Costly Social Learning and Rational Inattention

Srijita Ghosh
Dept. of Economics, NYU

September 19, 2016

Abstract

We consider a rationally inattentive agent with Shannon's relative entropy cost function. The agent also has a choice to observe the actions of agents from earlier generation subject to some cost, which is called the cost of social learning in the paper. We characterize the equilibrium based on the marginal cost of relative entropy referred to as private learning in the model. Given any non-concave cost of social learning function we show that as the marginal cost of private learning increases from zero, the number of agents observed from earlier generation does not change monotonically. For very high marginal cost of private learning, no learning of any type becomes optimal choice. To illustrate we also consider a special case where cost of observing up to some fixed number of agents, $\bar{c}$, is zero and very high for any higher number of observations. Even under this cost structure some agents would optimally choose strictly less than $\bar{c}$ agents to observe in equilibrium. Contrary to the herding literature, we have found that the private and social learning would be complements for higher marginal cost of private learning. We also find that improving social connectivity as measured by our cost of social learning may not be welfare improving always.

1 Introduction

To rationalize a suboptimal behavior of an economic agent which is most common in real world, economists often propose information friction as a plausible explanation, namely, they argue that an agent is choosing a suboptimal option because the agent does not perfectly know the payoff from all available alternatives. In the long tradition of incorporating information in choice problems many different forms of information frictions has been considered based on the particular problem at hand. E.g, the agent *does not have a choice to learn* and he only receives some information exogenously. Or the choice of acquiring information is *subject to a cost function* and these costs may be psychological or physical in nature. The physical cost is often due to the cost of buying the information and the psychological cost may reflect some mental cost of attention, memory or cognition required to process the information.

Similarly many different forms of acquiring information has also been considered in the literature. Agents may do experiments or take tests which often are constrained by the mental capacity or time available. In these experiments agents get some private signals about their payoffs or other variables of interest and use these signals in their choice problem. Instead of learning on their own, agents can directly observe the action of other

players in the economy and imitate them. Also instead of actual choice some *publicly observed signals* about the actions taken by other agents or the underlying state in the economy can be a source of information. Different information acquisition structures would give rise to different types of behavior of a rational agent and depending on the particular case, one information structure may be more natural to assume than any other.

If we broadly classify learning into two types, private learning, where agents learn on their own and gets a private signal and social learning, where agents learn from observing others or a public signal observable by everyone then we want to understand how these two types of learning interact? If a society becomes more connected it should become easier to obtain more information via social learning. But what would be the welfare implication of such an increased connectivity?

In this context of interaction between two models we ask the following questions. Suppose an agent is faced with a one-time choice between finite alternatives and wants to learn about the payoffs from different actions where learning is subject to some form of information friction. If the agent has options to both learn on his own and observe others' actions then how would he *optimally choose to learn*? Would a social planner choose the same amount of learning for him? Most importantly how does this individual choice get affected by the amount of information available in society? Specifically how would the structure of the society, namely the connectivity between agents in the society affects the choice problem of an individual agent?

But to answer these questions we need the following ingredients: first, a mechanism for private learning, second a mechanism for social learning and third a model of social connectivity. The psychology literature for several decades has studied private learning models. Economists have borrowed several ideas from this literature and has constructed two major learning model, namely *reinforcement learning* where agents act according to their experience from past choices to choose the alternatives that has paid well in past and *belief learning* where agents update their beliefs based on acquired information and act accordingly. For a survey of such learning models in economics refer [5].

Here we consider a one-time action by an agent and would use the belief learning approach where agents would update their posterior belief about their payoff based on some signal. Since our interest is to understand the trade-off between two types of learning we need to impose a cost structure on private learning, otherwise the agents would fully learn about their types and take the right choice always. The cost of private learning is borrowed from the *rational inattention* literature where the cost of learning is understood as a cognitive cost which is often the case with belief learning. The other option would have been to impose a costly and not fully informative signal structure where agents learn only partially. But the cost structure in the rational inattention literature helps us to assume away any particular complicated form of costly signal structure.

The *rational inattention* literature considers the discrete choice problem faced by an individual decision maker when information is costly to acquire. Following Sims (2003) [18], the cost of information has been modeled as a function of the *Shannon's relative entropy* between the prior belief and the expected posterior belief of the individual. This helps to abstract away from the detail modeling of the signal structure (as shown by Matejka and McKay(2014)) . Matejka and McKay (2014) [17] also showed that using a liner function of Shannon's relative entropy cost (where no information has zero cost) the optimal choice of a decision maker who is maximizing his expected utility takes the form of *multinomial logit*. Several other papers in the literature, e.g., Caplin and Dean (2015) [7], Caplin and Martin (2015) [10], Caplin ,Dean and Leahy (2016) [8] etc

has attempted to provide behavioral assumptions for the relative entropy cost function. More recently paper by Caplin, Leahy and Matejka(2015) [9], tried to combine social learning to generate a prior belief, to a model of rational inattention. In their paper any agent in period $t > 0$ gets to see the market share of every other generations before him without paying any cost. They found that observing the market share of many commodities affects the private learning and subsequently the optimal behavior of the agent in the model.

For the social learning mechanism we borrow from the *herding* literature. The *herding* literature following Banerjee (1992) [2] and Bikhchandani, Hirshleifer and Welch(1992) [3] considers an individual choice problem in an economy where agents make decisions sequentially. Agents are exogenously and randomly given with a payoff-relevant signal that they can use as their private information. All agents are ex-ante identical and any individual would observe the actions taken by all previous agents. Given everyone has an equally informative signal structure the actions of the previous agents also act as a source of information. But since, agent's action is not a sufficient statistic of his information, observing just the actions of some previous agents may generate some sub-optimality in the behavior of later agents. In equilibrium, agents might ignore their own private signal and choose to follow others' action blindly (the phenomenon known as "*herding*") even when it is not be optimal to do so. This generates the *herd externality* in the economy. In the model, *private signal* and *actions of others*, the two sources of information act as *substitutes*.

The social learning protocol used in this model is even simpler. Any agent in current generation can observe only the actions taken by other agents in only the last generation. So given the total number of agents are constant every period the maximum number of people that an agent can observe remains constant over time, which simplifies the model to a great extent. Notice that agents don't observe any noisy signals about the choices or information of others but observes their actual choices. Thus a later generation agent while updating his belief takes into account the possibility of mistakes made by earlier generation due to incomplete learning given the assumption that the agents are drawn randomly from a large enough population.

The final part required to close this model is the social connection structure. One of the most common way of imposing a social connection structure is to introduce an underlying network. But to avoid the complexity of a formal network model while retaining the notion of importance of social connectivity we introduce a cost function associated with social learning. Social learning in the model simply means observing the action of others from previous generation and putting a cost structure into it would mean agents need to incur some cost to observe actions of other agents. This is similar to the idea of social connectivity. In a more connected society it should easier for an agent to observe others which we represent with a lower cost of social learning. Also unlike herding literature, whether or not an agent would observe socially and when he decides to observe then how many agents he would observe from the previous generation is a choice of the individual but subject to a constraint imposed by the available technology of social connectivity, namely the cost of social learning function. Also we use the assumption that agents have to decide how many other agents to observe ex-ante before he actually observes anyone and after he choice is made the chosen number of agents are drawn randomly from earlier generation. This protocol is called block learning.

The main difference between using a network and this costly social learning model is the loss the heterogeneity across agents in this model. In our baseline model we assume

the cost of social learning function is same for everyone, so everything else being the same all agents would choose to observe the same number of agents. That is usually not the case in a network as agents have different degrees and can possibly observe different number of agents based on his degree centrality in the network. Instead of thinking the internal structure of the connectivity, our cost function actually refers to the technological level of connectivity available to a generation of agents and hence all agents in the same generation face the same constraint. Also once agents optimally decide how many agents to look at the sample is drawn randomly from the earlier generation. This gives anonymity of the observed agents which is a major simplification over a network structures where agents can only observe their neighbors who are not anonymous in their identity. Later we add an extension of the model where agents have different cost of social learning function which introduces the heterogeneity whereas keeping the anonymity of the observations.

Finally we combine these three parts to form a model and formalize the following questions. Consider an infinite period economy where a fixed number of agents enter every period and face a one-time choice between a finite number of actions, where the payoff from each action is not known to the agents. The agents are *rationally inattentive* and also has the option to observe the actions of agents in previous generation subject to a cost function, namely the "*cost of social learning*". In this model what would be the optimal choice of a rational agent who wants to maximize his expected payoff? Would herding still remain as a possible equilibrium? Would social and private learning remain *substitutes* as assumed in the herding literature or would they be *complements*? Moreover what would the policy implications of the model?

Surprisingly, we have found that the amount of social learning does not change *monotonically* with an increase in the marginal cost of private learning as would be the case if they were substitutes. For moderately high marginal cost of private learning, the two types of learning rather act as complements, furthermore when cost of private learning is very high, instead of choosing social learning entirely, the agents do not learn in any form at all. This implies that the policy suggestion coming from herding literature to restrict the learning of some agents initially to avoid herding is no more an optimal policy and crucially depends on the relative level of private learning technology and connectivity available in the society.

The rest of the paper is arranged as follows. Section 2 formally describes the two cost structures and sets up the baseline model. In section 3 we solve the agent's optimization problem and show the nonmonotonicity result. Section 4 discusses an interesting and instructive example where the social cost function takes the form of a capacity constraint. In this case agents can observe upto $\bar{n}$ many actions free of cost and no further observation is possible. This is closest to the idea of network structure without the heterogeneity, as if the degree of each agent is same in the economy. Section 5 discusses the following important extensions. First we introduce a payoff relevant aggregate state and see how that changes the incentive of social learning and whether the main result remains true. Then we introduce two forms of heterogeneity, namely heterogeneous private cost of learning and heterogeneous cost of social learning. The first reflects the difference of ability in the economy and the second one makes it similar to a network with a varied degree distribution. We show that in both cases the main result still holds true. Finally we consider a variant of the social learning protocol in which instead of block learning we use sequential learning protocol where agents chooses the number of observations sequentially based on his available information. Section 6 defines the steady state of this economy and discusses the herding behavior in a steady state. Section 7 concludes and

the main proofs are given in the Appendix.

2 Model

2.1 Environment

Consider an infinite horizon economy in discrete time, i.e. $t \in \{0, 1, \ldots, \infty\}$. At each period t a large but finite number of agents, N, enter the economy. At any period $t \geq 0$ when an agent enters the economy he chooses to learn, takes an irreversible action and leaves the economy never to come back again.

Let A be the finite set of actions that an agent can choose from. An agents doesn't know his idiosyncratic payoffs from taking any action $i \in A$. Let Ω be the set of all possible strict rankings of payoffs in A, hence Ω would be a finite state space. Let $\omega \in \Omega$ be a typical element in Ω which would be called the type of an agent. In the rest of our analysis we would assume $A = \{a, b\}$. Hence, $\Omega = \{\omega_1, \omega_2\}$ where $\omega_1 \equiv a \succ b$ and $\omega_2 \equiv b \succ a$. Let Γ be the set of possible distributions over Ω, that is $\Gamma \equiv \Delta(\Omega)$.

Let $\Delta(\Gamma)$ denote the set of all possible distributions over Γ. At any period $t \geq 0$ agents enter with a common prior $\gamma \in \Delta(\Gamma)$[1]. After entering, the agent tries to learn about his own type (idiosyncratic payoff) and then takes an irreversible one time action from the set A. Let μ^* denote the true distribution of types where $\mu^* \in int(\gamma)$.

Let $u : A \times \Omega \to \mathbb{R}$ be the state dependent utility function. Assume that agents are Bayesian expected utility maximizer. Let the payoff of different types of agents be given by

$$\begin{aligned} u(a, \omega_1) &= u(b, \omega_2) = \bar{u} \\ u(a, \omega_2) &= u(b, \omega_1) = \underline{u} \end{aligned} \tag{1}$$

where $\bar{u} > \underline{u}$, so type ω_1 gets a higher payoff from action a and type ω_2 gets a higher payoff from action b. Define $\Delta u = \bar{u} - \underline{u}$, the gain in payoff by matching over mismatching the state, which is assumed to be symmetric for both types.

2.2 Costly learning

Agents have two possible choices of learning, viz, private and social learning. Both these types of learning are costly to incur and also are different in nature as, the social learning gives information regarding μ^*, the true distribution of types, and private learning is targeted towards the own type, ω of the agent.

2.2.1 Private learning

The way this model sets up the private learning problem is similar to the recent literature on rational inattention as discussed in the introduction (refer to [17], [9]). This implies an agent can learn privately in the following way, namely if he chooses to learn privately

[1]We assume that γ is not biased towards an alternative, i.e. there exists a $\tilde{\lambda}$ such that $\forall \lambda \leq \tilde{\lambda}$, where λ be the marginal cost of private learning(as discussed later), and for all $0 \leq n \leq N$ if an agent observes $0 \leq x_n \leq n$ many action a (or b) the belief about $E\left(Pr(\omega_1) | \gamma_0, n\right)$(generated by Bayesian updating) will be on both sides of $\mu = 1/2$.

about the idiosyncratic state, i.e their own type ω, then he needs to pay some cost to update his posterior belief about ω, where the updating is done using Bayes rule.

The cost of private learning is given by Shannon's relative entropy cost of information . Let $P\left(i,\omega|\mu\right)$ be the posterior probability of choosing action $i \in A$ when type is $\omega \in \Omega$ and prior $\mu \in \Gamma$. Define $P\left(i|\mu\right) \equiv \sum_{\omega\in\Omega} \mu\left(\omega\right) P\left(i,\omega|\mu\right)$ as the prior probability of choosing action $a \in A$. The cost of private learning is given by,

$$C\left(\lambda,\mu\right) = \lambda \left\{ \underbrace{\sum_{\omega\in\Omega} \mu\left(\omega\right) \sum_{a\in A} P\left(a,\omega|\mu\right) \ln P\left(a,\omega|\mu\right)}_{\text{expected entropy of the posterior distributions}} - \underbrace{\sum_{a\in A} P\left(a|\mu\right) \ln P\left(a|\mu\right)}_{\text{entropy of the prior distribution}} \right\} \tag{2}$$

where $\lambda \in [0,\infty]$ be the marginal cost of private learning.

Instead of modeling as a choice over signal structures we assume the agents can directly choose a distribution over posterior distribution. As mentioned in [17] this is true because of Blackwell informativeness of different signal structures and the proposition 1 in [14]. The logic behind this equivalence is that an agent would never choose a signal structure where two distinct signal would give the same recommendation in terms of action choice due to Blackwell informativeness criterion which says that dropping one of the signals from the signal structure would be welfare improving. This implies for every signal in the chosen structure we can assign an unique payoff attached to it which helps us to use the result from [14]. Their result tells us we can interchangeably use signal structure and distribution of posteriors.

Given that we are using a belief learning model this formulation helps us to abstract away particular cost and signal structure and gives us a more general setup in terms of the choice of signal structure for private learning mechanisms.

2.2.2 Social learning

The social learning mechanism is similar to that of the herding literature with some major innovations. Similar to the herding literature an agent can observe the action but not the information or belief of the agents who has already decided on their choice. To fit our model we use agents from earlier generation and restrict that the agents can only observe the action of the agents form the preceding generation and no other generations.

Thus any agent at any period $t \geq 1$ can observe the action of $t-1$ generation agents subject to a cost, which would be called the cost of social learning. This is the major difference from the existing literature. The introduction of cost not only means that the agent would choose how many people to observe but also given that the agents are Bayesian expected utility maximizer the solution would be same as the constrained Pareto problem which means if there is any herding in the model that would be optimal herding unlike previous models in the literature.

The cost of social learning $c\left(n\right)$, where n be the number of agents of generation $t-1$ that an agent in generation t observes, has the following properties,

$$\begin{aligned} c\left(n\right) \geqslant 0, \quad 0 \leqslant n \leqslant N, n \in \mathbb{N} \quad \text{and } C\left(1\right) \leqslant \underline{u}, \quad C\left(N\right) > \bar{u} \\ c\left(n\right) \leq c\left(n+1\right), \quad 0 \leqslant n \leqslant N-1, n \in \mathbb{N} \\ c\left(i\right) - c\left(i+1\right) \leqslant c\left(i+1\right) - c\left(i\right) 1 \leqslant i \leqslant N-1 \end{aligned} \tag{3}$$

An agent can either choose n sequentially, i.e, after observing each individual he would decide whether he wants to observe one more individual and if he decides to observe then another observation is drawn from the period $t-1$ population without replacement; or the agent can choose n in a block, where before observing any action from any agent in period $t-1$ the agent decides how many people to observe and given his decision a sample of chosen size is drawn randomly from $t-1$ population. We assume for any learning protocol he would pay the same cost $c(n)$ after observing n actions at the end of the learning process. So, if he observes n many people under sequential learning he would pay $c(n)$ and not $nc(1)$[2].

Once an agent in generation t observes n many agents from generation $t-1$, the distribution of actions of the n agents works as a signal about μ^* for the agent. If he observes that x many people out of n had chosen action a, then the distribution of actions would be denoted by x_n. Given prior γ and observed distribution of action x_n, the agent would update his posterior about Γ, by considering how the agents in the earlier generation would behave under different distribution of types and the probability of mismatching the state by the $t-1$ period agents. The probability of mismatching would vary with the amount of social and private learning undertaken by generation $t-1$. We will call this error probability where the error is due to mismatch and not an error due to bounded rationality and this probability would be an important variable for the rest of the analysis.

