ESSAYS ON CONSTRAINED INFORMATION ACQUISITION

by

Taishi Sassano

Department of Economics
Duke University

Date: ______________________

Approved:

____
Philipp Sadowski, Supervisor

____
Philipp Sadowski

____
Todd Sarver

____
Arjada Bardhi

____
Cosmin Ilut

Dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Economics
in the Graduate School of
Duke University

2023

ABSTRACT

ESSAYS ON CONSTRAINED INFORMATION ACQUISITION

by

Taishi Sassano

Department of Economics
Duke University

Date: ______________________________
Approved:

____
Philipp Sadowski, Supervisor

____
Philipp Sadowski

____
Todd Sarver

____
Arjada Bardhi

____
Cosmin Ilut

An abstract of a dissertation submitted in partial fulfillment of the
requirements for the degree of Doctor of Philosophy
in the Department of Economics
in the Graduate School of
Duke University

2023

Abstract

This dissertation explores the theory of constrained information acquisition and its application. I will provide theoretical models to study how economic agents who are facing limited ability to process information acquire information and how the information acquisition affects the trade mechanism in binary trades.

First, I study a dynamic information acquisition problem of an individual who faces a limited ability to process information. The individual acquires information to predict a payoff-relevant state such as the market condition in two periods. I characterize the optimal information structure and discuss how we can identify the discount factor and the information capacity. Under the optimal information acquisition, he uses three types of learning strategies depending on how certain he is about the state. In addition, I will provide two ways to identify the discount factor and the information capacity from the observed data.

Second, I will explore an application of the costly information acquisition model. I study a trade model where buyers costly acquire information about a good and a seller offers menus of an upfront fee and a strike price to screen the buyers' ability to process information in order to differentiate the buyers. I first provide the buyer's optimal information acquisition problem and then the seller's optimal selling mechanism. The willingness to pay for the strike price depends on the ability to process information. The seller offers a higher strike price to a buyer with higher ability and a higher participation fee to extract the surplus.

Acknowledgements

I would like to thank my committee for continued support and encouragement: Dr. Philipp Sadowski, my committee chair; Dr. Todd Sarver; Dr. Arjada Bardhi; and Dr. Cosmin Ilut. I offer my sincere appreciation for the learning opportunities provided by committee.

My completion of this project could not have been accomplished without the support of faculty members and cohort of economics Ph.D. program at Duke university. Especially, I am grateful have a joint work with Yonggyun Kim. It provided me a great opportunity to explore applications and implications of my project.

Finally, I wish to thank my loving and supportive parents for support and encouragement. Without the support, I could not have continued to study in this program. It was a great comfort and relief to have your encouragement.

Contents

List of Figures

Chapter 1

Introduction

Information acquisition is ubiquitous in decision making, such as investment decisions and purchasing decisions. Access to the information helps people to make better decisions. On the other hand, economic agents cannot always process all the available information. This implies they are facing a cost or constraint in processing information. For example, people usually read only main topics on the newspaper even if they can read other articles for free. This information constraint becomes more and more important because, people are now more likely to be exposed to larger amount of information, due to the recent advancement of information technologies. Hence, it is important to understand the effect of information on the behavior and welfare of economic agents.

This dissertation explores a theory of constrained information acquisition and its application. I will provide theoretical models to study how economic agents who are facing limited ability to process information acquire information and how the information acquisition affects the trade mechanism in binary trades.

In Chapter 2, I study a dynamic information acquisition problem of an individual who faces a limited ability to process information. The individual acquires information to predict a payoff-relevant state such as the market condition in two periods. I characterize the optimal information structure and discuss how we can identify the discount factor and the information capacity. Under the optimal information acquisition, he uses three types of learning strategies (gradual learning, immediate decision and mixing) depending on how certain he is about the state. In addition, I will provide two ways to identify the discount factor and the information capacity from the observed data.

Chapter 3 is based on a joint work with Yonggyun (YG) Kim. I will explore an application of the costly information acquisition model. I study a trade model where buyers costly acquire information about a good and a seller offers menus of an upfront fee and a strike price to screen the buyers' ability to process information in order to differentiate the buyers. I first provide the buyer's optimal information acquisition problem and then the seller's optimal selling mechanism. The willingness to pay for the strike price depends on the ability to process information. The seller offers a higher strike price to a buyer with higher ability and a higher participation fee to extract the surplus.

My contributions in this joint work are, (1) characterization of the buyer's optimal information acquisition, and also (2) an illustrative example that shows the seller's optimal selling scheme. The baseline model of seller's selling mechanism (which includes an upfront fee and a strike price) and numerical examples are the contributions of my coauthor, YG Kim.

Chapter 4 provides an overview of related literature. I will relate this thesis to the literature on dynamic information acquisition, rational inattention, information design and trade mechanism and information acquisition.

Lastly, Chapter 5 provides the conclusion and discussions of this dissertation.

Chapter 2

Optimal Stopping under Capacity Constrained Information Acquisition

2.1 Introduction

Information acquisition plays a central role in decision making. Richer information allows better decisions. For example, an investor acquires information about market conditions before she decides to invest. The important trade-off is that the more precise information a decision maker acquires, the longer time it takes. In a standard information acquisition problem, a decision maker (DM) must choose how much information to acquire and when to stop the information acquisition.

I study a 2-period dynamic information acquisition problem where the DM chooses how frequently and how much she receives information. In each period, the DM can choose either to predict the unknown, payoff relevant state (e.g. market condition), or to wait for more information about the state. Following the standard rational inattention (RI) model, I assume a DM faces a limitation in processing information. Specifically, I assume that the amount of expected uncertainty reduction is limited in each period. Hence, the signal must arrive less frequently if it is precise, or conversely, it must be less precise if it arrives frequently. This model can capture a trade-off between temporal discounting and the quality of choices as well as a trade-off between the frequency and accuracy of signals.

The first main result characterizes the optimal information acquisition process in this problem. The optimal signal structure gives either an informative signal which induces the DM to decide, or a less informative signal which keeps her waiting. The DM acquires information in three different ways depending on the prior: (1) always wait, (2) always decide, and (3) mix .

The optimal information acquisition process is not unique, in the sense that there can be multiple optimal solutions that differ in the likelihood of receiving confirmatory and contradictory information. However, the accuracy of information and the frequency of signal arrival are unique. In terms of observable behavior, optimal information acquisition, therefore, generates a unique prediction for the accuracy of choices and the stopping probability (or the hazard rate). Importantly, accuracy and hazard rate are the behavioral data most often observed in applications and in experimental studies, for instance in Baron et al. (2015), Johnson et al. (2016), and Bhui (2019).

I also discuss how the DM's choice varies depending on the discount rate and the information capacity. I show that a higher discount factor always results in larger incentives to wait and acquire information. In contrast, the effect of the information capacity is non-monotonic.

Given the properties of the optimal signal structure, I can demonstrate how the information capacity (and discount rate) can be identified from choice behavior. For this purpose, I consider a variation of the original stopping problem in which the researcher can choose the frequency at which the DM can make a decision. By varying the frequency, the researcher can obtain richer data from a single DM. The main identification result establishes that both the capacity and the discount rate can be identified from such data. It is also shown that the capacity can be identified from just a limited amount of data.

This paper is organized as follows. The remainder of this section reviews related literature. Section 2 describes the DM's problem of choosing information structures and actions. Section 3 presents the main results. First, I provide the characterization of the optimal signal structure and comparative statics results. Then I discuss the connection between these results and the findings in experimental studies. Section 4 discusses the identification of the discount factor and the capacity from choice data. Section 5 discusses the results and presents potential next steps. Proofs of all the results are provided in the Appendix.

Literature Review

Rational inattention theory, developed by Sims (2003) considers the behavior of information constrained agents. This theory has been developed further in recent studies (Caplin and Dean, 2015; Matějka and McKay, 2015), and Steiner et al. (2017) extended the framework to a general dynamic model. While the primary goal in these papers is to characterize optimal stochastic choice, my main objective is to characterize the optimal information acquisition process.

To understand the DM's information acquisition process, many researchers (Arrow et al., 1949; Wald, 1947) have studied *optimal stopping problems*. More recently, Fudenberg et al. (2018) investigated the optimal stopping problem under the drift diffusion model. They found that when the DM makes a decision quickly, the quality of the choice tends to be high. Their theoretical prediction has been tested by an experimental study (Bhui, 2019). While they assumed that the information structure is exogenous, recent research has studied environments where the DM can choose the signal structures under information constraints. Che and Mierendorff (2019) investigated the optimal allocation of attention over different information sources. They characterized the optimal stopping rule and how the expected waiting time and the accuracy of information change depending on the prior belief.

In contrast, by modeling the information acquisition process explicitly, Zhong (2022) studied the information acquisition problem allowing general information structures in continuous time. In his model, the individual can choose any signal process as an information source subject to an information cost. In this environment, he showed that an optimal signal process is a Poisson process, which confirms the individual's prior belief. He also described the trade-offs in his setting. In addition to a standard stopping time trade-off, the individual faces a trade off between accuracy of the signal and the frequency of receiving information. Since the individual faces an information constraint, she cannot process all the information available.

While many researchers have modeled the DM's information acquisition process in continuous time, Steiner et al. (2017) studied the stopping problem in discrete time with a different cost structure. They found that an individual optimally stops the learning at a constant hazard rate. Since their analysis is limited to the case of uniform priors, the DM's behavior remains an open question with non-uniform prior beliefs. The qualitative difference between discrete and continuous time solutions arises in the non-uniform prior case.

(De Oliveira et al., 2017) discussed the identification of information costs using a set of menu choice data. In many situations, such as the optimal stopping problem, the available data is limited. For example, Bhui (2019) conducted an experimental study to discover the relationship between speed and accuracy of decisions. In his experiment, the observable data were the actual choice and response time. In my study, I will consider the identification of the model parameter using this type of data.

2.2 Model

2.2.1 States, Actions, and Payoffs

In each period $t = 1, 2$, a DM chooses $a_t \in \{0, 1, w\}$ to predict the unknown state $\theta \in \{0, 1\}$. If she chooses either $a_t = 0$ or 1 and matches the state, she will obtain a payoff of 1, and if she fails to match the state, she will get a payoff of 0. In both cases, the game ends. The DM can also choose w to postpone the decision and collect information, which gives a flow payoff of 0. She maximizes the discounted expected payoff, that is, for a pair of actions $a = (a_1, , a_2,)$ and the true state $\theta \in \{0, 1\}$

$$U(a) = E\left[u(a^1, \theta) + \delta u(a^2, \theta)\right],$$

where δ is the discount rate and, the stage payoff is given by

$$u(a^t, \theta) = \begin{cases} 1 & \text{if } a^t = (w^{t-1}, \theta) \\ 0 & \text{if otherwise} \end{cases}.$$

2.2.2 Information Acquisition

Suppose the DM initially has a prior belief p_0. The DM can wait for information about the true state θ by designing any signal structure subject to an information constraint. Instead of analyzing the signal structure itself, I will follow the belief approach: a DM directly chooses a distribution over posteriors under two constraints, which are the Bayesian constraint and the information constraint.

The Bayesian constraint requires that the expected posterior probability equals the prior. In addition, the information constraint requires that the expected reduction of uncertainty has to be less than a fixed amount. Let $\pi(p_{t+1}|p_t)$ be the probability that the belief is updated from p_t to p_{t+1} and $\tau(p_t)$ be a probability distribution over posteriors p_{t+1} given the prior belief p_t. The DM faces the following two constraints:

1. Bayesian Constraint (BC):

$$\sum_{p'_{t+1} \in supp(\tau)} \pi(p'_{t+1}|p_t) p'_{t+1} = p_t.$$

2. Information Constraint (IC): For a capacity $\kappa > 0$,

$$I(\theta; p_{t+1}|p_t) = H(p_t) - \sum_{p'_{t+1} \in supp(\tau)} \pi(p'_{t+1}|p_t) H(p'_{t+1}) \leq \kappa.$$

   where H is an quadratic function, i.e.

$$H(p_t) = -2(p_t - 1/2)^2.$$

This quadratic function is identical to the quadratic approximation of the information entropy function up to a constant. The entropy function is commonly used to measure uncertainty in the rational inattention literature and has appealing properties. For example, when random variables are independent, the entropy of these random variables is just a sum of the entropies of the individual random variables. In the context of information acquisition, this implies that the uncertainty reduction from a sequence of random signals is equal to the sum of the reduction from individual signals.

