ELSEVIER

Journal of Development Economics 73 (2004) 185–213

JOURNAL OF Development ECONOMICS

www.elsevier.com/locate/econbase

Social learning in a heterogeneous population: technology diffusion in the Indian Green Revolution

Kaivan Munshi

Department of Economics, Brown University, Box B/64 Waterman Street, Providence, RI 02912, USA
NBER, USA

Received 1 June 2001; accepted 1 March 2003

Abstract

Information flows are weaker in a heterogeneous population when the performance of a new technology is sensitive to unobserved individual characteristics, preventing individuals from learning from neighbors' experiences. This characterization of social learning is tested with wheat and rice data from the Indian Green Revolution. The rice-growing regions display greater heterogeneity in growing conditions and the new rice varieties were also sensitive to unobserved farm characteristics. Wheat growers respond strongly to neighbors' experiences, as expected, while rice growers do not. Rice growers also appear to experiment more on their own land to compensate for their lack of social information.

JEL classification: D80; O30
Keywords: Social learning; Technology diffusion

1. Introduction

There has been much interest in social learning, as an explanation for the diffusion process, in recent years (Banerjee, 1992; Bikhchandani et al., 1992 are early contributions to this literature). Social learning essentially describes a process by which an individual learns from his neighbors' experiences (their previous decisions and outcomes) about a new technology. This learning process places restrictions on the amount of information

E-mail address: munshi@brown.edu (K. Munshi).

0304-3878/$ - see front matter
doi:10.1016/j.jdeveco.2003.03.003

that can be generated within the community over time, giving rise to natural lags in adoption.[1]

Social learning is evidently weak, and diffusion rates will be slow, if the individual is unable to observe his neighbors' experiences perfectly. But this does not explain why some innovations diffuse faster than others, even when social information is readily available. It also does not explain why individuals or communities sometimes appear to persistently ignore their neighbors' (positive) experiences.[2] The point that I make in this paper is that it is not enough to observe your neighbors' decisions and their subsequent outcomes. The fact that a new technology worked well for a neighbor does not imply that it will work well for you, if characteristics that determine its performance vary in the population. Ellison and Fudenburg (1993) use this argument to justify simple rules of thumb where individuals learn from similar neighbors only, slowing down the rate of diffusion. The individual could in principle do better than that by controlling for differences between his own and his neighbors' characteristics when learning from their experiences, but only to the extent that these characteristics are observed. Social learning breaks down if unobserved, or imperfectly observed, individual characteristics are important determinants of neighbors' outcomes.

My objective in this paper is to show that social learning will be weaker in a heterogeneous population, particularly when the performance of the new technology is sensitive to neighbors' unobserved characteristics. This supplements previous empirical work that tests the general view that technology diffusion may be driven by an underlying social learning process. Besley and Case (1994) are the first to attempt that I am aware of to test for the presence of information externalities in the economics literature. Subsequently, Foster and Rosenzweig (1995) made an important contribution to this literature, providing support for the presence of social learning in Indian agriculture. It is not surprising that this previous work has focused on agriculture, since it provides a natural setting to test for the presence of social learning. Agricultural production is a relatively simple and well-understood process. The relevant social unit, for the purpose of social learning, is geographically determined in this case as the set of farmers in the village. The set of neighbors is also stable over time, at least in the rural Indian context that these studies consider. I follow these studies and situate my analysis in Indian agriculture as well.

We study a particular episode in recent history in this paper: the adoption of new agricultural technology in the Indian Green Revolution. The Green Revolution is associated with the introduction of high-yielding varieties (HYVs) of wheat and rice in

[1] In contrast, the traditional explanation for this lag focused on heterogeneity in the population (Griliches, 1957; Mansfield, 1968). The idea here was that some people may be more receptive to new ideas and innovations (Rogers, 1962) or, alternatively, that individuals face different economic opportunities which cause them to adopt at different speeds (Griliches, 1957). Note that social learning generates lags in adoption even in a homogeneous population when information about a new technology is imperfect.

[2] Ryan and Gross (1943), in an influential study that spawned an enormous diffusion literature in rural sociology, estimated that it took 14 years before hybrid seed corn was completely adopted in two Iowa communities. Similar lags have been observed in the industry as well; for instance, a 40-year gap appears between the first success of the tunnel oven and its general use in the pottery industry (Carter and Williams, 1957).

the late 1960s, which was followed by their widespread adoption, dramatically increasing farm productivity and rural incomes throughout the developing world.

The rice growing areas of Peninsular India are characterized by wide variation in soil characteristics. In contrast, conditions are fairly uniform in the Northern Plains, where wheat is grown traditionally (ICAR, 1978, 1985). Consistent with this view, we will see in Section 2 that crop yields (measured as crop profits per unit of land) are substantially more variable among the farmers in our sample that reside in the rice-growing areas of the country, as compared to those that reside in the wheat-growing areas.

Further, technological differences between the wheat and rice HYVs would have accentuated the differences between crops that I have described above. The early rice HYVs were quite sensitive to soil characteristics such as salinity, as well as to managerial inputs, which are difficult to observe. The rice grower would thus have found it difficult to control for differences between his own and his neighbors' characteristics when learning from their experiences. For wheat, two of the original six varieties imported from Mexico, Kalyan Sona and Sonalika proved to be extremely robust to growing conditions. This feature of the HYV technology, together with the generally uniform conditions in wheat-growing areas of the country, would have resulted in conditions that were ideal for social learning. As we would expect, while slow diffusion rates were initially observed in the rice growing areas of the country, Kalyan Sona and Sonalika spread rapidly and were ultimately adopted in areas that did not even grow wheat traditionally.

I test the link between unobserved heterogeneity and social learning more formally by estimating the grower's response to his neighbors' decisions and outcomes, separately by crop. We will see that wheat growers place relatively more weight on their neighbors' past acreage allocations and yield realizations, and relatively less weight on their own past decisions, consistent with the view that social learning was smoother for wheat. The same result is obtained independently with both farm-level and district-level data.

While these results are very encouraging, they do not conclusively establish the presence of social learning. To begin with, suppose that each grower bases his acreage decision in part on an (unobserved) information signal that he receives in each period. The neighbors' past decisions could then proxy for these unobserved individual information signals, if they are correlated across growers in the village and over time. When markets function imperfectly, access to scarce inputs will jointly determine acreage allocations and yield realizations. If these input constraints are serially dependent, then a spurious correlation between the grower's current acreage decision and neighbors' past acreage decisions and yield realizations could once more be obtained. Finally, high yields could relax the grower's liquidity constraint when credit markets function imperfectly, creating a direct link between lagged yields and current decisions.

We can avoid the spurious correlation associated with unobserved information signals by focussing on the grower's response to lagged yield realizations, which are observed by the econometrician. But if access to credit and other scarce inputs varied systematically across the wheat- and rice-growing areas of the country, then the observed difference in yield effect across crops could still be explained without appealing to differences in underlying social learning. My solution in this paper to provide additional support for the presence of social learning is to focus on districts that grow both wheat and rice. We will obtain the same pattern described earlier with this restricted sample of districts; HYV wheat

acreage responds strongly to past acreage allocations and yield realizations in neighboring districts, while HYV rice acreage does not, ruling out spatial variation in prices and access to scarce inputs as an alternative explanation for the observed differences across crops.

One advantage of working with an agricultural application is that the crops that we study are similar along many dimensions. Wheat and rice are staple cereals in the Indian diet and these crops typically dominate the grower's investment (acreage) portfolio. The mode of production for the two crops is similar, and given the organization of production in Indian agriculture, with all the plots in the village located in a single area, neighbors' past acreage decisions and their subsequent yield realizations are readily observable for both crops. Finally, the high-yielding varieties of wheat and rice were introduced around the same time and by the same government agencies.

Nevertheless, I do not wish to argue that the distinct diffusion patterns for these crops were driven entirely by differences between wheat and rice along a single learning dimension. HYV wheat provided a much higher return than the traditional technology that it replaced. It was also a relatively stable technology, associated with fairly certain yields. It is well known that an innovation with these characteristics will diffuse more rapidly.[3] The learning dimension that we focus on in this paper supplements the more standard explanations for the variation in diffusion rates across innovations. At the same time, the importance of the point that I am trying to make should not be minimized. As I note below, access to social information appears to have had a significant impact on the grower's investment behavior, by changing his incentive to experiment on his own land.

If the view that rice growers are informationally disadvantaged is correct, then we would expect such growers to compensate for their lack of social information by experimenting on their own land. A nice feature of the agricultural application is that investment is completely divisible, for both crops. The grower can thus choose the precise level of HYV acreage that is optimal for him. The empirical analysis concludes with the observation that rice growers who do adopt HYV allocate a greater share of their land to the new technology than comparable wheat growers. This is despite the fact that the likelihood of adoption is significantly higher for wheat growers. I will show in Section 5 that these patterns in the data are consistent with increased experimentation among rice growers, once we allow for forward-looking behavior in HYV acreage allocations. In contrast, if diffusion rates were faster for wheat only because the new technology provided a higher relative return for that crop, or was more certain, then wheat adopters would have allocated more land to HYV as well. The fact that they did not suggests that access to information mattered a great deal in this environment.

While the specific focus of this paper is on Indian agriculture, the point that I make has wider applicability. It may explain why individuals or communities sometimes appear to persistently ignore their neighbors' positive experiences with regard to migration, occupational choice, educational attainment and other household decisions. The problem in this case is that individual ability, which would have an impact on subsequent outcomes, may be difficult to observe or even characterize. For example, Manski (1993b) considers the problem faced by an individual attempting to learn the returns to schooling from his

[3] Rogers (1962, pp. 124), Mansfield (1968, pp. 119) and more recently Ellison and Fudenburg (1993) in a social learning context, all make this point.

neighbors' experiences. The outcomes of those who did and did not attend school cannot be compared without controlling for unobserved ability, which determines both schooling decisions and subsequent occupational outcomes. Thus, the standard problem of selection bias faced by the econometrician, attempting to infer the returns to schooling, applies to the individual attempting to learn from his neighbors as well. As with our agricultural application, social learning will be restricted in a heterogeneous neighborhood, particularly when occupational outcomes depend strongly on unobserved ability. While there are obviously many factors that prevent individuals from changing their behavior, the information problem that Manski describes, and which I attempt to identify in this paper, may often play an important role as well.

The paper is organized in six sections. Section 2 provides an account of the varietal improvements for wheat and rice during the Green Revolution. The farm-level sample is also described at this point, to check that the patterns in the data across the two crops are broadly consistent with the institutional background and our view of the diffusion process. Section 3 describes the social learning process, starting with the grower's acreage allocation decision when information about the new technology is imperfect. Section 4 bridges the theory and the estimation, discussing the empirical specification and potential bias that could arise from assumptions that were made in the previous section. Section 5 describes the empirical results and Section 6 concludes the paper.

