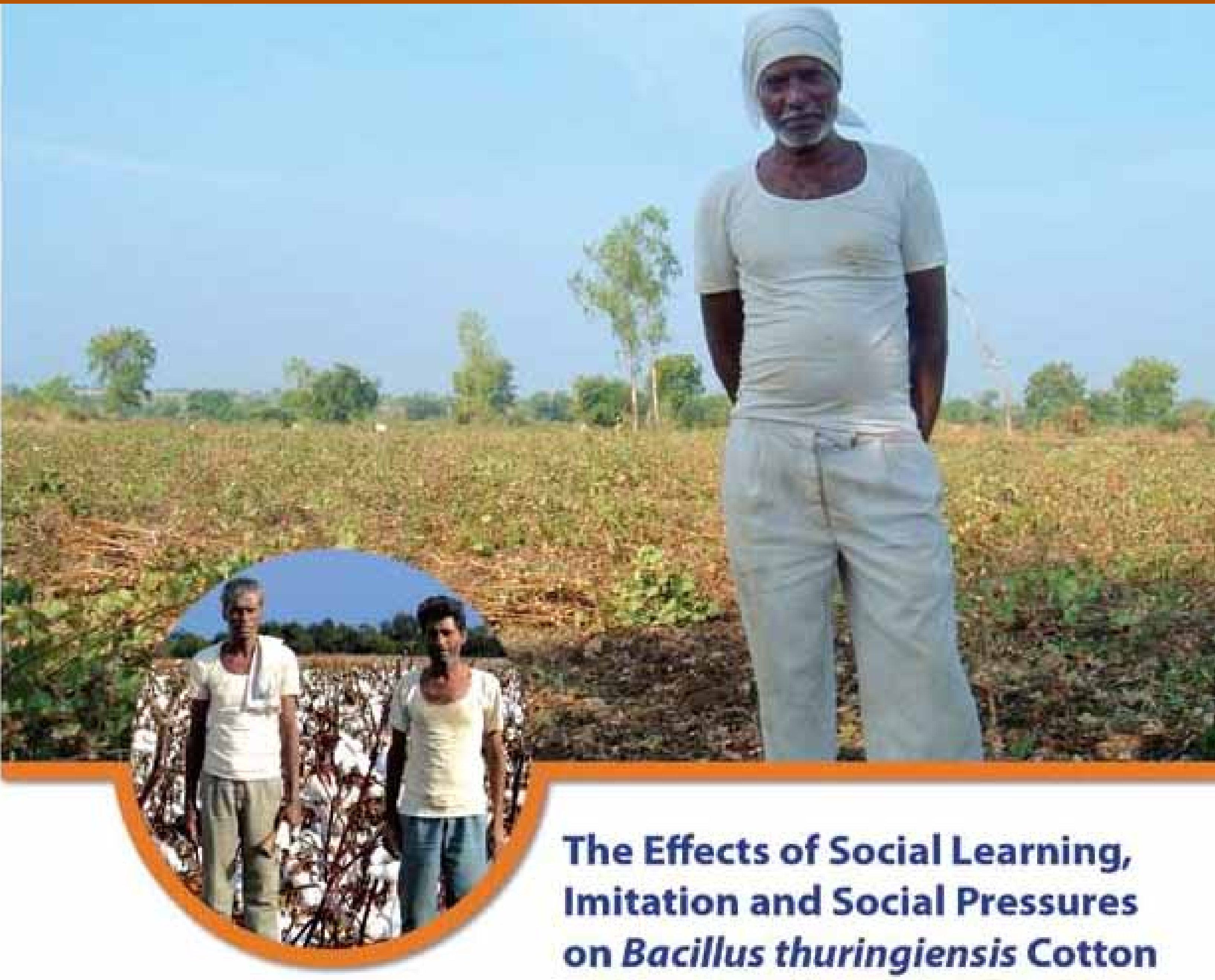

The Effects of Social Learning, Imitation and Social Pressures on *Bacillus thuringiensis* Cotton Technology Adoption Decisions in India

Research Bulletin no. 24

ICRISAT
Science with a human face

International Crops Research Institute for the Semi-Arid Tropics

Citation: Annemie Maertens. 2009.The Effects of Social Learning, Imitation and Social Pressures on *Bacillus thuringiensis* Cotton Technology Adoption Decisions in India. Research Bulletin no. 24. Patancheru 502 324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics. 28 pp. ISBN 978-92-9066-517-5. Order code: RBE 024.

The Effects of Social Learning, Imitation and Social Pressures on *Bacillus thuringiensis* Cotton Technology Adoption Decisions in India

Research Bulletin no. 24

Annemie Maertens

ICRISAT Science with a human face
International Crops Research Institute for the Semi-Arid Tropics
Patancheru 502 324, Andhra Pradesh, India

Acknowledgements

The data for this working paper were gathered in India during August 2007-November 2008 in collaboration with the International Crops Research Institute for 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, an 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 to the American Institute of Indian Studies for their administrative support, the Indian Statistical Institute in New Delhi for their hospitality and the Department of Applied Economics and Management and Chris Barrett for financial support.

Thanks are also due to the several enumerators and research assistants who aided me in my work: Sanjit Anilesh, Shraavya Bhagavatula, Pramod Bangar, Sana Butool, VD Duche, Shital Duche, Madhav Dhere, Anand Dhumale, Shilpa Indrakanti, Navika Harshe, Sapna Kale, Jessica Lebo, Labhesh Lithikar, Nishita Medha, Ramesh Babu Para, Abhijit Patnaik, Gore Parmeshwar, Amidala Sidappa, Nandavaram Ramakrishna, K Ramanareddy, PD Ranganath, and Arjun Waghmode.

I have benefited considerably from discussions with Cynthia Bantilan, Reena Badiani, Chris Barrett, Kaushik Basu, Larry Blume, AV Chari, VK Chopde, William R Coffman, Stephen Coate, Stefan Dercon, Neil Forrester, Ronald Herring, George Jakubson, Matin Qaim, KV Raman, KPC Rao, Mohan Rao, Nageswara Rao, Paulo Santos, Frank Shotkosky, and Kumaracharyulu and the comments of all the seminar participants at GT-IMPI-ICRISAT (December 2008), AEM-Cornell University (throughout 2007-2009) and ISI-New Delhi (December 2008) and one anonymous referee.

Finally, I'd like to mention that this is work in progress and that any remaining errors and omissions are my own.

Contents

The Effects of Social Learning, Imitation and Social Pressures on *Bacillus thuringiensis* Cotton Technology Adoption Decisions in India

Introduction

When new agricultural technologies are introduced, 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 (Besley and Case 1993; Feder et al. 1985; Sunding and Zilberman 2001). While prices, income and individuals' 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 each other, such as social learning, social pressure and imitation.

Farmers learn about prices of inputs and outputs, use of inputs and expected yields given these inputs from own experimentation, the media, input dealers and each others' experimentation, the latter being referred to as social learning (Baird 2003; Bandiera and Rasul 2006; Conley and Udry 2005; Foster and Rosenzweig 1995; Moser and Barrett 2006). In addition, a farmer's choice could be directly influenced by the choices of other farmers through social pressures, ie, deviation from standard agricultural practices entails a non-monetary cost, for instance, social exclusion (Appadurai 1989; Moser and Barrett 2006; Rogers 1964; Vasavi 1994). Finally, farmers can influence one another through imitation (Bandiera and Rasul 2006; Pomp and Burger 1995; Rogers 1964). In this regard, I distinguish behavioral imitation, motivated through a keeping-up-with-the-Joneses attitude, from imitation motivated through a learning story. In the latter, a farmer, despite not observing the outcome of another farmer's experimentation, makes inferences about the profitability of the technique if he observes another farmer using it.

These studies face three main challenges. First, one needs to separate the social interaction effects from the correlated effects, ie, 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 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 social pressures appear to impair the adoption of new technologies, giving information to a few selected farmers per village will most likely not be a successful strategy in tipping the equilibrium. 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.

This paper builds on the existing literature. Using the adoption process of *Bacillus thuringiensis* (Bt) cotton in rural India as an example, I show how one can distinguish between the influences of three different social interaction processes: social learning, imitation and social pressures. 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). In India, in years of severe pest pressure, the net-revenue density function of Bt cotton First Order Stochastically Dominates (FOSD) its non-Bt counterpart (Qaim and Zilberman 2003).