2.3 Time 0 agents

The $t=0$ agent has only the option of learning privately, so after he enters the economy with prior $\gamma \in \Delta(\Gamma)$, he chooses a distribution of posteriors to maximize the expected payoff. Hence, the optimization problem of a $t=0$ agent is given by,

$$V(A,\gamma) = \max_{P\left(i,\omega|\mu\right)_{i\in A,\omega\in\Omega}} \sum_{\omega\in\Omega} \gamma(\omega) P\left(i,\omega|\gamma\right) u(i,\omega) - C(\lambda,\gamma). \tag{4}$$

where $\gamma(\omega)$ be the expected probability of state ω under $\gamma \in \Delta(\Gamma)$. Following Matejka and McKay (2015),the solution to the agent's optimization problem is similar to a logit model of random utility. Hence for time $t \geq 0$ agents, the posterior probability of choosing action i would be

$$P\left(i,\omega|\gamma\right) = \frac{P\left(i|\gamma\right) e^{\frac{u(i,\omega)}{\lambda}}}{\sum_{j\in A} P\left(j|\gamma\right) e^{\frac{u(j,\omega)}{\lambda}}} \quad \forall i \in A, \omega \in \Omega \tag{5}$$

The Bayesian plausibility implies given their prior γ,

$$\sum_{\omega\in\Omega} \gamma(\omega_i) \frac{\exp\left(u(i,\omega)/\lambda\right)}{\sum_{j\in A} P\left(j|\gamma\right) \exp\left(u(j,\omega)/\lambda\right)} \leq 1 \quad \forall \quad i \in A. \tag{6}$$

The inequality holds with equality if $P\left(i|\gamma\right) > 0$.

Using equation 6 for both $a, b \in A$ we get,

[2]The assumption would not be restrictive if learning cost is linear or in the form of a capacity constraint.

$$P\left(a|\gamma\right)=\begin{cases}\frac{\gamma(\omega_1)\exp\left(\bar{u}/\lambda\right)-\left(1-\gamma(\omega_1)\right)\exp\left(\underline{u}/\lambda\right)}{\exp\left(\bar{u}/\lambda\right)-\exp\left(\underline{u}/\lambda\right)} & \text{if} \quad e^{-\Delta u/\lambda}\leq\frac{\gamma(\omega_1)}{1-\gamma(\omega_1)}\leq e^{\Delta u/\lambda}\\ 1 & \text{if } \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)}>e^{\Delta u/\lambda}\\ 0 & \text{if } \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)}<e^{-\Delta u/\lambda}\end{cases} \tag{7}$$

And the posterior probability of choosing actions in different states can be obtained by combining equation 7 and equation 5. Two observations to be made here, first even though there is no social learning the time $t=0$ agents do not always learn perfectly about their types and hence the observed distribution of action contains both heterogeneity of idiosyncratic payoff and mistakes in the process of private learning.

Second, when the agent learns about his own type this would shift his belief about the economy as well and would shift γ, but since we have assumed that N is large enough the deviation in γ due to only one observation would be very small so we would ignore it. Alternatively we can consider a more general state space say $W=\Omega\times\Gamma$, so every time the agent observes a signal we would simultaneously update both the components of W. The posterior and prior probability of choosing an action is based on the bigger state space and so is the cost of private learning. The cost of social learning is not affected though since it doesn't depend on the state space. Since ex-ante the agent doesn't know what signal he is going to observe, while obtaining the prior probability he would use the expected belief of γ rather than the actual belief at that point. By Bayesian plausibility the expected γ would coincide with his prior ("prior" to private learning) and hence ex-ante the expected posterior probability would be the same as well. In that model all the subsequent analysis would still hold but for simplicity in the rest of the paper we will not explicitly consider the change in belief over Γ due to private learning by assuming N to be large enough.

Let us define $\epsilon_0^a=P\left(a,\omega_2|\gamma\right)$ and $\epsilon_0^b=P\left(b,\omega_1|\gamma\right)$ as the corresponding probabilities of making mistake when choosing a and b at time $t=0$ by type ω_2 and type ω_1 agents respectively. Since it is common knowledge that agents are Bayesian expected utility maximizer with same cost of private learning and hence every agent chooses the same distribution of posteriors and hence in the next generation, any $t=1$ agent knows ϵ_0^i.

2.4 Time $t\geq 1$ agents

2.4.1 Optimal Learning Protocol

Any $t\geq 1$ period agent has two different choices for learning, namely social and private learning. In this section we discuss the possible protocols of learning. The first issue is about the sequencing of learning. As no restriction has been imposed on the agent as to whether to choose one type of learning first then the other, the agent has the option to mix the two types of learning in any possible way. For example, he may decide to learn privately first and then observe some individuals from earlier generation, or he may choose to alternate private and social learning starting with any one of the two upto any number of repetitions.

Given the large number of possible sequencing between two types of learning it is of value to look at which type of sequencing protocol would survive the optimality argument. The following lemma shows the optimal sequencing would always be of the form: *first social learning then private learning.* To prove that, it is sufficient to show for any

sequencing protocol where some social learning is chosen after a step of private learning, an agent would be better off by learning socially before private learning. That would imply only one learning protocol would survive in any equilibrium, namely, first social learning then private learning.

Lemma 1. *Any agent at period $t \geqslant 1$ would optimally choose to learn socially first then privately.*

Prof of lemma 1 is given in the appendix $A.1$. Lemma 1 gives the optimal sequence of learning, so in the agent's optimization problem we can use the γ'_{x_n} as the interim belief prior to private learning after observing x_n many action as out of n many observations.

Next we consider the choice between block and sequential learning. We start by defining the the two learning concepts in the context of the model. If an agent is doing block learning then before any learning takes place he would optimally choose a value of n^* and then observe n^* many agents from earlier generation who are chosen randomly. This would generate a distribution of beliefs over Γ, and then he learn privately using the optimality conditions based on the updated belief. If an agent chooses to do sequential learning then he doesn't need to choose a value of n ex ante but after each random observation from earlier generation he can decide to stop sampling or keep on sampling from earlier generation.

Since both types of learning are costly and $c(N) > \bar{u}$, the agent would never choose to learn fully under any learning protocol. Also given the nature of private cost of learning if an agent is confident enough then he would not learn privately at all and choose one action for sure. This along with the weakly convex cost of social learning function implies that there exists a level of belief (for both actions) such that if an agent has a belief above that level about occurrence of any of the states then the marginal benefit from learning would be less than the marginal cost of learning of any type. This implies the optimal strategy under any protocol would be of the form of a cutoff belief for social learning conditional on the number of agents being already observed.

Under sequential learning this means the agent would use a stopping rule for social learning based on the belief conditional on the number of observations and the cost associated with it. But if there is only one such cutoff belief that describes the stopping rule and agent keeps on learning until he reaches that belief then it would not be an optimal strategy. This is true because cost of social learning function is weakly convex and given any cutoff belief the probability that the agent would never reach to that belief with less than N many observations is non-zero. Since $c(N) > \bar{u}$ he would be better of to stop social learning before he reaches his cut-off and start learning privately. Hence the optimal strategy of an agent would be to choose a distribution of cutoff levels of beliefs conditional on the number of observations rather than one single cutoff belief to define the stopping rule that would maximize the expected payoff subject to the cost of social learning.

But the choice of distribution of beliefs under sequential learning is constrained by the choice of n and the parameters of the model. So, to choose a distribution of belief from the restricted set of distribution available via choice of n he would choose a distribution of n to achieve maximum expected value under the belief subject to the social cost function. For example, he might choose the following strategy: to stop at $n = n_1$ if belief reaches γ_1 after observing n_1 many people otherwise choose $n = n_1 + 1$ and stop there. Since one can calculate the probability of reaching γ_1 after n_1 observations this gives an implied distribution over n.

Since for sequential learning case, we need to consider the agents problem at each possible belief and observation pair it would be significantly more difficult to solve. Hence for sake of simplicity in the rest of the paper we assume the agents are doing block learning. The sequential learning case is discussed in some detail in the extension and the results have a similar structure to the block learning case.

Since we assumed social learning is done in a block then ex-ante before obtaining any social and private information, the agent would decide how many agents to look at from earlier generation ,i.e. choose $n \leq N$. For every interim belief after observing n many agents, $\gamma'_{x_n} \in \Delta(\Gamma)$, how much to learn privately. Once the agent chooses n, then he would observe n agents from earlier generation drawn randomly from the total population, i.e. the agent can't choose whom to look at. Thus the agent's problem becomes,

$$W(A,\gamma) = \max_n E_{\gamma_{x_n}}\left[V\left(A,\gamma'_{x_n}\right)\right] - c(n) \tag{8}$$

where γ'_{x_n} would be the belief over Γ after observing x_n many people taking action a out of n many people at the time of private learning and agents are Bayesian.

To update their belief based on observation of actions by agents in earlier generation, the agent need to know the probability of making mistake by the earlier generation. In the next section we discuss the order of beliefs generated by social learning.

2.4.2 Social learning and order of beliefs

Suppose an agent i at time t observes n agents from generation $t-1$, then he will update his belief over $\Delta(\Gamma)$ via Bayes rule. Since Ω has only two elements, wlog, we can denote any distribution $\mu \in \Gamma$ by the probability of type ω_1. If the agent's observed sample is x_n, i.e. x out of n agents chose action a then the posterior probability of any distribution $\mu \in \Gamma$ in the support of $\gamma \in \Delta(\Gamma)$ by Bayes rule would be,

$$P\left(\mu|\gamma, x_n\right) = \frac{P\left(x_n|\mu\right) P\left(\mu|\gamma\right)}{\int_{\nu\in\gamma} P\left(x_n|\nu\right) P\left(\nu|\gamma\right)} \tag{9}$$

and the probability of all $\mu \notin supp(\gamma)$ would remain zero.

But to calculate the $P\left(x_n|\gamma,\mu\right)$, the agent needs to know the probability of mismatches by earlier generation. This probability would be different for different generations. For a time $t=1$ agent the problem is simpler as he knows a $t=0$ agent had done only private learning and the prior γ and marginal cost of learning λ is same across all generation. All later generations need to take into consideration what was the optimal level of social learning in earlier generations (n^*_{t-1}), what distribution of interim beliefs over $\Delta(\Gamma)$ n^*_{t-1} can generate, for each such distribution over $\Delta(\Gamma)$ what would be the probability of mismatch and need to take an expectation over probability of mismatches given their prior γ.

To illustrate, let us assume that an agent i observes a total of 5 agents from earlier generation among which all 5 agents had chosen action a. The $t=1$ agents know that this is only due to private learning so ϵ^i_0 would be the probability of making mistake when taking action $i \in A$. Now, all 5 would choose action a, if all 5 were actually type ω_1 and no one made a mistake, 4 of them were type ω_1 and the ω_2 type made a mistake, and so on upto all 5 were type ω_2 and all made a mistake.

First we consider case of a $t=1$ agent, who knows that the earlier generation agents only learn privately and the probability of making mistake is $\epsilon^a_0 = P(a,\omega_2)$ and $\epsilon^b_0 = P(b,\omega_1)$, then the probability of observing x_n, given prior μ would be

$$P\left(x_n|\mu\right) = \sum_{k=0}^{n}\sum_{j=k^*}^{k^{**}} \binom{n}{x_n - 2j + k} \mu^{x_n - 2j + k} \left(\epsilon_0^a\right)^j \left(1-\epsilon_0^b\right)^{x_n - j} \left(1-\mu\right)^{n - x_n - k + 2j} \left(\epsilon_0^b\right)^{k-j} \left(1-\epsilon_0^a\right)^{n - x_n - k + j} \tag{10}$$

where

$$k^* = \begin{cases} 0 & \text{if} \quad k < \min\left\{x_n, n - x_n\right\} \text{ or} \quad x_n \leq k < n - x_n \\ k - n + x_n & \text{if} \quad k \geq \max\left\{x_n, n - x_n\right\} \text{ or} \quad n - x_n \leq k < x_n \end{cases}$$

and

$$k^{**} = \begin{cases} k & \text{if} \quad k \leq \min\left\{x_n, n - x_n\right\} \text{ or} \quad n - x_n \leq k < x_n \\ x_n & \text{if} \quad k > \max\left\{x_n, n - x_n\right\} \text{ or} \quad x_n \leq k \leq n - x_n \end{cases}$$

which uses the probability of making mistakes and the belief that the distribution is generated by $\mu \in \Gamma$. Plugging the value obtained from equation 10 into equation 9 we can calculate $P\left(\mu|\gamma, x_n\right)$ for every $\mu \in \gamma$, and can update the belief to γ'_{x_n}.

Agents in period $t > 1$ know that the earlier generation had a chance of social learning. Since in a generation all agents are ex-ante identical, everyone in one generation would choose the same n, but the same choice of n might lead to different beliefs. Also, the probability of error would be different for each realization of n sample. Consider a period t agent who knows that period $t-1$ agents optimally chose to observe m people from generation $t-2$. Let X_m denote all possible sample distribution for sample size m. Let the error probabilities[3] be $\epsilon^a_{x_m,t} = P\left(a, \omega_2|\gamma, x_m, t\right)$ and $\epsilon^b_{x_m,t} = P\left(b, \omega_1|\gamma, x_m, t\right)$ after observing $x_m \in X_m$ in period t. Using the prior γ, one can calculate the distribution over X_m, which gives an implied distribution let f^a_γ and f^b_γ over $\epsilon^a_{x_m,t}$ and $\epsilon^b_{x_m,t}$ respectively. Let us define ϵ^a_t and ϵ^b_t as $\epsilon^i_{m,t} = \int_{x_m \in X_m} \epsilon^i_{x_m,t} df^i_\gamma$, i.e., the expected probability of making mistake by choosing i in period $t-1$ after observing m many agents from generation $t-2$.

Since all agents are ex-ante identical in generation $t-1$ and the sample is drawn randomly, the agent in period t would use $\epsilon^a_{m,t}, \epsilon^b_{m,t}$ as the probability of making mistake and use the same rule as $t = 1$ agents in 10 but replacing ϵ^i_0 by $\epsilon^i_{m,t}$, for $i \in A$ when he knows the value of m for earlier generation.

Since $t = 0$ agents choose only private learning, n^*_1 is common knowledge and iterating the argument and using the fact that all agents are ex-ante identical, given $c\left(n\right)$ and λ the sequence of optimal choice of n^*_t and the corresponding $\epsilon^i_{n^*_t}$ for $i \in A$ would also be common knowledge to all generations. Hence, the error probabilities, $\epsilon^i_{n^*_{t-1}}$ for every generation is well-defined. The following remark explains the argument.

Remark 1. *Note that to obtain $\epsilon^i_{n^*_{t-1}}$ for $i = a, b$ any agent in period t needs to know the entire history of $\epsilon^i_{n^*_{s-1}}$, $\forall s \leq t$. But agents are ex-ante identical and chooses n^* in a block means that every agent in a single generation would have the same n^*_s and hence the same $\epsilon^i_{n^*}$ would apply to all agents in generation s and the economy starts at period $t = 0$ with ϵ^i_0. Which means given γ, the common prior and λ, the marginal cost of private learning, the two parameters that determine ϵ^i_0 would be sufficient to generate the deterministic*

[3]Error of mismatching and not behavioral error.

time path of the optimal choice n_s^* and hence the time path of $\epsilon^i_{n^*_{s-1}}$. So given γ and λ the error probabilities $\epsilon^i_{n^*_{t-1}}$ would be a function of t only. So $\epsilon^i_{n^*_{t-1}}$ would be same for every agent in period t and also would be common knowledge since both γ and λ are common knowledge.

2.4.3 Private Learning

Given lemma 1, we know agents first learn socially then with the updated belief γ'_{x_n} they learn privately. Following equation 5 the optimal private learning of an agent in any period $t \geq 1$ would be same as a $t = 0$ agent, except with a different interim belief over Γ,

$$P\left(i, \omega | \gamma'_{x_n}\right) = \frac{P\left(i|\gamma'_{x_n}\right) e^{\frac{u(i,\omega)}{\lambda}}}{\sum_{j \in A} P\left(j|\gamma'_{x_n}\right) e^{\frac{u(j,\omega)}{\lambda}}} \quad \forall i \in A \tag{11}$$

Note that the γ'_{x_n} doesn't have a time dimension because the only way different generation would be different in their behavior is through social learning and γ'_{x_n} already captures the differences via social learning. Hence the agent with belief γ'_{x_n} chooses to learn privately only if, $e^{-\Delta u/\lambda} \leq \frac{\gamma'_{x_n}(\omega_i)}{1-\gamma'_{x_n}(\omega_i)} \leq e^{\Delta u/\lambda}$. For any other value of γ'_{x_n} he would choose one action for sure. Now that we have discussed the optimal learning under both social and private separately, we can describe the optimization problem faced by the agent.

3 Agent's optimization

Since the agent knows the optimal amount of private learning given any γ'_{x_n} he would choose the optimal value of n such that expected utility is maximized. For each n, the agent would calculate all possible distribution of interim beliefs γ'_{x_n} following n observations. For each such γ'_{x_n} then he calculates his expected utility using private learning and chooses the n that maximizes his payoff after subtracting the cost of social learning. So the agent's problem can be written as,

$$W(A, \gamma) = \max_n \left(E_{\gamma'_{x_n}} \left[\max_{\substack{P\left(i,\omega|\gamma'\right) \\ i \in A, \omega \in \Omega}} \sum_{\omega \in \Omega} \gamma'_{x_n}(\omega) P\left(i, \omega | \gamma'_{x_n}\right) u(a, \omega) - C\left(\lambda, \gamma'_{x_n} | \gamma\right) \right] - c(n) \right) \tag{12}$$

where the expectation is taken over the distribution $\gamma'_{x_n} \in \Delta(\Gamma)$ that can be generated after observing n agents when the prior belief is γ.