2. Wheat and rice in the Indian Green Revolution

Dwarf varieties of wheat and rice have a higher grain-to-straw ratio and are more responsive to inputs such as fertilizer and irrigation. This improved HYV technology was first introduced in India in the late 1960s, resulting in enormous increases in agricultural productivity over the subsequent decades.[4]

Starting with wheat, two of the original varieties imported from Mexico, Sonora 64 and Lerma Rojo 64A were found to be particularly suited to Indian conditions. They were crossed with local varieties to yield the first generation of HYVs released for mass distribution in 1967. Of the five varieties released in that year, Kalyan Sona and Sonalika proved to be extremely popular and were subsequently adopted throughout the wheat-growing areas of the country.

Wheat is grown in three main regions in India: the Northwest Plains, the Northeast Plains and the Central Peninsular zone.[5] Four of the five varieties released in 1967, Sonalika, Kalyan Sona, Safed Lerma and Chhoti Lerma, were very robust to variation in growing conditions and were adopted in all three regions listed above. One more multi-zone variety Sharbati Sonora was released in 1968. Thereafter, all the new varieties introduced over the next decade were designed for a specific zone or state. Of the 68

[4] The discussion that follows is drawn from Evenson and David (1993) and ICAR (1978,1985).

[5] The Northwest Plains include the states of Punjab, Haryana, (western) Uttar Pradesh, Delhi and Rajasthan. The Northeast Plains include (eastern) Uttar Pradesh, Bihar, Orissa and Bengal. The Central Peninsular zone covers Madhya Pradesh and Gujarat. The states of Punjab, Haryana, Uttar Pradesh, Madhya Pradesh, Rajasthan and Bihar account for 85% of India's wheat production.

varieties released over the 1967–1977 period, 5 were multi-zonal, 34 were tailored to a single zone and 29 were state specific. However, the multi-zonal varieties completely dominated, in terms of overall coverage. For instance, Sonalika accounted for 65% and Kalyan Sona another 20% of the total supply of government-certified HYV seed in 1977–1978. No other variety accounted for more than 2.5% of seed supply that year. As a measure of their dominance, these two varieties even spread to areas that had traditionally grown rice, such as eastern Uttar Pradesh, Bihar and other states in eastern India.

In sharp contrast with the smooth diffusion of HYV wheat, the story for HYV rice is one of setbacks and disappointments. The original semi-dwarf varieties Taichung Native 1 and IR8 were imported from the International Rice Research Institute (IRRI) in the Philippines. The initial objective of the All-India Coordinated Rice Improvement Project (ACRIP) was to identify rice varieties that would yield well throughout the country. Unfortunately, the first HYVs released for mass distribution, Padma and Jaya, did not perform nearly as well as expected. They were found to be unsuitable in a variety of stress conditions such as water logging, salinity and drought. They were also found to be susceptible to a number of pests and diseases. Rice is for the most part grown during the monsoon season, when pests and diseases are rampant. High levels of fertilizer use with HYV cultivation apparently aggravated these problems.

Indian agricultural scientists realized very early that the wheat experience would not be replicated with rice. Over the subsequent decades, the thrust of the research effort was to develop HYVs that were suited to specific local conditions. The local *indicas* are relatively insensitive to the stress factors listed above. Imported varieties were crossed with the local varieties over successive generations, ultimately generating a huge number of HYVs that were both high yielding as well as robust to local conditions. As a measure of the increase in varietal diversity over time; 28 varieties were released in the 1966–1970 period, 73 over the 1971–1975 period, 98 over the 1976–1980 period, 117 over the 1981–1985 period and 151 over the 1986–1991 period. Of the 122 varieties introduced over the 1966–1978 period, only 21 were released by the Central Variety Release Committee for All-India distribution. Certainly, no single variety enjoyed a popularity that was comparable to the dominance of Sonalika and Kalyan Sona for wheat.

The discussion above provides us with two stylized facts which we exploit in the subsequent analysis. First, the early rice varieties were particularly sensitive to characteristics associated with the organic composition and other features of the soil. In contrast, the performance of HYV wheat was remarkably robust to growing conditions. We would thus expect information flows to have been generally smoother across wheat growers, particularly in the early years of the Green Revolution. Second, technological fine-tuning by cross-breeding with local *indicas* made subsequent rice varieties less sensitive to growing conditions. While this might have increased social learning among rice growers, such learning would have been limited in scope since the later rice HYVs were designed for specific local conditions.

Before proceeding to a more formal discussion on social learning, we first check to see whether the basic patterns in the data are consistent with this view of distinction between crops.

The farm-level data, obtained from the National Council of Applied Economic Research (NCAER) Additional Rural Incomes Survey (ARIS), provide information on

acreage allocation, crop incomes and farm characteristics for a sample of growers over a 3-year period, 1968–1970. Foster and Rosenzweig (1995) use the same data for their analysis. I mentioned earlier that the first HYV rice and wheat varieties were introduced in 1966 and 1967, so we are observing adoption behavior literally at the onset of the Green Revolution.

A major disadvantage of the farm-level data is that crop-specific HYV allocations are unavailable. However, we are in a position to identify the crops that are planted by each grower. The farm-level analysis thus partitions the data into wheat and rice villages. Wheat villages allocate more land to wheat than to rice, while the rice villages allocate more land to that crop. We will later need HYV yields in the first two sample years to identify social learning. Excluding villages that allocate no land to HYV in these years, we are left with 1006 growers residing in 92 villages. 32% of the land in wheat villages is allocate to that crop, while only 5% of the land in those villages is allocated to rice. In contrast, 63% of the land is allocated to rice and 10% to wheat in the rice villages. The bulk of the HYV allocation in these villages can thus be attributed to a single crop.

The first panel in Table 1 describes the likelihood of adoption, separately for wheat and rice villages. We see in columns 1–2 that adoption rates are generally increasing over time, with a steeper trajectory for the wheat villages. Since HYV is often grown on irrigated land, I subsequently restrict attention to growers owning some amount of irrigated land to rule out the possibility that the observed differences between crops are driven by underlying differences in the access to irrigation. Comparing adoption rates in columns 3 and 4 with what we saw earlier in columns 1–2, we see that the basic difference between crops continues to be obtained.

The second panel in Table 1 describes acreage allocated to HYV by adopters. Acreage allocations are increasing over time, except for a setback in year 2. The most striking feature of this panel is that adopters in wheat villages allocate less land to HYV than adopters in the rice villages, despite the fact that the average landholding is significantly larger in wheat villages and, as we saw above, that growers in these villages are more likely to adopt HYV. The same pattern across crops is obtained when we restrict attention to growers with positive levels of irrigated land in columns 3–4 of panel B.

These differences between crops are accentuated in panel C of Table 1, which describes the share of total land that is allocated to HYV by adopters. Growers in rice villages were seen to allocate more acreage (in absolute terms) to HYV above, despite holding less land on average, so it is not surprising that the gap between crops now widens. We see in panel C that growers in rice villages who do adopt allocate up to three times more of their land to HYV than comparable growers in wheat villages. Replacing gross-cropped area with gross-irrigated area as the measure of the scale of production in columns 3–4, we see that roughly the same pattern across crops is obtained. Clearly, the difference between crops is not due to a differential access to irrigation.

An important feature of our characterization of the growing environment is that rice growers must deal with much more heterogeneity, some of which will be unobserved, when learning from their neighbors. One obvious implication of this distinction between crops is that crop yields should vary more among the rice growers.

The coefficient of variation of crop yield, computed separately within the wheat and rice villages, is presented next in Table 1. The coefficient of variation across these villages

Table 1
Descriptive statistics: farm-level data

Mean landholding (GCA/GIA)	Mean (standard errors)			
	All growers		Growers with irrigation	
	Rice villages 3.77 (0.07)	Wheat villages 6.18 (0.12)	Rice villages 2.92 (0.07)	Wheat villages 4.56 (0.09)
	(1)	(2)	(3)	(4)
Panel A: Proportion of growers that adopted HYV				
Year 1	0.26 (0.01)	0.29 (0.02)	0.33 (0.02)	0.38 (0.02)
Year 2	0.25 (0.01)	0.40 (0.02)	0.35 (0.02)	0.48 (0.02)
Year 3	0.31 (0.01)	0.49 (0.02)	0.46 (0.02)	0.59 (0.02)
Panel B: HYV acreage allocated by adopters				
Year 1	0.23 (0.01)	0.19 (0.01)	0.21 (0.01)	0.19 (0.01)
Year 2	0.16 (0.01)	0.13 (0.04)	0.16 (0.01)	0.13 (0.004)
Year 3	0.50 (0.01)	0.41 (0.01)	0.50 (0.02)	0.40 (0.01)
Panel C: Share of land allocated to HYV by adopters				
Year 1	0.18 (0.02)	0.05 (0.004)	0.15 (0.02)	0.06 (0.01)
Year 2	0.09 (0.01)	0.03 (0.002)	0.10 (0.01)	0.04 (0.004)
Year 3	0.33 (0.03)	0.12 (0.01)	0.34 (0.03)	0.12 (0.01)
	Coefficient of variation			
	Within village		Between village	
	Rice villages	Wheat villages	Rice villages	Wheat villages
	(1)	(2)	(3)	(4)
Yield variation	0.70 (0.06)	0.50 (0.04)	2.45	0.49

Land ownership and HYV allocation are measured in hectares.
Yield is measured as the profit per unit of land.
Landholding is measured as the gross-cropped area (GCA) in columns 1–2 and gross-irrigated area (GIA) in columns 3–4.
Rice villages allocate more land to rice than wheat. Wheat villages allocate more land to wheat than to rice.
Panel C computes share of GCA in columns 1–2 and share of GIA in columns 3–4.
The within-village coefficient of variation is computed for each village-year.
The reported statistic is the mean (standard error) of the coefficient of variation for wheat and rice villages.
The between-village coefficient of variation is computed using the mean yield in each village.
Equality of means for wheat and rice can be rejected at the 5% significance level for all the statistics reported in this Table with two exceptions: row 1 of panel A (column 1 vs. column 2) and row 1 of panel B (column 3 vs. column 4).

is also provided, separately for each crop. As expected, the coefficient of variation within the village is significantly higher for rice than for wheat. The between-village variation increases sharply for rice while remaining very steady for wheat, consistent with the view that growing conditions are generally uniform in the wheat-growing areas of the country. Note that the crop yield is computed as the crop profit per unit of land. Given the nature of HYV technology for wheat and rice, the difference between crops that we have described would, if anything, be accentuated if we had computed variance statistics with HYV yields rather than crop yields.

This section provides a useful preview to the estimation results that we present later, where we find that social learning is absent among rice growers. Consistent with that result, diffusion rates are seen to be slower here among rice growers. We would also expect rice growers to experiment more on their own land, to compensate for the inferior social information that they receive. Indeed, we saw above that rice adopters allocate more land to HYV than wheat adopters, despite the fact that rice growers as a group are less likely to adopt the new technology and despite the fact that landholdings are significantly smaller on average for the rice growers.