The remainder of this article is structured as follows. The next section gives some background information on Bt cotton technology. This is followed by an introduction to the data collected, outline of the model and the empirical identification strategy. This is followed by a discussion of the results and conclusions.

Background

India is one of the largest producers of cotton in the world. In 2008, it produced 5,443 thousand metric tons compared to 2,985 thousand metric tons in the USA. The average cotton yield in 2008 in India was 579 kg/ha versus 951 kg/ha in the USA and 1,325 kg/ha in China (Table 1). However, yields in India have increased substantially since the early 90s (Figure 1). The main cotton producing states in India are Gujarat, Maharashtra, Punjab and Andhra Pradesh, producing over 80% of the cotton in India, with an average yield of 625 kg/ha and 253 kg/ha, 750 kg/ha and 381 kg/ha ginned cotton in 2006-07[1].

Table 1. Basic cotton statistics of China, USA and India.

Country	Cotton production[1] 1985	1990	1995	2000	2008	Cotton exports[1] 2008	Cotton yields[2] 2008
China	4,137	4,507	4,768	4,420	7,947	16	1,325
USA	2,924	3,376	3,897	3,742	2,985	2,830	951
India	1,964	1,989	2,885	2,380	5,443	1,328	579

1. In thousand metric tons.
2. In kg/ha (ginned) cotton.

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

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% 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 of 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).

As a response to 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 introduced into several of Mahyco's hybrid breeding lines during the 1990s. In 2002, the Genetic Engineering Approval Committee (GEAC), India, approved the commercial

1. Source: INDIASTAT. Selected state-wise estimated yield of cotton in India in 2006-07. Note the all-India average is 421 kg/ha.

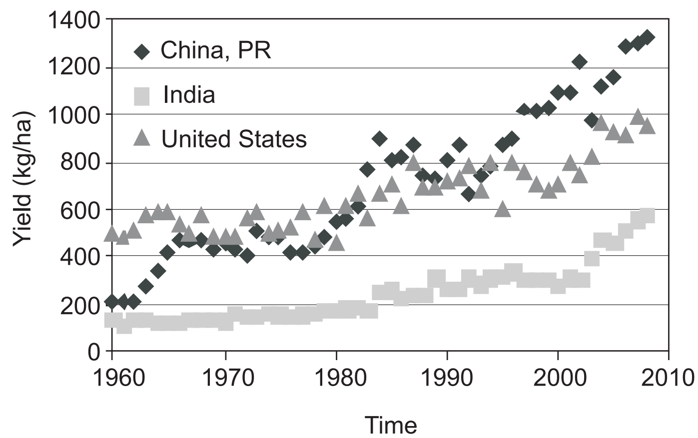

Figure 1. Evolution of cotton yield in China, India and USA.
Source: INDIASTAT.

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. The effect on pesticide use, yield and hence profits from switching from any non-Bt to any Bt cultivar can therefore be decomposed into two effects: the Bt effect and the germplasm effect.

Data from trials on Bt cotton and its isogenic non-Bt counterpart show that in high bollworm pressure years, the Bt net-revenue density function FOSD the non-Bt function[3].

If the bollworm pressure is low (and hence few pesticides need to be used) and the price of the Bt seed is high, this result might not hold. Pemsla et al. (2004), using farm data of non-isogenic Bt and non-Bt cultivars in Karnataka during 2002-03, conclude that the profitability density function of non-Bt FOSD the Bt function.

Data

The three villages selected for this study are part of the Village Level Studies (VLS) program of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT). In this program, ICRISAT followed up 300 randomly selected households from six villages between 1975 and 1985

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.

3. Data from Qaim and Zilberman (2003) from trials done in 2001-02, a high bollworm pressure year. I am grateful to Matim Qaim for sharing the dataset with me.

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[4].

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, to make the sample representative for each village in terms of land-holding size[5].

Between August 2007 and November 2008, I collected an additional round of data in three out of six villages: Aurepalle in Mahbubnagar district in Andhra Pradesh and Kanzara and Kinkheda in Akola district in Maharashtra. I covered 246 VLS farmers, 20 progressive farmers[6] and 3 village leaders using a combination of experiments and questionnaires, including quantitative and qualitative, objective and subjective, and closed and open-ended 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[7] and the District Collector's office.

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[8]. 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 distribution 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, eliciting the farmer's maximum 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 Lybbert et al. (2007), asks the farmer to construct four yield density functions of two Bt cultivars and two non-Bt cultivars of their choice conditional on the plot characteristics of their own plots[9]. The plot-level questionnaire includes questions on agricultural inputs used per plot and outputs produced, including prices.

Social networks were measured in three different ways. The first was 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 links of the farmer, but, as we know little about these contacts, it provides little information on the nature of these links.

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

5. 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).

6. I identified these "progressive farmers" in collaboration with village leaders and VLS investigators who have been living in the village since 2001.

7. Mandal refers to the third-level administrative area in Andhra Pradesh, below state and district. The equivalent in other states is tehsil (or taluka).

8. This section overlaps with the data collected in the 2001-07 ICRISAT VLS.

9. More specifically, we first elicited the minimum and maximum and then used 20 stones to form a conditional yield distribution. During the interview, we explicitly linked the number of stones to percentages. See also Delavande et al. (2008) for the various methods that can be used.

The second method was a random matching within the sample experiment based on Conley and Udry (2005) and Santos and Barrett (2007). Through this method, I elicited the characteristics of the relationship between two randomly drawn respondents, both members of the VLS sample. Each VLS respondent was matched up with six randomly drawn VLS respondents and four fixed progressive farmers. The questions included: 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 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 might incorrectly represent the population of information contacts of the respondent.

Third, the respondents were asked who they 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, a section on the information obtained since 2001-02 from contacts with extension agents, NGOs, input dealers and ICRISAT was included.

The questionnaire included one direct question regarding social pressures ("when you adopted Bt cotton, did you experience any resistance from fellow farmers, relatives or others?"). One plausible explanation underlying social pressures with regard to Bt cotton is the (incorrectly) perceived negative externalities of the technique on other farmers' crops and livestock. As such, 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 were collected.

These data were added to six rounds of the ICRISAT-VLS data (2001-07) containing household composition, landholding, per-plot agricultural input 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. A word on sample attrition (Appendix A). Of the 199 households covered in 2001-02 by the VLS, 92% were still in sample in 2007-08. It is these 184 households plus some of their split-offs and newly added households, a total of 246 households that I interviewed. Of these, 71% had data for all seven cropping years. Of the 29% of the households who are included in the sample from a later date, 30% were households that had split-off from the sample households during 2001-08. These split-off households are included from their date of split-off.

Table 2 gives selected descriptive statistics of the three villages. Aurepalle, with 925 households, 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 are prone to erosion. Kanzara and Kinkheda, with 319 and 189 households respectively, are located in the somewhat less drought-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 yields if properly managed. The average education level of the respondent (ie, the main decision-maker with regard to agriculture) is low, especially in Aurepalle (2.31 years). The average size of a household is between 4 and 5 members in all three villages.

Table 3 summarizes the main cotton-related statistics. In Aurepalle, Kanzara and Kinkheda, 60%, 86% and 82% respectively of the sampled households have farmed cotton in (any of) the last seven years. Of these, 76%, 48% and 11% respectively have adopted Bt cotton at any point in time in the last seven

Table 2. Descriptive statistics of Aurepalle, Kanzara and Kinkheda villages.

	Aurepalle	Kanzara	Kinkheda
Number of households in village	925	319	189
Number of households in sample	128	63	55
Soils[1]	Alfisols	Vertisols	Vertisols
Average rainfall (mm/year)[2]	542	1140	1052
Distance to nearest town (km)[3]	10; 12	9	12
Phone service in the village (year of establishment)[3]	1978; 2003	2002	1993
Average education level of respondent (in years)	2.31	6.61	6.89
Average number of household members	4.23	4.87	4.5
Average yearly income (Rs)[4]	43,543	53,720	38,087

1. Source: Walker and Ryan (1990) using the USDA soil taxonomy system.
2. In 2005-07.
3. The two figures for Aurepalle refer to the main village of Aurepalle and the sub-village of Nallavaripalli.
4. In 2004-05.

