Essays on information acquisition

Weijie Zhong

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2019

Abstract

Essays on information acquisition

Weijie Zhong

This dissertation studies information acquisition when the choice of information is fully flexible. Throughout the dissertation, I consider a theoretical framework where a decision maker (DM) acquires costly information (signal process) about the payoffs of different alternatives before making a choice. In Chapter 1, I solve a general model where the DM pays a cost that depends on the rate of uncertainty reduction and discounts delayed payoffs. The main finding is that the optimal signal process resembles a Poisson signal — the signal arrives occasionally according to a *Poisson process*, and it drives the inferred posterior belief to jump discretely. The optimal signal is chosen to confirm the DM's prior belief of the most promising state. Once seeing the signal, the decision maker is discretely surer about the state and stops learning immediately. When the signal is otherwise absent, the decision maker becomes gradually less sure about the state, and continues learning by seeking more precise but less frequently arriving signals. In Chapter 2, I study the sequential implementation of a target information structure. I characterize the set of decision time distributions induced by all signal processes that satisfy a per-period learning capacity constraint on the rate of uncertainty reduction. I find that all decision time distributions have the same mean, and the maximal and minimal elements by mean-preserving spread order are exponential distribution and deterministic distribution. The result implies that when the time preference is risk loving (e.g. standard or hyperbolic discounting), Poisson signal is optimal since it induces the riskiest exponential decision time distribution. When time preference is risk neutral (e.g. constant delay cost), all signal processes are equally optimal. In Chapter 3, I relax the assumption on information cost by assuming that the measure of signal informativeness is an indirect measure

from sequential minimization. I first show that an indirect information measure is supported by sequential minimization *iff* it satisfies: 1) monotonicity in Blackwell order, 2) sub-additivity in compound experiments and 3) linearity in mixing with no information. Then I study a dynamic information acquisition problem where the cost of information depends on an indirect information measure and the delay cost is fixed (the DM is time-risk neutral). The optimal strategy is to acquire Poisson type signals. The result implies that when the cost of information is measured by an indirect measure, Poisson signals are intrinsically cheaper than other signal processes. Chapter 4 introduces a set of useful technical results on constrained information design that is used to derive the main results in the first three chapters.

Table of contents

List of Figures

Acknowledgments

I would like to thank my dissertation committee members, Yeon-Koo Che, Navin Kartik, Qingmin Liu, Andrea Prat and Mark Dean. First of all, I would like to express my deepest gratitude to my principle advisor Yeon-Koo Che for his constant guidance and support, and for serving as my role model of a truly insightful, rigorous and diligent economist. I am indebted to my advisor Navin Kartik for his invaluable directions and comments throughout my research. He expanded my intellectual curiosity and academic interest in economic theory. I am extremely grateful to Qingmin Liu, who exerted tremendous effort to help me build up my skills and shape my taste for research. I am also grateful to Andrea Prat, for delivering an inspiring economic theory course which attracted me to explore the field, and for his excellent feedbacks at all stages of my works. I would also like to thank Mark Dean who provided insightful feedbacks to my work and deepened my understanding of economic theory.

It was my great pleasure collaborating with my coauthors Teddy Kim, Konrad Mierendorff and Xianwen Shi, together with Yeon-Koo Che, Navin Kartik and Qingmin Liu. I sincerely thank them not only for contributing significantly to the projects, but also for tolerating my ignorance and sluggish, and turning me into a better scholar.

The entire theory group at Columbia university has unanimously supported me throughout my doctoral study. In particular, I want to thank Marina Halac and Pietro Ortoleva for discussing the contents of this dissertation with me for numerous times and for giving all kinds of useful advice beyond just research. I received very useful feedbacks on my research and advices for my career from Xiaosheng Mu, Jacopo Perego and Evan Sadler, to whom I am also extremely grateful.

For helpful comments and discussions on this dissertation, I am also grateful to Sylvain Chassang, Johannes Hörner, Jakub Steiner, Philipp Strack, Tomasz Strzalecki, Andrzej Skrzypacz as well as participants at conferences and seminars, and visitors to Columbia University. In particular, Benjamin Golub, Yingni Guo, Shaowei Ke, Jonathan Libgober, Heng Liu, Harry Di Pei, Daniel Rappoport, Xingye Wu, Ming Yang and Jidong Zhou helped me a lot both in my research as scholars and on a personal level as friends.

On a personal level, I want to thank my schoolmates and friends for being encouraging and supportive regardless of my bad personality: Narisu Bai, Dawei Ding, Jiayin Hu, Ang Li, Xuan Li, RC Xizhi Lim, Yifeng Liu, Yifeng Luo, Bo Qin, Qiuying Qu, Shuaiwen Wang, Peifan Wu, Danyan Zha, Qing Zhang, Yi Zhu. Finally, I would like to thank my parents and other friends. It is impossible to finish this dissertation without their support.

Dedication

I dedicate the dissertation to my girlfriend Mengting Gu,
who has provided me unconditional love and support,
who hedged my risks when I chose to pursuit a risky career,
who guided me in making the most difficult decisions,
and who induced a positive Poisson jump in my life.

Introduction

This dissertation considers the following question: what is the optimal way to acquire information over time to learn about the payoffs of different options? This is a very classic question that has been extensively studied in the literature starting from Wald (1947) and Arrow, Blackwell, and Girshick (1949). However, we still do not have a complete answer to this question, as the conventional approaches have been searching within very limited types of information, e.g. many models consider only Brownian motion type information. Typically, papers in the literature study the optimal choice of "when to stop learning" taken a specific process of information as given, or the optimal control of a specific parameter of a given parametric family of information processes.

The goal of this dissertation is to answer the question by searching among *all* conceivable types of information, and completely endogenize the information acquisition strategy.

The practical motivation for permitting such *flexibility* in the type of information is that in practice the process of information acquisition can often be controlled in multiple aspects. The rapid development in statistics, data science and computer science is making information acquisition increasingly more flexible. For example, nowadays if a tech company wants to figure out the market's response to an internal innovation, it can launch an A-B test on an online marketing platform, fine-tune hundreds of parameters of the test design and change them adaptively when data arrives. Another example is that FDA recently published its guidance of *adaptive design for clinical trials* (see FDA (2018)). The guidance states that clinical trial designs with adaptive sample size, adaptive dose selection and response-adaptive randomization might improve the efficiency of the trials. In these examples, there is no a priori reason why some ad hoc restrictions on the type of information, e.g. the acquired data is generated from a normal distribution, are satisfied.

The optimal information acquisition process might generate skewed and fat-tailed data, which can only be fully covered in a completely flexible information acquisition model.

The theoretical framework for the entire dissertation is a sequential decision making model building upon the Wald framework. I consider a decision maker who makes a one-time choice from a set of actions, whose payoffs depend on a state unknown to the decision maker. The state is initially selected by the nature and remains fixed over time. At any instant of time, the decision maker chooses whether to *stop* learning and select an action or *continue* learning by choosing *nonparametrically* an informative signal structure for the next moment of time. Both delaying the decision and acquiring information are costly. Of course, hardly any prediction can be made in a model with such generality. I will proceed by solving this optimization problem, keeping the full generality in the decision problem and information acquisition, but imposing three different sets of more restrictive assumptions on the cost of delay and the cost of information in the following three chapters.

In Chapter 1, entitled "*Optimal dynamic information acquisition*", I study the case that (i) the decision maker discounts delayed utilities in a standard way, (ii) the cost of information depends on how fast the uncertainty about the unknown state is decreasing (also known as posterior separability). The goal of Chapter 1 is to fully solve for the optimal dynamic information acquisition strategy in a fairly general model with standard assumptions (discounting and posterior separable cost structure).

There are two main results. The first result states that although the model is non-parametric and allows fully flexible strategies, the optimal information acquisition strategy modeled as the induced posterior belief process can be restricted to a simple *jump-diffusion process* without loss. The second result fully characterizes the optimal belief process, which involves only a compensated Poisson jump process almost surely. In other words, it is optimal to conduct experiments that generate skewed and fat-tailed data.

Such experiment can be a stress test against the most promising state: Passing the test is rare but a pass is a conclusive proof that the state is very likely and a corresponding action should be adopted immediately. Otherwise, failing the test does not immediately end the test. I also show that conditional on failures, the future tests have higher difficulty — passing rate is lower but a pass is more precise.

The analysis in Chapter 1 illustrates that the optimality of Poisson type signal processes is a joint implication of the two assumptions in the model: exponential discounting and the information cost structure. Discussion in Appendix A.1.4.1 suggests that the posterior separability assumptions is essentially a *neutrality* condition: learning a target information structure through all equally costly strategies takes the same amount of time on average. To further understand the roll played by the two assumptions, I generalize each of them in the following two chapters.

In Chapter 2 on "*Time preference and information acquisition*", I keep the assumption on information cost and generalize the cost of delay to general convex or concave time cost. To get tractability in the model with further generality, I impose additional restrictions that (i) the flow cost of information acquisition is fixed (ii) the target decision rule is fixed. These restrictions shut down the dynamics of target decision rule and flow cost level, and highlight the implication of information on decision time. The main result of Chapter 2 is that for all convex time cost functions, the optimal dynamic information acquisition strategy is a *Poisson signal process* that either implements the target decision rule at a Poisson rate or generates no information with large probability. For any concave time cost, the optimal dynamic information acquisition strategy is a *pure accumulation* strategy that only accumulates information but makes no decision until a deterministic date. Noticing that the neutrality condition makes all information acquisition strategy equally efficient on average. So the key implication of difference strategies is that the Poisson signal process induces decision in a *riskiest* way on the dimension of time: decision is either taken very early on or there is a long delay. On the contrary, the pure accumulation strategy minimizes time-risk involved in decision making.

Chapter 2 reveals a key implication of information acquisition: it determines the risk in the decision making time. Therefore, under the neutrality condition (posterior separability assumption), all information acquisition strategies induce the same expected decision time and they only differ in the risks. Then, the preference on information acquisition strategies is solely pin down by the preference on time risk.

To deepen our understanding about the cost of information, I generalize the assumption on information cost in Chapter 3 on "*indirect information measure and dynamic learning*". I assume that (i) the cost of delay is linear in time (time-risk neutral) and (ii) the cost of information depends on an *indirect information measure*. An indirect information measure takes an arbitrary cost function of information as primitive, and for each signal structure derives the minimized expected total cost from a sequence of signal structures that replicates the original signal structure. In other words, the assumption I put on the cost of information is essentially that (i) I allow within period sequential minimization of information measure, (ii) there is increasing marginal cost to the information measure per period. The main result of Chapter 3 is that the optimal signal process is a direct compound Poisson signal: signal arrives according to a Poisson counting process and the arrival of signal suggests the optimal action directly, where the optimal action profile can be solved in an equivalent static rational inattention problem.

The analysis in Chapter 3 suggests that Poisson type information acquisition is not only the "riskiest" when we restrict the information cost to satisfy neutrality i.e. all learning strategies to be equally fast, it is also the "fastest" when we relax such restrictions on information cost, as long as the cost can be justified by within period information measure minimization.

Chapter 4 introduces a set of useful technical results on constrained information design, which are used to characterize the optimal strategies in Chapters 1 and 3. I characterize the set of all combinations of expected value of finite objective functions from designing information. I show that the set is compact, convex and can be implemented by signal structures with finite support when the state space is finite. Moreover, the set as a correspondence of prior belief is continuous. Based on this result, I develop a concavification method of Lagrangian that works with general constrained optimization. Other applications of the results include persuasion of receivers with outside options and screening using information.

Chapter 1

Optimal dynamic information acquisition

1.1 Introduction

When individuals make decisions, they often have imperfect information about the payoffs of different alternatives. Therefore, the decision maker (DM) would like to acquire information to learn about the payoffs prior to making a decision. For example, when comparing new technologies, a firm may not know the profitability of alternative technologies. The firm often spends a considerable amount of money and time on R&D to identify the best technology to adopt. One practically important feature of the information acquisition process is that the choice of "what to learn" often involves considering a rich set of salient aspects. In the previous example, when designing the R&D process, a firm may choose which technology to test, how much data to collect and analyze, how intensive the testing should be, etc. Other examples include investors designing algorithms to learn about the returns of different assets, scientists conducting research to investigate the validity of different hypotheses, etc.