3. Social learning and the adoption of new technology

This section describes the social learning process more formally. We begin with the special case where information about the new technology is perfect. Here, the grower moves to his optimal acreage immediately and social learning is absent. Thereafter, we consider the more interesting case with imperfect information. The grower's objective is now to learn the value of the HYV yield that he should expect on his land, which in turn determines the acreage that he will allocate to the new technology. Neighbors' previous decisions and outcomes (yield realizations) are shown to provide real information in this case. This discussion will ultimately lead to the empirical specification that we use in the paper to test for the presence of social learning. It will also motivate the difference in social learning between wheat and rice that we expect to see in the data.

I made a number of simplifying assumptions as we proceed to derive the empirical specification. Potential bias that could arise from these assumptions will be discussed in Section 4.

3.1. Acreage allocation with perfect information

We consider a simple model of investment in which there are two technologies: a new risky HYV technology and a safe traditional technology. This traditional technology includes traditional varieties of wheat or rice, as well as other crops that are grown in the area. The traditional technology provides a certain yield y_{TV}, where the yield is defined as the profit per unit of land. The assumption that y_{TV} is constant is made for convenience, and I will relax this assumption later in Section 4 when discussing the identification of social learning. The new technology provides a higher yield than the traditional technology but is risky. The expected yield from HYV depends on the quality of the grower's soil and the inputs applied to the new crop. We assume constant returns to scale, which is generally considered to be a reasonable description of agricultural production in India, so input choice can be treated independently of the acreage decision. When markets function smoothly, the choice of inputs will depend on input and output prices, as well as the quality of the soil. The yield from the new HYV technology for grower i in period t is consequently specified as

$$y_{it} = y(Z_i) + \eta_{it} \tag{1}$$

where $y(Z_i)$ is the expected yield and Z_i is a vector of soil characteristics and prices. This would include the price of competing crops as well if we expanded the portfolio of crops that was available to the grower. Notice that we are implicitly ignoring the possibility that prices could vary over time in Eq. (1). I will relax the assumption that prices are constant, as well as the assumption that markets function smoothly, later in Section 4. η_{it} is a mean zero serially independent disturbance term with variance λ_i^2. This term measures deviation from the true yield $y(Z_i)$, which is obtained under normal growing conditions: $E_i(\eta_{it}|y(Z_i))=0$. η_{it} is determined by a combination of rainfall, temperature and soil composition. While the farmer may be aware of some of the individual factors that contribute to the growing conditions in a given year, their net effect on the yield and, hence, η_{it} cannot be observed.

The grower's expected yield $y(Z_i)$ is known with certainty when information is perfect. The grower uses this to choose the optimal acreage A_{it}^* to be allocated to HYV. In general, A_{it} is chosen to maximize the expected utility from the grower's current profit V_{it} in each period, where

$$V_{it} = y_{it}A_{it} + y_{\mathrm{TV}}(\mathrm{GCA}_i - A_{it}) \tag{2}$$

and GCA_i is the gross-cropped area available for cultivation.

The yield from the traditional variety y_{TV} is constant, so the outcome V_{it} is a linear function of a single random variable y_{it}. Under these conditions, Meyer (1987) shows that expected utility maximization is equivalent to the mean–standard deviation approach. With this approach, the grower chooses A_{it} to maximize a function of the expected outcome and the standard deviation (or variance) of the outcome only. It is well known that the mean–variance model of asset–choice can also be applied if the outcome is normally distributed, which requires that $\eta_{it} \sim N(0, \lambda_i^2)$ in our case. Using either of these justifications, the grower chooses A_{it} to maximize $V(Y, \Sigma)$, where $Y \equiv (y(Z_i) - y_{\mathrm{TV}})A_{it} + y_{\mathrm{TV}} \cdot \mathrm{GCA}_i$, $\Sigma \equiv \lambda_i A_{it}$. Since the grower is risk averse, $V(Y, \Sigma)$ is increasing in Y and decreasing in Σ.

Assuming that an interior solution is obtained, and ignoring second-order effects, an increase in $y(Z_i) - y_{\mathrm{TV}}$ makes the new technology more attractive, with a corresponding increase in A_{it}^*.[6] By a similar argument, an increase in λ_i reduces A_{it}^*. The optimal acreage is an increasing function of $y(Z_i) - y_{\mathrm{TV}}$ and a decreasing function of λ_i, both of which are constant over time, so the time subscript on A_{it}^* can be ignored.

$$A_i^* = A(y(Z_i) - y_{\mathrm{TV}}, \lambda_i). \tag{3}$$

Proposition 1. *With perfect information about the new technology, the grower arrives at his optimal acreage immediately and there is no role for social learning.*

[6] The first-order condition is obtained in this case as,

$$\frac{\partial V}{\partial A_{it}} = V_Y(y(Z_i) - y_{\mathrm{TV}}) + V_\Sigma \lambda_i = 0.$$

$(\partial V/\partial A_{it})$ is increasing in $y(Z_i) - y_{\mathrm{TV}}$ if we ignore the second-order effects of $y(Z_i) - y_{\mathrm{TV}}$ on V_Y and V_Σ. We are also assuming that the usual second-order condition is satisfied.

3.2. Acreage allocation with imperfect information

The grower now has incomplete knowledge of the new technology. Specifically, we will assume that the expected yield $y(Z_i)$ is uncertain. Thus $\Sigma \equiv (\lambda_i + \sigma_{it})A_{it}$, where σ_{it}^2 is the variance of the grower's expected yield estimate. The expected yield, which was previously known with certainty, is also replaced by its estimated value $\hat{y}_{it}$. With these modifications, Eq. (3) is rewritten as

$$A_{it} = A(\hat{y}_{it} - y_{\mathrm{TV}}, \lambda_i, \sigma_{it}) \tag{4}$$

where the chosen acreage A_{it} is increasing in $\hat{y}_{it} - y_{\mathrm{TV}}$ and decreasing in λ_i, σ_{it}. Notice from Eq. (4) that the grower does not arrive at his optimal acreage immediately. We will see below that $\hat{y}_{it}$ slowly converges to $y(Z_i)$ as the grower receives more information over time, accompanied by a corresponding convergence in A_{it} to A_i^*.

The next step in the analysis is to characterize the process by which $\hat{y}_{it}$ converges to $y(Z_i)$. The σ_{it} term in Eq. (4) is very important here since the risk-averse grower is always interested in minimizing the uncertainty in $y(Z_i)$; recall that $V(Y,\Sigma)$ is decreasing in Σ and hence in σ_{it}. He would thus like to utilize all the information about $y(Z_i)$ that is available to him. Three sources of information are available to the grower. First, he receives an exogenous information signal, perhaps from the local extension agent, about the value of his expected yield in each period. We assume that this signal provides an unbiased estimate of the yield that the grower should expect on his own land. We will also assume that all growers in the village receive signals of equal precision. Second, he can use his neighbors' decisions to infer the signals that they received.[7] Third, he can learn directly from his own and his neighbors' yield realizations.

The timing of receipt of the alternative sources of information is as follows. At the beginning of a period, the grower receives his private information signal. In a Bayesian setting, the grower combines that signal with his prior at the beginning of the period to compute the best estimate of his expected yield. This in turn determines the acreage that he allocates to HYV in that period. Subsequently, he observes his neighbors' acreage decisions, which reveal the signals that they received, as well as all the yields that are realized in the village. With social learning, the new information from neighbors' decisions as well as their yield realizations is used to update the grower's prior about the value of his expected yield for the next period.

The grower's ability to learn from his neighbors is complicated by the fact that they may not share the same expected yield. The signal that a neighbor receives or his subsequent yield realization can then no longer be directly applied to the grower's own land. He must condition for differences between his own and his neighbor's characteristics when learning from him. This may not be possible if these characteristics are unobserved, in which case the grower may prefer to ignore information from his neighbors if the yield is sufficiently sensitive to the unobserved characteristics.

[7] We assume that the grower can invert his neighbors' decision functions to precisely determine the signal that they received. Banerjee (1992) shows that all the growers in the village could converge to the wrong true yield, and ultimately allocate a sub-optimal level of land to HYV, if this were not the case.

I consider two canonical cases for social learning in the discussion that follows. I will assume in Section 3.2.1 that the expected yield is the same for all growers in the village: $y(Z_i)=y$. This is the ideal situation for social learning, and we will see that the grower's acreage decision is determined by the mean acreage and the mean yield in the village in this case. Subsequently, in Section 3.2.2, I will consider the case where the expected yield depends on the grower's unobserved characteristics. As noted above, social learning could break down if the yield is sufficiently sensitive to these characteristics. This will allow us to derive implications for social learning that distinguish between wheat and rice. Specifically, we would expect the wheat growers to respond strongly to mean acreage and mean yield in their village, while these social effects could be absent among the rice growers.

In describing the social learning process above, I assumed that the uncertainty in the grower's expected yield σ_{it}^2 was exogenously determined by the information that was made available to him up to the beginning of period t. This allowed us to continue to assume that the grower maximized the expected utility from his current profit, or its equivalent $V(Y,\Sigma)$. In practice, the grower could supplement this information by planting more HYV on his own land. Additional acreage in the current period, with the yield "observations" that it provides, reduces the uncertainty in the expected yield in the future. The grower's maximization problem is more complicated in this case since he must account for the effect of his acreage allocation on future profits. Once we allow for social learning, there is also a strategic aspect to such forward-looking behavior, since the incentive to experiment depends on the amount of information that the grower expects to receive from his neighbors in the future. I show in the Appendix A, and in Section 4, that while the incorporation of experimentation in the model introduces additional terms associated with the neighbors' characteristics and access to information in the acreage equation, this does not create new problems for the identification of social learning.

3.2.1. Social learning when the expected yield is constant across growers

Before discussing the process by which the grower learns his true expected yield, I first proceed to rewrite acreage Eq. (4) in a manner that is consistent with the linear specification that I will ultimately estimate.

$$A_{it} = \pi_0 + \pi_1 \hat{y}_{it} + g(X_i, \sigma_{it}) \tag{5}$$

where X_i includes y_{TV} and λ_i from Eq. (4). It also reflects the grower's risk aversion, which determines the weight that is placed on λ_i, σ_{it} in that equation.

