Mutlidimensional and Selective Learning

A case study of Bt cotton farmers in India*

Srijita Ghosh†

November 2018

JOB MARKET PAPER

Click here for the latest version

Abstract

Most production technologies require using an optimal combination of multiple inputs. Farmers need to choose the best combination of seeds, fertilizers, pesticides etc. to maximize yield. They can learn about the production function by observing the conditional productivity of combinations of inputs (*cell*) or by the marginal productivity of each input across cells (*average*), where both types of learning are costly. I characterize the optimal learning strategy: observing an average is optimal for higher uncertainty and observing a cell is optimal for lower uncertainty. In a *sequential* learning problem with an optimal stopping time the optimal learning strategy is to start with observing averages and then switch permanently to observing cells. Depending on the uncertainty of averages, learning about averages only can be optimal, at the cost of a higher probability of error ("selective learning"). Selective learning describes the behavior of Indian cotton farmers when they switched to pest-resistant Bt seeds, as they did not reduce their pesticide use sufficiently. This informs about optimal extension policies (what *type* of information) for various types of production function. I also show that the learning mechanism in a laboratory setting predicts the behavior of subjects in the lab.

Keywords: Multidimensional learning, Selective Learning, Agricultural Technology
JEL Code: D83, D90, C91, Q16

*I am indebted to my advisors Andrew Caplin, Martin Rotemberg and Raquel Fernandez for their invaluable guidance and support throughout the project. I also greatly benefited from the comments and suggestions of Andrew Schotter, Debraj Ray, Laurent Mathevet, Dilip Abreu, David Pearce, Boyan Jovanovic, Alessandro Lizzeri, Ariel Rubinstein, Sylvain Chassang, Erik Madsen, Sharon Traiberman and Saravana Ravindran. I am also grateful to seminar participants at Micro theory workshop NYU, Applied micro student lunch NYU, ACEGD ISI Delhi 2017, NEUDC Conference 2018 for their comments and suggestions. Financial support from NYU CESS Lab is gratefully acknowledged. I thank Anwar Ruff for his technical support for the experiment. All mistakes are my own.

†Economics Department, New York University, email: sg3642@nyu.edu

1 Introduction

Farmers need to choose seed variety, fertilizer, pesticide, level of irrigation etc to maximize yield, managers of small firms need to decide the assignment of labor between tasks and machines to minimize the disruption in the production process (Hanna et al. (2014), Bloom et al. (2013) etc). In general, if the productivity of inputs is correlated then learning about the true production function becomes very difficult due to a large number of possibilities created by the multiplicity of inputs.

In response producers often do not try to learn about the overall optimal input mix. One reason for this is that it is often possible to learn about the marginal(or *average*) productivity of a single input(where the average is taken over all possible combination of other inputs). For example, instead of learning about the resulting yield from each combination of seed variety, fertilizer, pesticide, and irrigation level the farmer can learn about the average productivity of each seed variety, picking the one which is best on average.

In most standard learning model the decision maker (DM henceforth) can acquire information about the payoff function prior to the choice of action subject. The learning strategy can involve sequential sampling, choice of posterior precision, partitioning the observed data into categories subject to implicit or explicit cost of learning (Hébert and Woodford (2017), Caplin and Dean (2015), Fryer and Jackson (2008)).

However, learning average productivity of one input (*average* henceforth) and learning productivity of input combination (*cell* henceforth) are not necessarily equally or similarly informative about the payoff function. Hence the learning problem of a producer who has access to both types of information, both of which is costly, involves a novel trade-off that is not present in traditional learning problems with only one type of information.

If a farmer learns about the average productivity for each of the input and chooses a combination of inputs generating highest average productivity he would not necessarily choose the best combination of inputs. This is because learning about averages are not necessarily fully informative about the payoff function. This implies if the farmer relies too much on averages he will have a higher probability of error. On the other hand, since averages are informative about a subset of inputs, the farmer does not need to learn about every possible combination and save the cost of learning.

Selective learning is defined as a learning strategy where the producer only acquires information by observing averages and never observes any specific input combinations. This definition is similar to earlier definitions of selective attention in the literature where a producer or DM learn only about the payoff relevance of a subset of required input Schwartzstein (2014)). Note that, selective learning can lead to a higher probability of error, i.e, lower gross productivity even though it is optimal for the farmer since the cost of learning goes down.

The research question to be addressed in this paper is as follows: what is the optimal learning strategy for a producer when he can choose to observe conditional productivity of each input combinations (namely, what is the best combination of seed and pesticide) along with marginal productivity of each input (namely, what is best seed variety on *average*) separately. Given the optimal strategy, I want to find the condition under which *selective learning*, i.e, learning only about averages, even at the cost of a higher probability of error is optimal for the producer.

In this paper, I consider a case study of *selective learning* farmers. In 2002 a new variety of seed, Bt cotton, was introduced in India. Bt cotton is a genetically modified seed variety that is resistant to cotton bollworm, the major pest for cotton. Prior to 2002, pesticide use consisted of around 40% of the total input cost on average for Indian cotton farmers. Bt cotton was genetically designed to reduce the cost of pesticide. Bt cotton was also better along with other dimensions, for example, the average yield with Bt cotton has been higher than the non-Bt hybrids.

After the introduction in 2002 within ten years (2012-13) almost 95% of farmers adopted Bt cotton variety. Also, farmers rely on their prior belief about optimal pesticide use from the non-Bt period (2002-03) to decide the optimal level of pesticide even after switching to Bt cotton. However, I find that the cotton farmers are continuing the high pesticide input, resulting in lower net profit.

For most developing countries, especially India adoption of technology has remained one of the major policy challenge (Foster and Rosenzweig (2010)). In the Bt cotton case, however, the adoption rate of the new technology was significantly high but conditional on adopting the farmers are not optimally adjusting their pesticide use. Even though the lack of optimal implementation has been a common feature for many developing economy contexts as well (World Bank Report 2008 refers this as *management gap*). But the Bt cotton case is special since instead of underuse of key inputs farmers are overusing inputs leading to lower profitability. Thus this behavior can not be explained by financial constraints or other supply-side constraints. Thus the importance of imperfect learning becomes especially relevant to the Bt cotton case study.

I consider a DM who chooses two inputs, each with finitely many levels (n), to maximize expected output. Before choosing an optimal condition the DM can learn about the production by observing *cell* (input combinations) or *avearge* (marginal productivity of one input). I consider a *sequential* learning problem where in each round DM can choose to observe a cell or an average conditional on prior information. Finally, DM chooses an optimal stopping time for the learning strategy and chooses an input combination to get the final payoff.

The optimal learning strategy is characterized by the uncertainty of the belief, where the uncertainty is measured by Shannon entropy of the belief distribution. I find the optimal choice in any round of learning is to observe an average for higher uncertainty and a cell for lower uncertainty (whenever learning is optimal). This implies the optimal learning strategy for the sequential learning problem is to start with observing averages and then make a one-time switch to observing cells. Also, there exists a combination of prior belief and cost of learning such that only one type of learning is optimal.

However, this result rests on the assumption that by observing a cell the uncertainty of the belief over possible payoff functions do not go up. Since I do not restrict the possible prior beliefs the DM can have, there exists some beliefs where learning about a cell indeed can increase uncertainty. To deal with this issue for the main result I exclude these types of priors.

Moreover, I find that when the DM is choosing to observe a cell, the optimal strategy is to choose the cell that has highest one-period ahead expected payoff, i.e., the DM does not need to investigate the entire path of learning following the observation of the cell. Thus when observing cells the DM mostly *exploits* his available information. On the other hand, if the DM chooses to observe an average it is optimal for him to choose the average that reduces the uncertainty of the belief the most. Hence, while observing averages the DM wants to *explore* the possible payoff function.

Note that, this finding is different than standard learning by doing or Bandit problems in learning where the DM also faces a trade-off between exploiting, using the information he has about the payoff function and exploring, learning or experimenting to increase the precision of the posterior belief (Jovanovic and Nyarko (1996), Francetich and Kreps (2016a), Francetich and Kreps (2016b)). The separability of cells and averages results into this finding.

Also, I find that learning is only optimal if uncertainty is not too high or too low. Too high uncertainty discourages the DM to start learning at all and for too low uncertainty the learning cannot significantly affect the final payoff. This result is similar to rational inattention problems and optimal stopping problems (Drift-diffusion models) in learning (Matějka and McKay (2015), Fudenberg et al. (2017)).

Given the optimal learning strategy, *selective learning* is optimal when averages are sufficiently informative about the productivity of combination of inputs, i.e., observing only averages reduces the significant level of uncertainty and also cost of learning cell is not too low. For instance, in the Bt cotton example, on average Bt cotton gives significantly higher yield than non-Bt hybrids, so, observing only the average yield of the seed varieties reveals more information compared to the (hypothetical) case where the effect of Bt cotton on yield could only be observed in conjunction with lower pesticide use. This reduces the incentive of the farmers to observe the payoff from specific seed, pesticide combination leading to *selective learning*.

Since the optimal learning strategy requires answering *what* to learn about, reducing the cost of one type of learning can distort the incentive to learn and reduce the precision of learning as well. This generates the policy implication that a demand-driven extension policy where farmers decide what to learn about and subsequently information is provided at a lower cost by extension worker may not be able to resolve the problem of selective learning and can possibly make it worse. However, this implication rests on the assumption that the farmers have more than one possible learning technology available but does not use the multidimensionality of the payoff function.

Exploiting the multidimensionality, the payoff functions can be categorized into groups where selective learning would be optimal and/or would reduce the payoff (compared to learning only about the productivity of input combinations). In summary, a technological change that increases the average (but not high enough to reach the maximum value for sure) but also changes the correlation between the two inputs would lead payoff relevant selective learning. The higher average yield of Bt cotton combined with a change in the level of pesticide use fits the criteria. Whereas, other technologies, such as drought-resistant rice in India that does not improve yield on average but only affects yield for a specific combination of input choices would not display selective learning behavior.

Finally, since observing the learning behavior of farmers is not possible in the survey data, I test the learning mechanism in a laboratory setting. The subjects are making a decision in a multi-dimensional learning setup where they have the option to learn about the average payoff for each dimension as well. The optimal learning strategy as predicted by the theory would be to observe averages first followed by cells. Also observing only averages or only cells can be optimal.

I find that almost 85% of the subjects choose one of the three strategies. More specifically, around 24% of chosen strategies exhibit selective learning, i.e, learning about only the averages. Also, *selective learning* leads to significantly lower payoff compared to the average.

I consider several treatments by varying the uncertainty of the prior belief keeping expected payoff same. However, the proportion of choices satisfying the optimal learning strategy is robust to the changes in uncertainty in prior. This is also consistent with the theory.

The rest of paper is organized as follows: in the next section (2), I describe the case study of selective learning by cotton farmers in India. Section 3 describes the learning problem. Section 4 solves optimal learning strategy for the DM and discusses the condition for optimality of selective learning. Section 5 considers an extension with finitely many levels of output and 6 discusses the implications for reducing the cost of the two types of learning. Section 7 discusses the laboratory experiment. Section 8 discusses the two main counterfactual policy implication. Section 9 discusses the relevant literature and section 10 concludes. All proofs and tables are in the appendix.

2 Case Study: Bt cotton in India

2.1 Data

I use data from AICRP (All India Coordinated Research Project) on Cotton field demonstration and field experiment. In 2012-13 agriculture year, there are 16 cotton research centers across ten states conducted field experiment on several varieties of cotton and also held field demonstrations to provide information to farmers in the region.

I also use NSS 2012-2013 dataset (Round 70, schedule 33) and NSS 2002-03 (Round 59, schedule 33). The schedule 33 is the "Situation Assessment Survey of Agricultural Households" that has been conducted across India in two rounds for the agricultural year 2012 -13. Round 1 is conducted in July-December 2012 and round 2 is conducted in January - June 2013. Since cotton is a *Kharif* crop, i.e, cultivated during the rainy season, in India I only consider the visit 1 of schedule 33 survey.

Considering only the farmers who produce cotton as only crop or one of the major crops there are 2334 households. For these households information about farm level variables such as expenditures on different inputs(in Rs), value of output(in Rs), quantity produced(in Kg), cultivated land(in Ha), household level variables, e.g., total land holding(in Ha), number of members, social group etc and other production related characteristics e.g., access to various sources of information, knowledge about minimum support price, outstanding loans, crop insurance and crop loss etc are considered for analysis. For a detailed discussion on the control variables, see appendix, table 6.

Also, to construct a measure for prior belief, where prior refers to pre-Bt belief I consider NSS (round 59, schedule 33). For this dataset as well the farm and household level information has been provided. Also, other factors such as access to information, crop insurance has been reported for each household.

Cotton is one of the major cash crops in India, which is also the main raw material for the textile industry. The following map shows the proportion of cotton in the total agricultural production in the agriculture year 2012-13.

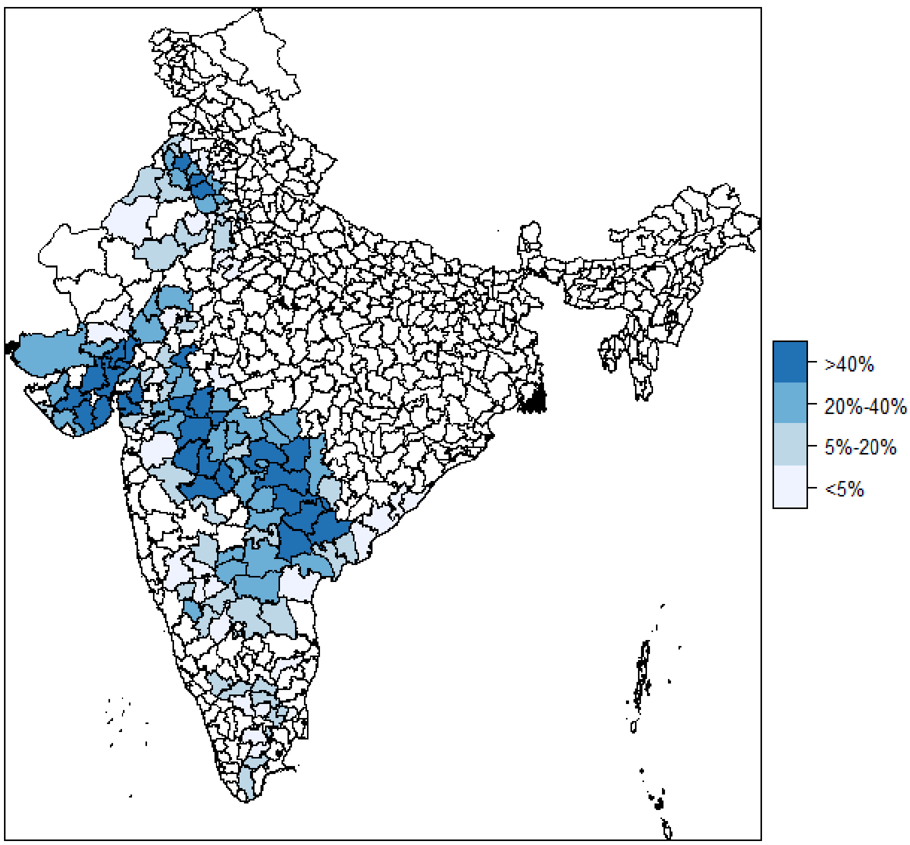


Figure 1: Share of cotton in total Agricultural Productivity

2.2 Empirical strategy

In 2002, a new type of cotton seed, Bt (*Bacilus thuringiensis*) cotton was commercially introduced in the Indian. Bt is a genetically modified pest resistant seed variety that is engineered by placing a gene from the bacterium *Bacilus thuringiensis* into the gene of the cotton to make the seed produce a pesticide that kills the larvae of the major pest of cotton, *bollworm*. The farmers have rapidly adopted this new seed variety as shown below in figure .

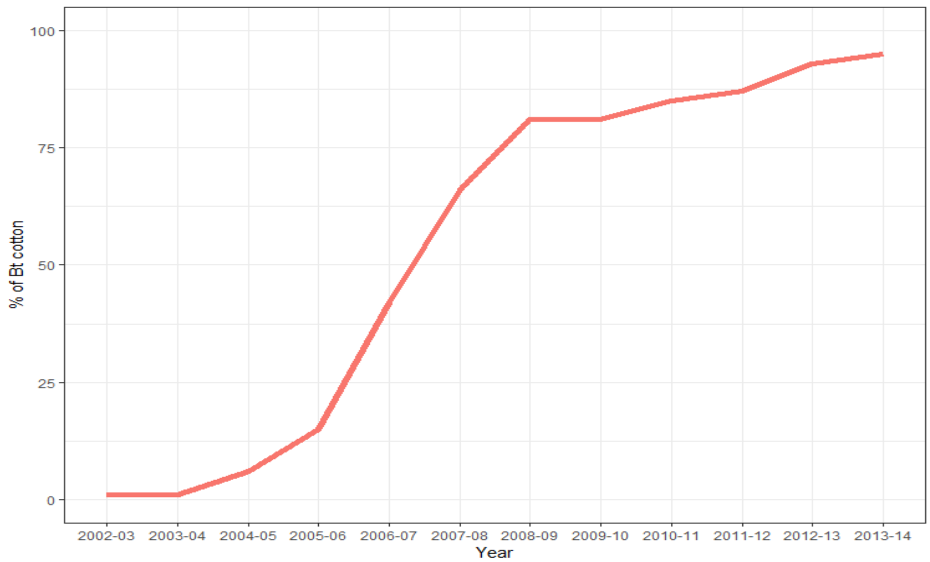


Figure 2: Rate of adoption for Bt cotton in India (Source: Choudhary and Gaur, 2014)

For this case study I will make three main observations, first, Bt cotton indeed increased the average yield of cotton, second, cotton farmers in India overused pesticide leading to a reduction in net profit in the 2012-13 survey year and third, the net profit after adopting Bt is negatively affected by prior use of pesticide use through the correlation of prior and current pesticide use.

2.2.1 Yield Difference

For each district growing cotton in 2012-13, I consider the nearest AICCIP center for the measurement of expected yield. Using the 2001-02 and 2002-03 I construct a measure of optimum expected yield (of seed cotton in Kg/Ha) as follows:

$$Exp\_yield_{2002,d} = (\max_j yield_{2002,d,j} + \max_j yield_{2001,d,j})/2$$

where $yield_{t,d,j}$ denotes the yield of seed type j under the optimal condition (as used by AICRP) in district d in year t. Thus for each district d the expected yield is the two year average of maximum yield for any variety of non-Bt cotton. Note that Bt cotton was introduced commercially in 2002.

For 2012-13 and 2013-14 I similarly construct a measure of expected yield as follows,

$$Exp\_yield_{2012,d} = (\max_j yield_{2012,d,j} + \max_j yield_{2013,d,j})/2$$

i.e, the two year average of maximum yield under the optimal condition for Bt varieties only. The difference between the two averages for each district $Exp\_yield_{2012,d} - Exp\_yield_{2002,d}$ gives the measure of yield difference between Bt and non-Bt cotton where both are cultivated under optimal condition. The map below shows the regional distribution of yield difference. A darker shade of green refers to higher yield difference (The details of the yield data is given in the appendix).

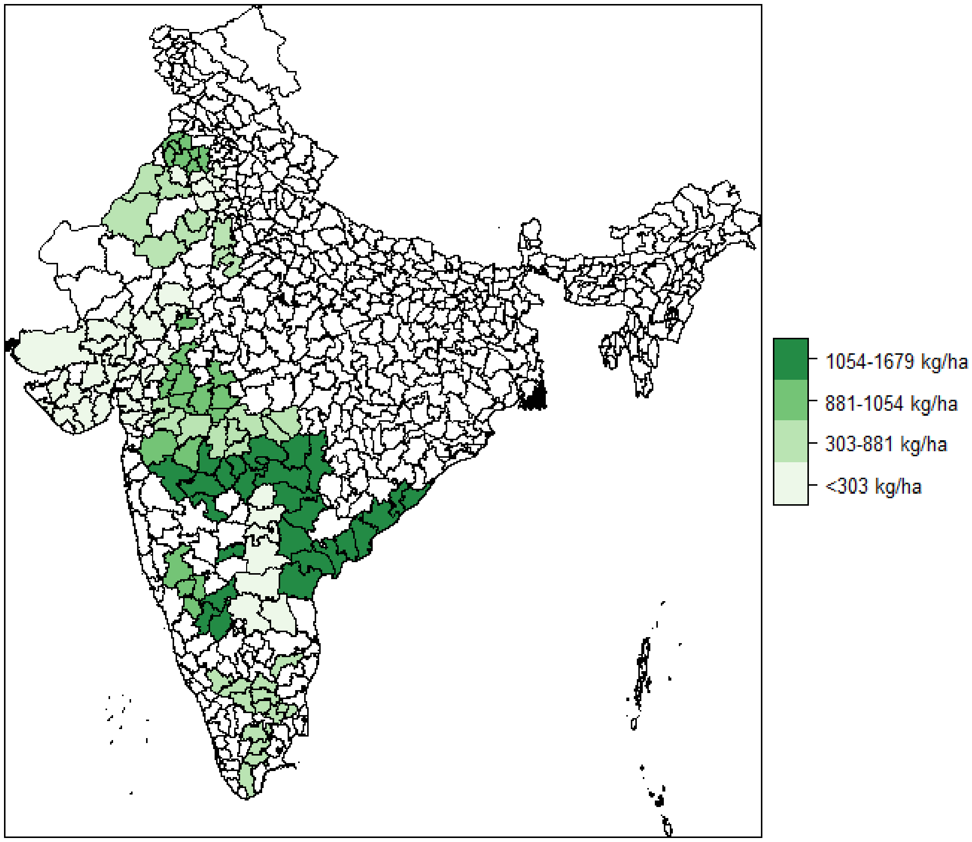


Figure 3: Difference in yield (kg/ha) across districts

2.2.2 Overuse of Pesticide

Since NSS data only contains the cost of input and value in Rs, to show that the farmers are overusing the pesticide first a cross-section regression to show the impact of pesticide use on various outcome variable using the following specification:

$$Y_{2012,j} = \beta_0 + \beta_1 Pesticide_{2012,j} + \mu_D + X_j \tag{1}$$

where $Y_{j,2012}$ refers to the outcome variable of interest, namely net profit (defined as total value-total cost), total value of output, total quantity of output and probability of crop loss. All response variables are reported per unit of land, measured in Ha. Also μ_D denotes the district fixed effect and X_j denote all other control variable for household j (ref table 6 for details).

2.2.3 Panel 1: pre and post cotton farmers

However, since table 8 uses cross-section data for 2012-13, it is possible that the farmers were using pesticide in anticipation of the risk of pest attack because cotton plants are believed to be susceptible to various types of pests. To test for that hypothesis we consider cotton farmers from two periods, before and after the introduction of Bt. I consider a difference-in-difference specification to consider whether the impact of pesticide over net profit had changed with the use of Bt seeds.

To that extent next I construct a pseudo panel data by combining the data from both NSS 59 and NSS 70 round for cotton farmers. . Using the constructed panel with cotton farmers I consider the following specification

$$Net\_Profit_j = \beta_0 + \beta_1 Pesticide_j + \beta_2 Pesticide_j * Post + \mu_D + Year + X_j + \epsilon_{djt}$$

where Net_Profit_j measures the net profit for household j either 2002 or 2012 survey year. The post variable is an indicator that the household data is from the 2012-13 agricultural year, i.e, using Bt cotton. Also, μ_D and $Year$ represent districts and year fixed effect.

2.2.4 Panel 2: Cotton producing households in *Kharif* and *Rabi* seasons

Since the constructed panel does not follow the same households across the two survey periods if some household level factors have been affecting the impact of pesticide on net profit then the earlier panel cannot address. To control for household level factors affecting the pesticide use I consider a panel data of farmers in 2012-13 who produces cotton in the rainy season and some other crop in the winter season.