2.3 Main Results

This section provides a characterization of the optimal distribution of posteriors and the resulting hazard rate (or equivalently, stopping probability). In this paper, I will say 'accuracy of a posterior' to describe how certain a DM is about the state after receiving a signal (more concretely, the distance between the posterior and 1/2).

2.3.1 Optimal Signal

$t = 2$ (Terminal Period)

The DM has no incentive to wait. Since every deciding posterior has the same accuracy, the optimal information structure includes two posteriors that are located symmetrically around $p = 1/2$. The accuracy of the optimal posteriors can be solved from the information constraint:

$$p^2 = V^2(p) = \min\{\frac{1}{2} + \sqrt{\frac{\kappa}{2} + \left(p - \frac{1}{2}\right)^2}, 1\} \tag{2.1}$$

Since the DM decides for sure, this is the optimal value at period T.

$t = 1$ (Initial Period)

The DM's problem is formulated as follows:

$$\begin{aligned} V^1(p) &= \max_{a \in \{0,1,w\}, \tau} E\left[u(a,\theta) + \delta V^2(p')\right] \\ &= \max_{\tau} E_\tau \left[\max\{p', 1-p', \delta V^2(p')\}\right], \end{aligned}$$

subject to (BC) and (IC). Here, τ denotes a distribution over posteriors and $p' \in supp(\tau)$ is a posterior.

Before moving to the main results, it is useful to consider a graphical illustration of this problem. The following lemma provides two properties which help with that.

Lemma 1. The following two properties hold.

1. The set of beliefs $[0, 1]$ is partitioned into three convex sets: there exists $p^* \in [1/2, 1]$ such that

$$[0, 1-p^*) = \{p | 1-p = \max\{p, 1-p, \delta V(p)\}\},$$

$$[1-p^*, p^*) = \{p | \delta V(p) = \max\{p, 1-p, \delta V(p)\}\},$$

$$[p^*, 1] = \{p | p = \max\{p, 1-p, \delta V(p)\}\}.$$

2. There exists $p^{**} \in [1/2, 1]$ such that

   - $V(p) = \delta$ for $p \in [0, 1-p^{**}] \cup [p^{**}, 1]$,
   - $V(p)$ is convex in $(1-p^{**}, p^{**})$.

Proof. See Appendix.

The first property says that the set of beliefs can be divided into three convex parts: those in the left induces the DM to choose $a_t = 0$, those in the right induces her to choose $a_t = 1$ and those in the middle induces her to wait. The reason why each part is convex is that the slope of the value function V is bounded above by one, while that of the expected payoff from a decision is always one. These properties ensures that the continuation value curve and the curve of expected payoff from a decision intersect at most once. The second property says that the continuation value is a convex function for intermediate beliefs and is a constant for extreme beliefs. Note that the entropy function H is concave and thus $-H$, which appears on the left hand side of the information constraint, is convex. This convexity is inherited by the value function. If the prior is sufficiently informative, then remaining uncertainty is so small that the DM can choose a fully informative signal and obtain payoff 1 in the next period. Hence, the continuation value is constant, δ, for extreme beliefs.

Figure 2.1 illustrates the results above. The solid curve illustrates the expected payoff from an optimal action for each belief p. When the DM is sure about the state (i.e. $p \in [0, 1-p^*] \cup [p^*, 1]$), she will stop and decide. Otherwise, she will choose to wait for additional information. The DM's problem is to choose an optimal distribution over the optimal payoffs subject to the information constraint. This is closely related to the information design problem (Kamenica and Gentzkow, 2011).

Parameter Restriction

Let $\tilde{p}$ be the smallest prior accuracy that allows the DM to choose a fully informative signal. This can be derived from IC:

$$H(\tilde{p}) - H(1) = \kappa$$

$$-2(\tilde{p} - 1/2)^2 + 2(1 - 1/2)^2 = \kappa$$

$$\tilde{p} = 1/2 + \sqrt{\frac{1-2\kappa}{4}}$$

If $\tilde{p} \geq \delta$, then exhausting IC is not unique optimum. For example, at prior $\tilde{p} - \varepsilon$, it is sufficient to jump to $\tilde{p}$ since the DM can choose a fully informative signal in the next period

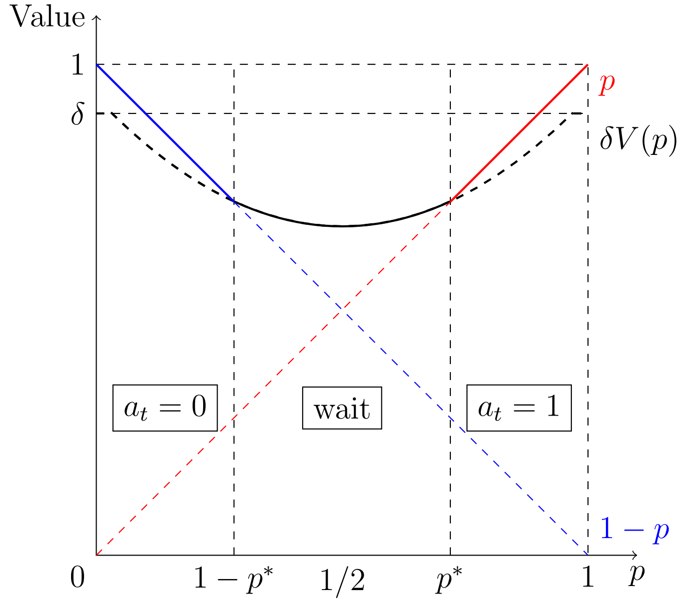

Figure 2.1: Expected payoffs from optimal actions

(see Figure 2.2). I will assume $\tilde{p} \leq \delta$ to simplify the analysis (Figure 2.3.1).

Next lamma shows that if this assumption holds for some parameters (δ, κ), then it still holds as the period length shrinks. This property guarantees that the same analysis holds if the period length changes.

Lemma 2. If $\tilde{p} \geq \delta$ holds for some parameters (δ, κ), then it still holds as the period length shrinks.

Proof: See Appendix.

Proposition 1 characterizes the optimal signal structure in this problem. Let $p_d^*(p_w^*)$ denote the optimal posterior accuracy that induces the DM to decide (wait) and τ_d^* denote the probability of deciding.

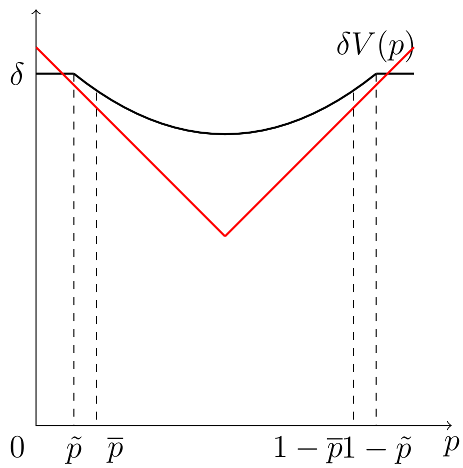

Figure 2.2: $\tilde{p} \leq \delta$

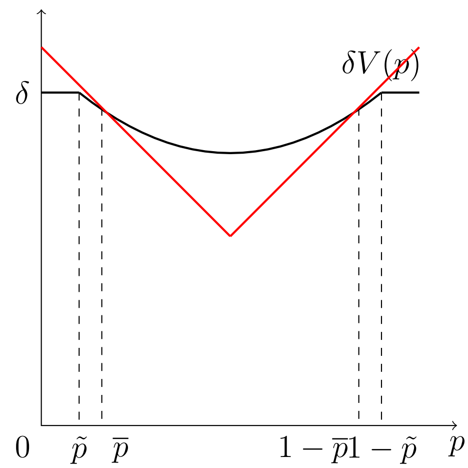

Figure 2.3: $\tilde{p} \geq \delta$

Proposition 1. Define

$$g := \frac{-1 + \sqrt{1 + 2\kappa\frac{1+\delta}{1-\delta}}}{2(1+\delta)}.$$

There are two thresholds

$$\overline{p} = \frac{1}{2} + \sqrt{g^2 - \frac{\kappa}{2}}$$

$$\underline{p} = \frac{1}{2} + \delta^2\sqrt{g^2 - \kappa}$$

such that the optimal distribution over posteriors is described as follows:

1. If $|p - 1/2| \geq |\overline{p} - 1/2|$, then the DM stops for sure, i.e. the there are two posteriors in the support:

$$(p_d^* - \frac{1}{2})^2 = \frac{\kappa}{2} + \left(p - \frac{1}{2}\right)^2. \tag{2.2}$$

2. If $|\overline{p} - 1/2| \geq |p - 1/2| \geq |\underline{p} - 1/2|$, then the DM mixes, i.e. there are four posteriors in the support:

$$(p_d^* - 1/2)^2 = g^2 \tag{2.3}$$

$$(p_w^* - 1/2)^2 = -\kappa/2 + \delta^2 g^2 \tag{2.4}$$

$$\tau_d^* = \frac{\kappa + (p - 1/2)^2 - \delta^2 g^2}{(1 - \delta^2)g^2 + \kappa/2} \tag{2.5}$$

3. If $|\underline{p} - 1/2| \geq |p - 1/2|$, then he DM waits for sure, i.e. there are two posteriors in the support:

$$(p_w^* - \frac{1}{2})^2 = \frac{\kappa}{2} + \left(p - \frac{1}{2}\right)^2 . \tag{2.6}$$

Proof. See Appendix.

Proposition 1 characterizes the optimal information structure. Depending on the prior, the optimal information structure changes in three phases. When the DM is not sure about the state($|p - 1/2| \leq |\underline{p} - 1/2|$), he will wait for additional information with probability 1 (gradual learning). When the DM is sure about the state($|p - 1/2| \geq |\overline{p} - 1/2|$), he will decide with probability 1 (immediate decision). At the intermediate prior, he will mix the gradual learning and the immediate decision. The movement of the belief can be illustrated as follows.

Case 1 : Gradual learning ($|p - 1/2| \leq |\underline{p} - 1/2|$)

There are two possible posteriors $p_w, 1 - p_w$ that are located symmetrically around 1/2. Both of these two posterior induce the DM to wait for additional information.

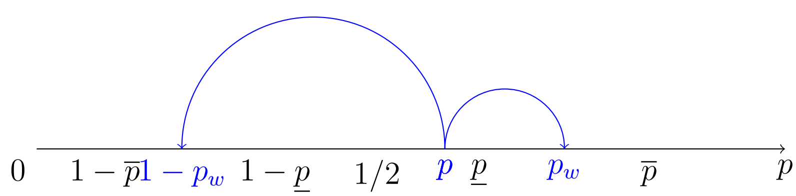

Figure 2.4: Gradual Learning

Case 2 : Immediate decision ($|p - 1/2| \geq |\overline{p} - 1/2|$)

There are two possible posteriors $p_d, 1 - p_d$ that are located symmetrically around 1/2. Both of these two posterior induce the DM to decide.