I now proceed to describe how $\hat{y}_{it}$ is determined. Expected yields are assumed to be constant across growers in the village so neighbors' information signals and yield realizations provide an unbiased estimate of the grower's own yield, without modification. The information signals that arrive in the village in a given period and the yields that are subsequently realized are pooled together by each grower and applied to update his prior in the next period. All growers in the village begin with a common prior $\hat{y}_0$ in period 0. Since they utilize the same information to update their priors in each period, all growers

have a common prior in subsequent periods as well. Over time these beliefs converge to the true yield y.[8]

Following the timing of various information sources that I outlined earlier, each grower combines the common prior at the beginning of period t, $\hat{y}_t$, with the private information signal that he receives at the beginning of that period, u_{it}, to arrive at his best estimate of the expected yield on his land. Beliefs at the time of planting thus vary across growers in the village in each period. It is straightforward to verify that $\hat{y}_{it}$ will be a weighted average of $\hat{y}_t$ and u_{it}, when η_{it}, u_{it} are uniformally distributed in a Bayesian setting.

$$\hat{y}_{it} = \alpha\hat{y}_t + (1-\alpha)u_{it}. \tag{6}$$

The next step in describing the learning process is to study how $\hat{y}_t$ is determined. For this we need to go back one period in time. After all the growers in the village have made their decision in period $t-1$, we will see below that the mean information signal that they received in that period $\bar{u}_{t-1}$ can be extracted from their decisions. Since no one is systematically misinformed, the mean signal provides more information about the expected yield than any individual signal. Thus, $\bar{u}_{t-1}$ supersedes u_{it-1} when it becomes available at the end of period $t-1$ and so is used to compute each grower's prior for the next period $\hat{y}_{\mathrm{t}}$. Of course, growers' subsequent yield realizations also appear as an additional source of information by the end of period $t-1$. Applying Bayes' Rule once more, the expression for $\hat{y}_t$ is consequently obtained as

$$\hat{y}_t = (1-\beta-\gamma)\hat{y}_{t-1} + \beta\bar{u}_{t-1} + \gamma\bar{y}_{t-1} \tag{7}$$

where $\hat{y}_{t-1}$ captures all of the information about the expected yield that was received in the village up to the beginning of period $t-1$ in Eq. (7). The village means, $\bar{u}_{t-1}$, $\bar{y}_{t-1}$, represent the new information that became available in that period.[9]

Substituting from Eq. (6) in Eq. (5),

$$A_{it} = \pi_0 + \pi_1\alpha\hat{y}_t + \pi_1(1-\alpha)u_{it} + g(X_i, \sigma_{it}). \tag{8}$$

Next, substituting the expression for $\hat{y}_t$ from Eq. (7) in the equation above,

$$A_{it} = \pi_0 + \pi_1\alpha(1-\beta-\gamma)\hat{y}_{t-1} + \pi_1\alpha\beta\bar{u}_{t-1} + \pi_1\alpha\gamma\bar{y}_{t-1} + \pi_1(1-\alpha)u_{it} + g(X_i, \sigma_{it}). \tag{9}$$

[8] We could easily relax the assumption that there is a degenerate distribution of beliefs in period 0, but this would complicate the derivation that follows without adding any intuition to our understanding of the learning process. The grower's prior belief in period t would now be replaced by two terms; z_{i0} the belief in period 0 and $\hat{y}_t$ which represents all the information that subsequently became available in the village. As long as the mean period 0 belief $\bar{z}_{i0}$ is common knowledge, the derivation that we describe below will follow through, leaving only two additional z_{i0}, $\bar{z}_{i0}$ terms in the residual of the acreage regression. These terms are just like X_i, which also lies in the residual of the acreage regression, and so create no new problems for the identification of social learning.

[9] The grower will actually place more weight on his prior as he grows more confident about the yield level on his land. Thus, α will be increasing, while β and γ are decreasing over time. We ignore the time subscripts on α, β, γ to simplify the exposition, but we will return to this point later when we study changes in learning over time. Note that we only take expectations across growers in the village at a given point in time in the discussion that follows, and so this simplification does not affect the derivation of the acreage function in any way.

While Eq. (9) completely describes the grower's learning process, $\hat{y}_{t-1}$, $\bar{u}_{t-1}$ are not directly observed by the econometrician. We therefore proceed to derive expressions for these terms as functions of observed variables to test for the presence of social learning. We begin with the $\bar{u}_{t-1}$ term. Lagging Eq. (8) and then taking expectations across all growers in the village,

$$\bar{u}_{t-1} = \frac{1}{\pi_1(1-\alpha)}[\bar{A}_{t-1} - \pi_0 - \pi_1\alpha\hat{y}_{t-1} - \bar{g}(X_i, \sigma_{it-1})].$$

$\hat{y}_{t-1}$, $\bar{g}(X_i,\sigma_{it-1})$ are common knowledge at the beginning of period $t-1$, so each grower can invert the (linear) acreage function to compute $\bar{u}_{t-1}$ from the observed acreage decisions $\bar{A}_{t-1}$ in period $t-1$.[10]

Substituting the expression for $\bar{u}_{t-1}$ derived above in Eq. (9),

$$A_{it} = \pi_0\left(1 - \frac{\alpha\beta}{1-\alpha}\right) + \pi_1\left(1 - \gamma - \frac{\beta}{1-\alpha}\right)\hat{y}_{t-1} + \frac{\alpha\beta}{1-\alpha}\bar{A}_{t-1} + \pi_1\alpha\gamma\bar{y}_{t-1} + \xi_{it}. \tag{10}$$

where $\xi_{it} \equiv \pi_1(1-\alpha)u_{it} + g(X_i,\sigma_{it}) - (\alpha\beta/1-\alpha)\bar{g}(X_i,\sigma_{it-1})$ consists of terms that are unobserved by the econometrician. Now we only need to express $\hat{y}_{t-1}$ in terms of observed variables to complete the derivation of the acreage function. Lagging Eq. (8) by one period,

$$\hat{y}_{t-1} = \frac{1}{\pi_1\alpha}[A_{it-1} - \pi_1(1-\alpha)u_{it-1} - g(X_i, \sigma_{it-1})].$$

The important point here is that all growers have the same prior at the beginning of period $t-1$. So each grower's acreage decision A_{it-1} provides information (to the econometrician) about the $\hat{y}_{t-1}$ term. Substituting the expression for $\hat{y}_{t-1}$ in Eq. (10), we finally arrive at a specification of the acreage function that we can take to the data,

$$A_{it} = \pi_0(\alpha+\beta) + \left(1 - \gamma - \frac{\beta}{1-\alpha}\right)A_{it-1} + \frac{\alpha\beta}{1-\alpha}\bar{A}_{t-1} + \pi_1\alpha\gamma\bar{y}_{t-1} + \epsilon_{it} \tag{11}$$

with $\epsilon_{it} = \xi_{it} - (1-\gamma-(\beta/1-\alpha))[\pi_1(1-\alpha)u_{it-1} + g(X_i,\sigma_{it-1})]$ consisting of all terms that continue to be unobserved by the econometrician.

Eq. (11) is very intuitive. When information is pooled efficiently within the village, A_{it-1} contains all the information about the expected yield that was available at the beginning of period $t-1$; specifically, the entire history of information signals and yield realizations up to that time. Conditional on A_{it-1}, $\bar{A}_{t-1}$ represents the new information that was received by the village in period $t-1$ through the exogenous signals. Similarly, $\bar{y}_{t-1}$

[10] We implicitly assume here that an interior solution to the acreage decision is always obtained. If we allow for the possibility that acreage decisions are censored at zero, then the grower would need information about the latent variable A^*_{jt-1} to correctly infer neighbor j's signal u_{jt-1} when $A_{jt-1}=0$. The unweighted average of the neighbors' previous period acreage decisions $\bar{A}_{t-1}$ will most likely no longer be a sufficient statistic to describe learning from past decisions. But the grower's yield response, which is the focus of our identification strategy, would be unaffected by such misspecification of the acreage function.

represents the information that was obtained from the yield realizations in that period. Note that the grower's own lagged yield y_{it-1} does not enter independently in Eq. (11) since it is superseded by the mean yield in the village $\bar{y}_{t-1}$.

Proposition 2. *When expected yields are constant in the village, the grower's acreage decision is determined by his lagged decision and the mean of his neighbors' previous decisions and yield realizations.*

3.2.2. Social learning when the expected yield depends on growers' characteristics

The previous scenario that we discussed was the most suitable for social learning. Neighbors' signals and yields provide an unbiased estimate of the grower's own expected yield and so can be utilized without modification. In fact, information from a neighbor is as useful as the information that the grower receives himself. However, the expected yield will more generally depend on the farmer's characteristics. The grower could condition for differences between his own and his neighbors' observed characteristics when learning from them. But the prospects for social learning decline immediately once we allow for the possibility that some of these characteristics may be unobserved, or imperfectly observed.

The grower may be prepared to accept transitory errors in his yield estimates; associated with the η disturbance term in Eq. (1) or noise in the exogenous information signals. However, mistakes that arise because he is unable to control for differences between his own and his neighbors' characteristics when learning from their yields are persistent, and therefore more serious. Take the case where all the neighbors' characteristics are unobserved by the grower. He now has two choices. He could rely on his own information signals and yield realizations, ignoring information from his neighbors. Consistent but inefficient estimates of the expected yield would be obtained with such individual learning. Alternatively, he could continue to utilize information from his neighbors, measured by the mean acreage and the mean yield, as before. The efficiency of his estimates increases with social learning since more information is being utilized, but some bias will inevitably be introduced since the grower cannot control for variation in the underlying determinants of the yield when learning from his neighbors. The grower will ultimately choose between individual learning and social learning on the basis of the trade-off between bias and efficiency.

Proposition 3. *The grower will choose individual learning if the population is heterogeneous and the yield is sufficiently sensitive to unobserved characteristics; otherwise, he will prefer to learn from his neighbors.*

The preceding result helps us predict the differences between crops that we should expect to see in the data. Wheat roughly corresponds to the case with uniform yields within the village, so we would expect the grower's acreage decision to respond strongly to his neighbors' past decisions and yield realizations. In contrast, the rice-growing areas of the country are more heterogeneous than the wheat-growing areas, and we know that HYV rice yields were particularly sensitive to unobserved inputs and soil characteristics. If the yield was sufficiently sensitive to these unobserved inputs, then

Proposition 3 tells us that the rice growers may have preferred to ignore information from their neighbors.

4. Empirical specification and potential bias in estimation

Based on Eq. (11), the acreage function that we estimate, separately for wheat and rice, is specified as

$$A_{it} = \prod_0 + \prod_1 A_{it-1} + \prod_2 \bar{A}_{t-1} + \prod_3 \bar{y}_{t-1} + \epsilon_{it}. \tag{12}$$

Given the institutional background, we would expect wheat growers to respond strongly to their neighbors' decisions and yields, whereas social learning should be weak or absent among the rice growers. In terms of the parameters of the model, β and γ should be large for wheat, and close to zero for rice. Looking back at Eq. (11), this tells us that Π_2, Π_3 should be large for wheat, while Π_1 should be large for rice; growers place relatively more weight on their own decisions when social learning is weak.

