and Kumara D. Charyulu. 2007. "Changes in Agriculture and Village Economies." Research Bulletin no 21, GT-IMPI, ICRISAT.

Rogers, Everett. 1962. Diffusion of Innovations. New York: Free Press.

Saha, Atanu, C. Richard Shumway, and Hovav Talpaz. 1994. "Joint Estimation of Risk Preference Structure and Technology Using Expo-Power Utility." American Journal of Agricultural Economics, 76:2, pp. 173-84.

Santos, Paulo and Christopher B. Barrett. June 2007. "Understanding the Formation of Social Networks." Working Paper, Department of Applied Economics and Management, Cornell University

Scott, John. 1991. Social Network Analysis: A Handbook. London - Newburry Park - New Delhi: Sage Publications.

Singh, R.P., Hans P. Binswanger, and N.S. Jodha. 1985. "Manual of Instructions for Economic Investigators in ICRISAT's Village Level Studies (Revised)." GT-IMPI, ICRISAT.

Stephens, Emma C. August 2007. "Feedback Relationships between New Technology Use and Information Networks: Evidence from Ghana." Working Paper, Department of Economics, Pitzer College.

Stone, Glenn Davis. 2007. "Agricultural Deskilling and the Spread of Genetically Modified Cotton in Warangal." Current Anthropology 48:1, pp. 67-103.

Sunding, David and David Zilberman. 2001. "The Agricultural Innovation Process: Research and Technology Adoption in a Changing Agricultural Sector," in Handbook of Agricultural Economics. Bruce L Gardner and Gordon C. Rauser eds: Elsevier Science, pp. 207-61.

Vasavi, A. R. 1994. "'Hybrid Times, Hybrid People': Culture and Agriculture in South India." Man, 29:2, pp. 283-300.

Walker, Thomas S. and James G. Ryan. 1990. Village and Household Economies in India's Semi-Arid Tropics. Baltimore and London: John Hopkins University Press.

Appendix A

Equation 6.1 rewrites 3.8 in the recursive formulation:

$$V(W_t, K_t) = \max_{\left\{c_t, d_t, \{L_{kt+1}, A_{kt+1}, x_{kt+1}\}_{\forall k}\right\}} [u(c,s)] + \delta E\left[V(W_{t+1}, K_{t+1})\right] \tag{6.1}$$

Assume differentiability of the value function with respect to the two state variables. The First Order Conditions with respect to $L_{bt,t+1}$, $A_{bt,t+1}$ and $\mathbf{x}_{bt,t+1}$, respectively, are:

$$\begin{aligned}
L\delta.E\left[\frac{\partial V_{t+1}}{\partial W_{t+1}}.\frac{\partial W_{t+1}}{\partial L_{bt,t+1}} + \frac{\partial V_{t+1}}{\partial K_{t+1}}.\frac{\partial K_{t+1}}{\partial L_{bt,t+1}}\right] &= \lambda_L && (6.2)\\
\left[\frac{\partial u_t}{\partial s_t}\frac{\partial s_t}{\partial \Delta_{it}} + \left\{\frac{\partial s_t}{\partial \Delta_{ijt}}\right\}_{\forall_j}\right] + \delta.E\left[\frac{\partial V_{t+1}}{\partial W_{t+1}}.\frac{\partial W_{t+1}}{\partial A_{bt,t+1}} + \frac{\partial V_{t+1}}{\partial K_{t+1}}.\frac{\partial K_{t+1}}{\partial A_{bt,t+1}}\right] &= \lambda_A && (6.3)\\
[-\mathbf{p}_i] + \delta.E\left[\frac{\partial V_{t+1}}{\partial W_{t+1}}.\frac{\partial W_{t+1}}{\partial \mathbf{x}_{bt,t+1}} + \frac{\partial V_{t+1}}{\partial K_{t+1}}.\frac{\partial K_{t+1}}{\partial \mathbf{x}_{bt,t+1}}\right] &= 0 && (6.4)
\end{aligned}$$

