Learning How People Learn

Guo Bai

January 10, 2023

Abstract

This paper investigates whether it is possible for a designer to learn how a decision maker learns from the information. The decision maker observes signals about an unknown payoff-relevant state, learns about the state from the signals and then takes an action at each time. A designer designs the payoff the decision maker gets, observes the decision maker's signals and actions, and then tries to learn how this decision maker learns from the information. This paper investigates a starting point of this problem: given the common knowledge that the decision maker is Bayesian but the prior is unknown, is it possible for the designer to learn the decision maker's prior? This paper studies the probability of the designer learning the decision maker's prior and characterises the optimal payoff structure that maximises this probability.

1 Introduction

When a payoff-relevant state is unknown, a decision maker (DM) learns from information about the state before making decisions. For example, before buying a product, the consumers watch advertising and introduction of the product, search the internet for reviews about the product, and then decide whether to buy or not. Investors read the financial report and news of a company to check if the company is worth investing in and then decide whether to invest or not. Chain restaurants do market research and use the data from the market research to see whether it is profitable to open a new restaurant. The DM collects information, processes information and learns from the information.

How does a DM learn from the information? One of the canonical theories is that the DM has some prior belief about the unknown state. If the DM receives some information and signals about the unknown state, then, she updates her prior belief based on the new observations. The process that the DM forms a prior belief and then updates her belief based on new observations is referred to as 'learning from the information' in this chapter. Mainstream economic theories impose Bayesian DM assumptions where the DM updates the belief using Bayes rule given some prior. However, the psychology and behavioural economic literature suggest that people process information systematically departure from using Bayes rule (see Tversky & Kahneman (1974), Rabin (1998) and Camerer (1998)). Behavioural economists have investigated and developed theories about alternative belief updating rules. Epstein et al. (2010) provides a non-Bayesian updating rule to capture the underreaction and overreaction to signals. Rabin (1998) analyses the bias in beliefs when the order of the signals matters in how people infer future signals. Besides the theories about the non-Bayesian updating rules, there are some ad hoc tests used to understand how a DM learn from the information. For example, Angrisani et al. (2017) studies how people process information in a social learning network. The literature infers that it is unclear how a DM learn from the information. In addition, learning how people learn from the information requires further studies.

The question of interest is: By observing what a DM sees and what they do, is it possible that a third party could learn how this DM learns from the information they received? For example, consider the experimental economists designing an experiment to learn how the experiment participant processes the information. If the designer sees the information received by the participant and the actions taken by the participant, then, is it possible for the designer to learn about the participant's beliefs? Theoretically speaking, answering this question provides a theoretical foundation for the experiments that are interested in detecting how people learn from the information. In addition, answering this question has applications in industry. Firms can benefit from learning how their consumers react to the information they provide. They can design advertising and free trials better to attract targeted consumers. Central banks can benefit from knowing how the public processes the news they release. They then can do better forward guidance in terms of the new monetary policy.

In this chapter, I investigate the question in the following situation. Suppose it is common knowledge that the DM updates their beliefs using Bayes rule but the prior is unknown, I investigate whether it is possible for the designer to learn the DM's prior belief by observing the signals she received and the actions she took. Learning the DM's prior can be considered a starting point for learning how the DM learns. Given that the

DM updates beliefs using Bayes rule, if the state is binary, all the designer needs to learn is one parameter. If the state is finite but not binary, the designer just needs to learn a distribution with a finite number of parameters. This is the least demanding situation in terms of what the designer needs to learn. If the designer cannot learn how the DM learns in this situation, then, it is very likely that the designer is not able to learn in a more demanding environment. It is hence considered a starting point.

Whether the designer can learn the DM's prior depends on the variation in the DM's actions and the informativeness of the signals. The variation in the DM's actions is important. For example, if the DM always takes the same action regardless of the beliefs, then, the designer can learn nothing about the DM. If the DM's action is reporting the beliefs directly, it is likely that the designer can learn how the decision maker learns with sufficiently many observations. The informativeness of the signals about the state matters as well. If the signals are very informative, the decision-maker learns the state very fast. Then, there will not be too many variations in the DM's actions eventually. It is likely that the designer cannot learn anything after a certain point. If the signals are not very informative, the DM's beliefs fluctuate a lot. The fluctuation of the beliefs then may induce variations in the actions. Then it is more likely that the designer can learn how the DM learns.

Related literature This paper is related to the experimental economics papers about detecting the bounded rationality of individuals. Angrisani et al. (2017) design the experiment to detect how individuals to update beliefs in a social learning network. The agents in their model act sequentially after observing statements from the neighbours and a private signal about the state. The agents' actions in their model are continuous. This paper is different from Angrisani et al. (2017). This paper focuses on asking the question of whether the experiment designer can learn how the agents update their beliefs by observing their actions and the signals. The focus is not on how the individual interprets the signals from different sources. Augenblick & Rabin (2018) examines the time-inconsistency preference of individuals. Their paper focuses on the experiment designer observing the actions to detect the form of the individual's preferences. The aim of their paper is different from this one. This paper is mainly about beliefs rather than preferences.

