Too Costly To Follow Blindly: Endogenous Learning and Herding

Srijita Ghosh *†

March 2024

Abstract

I analyze the impact of endogenizing social and private learning in a *herding* problem. Private learning is modeled à la *rational inattention* literature. I find a non-monotone relationship between social and private learning. They are substitutes when private learning is sufficiently cheap and become complement for higher private learning costs and eventually become uninformative. This happens because an increase in private learning costs makes social learning less informative. As an implication, only the reduction of the cost of private learning unambiguously increases welfare contrary to the *herding* result, where restricting social learning initially is optimal.

Keywords: Herding, Rational Inattention, Social Learning
JEL Code: D82, D83, D89

*I greatly benefited from the comments and suggestions of Andrew Caplin, Debraj Ray, David Pearce, and Parikshit Ghosh. I am also grateful to seminar participants at EEA 2019, ACEGD ISI Delhi 2016, and SERI 2020 for their comments and suggestions. All mistakes are my own.

†Economics Department, Ashoka University, email: srijita.ghosh@ashoka.edu.in

1 Introduction

This paper explores the role of observational learning, also known as social learning on private learning behavior when agents have heterogenous preferences and learning is costly. Starting with Banerjee (1992) and Bikhchandani et al. (1992) a vast literature on observational learning has explored how observational learning interacts with private information and under what conditions it leads to complete learning, i.e., private learning does not stop as the time horizon goes to infinity. Across various settings with homogenous (Banerjee (1992), Bikhchandani et al. (1992)) or heterogenous (Smith and Sørensen (2000)) preferences, with cost of private learning ((Hendricks et al. (2012), Mueller-Frank and Pai (2016), Ali (2018), Burguet and Vives (2000)) or cost of observational learning (Kultti and Miettinen (2006)) there are two robust findings. First, complete learning does not pccur except in a few special cases since agents free ride on social information by observing other agents. Second, as agents ignore their private information, they follow the actions of their predecessors that leads to an inefficient equilibrium, a phenomenon knows as *herding* in the literature.

Both these two finding across various settings is based on the observation that social learning becomes a substitute for private information which allows agents to free ride on social information and discard private information. Moreover, if private learning is costly, that creates a incentive to not learn privately, unless the cost of sufficiently low. By adding cost for observational learning as well I want to answer two main questions. First, does the two types of learning remain substitute for all ranges of relevant parameter or can they become complements as well? Second, under what conditions complete learning occurs and in absence of complete learning, does herding remain an equilibrium when both types of learning is costly.

In this model, a short-lived agent decide the optimal level of private information acquisition and social learning to make a decision in a binary choice context. Agents are heterogenous in their preference type. As they enter the economy the agents do not know their type but have a common prior over possible distribution of types. Based on a critical observation made by several earlier papers (Smith and Sørensen (2000),Hendricks et al. (2012), Mueller-Frank and Pai (2016), Burguet and Vives (2000)) that under heterogeneous preference, social and private learning are informative about different payoff relevant components, I assume that social learning informs the agent about the distribution of possible types of preferences however, is silent about the idiosyncratic preference type, whereas private learning informs the agent only about the idiosyncratic preference type.

Existing literature have modeled the cost of private learning in three major ways, namely, search cost (Hendricks et al. (2012), Mueller-Frank and Pai (2016)), cost of buying precision of (normal) signals (Burguet and Vives (2000), Bobkova and Mass (2022)) and cost of buying experiments (Ali (2018)). Under search cost, the models require a restrictive assumption that the agent cannot buy a product he has not searched. Whereas in the models with cost of purchasing precision requires a well-defined signal structure, e.g., signals following normal distribution. Following similar ideas I cosnider a model where agents buy experiments directly and instead of explicitly modeling the signal stricture I assume the cost is a function of the posterior belief disctribution generated by the experiments. The specific cost structure is modeled following the eentropy cost constraints in the Rational Inattention (RI) literature. For the social learning, similar to the existing literature I assume that it takes the form of observing actions chosen by predecessors. Furthemore, I assume observational cost is weakly increasing and weakly convex in the number of observations. Kultti and Miettinen (2006) assumes a linear cost of social learning, which is a special case of this model. Both types of learning being costly is only considered in Bobkova and Mass (2022) so far. However, they do not explicitly model the cost structure for the two types of learning and assume that agents have a total budget for learning. This model differs significantly, since first, there is no constraint on total budget for learning and second, since the two types of learning happens in completely different ways, I assume very different cost structure for each of them.

The main findings of the paper are as follows: first, I show that private learning and social learning are not substitutes for the entire range of relevant parameters. More specifically, given a cost of social learning function as the marginal cost of private learning increases, there exists athreshold value beyond which the two types of learning become complements. Second, contrary to the existing literature, complete learning occurs even at non-zero cost of private learning, moreover herding is not always an equilibrium in absence of complete learning. Given a cost of social learning function, herding becomes an equilibrium only for intermediate levels of marginal cost of private learning.

The standard trade-off through free riding on obsevational learning is still present in this model, that generates the substitutatbility of two types of learning. However, the cost structure generates one additional trade-off. As the cost of private learning increases, switching to more observation learning may not always be optimal. This is beacuse if it harder to learn individullay (due to higher cost of private learning), the information content of the observational learning would also be lower. A Bayesian should consider this and choose his learning strategy accordingly. When the second effect dominates, the two types of learning become complements.

In contrast to the existing literature, I find a range of values for the marginal private cost of learning parameter, where complete learning occurs. To get completeness of learning we do not need any restrictions on the boundedness of signals (e.g. Smith and Sørensen (2000)), or presence of public information (e.g., Hendricks et al. (2012)). This result also differes from Burguet and Vives (2000), Mueller-Frank and Pai (2016), andAli (2018), all of whom requires of the private learning cost to be sufficiently small. In their models complete learning is only possible if marginal cost of learning is zero at zero (Burguet and Vives (2000)) or if the cost of private learning goes is not bounded away from zero Mueller-Frank and Pai (2014), or beliefs can be changed for small enough cost Ali (2018).

In contrast, Goeree et al. (2006) considers a model with richness of private values such that learning incentive does not go to zero unless public beliefs are degenerate. In this model, public beliegs never become degenrate, but because of private learning can stop as it gets sufficiently costly. This finding is a common feature of learning models in *Rational Inattention* literature. Following Sims (2003), Matějka and McKay (2015) solve a discrete choice problem under the assumption that paing attention or learning is cognitively costly. They modeled the cost of learning as a linear function of the *Shannon's relative entropy* between the prior and the posterior belief and showed the optimal stochastic choice takes the form of multinomial logit. Caplin and Dean (2015) gives axiomatic characterization for costly information acquisition problems. Caplin et al. (2019) showed that rationally inattentive behavior implies existence of endogenous consideration sets, i.e., the agent doesn't always choose each possible action with positive probability even under non-degenrate belief. Caplin et al. (2015) explored the role of exogenous social learning of market share in a model of rational inattention. They found that observing that the externality as found in the observational learning continues to play a role in optimal learning strategy and leads to inefficiency. This paper is closest to Caplin et al. (2015) except I assume both types of learning are endogenous and costly. This allows me to explore the substitutability of complementarity between two types of learning.

The rest of the paper is arranged as follows. Section 2 describes the two cost structures and sets up the baseline model. In section 3, I solve the agent's optimization problem and find conditions for complete learning, and section 4 concludes.

2 Model

2.1 Environment

Timing: Consider an infinite horizon economy in discrete time, i.e. $t \in \{0, 1, \ldots, \infty\}$. At each period $t \geq 0$ a large but finite number of short-lived agents, N, enter the economy choose a learning strategy, take an irreversible action, and leave the economy never to come back again.

Action and Type: Let $A = \{a, b\}$ be the set of actions. We assume that agents have heterogenoeus preference over the actions. There are two possible

preference type denoted by $\Omega = \{\omega_1, \omega_2\}$, where $\omega_1 \equiv a \succ b$ and $\omega_2 \equiv b \succ a$ [1].
Belief: We assume that the agent does not know his preference type ω_i. Let $\Gamma \equiv \Delta(\Omega)$ be the set of possible distributions over Ω and $\Delta(\Gamma)$ denote the set of all possible distributions over Γ. At any period $t \geq 0$ agents enter with a common prior $\gamma_0 \in \Delta(\Gamma)$, i.e., a belief over possible type distributions. The agents are thus unaware of the true data genereating process of types in the population and the prior belief assigns probabilities to different possible distributions. Let μ^* denote the true distribution of types in Ω where $\mu^* \in int(\gamma_0)$[2].
Payoff: After entering, the agent can learn about his own type ω_i and chooses an alternative $i \in A$. Let $u : A \times \Omega \to \mathbb{R}$ be the type dependent utility function that maps the preference type of each agent into utilities. For simplicity consider a symmetric utility function for both types,

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

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 action and the type. Assume that agents are Bayesian expected utility maximizers.

2.2 Costly learning

Agents can learn two ways, either by gathering information privately or by observing choices of other agents who entered the economy before them. Since agents are of heterogenous types, social learning is informative about the distribution of types $\mu \in \Gamma$, and private learning is informative about one's own type, $\omega \in \Omega$.