With λ_L the Lagrange multiplier on the labor constraint and λ_A the Lagrange multiplier on the land constraint. The Envelope Conditions with respect to the state variables W and K_k, respectively, are:

$$\begin{aligned}
\frac{\partial V_t}{\partial W_t} &= \frac{\partial u_t}{\partial c_t} && (6.5)\\
\frac{\partial V_t}{\partial K_{bt,t}} &= \delta.\frac{\partial u_{t+1}}{\partial c_{t+1}}.\left[p_{bt}.\frac{\partial y_{bt,t+1}}{\partial K_{bt,t+1}}\right] && (6.6)
\end{aligned}$$

Note that:

$$\begin{aligned}
\frac{\partial K_t}{\partial L_{bt,t}} &= \frac{\partial K_t}{\partial y_{bt,t}}.\frac{\partial y_{bt,t}}{\partial L_{bt,t}} \text{ and } \frac{\partial W_t}{\partial L_{bt,t}} = p_{bt}.\frac{\partial y_{bt,t}}{\partial L_{bt,t}} && (6.7)\\
\frac{\partial K}{\partial A_{bt,t}} &= \frac{\partial K}{\partial y_{bt,t}}.\frac{\partial y_{bt,t}}{\partial A_{bt}} \text{ and } \frac{\partial W_t}{\partial A_{bt,t}} = p_{bt}.\frac{\partial y_{bt,t}}{\partial A_{bt,t}} && (6.8)\\
\frac{\partial K}{\partial \mathbf{x}_{bt,t}} &= \frac{\partial K}{\partial y_{bt,t}}.\frac{\partial y_{bt,t}}{\partial \mathbf{x}_{bt,t}} \text{ and } \frac{\partial W_t}{\partial \mathbf{x}_{bt,t}} = p_{bt}.\frac{\partial y_{bt,t}}{\partial \mathbf{x}_{bt,t}} && (6.9)
\end{aligned}$$

Appendix B

The theoretical model follows the following notation. upper-case letters for stocks and lower-case letters for flows. Letters written in bold denote vectors. Four kinds of subscripts are employed, the first subscript denotes the individuals i, the second subscript denotes the other farmer j, the third subscript denotes the crop-cultivar k, and the fourth subscript denotes time t. Note that subscript i is implied unless otherwise noted.

Notation	Description
u	per-period utility function
c	per-period consumption
s	per-period non-material satisfaction
L	labor
A	land
K	knowledge
W	wealth
N_i	set of farmers to whom i is connected
α_{ij}	parameter capturing the importance of individual j for farmer i in determining the social norm
y_k	yield of crop-cultivar k
ϵ_k	stochastic component of the yield function of crop-cultivar k follows $N(0, \sigma_k)$
p_i	vector of prices of variable inputs
p_k	price of crop-cultivar k
d	net-borrowings
r	gross-interest rate

Appendix C

The risk experiment is based on Lybbert and Just (2007). The experiment consists of a series of hypothetical farming seasons. I use Fisher Price building blocks, vertically stacked up, to represent yield distributions of cotton (seedcotton, in quintal per acre). Each block represents 5%. Green blocks represent high yield (8 Q/acre), yellow blocks represent medium yield (6 Q/acre) and red block represent low yield (4 Q/acre). I start with two trial distributions to learn the game and then do the four experimens outlined in Table 1, in that order. For each experiment I ask the farmer how much they would be maximum willing to pay for a bag of cotton seed that gives this particular yield distribution (a bag sufficient that suffices to sow one acre of cotton in monoculture).

Table C1: Risk Experiment, Set-Up

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
4 Q/acre	25	30	30	10
6 Q/acre	50	40	30	55
8 Q/acre	25	30	40	35
Average	6.00	6.00	6.20	6.50
Variance	2.00	2.40	2.76	1.55

The first distribution is the base line distribution. The second distribution has the same average, but a higher variance than the first distribution. The third distribution has a higher average yield than the first one, but also a considerably higher variance. The fourth distribution first-order stochastically dominates the first distribution.