In the next panel we consider the same households that produced cotton in 2012-13 agriculture year and use their pesticide use for other crops during the *Rabi* (winter) season. The "cotton" variable is a dummy for cotton produced by the household during rainy season. The difference-in-difference specification is given as follows:

$$\begin{aligned} Net\_Profit_{2012,j} = \beta_0 + \beta_1 Pesticide_{2012,j} + \beta_2 Pesticide_{2012,j} * Cotton_{2012} \\ + \mu_D + \mu_H + \mu_{Season} + X_{2012,j} + \epsilon_{djs} \end{aligned}$$

where μ_D, μ_H and μ_{Season} represents district FE, household FE and season FE respectively.

subsubsectionPanel 3: Household and time fixed effect

Finally we combine the two panel, namely cotton farmers in 2002-03 and 2012-13 and also the same cotton farmers in both *Kharif* and *Rabi* season. In case of a differential trend for some farm or household level characteristics that affects the pesticide use I consider a tripple difference specification as follows:

$$\begin{aligned} Net\_Profit_j = \beta_0 + \beta_1 Pesticide_j + \beta_2 Pesticide_j * Post + \beta_3 Pesticide_j * Cotton \\ + \beta_4 Pesticide_j * Post * Cotton + X_j + \epsilon_{djst} \end{aligned}$$

where $Pesticide_j * Post$ denotes the impact of pesticide on all crops in 2012-13 agricultural year, $Pesticide_j * Cotton$ separates the impact of pesticide by all cotton farmer in both periods and finally $Pesticide_j * Post * Cotton$ separates the impact for Bt cotton in 2012-13.

2.2.5 Role of Prior

Finally, for the third observation, I first construct a measure of prior belief of farmers in 2012-13. Since NSS schedule 33 is not a panel data, instead of considering the same household for prior belief an average use of pesticide in the same district as household j is pre-Bt period 2002-03 is used as a proxy for prior.

Using this definition of prior the impact of prior on pesticide use and net profit is measured using the following specification:

$$Net\_profit_{j,2012} = \beta_0 + \beta_1 * Prior + \beta_2 yield\_diff * X_{j,d} + \epsilon_j$$

One possible concern would be whether the districts with high prior use of pesticides have different suitability for Bt cotton since Bt cotton is a pest-resistant variety. However, in that case, the apparent dependence of net profit on prior would be due to the unobserved suitability of Bt cotton in these districts. To control for that in column 4 the yield difference for each district is added. Since yield difference measures the difference in yield for Bt and non-Bt cotton under the optimal farming condition and systemic difference related to pest resistance would be controlled for by this variable.

2.3 Results

Table 8 (in appendix) reports the impact of pesticide on net profit, total quantity, and crop loss. The variable pesticide measures the cost of pesticide per unit of land (i Ha.). The average level of pesticide use for the sample is Rs. 5856 with a standard deviation of Rs 9942. Column 1 uses the net profit per unit of land, column 2 uses the total value of output per unit of land, measured in Rs/ha and the third column uses the total quantity of output produced measured in Kg/ha. Finally, column 4 uses the probability of crop loss.

Table 8 shows that for Re. 1 increase in the cost of pesticide the net profit decreased by Rs. .874 in Ha controlling for X_j and μ_D. Even though the total value (in Rs/Ha) and total quantity (in Kg/Ha) increases but the probability of crop loss does not change with pesticide, which is consistent with the fact that Bt seed is pest resistant.

As shown in table 9 the impact of pesticide use on net profit in 2002-03 is not significantly different from zero but for Bt cotton use of pesticide reduces profitability and is similar to table 8. We can see the same pattern when instead of using net profit we use total value of production (in Rs/ha) or total quantity produced (in kg/ha) (see table 10). This implies that the impact of pesticide on net profit in post Bt period is not due to some characteristics of cotton farming.

In table 10 the impact of pesticide on the total value of output and the total quantity of output for the same constructed panel of cotton farmers. Note that pesticide use does not significantly change the total value or quantity of output.

Table 11 shows the impact of pesticide use on net profit in the houshold panel. As can be seen from the table pesticide use negative impact net profit only for cotton controlling for the household level unobseved factors whereas for other crops the impact is siginificantly positive.

As shown in table 12 use of pesticide has a significantly negative impact on net profit but this effect is due to cotton mostly since the coefficient for pesticide*post*other is statistically significantly positive. The same result holds true if we consider the total value of output per Ha instead of net profit.

This suggests that the Bt cotton farmers are most likely using more pesticide than optimal level and justifies the second claim for the case study. The diagram below shows the relationship between the pesticide use and net profit earned by the household for 2012-13 cotton farmers from NSS round 70.

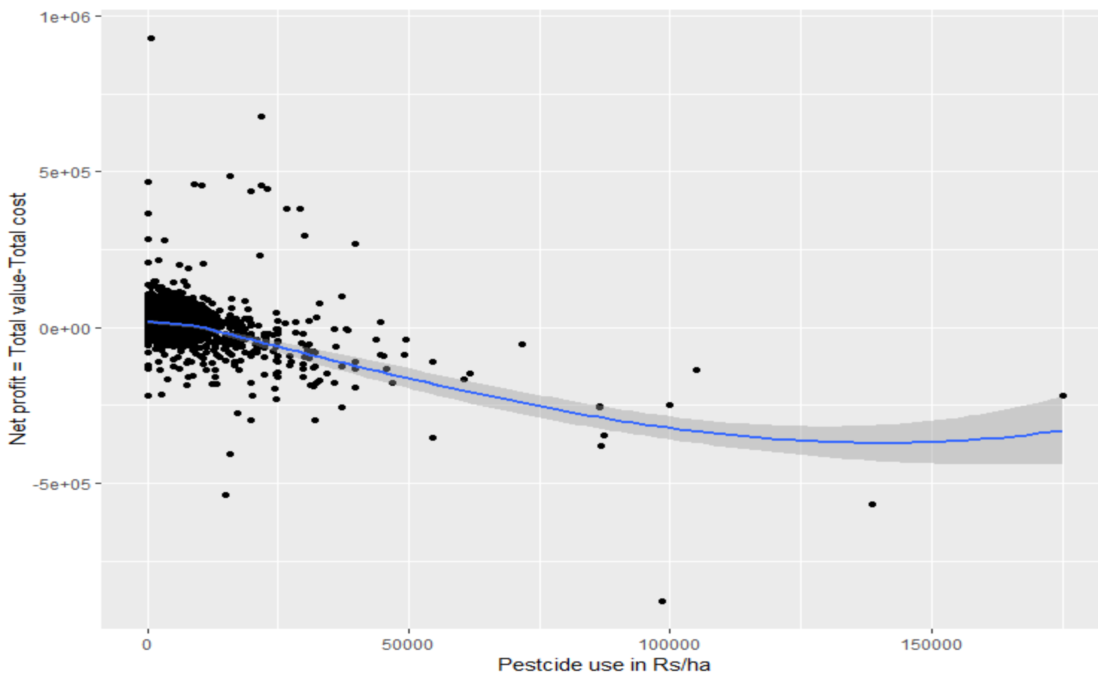


Figure 4: Pesticide use and net profit

Table 13 shows the relationship between prior use pesticide in the same district in 2002-03 and the post-Bt use of pesticide in 2012-13. The level of pesticide use is highly correlated between the time periods.

Table 14 shows the relationship between prior pesticide use and net profit in 2012-13. Furthermore, column 4 of table 14 shows that high prior in a district reduces the net profit in 2012-13 even though the relation is not significant (p-value .2). If farmers rely on their prior belief, more than what would be optimal, to decide on current pesticide use then a higher value of expected average pesticide use would lead to lower profit, due to increase in cost.

In the next section, I build a sequential learning problem for a DM that resembles the choice of the farmers in Bt cotton context. Our goal is to solve for optimal learning and check whether the case study satisfies the condition of selective learning as found in theory.

3 Model

3.1 Primitives

Consider a DM with one unit of indivisible land producing output Y with two inputs, S and P. For simplicity, output Y can take only two possible values. Each of the two inputs has finite n possible levels, namely $S = \{s_1, s_2, \ldots s_n\}$ and $P = \{p_1, p_2, \ldots p_n\}$. Let $\mathbf{A} = S \times P$ denote the set of possible comination, i.e, action/decision space.

Let $\pi : \mathbf{A} \rightarrow Y$ denote the unknown payoff function that maps each combination of inputs to the corresponding level of Y. For a typical element in $\mathbf{A}$, say (s_i, p_j) let $\pi(s_i, p_j)$ denote the payoff from using input combination (s_i, p_j). Note that, the payoff function can be represented in a $n \times n$ matrix where each cell denotes the payoff from a combination of S and P.

As mentioned in the introduction the DM would also have access to information about the marginal payoff from each input separately. The marginal distribution of payoff from S is represented by $\pi(s_i)$ and in the payoff matrix, this is shown as the final column. Similarly, the final row of $\pi(p_i)$'s denote the marginal distribution of payoff from P.

	p_1	$\ldots$	p_n
s_1	$\pi(s_1, p_1)$	$\ldots$	$\pi(s_1, p_n)$
$\ldots$	$\ldots$	$\ldots$	$\ldots$
s_n	$\pi(s_n, p_1)$	$\ldots$	$\pi(s_n, p_n)$

(a) Payoff matrix

	p_1	$\ldots$	p_n	S
s_1	$\pi(s_1, p_1)$	$\ldots$	$\pi(s_1, p_n)$	$\pi(s_1)$
$\ldots$	$\ldots$	$\ldots$	$\ldots$	$\ldots$
s_n	$\pi(s_n, p_1)$	$\ldots$	$\pi(s_n, p_n)$	$\pi(s_n)$
P	$\pi(p_1)$	$\ldots$	$\pi(p_n)$	

(b) Payoff matrix with average

For the marginal distribution let us assume a uniform distribution over the other input. This implies $\pi(s_i)$ denote the average value of payoff from row i and $\pi(p_j)$ denote the average value of payoff from column j.

Let Ω denote the state space, where each state is a realization of the payoff matrix. For an $n \times n$ matrix the state space Ω would thus contain $2^{n \times n}$ states.

The objective of the DM is to make a one-time choice of input combination to maximize his expected payoff. The DM can possible choose a mixed strategy over multiple combinations. The DM can also choose only one input in which the other input is chosen uniformly at random by nature.

3.2 Information

3.2.1 Prior

At the beginning of the choice problem the true state is chosen according data generating process $\mu^* \in \Delta(\Omega)$. The DM may or may not know the true DGP. The prior belief of the DM is given by μ_0 where $supp(\mu^*) \subset supp(\mu_0)$. This implies even though the DM may not know the true payoff state he can learn about the true state by acquiring information.

Note that, the prior belief defined over the prior, which is a $2^{n\times n}$ dimensional object. The prior can also be thought of as a distribution over possible payoff matrices. The two inputs can be correlated in any possible ways, i.e, no restrictions have been imposed on the prior. Also, each possible correlation between the two inputs or input combinations can be represented by a possible prior belief.

Some examples include: the two inputs are *substitutes* or *complements*. This can be represented by a prior belief that only puts positive probability on payoff matrices where only the cells along some diagonal generate $Y = 1$ and all other cells generate $Y = 0$. The following two tables illustrate:

	a_{21}	a_{22}	a_{23}	
a_{11}	0	1	0	1/3
a_{12}	1	0	0	1/3
a_{13}	0	0	1	1/3
	1/3	1/3	1/3	

(a) Example: Substitute/Complement I

	a_{21}	a_{22}	a_{23}	
a_{11}	0	0	1	1/3
a_{12}	0	1	0	1/3
a_{13}	1	0	0	1/3
	1/3	1/3	1/3	

(b) Example: Substitute/Complement II

3.2.2 Learning Tenchnology

The DM faces a sequential learning problem with a optimal stopping time. The DM has two possible sources of information, namely observing a cell and observing an average in the payoff matrix. The cost of observing a cell is c_l, cost of observing an average is c_a and $\delta = c_l - c_a$ denotes the different in two types of cost. There is no noise in observation, i.e, if the DM chooses a cell (s_i, p_j) he observes either 0 if $\pi(s_i, p_j) = 0$ and 1 if $\pi(s_i, p_j) = 0$. Similarly if he chooses average s_i (or p_j) and there are k many cells with payoff $Y = 1$ in row i (or column j) then he observes k/n.

Assumption 1. *No learning is costless.*

Since a cell or an average can take finitely possible values observing a cell or an average partitions the state space, i.e, set of possible payoff matrices into several blocks. For example if the DM chooses to observe cell (s_i, p_j) then the resulting partition contains two blocks, $B_{0,(i,j)}$ and $B_{1,(i,j)}$, where in block $B_{0,(i,j)}$ every payoff matrix has $\pi(s_i, s_j) = 0$ and in block $B_{0,(i,j)}$ every payoff matrix has $\pi(s_i, s_j) = 1$.

Thus observing multiple cells and averages can be expressed as a sequence of partitions of the state space. Let $\mathcal{P}$ denote a typical sequence of partition generated by observing cells and/or averages in the payoff matrix. Let $\gamma(\mathcal{P})$ denote a distribution over possible sequence of partitions. The DM can observe any number of cells or averages and in any possible sequence. So, after observing a cell or an average the DM can condition his choice of learning is next round based on the information already revealed. His learning strategy can thus be described by $\gamma(\mathcal{P})$ as a conditional sequence of partitions.

3.2.3 Updating Procedure

The DM is Bayesian, i.e, given prior belief μ and observation from a cell or an average the DM updates the belief over possible states using Bayes law. This implies in round t, i.e., after observing t cells or averages if the belief of the DM is μ_t and $\gamma(\mu_{t+1})$ denote the distribution of possible beliefs following learning strategy $\gamma(\mathcal{P})$ then by Bayes consistency,

$$\mu_t(\omega) = E(\mu_{t+1}(\omega)|\gamma_t)$$

Also, since informtion acquisition happens through partitioning the state space, Bayes updating implies if two states ω_k and ω_l are in the same block of the partition, i.e, observed cell of average have the same value in both states, then the following consistency condition holds true,

$$\frac{\mu_t(\omega_k)}{\mu_t(\omega_l)} = \frac{\mu_{t+1}(\omega_k)}{\mu_{t+1}(\omega_l)}$$

where μ_t and μ_{t+1} denote the t^{th} and $t+1^{th}$ round beliefs resp. This condition would be denoted as *partition consistency* condition.

3.2.4 Uncertainty of Belief

The optimal learning strategy would be characterised by the uncertainty of the belief of the DM. The uncertainty of belief at any round t is measured using the Shannon entropy of the distribution, namley,

$$H(\mu_t) = \sum_{\omega \in \Omega} \mu_t(\omega) \ln \mu_t(\omega)$$

Since Shannon entropy is separable in possible states, it will respect the properties of partitioning the state space. For example, suppose in round t the DM has belief t and is deciding to observe (s_i, p_j). Let $B_{0,(i,j)}$ and $B_{1,(i,j)}$ denote the blocks where $\pi(s_i, p_j)$ takes the values 0 and 1 respectively. Then the uncertainty of the belief μ_t can be written as

$$H(\mu_t) = \sum_{\omega \in B_{0,(i.j)}} \mu_t(\omega) \ln \mu_t(\omega) + \sum_{\omega \in B_{1,(i.j)}} \mu_t(\omega) \ln \mu_t(\omega)$$

Let the probability of $\pi(s_i, p_j)$ being 1 for belief μ_t was $\mu_{t,(i,j)}$. Further suppose the observed value of $\pi(s_i, p_j)$ is 0, then the uncertainty of the posterior belief μ_{t+1} can be written as

$$
\begin{aligned}
H(\mu_{t+1}) &= \sum_{\omega \in \Omega} \mu_{t+1}(\omega) \ln \mu_{t+1}(\omega) \\
&= \sum_{\omega \in B_{0,(i,j)}} \frac{\mu_t(\omega)}{1-\mu_{t,(i,j)}} \ln \frac{\mu_t(\omega)}{1-\mu_{t,(i,j)}} \\
&= \frac{1}{1-\mu_{t,(i,j)}} \left[\underbrace{\sum_{\omega \in B_{0,(i.j)}} \mu_t(\omega) \ln \mu_t(\omega)}_{\text{Residual uncertainty}} - \underbrace{(1-\mu_{t,(i,j)}) \ln 1 - \mu_{t,(i,j)}}_{\text{resolved uncertainty}} \right]
\end{aligned}
$$

Note that after observing cell (s_i, p_j) the uncertainty in belief can indeed increase if $\mu_{t,(i,j)}$ is sufficiently large and residual uncertainty in $B_{0,(i,j)}$ is sufficiently high. The residual uncertainty can also be expressed as $(1-\mu_{t,(i,j)})H(\mu_{t+1}|B_{0,(i,j)})$ where $H(\mu_{t+1}|B_{0,(i,j)})$ is the uncertainty of belief μ_{t+1} if $\pi(s_i, p_j) = 0$.

4 Decision Problem

The decision problem of the DM is as follows: the DM enters with prior belief μ_0 and decides whether to learn or not and decides to learn then which cell or average to observe. Given the observation, he updates his belief to μ_1 and again decides whether and how to learn. Similarly for every round t and belief μ_t the DM decides whether and how to learn. If he chooses to stop learning then he chooses a mixed strategy over input combinations. The payoff is realized and the DM exits.

Given the learning strategy and updating rules the DM's value function is given by,

$$
W(\mu_0) = \max_{\gamma(\mathcal{P})} E(\pi(s_i, p_j) - c(\mathcal{P})|\mu_0|\gamma, \mu_0) \tag{DP}
$$

where the DM chooses a learning strategy $\gamma(\mathcal{P})$ to maximize the expected payoff from input choices net of the cost of learning. Note that given the learning optimal learning strategy the DM would optimally choose the input combination with highest probability of generating $Y = 1$, hence the decision problem is equivalent to solving for the optimal learning strategy.

Instead of solving the $t = 0$ problem the DM's problem can be transformed to a recursive problem.

Lemma 1. *Given any prior μ_0 DM's problem (DP) can be written in recursive form as*

$$
V(\mu_t) = \max_{p \in \mathbf{P}} -c(p) + E_{\mu_t} V(\mu_{t+1})
$$

where $\mathbf{P}$ refers to all possible cells and averages inclduing no learning.

The proof is given in the appendix. The main intuition is as follows: in each round, the DM only finds a subset that contains the true state. Since intersection is commutative, the order of partitioning does not affect the Bayesian updating process. Thus given a belief μ_t the incurred cost of learning being a sunk cost the history of partitions is irrelevant for updating.

The characterization of optimal learning policy would be solved in two parts. First, given a belief μ_t the optimal learning choice would be characterized and given the solution of the recursive problem the optimal learning strategy for $t = 0$ problem would be derived.

For the recursive problem to obtain the optimal learning choice three questions need to be answered:

1. When to learn?

2. Whether to observe a cell or an average?

3. Which cell or average to observe?

The recursive problem will be solved backward. First, we will restrict to the problem where the DM can only learn about cells. Given the DM decides to learn by cells first I will find the optimal cell to observe then check the condition under which learning is optimal. Similarly, DM will be restricted to only learn about averages where I will find the optimal average to observe and find the condition under which learning is optimal. Finally, I will find the condition under which observing a cell or average or no learning is optimal when the DM has the choice of both types of learning.

4.1 Learning by Cells only

Suppose the DM has only access to information from cells. Observing a cell (s_i, p_j) partitions the state space Ω into two blocks, namely $B_{0,(i,j)}$ where $\pi(s_i, p_j) = 0$ and $B_{1,(i,j)}$ where $\pi(s_i, p_j) = 1$. Since the objective of the DM is to maxmize expected payoff, after observing $\pi(s_i, p_j) = 1$ no further learning is optimal.

In case $\pi(s_i, p_j) = 0$ however if the DM chooses not to learn anymore he does not necessarily get $Y = 0$, since $\pi(s_i, p_j) = 0$ only implies he would not choose cell (s_i, p_j) optimally. But there may exist another cell (s_k, p_l) such that when $\pi(s_i, p_j) = 0$ cell (s_k, p_l) has a positive probbaility of generating $Y = 1$ or $\pi_{t,(k,l)} > 0$, where $\pi_{t,(k,l)}$ denotes the probability of cell (s_k, p_l) to generate $Y = 1$.

The following definition formalizes this notion of expected payoff from observing only one cell and no further. The notation $\pi_{t+1|(i,j)=0}$ is used to denote the expected payoff from choosing the cell with $\pi_{t+1,(k,l)}$ when $\pi(s_i, p_j) = 0$ is observed.

Definition 1. *Suppose cell (i, j) generates $Y = 1$ with probability $\pi_{t,(i,j)}$ and in case it generates $Y = 0$, without further learning is the expected payoff is $\pi_{t+1|(i,j)=0}$ then the one-period expected payoff from cell i is given by*

$$\pi_{t,(i,j)} + (1 - \pi_{t,(i,j)})\pi_{t+1|(i,j)=0}.$$

As described in the section before the optimal learning strategy will be characterized by uncertainty of the belief as measured by Shannon entropy. However, learning is only optimal if $\pi(s_i, p_j) = 0$. The remaining uncertainty in this case is relevant for further gain in learning. The following definition formalizes the notion of rsidual uncertainty when $\pi(s_i, p_j) = 0$.

Definition 2. *If observing cell (i, j) partitions the state space into blocks $B_{1,(i,j)}$ and $B_{0,(i,j)}$ such that in B_1 all states have $\pi(s_i, p_j) = 1$ and in $B_{0,(i,j)}$ all states have $\pi(s_i, p_j) = 0$ then the information content in cell (i, j) is given by*

$$R^c_{t,(i,j)}(\mu) = \frac{\sum_{\omega \in B_{0,(i,j)}} \mu_t(\omega) \ln \mu_t(\omega)}{H(\mu_t)}.$$

Given the definition of one -period expected payoff and information content the optimal learning strategy in the recursion problem is given in the following lemma.

Lemma 2. *For any belief μ_t at any round t it is always optimal for the DM*

1. *To uncover the cell with the highest one-period expected payoff,*
2. *If more that one cell has same myopic expected payoff then uncover the cell with lowest $R^c(\mu)$,*
3. *In case the cells have same $R^c(\mu)$ as well, uncover any such cell at random.*

The proof is given in the Appendix. The main intuition is as follows: a higher value of one-period expected payoff can happen either due to a higher probability of generating $Y = 1$ or when cells are sufficiently negatively correlated, i.e, observing $\pi(s_i, p_j) = 0$ gives information about a cell (k, l) such that $\pi_{t+1,(k,l)} > \pi_{t,(i,j)}$, In both cases observing the cell with highest one-period expected payoff generates a sequence of cells where a cell with a higher probability of $Y = 1$ is observed earlier. This reduces the expected cost of learning and hence is optimal.

Given lemma, 2, the information content of all cells in a payoff matrix can be defined by the information content of the cell that would be chosen optimally if learning a cell is optimal.

Definition 3. *For any round t the informativeness of cell is given by the information content $R^c_{i,j}(\mu_{t+1})$ of cell (i, j), where cell (i, j) has the highest myopic expected payoff in round t. This will be denoted as the information content R^c_t.*

The following lemma characterizes the optimality of learning in terms of uncertainty of belief.

Lemma 3. *For any round t given expected payoff π_t, cost of learning c_l and information content of the optimal cell, learning by cell R^c_t is optimal only if the uncertainty of the belief μ_t as measured by the entropy $H(\mu_t)$ lies within the interval $[\underline{H}^c, \bar{H}^c]$.*

The proof of the lemma is given in the appendix. The main intuition is as follows: higher uncertainty has two opposing effects on the value of learning. Higher uncertainty implies higher expected cost since more cells need to be observed. If uncertainty if very low on the other hand then the gain from learning is small since very little new information can be revealed by learning. This generates the interval on uncertainty where learning is optimal.

4.2 Learning by Averages only

Let $a_{i,j}$ denote the j^{th} level of input i where $i \in \{S, P\}$ and $j \in \{1, \ldots, n\}$. There are some major differences between cells and averages. First, averages partitions the state space into potentially more than two block, at most one $n+1$ blocks $(0, 1/n, \ldots, 1)$. However, observing only averages does not always generate $Y = 1$ for sure.

The following definition formalizes the one-time expected payoff from observing an average $a_{i,j}$. Note that the expected payoff when $\pi(a_{i,j}) = k/n$ is not necessarily k/n but rather bounded below by k/n for $0 \leq k \leq n$.

Definition 4. *For any belief μ_t in round t the one-period expected payoff from uncovering any average $a_{i,j}$ is given by*

$$\sum_{p} \mu_{t,(i,j),p} \pi_{t+1,(i,j)=p}$$

where $\mu_{t,(i,j),p}$ denotes the probability of average i generating the fraction p and $\pi_{t+1,(i,j)=p}$ denotes the expected payoff from choosing the cells with highest expected payoff upon observing p from average $a_{i,j}$.