The statistical problem that arises immediately with this regression is that A_{it-1} and u_{it-1} (which appears in ϵ_{it} from Eq. (11)) are mechanically correlated; this is easy to see by lagging Eq. (8) by one period. $\bar{A}_{t-1}$ will also be correlated with u_{it-1} if information signals are correlated across neighbors in the village.[11] We have no reason to expect that this bias will vary across crops, and so the difference in the Π_2 estimates for wheat and rice could still be attributed to underlying differences in social learning. Nevertheless, the discussion on identification will shift focus to $\bar{y}_{t-1}$, since the link between individual acreage allocations and lagged yield realizations in the village provides us with a cleaner and more direct test of social learning.

We will later see that strong yield effects are obtained for wheat growers ($\hat{\Pi}_3>0$), while these effects are absent among the rice growers ($\hat{\Pi}_3=0$). While these results are very promising, the challenge for me is to convince the reader that the positive and significant yield effect that we obtain for wheat growers does in fact reflect an underlying social learning process. The discussion on identification that follows will thus concentrate on the wheat regression, and assume that all growers in the village share a common yield y. We assume that each grower receives unbiased information signals about the yield on his land; taking expectations over time $E_t(u_{it})=y$. Recall from Eq. (1) that $\bar{y}_{t-1}=y+\eta_{t-1}$; taking expectations over time once again, $E(\bar{\eta}_{t-1})=0$ and so u_{it} and $\bar{y}_{t-1}$ both have mean y. The regression analysis exploits variation in $\bar{y}_{t-1}$ across villages and over time to identify social learning. Following the discussion above, u_{it} in the residual of the acreage regression, and information signals more generally, will be mechanically correlated with $\bar{y}_{t-1}$, biasing the Π_3 estimate.

[11] Following Manski (1993a), neighbors' past acreage allocations will be spuriously correlated with the grower's current acreage decision if any unobserved determinant of the acreage decision is correlated across neighbors and over time.

One solution to the mechanical correlation described above is to study acreage response to changes in yield realizations in the village. Given the sequential nature of the social learning process, neighbors' yield realizations in period $t-1$ alone are used to update the grower's prior at the beginning of the next period. Since the econometrician has access to yield realizations over multiple periods, $\bar{y}_{t-1}$ can be replaced by $\bar{y}_{t-1}-\bar{y}_{t-2}$ to implement this alternative test of social learning. $\bar{y}_{t-1}-\bar{y}_{t-2}=\bar{\eta}_{t-1}-\bar{\eta}_{t-2}$, which is uncorrelated with u_{it}, u_{it-1}, and so avoids the mechanical correlation described above. Note that this is a relatively stringent test of social learning since we are studying whether the grower responds to the noise component η_{t-1} in his neighbors' past yields.[12] While the econometrician may not observe this noise, it is still possible to obtain an estimate of this noise, as described above.[13]

The district-level data provide HYV yields directly, so the procedure described above follows through without modification. But HYV yields are unavailable with the farm-level data. Our measure of the yield, from Eq. (2), will be

$$\frac{V_{it}}{A_{it}} = y + \eta_{it} + y_{\mathrm{TV}}\left(\frac{\mathrm{GCA}_i}{A_{it}} - 1\right).$$

Lagging one period and taking expectations across growers (who allocate land to HYV) in the village,

$$E_{t-1}\left(\frac{V_{it-1}}{A_{it-1}}\right) = y + \bar{\eta}_{t-1} + y_{\mathrm{TV}}E_{t-1}\left(\frac{\mathrm{GCA}_i}{A_{it-1}} - 1\right).$$

$E_{t-1}(V_{it-1}/A_{it-1})-E_{t-2}(V_{it-2}/A_{it-2})$ gives us exactly what we want, $\bar{\eta}_{t-1}-\bar{\eta}_{t-2}$, except for the additional $\Delta \equiv E_{t-1}(\mathrm{GCA}_i/A_{it-1}-1)-E_{t-2}(\mathrm{GCA}_i/A_{it-2}-1)$ term. Looking back at Table 1, HYV allocations were fairly stable for both crops from Year 1 to Year 2. Δ will be small, and so omission of this additional term should have little impact on the Π_3 estimate. GCA_i/A_{it-1}, GCA_i/A_{it-2} are also observed by the econometrician, so we can include Δ directly in the acreage regression, under the maintained assumption that y_{TV} is constant across villages and over time.[14]

[12] The estimated yield effect will change when we replace $\bar{y}_{t-1}$ with $\bar{y}_{t-1}-\bar{y}_{t-2}$ for two reasons. First, the correlation with unobserved information signals is avoided when we difference the yield. Second, an additional $\bar{y}_{t-2}$ term appears in the residual of the acreage regression when we difference the yield. This term will be negatively correlated with $\bar{\eta}_{t-1}-\bar{\eta}_{t-2}$, so we obtain conservative estimates of the yield effect with this alternative estimation strategy.

[13] Why does the grower not go through a similar exercise to remove the noise from his neighbors' yield realizations? Unlike the econometrician, who is only interested in determining whether or not the grower learns from his neighbors' yields, the grower is interested in a particular y component of the yield realization. Purging the yield realization of the noise removes all the useful information about y as well in this case. To illustrate this point, the grower's best estimate of the noise in period $t-1$ is $\bar{y}_{t-1}-\hat{y}_{t-1}$, where $\hat{y}_{t-1}$ is the prior at the beginning of that period. Subtracting the noise estimate from $\bar{y}_{t-1}$, the grower is back to where he started and nothing will be learned about y in period $t-1$.

[14] We could easily relax the assumption that y_{TV} is constant over time by including $E_{t-1}(\mathrm{GCA}_i/A_{it-1}-1)$, and the corresponding term for period $t-2$, separately in the acreage regression. But we must maintain the assumption that y_{TV} does not vary across villages in a given year (I will return to this point below).

Moving ahead to other potential sources of bias, we next consider the omission of σ_{it}, σ_{it-1} in Eq. (12). Note that we do not need to worry about the X_i term once we focus on the differenced yield. I mentioned earlier that it is the σ_{it}^2 term that gives rise to social learning, since the risk-averse grower wants to reduce uncertainty in y as much as possible. This term also contributes to the commonly observed increase in the aggregate adoption of an innovation, since it declines in value over time as more information is progressively received by each grower. While σ_{it}^2 obviously plays an important role in the learning process, its omission from the acreage regression does not bias our estimates of the yield effect. σ_{it}^2 is decreasing in the number of yield "observations" that the grower has access to. With social learning, σ_{it}^2 is therefore a function of past HYV allocations in the village. While this means that $\sigma_{it}^2(\bar{A}_0,\ldots,\bar{A}_{t-1})$ will be correlated with the acreage terms, A_{it-1}, $\bar{A}_{t-1}$, it will still be orthogonal to $\bar{\eta}_{t-1}$.[15]

I conclude the discussion on bias by considering additional omitted variables that could arise due to assumptions that were made in Section 3. In that section, we ignored market imperfections, which give rise to input constraints, as well as price fluctuations. This allowed us to specify the expected yield as $y(Z_i)$ in Eq. (1), which corresponds to the constant y when the expected yield is assumed to be invariant across growers in the village. Once these assumptions are relaxed, y must be replaced by $y+\epsilon_{t-1}$, where ϵ_{t-1} is an input or price shock in period $t-1$. These shocks affect both the extent of cultivation (acreage) as well as its intensity (yield). Thus, they could, in principle, affect the acreage decision and yields in period $t-1$, A_{it-1}, $\bar{A}_{t-1}$, $\bar{y}_{t-1}$, as well as current acreage A_{it} if they are serially correlated, giving rise to a spurious yield effect.

Differencing the yield does not help in this case since $\bar{\eta}_{t-1}+\epsilon_{t-1}$ continues to be correlated with the omitted determinant of the current acreage when the shocks are serially correlated. Moreover, the shocks that we discussed above are all received prior to planting. Suppose instead that the village received a weather shock after planting. Now the differenced yield would be orthogonal to the other regressors in the acreage equation but could still be correlated with unobserved determinants of the acreage in the next period through changes in the grower's liquidity constraint.

As noted in the Introduction, my solution in this paper to provide additional support for the presence of social learning is to show that the same pattern of estimated acreage and yield effects, across crops, is obtained with the full sample of districts and with a restricted sample of districts that allocate land to both HYV wheat and rice. By comparing the acreage and yield response, across crops, for the same set of districts, we can effectively rule out the alternative explanation that the estimated differences between crops are driven by differences in access to scarce resources across the wheat- and rice-growing areas of the country. In addition, we will include lagged farm incomes directly in the acreage regression to control for the grower's access to credit.

[15] Acreage decisions in period $t-1$ are made before yields are realized in that period, so σ_{it}^2 will be uncorrelated with $\bar{\eta}_{t-1}$. But σ_{it}^2 will be correlated with $\bar{\eta}_{t-2}$, and hence with the differenced yield. Looking back at Eq. (10), it is straightforward to verify that this correlation only leads to conservative estimates of the yield effect. The preceding argument also tells us that σ_{it-1}^2 will be uncorrelated with both $\bar{\eta}_{t-1}$, $\bar{\eta}_{t-2}$. I did not consider forward-looking behavior when characterizing the learning process in Section 3. But I show in the Appendix A that forward-looking behavior only introduces an additional λ_j, σ_{jt} term for each neighbor j in the grower i's acreage function. Following the discussion above, this does not create any new problems for estimation.

5. The empirical analysis

This section explores the presence of social learning in the data. I begin by estimating the acreage equation with farm-level data. Thereafter, I repeat the analysis at the district level. Social learning is noticeably stronger among wheat growers with both datasets. Subsequently, we discuss patterns in the data, observed earlier in Table 1, that suggest that informationally disadvantaged rice growers are more likely to experiment on their own land.

5.1. Estimating the acreage equation with farm-level data

We have 3 years of farm-level data. Since two lags are required to compute $\bar{y}_{t-1}-\bar{y}_{t-2}$, the farm-level regressions are run with data from year 3 only. Given our measure of the HYV yield, V_{it}/A_{it}, the sample must also be restricted to villages that allocated some amount of land to HYV in the first 2 years. We begin with a preliminary regression in Table 2, columns 1–2, that includes $\bar{y}_{t-1}$ directly as a regressor. Growers in wheat villages respond strongly to their neighbors' lagged acreage decisions and yield realizations, while growers in rice villages place relatively more weight on their own lagged decisions.[16] Notice, in particular, that yield effects are completely absent for rice. Following the discussion in the previous section, these results are consistent with the view that social learning was strong for wheat, and weak or absent for rice.