Mathematically, this paper is using the idea of a sequential probability ratio test by Wald (1945). The sequential probability ratio test considers the odds ratio as a function of the probability of each observation and the number of observations. The sequential probability ratio test is for hypothesis testing. The aim of the test is to decide which hypothesis is correct in the shortest time period, i.e. with the smallest amount of observations. There are two thresholds: a lower one and a higher one. The sequential probability ratio test ends if the odds ratio falls below the lower threshold or jumps above the higher threshold. The stopping time of the test follows Wald's identity. It allows us to calculate the probability that the odds ratio hits the two thresholds. This paper uses Wald's identity to calculate the probability of the designer learning the prior.

2 The Model

2.1 Model Setup

There are two players: a designer and a decision maker (DM). Time $t = 0, 1, 2, \ldots$ is discrete and potentially infinite. The state is drawn from the set $\Theta = \{\theta, \theta'\}$ at time $t = 0$ and is constant over time. Both the designer and the DM are uninformed about the state. The designer's prior belief is denoted $\mu_0 = \Pr(\theta) > 0$ and the DM's prior belief is denoted $p_0 = \Pr(\theta) > 0$. The DM's prior belief is private information and is referred to as the DM's type. Let P denote the set of the DM's type. I discuss two cases in this chapter: $P = \{\underline{p_0}, \bar{p_0}\}$ and $P = (0, 1)$. The designer believes that p_0 follows a distribution Π. The DM observes a signal and then takes an action at each time t to maximise her time-t utility. The designer designs the DM's payoff function and observes the DM's actions. The designer's objective is to learn the DM's type.

The designer designs the payoff function $U : A \times \Theta \to \mathbb{R}^+$ at $t = 0$. At each time $t > 0$, the DM observes a signal $s_t \in \{1, 2, , \ldots, S\} = \mathcal{S}$ and takes an action $\alpha_t \in A = \{0, 1\}$. If the state is θ, the DM observes signal $s_t = s$ with probability $\Pr(s_t = s|\theta) = a_s > 0$ where $\bar{a} := (a_s)_{s\in\mathcal{S}}$ is the distribution of the signals in state θ. If the state is $\theta^{'}$, the DM observes signal $s_t = s$ with probability $\Pr(s_t = s|\theta) = b_s > 0$ where $\bar{b} := (b_s)_{s\in\mathcal{S}}$ is the distribution of the signals in state $\theta^{'}$. The signal structure is exogenous. Let $h_t = (s_0, s_1, \ldots, s_{t-1}) \in s^t$ denote the history of signals. The DM observes h_t and forms a posterior belief $p_t = \Pr(\theta|h_t)$ using Bayes rule. The DM takes an action $\alpha_t \in A = \{0, 1\}$ at each time t to maximise her time t expected payoff. The payoff at time t is not observed by the DM until the experiment ends.

The timing is as follows. At time $t = 0$, the designer designs the utility function U. At time $t > 0$, the DM observes the signal s_t, updates her belief, and then takes an action α_t. The designer observes the action α_t. If the designer learns the type of DM, the experiment ends. If the designer does not learn the type of DM, the experiment continues to the next period.

Given the binary actions and binary states, instead of designing the utility function $U : A \times \Theta \to \mathbb{R}^+$, the designer's action can be simplified to choosing a threshold

$$r := \frac{U\left(0, \theta'\right) - U\left(1, \theta'\right)}{U\left(1, \theta\right) - U\left(0, \theta\right) + U\left(0, \theta'\right) - U\left(1, \theta'\right)} \in (0, 1)$$

such that it is optimal for the DM to take action $\alpha_t = 1$ if and only if $p_t > r$.

2.2 An Example

The designer has two kinds of biased coins: A and B. The name of the coin is the 'state' in the model. Coin A comes up Heads with probability $\frac{2}{3}$ and Tails with probability $\frac{1}{3}$. Coin B comes up Heads with probability $\frac{1}{3}$ and Tails with probability $\frac{2}{3}$. These two kinds of biased coins are in a non-transparent box. The composition of the coins inside the box is unknown. At the beginning of the game, the DM draws a coin from the box. Then, the DM tosses the coin and guesses which coin it is. Since the draw is random and the composition of the coin is unknown, the state is unknown and the DM's prior belief is private.

The designer designs the DM's payoff. The payoff function is announced before the coin is drawn. An example of the payoff function is as follows. If the coin is A and the DM guesses it correctly, then, the DM gets M. If the coin is A and the DM guesses it wrong, then, she gets m. If the coin is B and the DM guesses it correctly, then, she gets N. If the coin is B and the DM guesses it wrong, then, she gets n. The designer wants to learn the DM's prior by observing the coin toss results and the DM's guesses.

The timing of the payment guarantees that the DM does not learn the state from the payoff. If there is only one period, after the guess, the designer and the DM check the name of the coin, and then the designer pays the DM accordingly. If there is more than one period, the coin is tossed and the DM makes guesses at each time. The payment is made at the end of the experiment. The experiment ends after a predetermined time period T (if the time horizon is finite), after the designer learns the type of the DM (if the time horizon is infinite), or continues forever if there is no learning. Since the state is unknown to both the DM and the designer, the designer records the DM's guess at each time and calculates the payoffs at the end of the experiment when the state is revealed.

2.3 Discussion

I make the assumptions that the designer can design the DM's payoffs and DM's actions are observable. These assumptions make it easier for the designer to learn the DM's type. The purpose of this chapter is to check whether the designer can learn in this environment. If the designer cannot learn in this case, then, it is even harder to learn when actions are noisy or the utility function is exogenous.