Similar to the cell problem we also define the information content of the average when no cell with $Y = 1$ is obtained by learning, since further learning optimal only in that case.

Definition 5. *For any period t let $\pi^a_{i,j}$ denote the probability of obtaining a cell with $Y = 1$ after observing average $a_{i,j}$. Then the information content of average $a_{i,j}$ is denoted by the ratio of the expected uncertainty in μ_{t+1} and uncertainty in period t,*

$$R^a_{t,(i,j)} = \frac{\sum_{\omega|Y<1} \mu_t(\omega) \ln \mu_t(\omega)}{H(\mu_t)}$$

Note that a higher value of $\pi^a_{i,j}$ would reduce the value of $R^a_{t,(i,j)}$, i.e, increase the information content for average $a_{i,j}$.

The follwing lemma finds the optimal avearge to observe in the recursive problem.

Lemma 4. *For any round t and any belief μ_t the optimal strategy for the DM is*

1. *Observer average $(a_{i,j})$ with lowest $R^a_{t,(i,j)}$, i.e, highest information content*
2. *If two averages have same $R^a_{t,(i,j)}$ then the average with highest myopic expected payoff*
3. *If $R^a_{t,(i,j)}$ for all average (i,j) then no average would be uncovered*

The proof of the lemma is given in the appendix. The main intuition is as follows: higher information content either implies a higher probability of obtaining a cell with $Y = 1$ or the reduction in uncertainty is higher in case no $Y = 1$ cell is observed. In both cases, the expected cost of learning decreases increasing the benefit of learning about $a_{i,j}$.

Similar to the cells we can define the information content of averages by the optimal average choice.

Definition 6. *Informativeness of average is defined as the value of information content* R_t^a *for average* i *that has the lowest* $R_{t,(i,j)}^a$*, i.e, the optimal average to uncover in period* t*.*

Also, similar to lemma 3 the following lemma characterizes the optimal learning strategy.

Lemma 5. *For any round* t *given expected payoff* π_t*, cost* c_a *and level of informativeness* R_a*, the DM would choose to learn if the uncertainty of the belief, namely* $H(\mu_t)$ *is within the interval,* $[\underline{H}^a, \bar{H}^a]$*.*

The proof of the lemma is given in appnedix and the main intuition is same as lemma 3.

4.3 Both types of Learning

Finally, we consider the recursive problem of the DM where he has the choice of observing both cells and averages. Note that the optimal cell and average to be chosen is given by lemma 2 and 4.

Lemma 6. *At any round* t *given expected payoff* π_t*, information content for cell and average* R^c *and* R^a *and cost difference* $\delta = c_l - c_a$ *the optimal learning strategy in round* t *is as follows: for* $\underline{H}_l \leq \underline{H}_h \leq \overline{H}_l \leq \overline{H}_h$

1. *If uncertainty is in the interval* $(\overline{H}_l, \overline{H}_h)$ *then it is optimal for the DM to uncover an average,*
2. *If uncertainty is in between* $(\underline{H}_l, \underline{H}_h)$ *then DM optimally chooses to uncover a cell in the matrix*
3. *No learning everywhere else.*

Lemma 6 shows that for a higher level of uncertainty observing average is optimal and for lower uncertainty observing cell is optimal.

The proof of the lemma is given in the appendix. The main intuition is as follows: a lower level of uncertainty implies either there are few cells to be observed or cells are more correlated. In both cases, the gain from learning for a lower uncertainty belief is higher in case of observing a cell than an average.

4.4 Full Model

Finally, we consider the $t = 0$ problem of the DM where he chooses a learning strategy to maximize expected payoff. The optimal strategy is characterized by the uncertainty of the belief. But before analyzing the optimal learning strategy we need to consider one type of belief that changes the relative information content of the cell and average when a cell is observed. The following definition of crucial cells formalizes this notion.

Definition 7. *We would characterize a cell* (s_i, p_j) *or subset of cells as crucial if one of the condition is true*

- *$H(\mu_{t+1}|B_{0,(i,j)}) > H(\mu_t)$ for some cell (s_i, p_j) in the subset, i.e, in case $Y = 0$ is observed uncertainty increases.*

- *$E(R^a_{t+1}|B_{0_{i,j}}) < R^a_t$ for some cell (s_i, p_j) in the subset, i.e„ in case $Y = 0$ the optimal average has higher information content.*

The followig theorem describes the optimal learning strategy of the DM in terms of ucertainty of the belief. Note that, no restrictions have been imposed on the value of δ, i.e., the cost difference.

Theorem 1. *If for a given belief μ if no crucial cells exists then there exists two cutoff values of $\delta = c_l - c_a \geq 0$, namely, $d\bar{e}lta, \underline{\delta} \geq 0$ given c_a such that the optimal learning strategy is as follows:*

1. *For $\delta \in [\underline{\delta}, \bar{\delta}]$ start by observing averages and then switch to observing cells.*

2. *If $\delta < \underline{\delta}$ then observe only cells.*

3. *If $\delta > \bar{\delta}$ then observe only averages.*

The proof of the theorem is given in the appendix. The proof is an application of lemma 6 in the optimal stopping problem instead of the recursive problem. Without any crucial cell being present the uncertainty reduces as the DM learns more and hence cells become more profitable. This generates a one-time switch from average to cell and not the other way round.

4.5 Selective Learning

Finally going back to the motivating case study of cotton farmers in India the following corollary gives the condition under which selective learning is optimal.

Corollary 1. *Learning only about averages not cells is optimal if*

1. *there are no crucial cells*

2. *R^a is sufficiently low, i.e, averages are sufficiently informative and*

3. *c_l, cost of uncovering cells are sufficiently high*

The proof of the corollary is given in the appendix. Note that a higher cost of learning about cells would trivially make selective learning optimal but a high c_l is not necessary to generate selective learning. If averages are sufficiently informative then the interval in which learning about average is optimal is larger and also the updating changes the belief significantly. This implies observing an average with high information content can reduce uncertainty sufficiently such that no further learning is optimal.

Selective learning does not necessarily lead to a lower expected payoff. But in case a cell with $Y = 1$ is not obtained, the selective learning strategy involves a positive probability of making an error in choice. The value function of the DM consists of the gross payoff that increases with learning and cost of learning that also increases with learning but lower the net payoff. Selective learning makes the DM better off by the lower cost of learning sacrificing gross payoff, however. This has important policy implications.

4.6 Discussion of assumptions

The main assumption in this model is that when learning about the production function the DM can either learn about the conditional productivity of an input combination or he can learn about the marginal productivity of an input separately. Observing marginal productivity may not be feasible in all possible cases of multidimensional learning, but the Bt cotton case study discussed here ensures that farmers indeed had access to information about best seed quality per se. However, the model can be applied to any general learning environment where this assumption holds true.

There is one strict assumption about the production function that has been made to simplify the analysis. The first one is that there are only two levels of output, 0 and 1. For most agricultural setting or general production function, this assumption would not hold true. In the next section, I discuss an extension of the model where output Y can take finitely many possible values. There is no restriction on how these values are arranged on $\mathbf{R}$ but the assumption of finiteness is required. The main result from lemma 6 that higher uncertainty makes averages more profitable and lower uncertainty makes cells more profitable goes through.

Another simplifying assumption is that the DM can perfectly observe the payoff from each cell or average upon paying a fixed cost. In the case study discussed here, this assumption may not be very restrictive. Farmers do not learn if they don't avail any sources of information. But most sources of information give a reliable and true estimate about productivity. However, in most learning models the DM cannot always obtain a perfect signal about the state. To address this, I consider another extension (see Appendix) where the DM obtains a noisy signal about the payoff from a cell or an average and the cost of learning increases with the precision of the signal. The main result from lemma 6 and theorem 1 still holds true.

5 Extension: Finitely many values of Y

5.1 Primitives

In the baseline model, the output Y can take only two possible values, namely 0 and 1. In this section this assumption is relaxed. Consider the same primitives and learning as the baseline model with one modification. The set of possible values of output Y can take n possible values given by $\{0, x_1, \ldots, x_{m-2}, 1\}$, where $m < \infty$ where 0 and 1 correspondingly normalize the lower and upper bound on the possible values of Y. The intermediate $n-2$ values can be distributed any possible way along the $[0,1]$ interval and are known to the DM.

Example of possible values of Y are $\{0, .5, .8, .95, 1\}$, $\{0, .25, .5, 0.75, 1\}$. In the first case the values of Y are not equidistant, henec for a given set of values of costs of learning c_l and c_a, the incentive to learn for different value might not be the same. For example if $c_l, c_a > .05$ then the DM would never to choose to learn about another cell and average when he observes a cell with $\pi(s_i, p_j) = .95$. However, in case of the second example, given the unifrom difference between possible value of Y the incentive to learn at any level remains the same.

5.2 Properties of Belief

Note that, if x_k be the smallest value of Y such that $1 - x_k < c_l$ and $1 - x_k < c_a$ then no learning is optimal after observing any payoff in $\{x_k, x_{k+1}, \dots 1\}$. This is true for any possible prior belief μ_0. However, for some beliefs if the DM knows the maximum possible value of output is $x_l < 1$ then no learning becomes optimal for even smaller values of Y. For a given value of μ_0 let $x_{\bar{k}}$ denote the minimum possible value of Y such that the DM would never learn about any other cell or average when he observe the payoff $x_{\bar{k}}$.

The one period expected payoff from observing the cell (s_i, p_j) in round t now becomes

$$\mu_{t,x_{\bar{k}}}\pi_{t,(i,j)} + (1 - \mu_{t,x_{\bar{k}}})\pi_{t+1|(i,j)<x_{\bar{k}}}$$

where $\mu_t, x_{\bar{k}}$ denote the probability that the cell (s_i, p_j) would generate a payoff higher than $x_{\bar{k}}$, $\pi_{t,(i,j)>x_{\bar{k}}}$ denote the expected payoff if the cell generate a payoff higher than $x_{\bar{k}}$ and $\pi_{t+1|(i,j)<x_{\bar{k}}}$ denote the expected payoff from choosing the cell with highest expected payoff conditional on observing $\pi(s_i, p_j) < x_{\bar{k}}$. For $m = 2$ case the highest payoff was 1 the first term in the expression only involved the probability of observing $Y = 1$, but with $m > 2$ this is not the case anymore.

The one period expected payoff from observing the average $a_{i,j}$ where $i \in \{S, P\}$ and $j \in \{1, \dots\}$ remains as,

$$\sum_p \mu_{t,(i,j),p}\pi_{t+1,(i,j)=p}$$

where $\pi_{t+1,(i,j)\geq p}$ denote the expected payoff from choosing the cell(s) with highest expected payoff given the average $a_{i,j}$ generates p. There are two opposing impact on the informativeness of average as m increases. First, with higher m the possible values of an average $a_{i,j}$ are also greater, thus averages can be informative about the level of payoff, i.e, possible maximum payoff from learning. On the other hand the same average value can be generated by mutiple combinations of x_k, which makes the averages less informative.

For the information content, note that, observing a cell now partitions the state space into m possible blocks. The relevant residual uncertainty involves only the values of $x_k < x_{\bar{k}}$. Let $B_{k,(i,j)}$ denote the block that contains all states where $\pi(s_i, p_j) = x_k$. Then the informativenes of cell (s_i, p_j) is given by

$$R^c_{i,j}(\mu) = \frac{\sum_{\omega \in \cup_{k<\bar{k}} B_{k,(i,j)}} \mu(\omega) \ln \mu(\omega)}{H(\mu)}$$

Observing an average also generates more than $n + 1$ possible partitions. Let M denote the number of possible of partitions generated by a set of possible values in $\{0, x_1, \dots, x_{m-2}, 1\}$. The information content for average $a_{i,j}$ is now given by,

$$R^a_{t,(i,j)} = \frac{\sum_{\omega|Y<x_{\bar{k}}} \mu_t(\omega) \ln \mu_t(\omega)}{H(\mu_t)}$$

For this model, first I will consider the recursive model. Note that, since the DM is Bayesian and with $m > 2$ the learning strategy is still partitioning the state space into several blocks and cost of learning is sunk at every round the recursive form of the learning problem exists (simiar to proof of lemma 1). The following lemma characterizes the learning strategy in the recursive problem in terms of uncertainty of the belief in round t,

Lemma 7. *At any round t given expected payoff π_t, information content for cell and average R^c and R^a and cost difference $\delta = c_l - c_a$ the optimal learning strategy in round t is as follows: for $\underline{H}^m_l \leq \underline{H}^m_h \leq \overline{H}^m_l \leq \overline{H}^m_h$*

1. *If uncertainty is in the interval $(\overline{H}^m_l, \overline{H}^m_h)$ then it is optimal for the DM to uncover an average,*
2. *If uncertainty is in between $(\underline{H}^m_l, \underline{H}^m_h)$ then DM optimally chooses to uncover a cell in the matrix*
3. *No learning everywhere else.*

The proof of the proposition is given in the Appendix. Even though it is similar to the proof of lemma 6 one concern is due to many possible values of Y, the same expected value or average can be generated by various combinations of such values, e.g. obtaining 1/2 for sure would have same expected payoff of obtaining 1 and 0, each half of the times. However, the uncertainty in these two scenarios would be different and the proof accounts for that.

Given lemma 7 it can be shown that the result of theorem 1 holds true even when $m > 2$.

Proposition 1. *The result of theorem 1 holds true for $m > 2$.*

Proof. Lemma 7 implies that if the DM would continue choosing cell if he starts by observing and uncertainty decreases and if he starts by observing averages then he can make a one-time switch to cells as uncertainty goes below the required cutoff.

Given μ_0 and the assumption of no crucial cell implies that the relative information content of cells and averages do not change, i.e, with learning the R^c and R^a can only reduce but averages do not become more informative compared to cells.

However, learning also changes π_t for every round. If π_t increases with learning then cells are more likely to be correlated and hence the net benefit from a cell is at least as large as that of averages. If π_t decreases with learning then either the cells are more positively correlated or more likely to generate a lower x_k since uncertainty also reduces (no crucial cells). If the DM observes cell for higher π_t then with a decrease in π_t and uncertainty of belief μ_t the cells become more informative if there is a higher correlation. A lower spread does not affect the informativeness of the average since multiple combinations of x_k can generate the same average value. Thus observing cells remain optimal. Hence, proved. □

6 Comparative Statics

In section 3, it has been shown that the optimal learning policy depends on the cost differential δ and the prior belief. The averages are more correlated, i.e, more informative about other cells or averages but may not guarantee a payoff of $Y = 1$ is observed just by themselves. On the other hand, the cells are less correlated (mechanically) and hence if the DM chooses only to observe cells the expected cost would be higher.

In this section, we will consider the impact of changing the cost of observing cell, c_l and average, c_a. In most extension program the intervention is to provide information to the farmers. In this model, this translates to reducing at least one type of cost of learning. However, the optimal policy characterization in this model would also depend on the relative change in the cost of the two types of learning instead of a standard reduction in the cost of learning.

Since the benefits of the two types of learning strategies are different, a change in the relative cost of the two types of learning can have a significant impact on the optimal learning outcome by affecting the incentive of the DM to learn. The following two corollary investigates this effect and shows that a disproportionate decrease in one type of cost of learning can indeed lower the level of learning for the DM.

Corollary 2. *For a given value of c_a, reducing the cost of learning cells, c_l can increase the probability of error, i.e, probability of choosing an input combination generating $Y = 0$ and hence reduce the expected gross payoff.*

Corollary 3. *For a given value of c_l, reducing the cost of learning averages, c_a can increase the probability of error, i.e, probability of choosing an input combination generating $Y = 0$ and hence reduce the expected gross payoff.*

The proofs are given in the appendix. The main intuition is as follows: reduction in cost can affect the net payoff in two ways, either the lower cost leads to more learning for the same type increasing the gross payoff or by reducing the cost of learning due to lower cost. However, the two types of cost need not be a complement to each other. If both types are costs are sufficiently and one cost is sufficiently reduced then the DM may want to substitute one type of learning with another if it sufficiently reduces the cost. However, this substitution can reduce the gross payoff since the two types of learning choices are differently informative about the payoff function.

Since any reduction in cost increases the net benefit an optimal strategy needs to ensure that the precision of posterior, i.e. the gross expected payoff does not reduce for any fixed budget for cost reduction. To illustrate, if before the cost reduction the DM has been using both types of information reduction in both costs would preserve the complementarity between the two types of learning whereas the reduction in only one type of cost disproportionately can make the two types substitutable.

In the latter case the impact of cost reduction policy works through only the cost of learning channel and not through the gross payoff. Moreover, the two effects are also going against each other, reducing the possible impact of the cost reduction policy. However, the former strategy would affect the net payoff through both the gross payoff and cost of learning channel.

Moreover, the optimal policy would always be dependent on the prior belief, more precisely, on the relative information content of the averages and cell in all possible rounds.

6.1 Social Learning

In the last section, we showed that lowering one type of cost of learning disproportionately can lead to higher probability of errors. This effect becomes even more severe when DMs learn socially from each other. Since the error of one farmer can actually increase the probability of error for other DMs who learn from him.

Let us consider two groups of DMs in the economy, namely *leaders* and *followers*. The group of leaders does not have access to social learning and can only learn by uncovering cells and averages. Each member from the group of followers can choose to learn socially from at most one DM from the group of leaders. Let N be the total number of DM including both groups and $\overline{Y}$ be the average payoff for the entire economy[1].

Let us further assume that the economy is divided into several neighborhoods where social learning is only possible within a neighborhood. All DMs in the same neighborhood has the same prior belief over possible states but not necessarily have the same true state [2]. Any leader in a neighborhood enters the economy at the beginning of $t = 0$, solves the decision problem and leaves at the end of $t = 0$. Any follower in a neighborhood enters the economy at the beginning of $t = 1$, has a choice to observe the chosen input combination for at most one leader that he chooses to observe at no cost, then solves the decision problem and leaves the economy at the end of period $t = 1$.

For a given DM i in period $t = 1$ and neighborhood n_1 let $\nu_{i,j}$ denote the probability with which a DM j in the leader group would have same state as i. We assume that $\nu_{i,j}$ is known for all leaders form period $t = 0$ for DM i. Since the objective of the DM i is to maximize expected payoff subject to cost of learning he would always choose to observe the DM j from $t = 0$ who has the highest $\nu_{i,j}$ over all j.

The leaders do not have any altruistic motives towards the followers and hence for both groups of DM the objective is to maximize his own expected payoff given his belief and costs of learning. The cost of learning is the same for all DMs in the same neighborhood. [3].

In this section, I want to explore the effect of changing the cost of learning on $\overline{Y}$ when social

[1] I have abstracted away from the price of the output in the calculation of payoff, however, an economy-wide increase in the probability of error would generate general equilibrium effect through price. In this paper, I will not consider the GE effect and hence I would not be able to do complete welfare analysis and will only focus on productivity assuming prices as given. This assumption would be justified if the price of the output is determined in a global market where the economy in question only contributes to a smaller extent.

[2] This would be true if the payoff functions for all DMs in a neighborhood is generated by the same data generating process which is known to the DMs

[3] An alternative assumption would be the DMs have heterogeneous costs of learning but the cost of learning for the leader is known to the follower

learning is feasible. Any policy experimentation would decrease the cost of learning for both groups of DMs. The following proposition shows the impact of such a policy of reducing the cost of one type of learning over the average yield in the economy.

Proposition 2. *If only one of the two costs of learning is reduced significantly then the average payoff* $\overline{Y}$ *can decrease even though the DMs enjoys higher net payoff.*

In this model, we have assumed away the mechanism through which the producers learn socially and divided the producers into groups of leaders or followers. In the literature as well in the case study there are multiple ways DMs learn socially. For the agriculture sector in developing countries in many cases, the farmers learn from earlier generation but with a change in technology that is less likely. For any new technology, it is usually the case that a group of farmers adopts the technology earlier than others and all other farmers learn from the experience of the early adopters.

7 Experiment

In this section, I want to test the theoretical prediction in a laboratory setting. The goal of this experiment is to test whether, in a multi-dimensional learning problem where the subjects have the option of learning about both cells and averages whether they choose the optimal strategy, namely, they observe the averages first and then make a one-time switch to observing cells before choosing an action.

Also, the subjects can choose to learn about only averages or only cells. In case the subjects choose only to learn about averages, i.e, a *selective learning* strategy I want to test whether the payoff is indeed lower than the average over all possible strategy. This would show that selective learning is payoff relevant for subjects, a similar result found in the Bt cotton example discussed in section 2.

In the experiment, the subjects face a learning problem with two dimensions. The environment replicates the learning technology described in the section 3. The subjects face a sequential learning problem over the cells and averages and can choose to stop to go the choice task whenever they want. No learning at every stage is possible. The experiment records the learning choices, action choices conditional on learning and the payoff for each subject.

Given the laboratory setting the prior is the same as the true data generating process and known to both the subjects and the experiment. The realization of the true payoff matrix is also observed by the experimenter, which implies the mistake can be quantified in terms of payoff. The main treatment considered for the experiment is the different priors. However, none of the priors contain any *crucial condition* (a necessary condition for the theoretical result). This implies the optimal strategy should not be affected by the change in the uncertainty of the prior.

One important feature of the theoretical model is the DM can choose a mixed strategy to maximize his payoff. To replicate the similar choice set for the subjects they are asked to choose a combination of more than one time. Since there are in total of nine possible colors and shapes I allow them to choose nine times in total. It has been communicated to the subjects that they can choose the exact same combination all nine times. Also, in the introduction with a simple 2×2 example, I explain how choosing multiple combinations they can generate a probability distribution over choices without explicitly stating that this is a mixed strategy.

7.1 Experiment Design

I have used the otree software to design my experiment. Each round of the experiment consists of two tasks, namely a learning task, followed by a choice task. The subject can play any number of rounds in 20 minutes. The total payoff is cumulative, i.e, the sum of payoff from all rounds.

7.1.1 Learning Task

For the learning task, the subject is shown a 9×9 matrix where each row represents a shape and each column represent a color. Each cell denotes a shape-color combination and is covered initially(see Appendix for screenshot). The subjects can open any shape-color combination or any shape or color separately from the matrix. A color-shape combination would be referred to as a cell and a single color of shape would be referred as average for the rest of the section.

By opening a cell the subjects can learn whether the payoff from the cell is 1 or 0 and by opening an average they can learn how many ones are there in the chosen row or column. However, learning is not free for the subject. After opening a cell or an average the subjects need to solve a problem to learn about the value in the corresponding cell or average.

If subjects choose to open a cell he observes a numerical problem, e.g.,

$$\text{Solve: } 0.15235034 + 0.24505243 - 0.14855077$$

Additionally, they are also given with a cutoff value, say 0.248096419675. If the solution of the problem is greater than the cutoff value then the cell gives payoff of 1, i.e, equivalent to \$1 and 0 otherwise. Note that, there is no noise in the information provided, i.e if the solution of the cell problem is above the cutoff the payoff is 1 for sure.

If subjects choose to observe an average they observe a numerical problem similar to:

The seventh digit represents the number of ones in the row: .2688798354

The subjects can open any number of cells and averages and any number of times. he chooses to solve the problem he can store his solution and the value becomes available when he makes the decision for the next round of learning. However, the subject is not given ant feedback on whether his solution is correct or not. The subject decides when to finish the learning task and move on to the choice task.

Additionally, at each stage, the subjects can choose to not learn at all. For example, the first time the subjects see the 9×9 matrix they have an option to pass to go directly to the choice task. Also after each observation, they have to choose between continuing to learn and move on to the choice task.

7.1.2 Choice Task

Once the subjects choose to move to the choice task they need to make nine choices. Each choice is a shape-color combination. However, there is no restriction on how the subjects can choose this shape-color combination. For example, he can choose the same combination all nine times, nine different combinations, three combinations each three times and so on.

The subject can also choose the average row or the average column. However, if he chooses the average row and the average column together then the resulting payoff is 0 for sure and he is given a warning.

Given the subject's shape-color combination choices one shape-color from the list of nine combinations is chosen at random with uniform probability. The payoff of the subject is the payoff of the cell thus chosen. If the subject has chosen an average instead then one of the combinations from the chosen row or column would be selected at random for the payoff.