To capture such richness, in this chapter, I consider a DM who can choose "what to learn" in terms of *all* possible aspects, as well as "when to stop learning". The main goal is to obtain insight into dynamic information acquisition without restriction on what type of information can be acquired. In contrast to my approach, the classic approach is to focus on one aspect while leaving all other aspects exogenously fixed. The seminal works by Wald (1947) and Arrow, Blackwell, and Girshick (1949) study the choice of "when to stop" in a stopping problem with all aspects of the learning process being exogenous. Building upon the Wald framework, Moscarini and Smith (2001) endogenize one aspect of learning, the *precision*, by allowing the DM to control a precision parameter of a Gaussian signal process. Che and Mierendorff (2016) endogenize another aspect of learning, the *direction*, by allowing the DM to allocate limited attention to different news sources, each biased in a different direction. Here, by allowing all learning aspects to be endogenous, the current chapter contributes by studying which learning aspect(s) is(are) endogenously relevant for the DM and how the optimal strategy is characterized in terms of these aspects.

In the model, the DM is to choose from a set of actions, whose payoffs depend on a state unknown to the DM. The state is initially selected by nature and remains fixed over time. At any instant of time, the DM chooses whether to stop learning and select an action or to continue learning by *nonparametrically* choosing the evolution of the belief process. The choice of a nonparametric belief process models the choice of a dynamic information acquisition strategy with no restriction on any aspect. I introduce two main economic assumptions. (i) The DM discounts delayed payoffs. (ii) Learning incurs a flow cost, which depends convexly on how fast the uncertainty about the unknown state is decreasing. The main model is formulated as a stochastic control-stopping problem in continuous time.

The main result shows that the optimal strategy is contained in a simple family characterized by a few endogenously relevant aspects (Theorem 1.1) and fully solves for the optimal strategy in these aspects (Theorems 1.2 and 1.3). Specifically, the first result states that although the model is nonparametric and allows for fully flexible strategies, the belief process can be restricted to a simple *jump-diffusion process* without loss. In other words, a combination of a *Poisson signal*—a rare and substantial breakthrough that causes a jump in belief—and a *Gaussian signal*—frequent and coarse evidence that drives belief diffusion—is endogenously optimal. A jump-diffusion belief process is characterized by four parameters: the *direction*, *size* and *arrival rate* of the jump, and the flow variance of the diffusion. The four parameters represent four key aspects of learning: the *direction*, *precision* and *frequency* of the Poisson signal, and the precision of the Gaussian signal. The first result suggests that the DM need consider only the trade-offs among these aspects; any other aspect is irrelevant for information acquisition.

The second result fully characterizes the parameters of the optimal belief process. I find that the Poisson signal strictly dominates the Gaussian signal almost surely, i.e. no resources should ever be invested in acquiring the Gaussian signal. The optimal Poisson signal satisfies the following qualitative properties in terms of the three aspects and the stopping time:

- ***Direction***: The optimal direction of learning is *confirmatory*– the arrival of a Poisson signal induces the belief to jump toward the state that the DM currently finds to be most likely. As an implication of Bayes rule, the absence of a signal causes the belief to drift gradually towards the opposite direction, namely, the DM gradually becomes less certain about the state.

- ***Precision***: The optimal signal precision is *negatively related* to the continuation value. Therefore, when the DM is less certain about the state, the corresponding continuation value is lower, which leads the DM to seek a more precise Poisson signal.

- ***Frequency***: The optimal signal frequency is *positively related* to the continuation value. In contrast to precision, the optimal signal frequency decreases when the DM is less certain.

- ***Stopping time***: The optimal time to stop learning is immediately after the arrival of the Poisson signal. Therefore, the breakthrough happens only once at the optimum. Then, the DM stops learning and chooses an optimal action based on the acquired information.

The optimal strategy is very heuristic and easy to implement. In the previous example, the firm can choose the technology to test, as well as the test precision and frequency. As a result, the optimal strategy is implementable. The optimal R&D process involves testing the most promising technology. The optimal test is designed to be difficult to pass, so good news comes infrequently, as in a Poisson process. A successful test confirms the firm's prior conjecture that the technology is indeed good and the firm immediately adopts the technology. Otherwise, the firm continues the R&D process. No good news is bad news, so the firm becomes more pessimistic about the technology and revises the choice of the most promising technology accordingly. The future tests involve higher passing thresholds and lower testing frequency. As illustrated by the example, although this chapter studies a benchmark with fully flexible information acquisition, the optimal strategy applies to more general settings where information acquisition is *not* fully flexible, but involves these salient aspects.

The main intuition behind the optimal strategy is a novel *precision-frequency trade-off*. Consider a thought experiment of choosing an optimal Poisson signal with fixed direction and cost level. The remaining two parameters—precision and frequency—are pinned down by the marginal rate of substitution between them. Importantly, the trade-off depends on the continuation value. Due to discounting, when the continuation value is higher, the DM loses more from delaying the decision. Therefore, the DM finds it optimal to acquire a signal more frequently at the cost of lowering the precision to avoid costly delay. In other words, the marginal rate of substitution of frequency for precision is increasing in the continuation value. As a result, frequency (precision) is positively (negatively) related to the continuation value.

In addition to precision and frequency, this intuition also explains other aspects. First, the Gaussian signal is equivalent to a special Poisson signal with close to zero precision and infinite frequency. The previous intuition implies that infinite frequency is generally suboptimal except when the continuation value is so high that the DM would like to sacrifice almost all signal precision. As a result, the Gaussian signal is strictly suboptimal except for the non-generic stopping boundaries. Second, for any fixed learning direction, Bayes rule implies that the absence of a signal pushes belief away from the target direction; to ensure the same level of decision quality the signal precision should increase over time to offset the belief change. By acquiring a confirmatory signal, the DM becomes more pessimistic and, consequently, more patient over time. Therefore she can reconcile both incentives through reducing the signal frequency and increasing the signal precision. By contrast, if the DM acquires a contradictory signal, she becomes more impatient over time and prefers the frequency to be increasing. The two incentives become incongruent, thus, learning in a confirmatory way is optimal.

This intuition suggests that the crucial assumption for the optimal strategy is discounting — discounting drives the key precision-frequency trade-off. This observation highlights the deep connection between dynamic information acquisition and the DM's attitude toward time-risk. Discounting implies that the DM is risk loving toward payoffs with uncertain resolution time, as the exponential discounting function is convex. Intuitively, the riskiest information acquisition strategy is a "greedy strategy" that front-loads the probability of success as much as possible, at the cost of a high probability of long delays. The confirmatory Poisson learning strategy in this chapter exactly resembles a greedy strategy. The key property of the strategy is that all resources are used in verifying the conjectured state directly and no intermediate step occurs before a breakthrough. By contrast, alternative strategies, such as Gaussian learning and contradictory Poisson learning, involve accumulating substantial intermediate evidence to conclude a success. The intermediate evidence in fact hedges the time risk: the DM sacrifices the possibility of immediate success to accelerate future learning.

Extensions of the main model further illustrate the role played by each key assumption. The first extension replaces discounting with a fixed flow delay cost. In this special case, all dynamic learning strategies are equally optimal, as the crucial precision-frequency trade-off becomes value independent. This extension also illustrates that all learning strategies in the model are equally "fast" on average and differ only in "riskiness". This result further illustrates that the preference for time risk pins down the optimal strategy. Second, I consider general cost structures and find that the (strict) optimality of a Poisson signal over a Gaussian signal is surprisingly robust: it requires a minimal *continuity* assumption. Third, I study an extension where the flow cost depends linearly on the uncertain reduction speed. In this special case, learning has a constant return to signal frequency. As a result, the optimal strategy is to learn infinitely fast, that is, acquire all information at period zero.

This chapter provides rich implications by allowing learning to be flexible in all aspects. First, the main results highlight the optimality of the Poisson signal compared to the widely adopted diffusion models. Specifically, the diffusion models are shown to be justified only under the lack of discounting. Second, the characterization of the optimal strategy unifies and clarifies insights from some existing results. In these results, although the DM is limited in her learning strategy, she actually implements the flexible optimum whenever feasible and approximates the flexible optimum when infeasible. Moscarini and Smith (2001)'s insight that the "intensity" of experimentation increases in continuation value carries over to my analysis. I further unpack the design of experiment and show that higher "intensity" contributes to faster signal arrival but lower signal precision. Che and Mierendorff (2016) make same prediction about the learning direction as that of my analysis when the DM is uncertain about the state. But they predict the opposite when the DM is more certain about the state– the DM looks for a signal contradicting the prior belief. I clarify that the contradictory signal is an approximation of a high-frequency confirmatory signal when the DM is constrained in increasing the signal frequency.

The rest of this chapter is structured as follows. The related literature is reviewed in Section 1.2. The main continuous-time model and illustrative examples are introduced in Section 1.3. The dynamic programming principle and the corresponding Hamilton-Jacobi-Bellman (HJB) equation are introduced in Section 1.4. I analyze an auxiliary discrete-time problem and verify the HJB equation in Section 1.5. Section 1.6 fully characterizes the optimal strategy and illustrates the intuition behind the result. In Section 1.7 I discuss the key assumptions used in the model. Section 1.8 explores the implications of the main model on response time in stochastic choice and on a firm's innovation. Further discussions of other assumptions are presented in Appendix A.1, and key proofs are provided in Appendix A.2. All the remaining proofs are relegated to Appendix B.

1.2 Related literature

1.2.1 Dynamic information acquisition

This chapter is closely related to the literature about acquiring information in a dynamic way to facilitate decision making. The earliest works focus on the duration of learning. Wald (1947) and Arrow, Blackwell, and Girshick (1949) analyze a *stopping problem* where the DM controls the decision time and action choice given exogenous information. Moscarini and Smith (2001) extend the Wald model by allowing the DM to control the precision of a Gaussian signal. A similar Gaussian learning framework is used as the learning-theoretic foundation for the drift-diffusion model (DDM) by Fudenberg, Strack, and Strzalecki (2018). Following a different route, Che and Mierendorff (2016), Mayskaya (2016) and Liang, Mu, and Syrgkanis (2017) study the sequential choice of information sources, each of which is prescribed exogenously.

Other frameworks of dynamic information acquisition include sequential search models (Weitzman (1979), Callander (2011), Klabjan, Olszewski, and Wolinsky (2014), Ke and Villas-Boas (2016) and Doval (2018)) and multi-arm bandit models (Gittins (1974), Weber et al. (1992), Bergemann and Välimäki (1996) and Bolton and Harris (1999)). These frameworks are quite different from my information acquisition model. However, the forms of information in these models are also exogenously prescribed, and the DM has control over only whether to reveal each option.

Compared to the canonical approaches, the key new feature of my framework is that the DM can design the information generating process nonparametrically. In a similar vein to this chapter, two concurrent papers Steiner, Stewart, and Matějka (2017) and Hébert and Woodford (2016) model dynamic information acquisition nonparametrically; however they focus on other implications of learning by abstracting from sequentially smoothing learning. In Steiner, Stewart, and Matějka (2017) the linear flow cost assumption makes it optimal to learn instantaneously, whereas in Hébert and Woodford (2016), the no-discounting assumption makes all dynamic learning strategies essentially equivalent.[1] By contrast, the main focus of this chapter is on characterizing the optimal way to smooth learning. I analyze the setups of these two papers as special cases in Sections 1.7.1 and 1.7.3.

A main result of this chapter is the endogenous optimality of Poisson signals. Section 1.7.2 shows a more general result: a Poisson signal dominates a Gaussian signal for generic cost functions that are continuous in the signal structure. This result justifies Poisson learning models, which are used in a wide range of problems, e.g., Keller, Rady, and Cripps (2005), Keller and Rady (2010), Che and Mierendorff (2016), and Mayskaya (2016); see also a survey by Hörner and Skrzypacz (2016).