From the results of these experiments, I calculate a risk-aversion parameter as follows. Assume that the farmer solves the following maximization problem 6.10 - **??**, with φ denoting the probabilities, WTP, Willingness-to-Pay, m, initial wealth, p, price of cotton, Q, yield outcome, R, a minimum subsistence level.[15]

$$\max_{WTP} \varphi_1.u\left[m - WTP + p.Q_1\right] + \varphi_2.u\left[m - WTP + p.Q_2\right] + \varphi_3.u\left[m - WTP + p.Q_3\right] \quad (6.10)$$

[15] Alternatively, R can also be interpreted as a credit constraint.

$$s.t.m - WTP + p.Q_3 \geqslant R \tag{6.11}$$

The Kuhn-Tucker First Order Conditions of this problem are, with $\mu \geqslant 0$ the multiplier on the constraint:

$$\varphi_1.u'\left[m - WTP + pQ_1\right] + \varphi_2.u'\left[m - WTP + pQ_2\right] + \varphi_3.u'\left[m - WTP + pQ_3\right] \leq 0 \tag{6.12}$$

$$\mu.\left[m - WTP + pQ_3 - R\right] = 0 \tag{6.13}$$

$$\left[m - WTP + pQ_3 - R\right] \geqslant 0 \tag{6.14}$$

Thus, there are two possible solutions, a corner solution and an interior solution. In the latter case, the solution is characterized by:

$$\varphi_1.u'\left[m - WTP + pQ_1\right] + \varphi_2.u'\left[m - WTP + pQ_2\right] + \varphi_3.u'\left[m - WTP + pQ_3\right] = 0 \tag{6.15}$$

Adopting a constant relative risk aversion utility function, $u = x^{\beta}$, 6.15 can be rewritten as:

$$\varphi_1.\beta.\left[m - WTP + p.Q_1\right]^{\beta-1} + \varphi_2.\beta.\left[m - WTP + p.Q_2\right]^{\beta-1} + \varphi_3.\beta.\left[m - WTP + p.Q_3\right]^{\beta-1} = 0 \tag{6.16}$$

From the experiment WTP, the outputs (Qs) and the probabilities (φs) are known. From the data m and p are known. For each experiment, expression 6.16 can be (numerically) solved for β within the interval (0,1). Averaging the resulting parameters and calculate the relative risk aversion coefficient:

$$R_r \equiv -c.\frac{u''(c)}{u'(c)} = -(\hat{\beta} - 1) \tag{6.17}$$

In the case of a corner solution,

$$\left[m - WTP + pQ_3 - R\right] = 0 \tag{6.18}$$

Plugging in this into 6.12, one obtains:

$$\varphi_1.u'\left[m - WTP + p.Q_1\right] + \varphi_2.u'\left[m - WTP + p.Q_2\right] + \varphi_3.u'\left[R\right] \leq 0 \qquad (6.19)$$

Calculating R from 6.18 and (numerically) solving 6.19 one can obtain a range for β. Note that as m, p, Q_3 and R are all fixed for a certain household, the constraint is either always binding or never binding for that household. As such, one can recognise a corner solution by testing for $WTP_1 = WTP_2 = WTP_3 = WTP_4$.

Table C2 shows the average willingness to pay for the four lotteries in each of the villages. Note the difference between Aurepalle and Kanzara on the one hand and Kinkhed on the other, across all experiments. Recall that the first experiment second-order stochastically dominates the second experiment and any risk-averse farmer should be willing to pay less for the second. The results indicate that this is often not the case.

Table C2: Average Willingness-to-Pay

	Experiment 1	Experiment 2	Experiment 3	Experiment 4
Aurepalle	495	546	648	743
Kanzara	640	644	753	516
Kinkhed	1142	1370	1597	1488

Appendix D

Define HH as the number of households in the village, VLS as the number of households sampled by the VLS in the village, C as the number of contacts (housholds) farmer i has in the village and S the number of cards drawn during the sample-within-sample game. Given HH, VLS, C and S, what is the probability of drawing x contacts of farmer i during the sample-within sample game?