[1]For sake of simplicity, we only consider strict preference rankings, since under indifference the agent does not have any incentive to learn

[2]This is to ensure agents assign positive probabilities on the true distribution, thus it would be possible for the Bayesian agents to learn about the true distribution

2.2.1 Private learning

Signal Structure: Let S denote the set of possible signals that the agent can observe when learning privately. Let $\pi(s,\omega) : S \times \Omega \to [0,1]$ be an information structure that the agent chooses. Given γ_0, $\pi(s,\omega)$ generates a distribution of posterior beliefs $\gamma_\pi \in \Delta(\Gamma)$[3]. Then for any prior belief $\mu \in \Gamma$, the posterior belief for any state ω given signal s by Bayes' Law would be

$$Pr(\omega|s) = \frac{\pi(s|\omega)\mu(\omega)}{\sum_{\omega'} \pi(s|\omega')\mu(\omega')}. \tag{2}$$

Posterior probability of types: If two distinct signals generate the same posterior belief then they are equally Blackwell informative (Blackwell et al. (1953)). Since more signals are weakly more costly it is optimal to choose a unique signal to generate any posterior belief. Hence, the posterior probability of any type is

$$\gamma_\pi(Pr(\omega|s)) = \sum_\omega \pi(s|\omega)\mu(\omega) = P_\pi(s). \tag{3}$$

where $P_\pi(s)$ denote the probabilit of observing signal s under the signal structure π. By Bayes Plausibility (refer Kamenica and Gentzkow (2011), Matějka and McKay (2015))

$$\sum_s P(\omega|s)\gamma_\pi(P(\omega|s)) = \mu(\omega). \tag{4}$$

Choosing any γ is hence equivalent to choosing an information structure$\pi(s|\omega)$ (refer Matějka and McKay (2015),)[4]

Posterior probability of actions: Let $P(i,\omega|\mu)$ be the conditional (posterior) probability of choosing action $i \in A$ when type is $\omega \in \Omega$ and prior $\mu \in \Gamma$ and $P(i|\mu) \equiv \sum_{\omega\in\Omega} \mu(\omega) P(i,\omega|\mu)$ be the unconditional (prior) probability of choosing action $a \in A$. By similar logic as before, suppose the same

[3]Note that since agents enter with a prior belief over $\Delta(\Gamma)$, we get for any $\mu \in supp(\gamma)$, $\pi(s,\omega|\mu) = \pi(s,\omega)$, i.e., the signal structure is independent of the distribution $\mu \in \Gamma$. Moreover, this implies private learning is only informative about Ω and not Γ.

[4]$\pi(s|\omega) = \frac{P(\omega|s)\gamma(P(\omega|s))}{\mu(\omega)}$.

action can be chosen following two distinct posterior beliefs. Then they are equally Blackwell informative (Blackwell et al. (1953)). Thus each posterior would lead to the choice of unique action. Hence choosing a distribution of the conditional probability of actions is equivalent to choosing an information structure.

$$P(a,\omega) = \sum_{\gamma_\pi} \pi(\gamma|\omega) Pr(a|\gamma) = \pi(\gamma|\omega) \tag{5}$$

Cost of Private Learning: The cost of private learning is given by Shannon's relative entropy between the prior and the posterior probability of choice(Cover and Thomas (2012)). The cost function is given by,

$$C(\lambda,\mu) = \lambda \left\{ \underbrace{\sum_{\omega\in\Omega} \mu(\omega) \sum_{a\in A} P(a,\omega|\mu) \ln P(a,\omega|\mu)}_{\text{expected entropy of the posterior distributions}} - \underbrace{\sum_{a\in A} P(a|\mu) \ln P(a|\mu)}_{\text{entropy of the prior distribution}} \right\} \tag{6}$$

Since $P(i|\mu) = \sum_\omega P(i,\omega|\mu)$ we can rewrite the expression as

$$\begin{aligned} C(\lambda,\mu) &= \lambda \sum_{\omega\in\Omega} \mu(\omega) \left\{ \sum_{a\in A} P(a,\omega|\mu) \ln P(a,\omega|\mu) - P(a,\omega|\mu) \ln P(a|\mu) \right\} \\ &= \lambda E_\omega D(P(a|\mu)||P(a,\omega|\mu)) \end{aligned} \tag{7}$$

where $\lambda \in [0,\infty]$ be the marginal cost of private learning and $D(p||q)$ denote the relative entropy between p and q [5] . We assume λ is same for all agent and is common knowledge.

2.2.2 Social learning

Social Learning protocol: ollowing the existing literature on observational learning I assume that agents can only observe the action of their predecssors,

[5]The relative entropy between two distribution $p(x)$ and $q(x)$ is given by, $\sum_x p(x) \ln \frac{p(x)}{q(x)}$.

no other information, e.g., their belief or payoff can be observed. Any agent at any period $t \geq 1$ can observe the action of any agent from $t-1$ generation or before subject to a cost. Once the t^{th} generation agent decides n, he randomly picks n agents from generation $t-1$ or before and observes their chosen actions. For simplicity we assume n is chosen before observing any agent from any previous generation, i.e., they are observed in a block, which means the entire data would be observed together. In the existing literature of social learning this is common learning protocol assumed, however instead of assuming that the agent can observe every other agents from all previous generation we consider agents will endogenously choose to observe n many agents due to the cost of social learning.

Cost of Social learning: FThe cost of social learning function is given by $c : \{1, \ldots, N\} \to \mathbb{R}_+$ and a typical instant would be written as $c(n)$ where n be the number of previous generation agents that an agent in generation t observes. The cost function has the following properties,

$$\begin{aligned} c\left(n\right) &\in (0, \infty), \quad 0 < n \leq N, \\ c\left(n\right) &\leq c\left(n+1\right), \\ c\left(n\right) - c\left(n+1\right) &\leq c\left(n+1\right) - c\left(n\right), \quad 0 \leqslant n \leqslant N-1 \\ c\left(N\right) &> \bar{u} \end{aligned} \tag{8}$$

i.e., $c(n)$ is positive, finite, weakly monotone, weakly convex, and observing everyone is never optimal. We assume that every agent in the economy faces the same cost of social learning.

Belief: Since any agent in this economy can only choose one of two actions, a and b, the information from social learning can be summarized by the number of observations where an agent had chosen action a. Let x_n denote the number of action a chosen by n agents observed by the decision maker. The Bayesian agent, given a belief γ updates her belief to $\gamma_{x_n} \in \Delta(\Gamma)$ upon observing x_n. Here, he accounts for the possible mismatch between type and action chosen by his predecessors.

2.3 Benchmark Case: Only Private Learning

Let us consider the benchmark case where there is no observational information available to the decision maker. Let $\mu_0(\omega) = E_{\gamma_0}(\mu(\omega))$ denotes the expected probability of preference types ω given prior belief γ_0. For example, if γ_0 denotes a uniform distribution over $[0, 1]$, i.e., the completely uninformation prior then $\mu_0 = 0.5$ denote the expected probability of an agent being type ω_1 (or ω_2).

The optimization problem with only private learning is given by,

$$V(A, \mu_0) = \max_{P(i,\omega|\mu_0)} \sum_{\omega \in \Omega} \mu_0(\omega) P(i, \omega|\mu_0) u(i, \omega) - C(\lambda, \mu_0). \tag{9}$$

where the first term on RHS denote the expected payoff from choosing the optimal action and the second term refers to the cost of learning incurred.

For notational simplicity let us denote $z(i, \omega) = \exp(u(i, \omega)/\lambda)$, $\underline{\lambda} = \exp(\underline{u}/\lambda)$, and $\overline{\lambda} = \exp(\overline{u}/\lambda)$. Following Matějka and McKay (2015), the solution to the agent's optimization problem would be

$$P(i, \omega|\mu_0) = \frac{P(i|\mu_0) z(i, \omega)}{\sum_{j \in A} P(j|\mu_0) z(j, \omega)} \quad \forall i \in A, \omega \in \Omega \tag{10}$$

The Bayesian plausibility implies given their prior γ_0,

$$\sum_{\omega \in \Omega} \mu_0(\omega_i) \frac{z(i, \omega)}{\sum_{j \in A} P(j|\gamma) z(j, \omega)} \leq 1 \quad \forall \quad i \in A. \tag{11}$$

The inequality holds with equality if $P(i|\gamma) > 0$. Note that, $P(i|\gamma)$ denote the unconditional probability of choosing action $i \in A$ given prior belief γ. If the prior belief is high enough for one of two possible types, due to the cost of private learning, it can be optimal for the agent to not learn and choose only one action with probability 1. Later we will show in that case learning is not complete.

Using equation 11 for both $a, b \in A$ we get,

$$P(a|\mu_0) = \begin{cases} \frac{\mu_0(\omega_1)\bar{\lambda} - \mu_0(\omega_2)\underline{\lambda}}{\bar{\lambda} - \underline{\lambda}} & \text{if} - \Delta u/\lambda \le \ln \frac{\mu_0(\omega_1)}{\mu_0(\omega_2)} \le \Delta u/\lambda \\ 1 & \text{if } \ln \frac{\mu_0(\omega_1)}{\mu_0(\omega_2)} > \Delta u/\lambda \\ 0 & \text{if } \ln \frac{\mu_0(\omega_1)}{\mu_0(\omega_2)} < -\Delta u/\lambda \end{cases} \tag{12}$$

Thus the posterior probability of choosing actions for different types can be obtained by combining equation 10 and 12.

Before we solve the private learning probelm, let us consider one key property of the relative Shannon entropy. This cost function belongs to a broader class of cost function called the Uniform Posterior Separable (UPS) cost functions. One defining characteristic of UPS cost function is Likelihood Invariant Posterior (LIP)[6]. In a two action-two state symmetric payoff choice problem, like this model, LIP implies there exists a cutoff value of belief over Ω, say μ', such that private learning is only optimal if $\mu \in [1-\mu', \mu']$. For any value of μ where private learning is optimal posterior belief lies in the set $\{1-\mu', \mu'\}$, i.e., irrespective of the prior belief over Ω the posterior belief is always the same (given the same signal realization).

Lemma 1. *The private learning value function $V(\mu)$ is convex in $\left[\frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}}, \frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}\right]$ and linear for $\mu \in [0, \frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}}) \cup (\frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}, 1]$ as shown in figure 1. Moreover, the learning strategy satisfies the Likelihood Invariant Property (LIP).*

Proof. Let us use the notation $p_a = P(a|\mu)$, then the posterior probability of type ω upon observing signal a is given by,

$$\gamma_\pi(i\omega) = \frac{P(i,\omega)\mu(\omega)}{\sum P(i,\omega)\mu(\omega)}, \tag{13}$$

plugging in the values of $P(i, \omega|\mu)$ we get,

$$\gamma_\pi(a, \omega_1) = \frac{p_a \mu \bar{\lambda}(p_a \underline{\lambda} + (1-p_a)\bar{\lambda})}{p_a \mu \bar{\lambda}(p_a \underline{\lambda} + (1-p_a)\bar{\lambda}) + (1-\mu)p_a \underline{\lambda}(p_a \bar{\lambda} + (1-p_a)\underline{\lambda})}$$

[6] Refer Caplin et al (2019) for more details.