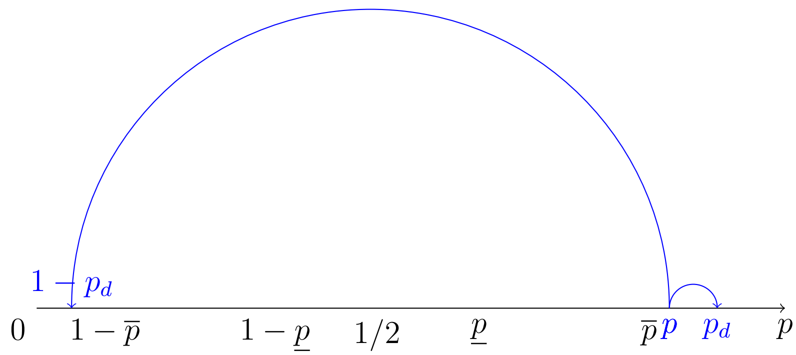

Figure 2.5: Immediate Decision

Case 3 : Mix ($|\overline{p}-1/2| \geq |p-1/2| \geq |\underline{p}-1/2|$)

There are four possible posteriors $p_w, 1-p_w, p_d, 1-p_d$ that are located symmetrically around 1/2. p_w and $1-p_w$ induce the DM to wait for additional information and p_d and $1-p_d$ induce the DM to decide. This case can be split into two subcases:

Case 3-1 The waiting posterior is more accurate than the prior ($|p_w^* - 1/2| \geq |p-1/2|$)

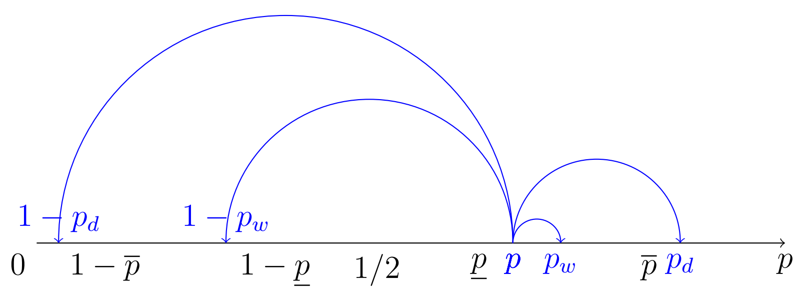

Figure 2.6: Mix: Case 1

Case 3-2 The waiting posterior is less accurate than the prior ($|p_w^* - 1/2| \leq |p-1/2|$)

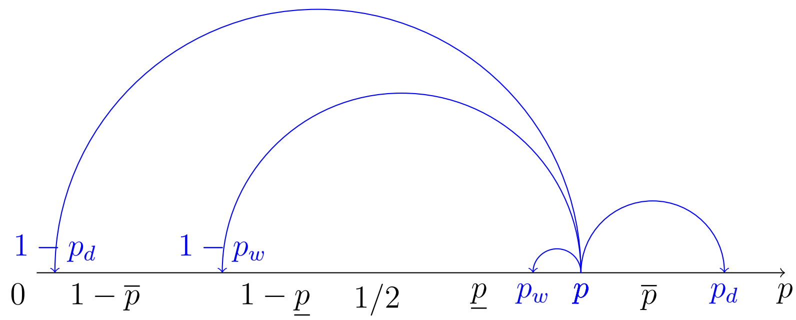

Figure 2.7: Mix: Case 2

As in a result for the continuous time case (Zhong, 2022), the optimal signal has a Poisson-type information structure in the mixing region. That is, the DM receives either an accurate signal which allows her to decide, or an less accurate signal which keeps her waiting for additional information. However, while signals give only confirmatory information in Zhong's solution, the informative signal can be contradictory in this discrete setting. The contradictory information also exists in other two cases (the case of gradual learning and immediate decision).In many situations, such as investment decisions, it seems plausible that the DM will search both confirmatory and contradictory information to make better decisions.

Proof Sketch

The DM's problem can be solved using Lagrangian:

$$\mathcal{L} = \sum_{i=1}^{n} \tau_i \max\{p_i, 1-p_i, \delta V_2(p_i)\} + \lambda \left(\kappa - \left(H(p) - \sum_i \tau_i H(p_i)\right)\right)$$

$$+ \mu \left(p - \sum_i \tau_i p_i\right) + \nu \left(1 - \sum_i \tau_i\right)$$

$$= \sum_{i=1}^{n} \tau_i \left(\max\{p_i, 1-p_i, \delta V_2(p_i)\} + \lambda H(p_i) - \mu p_i - \nu\right) + \lambda(\kappa - H(p)) + \mu p + \nu.$$

Although the DM actually faces the information constraint, we can view this problem as a costly information acquisition problem for a fixed multiplier of the information constraint λ. Hence, I will first solve this problem for each value of λ. In the proof, I will solve relaxed problem where the DM faces only the information constraint and show that the solution to the relaxed problem actually satisfies the Bayesian plausibility. Then, each problem can be solved by concavifying two functions: the expected value from deciding $v_d(p)$and that from waiting. $v_w(p)$, where

$$v_d(p) := \max\{p, 1-p\} + \lambda H(p),$$

$$v_w(p) := \delta V_2(p) + \lambda H(p).$$

These two values are illustrated in Figure 2.8,2.9 and 2.10. First, note that the value of λ is determined by the prior belief for a fixed parameters (δ, κ). By the concavity of the information measure, the shadow value λ of the information capacity is higher when the prior is less accurate. It can be shown that λ is decreasing in the prior accuracy in optimal. Figure 2.8 shows the maximum value from deciding is higher than that from waiting when λ is low. In other words, when the DM is sure about the state, then the optimal information strategy is immediate decision. In contrast, Figure 2.9 illustrates that the value from waiting is higher when the λ is low (or, equivalently the DM is unsure about the state). Figure 2.10 illustrates the threshold case where the maximum value from each action is the same. In this case, the DM will mix the strategies so that the information constraint and Bayesian plausibility are both satisfied.

Next, for the optimal information structure given λ, I will derive prior beliefs that support λ. In the waiting region and deciding region, λ is decreasing in the prior accuracy since the marginal cost of posterior accuracy becomes larger as the prior becomes more accurate. Hence the DM will wait for sure when the prior is close to 1/2 and decide for sure when the prior is far from 1/2. However, it is not possible to directly switch the strategy from gradual learning to immediate decision. If the DM switches the strategy at the threshold λ, then the set of possible posteriors changes from $\{p_w, 1-p_w\}$ to $\{p_d, 1-p_d\}$. (See Figure 2.10) Since the information constraint must be exhausted in optimal, this jump cannot be feasible. For this reason, there is a set of prior beliefs that support the same threshold λ, which I call "mixing region". This region is necessary for the smooth transition from the gradual learning to the immediate decision.

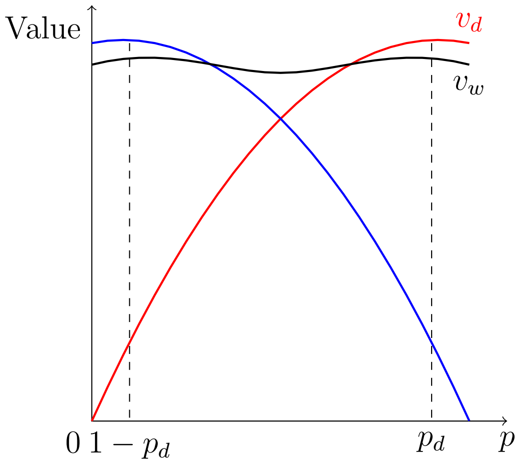

Figure 2.8: Low λ

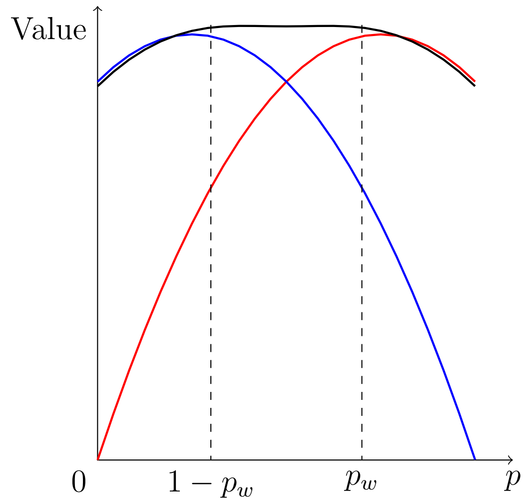

Figure 2.9: High λ

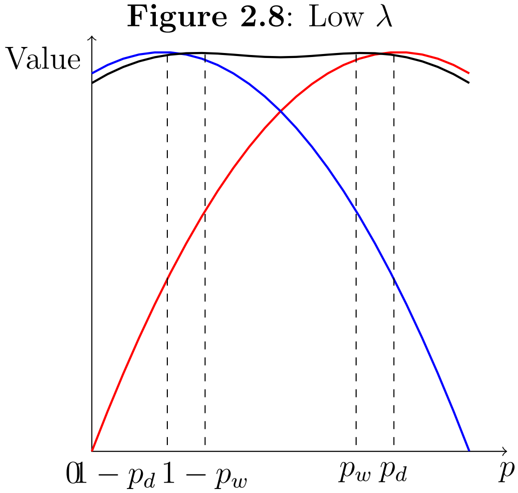

Figure 2.10: Threshold λ

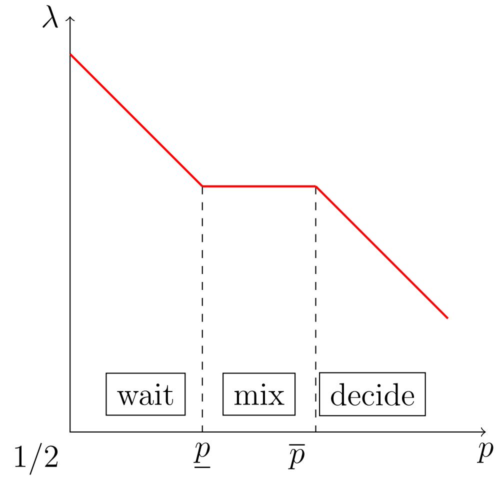

Figure 2.11: Relationship between p and λ

The next proposition describes in detail properties of the optimal information structure given in Proposition 1.

Proposition 2. Threshold beliefs $\overline{p}$ and $\underline{p}$ are increasing in δ and κ.

Proof. See Appendix.

Proposition 2 states that (1) the waiting region expands and the deciding region shrinks as the discount factor increases (2) the two regions changes non-monotonically in the capacity κ.

Indeterminacy

Although Proposition 1 describes the shape of the solution, we still have indeterminacy in mixing region. Given an optimal signal structure, we can move the weight from one accurate posterior to the other accurate posterior and also from a less accurate posterior to the other less accurate posterior so that Bayesian plausibility condition holds. Since the accuracy of signals and the stopping probability does not change, the information constraint is still satisfied, and the expected payoff is still the same. Hence, the new signal structure is also optimal, and this implies that the optimal solution is not unique.

Nonetheless, the accuracy and the stopping probability are unique. These two types of data are what researchers observe in many experimental studies and they play an important role in later chapters.

2.4 Identification

In this section, I will demonstrate how the parameters (δ, κ) in the optimal stopping problem can be identified from choice data (i.e. expected waiting time and choice accuracy). In the previous setting, the parameters (δ, κ) cannot be identified from the choice data considered so far since multiple pairs of parameters can generate the same choice. For this reason, I now discuss additional assumption and a richer set of choice data to provide two possible ways to identify the parameters.

2.4.1 Assumption on the Prior Belief

Without any assumption, a researcher can observe two types of data (i.e. expected waiting time and choice accuracy) and there are three unknowns (δ, κ and the prior p). This implies the parameters cannot be identified from the choice data. For this reason, I will assume that the subject's prior is $p = 1/2$. This seems natural in the context of experimental studies where the subjects is asked to do simple tasks to make decisions (e.g. a random dot motion discrimination task; Bhui (2019)). With the assumption on the prior p, we have two unknowns and two equations to identify the parameters.