Table 3. Basic cotton related descriptive statistics.

	Aurepalle	Kanzara	Kinkheda
Households that farm cotton (%)[1]	60	86	82
Cotton farmers that have adopted Bt cotton (%)[1]	76	48	11
Cotton farmers that have never heard of Bt cotton (%)	4	0	0
Average perceived mean of Bt cotton distribution (Q/acre)[2]	7.60	6.82	5.45
Average perceived mean of non-Bt cotton distribution (Q/acre)[2]	4.79	3.88	3.72
Average number of farmers known to cultivate Bt cotton[3]	85	38	34
Respondents with safety concerns regarding Bt cotton (%)	13	51	53

1. In 2001-08.
2. In 2007-08, 1 quintal = 100 kg.
3. In 2007-08.

years. Dis-adoption rates are low. In Kinkheda, there was no dis-adoption. In Aurepalle, only 2% of the cotton farmers dis-adopted in 2005-06. In Kanzara, between 2% to 7% of the cotton farmers dis-adopted in any given year. The farmers who dis-adopted, except for three of them who stopped farming cotton altogether, re-adopted Bt cotton the next year.

Figure 2 plots the number of farmers cultivating Bt cotton as a percentage of the total number of cotton farmers (the "adoption" curves). The adoption process was markedly different in the three villages. While Kanzara displayed a smooth adoption process, Aurepalle showed a sudden jump in 2005 and Kinkheda a reluctant take-off in the same year. In Andhra Pradesh, two out of three Mahyco cultivars were not approved by GEAC during 2002-03 and 2003-04[10]. This might be one of the causes of the delayed adoption in Aurepalle. The current adoption rate in Aurepalle, Kanzara and Kinkheda stand at 98%, 50% and 14% respectively.

Turning to current knowledge about Bt cotton in Table 3 in Aurepalle, Kanzara and Kinkheda, 23%, 5% and 0% respectively of the respondents had never heard of Bt cotton at the time of the interview in 2008. This figure is much lower among cotton farmers, 4%, 0% and 0%, respectively. Assuming a step-wise distribution with the minimum and maximum of the distribution as specified by the

10. See http//www.envfor.nic.in/divisions/csurv/geac/biosafety.html

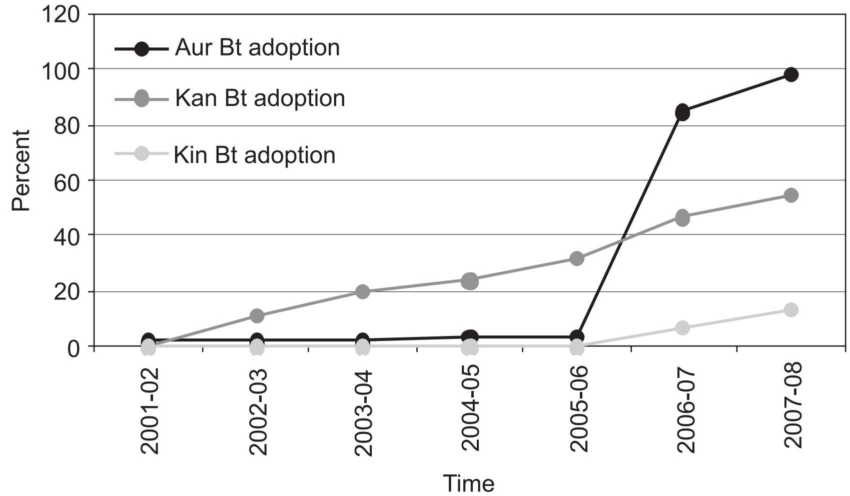

Figure 2. Number of farmers adopting Bt cotton.

respondent, I calculated the mean of the perceived conditional yield distribution of Bt cotton and non-Bt cotton for each respondent. These reflect the current beliefs about Bt cotton and non-Bt cotton for each respondent. One can see that, on average, the mean yield is perceived to be higher for Bt cotton versus non-Bt cotton. This difference is smaller in Kinkheda compared to Kanzara and Aurepalle. In 2007-08, the respondents in Aurepalle, Kanzara and Kinkheda, on average, knew 85, 38 and 34 farmers respectively who cultivated Bt cotton in 2007-08[11].

Concerns about health and environmental implications of this new technology might be one of the factors driving social pressures. In Aurepalle, Kanzara and Kinkheda, 13%, 51% and 53% respectively 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 it", "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".

Regarding the structure of information networks, Table 4 presents some selected statistics. In Aurepalle, Kanzara and Kinkheda, on average, 78%, 85% and 57% respectively of the Bt farmers known to the respondent lived in the village. The next five rows present the results from the random matching within sample game. Recall that each respondent draws 6 name cards of VLS respondents and is given a set of 4 fixed cards with names of progressive farmers. Denote the individual on the card by x. The next 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 Kinkheda, literally everyone knows everyone else. The next column gives, conditional on knowing x, the number of times that the respondent thought (for sure) that x was a farmer, again as a percentage. Similarly, the following column indicates, conditional on x thought to be a farmer, the number of times that the respondent thought (for sure) that x was a cotton farmer. The last 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

11. In case of a range given by the respondent, the average was taken. If the respondent answered "don't know", it was assumed they knew no Bt cotton farmers. Note that a "don't know" can be interpreted in two ways: (1) I do not remember how many farmers I knew that year that adopted Bt cotton, or (2) in that year, I did not know how many farmers adopted Bt cotton. From my conversations on the field, it appears that the second interpretation was far more common.

Table 4. Structure of information networks in Aurepalle, Kanzara and Kinkheda villages.			
	Aurepalle	Kanzara	Kinkheda
Bt farmers known to live in the village (%)	78	85	57
Known *x*?	88	99	100
Does *x* farm?	82	84	92
Does *x* farm cotton?	57	70	90
Is there a "link"?	29	42	63

non-Bt cotton and the yield obtained by *x*. Once again, in Kinkheda, farmers not only knew each other by name, but also by cultivar and yield.

Model

This section outlines a simple 2-period agricultural household model. The notations used in this model are outlined in Appendix B. Consider a representative farmer, denoted with subscript *i*. Subscript i is implied unless otherwise noted. Time is indexed by subscripts $t \in \{1,2\}$. Subscripts b and *n*, respectively, refer to Bt cotton and non-Bt cotton. The farmer possesses land, denoted *A*, assumed to be fixed over time. The farmer's per-period well-being *u* is a function of the farmer's consumption, denoted by *c*, and a non-material satisfaction term, denoted by *s*. I assume that this utility function takes the following shape:

$$u_t \equiv \ln c_t - s_t \tag{1}$$

The satisfaction term is included to capture social pressures and (behavioral) imitation effects and is defined as in 2 and 3 with $A_{b,t}$, A_t, $A_{PF,t}$ denoting, respectively, farmer i's acreage under Bt cotton, the village average acreage under Bt cotton and a progressive farmer's, denoted *PF*, acreage under Bt cotton at time *t*. I denotes an indicator function taking the value of 1 if $A_{b,t}>0$ and zero otherwise.

$$s_{i1} \equiv \left[\left(A_{b,1} - \overline{A}_1\right) I_{A_{b,1}>0}\right]^2 + \left[A_{b,1} - A_{PF,1}\right]^2 \tag{2}$$

$$s_{i2} \equiv \left[\left(A_{b,2} - \overline{A}_2\right) I_{A_{b,2}>0}\right]^2 + \left[A_{b,2} - A_{PF,2}\right]^2 + \theta \left[\left(A_{b,1} - \overline{A}_1\right) I_{A_{b,1}>0}\right]^2 \tag{3}$$

The first terms in 2 and 3 represent social pressure and the second term captures (behavioral) imitation effects. Note that social pressures can have lingering effects in the next period, as captured in the third term by θ.