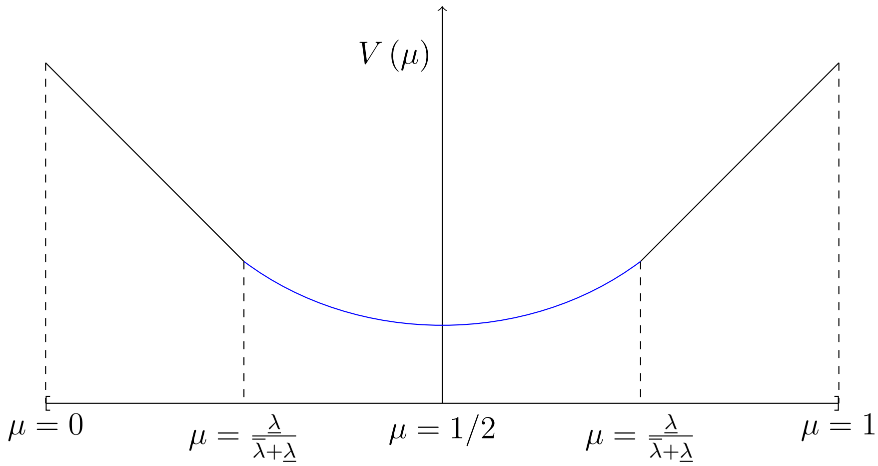

Figure 1: The value function $V(\mu)$

Plugging in the values of p_a we get,

$$\gamma_\pi(a, \omega_1) = \frac{\overline{\lambda}}{\overline{\lambda} + \underline{\lambda}}$$

Similarly, we can show $\gamma_\pi(b, \omega_1) = \frac{\underline{\lambda}}{\overline{\lambda}+\underline{\lambda}}$. Thus $\gamma_\pi(i, \omega)$ is independent of $\mu \in \Gamma$ for all $\mu \in [\frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}}, \frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}}]$, i.e., the learning strategy satisfies LIP.

To find the shape of the value function $V(\mu)$ we plug in the values of the posterior choice probabilities $P(i,\omega|\mu)$ and get,

$$\begin{aligned} V(\mu) = \bar{u}\left[\frac{\mu p_a\bar{\lambda}+(1-\mu)(1-p_a)\bar{\lambda}}{p_a\bar{\lambda}+(1-p_a)\underline{\lambda}}\right] + \underline{u}\left[\frac{(1-\mu)\,p_a\underline{\lambda}+\mu(1-p_1)\underline{\lambda}}{(1-p_a)\,\bar{\lambda}+p_a\underline{\lambda}}\right] \\ -\lambda\left[\mu\left\{\frac{p_a\bar{\lambda}}{p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}}\ln\frac{p_a\bar{\lambda}}{p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}}\right.\right. \\ \left.+\frac{(1-p_a)\,\underline{\lambda}}{p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}}\ln\frac{(1-p_a)\,\underline{\lambda}}{p_a\bar{\lambda}+(1-p_a)\,\underline{\lambda}}\right\} \\ (1-\mu)\left\{\frac{p_a\underline{\lambda}}{(1-p_a)\,\bar{\lambda}+p_a\underline{\lambda}}\log\frac{p_a\underline{\lambda}}{(1-p_a)\,\bar{\lambda}+p_a\underline{\lambda}}\right. \\ \left.+\frac{(1-p_a)\,\bar{\lambda}}{(1-p_a)\,\bar{\lambda}+p_a\underline{\lambda}}\ln\frac{(1-p_a)\,\bar{\lambda}}{(1-p_a)\,\bar{\lambda}+p_a\underline{\lambda}}\right\} \\ \left.-p_a\log p_a-(1-p_a)\ln(1-p_a)\right]. \quad (14) \end{aligned}$$

The first two components denote the benefit (B hereafter) from private learning and the rest of the components together (C hereafter) denote the cost of learning.

Given λ, $V(\mu)$ 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

$$p(a)=\begin{cases}1 & \text{if}\quad \frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}<\mu\leq 1\\ 0 & \text{if}\quad 0\leq\mu\leq\frac{\underline{\lambda}}{\bar{\lambda}+\underline{\lambda}}\end{cases}\Rightarrow V'_\mu=\begin{cases}\Delta u>0 & \text{if}\quad \frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}\leq\mu<1\\ -\Delta u<0 & \text{if}\quad 0<\mu\leq\frac{\underline{\lambda}}{\bar{\lambda}+\underline{\lambda}}.\end{cases}$$

The cutoffs are differentiable in λ, $\frac{d\frac{\bar{\lambda}}{\bar{\lambda}+\underline{\lambda}}}{d\lambda}=-\frac{d\frac{\underline{\lambda}}{\bar{\lambda}+\underline{\lambda}}}{d\lambda}=-\frac{(\bar{u}-\underline{u})}{\lambda^2}\frac{\bar{\lambda}\underline{\lambda}}{\left(\bar{\lambda}+\underline{\lambda}\right)^2}<0$ 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 $\left(\frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}}\frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}\right)$, $p_a\in(0,1)$. Let us rewrite $V(\mu)$ for the region

$\left(\frac{\underline{\lambda}}{\underline{\lambda}+\bar{\lambda}} \frac{\bar{\lambda}}{\underline{\lambda}+\bar{\lambda}}\right)$ by putting the value of $p_a = \frac{\mu\bar{\lambda}-(1-\mu)\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}}$. Thus we get,

$$\begin{aligned} V(\mu) = \frac{\bar{u}\bar{\lambda}+\underline{u}\underline{\lambda}}{\bar{\lambda}+\underline{\lambda}} + \lambda \Bigg[& \frac{\mu\bar{\lambda}-(1-\mu)\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}} \ln \frac{\mu\bar{\lambda}-(1-\mu)\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}} + \frac{(1-\mu)\bar{\lambda}-\mu\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}} \ln \frac{(1-\mu)\bar{\lambda}-\mu\underline{\lambda}}{\bar{\lambda}-\underline{\lambda}} \\ & - \left(\frac{\mu\bar{\lambda}^2-(1-\mu)\bar{\lambda}\underline{\lambda}}{\bar{\lambda}^2-\underline{\lambda}^2} \ln \frac{\mu\bar{\lambda}^2-(1-\mu)\bar{\lambda}\underline{\lambda}}{\bar{\lambda}^2-\underline{\lambda}^2} + \frac{(1-\mu)\bar{\lambda}\underline{\lambda}-\mu\underline{\lambda}^2}{\bar{\lambda}^2-\underline{\lambda}^2} \ln \frac{(1-\mu)\bar{\lambda}\underline{\lambda}-\mu\underline{\lambda}^2}{\bar{\lambda}^2-\underline{\lambda}^2} \right) \\ & - \left(\frac{\mu\bar{\lambda}\underline{\lambda}-(1-\mu)\underline{\lambda}^2}{\bar{\lambda}^2-\underline{\lambda}^2} \ln \frac{\mu\bar{\lambda}\underline{\lambda}-(1-\mu)\underline{\lambda}^2}{\bar{\lambda}^2-\underline{\lambda}^2} + \frac{(1-\mu)\bar{\lambda}^2-\mu\bar{\lambda}\underline{\lambda}}{\bar{\lambda}^2-\underline{\lambda}^2} \ln \frac{(1-\mu)\bar{\lambda}^2-\mu\bar{\lambda}\underline{\lambda}}{\bar{\lambda}^2-\underline{\lambda}^2} \right) \Bigg] \end{aligned}$$

Simplifying further we get,

$$\begin{aligned} V(\mu) &= \underbrace{\frac{\bar{u}\bar{\lambda}+\underline{u}\underline{\lambda}}{\bar{\lambda}+\underline{\lambda}}}_{B(\mu)} + \underbrace{\lambda \left[\ln(\bar{\lambda}+\underline{\lambda}) - \frac{\bar{u}\bar{\lambda}}{\lambda(\bar{\lambda}+\underline{\lambda})} - \frac{\underline{u}\underline{\lambda}}{\lambda(\bar{\lambda}+\underline{\lambda})} + \mu\ln\mu + (1-\mu)\ln(1-\mu) \right]}_{-C(\mu)} \\ &= \lambda \left[\ln(\bar{\lambda}+\underline{\lambda}) + \mu\ln\mu + (1-\mu)\ln(1-\mu) \right] \end{aligned}$$

This implies,

$$V'(\mu) = \lambda \ln \frac{\mu}{1-\mu} \begin{cases} \geq 0 & \text{if} \quad \mu \geq 0.5 \\ < 0 & \text{if} \quad \mu < 0.5 \end{cases}$$

Since the value function is symmetric in μ around $\mu = 0.5$ consider only $\mu \geq 1/2$ region. Since $V'(\prime) = 0$ only for $\mu = 1/2$ and $V''(\mu) = \lambda[\frac{1}{\mu} + \frac{1}{1-\mu}] > 0$, given any λ, V attains global minima at $\mu = 0.5$.

Similarly, differentiating $V(\mu)$ w.r.t. λ we get,

$$V'_\lambda = \ln(\bar{\lambda}+\underline{\lambda}) + \mu\ln\mu + (1-\mu)\ln(1-\mu) - \frac{\bar{u}\bar{\lambda}+\underline{u}\underline{\lambda}}{\lambda(\bar{\lambda}+\underline{\lambda})}$$

Thus $V'_\lambda < 0$ if and only if $V(\mu) < B(\mu)$, which is true for every $\lambda > 0$.

Furthermore,

$$\frac{\partial V'_\mu}{\partial \lambda} = \ln \frac{\mu}{(1-\mu)} \begin{cases} \geq 0 & \text{if} \quad \mu \geq 0.5 \\ < 0 & \text{if} \quad \mu < 0.5 \end{cases}$$

Figure 1 illustrates the shape of the value function. The function is concave in the blue region and linear outside. Moreover if the prior belief lies in the blue region, by LIP the posterior belief about state ω_1 will lie in either end points of the blue region. □

Thus under any prior belief μ we find that agents do not always learn perfectly about their types. Let $\epsilon^a(\mu) = P\left(a, \omega_2|\mu\right)$ and $\epsilon^b(\mu) = P\left(b, \omega_1|\mu\right)$ be the corresponding mismatch probabilities when choosing a and b type ω_2 and type ω_1 agents respectively. Using lemma 1 we can write the mismatch probabilities as,

$$\epsilon^a(\mu) = P(a, \omega_2|\mu) = \frac{\underline{\lambda}(\mu\overline{\lambda} - (1-\mu)\underline{\lambda})}{(1-\mu)(\overline{\lambda}^2 - \underline{\lambda}^2)} \tag{15}$$

$$\epsilon^b(\mu) = P(b, \omega_1|\mu) = \frac{\overline{\lambda}((1-\mu)\overline{\lambda} - \mu\underline{\lambda})}{\mu(\overline{\lambda}^2 - \underline{\lambda}^2)} \tag{16}$$

Thus for all agents with the same belief μ, since the cost of learning is same and all agents are Bayesian expected utility maximizers, they will have the same mismatch probability. Provided μ is known, any Bayesian agent would be know the mismatch probabilities of any agent they observe.