Proposition 3. Assume the prior belief is $p = 1/2$. The discount factor and the capacity (δ, κ) can be identified from the choice accuracy and the waiting time.

2.4.2 Varying Period Length

In the previous section, the DM is assumed to make a decision in each unit of time. Another way to identify the parameters is to consider situation where the researcher can control the frequency of decision making. That is, if a researcher sets Δ as a period length, then the DM can make a decision in each Δ unit of time. In that case, the DM can process $\Delta \cdot \kappa$ unit of information in each period, and the discount rate is δ^{Δ}. By changing the period length $\Delta > 0$, the researcher can observe choice data generated by $(\delta^{\Delta}, \kappa \cdot \Delta)$ from a single DM in an experiment.

I assume that the DM's prior belief satisfies $|p - 1/2| < |\overline{p} - 1/2|$. This assumption implies that the choice accuracy is determined by the parameters (δ, κ).

Proposition 4. Suppose the choice accuracy is observed for period lengths Δ_1, Δ_2. Then the discount factor and the capacity (δ, κ) can be identified.

Starting waiting or mixing region, the resulting choice accuracy is uniquely pinned down given parameters. If we observe choice accuracy for two distinct period length, then we have two unknown and two equations to identify the parameters.

Remark. Unlike the information entropy function, the quadratic function does not have the linearity. Under entropy function, the total amount of information from independent signals is just a sum of the amount of information from individual signals. Since the quadratic function does not have this property, varying the period length might affect the optimal solution. However, since my quadratic function is a quadratic approximation of the entropy function, they share similar properties such as concavity. Hence, the optimal solutions under the two cost function should not drastically different.

2.5 Conclusion

This study investigates the optimal information acquisition process in discrete time when a DM can design the information structure subject to an information capacity constraint.

The main finding is that the DM uses three ways of learning (always stop, always wait and mix) depending on the prior. This optimal solution is qualitatively different from the continuous-time solution in an existing study. First, contradictory information is always received with positive probability. Second, if initial prior is uncertain, then the resulting choice accuracy can be uniquely pinned down from the parameters.

In addition, the parameters can be identified from the observed choice. Assuming the initial prior is sufficiently uncertain, we can identify the parameters by observing the choice accuracy for two distinct period lengths.

2.6 Appendix

Lemma A1.

$$g := \frac{-1+\sqrt{1+2\kappa\frac{1+\delta}{1-\delta}}}{2(1+\delta)} = \frac{-1+\sqrt{1-2\kappa+\frac{4\kappa}{1-\delta}}}{2(1+\delta)}$$

is increasing in δ.

Proof. Taking the derivative, we have

$$\frac{dg}{d\delta} = \frac{\frac{\frac{4\kappa}{(1-\delta)^2}}{2\sqrt{1-2\kappa+\frac{4\kappa}{1-\delta}}}(1+\delta) - (-1+\sqrt{1-2\kappa+\frac{4\kappa}{1-\delta}})}{2(1+\delta)^2}$$

$$= \frac{\frac{\frac{4\kappa(1+\delta)}{(1-\delta)^2} - 2(1-2\kappa+\frac{4\kappa}{1-\delta})}{2\sqrt{1-2\kappa+\frac{4\kappa}{1-\delta}}} + 1}{2(1+\delta)^2}$$

$$= \frac{\frac{\frac{4\kappa\{(1+\delta)-2(1-\delta)\}}{(1-\delta)^2} - 2(1-2\kappa)}{2\sqrt{1-2\kappa+\frac{4\kappa}{1-\delta}}} + 1}{2(1+\delta)^2}$$

$$> \frac{\frac{-2}{2\sqrt{1-2\kappa+\frac{4\kappa}{1-\delta}}} + 1}{2(1+\delta)^2}$$

$$> \frac{-1+1}{2(1+\delta)^2}$$

$$= 0$$

Hence, it is shown that the deriative of g with respect to δ is positive. ■

Proof of Lemma 2

Let $\Delta > 0$ be a period length and define

$$\delta_\Delta := \delta^\Delta, \quad \text{and} \quad \kappa_\Delta := \Delta\kappa.$$

It is sufficient to show that

$$\tilde{p} = 1/2 + \sqrt{\frac{1-2\Delta\kappa}{4}} \geq \delta^\Delta \quad \forall \Delta \leq 1.$$

Taking derivative of each side with respect to δ, we have

$$1/2(1/4-\Delta\kappa/2)^{-1/2}(-\kappa/2) \;\text{ and }\; \delta^{\Delta}\ln\delta.$$

Note that the derivatives of both $\tilde{p}$ and δ^{Δ} are 0 at $\Delta=0$. and negative for all $\Delta\leq 1$. Since the derivative of $\tilde{p}$ is concave and that of δ^{Δ} is convex in Δ, the former is greater than the latter. ■

Proof of Proposition 1.

$$\begin{aligned}\mathcal{L} &= \sum_{i=1}^{n}\tau_i\max\{p_i,1-p_i,\delta V_2(p_i)\}+\lambda\left(\kappa-\left(H(p)-\sum_i\tau_iH(p_i)\right)\right)\\ &+\mu\left(p-\sum_i\tau_ip_i\right)+\nu\left(1-\sum_i\tau_i\right)\\ &=\sum_{i=1}^{n}\tau_i\left(\max\{p_i,1-p_i,\delta V_2(p_i)\}+\lambda H(p_i)-\mu p_i-\nu\right)+\lambda(\kappa-H(p))+\mu p+\nu.\end{aligned}$$

FOC:

$$(a_t=1):\tau_1(1+\lambda H'(p_1)-\mu)=0, \tag{2.7}$$

$$(a_t=0):\tau_0(-1+\lambda H'(p_0)-\mu)=0, \tag{2.8}$$

$$(a_t=w):\tau_w(\delta V_T'(p_w)+\lambda H'(p_w)-\mu)=0, \tag{2.9}$$

$$(\tau_1):p_1+\lambda H(p_1)-\mu p_1-\nu\leq 0 \text{ with eq. if } \tau_1>0, \tag{2.10}$$

$$(\tau_0):1-p_0+\lambda H(p_0)-\mu p_0-\nu\leq 0 \text{ with eq. if } \tau_0>0, \tag{2.11}$$

$$(\tau_w):\delta V_T(p_w)+\lambda H(p_w)-\mu p_w-\nu\leq 0 \text{ with eq. if } \tau_w>0. \tag{2.12}$$

$$\tag{2.13}$$

1. Case 1: $\tau_w=0$. (Immediate decision, symmetric solution $\mu=0$)

   - Case 1-1: Fully informative sigal is available.
     $p_1=1, p_0=0$.

- Case 1-2: Fully informative sigal is not available.

  The problem is the same as the one in period $t = T$. The solution is

$$p^*_{T-1} = \min\{\frac{1}{2} + \sqrt{\frac{\kappa}{2} + \left(p - \frac{1}{2}\right)^2}, 1\}.$$

$\tau_w = 0$ is optimal if the value of deciding is (weakly) greater than that of waiting, i.e.

$$\delta V_T(p_w) + \lambda H(p_w) \leq \nu = p_1 + \lambda H(p_1).$$

Let $\beta(p) := \sqrt{\kappa/2 + (p - 1/2)^2}$.

(a) FOC

$$\begin{aligned}
\delta V'_T(p) + \lambda H'(p) &= \delta \frac{p - 1/2}{\beta(p)} - 4\lambda(p - 1/2) \\
&= (p - 1/2)\left(\frac{\delta}{\beta(p)} - 4\lambda\right)
\end{aligned}$$

The solution is $p^* = 1/2$ or,

$$\begin{aligned}
\frac{\delta}{\beta(p^*)} &= 4\lambda \\
\beta(p^*) &= \frac{\delta}{4\lambda} \\
\sqrt{\kappa/2 + (p^* - 1/2)^2} &= \frac{\delta}{4\lambda} \\
(p^* - 1/2)^2 &= \frac{\delta^2}{16\lambda^2} - \kappa/2
\end{aligned}$$

(b) Check the optimality condition.

Here,

$$\begin{aligned}
\delta V_T(p^*) + \lambda H(p^*) &= \frac{\delta}{2} + \lambda/2 + \frac{\delta^2}{8\lambda} + \kappa\lambda \\
p_1 + \lambda H(p_1) &= \frac{1}{4\lambda} + 1/2 + \lambda\left(1/2 - 2(\frac{1}{4\lambda} + 1/2 - 1/2)^2\right) \\
&= \frac{1}{8\lambda} + 1/2 + \lambda/2
\end{aligned}$$

Hence, we have

$$\delta V_T(p_w) + \lambda H(p_w) \leq \nu = p_1 + \lambda H(p_1)$$

$$\frac{\delta}{2} + \lambda/2 + \frac{\delta^2}{8\lambda} + \kappa\lambda \leq \frac{1}{4\lambda} + 1/2 + \lambda/2 - \frac{1}{8\lambda}$$

$$\frac{\delta}{2}\lambda + \frac{\delta^2}{8} + \kappa\lambda^2 \leq \frac{1}{4} + \lambda/2 - \frac{1}{8}$$

$$\lambda^2 - \frac{1-\delta}{2\kappa}\lambda - \frac{1-\delta^2}{8\kappa} \leq 0$$

This inequality shows the range of λ where the DM will wait for sure. In terms of prior p,

$$(\frac{1}{4\beta(p)})^2 - \frac{1-\delta}{2\kappa}\frac{1}{4\beta(p)} - \frac{1-\delta^2}{8\kappa} \leq 0$$

$$(1-\delta^2)\beta^2(p) + (1-\delta)\beta(p) - \kappa/2 \geq 0.$$

Since $\beta(p) \geq 0$,

$$\beta(p) \geq \frac{-(1-\delta) + \sqrt{(1-\delta)^2 + 2\kappa(1-\delta^2)}}{2(1-\delta^2)}$$

$$\sqrt{\kappa/2 + (p-1/2)^2} \geq \frac{-(1-\delta) + \sqrt{(1-\delta)^2 + 2\kappa(1-\delta^2)}}{2(1-\delta^2)}$$

$$(p-1/2)^2 \geq \left(\frac{-(1-\delta) + \sqrt{(1-\delta)^2 + 2\kappa(1-\delta^2)}}{2(1-\delta^2)}\right)^2 - \frac{\kappa}{2}.$$

Let $\overline{p}$ be this threshold belief. The right hand side is non-monotonic in κ and increasing in δ.