We repeat the exercise described above in Table 2, columns 3–4, but replace lagged yield $\bar{y}_{t-1}$ with the differenced yield $\bar{y}_{t-1}-\bar{y}_{t-2}$. The point estimates of the yield effect do decline for wheat, but otherwise, the results are entirely unchanged from columns 1–2. The acreage coefficients are very stable for both crops, and the yield effect for rice continues to be absent.[17]

HYVs must be cultivated on irrigated land to perform near their potential, and hence, access to irrigation has been seen to be an important constraint to HYV adoption. Including gross-irrigated area as an additional control in Table 2, columns 5–6, we see that irrigation has a strong effect on HYV acreage allocation for both crops. But the remaining coefficients are unchanged from columns 4–5.

We complete the basic discussion on the farm-level regressions by including $\Delta \equiv E_{t-1}$ $(\mathrm{GCA}_i/A_{it-1}-1)-E_{t-2}$ $(\mathrm{GCA}_i/A_{it-2}-1)$ as an additional control in Table 2, columns 7–8.

[16] The grower's (or the district's) own lagged acreage is excluded when computing the neighbors' lagged acreage in all the regressions that are presented in this paper.

[17] Note that we could drop the acreage terms from the regression without noticeably affecting the (differenced) yield effects. The differenced yield coefficient in column 4 declines from 0.158 to 0.148, while the differenced yield coefficient in column 3 continues to be tiny, when the acreage terms are excluded from the regression. I assume, as do Besley and Case (1994), that growers are aware of the optimal level of input use on their plots: the expected yield can therefore be treated as exogenous. In contrast, Foster and Rosenzweig assume that optimal use is learned over time: the yield is endogenous and increasing in neighbors' lagged acreage allocations. They specify the lagged yield $\bar{y}_{t-1}$ to be the true yield y minus an error term $\epsilon(\bar{A}_0,\ldots,\bar{A}_{t-2})$, which is declining in past HYV allocations in the village. Differencing the yield now no longer leaves us with $\bar{\eta}_{t-1}$. The additional $\epsilon(\bar{A}_0,\ldots,\bar{A}_{t-2})$ term must still be accounted for. Since this term would be correlated with A_{it-1}, $\bar{A}_{t-1}$, the stability of the differenced yield coefficient allows us to rule out another potential source of spurious correlation.

Table 2
Farm-level regressions

Dependent variable:	Acreage allocated to HYV							
Village:	Rice	Wheat	Rice	Wheat	Rice	Wheat	Rice	Wheat
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lagged own acreage	0.756	0.272	0.752	0.245	0.686	0.200	0.679	0.203
	(0.217)	(0.158)	(0.217)	(0.155)	(0.215)	(0.153)	(0.216)	(0.152)
Lagged village acreage	1.037	2.232	1.008	2.074	0.935	1.959	0.918	1.963
	(0.309)	(0.332)	(0.307)	(0.333)	(0.302)	(0.350)	(0.297)	(0.353)
Lagged village yield	0.098	0.276	–	–	–	–	–	–
	(0.239)	(0.108)						
Differenced village yield	–	–	0.022	0.158	0.087	0.176	−0.387	0.159
			(0.139)	(0.035)	(0.134)	(0.037)	(0.366)	(0.054)
Gross irrigated area	–	–	–	–	0.014	0.008	0.014	0.008
					(0.004)	(0.004)	(0.004)	(0.004)
Δ	–	–	–	–	–	–	0.0004	0.00001
							(0.0003)	(0.00004)
R-squared	0.267	0.415	0.266	0.427	0.285	0.444	0.295	0.443
Number of observations	593	413	593	413	593	413	593	413
Number of villages	53	39	53	39	53	39	53	39

Standard errors in parentheses account for correlated residuals within the village and heteroscedasticity.
Acreage is measured in hectares. Yield is measured as the ratio of farm profit to HYV acreage.
Village yield is the average across farmers in the village. Lagged yield is the year 2 statistic. Differenced yield is the year 2 statistic minus the year 1 statistic.
Δ is the bias term that must be included because HYV yields are not directly observed.
Rice villages allocate more land to rice than wheat. Wheat villages allocate more land to wheat than to rice.
Columns 1–2: lagged yield, without irrigation control.
Columns 3–4: differenced yield, without irrigation control.
Columns 5–6: differenced yield, with irrigation control.
Columns 7–8: differenced yield, with irrigation control and bias term.

The rice yield effect is somewhat unstable in this regression—it is negative and large in absolute magnitude—but imprecisely estimated. In contrast, the wheat yield effect continues to be positive and significant, and is very similar to what we obtained in column 4 and column 6.[18] The acreage coefficients for both crops remain very stable across all the specifications in Table 2.

When the true mean is constant in the village, the grower's own yield is superseded by the village mean $\bar{y}_{t-1}$. This is why y_{it-1} did not appear in Eqs. (11) and (12). However, y_{it-1} does belong in the acreage equation when social learning is weak or absent. Indeed, it is the possibility of observing more yield realizations that motivates the grower to experiment on his own land, to compensate for his lack of social information. To avoid the possibility that neighbors' lagged yield realizations simply proxy for the grower's own

[18] I also experimented with a more flexible specification, which includes both terms in Δ separately as regressors; this effectively allows y_{TV} to vary over time. The results are very similar to what we obtain in columns 7–8. Suppose, instead, that y_{TV} varied over time and across villages. Then we would have an additional yield term in the residual of the acreage regression, which could be correlated with the differenced HYV yield if HYV and traditional yields are correlated. But note that the only direct role for traditional yield in the acreage regression would be to relax the grower's liquidity constraint, and we have already discussed that particular source of bias.

Table 3
Farm-level regressions—robustness tests

Dependent variable:	Acreage allocated to HYV					
Model:	OLS				Tobit	
Village:	Rice (1)	Wheat (2)	Rice (3)	Wheat (4)	Rice (5)	Wheat (6)
Lagged own acreage	–	–	0.756	0.094	1.316	0.345
			(0.201)	(0.160)	(0.312)	(0.247)
Lagged village acreage	0.469	3.615	0.886	2.072	1.141	2.638
	(0.588)	(0.667)	(0.290)	(0.347)	(0.376)	(0.316)
Differenced village yield	0.072	0.190	0.088	0.175	0.227	0.363
	(0.150)	(0.044)	(0.132)	(0.030)	(0.170)	(0.067)
Gross irrigated area	0.018	0.012	0.021	0.001	0.029	0.015
	(0.010)	(0.006)	(0.007)	(0.004)	(0.007)	(0.004)
Lagged own income	–	–	−0.011	0.008	–	–
			(0.006)	(0.002)		
R-squared	0.018	0.378	0.297	0.462	0.165	0.474
Number of observations	340	173	593	413	593	413
Number of villages	53	39	53	39	53	39

Standard errors in parentheses account for correlated residuals within the village and heteroscedasticity.
Acreage is measured in hectares. Yield is measured as the ratio of farm profit to HYV acreage.
Village yield is the average across farmers in the village. Differenced yield is year 2 statistic minus the year 1 statistic.
Differenced traditional–HYV ratio is the bias term that is included because HYV yields are not directly observed.
Rice villages allocate more land to rice than wheat. Wheat villages allocate more land to wheat than to rice.
Columns 1–2: sample restricted to farmers who did not allocate land to HYV in year 2.
Columns 3–4: lagged income included as an additional regressor.
Columns 5–6: Tobit model.

lagged yield in the acreage regression, we restrict attention to growers who allocated no land to HYV in year 2 in Table 3, columns 1–2. These growers have no yield realizations of their own to learn from, yet we see that wheat growers continue to respond to their neighbors' yields, while the yield effect for rice is absent. The response to neighbors' lagged decisions is also much larger for wheat than for rice, consistent with the patterns we observed before.[19]

We noted earlier that high yields could affect the grower's subsequent acreage decision by relaxing his liquidity constraint. To control for the grower's access to credit, we include lagged income as an additional regressor in Table 3, columns 3–4. Lagged income has a positive and significant effect on wheat acreage. But the lagged acreage and yield coefficients, for both crops, are unaffected by the inclusion of this additional control.

Finally, when describing the grower's investment decision in Section 3, I assumed that an interior solution was always obtained. In practice, the risk-averse grower would allocate no acreage to HYV if the uncertainty associated with the new technology outweighed the additional mean return that it provided. If we assume that the terms in the residual are

[19] The same result is obtained when the sample is restricted to growers who allocated no land to HYV in years 1 and 2.

normally distributed, then the Tobit model can be applied to estimate the acreage regression. Eq. (12) is rewritten as,

$$A_{it}^{*} = \prod\nolimits_{0} + \prod\nolimits_{1} A_{it-1} + \prod\nolimits_{2} \bar{A}_{t-1} + \prod\nolimits_{3} \bar{y}_{t-1} + \epsilon_{it}$$

$A_{it}=A_{it}^{*}$ if $A_{it}^{*}>0$, $A_{it}=0$ otherwise.

Estimates with the Tobit model are presented in columns 5–6 of Table 3.[20] The usual pattern of coefficients, across crops, continues to be obtained.

5.2. Estimating the acreage equation with district-level data

District-level data are available for all 270 districts in the major agricultural states of the country over an extended time period. The data are compiled from official Government of India sources. These include Directorate of Economics and Statistics publications and Indian Agricultural Statistics. Evenson et al. (1994) provide a detailed description of the data and its construction.

Acreage in the district-level regressions is measured by the amount of land allocated to HYV wheat or rice in the district. Neighbors' lagged acreage is measured by the average amount of land allocated to HYV (for the relevant crop) in the geographically contiguous districts in the previous year. We will also experiment with average HYV in the state as an alternative measure of the neighbors' acreage decisions. The district-level data cover the 1969–1985 period, so a full set of year dummies will be included in all the regressions.[21] Finally, HYV yields, measured as the output per unit of land separately for wheat and rice, are available at the state level. As with the farm-level regressions, neighbors' lagged yield $\bar{y}_{t-1}$ will be replaced by the differenced yield $\bar{y}_{t-1}-\bar{y}_{t-2}$ in most of the district-level regressions.

We implicitly assume that a representative grower makes decisions in the district-level regressions and that he is learning from decisions and outcomes in the neighboring districts. But such aggregation of the underlying farm-level learning process could bias the estimated coefficients on the district's own lagged acreage as well as neighbors' lagged acreage. The distortion that arises here is that shocks to a district's acreage fail to be completely captured by its neighbors in the subsequent period. To see this, consider the impact of an exogenous information signal that arises in the far corner of a district. Since the signal travels from farm to farm as it propagates through the district, it may take many periods before it is received by growers in the adjacent district. The signal would also continue to have an impact for multiple periods once it reached the adjacent district, as it spread through the farms there. This could generate a downward bias on the neighbors'

[20] The standard errors in Table 2 and columns 1–4 of Table 3 correct for heteroscedasticity and correlated residuals within the village. However, the Tobit estimates do not account for heteroscedasticity, and it is well known that biased estimates could be obtained in this case. The results in columns 5–6 should therefore be treated with caution.