There will two types of cost attached to choosing optimal n, the explicit one is the cost of observing more people, the implicit one would be, if more agents are being observed, then the sample size would increase, for a Bayesian this implies that after observing the same proportion of people taking action i, he would be able to shift his belief further away from the prior in favor of action i compared to a smaller sample. But given the cost of private learning there exists a level of prior, not a completely informative one, for which no private learning is optimal. This means if he learns too much socially later he would optimally choose to not learn privately at all. This would affect the probability of

making mistake by the agent and he would internalize the cost of such a mistake. Thus there may exist a private cost function for which even with a very low/zero cost of social learning, agents would be reluctant to learn socially with the anticipation that it may affect their private learning behavior and hence their expected payoff.

Result 1. *If period $t = 1$ agents choose not to learn socially then no agent in any $t > 1$ generation would choose social learning.*

Proof. If at $t = 1$, agents choose $n^* = 0$, then for generation $t = 2$, the error probabilities would remain ϵ_0^i, which are same as that of $t = 1$ generation. Hence, if $n^* = 0$ was optimal for $t = 1$ generation, it would remain optimal for $t = 2$.

If for any generation $s > 2$, $n^* = 0$ was optimal then for the next generation error probabilities would be ϵ_0^i, same as $t = 1$ generation, hence their optimal choice would be $n^* = 0$. Thus by induction for all $t \geq 1$, $n^* = 0$ would be the optimal solution. □

Result 2. *For every common prior γ, except the uniform prior i.e. where $\gamma(\omega_i) = 1/2$ for $i = a, b$, there exists a $\bar{\lambda}_\gamma < \infty$, such that $\forall \lambda > \bar{\lambda}_\gamma$ there will be no learning in any generation $t \geq 0$.*

Proof. We build this proof in two steps, first we show that if it is not optimal for $t = 0$ generation agents to learn privately then it is not optimal for any later generation to learn either privately or socially. Then we show that under the given condition it is not optimal for generation $t = 0$ to learn privately.

For the first part, suppose $t = 0$ generation agent don't learn privately at all this implies every one of them would have a log-likelihood ratio either $\log \frac{\gamma'(\omega_1)}{1-\gamma'(\omega_1)} > \Delta u/\lambda$ or $\log \frac{\gamma'(\omega_1)}{1-\gamma'(\omega_1)} < -\Delta u/\lambda$ since all agents are ex-ante identical. In the first case everyone chooses action a, in the second everyone chooses action b. WLOG, we can consider the case where everyone chooses a since the other case would be similar.

Now consider an agent i in generation $t = 1$. Given the common prior γ, $P\left(a|\gamma\right) = 1$, if he decides to learn socially and chooses $n^* > 0$, then he would only end up observing agents with only action a. Now the error probabilities given λ and γ when $t = 0$ don't learn and choose a for sure is given by $\epsilon_0^a = 1$ and $\epsilon_0^b = 0$. Putting these values in equation 10 and considering that $x_n = n$ for any choice of n^* we get

$$P\left(x_n|\mu\right) = \sum_{k=0}^{n} \binom{n}{n-k} \mu^{n-k} \left(1-\mu\right)^k = 1; \quad \forall \mu \in \gamma \tag{13}$$

Using this value in equation 9 we get,

$$P\left(\mu|\gamma, x_n\right) = \frac{P\left(\mu|\gamma\right)}{\int_{\nu\in\gamma} P\left(\nu|\gamma\right)} = P\left(\mu|\gamma\right) \tag{14}$$

Hence the posterior belief after social learning would be same as the prior belief. As social learning is costly, the optimal choice would be $n^* = 0$. Now using result 1 we can argue that $n^* = 0$ would be optimal for all $t > 1$ generation as well.

Also if it is not optimal for any agent in period $t = 0$ to learn privately, it will not be beneficial for a period $t = 1$ agent to do so given the optimal choice be $n^* = 0$. As absent social learning an agent in period $t = 1$ has same optimization problem as that of an agent in period $t = 0$ and for the $t = 0$ agent it was optimal not to do any private learning. But any agent in later generations, $t > 1$ would also have same optimization

problem as $t = 0$ agents as $n^* = 0$ for any generation. Hence, the optimal amount of private learning for any agent in generation $t > 1$ would also be zero.

Given this we just need to verify whether there exists any such $\bar{\lambda}_\gamma < \infty$ such that for all $\lambda > \bar{\lambda}_\gamma$, $t = 0$ agents don't learn. Now the condition for no private learning is given by

$$\log \frac{1-\gamma'(\omega_1)}{\gamma'(\omega_1)} > \Delta u/\lambda \text{ or } \log \frac{1-\gamma'(\omega_1)}{\gamma'(\omega_1)} < -\Delta u/\lambda \tag{15}$$

$$\Rightarrow \lambda > \bar{\lambda}_\gamma = \max\left\{\Delta u/\log \frac{1-\gamma(\omega_1)}{\gamma(\omega_1)}, \Delta u/\log \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)}\right\} \tag{16}$$

Since, $\gamma(\omega_i) \neq 1/2$, this implies

$$\log \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)} \neq 0 \quad \text{and} \log \frac{1-\gamma(\omega_1)}{\gamma(\omega_1)} \neq 0$$

Hence, $\bar{\lambda}_\gamma < \infty$. □

Remark 2. *As $\gamma(\omega_1)$ increases (or decreases) towards 1 (or 0), the denominator in the first term of the max function would be negative (or positive) and also decreasing (or increases) with increase in $\gamma(\omega_1)$ and the denominator of the second term is positive(or negative) and increasing (or decreases). So for a high (low) $\gamma(\omega_1)$ the second (or first) term becomes the maximum and it is decreasing in increase(or decrease) in $\gamma(\omega_1)$. Hence $\bar{\lambda}(\gamma)$ decreases as the prior becomes more biased towards any one action. This is intuitive in the sense that if agents are more or less sure about which action to take then there is no point in wasting on learning a little bit more when marginal cost of private learning is high enough.*

Remark 3. *In the proof of result 2, we showed that if no agent learns privately in period $t = 0$ then no future agent would want to learn, but this statement is not necessarily true for any other generation $t > 0$. This is because, no learning in $t = 0$ implies same action is chosen by all agents but no private learning in any other $t > 0$ doesn't guarantee that as they have the option of learning socially. Then future generation might exploit the heterogeneity in action choice and some agents may end with diffused enough beliefs to learn privately.*

Theorem 1. *Given the social learning cost function in 3 and the prior γ, there exist $0 \leqslant \lambda^* < \lambda^j < \lambda^{**} \leqslant \infty$, such that*

1. *For all $\lambda \leqslant \lambda^*$, the optimal level of social learning at any period $t \geqslant 1$, $n_t^*(\lambda_1) \leqslant n_t^*(\lambda_2)$, where $\lambda_1 \leqslant \lambda_2$, i.e. optimal social learning is non-decreasing in marginal cost of private learning or social and private learning are "substitutes".*

2. *For all $\lambda \in [\lambda^*, \lambda^j) \cup (\lambda^j, \lambda^{**}]$, the optimal level of social learning at any period $t \geqslant 1$, $n_t^*(\lambda_1) \geqslant n_t^*(\lambda_2)$ where $\lambda_1 \leqslant \lambda_2$ and either $\lambda_1, \lambda_2 \in [\lambda^*, \lambda^j)$ or $\lambda_1, \lambda_2 \in (\lambda^j, \lambda^{**}]$, i.e optimal social learning is non-increasing in marginal cost of private learning or the social and private learning are "complements".*

3. *For any $t \geqslant 1$, $\lim_{\lambda^j -} n_t^*(\lambda) < \lim_{\lambda^j +} n_t^*(\lambda)$, i.e. the optimal n_t^* takes an upward jump at λ^j.*

4. *For all $\lambda > \lambda^{**}$, the optimal social learning is $n_t^* = 0$ at any period $t \geqslant 1$,i.e. the social learning becomes completely uninformative.*

The proof of the theorem is given in Appendix $A.2$. The intuition is as follows, the social learning imposes two types of cost, one physical cost of observing agents another the cost due to shifting beliefs against one's idiosyncratic type.

The informativeness of the social learning depends on the cost of private learning, since agents observe just the actions and not the beliefs of the agents from earlier generation. The action taken by an agent reflects the private learning of the individual, hence, if private learning is costly then the next generation agents can't learn a lot by observing the actions of the earlier generation. In the extreme case, where all agents take the same action in the period $t = 0$, no learning is beneficial since learning is costly and completely uninformative. Thus for very high λ, the physical cost of social learning becomes more relevant and the two types of learning work as complement.

For very low λ, the situation is opposite since the incremental benefit of having a less diffused prior is significant as agents learn a lot privately. This also means social learning can shift interim belief a lot, hence reduce the need for subsequent private learning. Thus the two types of cost become substitute of each other.

By assumption $c(N) > \bar{u}$, which implies there exists a level of $\bar{n} < N$ such that agents will never choose $n^* > \bar{n}$. This restricts the possible interim belief distribution generated by learning socially and leads to the surprising result that the optimal social learning takes an upward jump. This happens due to the nature of the cost of private learning function. Since higher λ means less precise posterior, then reducing the cost of private learning by learning a little less may matter a lot in terms of net payoff. This generates a value function that is decreasing in interim belief over some intermediate (between 1/2 and 1) level of interim belief μ. Since agents would try to avoid the loss due to excess social learning, there is a discrete jump in the optimal policy function.

For a high enough λ to the right of the cutoff, agents rather prefer to learn socially as much as possible and be biased in one direction in order to attain higher ex-ante expected payoff by lowering the cost incurred in private learning. For a λ just below the cutoff it would be better to restrict social learning, i.e. n^* and not end up in the decreasing part of the value function which explains the jump in the policy function.

4 A Special Case

In this section we will consider a special case of the general model discussed in the last section. It will be special in two ways. First, let the common prior at any $t \geqslant 0$ be $\gamma \sim U[0,1]$, which means $\gamma(\omega_1) = E(\mu|\gamma) \equiv \mu_0 = 1/2$. As we have already seen in result 2, in this case the social learning behavior for very high λ is different from any other prior, i.e., it encourages more learning of any form. Second, we put some structure on the social cost function that encourages more social learning.

The rest of the problem remains the same. The state dependent utility function is as give in 1 and the private cost of information is Shannon's mutual entropy with a marginal cost $\lambda \in [0, \infty]$. But the social cost of learning takes the form of a capacity constraint. For $\bar{c} \in \{1, 2, \ldots N-1\}$ the cost of social learning is given by,

$$c(n) = \begin{cases} 0 & \text{if} \quad n \leq \bar{c} \\ M >> \bar{u} & \text{if} \quad n > \bar{c}. \end{cases}$$

So individuals can observe upto $\bar{c}$ many people with no physical cost but to observe an extra agent they need to incur a huge cost M which is higher than the maximum payoff

generated by the choice problem. In terms of the notation used earlier $\bar{n} = \bar{c}$. Since the social cost function has a form of capacity constraint as $\bar{c}$ increases the social cost of learning essentially decreases as it becomes easier to observe more and more agents.

The motivation behind the specific type of cost function comes from the social learning in network literature. In the learning in network literature it is often described that an agent can only learn from his neighbors. In our setting there is no well-defined underlying network structure and observed agents are chosen randomly from the entire economy. But it would be analytically same problem as analyzing a network where all agents have equal number of neighbors, $\bar{c}$ and have no prior knowledge about how agents in their neighborhood are behaving and the distribution of types is independent of the location in the network. Then if an agent is restricted to choose only from his neighbors and observes some $n \leq \bar{c}$ many earlier generation agents he would consider them to be drawn randomly from the entire economy.

The interim and posterior probability of choosing an action is same as the general case and for the social learning each period if an agent chooses $n^* \leq \bar{c}$ then the ex-ante expected payoff given by,

$$W(A,\gamma) = \max_{n\leq\bar{c}, P\left(i,\omega|\gamma'_{x_n}\right)_{i\in A,\omega\in\Omega}} E_{\gamma'_{x_n}}\left[\sum_{\omega\in\Omega}\gamma'_{x_n}(\omega)\,P\left(i,\omega|\gamma'_{x_n}\right)u(i,\omega) - C\left(\lambda,\gamma'_{x_n}\right)\right]$$

where γ'_{x_n} denote the distribution of interim beliefs after social learning. And we have the following interesting result using theorem 1,

Theorem 2. *Given the cost of social learning $\bar{c} > 0$,*

1. *The optimal social learning is always positive, $n^* > 0$,*
2. *There exists $0 < \lambda' < \lambda^j < \infty$ such that*
   - *i. If $\lambda \leqslant \lambda'$ or $\lambda \geqslant \lambda^j$ then $n^* = \bar{c}$*
   - *ii. If $\lambda \in \left(\lambda', \lambda^j\right)$ then $n^* < \bar{c}$*

The proof of the theorem is given in Appendix *A*.3. The result is counter-intuitive for the following reason, for an intermediate range of values of λ, the agents are optimally choosing not to observe everyone they can, where observing more agents gives more information (in a Bayesian world) at no physical cost and the alternative way of learning is costly. The intuition behind the result is agents take into consideration the possible biases in choice induced by social learning and restricts their amount of social learning so as not to be overwhelmed by social learning and reduce private learning which is directed towards their own type. In other words, the agent internalizes the "herding" externality and chooses n optimally.

5 Extensions

5.1 Aggregate State

In this section we consider an aggregate state space along with the idiosyncratic state space. Let S denote finite aggregate state space. Without loss of generality let $S = \{h, l\}$

with the notion that h be the high state and l be the low state of the economy. The state dependent utility function for $s = h, l$ is given by

$$\bar{u}_s = u(a, \omega_1, s) > u(b, \omega_1, s) = \underline{u}_s$$
$$\bar{u}_s = u(b, \omega_2, s) > u(a, \omega_2, s) = \underline{u}_s$$

This implies the order of preference for both type of agents namely ω_1 and ω_2 remains the same and symmetric as before but magnitude of the difference depends on the aggregate state.

Let P_S denote the true transition probability matrix of the aggregate state space S, where P_S is common knowledge but an agent in period $t > 0$ doesn't know the realized aggregate state in period $t-1$. Every agent in period $t > 0$ enters with a common belief $\pi_0 \in \Pi \equiv \Delta(S)$ about the last period aggregate state and for $t = 0$ agents, let $\pi^0 \in \Pi$ be the common prior belief about aggregate state in period $t = 0$ where π^0 is more informative than π_0 in the sense for different aggregate states in period $t = 0$ nature chooses a π^0 closer to the truth than π_0 for any aggregate state s. Everyone knows π_0 but the realized value of π^0 is only observed by $t = 0$ agents.

Apart from the learning channel for the idiosyncratic state the agents can also choose to learn about the aggregate state. We assume that agents don't learn about aggregate state via private learning, hence the only way to learn about the aggregate state is via social learning which gives information about the aggregate state in period $t-1$ and using P_S belief about period t is formed. We further assume that the aggregate state S is independent of distribution of idiosyncratic state, which simplifies the analysis of belief formation.

Note that the aggregate state is not a "true" static feature of the economy, rather a dynamically changing state. This also includes the usual "true" state notion of aggregate state if P_S is an identity matrix ($\mathbb{I}_{2\times 2}$) and nature chooses a state at $t = 0$ with some probability distribution $\pi_{nature} \in \Delta(S)$ at the beginning of period $t = 0$. The common belief π_0(or π^0 for $t = 0$) may or may not be same as π_{nature} and the dynamics of evolution of belief based on social learning from earlier generation can be analyzed.

5.1.1 Belief Formation

By assumption the aggregate state is independent of the distribution of the idiosyncratic state, so the idiosyncratic type of a person is not informative of the aggregate state hence lemma 1 still holds true.

Since the agent doesn't learn about aggregate state privately, the belief about the aggregate state prior to private learning would not be updated in the process of private learning. Let $\pi_{x_n} \in \Pi$ be the belief about the aggregate state prior to private learning if he observes x_n many a's out of n observations from $t-1$ generation. Then the expected utility from choosing action i would be given by

$$u_{\pi_{x_n}}(i, \omega) = u(i, \omega, g)\, Pr\left(g|\pi_{x_n}\right) + u(i, \omega, b)\, Pr\left(b|\pi_{x_n}\right) \tag{17}$$

Hence, posterior probability of choosing action i after private learning would still be given by 5 where the $u(i, \omega)$ would be as replaced by $u_{\pi_{x_n}}(i, \omega)$ as defined in equation 17. So the analysis regarding private learning won't change.