In the example, the DM announces the 'guess' each time and the designer observes this guess when it is announced. This is the case that the action is observable. It may not always be the case that the action is observable. Sometimes, only the noisy version of the action is observable. For example, suppose the DM cannot talk but has two coins, A and B, in her pocket. After tossing the coin drawn from the box, the DM cannot announce her guess but she can pick up the corresponding coin in her pocket. Then, the designer can flip the coin picked by the DM, observes a signal of the DM's guess and then infers the DM's guess, The observable action assumption makes it easier for the designer to learn. During the experiment, what the designer observes are a series of actions taken by the DM and a series of signals about the state. There is no extra level of uncertainty compared to the noisy version of the actions. In this section, I consider the case with observable actions to check if the designer can learn the type of the DM. If the designer cannot learn the type of the DM in this case, it is highly possible that the designer cannot learn when the observed actions are noisy.

In the example, the designer can choose the payoffs the DM gets. This makes it easier for the designer to learn. Consider the case that the designer cannot choose the payoff function in the coin example. Suppose that the DM gets 1 if she makes a correct guess and 0 otherwise. Assume that the DM is risk neutral. Then the DM would guess A if she believes that the coin is A with a probability of at least a half. Suppose that the DM has the prior that the coin is A with probability 0.8, and if the coin comes up heads more than tails, it is likely that the DM's belief is always bigger than a half. The designer then keeps observing the A guesses. The designer may not learn whether the DM has a prior 0.8 or the DM has other priors bigger than 0.8. If the DM with prior 0.8 guesses A all the time, the DM who has a prior bigger than 0.8 would guess A as well. The designer then cannot learn. However if the designer can choose the payoff function, the designer has the ability to induce actions that are favourable in terms of learning. Suppose that the designer somehow believes that the DM has the prior that the coin is A with a high probability. The DM believes that the coin is A with a probability of 0.8. Then, the designer may learn the prior of the DM if the designer chooses the payoff function as follows. If the coin is A and the DM guesses it correctly, the DM gets 1; if the coin is B and the DM guesses it correctly, the DM gets 3; if the guess is wrong, the DM gets 0. Given this payoff function, the DM will guess B if she believes that the probability that the coin is A is smaller than $\frac{3}{4}$. The payoff of guessing B correctly is big enough so that it is worth taking the risk. Consider the action of the DM who believes that the coin is A with a probability of 0.8. Before the first signal, the DM guesses A because the prior is 0.8. After the first signal, if the signal is tails, the DM believes that the coin is A with probability $\frac{2}{3}$ and then guesses B. Then, the designer can back up the prior belief of the DM. The designer can back up that the prior of the DM is in the range $(\frac{3}{4}, \frac{6}{7})$.

3 Analysis

This section analyses the model. I first show that observing a switch of action helps the designer to learn the DM's prior. Then, I show that when there are two types of the DM, that is, $P = \{\underline{p_0}, \bar{p_0}\}$, the designer learns the DM's type immediately. Then, I show that when $P = (0, 1)$, the designer does not learn the DM's type for certain. I characterise the designer's optimal payoff function design that maximises the probability of learning in this case.

3.1 Switch of Actions

In this section, I show that observing a switch of action helps the designer to learn the prior of the DM. Without imposing assumptions on P, I compute the probability of observing a switch of action conditional on the state θ and θ' given a threshold r.

Although the time in this model is infinite, the real time for the designer to learn the type of the DM is actually finite. This is because a Bayesian DM learns the state after observing sufficiently many signals. Therefore, given any type $p_0 \in (0, 1)$, the sequences of the DM's actions eventually converge. The sequence of the DM's actions converges to 1 if the state is θ and converges to 0 if the state is θ'. As a result, the designer must learn the DM's prior faster than the DM learns the state.

In order to learn the type of the DM, there must be variations in the different types of DMs' actions. The designer learns the type of the DM if the sequences of actions are different for different types of the DM after they receive the same sequence of the signals. Consider the case that there are two types of the DM. After receiving the same sequence of signals, if one type of the DM takes a sequence of actions $(0, 0, 0, 0, \dots, 0, 0, \dots)$ and the other type of the DM takes a sequence of actions $(0, 0, 1, 1, \dots, 0, 0, \dots)$, the designer can learn the type of the DM. Consider the case that there are more than two types of the DM. The designer learns the type of the DM if different types of the DM switch their actions and first switch their actions at different time after observing the same sequence of signals.

Suppose the designer chooses a threshold r. The DM takes action 1 if the belief about the state being θ is greater than or equal to r. Consider the DM has a prior p_0 smaller than r. The DM first takes action 0. After receiving a sequence of signals, the DM first switches the action from 0 to 1 if the belief about the state goes from below the threshold r to above the threshold r. If the state is θ, this switch of actions happens with probability 1 as the sequence of actions converges.

If the DM has the prior p_0 larger than r. The DM first takes action 1. The designer can observe a switch of action if the DM's belief about the state being θ falls below the threshold r. Since the DM is Bayesian the belief converges to 1 if the state is θ, the designer may or may not observe a switch of action. The probability of the DM's belief falling below r is positive but not 1. This probability can be derived given the value of r using Wald's Fundamental Identity. The probability of observing a switch of action is summarized in the lemma below.