2.4 Social learning and Order of Beliefs

Suppose an agent i at time t observes n agents from generations $t-1$ or before, then he will update his belief over $\Delta\left(\Gamma\right)$ via Bayes rule. 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 $\mu \in supp(\gamma_0)$ is,

$$P(\mu|\gamma, x_n) = \frac{P(x_n|\mu)\, P(\mu|\gamma)}{\int_{\nu \in supp(\gamma)} P(x_n|\nu)\, P(\nu|\gamma)} \tag{17}$$

and zero otherwise. Here $P(\mu|\gamma)$ and $P(\nu|\gamma)$ are given by the prior belief of the agent, however, to calculate $P(x_n|\mu)$ he needs to consider the choices of agents in previous generation. Specifically, the agent needs to know the posterior choice probabilities of the earlier generation. The posterior choice probabilities depend on both the social and private learning. Given both types of learning technologies are common knowledge, any agent in t generation can infer how many agents have been observed by $s = 1, \ldots t-1$ generation but no other information, i.e., who they observed, what was the action chosen by these agents etc. Thus from the point of view of generation t agent any agent from earlier generations had undertaken private learning with a prior belief given by $E(\mu|\gamma_0)$. Thus all agents from any previous generation is considered as identical.

Given this prior belief let us consider the probability that any $t \geq 1$ would observe x_n many action a out of n observations. Before calculating this, let us consider a simple example where the agent observes 3 actions from earlier generation and 2 of them are a's. Since the private learning technology is common knowledge the period t agent can infer with what likelihood the previous generation agents chose a when indeed they are of type ω_1 and when an ω_2 type agent mistakely chosen a. The following represents the various possibilities of agents in previous generation mismatching their types and actions and the corresposnding probabilities, assuming μ is the prior probability of being type ω_1.

- **No Mistakes**: Both agents choosing a are of type ω_1 and the agent choosing b are of type ω_2

$$Prob = 3\mu^2(1-\mu)(1-\epsilon_b)^2(1-\epsilon_a)$$

- **One Mistake**:

- Only one agent choosing a is of type ω_1, the other two agents are of type ω_2

$$Prob = 3\mu(1-\mu)^2(1-\epsilon_b)\epsilon_a(1-\epsilon_a)$$

- All agents are of type ω_1

$$Prob = \mu^3(1-\epsilon_b)^2\epsilon_b$$

- **Two Mistakes**:
  - Only one agent choosing a is of type ω_1 but the agent choosing b is also of type ω_1

$$Prob = 3\mu^2(1-\mu)(1-\epsilon_b)\epsilon_a\epsilon_b$$

  - All agents are of type ω_2

$$Prob = (1-\mu)^3\epsilon_a^2(1-\epsilon_b)$$

- **Three Mistakes**: Every agent mismatches action and type

$$Prob = 3\mu(1-\mu)^2\epsilon_a^2\epsilon_a$$

Genralizining this over any n and x_n we get for any $t \geq 1$ agent, the probability of observing x_n, given prior μ would be,

$$P(x_n|\mu) = \sum_{k=0}^{n}\sum_{j=k^*}^{k^{**}} \binom{n}{x_n - 2j + k} \mu^{x_n-2j+k} (\epsilon_a)^j (1-\epsilon_b)^{x_n-j} (1-\mu)^{n-x_n-k+2j} (\epsilon_b)^{k-j} (1-\epsilon_a)^{n-x_n-k+j} \quad (18)$$

where, k denote the total number of possible mismatches, either ω_1 choosing b or ω_2 choosing a and j denote the number of ϵ_a type mismatch. Thus $x_n - 2j + k$ denote the number of true type a agents.

The bounds for j are given as follows,

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

and

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

where both expressions consider the possibility that the total number of mismatch can be so large, everyone who chooses a (or b) might be mismatching. Plugging the value obtained from equation 18 into equation 17 we can calculate $P(\mu|\gamma, x_n)$ for every $\mu \in \gamma$, and can update the belief to γ_{x_n}, where $\gamma_{x_n} \in \Delta(\Gamma)$ denote the interim belief over Ω after observing x many agents choosing a out of n randomly observed agents. .

2.5 Optimal Learning Protocol

All agents in period $t = 0$ can obly gather information privately. For any $t \geq 1$ period agent there are two different choices for learning, namely social and private learning. By social learning the agent observes actions of previous generations and update their belief over $\Delta(\Omega)$ and by private learning they update their beilef over own type $\omega \in \Omega$. The following lemma shows, that optimal sequencing would always be of the form: *first social learning then private learning.*

Lemma 2. *Any agent in period $t \geqslant 1$ would optimally choose to learn by observing others first followed by learning privately.*

Proof. Suppose not. Consider an agent with belief $\gamma_1 \in \Gamma$ such that private learning is optimal at every $\mu \in supp(\gamma_1)$. WLOG, let us assume $E(\mu|\gamma_1) > 0.5$, in expectation the prior belief was biased towards state ω_1. Let us consider strategy 1:

- given γ_1 as the prior belief learn privately by choosing a signal structure $\pi_1(s, \omega)$
- observe n agents and update belief to $\tilde{\gamma} \in \Gamma$ (n is chosen optimally)
- given $\tilde{\gamma}$ learn privately if need be

I will show the agent can be made better off by choosing an alternate strategy, call it strategy 2:

- observe n agents and update belief to $\tilde{\gamma} \in \Gamma$ (n is chosen optimally)
- given $\tilde{\gamma}$ learn privately if need be

Since private learning is only informative about Ω, in both the two cases the optimal choice of n would be the same. Choice of n and subsequent observations changes belief over $\gamma \in \Delta(\Omega)$ whereas $\pi(s, \omega)$ changes belief over $\omega \in \Omega$. Upon observing n agents and changing his belief to $\tilde{\gamma}$ the agent will adjust his belief over Ω based on the signal relatization using $\pi_1(s, \omega)$. Let us assume WLOG $E(\mu|\tilde{\gamma}) \geq 0.5$

Moreover, since all agents are Bayesian the sequence of learning does not affect belief. This implies unless $\gamma_1 = \tilde{\gamma}$, at the end of step 2 in strategy 1 the agent's belief would not coincide with the belief under strategy 2.

This generates two possibilities, either $E(\mu|\gamma_1) > E(\mu|\tilde{\gamma})$ or $E(\mu|\gamma_1) < E(\mu|\tilde{\gamma})$ [7]. If $E(\mu|\gamma_1) > E(\mu|\tilde{\gamma})$, with $\pi_1(s, \omega)$, the agent's posterior belief would be in $(\frac{\underline{\lambda}}{\overline{\lambda}+\underline{\lambda}}, \frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}})$ and it will be optimal for his to learn. However, with $E(\mu|\gamma_1) < E(\mu|\tilde{\gamma})$ no further learning would be needed.

Under the second strategy of first social learning and then private learning, in case $E(\mu|\gamma_1) > E(\mu|\tilde{\gamma})$, the total cost incurred would be identical to that of the strategy 1, since the agent would choose a signal structure $\pi_2(s, \omega)$ such that his updated belief about ω is $\left\{\frac{\underline{\lambda}}{\overline{\lambda}+\underline{\lambda}}, \frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}}\right\}$. However, if $E(\mu|\gamma_1) < E(\mu|\tilde{\gamma})$

[7] In case of equality both strategies would be equally costly.

then the agent only incurs

$$\begin{aligned}C_2 &= \lambda E_\omega D(P\left(a,\omega|E(\mu|\tilde{\gamma})\right)||P\left(a|E(\mu|\tilde{\gamma})\right)) \\ &< \lambda E_\omega D(P\left(a,\omega|E(\mu|\gamma_1)\right)||P\left(a|E(\mu|\gamma_1)\right)) = C_1.\end{aligned}$$

Thus in expectation the agent can save some amount of cost of private learning under the second strategy making it a strictly better strategy for him. Hence, proved. □

The main intuition is as follows: social learning is informative about the distribution of types and private learning is informative about the idiosyncratic type. Learning socially first changes the belief over the data generating process of types or type distribution which subsequently changes the prior belief over DM's idiosyncratic type. Since agents are Bayesian and order of updating doesn't affect posterior belief, this cannot increase the cost of learning vis-a-vis a strategy when some private learning is done before social learning. But after some social learning is done there is a positive probability the DM would choose to learn less privately, in that case, if the private learning was undertaken before social learning it gives rise to a sunk cost that could have been avoided. Thus it is weakly better to learn socially first, privately later.

2.6 Value Function

Since agents in period $t = 0$ does not have the option to observe other agents' actions, their value function only consists of private learning and is given by equation 9.

Given lemma 2, we know agents first learn socially then with the updated belief γ'_{x_n} they learn privately. Following equation 10, the optimal private learning of an agent in any period $t \geq 1$ would be the same as a $t = 0$ agent, except with a different interim belief over Γ. Thus for any generation $t \geq 1$ agents, the value function would be,

$$W\left(A,\gamma\right) = \max_n \sum_{\mu \in supp(\gamma_{x_n})} V\left(A,\mu\right)\gamma_{x_n}(\mu) - c\left(n\right) \tag{19}$$

where $V(.)$ denotes the net expected payoff following private learning and the agent optimally chooses n to maximize the V net of the cost of social learning. Furthermore, the agent with belief γ_{x_n} chooses to learn privately only if, $-\Delta u/\lambda \leq \ln \frac{\mu_{x_n}(\omega_1)}{\mu_{x_n}(\omega_2)} \leq \Delta u/\lambda$ where $\mu_{x_n}(\omega) = \sum_{\mu \in supp(\gamma_{x_n})} \gamma_{x_n}(\mu)\mu(\omega)$. For any other value of γ_{x_n}, he would choose one action for sure.