Having nine choices serves two purposes for the experiment. First, this gives the subjects an easy way to randomize over combinations. Nine choices ensure the subject can randomize over the entire 9×9 matrix. Second, the mixed strategy also generates the posterior belief of the subject.

7.1.3 Prior

Three possible matrices were chosen for the experiment, as described below. Note that, in each case, the cells with 1 are chosen randomly using uniform distribution and subjects were informed about this.

i. exactly 72 cells generating 0 and rest 9 cells generate 1

ii. exactly 54 cells generating 0 and rest 27 cells generate 1

iii. exactly 36 cells generating 0 and rest 45 cells generate 1

All prior beliefs were generated using these possibilities. The subjects were given the true data generating processes as their prior and the realized state is known to the experimenter. Combined with this the fact that the subjects make multiple choices the experiment generates state-dependent stochastic choice data.

There are three possible priors considered, each with the same expected payoff but varying degree of uncertainty. Below are the three prior is descending order of uncertainty:

1. All three possibilities have equal probability, namely 1/3 each

2. With 50% chance there are 54 zeros and each of the other two possibilities have 25%

3. With 50% chance there are 72 zeros and with 50% chance there are 36 zeros

7.2 Results

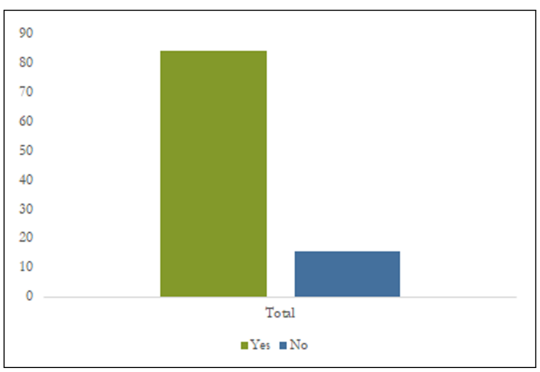
I have run 4 sessions with 35 subjects in the NYU CESS lab with NYU undergraduate students. On average in 20 minutes each subject played $6-7$ rounds. Excluding the first round data and pooling subjects together we have 192 observations, i.e, 192 rounds of the experiment.

In the introduction, before the subjects started the main experiment the possibilities and types of problems were explained to the subject. To ensure that the subjects understood the task a quiz consisting of ten questions were administered. The subjects who scored below 4 out of 10 in the quiz were to be not allowed to participate in the main experiment. All my subjects were able to finish the quiz to take the main experiment.

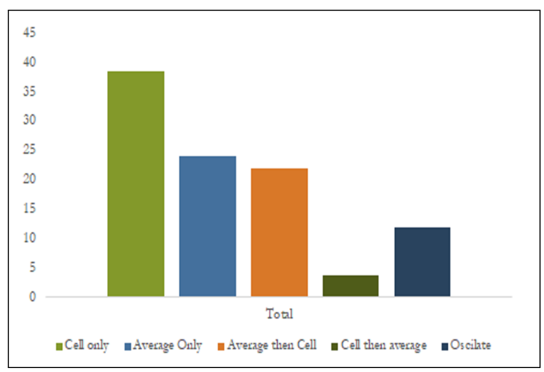
Optimal Strategy: There are only three strategies that can be chosen optimally by the subjects, namely, i. observe only cells, ii. only averages and iii. start with average and switch to cell permanently.

Violation: A violation of the theoretical prediction happens when the subject chooses to i. observe cells first and a permanent switch to average, ii. oscillate between observing cells and averages.

The first two graphs below show the percentage of rounds (pooling across subjects) that satisfy theoretical prediction.



(a) Percentage of rounds with no violation



(b) Chosen learning strtegies in percentage

The left panel shows the observed behavior in around 84% of the round did not violate the theoretical prediction. The right panel shows the same data when broken down into chosen strategies. The first three strategies are optimal and the last two are suboptimal. All differences in the proportion of rounds with any one type of optimal strategy against any one type of suboptimal strategy are statistically significant (See appendix table 16*a*). Note that, significantly more rounds involve choosing averages then a one-time switch to cell compared to the two suboptimal strategies.

Also, in around 24% rounds the subjects uses a strategy is to observe only averages. This is the *selective learning* strategy described in the model. The average payoff from this strategy is given by 0.4265. In comparison, the expected payoff from all strategies combined is 0.6094 for a round in the data, which is statistically significantly higher. Thus selective learning is payoff relevant to the subjects. The same result holds true if broken down across different prior beliefs.

Figure 10 shows the pattern of optimal strategy choice when broken down by prior. Across all priors, the proportion of rounds that satisfy the theoretical prediction is above 80%. However, between prior, there is not much of a variation even though with more uncertainty the subjects are more likely to choose an optimal strategy.

Figure 11 further breaks down the figure 10 into all possible strategy choices for all priors. There are significantly more rounds where subjects choose one of the optimal strategies compared to any suboptimal strategies (the detailed t-stats are given in the appendix).

Thus I show that the theoretical prediction of optimal learning mechanism is validated in the experiment. A significant proportion of choices exhibit selective learning and is indeed payoff relevant.

8 Policy Implication

The objective of this paper is to understand an imperfect learning behavior of cotton farmers in India. The model and the following analysis if optimal strategy helps us to understand the learning mechanism that can lead to the observed type of imperfect learning behavior. However, the analysis would remain incomplete without discussing the policy implications in such a learning environment.

In this section, I will consider two specific policy implications of the model. The first one is true for any costly learning problem where the DM needs to choose *what* to learn about before deciding how much to learn about. Since in this model the DM has a choice between marginal and conditional productivity to learn about, the question of *what* to learn about becomes relevant.

The other policy is specific to the multidimensional structure of the model. Given assumptions about the payoff functions (namely there are finitely many levels of inputs), I can characterize the type of payoff function that would lead to selective learning with a lower level of precision. If we consider the policymaker has a choice between providing different types of information, namely, conditional or marginal productivity and providing two types of policies has different costs attach to it then characterizing these payoff functions is informative about the extension policy. The extension program can target the type of technological change where providing only marginal information does not benefit the farmers and only conditional information would increase payoff for all types of farmers or the type of technology for which providing any type of information is equally effective and so on.

8.1 Policy 1: Demand-driven Extension

After the success of the green revolution as the gains from extension programs started to wane a demand-driven approach to extension program replaced the existing training-and-visit (T&V) approach to extension program. After the pilot study between 1995 and 2003 across 28 districts in India, the new approach to extension services in the form of a central project ATMA (Agricultural Technology Management Agency) has been introduced in 2005. The major change in the approach of this new extension program is that, instead of training the farmers what is the best way of producing a crop the new program allowed the input of the farmers on what information to be provided to them.

Even after more than 10 years of implementation, the impact of ATMA has been a mixed experience (see Glendenning and Babu 2011 for a survey of the performance of ATMA after 5 years of implementation). Even though it may be true that a demand-driven approach that takes input from farmers would be more approachable for the farmers and would increase the success rate of extension through a higher rate of penetration but depending on how farmers learn there can be severe drawbacks of this type of policy.

In the context of this model a demand-driven extension policy like ATMA can reduce the cost of learning for either or both marginal and conditional productivity depending on the learning constraints the farmers have been facing. For example, if the extension service enables the farmer to ask questions about specific topics and learn from a server or an agriculture expert using telecommunication technology (like mobile phone connections) then depending on the kind of questions the farmer can ask would determine the type of information becomes cheaper for the farmers. This leads to the following result that uses the observations from the comparative statics result.

Result 1. *An introduction to demand-driven extension policies can reduce the rate of learning for farmers when the cost of learning prior to the intervention was sufficiently high.*

This result follows from corollary 2 and 3 jointly. If initially, the cost of learning is too high such that the DM chooses both types of learning because learning about one type of information was not sufficiently informative given the cost and after the reduction of cost the farmer switches to only marginal or conditional learning then a reduction in cost can lead to a lower level of learning. The mechanism is as follows: a lower cost of learning for any type of learning would imply the DM would want to learn more by using that type of information. If the obtained information from one source is sufficient then the DM would not use other types of learning. This would lead to lower cost of learning but also lower gross payoff since the level of learning goes down.

8.2 Policy 2: Categorization of Technology

To categorize the type of technology that would be more suitable for selective learning and/or lead to lower precision of learning let us start by considering a 3×3 model of payoff function, i.e, there are 3 levels of S and P and the output Y takes two possible values 0 and 1.

First, we will categorize the production technology based on the expected payoff given prior belief μ_0. For simplicity, we can do so by considering the possible number of 0s present in the payoff matrix. Given a level of expected payoff, we can further classify the production functions by the correlation between two inputs as well, which will determine the impact of selective learning. For illustration purpose suppose the cost of learning about a cell $c_l < 0.33$, i.e, if the DM knows that there are exactly 2 out 3 cells in a row or column in a payoff matrix then under a uniform belief (no further information about the payoff matrix) it is profitable for the DM to learn about one more cell.

Good States : In a 3×3 payoff matrix if there exists at least one row or column that constitutes of only $Y = 1$ then the payoff function would belong to the set of good states. This can happen if either there are at least 7 cells with $Y = 1$ or there are exactly 6 cells with $Y = 1$ but not all rows or column give the same expected payoff. For these production function, the DM can perfectly learn about a choice of a cell with $Y = 1$ just by observing averages, expected payoff loss due to selective learning is determined by the minimum of the two types of costs.

Bad States : In a 3×3 payoff matrix if there exists at least one row or column that constitutes of only $Y = 0$ then the payoff function would belong to the set of bad states. This can happen if there are at most 2 cell with $Y = 1$ or 3 cells with $Y = 1$ but not all cells or columns gave the same payoff. In this case, however, observing only averages is not perfectly informative about a cell with $Y = 1$ but given $c_l < .33$ the expected loss is bounded by the lowest of the two types of costs.

Medium states with change in correlation : Thus selective learning would only be payoff relevant, i.e, the loss would not be bounded by the lowest of the two costs (the DM can learn better by observing only cells) if the payoff function is neither good or bad state, which will be denoted as medium state. However, not every type of payoff function would lead to payoff relevant selective learning. The following matrix illustrates,

	p_1	p_2	p_3	S
s_1	0	0	1	1/3
s_2	0	0	1	1/3
s_3	1	1	0	2/3
P	1/3	1/3	2/3	

Table 3: Selective learning I: true payoff function

In this example the payoff loss due to selective learning is not bounded by the lowest of the two costs. Even if cost of observing averages goes to zero, $c_a = 0$ and the DM observes all the averages he cannot perfectly learn about the true production function. For uniform prior belief over all cell upon observing the posterior belief of obtaining $Y = 1$ for each cell would be as follows: If $c_l > 0.2$

	p_1	p_2	p_3	S
s_1	0.2	0.2	0.6	1/3
s_2	0.2	0.2	0.6	1/3
s_3	0.6	0.6	0.8	2/3
P	1/3	1/3	2/3	

Table 4: Selective learning II: posterior belief

no further learning is optimal here, the DM would choose (s_3, p_3) resulting in $Y = 0$ payoff. In this case the expected loss due to selective learning is not bounded by the lowest of the two costs.

There are two main features of the production function in the earlier example, one, there exists one row and one column that generates strictly higher expected payoff compared to other levels of the same input and two, the correlation between the two input changes for the highest average payoff level of inputs. This feature can be generalized as well using corollary 1. The row or the column with strictly higher average payoff ensures that sufficient uncertainty is reduced but two facts, namely, not observing $Y = 1$ for sure and correlation can change for the possible states belonging to the same block that is generated by observing the averages ensure that the residual uncertainty is not zero, i.e., selective learning is payoff optimal. If the correlation between the two inputs were to remain constant then no residual uncertainty is left which is equivalent to knowing a cell with $Y = 1$.

For more than two levels of values of Y, a similar result can be obtained as shown in the example below. Suppose Y can take four possible values $Y \in \{0, 0.25, 0.5, 0.75, 1\}$. Suppose pesticide can take four possible values $\{0, 1, 2, 3\}$ corresponding to *low*, *medium*, *high* and *very high* level of use. I will compare the expected payoff from using the best variety of existing seed, say Non_bt seed to a new seed variety, namely Bt seeds.

Panel 6*a*, the payoff for the example is plotted. The average payoff from Non_bt is .5 and the average payoff from Bt is .625, thus learning about only average would make Bt a better choice. The best average level of pesticide is 2 but the combination of Bt seed with pesticide level 2 generates $Y = 0.5$ which is half the maximum possible value of Y that can be obtained.

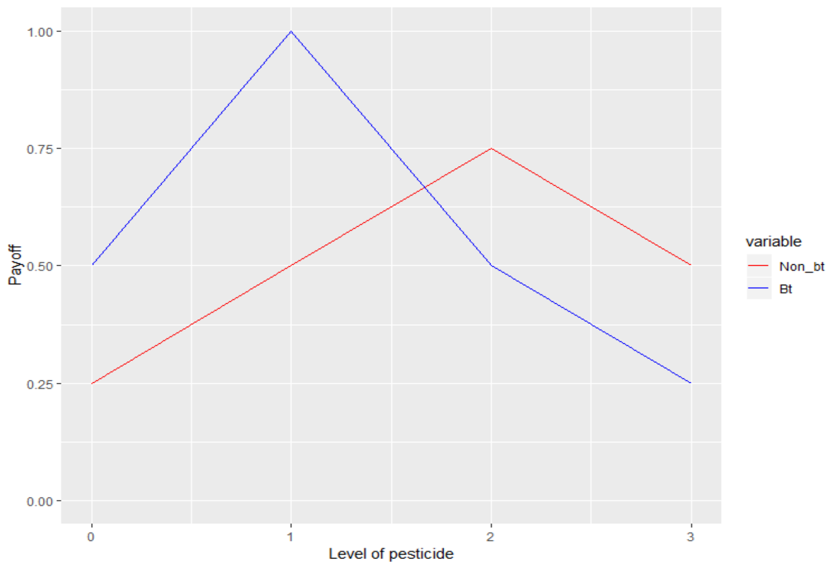
Panel 6*b*, I have plotted the net profit and pesticide use for the two rounds 2002 and 2012 of data for cotton farmers. Note that, 2012 is a proxy for Bt cotton use and 2002 is a proxy for Non-Bt use. Note that, the pattern in the data is same as the pattern in the example, i.e, Bt generates higher average payoff but the optimal level of pesticide conditional on Bt shifts to the left.

In comparison, if the technology is such that the average level remains the same but the gain is by changing the input combinations then selective learning would not be optimal, since learning about averages would not be sufficiently informative. Such an example is *Sahbhagi Dhan* rice variety in India. It is a drought-tolerant rice variety that only generates higher payoff for a dryer and hotter condition. Thus the yield is lower in a normal year the yield is lower (see Anantha et al 2015) but higher in the drought year. This model predicts selective learning cannot be an optimal strategy for this technology since the payoff gain can only be achieved by learning about the interaction between the seed and level of irrigation.

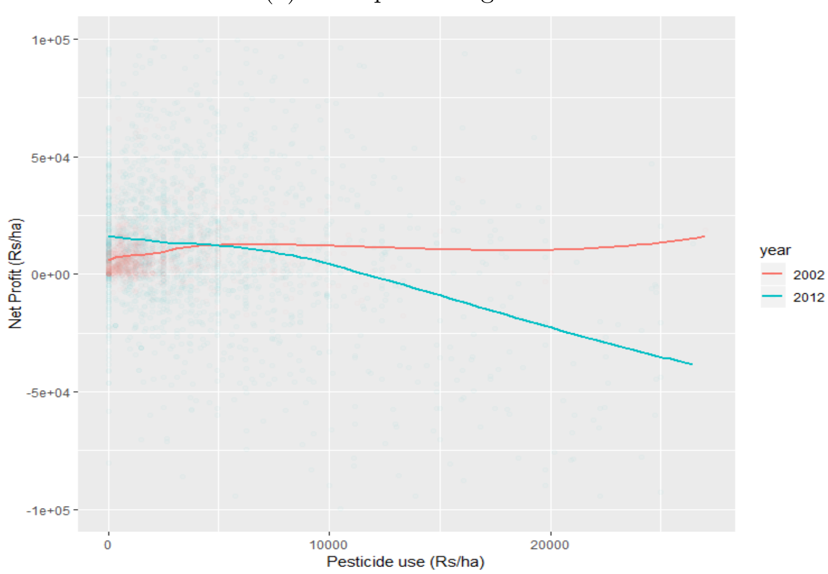
9 Literature

Since low productivity in agriculture is a major concern in most developing countries understanding the reason for the sub-optimal use of inputs and/or technology has been an important research question in development economics. The main reasons for the sub-optimal use of inputs as discussed in the literature are: financial constraints[4], supply-side constraints and learning constraints(Duflo et al. (2008),Duflo et al. (2011), RIRDC report 2003). See De Janvry et al. (2017) for a survey of field experiment targeted to affect the sub-optimal choice of farmers in developing countries.

[4] Lambrecht et al. (2014), Karlan et al. (2014)



(a) Example: change in correlation



(b) Bt cotton: Net profit and pesticide use

Constraints to learning can take many forms starting from lack of information or costly experimentation/processing or cognitive limitation faced by the farmers. If the information is costly to acquire or requires some costly experimentation the farmers may optimally choose not to learn about the optimal practices. One of the major focus of agricultural extension services thus have been to successfully transmit information to the farmers (Kondylis et al. (2017), Meijer et al. (2015) and Glendenning et al. (2010) for Indian context).

To mitigate learning constraints farmers often learn from the experience of their neighbors. Social learning can change adoption/optimal choice behavior for new farmer based on the outcome of the more experienced farmers [5]. However, social learning mechanisms are often not effective in the case of agriculture due to the idiosyncratic nature of the agricultural production function.

Constraint to learning can be a result of the cognitive limitation of the farmers. As discussed in Mullainathan and Shafir (2013) the poorer farmers face a stricter cognitive constraint as a result of poverty. Also especially in India where the average rate of literacy is around 80% a significant proportion of farmers do not have the necessary education to systematically learn via experiments or from other formal resources (govt websites, newspaper etc.).

In terms of research question, this paper is closest to Schwartzstein (2014) where the author introduces the concept of selective attention, i.e., paying attention to only input in a production process. The first major difference is that in Schwarzstein 2014 the objective of the DM is to correctly predict the state rather than maximizing the expected payoff. The second major difference is that the DM in Schwarzstein 2014 does not choose whether to learn about both inputs or not since the cognitive constraint he faces is exogenously given. This leads to different policy implications. For example, reducing the cost of learning in Schwarzstein 2014 can never increase the probability of mistake.

The main mechanism of the decision problem in this paper is in the strand of recent attention literature. In the *Rational Inattention* literature the DM faces a decision problem where he has the choice of learning which is costly. The cost of learning/attention is treated as a cognitive cost that affects the net utility of the DM. Several types of deviations of behavior from rationality have been explained as a result of the cognitive cost restricting the possible choices of the DM, which is precisely the same approach I follow in this paper. See Sims (2003), Matějka and McKay (2015), Caplin and Dean (2015), Caplin et al. (2017),Woodford (2014), Hébert and Woodford (2017), Steiner et al. (2017), Matějka and Tabellini (2017), Gabaix (2014).

Apart from Matějka and Tabellini (2017) no other paper consider the multidimensional nature of the decision problem which is the key focus of this paper. In Matejka and Tabellini 2015 even though there are multiple dimensions, they are additively combined to derive the final payoff. We do not require any such structure of payoff function.

Also, even though the DM in this model faces a one-time decision problem the recursive approach of the problem is similar to Woodford (2014) and Hébert and Woodford (2017). However, I emphasize the decision what or how to learn also along with when to stop learning.

[5] Conley and Udry (2001), Beaman et al. (2014a), Cai et al. (2015)

Besides the rational inattention literature that mainly focuses on deviations of behavior from a rational standard other literature on financial decision making has discussed the role multiple dimensions of a decision-making problem. See Van Nieuwerburgh and Veldkamp (2009), Van Nieuwerburgh and Veldkamp (2010), Mondria (2010), Peng and Xiong (2006) for the portfolio choice problem where the DM also needs to learn about the payoff from several assets. Every model in this literature assumes some specific form of relationship between the assets (in terms of correlation) and special cost functions for learning. No such assumption on prior is made in this paper.

The feature of the trade-off between cell and average learning strategy in this model is however similar to *coarse categorization* models of prediction (See Fryer and Jackson (2008), Mullainathan et al. (2008), Mohlin (2014) etc). In this models, the DM wants to predict an outcome variable based on a set of input variables. The DM partitions the set of possible input variables and predicts the outcome variable for the entire block in the partition.

There are two major differences between the coarse categorization/partition models and our model. First, Instead of choosing a finer or coarser partition as in the coarse categorization models, the DM in this paper is choosing between two types of partition, one based on unconditional learning and the other on conditional learning which generates the depth-breadth trade-off.

Second, in all these categorization models the objective of the DM is to predict the outcome but in our model, the DM wants to maximize the payoff and learning about the relationship between input and output variable is instrumental in maximizing payoff. As noted by Van Nieuwerburgh and Veldkamp (2010) in an otherwise similar environment the learning strategy would be very different if the DM is maximizing payoffs instead of predicting only.

Some recent papers have explored the implications of multi-dimensional learning in a variety of context. In Richter (2017) the DM wants to choose an object out of n options and each object has n characteristics. The objective of the DM is to choose the object with the highest sum of attribute values. They find that DM would do a breadth search, learn about one characteristic for multiple objects when n is small and the DM would do a depth search, learn about one object but all characteristics when n is large. The major difference in my paper is that the DM maximizes the payoff from one combination and not the sum across a row, this reduces the incentive to learn overall but increases the payoff from learning about average (equivalent to single characteristics breadth search).

Liang et al. (2017) considers a model of dynamic information acquisition from many correlated information sources. They find for general information structure after a finite period the optimal strategy is myopic, i.e, to choose the signal that maximally reduces uncertainty. In this paper I get a similar result regarding reducing maximum uncertainty being a component of optimal strategy but the signal structure here is significantly different since the correlation in signals, i.e, observations of cells and averages is implicit in the prior, which makes it easier for the DM to consider the benefit from using the correlation. In fact, the equivalent myopic strategy in my model would incorporate the correlation between observation from averages and cells.

Leme and Schneider (2018) studies a problem of contextual search, which is a multidimensional search problem in contextual decision-making. The objective of the DM is to find the location of a point in a n dimensional object by partitioning the object into two blocks. They find an algorithm that divides the object in either half lengthwise (in n-dimension) or divides the object in half measured by an intrinsic volume (in $n-1$ dimension, e.g. half area wise for a regular $2-D$ object) significantly improves the speed of convergence over the algorithm that only considers partitioning in n dimensional space. The role of average in this model is similar to a lower dimensional partitioning in the current model. However, since the DM in my model needs to find only one combination and there can be more than one optimal combination, the incentive to learn is lower in my model.

10 Conclusion

In this paper, I have built a model of multi-dimensional learning in a production environment where the payoff-maximizing producer has an additional access to information about the marginal productivity of each input besides the information about the productivity of each input combination. In a sequential learning problem with optimal stopping choice when both types of learning are costly the optimal policy is to start by learning about averages then permanently switch to learning about cells as the uncertainty of the belief over production function, as measured by Shannon entropy, goes below a threshold.

The optimal cell to observe is the one that maximizes the payoff from one-period learning only and the optimal average to observe is the one that reduces the uncertainty of belief the most. Moreover, observing only cell or only average can also be optimal for some prior and cost difference. More specifically if information content on average is significantly high then observing only averages, i.e, selectively learning about an input becomes optimal. However, selective learning can significantly reduce the payoff of the DM even though the net benefit increases due to lower cost.

In a case study with cotton farmers in India, I show evidence of selective learning and justify this behavior in light of the model. This has significant policy implications for agricultural extension programs in India. According to the model reducing one type of cost, disproportionately can lead to a lower level of learning even when the DM is better off in terms of net payoff. If learning has positive externality then the optimal policy needs to involve reducing the cost of both types of learning so that the incentives to learn is not distorted.

Finally, I test the mechanism of learning in a laboratory setting where the subjects face a multidimensional, 9×9 payoff matrix choice with the additional choice of learning about the average payoff from each row and column. I find that in almost 85% of rounds the subjects' behavior satisfy the theoretical prediction. Moreover, they choose all three types of optimal learning strategy. Selective learning happens for a significant fraction of rounds and is payoff relevant, i.e., the average payoff from selective learning rounds and significantly lower compared to the average payoff across all strategies.