Let $v^* < 0$ satisfy $\sum_{s\in\mathcal{S}} a_s(\frac{a_s}{b_s})^{v^*} = 1$, let $u^* > 0$ satisfy $\sum_{s\in\mathcal{S}} b_s(\frac{a_s}{b_s})^{u^*} = 1$. and let $k := \ln\frac{r}{1-r} - \ln\frac{p_0}{1-p_0}$. Proposition 1 characterises the probability of observing a switch of action in state θ and $\theta^{'}$ given the type of the DM p_0 and the threshold r.

Proposition 1. *Assume that θ is the underlying state and p_0 is the type of the DM.*

If the designer chooses $r > p_0$, the switch of actions is observed w.p. 1.

If the designer chooses $r < p_0$, the switch of actions is observed w.p. $\frac{1}{e^{v^*k}}$.

If the designer chooses $r = p_0$, the switch of actions is observed w.p. $\frac{1}{e^{v^*\eta}}$ where $\eta = \lim_{r\to p_0} k$.

Assume that θ' is the underlying state and p_0 is the type of the DM.

If the designer chooses $r < p_0$, the switch of actions is observed w.p. 1.

If the designer chooses $r > p_0$, the switch of actions is observed w.p. $\frac{1}{e^{u^*k}}$.

If the designer chooses $r = p_0$, the switch of actions is observed w.p. $\frac{1}{e^{u^*\eta}}$ where $\eta = \lim_{r\to p_0} k$.

3.1.1 A Special Case: Binary Signals

The probability of observing a switch of action requires solving for v^* or u^* numerically. It is hard to compute the exact values when there are many signals. In this section, I consider a special binary-signal case.

There are two signals $\mathcal{S} = \{1, 2\}$ with distributions $(a, 1-a)$ in state θ and $(1-a, a)$ in state θ'. I assume $a > 0.5$ so that it is more likely to receive signal 1 in state θ and more likely to receive signal 2 in state θ'. When a is very close to a half, i.e. when the signal is very uninformative, the belief of the DM takes very small steps after receiving one more signal. In the limit, it is like the case with continuous times. The probability of learning is easier to compute given this signal structure.

Proposition 2 characterises the probability of observing a switch of action when $\mathcal{S} = \{1, 2\}$. Let $B := \frac{\ln \frac{r}{1-r} - \ln \frac{p_0}{1-p_0}}{\ln \frac{a}{1-a}}$.

Proposition 2. *Assume that θ is the underlying state and p_0 is the type of the DM.*

If the designer chooses $r > p_0$, a switch of actions is observed with probability 1.

If the designer chooses $r < p_0$, a switch of actions is observed with probability $(\frac{1-a}{a})^{-B}$.

If the designer chooses $r = p_0$, a switch of actions is observed with probability $\frac{1-a}{a}$.

Assume that θ' is the underlying state and p_0 is the type of the DM.

If the designer chooses $r \leq p_0$, a switch of actions is observed with probability 1.

If the designer chooses $r > p_0$, a switch of actions is observed with probability $(\frac{1-a}{a})^{B}$.

In the following sections, I discuss the designer's optimal choice of the threshold r and the probability of learning the DM's type when the signals are binary.

3.2 Two types of DM

When there are two types of DM, the designer can learn the type of the DM immediately.

Theorem 1. *When $P = \{\underline{p_0}, \bar{p_0}\}$, the designer can learn the type of the DM immediately by choosing $r^* \in \left(\underline{p_0}, \bar{p_0}\right]$.*

By choosing the threshold r^*, the DM with prior $\underline{p_0}$ takes action 1 and the DM with prior $\bar{p_0}$ takes action 0. The designer learns the type of the DM immediately.

3.3 Continuous Types

In this section, I consider the case when $P = (0, 1)$ and when Π is uniform. This case describes the situation that the designer knows nothing about the DM's prior. The designer does not know the potential prior the DM may have and the designer does not know which prior is more likely. The continuously uniformly distributed prior can be considered as a conservative assumption the designer has about the DM. In this case, the designer does not learn the type of the DM with probability one. The objective of the designer is to maximise the probability of learning the prior. There are three factors that affect the choices of the threshold r when the designer wants to maximise the probability of learning the prior of the DM. The three factors are the designer's prior belief about the state, the designer's prior belief about the type of the DM, and the distances between the priors of the DM.

The designer's prior belief about the state determines whether the threshold r is closer to 0 or closer to 1. If the designer believes that the state is more likely to be θ, the designer believes that it is more likely that the beliefs of the DM are going up to 1 regardless of the types of the DM. Choosing the threshold r closer to 1 can maximise the probability of learning the type of the DM in the state θ. However, if the designer believes that the state is more likely to be θ', the designer believes that it is more likely that the beliefs of the DM are going down to 0 regardless of the types of the DM. Choosing the threshold r closer to 0 can maximise the probability of learning the type of the DM in the state θ'.

The designer's prior belief about the type of the DM determines whether the designer wants to choose the threshold r closer to a certain prior. The closer the threshold r is to a range of certain priors, the higher the probability is for the designer to learn those priors when the state is unknown.

The distances between the priors matter. Let the lowest prior the DM can have is $\underline{p_0}$ and the highest prior the DM can have is $\bar{p_0}$. If the distribution of the priors has a smaller variance, it is possible that the designer is more willing to choose the threshold $r \in [\underline{p_0}, \bar{p_0}]$. If the distribution of the prior has a larger variance, it is possible that the designer is more willing to choose the threshold to be $\underline{p_0}$ or $\bar{p_0}$.