3 Results

3.1 Optimal Learning Strategy

The following lemma characterizes the relationship between cost of private learning and the informativeness of social learning. Since agents are Bayesian, they correctly predict that as the cost of private learning goes up the observed action of their predecessors are less likely to match their type. Thus the updated belief post observing the same signal cannot change belief by much.

Lemma 3. *As the marginal cost of private learning λ increases, the informativeness of social learning decreases.*

Proof. Since the DM's objective is to match his type with his chosen action, the relevant statistic for him is the $E_\gamma(\mu) = \int_0^1 \mu f(\mu) d\mu$. Let $\mu_a = E_\gamma(\mu|a, \mu_0)$ and $\mu_b = E_\gamma(\mu|b, \mu_0)$ denote the updated belief after observing action a and b respectively given prior belief μ_0. We want to show $|E_\gamma(\mu|i, \mu_0) - \mu_0|$ is decreasing in λ. This implies we need to show,

$$\frac{\partial \mu_a}{\partial \lambda} \leq 0 \quad \text{and} \quad \frac{\partial \mu_b}{\partial \lambda} \geq 0 \quad \forall \lambda \in (0, \lambda^{**})$$

Given equation 17, we can write,

$$\mu_a = \int_0^1 \frac{\mu(1-\epsilon_a-\epsilon_b)+\epsilon_a}{\mu_0(1-\epsilon_a-\epsilon_b)+\epsilon_a} \mu f(\mu) d\mu$$
$$\mu_b = \int_0^1 \frac{(1-\epsilon_a)-\mu(1-\epsilon_a-\epsilon_b)}{(1-\epsilon_a)-\mu_0(1-\epsilon_a-\epsilon_b)} \mu f(\mu) d\mu$$

Thus,

$$\frac{d\mu_a}{d\lambda} = -\frac{(\int_0^1 \mu^2 f(\mu) d\mu - \mu_0^2)(\epsilon_a \frac{d\epsilon_b}{d\lambda} + (1-\epsilon_b)\frac{d\epsilon_a}{d\lambda})}{(\mu_0(1-\epsilon_a-\epsilon_b)+\epsilon_a)^2} \leq 0$$
$$\Rightarrow \epsilon_a \frac{d\epsilon_b}{d\lambda} + (1-\epsilon_b)\frac{d\epsilon_a}{d\lambda} \geq 0$$

Plugging in the values of ϵ_i we get,

$$(\mu_0\bar{\lambda}^2 - (1-\mu_0)\bar{\lambda}\underline{\lambda})\Big[2(\bar{u}\bar{\lambda}^2 - \underline{u}\underline{\lambda}^2)(\mu_0\bar{\lambda}\underline{\lambda} - (1-\mu_0)\underline{\lambda}^2) - (\bar{\lambda}^2 - \underline{\lambda}^2)(\mu_0(\bar{u}+\underline{u})\bar{\lambda}\underline{\lambda} - 2(1-\mu_0)\underline{u}$$
$$+(\mu_0\bar{\lambda}\underline{\lambda} - (1-\mu_0)\underline{\lambda}^2)(2(\bar{u}\bar{\lambda}^2 - \underline{u}\underline{\lambda}^2)((1-\mu_0)\bar{\lambda}\underline{\lambda} - \mu_0\underline{\lambda}^2) - (\bar{\lambda}^2 - \underline{\lambda}^2)((1-\mu_0)(\bar{u}+\underline{u})\bar{\lambda}\underline{\lambda} - 2\mu_0\underline{u}\underline{\lambda}^2))$$

Simplifying we get,

$$(\mu_0\bar{\lambda} - (1-\mu_0)\underline{\lambda})^2 > 0$$

which would be true for all values of $\lambda \geq 0$. Similarly for μ_b we find,

$$\frac{d\mu_b}{d\lambda} = \frac{(\int_0^1 \mu^2 f(\mu) d\mu - \mu_0^2)((1-\epsilon_a)\frac{d\epsilon_b}{d\lambda} + \epsilon_b\frac{d\epsilon_a}{d\lambda})}{(1-\epsilon_a-\mu_0(1-\epsilon_a-\epsilon_b))^2} \geq 0$$
$$\Rightarrow (1-\epsilon_a)\frac{d\epsilon_b}{d\lambda} + \epsilon_b \frac{d\epsilon_a}{d\lambda} \geq 0$$

Plugging in the values of ϵ_i we get,

$$((1-\mu_0)\bar{\lambda}^2 - \mu_0\bar{\lambda}\underline{\lambda})(2(\bar{u}\bar{\lambda}^2 - \underline{u}\underline{\lambda}^2)((1-\mu_0)\bar{\lambda}\underline{\lambda} - \mu_0\underline{\lambda}^2) - (\bar{\lambda}^2 - \underline{\lambda}^2)((1-\mu_0)(\bar{u}+\underline{u})\bar{\lambda}\underline{\lambda} - 2\mu_0\underline{u}$$
$$+((1-\mu_0)\bar{\lambda}\underline{\lambda} - \mu_0\underline{\lambda}^2)(2(\bar{u}\bar{\lambda}^2 - \underline{u}\underline{\lambda}^2)(\mu_0\bar{\lambda}\underline{\lambda} - (1-\mu_0)\underline{\lambda}^2) - (\bar{\lambda}^2 - \underline{\lambda}^2)(\mu_0(\bar{u}+\underline{u})\bar{\lambda}\underline{\lambda} - 2(1-\mu_0)\underline{u}\underline{\lambda}^2))$$

Simplifying we get,

$$((1-\mu_0)\bar{\lambda} - \mu_0)\underline{\lambda})^2 > 0$$

which would also be true for all values of $\lambda \geq 0$. □

The following theorem characterizes the relationship between optimal private and social learning obtained from solving the optimization problem in equation 19.

Theorem 1. *Given the social learning cost function in 8 and the prior* γ_0*, there exist* $0 < \lambda^* < \lambda^{**} < \infty$*, such that*

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

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

3. *For all* $\lambda > \lambda^{**}$*, not learning is optimal.*

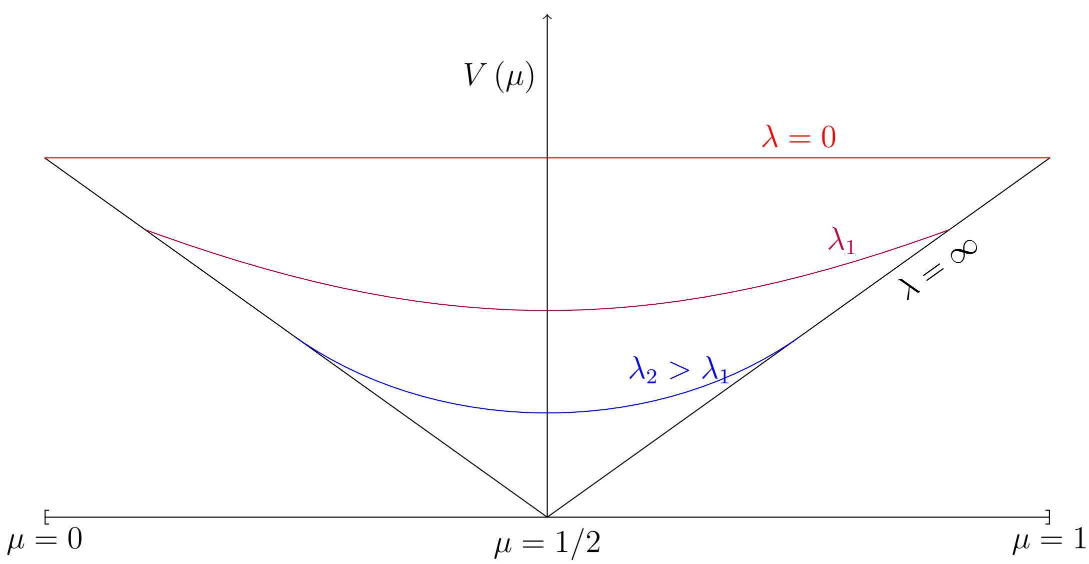

Figure 2: The value function $V(\mu)$ for different λs

Before proving the theorem formally, let me discuss the main intuition behind the proof using the shape of the interim value function following social learning. The proof is constructed in five steps. In the first step, I show that for a high enough private cost of learning, all forms of learning become uninformative since early generations make decisions based on the prior belief alone. Thus all the following generations would choose no social learning as well, stopping all forms of learning in the economy.

In step two, I establish the shape of the interim value function (refer to figure 2) and show how the shape of the value function changes with the value of λ. In the limiting case of $\lambda = 0$, the agent learns perfectly and the value function is given by a horizontal line, and when $\lambda \to \infty$ the value function reflects no private learning which leads to choosing an action based on prior only, where value function becomes a linear function of the interim belief μ.

In step three, I show that social learning shifts the interim belief μ of the agent. I divide the value function over three different regions of interim belief μ (when $\mu \geq 1/2$, since the other side is symmetric). Then I establish the agent will always choose social learning such that his interim belief lies in the first of the third region where the value function is increasing in μ. This creates a jump in the optimal choice of social learning.

Finally, steps four and five explore the impact of different λ on the learning strategy. The value of λ puts a bound on the maximum possible shift away from the prior belief due to social learning. This uses the intuition that agents are Bayesian, and they update that for high level λ observing the action of pervious period agent does not reflect much of private learning by the previous period agents. Thus an increase in λ has two impacts, one, it makes private learning relatively more expensive and second, it also makes social learning relatively uninformative, i.e., social learning cannot move the interim belief much. When the cost of private learning is sufficiently low, the interim value function is flatter and the effect on loss of informativeness is relatively small since everyone in the economy already chooses a high level of private learning. This makes the two types of learning substitute for each other for low levels of private cost of learning. As the cost of private learning increases the substitutability component becomes relatively smaller, making the two types of learning complementary.

Proof. Given lemma 2, we can solve the optimization problem backward. First, for any intermediate belief $\mu \in supp(\gamma)$ the optimal private learning generate

$V(\mu)$, then given $V(\mu)$, n^* is chosen to maximize $W(\gamma)$.