Bibliography

Beaman, L., Karlan, D., Thuysbaert, B., and Udry, C. (2014a). Self-selection into credit markets: Evidence from agriculture in mali. Technical report, National Bureau of Economic Research.

Beaman, L., Magruder, J., and Robinson, J. (2014b). Minding small change among small firms in kenya. *Journal of Development Economics*, 108:69–86.

Bloom, N., Eifert, B., Mahajan, A., McKenzie, D., and Roberts, J. (2013). Does management matter? evidence from india. *The Quarterly Journal of Economics*, 128(1):1–51.

Cai, J., De Janvry, A., and Sadoulet, E. (2015). Social networks and the decision to insure. *American Economic Journal: Applied Economics*, 7(2):81–108.

Caplin, A. and Dean, M. (2015). Revealed preference, rational inattention, and costly information acquisition. *The American Economic Review*, 105(7):2183–2203.

Caplin, A., Dean, M., and Leahy, J. (2017). Rationally inattentive behavior: Characterizing and generalizing shannon entropy. Technical report, National Bureau of Economic Research.

Chen, D. L., Schonger, M., and Wickens, C. (2016). otree—an open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9:88–97.

Conley, T. and Udry, C. (2001). Social learning through networks: The adoption of new agricultural technologies in ghana. *American Journal of Agricultural Economics*, 83(3):668–673.

De Janvry, A., Sadoulet, E., and Suri, T. (2017). Field experiments in developing country agriculture. In *Handbook of Economic Field Experiments*, volume 2, pages 427–466. Elsevier.

Duflo, E., Kremer, M., and Robinson, J. (2008). How high are rates of return to fertilizer? evidence from field experiments in kenya. *American economic review*, 98(2):482–88.

Duflo, E., Kremer, M., and Robinson, J. (2011). Nudging farmers to use fertilizer: Theory and experimental evidence from kenya. *American Economic Review*, 101(6):2350–90.

Foster, A. D. and Rosenzweig, M. R. (2010). Microeconomics of technology adoption. *Annu. Rev. Econ.*, 2(1):395–424.

Francetich, A. and Kreps, D. M. (2016a). Choosing a good toolkit, i: Formulation, heuristics, and asymptotic properties.

Francetich, A. and Kreps, D. M. (2016b). Choosing a good toolkit, ii: Simulations and conclusions.

Fryer, R. and Jackson, M. O. (2008). A categorical model of cognition and biased decision making. *The BE Journal of Theoretical Economics*, 8(1).

Fudenberg, D., Strack, P., and Strzalecki, T. (2017). Stochastic choice and optimal sequential sampling.

Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129(4):1661–1710.

Glendenning, C. J., Babu, S., and Asenso-Okyere, K. (2010). Review of agricultural extension in india. *Are farmers information needs being met.*

Hanna, R., Mullainathan, S., and Schwartzstein, J. (2014). Learning through noticing: Theory and evidence from a field experiment. *The Quarterly Journal of Economics*, 129(3):1311–1353.

Hébert, B. and Woodford, M. (2017). Rational inattention and sequential information sampling. Technical report, National Bureau of Economic Research.

Jovanovic, B. and Nyarko, Y. (1996). Learning by doing and the choice of technology. *Econometrica*, 64(6):1299.

Karlan, D., McConnell, M., Mullainathan, S., and Zinman, J. (2016). Getting to the top of mind: How reminders increase saving. *Management Science*, 62(12):3393–3411.

Karlan, D., Osei, R., Osei-Akoto, I., and Udry, C. (2014). Agricultural decisions after relaxing credit and risk constraints. *The Quarterly Journal of Economics*, 129(2):597–652.

Kondylis, F., Mueller, V., and Zhu, J. (2017). Seeing is believing? evidence from an extension network experiment. *Journal of Development Economics*, 125:1–20.

Lambrecht, I., Vanlauwe, B., Merckx, R., and Maertens, M. (2014). Understanding the process of agricultural technology adoption: mineral fertilizer in eastern dr congo. *World Development*, 59:132–146.

Leme, R. P. and Schneider, J. (2018). Contextual search via intrinsic volumes. *CoRR*, abs/1804.03195.

Liang, A., Mu, X., and Syrgkanis, V. (2017). Dynamic information acquisition from multiple sources. *arXiv preprint arXiv:1703.06367*.

Matějka, F. and McKay, A. (2015). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98.

Matějka, F. and Tabellini, G. (2017). Electoral competition with rationally inattentive voters.

Meijer, S. S., Catacutan, D., Ajayi, O. C., Sileshi, G. W., and Nieuwenhuis, M. (2015). The role of knowledge, attitudes and perceptions in the uptake of agricultural and agroforestry innovations among smallholder farmers in sub-saharan africa. *International Journal of Agricultural Sustainability*, 13(1):40–54.

Mohlin, E. (2014). Optimal categorization. *Journal of Economic Theory*, 152:356–381.

Mondria, J. (2010). Portfolio choice, attention allocation, and price comovement. *Journal of Economic Theory*, 145(5):1837–1864.

Mullainathan, S., Schwartzstein, J., and Shleifer, A. (2008). Coarse thinking and persuasion. *The Quarterly journal of economics*, 123(2):577–619.

Mullainathan, S. and Shafir, E. (2013). *Scarcity: Why having too little means so much*. Macmillan.

Peng, L. and Xiong, W. (2006). Investor attention, overconfidence and category learning. *Journal of Financial Economics*, 80(3):563–602.

Richter, S. G. L. P. M. (2017). Breadth versus depth.

Rivera, W. and Alex, G. (2004a). *Demand-driven Approaches to Agricultural Extension: Case Studies of International Initiatives*. World bank.

Rivera, W. and Alex, G. (2004b). *Privatization of extension systems: Case studies of international initiatives*. World bank.

Rivera, W. and Alex, G. (2018). Agriculture and rural development discussion paper 8 extension reform for rural development.

Schwartzstein, J. (2014). Selective attention and learning. *Journal of the European Economic Association*, 12(6):1423–1452.

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690.

Steiner, J., Stewart, C., and Matějka, F. (2017). Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica*, 85(2):521–553.

Van Nieuwerburgh, S. and Veldkamp, L. (2009). Information immobility and the home bias puzzle. *The Journal of Finance*, 64(3):1187–1215.

Van Nieuwerburgh, S. and Veldkamp, L. (2010). Information acquisition and under-diversification. *The Review of Economic Studies*, 77(2):779–805.

Woodford, M. (2014). Stochastic choice: An optimizing neuroeconomic model. *The American Economic Review*, 104(5):495–500.

A Appendix: Data

The major producers of cotton are ten Indian states divided into three regions, Northern, Central, and Southern. The Northern region consists of the states of Punjab, Haryana, and Rajasthan, Central region of Gujarat, Madhya Pradesh and Maharastra and Southern region of Karnataka, Telengana, Andhra Pradesh and Tamil Nadu. The three regions differ significantly in terms of irrigation, type of cotton produced and the length of the crop cycle (See table 5).

Particulars	Northern	Central	Southern
States	Punjab, Haryana, Rajasthan	Gujarat, Madhya Pradesh, Maharashtra	Karnataka, Telengana, Andhra Pradesh, Tamil Nadu
Irrigation	Mostly Irrigated	Irrigated and Rainfed	Irrigated and Rainfed
Planting Season	April-May	June-July	July-August
Harvesting Season	October-November	November- April	December- March
Variety	Medium and Short Staple	Medium and Long Staple	Long and Extra-Long Staple

Table 5: Regional Variation in Cotton Production in India

NSS schedule 33 does not report the actual learning content of the farmer but records various information sources. Table 7 shows the average learning of the farmer. Note that all sources of information does not necessarily give a similar type of information. For example, government extension services provide information about specific input combination to maximize the total payoff whereas input dealers or media often provide information about one input (marginal distribution) separately.

Variable	Description	Unit	Summary statistics
	A. Input		
Seed	Cost of seed per ha	Rs	mean = 4584.465, sd = 6281.258
Irrigation	Cost of irrigation per ha	Rs	mean = 478.7396, sd = 2001.767
Fertiliser	Cost of chemical fertiliser per ha	Rs	mean = 11269.87, sd = 15881.04
Manure	Cost of manure per ha	Rs	mean = 1140.944, sd = 5098.996
Area_irr	Percentage of irrigated area	%	mean = 0.4034814, sd = 0.488216
Land	Total cultivated land (in logs)	Ha.	mean = 1.322927, sd = 1.367415
	B. Household		
Land Possession	Total land possession (own or rented)	Ha.	mean = 2.657479 , sd = 2.614432
Ration Card	1 = Antodaya (poorest) 2= BPL(below poverty line), 9 = other	%	1: 3.1%, 2: 40.5%, 9: 48.6%
MSP	Aware of Minimum Support Price	%	aware : 21.7%
Insurance	1: insured only with loan, 2: insured, 3: not insured	%	1: 10%, 2: 1.6%, 3: 88.4%
Household size	in logs	N/A	mean = 5.16, sd = 2.33
Age	Age of the HH Head(in logs)	Years	mean = 49.75, sd = 13.03
	C. Information		
Extension Agents	Accessed information at least once	%	8.2%
Private	same	%	13.3%
Other progressive farmers	same	%	30.8%
Media	same	%	28%

Table 6: Control Variable

Source of Information	% accessed at least once
Extension Agents	8.2%
Private	13.3%
Other progressive farmers	30.8%
Media	28%

Table 7: Source of Information

	Dependent variable:			
	Net Profit	Total Value	Total Quantity	Crop Loss (=1)
	(1)	(2)	(3)	(4)
Pesticide	-0.874^{***}	0.013^{**}	0.345^{*}	-0.0000008
	(0.195)	(0.005)	(0.185)	(0.000006)
Observations	2,144	2,141	2,144	2,144
Mean	7282	54311	1383	
Sd	69465	64301	1687	
Adjusted R^2	0.413	0.331	0.384	.360

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Respondents from 2012-13 agriculture year who produced cotton as the only crop or major crop during the Kharif season. All input variables are measured in Rupees. Net profit is defined as Total Value (in Rs) minus total cost of all inputs(in Rs). Total quantity is measured in Kg. 113 districts from 10 states are considered and district FE added. all other inputs, farm level characteristics are controlled for.

Table 8: Effect of pesticide

	Dependent variable:			
	Net Payoff = Total Value - Total Cost			
	(1)	(2)	(3)	(4)
Pesticide	−0.038	−0.085	−0.091	−0.414
	(0.381)	(0.409)	(0.410)	(0.465)
Pesticide* post	−1.315***	−1.231***	−1.216***	−0.843*
	(0.391)	(0.419)	(0.420)	(0.478)
Household	No	Yes	Yes	Yes
Information	No	No	Yes	Yes
Post*X	No	No	No	Yes
Observations	4,005	3,886	3,881	3,881
R^2	0.373	0.371	0.372	0.379
Adjusted R^2	0.347	0.343	0.343	0.348
Note:	*p<0.1; **p<0.05; ***p<0.01			

Respondents who produced cotton in relevant agriculture year from both 2002-03 and 2012-13 survey years are considered. District and Year fixed effects are included. Fourth column also includes the interaction term for each explanantory variable(not reported). District FE and Year FE is included in all columns.

Table 9: Panel 1: net payoff

	Dependent variable:			
	Total Value of Output		Total Quantity of Output	
	(1)	(2)	(3)	(4)
Pesticide	0.364	0.066	0.062***	0.031*
	(0.384)	(0.436)	(0.016)	(0.018)
Pesticide* post	−0.218	0.114	−0.052***	−0.019
	(0.393)	(0.448)	(0.016)	(0.018)
Post*X	No	Yes	No	Yes
Observations	3,881	3,881	3,878	3,878
R^2	0.455	0.458	0.225	0.235
Adjusted R^2	0.430	0.431	0.190	0.197

Note: *p<0.1; **p<0.05; ***p<0.01

Respondents who produced cotton in relevant agriculture year from both 2002-03 and 2012-13 survey years are considered. District and Year fixed effects are included. Fourth column also includes the interaction term for each explanantory variable(not reported).

Table 10: Panel 1: Impact of pesticide on total value and total quantity of output

	Dependent variable:			
	Net profit		Total Value	
	(1)	(2)	(3)	(4)
Pesticide	6.398***	3.469***	7.301***	3.960***
	(0.633)	(0.724)	(0.605)	(0.715)
Pesticide*Cotton	−7.421***	−4.478***	−7.136***	−3.744***
	(0.612)	(0.724)	(0.585)	(0.717)
Cotton* X	No	Yes	No	Yes
Observations	3,405	3,405	3,405	3,405
R^2	0.782	0.816	0.797	0.832
Adjusted R^2	0.403	0.493	0.445	0.536

Note: *p<0.1; **p<0.05; ***p<0.01

Households that produce cotton in Kharif season of 2012-13 are considered. Other crop refers to all other crops produced by same households in Rabi season of 2012-13. Column (2) and (4) includes all interaction terms for all explanatory variables. District FE, Season FE and Household FE are included.

Table 11: Panel 2: Net profit and Total Value

	Dependent variable:			
	Net profit			
	(1)	(2)	(3)	(4)
Pesticide*post*Cotton	-3.357^{***}	-3.352^{***}	-3.345^{***}	-3.303^{***}
	(0.679)	(0.679)	(0.679)	(0.882)
Farm	No	Yes	Yes	Yes
Information	No	No	Yes	Yes
Interaction	No	No	No	Yes
Observations	6,960	6,960	6,960	6,960
R^2	0.280	0.280	0.282	0.335
Adjusted R^2	0.263	0.263	0.264	0.315

Note: *p<0.1; **p<0.05; ***p<0.01

Cotton farmers from both 2002-03 and 2012-13 survey years are incldued. Other crop refres to all other crops produced by the same households in same agriculture year but during Rabi season. All othet inputs,. District FE, Year FE and Season FE added. Column 4 includes the interaction term for all explanatory variables. Interaction includes for all explanatory variable interaction with post, cotton and post× cotton separately.

Table 12: Full Panel: Net profit

	Pesticide		
	(1)	(2)	(3)
Prior	0.537***	0.537***	0.530***
	(0.112)	(0.113)	(0.116)
X_1	No	Yes	Yes
X_2	No	No	Yes
Observations	2,148	2,148	2,140
Adjusted R^2	0.544	0.549	0.552
Note:			*p<0.1; **p<0.05; ***p<0.01

Prior refers to district level average pesticide use in 2002-03. Region FE included.

Table 13: Prior and Pesticide Use

	Net Profit			
	(1)	(2)	(3)	(4)
Prior	−2.204*	−2.332*	−1.669	−2.135
	(1.285)	(1.298)	(1.323)	(1.354)
X_1	No	Yes	Yes	Yes
X_2	No	No	Yes	Yes
Yield Difference	No	No	No	Yes
Observations	2,148	2,148	2,148	2,140
Adjusted R^2	0.323	0.329	0.337	0.335
Note:				*p<0.1; **p<0.05; ***p<0.01

Table 14: Prior and Net Profit

B Appendix: proofs

B.1 Proof of lemma 1

Proof. We want to show that given any prior belief μ_0 the DM's optimal choice of observing a cell at any stage t where the DM's belief is given by μ_t depends only on μ_t and not the history of μ_s for $s \leq t$. This is sufficient because if μ_s alone determines the choice of cell p then the DM's problem can be written as optimal choice of p given μ_t which is described in the value function $V(\mu_t)$.

The strategy of observing any cell generates a unique partition of the state space Ω. For example, if the DM chooses to observe a cell in the matrix, it can have two possible values, namely 0 or 1. Thus observing the cell is equivalent to partitioning the state space into two blocks, one where the cell takes value 1 and the other where it takes 0. Whereas if the DM observes an average it can take values in $\{0, 1/n, 2/n, \ldots, 1\}$.

Once the DM observes the realization in a given cell he chooses the block of cells that contains the true state. For example, if he observes an average where there is only 1 cell that generates $Y = 1$ then he would learn that the true state belongs to the block of states where the average takes the value $1/n$. If now he observes another cell from the same row (or column) then that cell can take value 0 or 1, which would further divide the new subset into two blocks and the DM would learn which block contains the true state.

Thus any learning strategy can be described as a sequence of partitions where the DM learns that true state belong to the intersection of relevant blocks. Given the partition consistency condition, the DM cannot further update his belief about the states in the block without further observation.

Now suppose there are two histories that generate the same belief μ' but the optimal strategy following them would be different. Note that, for both two histories, say μ^1 and μ^2 the resulting subset that contains the true state under μ' must be the same otherwise the belief μ' cannot be same.

This implies there are three possible ways in which the two histories are different. First, the same cells are observed but in a different order. Since the intersection of sets is commutative and each cell generates a unique partition of Ω the order of partition is irrelevant for decision making.

Second, under one of the histories, say μ^1 one cell has been observed that was not payoff relevant, i.e, the updating did not reduce any uncertainty, but this would imply the strategy chosen under μ^1 was not optimal.

Third, there exist multiple sequences of cells observing which would generate the same partition. In that case, the DM must be indifferent about choosing any of two sequence which again implies it cannot be relevant for future decision making. Otherwise ex-ante the DM would not choose one of the two histories optimally.

Thus any t^{th} round belief μ_t denotes a unique subset of states that contains the true state and hence the optimal policy only depends on the belief μ^t and not on the entire history.

□

B.2 Correlation of cells and Uncertainty

Result 2. *For any two beliefs μ_1 and μ_2 if all else being equal if there exists a pair of cells (s_i, p_j) and (s_k, p_l) such that the two cells are more correlated under μ_2 then $H(\mu_1) > H(\mu_2)$, i.e, μ_2 would have higher uncertainty.*

Proof. All else being there are four possibilities where μ_1 and μ_2 differ. They are as follows:

1. $\pi(s_i, p_j) = 1$ and $\pi(s_k, p_l) = 1$
2. $\pi(s_i, p_j) = 1$ and $\pi(s_k, p_l) = 0$
3. $\pi(s_i, p_j) = 0$ and $\pi(s_k, p_l) = 1$
4. $\pi(s_i, p_j) = 0$ and $\pi(s_k, p_l) = 0$

Higher negative correlation implies the probability of possibility 2 and 3 are higher than that of 1 and 4 and opposite for positive correlation. However, in both cases, the possible values are less spread since once two out of the four possibilities are more likely. For a lower value of correlation the probability of all four possibilities are closer. If all else are equal then the belif μ_2 where there is greater correlation the uncertainty is also lower.

Note that, this result does not depend on number of possible values of Y. If there are more than two possible values of Y then one strict subset of possibilities for the payoffs of the two cells have a higher probability compared to other possiblilities. This implies the uncertainty would be lower when correlation is higher. □

B.3 Proof of lemma 2

Proof. As noted before observing a cell (s_i, p_j) partitions the state space into two blocks, namely $B_{0,(i,j)}$ where $\pi(s_i, p_j) = 0$ and $B_{1,(i,j)}$ where $\pi(s_i, p_j) = 1$. Further learning is only optimal if the observed cell generates $Y = 0$ otherwise the DM would choose the observed cell. Let π_t denote the expected payoff in round t with belief μ_t, i.e., π_t is the probability of the cell(s) that has the highest probability of generating $Y = 1$ without further learning. To prove the lemma two observations are to be made.

Observation 1: The order of cells observed does not affect the incentive to learn. In each round, the DM finds the intersection of Blocks, of partitions generated by observing cells, where the true state lies. Since intersection is commutative the order of observation does not change the posterior belief. In this lemma, the objective is to compare two strategies where the order of observing cells are rearranged, since given a belief μ_t in any period the same set of cells are to be observed when learning is optimal. Even though rearranging the cells generate the same posterior, the order of observation affects the expected cost as the DM stops learning once he finds a cell with payoff $Y = 1$.

Observation 2: Lower $R^c_{i,j}(\mu_t)$ leads to lower expected cost of learning. For a given belief μ_t the denominator in $R^c_{i,j}(\mu_t)$ remains the same for all cells. The numerator is lower is either $\pi_{t,(i,j)}$, i.,e., the total probability of all states in $B_{1,(i,j)}$ is lower or the residual uncertainty is lower. This implies if a cell with lower $R^c_{i,j}(\mu_t)$ either it has a higher probability of generating $Y = 1$ or in case it generates $Y = 0$, the cell is correlated with another subset of cells that has a higher probability of generating $Y = 1$. In both cases, fewer cells need to observe and hence the expected cost is lower.

The objective is to show that a higher one-period expected payoff implies a higher net value from learning. Let us start with two cells, namely (s_i, p_j) and $(s_{i'}, p_{j'})$ where $\pi_{t,(i,j)}$ takes the highest value and $\pi_{t,i',j'}$ has the second highest value in the payoff matrix. There are three possibilities that need to be considered.

Case 1:

$$\pi_{t,(i,j)} + (1 - \pi_{t,(i,j)})\pi_{t+1|(i,j)=0} < \pi_{t,(i',j')} + (1 - \pi_{t,(i',j')})\pi_{t+1|(i',j')=0} \tag{2}$$

This implies $\pi_{t+1|(i,j)=0} < \pi_{t+1|(i',j')=0}$, i.e, after observing cell (i', j') to be generating $Y = 0$ the DM finds a cell (k, l) such that

$$\pi_{t+1,(k,l)} > \pi_{t+1,(i,j)}.$$

If the DM observes (i, j) first then the possible sequence of learning would be given by (i, j) followed by (i', j'), followed by (k, l). The alternate strategy is to start with (i', j'), followed by (k, l) followed by (i, j). By observation 1 the rest of the learning startegy will not be affected after any of these two strategies. Note that, under strategy 2 the DM observes a cell (k, l) with highest probability of generating $Y = 1$ is later period. The second would be better if

$$c_l(1 + \pi_{t,i',j'} + \pi_{t,i',j'}\pi_{t+1,(k,l)}) < c_l(1 + \pi_{t,(i,j)} + \pi_{t,(i,j)}\pi_{t,i',j'})$$
$$\pi_{t,i',j'}(1 + \pi_{t+1,(k,l)}) < \pi_{t,(i,j)}(1 + \pi_{t+1,i',j'})$$

$$\pi_{t+1,i',j'}(\pi_{t+1,(k,l)} - \pi_{t,(i,j)}) < \pi_{t,(i,j)} - \pi_{t+1,(i',j')} \tag{3}$$

Note that, in this example $\pi_{t+1|(i,j)=0} = \pi_{t+1,(i',j')}$ and $pi_{t+1|(i',j')=0} = \pi_{t+1,(k,l)}$, so rearraging equation **??** we get

$$\pi_{t,(i',j')}(\pi_{t+1,(k,l)} - \pi_{t+1,(i,j)}) < \pi_{t+1,(k,l)} - \pi_{t,(i,j)} \tag{4}$$

This is true $\pi_{t+1,(i,j)} = \pi_{i',j'}$ otherwise observing (s_i, p_j) is informative about $(s_{i',j'})$ and observing (s_i, p_j) would be strictly better. Note that inequality 4 would imply inequlaity 3 if $\pi_{t+1,(k,l)} < 3\pi_{t,(i',j')}$ as $\pi_{t,(i,j)} \geq \pi_{t,(i',j')}$. Since $\pi_{t,(k,l)} \leq \pi_{t,(i',j')}$ the condition $\pi_{t+1,(k,l)} < 3\pi_{t,(i',j')}$ holds true, i.e, if not the DM would choose to observe (k, l) before (i', j'). Thus the second strategy is optimal.

Case 2:

$$\pi_{t,(i,j)} + (1-\pi_{t,(i,j)})\pi_{t+1|(i,j)=0} < \pi_{t,(i',j')} + (1-\pi_{t,(i',j')})\pi_{t+1|(i',j')=0} \tag{5}$$

This can happen if after observing $\pi(s_i, p_j) = 0$ the DM finds a cell (k,l) such that $\pi_{t+1,(k,l)} > \pi_{t+1,(i',j')}$. In this case (i,j) has a higher probability of generating $Y=1$ and also more informative, so observing (i,j) before (i',j') reduces the expected cost of learning.