2. Case 2: $\tau_1 = \tau_0 = 0$ (Gradual learning).

- Optimal gradual learning strategy:

$$\max_{\tau_i,p_i}\sum_{i=1}^{n}\tau_i\delta V_T(p_w^i)$$

$$s.t. H(p)-\sum_i \tau_i H(p_w^i)\le\kappa$$

$$\sum_i \tau_w^i p_w^i = p$$

$$0\le \tau_w^i\le 1\ \ \forall i,\ \sum_i \tau_w^i = 1$$

$$0\le p_w^i\le 1\ \ \forall i.$$

- FOC:

$$\tau_w^i\left(\delta V_T'(p_w^i)+\lambda H'(p_w^i)-\mu\right)=0. \tag{2.14}$$

$$\left(\frac{\delta}{\beta(p_w^i)}-4\lambda\right)(p_w^i-1/2)=\mu. \tag{2.15}$$

- I will find a solution with $\mu = 0$ and confirm that it satisfies the Bayesian plausibility.

$$p_w=1/2,\ (p_w-1/2)^2=\frac{\delta^2}{16\lambda^2}-\frac{\kappa}{2}.$$

3. SOC

$$\frac{\partial^2\mathcal{L}}{\partial(p_w^i)^2}(p_w^i)=\delta\frac{\beta(p_w^i)-(p_w^i-1/2)\beta'(p_w^i)}{\beta(p_w^i)^2}-4\lambda \tag{2.16}$$

$$\frac{\partial^2\mathcal{L}}{\partial(p_w^i)^2}(1/2)=\delta\sqrt{\frac{2}{\kappa}}-4\lambda \tag{2.17}$$

$$\begin{cases}>0 & (\lambda<\frac{\delta}{4}\sqrt{\frac{2}{\kappa}})\\ <0 & (\lambda>\frac{\delta}{4}\sqrt{\frac{2}{\kappa}}).\end{cases} \tag{2.18}$$

$$\frac{\partial^2\mathcal{L}}{\partial(p_w^i)^2}(p_w^{*i})=\delta\frac{\frac{\delta}{4\lambda}-(\frac{\delta^2}{16\lambda^2}-\frac{\kappa}{2})\frac{4\lambda}{\delta}}{\frac{\delta^2}{16\lambda^2}}-4\lambda \tag{2.19}$$

$$\begin{cases}>0 & (\lambda>\frac{\sqrt{\delta^3}}{4}\sqrt{\frac{2}{\kappa}})\\ <0 & (\lambda<\frac{\sqrt{\delta^3}}{4}\sqrt{\frac{2}{\kappa}}).\end{cases} \tag{2.20}$$

Since

$$\frac{\delta}{4}\sqrt{\frac{2}{\kappa}} > \frac{\sqrt{\delta^3}}{4}\sqrt{\frac{2}{\kappa}},$$

$p_w^i = 1/2$ and $(p_w - 1/2)^2 = \frac{\delta^2}{16\lambda^2} - \frac{\kappa}{2}$ can not be optimal at the same time. In addition, if $p_w^i = 1/2$ is received with probability 1, then the information constraint can not be exhausted. Hence the only solution candidate is the symmetric two-sided jump ($(p_w - 1/2)^2 = \frac{\delta^2}{16\lambda^2} - \frac{\kappa}{2}$) within the waiting region. If this is true, then it must exhaust the information capacity:

$$H(p) - H(p_w^i) = \kappa$$

$$-2(p-1/2)^2 + 2(\frac{\delta^2}{16\lambda^2} - \frac{\kappa}{2}) = \kappa$$

$$\frac{1}{\lambda^2} = 16\frac{(p-1/2)^2+\kappa}{\delta^2}$$

$$\lambda = \frac{\delta}{4\sqrt{(p-1/2)^2+\kappa}} = \frac{\delta}{4\gamma(p)}$$

Hence,the optimal waiting posterior is given by

$$(p_w^i - 1/2)^2 = \frac{\delta^2}{16}\cdot 16\frac{(p-1/2)^2+\kappa}{\delta^2} - \frac{\kappa}{2}$$
$$= \frac{\kappa}{2} + (p-1/2)^2$$

It is clear that $p < p_w^i$. The optimal payoff from gradual learning is

$$\delta V_T(p_w^i) + \lambda H(p_w^i)$$
$$= \delta\left((\frac{1}{2} + \sqrt{k/2 + (p_w^i - \frac{1}{2})^2}\right) + \frac{\delta}{4\sqrt{(p-\frac{1}{2})^2+\kappa}}\left((\frac{1}{2} - 2(p_w^i - \frac{1}{2})^2\right)$$
$$= \frac{\delta}{2} + (\delta^{1/2} - \frac{1}{2})\sqrt{(p-1/2)^2+\kappa} + \delta\frac{\kappa+1/2}{4\sqrt{(p-1/2)^2+\kappa}}.$$

4. Condition for the optimality of gradual learning

$$p_1 + \lambda H(p_1) \leq \nu \leq \delta V_T(p_w^i) + \lambda H(p_w^i).$$

For $(p_w - 1/2)^2 = \frac{\delta^2}{16\lambda^2} - \frac{\kappa}{2}$,

$$1/2 + 1/4\lambda + \lambda/2 - 1/\lambda\delta \leq \delta/2 + \frac{\delta^2}{8\lambda} + \lambda/2 + \kappa\lambda.$$

$$\lambda^2 - \frac{1-\delta}{2\kappa}\lambda - \frac{1-\delta^2}{8\kappa} \geq 0.$$

In terms of prior p,

$$(\frac{\delta}{4\gamma(p)})^2 - \frac{1-\delta}{2\kappa}\frac{1}{4\gamma(p)} - \frac{1-\delta^2}{8\kappa} \geq 0$$

$$(1-\delta^2)\gamma^2(p) + \delta(1-\delta)\gamma(p) - \frac{\delta^2\kappa}{2} \leq 0.$$

Since $\gamma(p) \geq 0$,

$$0 \leq \gamma(p) \leq \frac{-\delta(1-\delta) + \sqrt{\delta^2(1-\delta)^2 + 2\kappa\delta^2(1-\delta^2)}}{2(1-\delta^2)}$$

$$0 \leq \sqrt{\kappa + (p-1/2)^2} \leq \frac{-\delta(1-\delta) + \delta\sqrt{(1-\delta)^2 + 2\kappa(1-\delta^2)}}{2(1-\delta^2)}$$

$$(p-1/2)^2 \leq \delta^2\left(\frac{-(1-\delta) + \sqrt{(1-\delta)^2 + 2\kappa(1-\delta^2)}}{2(1-\delta^2)}\right)^2 - \kappa.$$

Let $\underline{p}$ be the threshold belief determined by the inequality above.

5. Case 3; $\tau_0, \tau_1, \tau_w > 0$. In this case, the solution is symmetric around half, which implies $\mu = 0$. FOCs are reduced to

$$(a_t = 1) : 1 + \lambda H'(p_1) = 0, \tag{2.21}$$

$$(a_t = 0) : -1 + \lambda H'(p_0) = 0, \tag{2.22}$$

$$(a_t = w) : \delta V_T'(p_w) + \lambda H'(p_w) = 0, \tag{2.23}$$

$$(\tau_1) : p_1 + \lambda H(p_1) - \nu = 0, \tag{2.24}$$

$$(\tau_0) : 1 - p_0 + \lambda H(p_0) - \nu = 0, \tag{2.25}$$

$$(\tau_w) : \delta V_T(p_w) + \lambda H(p_w) - \nu = 0. \tag{2.26}$$

From the F.O.Cs, we can solve p_1, p_0, p_w as a function of λ.

$$(a_t = 1) : 1/\lambda = -H'(p_1) = 4p_1 - 2, \tag{2.27}$$

$$(a_t = 0) : 1/\lambda = H'(p_0) = -4p_0 + 2, \tag{2.28}$$

$$(a_t = w) : 1/\lambda = -\frac{H'(p_w)}{\delta V_T'(p_w)} = 4/\delta\sqrt{\kappa/2 + (p_w - 1/2)^2}, \tag{2.29}$$

$$\implies (p_w - 1/2)^2 = \left(\frac{\delta}{4\lambda}\right)^2 - \kappa/2. \tag{2.30}$$

$$(\tau_1, \tau_w) : \delta V_T(p_w) + \lambda H(p_w) = p_1 + \lambda H(p_1) \tag{2.31}$$

$$\implies \delta\left(\frac{1}{2} + \sqrt{\frac{\kappa}{2} + (p_w - 1/2)^2}\right) + \lambda\left(1/2 - 2(p_w - 1/2)^2\right) \tag{2.32}$$

$$= p_1 + \lambda\left(1/2 - 2(p_1 - 1/2)^2\right) \tag{2.33}$$

$$\frac{\delta}{2} + \frac{\delta^2}{8\lambda} + (1/2 + \kappa)\lambda = \frac{1}{8\lambda} + \frac{\lambda}{2} + \frac{1}{2}. \tag{2.34}$$

$$\kappa\lambda - \frac{1-\delta}{2}\lambda - \frac{1-\delta^2}{8} = 0. \tag{2.35}$$

Since λ is independent of p. ,all the posteriors $(p_1, 1-\lambda, p_w, 1-p_w)$ in the support must be independent of λ. This implies that at any prior in the "mix" region, all the postriors stay constant. The DM changes the weights to make IC and BC satisfied. In the mix region,

$$(p_1 - 1/2)^2 = \left(\frac{-(1-\delta) + \sqrt{(1-\delta)^2 + 2\kappa(1-\delta^2)}}{2(1-\delta^2)}\right)^2 \tag{2.36}$$

$$(p_w - 1/2)^2 = -\kappa/2 + \delta^2\left(\frac{-(1-\delta) + \sqrt{(1-\delta)^2 + 2\kappa(1-\delta^2)}}{2(1-\delta^2)}\right)^2 \tag{2.37}$$

$$\tau_1 = \frac{\kappa + (p - 1/2)^2 - \delta^2\left(\frac{-(1-\delta)+\sqrt{(1-\delta)^2+2\kappa(1-\delta^2)}}{2(1-\delta^2)}\right)^2}{(1-\delta^2)\left(\frac{-(1-\delta)+\sqrt{(1-\delta)^2+2\kappa(1-\delta^2)}}{2(1-\delta^2)}\right)^2 + \kappa/2} \tag{2.38}$$

Feasibility

- In case 1 and case 2, IC and BC are clearly satisfied.

- In case 3, it is clear that IC and BC are satisfied as long as all the posteriors are more accurate than the prior. However, it is also possible that the waiting posterior is less accurate than the prior. In that case, it is sufficient to show that

$$\tau_1 p_1 + (1 - \tau_1) p_w \geq p,$$

or equivalently,

$$g(p) := \tau_1 p_1 + (1 - \tau_1) p_w - p$$

is decreasing in p, since the above inequality is satisfied in the largest belief $\overline{p}$ in mixing region.

$$\begin{aligned}
\frac{d}{dp}\left(\tau_1 p_1 + (1-\tau_1)p_w - p\right) &= \frac{2(p-1/2)}{(1-\delta^2)g + k/2}(\sqrt{g} - \sqrt{d^2 g - k/2}) - 1 \\
&= \frac{2(p-1/2)(\sqrt{g} - \sqrt{d^2 g - k/2})}{(\sqrt{g} + \sqrt{d^2 g - k/2})(\sqrt{g} - \sqrt{d^2 g - k/2})} - 1 \\
&= \frac{2(p-1/2)}{(\sqrt{g} + \sqrt{d^2 g - k/2})} - 1 \\
&= \frac{2\sqrt{g - k/2}}{\sqrt{g} + \sqrt{d^2 g - \kappa/2}} - 1 \\
&= \frac{2\sqrt{1 + \frac{\kappa}{2g}}}{1 + \sqrt{\delta^2 - \frac{\kappa}{2g}}} - 1 \\
&< \frac{2}{1+1} - 1 \quad \text{(by Lemma A1.)} \\
&= 0.
\end{aligned}$$

Hence, $g(p)$ is monotonically decreasing in the mixing region and $g(\overline{p}) > 0$. This implies that $g(p) \geq 0$ for any prior belief in the mixing region. Hence, we can find a probability distribution over the posteriors that satisfies the Baysian plausibility. ■

Proof of Proposition 2

Given the functional forms

$$\overline{p} = \frac{1}{2} + \sqrt{g^2 - \frac{\kappa}{2}},$$

$$\underline{p} = \frac{1}{2} + \delta^2 \sqrt{g^2 - \kappa},$$

it is clear that $\overline{p}, \underline{p}$ are increasing in δ and κ by Lemma A1. ■

Chapter 3

Screening Information-Processing Information

This chapter is based on a joint work with Yonggyun(YG) Kim.

My contributions in this joint work are, (1) characterization of the buyer's optimal information acquisition, and also (2) an illustrative example that shows the seller's optimal selling scheme. The baseline model of seller's selling mechanism (which includes an upfront fee and a strike price) and numerical examples are the contributions of my coauthor, YG Kim.

3.1 Introduction

Due to the recent advancement of information technology, consumers can acquire more information about goods and services. They choose what to learn since information acquisition is usually costly. One of the issues is how consumers' information acquisition affects trading.