Turning to the production side of the model, recall that the effect of pesticide depends on bollworm pressure. Denote bollworm pressure by $q \in \{q_L, q_H\}$, referring to high and low pressure respectively. The perceived probability of a high bollworm pressure year is Q and the farmer makes his decision regarding pesticide use in the middle of each period, ie, after *q* is revealed. Denote prices of cotton, pesticides and seed, respectively, by *p*, p_x and p_s pesticides by *x* and unexpected shocks caused by weather fluctuations as ε, an *i.i.d* mean *1* random variable. Now, farm profits are defined as follows:[12]

12. For reasons of tractability I opt for a CRS Cobb-Douglas function.

$$\pi\left(\alpha(q_t),\beta(q_t),\varepsilon_t\right) \equiv p\left[\alpha(q_t)A_{b,t}^{\beta_b(q_t)}x_{b,t}^{1-\beta_b(q_t)}\varepsilon_{b,t} + (A-A_{b,t})^{\beta_n(q_t)}x_{n,t}^{1-\beta_n(q_t)}\varepsilon_{n,t}\right] \\ - p_x(x_{b,t}+x_{n,t}) - p_{s,b}A_{b,t} - p_{s,n}(A-A_{b,t}) \quad (4)$$

The object of learning in 4 is $\alpha(q)$ and $\beta_b(q)\subset[0,1]$, the parameters of the Bt cotton production function. Denote the prior beliefs on these as $F_0(\propto(q),\ \beta_b(q))$.

Note that $\beta_n(q)\subset[0,1]$ is assumed to be known by the farmer. Beliefs are updated at the beginning at each period, before decisions regarding acreage are made, using the following rule, where N_i denotes all the members of farmer i's network:

$$F_t(\alpha(q),\beta_b(q)) = L\left[F_{t-1}\left(\alpha(\cdot),\beta_b(\cdot)\right),\right)\sum_{\forall j\in N_i} I_{A_{j,b,t>0}}, \sum I_{A_{i,t-1}}, \sum_{\forall j\in N_i} I_{A_{j,b,t-1>0}}\right] \quad (5)$$

Recognizing the abuse of notation in 5 with regard to the indicator function I, the second term refers to indirect learning by observing the current behavior of one's contacts, the third to the information gathered through own experimentation and the last to information gathered from the farmer's contacts' experimentation.

At the beginning of period 1, the farmer will decide on $A_{b,1}$ taking into account the expected use of pesticides and the effect of the choice of $A_{b,1}$ on future beliefs. pressurealizing the price of the consumption good to 1 and excluding all other forms of income, ie, $c_t \le \pi_t(\cdot)$ the farmer solves the following decision problem at the beginning of period 1, where δ denotes the discount factor:

$$\begin{aligned} &\max_{A_{b,1}} E_1\left[\ln\pi_1(\cdot) - s_1(\cdot) + \delta\ \ln\left[\pi_2(\cdot) - s_2(\cdot)\right]\right] \\ &s.t.\ 0 \le A_{b,1} \le A \\ &s.t.\ x_{b,1} = x_{b,1}(q, A_{b,1}) \\ &s.t.\ x_{n,1} = x_{n,1}(q, A_{b,1}) \\ &s.t.\ A_{b,2}, x_{b,2}, x_{n,2}\ \textit{optimally chosen} \end{aligned} \quad (6)$$

where E_1 is the expectation with regard to Q, $\alpha(q)$, $\beta_b(q)$, ε, $\overline{A}_1$, $A_{PF,1}$ and $\{A_{j,1}\}_{\forall j\in Ni}$. I abstract from the belief formation process with regard to A_1, $A_{PF,1}$ and $\{A_{j,1}\}_{\forall j\in Ni}$ and do not impose correct expectations in this regard. Denote the optimal solution of 6 with $A_{b,1}$*. Note that conditional on $0<A_{b,1}<A$ and $x_{b,1}>0$, a change in $A_{b,1}$ or $x_{b,1}$, everything else equal, will not result in a change in the expected change of beliefs. The demand for pesticides, $x_{b,1}(q,A_{b,1})>0$ (and similarly $x_{n,1}(q,A_{b,1})>0$) is derived from 6,with $Y_{b,1}$ the production of Bt cotton:[13]

$$E_1\left[\frac{1}{\pi_1(\cdot)}\left[\frac{p(1-\beta_b(q))Y_{b,1}(\cdot)}{x_{b,1}} - p_x\right]\right] = 0 \quad (7)$$

Note that if the various random variables are independent, one can derive the demand for pesticides conditional on bollworm pressure and acreage under Bt cotton:

13. Note that if $0<A_{b,1}{}^{**}<A$ then $x_{b,1}{}^{**}>0$ and $x_{n,1}{}^{**}>0$.

$$\pi_{b,1}(q_1, A_{b,1}) = A_{b,1}\left[\frac{(1-E_1\beta_b(q_1)).E_1\alpha(q_1)}{p_x}\right]^{\frac{1}{E_1\beta_b(q_1)}} \quad (8)$$

The First Order Condition of 6 defines the candidate (interior) optimal solution $A_{b,1}**$:

$$E_1\left[\frac{1}{\pi_1 **}\left[\begin{bmatrix}\frac{p\beta_b Y_{b,1}}{A_{b,1}} + \frac{p(1-\beta_b)Y_{b,1}**}{x_{b,1}}\frac{\partial x_{b,1}}{\partial A_{b,1}} - \frac{p\beta_n Y_{n,1}**}{(A-A_{b,1}**)} \\ \frac{p(1-\beta_n)Y_{n,1}**}{x_{n,1}}\frac{\partial x_{n,1}}{\partial A_{b,1}} - p_{s,b} + p_{s,n}\end{bmatrix} \\ -2(1+\delta\theta)(A_{b,1}**-\bar{A}_1) - 2(A_{b,1}**-A_{PF,1})\right]\right] = 0 \quad (9)$$

where $x_{b,1}=x_{b,1}(q,A_{b,1})$ and $x_{n,1}=x_{n,1}(q,A_{b,1})$. Denote $x_{b,1}**=x_{b,1}(q,A_{b,1}**)$ and $x_{n,1}** = x_{n,1}(q,A_{b,1}**)$. The derivative of $A_{b,1}**$ with respect to Q, the perceived probability of a high bollworm pressure year, is:

$$\frac{dA_{b,1}**}{dQ} = \frac{-1}{SOC}E_1\left[\begin{bmatrix}\frac{p\beta_b Y_{b,1}**}{A_{b,1}**} + \frac{p(1-\beta_b)Y_{b,1}**}{x_{b,1}}\frac{\partial x_{b,1}}{\partial A_{b,1}} \\ -\frac{p\beta_n Y_{n,1}**}{(A-A_{b,1}**)} + \frac{p(1-\beta_n)Y_{n,1}**}{x_{n,1}}\frac{\partial x_{n,1}}{\partial A_{b,1}}\end{bmatrix}_{q=q_H} \\ \begin{bmatrix}\frac{p\beta_b Y_{b,1}**}{A_{b,1}**} + \frac{p(1-\beta_b)Y_{b,1}**}{x_{b,1}}\frac{\partial x_{b,1}}{\partial A_{b,1}} \\ -\frac{p\beta_n Y_{n,1}**}{(A-A_{b,1})} + \frac{p(1-\beta_n)Y_{n,1}**}{x_{n,1}}\frac{\partial x_{n,1}}{\partial A_{b,1}}\end{bmatrix}_{q=q_L}\right] \quad (10)$$

If Bt cotton in a low bollworm pressure year has little effect but Bt cotton increases the marginal productivity of pesticides in a high bollworm pressure year (and the latter effect is not offset by a lower marginal productivity of land of Bt cotton), the effect of an increase in Q is positive. The derivative with respect to $E_1\bar{A}_1$ and $E_1A_{PF,1}$ are respectively:

$$\frac{dA_{b,1}**}{dE_1\bar{A}_1} = \frac{-1}{SOC}\left[2(1+\delta\theta)\right] > 0 \quad (11)$$

$$\frac{dA_{b,1}**}{dE_1A_{PF,1}} = \frac{-1}{SOC}2 > 0 \quad (12)$$

Equations 11 and 12, respectively, show how changes in social pressures and (behavioral) imitation effect optimal land allocation. An increase in the village average acreage under Bt cotton will increase the optimal acreage under Bt cotton. Similarly, an increase in the progressive farmer's acreage under Bt cotton will increase the optimal acreage under Bt cotton.