Case 3:

$$\pi_{t,(i,j)} + (1-\pi_{t,(i,j)})\pi_{t+1|(i,j)=0} = \pi_{t,(i',j')} + (1-\pi_{t,(i',j')})\pi_{t+1|(i',j')=0} \tag{6}$$

In this case after observing (i',j') $\pi_{t+1,(i,j)}$ remains the highest, i.e., observing (s_i,p_j) remains optimal. Given the belief μ_t here the cell with lower $R^c_{i,j}$ would generate a lower expected cost given observation 2. If $\pi_{t,(i,j)}$ is sufficiently high then $R^c_{i,j} < R^c_{i',j'}$ and observing (s_i,p_j) is optimal. Otherwise the residual uncertainty by observing $(s_{i'},p_{j'})$ dominates and makes $(s_{i'},p_{j'})$ the optimal choice.

Finally, if both $\pi_{t,(i,j)} = \pi_{t,i',j'}$ and $R^c_{i,j} = R^c_{i',j'}$, observing any order of observation would generate the same belief and have same expected cost hence the DM would be indifferent between them. □

B.4 Proof of lemma 3

Proof. Let cell (s_i,p_j) be the optimal cell to observed, i.e., satisfies the condition given in lemma 2. Let $E_t(V|(i,j))$ denote the expected payoff from observing cell (s_i,p_j), then learning is optimal only if,

$$E_t(V|(i,j)) - \pi_t \geq c_l$$

where π_t denotes the expected payoff from no further learning is round t. The proof consists of two steps. First, how $H(\mu_{t+1}|B_{0,(i,j)})$ characterizes the optimal learning and second $H(\mu_{t+1}|B_{0,(i,j)})$ and $H(\mu_t)$ has a one-to-one relationship given π_t and $R^c_{i,j}$.

First note that given π_t and c_l the value of learning depends only on $E_t(V|(i,j))$. There are two possibilities, namely (s_i,p_j) generates $Y=1$ which happens with probability at most π_t given lemma 2. However, if (s_i,p_j) generates $Y=0$ then the value of learning depends on further possibility of learning given the updated belief. The objective is to characterize the possibility of further learning given $\pi(s_i,p_j)=0$ by $H(\mu_{t+1}|B_{0,(i,j)})$.

There are two opposing effects of higher $H(\mu)$, namely if high $H(\mu)$ implies there are a greater number of possible states then the expected cost of learning is higher then higher uncertainty $H(\mu)$ leads to a smaller value of learning. Whereas if high $H(\mu)$ implies that the belief can possibly change significantly then higher uncertainty $H(\mu)$ implies a higher value of learning.

To show this consider the two extreme cases, if μ is close to uniform then the payoff from different cells are independent, i.e, observing one cell does not change the belief over the payoff from other cell. If μ reduces from the uniform belief then the cells become more informative about each other and this would decrease the expected cost of learning. On the other hand, if uncertainty is so small such that observing a new cell cannot change the belief further then there is no value in learning. If another belief μ has ϵ higher uncertainty then the value from learning increases.

Given the same π_t and R^c let us consider three possible beliefs μ_1, μ_2 and μ_λ that can be generated when $\pi(s_i, p_j) = 0$ where $H(\mu_\lambda) = \lambda H(\mu_1) + (1-\lambda)H(\mu_2)$. WLOG let $E_{\mu_1}V > E_{\mu_2}V$, i.e, the value from learning for belief μ_1 is higher. To show that $E_\mu V$ is quasi-concave in $H(\mu)$ it needs to be shown that $E_{\mu_\lambda}V \geq E_{\mu_2}V$, i.e., the expected gain from μ_λ is higher than the lower value.

For any two beliefs μ and ν, $H(\mu) > H(\nu)$ if either μ gives positive probability to more states and/or μ is closer to uniform. Let $H(\mu_1) > H(\mu_2)$, then $H(\mu_\lambda) > H(\mu_2)$. Since μ_1 has a higher value from learning it must be the case that the effect of having more states is dominated by the possibility of further learning. Since $H(\mu_\lambda)$ lies in between $H(\mu_1)$ and $H(\mu_2)$, the expected number of cells to be observed to obtain $Y = 1$ is bounded by the expected number of cells by that of μ_1 and the expected change in belief is at least as high as μ_2. This implies that the net benefit from observing is at least as much as high as $E_{\mu_2}V$.

Similarly, if $H(\mu_1) < H(\mu_2)$, then $H(\mu_\lambda) < H(\mu_2)$. In this case, the effect of the lower expected cost of learning dominates and since μ_λ would have a lower cost of learning than μ_2 but more information content than μ_1 implies the gain from learning is at least as much as $E_{\mu_2}V$.

The quasi-concavity of the expected value in $H(\mu_{t+1}|B_{0,(i,j)})$ implies that for given c_l and π_t there exists an interval in $H(\mu_{t+1}|B_{0,(i,j)})$, such that learning is only optimal but not optimal when $H(\mu_{t+1}|B_{0,(i,j)})$ is too high or too low.

Finally, note that

$$H(\mu_{t+1}|B_{0,(i,j)}) = \frac{1}{1-\mu_t}\left[R^c H(\mu_{t,(i,j)}) + \ln\left(1-\mu_{t,(i,j)}\right)\right].$$

Thus if $\mu_t = \mu_{t,(i,j)}$ then there is a one-to-one relationship between $H(\mu_{t+1}|B_{0,(i,j)})$ and $H(\mu_t)$ given π_t and R^c. Even if $\mu_t > \mu_{t,(i,j)}$ then $R^c = R^c_{i,j}$ and hence two beliefs with same $H(\mu_t)$ cannot degerate two different $H(\mu_{t+1}|B_{0,(i,j)})$. Hence there is a one-to-one relationship between $H(\mu_t)$ and $H(\mu_{t+1}|B_{0,(i,j)})$ and thus the interval in $H(\mu_{t+1}|B_{0,(i,j)})$ can also be expressed as $H(\mu_t)$ given π_t and R^c. Hence, proved. □

B.5 Proof of lemma 4

Proof. When a DM observes an average $a_{i,j}$ where $i \in \{S, P\}$ and $j \in \{1, \ldots, n\}$ the state space is partitioned into several blocks, one for each possible value of the average $a_{i,j}$. Given the definition of $R^a_{i,j}$ between two averages $a_{i,j}$ and $a_{k,l}$, average $a_{i,j}$ would have lower $R^a_{i,j}$ if either the probability of generating $Y = 1$ is higher for $a_{i,j}$ or given that $Y = 1$ is not generated the resulting uncertainty as measured by $H(\mu_{t+1}|B_{k/n,(i,j)})$ is lower, where $H(\mu_{t+1}|B_{k/n,(i,j)})$ is the entropy of the belief μ_{t+1} given μ_t and $\pi(a_{i,j}) = k/n$. Let $\pi^a_{1,(i,j)}$ denote the probability that observing average $a_{i,j}$ would informa about a cell with $Y = 1$.

Step 1: Let us first consider the case where there are two averages $a_{i,j}$ and $a_{k,l}$, both have the same lowest $R^a_{i,j}$ but $a_{i,j}$ have a higher expected payoff. Higher expected payoff implies fewer cells to observe and hence a lower expected cost of learning. Hence observing $a_{i,j}$ would be optimal.

Step 2: To show that observing the average with lowest $a_{i,j}$ is optimal it is sufficient to show that for any two averages $a_{i,j}$ and $a_{k,l}$ if $R^a_{i,j} < R^a_{k,l}$ then $E_{\mu_t}(V|a_{i,j}) > E_{\mu_t}(V|a_{k,l})$. Let us consider two such averages and two possible cases.

Case 1: Suppose $\pi^a_{t,(i,j)} < \pi^a_{t,(k,l)}$ but $R^a_{t,(i,j)} < R^a_{t,(k,l)}$. Consider a hypothetical average value $a_{i',j'}$ such that the $\pi^a_{k,l} = \pi^a_{i',j'}$ but $R^a_{i,j} = R^a_{t,(i',j')}$. Then observing $a_{i',j'}$ generated lower expected cost than $a_{k,l}$, i.e.,

$$E_{\mu_t}(V|a_{i',j'}) \geq E_{\mu_t}(V|a_{k,l}).$$

Also between $a_{i',j'}$ and $a_{k,l}$ since $\pi^a_{t,(i',j')} > \pi^a_{t,(i,j)}$ for $R^a_{t,(i,j)} = R^a_{t,(i',j')}$ to hold true $\pi^a_{t,(i',j')}$ has to be sufficiently smaller than $(1 - \pi^a_{t,(i',j')})$ i.e, the case where $\pi(a_{\pi^a_{t,(i',j')}}) < 1$ dominates. Also, the expected payoff if $\pi(a_{t,(i',j')}) < 1$ is lower than the expected payoff when $\pi(a_{t,(i,j)}) < 1$,. These two observations imply the expected payoff for $a_{i',j'}$ has to be lower than that of $a_{i,j}$. Given step 1 it would thus be optimal to observe $a_{i,j}$ over $a_{i',j'}$, i.e,

$$E_{\mu_t}(V|a_{i',j'}) \leq E_{\mu_t}(V|a_{i,j}).$$

Combining the two results we get,

$$E_{\mu_t}(V|a_{k,l}) \leq E_{\mu_t}(V|a_{i,j})$$

i.e, observing $a_{i,j}$ is optimal.

Case 2: Suppose $\pi^a_{i,j} < \pi^a_{k,l}$ but $R^a_{i,j} \geq R^a_{k,l}$. This implies average $a_{i,j}$ has a higher probability of generating $Y = 1$ and also average $a_{i,j}$ is not significantly more informative than $a_{k,l}$. This implies observing $a_{k,l}$ would lead to lower cost of learning and hence observing $a_{k,l}$ would be the optimal strategy.

Hence, proved. □

B.6 Proof of lemma 5

Proof. Proof of this lemma is similar to that of the proof of lemma 3. There are two steps to the proof. First, the optimal learning strategy can be characterized by an interval on $H(\mu_{t+1}|Y < 1)$

given expected payoff π_t,information content R^a and cost of learning c_a and second, there is a one-to-one relationship between $H(\mu_{t+1}|Y<1)$ and $H(\mu_t)$.

Proof of the first step is similar to lemma 3. For higher residual uncertainty $H(\mu_{t+1}|Y<1)$, i.e, when learning is optimal, the expected number of averages to be observed is higher and for lower residual uncertainty $H(\mu_{t+1}|Y<1)$ the expected change in payoff due to learning is smaller. Due to these two opposing effects the value function is quasi-concave in $H(\mu_{t+1}|Y<1)$, This implies that learning is optimal within an interval in $H(\mu_{t+1}|Y<1)$.

To prove the second step let us note that the residual uncertainty can be written as

$$H(\mu_{t+1}|Y<1) = \frac{1}{1-\pi^a_{t,(i,j)}}\left[R^a H(\mu_t) + \ln\left(1-\pi^a_{t,(i,j)}\right)\right].$$

Hence given R^a there is a one-to-one relationship between $H(\mu_{t+1}|Y<1)$ and $H(\mu_t)$ for $\pi^a_{t,(i,j)} \leq \pi_t$. Hence, proved. □

B.7 Proof of lemma 6

Proof. Since the DM now has access to both cells and averages, learning is optimal if at least one type of learning is optimal. Lemma 3 and 5 implies that there exists an upper and a lower bound on $H(\mu_t)$ such that learning is not optimal above the upper bound and below the lower bound. Note that, none of these bounds are strict, i.e., it is possible that learning is always optimal for some cost difference δ and belief μ.

Given the upper and lower bound it is sufficient to show that given expected payoff π_t in round t for any two belief μ_1 and μ_2 such that $H(\mu_1) > H(\mu_2)$ if at μ_1 it is optimal to observe cell then for μ_2 it would also remain optimal to observe cell. Also, if at μ_2 observing an average was optimal then it would remain so under μ_1 as well. This would generate a cutoff in $H(\mu_t)$ in between the upper and lower bound such that observing average is not optimal below the cutoff and observing cell is not optimal above the cutoff, given the quasi-concavity of the value function for each type of learning in $H(\mu_t)$.

Consider two such belief μ_1 and μ_2 with $H(\mu_1) > H(\mu_2)$ given an expected payoff π_t. There are three possible cases. First, $\pi_{t,1} = \pi_{t,2}$ but $\pi_{t,(i,j),1} > \pi_{t,(i,j),2}$, i.e, even though the expected payoff is same under both beliefs, under μ_1 there is a higher probability of fnding a cell with $Y = 1$ with the optimal cell chosen. In other words, under μ_2 the DM chooses a cell (i', j') has does not have the highest probability of generating $Y = 1$. Second, $\pi_{t,1} = \pi_{t,2}$ but $\pi_{t,(i,j),1} < \pi_{t,(i,j),2}$, the expected payoff from no learning would be same under both beliefs but the optimal cells to be observed generates possibly lower probability of generating $Y = 1$. Third, $\pi_{t,(i,j),1} = \pi_{t,(i,j),2}$ but $\pi_{t+1,2|(i,j)=0} > \pi_{t+1,1|(i,j)=0}$, i.e, under μ_2 if $\pi(s_{i',p_{j'}}) = 0$ is observed then the expected payoff is higher than a similar situation under the case of μ_1. Fourth, $\pi_{t,(i,j),1} = \pi_{t,(i,j),2}$ but $\pi_{t+1,2|(i,j)=0} > \pi_{t+1,1|(i,j)=0}$, i.e, opposite of case 2. The resulting strategy would be different in each of these cases each the incentive to learn is different. Hence, we will consider them separately.

Case 1: Since $\pi_{t,(i,j),1} > \pi_{t,(i,j),2}$ but $\pi_{t,1} = \pi_{t,2}$ and $H(\mu_{t+1,2}|B_{0,(i,j)}) < H(\mu_{t+1,1}|B_{0,(i,j)})$, there exists at least a pair of cells (s_i, p_j) and (s_k, p_l) such that the negative correlation among them are significantly higher under μ_2 making (s_i, p_j) (WLOG) the optimal cell to be observed under μ_2. Since a higher negative correlation implies informative of cells increase more than that of averages, if the DM was observing cells under μ_1 he would also find it optimal to observe cells under μ_2 as well.

Case 2 : If $\pi_{t,(i,j),1} < \pi_{t,(i,j),2}$ but $\pi_{t,1} = \pi_{t,2}$ then there must exist a pair of cells under μ_1 that are more correlated and generates atleast $\pi_{t,1}$ compared to μ_2. But this would imply that $H(\mu_2) > H_{\mu_1}$ since more correlation with same expected payoff implies lower uncertainty. Thus this condition cannot be satisfied.

Case 3: $H(\mu_{t+1,2}|B_{0,(i,j)}) < H(\mu_{t+1,1}|B_{0,(i,j)})$ and $\pi_{t+1,2|(i,j)=0} > \pi_{t+1,1|(i,j)=0}$, i.e, under μ_2 the residual uncertainty in case $\pi(s_i, p_j) = 0$ is lower and μ_2 also generates a higher expected payoff without any further learning. This is similar to case 1, i.e., there exists a pair of cells $(s_{i'}, p_{j'})$ and (s_k, p_l) such that the negative correlation between the two cells are greater under μ_2. This increases the gain from observing cell under μ_2. Hence, if the DM was observing cells under μ_1 then he would also observe cells under μ_2.

Since the possible gain from an average is bounded by the gains from cells as expected gain from cells increase the gain from average increases but is bounded by the change in payoff from cells as $\pi_{t+1,2|(i,j)=0} > \pi_{t+1,1|(i,j)=0}$, i.e, the probability of obtaining $Y = 1$ by observing cells increases. If it was optimal to observe cells before it would remain so after the gain from observing cells increase.

Case 4: $H(\mu_{t+1,2}|B_{0,(i,j)}) < H(\mu_{t+1,1}|B_{0,(i,j)})$ but $\pi_{t+1,2|(i,j)=0} < \pi_{t+1,1|(i,j)=0}$, i..e, the residual uncertainty in case $\pi(s_i, p_j) = 0$ is lower but μ_2 generates a lower expected payoff without any further learning. In this case the expected gain from observing a cell reduces due to a lower probability of obtaining $Y = 1$.

There are two possible scenarios where $H(\mu_{t+1,2}|B_{0,(i,j)}) < H(\mu_{t+1,1}|B_{0,(i,j)})$ but $\pi_{t+1,2|(i,j)=0} < \pi_{t+1,1|(i,j)=0}$. First, there exists at least one pair of cells $(s_{i'}, p_{j'})$ and (s_k, p_l) such that the two cells are more positively correlated under μ_2 than under μ_1. Second, the level of correlation is same for all pairs of cells under μ_1 and μ_2 but the belief is less diffused under μ_2, i.e, when the DM is relatvely more sure about the best choice of cell and in case that cell generates $Y = 0$ there are not many other cells that would generate $Y = 1$ with a higher probability.

In the first case, the cells are more informative but the informativeness of the averages can increase as well. If the pair of cells $(s_{i'}, p_{j'})$ and (s_k, p_l) belong to different column and rows then the change in uncertainty would not change the relative informativeness of cell and average. Hence, if it was optimal to observe the cell under μ_1 with an increase in correlation of cell it would remain so.

If however, $(s_{i'}, p_{j'})$ and (s_k, p_l) belongs to same row or same column the informativeness of cell and average both increases. But since the two cells are more correlated in the case where the observed cell generates $\pi(s_{i'}, p_{j'}) = 0$ observing another cell become less profitable. But if the DM observe avearges he cannot distinguish perfectly between the two possibilities where $\pi(s_{i'}, p_{j'}) = 0, \pi(s_k, p_l) = 1$ and $\pi(s_{i'}, p_{j'}) = 1, \pi(s_k, p_l) = 0$. In these two cases observing cell would generate weakly higher payoff by reducing expected cost of learning. For the other two possibilities the cell and average are equally informative. Thus observing cell becomes more profitable under μ_2 where the pair of cells $(s_{i'}, p_{j'})$ and (s_k, p_l) are more correlated.

In the second case, the net benefit from learning decreases as the uncertainty of the belief is lower and expected payoff in case $Y = 0$ is lower as well. In case the cells for which the probability of generating $Y = 1$ changes (either increases or decrease) then the informativeness of the cell and the average change similarly. This is because averages take one values with higher probability and one of the cells generate $Y = 1$ with probability but every other cell generates $Y = 1$ with lower probability. But the increase in benefit from averages is bounded by the benefit from cells unless $\mu_{t+1,(i,j)=0} = 0$ and costs do not change. This means if the DM was observing cells at μ_1 then observing cells remain optimal under μ_2 as well. In case $\mu_{t+1,(i,j)=0} = 0$ no learning is optimal at μ_2.

If the cells that have a different probability of generating $Y = 1$ belong to different rows and column then relative information content between cells and averages do not change between μ_1 and μ_2. However, the benefit from learning decreases overall, hence the DM either continue observing cells or stop learning. Hence, for μ_2 observing a cell is optimal if it was optimal under μ_1.

On the other hand, if uncertainty increases from μ_2 to μ_1, then it must be true that under μ_1 are either less correlated or the belief is more diffused, i.e, more cells can generate $Y = 1$ probability with similar probability. In both cases $\pi_{t,(i,j)=0}$ is lower under μ_1. However, under μ_1, the set of possible values for the average increases, which implies the averages can resolve more uncertainty under μ_1 (μ_1 also has more uncertainty). Thus averages are relatively more informative when $\pi_{t,(i,j)=0}$ is lower making observing average more profitable under μ_1 as well.

However, this does not guarantee that there exists valued of $H(\mu_t)$ such that learning by cell or average is at all. Also, if the average is not informative and δ is sufficiently large then it is possible that there exists values of $H(\mu_t)$ for which learning an average has become uninformative but observing a cell has a prohibitive. This generates the two cutoffs in the middle. Again all these cutoffs need not be strict, the length of any of the intervals can go to zero. □

B.8 Proof of theorem 1

Proof. For this proof, we consider the optimal learning strategy of the DM. Lemma 6 describes the optimal choice for the recursive learning problem. Thus the proof of this theorem would be an application of lemma 6.

Lemma 6 implies the if $H(\mu_0) \in (\overline{H}_l(\pi_0), \overline{H}_h)(\pi_0)$ observing an average is optimal and for $H(\mu_0) \in (\underline{H}_l(\pi_0), \underline{H}_h)(\pi_0)$ obseving a cell in round t is optimal. The assumption of no crucial cells implies that observing a cell cannot increase the uncertainty of the belief in the following round and observing a cell cannot increase the information content of averages for the payof matrix.

Thus if the DM starts with observing cells then observing cells will continue to be optimal since unceratinty of belief μ_t for any subsequent round $t > 0$. If the DM starts with observing an average in round $t = 0$ then it is possible that there exists a round s such that $H(\mu_s) \in (\underline{H}_l(\pi_s), \underline{H}_h)(\pi_s)$ since $H(\mu_s) < H(\mu_0)$. In this case the DM will switch to observing cells.

However, the optimal intervals depend on the value of π_t. We need to check whether a DM who observes a cell that reduces π_t can switch to observing averages in the next round. Suppose there are two round t and $t+1$ such that $\pi_t > \pi_{t+1}$. Since $H(\mu_t) < H(\mu_{t+1})$(by assumption of no crucial cell) the remaining cells at $t+1$ has a lower probability of generating $Y = 1$.

This would imply the gain from observing an average or a cell decreases between t and $t+1$ due to a lower probability of obtaining $Y = 1$. However, a lower level of uncertainty would imply the averages cannot take many possible values reducing the information content of the averages. Since cells always take only two possible values, these would not be the case with cells. Thus if the DM was observing at period t it would optimal for him to observe cell in period $t+1$ as well.

This implies if a DM starts by observing a cell he would continue observing cells in the subsequent period and if the DM starts with averages he can continue observing only averages or switch to cell.

Note that, if c_a is sufficiently high, i.e., δ is sufficiently low then $(\overline{H}_l, \overline{H}_h)$ is smaller compared to $(\underline{H}_l, \underline{H}_h)$ for all π_t, similarly if c_l is sufficiently high, i.e., δ is sufficiently high then $(\underline{H}_l, \underline{H}_h)$ is smaller compared to $(\overline{H}_l, \overline{H}_h)$ for all π_t. In the former case observing only cells would be optimal and in the later case observing averages is optimal. □

B.9 Proof of corollary 1

Proof. Givem lemma 6 the DM would start by observing averages only if $H(\mu_0) \in (\overline{H}_l(\pi_0), \overline{H}_h)(\pi_0)$. However observing only averages would be optimal if either one of the three possible cases hold true,

Case 1 For some t DM discover a cell with $Y = 1$ by observing an average and for all $s \leq t$, $H(\mu_s) \in (\overline{H}_l(\pi_s), \overline{H}_h)(\pi_s)$. However in this observing only averages guarantess $Y = 1$, i.e., selective learning is not payoff relevant.

Case 2 For some t DM observes an average and updates his belief to μ_{t+1} such that $H(\mu_{t+1}) < \underline{H_l}(\pi_{t+1})$, i.e, no further learning is optimal and and for all $s \leq t$, $H(\mu_s) \in (\overline{H}_l(\pi_s), \overline{H}_h)(\pi_s)$. In this case average is sufficiently informative and c_l is not too small. Note that, here observing only averages does not guarantee $Y = 1$.

Case 3 Finally for some t the DM observes an average such that $H(\mu_{t+1}) \in (\underline{H_h}(\pi_{t+1}, \overline{H}_l(\pi_{t+1}))$ and for all $s \leq t$, $H(\mu_s) \in (\overline{H}_l(\pi_s), \overline{H}_h)(\pi_s)$. This can happen only if c_l is sufficiently high even though averages are not very informative. □

B.10 Proof of corollary 2

Proof. From theorem 1, we know given c_a if c_l is reduced, i.e, δ is reduced then the DM can change his strategy would become more likely to observe cells. Note that, if before the change in cost the DM was only observing cells then the reduction in c_l can only increase the incentive to learn for the DM. Since observing more cells increases the probability of finding a cell with a payoff of $Y = 1$ a reduction in cost would only reduce the probability of making mistake.

Hence we will only need to consider the cases where the DM was observing only averages or observing both cells and averages. We have shown that observing averages does not necessarily reveal the true state, hence due to reduction the δ if the DM chooses to observe cells as well then the precision of posterior cannot reduce.