1.2.2 *Rational inattention*

This chapter is a dynamic extension of the static rational inattention (RI) models, which consider the flexible choice of information. The entropy-based RI framework is first introduced in Sims (2003). Matějka and McKay (2014) study the flexible information acquisition problem using an entropy-based informativeness measure and justify a generalized logit decision rule. Caplin and Dean (2015) take an axiomatization approach and

[1] Steiner, Stewart, and Matějka (2017) assume the decision problem to be history dependent. Therefore, non-trivial dynamics remain in the optimal signal process. However, the dynamics are a results of the history dependence of the decision problem rather than the incentive to smooth information. In the dynamic learning foundation of Hébert and Woodford (2016), all signal processes are equally optimal because of a key no-discount assumption. They select a Gaussian process exogenously to justify a neighbourhood-based static information cost structure.

characterize decision rules that can be rationalized by an RI model. On the other hand, this chapter also serves as a foundation for RI models, as it characterizes, in detail, how the reduced-form decision rule is supported by acquiring information dynamically. In several limiting cases, my model completely reduces to a standard RI model.

The RI framework is widely used in models with strategic interactions (Matějka and McKay (2012), Yang (2015a), Yang (2015b), Matějka (2015), Denti (2015), etc). My work is different from these works as no strategic interaction is considered and the focus is on repeated learning. Despite the strategic component, Ravid (2018) also studies a dynamic model with repeated learning. In Ravid (2018), an RI buyer learns sequentially about the offers from a seller and the value of the object being traded. Similar to the DM in my model, the buyer systematically delays trading in equilibrium, and the stochastic delay resembles the arrival of a Poisson process.[2] However, in Ravid (2018), the delay is an equilibrium property that ensures the buyer's strategy is responsive to off-path offers. By contrast, the stochastic delay in my work is a property of an optimally smoothed learning process.

I use the reduction speed of uncertainty as a measure of the amount of information acquired per unit time. This measure captures the *posterior separability* from Caplin and Dean (2013). The posterior separable measure nests *mutual information* (introduced in Shannon (1948)) as a special case and is widely used in Gentzkow and Kamenica (2014), Clark (2016), Matyskova (2018), Rappoport and Somma (2017), etc. I provide an axiomatization for posterior separability based on the chain rule in Appendix A.1.4.1. Caplin, Dean, and Leahy (2017) axiomatize (uniform) posterior separability based on behavior data. Morris and Strack (2017) provide a dynamic foundation for posterior separability based on implementing an information structure with Gaussian learning. In addition to

[2]Precisely speaking, in the analysis of Proposition 2, Ravid (2018) shows that when quality is deterministic, the delay time distribution is exponential, which is the same as the stopping time induced by a Poisson signal process.

axiomatizing posterior separability, Frankel and Kamenica (2018) relates to my work in another interesting way. The *valid measure of information* defined in their paper coincides with the uncertainty reduction speed per unit arrival rate of a Poisson signal derived in this chapter.

1.2.3 Information design

In this chapter, I use a belief-based approach to model the choice of information. This approach is widely used for studying Bayesian persuasion models (Kamenica and Gentzkow (2011), Ely (2017), Mathevet, Perego, and Taneva (2017), etc.). An important methodology in this literature is the concavification method developed in Aumann, Maschler, and Stearns (1995) (based on Carathéodory's theorem). An alternative approach to model information is the direct signal approach [3] used in both information design problems, such as Bergemann and Morris (2017), and rational inattention problems. However, neither of the two methods applies to my dynamic information acquisition problem. I take the belief-based approach as in Bayesian persuasion models, but utilize a generalized concavification method developed in Chapter 4.

1.2.4 Stochastic control

Methodologically, this chapter is closely related to the theory of continuous-time stochastic control. The early theories study control processes measurable to the natural filtration of Brownian motion (see Fleming (1969) for a survey). The application of Bellman (1957)'s dynamic programming principle leads to the HJB equation characterization of the value function. On the contrary, the main stochastic control problem of this chapter has general martingale control process, which is a variant of the (semi)martingale models of stochastic control studied in Davis (1979), Boel and Kohlmann (1980), Striebel (1984), etc. However,

[3] This approach applies to settings where without loss of generality we can restrict the problem to considering only signals that are direct recommendations of actions.

none of the existing theories are sufficiently general to nest the stochastic control problem studied in this chapter. I introduce an indirect method that proves a verification theorem for a tractable HJB equation.

1.3 Model setup

The main model is a continuous-time stochastic control problem. A DM chooses an irreversible action at an endogenous decision time. The DM can control the information received before the decision time in a flexible manner, bearing a cost on information.

Decision problem: Time $t \in [0, +\infty)$. The DM discounts the delayed utility with rate $\rho > 0$. The DM is a vNM expected utility maximizer with Bernoulli utility associated with action-state pair $(a, x) \in A \times X$ at time t being $e^{-\rho t} u(a, x)$. Both the action space A and the state space X are finite. The DM holds a prior belief $\mu \in \Delta(X)$ about the state. Define $F(\nu) \triangleq \max_{a \in A} E_\nu[u(a, x)]$ given belief $\nu \in \Delta(X)$.

Information: I model information using a belief-based approach. A distribution of posterior beliefs is induced by an information structure according to Bayes rule *iff* the expectation of posterior beliefs is equal to the prior. Hence, in a static environment the choice of information can be equivalently formulated as the choice of a distribution of posterior beliefs (see Kamenica and Gentzkow (2011) for example). Extending this formulation to the dynamic environment studied here, I assume that the DM chooses the entire posterior belief process $\langle \mu_t \rangle$ in a nonparametric way. Now Bayes' rule should be satisfied at every instant of time—$\forall s > t$, the expectation of μ_s is μ_t. Thus, I restrict $\langle \mu_t \rangle$ to be a martingale, with $\langle \mathcal{F}_t \rangle$ as its natural filtration. A formal justification that choosing a belief martingale is equivalent to choosing a dynamic information structure is provided in Appendix A.1.4.

It is useful to define the following operator $\mathcal{L}_t$ for any $\langle\mu_t\rangle$ and $f : \Delta(X) \to \mathbb{R}$:

$$\mathcal{L}_t f(\mu_t) = E\left[\frac{\mathrm{d}f(\mu_t)}{\mathrm{d}t}\middle|\mathcal{F}_t\right] \triangleq \lim_{t'\to t^+} E\left[\frac{f(\mu_{t'}) - f(\mu_t)}{t'-t}\middle|\mathcal{F}_t\right]$$

By definition, $\mathcal{L}_t f$ captures the expected *speed* at which $f(\mu_t)$ increases. Let $\mathscr{D}(f)$ be the domain of $\langle\mu_t\rangle$ on which $\mathcal{L}_t f(\mu_t)$ is well defined.[4] For well-behaved Markov process $\langle\mu_t\rangle$ and $C^{(2)}$ smooth f, $\mathcal{L}f$ is the standard *infinitesimal generator* (subscript t omitted).

Cost of information: I assume that the flow cost of information depends on how fast the information reduces uncertainty. The flow cost of information is $C(I_t)$, where:

Assumption 1.1. *$I_t = -\mathcal{L}_t H(\mu_t)$, where $H : \Delta(X) \to \mathbb{R}$ is concave and continuous.*

I call H an *uncertainty measure*—because H is concave *iff* $E[H(\mu)]$ captures the Blackwell order on the belief distribution. By Assumption 1.1, I_t is the *speed* at which uncertainty falls when the belief updates. I call I_t the *(flow) informativeness measure*. One example of H is the entropy function $H(\mu) = -\sum \mu_x \log(\mu_x)$. Revelation of information reduces entropy; hence, the entropy reduction speed is a natural measure of the amount information. Assumption 1.1 is the main technical assumption in my analysis. I generalize this assumption in Section 1.7.2. For further discussions, see Appendix A.1.4, where I show that it is the continuous-time analog of "posterior separability" and provide an axiom for posterior separability.

Stochastic control: The DM solves the following stochastic control problem:

$$V(\mu) = \sup_{\langle\mu_t\rangle\in\mathbb{M},\tau} E\left[e^{-\rho\tau}F(\mu_\tau) - \int_0^\tau e^{-\rho t}C(I_t)\mathrm{d}t\right] \tag{1.1}$$

[4]Formally, $\langle\mu_t\rangle \in \mathscr{D}(f)$ if the uniform limit (w.r.t t) exists *almost surely*. Let $\mathscr{D} = \bigcap_{f\in C(\Delta X)} \mathscr{D}(f)$. $\mathscr{D}$ contains all Feller processes, whose transition kernels are *stochastically continuous* w.r.t. t and *continuous* w.r.t. state μ. However, $\mathscr{D}$ is much more general than Feller processes as it allows the transition kernel to be discontinuous in state μ.

where $\mathbb{M}$ is the set of all martingales $\langle \mu_t \rangle$ in $\mathscr{D}(H)$ with cadlag[5] path and satisfying $\mu_0 = \mu$, and τ is a $\langle \mathcal{F}_t \rangle$-measurable stopping time.[6]

The objective function in Equation (1.1) is fairly standard in canonical information acquisition problems. The DM acquires information that affects $\langle \mu_t \rangle$ and chooses stopping time τ to maximize the expected stopping payoff $E\left[e^{-\rho\tau}F(\mu_\tau)\right]$ less the total information cost $E\left[\int_0^\tau e^{-\rho t}C(I_t)\mathrm{d}t\right]$. The novel feature is that the DM is allowed to fully control $\langle \mu_t \rangle$, in contrast to canonical models, where the DM controls only a few parameters determining $\langle \mu_t \rangle$. The nonparametric control of the belief process exactly captures the flexible design of information by the DM.

I make the following assumption on the cost function $C(I)$ to generate incentive to smooth learning over time.

Assumption 1.2. $C : \mathbb{R}^+ \to \overline{\mathbb{R}}^+$ *is weakly increasing, convex and continuous.* $\lim\limits_{I\to\infty} C'(I) = \infty$.

The increasing and continuous cost function assumption is standard. The convexity of $C(I)$ and the condition $\lim C'(I) = \infty$ give the DM strict incentive to smooth the acquisition of information. Given Assumption 1.2, if the DM acquires all information immediately then uncertainty falls at infinite speed and the marginal cost $C'(I)$ is infinite, hence suboptimal.[7] I solve a special case violating Assumption 1.2 in Section 1.7.3, where I assume C to be linear. In this case the optimal strategy is to acquire all information at $t = 0$ (a static strategy).

[5]cadlag: $\mu_t : t \mapsto \Delta(X)$ is right continuous with left limits. Note that assuming martingale $\langle \mu_t \rangle$ being cadlag can be weakened to assuming $\langle \mathcal{F}_t \rangle$ being right continuous (see the martingale modification theorem in Lowther (2009)).

[6]I postpone the formal definition of integrability in Equation (1.1) to Section 1.5.1. For now, assume that the integral is well defined for all admissible strategies. Further discussions in Remark A.2 provide a formal justification that ignoring the integrability is innocuous.

[7]A weaker sufficient condition can guarantee information smoothing: $\sup_I \bar{\lambda} I - C(I) > \rho \sup F$, where $\bar{\lambda} = \underline{\lim}_{I\to\infty} \frac{C(I)}{I}$. This condition explicitly states that when I is sufficiently large, C is sufficiently convex that the utility gain from smoothing information dominates the loss from waiting longer. All the following theorems in this chapter are proved under this weaker condition.

In Example 1.1, I present a few examples of canonical Wald-type sequential learning models, each of which is a variant of Equation (1.1) with additional constraints on the set of admissible belief processes. Example 1.1 first illustrates how different learning technologies can be systematically compared under the same framework with an entropy-based cost function. The comparison also illustrates why a fully flexible learning framework is useful.

Example 1.1. Let the state be binary $X = \{l, r\}$. The prior belief of state $x = r$ is $\mu \in (0, 1)$. $A = \{L, R\}$. The DM wants to choose an action that matches the state: $u(L, l) = u(R, r) = 1$; $u(L, r) = u(R, l) = -1$. The discount rate $\rho = 1$, H is the standard entropy function: $H(\mu) = -\mu \log(\mu) - (1-\mu)\log(1-\mu)$, and the information cost $C(I) = \frac{1}{2}I^2$.