Upon observing n agents, an agent in period $t > 0$ would update his belief about both the aggregate state in period $t-1$ and the distribution of types in the economy. For $t = 1$

the error probabilities remain the same except it would be aggregate state dependent, namely $\epsilon_0^{i,s}$, then for $i = a, j = 2$ and $i = b, j = 1$ and $s = \{g, b\}$ we have,

$$\epsilon_0^{i,s} = P\left(i, \omega_j | \mu, \pi = Pr\left\{s\right\} = 1\right) \tag{18}$$

For any generation $t > 1$ the error probabilities are again the expected error probabilities given n and it also uses the same conditioning on aggregate state as in equation 18. Thus the probability of observing x_n for $s = h, l$ would be given by

$$\begin{aligned} P\left(x_n | \mu, s\right) = \sum_{k=0}^{n} \sum_{j=k^*}^{k^{**}} \binom{n}{x_n - 2j + k} \mu^{x_n - 2j + k} \left(\epsilon_t^{a,s}\right)^j \left(1 - \epsilon_t^{b,s}\right)^{x_n - j} \\ (1-\mu)^{n - x_n - k + 2j} \left(\epsilon_t^{b,s}\right)^{k-j} \left(1 - \epsilon_t^{a,s}\right)^{n - x_n - k + j} \end{aligned} \tag{19}$$

where the error probabilities uses similar conditioning as in equation 18 for any $t \geq 1$, otherwise same as before. Then using independence a Bayesian agent would update his belief as follows,

$$P\left(\mu, s | \gamma, \pi_0, x_n\right) = \frac{P\left(x_n | \mu, s\right) P\left(\mu | \gamma\right) P\left(s | \pi_0\right)}{\int_{\substack{\nu \in \gamma \\ s \in S}} P\left(x_n | \nu\right) P\left(\nu | \gamma\right) P\left(s | \pi_0\right)}, \quad \text{for } s = h, l, \mu \in \gamma \tag{20}$$

Hence, the belief about the aggregate state of period $t-1$ would be,

$$P\left(s | \gamma, \pi_0, x_n\right) = \int_{\mu \in \gamma} P\left(\mu, s | \gamma, \pi_0, x_n\right) \tag{21}$$

Then using the transition probability matrix, P_s the agent would form π_{x_n}, the belief about the aggregate state in period t. Given that the agent would learn privately and choose an action to maximize ex-ante expected utility.

5.1.2 Agent's Optimization

Given the belief π_{x_n} we can construct $u_{\pi_{x_n}}(i, \omega)$ using equation 17, then the $V(\mu)$ remains same as before except the state dependent utilities are expected utilities over aggregate state. Since, V function remains the same all the analysis about the shape of V still holds true. The only difference being, when agents choose n then, it generates a distribution of beliefs over π, aggregate state for all possible x_n and a change in π would generate a different expected utility and hence a different level of V. So by choosing n agents not only move along V but also V is shifted.

Thus the optimization problem becomes,

$$W^S(A, \gamma) = \max_n E_{\gamma'_{x_n}} \left[E_{\pi_{x_n}} \left(V\left(A, \gamma'_{x_n}\right) \right) \right] - c(n) \tag{22}$$

which is same as 8 except the $V\left(A, \gamma'_{x_n}\right)$ is replaced by the expected $V\left(A, \gamma'_{x_n}\right)$, where the expectation is over π_{x_n}, the posterior probability distribution of aggregate states after observing x_n. With this modification the qualitative results of initial non-decreasing, followed by non-increasing along with a jump in the non-decreasing part level of social learning for different levels of λ i.e λ^* and λ^j hold true, but the cutoffs would be determined differently. For determining λ^j, we would use the expected value function given

any n as in equation 22 and the rest of the argument goes through since V has similar shape at all possible level of π. Also, the $\lim_{\lambda\to 0} V'_\mu \to 0$, hence $\lambda^* \geq 0$ also exists.

The more surprising result is given in the following proposition,

Proposition 1. *If $\gamma(\omega_i) \neq 1/2$, there exists a $\lambda^{**} < \infty$, such that for all $\lambda > \lambda^{**}$, the optimal social learning at period $t \geq 1$ is zero.*

Proof. The proof uses the similar idea of result 2. Let's start by showing for any other γ, π_0, there exists λ^{**} such that for all $\lambda > \lambda^{**}$, $P\left(a|\gamma\right) = 1(\text{or } 0)$. Now the prior probability of choosing action a for an agent in $t = 0$ is given by,

$$P\left(a|\gamma,\pi^0\right) = \begin{cases} \frac{\gamma(\omega_1)\exp\left(u^{\pi_0}(a,\omega_1)/\lambda\right) - \left(1-\gamma(\omega_1)\right)\exp\left(u_{\pi_0}(b,\omega_1)/\lambda\right)}{\exp\left(u_{\pi^0}(a,\omega_1)/\lambda\right) - \exp\left(u_{\pi_0}(b,\omega_1)/\lambda\right)} & \text{if} \quad e^{-\Delta^{\pi_0}u/\lambda} \leq \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)} \leq e^{\Delta_{\pi_0}u/\lambda} \\ 1 & \text{if } \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)} > e^{\Delta_{\pi^0}u/\lambda} \\ 0 & \text{if } \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)} < e^{-\Delta_{\pi^0}u/\lambda} \end{cases} \tag{23}$$

Hence, for $\log\frac{\gamma'(\omega_1)}{1-\gamma'(\omega_1)} > \Delta^{\pi_0}u/\lambda$ or $\log\frac{\gamma'(\omega_1)}{1-\gamma'(\omega_1)} < -\Delta_{\pi^0}u/\lambda$, i.e if

$$\lambda > \Delta_{\pi^0}u/\log\frac{\gamma'(\omega_1)}{1-\gamma'(\omega_1)} \text{ or } \lambda > \Delta_{\pi^0}u/\log\frac{1-\gamma'(\omega_1)}{\gamma'(\omega_1)}$$

the optimal level of private learning by agent in period $t = 0$ is zero. Now define $\lambda^{**} = \max\left\{\max_{\pi^0\in\Pi_0}\Delta_{\pi^0}u/\log\frac{\gamma'(\omega_1)}{1-\gamma'(\omega_1)}, \max_{\pi^0\in\Pi_0}\Delta_{\pi^0}u/\log\frac{1-\gamma'(\omega_1)}{\gamma'(\omega_1)}\right\}$[4] given μ, then for all $\lambda > \lambda^{**}$, agents in period $t = 0$ don't learn privately. If $\Delta_{\pi^0}u$ is finite for all $\pi^0 \in \Pi_0$ then by assuming $\mu \neq 1/2$, we ensure $\lambda^{**} < \infty$.

Since Π_0 is common knowledge, period $t = 1$ agent knows that for $\lambda > \lambda^{**}$, the $t = 0$ would always choose action $a(\text{or } b)$, hence social learning is completely uninformative about both aggregate and idiosyncratic state. So it is optimal to choose $n^* = 0$ for $t = 1$ generation agents.

Using similar argument as in result 2, we can conclude that if it is optimal for $t = 1$ agents to not learn socially then it is optimal for any $t > 1$ agents to not learn socially either. Hence, proved. □

The result is surprising, because even after introducing a payoff relevant aggregate state which can only be learned by social learning, there still exists a level λ such that for any higher marginal cost of private learning an agent optimally chooses zero social learning. The intuition behind the result is that a high level of λ makes any agent to stick with their prior and no learning at all. Hence, any behavior becomes completely uninformative for next generation which makes zero social learning optimal.

5.2 Heterogeneous Cost of Private learning

To the baseline model of section 2 now we add heterogeneity in the cost of private learning. Everything else being the same let $\lambda \sim F(\lambda)$, instead of λ being constant for all agents in the economy, for all $t \geqslant 0$, where the distribution F is common knowledge but while observing the action of an agent in period $t-1$, a t period agent can't infer the corresponding λ. We further assume that F in independent of type distribution.

[4]Since utility is bounded and $\gamma(\omega_i) \neq 1/2$ the maximum always exists.

Given μ, period $t = 1$ agent knows that different λs in period $t = 0$ would choose different levels of private learning. Let ϵ^i_λ denote the error probability of a $t = 0$ agent when the cost of private learning is λ, given $\lambda \in supp\,(F)$. Then define the expected error probability after observing any agent taking action i from generation $t = 0$ as

$$\epsilon^i_{0,F} = \int_{\lambda \in F} \epsilon^i_\lambda dF \tag{24}$$

If the earlier error probabilities are replaced by $\epsilon^i_{0,F}$ as defined in 24, the optimization problem for the agent in $t = 1$ remains same as before. Hence, the optimal solution to the problem would be same as before for any $\lambda \in F$.

The problem for $t > 1$ generation would be different since, the optimal level of n^* is different for different λ. The generation $t = 2$ agent would know the optimal n^*_λ for each $\lambda \in F$ and hence the error probability would be different from the baseline model. Let $X_{n^*_\lambda}$ denote all possible sample distribution for sample size n^*_λ. Let the error probabilities be $\epsilon^a_{x_{n^*_\lambda},\lambda,t} = P\left(a, \omega_2|\gamma, x_{n^*_\lambda}, t, \lambda\right)$ and $\epsilon^b_{x_{n^*_\lambda},\lambda,t} = P\left(b, \omega_1|\gamma, x_{n^*_\lambda}, \lambda, t\right)$ after observing $x_{n^*_\lambda} \in X_{n^*_\lambda}$ in period t by an agent with marginal cost of private learning being $\lambda \in F$. Using the prior γ the distribution over $X_{n^*_\lambda}$ can be obtained for each $\lambda \in F$ and let γ generates an implied distribution f^a_γ and f^b_γ over $\epsilon^a_{x_{n^*_\lambda},\lambda,t}$ and $\epsilon^b_{x_{n^*_\lambda},\lambda,t}$ respectively. Using independence between γ and F let us define $\epsilon^a_{t,F}$ and $\epsilon^b_{t,F}$ as $\epsilon^i_{t,F} = \int_{\lambda \in F} \int_{x_{n^*_\lambda} \in X_{n^*_\lambda}} \epsilon^i_{x_{n^*_\lambda},\lambda,t} df^i_\gamma dF$ as the expected error probability by choosing i in period $t-1$ after observing n^*_λ many agents from generation $t-2$ when the marginal cost of private learning is $\lambda \in F$. Given $\epsilon^i_{t,F}$ at any period $t > 2$ the error probabilities can be generated recursively.

The new error probabilities are the expected error probabilities over F. For each $\lambda \in F$ we calculate the expected error probability and then take expectation over the expected probability wrt F, to get the new error probabilities. Now F being common knowledge the path of $n^*_{\lambda,t}$ is also common knowledge for each λ and for each generation t by the recursive nature of the problem. Thus for $t > 2$ the error probabilities are well defined. Hence, given these new error probabilities the optimization problem for any λ is same as before and all the results still hold true.

5.3 Heterogeneous Cost of Social Learning

If instead of having different λs, the agents have different social cost functions $c_\alpha \sim G$ where c_α belongs to the set of all functions satisfying the conditions given in 3. In this case the problem would not be very different when G is common knowledge and is also independent of distribution of types. In that case the period $t = 0$ agents are still identical since they don't learn socially. So, ϵ^i_0 would not change. For agents in $t \geqslant 1$ the heterogeneity would be relevant if $n^* > 0$ for some cost types. Otherwise, it would be same as the baseline model. So the only interesting case is when $n^*_t > 0$ for some t and some α, where the generation t agents would have different error probabilities.

Consider the case when $n^*_t > 0$ for a positive measure set of cost types at some period $t > 0$, i.e. an agent with cost function c_α chooses $n^*_\alpha \geqslant 0$ (with strict inequality for a positive measure αs). Let $X_{n^*_\alpha}$ denote all possible sample distributions for sample size n^*_α, then the error probabilities would be given by $\epsilon^a_{x_{n^*_\alpha},t} = P\left(a, \omega_2|\gamma, x_{n^*_\alpha}, t\right)$ and $\epsilon^b_{x_{n^*_\alpha},t} = P\left(b, \omega_1|\gamma, x_{n^*_\alpha}, t\right)$. Using independence between γ and G let us define ϵ^a_t and ϵ^b_t as $\epsilon^i_t = \int_{C_\alpha \in G} \int_{x_{n^*_\alpha} \in X_{n^*_\alpha}} \epsilon^i_{x_{n^*_\alpha},t} df^i_\gamma dG$ as the expected probability of making mistake by

choosing i in period $t-1$ after observing n_α^* many agents from generation $t-2$ when the cost of social learning is $c_\alpha \in G$. Using this new error probabilities the problem remains the same and hence all the results still hold true.

5.4 Sequential Learning

Throughout the paper we assumed that agents are using block learning. But in this section we consider the case of sequential social learning. As we discussed earlier under sequential learning agents choose a stopping strategy conditional on belief and number of observations instead of choosing only one value of n, hence we cannot rewrite similar statements to that of theorem 1 for the sequential learning case.

To prove a similar result as that of theorem 1 we need to consider the entire support of the stopping strategy which gives a nonempty set of values of n. Let us define that set of optimal values of n at period $t > 0$ to be $\mathbb{N}_t$. Let n_{min}^t denotes the minimum value of n in the set $\mathbb{N}_t$. Under sequential strategy we can write a similar proposition as that of theorem 1.

Theorem 3. *Under sequential social learning , there exist a set of cutoff values of λ, namely, $0 \leqslant \lambda_s^* \leq \lambda_s^i < \lambda_s^d \leq \lambda_s^j < \lambda^{**} \leqslant \infty$, such that*

1. *For all $\lambda \leqslant \lambda_s^*$, the minimum level of social learning at any period $t \geqslant 1$ in the optimal set $\mathbb{N}_t$ is such that $n_{min}^t(\lambda_1) \leqslant n_{min}^t(\lambda_2)$, where $\lambda_1 \leqslant \lambda_2$, i.e. n_{min} is non-decreasing in marginal cost of private learning.*

2. *For all $\lambda \in \left[\lambda_s^i, \lambda_s^d\right)$, the minimum level of optimal social learning at any period $t \geqslant 1$ is such that, $n_{min}^t(\lambda_1) \geqslant n_{min}^t(\lambda_2)$ where $\lambda_1 \leqslant \lambda_2$ and $\lambda_1, \lambda_2 \in \left[\lambda_s^i, \lambda_s^d\right)$,i.e., n_{min}^t is non-decreasing in marginal cost of private learning.*

3. *For any $t \geqslant 1$, $\lim_{\lambda^j -} n_{min}^t(\lambda) < \lim_{\lambda^j +} n_{min}^t(\lambda)$, i.e. the minimum optimal level of social learning n_{min}^t takes an upward jump at λ_s^j.*

4. *For all $\lambda > \lambda_s^{**}$, the optimal social learning set is singleton, specifically $\mathbb{N}_t = \{0\}$ at any period $t \geqslant 1$,i.e. the social learning becomes completely uninformative.*

The proof of the theorem is given in the Appendix, *A*.4. Note that this theorem though similar in spirit but is much weaker than theorem 1. Instead of analyzing the whole set of n we are only able to discuss the n_{min} which is a much weaker statement. But overall the nature of the equilibrium is similar. Initially the two types of learning works as substitutes and eventually becomes complement. Also for middle range of values of λ the agents restrict their choices of n and as λ starts to increase even more the number of observations increases suddenly. Instead of thinking in terms of n_{min} if we think in terms of highest belief in terms of μ (which most often coincides with the choice of minimum n), the two theorems are more similar than they appear here.

6 Steady State and Herding

As we motivated the problem in the introduction by referring to herding behavior, in this section we would discuss whether herding remains as an equilibrium under any social and private cost function. In general, herding is referred to a steady state were all agents ignore their own private information and follow the actions of the previous agents blindly.

In the following section first we formally define the staedy state for this economy and explore the possibility of a herding steady state.

6.1 Distribution of Actions

To analyze what happens to the economy in the long run, we need to define a steady state in this context. A steady state in this economy would be such that that distribution of actions would become stationary. Let M be the set of all possible shares of actions. Given $A = \{a, b\}$ and $\Omega = \{\omega_1, \omega_2\}$, M is also finite. In this case $M = \left\{i/N | i \in \mathbb{N}, 1 \leq i \leq N\right\}$ denotes all possible shares of action a. Let m_t denote the realized distribution of actions at t.

Since at $t = 0$ agents get no social information, given μ^* and λ, the initial distribution of m_0 is given by

$$Pr\left\{m_0 = n/N\right\} = \sum_{k=0}^{N} \sum_{j=k^*}^{k^{**}} \binom{N}{n-2j+k} \mu^{n-2j+k} \left(\epsilon_0^a\right)^j \left(1-\epsilon_0^b\right)^{n-j} \left(1-\mu\right)^{N-n-k+2j} \left(\epsilon_0^b\right)^{k-j} \left(1-\epsilon_0^a\right)^{N-n-k+j} \quad (25)$$

where k^* and k^{**} is defined as before.

For $t = 1$, given λ and $\bar{c}$ let n_1^* be the optimal level of social learning. Then using the similar approach we can calculate the probability distribution m_1 over M. For that we use m_0 and n_1^* to obtain all possible γ'_{x_n} after social learning the the error probabilities. For $t = 2$ however the problem would not be identical since ϵ_1^i is not necessarily equal to ϵ_2^i, hence n_1^* may not be equal to n_2^*. Given λ and γ let us define $\sum_{i=a,b} \epsilon_{max}^i = \max_n \sum_{i=a,b} \epsilon_n^i$ and $\sum_{i=a,b} \epsilon_{min}^i = \min_n \sum_{i=a,b} \epsilon_n^i$ as the maximum and minimum possible sum of error probabilities, given λ. For any generation t the sum of error probabilities would be in between these two values when λ is given.

6.2 Steady State

Let n_{max}^* and n_{min}^* be the corresponding optimal levels of social learning for $\sum_{i=a,b} \epsilon_{max}^i$ and $\sum_{i=a,b} \epsilon_{min}^i$ respectively given λ. Since social learning is costly and higher $\sum_{i=a,b} \epsilon^i$ implies the smaller change in the posterior, $n_{max}^* \leqslant n_{min}^*$. Hence, there are two possibilities, either $n_{max}^* = n_{min}^*$ or $n_{max}^* < n_{min}^*$. In the first case, the optimal n^* for any generation t would be given by $n_{max}^* = n_{min}^* = n^*$.