[21] The district-level regressions restrict attention to districts that allocate land to HYV wheat or rice at some point over the full 1969–1985 sample period and to district-years in which some amount of wheat or rice is planted. Thus, we discard districts that are completely unsuitable for HYV cultivation from the analysis.

lagged acreage, while the coefficient on the district's own lagged acreage would be biased in the opposite direction.

Using state-level yields to measure the neighbors' past outcomes creates similar aggregation problems. If yields are uniform within states, but vary across larger geographical areas, then state yields will provide us with a good measure of neighbors' yields. We would expect this to be the case with wheat. But if the scope of social learning is limited, as with rice, then yield realizations at the state level could fail to reveal the presence of social learning.

Keeping in mind the limitations of district-level analysis discussed above, the basic regression result in Table 4, columns 1–2, broadly matches the farm-level results presented earlier in Tables 2 and 3. Wheat growers place relatively more weight on their neighbors' lagged decisions, but the difference between crops is not as pronounced as it was with farm-level data. More importantly, a strong yield effect (significant at the 10% level) is obtained for wheat, while this effect is absent for rice. These differences between crops are accentuated when we include gross cropped area as an additional control in the acreage regression in Table 4, columns 3–4. Irrigation has a positive and significant effect on HYV acreage for both crops. But notice now that rice growers place substantially more weight on their own lagged decisions, consistent with the view that social learning is weaker for that crop, and that the yield effect for wheat is now more precisely estimated while the corresponding effect for rice remains absent.

Table 4
District-level regressions

Dependent variable:	Acreage allocated to HYV					
Crop:	Rice (1)	Wheat (2)	Rice (3)	Wheat (4)	Rice (5)	Wheat (6)
Lagged district acreage	0.994	0.981	0.966	0.881	0.966	0.881
	(0.015)	(0.022)	(0.022)	(0.045)	(0.022)	(0.044)
Lagged neighbors mean acreage	0.021	0.045	0.026	0.040	0.027	0.041
	(0.020)	(0.017)	(0.020)	(0.017)	(0.020)	(0.017)
Differenced state yield (1969–85)	−0.0001	0.012	−0.001	0.011	–	–
	(0.002)	(0.007)	(0.002)	(0.005)		
Differenced state yield (1969–77)	–	–	–	–	0.002	0.013
					(0.001)	(0.007)
Differenced state yield (1978–85)	–	–	–	–	−0.008	0.008
					(0.007)	(0.004)
Gross irrigated area	–	–	0.022	0.051	0.022	0.051
			(0.009)	(0.018)	(0.009)	(0.017)
Q Statistic	0.434	0.111	0.389	0.005	0.388	0.005
R-squared	0.935	0.933	0.937	0.938	0.937	0.938
Number of observations	2249	2499	2249	2499	2249	2499

Standard errors in parentheses account for correlated residuals within the state and for heteroscedasticity.
All regressions are estimated with a full set of year dummies.
$Q \sim X_1^2$ under H_0: no serial correlation. Critical value above which the null is rejected at the 5% level is 3.84.
Columns 1–2: estimate yield effect over the full sample period (1969–1985), without irrigation control.
Columns 3–4: estimate yield effect over the full sample period (1969–1985), with irrigation control.
Columns 5–6: study stability of the yield effect, with irrigation control.

The time span covered by district-level data is long enough to study the stability of the yield effects over time. While we treated the learning weights α, β, γ as constants in Section 3 to simplify the exposition, we noted that α would be increasing, and β, γ decreasing, over time. This implies in turn that the yield effect Π_3 should be decreasing over time. Partitioning the time series into two equal periods, 1969–1977 and 1978–1985, in Table 4, columns 5–6, we see that wheat yield effects remain strong in both periods, although the point estimate does decline over time. In contrast, yield effects for rice are completely absent in both periods. It is possible that state-level yield is just to aggregate a statistic to identify underlying social learning for rice, particularly in the later years of the Green Revolution. Recall that rice HYVs were designed for specific local conditions, and so social learning would have been limited in scope for this crop.[22]

One advantage of the district-level analysis is that both crops are cultivated in many districts. We now proceed to estimate the acreage regression with a restricted sample of district years in which HYV yields for both wheat and rice are available. Yields for both crops are included in the acreage regression in Table 5, columns 1–2. Concentrating on the yield effects, we see that the usual pattern across crops is obtained; yield effects are much stronger for wheat than for rice. It is also reassuring to observe that cross-crop yield effects are absent.[23]

An alternative strategy to control for unobserved district-level variables is to include fixed effects in the acreage regression. The differenced yield is now replaced by the lagged yield in Table 5, columns 3–4; with fixed effects, we are effectively differencing from the average district yield over the sample period. The coefficient on the district's lagged acreage, which was previously most likely capturing both state dependence and heterogeneity, declines as expected. The neighbors' acreage effect is now absent, but this might be due to the bias that appears when the lagged decision is included in fixed effect regressions. But notice that the yield effects are very similar to what we obtained previously; they are strong and significant for wheat, and absent for rice.

Continuing with these robustness tests, we replace the set of contiguous districts with the state when computing neighbors' lagged acreage in Table 5, columns 5–6. Neighbors' acreage and yield are now computed at the same level of aggregation in this specification. Comparing the estimates in columns 5–6 with what we obtained earlier in Table 4, columns 3–4, the same pattern of coefficients is obtained; wheat growers place relatively more weight on their neighbors' lagged acreage and yields. We conclude the district-level analysis by reporting Tobit estimates of the acreage regression in Table 5, columns 7–8.

[22] I also experimented with an alternative specification that allowed the coefficients on lagged district acreage and neighbors' lagged acreage to vary over time as well. While the estimated yield effects are very similar to what we saw in columns 5–6, a similar decline in the acreage coefficients over time for wheat is not observed.

[23] The results just described compare the wheat and rice yield response for the same set of districts. This helps rule out the possibility that access to scarce inputs, such as credit and fertilizer, varies systematically across the wheat- and rice-growing areas of the country, driving the differences between crops that we observe in the data. But these results could still be generated by crop-specific differences across districts and over time. It is reassuring to note that including output prices, measured at the state level, in the acreage regression has absolutely no effect on the estimated acreage or yield effect for both crops. The price effects are actually negative, which is most likely because acreage (output) and prices are jointly determined.

Table 5
District level regressions – robustness tests

Dependent variable:	Acreage allocated to HYV							
Model:	OLS						Tobit	
Neighbors:	Contiguous districts				State		Contiguous district	
Crop:	Rice (1)	Wheat (2)	Rice (3)	Wheat (4)	Rice (5)	Wheat (6)	Rice (7)	Wheat (8)
Lagged district acreage	1.008 (0.023)	0.903 (0.050)	0.627 (0.065)	0.659 (0.100)	0.984 (0.020)	0.879 (0.045)	0.967 (0.009)	0.893 (0.012)
Lagged neighbors mean acreage	0.022 (0.030)	0.050 (0.032)	0.073 (0.079)	−0.016 (0.079)	−0.021 (0.015)	0.039 (0.015)	0.038 (0.012)	0.038 (0.012)
Differenced state yield (rice)	0.002 (0.003)	−0.002 (0.003)	–	–	−0.0001 (0.002)	–	−0.001 (0.001)	–
Differenced state yield (wheat)	−0.002 (0.003)	0.009 (0.005)	–	–	–	0.012 (0.005)	–	0.009 (0.002)
Lagged state yield	–	–	−0.001 (0.002)	0.014 (0.007)	–	–	–	–
Gross irrigated area	0.012 (0.003)	0.037 (0.016)	0.130 (0.048)	0.114 (0.025)	0.023 (0.009)	0.053 (0.018)	0.024 (0.003)	0.047 (0.004)
Q Statistic	0.720	0.091	0.098	0.021	0.428	0.004	0.377	0.009
R-Squared	0.938	0.957	0.953	0.951	0.937	0.938	0.235	0.239
Number of observations	1534	1534	2249	2499	2257	2499	2249	2499

Standard errors in parentheses account for correlated residuals within the state and for heteroscedasticity (except columns 5–6).
All regressions are estimated with a full set of year dummies.
$Q \sim X_1^2$ under H_0: no serial correlation. Critical value above which the null is rejected at the 5% level is 3.84.
Columns 1–2: includes all state years which allocate land to both HYV rice and HYV wheat.
Columns 3–4: include district fixed effects.
Columns 5–6: the set of neighbors is specified to be all districts in the state (except the district itself).
Columns 7–8: tobit model.

The basic differences between crops continue to be obtained; in particular, while strong yield effects are obtained for wheat, these effects are absent for rice.

5.3. Extending the analysis to permit experimentation

Our view in this paper is that information flows across neighbors are restricted among rice growers. Can the rice grower substitute in some way for this missing information? We previously assumed that the variance in the expected yield estimate, σ_{it}^2 in Eq. (4), was exogenous. The risk-averse grower utilized all the information that was available to him, from external sources and his neighbors, to reduce σ_{it}^2. But he made no attempt to manipulate σ_{it}^2 by experimenting with the new technology. One possibility for the informationally disadvantaged rice grower is to experiment more on his own land, since allocating acreage to HYV makes more yield realizations available, reducing σ_{it}^2 in the future.

How would we identify experimentation in the data? To derive implications for HYV adoption and acreage allocation with experimentation, I extend the grower's investment

problem presented in Section 3 (as noted in the previous section, this extension to the model does not affect the empirical analysis discussed earlier in any way). The grower now takes account of the effect of his acreage decision on future profits. Since neighbors' decisions provide free information, through their subsequent yield realizations, there is also a strategic element to the decision-making process. Our main result, which is derived more formally in the Appendix A, is that a grower without access to social information will experiment more on his own land. However, this additional individual information will not compensate him entirely for the information that he would have received from his neighbors.

The intuition for this result is quite straightforward. Experimentation is costly since the grower must bear all the risk on the acreage that he allocates to HYV. In contrast, with social learning, the grower can free-ride on his neighbors' yield observations. I show that as long as the gain from experimentation is declining at the margin, the informationally disadvantaged grower will never fully compensate for his lack of social information. Once we introduce heterogeneity in the population with regard to the acreage allocation decision, and allow for the possibility that a grower could choose not to allocate any land to HYV, this implies that growers with access to neighbors' yields will be more likely to adopt the new technology. This explains a striking stylized fact that we observed in Table 1; rice growers are less likely to adopt HYV, but will allocate more acreage to the new technology when they do adopt, presumably because their incentive to experiment is sufficiently strong.

6. Conclusion

Information should flow less smoothly in a heterogeneous population, particularly when the performance of a new technology is sensitive to unobserved or imperfectly observed individual characteristics. The problem in this case is that the individual will be unable to condition for differences between his own and his neighbors' characteristics when learning from their experiences.