I consider three simple heuristic learning technologies: Gaussian learning, perfectly revealing breakthroughs and partially revealing evidence. A DM who uses a specific learning technology is modeled by restricting the admissible control set $\mathbb{M}$ to include only the corresponding family of processes. In each case, the DM controls a parameter that represents one aspect of learning.

1. *Gaussian learning*: The signal follows a Brownian motion whose drift is the true state, and whose variance is controlled by the DM. Therefore, the posterior belief follows a diffusion process (Bolton and Harris (1999)), so the set of admissible controls are:

$$\mathbb{M}_D = \{\langle \mu_t \rangle | \mathrm{d}\mu_t = \sigma_t \mathrm{d}W_t\}$$

The DM controls the signal precision $\langle \sigma_t \rangle$. According to Ito's lemma, $I_t = -\frac{1}{2}\sigma_t^2 H''(\mu_t) = \frac{\sigma_t^2}{2\mu_t(1-\mu_t)}$. This problem is studied in Moscarini and Smith (2001)[8], where the value func-

[8]With "belief elasticity" defined as $\mathscr{E}(\mu) = \mu(1-\mu)$ in my model.

tion is characterized by HJB:

$$\rho V_D(\mu) = \sup_{\sigma>0} \frac{1}{2}\sigma^2 V_D''(\mu) - \frac{1}{2}\left(\frac{\sigma^2}{2\mu(1-\mu)}\right)^2$$

The solution $V_D(\mu)$ is plotted as the blue curve in Figure 1.1. The shaded region is the experimentation region and the non-shaded region is the stopping region.

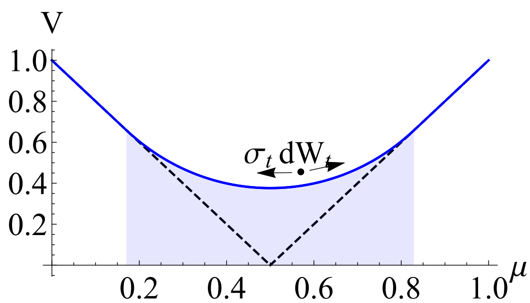


Figure 1.1: Incremental information

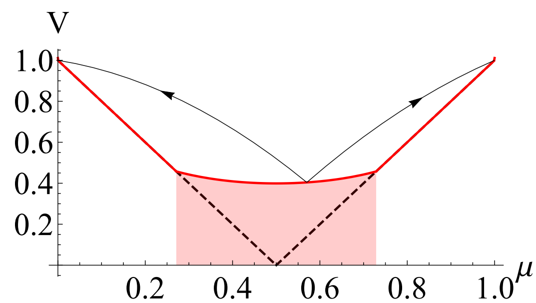


Figure 1.2: Breakthroughs

2. *Breakthroughs*: The DM observes breakthroughs that perfectly reveal the true state with arrival rate λ_t. Then, belief follows a Poisson process that jumps to 1 if the state is r and to 0 if the state is l. The set of admissible control is:

$$\mathbb{M}_B = \left\{\langle\mu_t\rangle | \mathrm{d}\mu_t = (1-\mu_t)\mathrm{d}J_t^1(\lambda_t\mu_t) + (0-\mu_t)\mathrm{d}J_t^0(\lambda_t(1-\mu_t))\right\}$$

$\langle J_t^i(\cdot)\rangle$ are independent Poisson counting processes with Poisson rate $(\cdot)$. The DM controls the signal frequency $\langle\lambda_t\rangle$. The Entropy reduction speed is $\lambda_t H(\mu)$. The HJB equation is as follows:

$$\rho V_B(\mu) = \sup_{\lambda>0} \lambda(\mu F(1) + (1-\mu)F(0) - V_B(\mu)) - \frac{1}{2}(\lambda H(\mu))^2$$

The solution V_B is plotted as the red curve in Figure 1.2. The two arrows show the belief jumps induced by breakthroughs at μ.

3. *Partially revealing evidence*: The DM allocates one unit of total attention to two news sources, each revealing one state with arrival rate $\gamma = 2$. Then belief follows a compensated Poisson process, and the set of admissible belief processes is:

$$\mathbb{M}_P = \left\{ \langle \mu_t \rangle \middle| \begin{aligned} \mathrm{d}\mu_t =& (1-\mu_t)(\mathrm{d}J_t^1(\alpha_t\gamma\mu_t) - \alpha_t\gamma\mu_t\mathrm{d}t) \\ &+ (0-\mu_t)(\mathrm{d}J_t^0((1-\alpha_t)\gamma(1-\mu_t)) - (1-\alpha_t)\gamma(1-\mu_t)\mathrm{d}t) \end{aligned} \right\}$$

$\langle J_t^i(\cdot)\rangle$ are independent Poisson counting processes with Poisson rate $(\cdot)$. The DM controls $\langle\alpha_t\rangle$, the attention allocated to the signal revealing state r. This control process is identical to that in Che and Mierendorff (2016). Applying their analysis, optimal α_t is a bang-bang solution, and the HJB equation is:

$$\begin{aligned} \rho V_P(\mu) = \max\Big\{ & \gamma\mu\big(F(1) - V_P(\mu) - V_P'(\mu)(1-\mu)\big) - \frac{1}{2}\big(\gamma\mu(H(\mu) + H'(\mu)(1-\mu))\big)^2, \\ & \gamma(1-\mu)\big(F(0) - V_P(\mu) - V_P'(\mu)(0-\mu)\big) - \frac{1}{2}\big(\gamma(1-\mu)(H(\mu) + H'(\mu)(0-\mu))\big)^2 \Big\} \end{aligned}$$

The solution V_P is plotted as the black curve in Figure 1.3. The optimal strategy is qualitatively the same as in Che and Mierendorff (2016). In the deep gray region, optimal learning direction is *confirmatory*: the arrival of news reveals the a priori more likely state (represented by solid arrows). In the light gray region, optimal learning direction is *contradictory*: the arrival of news reveals the a priori less likely state (dashed arrows).

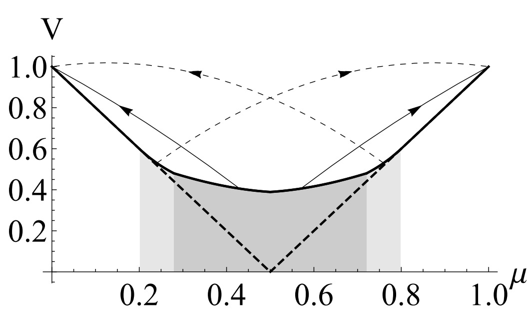


Figure 1.3: Partially revealing evidence

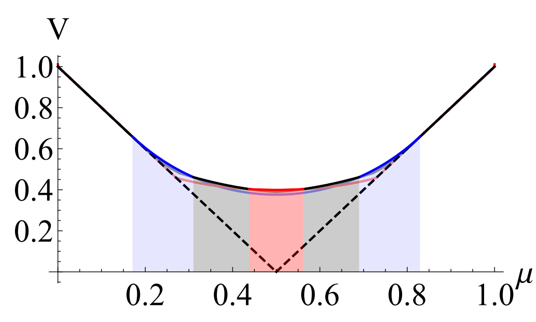


Figure 1.4: Comparison

In this example, the three learning technologies are analyzed for the same underlying decision problem and the same entropy cost function. Therefore, the utilities are directly comparable. I plot all three value functions in Figure 1.4 and use differently colored regions to illustrate the order of utility. Each color corresponds to a learning strategy being optimal: blue—Gaussian learning, red—breakthroughs, and gray—confirmatory evidence.[9] As shown in Figure 1.4, allowing the DM to use a rich set of strategies improves the decision-making quality.

More interestingly, there appears to be a pattern when optimizing in different aspects. When the prior belief is highly uncertain, a fully revealing Poisson signal that can bring the DM directly to a conclusion is optimal. When the prior belief is quite uncertain but asymmetrically in favor of one state, allocating attention to the more promising direction becomes optimal. When the prior belief is very certain, an imprecise but frequent Gaussian signal becomes optimal. The formal analysis for fully flexible information acquisition in Section 1.6 illustrates that this pattern is systematic: the optimal direction, precision and frequency of learning are exactly the relevant aspects and are closely related to the location of the prior belief.

1.3.1 Motivation for a flexible model

Example 1.1 implies that single-aspect models are insufficient for modeling a dynamic information acquisition problem with a rich strategy set. For instance, the model considering only partially revealing evidence predicts that seeking contradictory evidence is generally optimal when the belief is uncertain. However, further analysis shows that this prediction is misleading when Gaussian signals are also feasible. Studying a model where information acquisition is flexible in *all* aspects enables us to obtain insights about information acquisition without interference from any ad hoc restriction. Such insights

[9] In this example, whenever contradictory learning dominates confirmatory learning, contradictory learning is dominated by Gaussian learning, thus, contradictory learning is not optimal in any region.

include which aspect(s) is(are) endogenously salient for information acquisition and how each of these aspects is determined by the DM's incentives.

Although the results are derived in a fully flexible model, they apply to much more general settings where information acquisition is *not* flexible in all aspects. First, all results directly apply to all settings where information acquisition is flexible in those endogenously salient aspects, as all other aspects are redundant for implementing the unconstrained optimum. Second, even for settings where some of the relevant aspects are constrained, the intuitions from the flexible model identify the DM's most important incentive and how the hypothetically ideal strategy might be approximated by adjusting other aspects. In fact, the analysis of the flexible model in Sections 1.4 and 1.6 shows that the set of endogenously salient aspects is quite small, and the optimal strategy satisfies very simple qualitative properties in these aspects. Therefore, the findings of this chapter are useful in a very wide range of settings.

1.4 Dynamic programming and HJB equation

Solving Equation (1.1) is not an easy task due to the abstract strategy space. To the best of my knowledge, no general theory applicable to this stochastic control problem exists. The most closely related problems are studied in a set of remarkable papers on the martingale method in stochastic control (Davis (1979),Boel and Kohlmann (1980),Striebel (1984)). These papers introduce abstract formulations of stochastic control problems with general (semi)martingale control processes. The problems have finite horizon and specific objective functions; hence, they do not nest Equation (1.1).

Nevertheless, it is useful to introduce the general dynamic programming principle and HJB characterization. On the basis of the intuition of dynamic programming, the conjecture that $V(\mu_t)$ satisfies the following HJB is reasonable:

$$\max\Big\{\underbrace{F(\mu_t)-V(\mu_t)}_{\text{stopping value}},\ \underbrace{-\rho V(\mu_t)}_{\text{discount}}+\sup_{\mathrm{d}\mu_t}\Big\{\underbrace{\mathcal{L}_t V(\mu_t)}_{\text{continuation value}}-\underbrace{C(-\mathcal{L}_t H(\mu_t))}_{\text{control cost}}\Big\}\Big\}=0 \tag{1.2}$$

HJB Equation (1.2) is conceptually the same as the standard HJB equation. Recall the definition for operator $\mathcal{L}_t$, $\mathcal{L}_t V(\mu_t)$ is the flow utility gain from continuing. The exact form of $\mathcal{L}_t V$ and $\mathcal{L}_t H$ depends on the probability space, the filtration and the control process in the neighbourhood of t (which are summarized by the symbol $\mathrm{d}\mu_t$). Therefore, Equation (1.2) essentially states the dynamic programming principle: at any instance when the control is chosen optimally, either stopping is optimal (the first term is 0) or continuing is optimal and the net continuation gain equals the loss from discounting (the second term is 0).