In the second case however that may not be true. Since both n_{max}^* and n_{min}^* are bounded by 0 and $\bar{n}$, there are few possibilities to be considered. First, n^* keeps on increasing over time$(n_t^* \geqslant n_{t+1}^*)$ until $n^* = n_{min}^*$ and stays there forever. Second n^* keeps on decreasing over time$(n_t^* \leqslant n_{t+1}^*)$ until $n^* = n_{max}^*$ and stays there forever. Third there exists a $n \in (n_{max}^*, n_{min}^*)$ such that if $n_t^* = n$ then $n_{t+1}^* = n$. Lastly, there exists a set of $n_j \in [n_{max}^*, n_{min}^*]$ such that n^* switches from one n_j to another in a deterministic way. The steady state in this economy would be defined as the stationary distribution of the actions m in the economy.

Except the last possibility of the second case where n^* changes within a set, we have a n^* and $t^* < \infty$ such that $\forall t \geqslant t^*$, the optimal level of social learning is n^*.

Result 3. *A sufficient condition for the existence of a steady state in this economy is that there exists a $t^* \leq \infty$ such that $n_{t^*}^* = n_{t^*+1}^* = n$ for some $0 \leq n \leq N$.*

Proof. Assume that there exists a $t^* \leq \infty$ such that $n^*_{t^*} = n^*_{t^*+1} = n$ for some $0 \leq n \leq N$. Then starting at t^* the optimal solution of all future generation would be same, this is true because if $t^* + 1$ generation agent find it optimal to choose n when t^* chooses n, then $t^* + 2$ generation would also find it optimal to choose n as they face the exact same problem as of $t^* + 1$ agents. Iterating the logic every future generation would choose n. Given n, X_n denotes the set of all possible values of x_n and for each such x_n the posterior probability of choosing an action would be same. Given the recursive nature of the problem thus we can define $P_{ij} = Pr\left\{M_{t+1} = j | M_t = i\right\}$ as the transition probability matrix over states in M. This is well defined because for all possible values of $i \in M$, given n we can generate the distribution over X_n and thus calculate the probability of being at different states of M in next period. Since λ and γ is common knowledge, the sequence of n^*_t is deterministic, hence we can calculate the distribution m_{t^*}. With m_{t^*} as the initial distribution and P_{ij} as the transition probability matrix the Markov chain over state space M would give the distribution of action at any $t \geq t^*$.

Now by definition the transition probability matrix is aperiodic. So given M is finite it is sufficient to show that P_{ij} is irreducible for the existence (and also uniqueness) of a steady state. For any $\lambda \in \left[0, \lambda^j\right)$, agents would always learn privately as they remain before the final increasing part of the value function where learning stops. Hence, $P_{ij} \in (0, 1)$ for any $i, j \in M$ for any such λ, hence a steady state exists.

For $\lambda \in \left[\lambda^j, \lambda^{**}\right]$, there exists a value of x^*_n such that for $x_n \geqslant x^*_n$, the optimal choice is to not learn privately, which implies there exists a state $m_a \leq 1$ such that for all $i \geqslant m_a$, $P_{iM} = 1$, hence $i = M$ is a deterministic steady state distribution of actions, as $P_{MM} = 1$. Similar argument would hold for action b as well. Finally for all $\lambda > \lambda^{**}$, the optimal choice of action is determined by the prior, say a wlog, then everyone taking action a would be the unique steady state. Also note that in this case $n^*_t = 0$ for all $t > 0$. Hence a steady state exists for all possible λ under the above mentioned condition in the statement of the result. □

In case where n^* oscillates within a bounded set in a deterministic way, there may not exist any steady state but the set of distribution over M given n_j in the set would generate a cycle of choice of n in the economy.

6.3 Herding

With the definition of steady state, we can now define herding as a steady state where $n^* > 0$ and there exists $n_h \leqslant N$ such that if $n \geqslant n_h$ many agents take action a (or b) in some period $t \geqslant 0$, then all agents in all periods $s \geqslant t + 1$ would only choose action a (or b).

Proposition 2. *For all $\lambda \in [0, \infty] \setminus \left[\lambda^j, \lambda^{**}\right]$ there doesn't exist any herding steady state, almost surely.*

Proof. Using theorem 1 we know for all $\lambda > \lambda^{**}$, the optimal social learning is zero, hence herding is not an equilibrium for all such λs.[5]

For all $\lambda < \lambda^j$, the $\bar{\mu}_n$ (on either side of 1/2), is in the increasing part of the value function $V(\mu)$, which implies for all possible values of x_n the optimal level of private learning is not zero. Since private learning is informative about idiosyncratic state, the

[5] For $E(\mu|\gamma) = 1/2$, λ^{**} becomes ∞ but still the result holds true.

probability of everyone choosing action a (or b) after private learning would be zero for any $\mu^* \in int(\gamma)$. Hence, there is no herding steady state with probability 1.

But for $\lambda \in \left[\lambda^j, \lambda^{**}\right]$, the $\bar{\mu}_n$ goes to the final increasing section, hence the probability of no private learning by any agent is strictly positive after observing a high (or low) enough x_n. Also the probability of everyone taking the same action in a generation is strictly positive. Since, μ and λ are such that agents choose to learn socially, this would imply if all agents in a generation take the same action then every agent in the next generation would also choose the same action. Thus, herding can't be ruled out as a possible steady state. □

6.4 A special case

Which of the four cases would occur given λ and μ would depend on the social cost function. For our special in section 4, when $\mu = 1/2$ and the social cost function takes the form of capacity constraint, we have the following results:

Proposition 3. *If $\lambda \leqslant \lambda'$ then there exists an unique steady state with $n^* = \bar{c}$ with no herding.*

Proof. Using theorem 2, we know for $\lambda \leqslant \lambda'$ the optimal $n^* = \bar{c}$. Since n^* is same for all generation $t \geqslant 1$ by result 3 a steady state exists.

For uniqueness we want P_{ij} to be aperiodic and irreducible. By definition of the problem P_{ij} is aperiodic. So we need to check whether P_{ij} is irreducible. Since $\mu = 1/2$, after observing $\bar{c}$ many agents for all $\lambda \leqslant \lambda'$ the belief would be in the initial increasing zone, hence, private learning is always positive for any possible value of M_t. So $0 < P_{ij} < 1$ for all i, j, hence there exists an unique steady state with $n^* = \bar{c}$.

Since, even with the most extreme belief of observing everyone taking action a (wlog), agents learn privately, then herding would be an equilibrium only if there exists some t^* such that for all $t \geqslant t^*$, every agent with every possible $x_{\bar{c}}$ and learning privately chooses action a. Given $\lambda \leqslant \lambda'$ and $\bar{c}$, the error probabilities ϵ_t^i for any $t > 0$ are strictly less than one because of positive private learning and the true distribution of types being $\mu^* \in int(\gamma)$. Thus the probability that every agent chooses one action for all future periods goes to zero for large enough N, as assumed here. Which implies the stationary distribution of the Markov chain can't be a mass point at 0 or 1. Hence, herding is not a steady state of the economy. □

Proposition 4. *If $\lambda \geqslant \lambda^j$ then there exists herding equilibrium where $n^* = \bar{c}$ for all $t > 0$ and everyone chooses action a (or b).*

Proof. Using theorem 2, we know for $\lambda \geqslant \lambda^j$ the optimal $n^* = \bar{c}$, hence by result 3 a steady state exists. Since λ is high enough there exist $\bar{x}_{\bar{c}}$ (or $\underline{x_{\bar{c}}}$) such that for all $x_{\bar{c}} \geqslant \bar{x}_{\bar{c}}$ (or $x_{\bar{c}} \leqslant \underline{x_{\bar{c}}}$) the optimal choice would be action a (or b) without private learning. Hence, everyone choosing action a (or b) after some $n^* > 0$ would be an absorbing state.Thus herding equilibrium exists. □

6.5 Welfare Implication

In this model the constrained planner's problem is same as the agent's optimization problem. Hence, given λ and $c(n)$, the planner would choose the same level of social and private learning as under the decentralized equilibrium. Thus unlike the herding literature

a la Banerjee(1992) and BHW (1992) this model doesn't have any herd externality. So, restricting some agents in earlier generation to only learn privately would not be welfare improving and might actually decrease the welfare. The reason for the loss being the cost of private learning which is absent in the herding literature.

Also in the model herding can be a possible equilibrium but it would be nonetheless Pareto optimal. Since herding can occur at a sufficiently high level of λ, the cost of private information is significantly high, hence it might be optimal to not investigate privately incurring large cost and to follow the earlier generations blindly.

7 Conclusion

To summarize, this paper constructs and solves a model of individual stochastic choice where agents are rationally inattentive and face a costly social learning function. We showed that for such an agent the optimal choice of social learning is not monotonic in the marginal cost of private learning and it can be welfare improving for such an agent to observe the action of other agents in the economy even under the fear of herding based on the relative cost of private and social learning.

We consider that the agents are rational Bayesian expected utility maximizer. Thus when updating following the observation from the society they take into account the possibility of errors in the choices made by other agents in the previous generation and thus internalize the herd externality. This way we solve the constrained Pareto optimal for the social planner and thus if there is herding in equilibrium then it would be optimal in our model.

Since the solution of the model is first best in terms of constrained Pareto efficiency the only policy implication would be to reduce the cost of learning. For example a welfare improving policy would be to reduce λ, as reducing λ would unambiguously increase the ex-ante expected value of the agent. On the other hand restricting or encouraging the social learning by changing the cost of social learningi.e., changing the technology of connectivity would have different effect depending on the level of λ and the cost of social learning. This brings us back to our initial question of whether improved connectivity in a society necessarily better in terms of welfare? This model says as a society becomes more and more connected unless the technology of private learning is also improving it would rather become worse possibly for the agents in the economy in terms of welfare.

One of the surprising result from our example is that even when the marginal cost of observing agents from earlier generation is zero and the distribution of actions of earlier generation contains relevant information in the form that upon observing the distribution of actions an agent can update his belief, he may optimally choose not to observe everyone possible since later that would affect his incentive to learn privately. Thus the apparent "time inconsistency", i.e. not using a cheaper learning method in stage 1 since it would affect the stage 2 behavior is optimal in this case and the reason behind this counter-intuitive result is that the role of the two types of learning are different in the model.

The social learning is done mainly to reduce the next stage cost of private learning since it does not directly give information about the agent's idiosyncratic type and rather affects it indirectly by changing the belief about the society. Whereas private learning is more important for an agent in order to reveal his own type and he would rather choose a lower level of social learning to not harm his incentives for private learning. This clearly highlights the asymmetry in the private information vs observing others' action, which

is at best a noise signal about their types. Unlike herding literature in this model agents choose optimal level of both types of learning taking this asymmetry into consideration.

The major motivation behind the social cost function was that the agents are born in a network and can only observe their neighbors which restricts their choice of social learning differently based on their degree. But if in a network the place an agent is born has a correlation with who he can observe, then the lemma 1 breaks down and we can't use the analysis discussed in this paper.

But consider an exogenously given network such that at any time $t \geq 0$ agents are born randomly on the nodes and each node has equal probability of being of any type. At any time $t > 1$ a node would have one agent who can observe all the earlier generation agents in his neighborhood and he puts equal weight on every node (he does not observe his own node). Also assume that the distribution of degrees in the network are common knowledge but agents can't observe the actual degrees of his neighbors. In that framework this model would give same result. Then it would be the case of heterogeneous social cost function (as discussed in the extensions) as different agents would have different level of $\bar{c}$, depending on how many neighbors they have. Though the choice problem of the individual would be the same, the dynamics of that economy would be different and how the characteristics of the network would affect the steady state would be worth exploring.

Another motivating question for this paper was how do the two types of learning interact, more specifically whether private and social learning are substitutes or complements, i.e., when cost of private learning goes up would there be more or less social learning? Given the characterization of learning behavior in the model we showed that with any weakly convex social cost function when private learning is relatively cheap, the two types of learning are substitutes and for moderately higher cost of private learning they would become complements. Hence, the notion of "blind following blind" doesn't work. This is because when agents know that they themselves won't learn privately a lot, they would also know that other agents' action would be equally uninformative about their idiosyncratic state but they get to save some by not observing others. Hence they would use that argument to choose optimally in a costly learning framework. In the extreme case when the cost of learning becomes too high no one would learn anything and take an action as dictated by their prior.

In all the extensions the main assumption is that the agents have common knowledge about how the aggregate state evolves or the type of heterogeneity of the population and this is common knowledge. In case when there is a difference in belief in the economy regarding the aggregate process or the heterogeneity of the population, this results need not hold true. As upon observing an agent's action, to calculate the error probabilities the correlation structure has to be considered and it would be a much more difficult problem to solve.

References

[1] Venkatesh Bala and Sanjeev Goyal. Conformism and diversity under social learning. *Economic theory*, 17(1):101–120, 2001.

[2] Abhijit V Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, pages 797–817, 1992.

[3] Sushil Bikhchandani, David Hirshleifer, and Ivo Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of political Economy*, pages 992–1026, 1992.

[4] David Blackwell et al. Equivalent comparisons of experiments. *The annals of mathematical statistics*, 24(2):265–272, 1953.

[5] Thomas Brenner. Agent learning representation: advice on modelling economic learning. *Handbook of computational economics*, 2:895–947, 2006.

[6] Steven Callander and Johannes Hörner. The wisdom of the minority. *Journal of Economic theory*, 144(4):1421–1439, 2009.

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

[8] Andrew Caplin, Mark Dean, and John Leahy. Rational inattention, optimal consideration sets, and stochastic choice. Technical report, NYU working paper, 2016.

[9] Andrew Caplin, John Leahy, and Filip Matějka. Social learning and selective attention. Technical report, National Bureau of Economic Research, 2015.

[10] Andrew Caplin and Daniel Martin. A testable theory of imperfect perception. *The Economic Journal*, 125(582):184–202, 2015.

[11] Thomas M Cover and Joy A Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[12] Jacques Crémer. A simple proof of blackwell's "comparison of experiments" theorem. *Journal of Economic Theory*, 27(2):439–443, 1982.

[13] Henrique De Oliveira, Tommaso Denti, Maximilian Mihm, and Kemal Ozbek. Rationally inattentive preferences. *Available at SSRN 2274286*, 2014.

[14] Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *The American Economic Review*, 101(6):2590–2615, 2011.

[15] Ilan Lobel and Evan Sadler. Information diffusion in networks through social learning. *Theoretical Economics*, 10(3):807–851, 2015.

[16] Paola Manzini and Marco Mariotti. Stochastic choice and consideration sets. *Econometrica*, 82(3):1153–1176, 2014.

[17] Filip Matejka and Alisdair McKay. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *The American Economic Review*, 105(1):272–298, 2014.

[18] Christopher A. Sims. Implications of rational inattention. *Journal of Monetary Economics*, 50(3):665 – 690, 2003. Swiss National Bank/Study Center Gerzensee Conference on Monetary Policy under Incomplete Information.

[19] Michael Woodford. Inattentive valuation and reference-dependent choice. *Unpublished Manuscript, Columbia University*, 2012.

A Appendix

A.1 Proof of Lemma 1

Proof. Suppose not. Suppose an agent i in period $t \geq 1$ observes n many agents from generation $t-1$ but instead of following the first social then private rule he decides to observe $n - n_1$ many agents first and update his belief over Γ for some n_1, where $0 < n_1 \leq n$. Then he chooses optimal level of private learning given the updated belief over Γ and the marginal cost of private learning λ. After the private learning is done then again he observes n_1 many agents, where $n \geqslant n_1$ and $n_1 > 0$. Note that this is more general than a sequencing protocol with just first private then social learning but for $n = n_1$ this protocol coincides with the protocol to learn privately first then socially.

Let before the private learning his belief over Γ was γ_1 and given γ_1 his belief over his own Ω was μ_1. Private learning would change his belief from μ_1 to μ_2. Now, once he observes the rest n_1 many agents his belief about $\Delta(\Gamma)$ changes from γ_1 to γ_2. Given his private learning information structure this changes his belief over Ω from μ_2 to μ_3. The cost of private learning is a sunk cost but in the light of the new information the agent now has a different posterior probability of choosing action $i \in A$. Since the amount of private learning is based on γ_1 and not γ_2 this means the marginal benefit of private learning coming from a changed belief μ_2 may be different from marginal cost of private learning which is based on μ_1. In the case when marginal cost is greater than the marginal benefit this protocol would not be an optimal protocol.

In the rest of the proof we discuss the cases when the marginal cost of private learning is higher than the marginal benefit and show that these cases would happen with positive probability. On the other hand if the marginal benefit is higher the agent would optimally choose further private learning to update μ_2. In that case he is indifferent between this protocol and the protocol where he observes all n agents first and then learn privately since in both cases the final amount of private learning is determined by γ_2 and hence the same.

Let us define x_{n-n_1} as the number of agents from period $t-1$ who had taken action a out of the first $n-n_1$ observations and x_{n_1} denotes the same for the rest n_1 many agents. Combining the two we can define x_n as the number of agents choosing action a in the entire sample of size n, which means $x_n = x_{n-n_1} + x_{n_1}$. In the case where i would have chosen to do more private learning after x_{n_1} compared to x_n, he incurs a loss due to the sunk cost related to excess private learning. We want to show that the probability of the event where i learns more following x_{n-n_1} than following x_n, i.e., excess private learning, would be non-zero which would imply that the expected benefit from switching to the first social then private protocol is strictly positive. We would prove this by contradiction supposing the probability to be zero.