Step 1: No learning above λ^{**} By lemma 1, private learning is not optimal if $\mu \in [0, \frac{\underline{\lambda}}{\overline{\lambda}+\underline{\lambda}}) \cup [\frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}}, 1]$. Thus social learning is not informative, because agents choose according to their common prior. Consider $\lambda^{**} = \max\left\{-\ln\frac{\mu}{1-\mu}/\Delta, \ln\frac{\mu}{1-\mu}/\Delta u\right\}$, then no learning is optimal for $\lambda > \lambda^{**}$.

Step 2: Interaction between private and social learning There is a relationship between the level of precision of private and social learning. The parameter λ, the marginal cost of private learning function indirectly captures the level of precision in private learning. For $\lambda \to 0$, the marginal cost of private learning being too low, the DM makes almost no mistake and can effectively match his action with state. However, for $\lambda \to \infty$ the opposite happens. To capture the level of precision in social learning, let us consider the expression $\epsilon_a + \epsilon_b$, i.e., the total probabilty of action and state mismatch.

We can show,

$$\frac{\partial \epsilon_a + \epsilon_b}{\partial \lambda} > 0$$

$$\Rightarrow \frac{2\bar{u}}{\lambda}\exp\frac{2\bar{u}}{\lambda} - \frac{2\underline{u}}{\lambda}\exp\frac{2\underline{u}}{\lambda} > \frac{\bar{u}+\underline{u}}{\lambda^2}(\exp\frac{2\bar{u}}{\lambda} - \exp\frac{2\underline{u}}{\lambda})$$

$$\Rightarrow \frac{\bar{u}+\underline{u}}{\lambda^2}\exp\frac{2\bar{u}}{\lambda} > -\frac{\bar{u}-\underline{u}}{\lambda^2}\exp\frac{2\underline{u}}{\lambda}$$

For any $\bar{u} > \underline{u}$ and $\lambda > 0$ the last inequality always holds. This implies the higher the marginal private cost of learning the higher would be probability of mismatch in observed action of predecessor. We claim this will create a possible upper bound on the level of social learning given any value of λ.

To illustrate further, let us consider the extreme example, where the DM has observed n actions all of which are action a. For the purpose of illustration let us consider the prior belief over $\Delta(\Omega)$ is uniform. Then the posterior belief over $\Delta(\Omega)$ given this observation would be,

$$Pr(\mu|x_n = n) = \frac{((\mu(1-\epsilon_a-\epsilon_b)+\epsilon_a)^n)(1-\epsilon_a-\epsilon_b)(n+1)}{(1-\epsilon_b)^{n+1} - ((2\mu_0-1)(1-\epsilon_a-\epsilon_b)-\epsilon_a)^{n+1}}$$

This implies the expected value would be,

$$E(\mu|x_n = n) = \frac{(1-\epsilon_a-\epsilon_b)^{n+1}(n+1)}{(1-\epsilon_b)^{n+1} - ((2\mu_0-1)(1-\epsilon_a-\epsilon_b)-\epsilon_a)^{n+1}} \sum_{k=0}^{n} \binom{n}{k} (\frac{\epsilon_a}{1-\epsilon_a-\epsilon_b})^{n-k} \frac{1-(2\mu_0-1)^{k+2}}{k+2}$$

If $\bar{u} > \underline{u}$, then we get $E(\mu|x_n = n)$ decreases with λ. Note that, under social learning protocol described here, $E(\mu|x_{n=n})$ becomes the interim belief of the DM over Ω at the time of private learning. Thus a decrease in this value would lead to a less precise *prior* for the DM while undertaking the private learning.

The main point discussed in the above example carries through in other cases as long as $\epsilon_a + \epsilon_b$ increases in λ. Given the shape of the value function $V(\mu)$, this creates a negative relationship between the cost of private learning and the benefit obtained from social learning.

Step 3: Substitute and Complement To explore the substituatibility of complementarity of the private and social learning, let us first explain the role of change in λ on decision making. WLOG let us consider $\mu_0 \geq 0.5$, the analysis would be symmetric for $\mu_0 < 0.5$. In step 1, we have established that $V'_\lambda < 0$ for all $\lambda > 0$. Furthermore,

$$\frac{\partial^2 V}{\partial \lambda^2} = \frac{(\bar{\lambda}+\underline{\lambda})(\bar{u}^2\bar{\lambda}+\underline{u}^2\underline{\lambda}) + (\bar{u}\bar{\lambda}+\underline{u}\underline{\lambda})(\lambda(\bar{\lambda}+\underline{\lambda}) - \bar{u}\bar{\lambda} - \underline{u}\underline{\lambda})}{\lambda^3(\bar{\lambda}+\underline{\lambda})^2} - \frac{\bar{u}\bar{\lambda}+\underline{u}\lambda}{\lambda^2(\bar{\lambda}+\underline{\lambda})} > 0$$

$$\Leftrightarrow \bar{\lambda}\underline{\lambda}(\bar{u}-\underline{u})^2 > 0$$

Thus $V'_{\lambda\lambda} > 0$ for all $\lambda > 0$. This implies the fall in $V(\lambda)$ lowers as λ increases.

The decrease in $V(\lambda)$ due to an increase in λ can be compensated by an increase in intermediate belief μ(WLOG we have assumed $\mu_0 \geq 0.5$). This is because,

$$V'_\mu = \lambda \ln \frac{\mu}{1-\mu} \geq 0$$

for the range of $\mu \geq 0$. Also, $V'_{\mu\lambda} = \ln \frac{\mu}{1-\mu} > 0$ for the relevant range. Note that this is increasing in μ. Thus if social learning can change the intermediate belief significantly, it will increase the net benefit from private learning.

Before private learning is undertaken the DM can change intermediate belief by social learning. In step 3 we have shown $E(\mu|x_n = n)$ increases in n. Furthermore $E(\mu|x_{n=n})_{\lambda,n} < 0$, i.e., as λ increases the change in intermediate belief due an increase in n when $x_n = n$ is smaller. This is true for observation of x_n that strengthens the evidence for either state. However, for the value of x_n that weakens evidence against both, i.e., make the intermediate belief more diffused that is necessarily not the case. As a result in expectation $E_n V(\mu)$ increases more with an increase in n for lower values of λ.

Combining the three effects together we can conclude as λ increases from 0, there is a sharp decline in net value from private learning $V(\mu)$, whereas given λ sufficiently small, the same change in n can increase $E(\mu|x_n)$ more significantly. Furthermore, for any given level of n if social learning changes by Δn,

$$\lim_{\lambda \to 0} \frac{\Delta EV(\mu)}{\Delta n} \to \infty$$

Moreover, since $V'(\lambda\lambda) > 0$, as λ increases the loss in net value of private learning is lower. Also, a higher λ implies $E(\mu|x_n)$ cannot change much, thus restricting the benefit from social learning. Given step 1, as $\lambda \to \lambda^{**}$, the cost of increase in social learning would start to dominate the gain. In the extreme case, for any given n and change of social learning Δn

$$\lim_{\lambda \to \lambda^{**}} \frac{\Delta EV(\mu)}{\Delta n} \to 0$$

Since $c(n) \in (0, \infty)$ and $V(\mu)$ is continuously differentiable in λ and $V'_\lambda < 0$, by intermediate value theorem, for any n there exists a $\lambda^*(n) \in (0, \lambda^{**})$ such that

$$\frac{\Delta W(\lambda, n)}{\Delta n} \begin{cases} = 0 & \text{for} \quad \lambda = \lambda^*(n) \\ > 0 & \text{for} \quad \lambda < \lambda^*(n) \\ < 0 & \text{for} \quad \lambda > \lambda^*(n) \end{cases}$$

This implies as λ increases from 0, the optimal level of social learning would be non-decreasing. As n increases (or remains the same), a further increase in λ would reduce the benefit from social learning. Moreover, given $c(n)$ is weakly monotone and weakly convex in n, $\lambda^*(n)$ would be decreasing in n. Thus there exists a $\lambda^* < \lambda^{**}$ for some n, beyond which an increase in n with an increase in λ would not be profitable and n would be non-increasing in λ for all $\lambda > \lambda^*$.

This implies the two types of learning would be substitutes for $\lambda < \lambda^*$ and complements for $\lambda > \lambda^*$. This proves part 1 and 2 of the theorem. □

3.2 Herding

In previous literature (e.g., Banerjee 1992) agents were choosing action sequentially and later agents could observe actions of all the previous agents. In this case, herding was defined as an equilibrium where agents ignore their private signal and choose an action a if all previous agents have chosen action a. In our framework also agents enter sequentially and can observe the action choice of previous agents. However, in each generation there are multiple agents and future generations typically observe only a fraction of them, subject to cost of social learning. For the purpose of comparison we modify the notion of herding in our model.

Definition 1. *An equilibrium in this economy is considered is considered a Herding equilibrium if after all $t = 0$ agents choose the same action a_i, every agent from all future generations $t \geq 1$ would choose action a_i following social learning with $n^* > 0$ and no further private learning will be undertaken.*

Since the *if* condition involves all $t = 0$ agents choose the same action, it would be easiest for herding to take place. Since all agents in $t = 1$ will observe action a_i is being chosen. As discussed earlier this will move the interim belief to the furthest.

Theorem 2. *Given the social learning cost function $c(n)$ is sufficiently small, $\exists \underline{\lambda}_H, \overline{\lambda}_H$ such that herding would be an equilibrium if and only if $\lambda \in [\underline{\lambda}_H, \overline{\lambda}_H]$. Moreover, $\underline{\lambda}_H \in (0, \overline{\lambda}_H]$ and $\overline{\lambda}_H \in [\lambda^*, \lambda^{**})$.*

Proof. We will prove this theorem in two steps. In the first step we will show for any parameter values, a necessary and sufficient condition for herding is given by $\mu_n \equiv E(\mu | x_n = n) \in R_2$. The second step will find conditions on λ such that $\mu_{n^*(\lambda)} \in R_2$.