For a simple example, let $\mathbb{M}$ be a family of Markov jump-diffusion belief processes, characterized by the following SDE:

$$\mathrm{d}\mu_t=\underbrace{(\nu(\mu_t)-\mu_t)(\mathrm{d}J_t(p(\mu_t))-p(\mu_t)\mathrm{d}t)}_{\text{compensated Poisson part}}+\underbrace{\sigma(\mu_t)\mathrm{d}W_t}_{\text{Gaussian diffusion}} \tag{1.3}$$

where $(p,\nu,\sigma):\mu_t\mapsto\mathbb{R}^+\otimes\Delta(\mathrm{Supp}(\mu))\otimes\mathbb{R}^{|\mathrm{Supp}(\mu)|-1}$ are control parameters, $J_t(\cdot)$ is a Poisson counting process with Poisson rate $(\cdot)$, and W_t is a standard one-dimensional Wiener process. Note that this example also nests all three families of strategies in Example 1.1 as special cases[10]. Itô's lemma implies an explicit form for the infinitesimal generator:

[10]The admissible control sets in the second and third cases in Example 1.1 are not exactly nested in Equation (1.3). However, they can be viewed as mixed strategies of pure Poisson-jump processes defined by Equation (1.3).

$$\mathcal{L}V(\mu) = \underbrace{p(V(\nu) - V(\mu) - \nabla V(\mu)(\nu - \mu))}_{\text{flow value of Poisson jump \& drift}} + \underbrace{\frac{1}{2}\sigma^T \mathrm{H}V(\mu)\sigma}_{\text{flow value of diffusion}}$$

where ∇ and H are the gradient and Hessian operators, respectively. By replacing $\mathcal{L}$ in Equation (1.2) with its explicit expression, we obtain a parametrized HJB Equation (1.4):

$$\begin{aligned}\rho V(\mu) = \max\Big\{\rho F(\mu), \sup_{p,\nu,\sigma} & \; p(V(\nu) - V(\mu) - \nabla V(\mu)(\nu - \mu)) + \frac{1}{2}\sigma^T \mathrm{H}V(\mu)\sigma \\ & - C\Big(p(H(\mu) - H(\nu) + \nabla H(\mu)(\nu - \mu)) - \frac{1}{2}\sigma^T \mathrm{H}H(\mu)\sigma\Big)\Big\}\end{aligned} \tag{1.4}$$

On the other hand, when $\mathbb{M}$ is the jump-diffusion family, the jump-diffusion control theory (see textbooks, e.g., Hanson (2007)) provides a *verification theorem* that proves that the value function for Equation (1.1) is exactly characterized by HJB Equation (1.4).

This simple example illustrates how a specific stochastic control problem relates to an HJB equation. Now, consider the general problem Equation (1.2) without any restriction on the admissible belief process. First, we require a verification theorem stating that the HJB Equation (1.2) characterizes the solution of Equation (1.1). Second, a representation theorem for the abstract operator $\mathcal{L}_t$ is also necessary to make Equation (1.2) practically tractable. The existing theories on martingale methods have little power for both tasks.[11] In Theorem 1.1, I achieve both goals by showing that the solution of Equation (1.1) is characterized by a simple parametric HJB equation:

Theorem 1.1. *Assume H is strictly concave and $C^{(2)}$ smooth on interior beliefs in $\Delta(X)$, Assumptions 1.1 and 1.2 are satisfied. Let $V(\mu) \in C^{(1)}\Delta(X)$ be a solution[12] to HJB Equation (1.4);*

[11] First, the existing martingale methods verify the HJB equation for different sets of problems that do not cover this specific problem. Moreover, the martingale method only states the existence of such $\mathcal{L}_t V$ (for example theorem 4.3.1 of Boel and Kohlmann (1980)) and does not provide an explicit representation. This issue is considered to be the main drawback of the martingale method (see discussions in Davis (1979)).

[12] The $C^{(1)}$ solution to the second-order ODE is not well defined. To be precise, V is a viscosity solution (see Crandall, Ishii, and Lions (1992)). In the viscosity solution, $\sigma^T \mathrm{H}V(\mu)\sigma$ is replaced by $D^2 V(\mu,\sigma)||\sigma||^2$,

then $V(\mu)$ solves Equation (1.1).

Theorem 1.1 first states that $V(\mu)$ is characterized by a HJB equation. More surprisingly, Theorem 1.1 also states that the HJB is exactly Equation (1.4). As a direct corollary, Equation (1.1) can be solved by considering only the family of Markov jump-diffusion processes characterized by SDE (1.3). The compensated Poisson jump part and Gaussian diffusion part in SDE (1.3) each represents a simple learning strategy.

- ***Poisson learning***: The DM uses *Poisson learning* or acquires a *Poisson signal* when a compensated Poisson part exists in the belief process. A Poisson jump in the belief process can be induced by observing non-conclusive news whose arrival follows a Poisson process. The compensating belief drift is induced by observing no news arriving. The control variables for Poisson learning are (p, ν), which represent three endogenously relevant aspects of Poisson learning. The arrival rate p represents the *frequency* of learning. The direction of belief jump represents the *direction* of learning. The magnitude of belief jump represents the *precision* of learning.

- ***Gaussian learning***: The DM uses *Gaussian learning* or acquires a *Gaussian signal* when a diffusion part exists in the belief process. Gaussian diffusion in the belief process can be induced by observing the realization of a Gaussian process, with state x being the unobservable drift. The flow variance σ represents the signal precision.

Equation (1.4) suggests that to determine the optimal strategy in all relevant aspects, the DM considers four types of trade-offs : (i) the standard continuing-stopping trade-off in optimal stopping problems, captured by the outer-layer maximization; (ii) the information cost-utility gain trade-off, which determines the total cost spent on learning;

where $D^2V(\mu,\sigma) = \overline{\lim_{\delta\to 0}} 2\frac{V(\mu+\delta\sigma)-V(\mu)-\nabla V(\mu)\delta\sigma}{\delta\|\sigma\|^2}$.

(iii) the Poisson-Gaussian trade-off, which determines the proportion of cost allocated to the Poisson signal (p, ν) and the Gaussian signal σ; (iv) the precision-frequency trade-off, which determines the marginal rate of substitution of signal frequency for precision. These trade-offs, especially the precision-frequency trade-off, will be discussed in detail to characterize the solution to Equation (1.4) in Section 1.6.

The proof of Theorem 1.1 uses an indirect method. I characterize Equation (1.1) as the limit of a series of auxiliary discrete-time problems. The discrete-time analyses are presented in Section 1.5. Readers interested in the solution of HJB Equation (1.4) can jump to Section 1.6.

1.5 The auxiliary discrete-time problem

In this section, I introduce the steps for proving Theorem 1.1 using an auxiliary discrete-time problem. First, in Section 1.5.1 I introduce a discrete-time stochastic control problem that converges to the continuous-time problem. Then I characterize the Bellman equation for the discrete-time problem in Section 1.5.2. In Section 1.5.3, I introduce a key lemma that links all the discrete-time analyses and proves Theorem 1.1.

1.5.1 Discrete-time problem

I consider a stochastic control problem that is a discrete-time analog of Equation (1.1). Then I illustrate the discretization of the original problem. The discretization serves as a useful intermediary showing that the discrete-time problem converges to the continuous-time problem.

Decision problem: The primitives (A, X, u, μ, ρ) are the same as those in Section 1.3. Time is discrete $t \in \mathbb{N}$, and the period length $dt > 0$. The payoff delayed by t periods is discounted by $e^{-\rho dt \cdot t}$.

Information: The DM chooses the posterior belief process $\langle \widehat{\mu}_t \rangle$ in a nonparametric way. $\langle \widehat{\mu}_t \rangle$ is restricted to be a martingale. Let $\langle \widehat{\mathcal{F}}_t \rangle$ be the natural filtration of $\langle \widehat{\mu}_t \rangle$.

Cost of information: Define $C_{dt}(I) \triangleq C(\frac{I}{dt})dt$. The per-period cost of information is assumed to be $C_{dt}(E[H(\hat{\mu}_t) - H(\hat{\mu}_{t+1})|\hat{\mathcal{F}}_t])$. Note that this is exactly the finite-difference analog of the flow cost $C(-\mathcal{L}_t H(\mu_t))$ in the continuous-time problem.

Optimization problem: The DM solves the following stochastic control problem:

$$V_{dt}(\mu) = \sup_{\langle\hat{\mu}_t\rangle \in \widehat{\mathbb{M}}, \hat{\tau}} E\left[e^{-\rho dt \cdot \hat{\tau}} F(\hat{\mu}_{\hat{\tau}}) - \sum_{t=0}^{\hat{\tau}-1} e^{-\rho dt \cdot t} C_{dt}\left(E\left[H(\hat{\mu}_t) - H(\hat{\mu}_{t+1})|\hat{\mathcal{F}}_t\right]\right)\right] \tag{1.5}$$

where $\widehat{\mathbb{M}}$ is the set of discrete-time martingales satisfying $\hat{\mu}_0 = \mu$, and τ is a $\langle\hat{\mathcal{F}}_t\rangle$–measurable stopping time. Note that in this section, all discrete-time stochastic processes and random variables are labeled with "hat" to differentiate them from continuous-time processes.

The purpose of analyzing the discrete-time problem is to characterize the continuous-time value function $V(\mu)$. Therefore, the first step is to show that $V_{dt}(\mu)$ approximates $V(\mu)$. To study the relation between $V_{dt}(\mu)$ and $V(\mu)$, let us discretize the objective function in Equation (1.1). For any admissible strategy $(\langle\mu_t\rangle, \tau)$, consider the Riemann sum:

$$W_{dt}(\mu_t, \tau) = \sum_{i=1}^{\infty} \text{Prob}(\tau \in [(i-1)dt, idt]) E\left[e^{-i\rho dt} F(\mu_{idt}) - \sum_{j=0}^{i-1} e^{-j\rho dt} C\left(I_{jdt}\right) dt\right]$$

where $I_{jdt} = E\left[\frac{H(\mu_{jdt}) - H(\mu_{(j+1)dt})}{dt}|\mathcal{F}_{jdt}\right]$. The objective function in Equation (1.1) is defined in the notion of the Riemann-Stieltjes integral as $\lim_{dt\to 0} W_{dt}(\mu_t, \tau)$. I call the martingale $\langle\mu_t\rangle$ *integrable* if the limit $\lim_{dt\to 0} W_{dt}(\mu_t, \tau)$ exists.[13] Unless otherwise stated, $\mathbb{M}$ is restricted to contain integrable processes, an innocuous restriction that enables me to avoid technical discussions of integrability.[14] Then it follows that $V(\mu) = \sup_{\langle\mu_t\rangle \in \mathbb{M}, \tau} \lim_{dt\to 0} W_{dt}(\mu_t, \tau)$.

[13]The standard definition for integrability also requires the limit to exist uniformly for all alternative nonuniform discretizations of the time horizon and all alternative measurable stopping times. Here I use the weaker integrability requirement for notational simplicity. The optimal strategy actually satisfies the stronger integrability requirements, so the current definition can be used without loss. The discretization of $\langle I_t\rangle$ is WLOG given the uniform convergence in the definition of $\mathscr{D}(H)$.

[14]The detailed discussion of why restricting belief to be integrable is innocuous is in Remark A.2.

Now, consider the relation between W_{dt} and V_{dt}. I argue that the objective function in Equation (1.5) is equivalent to $W_{dt}(\mu_t, \tau)$. This result can ben verified by noting that if $(\langle \mu_t \rangle, \tau)$ and $(\langle \widehat{\mu}_t \rangle, \widehat{\tau})$ jointly satisfy $\widehat{\mu}_t = \mu_{t \cdot dt}$ and $\widehat{\tau} = \lceil \tau / dt \rceil$, then:

$$W_{dt}(\mu_t, \tau) = E\left[e^{-\rho dt \widehat{\tau}} F(\widehat{\mu}_{\widehat{\tau}}) - \sum_{t=0}^{\widehat{\tau}-1} e^{-\rho dt \cdot t} C_{dt}\Big(E\Big[H(\widehat{\mu}_t) - H(\widehat{\mu}_{t+1}) | \widehat{\mathcal{F}}_t \Big] \Big) \right]$$

Given feasible strategy $(\langle \mu_t \rangle, \tau)$, such $(\langle \widehat{\mu}_t \rangle, \widehat{\tau})$ can be constructed by simply discretizing the continuous-time strategy. Given feasible strategy $(\langle \widehat{\mu}_t \rangle, \widehat{\tau})$, such $(\langle \mu_t \rangle, \tau)$ can be constructed by the Kolmogorov extension theorem. Therefore, it follows that $V_{dt}(\mu) = \sup_{\langle \mu_t \rangle \in \mathbb{M}, \tau} W_{dt}(\mu_t, \tau)$. Now that both V and V_{dt} are characterized using W_{dt}, W_{dt} can be used as an intermediary to link V and V_{dt}:

$$\begin{cases} V(\mu) = \sup\limits_{\langle \mu_t \rangle, \tau} \lim\limits_{dt \to 0} W_{dt}(\mu_t, \tau) \\ \\ \lim\limits_{dt \to 0} V_{dt}(\mu) = \lim\limits_{dt \to 0} \sup\limits_{\langle \mu_t \rangle, \tau} W_{dt}(\mu_t, \tau) \end{cases}$$

Clearly, V and $\lim V_{dt}$ are obtained by taking the limit of W_{dt} in different orders. Therefore, V_{dt} approximates V when the two limits are interchangeable, which is indeed true as proved in Lemma 1.1:

Lemma 1.1. *Given Assumption 1.1, $\forall \mu \in \Delta(X)$, $\lim_{dt \to 0} V_{dt}(\mu) = V(\mu)$.*

1.5.2 Discrete-time Bellman equation

Equation (1.5) is a discrete-time sequential optimization problem with bounded payoffs and exponential discounting. Therefore, standard dynamic programming theory applies and provides the Bellman equation that characterizes V_{dt}.