The derivative of $A_{b,1}$** with respect to $E_1\alpha(q)$ and $E_1\beta_{b,1}(q)$ are (assuming independence among the various random variables):

$$\frac{dA_{b,1}**}{dE_1\alpha(q)} = \frac{-1}{SOC}\,p\left[\frac{E_1\beta_b E_1 Y_{b,1}**^2}{A_{b,1}**\;E_1\alpha(q)} + \frac{(1-E_1\beta_b)E_1 Y_{b,1}**^2}{x_{b,1}**\;E_1\alpha(q)}\,\frac{\partial x_{b,1}}{\partial A_{b,1}}\right] \tag{13}$$

$$\frac{dA_{b,1}**}{dE_1\beta_b(q)} = \frac{-1}{SOC}\,p\left[\frac{E_1\beta_b E_1 Y_{b,1}**^2\left[\ln A_{b,1}** - \ln x_{b,1}**\right]}{A_{b,1}**} + \frac{(1-E_1\beta_b)E_1 Y_{b,1}**^2\left[\ln A_{b,1}** - \ln x_{b,1}**\right]}{x_{b,1}**}\,\frac{\partial x_{b,1}}{\partial A_{b,1}}\right] \tag{14}$$

Note that 13 is strictly positive, ie, if the expected value of the technology parameter of the Bt cotton production function increases, the acreage under Bt cotton increases as well. The sign of equation 14 is ambiguous and will depend on the sign of [ln$A_{b,1}$**–ln$x_{b,1}$**]. Note that a change in $A_{j,1}$ can result in a change in beliefs at the time t regarding the parameters of the Bt cotton production function following 5.

To conclude this section, note that at $A_{b,1}$=0, the objective function of 6 is discontinuous. At that point the own experimentation term in 5 is zero and the social pressure term in $s_1(\cdot)$ drops out, boosting utility upwards. Thus, in order to determine the optimal choice of $A_{b,1}$, denoted $A_{b,1}$*, one needs to compare the expected discounted utility at $A_{b,1}$** as defined by 9 and the expected discounted utility in case of non-adoption as defined by:

$$\begin{aligned}
&E_1\left[\ln(A^{\beta_n(q)} x_{n,1}^{1-\beta_n(q)}\varepsilon_{n,1} - p_x x_{n,1} - p_{s,n}A) - A_{PF,1}{}^2 + \delta\left[\ln\pi_2(\cdot) - s_2(\cdot)\right]\right]\\
&s.t.\; x_{n,1} = x_{n,1}(q,A)\\
&s.t.\; A_{b,2}, x_{b,2}, x_{n,2}\; optimally\, chosen
\end{aligned} \tag{15}$$

Note that as one shifts from $A_{b,1}$=0 to a small, but positive $A_{b,1}$, one affects 5. For simplicity's sake, imagine that the farmer updates his beliefs in a (Gaussian) Bayesian fashion (Chamley 2004). In this case, even though one cannot predict the change in expected value of $\alpha(q)$ and $\beta_b(q)$, one can show that, no matter the outcome of the experiment, the variance on $\alpha(q)$ and $\beta_b(q)$ will go down. Any risk-averse decision-maker should therefore ex-ante attribute a higher expected utility to a decision problem in which one has access to the information of one experiment more, everything else equal. But this also implies that this positive effect of experimentation can be dampened by expected information coming from the experimentation of other farmers. As such, one can expect some degree of free-riding on others' experimentation.

Empirical strategy

Following the model outlined above, the reduced form function for $A_{b,1}$** is:

$$A_{b,1} = A\left[A, p, p_x, p_{s,b}, p_{s,n}, E_1\overline{A}_1, E_1 A_{PF,1}, \delta, \theta, F_1(\alpha(q), \beta_b(q)), \beta_n(q), F_\varepsilon\right] \tag{16}$$

The data concerning the 2008-09 cropping year provide planned acreage under Bt and non-Bt cotton ($A_{b,1}$,A), information on time preferences (δ), current beliefs regarding the performance of Bt and non-Bt cotton ($F_1(\alpha(q),\beta_b(q))$, $\beta_n(q)$ unconditional on q and ε but conditional on land quality and the number of insecticide sprays thought to be needed) and expected prices of Bt cotton and non-Bt cotton seeds ($p_{s,b}$ and $p_{s,n}$). I assume that the social norm effect (θ) and the prices that farmers face in terms of cotton and pesticides are well-known village-level variables. Note that the theoretical model assumes no individual differences in the specification of the utility function 1 and the satisfaction function 2. In addition, the theory model ignored differences in soil quality, wealth and non-cotton related income sources. In practice, it will be useful to control for these variables as they might be driving spurious correlation. In the first specification 17, I assume that all individuals have the same beliefs regarding $E_1\overline{A}_1$ and $E_1A_{PF,1}$. Note that in this case, one is unable to distinguish the social pressure, (behavioral) imitation and village-level free-riding effects from the other village-level effects.

$$A_{plan,i,j} = \gamma_1 M_{b,i} + \gamma_2.V_{b,i} + \gamma_3.M_{n,i} + \gamma_4.V_{n,i} + \gamma_5.SP_{n,i} + \gamma_6.X_{ij} + \gamma_7 VILL_{ij} + \mu_{ij}$$
$$VILL_{ij} \rightarrow p_x, p, E_1\overline{A}_1, E_1A_{PF,1}, \theta \qquad (17)$$

where *M* denotes mean, *V* denotes variance, *SP* denotes number of (insecticide) sprays needed per season, *VILL* denotes a village-specific fixed effect and X_{ij} denotes the individual level controls, ie, land, Bt/non-Bt cotton seed prices, land, (self-perceived) time preferences and (self-perceived) risk aversion. In the second specification 18, I control for social pressures using the perceived safety concerns regarding Bt cotton of other village farmers. The variable “SAFE” is constructed as the average of the answers to: “To what degree do you think that other village farmers think that Bt cotton is hazardous for (1) animal health, (2) for human health, (3) for the environment?”.

$$A_{plan,i,j} = \gamma_1 M_{b,i} + \gamma_2.V_{b,i} + \gamma_3.M_{n,i} + \gamma_4.V_{n,i} + \gamma_5.SP_{n,i} + \gamma_6.X_{ij}$$
$$+ \gamma_7 VILL_{ij} + \gamma_8 SAFE_{ij} + \mu_{ij} \qquad (18)$$
$$VILL_{ij} \rightarrow p_x, p, E_1A_{PF,1}, \theta$$

The third specification 19 adds the variable “NUMBER”, ie, the number of villagers that the farmers believe will plant Bt cotton in June 2008. This variable, arguably, captures the combined social pressure from non-safety origins, (behavioral) imitation and free rider effects. Note that this specification controls for social pressures originating from safety concerns.

$$A_{plan,i,j} = \gamma_1 M_{b,i} + \gamma_2.V_{b,i} + \gamma_3.M_{n,i} + \gamma_4.V_{n,i} + \gamma_5.SP_{n,i} + \gamma_6.X_{ij}$$
$$+ \gamma_7 VILL_{ij} + \gamma_8 SAFE_{ij} + \gamma_9 NUMBER_{PF} + \mu_{ij} \qquad (19)$$
$$VILL_{ij} \rightarrow p_x, p, \theta$$

The fourth and final specification 20 uses data from the random matching within sample game and adds the percentage of progressive farmers that cultivate Bt cotton as a percentage of all progressive farmers as perceived by the respondent, denoted by “*PF*”. This variable captures (behavioral) imitation and free rider effects.

$$A_{plan,i,j} = \gamma_1 M_{b,i} + \gamma_2 . V_{b,i} + \gamma_3 . M_{n,i} + \gamma_4 . V_{n,i} + \gamma_5 . SP_{n,i} + \gamma_6 . X_{ij}$$
$$+ \gamma_7 VILL_{ij} + \gamma_8 SAFE_{ij} + \gamma_{10} E_{ij} A_{PF} + \mu_{ij} \qquad (20)$$

$$VILL_{ij} \rightarrow p_x, p, \theta$$

The quasi-panel data covering 2001-02 up to 2007-08 do not provide information on beliefs. Adapting 5 to a panel data context:

$$F_t(\alpha(q), \beta_b(q)) = L\left[F_0(\alpha(\cdot), \beta_b(\cdot)), \sum_{\forall j \in N_i} I_{A_{j,b,t}>0}, \sum_{\tau=1}^{t-1} \sum I_{A_{i,b,\tau}>0}, \sum_{\tau=1}^{t-1} \sum_{\forall j \in N_i} I_{A_{j,b,\tau}>0}\right] \qquad (21)$$

Assuming that all respondents start with the same prior beliefs, one can study beliefs formation through the following specification:

$$F_t(\alpha(q), \beta_b(q)) = \gamma_9 \sum_{\forall j \in N_i} I_{A_{j,b,t}>0} + \gamma_{10} \sum_{\tau=1}^{t-1} \sum I_{A_{i,b,\tau}>0} + \gamma_{13} \sum_{\tau=1}^{t-1} \sum_{\forall j \in N_i} I_{A_{j,b,\tau}>0}$$
$$+ \gamma_{14} X_{ij} + \gamma_{15} VILL_{i,j,t} + \mu_{i,j,t} \qquad (22)$$

where $X_{i,j,t}$ stands for individual level effect such as education and age and $VILL_{i,j,t}$ can include either village fixed effects or a measure of village-level information flow into the village through extension agents, etc.