Observe that the agent would learn more following x_n only if after observing n agents his belief moves towards uniform distribution over Γ. WLOG, assume x_{n-n_1} induced that ω_1 is more likely, $Pr(\omega_1) > 1/2$, then in the bigger sample n the proportion of people taking action a in earlier generation, x_n has to be such that $x_n/n < x_{n-n_1}/n - n_1$ always for the probability to be zero, as agents are Bayesian and observing the same proportion of people taking action a in a larger sample gives more evidence towards ω_1.

For the probability to be zero, after every possible x_n the agent would have to chose more private learning. Define $\underline{x_n}(x_{n-n_1})$ to be the highest number of agents choosing action a out of n, such that after observing $\underline{x_n}(x_{n-n_1})$ the belief about $\gamma(\omega_1)$ would

be weakly less than as the belief after observing x_{n-n_1} out of $n-n_1$, which implies $\underline{x_n}(x_{n-n_1})/n < x_{n-n_1}/n-n_1$. The inequality holds true because otherwise given the prior that ω_1 is a more likely state, the agent would update to a belief more biased towards a. So to learn more after observing all n actions, the agent has to observe x_{n_1} such that $x_{n-n_1} \leq x_n \leq \underline{x_n}(x_{n-n_1})$. If $\underline{x_n}(x_{n-n_1}) < x_{n-n_1}$, then the probability of learning more after x_{n-n_1} compared to x_n, i.e., excess private learning becomes one.

Let us then consider the case, where $\underline{x_n}(x_{n-n_1}) > x_{n-n_1}$. The probability that x_n lies in this region is given by,

$$P\left(x_{n-n_1} \leq x_n \leq \underline{x_n}(x_{n-n_1})\right) = \sum_{j=0}^{\underline{x_n}-x_{n-n_1}} \binom{n}{j} \left(P_{t-1}(a)\right)^{x_{n-n_1}+j} \left(1-P_{t-1}(a)\right)^{n-x_{n-n_1}-j}$$

where $P_{t-1}(a)$ be the expected probability of choosing action a by $t-1$ generation agents. The probability on the LHS would always be less that one if $P_{t-1}(a) > 0$. But if, $P_{t-1}(a) = 0$, then $x_n = x_{n-n_1} = 0$ always for any $n, n_1 \geq 0$. Hence the agent would always get more and more evidence towards choosing action b and the updated belief can never go towards uniform belief. The Bayesian agent would never learn more after getting even stronger evidence towards the same action from n observations as after $n-n_1$ observations. Hence the LHS probability is strictly less than one, i.e., sometimes agent i would have done excess private learning after x_{n-n_1}.

As mentioned earlier, in the case where he has had learned more he incurs some loss which can be avoided if he had observed all n agents before doing any private learning. In the case where he needs to do more private learning the sequence is irrelevant. Hence, he would be better off by doing social learning first and then private learning. □

A.2 Proof of Theorem 1

Proof. To prove this theorem we first consider the optimal choice of an agent at any time $t \geq 1$.Since the optimization problem is identical for every generation $t \geqslant 1$, except for the ϵ_t^i, for $i = a, b$; we suppress the time subscript and solve the problem for any $t \geq 1$[6]. The level of n^*. i.e. the optimal choice of n would depend on ϵ^i but the qualitative results will not be affected.

Given lemma 1, we can solve the optimization problem in two stages. The agent's objective is to maximize ex-ante expected payoff given prior γ and the two cost functions, equation 2 and equation 3. In the first stage he would optimally choose n^*, which would generate a distribution $\gamma'_{n^*} \in \Delta(\Gamma)$ of intermediate beliefs over Γ. Given any intermediate belief $\mu \in supp(\gamma'_{n^*})$ the agent would then optimally choose the level of private learning that would generate a distribution of posteriors for each μ. We would solve this problem backward, namely first we would find out the optimal level private learning given μ and then we would analyze how the interim expected value changes with μ. Given how the value function changes with μ, we would choose n^* to maximize ex-ante expected value function.

Step 1: The solution to the stage 2 problem is already discussed in section 2.2.1,i.e., given an interim belief $\mu \in \gamma'_{n^*}$ what is the posterior probability of choosing an action $i \in A$.

[6]Note that we already have the solution for $t = 0$ period, as given in section 2.3

Step 2: Given the solution to the stage 2 problem now we consider how the interim expected value changes with interim belief, $\mu \in \gamma'_{n^*}$. Let $\mu = Pr(\omega_1)$ denote the interim belief of an agent before private learning and p_a be the probability of choosing action a given μ prior to any private learning. Then using the optimal choices as described in equation 5 and the parameters of the model, $\bar{u}, \underline{u}, \lambda$, we can write the interim value function given belief μ to be,

$$\begin{aligned} V(\mu) = \bar{u}\left[\mu\frac{p_a\bar{\lambda}}{p_a\bar{\lambda}+(1-p_a)\underline{\lambda}} + (1-\mu)\frac{(1-p_a)\bar{\lambda}}{(1-p_a)\bar{\lambda}+p_a\underline{\lambda}}\right] \\ + \underline{u}\left[(1-\mu)\frac{p_a\underline{\lambda}}{(1-p_a)\bar{\lambda}+p_a\underline{\lambda}} + \mu\frac{(1-p_a)\underline{\lambda}}{p_a\bar{\lambda}+(1-p_a)\underline{\lambda}}\right] \\ - \lambda\left[\mu\left\{\frac{p_a\bar{\lambda}}{p_a\bar{\lambda}+(1-p_a)\underline{\lambda}}\log\frac{p_a\bar{\lambda}}{p_a\bar{\lambda}+(1-p_a)\underline{\lambda}} + \frac{(1-p_a)\underline{\lambda}}{p_a\bar{\lambda}+(1-p_a)\underline{\lambda}}\log\frac{(1-p_a)\underline{\lambda}}{p_a\bar{\lambda}+(1-p_a)\underline{\lambda}}\right\}\right. \\ (1-\mu)\left\{\frac{p_a\underline{\lambda}}{(1-p_a)\bar{\lambda}+p_a\underline{\lambda}}\log\frac{p_a\underline{\lambda}}{(1-p_a)\bar{\lambda}+p_a\underline{\lambda}} + \frac{(1-p_a)\bar{\lambda}}{(1-p_a)\bar{\lambda}+p_a\underline{\lambda}}\log\frac{(1-p_a)\bar{\lambda}}{(1-p_a)\bar{\lambda}+p_a\underline{\lambda}}\right\} \\ \left. -p_a\log p_a - (1-p_a)\log(1-p_a)\right] \quad (26) \end{aligned}$$

where, $\bar{\lambda} = \exp(\bar{u}/\lambda)$ and $\underline{\lambda} = \exp(\underline{u}/\lambda)$ and p_a is the shorthand for $P(a|\mu)$, i.e., the prior(interim) to private learning probability of choosing action a.

Given λ the interim value function $V(.)$ is continuous in μ for $\mu \in [0,1]$ and continuously differentiable wrt μ in the open set $(0,1) \cap \left\{\frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}}, \frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}\right\}^C$. Since for $\mu > \frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}$ the agents don't learn privately any more and takes the action a for sure the derivative of the value function for $\mu \in \left(\frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}, 1\right)$ is given by

$$V'_\mu = \bar{u} - \underline{u} > 0.$$

Similarly, for $\mu < \frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}}$ an agent chooses action b for sure, hence the derivative of the value function for $\mu \in \left(0, \frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}}\right)$ is given by

$$V'_\mu = -(\bar{u} - \underline{u}) < 0$$

Note that the derivative of the value function at extreme interim beliefs are constant, since agents stop private learning and choose the prior favorable action. Also the cutoffs are differentiable wrt to λ,

$$\frac{d\frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}}{d\lambda} = -\frac{d\frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}}}{d\lambda} = -\frac{(\bar{u}-\underline{u})}{\lambda^2}\frac{\bar{\lambda}\underline{\lambda}}{\left(\bar{\lambda}+\underline{\lambda}\right)^2} < 0$$

hence, $\frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}(\frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}})$ is decreasing(increasing) in λ. In the limit when $\lambda \to \infty$ the value function $V(\mu)$ becomes piecewise linear in $[0,1]$ with a kink at $1/2$.

In the region, $\left(\frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}}\frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}\right)$ as the agent is learning privately the derivative of the value function is given by,

$$
\begin{aligned}
V'_\mu = \\
&\underbrace{\frac{p_a\bar{\lambda}}{p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}}\left[\bar{u}-\lambda\log\frac{p_a\bar{\lambda}}{p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}}\right]}_{(1)}-\underbrace{\frac{(1-p_a)\,\bar{\lambda}}{(1-p_a)\,\bar{\lambda}+p_a\underline{\lambda}}\left[\bar{u}-\lambda\log\frac{(1-p_a)\,\bar{\lambda}}{(1-p_a)\,\bar{\lambda}+p_a\underline{\lambda}}\right]}_{(2)} \\
&-\underbrace{\frac{p_a\underline{\lambda}}{(1-p_a)\,\bar{\lambda}+p_a\underline{\lambda}}\left[\underline{u}-\lambda\log\frac{p_a\underline{\lambda}}{(1-p_a)\,\bar{\lambda}+p_a\underline{\lambda}}\right]}_{(3)}+\underbrace{\frac{(1-p_a)\,\underline{\lambda}}{p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}}\left[\underline{u}-\lambda\log\frac{(1-p_a)\,\underline{\lambda}}{p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}}\right]}_{(4)} \\
&+\underbrace{\mu\frac{\bar{\lambda}\underline{\lambda}}{\left(p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}\right)^2}\frac{\bar{\lambda}+\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}}\left[\bar{u}-\underline{u}-\lambda\log\frac{p_a\bar{\lambda}}{(1-p_a)\,\underline{\lambda}}\right]}_{(5)} \\
&-\underbrace{(1-\mu)\,\frac{\bar{\lambda}\underline{\lambda}}{\left(p_a\underline{\lambda}+(1-p_a)\,\bar{\lambda}\right)^2}\frac{\bar{\lambda}+\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}}\left[\bar{u}-\underline{u}-\lambda\log\frac{(1-p_a)\,\bar{\lambda}}{p_a\underline{\lambda}}\right]}_{(6)}-\underbrace{\lambda\frac{\bar{\lambda}+\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}}\log\frac{1-p_a}{p_a}}_{(7)}
\end{aligned}
\tag{27}
$$

To obtain the optimal interim probability distribution we first want to analyze the sign of the derivative of the value function. For that, it would suffice to look at $\mu \geqslant 1/2$ since the value function is symmetric in μ around $\mu = 1/2$. Now rearranging terms,

$$
(5)-(6)-(7)=\underbrace{\lambda\frac{\bar{\lambda}+\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}}\log\frac{p_a}{1-p_a}}_{\substack{\geq 0 \text{ if } \mu\geq 1/2\\ \leq 0 \text{ if } \mu\leq 1/2}}\left[1-\frac{\mu\bar{\lambda}\underline{\lambda}}{\left(p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}\right)^2}-\frac{(1-\mu)\,\bar{\lambda}\underline{\lambda}}{\left(p_a\underline{\lambda}+(1-p_a)\,\bar{\lambda}\right)^2}\right]
$$

The first term is strictly positive for $\mu \in \left(1/2, \frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}\right)$, and simplifying the second term inside the bracket and plugging the value of p_a (refer equation 7) we get,

$$
\left[1-\frac{\mu\bar{\lambda}\underline{\lambda}}{\left(p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}\right)^2}-\frac{(1-\mu)\,\bar{\lambda}\underline{\lambda}}{\left(p_a\underline{\lambda}+(1-p_a)\,\bar{\lambda}\right)^2}\right]=\underbrace{1-\frac{\bar{\lambda}\underline{\lambda}}{\left(\bar{\lambda}+\underline{\lambda}\right)^2}\left(\frac{1}{\mu}+\frac{1}{1-\mu}\right)}_{\substack{\geq 0 \text{ if } 1/2\leqslant\mu<\bar{\lambda}/\bar{\lambda}+\underline{\lambda}\\ \leq 0 \text{ if } \underline{\lambda}/\bar{\lambda}+\underline{\lambda}<\mu\geqslant 1/2}}
$$

which is also strictly positive for $\mu \in \left(1/2, \frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}\right)$, and hence $(5)-(6)-(7)$ is strictly positive in the relevant range.

Now we simplify the first block in the expression of derivative of the value function (V'_μ) and rewrite it as,

$$\big((1)+(4)\big)-\big((2)+(3)\big)=\underbrace{\frac{\bar{\lambda}\underline{\lambda}\,(2p_a-1)\,(\bar{u}-\underline{u})}{\big(p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}\big)\,\big(p_a\underline{\lambda}+(1-p_a)\,\bar{\lambda}\big)}}_{(1')}+\underbrace{2\lambda}_{>0}$$

$$+\underbrace{(\bar{u}-\underline{u})\,\frac{\bar{\lambda}\left(p_a^2+(1-p_a)^2\right)+2p_a\,(1-p_a)\,\underline{\lambda}}{\big(p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}\big)\,\big(p_a\underline{\lambda}+(1-p_a)\,\bar{\lambda}\big)}}_{(2')}-\underbrace{\frac{\bar{\lambda}\underline{\lambda}\,(2p_a-1)}{\big(p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}\big)\,\big(p_a\underline{\lambda}+(1-p_a)\,\bar{\lambda}\big)}\log\frac{p_a}{1-p_a}}_{(3')}$$

Since $(2')$ is always positive we combine $(1')$ and $(3')$ to obtain,

$$\big(1'\big)-\big(3'\big)=\underbrace{\frac{\bar{\lambda}\underline{\lambda}\,(2p_a-1)}{\big(p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}\big)\,\big(p_a\underline{\lambda}+(1-p_a)\,\bar{\lambda}\big)}}_{\geqslant 0 \text{ for } \mu\geqslant 1/2}\left(\bar{u}-\underline{u}-\lambda\log\frac{p_a}{1-p_a}\right)$$

where the second term inside the bracket can be rewritten as

$$\bar{u}-\underline{u}-\lambda\log\frac{p_a}{1-p_a}=\lambda\log\frac{\frac{\bar{\lambda}}{\underline{\lambda}}\left(\mu\frac{\bar{\lambda}}{\underline{\lambda}}-(1-\mu)\right)}{(1-\mu)\,\frac{\bar{\lambda}}{\underline{\lambda}}-\mu}$$

hence, combining the two terms we get when $\mu>1/2$,

$$\big(1'\big)-\big(3'\big)=\begin{cases}\leq 0 & \text{if } \frac{\bar{\lambda}}{\underline{\lambda}}\in\left[\frac{\mu-\sqrt{2\mu-1}}{1-\mu},\,\frac{\mu+\sqrt{2\mu-1}}{1-\mu}\right]\\ >0 & \text{otherwise}\end{cases}$$

Since $\bar{\lambda}>\underline{\lambda}$, $(1')-(3')>0$ at $\mu=1/2+\epsilon$ $(\mu\downarrow 1/2)$, $(1')-(3')<0$ at $\mu=\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}-\epsilon$ $(\mu\uparrow\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}})$ where $\epsilon\to 0$, and $(1')-(3')=0$ at $\mu=\frac{\bar{\lambda}^2+\underline{\lambda}^2}{\left(\bar{\lambda}+\underline{\lambda}\right)^2}<\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$. This implies $(1')-(3')$ starts as a positive term when $\mu\downarrow 1/2$ and eventually becomes negative before μ reaches $\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$. That means the expression $\big((1)+(4)\big)-\big((2)+(3)\big)$ starts as positive for $\mu\downarrow 1/2$ but to determine the sign of the expression near $\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$ we need to look at the sign of the derivative V'_μ near $\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$(to the left of it). But V is not differentiable at $\mu=\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$ and becomes linear for $\mu>\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$, hence we only look at the left derivative of the value function at $\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$ and get,

$$\lim_{\epsilon\to 0}V'\left(\mu=\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}-\epsilon\right)=\underbrace{(\bar{u}-\underline{u})\left[1+\frac{\bar{\lambda}+\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}}\lim_{\epsilon\to 0}\left(\mu\frac{\underline{\lambda}}{\bar{\lambda}}-(1-\mu)\,\frac{\bar{\lambda}}{\underline{\lambda}}\right)\right]}_{(*)}$$

$$\underbrace{-\lambda\frac{\bar{\lambda}+\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}}\lim_{\epsilon\to 0}\left[\mu\frac{\underline{\lambda}}{\bar{\lambda}}\log\frac{p_a\bar{\lambda}}{(1-p_a)\,\underline{\lambda}}-(1-\mu)\,\frac{\underline{\lambda}}{\bar{\lambda}}\log\frac{(1-p_a)\,\bar{\lambda}}{p_a\underline{\lambda}}+\log\frac{1-p_a}{p_a}\right]}_{(**)}$$

where

$$(*) = -\lim_{\epsilon\to 0}\epsilon\frac{\bar{u}-\underline{u}}{\bar{\lambda}\underline{\lambda}\left(\bar{\lambda}-\underline{\lambda}\right)}\left(\bar{\lambda}^2+\underline{\lambda}^2\right)\left(\bar{\lambda}+\underline{\lambda}\right)=0$$

as the term multiplying ϵ doesn't depend on ϵ. For the other term $(**)$ since

$$\lim_{\epsilon\to 0}\mu\frac{\underline{\lambda}}{\bar{\lambda}}+(1-\mu)\frac{\bar{\lambda}}{\underline{\lambda}}=1$$

we can rewrite $(**)$ as,

$$(**) = \lambda\frac{\bar{\lambda}+\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}}\lim_{\epsilon\to 0}\left[\left(\mu\frac{\underline{\lambda}}{\bar{\lambda}}+(1-\mu)\frac{\bar{\lambda}}{\underline{\lambda}}\right)\log\frac{\bar{\lambda}}{\underline{\lambda}}-\left(\mu\frac{\underline{\lambda}}{\bar{\lambda}}+(1-\mu)\frac{\bar{\lambda}}{\underline{\lambda}}-1\right)\log\frac{1-p_a}{p_a}\right]$$

$$=\lambda\frac{\bar{\lambda}+\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}}\log\frac{\bar{\lambda}}{\underline{\lambda}}>0$$

using the convention $0\log 0=0$.