We test this view of the social learning process with data from the Indian Green Revolution. Wheat and rice are similar crops in many ways, and new high yielding varieties for both crops were introduced at the same time in the late 1960s. However, the rice-growing regions of the country are characterized by much greater heterogeneity in growing conditions. The new HYV technology for rice was also found to be particularly sensitive to conditions which are difficult to observe. We find, as expected, that social learning is weaker among rice growers. The same result is obtained independently with farm- and district-level data, explaining in part the relatively rapid diffusion of HYV wheat. Consistent with this view, rice growers also appear to experiment more on their own land, presumably to compensate for the lack of social information.

What can the government do to reduce this information problem? Information was provided to growers in India through what is known as the Training and Visit (T&V) system of agricultural extension (Feder and Slade, 1986 provide details of this system). Under the T&V system, extension workers focus their attention on a small group of contact farmers in each village. The implicit assumption here is that information will propagate from these

farmers through the rest of the village. This system evidently worked very well with wheat. For rice, we would expect that the few contact workers in each village had little impact on HYV adoption. In general, it may be necessary to invest in more concentrated external information programs when the flow of social information is restricted.

Acknowledgements

This is a substantially revised version of Chapter 1 of my PhD dissertation "Three Essays on Institutions and Economic Development," Department of Economics, MIT, 1995. I thank Abhijit Banerjee, Esther Duflo, Charles Manski, Jack Porter, Mark Rosenzweig, Chris Udry, two anonymous referees and numerous seminar participants for valuable comments and suggestions. Robert Evenson very graciously provided me with the district data and Andrew Foster and Mark Rosenzweig were kind enough to offer me the farm-level data used in this paper. I am responsible for any errors that may remain.

Appendix A. Acreage allocation with experimentation

To derive implications for HYV adoption and acreage allocations with experimentation, I extended the grower's investment problem presented earlier in Section 3. To keep the exposition simple, we will study a two-person two-period world. We will also ignore the role of information signals in the discussion that follows; yield realizations are the only source of information about the expected yield.

We begin with the second and final period T, in which there is no experimentation. Grower i chooses his acreage to maximize $V(Y_T,\Sigma_T)$ as usual, where $Y_T \equiv A_{iT}\hat{y}_{iT}$, $\Sigma_T \equiv \sigma_{iT}A_{iT}$. Notice that the yield from the traditional variety y_{TV} and the natural variation in the yield λ_i have been suppressed to simplify the exposition. The variance in the expected yield estimate $\sigma_{iT}^2(A_{iT-1}+A_{jT-1})$ is decreasing in the number of yield "observations" made available to the grower after period $T-1$. Both growers have the same true yield y, so A_{iT-1} and the neighbor's acreage A_{jT-1} provide perfectly substitutable information.

With this set-up, A_{it} satisfies the first-order condition

$$V_{Y_T}\hat{y}_{iT} + V_{\Sigma_T}\sigma_{iT} = 0 \tag{13}$$

from which we obtain an expression analogous to Eq. (4),

$$A_{iT} = A(\hat{y}_{iT}, \sigma_{iT}(A_{iT-1} + A_{jT-1})). \tag{14}$$

A similar expression is derived for neighbor j. Working back one period, the decision problem becomes more complex. The grower chooses A_{iT-1} to maximize $V(Y_{T-1},\Sigma_{T-1})+\delta V(Y_T,\Sigma_T)$, where δ is a discount factor. The corresponding first-order condition is now obtained as

$$V_{Y_{T-1}}\hat{y}_{iT-1} + V_{\Sigma_{T-1}}\sigma_{iT-1} + \delta V_{\Sigma_T}A_{iT}\frac{\partial \sigma_{iT}}{\partial A_{iT-1}} = 0 \tag{15}$$

where σ_{iT-1} is the grower's (exogenous) prior about the yield uncertainty, which is assumed to be common knowledge in the village.[24] σ_{jT-1} is the corresponding prior for neighbor j.

It is the third term on the left-hand-side of Eq. (15) that reflects the forward-looking nature, as well as the strategic element, of the decision-making process. Since $V_{\Sigma T}<0$ and $(\partial\sigma_{iT}/\partial A_{iTt-1})<0$, this term is positive. This implies, in turn, that $V_{Y_{T-1}}\hat{y}_{iT-1}+V_{\Sigma_{T-1}}\sigma_{iT-1}<0$. The optimal acreage in period $T-1$ is higher than it would be without experimentation.

The grower's incentive to experiment will depend on the information that he expects to receive from his neighbor's yield realizations. Grower i treats A_{jT-1} as parametric when making his decision in period $T-1$. Applying the implicit function theorem to the first-order condition, Eq. (15),

$$\left[\frac{\partial}{\partial A_{iT-1}}\left(V_{Y_{T-1}}\hat{y}_{iT-1}+V_{\Sigma_{T-1}}\sigma_{iT-1}\right)+\frac{\partial F}{\partial A_{iT-1}}\right]\mathrm{d}A_{iT-1}+\frac{\partial F}{\partial A_{jT-1}}\mathrm{d}A_{jT-1}=0 \qquad (16)$$

where $F\equiv\delta V_{\Sigma_T}A_{iT}(\partial\sigma_{iT}(A_{iT-1}+A_{jT-1})/\partial A_{iT-1})$ is the gain from experimentation.

Since A_{iT-1} and A_{jT-1} are perfect substitutes as a source of information in period T, $(\partial F/\partial A_{iT-1})=(\partial F/\partial A_{jT-1})$. Further, we noted that A_{iT-1} was above the optimal level corresponding to maximization of $V(Y_{T-1},\Sigma_{T-1})$ alone. $\partial(V_{Y_{T-1}}\hat{y}iT-1+V_{\Sigma_{T-1}}\sigma_{iT-1})/\partial A_{iT-1}$ must therefore be negative. Finally, F represents the gain from experimentation, which we saw was positive. It seems reasonable to assume that the marginal gain from experimentation declines as more information is accumulated, $(\partial F/\partial A_{iT-1})=(\partial F/\partial A_{jT-1})<0$. Under these conditions, it is easy to verify from Eq. (16) that $-1<(\mathrm{d}A_{iT-1}/\mathrm{d}A_{jT-1})<0$. An increase in neighbor j's acreage lowers grower i's incentive to experiment, but the reduction is less than one-to-one.

Substituting the expression for A_{it} from Eq. (14) in the first-order condition, Eq. (15), we obtain

$$A_{iT-1}=A(\hat{y}_{iT-1},\sigma_{iT-1},\hat{y}_{iT},\sigma_{iT}(A_{iT-1}+A_{jT-1})) \qquad (17)$$

A similar expression is obtained for A_{jT-1}. We could, in principle, solve these equations simultaneously to derive the Markov Perfect Equilibrium of this game. It is not necessary, however, to derive the equilibrium levels of A_{iT-1} and A_{jT-1}. We are interested in comparing the behavior of wheat and rice growers. Since the wheat grower learns from his neighbors, $A_{jT-1}>0$ in equilibrium. In contrast, social learning is effectively absent for rice, which is equivalent to specifying $A_{jT-1}=0$ for that crop. The discussion above tells us that rice growers will therefore experiment more on their own land, so A_{iT-1} will be larger for them everything else being equal. However, the total amount of information available, measured by $A_{iT-1}+A_{jT-1}$, will be greater for wheat.

[24] The $(\partial V(Y_T,\Sigma_T)/\partial A_{iT})(\partial A_{iT}/\partial A_{iT\text{-}1})$ term drops out of the first-order condition above since $(\partial V(Y_T,\Sigma_T)/\partial A_{iT})=0$ from Eq. (13).

I conclude this discussion by deriving the acreage function, corresponding to Eq. (5), with forward-looking behavior. Substituting the corresponding equation for A_{jT-1} in Eq. (17) above, we obtain

$$A_{iT-1} = A(\hat{y}_{iT-1}, \hat{y}_{jT-1}, \hat{y}_{iT}, \hat{y}_{jT}, \sigma_{iT-1}, \sigma_{jT-1}) \tag{18}$$

Since all the growers in the village have the same true yield y, grower i's best estimate of $\hat{y}_{jT-1}$, $\hat{y}_{jT}$, $\hat{y}_{iT}$ in period $T-1$ is just $\hat{y}_{iT-1}$. The acreage function can therefore be simplified as

$$A_{iT-1} = A(\hat{y}_{iT-1}, \sigma_{iT-1}, \sigma_{jT-1}) \tag{19}$$

Leaving aside y_{TV} and λ_i, the only difference between this expression and Eq. (5), which is derived without forward-looking behavior, is the addition of the σ_{jT-1} term.

References

Banerjee, A.V., 1992. A simple model of herd behavior. The Quarterly Journal of Economics 117 (3), 797–817.

Besley, T., Case, A., 1994. Diffusion as a learning process: evidence from HYV cotton, Princeton University, RPDS Discussion Paper #174.

Bikhchandani, S., Hirshleifer, D., Welch, I., 1992. A theory of fads, fashion, custom, and cultural change as informational cascades. Journal of Political Economy 100 (5), 992–1026.

Carter, C.F., Williams, B.R., 1957. Industry and Technical Progress. Oxford Univ. Press, London.

Ellison, G., Fudenburg, D., 1993. Rules of thumb for social learning. Journal of Political Economy 101 (4), 612–643.

Evenson, R., David, C., 1993. Adjustment and Technology: The Case of Rice Development Centre Studies, OECD, Paris.

Evenson, R.E., Pray, C.E., Rosegrant, M.W., 1994. Sources of agricultural productivity growth in India. IFPRI Research Report.

Feder, G., Slade, R.H., 1986. A comparative analysis of some aspects of the training and visit system of agricultural extension in India. Journal of Development Studies, 22.

Foster, A.D., Rosenzweig, M.R., 1995. Learning by doing and learning from others: human capital and technical change in agriculture. Journal of Political Economy 103 (6), 1176–1209.

Griliches, Z., 1957. Hybrid corn: an exploration in the economics of technological change. Econometrica 25 (4).

Indian Council of Agricultural Research, 1978. Wheat Research in India, New Delhi.

Indian Council of Agricultural Research, 1985. Rice Research in India, New Delhi.

Mansfield, E., 1968. The Economics of Technological Change. W.W. Norton, New York.

Manski, C.F., 1993a. Identification of endogenous social effects: the reflection problem. Review of Economic Studies, 60.

Manski, C.F., 1993b. Adolescent econometricians: how do youth infer the returns to schooling. In: Clotfelter, C.T., Rothschild, M.T. (Eds.), Studies of Supply and Demand in Higher Education. University of Chicago Press, Chicago.

Meyer, J., 1987. Two-moment decision models and expected utility maximization. The American Economic Review 77 (3), 421–430.

Rogers, E.M., 1962. Diffusion of Innovations. The Free Press, New York.

Ryan, B., Gross, N.C., 1943. The diffusion of hybrid seed corn in two Iowa communities. Rural Sociology, 8.