Step 1: WLOG let us assume all agents in $t = 0$ choose action a. This would imply for any agent in $t = 1$ generation the observation due to social learning would be given by $x_{n^*(\lambda)=n^*}$, where n^* denote the optimal level of n chosen for any given λ. If all these agents also choose according to the social learning observation, namely action a then for all agents in period t at optimal $n^*(\lambda)$ the observation would also be given by $x_{n^*(\lambda)} = n^*$. Using similar logic, for every generation the optimal choice of social learning would generate $x_{n^*(\lambda)} = n^*$.

Given this observation let us consider the expected posterior belief after observing $x_{n^*(\lambda)} = n^*$, which we will denote as μ_n (slightly abusing the notation). Let us consider a set of parameters such that $\mu_n \in R_2$. By definition $\mu_n \geq \frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}}$. Since $p_a = 1$ in this case, the agent will optimally decide not to undertake any private learning and choose action a.

Now let us consider a set of parameters where $\mu_n \in R_1$. In this region it is optimal for agents in period 1 to undertake private learning. Since for any $\lambda \in (0, \infty)$, $\epsilon_b > 0$ for $\mu_n \in R_1$ with positive probability some agents will choose b in period $t = 1$. This would imply for any agent in $t = 2$ the expected posterior belief would also remain in region 1, triggering private learning. By similar logic, with strictly positive probability there will always be some agents in any generation $t \geq 1$ who will choose action b. Thus $\mu_n \in R_2$ is a necessary and sufficient condition for herding.

Step 2: To find the range of λ given social cost function $c(n)$ let us write the condition $\mu_n \in R_2$ in terms of model parameters.

$$\frac{(1-\epsilon_a-\epsilon_b)^{n+1}(n+1)}{(1-\epsilon_b)^{n+1}-((2\mu_0-1)(1-\epsilon_a-\epsilon_b)-\epsilon_a)^{n+1}}\sum_{k=0}^{n}\binom{n}{k}(\frac{\epsilon_a}{1-\epsilon_a-\epsilon_b})^{n-k}\frac{1-(2\mu_0-1)^{k+2}}{k+2} \geq \frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}}$$

The shape of the RHS of the inequality is given as follows,

$$\frac{\partial \frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}}}{\partial \lambda} = -\frac{\overline{\lambda}\underline{\lambda}(\overline{u}-\underline{u})}{\lambda^2(\overline{\lambda}+\underline{\lambda})^2} < 0$$

i.e., RHS in decresasing in λ. Moreover,

$$\frac{\partial^2 \frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}}}{\partial \lambda^2} = \frac{\overline{\lambda}\underline{\lambda}(\overline{u}-\underline{u})((\overline{u}+\underline{u})(\overline{\lambda}+\underline{\lambda}))-2(\overline{\lambda}\overline{u}+\underline{\lambda}\underline{u}-\lambda(\overline{\lambda}+\underline{\lambda}))}{\lambda^4(\overline{\lambda}+\underline{\lambda})^3}$$

The sign of the above expression is given by,

$$\frac{\partial^2 \frac{\overline{\lambda}}{\overline{\lambda}+\underline{\lambda}}}{\partial \lambda^2} \leq 0$$
$$\Rightarrow (\overline{u}+\underline{u}+2\lambda)(\overline{\lambda}+\underline{\lambda}) \leq 2(\overline{\lambda\overline{u}}+\underline{\lambda u})$$

which happens for sufficiently small λ given $\overline{u}$ and $\underline{u}$. This implies the RHS takes an inverted S-shape and is downward sloping everywhere.

Similarly we can consider the shape of LHS as well. As mentioned in step 3 of the proof of theorem 1 the LHs isdecrasing in λ. Furthermore, LHS also assumes inverted S-shape where the curvature depends on n. For amy feasible n, as $\lambda \to 0$, $LHS < RHS$. Also, at $\lambda = \lambda^{**}$, since $\epsilon_a = 1$ and $\epsilon_b = 0$,

$$\begin{aligned} P(\mu|x_n = n) &= \frac{P(x_n = n|\mu)f(\mu)}{\int_0^1 P(x_n = n)|\nu)f(\nu)d\nu} \\ &= \frac{(\mu(1-\epsilon_b)+(1-\mu)\epsilon_a)^n f(\mu)}{\int_0^1(\nu(1-\epsilon_b)+(1-\nu)\epsilon_a)^n f(\nu)d\nu} \\ &= \frac{f(\mu)}{\int_0^1 f(\nu)d\nu} = f(\mu) \end{aligned}$$

Thus,

$$LHS = E(\mu|x_n = n) = \mu_0 = \frac{\overline{\lambda}}{\overline{\lambda} + \underline{\lambda}} = RHS.$$

This implies LHS intersects RHS at least once in the relevant range of λ for any n. However, as n increases, the LHS curve shifts upward, while the RHS is independent of n. Furthemore, LHS is flatter than RHS and becomes steeper with an increase in n. This implies there exists an $\bar{n}$, such that for all $n \geq \bar{n}$, LHS intersects RHS twice, first from below at a value lower than λ^{**} and then from above at λ^{**}. For all $n < \bar{n}$ LHS intersects RHS only once at λ^{**}, from below.

We would consider the cost of learning function $c(n)$ to be not sufficiently small if for every $\lambda \in (0, \lambda^{**})$, the optimal choice $n^*(\lambda) < \bar{n}$, i.e., $LHS < RHS$ for all relevant values of λ. For the rest of the proof let us consider a social learning cost function where there exists a $n^*(\lambda) \geq \bar{n}$ for some $\lambda \in (0, \lambda^{**})$.

Given $n^*(\lambda) \geq \bar{n}$ for some λ, let us define $\underline{\lambda}_H$ as the minimum value of λ such that $LHS \geq RHS$. Similarly let define $\overline{\lambda}_H$ as the maximum value of λ such that $LHS \geq RHS$.

Suppose $\overline{\lambda}_H = \lambda_1 < \lambda^*$, then an increase in λ would weakly increase the optimal n^*. Since LHS shifts up with n and intersects the RHS twice before λ^{**}, there would exist $\lambda_2 > \lambda_1$ such that $LHS \geq RHS$ at λ_2, thereby contradicting λ_1 being the highest such value of λ. At λ^{**} the optimal choice of $n^*(\lambda^{**}) = 0$. In this case, the LHS becomes a straight line intersecting RHS at λ^{**}, thus $\overline{\lambda}_H \in [\lambda^*, \lambda^{**})$.

Let us consider any $\lambda \in [\underline{\lambda}_H, \lambda^*]$. In this range the two type of learning are substitutes, i.e., as λ increases $n^*(\lambda)$ also increases weakly. Hence, if at $\underline{\lambda}_H$, $LHS \geq RHS$, for a higher λ with a weakly higher n^* $LHS \geq RHS$ would remain true.

For any $\lambda \in [\lambda^*, \overline{\lambda}_H]$, the two types of learning are complements, i.e., if λ decreases $n^*(\lambda)$ weakly increases. Thus if $LHS > RHS$ at $\overline{\lambda}_H$, it would also be true for any $\lambda \in [\lambda^*, \overline{\lambda}_H]$. Thus for any $\lambda \in [\underline{\lambda}_H, \overline{\lambda}_H]$ herding would be an equilibrium. $\square$

Theorem 2 shows that herding would be an equilibrium even when the two types of learning are complements, since $\overline{\lambda}_H \in [\lambda^*, \lambda^{**})$. This has important policy implications.

Consider the policy where the first few generations are not allowed to learn socially. This is a welfare-improving policy in the herding literature since this encourage using own private signal, which in turn increases the probability of choosing the correct action over time.

But given 2 since herding can be optimal when the two types of learning are complements, reducing social learning would reduce the level of private learning and can reduce the net expected payoff of the agent. Thus this is not unambiguously a welfare-improving policy.

Corollary 1. *Herding is more likely for a less precise prior belief, i.e., when μ_0 is closer to 0.5.*

Proof. As μ increases LHS becomes flatter and also

$$\frac{\partial \lambda^{**}}{\partial \mu_0} = -\frac{(\overline{u} - \underline{u})}{\mu_0 \ln\left(\frac{\mu_0}{1-\mu_0}\right)} < 0.$$

This implies as μ_0 increases for a given set of parameters $\overline{u}$ and $\underline{u}$ and any value of n, a smaller range of λ makes $LHS \geq RHS$, i.e., $\mu_n \in R_2$. Thus by theorem 2, for a larger μ_0 herding is possible for a smaller range of λ, making is less likely. □

Note that it is not obvious a priori that with an increase in μ_0, herding would be less likely, since a higher μ_0 on one hand reduces the probability of any learning by reducing λ^{**} but, makes the private learning cheaper at any λ thereby decreasing the dependence on social learning. However, since individual agents are learning more under higher μ_0 in period $t = 0$, herding becomes less harmful in terms of payoff loss for future generation.

3.3 Discussion of Assumptions

One of the most crucial assumptions of the paper is the private learning technology. The assumption of Shannon mutual entropy simplifies the structure of the optimization problem and allows me to describe the optimal learning strategy, however, this simplification preserves several important features of any private learning technology. First, the mutual entropy cost function belongs to a larger class of cost functions, namely, Posterior Separable (PS) (refer Caplin et al. (2022)) which allows the cost of learning to be dependent on the posterior only. Thus we do not need to specify the form of information structure that generates the posterior. Second, the cost function allows the cost to be increasing in *precision* without any distributional assumption on the prior or signal structure. Third, the cost of the learning strategy depends on the prior. This captures the notion that with a sufficiently confident prior the cost of learning becomes relatively more expensive for further learning which makes it easier for a herding equilibrium to exist. Assuming other learning technologies can make the problem intractable or uninteresting.

For simplification, I have made several other assumptions about learning protocols. Assuming the payoff only depends on idiosyncratic states implies social learning is only informative about the distribution of types. The assumptions of the homogeneous private and social cost of learning allow the agents in a later generation to update their information upon social learning from previous generations. Also, for social learning, we assumed the protocol of block learning where n is chosen before any observation. This assumption generates lemma 1. In the appendix, we relax all these assumptions. We consider four extensions, namely, aggregate state affecting payoff, heterogenous cost of private and social learning, and sequential learning protocol. Under suitable adjustments, all these extensions preserve the main result of the paper.