Hence, $(*)-(**)<0$ so the value function $V(.)$ is decreasing to the left of the cutoff. Also $V'_\mu\left(\mu=1/2\right)=0$ and $V'_\mu|_{\mu=1/2+\epsilon}>0$. This implies the value function attains minimum at $\mu=1/2$ and there exists a $\mu_{max}\in\left(\frac{\bar{\lambda}^2+\underline{\lambda}^2}{\left(\bar{\lambda}+\underline{\lambda}\right)^2},\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}\right)$ such that the value function attains an interior maximum at μ_{max} given λ and the value function is decreasing near the cutoff $\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$. Although the value function is not differentiable and has a kink at $\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$, but we have shown it attains a local minimum at this point, since the left derivative is negative and the right derivative is positive. But also,

$$V\left(\mu=1/2\right)=\frac{\bar{u}\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}+\frac{\underline{u}\underline{\lambda}}{\bar{\lambda}+\underline{\lambda}}-\lambda\left[\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}\log\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}+\frac{\underline{\lambda}}{\bar{\lambda}+\underline{\lambda}}\log\frac{\underline{\lambda}}{\bar{\lambda}+\underline{\lambda}}+\log 2\right]$$
$$<\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}\left(\bar{u}+\underline{u}\right)=V\left(\mu=\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}\right)$$

which implies V attains global minima at $\mu=1/2$[7].

As λ increases, since

$$\frac{d\frac{\bar{\lambda}^2+\underline{\lambda}^2}{\left(\bar{\lambda}+\underline{\lambda}\right)^2}}{d\lambda}=\frac{\bar{u}-\underline{u}}{\lambda^2}\left(\frac{\bar{\lambda}}{\underline{\lambda}}+1\right)\frac{\bar{\lambda}}{\underline{\lambda}}\left(1-\frac{\bar{\lambda}}{\underline{\lambda}}\right)<0$$

and $\frac{d\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}}{d\lambda}<0$ then the maximizer μ_{max} would also decrease with λ since the interval in which μ_{max} lies shifts to the left towards $1/2$.

Let an agent chooses n in the first step which generates a distribution of intermediate beliefs $\{\mu_i^n\}_{i=0}^n=Pr(\omega_1)$ where μ_i^n denotes the expected probability of being type ω_1

[7]Since V is symmetric in μ around $1/2$.

after observing i many action as out of total n observations, i.e. $\mu_i^n = E\left[pr\left(\omega_1\right)|\gamma_{x_n}\right]$. Then the agents ex-ante expected value is given by

$$W_n\left(A,\gamma\right) = \sum_{i=0}^{n} V\left(\mu_i^n\right) Pr\left(\mu_i^n|\gamma\right) - c\left(n\right) \tag{28}$$

Step 3: Apart from the constraint on the final posterior distribution of beliefs that the marginal cost of private learning λ poses on the amount of private learning by restricting positive private learning only upto a cutoff value of interim belief $\mu < 1$, λ also affects the error probabilities $\left\{\epsilon^i\right\}_{i\in A}$. Given the Bayesian updating rule and prior γ, if an agent observes n many agents from earlier generation then the error probabilities pose a limit on how much the posterior can deviate from the prior belief. Let us consider the extreme case, where an agent observes n many agents and out of this n many observations everyone had chosen action a, then by equation 10, we have

$$\begin{aligned} P\left(x_n = n|\mu\in\gamma\right) &= \sum_{k=0}^{n} \binom{n}{n-k} \mu^{n-k}\left(\epsilon^a\right)^k\left(1-\mu\right)^k\left(1-\epsilon^b\right)^{n-k} \\ &= \left(\epsilon^a + \mu\left(1-\epsilon^a-\epsilon^b\right)\right)^n \end{aligned} \tag{29}$$

whereas if every n agents had taken action b, then

$$\begin{aligned} P\left(x_n = 0|\mu\in\gamma\right) &= \sum_{k=0}^{n} \binom{n}{k} \mu^{k}\left(\epsilon^a\right)^{n-k}\left(1-\mu\right)^{n-k}\left(1-\epsilon^b\right)^{k} \\ &= \left(\epsilon^b + \left(1-\mu\right)\left(1-\epsilon^a-\epsilon^b\right)\right)^n \end{aligned} \tag{30}$$

and the expected probability of being type ω_1 under new belief γ_n' generated by Bayes rule would be

$$E\left(\mu|x_n,\gamma\right) = \frac{\int_{\mu\in\gamma}\mu\left(\epsilon^a+\mu\left(1-\epsilon^a-\epsilon^b\right)\right)^n d\gamma}{\int_{\mu\in\gamma}\left(\epsilon^a+\mu\left(1-\epsilon^a-\epsilon^b\right)\right)^n d\gamma}$$

If $\epsilon^a + \epsilon^b = 1$, then $E\left(\mu|x_n,\gamma\right) = E\left(\mu|\gamma\right)$, i.e. the prior is same as the posterior. Now, $\epsilon^a + \epsilon^b$ can be used as measure of mismatch error and the highest value it can take is 1. Also

$$\frac{d\left(\epsilon^a+\epsilon^b\right)}{d\lambda} = 2p_a\left(1-p_a\right)\frac{\bar{\lambda}\underline{\lambda}}{\lambda}\log\frac{\bar{\lambda}}{\underline{\lambda}} > 0 \quad \forall p_a,\bar{\lambda},\underline{\lambda}. \tag{31}$$

That means as λ goes up, the maximum value of the posterior given n observations is closer to the prior. Thus λ gives another bound on γ_{x_n}' via social learning channel.

Step 4: Now we would consider the interaction between the bound posed by social learning and the shape of the value function wrt μ. Given a common prior γ let us define $\bar{\mu}^n$ as the furthest Bayesian posterior (interim in the model) belief about Γ from prior γ after observing n many agents from earlier generation. Since the furthest from prior belief is generated by the extreme cases where everyone is the sample had chosen a or everyone had chosen b, $\bar{\mu} = \max\left\{\mu_n^n, \mu_0^n\right\}$. Let us assume that $\bar{\mu} = \mu_n^n$ [8].

[8]Here it is assumed that $E\left(\mu|\gamma\right) \geqslant 1/2$. Later in note 2 we would show that this is without loss of generality.

Fix λ and γ, then given the shape of the interim value function wrt μ and $\bar{\mu}^n$ there are three possibilities,

1. $0 \leqslant \bar{\mu}^n \leqslant \mu_{max}$, i.e the highest μ an agent can have after observing everyone taking action a is in the initial increasing part of $V(\mu)$ function,

2. $\mu_{max} < \bar{\mu}^n \leqslant \frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$, i.e the highest μ an agent can have after observing everyone taking action a is in the decreasing part of $V(\mu)$ function,

3. $\bar{\mu}^n > \frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$, the highest μ an agent can have after observing everyone taking action a is in the final linear increasing part of $V(\mu)$ function.

Next we would argue that no agent would choose an n such that, $\bar{\mu}^n$ lies in the decreasing part of $V(\mu)$. Here we only consider one side of the value function namely $\mu \geq 1/2$ but the other part would be symmetric. Now, since agents are Bayesian, a bigger sample, (say n_2) with all a action would move their interim belief μ further towards 1 compared to a smaller n, (say n_1) where $n_1 < n_2$. Also, n_2 spreads the distribution of μ_x^n more than n_1. Since we consider only $\mu \geq 1/2$ case then the agent has to observe at least more than half of the observations are choosing a. And upon observing $\left\lceil \frac{n_1}{2} \right\rceil \leqslant x_{n_2} \leqslant n_1$ i.e., more than half of n_1 but less than n_1 many people taking action a, the belief generated by n_2 $\mu_{x_{n_2}}^{n_2}$ would be lower than $\mu_{x_{n_2}}^{n_1}$.

This means if the optimal n, say n_2 is such that $\bar{\mu}^{n_2}$ is in the decreasing zone of the value function, then there exists $n_1 < n_2$, such that reducing the optimal observation to n_1 would bring $\bar{\mu}^{n_1}$ back to increasing part of $V(\mu)$ and μ_{n_1} is closest to μ_{max} for all possible choice of $n_1 \leq n_2$. Now to generate a belief higher than $\bar{\mu}_{n_1}$ the agent needs to observe x_{n_2} which is atleast greater than n_1, which implies the $Pr\left(\mu_{x_{n_2}} > \bar{\mu}_{n_1}\right) \leq Pr\left(\mu_{x_{n_1}} = \bar{\mu}_{n_1}\right)$ and all such events generate a lower payoff than $\bar{\mu}_{n_1}$. For all other $\left\lceil \frac{n_1}{2} \right\rceil \leqslant x_{n_2} \leqslant n_1$, the resulting μ_x will be higher under n_1. This means for all such x_{n_2}, choosing n_1 instead of n_2 would generate a higher payoff. [9] This implies the expected benefit from observing n_1 many agents is weakly higher than that of n_2.

Since $c(n_2) \geqslant c(n_1)$ but $\sum_{i=0}^{n_2} V\left(\mu_i^{n_2}\right) Pr\left(\mu_i^{n_2}|\gamma\right) \leqslant \sum_{i=0}^{n_1} V\left(\mu_i^{n_2}\right) Pr\left(\mu_i^{n_1}|\gamma\right)$, the agent would be better off by choosing $n_1 < n_2$. Hence no agent would choose n^* such that $\bar{\mu}^{n^*}$ falls in the decreasing section of $V(\mu)$. So an agent would either choose an n that restricts him to the initial increasing part of $V(\mu)$ or a high enough n that would take him to the final linear increasing part.

For the agent who would choose an n that takes him to the final increasing part of $V(\mu)$ the choice of n would have to be high enough such that not only $\bar{\mu}^n < \frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$ but also $V(\bar{\mu}^n) \geqslant V(\mu_{max})$, i.e. the maximum possible value is atleast greater that the interior maximum, otherwise deceasing n to remain in the initial increasing part would make him better off, by saving the cost of observing the extra agents. Let λ_n^j denote the minimum level of marginal cost of private learning given n, such that the expected value of choosing an n that restricts the interim belief to the initial increasing zone is weakly smaller than that of going to the final increasing zone.

[9] Similar argument holds for other values of x_n, since $V(\mu)$ is symmetric and n_1 spreads the interim distribution of beliefs lesser than n_2. If max in both side of $\mu = 1/2$ falls in decreasing section then choose n_1 to bring the maximum of the two ($\bar{\mu}^n$) in the increasing section, which will automatically bring the other (minimum of the two) to the increasing section.

Then for $\lambda < \lambda_n^j$ and $\bar{\mu}^n(\lambda) > \frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$ there exists n_1^λ such that $\bar{\mu}^{n_1^\lambda} < \mu_{max}(\lambda)$ and $\forall n$ with $\bar{\mu}^n(\lambda) > \frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$, we have

$$\sum_{i=1}^{n_1^\lambda+1} V\left(\mu_i^{n_1^\lambda}\right) Pr\left(\mu_i^{n_1^\lambda}|\gamma\right) - c\left(n_\lambda^1\right) \geq \sum_{i=1}^{n+1} V\left(\mu_i^n\right) Pr\left(\mu_i^n|\gamma\right) - c(n)$$

and for $\lambda \geqslant \lambda_n^j$, $\exists n$ s.t. $\bar{\mu}^n(\lambda) > \frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}$ and $\forall n_1^\lambda$ with $\bar{\mu}^{n_1^\lambda} < \mu_{max}(\lambda)$, we have

$$\sum_{i=1}^{n_1^\lambda+1} V\left(\mu_i^{n_1^\lambda}\right) Pr\left(\mu_i^{n_1^\lambda}|\gamma\right) - c\left(n_\lambda^1\right) < \sum_{i=1}^{n+1} V\left(\mu_i^n\right) Pr\left(\mu_i^n|\gamma\right) - c(n)$$

Hence, for $\lambda < \lambda_n^j$, it would be optimal to reduce n and stay in the initial increasing zone and for $\lambda \geqslant \lambda_n^j$, choosing n to go to the final increasing part of the value function would be optimal.

Note 1: Note that since lower λ has a higher μ_{max}, given n and γ, the continuity of the V function with respect to λ and Bayesian updating in stage 1, we can find two thresholds on λs, namely $\lambda_{n,\gamma}^1$ and $\lambda_{n,\gamma}^2$ such that for $\lambda \leqslant \lambda_{n,\gamma}^1$, $\bar{\mu}^n$ would be in the initial increasing part, for $\lambda_{n,\gamma}^1 < \lambda \leqslant \lambda_{n,\gamma}^2$, $\bar{\mu}^n$ would be in the decreasing part and all $\lambda > \lambda_{n,\gamma}^2$, $\bar{\mu}^n$ would be in the final increasing part of the value function.

Note 2: For step 4 we assumed $E\left(\mu|\gamma\right) \geqslant 1/2$. In the other case, $\bar{\mu}^n$ would be the belief generated by observing n many agents taking action b, since by definition $\bar{\mu}^n = \max\{\bar{\mu}_n^n, \bar{\mu}_0^n\}$. Step 4 would be WLOG since if the maximum of the two is bounded by choice of n in the initial increasing part of $V(\mu)$ then the minimum would automatically be bounded in the initial increasing part of the value function as higher n spreads the belief away from $E\left(\mu|\gamma\right)$[10]. So considering the case of $E\left(\mu|\gamma\right) \geqslant 1/2$ is indeed without loss of generality for step 4. Since step 5 would use result from step 4, we would continue with our assumption without loss of generality.

Step 5 By definition of the social cost function as given in equation 3, there exists $\bar{n} \leqslant N$ such that $c(\bar{n}-1) \leqslant \bar{u} \leqslant c(\bar{n})$, which means agents will never observe more than $\bar{n}$ many people. Let us assume (for now) that $\bar{n}$ is such that the set of λs in all three regions using $\bar{\mu}^{\bar{n}}$ as discussed in step 4 is non-empty. We will discuss all other possibilities later.

Now consider $\lambda < \lambda_{\bar{n}}^j$ such that $\bar{\mu}^{\bar{n}} > \mu_{max}(\lambda)$. For all these λ, the optimal $n < \bar{n}$ and $\bar{\mu}^n \leq \mu_{max}(\lambda)$, since reducing n to the initial increasing part of $V(\mu)$ is better for the agent. And for all $\lambda \geqslant \lambda_{\bar{n}}^j$ it would be better to increase n such that $\bar{\mu}^n$ lies in the final increasing section. Define, λ' as

$$\lambda'_{\bar{n}} = \max\left\{\lambda > 0 | \bar{\mu}^{\bar{n}}(\lambda) \leqslant \mu_{max}(\lambda)\right\} \tag{32}$$

Hence there are three different intervals,

[10]For $\lambda \geqslant \lambda_n^j$ we have already considered the expected value of the ex-ante value function so the assumption $E\left(\mu|\gamma\right) \geqslant 1/2$ doesn't have a bite and no λ would choose n such that $\bar{\mu}^n$ is in the decreasing part of $V(\mu)$, hence we have covered all possible cases.

1. $0 \leqslant \lambda \leqslant \lambda'_{\bar{n}}$ have $\bar{\mu}^{\bar{n}}$ in the initial increasing zone

2. $\lambda'_{\bar{n}} < \lambda < \lambda^j_{\bar{n}}$ have $\bar{\mu}^{\bar{n}} > \mu_{max}$ and would choose $n < \bar{n}$ to restrict to the initial increasing zone

3. $\lambda \geqslant \lambda^j_{\bar{n}}$, would choose as high n as possible subject to the social cost function to be in the final linear increasing zone.