We study a monopoly pricing problem where buyers are uncertain about their valuations for a good. A monopolistic seller posts participation fee and strike price, then the buyers can privately and costly learn about their valuations and decide whether to buy the good or not. Here, the seller provides the menus of pairs of prices to discriminate buyers based on the ability of processing information.

This type of offer can be widely observed in the real economy. For example, we can think of down payments in the housing market. Here, the buyer pays some amount in advance and then pays the rest after further inspection. While the buyer's willingness to buy the house clearly depends on the quality of the house, some buyers will better be able to assess the quality. There are many other examples, such as online shopping. For instance, online retailers may offer the same product as a "prime" product with free returns or as a regular product where returns involve shipping and return fees. The prime product with free return has advance payment while the regular product does. Another example is travel bookings, where the consumer can choose the timing of payment with full, partial or no advance payment. These examples motivate us to study price discrimination problems, where the seller offers a menu of participation fee, strike price pairs and discriminates the consumers depending not on their value, but on the amount of information they can process.

We show that the buyer seeks good (bad) news about the good if the strike price is high (low) in equilibrium. If the strike price is high, then the buyer is pessimistic about the trade and willing to trade only if she receives good news about the good. Hence, she focuses only on good news and reduces the cost of information acquisition. It seems natural that when checking a cheap house, a buyer will inspect more carefully. Given this buyer's strategy, the seller chooses prices to maximize revenue. We will show, in an example, that the optimal offer charges high (low) participation fee and low (high) strike price to a buyer whose information processing ability is high (low). This result is intuitive in the sense that if the information processing ability is high, then the gain from the information acquisition is also high. Hence, such a buyer's willingness to pay is also high. In the housing market example, a buyer with specialized knowledge has higher willingness to pay for the strike price since he has larger chance to receive detailed information about the house after inspection.

Our paper contributes a literature that studies how information acquisition affects trade mechanism (Ravid et al., 2022; Thereze, 2022). The main difference is that we also consider the seller's screening problem when the information processing ability is private information of the buyers. Guo et al. (2022) studied the mechanism for selling experience goods when a buyer learns his own value by consuming it. The closest study to ours is Guo et al. (2020), who studied the optimal discriminatory disclosure problem where the seller controls the information. In contrast, we consider situations where the buyer can flexibly acquire information. The seller's discriminatory offer is different from traditional price discrimination models in the sense that the seller discriminates the buyer based not on the valuation but on the information processing ability.

The costly information acquisition problem is studied in the literature of rational inattention (Sims, 2003). In this literature, it is commonly assumed that the information cost is proportional to the entropy reduction. In our model, the amount of uncertainty is measured by a quadratic function. This measure makes the model tractable, and it has similar properties as the entropy function in the sense that the quadratic function is proportional to the quadratic approximation of the entropy function.

The rest of this paper is organized as follows. Section 2 describes the model. Section 3 gives the main results. We first provide the buyer's optimal strategy. we then derive the seller's optimal strategy in an example. Section 4 concludes this chapter.

3.2 Model

We study a monopoly pricing problem where a seller tries to sell a good to buyers who are uncertain about their valuations. There are two states $\omega \in \Omega = \{0, 1\}$ that determine the buyer's valuation b_0, b_1 ($b_1 > b_0$). The seller's valuation is normalized to 0. In addition, there are two types $t \in \{1, 2\}$ of buyers that determines the (inverse measure of) ability of processing information k_1, k_2 ($\kappa_1 < \kappa_2$). Both of them share a prior belief $f \in \Delta(\Omega)$. While each buyer knows his type, the seller only knows the distribution over the types $g \in \Delta(\{1, 2\})$.

The monopolistic seller posts the offer (a, P) where a is an advance payment and P is a strike price. Then, the buyers can privately and costly learn about their valuations and decide whether to buy the good or not.

Information Acquisition

The buyer can flexibly acquire the information about the state with information cost. He can choose a set of possible signals and a signal structure $\pi : \Omega \to \Delta(S)$. Let q_i be the posterior after receiving a signal $s_i \in S$, i.e.

$$q_i := \frac{f\pi(s_i|\omega = 1)}{f\pi(s_i|\omega = 1) + (1 - f)\pi(s_i|\omega = 0)}.$$

Then, an information structure π induces a distribution over posteriors τ and τ_i is an ex ante probability that the buyer receives a posterior belief q_i, i.e.

$$\tau_i = f\pi(s_i|\omega = 1) + (1 - f)\pi(s_i|\omega = 0).$$

We will assume that the information cost function is quadratic: For a signal structure π, the amount of information is measured by a function C such that

$$
\begin{aligned}
C(\pi) &= \sum_i \tau_i \left((f - 1/2)^2 - (q_i - 1/2)^2\right) \\
&= f(1-f) - \sum_i \tau_i q_i (1 - q_i).
\end{aligned}
$$

For a belief p, $(p-1/2)^2$ measures how certain the buyer is about the state. The function C measures the expected reduction of uncretainty due to the information acquisition. The orverall cost is given by $\kappa C(\pi)$ where $\kappa > 0$ is the parameter that represents the ability of processing information. Our assumption requires that the cost of information acquisition is proportional to the expected reduction of uncretainty.

The assumption of quadratic cost enhances tractability. In addition, it is considered as a quadratic approximation of entropy function that is widely used in the literature.

This game proceeds as follows:

Timing of the game

1. The seller posts a menu of pairs of an advance payment and a strike price (a, P);
2. The buyer chooses a deal (a, P);
3. The buyer chooses to an information structure π with an information cost $\kappa C(\pi)$.
4. The buyer receives a signal from π, then decides to accept the offer or reject.

3.3 Results

We will discuss the buyer's information acquisition problem given the seller's offer and then the seller's optimal offer.

3.3.1 Buyer's Problem

Given the seller's offer (a, P), the buyer's optimal information acquisition strategy solves the following problem[1]:

$$\max_{X,\pi} \sum_{\omega\in\Omega} \sum_{k=1}^{N} (b_\omega - P) \cdot X(s_k) \cdot \pi(s_k|\omega) \cdot f(\omega) - \kappa C(\pi)$$

where $X(s_k)$ is the probability of trading given a signal s_k.

We can think of this problem as a costly persuasion problem (Gentzkow and Kamenica, 2014) where the buyer persuades herself. Following a standard approach in the Bayesian persuasion literture (Kamenica and Gentzkow, 2011), we will use a belief approach: instead of a signal structure the buyer directly chooses a distribution over posteriors. These two approaches are equivalent under the Bayesian plausibility condition, i.e.

$$f = \sum_{k=1}^{n} r(s_k) q_k.$$

This condition requires that the expected value of posteriors must be equal to the prior. This guarantees that the buyer updates his belief following Bayes rule.

Following the belief approach, we can rewrite the problem as follows:

$$\sum_{\omega\in\Omega} \sum_{k=1}^{N} (b(\omega) - P) \cdot X(s_k) \cdot \pi(s_k|\omega) \cdot f(\omega) - \kappa C(\pi)$$

$$= \sum_{\omega\in\Omega} \sum_{k=1}^{N} ((b(\omega) - P) \cdot X(s_k) \cdot q(\omega|s_k) - \kappa\left(f(1-f) - q_k(1-q_k)\right)) \cdot \tau(s_k)$$

$$= \sum_{k=1}^{N} ((E[b(\omega)|s_k] - P) \cdot X(s_k) - \kappa\left(f(1-f) - q_k(1-q_k)\right)) \cdot \tau(s_k)$$

$$= \sum_{k=1}^{N} ((q_k b(1) + (1-q_k)b(0) - P) \cdot X(s_k) - \kappa\left(f(1-f) - q_k(1-q_k)\right)) \cdot \tau(s_k).$$

[1]The advance payment a is omitted since it is a sunk cost in this problem

Given an information structure π, after observing s_k, the buyer chooses

$$X^*(s_k) = \begin{cases} 1 & \text{if } E[b(\omega)|s_k] > P \\ [0,1] & \text{if } E[b(\omega)|s_k] = P \\ 0 & \text{if } E[b(\omega)|s_k] < P \end{cases}$$

In this problem, the buyer's information structure can include infinitely many messages. However, the next lemma shows that the optimal information structure includes at most two message.

Lemma 1. The optimal signal structure includes at most two messages.

Proof 1. Let $X^*(s_k)$ be the optimal choice given a signal s_k. Define a partition of S as

$$S_0 = \{s \in S | X^*(s) = 0\}$$

$$S_1 = \{s \in S | X^*(s) = 1\}$$

$$S_I = \{s \in S | X^*(s) \in (0,1)\}.$$

Consider a new information structure (π', S') and a choice rule $X'(\cdot)$ such that $S' = \{s_0, s_1\}$, $X'(s_1) = 1, X'(s_0) = 0$,

$$\pi'(s_0|\omega) = \sum_{s \in S_0 \cup S_I} \pi(s|\omega),$$

$$\pi'(s_1|\omega) = \sum_{s \in S_1} \pi(s|\omega).$$

It is easy to check that the new information structure gives the buyer the same (material) payoff as the original one. On the other hand, if S has more than two elements, π' is a strict garbling of π, and thus $C(\pi') < C(\pi)$. If the information constraint is binding under π, the buyer can enjoy higher payoff by using the additional amount of information. If the information constraint is not binding under π, then π must be a fully informative signal that has two messages. In both cases, the optimal signal structure has at most two messages. ■

Lemma 1 allows us to focus on the cases with one and two message(s) in the information structure. The case with one message is the degenerate case, i.e., there is no information acquisition. In this case, the buyer makes a decision based on the prior. Now consider the case with two messages (s_0, s_1) that induce different decisions. It is without loss to assume that the buyer trades after receiving s_1 ($X(s_1) = 1$) and does not trade after receiving s_0 ($X(s_0) = 0$).

The original problem can be rewritten as follows.

$$\max_{q_1, q_0, r_1} (q_1 b_1 + (1 - q_1) b_0 - P - \kappa(f(1-f) - q_1(1-q_1)))) \cdot \tau_1$$

$$- \kappa \left(f(1-f) - q_0(1-q_0)\right)(1-\tau_1)$$

$$\text{s.t.} \quad \tau_1 q_1 + (1-\tau_1) q_0 = f. \qquad \text{(BP, } \lambda\text{)}$$

We can view this problem as a concavification problem (Gentzkow and Kamenica, 2014) with two functions:

$$v_1(q) := \left(q_1 b_1 + (1-q_1) b_0 - P - \kappa(f(1-f) - q_1(1-q_1)))\right).$$

$$v_0(q) := -\kappa(f(1-f) - q(1-q)).$$

Here, $v_i(q)$ is the expected payoff given signal s_i is received and the posterior is updated to q. Figure 3.1 illustrates the optial signal structure for a specific parameter ($P = 0.5, f = 0.3, \kappa = 1$).

Formally, we can construct the Lagrangian to solve this problem.