The following lemma characterises the upper and lower bound of the optimal threshold r given P and the signal structure.

Lemma 1. *When $P = (\underline{p_0}, \bar{p_0})$, the optimal threshold r^* satisfies the condition that*

$$\ln \frac{\underline{p_0}}{1-\underline{p_0}} - \ln \frac{a}{1-a} < \ln \frac{r}{1-r} < \ln \frac{a}{1-a} + \ln \frac{\bar{p_0}}{1-\bar{p_0}}.$$

This is a direct result from the discussion above. If the type of the DM p_0 is in the interval $[\underline{p_0}, \bar{p_0}]$, the designer never chooses the threshold r such that $\frac{\ln \frac{r}{1-r} - \ln \frac{\bar{p_0}}{1-\bar{p_0}}}{\ln \frac{a}{1-a}} > 1$, and the designer never chooses the threshold r such that $\frac{\ln \frac{r}{1-r} - \ln \frac{\underline{p_0}}{1-\underline{p_0}}}{\ln \frac{a}{1-a}} < -1$.

Next, I characterise the condition that the optimal threshold r^* satisfies. I show that the optimal threshold that maximises the probability of learning the prior is a function of μ_0 and it is increasing in μ_0. If the designer believes that the state is more likely to be θ, the designer will choose r^* closer to 1. It will allow the designer to learn more priors in state θ, but learn less in state θ'. If the designer believes that the state is more likely to be θ',the designer will choose the r^* closer to 0. It will allow the designer to learn more priors in state θ', but learn less in state θ.

Proposition 3. *The optimal choice of the threshold r^* that maximises the probability of learning the prior satisfies*

$$\frac{(r^*)^2}{(1-r^*)^2}\frac{\ln r^* + 1 - r^*}{\ln(1-r^*) + r^*} = \frac{1-\mu_0}{\mu_0}.$$

The optimal threshold r^* increases in μ_0.

4 Extensions

4.1 Finite Periods of Learning

Previous sections discuss the cases that the time horizon is infinite. I focus on discussing the probability of learning in the long run. In this section, I consider the case that the time $t = 0, 1, \dots, T$ is finite. The finite-time assumption is more relevant to real-life applications. For experimental economists who design experiments to detect the belief of the experiment candidates, the experiments cannot last forever. For firms that are doing trials to understand their customers, they cannot do an infinite number of trials

Proposition 4 characterises the probability of observing a switch of action in state θ and $\theta^{'}$ given the type of the DM, the threshold r and the final time period of the experiment T. Before presenting the proposition, I define $\Phi_B(s)$, $\Psi_B(s)$, $\Xi_B(s)$ and $\Omega_B(s)$ functions where

$$\Phi_B(s) := (\frac{1-(1-4a(1-a)s^2)^{\frac{1}{2}}}{2(1-a)s})^B,$$

$$\Psi_B(s) := (\frac{1-(1-4a(1-a)s^2)^{\frac{1}{2}}}{2as})^{-B},$$

$$\Xi_B(s) := (\frac{1-(1-4a(1-a)s^2)^{\frac{1}{2}}}{2as})^B,$$

and

$$\Omega_B(s) := (\frac{1-(1-4a(1-a)s^2)^{\frac{1}{2}}}{2(1-a)s})^{-B}.$$

Let $\phi_B(\tau) = \frac{\Phi_B^{(\tau)}(0)}{\tau!}$, $\psi_B(\tau) = \frac{\Psi_B^{(\tau)}(0)}{\tau!}$, $\xi_B(\tau) = \frac{\Xi_B^{(\tau)}(0)}{\tau!}$, and $\omega_B(\tau) = \frac{\Omega_B^{(\tau)}(0)}{\tau!}$ where $f^{(\tau)}(0)$ is the τ-th derivative of the function f evaluated at 0.

Proposition 4. *Assume that θ is the underlying state and p_0 is the type of the DM. If the designer sets $r > p_0$, the designer learns the type of the DM with probability $\sum_{\tau=1}^{T} \phi_B(\tau)$. If the designer sets $r < p_0$, the designer learns the type of the DM with probability $\sum_{\tau=1}^{T} \psi_B(\tau)$.*

Assume that θ' is the underlying state and p_0 is the type of the DM. If the designer sets $r > p_0$, the designer learns the type of the DM with probability $\sum_{\tau=1}^{T} \xi_B(\tau)$. If the designer sets $r < p_0$, the designer learns the type of the DM with probability $\sum_{\tau=1}^{T} \omega_B(\tau)$.

5 Conclusion

This chapter investigates whether a designer can learn how a DM learns an underlying state by observing some public signals and the DM's actions. This chapter analyses the cases in which the prior of the DM is unknown. The objective of the designer is to learn the DM's prior. This chapter shows that the key to learning the DM's prior is being able to observe a switch in the DM's action. If the designer can design the DM's payoffs, then, there is a positive probability that the designer can learn the DM's prior. This chapter also characterises conditions for the optimal payoff design that maximises this probability.

References

Angrisani, M., Guarino, A., Jehiel, P., Kitagawa, T. et al. (2017), Information redundancy neglect versus overconfidence: a social learning experiment, Technical report, Centre for Microdata Methods and Practice, Institute for Fiscal Studies.