Define event G_x as "x number of contacts of farmer i are drawn during the game" and event B_y as "y number of contacts of farmer i are part of the VLS", then:

$$P(B_y) = \frac{C_C^Y.C_{H-C}^{VLS-y}}{C_H^{VLS}} \tag{6.20}$$

$$P\left(G_x|\, B_y\right) = \frac{C_y^x.C_{VLS-x}^{S-x}}{C_{VLS}^S} \tag{6.21}$$

Using Bayes Theorem and the definition of average, one can caluculate that in Kinkhed, using S=6, one would on average only draw 1.41 contacts of the farmer i if C=40. In Kanzara this number is 0.71. During the game however, out of 6 contacts, the farmer knew often 4 to 5. This implies that the number of contacts (C) might be much higher than 40. In fact, using a linear model, this would imply that, in Kinkhed, the farmer knows 66% of the village. Another option would be is that the VLS sample actually created information links that were not there before the VLS started. However, the answers to the question "after becoming member of the VLS, how did your relationship change with person X?" indicate that the latter is not the case.

Appendix E

Continuing with the notation introduced in Appendix D, assume the behavior of farmer i, denoted b_i, is influences by individual characteristics (x_i) and the aggregate behavior all his contacts ($j \in N_i$). Call 6.22 the "true model", with parameters a_3, a_4 and a_5:

$$b_{it} = a_3 + a_4.x_i + a_5.\sum_{j \in N_i} b_{jt-1} + \gamma_{T,it} \text{ with } \gamma_{T,it} \text{ ˜ } N(0, \sigma_{\gamma_T}) \tag{6.22}$$

Assume that the set of contacts of farmer i all live in the village, or $N_i \subset HH$ The VLS samples around 15% of the households. Rewrite 6.22:

$$b_{it} = a_3 + a_4.x_i + a_5.\left[\sum_{j \in N_i \subset VLS} b_{jt-1} + \sum_{j \in N_i \not\subseteq VLS} b_{jt-1}\right] + \gamma_{T,it} \tag{6.23}$$

Redefining the error term:

$$\gamma'_{T,it} = \left[a_5 \sum_{j \in N_i \not\subseteq VLS} b_{jt-1} + \gamma_{T,it}\right] \tag{6.24}$$

And calculating the covariance between $\gamma'_{T,it}$ and $\sum_{j \in N_i \subset VLS} b_{jt-1}$:

$$Cov\left(\gamma'_{T,it}, \sum_{j \in N_i \subset VLS} b_{jt-1}\right) = a_5.Cov\left(\sum_{j \in N_i \subset VLS} b_{jt-1}, \sum_{j \in N_i \not\subseteq VLS} b_{jt-1}\right) \tag{6.25}$$

If the contacts of farmer i each make their decisions individually and no correlated effects are at work either, covariance **??** is zero. However, if the VLS samples the links of farmer i, but does not sample the links of these links who also happen to be links of farmer i that are included in the sample, this covariance will not be zero. A non-zero, positive, covariance will result in a biased estimate of a_5, overestimating the effect of social learning.

In the random matching within sample experiment one samples a subset of these links, i.e., $j \in (VLS \cap S \cap N_i)$. The error term now includes an additional term. While the chance of not sampling the links of the links goes up, the chance of these being links of farmer i that are included in the sample goes down. The relative importance of each of these mechanism depends on the underlying network structure.

$$b_{it} = a_3 + a_4.x_i + a_5. \sum_{j \in N_i \subset VLS, j \in S} b_{jt-1} + \left[a_5. \sum_{j \in N_i \subset VLS, j \notin S} b_{jt-1} + a_5. \sum_{j \in N_i \not\subseteq VLS} b_{jt-1} + \gamma_{T,it} \right] \tag{6.26}$$