Lagrangian:

$$\mathcal{L} = (q_1 b_1 + (1-q_1) b_0 - P - \kappa(f(1-f) - q_1(1-q_1)))) \cdot \tau_1$$

$$- \kappa \left(f(1-f) - q_0(1-q_0)\right)(1-\tau_1) + \lambda(f - \tau_1 q_1 - (1-\tau_1) q_0).$$

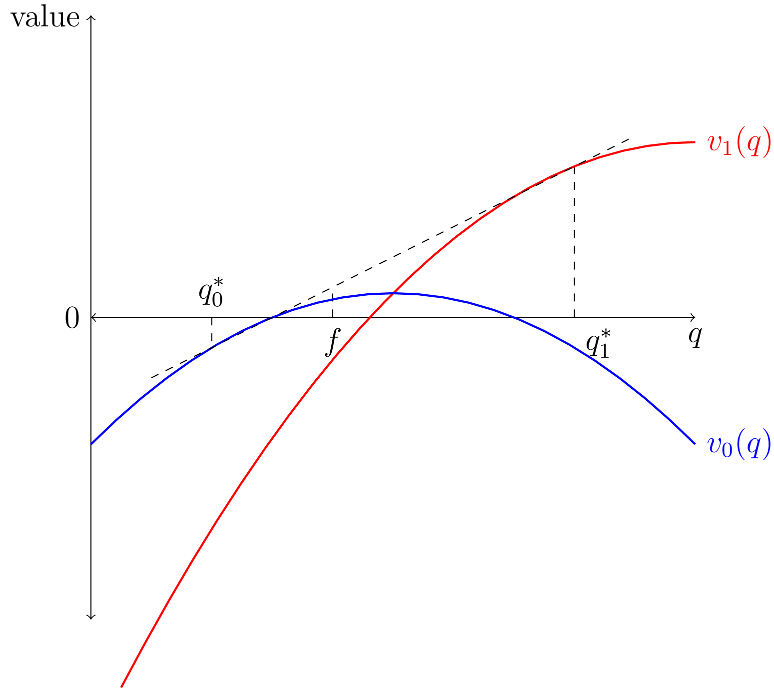

Figure 3.1: Concavification Problem

F.O.C.s:

$$q_1 : b_1 - b_0 - \kappa(2q_1 - 1) - \lambda = 0 \tag{3.1}$$

$$q_0 : -\kappa(2q_0 - 1) - \lambda = 0 \tag{3.2}$$

$$\tau_1 : q_1 b_1 + (1 - q_1)b_0 - P - \kappa(f(1-f) - q_1(1-q_1))) + \kappa\left(f(1-f) - q_0(1-q_0)\right) \tag{3.3}$$
$$- \lambda(q_1 - q_0) = 0.$$

Here, equations (3.1) and (3.2) require that the marginal value and cost of signal accuracy are the same. Equation (3.3) guarantees that the value from each signal must be the same.

Solving these conditions for q_1^*, q_0^* and r_1^*, we have

$$q_1^* = \frac{P - b_0}{b_1 - b_0} + \frac{b_1 - b_0}{4\kappa}$$

$$q_0^* = \frac{P - b_0}{b_1 - b_0} - \frac{b_1 - b_0}{4\kappa}$$

$$\tau_1^* = \frac{1}{2} + \frac{2\kappa}{b_1(1 - b_0)}\left(f - \frac{P - b_0}{b_1 - b_0}\right)$$

Figure 3.2 illustrates the optimal information acquisition strategy given the strike price P and the information cost parameter κ. When κ is small, the buyer can receive precise information with low cost. As κ becomes larger, she can only receive partial information. When P is high, the buyer will not buy unless she receives a strong indication of a good state. When P is low, the buyer will buy unless she receives a strong indication of a bad state. Hence, she seeks precise information for a bad state and less precise information for a good state. In this mannar, she acquires essential information with lowest possible information cost given the seller's offer (a, P) and her own ability to process information κ.

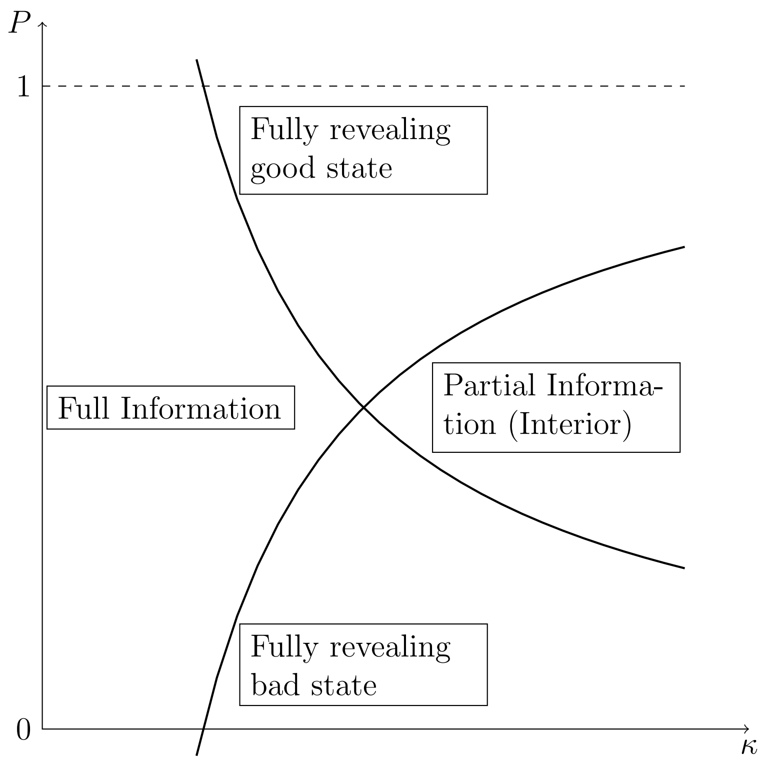

Figure 3.2: Optimal Information Acquisition with $b_1 = 1$ and $b_0 = 0$

Example: Suppose $b_1 = 1$ and $b_0 = 0$. The buyer's optimal strategy is reduced to

$$q_1^* = \min\{P + \frac{1}{4\kappa}, 1\}$$

$$q_0^* = \max\{P - \frac{1}{4\kappa}, 0\}$$

$$\tau_1^* = \frac{1}{2} + 2\kappa\,(f - P)$$

Optimal Value

Given the optimal solution, the buyer's payoff is given by

$$V^*(P,\kappa) = -\kappa f(1-f) + \tau^*(q_1^* b_1 + (1-q_1^*)b_0 - P + \kappa q_1^*(1-q_1^*)) + (1-\tau^*)\kappa q_0^*(1-q_0^*).$$

Figure 3.3 illustrates the buyer's indifference curve on $(\kappa, P)-$plane. The indifference curve is downward sloping since κ is the coefficient of the information cost. The slope becomes small as κ becomes small since the buyer with larger κ acquires information less actively. In addition, the marginal value of the strike price is τ^* by envelope theorem.

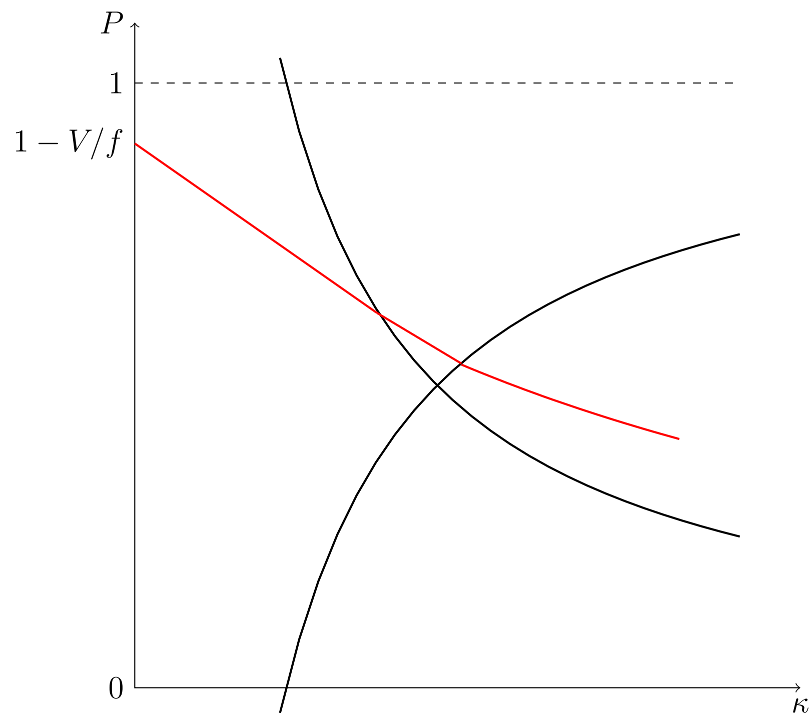

Figure 3.3: Buyer's Indifference Curve with $b_1 = 1$ and $b_1 = 0$.

3.3.2 Seller's Problem

The seller knows that there are two types of buyers: "high" type (type 1) has high information processing ability (low κ) and "low" type (type 2) has low ability (high κ). g is the fraction of the high type. The seller maximizes the revenue using the advance payment and the strike price. He faces the standard constraints: the participation constraint and the incentive compatibility constraint.

$$
\begin{aligned}
\max_{a_i, P_i} \quad & g(a_1 + \tau(P_1, \kappa_1)P_1) + (1-g)(a_2 + \tau(P_2, \kappa_2)P_2) \\
s.t. \quad & -a_i + V(P_i, \kappa_i) \geq 0 \qquad\qquad\qquad \forall i \qquad (PC_i) \\
& -a_i + V(P_i, \kappa_i) \geq -a_j + V(P_j, \kappa_i) \quad \forall i, j \quad (IC_i)
\end{aligned}
$$

Here, $\tau(P, \kappa)$ is the trade probability if the seller offers the strike price P to a buyer with *kappa*. It is easy to check that the participation constraint for type 1 buyer (IR1) slacks, that for type 2 (IR2) is binding, and the incentive compatibility constraint for type 1 is binding (IC1). In addition, (IC1) implies (IC2) if and only if $\min\{P_1, P_2\} \geq f$. Then, we can rewrite the seller's problem as follows:

$$
\begin{aligned}
\max_{a_i, P_i} \quad & g(a_1 + \tau(P_1, \kappa_1)P_1) + (1-g)(a_2 + \tau(P_2, \kappa_2)P_2) \\
s.t. \quad & -a_2 + V(P_2, \kappa_2) = 0 \qquad\qquad\qquad (PC2) \\
& -a_1 + V(P_1, \kappa_1) = -a_2 + V(P_2, \kappa_1) \qquad (IC1) \\
& f \leq \min\{P_1, P_2\}.
\end{aligned}
$$

We will solve a relaxed problem with (PC2) and (IC1) and confirm that the solution satisfied the last condition.

F.O.C.

$$
\begin{aligned}
P_1 : & g\left(\frac{d}{dP_1}\tau(P_1, \kappa_1)P_1 + r(P_1, \kappa_1)\right) + \mu\frac{d}{dP_1}V(P_1, \kappa_1) = 0 \\
P_2 : & (1-g)\left(\frac{d}{dP_2}\tau(P_2, \kappa_2)P_2 + \tau(P_2, \kappa_2)\right) + \mu\frac{d}{dP_2}V(P_2, \kappa_2) - \mu\frac{d}{dP_2}V(P_2, \kappa_1) = 0 \\
a_1 : & g - \mu = 0 \\
a_2 : & 1 - g - \lambda + \mu = 0
\end{aligned}
$$

Instead of providing a full characterization of the optimal pricing strategy, we will provide an example that illustrates the key intuition.

Example: Suppose $b_1 = 1$, $b_0 = 0$, $f = 0.2$, $g = 0.5$, $\kappa_1 = 1$ and $\kappa_2 = 2$. The seller's optimal offer is

$$P_1^* \simeq 0.4,$$

$$P_2^* \simeq 0.26,$$

$$a_1^* \simeq 0.025,$$

$$a_2^* \simeq 0.04.$$

This example suggests that the seller charges a low participation fee and a high strike price to the high type. The intuition is that the high type buyer has larger (expected) gain from information acquisition. Hence, she has larger willingness to pay for the strike price. This difference in the values of information acquisition allows the seller to screen the buyers' ability to process information.

3.4 Conclusion

We studied a monopoly pricing problem where a buyer costly acquire information in purchasing decision. We showed that if the strike price is high (low), the buyer seeks a strong indication of a good (bad) state and a relatively weak indication of a bad (good) state in order to reduce the information acquisition cost. Given this strategy, we provided an example that suggests the seller's optimal offer charges high (low) participation fee and low (high) strike price to a buyer whose information processing ability is high (low). In the example of the housing market, the buyer's knowledge of houses determines the ability to process information. Buyers who do not have sufficient knowledge of houses will prefer to pay high down payment and low strike price, while those who have sufficient knowledge will choose a low down payment and higher strike price. The fact that the two types of buyers prefer different contracts. opens the door to price discrimination against low types who have lower ability to process information. Our model allows investigating this type of price discrimination.