4 Conclusion

To conclude, this paper solves a model of individual stochastic choice where agents are rationally inattentive and face a costly social learning mechanism. The optimal choice of social learning is non-monotonic in the marginal cost of private learning. Herding can only happen for an intermediate level of private cost of learning where the two types of learning are complements. Restricting access to social learning in early periods does not necessarily improve welfare. In case the two types of learning are complements of each other, restricting access to social learning can increase the probability of mismatching actions to types since the optimal level of private learning also decreases. The only unambiguously welfare-improving policy is to lower the marginal cost of private learning.

In addition to the context of college major choice, this model can be applied to various discrete choice problems related to life-cycle decisions, e.g., choosing an appropriate career, deciding whether to get a college education or not, deciding whether to join labor force participation or not. Given the significance of these decisions on lifetime income and wealth accumulation and the significant role of social learning in all of them, appropriate policy measures can affect the quality of decision making of individual agents.

Bibliography

Ali, S. N. (2018). Herding with costly information. *Journal of Economic Theory 175*, 713–729.

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

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

Blackwell, D. et al. (1953). Equivalent comparisons of experiments. *The annals of mathematical statistics 24*(2), 265–272.

Bobkova, N. and H. Mass (2022). Two-dimensional information acquisition in social learning. *Journal of Economic Theory 202*, 105451.

Brenner, T. (2006). Agent learning representation: advice on modelling economic learning. *Handbook of computational economics 2*, 895–947.

Burguet, R. and X. Vives (2000). Social learning and costly information acquisition. *Economic theory 15*(1), 185–205.

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

Caplin, A., M. Dean, and J. Leahy (2019). Rational inattention, optimal consideration sets, and stochastic choice. *The Review of Economic Studies 86*(3), 1061–1094.

Caplin, A., M. Dean, and J. Leahy (2022). Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy 130*(6), 1676–1715.

Caplin, A., J. Leahy, and F. Matějka (2015). Social learning and selective attention. Technical report, National Bureau of Economic Research.

Cover, T. M. and J. A. Thomas (2012). *Elements of information theory*. John Wiley & Sons.

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

Goeree, J. K., T. R. Palfrey, and B. W. Rogers (2006). Social learning with private and common values. *Economic theory 28*, 245–264.

Hendricks, K., A. Sorensen, and T. Wiseman (2012). Observational learning and demand for search goods. *American Economic Journal: Microeconomics 4*(1), 1–31.

Kamenica, E. and M. Gentzkow (2011). Bayesian persuasion. *The American Economic Review 101*(6), 2590–2615.

Kultti, K. and P. Miettinen (2006). Herding with costly information. *International Game Theory Review 8*(01), 21–31.

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

Mueller-Frank, M. and M. M. Pai (2016). Social learning with costly search. *American Economic Journal: Microeconomics 8*(1), 83–109.

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

Smith, L. and P. Sørensen (2000). Pathological outcomes of observational learning. *Econometrica 68*(2), 371–398.

A Extensions

A.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.

A.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 **??** still holds .

Since the agent doesn't learn about aggregate state privately, the belief about the aggregate state before 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

$$u_{\pi_{x_n}}(i,\omega) = u(i,\omega,g)\, Pr(g|\pi_{x_n}) + u(i,\omega,b)\, Pr(b|\pi_{x_n}) \tag{20}$$

Hence, posterior probability of choosing action i after private learning would still be given by 10 where the $u(i,\omega)$ would be as replaced by $u_{\pi_{x_n}}(i,\omega)$ as defined in equation 20. 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,

$$\epsilon_0^{i,s} = P\Big(i, \omega_j | \mu, \pi = Pr\{s\} = 1\Big) \tag{21}$$

For any generation $t > 1$ the error probabilities are again the expected error probabilities given n and it also uses the same conditioning on the aggregate state as in equation 21. Thus the probability of observing x_n for $s = h, l$ would be given by

$$P\left(x_n|\mu, s\right) = \sum_{k=0}^{n} \sum_{j=k^*}^{k^{**}} \binom{n}{x_n - 2j + k} \mu^{x_n - 2j + k} \left(\epsilon_t^{a,s}\right)^j \left(1 - \epsilon_t^{b,s}\right)^{x_n - j} \left(1 - \mu\right)^{n - x_n - k + 2j} \left(\epsilon_t^{b,s}\right)^{k-j} \left(1 - \epsilon_t^{a,s}\right)^{n - x_n - k + j} \tag{22}$$

where the error probabilities uses similar conditioning as in equation 21 for any $t \geq 1$, otherwise same as before. Then using independence a Bayesian agent would update his belief as follows,

$$P\left(\mu, s|\gamma, \pi_0, x_n\right) = \frac{P\left(x_n|\mu, s\right) P\left(\mu|\gamma\right) P\left(s|\pi_0\right)}{\int_{\substack{\nu \in \gamma \\ s \in S}} P\left(x_n|\nu\right) P\left(\nu|\gamma\right) P\left(s|\pi_0\right)}, \quad \text{for } s = h, l, \mu \in \gamma \tag{23}$$

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

$$P\left(s|\gamma, \pi_0, x_n\right) = \int_{\mu \in \gamma} P\left(\mu, s|\gamma, \pi_0, x_n\right) \tag{24}$$

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.

A.1.2 Agent's Optimization

Given the belief π_{x_n} we can construct $u_{\pi_{x_n}}\left(i, \omega\right)$ using equation 20, then the $V\left(\mu\right)$ 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. 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,

$$W^S(A,\gamma) = \max_n E_{\gamma'_{x_n}} \left[E_{\pi_{x_n}} \left(V\left(A, \gamma'_{x_n}\right)\right)\right] - c(n) \tag{25}$$

which is same as 19 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 25 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 part 4 of theorem 1. Let's start by showing for any other γ, π_0, there exists λ^{**} such that for all $\lambda > \lambda^{**}$, $P(a|\gamma) = 1$(or 0). Now the prior probability of choosing action a for an agent in $t = 0$ is given by,

$$P\left(a|\gamma,\pi^0\right) = \begin{cases} \frac{\gamma(\omega_1)e^{\frac{u^{\pi_0}(a,\omega_1)}{\lambda}} - \left(1-\gamma(\omega_1)\right)e^{\frac{u_{\pi_0}(b,\omega_1)}{\lambda}}}{e^{\frac{u_{\pi^0}(a,\omega_1)}{\lambda}} - e^{\frac{u_{\pi_0}(b,\omega_1)}{\lambda}}} & \text{if } \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)} \in [e^{\frac{-\Delta^{\pi_0}u}{\lambda}}, e^{\frac{\Delta_{\pi_0}u}{\lambda}}] \\ 1 & \text{if } \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)} > e^{\Delta_{\pi^0}u/\lambda} \\ 0 & \text{if } \frac{\gamma(\omega_1)}{1-\gamma(\omega_1)} < e^{-\Delta_{\pi^0}u/\lambda} \end{cases} \tag{26}$$

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

$$\lambda^{**} = \max \left\{ \max_{\pi^0 \in \Pi_0} \Delta_{\pi^0} u / \log \frac{\gamma'(\omega_1)}{1 - \gamma'(\omega_1)}, \max_{\pi^0 \in \Pi_0} \Delta_{\pi^0} u / \log \frac{1 - \gamma'(\omega_1)}{\gamma'(\omega_1)} \right\}$$

[8] 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(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.

Thus, 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 counter-intuitive 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 an agent stick with their prior and no learning at all. Hence, any behavior becomes completely uninformative for the next generation which makes zero social learning optimal.

A.2 Heterogeneous Cost of Private learning

To the baseline model of section 2 now we add heterogeneity in the marginal cost of private learning. Everything else is the same let $\lambda \sim F(\lambda)$, instead of λ is 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.

[8] 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_λ^i 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

$$\epsilon_{0,F}^i = \int_{\lambda \in F} \epsilon_\lambda^i dF \tag{27}$$

If the earlier error probabilities are replaced by $\epsilon_{0,F}^i$ as defined in 27, the optimization problem for the agent in $t = 1$ remains the 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_γ^a and f_γ^b 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 $\epsilon^i_{t,F} = \int_{\lambda \in F} \int_{x_{n_\lambda^*} \in X_{n_\lambda^*}} \epsilon^i_{x_{n_\lambda^*},\lambda,t} df_\gamma^i dF$ 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. Since F is 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 the same as before.

To state a similar result as that of theorem 1 we first need to define an ordering over all possible $F(.)$. Let us consider the FOSD over the distribution of λ. Given any cost of social learning function $c(n)$, F_1 would generate the lower error probability than F_2 if F_2 FOSD F_1. This is true because the error probabilities are obtained by taking an expectation over different error probabilities for different λ and a higher λ always give higher error probability. If we replace the statement of Theorem 1 with F, in place of λ along with $FOSD$ all the results still hold.

A.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 8. In this case, the problem would not be very different when G is common knowledge and is also independent of the distribution of types. In that case, the period $t = 0$ agents are still identical since they don't learn socially. So, ϵ_0^i 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 the 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_t^a and ϵ_t^b as $\epsilon_t^i = \int_{C_\alpha \in G} \int_{x_{n_\alpha^*} \in X_{n_\alpha^*}} \epsilon^i_{x_{n_\alpha^*},t} df_\gamma^i dG$ 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.

A.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^t_{min} denotes the minimum value of n in the set $\mathbb{N}_t$. Under the sequential strategy, we can write a similar proposition as that of theorem 1.

Proposition 2. *Under sequential social learning , there exist a set of cutoff values of* λ, *namely,* $0 \leqslant \lambda^*_s \leq \lambda^i_s < \lambda^d_s \leq \lambda^j_s < \lambda^{**} \leqslant \infty$, *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* $n^t_{min}(\lambda_1) \leqslant n^t_{min}(\lambda_2)$, *where* $\lambda_1 \leqslant \lambda_2$, *i.e.* n_{min} *is non-decreasing in marginal cost of private learning.*

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

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

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.*

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 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 $V(\mu)$ for $\mu \geq 1/2$ as the other case would be symmetric. So a higher belief, i.e., a higher value of μ would mean a belief further away from the 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 $V(\mu)$ 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 the marginal gain is less than marginal loss, where the loss is due to extra cost, then n_{min} has to be associated with the 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 an 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 the 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 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 to 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.

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 sets 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 a marginal loss as it would imply the agent would never learn up to the final increasing section. Also for all other beliefs, 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 theorem 1, 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.

□