Lemma 1.2 (Discrete-time Bellman)**.** *V_{dt} is the unique solution in $C(\Delta X)$ of the following functional equation:*

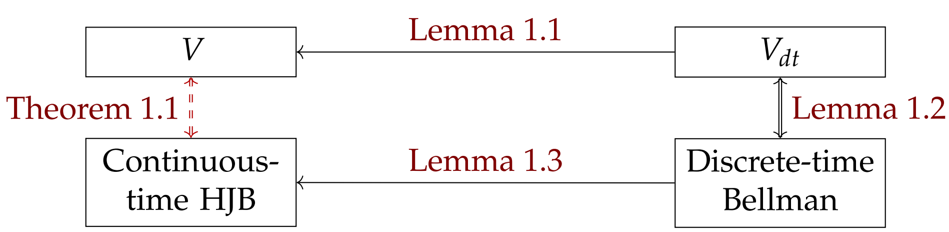
$$V_{dt}(\mu) = \max \left\{ F(\mu), \max_{p_i, \nu_i} e^{-\rho dt} \sum_{i=1}^{N} p_i V_{dt}(\nu_i) - C_{dt}\left(H(\mu) - \sum p_i H(\nu_i) \right) \right\} \tag{1.6}$$
$$s.t. \sum p_i \nu_i = \mu$$

where $N = 2|X|$, $p \in \Delta(N), \nu_i \in \Delta(X)$.

Equation (1.6) is a standard Bellman equation, except that it covers a restricted space of strategies. The choice of signal structure is restricted to have support size no larger than $2|X|$, while the original space contains signal structures with an arbitrary number of realizations. This simplification is based the generalized concavification methodology developed in Theorem 2 of Chapter 4. The standard concavification methodology is an application of the Carathéodory theorem to the graph of the objective function in the belief space.[15] Equation (1.6) involves an additional term $C_{dt}(H(\mu) - \sum p_i H(\nu_i))$, which makes the standard method inapplicable. The general method suggests that the maximum is characterized by concavifying a linear combination of V_{dt} and H.

1.5.3 Convergence and verification theorem

The following figure illustrates the roadmap for proving Theorem 1.1.



Theorem 1.1 is represented by the red dashed arrow on the left. The discrete-time problem's value function V_{dt} is the solution of the Bellman equation Equation (1.6) (the double arrow on the right, proved in Lemma 1.2). I have shown that V_{dt} converges to the continuous-time optimal control value V (the arrow on the top, proved in Lemma 1.1). In

[15]See Aumann, Maschler, and Stearns (1995) and Kamenica and Gentzkow (2011))

the next lemma, I show that solution of HJB Equation (1.4) is the limit of solution of Equation (1.6) (the arrow on the bottom, to be proved in Lemma 1.3). Therefore, the function solving HJB Equation (1.4) is the value function of the continuous-time stochastic control problem Equation (1.1).

Lemma 1.3. *Assume H is strictly concave and $C^{(2)}$ on interior beliefs, Assumption 1.2 is satisfied. Suppose $V(\mu) \in C^{(1)}$ is a solution to Equation (1.4). Then $V_{dt} \xrightarrow[dt \to 0]{L_\infty} V$.*

Lemma 1.3 proves that whenever Equation (1.4) has a solution, the solution is unique and coincides with the limit of solution to discrete-time problem Equation (1.6). Verification theorem Theorem 1.1 is a direct corollary of Lemmas 1.1, 1.2 and 1.3.

1.6 Optimal information acquisition

In this section I prove the existence of the solution to the continuous-time HJB Equation (1.4) and fully characterize the value and policy functions, assuming binary states and two forms of flow cost function: a hard cap and a smooth convex function. In both cases, the optimal strategies share the same set of qualitative properties. Then in Section 1.6.2, I discuss the key trade-offs in the optimization problem and provide the intuition for the optimal strategy. First, I introduce the assumptions for tractability:

Assumption 1.3.

1. *(Binary states):* $|X| = 2$.
2. *(Positive payoff):* $\forall \mu \in [0,1],\ F(\mu) > 0$.
3. *(Uncertainty measure):* $H''(\mu) < 0$ *and locally Lipschitz on* $(0,1)$, $\lim_{\mu \to 0,1} |H'(\mu)| = \infty$.

Assumption 1.3 comprises three parts. First, I restrict the state space to be binary. Therefore, the belief space is one dimensional, and I can use ODE theory to construct a candidate solution. Although the existence of the solution technically relies on the binary state assumption, the characterization generalizes to general state spaces, as discussed in

Appendix A.1.3. Second, I assume that the utility from decision making is strictly positive so that "delay forever" is strictly suboptimal. This restriction is made without loss of generality in the sense that we can always add a dummy "outside action" that gives ε payoff. Third, I assume that H is sufficiently smooth, strictly convex (which rules out free information) and satisfies an Inada condition (which guarantees a non-degenerate stopping region).

1.6.1 Main characterization theorem

Theorem 1.1 states that to characterize $V(\mu)$, it is sufficient to find a smooth solution to HJB Equation (1.4). I prove the existence of a solution and characterize the optimal strategy under Assumption 1.2-a or Assumption 1.2-b, two slightly stronger variants of Assumption 1.2.

Assumption 1.2-a (Capacity constraint)**.** *There exists c s.t.* $C(I) = \begin{cases} 0 & \text{when } I \leqslant c \\ +\infty & \text{when } I > c \end{cases}$

Assumption 1.2-a restricts the cost function C to be a hard cap: information is free when its measure is below capacity c and infinitely costly when it exceeds this capacity.[16] This condition forces the DM to smooth the information acquisition process over time.

Theorem 1.2. *Given Assumptions 1.1, 1.2-a and 1.3, there exists a quasi-convex value function $V \in C^{(1)}(0,1)$ solving Equation (1.4). Let $E = \{\mu \in [0,1] | V(\mu) > F(\mu)\}$ be the experimentation region. There exists policy function $\nu : E \to [0,1]$ satisfying:*

$$\rho V(\mu) = -c\frac{F(\nu(\mu)) - V(\mu) - V'(\mu)(\nu(\mu) - \mu)}{H(\nu(\mu)) - H(\mu) - H'(\mu)(\nu(\mu) - \mu)}$$

where $\nu(\mu)$ is unique a.e. and satisfies the following properties. $\exists \mu^* \in \arg\min V$ s.t.

[16] $\lim_{I\to\infty} C'(I)$ is not well defined with Assumption 1.2-a. However, it is not hard to see that Assumption 1.2-a still satisfies the weaker formulation discussed in Footnote 7. As a result, Theorem 1.1 applies with Assumption 1.2-a.

1. *Poisson learning:* $\rho V(\mu) > -c\frac{V''(\mu)}{H''(\mu)}\ \forall \mu \in E \backslash \mu^*$.
2. *Direction:* $\mu > \mu^* \implies \nu(\mu) > \mu$ *and* $\mu < \mu^* \implies \nu(\mu) < \mu$.
3. *Precision:* $|\nu(\mu) - \mu^*|$ *is decreasing in* $|\mu - \mu^*|$ *on each interval of* E.
4. *Stopping time:* $\nu(\mu) \in E^C$ *(a successful experiment lands in the stopping region).*

Theorem 1.2 proves the existence of a solution to Equation (1.4) and characterizes the optimal policy function. The theorem first states that the optimal value function is implemented by a *Poisson signal*, i.e., seeking a breakthrough that causes the belief to jump to $\nu(\mu)$. Moreover, property 1 states that the Gaussian signal is strictly dominated, except for at most one critical belief. Therefore, as discussed in Section 1.4, the optimal strategy is Poisson learning, which can be characterized by three aspects of learning and the stopping time.

Direction: Property 2 states that the optimal direction is *confirmatory*: when $\mu > \mu^*$, the DM holds a high prior belief for state 1 and acquires a signal whose arrival induces an even higher posterior belief $\nu(\mu)$ and vice versa for $\mu < \mu^*$.

Precision: Property 3 states that the optimal precision measured by $|\nu(\mu) - \mu^*|$ is *negatively related* to how certain the belief is (measured by $|\mu - \mu^*|$). Since $\mu^* \in \arg\max V$, the property equivalently states that precision is negatively related to the continuation value.

Frequency: With Assumption 1.2-a, frequency is automatically determined given the precision, according to $p(\mu) = -\frac{c}{H(\nu(\mu)) - H(\mu) - H'(\mu)(\nu(\mu) - \mu)}$. Thus, the optimal frequency is *positively related* to the continuation value.

Stopping time: Property 4 states that the image of ν is always in the stopping region. In other words, the optimal stopping time is exactly the signal arrival time.

By combining these properties, we can qualitatively determine the optimal learning dynamics. The DM seeks a signal that arrives according to a Poisson process. The arrival of the signal confirms the DM's prior belief and is sufficiently accurate to warrant an immediate action. Absent the arrival of a Poisson signal, the DM becomes less certain about the state, following Bayes' rule. The DM's continuation value decreases correspondingly; hence, she continues seeking a Poisson signal with lower frequency and higher precision.

Assumption 1.2-b (Convex cost). $C \in C^{(2)}\mathbb{R}^+$, $C(0) = 0$, $C'(I) \geqslant 0$, $C''(I) > 0$, $\lim_{I\to\infty} C'(I) = \infty$.

Assumption 1.2-b restricts the cost function C to be $C^{(2)}$ smooth and strictly convex: acquiring an additional unit of information is of strictly increasing marginal cost. The condition on $\lim C'(I)$ in Assumption 1.2 is retained. If we replace Assumption 1.2 with Assumption 1.2-b, we obtain the following characterization theorem:

Theorem 1.3. *Given Assumptions 1.1, 1.2-b and 1.3, there exists a quasi-convex value function $V \in C^{(1)}(0,1)$ solving Equation (1.4). Let $E = \{\mu \in [0,1] | V(\mu) > F(\mu)\}$ be the experimentation region. There $\exists$ policy functions $\nu : E \to [0,1]$ and $I \in C^{(1)}(E)$*[17] *satisfying:*

$$\rho V(\mu) = - I(\mu) \cdot \frac{F(\nu(\mu)) - V(\mu) - V'(\mu)(\nu(\mu) - \mu)}{H(\nu(\mu)) - H(\mu) - H'(\mu)(\nu(\mu) - \mu)} - C(I(\mu))$$

where ν and I are unique a.e. and satisfy the following properties. $\exists \mu^* \in \arg\min V$ s.t.