For future work, we can consider the situation where the seller can affect the information cost parameter κ. For example, the free return period allows the buyer to use the good and know more about the good before the trade is finalized. If the free return period is long, then she can acquire more information easily (which is interpreted as low κ). By modeling the seller's strategic choice of the period length, we can discuss the welfare effect of the length of the free return period. For example, in many countries, the government makes a regulation that guarantees the consumer's cooling-off period (free-return period). Our study can potentially help to assess the welfare effect of such policies.

Chapter 4

Overview of the Literature

4.1 Introduction

This chapter provide an overview of the related literature and the relationship between this dissertation and existing studies. This dissertation is about information constraints in economics and there are mainly four lines of literature that are closely related: (1) dynamic information acquisition, (2) rational inattention, (3)information design and (4) trade mechanisms and information acquisition.

Throughout this dissertation, I study environments where agents endogenously choose their learning strategies subject to the information capacity or cost. Since they cannot process all available information, they have to choose what to learn and how much to learn. While my approach allows us to analyze these rich trade-offs, it can complicate the analysis in many environments. Some recent studies have explored how they acquire the information and how it affects their behavior in limited environments (such as limited state spaces, payoff structures and action spaces). My study contributes to this line of study in the context of a dynamic information acquisition problem and a binary trade problem.

In chapter 2, I study a dynamic information acquisition problem of an individual who faces a limited ability to processing information. This chapter is in the line of the dynamic information acquisition literature and the rational inattention literature. In this problem, the individual faces the information capacity and thus cannot acquire sufficient information at once. Hence, he allocates the amount of information he acquires over time. I use the methodology from the information design literature.

Chapter 3 is based on joint work with Younggyn Kim and gives an application of the costly information acquisition model. This chapter is mainly about the trade mechanisms and information acquisition. In a simple binary trade model, I demonstrate how the seller's offer affects the buyer's information acquisition, and also how the seller screens the ability to process information to extract surplus. Since the cost of information is introduced, this is also related to the rational inattention literature.

4.2 Information Constraints in Economics

Dynamic Information Acquisition

Earlier studies (Arrow et al., 1949; Wald, 1947) have studied the optimal stopping problem where an individual chooses the timing to take an action based on sequentially received random signals. More recently, Fudenberg et al. (2018) studied the optimal stopping problem under the drift diffusion model, and found that when the DM makes a decision quickly, the quality of the choice tends to be high. Their findings are supported by empirical studies (Bhui, 2019). In these studies, they have explored a trade-off between the cost of waiting and the benefit from additional information.

While they assumed that the information structure is exogenous, recent studies have allowed the DM to choose the signal structures under information constraints to discuss what type of information he acquires, as well as how long he waits for the information. Che and Mierendorff (2019) introduced the allocation of attention over different information sources. They characterized the optimal stopping rule and how the expected waiting time and the accuracy of information change depending on the prior belief.

Zhong (2022), studied more general information acquisition problem where the individual can choose any information structure subject to an information cost in continuous time. By allowing flexible information acquisition, he discovered an important trade-off: In addition to a standard trade-off (between the cost of waiting and the benefit of additional information), the individual also faces a trade-off between accuracy of the signal and the frequency of receiving information.

The dynamic flexible information acquisition problem is generally complicated problem so that a large part of the analysis in Zhong (2022) is limited to two state, continuous-time case. The solution in more general case has not been discovered.

In chapter 2, I will study a dynamic (two period) flexible information acquisition problem wth information constraints. Discrete-time modeling allows us to discuss the consistency of my results to empirical findings and how we can identify the discount factor and capacity from the observed data. On the other hand, it also complicates the analysis since the technique for continuous time model cannot be applied. The main difference is that the contradictory information is essential in discrete time setting.

Rational Inattention

Due to the recent advancement of information technology (such as SNS), people are likely to be exposed to a large amount of information. In such cases, people cannot process all available information. Hence, they have to choose what and how much to learn.

Rational inattention theory, developed by Sims (2003) considers the behavior of information constrained individuals. In this literature, the measurement for the amount of information is introduced (for example, the information entropy function). In many rational inattention models, the DM is assumed to incur the cost depending on the amount of information he acquires, or to be able to acquire a limited amount of information in an unit of time. This theory have provided important implications on economic behaviors such as wage stickiness (Sims, 2003).

Recent studies (Caplin and Dean, 2015; Matějka and McKay, 2015) developed this theory further and Steiner et al. (2017) extended the framework to a general dynamic model with a constant cost flow.

In another line of study, (De Oliveira et al., 2017) discussed the identification of information costs using a set of menu choice data.

The DM in my study is also rationally inattentive to some information because of the information cost and the information capacity. In chapter 2, I study a dynamic flexible information acquisition problem where the cost of waiting is captured by the geometric discounting. I also discuss how we can identify the parameter (discount factor and the information capacity) from observed data. In chapter 3, the buyer is assumed to have a limited ability to process information. I discovered how the seller's offer affects the information acquisition of buyers and how the buyer's willingness to pay changes before and after the information acquistion.

Information Design

Information design literature has studied the effect of information provision on the individual's behavior. This type of situations is widely observed in the economy (e.g. a prosecutor and a judge, lobbyists, and salespeople).

The analysis is based on the Bayesian persuasion models (Kamenica and Gentzkow (2011), Gentzkow and Kamenica (2014)), where an individual (I call him a sender) tries to persuade another (I call her a receiver) to manipulate her action for his own benefit. More concretely, the sender commits to an information structure and fully reports the signal realization, and then the receiver chooses an action based on the realization. Kamenica and Gentzkow (2011) demonstrated how a prosecutor (sender) persuades a judge (receiver) to change her action, and then showed when the sender can gain from persuasion.

In chapter 2 and 3, the information acquisition is modeled as a self-persuasion problem, where an individual persuades himself to choose a better action. This study employs two important methodologies. First, in the belief-based approach, the DM directly chooses the distribution over posteriors. Second, the concavification method (developed in Aumann et al. (1995)) enables us to solve this type of problems graphically.

In chapter 2, since the DM faces the information constraint, the application of the concavification is not straightforward. I study the concavification method with a Lagrange multiplier of the information constraint. The multiplier can be interpreted as a shadow value of the information capacity and it is determined by current belief. Hence, I will discover how the multiplier determines the optimal information acquisition and how the prior belief determines the multiplier in order to answer how the DM acquires information given a prior.

In chapter 3, the buyer's information acquisition is a costly persuasion problem (Gentzkow and Kamenica, 2014) where the buyer persuades himself. Since the optimal information acquisition depends on the seller's offer, I can answer how the seller's offer affects the buyer's information acquisition.

Trade Mechanism and Information Acquisition

Recent studies investigate how information acquisition affects trade mechanism (Ravid et al., 2022; Thereze, 2022). They have introduced the information acquisition stage in trade models and discovered the mutual dependency between the information acquisition strategy and the pricing scheme.

We also address this problem in chapter 3. The main difference is that we also consider the seller's screening problem when the information processing ability is private information of the buyers. The closest study to ours is Guo et al. (2020), who studied the optimal discriminatory disclosure problem where the seller controls the information. In contrast, we consider situations where the buyer can flexibly acquire information subject to the cost. By combining the idea from the rational inattention and flexible information acquisition, the model describes how the buyer's ability to process information affects the optimal information acquisition and the willingness to pay before and after the information acquisition. The seller's discriminatory offer is different from traditional price discrimination models in the sense that the seller discriminates the buyer based not on the valuation but on the information processing ability.

Chapter 5

Conclusion

This dissertation explored a theory of constrained information acquisition and its application. I provided a theoretical analysis on the constrained information acquisition.

In Chapter 2, I studied how an individual with the information capacity acquire information. I characterize the optimal information structure and discuss how we can identify the discount factor and the information capacity.

In Chapter 3, I studied a monopoly pricing problem where a buyer costly acquire information. I first characterized the optimal information acquisition strategy given the seller's offer, and then discussed the optimal pricing scheme to screen the buyers' abilities to process information.

The economic effect of the constrained information acquisition has not been fully discovered. Although some recent studies has discovered the individual's optimal information acquisition strategy under information constraint and its economic implications, their analyses are limited to specific environments (e.g. limited state spaces or information sources). As a part of literature, this dissertation can also help to understand the effect of information constraints in the economy.

References

Arrow, K. J., D. Blackwell, and M. A. Girshick (1949): "Bayes and Minimax Solutions of Sequential Decision Problems," *Econometrica*, 17, 213–244.

Aumann, R. J., M. Maschler, and R. E. Stearns (1995): *Repeated games with incomplete information*, MIT press.

Baron, J., S. Scott, K. Fincher, and S. E. Metz (2015): "Why does the Cognitive Reflection Test (sometimes) predict utilitarian moral judgment (and other things)?" *Journal of Applied Research in Memory and Cognition*, 4, 265–284.

Bhui, R. (2019): "Testing Optimal Timing in Value-Linked Decision Making," *Computational Brain & Behavior*, 2, 85–94.

Caplin, A. and M. Dean (2015): "Revealed Preference, Rational Inattention, and Costly Information Acquisition," *American Economic Review*, 105, 2183–2203.

Che, Y.-K. and K. Mierendorff (2019): "Optimal Dynamic Allocation of Attention," *American Economic Review*, 109, 2993–3029.

De Oliveira, H., T. Denti, M. Mihm, and K. Ozbek (2017): "Rationally inattentive preferences and hidden information costs," *Theoretical Economics*, 12, 621–654.

Fudenberg, D., P. Strack, and T. Strzalecki (2018): "Speed, accuracy, and the optimal timing of choices," *American Economic Review*, 108, 3651–84.

Gentzkow, M. and E. Kamenica (2014): "Costly Persuasion," *American Economic Review*, 104, 457–62.

Guo, H., W. He, and B. Liu (2022): "Learning by Consuming: Optimal Pricing with Endogenous Information Provision," *arXiv preprint arXiv:2209.01453*.

Guo, Y., H. Li, and X. Shi (2020): "Optimal discriminatory disclosure," .

Johnson, E. D., E. Tubau, and W. De Neys (2016): "The Doubting System 1: Evidence for automatic substitution sensitivity," *Acta psychologica*, 164, 56–64.

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

Matějka, F. and A. McKay (2015): "Rational Inattention to Discrete Choices: A New Foundation for the Multinomial Logit Model," *American Economic Review*, 105, 272–98.

Ravid, D., A.-K. Roesler, and B. Szentes (2022): "Learning before trading: on the inefficiency of ignoring free information," *Journal of Political Economy*, 130, 346–387.

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

Steiner, J., C. Stewart, and F. Matějka (2017): "Rational Inattention Dynamics: Inertia and Delay in Decision-Making," *Econometrica*, 85, 521–553.

Thereze, J. (2022): "Screening Costly Information," .

Wald, A. (1947): "Foundations of a General Theory of Sequential Decision Functions," *Econometrica*, 15, 279–313.

Zhong, W. (2022): "Optimal Dynamic Information Acquisition," *Econometrica*, 90, 1537–1582.

Biography

Taishi Sassano is a sixth-year Ph.D. student at Duke University, department of economics. He received a bachelor's degree in Economics from University of Tsukuba in Japan and a master's degree in Economics from Hitotsubashi University in Japan.

He is interested in decision theory, behavioral economics, and game theory. His dissertation is about constrained information acquisition in economics.

He has received followship from Duke University (Duke Economics Department Fellowship, 2017-2023 ,Duke Graduate School Summer Research Fellowship, 2018-2021), and he has a publication (with coauthors) on Economics Bulletin.