Step 6: Let the optimal n for $\lambda = \lambda'_{\bar{n}}$ be n_1^*, then the optimality condition would give

$$\sum_{i=0}^{n_1^*} V\left(\mu_i^{n_1^*}\right) Pr\left(\mu_i^{n_1^*}|\gamma\right) - \sum_{i=0}^{n_1^*-1} V\left(\mu_i^{n_1^*-1}\right) Pr\left(\mu_i^{n_1^*-1}|\gamma\right) \geqslant c\left(n_1^*\right) - c\left(n_1^* - 1\right)$$

Note that $c(.)$ is same for all λs, but for $\lambda \to 0$,

$$\lim_{\lambda \to 0+} \sum_{i=0}^{n_1^*} V\left(\mu_i^{n_1^*}\right) Pr\left(\mu_i^{n_1^*}|\gamma\right) - \sum_{i=0}^{n_1^*-1} V\left(\mu_i^{n_1^*-1}\right) Pr\left(\mu_i^{n_1^*-1}|\gamma\right) = 0 \leqslant c\left(n_1^*\right) - c\left(n_1^* - 1\right)$$

So for $\lambda \to 0$ the optimal n would not be higher than n_1^*. Also for all $\lambda \leqslant \lambda'_{\bar{n}}$, the optimal n would lead to an interim belief μ such that $V'(\mu) > 0$. Now we make the following observations:

1. For a given $\mu < \bar{\mu}^{n^*}(\lambda)$, $\lim_{\lambda \to 0} V'_{\mu\lambda} \downarrow 0$, which implies there exists $\lambda''_{\bar{n}}$ such that V_μ becomes flatter with decrease in λ for all such μ and for all $\lambda \leqslant \lambda''_{\bar{n}}$ by continuity of V'_μ wrt λ. Since the social cost function $c(n)$ is same for all λ but a flatter V_μ means smaller change in the benefit from an increase in μ, n^* would be non-decreasing λ in this region of values of λ.

2. Since, for lower λ μ_{max} is higher, $\lim_{\lambda \to \lambda'_{\bar{n}}-} V'_{\mu\lambda} \leqslant 0$ for $\mu \in \mathcal{N}_\epsilon\left(\bar{\mu}^{n^*}\left(\lambda'_{\bar{n}}\right)\right)$, then the optimal n for a λ which is smaller but close to $\lambda'_{\bar{n}}$ would be not lower than n_1^*, i.e n^* would be non-increasing in λ in left neighborhood of λ'_n.

Combining these two observations and the fact that V'_μ is continuous in λ we get that there exists a $\lambda^* \in [\lambda'', \lambda']$[11] such that n^* is non-decreasing for $\lambda \leqslant \lambda^*$ and non-increasing for $\lambda \geqslant \lambda^*$. This proves the part (i) of the theorem[12].

Step 7: For $\lambda' < \lambda < \lambda^j$, the optimal strategy is to choose n^* such that $\bar{\mu}^{n^*}$ remains in the increasing part of $V(\mu)$. For lower λ, μ_{max} is higher (by note 1) and a higher n increases $\bar{\mu}^n$, thus the optimal n^* would be non-increasing in λ for $\lambda \in \left(\lambda', \lambda^j\right)$ [13]. Combining this with Step 6 we get n^* is non-increasing for $\lambda \in \left[\lambda^*, \lambda^j\right)$.

For $\lambda > \lambda^j_{\bar{n}}$, if λ is such that social learning is also informative, i.e., $\lambda \leqslant \lambda^{**}$ (See result 2) then the optimal strategy is to choose the maximum possible n given the cost

[11] Since given a social cost function, $\bar{n}$ is fixed, we drop the $\bar{n}$ subscript.

[12] If the social cost function is such that $c(\bar{n}) - c(\bar{n} - 1) = 0$, then $\lambda^* = \lambda'' = 0$

[13] This is because given n when a lower λ remains in the increasing part a higher λ might do to the decreasing part. So a higher λ would then decrease n further.

of social learning. Since $\epsilon^a + \epsilon^b$ is increasing in λ, i.e. $\bar{\mu}^n$ is decreasing in λ given n. But the cost function is same for all λ implies for $\lambda_1 < \lambda_2$, given n

$$\sum_{i=0}^{n} V_{\lambda_1}\left(\mu_i^n\right) Pr\left(\mu_i^n|\gamma\right) - c\left(n\right) \geqslant \sum_{i=0}^{n} V_{\lambda_2}\left(\mu_i^n\right) Pr\left(\mu_i^n|\gamma\right) - c\left(n\right)$$

This implies in this range of λ, i.e., $\lambda \in \left(\lambda^j, \lambda^{**}\right]$ also n^* is non-increasing in λ. This concludes the proof of part (ii) of the theorem.

Step 8: For $\lambda < \lambda^j$, the optimal choice was to reduce n such that $\bar{\mu}^n$ lies in the increasing part, whereas for $\lambda > \lambda^j$ the optimal choice of n is such that $\bar{\mu}^n$ in in the final linear increasing part of the $V\left(\mu\right)$. Since $\mu_{max}\left(\lambda\right) < \frac{\bar{\lambda}}{\lambda+\underline{\lambda}}$, and agents use Bayesian updating rule, $n^*\left(\lambda^j - \delta\right) < n^*\left(\lambda^j + \delta\right)$, for $\delta \to 0$. This proves the part (iii) of the theorem.

Using result 2, there exists $\lambda^{**} \leq \infty$ such that for $\lambda > \lambda^{**}$, the social learning is completely uninformative, hence, $n^* = 0$. This proves part (iv) of the theorem[14].

Step 9: Given a social cost function $c\left(n\right)$ and a common prior γ there are few other possibilities (refer step 5) that we haven't discussed yet. For this step also we will assume $E\left(\mu|\gamma\right) \geqslant 1/2$ without loss of generality since all results would go through if we use the definition of $\bar{\mu}^{\bar{n}}$ as the maximum of the two possibilities (as discussed in Note 2 earlier).

First, suppose $\bar{n}$ is small enough such that $\bar{\mu}^{\bar{n}}$ lies in the initial increasing section of $V\left(\mu\right)$ for every λ. But this is not possible unless $E\left(\mu|\gamma\right) = 1/2$ and $\bar{n} = 0$. This is true because for any other γ such that $E\left(\mu|\gamma\right) > 1/2$, the prior is atleast ϵ away from $1/2$ and as $\lambda \to \infty$ the initial increasing section shrinks to the point $1/2$, which implies even with $n = 0$, $\bar{\mu}^n = E\left(\mu|\gamma\right)$ would be in the final increasing part of $V\left(\mu\right)$. And for $E\left(\mu|\gamma\right) = 1/2$, if λ is large but finite ($\lambda < \infty$ and $\lambda \to \infty$) then any $n > 0$ would shift the belief to final increasing section for such a high λ ($\lambda \to \infty$) by similar logic. But $\bar{n} \geqslant 1$ by definition, so this can't happen. Hence, the set of λs in the final increasing section is not null.

Second, suppose $\bar{n}$ is large enough such that $\bar{\mu}^{\bar{n}}$ lies in the final increasing section for all λ. But this implies for $\lambda \to 0$, the belief after observing $\bar{n}$ many people taking action a (or b) has to go to 1 (or 0). For any finite N, with $\bar{n} \leq N$, this is not possible unless $E\left(\mu|\gamma\right) = 1$(or 0). Since the true distribution $\mu^* \in int\left(\gamma\right)$ by assumption, this isn't possible (almost surely). Hence, the set of λ in the initial increasing section is never empty. By continuity of the value function V in λ and step 4 the only possibility thus left is where all three sets of λs, i.e. $\bar{\mu}^{\bar{n}}$ is in initial increasing, decreasing and final increasing sections of $V\left(\mu\right)$ are nonempty. This completes the proof of the theorem.

□

A.3 Proof of Theorem 2

Proof. The proof of the theorem uses the results from theorem 1. In this particular case, since $E\left(\mu|\gamma\right) = 1/2$, $P\left(a|\gamma\right) = 1/2$ for generation $t = 0$. This means all agents in generation $t = 0$ learns privately for all $\lambda < \infty$, hence social learning is informative for any $t = 1$.

[14]For $E\left(\mu|\gamma\right) = 1/2$, the case not discussed in result 2, we would have $\lambda^{**} = \infty$.

Since $P\left(a|\gamma\right) = 1/2$, some learning is always optimal for all $\lambda < \infty$. Combining this with the fact that some social learning is informative and optimal for $t = 1$, it would be informative and optimal for any $t > 1$ as well. This is true because social learning can possibly shift the belief towards one action since there is heterogeneity in the behavior of the agents. Any shift away from $\mu = 1/2$ is better for an agent as the interim value function attains global minima at $\mu = 1/2$. This means $\lambda^{**} = \infty$, which proves the first part of the theorem.

For the second part, at λ', $V'|_{\bar{c}} \leqslant 0$ since $C\left(n\right) - C\left(n-1\right) = 0$ and agents are optimizing. This implies for all $\lambda < \lambda'$, the slope would be higher for $n = \bar{c}$, since $\bar{\mu}^{\bar{n}} < \mu_{max}$ for all these λ. Then using theorem 1 we have $\lambda^* = 0$, hence all $\lambda \leqslant \lambda'$ would choose $n^* = \bar{c}$, since the slope of the expected value function would be positive even in the extreme case of observing all $\bar{c}$ many agents choosing a (or b) but the cost function is constant at 0.

For $\lambda > \lambda^j$ also the incremental cost of observing one additional action is zero for any $n \leqslant \bar{c}$, but the change in the expected value is positive however small the shift in belief be, since the slope of the value function is positive. This implies for all $\lambda \geqslant \lambda^j$, the optimal $n^* = \bar{c}$.

Finally for $\lambda \in \left(\lambda', \lambda^j\right)$, we have $\bar{\mu}^{\bar{c}}$ in the decreasing part of the value function. Hence by step 4 of theorem 1 the optimal choice for all such λ would be an n^* such that $n^* < \bar{n}$, also n^* would be non-increasing in this part. This is because a a higher λ attains interior maximum at a belief closer to $1/2$, so a strictly higher n for a higher λ might lead to $\bar{\mu}^{\bar{c}} > \mu_{max}$, which is worse for the agent. This completes the proof of part 2. □

A.4 Proof of Theorem 3

Proof. As we have already discussed, under sequential learning agents choose a set of n conditional on belief in equilibrium. First we discuss the position of the n_{min} in terms of beliefs in an equilibrium for different values of λ. Then we use ideas from proof of theorem 1 to complete this proof. For the rest of the proof we would only consider the value function $V\left(\mu\right)$ for $\mu \geq 1/2$ as the other case would be symmetric. So a higher belief, i.e., higher value of μ would mean a belief further away from uniform belief which is a more informative belief as well.

Step 1: In this step we discuss the position of n_{min} for different values of λ.For notational simplicity we drop the time superscript. We know the value function $V\left(\mu\right)$ is C^2 in the domain $\left(\underline{\lambda}/\bar{\lambda} + \underline{\lambda}, \bar{\lambda}/\bar{\lambda} + \underline{\lambda}\right)$ and attains an interior minimum at $\mu = 1/2$ and an interior local maximum at μ_{max}. This implies the function is locally concave near μ_{max} and locally convex near $1/2$. If n_{min} is at some level of belief μ_1, which means an agent would optimally choose to top learning socially after observing n many actions when his belief is μ_1, then at μ_1 the marginal gain from observing one more action would be least among all choices of n. This is true because the cost function is weakly convex implies a higher n generates a higher increase in marginal cost. Since an agent would only choose to stop learn if marginal gain is less than marginal loss, where the loss is due to extra cost, then n_{min} has to be associated with lowest marginal gain. This gives us a natural candidate for n_{min} which is closest to the μ_{max} as the function attains local maxima at that point and hence would be flattest there.

But as we noted earlier in the proof of theorem 1 the cost of social learning function puts a restriction on how much an agent can learn by imposing a maximum value of

n, namely $\bar{n}$.Let us consider only those λs for which the maximum possible belief at $\bar{n}$ remains below μ_{max}. For a small enough λ in that range the $\bar{n}$ restricts the belief away from μ_{max} to a lower value. This means n_{min} may not be associated with the belief closest to μ_{max}. For any such λ, the marginal gain is thus lowest for a choice of n that keeps the belief closest to $1/2$ due to the locally convex nature of the value function near $\mu = 1/2$.

But for a high enough λ when $\bar{n}$ is such that there is a choice of n where the belief is very close to the μ_{max} then that n would generate lowest marginal gain and become the n_{min}. For intermediate value of λ the smallest μ would generate n_{min} if the marginal gain is lower at the smallest μ compared to the μ closest to μ_{max}.

Now consider the case where λ is such that the maximum possible belief lies in $\left(\mu_{max}, \bar{\lambda}/\bar{\lambda} + \underline{\lambda}\right)$. This implies that for all such λ as agent can have a belief in the decreasing part of the value function $V(\mu)$. But unlike the case of block learning an agent might choose an n such that he optimally ends up with a belief in the decreasing part. The reason behind this is as follows: under sequential learning an agent decides whether to observe another action standing at some belief and n combination so if the agent has a belief not very close to μ_{max} but such that observing one action would lead him to the decreasing part where the expected marginal gain is higher than the marginal loss then he would choose to observe one more n and would probably end at a belief in the decreasing part.

But again there is a cutoff belief lower than the maximum possible belief $\bar{\mu}$ under $\bar{n}$ in the decreasing part of the value function such that an agent would never choose to observe any more actions standing at that belief. The logic is similar to the one used in the proof of theorem 1. We know an agent would only observe an extra action if the expected marginal gain is higher. And also we know a higher n spreads the distribution of beliefs. Given $\bar{n}$, for all these λs the $\bar{\mu}$ would remain in the decreasing part and hence the marginal gain would become negative for a high enough $\mu \leq \bar{\mu}$ due to spreading of the distribution of belief. Since the cost function is weakly convex this implies the agents would only learn until the marginal gain is higher than the cost and that restricts the choice of n. Since a value further sway from μ_{max} in the decreasing section would more likely generate an even lower value on $V(\mu)$ because of an increased probability the cutoff must remain close enough to μ_{max}.

For all these λ the n_{min} would remain closest to μ_{max} because the of two reasons. First the value function is flattest near μ_{max} due to local concavity and second a higher belief in the decreasing section is restricted by a cutoff belief close to μ_{max}. The first condition implies no belief to the left of μ_{max} and further away from it would generate the n_{min} and both first and second part combined make sure that a belief further away in the decreasing section would not generate n_{min} due to local concavity of the value function and that fact that the cutoff would not be further away from μ_{max} which implies the local concavity argument still holds true.

For all the λs such that the $\bar{\mu}$ falls in the final increasing section the agent would only choose to learn upto a belief higher than $\bar{\lambda}/\bar{\lambda} + \underline{\lambda}$ only if the marginal gain is higher. As λ increases the $\bar{\mu}$ is further away from $\bar{\lambda}/\bar{\lambda} + \underline{\lambda}$ which implies the marginal gain becomes higher. This implies there exists a minimum value of λ, say λ_s^j such that an agent would start to choose to learn upto a belief that is higher than $\bar{\lambda}/\bar{\lambda} + \underline{\lambda}$, i.e. in the final increasing section.

For all $\lambda < \lambda_s^j$, the n_{min} remains the one corresponding to the belief closest to the μ_{max} but for $\lambda \geq \lambda_s^j$ that would not be the case. For these set of higher λs the n_{min} would be in the final increasing part. First of all the earlier candidate for n_{min} namely the one

closest to μ_{max} would not remain so because of the following reason: if a belief closer to μ_{max} has a lower n than that of the one in the final decreasing part, then the marginal gain from choosing another observation would be lower for the former compared to the latter since the marginal increase in cost is lower for the former. But we know for these λs the marginal gain to move into the final increasing section is higher than the marginal loss and they try to learn as much as possible which implies for the lowest n if it is near μ_{max} the marginal gain can't be lower than marginal loss as it would imply the agent would never learn upto the final increasing section. Also for all other belief the marginal gain is higher than that of the one closest to μ_{max} which increases the incentive to learn and hence the n_{min} would be in the final increasing section.

Step 2: Now that we have the position of n_{min} for different values of λ, we can prove the theorem. Let us start with very high values of λ. When λ is very high and above some threshold λ^{**}, as proved in the result 2, the social learning becomes completely uninformative because any agent in period 1 would know that period 0 agents have not done any private learning and would do no learning of any kind which would imply no later generation would learn as well. This proves the part 4 of the theorem.

Define λ_1 as the maximum value of λ such that n_{min} remains closest to 1/2. Using step 6 of theorem 1 when $\lambda < \lambda_1$ and close to 0, as λ increases the value function becomes steeper which implies the marginal gain from social learning at weakly higher for a higher λ near 1/2. This implies there exists a maximum value of λ say $\lambda_s^* \leq \lambda_1$ where n_{min} is non-decreasing. This proves part 1 of the theorem.

For any λ higher than λ_1 the n_{min} is closest to μ_{max}. Let λ_2 denote the maximum value of λ such that $\bar{\mu} \leq \mu_{max}$. Again using step 6 of the proof of theorem 1 which shows that for a choice of n that is close enough to μ_{max} optimally n would be non-increasing in λ. So there exists a minimum value of λ say $\lambda_s^i \geq \lambda_1$ such that for all $\lambda \in \left[\lambda_s^i, \lambda_2\right]$ the n_{min} would be non-increasing.

For $\lambda \in \left(\lambda_2, \lambda_s^j\right)$ the n_{min} still remains the one closest to μ_{max} and for low enough λ since $\bar{\mu}$ is smaller there exists a maximum value of λ say λ_s^d such that the n_{min} remains to the left of μ_{max},since the marginal gain from choosing to go the decreasing section is limited by $\bar{n}$. Thus for all such $\lambda \leq \lambda_s^d$ the n_{min} would be non-increasing using the step 6 of theorem 1. This completes the proof of the part 2 of the theorem.

Finally at λ_s^j the n_{min} shifts from near μ_{max} to the final increasing section, which implies n_{min} makes a upward jump at λ_s^j as a strictly higher belief corresponding to n_{min} can only be obtained by a strictly higher choice of n_{min} for sufficiently close λs in the neighborhood of λ_s^j (remember the derivative of the value function is continuous in λ). This proves the part 3 of the theorem and completes the proof of the theorem.

□