1. *Poisson learning: $\rho V(\mu) > \max_{\sigma} \frac{1}{2}\sigma^2 V''(\mu) - C(-\frac{1}{2}\sigma^2 H''(\mu))$ $\forall \mu \in E \backslash \mu^*$.*
2. *Direction: $\mu > \mu^* \implies \nu(\mu) > \mu$ and $\mu < \mu^* \implies \nu(\mu) < \mu$.*
3. *Precision: $|\nu(\mu) - \mu^*|$ is decreasing in $|\mu - \mu^*|$ on each interval of E.*
4. *Stopping time: $\nu(\mu) \in E^C$.*
5. *Intensity: $I(\mu)$ is increasing in $V(\mu)$.*

With the exception of property 5, the optimal strategy has the same set of properties as Theorem 1.2. Property 5 states that the informativeness measure I of the optimal signal

[17] Note that given ν, selecting I or p is equivalent. They uniquely pin down each other according to equation $I(\mu) = p(\mu)(-H(\nu(\mu)) + H(\mu) + H'(\mu)(\nu(\mu) - \mu))$.

is higher when the continuation value is higher. Since the belief process drifts downward the value function conditional on continuation, the DM invests less in information acquisition as time passes.

The intuition for property 5 is discussed in Moscarini and Smith (2001). The marginal gain from experimentation is proportional to the continuation value while marginal cost is increasing in I. Therefore, the optimal cost is increasing in the value function. This property is called "value-level monotonicity" in Moscarini and Smith (2001), where the level (flow variance of the diffusion process) is a parameter for both the cost and precision of a Gaussian signal. My analysis identifies this intuition separately from another important trade-off between signal precision and frequency. I refer to property 5 as "value-intensity monotonicity". Here I rename parameter I the *intensity* of learning, which is more intuitive and concise than "informativeness measure".

Examples

In this section, I first provide a minimal working example that illustrates Theorem 1.3 in Example 1.2. Then I provide supplementary examples to illustrate a rich set of implications of my model, including multiple phases of learning in Example 1.3 and learning from a one-sided search in Example 1.4.

Example 1.2. Consider the problem studied in Example 1.1. $F(\mu) = \max\{2\mu - 1, 1 - 2\mu\}$, $H(\mu) = -\mu \log(\mu) - (1-\mu)\log(1-\mu)$, $\rho = 1$, and $C(I) = \frac{1}{2}I^2$. No parametric assumption is placed on the set of admissible belief process.

The solution is presented in Figures 1.5 and 1.6. In Figure 1.5-(a), dashed lines depict $F(\mu)$, the blue curve depicts $V(\mu)$, and the blue shaded region is experimentation region E. Figure 1.5-(b) shows the optimal posterior $\nu(\mu)$ as a function of the prior. As stated in Theorem 1.3, the policy function is piecewise smooth and decreasing. The three arrows in Figure 1.5-(a) depict the optimal strategies prescribed at three different priors. The arrows start at the priors and point to the optimal posteriors. The blue curve in Figure 1.5-(c) shows the optimal intensity $I(\mu)$ as a function of the prior. Clearly, $I(\mu)$ is isomorphic to $V(\mu)$ in the experimentation region.

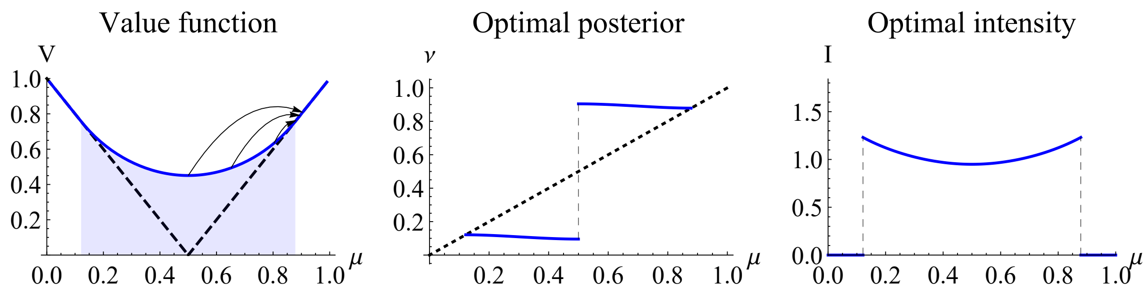


Figure 1.5: Value and policy functions

Figure 1.6 illustrates the dynamics of the optimal policy. Figure 1.6-(a) depicts the optimal belief process. Conditional on no signal arrival, the posterior belief drifts towards the critical belief level $\mu^* = 0.5$. In this example, two *phases* of learning occur (represented by different colors of shaded regions in Figure 1.6-(a)). In the first phase (blue region), the DM seeks a Poisson signal to confirm the most likely state. As time passes, the signal precision increases while signal frequency and learning intensity decreases (as in Figure 1.6-(b)&(c)). Eventually, the DM believes that the two states are equally likely and switches to the second phase (gray region). In the second phase, she seeks two signals that confirm each state in a balanced way such that before any signal arrives her posterior belief is stationary.

Recall the three learning technologies in Example 1.1. They approximate the full solution in Example 1.2. In general, the optimal signal is a confirmatory Poisson signal with varying precision and frequency. However, in Example 1.1, the precision and frequency of the confirmatory Poisson signal are exogenously fixed. Therefore, for very certain prior beliefs, the ideal high-frequency Poisson signal is approximated by a Gaussian signal. For very uncertain prior beliefs, the ideal signal is approximated by acquiring perfectly

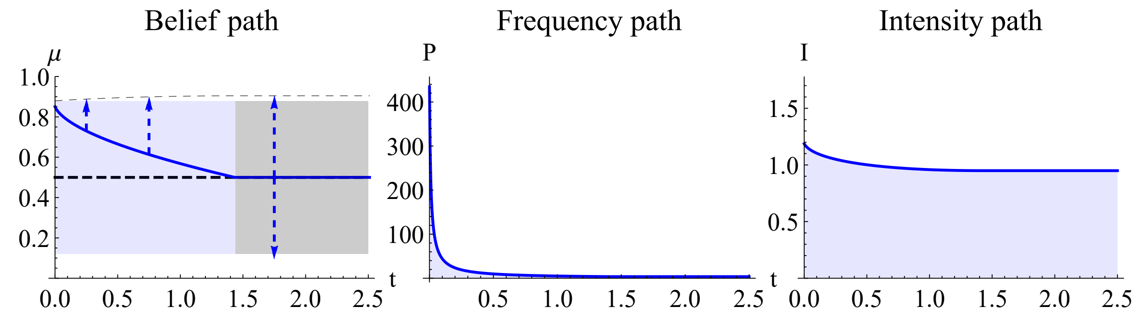


Figure 1.6: Dynamics of optimal policy

revealing breakthroughs with low frequency.

Example 1.3 (Multiple phases)**.** Figure 1.7 depicts an example with four actions, whose expected payoffs are represented by the four dashed lines in Figure 1.7-(a). The two blue dashed lines are called riskier actions, and the two red dashed lines are called safer actions. The upper envelope of the four lines is $F(\mu)$. The experimentation region contains three disjoint intervals. For the middle interval, in the red regions, the DM has a more extreme belief and searches for a signal that confirms a safer action (red arrow). In the blue region, the DM has a more ambiguous belief and searches for a riskier action (blue arrow). Figure 1.7-(c) depicts the optimal belief process with a prior belief in the red region. The experimentation follows three phases, the DM searches for a safer action in phase 1, searches for a riskier action in phase 2 and searches in a balanced way in phase 3.

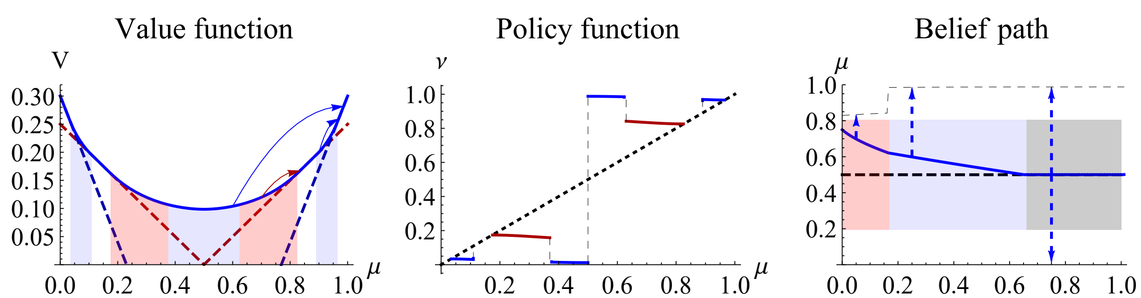


Figure 1.7: Example with four alternatives

Example 1.4 (One-sided search)**.** Figure 1.8 depicts an example where the optimal strategy includes only one-sided search. A safe action with deterministic payoff and a risky action whose payoff is higher than that of the safe action in state 1 exists. As illustrated in Figure 1.8-(a), both $F(\mu)$ and $V(\mu)$ are monotonically increasing. According to property 1, $\mu > \mu^*$ in the entire experimentation region E. Figure 1.8-(b) shows that the optimal strategy is always to search for a Poisson signal that induces a posterior belief higher than the prior. Figure 1.8-(c) shows that in this example, only one phase occurs. If no signal arrives before the belief reaches to the critical belief, the optimal solution is for the DM to stop learning and choose the safe action.

This example illustrates more precisely the definition of confirmatory evidence: the optimal belief jump is in the direction of a more *profitable* state. The profitability of a state depends jointly on its likelihood and the corresponding payoff of the actions. In this example, consider a prior belief less than 0.5. Although state 0 is more likely, since it is dominated by state 1 for any action, state 1 is unambiguously more profitable to learn about. Therefore, the optimal confirmatory evidence is always revealing state 1.

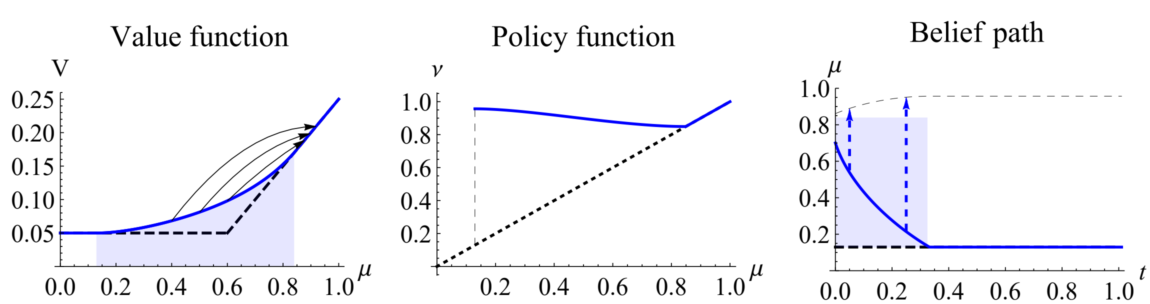


Figure 1.8: Example with one-sided search

1.6.2 *Proof methodology and key intuitions*

In Section 1.3, I introduce four types of trade-offs. Now, I discuss the trade-offs in detail and illustrate how they determine the optimal strategy in each salient aspect. I first derive a geometric characterization of the optimal policy in Section 1.6.2.1. Then, I discuss how the key trade-offs are represented by the geometric characterization and provide intuitions for the optimal policy. In Section 1.6.2.2, I present the sketch of a proof for Theorem 1.2.