Augenblick, N. & Rabin, M. (2018), 'An experiment on time preference and misprediction in unpleasant tasks'.

Camerer, C. (1998), 'Bounded rationality in individual decision making', *Experimental economics* **1**(2), 163–183.

Epstein, L. G., Noor, J. & Sandroni, A. (2010), 'Non-bayesian learning', *The BE Journal of Theoretical Economics* **10**(1).

Grimmett, G. & Stirzaker, D. (2001), *Probability and random processes*, Oxford university press.

Rabin, M. (1998), ‘Psychology and economics’, *Journal of economic literature* **36**(1), 11–46.

Tversky, A. & Kahneman, D. (1974), ‘Judgment under uncertainty: Heuristics and biases’, *science* **185**(4157), 1124–1131.

Wald, A. (1945), ‘Sequential tests of statistical hypotheses’, *The annals of mathematical statistics* **16**(2), 117–186.

A Proof of Proposition 1

First, consider the DM's belief. The DM forms the belief using the Bayes rule. Define the log odds ratio $\Lambda_t := \ln(\frac{p_t}{1-p_t})$. The following part shows that Λ_t follows a random walk. The odds ratio is

$$\frac{p_{t+1}}{1-p_{t+1}} = \frac{p_t}{1-p_t}\frac{\Pr(s_t|\theta)}{\Pr(s_t|\theta')} = \frac{p_0}{1-p_0}\prod_{\tau=0}^{t}\frac{\Pr(s_\tau|\theta)}{\Pr(s_\tau|\theta')} = \frac{p_0}{1-p_0}\prod_{s\in\mathcal{S}}(\frac{a_s}{b_s})^{\hat{t}_s}$$

where $\hat{t}_s$ is the number of s signals before time $t+1$. Taking the logarithms of the odds ratio, we have

$$\ln\frac{p_{t+1}}{1-p_{t+1}} = \ln\frac{p_t}{1-p_t} + \ln\frac{a_{s_t}}{b_{s_t}} = \ln\frac{p_0}{1-p_0} + \sum_{\tau=0}^{t}\ln\frac{a_{s_\tau}}{b_{s_\tau}}$$

Thus

$$\Lambda_{t+1} = \Lambda_t + \ln\frac{a_{s_t}}{b_{s_t}} \tag{1}$$

The log odds ratio follows a random walk with steps $\ln\frac{a_{s_t}}{b_{s_t}}$. If the underlying state is θ, the random walk takes steps with the probability determined by the probability distribution a.

Now consider the DM's action. The DM takes action $\alpha_t = 1$ if $p_t \geq r$ and $\alpha_t = 0$ otherwise. Therefore, the DM takes action $\alpha_t = 1$ if $\Lambda_t \geq \ln\frac{r}{1-r}$. If the DM is observed to switch action from 0 to 1 in period t, then we have $\Lambda_t \approx \ln\frac{r}{1-r}$. Thus,

$$\ln\frac{p_0}{1-p_0} + \sum_{\tau=0}^{t}\ln\frac{a_{s_\tau}}{b_{s_\tau}} \approx \ln\frac{r}{1-r}.$$

Therefore, observing a switch of action is equivalent to observing the type of the DM.

Next step is to characterise the probability of observing a switch of action given the choice of r. Let $x_t := \ln \frac{a_{s_t}}{b_{s_t}}$. Conditional on the state θ, $x_1, x_2, \dots$ are identically and independently distributed random variables. Let $y_t = x_1 + \cdots + x_t$. The sequence $y = \{y_\tau : \tau > 0\}$ is a random walk starting at the origin. Let $k := \ln \frac{r}{1-r} - \ln \frac{p_0}{1-p_0}$. The probability of observing a switch of action is equal to the probability that the random walk y hits the value k.

If $k > 0$, the probability that y hits the value k is 1. Conditional on θ, we have $\mathbb{E}_\theta[\Lambda_{t+1} \mid \Lambda_t] > \Lambda_t$. Therefore, Λ_t is a submartingale. According to the Martingale convergence theorem, if the state is θ, $\Lambda_t \to \infty$ almost surely. If the state is θ', $\Lambda_t \to -\infty$ almost surely.

Since the random walk y satisfies

$$y_t = \Lambda_{t+1} - \ln \frac{p_0}{1-p_0},$$

if the state is θ, the random walk y hits the value $k > 0$ with probability 1.

Now consider the situation that $k < 0$. Let $A > 0$. Define a stopping time τ to be the first time the random walk y exists the interval $[k, A]$. The stopping time τ is finite with probability 1 as the the random walk y tends to infinity. According to Wald's Identity, the stopping time τ satisfies

$$1 = \mathbb{E}[\frac{e^{vy_\tau}}{(\sum_s a_s(\frac{a_s}{b_s})^v)^\tau}] \text{ for } \forall v \neq 0 \text{ s.t. } \sum_s a_s(\frac{a_s}{b_s})^v \geq 1.$$

Choose $v^* < 0$ such that $\sum_s a_s(\frac{a_s}{b_s})^{v^*} = 1$, then,

$$1 = \mathbb{E}[e^{v^* y_\tau}].$$

Let τ_A and τ_k be the two stopping times that y_τ hits A and y_τ hits k, then

$$1 \approx \Pr(y_\tau = A)e^{v^* A} + \Pr(y_\tau = k)e^{v^* k}.$$

Notice that this is an approximation because when y_t hits k, it does not exactly equal k. The equality holds with an equal sign if either the steps the random walk takes are infinitely small (the continuous time case) or the random walk takes steps up and down of the equal sizes. Next section will discuss the case that the random walk takes steps up and down of the equal sizes. Use the fact that $\Pr(y_\tau = k) = 1 - \Pr(y_\tau = A)$ we can get that

$$\Pr(y_\tau = k) \approx \frac{1 - e^{v^* A}}{e^{v^* k} - e^{v^* A}}.$$

Let $A \to \infty$, we have

$$\Pr(y_\tau = k) \to \frac{1}{e^{v^* k}}.$$

When the state is $\theta^{'}$, the proof follows the same approach.