The reduced form functions for $A_{b,t}$** and $x_{b,t}$** following the theory model outlined in the previous section are:

$$A_{b,t} = A\left[A_t, p_t, p_{x,t}, p_{s,b,t}, p_{s,n,t}, E_t \overline{A}_t, E_t A_{PF,t}, \delta, \theta, F_t(\alpha(q), \beta_b(q)), \beta_n(q), F_\varepsilon\right] \qquad (23)$$

$$x_{b,t} = A\left[A_t, p_t, p_{x,t}, p_{s,b,t}, p_{s,n,t}, E_t \overline{A}_t, E_t A_{PF,t}, \delta, \theta, F_t(\alpha(q), \beta_b(q)), \beta_n(q), F_\varepsilon\right] \qquad (24)$$

If we would have information on $F_{i,t}(\cdot)$, this would suggest the following specification for $A_{b,i,j,t}/A_{i,j,t}$:

$$A_{b,i,j,t}/A_{i,j,t} = \gamma_{17}.p_{j,t} + \gamma_{18}.\delta_i + \gamma_{19}.RISK_i + \gamma_{20}.E_{i,t}\overline{A}_t$$
$$+\gamma_{21}.E_{i,t} A_{PF,t} + \gamma_{22}.F_{i,t}(\alpha(q), \beta(q)) \qquad (25)$$
$$+\gamma_{23}.VILL_j + c_i + \mu_{ijt}$$

where $RISK_i$ is the subjective measure of risk aversion, $VILL_{i,j}$ is a village fixed effect capturing the influence of $\beta_n(q)$ and θ and c_i is an unobservable individual fixed effect representing the effect of Q and F_ε. As I do not have information on past beliefs, I will use 21 to guide the specification of this term. First focus on the last two terms of 21. Essentially, one needs to answers two questions. First, which other farmers *j* does one select? Second, how does one map $(Y_{j,b,t}, A_{j,b,t}, x_{j,b,t})$ onto $F_{i,t}(\cdot)$? Note that the latter involves the question of aggregation across 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 mapping, one can replace the last term in 21 with:

$$F_{i,t} = \gamma_{14} \sum_{\tau=1}^{t-1} I_{A_{i,b,\tau}>0} + \gamma_{15} \sum_{\tau=1}^{t-1} I_{A_{j,b,\tau}>0} \qquad (26)$$

Admitting, this linear mapping is a rather crude method of capturing the increase in knowledge from one's own and others' experiences. 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*.[14]

As such, one needs to think about sampling the set of contacts of each farmer. Two of the most popular techniques are respondent-driven snowball sampling[15] and taking the "network of a sample". The first technique is useful when one is interested in properties of the network itself, but as it 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 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. 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 that one can take certain aspects of j's production and the relationship between i and j into account. In case of it, it is the knowledge about the profitability of Bt cotton that will drive the binary adoption decision, or:

$$F_{i,t} = \gamma_{16} \sum_{\tau=1}^{t-1} I_{A_{i,b,\tau}>0} + \gamma_{16} \sum_{\tau=1}^{t-1} \sum_{\forall j \text{ linked to } i} I_{A_{j,b,\tau}>0} + \gamma_{17} \sum_{\tau=1}^{t-1} INF_{\tau} \qquad (27)$$

where INF denotes a measure of the information coming from outside of the village. And "*j* linked to *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 pressure and (behavioral) imitation term, the (preliminary) analysis assume correct expectations, ie,

$$E_{i,t}\overline{A}_t = \overline{A}_t, E_{i,t}A_{PF,t} = \sum_{j=1}^{ji=4} (1/4) A_{PF,j,t} \qquad (28)$$

where in 28 I consider only the four most influential progressive farmers.

Results

Before discussing the results of the 2008-09 cropping year analysis following specifications 17, 18, 19 and 20, one remark is in order. The dependent variables of this analysis relate to the future plans of the respondent regarding Bt and non-Bt cotton cultivation. About 130 respondents mentioned however that they were not planning (or did not yet know their plans) to grow cotton at all in June

14. Due to time constraints, one cannot ask the farmer for information regarding their peer's behavior for all peers. 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 perceived behavior of *j* (Hogset and Barrett 2008).

15. This is a technique where existing study subjects recruit future subjects from among their acquaintances. Thus the sample group appears to grow like a rolling snowball.

2008. This issue also relates to the issue of landlessness in the sample. With imperfect land, credit or food crop markets, households who would wish to grow Bt cotton, and who happen to have little or no land, might end up not growing any cash crops at all. About 75% of the respondents claimed that they made their cropping pattern decision first and then, conditional on these, tried to lease (rent) land in/out (rather than the other way around).

In practice, the results of a probit analysis (using a dummy variable taking the value of one, if the respondent plans to grow Bt cotton next year and zero, if the respondent plans to grow non-Bt cotton or no cotton at all) including all Kanzara and Kinkheda households[16] , following specifications 17, 18, 19 and 20 indicate that the acreage of owned land is the only variable that was significantly different from zero at the 1% level (109 observations, results not shown).

Including Aurepalle in the analysis changed the results (241 observations, results not shown). In 17, the variables owned land (1%), M-Bt (5%), SP-Bt (1%), and SP-nonBt (1%) are all significant with the expected sign. In addition, the three village-level fixed effects are significant as well (Aurepalle at 1%, Kanzara and Kinkheda at 5%). Following specification 18, the SAFE variable is significant and with the expected sign at the 5% level. The other variables mentioned remain significant with the expected sign, apart from the Kanzara and Kinkheda fixed effects which lose their significance. Including the number of Bt cotton farmers known in addition to the SAFE variable in 19 does not change the results much, the variable NUMBER is not significant and the Aurepalle fixed effects loses its significance. Similarly, adding the progressive farmers' variable PF as in 20 does not change the results much. The PF variable is not significant, and the Aurepalle fixed effect again gains its significance at the 10% level.

As in Aurepalle, all respondents who had decided to include cotton in their cropping pattern for the year 2008-09 opted for Bt cotton. This difference in results is entirely driven by inter-village differences. The strong significance of the variables concerning the production function of Bt cotton point towards the importance of knowledge. The significance of the SAFE variable illustrates the importance of social pressures. While the lack of significance of the NUMBER and PF variables might be interpreted as a lack of (behavioral) imitation, this result might also indicate a problem in the selection of the set of progressive farmers in case of PF.

Table 5 presents the results of the 2008-09 cropping year analysis (OLS) following specifications 17, 18, 19 and 20 using planned acreage under Bt cotton as a percentage of planned total area under cotton as dependent variable and including only the households that were planning to cultivate cotton during 2008-09 (all villages). The results indicate that the perceived mean of the Bt cotton distribution and price of seed matters consistently over the different specifications, but not a lot. The SAFE variable is (almost) significant at the 10% level in specifications 18 and 19. The small but negative and significant NUMBER indicates that the free rider effect dominates the (behavioral) imitation effect. The large size of the village fixed effects indicates that the majority of the variation is between the villages and not captured by the social pressure and imitation variables. This difference cannot be justified by a difference in the price of cotton (the average price of cotton was Rs 19.54/kg, Rs 22.49/kg and Rs 21.9/kg in Aurepalle, Kanzara and Kinkheda respectively in 2007-08).