Thus the only possible case is where the DM was observing both averages and cells but switch to observing only cells then the precision of posterior can decrease. However, we need to show this would be the optimal choice for the DM. Consider the following case, where the DM was observing averages before cells to narrow down the possible rows or column to find the optimal cell. However, after a change in δ he only observes cells.

This is possible where observing averages are not very informative, i.e, the DM needs to observe many averages to be able to find the optimal row or column to observe. In this case, if the cost of cells become sufficiently small such that the DM would rather start with cells. Since the uncertainty reduced by averages are different than that by cells, it is possible under the former strategy the net gain from learning cells were higher when the DM narrowed down the row or column to observe compared to the later when the DM opens several cells directly reducing the uncertainty significantly. In this case, the DM may stop short of learning the true state.

Consider the decision problem where $n = 3$ and the prior belief of the DM is that exactly five out of nine cells contain $Y = 1$ where each cell is equally likely to generate $Y = 1$. If the cost of observing cell, $c_l \in (.278, .3)$ then it is optimal for the DM to start with observing average and switch to observing cell from a different row or column if the observed average is .33.

Now if the cost of learning cells decreases to a value between $c_l \in (.267, .277)$ then the DM would optimally choose to observe at most one cell for sufficiently high c_a. The resulting uncertainty in the second case would be higher and would lead to a lower expected gross payoff.

□

B.11 Proof of corollary 3

Proof. From theorem 1 we know given c_l if c_a reduces, i.e, δ increases the DM can switch learning strategies. First of all, if at the initial level of δ the DM was only observing only averages then any reduction in cost can only increase learning, i.e, the DM would uncover weakly more averages leading to a more precise posterior belief.

So, we only need to consider the case where DM was either observing only cells or observing averages followed by cells. In case the DM was only observing cells if after the cost reduction he starts to observe averages as well then also the posterior precision cannot decrease. Since averages reduce more uncertainty for a higher level of uncertainty switching from only observing cell to observing averages followed by cells allows the DM to narrow down his search of cells and thus weakly increases the posterior precision. The same argument applies when the DM continues to observe both cells and averages.

Thus we need to consider the case where an increase in δ leads the DM to observe only averages when he was previously observing both or only cells. Consider the case where the DM was previously observing both, namely starting with averages and narrowing the possibilities and then observe cells but after a change in cost only observes average.

Since averages are not necessarily fully informative, switching to only averages can reduce the precision of posterior, i.e, increase the probability. But this would only be optimal if the cost of only observing average strategy is significantly lower. If averages become significantly cheaper such that the DM finds it optimal significantly more averages than before, then for some prior it is possible that after observing this bigger set of averages the additional gain from observing any cell becomes much smaller. If the cost of observing cell c_l is sufficiently higher compared to the smaller gain in payoff then the DM would optimally choose to not learn about cells reducing the precision of posterior.

Consider the following decision problem where $n = 3$ and the prior belief of the DM is that exactly six out of nine cells contain $Y = 1$ and all cells are equally likely to generate $Y = 1$.

Suppose initially $c_a < .133$ and $c_l \leq .33$ but $\delta < .2$. In this case, optimal strategy is to observe the average and then observe a cell if the value of the average is .67. The final payoff, in this case, is 1 and there is no probability of mistake.

If now c_a reduces significantly whereas the c_l remains the same such that $\delta > .2$ then observing only averages become optimal. But in that case, if there are exactly two cells in each row and column that gives $Y = 1$ the expected gross payoff is less than 1 and the probability of mistake is .33.

□

B.12 Proof of proposition 2

Proof. Consider a follower DM i in period $t = 1$ who chooses to observe the action of a leader DM j. If the DM i has chosen input combination (s_i, p_j) then the follower DM i knows that the expected gain from opening another cell is less than c_l and the expected gain from another average is less than c_a since DM i knows the cost of learning for j.

Given the common prior the DM i can update his belief by partitioning the state space since he knows the true state must belong to the block where the expected gain from learning given the choice of input strategy is bounded by the costs c_l and c_a. Let us denote this updated belief by $\mu_0^{i,j}$. Let $\mu_0^{i,j}(\omega)$ denote the probability for any state ω upon social learning.

For example, let $n = 3$, i.e., there are 9 possible input combination and the common prior is the uniform belief over all the production models where exactly one cell generates $Y = 1$. Suppose the costs of learning is such that DM would not stop learning unless there are only three cells left to be opened. If the follower DM i learns that the leader DM j has chosen cell (s_k, p_l) then DM i infers that either DM j has uncovered the cell (s_k, p_l) and has learned that it generates $Y = 1$ for sure or the DM j has opened 6 cells excluding cell (s_k, p_l) all of which have generated $Y = 0$. Since the follower DM i does not observe the learning strategy of j his updated belief would be such that the state where cell (s_k, p_l) has probability .92 and all other eight states have probability .01 of generating $Y = 1$.

However, since DM j may not have the same payoff matrix as DM i, he needs to incorporate his belief $\nu_{i,j}$ about the closeness of the two DMs. Given $\nu_{i,j}$ the updated belief post social learning for any state ω would be

$$\mu_0^S(\omega) = \nu_{i,j}\mu_0^{i,j}(\omega) + (1 - \nu_{i,j})\mu_0(\omega)$$

where $\mu_0(\omega)$ deonte the common prior belief and μ_0^Sdenote the belief upon social learning.

Then the DM i would choose a decision optimally based on μ_0^S. Note that, since the uncertainty in μ_0^{ij} would be lower than μ_0 as μ_0^{ij} is the result of learning by DM j the belief μ_0^S would also have a lower uncertainty than μ_0.

Compared to the common prior μ_0 since μ_0^S has lower uncertainty the incentives to learning change. Let us consider the extreme case where $\nu_{i,j} = 1$, the leader j has the same state as that of follower DM i, then given the same cost of learning it would not be optimal for i to learn further given his belief after social learning. This implies by corollary 2 and 3 if one of the two costs reduces substantially such that the DMs in period $t = 0$ switches to using only one type of learning strategy form observing both average and cell sequentially. This would, in turn, imply that the follower DM i would also choose not to learn further and experience lower gross expected payoff as the probability of error propagates through μ_0^S.

In case $\nu_j < 1$, i.e., the two DMs do not share the same payoff function, if $\nu_{i,j}$ is sufficiently large similar effects would take place. This is because for sufficiently high $\nu_{i,j}$ under the updated belief μ_0^S the net gain from learning is lower than the net gain under μ_0 and hence social learning would reduce the learning incentive of the follower DM i and the probability of error would also propagate. If any ϵ cost of social learning is introduced such that the follower DM i would only observe the action of DM j only if it reduces his private cost of learning then the result holds true. □

B.13 Proof of lemma 7

Proof. The proof consists of four steps. First, given an expected payoff π_t in round t, if there are two possible beliefs μ_1 and μ_2 such that $H(\mu_1) > H(\mu_2)$ and at μ_1 it is optimal for the DM to observe a cell then it would remain optimal for the DM to observe cell for μ_2 as well. Second, if at μ_2 observing an average is optimal then it would remain so under μ_1 as well. This would generate a cutoff value of $H(\mu)$ such that given expected payoff π_t, observing cells are optimal for the DM only when the uncertainty of the belief is lower than this cutoff and observing averages would be optimal when the uncertainty of belief is higher than the cutoff.

Third, given expected payoff π_t and informativeness of cells R_t^c, observing cell is optimal only if the level of uncertainty is within an interval. Fourth, given expected payoff π_t and informativeness of averages R_t^a, observing average is optimal only if the level of uncertainty is within an interval. This would generate the upper and lower bound on uncertainty such that learning is optimal only within the range of the upper and lower bound.

Step 1 : Let μ_1 and μ_2 be two possible beliefs such that $H(\mu_1) > H(\mu_2)$ but $\pi_{t,1} = \pi_{t,2}$. There are four possible cases, three of which is similar to lemma 6.

Case 1: Let $\mu_{t,x_{\bar{k}}}\pi_{t,(i,j),1} > \mu_{t,x_{\bar{k}}}\pi_{t,(i,j),1}$, i.e., μ_1 has a higher probability of generating $Y \geq x_{\bar{k}}$. This implies under μ_2 the DM optimally observes a cell that generates a lower probability of generating $Y \geq x_{\bar{k}}$ than the cell with highest expected payoff. This is optimal only if the value of learning is higher under μ_2 when observed cell generates $Y \geq x_{\bar{k}}$. Hence, there exists a pair of cells $(s_{i'}, p_{j'})$ and (s_k, p_l) such that the negative correlation between the two cells are higher under μ_2. Since a higher correlation among cells makes cells more informative compared to averages, if the DM was observing cells for μ_1 he would continue observing cells under μ_2.

Case 2: Let $\mu_{t,x_{\bar{k}}}\pi_{t,(i,j),1} = \mu_{t,x_{\bar{k}}}\pi_{t,(i,j),1}$ but $x_{\bar{k},1} > x_{\bar{k},2}$, i.e, even though the expected payoff from observing the optimal cell is same the set of possible values under μ_1 takes more extreme value compared $]mu_2$. With $m = 2$ this can never be the case. This implies under μ_2 there is lower value from learning for both types. However, if the same expected payoff is generated different set of values of x_k the averages cannot be able to distinguish between possible cases. Thus informativeness of average remains the same under both the belief. If observing cells were optimal under μ_1 then the DM would optimally choose either to observe cells or no learning at all.

Case 3: Let $\mu_{t,x_{\bar{k}}}\pi_{t,(i,j),1} = \mu_{t,x_{\bar{k}}}\pi_{t,(i,j),1}$ and $x_{\bar{k},1} = x_{\bar{k},2}$, but $\pi_{t+1|(i,j)<x_{\bar{k}},2} > \pi_{t+1|(i,j)<x_{\bar{k}},1}$, i.e, if the DM observes $Y < x_{\bar{k}}$ then the value of learning is higher under μ_2. Similar to the logic of lemma 6 this can happen only if there exists at least a pair of cells $(s_{i'}, p_{j'})$ and (s_k, p_l) such that the negative correlation between the two cells are higher under μ_2 than under μ_1 making the cells more informative under μ_2.

Case 4: Let $\mu_{t,x_{\bar{k}}}\pi_{t,(i,j),1} = \mu_{t,x_{\bar{k}}}\pi_{t,(i,j),1}$ and $x_{\bar{k},1} = x_{\bar{k},2}$, but $\pi_{t+1|(i,j)<x_{\bar{k}},2} < \pi_{t+1|(i,j)<x_{\bar{k}},1}$, i.e, if the DM observes $Y < x_{\bar{k}}$ then the value of learning is lower under μ_2. Again similar to lemma 6 this happens when a pair of cells are more positively correlated or given a level of correlation the dispersion in probability of generating a cell with $Y > x_{\bar{k}}$ is higher under μ_2. In both the two cases the increase in informativeness of the avearge is bounded by that of the cells making cells optima under μ_2 as well.

Step 2 : Suppose observing averages are optimal under μ_2. If uncertainty increases for μ_1 then of the the three possibilities can happen, one, the correlation between the cells decrease under μ_1 but $x_{\bar{k}}$ is same, two, $x_{\bar{k}}$ is same but the belief is more diffused under μ_1 and three, $x_{\bar{k},1} > x_{\bar{k},2}$ but the expected value is same.

If the correlation between cells decrease then the informativeness of the cells reduces more making averages optimal under μ_1 as well. If the belief becomes more diffused given a level of correlation cells are relatively less informative and observing average remains optimal under μ_1. Finally if $x_{\bar{k},1} > x_{\bar{k},2}$ then relative informativeness of average remain same but cells become less informative under μ_2 (same as case 2 in step 1). If observing averages was optimal under μ_2 it would remain so under μ_1 as well.

Combining step 1 and step 2, there exists a cutoff value of $H(\mu)$ such that learning cells is optimal below it and learning averages is optimal above it given π_t.

Step 3 : Similar to the proof of lemma 3 it can be shown that the value of learning cells is quasi-concave in residual uncertainty of cell, $E_t(H(\mu_{t+1}|B_{k<\bar{k},(i,j)}))$ and given R_t^cthere is a one-to-one relationship between $E(H(\mu_{t+1}|B_{k<\bar{k},(i,j)}))$ and $H(\mu_t)$.

The proof quasi-concavity is similar to lemma 3, higher uncertaintyon one hand, increases the expected number of cells to be observed leading to a higher cost of learning but on the other hand, also implies higher spread in possible values. Additionally a higher spread in values can imply higher likelihood of getting a higher value x_k through learning. Thus there are two opposing effects of higher uncertainty. For any belief μ_λ such that $E(H(\mu_{t+1,\lambda}|B_{k<\bar{k},(i,j)})) \in [E(H(\mu_{t+1,2}|B_{k<\bar{k},(i,j)})), E(H(\mu_{t+1,1}|B_{k<\bar{k},(i,j)}))$ if $V(\mu_2) \geq V(\mu_1)$ then the the effect of higher expected cost dominates thus $V(\mu_\lambda) \geq V(\mu_1)$. In the other case, if $V(\mu_2) < V(\mu_1)$ then the effect of higher spread dominates and thus $V(\mu_\lambda) \geq V(\mu_2)$ for a given value of $R^c_{t,(i,j)}$.

For the second part the residual uncertainty can be written as

$$E(H(\mu_{t+1}|B_{k<\bar{k},(i,j)})) = \frac{1}{1-\mu_{t,x_k}}\left[R_t^c H(\mu_t) + \ln(1-\mu_{t,x_k})\right]$$

which gives a one-to-one relationship between $H(\mu_t)$ and $E(H(\mu_{t+1}|B_{k<\bar{k},(i,j)}))$ given R_t^c and π_t.

Step 4 : Similar to proof of lemma 5 and step 3 it can be shown that the value of learning cells is quasi-concave in residual uncertainty of average, $E(H(\mu_{t+1}|k<\bar{k}, a_{i,j}))$ and given R_t^a there is a one-to-one relationship between $E(H(\mu_{t+1}|k<\bar{k}, a_{i,j}))$ and $H(\mu_t)$. This step is similar to step 3 since the value of learning increases in uncertainty as there are higher probability of obtaining a higher x_k (given π_t) and lower in uncertainty as the expected number of averages to be observed increases with uncertainty. Finally, the residual uncertainty can be written as

$$E(H(\mu_{t+1}|k<\bar{k}, a_{i,j})) = \frac{1}{1-\mu_{t,x_k}}\left[R_t^a H(\mu_t) + \ln(1-\mu_{t,x_k})\right]$$

which gurantees a one-to-one relationship between $E(H(\mu_{t+1}|k<\bar{k}, a_{i,j}))$ and $H(\mu_t)$.

Combining steps 3 and 4 generate a lower and upper bound on learning. However, since the bound for cells depends on R^c and that of averages depend on R^a, it is possible that for some belief the uncertainty is such that it is lower than the lower bound for average but higher than the upper bound of cells leading to no learning at all. Also, it is possible that the lower bound for a cell is higher than average implying observing only average is always better or the upper bound for average is lower than the upper bound for cell implying learning by cells always optimal. Hence, proved. □

C Extension: Noisy Observations

In the baseline model the learning technology is such that when the DM observes a cell or an average he can observe the payoff without any noise. In this context I will consider an otherwise similar model but will relax the assumption of no noise in observation.

Suppose, whenever the DM observes a cell or average he need to also choose with how much precision he would observe different payoff values. Let $\gamma^c_{1,t,(i,j)}$ and $\gamma^c_{0,t,(i,j)}$ denote the corresponding probabilities of observing $Y = 1$ and $Y = 0$ when observing (s_i, p_j) in period t and the true payoff is $\pi(s_i, p_j) = 1$ and $\pi(s_i, p_j) = 0$ respectively. Note that, it is not assumed that $\gamma^c_{1,t,(i,j)} = \gamma^c_{0,t,(i,j)}$, i.e, the DM can choose different level of precision contingent on the state. If $\gamma^c_{1,t,(i,j)} > \gamma^c_{0,t,(i,j)}$ then in the state where $\pi(s_i, p_j) = 1$ the precision of belief is higher than the case when $\pi(s_i, p_j) = 0$

Similarly for averages, let $\gamma^a_{p,t,(i,j)}$ be the probability of observing $\pi(a_{i,j}) = p$ in period t for average $a_{i,j}$ where $i \in \{s, p\}$ and $j \in \{1, \ldots, n\}$ when indeed $\pi(a_{i,j}) = p$. Also the choice of γ^a can be different for different values of p.

The cost of learning now depends only the level of precision both cells and averages. If the DM chooses to observe a cell and $\gamma^c_{1,t,(i,j)}$ and $\gamma^c_{0,t,(i,j)}$ as precisions of posteriors when $\pi(s_i, p_j) = 1$ and $\pi(s_i, p_j) = 0$ respectively then the cost of learning is as follows:

$$K_c(\mu_t, \gamma^c_{1,t,(i,j)}, \gamma^c_{0,t,(i,j)}) = \lambda_c \left[H(\mu_t) - E_{\mu_t}(\mu_{t+1}|\gamma^c_{1,t,(i,j)}, \gamma^c_{0,t,(i,j)})\right]$$

i.e., the cost of learning is the difference in the level of uncertainty, as measured by Shannon entropy between the prior and the posterior belief conditional on the choice of $\gamma^c_{1,t,(i,j)}, \gamma^c_{0,t,(i,j)}$. Also, λ_c is the marginal cost of learning a cell and there is no fixed cost of learning about different cells. Since observing (s_i, p_j) only partitions the state space into two different blocks, the difference between the round t and round $t+1$ entropy can be written in terms of the outcome of the observed cell only.

$$\begin{aligned} H(\mu_t) - H(\mu_t|\gamma^c_{1,t,(i,j)}, \gamma^c_{0,t,(i,j)}, \pi(s_i, p_j) = 1) = -\Big[&\mu_{t,(i,j)} \ln \mu_{t,(i,j)} + (1 - \mu_{t,(i,j)}) \ln (1 - \mu_{t,(i,j)}) \\ &- \gamma^c_{1,t,(i,j)} \ln \gamma^c_{1,t,(i,j)} - (1 - \gamma^c_{1,t,(i,j)}) \ln (1 - \gamma^c_{1,t,(i,j)})\Big], \end{aligned}$$

$$\begin{aligned} H(\mu_t) - H(\mu_t|\gamma^c_{1,t,(i,j)}, \gamma^c_{0,t,(i,j)}, \pi(s_i, p_j) = 0) = -\Big[&\mu_{t,(i,j)} \ln \mu_{t,(i,j)} + (1 - \mu_{t,(i,j)}) \ln (1 - \mu_{t,(i,j)}) \\ &- \gamma^c_{0,t,(i,j)} \ln \gamma^c_{0,t,(i,j)} - (1 - \gamma^c_{0,t,(i,j)}) \ln (1 - \gamma^c_{0,t,(i,j)})\Big]. \end{aligned}$$

Similarly, if the DM chooses to observe an average $a_{i,j}$ with precision $\gamma^a_{p,t,(i,j)}$ for all $p \in$

$\{0, 1/n, \ldots, k/n, \ldots, 1\}$ then the cost of learning is given by

$$K_a(\mu_t, \gamma^a_{p,t,(i,j)}) = \lambda_a \left[H(\mu_t) - E_{\mu_t}(\mu_{t+1} | \gamma^a_{p,t,(i,j)}) \right]$$

Here, also the entropy depends only on the observed average and all possible values teh avearge can possibly take and there are at most $n + 1$ such possible values.

One important observation is that, given the cost is measured in entropy the DM would always choose to observe a cell or an average only once. This is bacuase Shannon entropy is linear in the posterior. If the DM breaks learning a cell (s_i, p_j) in two steps such that in the first step he chooses $\gamma^c_{1,t,(i,j)} = \gamma^c_{0,t,(i,j)} = p$ and in the second step he chooses $\gamma^c_{1,t',(i,j)} = \gamma^c_{0,t',(i,j)} = q$ then the total cost of learning in period t would be

$$\begin{aligned} K_c(\mu_t, \gamma^c_{1,t,(i,j)}, \gamma^c_{0,t,(i,j)}) &= \lambda_c \left[H(\mu_t) - E_{\mu_t}(\mu_{t+1} | \gamma^c_{1,t,(i,j)}, \gamma^c_{0,t,(i,j)}) \right] \\ &= -\lambda_c \left[\mu_{t,(i,j)} \ln \mu_{t,(i,j)} + (1 - \mu_{t,(i,j)}) \ln (1 - \mu_{t,(i,j)}) - p \ln p - (1-p) \ln (1-p) \right], \end{aligned}$$

where $\mu_{t,(i,j)}$ is the probability that cell (s_j, p_j) would generate $Y = 1$ given belief at round t and in period t' it would be

$$\begin{aligned} K_c(\mu_{t'}, \gamma^c_{1,t',(i,j)}, \gamma^c_{0,t',(i,j)}) &= \lambda_c \left[H(\mu_{t'}) - E_{\mu_{t'}}(\mu_{t'+1} | \gamma^c_{1,t',(i,j)}, \gamma^c_{0,t',(i,j)}) \right] \\ &= -\lambda_c \left[p \ln p + (1-p) \ln (1-p) - q \ln q - (1-q) \ln 1 - q \right]. \end{aligned}$$

If instead he has chosen q in round t the cost of learning would be

$$K_c(\mu_t, \gamma^c_{1,t,(i,j)}, \gamma^c_{0,t,(i,j)}) = -\lambda_c \left[\mu_{t,(i,j)} \ln \mu_{t,(i,j)} + (1 - \mu_{t,(i,j)}) \ln (1 - \mu_{t,(i,j)}) - q \ln q - (1-q) \ln (1-q) \right],$$

i.e, the sum of the cost distributed in two periods. Thus cost does not reduce if the learning problem is divided into two steps. The benefit of learning linear and depneds on the posterior belief at the time of choice from $\mathbf{A} = S \times P$. Thus the benefit would not also change if the learning is into more than rounds. However, between t and t' he may find another cell that generates $Y = 1$ and can save the cost by learning about the t round cell with desired level of precisio, namely, q in round t only. Thus the DM would always choose to observe one cell or one average only once.

Also, given learning is costly and the main objective of the DM is to maximize net expected payoff the DM would always choose a cell when he observes $Y = 1$ as payoff. Note that, unlike the baseline model this does not imply he would get a payoff of $Y = 1$ unless he chooses to learn perfectly about the cell, i.e, $\gamma_{1,t(i,j)} = 1$. Since the DM uses the Bayes rule to update the probability that $\pi(s_i, p_j) = 1$ given the DM observes the payoff to be 1 would be

$$P(\pi^o(s_i, p_j) = 1 | \pi^o(s_i, p_j) = 1) = \frac{\mu_{t,(i,j)} \gamma_{1,t,(i,j)}}{\mu_{t,(i,j)} \gamma_{1,t,(i,j)} + (1 - \mu_{t,(i,j)})(1 - \gamma_{0,t,(i,j)})}$$

where π^o denote the observed value of the payoff and similarly the probability of $\pi(s_i, p_j) = 0$ given the DM observes the payoff to be 0 would be

$$P(\pi^o(s_i,p_j)=0|\pi^o(s_i,p_j)=0)=\frac{(1-\mu_{t,(i,j)})\gamma_{0,t,(i,j)}}{(1-\mu_{t,(i,j)})\gamma_{0,t,(i,j)}+\mu_{t,(i,j)}(1-\gamma_{1,t,(i,j)})}.$$

In the baseline model for the recursive problem there were three questions needed to answer, namely, first, whether to learn, second, what to learn, cell or average and third, which cell or average to learn about. One additional question that needs to be answered in this model, namely, how much to learn about, what level of precision to be chosen?