If $k < 0$, the random walk y hits the value k with probability 1. From the proof for lemma 1, conditional on θ', Equation (1) implies that $\mathbb{E}_\theta[\Lambda_{t+1} \mid \Lambda_t] < \Lambda_t$. Therefore, $-\Lambda_t$ is a submartingale. If the state is θ', $\Lambda_t \to -\infty$ almost surely. The random walk y hits the value $k < 0$ with probability 1.

Consider the case that $k > 0$. Let $A < 0$. Define the stopping time τ to be the first time the random walk y exits the interval $[A, k]$. The stopping time τ is finite with probability 1. According to Wald's identity, the stopping time satisfies

$$1 = \mathbb{E}[\frac{e^{uy_\tau}}{(\sum_s b_s(\frac{a_s}{b_s})^u)^\tau}] \text{ for } \forall u \neq 0 \text{ s.t. } \sum_s b_s(\frac{a_s}{b_s})^u \geq 1.$$

Choose $u^* > 0$ such that $\sum_s a_s(\frac{a_s}{b_s})^{u^*} = 1$. Follow the same steps as in the proof for lemma 1. We have

$$\Pr(y_\tau = k) \to \frac{1}{e^{u^* k}}.$$

B Proof of Proposition 2

Consider the random walk y defined previously where $y_t = x_1 + \cdots + x_t$. Given the signal structure characterised above, if $s_t = 1$, we have $x_t = \ln \frac{a}{1-a}$, and if $s_t = 2$, we have $x_t = \ln \frac{1-a}{a}$. The random walk y now starts at the origin and takes steps up and down by equal amounts $\ln \frac{a}{1-a}$. If the threshold is r, let $B = \frac{\ln \frac{r}{1-r} - \ln \frac{p_0}{1-p_0}}{\ln \frac{a}{1-a}}$ be the corresponding steps the random walk y takes to hit the value $\ln \frac{r}{1-r} - \ln \frac{a}{1-a}$. The expressions 'the random walk y hits the value $\ln \frac{r}{1-r} - \ln \frac{a}{1-a}$' and the 'random walk y hits the step B' are the same.

We now have a random walk y starting at the origin taking steps up and down by equal amounts $\ln \frac{a}{1-a}$. If the state is θ, the random walk y taking steps up with probability a and steps down with probability $1-a$. If $r > p_0$, $B > 0$. The random walk y hits the step B with probability 1. If $r < p_0$, $B < 0$. The random walk y hits the step B with probability $(\frac{1-a}{a})^{-B}$.

If the state is θ', the random walk y taking steps up with probability $1-a$ and steps down with probability a. If $r > p_0$, $B > 0$. The random walk y hits the step B with probability $(\frac{1-a}{a})^B$. If $r < p_0$, $B < 0$. The random walk y hits the step B with probability 1.

C Proof of Proposition 3

Let $B = \frac{\ln \frac{r}{1-r} - \ln \frac{p_0}{1-p_0}}{\ln \frac{a}{1-a}}$ be the steps. If the state is θ, for $p_0 < r$, the designer learns the type with probability 1. For $p_0 > r$, the designer learns the type with probability $(\frac{1-a}{a})^{-B}$. If the state is θ', for $p_0 < r$, the designer learns the type with probability $(\frac{1-a}{a})^B$. For $p_0 > r$, the designer learns the type with probability 1. The probability of learning is

$$
\begin{aligned}
\Pr(Learning) &= \mu_0(\int_0^r 1dp_0 + \int_r^1 (\frac{1-a}{a})^{-B} dp_0) + (1-\mu_0)(\int_0^r (\frac{1-a}{a})^B dp_0 + \int_r^1 1dp_0) \\
&= \mu_0(r + \frac{r}{1-r}(-1 - \ln r + r)) + (1-\mu_0)(\frac{1-r}{r}(-r - \ln(1-r)) + 1 - r)
\end{aligned}
$$

Take the first order derivative with respect to r,

$$
\begin{aligned}
\frac{d\Pr(Learning)}{dr} &= \mu_0(\frac{d}{dr}\int_0^r 1dp_0 + \frac{d}{dr}\int_r^1 (\frac{1-a}{a})^{-B} dp_0) \\
&+ (1-\mu_0)(\frac{d}{dr}\int_0^r (\frac{1-a}{a})^B dp_0 + \frac{d}{dr}\int_r^1 1dp_0) \\
&= \ln\frac{1-a}{a}\frac{dB}{dr}(-\mu_0\int_r^1 (\frac{1-a}{a})^{-B} dp_0 + (1-\mu_0)\int_0^r (\frac{1-a}{a})^B dp_0) \\
&= \ln\frac{1-a}{a}(\ln\frac{a}{1-a})^{-1}\frac{1}{r(1-r)}[-\mu_0\frac{r}{1-r}(r - 1 - \ln r) \\
&- (1-\mu_0)\frac{1-r}{r}(r + \ln(1-r))] \\
&= \mu_0\frac{1}{(1-r)^2}(r - 1 - \ln r) + (1-\mu_0)\frac{1}{r^2}(r + \ln(1-r))
\end{aligned}
$$