Excluding Aurepalle from the analysis (42 observations, results not shown) results in highly unstable parameter estimates. Even after stripping down the model, excluding seed prices that never appear significant, only the following variables appear to influence the percentage Bt decision significantly

16. Except for the households whose respondent answered "don't know" with regard to their future plans. I did not include the variables with regard to time preferences and land quality in the probit regression.

Table 5. Planned Bt area as a percentage of total area (114 observations).				
	Spec 1	Spec 2	Spec 3	Spec 4
M-Bt	0.0500426***	0.061668***	0.0548215**	0.0707883***
M-nonBt	0.0061037	0.011781	0.0127545	0.0130178
V-Bt	-0.0178089	-0.0292495	-0.0252347	-0.0385381
V-nonBt	-0.0079551	-0.0401024	-0.0454785	-0.0437612
SP-Bt	0.0061186	0.0074114	0.0063498	0.000339
SP-nonBt	-0.0164662	-0.0166113	-0.0209205	-0.0144807
Btseedprice	-0.0002382***	-0.0002107**	-0.0001423	-0.0001951**
Nonbtseedprice	0.0004518***	0.0003795**	0.0003096**	0.000389**
Subjrisk	0.0436611	0.0356425	0.0333313	0.0490675
Dummyaur	0.5577296***	0.625275***	0.6746508***	0.6210452***
Dummykanz	0.280399	0.4833304*	0.5626444**	0.6071597**
Dummykink	0.1223878	0.3448538	0.3516863	0.3967136*
SAFE		-0.0769671	-0.0748849	-0.0811434*
NUMBER			-0.0000548**	
PF				-0.2110548

*** significant at 1% level.
** significant at 5% level.
* significant at 10% level.

across specifications: M-Bt, V-Bt, subjrisk and NUMBER. The sizes of the corresponding coefficients are however much larger compared to the ones presented in Table 6, with the M-Bt coefficient around 20 percentage points, the V-Bt coefficient around -17 percentage points and the subjrisk coefficient around 14 percentage points. The SAFE variable is not significant, but its coefficient is large and negative in all specifications concerned, around 15 percentage points. The NUMBER variable is significant at the 10% level but its effect negative but small, once again pointing towards a free rider effect.

None of the control variables for land quality (value per acre or perceived land quality) and time preferences are significant in these analyses. As such, I opt to exclude them from the final analyses.

Table 6 presents the (preliminary) results of the probit regressions using 26 to define the learning term and a second order term for both learning and social pressure[17]. As the experience term perfectly predicts adoption, ie, 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)[18].

The first two columns present the results of the regression without a year*aurepalle interaction terms. The second two columns add aurepalle*year fixed effects to the analysis. Both the probit coefficient and the marginal effects at the average are presented.

17. These regressions were taken in December on data that was only partially validated. So the results should be taken with a pinch of salt.

18. 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 years. I included only last year's experience interacted with five experience dummies.

Table 6. Panel data analysis - mapping 26 (pooled probit).				
	N=1074		N=917	
	Coefficient	dF/dx	Coefficient	dF/d/x
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
Learningsquare	-0.0000162***	-1.85E-06	-0.0000201***	-3.04E-06
Pressure	-1.187079	-0.1357407	-0.9748837	-1.48E-01
Pressuresquare	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	

*** significant at 1% level.
** significant at 5% level.
* significant at 10% level.

Note that having one acre of land (ie, 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 wages. Somewhat surprisingly the price of Bt cotton seed is significant only at the 5% level in the first regression (and has the "wrong" sign) and is not significant in the second regression. This could be explained by the way this price is measured, ie, 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 was measured in 2005, and income as such is not comparable across years, I opt not to include this variable[19].

The first and second order learning terms are significant in both regressions at the 1% level. Note that the sign of the second order learning term indicates a concave learning effect. The imitation term and its square were significant in both regressions. Its coefficient is large; one acre/PF more under Bt cotton increases the probability of adoption, respectively, 7 percentage points and 11 percentage points. In the second regression, the year*aurepalle fixed effects are significant and large for 2004 and 2006. The insignificance of the risk dummy in both regressions one and two is probably due to the rather crude measurement of risk aversion[20].

19. Before 2005, income was measured through a direct question to the respondent who was asked to recall his/her income one year back. From 2005 onwards, income was 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.

20. In the panel data analysis, 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 7 presents the results of the probit regression, using 27 to define the learning term and a second order for learning, social pressure and imitation, including outside information. Interestingly, the male wage was significant in the first regression. The first and second order terms of learning are significant in both regressions. Knowing both cultivar and yield of one additional farmer in the past increases the probability of adoption with 1.2 percentage points and 1.6 percentage points respectively. Knowing both cultivar and yield of ten more additional farmers would therefore increase the probability of adoption with 12 percentage points and 16 percentage points, respectively. Hearing from one additional outside source about Bt cotton increases the probability of adoption by 4 percentage points to 6 percentage points. The first order pressure term remains insignificant, but the second order was significant. The imitation term remains significant and relatively large, 6 percentage points to 11 percentage points. The year*aurepalle interaction dummies are significant for three years.

Table 7. Panel data analysis - mapping 27 (pooled probit).

	N=1074		N=917	
	Coefficient	dF/dx	Coefficient	dF/dx
Land	0.067993***	0.0069739	0.0714803***	0.0097137
Member	0.1471257***	0.0150905	0.1477782***	0.0200822
Pcotton	-0.0209988	-0.0021538	0.0023733	0.0003225
Pmale	0.0716727***	0.0073514	0.0520862	0.0070782
Pfemale	-0.0354244	-0.0036334	-0.056568	-0.007687
Pbtseed	0.0005451**	0.0000559	0.0002835	0.0000385
Learning	0.1097513***	0.011257	0.1074867***	0.0146068
Learningsquare	-0.0046689**	-0.0004789	-0.003843*	-0.000522
Outside	0.4196653***	0.0430444	0.4197865***	0.0570464
Outsidesquare	-0.0905735**	-0.00929	-0.0895567**	-0.01217
Pressure	-1.264985	-0.1297476	-1.162123	-0.157925
Pressuresquare	0.6085618*	0.0624193	0.4427852	0.0601718
Imitation	0.5969088***	0.061224	0.8576377***	0.1165477
Imitationsquare	-0.0419734***	-0.0043051	-0.054643**	-0.007426
Riskdummy	0.0040125	0.0004125	0.0984271	0.0139753
2004*aur			1.497136**	0.4056759
2005*aur			1.788732***	0.5088275
2006*aur			2.004588*	0.5820978
2007*aur			2.134493	0.626477
Constant			-6.735872	

*** significant at 1% level.
** significant at 5% level.
* significant at 10% level.

Conclusion

Preliminary results indicate that social learning matters. The size of the effect depends on how social learning is defined and from whom one learns. In this regard, the difference between learning from fellow farmers and learning from institutional sources is striking. In the panel data analysis, the social pressure effect seems to depend on the inclusion of an imitation term, pointing towards social pressures that are mainly driven by the adoption behavior of the progressive farmers. In the cross sectional analysis concerning the 2008-09 cropping year, a clear social pressure effect originating from safety and environmental concerns with regard to Bt cotton is present. In both panel data analysis and cross sectional analysis, the village fixed effects are significant and large. In the case of the panel data analysis, this might be due to the incorrect measurement of the price of Bt cotton seed in Andhra Pradesh.

As this is a working paper in progress, it might be useful to give a brief overview of future work.

First, the measurement of several variables will be improved upon. The prices of Bt cotton seeds should reflect the prices that farmers are facing in the market each year and a variable income should be included in both cases. In addition, the identity of the farmers j and the structure of the social learning term can be further investigated.

Second, the theoretical model as it stands does not contribute much to the analysis. As such, I will further expand the empirical identification strategy without relying on a specific structural model. The econometric model used should be refined to take into account the "time-effects" of the seemingly irreversible adoption decision, using a duration analysis and extended towards the pesticide decision using a system-based approach.

Third, several of the underlying assumptions need to be tested explicitly, most importantly with regard to the validity of the random matching within sample approach and the correct selection of progressive farmers for the imitation component.