The expected payoff of the DM given choices of $\gamma^c_{1,t,(i,j)}$ and $\gamma^c_{0,t,(i,j)}$ when he decides to observe a cell would be

$$\begin{aligned}E_{\mu_t}\pi_t &= P(\pi^o(s_i,p_j)=1|\pi(s_i,p_j)=1)\times 1+P(\pi^o(s_i,p_j)=0|\pi(s_i,p_j)=1)\times E_{\mu_{t+1,(i,j)=1}}(\pi_{t+1})\\ &+P(\pi^o(s_i,p_j)=1|\pi(s_i,p_j)=0)\times 0+P(\pi^o(s_i,p_j)=0|\pi(s_i,p_j)=0)\times E_{\mu_{t+1,(i,j)=0}}(\pi_{t+1})\end{aligned}$$

The cost of choosing $\gamma^c_{1,t,(i,j)}$ and $\gamma^c_{0,t,(i,j)}$ are the same and the DM would optimally choose $\gamma^c_{1,t,(i,j)}$ and $\gamma^c_{0,t,(i,j)}$ such that the marginal benefit equals the marginal cost. The marginal benefit from $\gamma^c_{1,t,(i,j)}$ is given by

$$\begin{aligned}\frac{\partial E_{\mu_t}\pi_t}{\partial\gamma^c_{1,t,(i,j)}} &= \frac{\mu_{t,(i,j)}(1-\mu_{t,(i,j)})(1-\gamma^c_{0,t,(i,j)})}{\left[\mu_{t,(i,j)}\gamma^c_{1,t,(i,j)}+(1-\mu_{t,(i,j)})(1-\gamma^c_{0,t,(i,j)})\right]^2}\left[1-E_{\mu_{t+1,(i,j)=1}}(\pi_{t+1})\right]\\ &+\frac{\mu(1-\mu_{t,(i,j)})\gamma^c_{0,t,(i,j)}}{\left[(1-\mu_{t,(i,j)})-\gamma^c_{0,t,(i,j)}+\mu_{t,(i,j)}(1--\gamma^c_{1,t,(i,j)})\right]^2}E_{\mu_{t+1,(i,j)=0}}(\pi_{t+1})\end{aligned}$$

and the marginal benefit from $\gamma^c_{0,t,(i,j)}$ is given by

$$\begin{aligned}\frac{\partial E_{\mu_t}\pi_t}{\partial\gamma^c_{0,t,(i,j)}} &= \frac{\mu_{t,(i,j)}(1-\mu_{t,(i,j)})\gamma^c_{1,t,(i,j)}}{\left[\mu_{t,(i,j)}\gamma^c_{1,t,(i,j)}+(1-\mu_{t,(i,j)})(1-\gamma^c_{0,t,(i,j)})\right]^2}\left[1-E_{\mu_{t+1,(i,j)=1}}(\pi_{t+1})\right]\\ &+\frac{\mu(1-\mu_{t,(i,j)})(1-\gamma^c_{1,t,(i,j)})}{\left[(1-\mu_{t,(i,j)})-\gamma^c_{0,t,(i,j)}+\mu_{t,(i,j)}(1--\gamma^c_{1,t,(i,j)})\right]^2}E_{\mu_{t+1,(i,j)=0}}(\pi_{t+1})\end{aligned}$$

Note that, $(1-\gamma^c_{0,t,(i,j)})<\gamma^c_{1,t,(i,j)}$ and $\gamma^c_{0,t,(i,j)}>(1-\gamma^c_{1,t,(i,j)})$. This implies if $\left[1-E_{\mu_{t+1,(i,j)=1}}(\pi_{t+1})\right]>E_{\mu_{t+1,(i,j)=0}}(\pi_{t+1})$ then the marginal benefit of $\gamma_{0,t,(i,j)}$ is higher, whereas for $\left[1-E_{\mu_{t+1,(i,j)=1}}(\pi_{t+1})\right]<E_{\mu_{t+1,(i,j)=0}}(\pi_{t+1})$ the marginal benefit of $\gamma_{1,t,(i,j)}$ is higher. Thus in the former case $\gamma_{0,t,(i,j)}>\gamma_{1,t,(i,j)}$ and in the later case $\gamma^c_{0,t,(i,j)}<\gamma^c_{1,t,(i,j)}$ for all $\lambda_c>0$.

In both the two equations the first term measures the extent of type I error multiplied by the loss in payoff due to the type I error, i.e, concluding $Y=0$ when indeed $Y=1$. Similarly, the second term measures the extent of type II error multiplied by the implied loss, i.e, concluding $Y=1$ when indeed $Y=0$. Thus if the loss due to type I error is higher the DM would choose a higher $\gamma^c_{0,t,(i,j)}$ and vice-versa.

A similar observation can be made about averages as well. The marginal benefit from changing the precision when $Y = p$ depends on the change in the probability of relevant error and the loss due to error. The higher the loss due to an error the more likely the DM would choose a γ_p^a that minimizes the error.

Another important observation is that optimal cell to be observed would now be characterized by highest one-period *net* expected payoff and the optimal average would be characterized by the highest uncertainty reducing average net of cost of learning. This follows directly from lemma 2 and lemma 4 with the addition that the cost of learning now depends on the level of uncertainty, thus observing two cells or two averages are not equally costly and the net benefit accounts for the cost difference.

Finally I want to show that an equivalent of lemma 6 also holds true for this specification. Before stating the lemma let us redefine the information content in terms of observed value of payoff rather than actual value of payoff.

$$R_{i,j}^{c,k}(\mu) = \frac{\sum_{\omega \in B_{0,(i,j)}} \mu(\omega|\pi^o(s_i, p_j) = 1) \ln(\mu(\omega|\pi^o(s_i, p_j) = 1))}{H(\mu)}.$$

$$R_{t,(i,j)}^{a,k} = \frac{\sum_{\omega|Y<1} \mu_t(\omega|\pi^o(a_{i,j}) = p) \ln \mu_t(\omega|\pi^o(a_{i,j}) = p)}{H(\mu_t)}$$

Also, the information content of cells and averages are defined as the information content of the optimal cell and optimal average respectively.

Lemma 8. *At any round t given expected payoff π_t, information content for cell and average $R^{c,k}$ and $R^{a,k}$ and marginal costs λ_c and λ_a, the optimal learning strategy in round t is as follows: for $\underline{H}_l \le \underline{H}_h \le \overline{H}_l \le \overline{H}_h$*

1. *If uncertainty is in the interval $(\overline{H}_l, \overline{H}_h)$ then it is optimal for the DM to uncover an average,*

2. *If uncertainty is in between $(\underline{H}_l, \underline{H}_h)$ then DM optimally chooses to uncover a cell in the matrix*

3. *No learning everywhere else.*

Proof. Similar to lemma 7 the proof of this lemma consists of two major steps, one for any two beliefs μ_1 and μ_2 such that $H(\mu_1) > H(\mu_2)$ it cannot be true that the DM observes average for μ_2 but cell for μ_1 and two, there is a lower and upper bound in terms of uncertainty such that DM would never learn outside the interval generated by the two bounds.

Step 1 : First, let us consider two beliefs μ_1 and μ_2 such that $H(\mu_1) > H(\mu_2)$. Suppose at μ_1 the DM chose to observe a cell, then it needs to be shown that for μ_2 also the DM would choose to observe a cell. The proof is similar to that of lemma 6, there are four possibilities, namely, $\pi_{1,t,(i,j)}^k > \pi_{2,t,(i,j)}^k$, $\pi_{1,t,(i,j)}^k < \pi_{2,t,(i,j)}^c$, $\pi_{1,t,(i,j)}^k = \pi_{2,t,(i,j)}^k$ with $\pi_{2,t+1,(i,j)=0} > \pi_{1,t+1,(i,j)=0}$ and finally $\pi^k_{1,t,(i,j)} = \pi^k_{2,t,(i,j)}$ with $\pi_{2,t+1,(i,j)=0} < \pi_{1,t+1,(i,j)=0}$. However, if $\pi_{1,t,(i,j)} < \pi_{2,t,(i,j)}$ given π_t, where π^k denotes the net expected payoff subject to the cost of learning. Also from lemma 6, it cannnot be the case that $\pi^k_{1,t,(i,j)} < \pi^c_{2,t,(i,j)}$ if $H(\mu_1) > H(\mu_2)$, so we need to consider only three possible cases.

Case 1 : $\pi^k_{1,t,(i,j)} > \pi^k_{2,t,(i,j)}$, i.e, there exists at least a pair of cells (s_i, p_j) and (s_k, p_l) such that the negative correlation between the cells are higher under μ_2 compared to μ_1 and the DM optimally chooses to observe one of them in round t. However, a higher negative correlation increases both $(1 - E_{\mu_{t+1,(i,j)=1}}(\pi_{t+1}))$ and $E_{\mu_{t+1,(i,j)=0}}(\pi_{t+1})$ increasing the marginal gain from learning about cells. Since the DM optimally chooses marginal benefit equal to marginal cost of learning, his net gain from learning cells can only increase under μ_2. The marginal gain from avearges can also increase in case the pair of cells are in the same row or column but the gain in avearges in bounded by the gain in cells. Thus if the DM chooses to observe cells under μ_1 then he would continue so under μ_2 as well.

Case 2 : $\pi^k_{1,t,(i,j)} = \pi^k_{2,t,(i,j)}$ with $\pi_{2,t+1,(i,j)=0} > \pi_{1,t+1,(i,j)=0}$. This case also implies that there exists a pair of cells (s_i, p_j) and (s_k, p_l) such that the negative correlation between the cells are higher under μ_2 compared to μ_1 and the DM optimally chooses to observe one of them in round t. The rest of the analysis is same as before.

Case 3 : $\pi^k_{1,t,(i,j)} = \pi^k_{2,t,(i,j)}$ with $\pi_{2,t+1,(i,j)=0} < \pi_{1,t+1,(i,j)=0}$. This implies either there exists a pair of cells (s_i, p_j) and (s_k, p_l) such that the positive correlation between the cells are higher under μ_2 compared to μ_1 and the DM optimally chooses to observe one of them in round t or the level of correlation between the two beliefs are the same for the observed cell but in case $\pi(s_i, p_j) = 0$ the probability of obtaining $Y = 1$ is lower.

In case the positive correlation increases, both the two values $(1 - E_{\mu_{t+1,(i,j)=1}}(\pi_{t+1}))$ and $E_{\mu_{t+1,(i,j)=0}}(\pi_{t+1})$ decreases, making the net gain from observing cells lower. If the pair of cells is in different rows and columns then the net gain from averages decreases in a similar way to the cell. If observing cells were optimal at μ_1 it would be so under μ_2 as well. However, if the pair of cells is in the same row or column then observing cells reduces the need for further learning compared to average. This is because if the DM observes one of them to be $Y = 0$ then learning about another cell generates very low net benefit. However, if the DM were to observe the averages he cannot differentiate perfectly which of the cells generates $Y = 0$ thus generating a greater need to learn. Thus if the DM were to choose cells at μ_1, he would continue doing so under μ_2 as well.

Finally, if the overall probability if finding another cell with $Y = 1$ conditional on $\pi(s_i, p_j) = 0$ then the incentive to learn decreases for both averages and cells in the similar proportion. If observing cells were optimal at μ_1 then the DM wither chooses to not to learn at all or continue observing cells since averages do not become relatively better.

A similar logic can be applied to show that if under μ_2 the DM was choosing averages then he would continue choosing averages under μ_1 as well. Since μ_1 has higher uncertainty one of the three cases can happen, either the negative correlation between a pair of cells that the DM would optimally observe is lower under μ_1 or the positive correlation for the same is lower or under μ_1 the probability of generating $Y = 1$ is higher conditional on $\pi(s_i, p_j) = 0$.

If the negative correlation is lower for μ_1 then the net gain for cells are lower by the same logic as case 1 and 2 and hence if the DM were to observe average under μ_2 he would not switch to cells under μ_1 since cells do not become relatively more informative.

In case positive correlation is higher or there are fewer possibilities under μ_2 that would generate $Y = 1$ the averages can take more possible values under μ_1 with flatter belief. If the average can take more possible values with flatter beliefs the reduction in uncertainty would be higher from observing an average under μ_1. Thus even if the net value from learning by cells due to a higher probability of generating $Y = 1$ under μ_1 the gain from averages come from learning the level of payoff, i.e, the possible number of cells generating $Y = 1$. Thus if the DM chooses to observe an average under μ_2 then he would continue so under μ_1 as well.

These observations generate a cutoff value of $H(\mu)$ such that observing averages is only optimal if the uncertainty is above the cutoff and observing cells are optimal are optimal below the cutoff.

Step 2 : Similar to lemma 3 there is two opposing impact of uncertainty on the benefit of observing a cell. For higher residual uncertainty the expected number of cells to be observed is higher and for lower residual uncertainty, observing one more cell cannot increase payoff by much. However, now the cost of learning also depends on the level of uncertainty where residual uncertainty refers to the level of uncertainty given the DM has observed $\pi(s_i, p_j) = 0$, Similar to lemma 3 there is a one-to-one relationship between residual uncertainty $H(\mu_{t+1}|\pi^o(s_i, p_j) = 1)$ and uncertainty $H(\mu_t)$ given $R_t^{c,k}$.

For very high uncertainty μ_t given a level of precision $\gamma_{1,t,(i,j)}$ and $\gamma_{0,t,(i,j)}$ is higher. This implies for high uncertainty the impact of higher cost of learning further reduces the net benefit from learning. On the other hand for a very low level of uncertainty since the Shannon entropy is convex in the posterior belief the marginal cost of updating beliefs become very high whereas the gain from learning is linear and low due to low uncertainty. These two observations combined generates an upper and a lower bound on the uncertainty of belief such that learning is optimal only for beliefs where uncertainty is within the interval.

Similarly following the steps of lemma 5, the level of residual uncertainty has two opposing effect on the benefit of learning and given $R_t^{a,k}$ there is a one-to-one relationship between the residual uncertainty $H(\mu_{t+1}|\pi^o(a_{i,j}) < 1)$ and $H(\mu_t)$. Also, the expected cost if higher for high uncertainty and the marginal cost is higher for lower uncertainty further strengthening the result from lemma 5. Thus there is an interval on the uncertainty of the belief, where learning about averages is only optimal if the uncertainty of the belief μ_t lies within the interval. Combining step 1 and step 2 all the four cutoffs can be generated. Hence, proved. □

Proposition 3. *The result of theorem 1 holds true for $\lambda_c, \lambda_a \geq 0$.*

Proof. Proof of the proposition follows directly from lemma 8 and the assumption that there are no crucial cells. The only concern is that as the DM uncovers more cells the expected payoff changes which can, in turn, reduce the cutoff value at which learning cell becomes optimal and thus even though the uncertainty is lower for later round the incentive to learn averages increase.

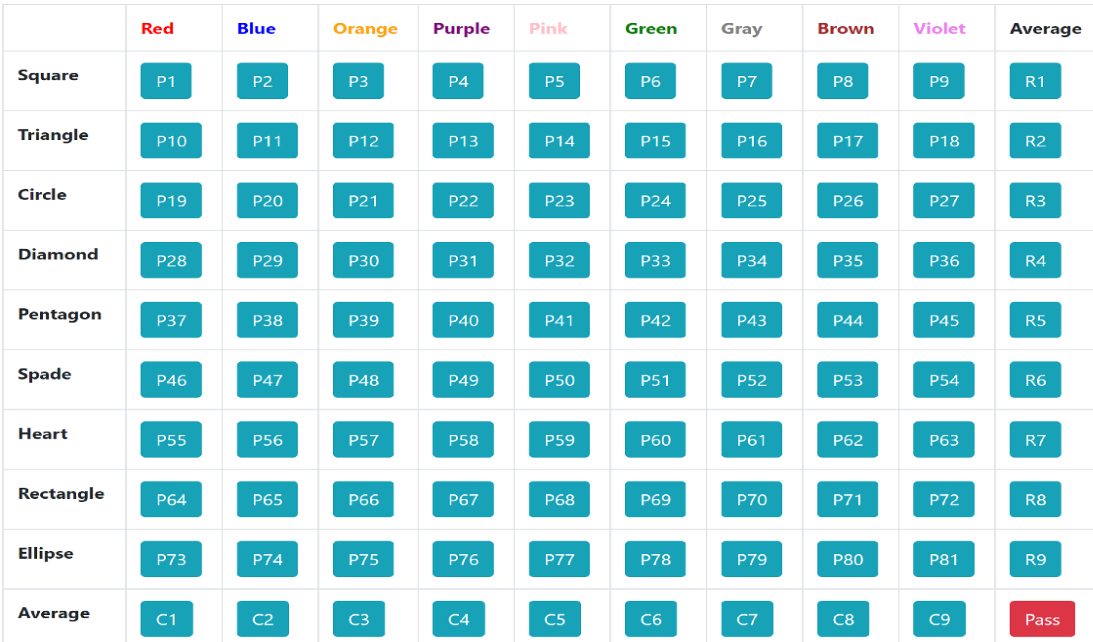
If the expected payoff increases with learning then the cells become more negatively correlated, i.e, observing $\pi^o(s_i, p_j) = 0$ becomes more informative. Even if the average becomes more informative as well, the net gain is bounded by the that of the cells. Thus the DM would not switch to observing averages optimally.

If the expected payoff decreases along with a decrease in uncertainty it must be the case that there are fewer possibilities where $Y = 1$ can be obtained. Thus the possible values that the averages can take decreases reducing the information content of the average. Also, the lower the probability of generating $Y = 1$ lowers the benefit of learning from both cells and averages. Since the net benefit from averages reduce due to two effects lower uncertainty and lower expected payoff whereas cells are only affected by lower expected payoff then DM would not switch to observing averages.

The convexity in the Shannon cost implies the drop in the cost of learning for averages cannot dominate the drop in uncertainty since averages contain more uncertainty than cells. Thus if the DM were choosing cells earlier he would continue so after a reduction in expected payoff and uncertainty. Hence, proved. □

D Appendix: Experiment

D.1 Learning Task



	Red	Blue	Orange	Purple	Pink	Green	Gray	Brown	Violet	Average
Square	P1	P2	P3	P4	P5	P6	P7	P8	P9	R1
Triangle	P10	P11	P12	P13	P14	P15	P16	P17	P18	R2
Circle	P19	P20	P21	P22	P23	P24	P25	P26	P27	R3
Diamond	P28	P29	P30	P31	P32	P33	P34	P35	P36	R4
Pentagon	P37	P38	P39	P40	P41	P42	P43	P44	P45	R5
Spade	P46	P47	P48	P49	P50	P51	P52	P53	P54	R6
Heart	P55	P56	P57	P58	P59	P60	P61	P62	P63	R7
Rectangle	P64	P65	P66	P67	P68	P69	P70	P71	P72	R8
Ellipse	P73	P74	P75	P76	P77	P78	P79	P80	P81	R9
Average	C1	C2	C3	C4	C5	C6	C7	C8	C9	Pass

Figure 7: Learning task: 9×9 matrix with averages

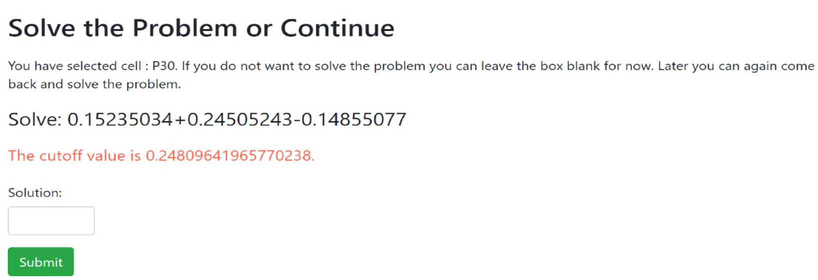


Figure 8: Learning task: sample cell question

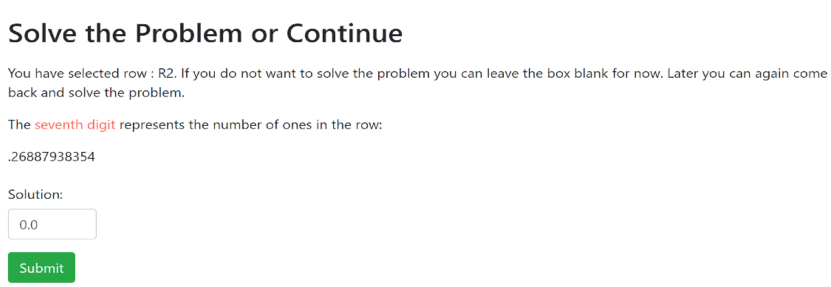


Figure 9: Learning task: sample average question

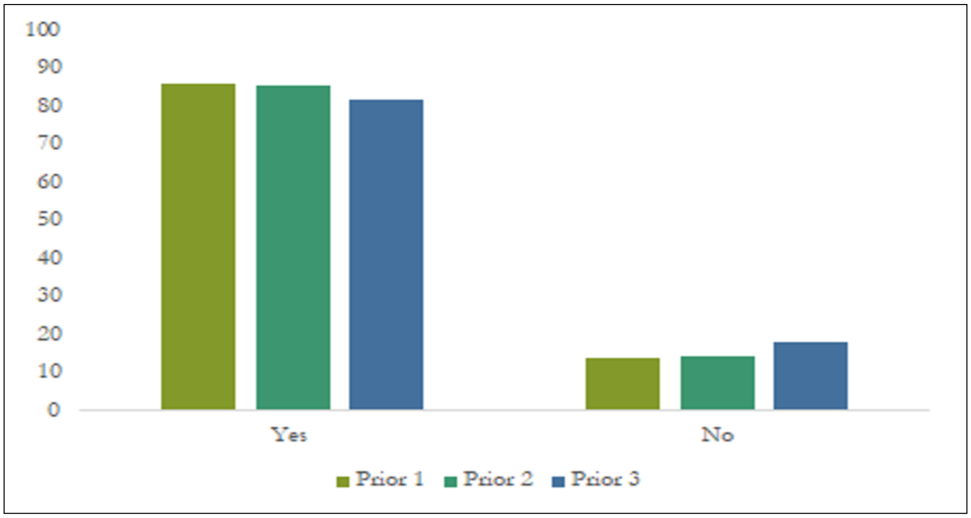


Figure 10: Percentage of rounds with no violation

	Satisfy-Violate
All priors	13.1182654
Prior 1	8.699706741
Prior 2	7.144350918
Prior 3	7.046855746

Table 15: One sided t-stat for proportion difference

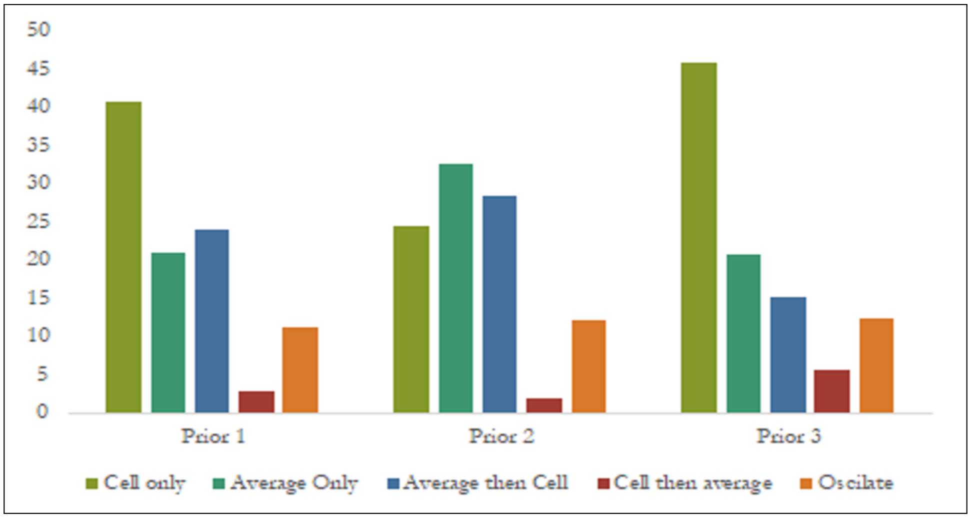


Figure 11: Chosen learning strategies in percentage

	cell then average	oscillate
cell only	8.053193547***	5.671363945***
average only	5.675196169***	2.968255872***
average then cell	5.28156649***	2.518898785***

(a) All prior: one-sided t-stat for difference is proportions

	cell then average	oscillate
cell only	9.780305957***	7.282821705***
average only	5.803699627***	2.788364241***
average then cell	6.411276004***	3.483703197***

(b) Prior 1: one-sided t-stat for difference is proportions

	cell then average	oscillate
cell only	7.217577097***	8.467816905***
average only	6.08747754***	2.994086108***
average then cell	6.77413281***	3.774260891***

(c) Prior 2: one-sided t-stat for difference is proportions

	cell then average	oscillate
cell only	9.908199093***	8.002215876***
average only	4.653925448***	2.328443077**
average then cell	3.257123621***	0.82966687

(d) Prior 3: one-sided t-stat for difference is proportions