Take the second order derivative with respect to r,

$$
\begin{aligned}
\frac{d^2\Pr(Learning)}{dr^2} &= -\mu_0\frac{-r^2 + 2r\ln r + 1}{r(1-r)^3} - (1-\mu_0)\frac{2\ln(1-r) - \frac{(r-2)r}{1-r}}{r^3} \\
&< 0
\end{aligned}
$$

The probability of learning is concave in $r \in (0, 1)$.

Therefore, there exists a $r^* \in (0, 1)$ such that the probability of learning is maximised.

Next, show r^* is increasing in μ_0. Take the first order condition, we have

$$\mu_0 \frac{1}{(1-r^*)^2}(r^* - 1 - \ln r^*) + (1-\mu_0)\frac{1}{(r^*)^2}(r^* + \ln(1-r^*)) = 0$$

Then we have

$$(1-\mu_0)\frac{1}{(r^*)^2}(r^* + \ln(1-r^*)) = \mu_0 \frac{1}{(1-r^*)^2}(\ln r^* + 1 - r^*)$$

Therefore,

$$\frac{\mu_0 \frac{1}{(1-r^*)^2}(\ln r^* + 1 - r^*)}{(1-\mu_0)\frac{1}{(r^*)^2}(r^* + \ln(1-r^*))} = 1$$

Thus

$$\frac{(r^*)^2}{(1-r^*)^2}\frac{\ln r^* + 1 - r^*}{r^* + \ln(1-r^*)} = \frac{1-\mu_0}{\mu_0}$$

We can write r^* as a function of μ_0. The left hand side is decreasing in r^*, the right hand side is decreasing in μ_0. Therefore r^* is increasing in μ_0.

D Proof of Proposition 4

Consider the random walk y defined in Proposition 1. Let $B = \frac{\ln \frac{r}{1-r} - \ln \frac{p_0}{1-p_0}}{\ln \frac{a}{1-a}}$. Let

$$\phi_B(\tau) = \Pr(y_1 \neq B, \ldots, y_{\tau-1} \neq B, y_\tau = B)$$

be the probability that the random walk y first hits the step B at the τ-th step when $B > 0$ with generating function

$$\Phi_B(s) = \sum_{\tau=1}^{\infty} \phi_B(\tau) s^n = \left(\frac{1 - (1 - 4a(1-a)s^2)^{\frac{1}{2}}}{2(1-a)s}\right)^B. \tag{2}$$

Let

$$\psi_B(\tau) = \Pr(y_1 \neq B, \ldots, y_{\tau-1} \neq B, y_\tau = B)$$

be the probability that the random walk y first hits the step B at the τ-th step when $B < 0$ with generating function

$$\Psi_B(s) = \sum_{\tau=1}^{\infty} \psi_B(\tau)s^n = (\frac{1-(1-4a(1-a)s^2)^{\frac{1}{2}}}{2as})^{-B}. \tag{3}$$

We have $\phi_B(\tau) = \frac{\Phi_B^{(\tau)}(0)}{\tau!}$ and $\psi_B(\tau) = \frac{\Psi_B^{(\tau)}(0)}{\tau!}$ where $f^{(\tau)}(0)$ is the τ-th derivative of the function f evaluated at 0.

Assume that θ is the underlying state. The random walk y takes steps up with probability a and steps down with probability $1-a$. The generating functions can be written as (2) and (3) [1]. Then summing ϕ_B and ψ_B over τ from $\tau = 1$ to $\tau = T$ gives us the probability of hitting the step B within T periods.

Similar results hold for assuming θ' being the underlying state. Let

$$\xi_B(\tau) = \Pr(y_1 \neq B, \ldots, y_{\tau-1} \neq B, y_\tau = B)$$

be the probability that the random walk y first hits the step B at the τ-th step when $B > 0$ with generating function

$$\Xi_B(s) = \sum_{\tau=1}^{\infty} \xi_B(\tau)s^n = (\frac{1-(1-4a(1-a)s^2)^{\frac{1}{2}}}{2as})^{B}. \tag{4}$$

Let

$$\omega_B(\tau) = \Pr(y_1 \neq B, \ldots, y_{\tau-1} \neq B, y_\tau = B)$$

be the probability that the random walk y first hits the step B at the τ-th step when $B < 0$ with generating function

$$\Omega_B(s) = \sum_{\tau=1}^{\infty} \omega_B(\tau)s^n = (\frac{1-(1-4a(1-a)s^2)^{\frac{1}{2}}}{2(1-a)s})^{-B}. \tag{5}$$

Assume that θ' is the underlying state. The random walk y takes steps up with probability $1-a$ and steps down with probability a. The generating functions can be written as (4) and (5). Then summing ξ_B and ω_B over τ from $\tau = 1$ to $\tau = T$ gives us the probability of hitting the step B within T periods.

[1]See Grimmett & Stirzaker (2001)