References

Appadurai A. 1989. Transformations in the culture of agriculture. Pages 173-186 *in* Contemporary Indian Tradition: Voices on Culture, Nature and the Challenge of Change (Carla MB, ed.). Washington and London: The Smithsonian Institution Press.

Asia-Pacific Consortium of Agricultural Biotechnology. 2006. Bt Cotton in India - A status report. New Delhi. 45 pp.

Baird S. 2003. Modeling technology adoption decisions: An analysis of high yielding variety seeds in India during the Green Revolution. Working Paper. University of California, Berkeley, USA: Department of Agricultural and Resource Economics. 34 pp.

Bandiera O and **Rasul I.** 2006. Social networks and technology adoption in northern Mozambique. Economic Journal 116:514:869-902.

Bantilan MCS, Anand Babu P, Anupama GV, Deepthi H and **Padmaja R.** 2006. Dryland agriculture: Dynamic challenges and priorities. Research Bulletin no. 20. Patancheru 502 324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics. 32 pp.

Besley T and **Case A.** 1993. Modeling technology adoption in developing countries. American Economic Review 83(2):396-402.

Conley TG and **Udry C.** 2005. Learning about a new technology: Pineapple in Ghana. Working Paper. Yale University, USA: Department of Economics. 54 pp.

Delavande A, Giné X and **McKenzie D.** 2008. Measuring subjective expectations in developing countries: A critical review and new evidence. Policy Research working paper World Bank no. WPS 4824. Washington DC, USA: World Bank. 35 pp.

Feder G, Just R and **Zilberman D.** 1985. Adoption of agricultural innovations in developing countries: A survey. Economic development and cultural change 32(2):255-297.

Foster AD and **Rosenzweig MR**. 1995. Learning by doing and learning from others: Human capital and technical change in agriculture. Journal of Political Economy 103(6):1176-1209.

Hogset H and **Barrett CB.** 2008. Social learning, social influence and projection bias: A caution on inferences based on proxy-reporting of peer behavior. Working Paper Model. University College, USA and Cornell University, USA. 51 pp.

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

Lybbert TJ, Barrett CB Mccpeak JG and **Luseno WK.** 2007. Bayesian herders: Updating of rainfall beliefs in response to external forecasts. World Development 35(3):480-497.

Lybbert TJ and **Just DR.** 2007. Is risk aversion really correlated with wealth? How estimated probabilities introduce spurious correlation. American Journal of Agricultural Economics 89(4):839-1224.

Moser CM and **Barrett CB.** 2006. The complex dynamics of smallholder technology adoption: The case of SRI in Madagascar. Agricultural Economics 35(3):373-388.

Pemsla D, Waibela H and **Orphalb J.** 2004. A methodology to assess the profitability of Bt-cotton: Case study results from the state of Karnataka, India. Crop Protection 23(12):1249-1257.

Pomp M and **Burger K.** 1995. Innovation and imitation: Adoption of cocoa by Indonesian smallholders. World Development 23(3):423-431.

Qaim M and **Zilberman D.** 2003. Yield effects of Genetically Modified crops in developing countries. Science 299: 7 February 2003:900-902.

Qaim M. 2003. Bt cotton in India: Field trial results and economic projections. World Development 31(12): 2115-2127.

Rao KPC and **Kumara Charyulu D.** 2007. Changes in agriculture and village economies. Research Bulletin no. 21. Patancheru 502 324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics. 136 pp.

Rogers E. 1964. Diffusion of Innovations. New York, USA: The Free Press. 79 pp.

Santos P and **Barrett CB.** 2007. Understanding the formation of social networks. Working Paper, Applied Economics and Management, Cornell University, USA. 40 pp.

Scott J. 1991. Social network analysis: A handbook. London - Newburry Park - New Delhi, India: Sage Publications. 205 pp.

Singh RP, Binswanger HP and **Jodha NS.** 1985. Manual of instructions for economic Investigators in ICRISAT's Village Level Studies (Revised). Patancheru 502 324, Andhra Pradesh, India: International Crops Research Institute for the Semi-Arid Tropics. 151 pp.

Sunding D and **Zilberman D.** 2001. The agricultural innovation process: Research and technology adoption in a changing agricultural sector. Pages 207-261 *in* Handbook of Agricultural Economics (Bruce LG and Gordon CR, eds.). Amsterdam, the Netherlands: Elsevier Science.

Vasavi AR. 1994. Hybrid times, hybrid people: Culture and agriculture in South India. Man 29(2):283-300.

Walker TS and **Ryan JG.** 1990. Village and household economies in India's Semi-Arid Tropics. Baltimore, USA and London, UK: John Hopkins University Press. 394 pp.

Appendices

Appendix A

Number of households included in the sample, by date.			
Year	Aurepalle	Kanzara	Kinkheda
From 2001-02	94	52	29
From 2002-03	0	0	0
From 2003-04	5	0	2
From 2004-05	0	0	0
From 2005-06	18	5	22
From 2006-07	11	6	1
From 2007-08	0	0	1
Total sample	128	63	55

Appendix B

The theoretical model follows the following notation. Three kinds of subscripts are employed, the first subscript denotes the individual i, the second subscript denotes the crop and the third subscript denotes time t. Note that subscript i is implied unless otherwise noted. Subscript "b" refers to Bt cotton and subscript "n" refers to non-Bt cotton.

Notation	Description
u	Per-period utility function
c	Per-period consumption
s	Per-period non-material satisfaction
θ	Lingering social pressure term
δ	Discount factor
A	Land
x	Pesticides
$\alpha(q)$ and $\beta_b(q)$	Parameters of Bt cotton production function distributed F((q), $_b$(q))
$B_n(q)$	Parameter of non-Bt production function
$L(\cdot)$	Learning function
N_i	Set of farmers to whom i is connected
Y	Production
ε	Stochastic component of the production function
p	Price of cotton
p_x	Price of pesticide
P_s	Price of seed
Q	Probability of a high bollworm pressure year (q_{H})
1-Q	Probability of a low bollworm pressure year (q_{L})

About ICRISAT

The International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) is a non-profit, non-political organization that does innovative agricultural research and capacity building for sustainable development with a wide array of partners across the globe. ICRISAT's mission is to help empower 644 million poor people to overcome hunger, poverty and a degraded environment in the dry tropics through better agriculture. ICRISAT belongs to the Alliance of Centers of the Consultative Group on International Agricultural Research (CGIAR).

Contact Information

ICRISAT-Patancheru
(Headquarters)
Patancheru 502 324
Andhra Pradesh, India
Tel +91 40 30713071
Fax +91 40 30713074
icrisat@cgiar.org

ICRISAT-Liaison Office
CG Centers Block
NASC Complex
Dev Prakash Shastri Marg
New Delhi 110 012, India
Tel +91 11 32472306 to 08
Fax +91 11 25841294

ICRISAT-Nairobi
(Regional hub ESA)
PO Box 39063, Nairobi, Kenya
Tel +254 20 7224550
Fax +254 20 7224001
icrisat-nairobi@cgiar.org

ICRISAT-Niamey
(Regional hub WCA)
BP 12404, Niamey, Niger (Via Paris)
Tel +227 20722529, 20722725
Fax +227 20734329
icrisatsc@cgiar.org

ICRISAT-Bamako
BP 320
Bamako, Mali
Tel +223 20 223375
Fax +223 20 228683
icrisat-w-mali@cgiar.org

ICRISAT-Bulawayo
Matopos Research Station
PO Box 776
Bulawayo, Zimbabwe
Tel +263 83 8311 to 15
Fax +263 83 8253, 8307
icrisatzw@cgiar.org

ICRISAT-Lilongwe
Chitedze Agricultural Research Station
PO Box 1096
Lilongwe, Malawi
Tel +265 1 707297, 071, 067, 057
Fax +265 1 707298
icrisat-malawi@cgiar.org

ICRISAT-Maputo
c/o IIAM, Av. das FPLM No 2698
Caixa Postal 1906
Maputo, Mozambique
Tel +258 21 461657
Fax +258 21 461581
icrisatmoz@panintra.com

www.icrisat.org

ISBN 978-92-9066-517-5 RBE 024 202-2009