1.6.2.1 Geometric representation and key trade-offs

A though experiment is useful to gain intuition. Fix the value function V and consider a simplified optimization problem:

$$\sup_{p\geqslant 0,\nu} p\big(V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)\big)-C\big(p\big(H(\mu)-H(\nu)+H'(\mu)(\nu-\mu)\big)\big) \tag{1.7}$$

Equation (1.7) is more restrictive than Equation (1.4). I assume that the DM acquires only a Poisson signal. Let us temporarily ignore the Gaussian signal. Define:

$$\begin{cases} U(\mu,\nu)=V(\nu)-V(\mu)-V'(\mu)(\nu-\mu) \\ J(\mu,\nu)=H(\mu)-H(\nu)+H'(\mu)(\nu-\mu) \end{cases}$$

The interpretation of $U(\mu,\nu)$ is the flow value per unit arrival rate from a Poisson signal with posterior ν. Similarly, $J(\mu,\nu)$ is the flow uncertainty reduction per unit arrival rate from the Poisson signal. Then Equation (1.7) can be rewritten as:

$$\begin{aligned} &\sup_{p\geqslant 0,\nu} p\cdot U(\mu,\nu)-C(p\cdot J(\mu,\nu)) \\ \overset{I\triangleq p\cdot J(\mu,\nu)}{\Longleftrightarrow} &\sup_{I\geqslant 0,\nu}\left(\frac{U(\mu,\nu)}{J(\mu,\nu)}\right)\cdot I-C(I) \end{aligned}$$

The problem is separable in choosing I and ν. The solution (ν^*, I^*) is characterized by:

$$\begin{cases} \nu^* \in \arg\max_{\nu} \frac{U(\mu,\nu)}{J(\mu,\nu)} \\ C'(I^*) = \max_{\nu} \frac{U(\mu,\nu)}{J(\mu,\nu)} \end{cases}$$

The optimal posterior ν^* maximizes $\frac{U(\mu,\nu)}{J(\mu,\nu)}$—the value to uncertainty reduction ratio. Let $\lambda = C'(I^*) = \max \frac{U(\mu,\nu)}{J(\mu,\nu)}$; then, $U(\mu,\nu) \leqslant \lambda J(\mu,\nu)$ and the equality holds at ν^*.[18] Define $G(\mu) = V(\mu) + \lambda H(\mu)$. I call $G(\mu)$ the *gross value function*. Then, the definition of U and V implies $U(\mu,\nu) - \lambda J(\mu,\nu) = G(\nu) - G(\mu) - G'(\mu)(\nu - \mu)$. Hence, $U(\mu,\nu) \leqslant \lambda J(\mu,\nu)$ implies that the gross value function has the following property:

$$\begin{cases} G(\nu) \leqslant G(\mu) + G'(\mu)(\nu - \mu) & \forall \nu \in [0,1] \\ G(\nu^*) = G(\mu) + G'(\mu)(\nu^* - \mu) \end{cases} \tag{1.8}$$

Equation (1.8) states that $G(\nu)$ is everywhere (weakly) below the tangent line of G at μ, except $G(\mu)$ and $G(\nu^*)$ touch the tangent line. The tangent line is linear (hence concave) and thus weakly dominates G's upper concave hull $\text{co}(G)$. Therefore, $G(\mu) = \text{co}(G)(\mu)$ and $G(\nu^*) = \text{co}(G)(\nu^*)$. See Figure 1.9 for a graphical illustration.

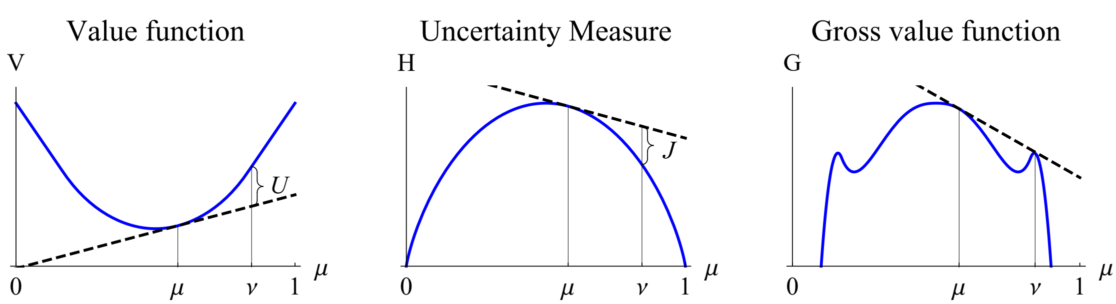


Figure 1.9: Concavification of the gross value function

[18] With Assumption 1.2-a, $I^* = c$ and $\lambda = \max \frac{U(\mu,\nu)}{J(\mu,\nu)}$ is the Lagrangian multiplier for constraint $I \leqslant c$.

Figure 1.9-(a) and Figure 1.9-(b) depict the value function V and the uncertainty measure H, respectively. Figure 1.9-(c) depicts the gross value function $G = V + \lambda H$, where λ is calculated for the prior μ. As discussed, G touches the upper concave hull at both μ and ν^*. When ν^* is unique, μ and ν^* are the two boundary points of the *concavified region* (the interval (μ, ν) on which $G < \text{co}(G)$).

Equation (1.8) is called a *concavification* characterization as it is an analog to the concavification method in Bayesian persuasion problems. The difference is that in a Bayesian persuasion problem, the boundary points of a concavified region are optimal posteriors, whereas in the current problem, the prior is also on the boundary of a concavified region. This property has clear economic meaning. G is called the gross value function because it integrates value function V and uncertainty measure H using marginal cost level λ. λ is a multiplier that captures the marginal effect of reducing uncertainty on flow cost. Therefore, solving:

$$\sup_{p \geqslant 0, \nu} p(G(\nu) - G(\mu) - G'(\nu)(\nu - \mu)) \tag{1.9}$$

is equivalent to solving Equation (1.7). Whether Equation (1.9) yields a positive payoff depends on whether $G(\mu) < \text{co}(G)(\mu)$. Suppose $G(\mu) < \text{co}(G)(\mu)$. Then, there is a strictly positive gain from information and Equation (1.9) is strictly positive. However, Equation (1.9) is linear in the signal arrival rate p. As a result the DM has incentive to increase p, which drives up marginal cost $C'(\cdot)$. Thus, when the optimum is reached, $C'(\cdot)$ (or λ) must be such that solving Equation (1.9) yields exactly zero utility: $G(\mu) = \text{co}(G)(\mu)$. This characterization illustrates that in the continuous time limit, information is smoothed such that uncertainty is reduced by only an infinitesimal amount at every instant of time.

Now, suppose that the HJB is satisfied, i.e., Equation (1.7) equals the flow discounting loss $\rho V(\mu)$. Then applying $I^* = p^* \cdot J(\mu, \nu^*)$ and $C'(I^*) = \frac{U(\mu,\nu^*)}{J(\mu,\nu^*)}$ to the HJB implies:

$$\begin{aligned} \rho V(\mu) &= p^* \cdot U(\mu, \nu^*) - C(p^* \cdot J(\mu, \nu^*)) \\ \implies \rho V(\mu) &= I^* C'(I^*) - C(I^*) \end{aligned} \tag{1.10}$$

Combining Equation (1.8) and Equation (1.10) dentifies the value function V and corresponding strategies p, ν.[19] Now, I analyze key trade-offs in the dynamic information acquisition problem by studying Equations (1.8) and (1.10).

1. Utility gain vs. information cost

Equation (1.10) illustrates the utility gain vs. information cost trade-off. Since C is a convex function, $IC'(I) - C(I)$ is increasing in I[20], that is, the optimal flow informativeness measure I is isomorphic in continuation value $V(\mu)$. This property is exactly the "value-intensity monotonicity" I introduced in Section 1.6.1.

The intuition for this property is simple. The marginal cost of increasing the intensity of the signal *proportionately* is $IC'(I)$. The marginal gain is obtained from increasing the arrival rate proportionately (keeping the signal precision fixed, as in the envelope theorem). Increasing the arrival rate by a unit proportion reduces the waiting time by the same proportion, so the marginal gain from increasing I by a unit proportion is discount ρV plus cost $C(I)$. At the optimum, the maginal cost equals the margina gain; therefore, we obtain Equation (1.10) and the flow informativeness is monotonic in value function.

If we consider the case with Assumption 1.2-a, then λ in Equation (1.8) is replaced by the shadow cost of increasing informativeness (see Footnotes 18 and 19). Equation (1.10) can be written as $\rho V(\mu) = c\lambda$. Although the intensity is fixed, in this case, a monotonicity between the shadow cost and value function remains.

[19] With Assumption 1.2-a, $C(I^*) = 0$ and $I^* = c$. Thereofre, $\rho V(\mu) = \lambda c$.

[20] $\frac{\mathrm{d}}{\mathrm{d}I}(IC'(I) - C(I)) = IC''(I) \geqslant 0$

In summary, by studying the utility gain vs. information cost trade-off, I established a monotonicity between the shadow/marginal cost λ and the continuation value $V(\mu)$. (I refer to both as the "value-intensity monotonicity" for notational simplicity.) Now that I characterized λ, we can proceed to Equation (1.8).

2. Precision vs. frequency

A novel trade-off characterized by Equation (1.8) is the precision vs. frequency trade-off. The value-intensity monotonicity determines I from the value function. Now, the DM allocates total intensity I to precision (parametrized by the size of belief jumps) and frequency (parametrized by the arrival rate of jumps). Equation (1.8) suggests that the optimal signal precision can be solved by concavifying the gross value function $G(\mu)$. In this section, I illustrate how this trade-off changes for different priors and explain the intuition.

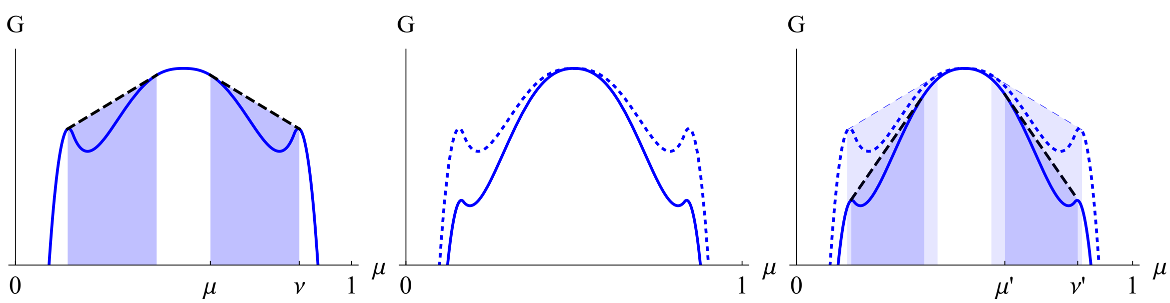


Figure 1.10: Precision-frequency trade-off

Figure 1.10 shows how varying λ affects the optimal jump size. In Figure 1.10-(a) the blue curve is $G(\mu)$, and the dashed curve is $\text{co}(G)$. I call the blue region, where $G(\mu) < \text{co}(G)(\mu)$, the *concavified region* and the white region, where $G(\mu) = \text{co}(G)(\mu)$, the *globally concave region*. The prior μ and optimal posterior ν are on the boundary of a concavified region. Consider $G_1 = V + \lambda_1 H$, where $\lambda_1 > \lambda$. Figure 1.10-(b) depicts both G (the dashed curve) and G_1 (the blue curve). Since G_1 is G plus a strictly concave function, any belief in the globally concave region of G is still in the globally concave region of G_1. As a result, as

λ increases, the white region is expands and the blue region contracts (see Figure 1.10-(c)). Thus, the prior and optimal posterior move closer together. Recall that λ is monotonic in V, which means the DM is more willing to choose a signal that induces shorter belief jump when the continuation value is higher.

The intuition for this property is as follows. When the DM is more certain about the state, the continuation value is higher; hence, the utility loss from discounting is higher. The DM wants to receive a signal more frequently to benefit from the high value sooner. In other words, the marginal rate of substitution of frequency for precision is increasing in the continuation value. In this analysis, the continuation value is isomorphic to λ, which controls the shape of G. The marginal rate of substitution of frequency for precision is exactly captured by the global concavity of the gross value function; thus, the analysis presented by Figure 1.10 exactly illustrates the intuition.

Confirming vs. contradicting: The analysis above determines the magnitude of the optimal belief jump. The optimal jump direction remains to be determined to pin down the optimal posterior. Now, I show that the precision-frequency trade-off also implies the optimality of confirmatory learning.

Let us hypothetically consider a belief μ at which jumping toward the right is optimal (weakly). In both panels of Figure 1.11, μ is the prior and ν_L, ν_R are optimal posteriors on each side of μ. Jumping to ν_R (the black arrow) is better than jumping to ν_L (the dashed black arrow). Let V be increasing around μ. Now consider the DM's incentive at μ_1 slightly larger than μ (in Figure 1.11-(a)). Although the corresponding optimal posteriors could also move, keeping them fixed at ν_L and ν_R has only a second-order effect on utility. We can compare ν_L and ν_R to pin down the optimal posterior for μ_1. Since $\mu_1 > \mu$, ν_R is closer to prior, and ν_L is farther from prior. Moreover, $V(\mu_1) > V(\mu)$ implies that the DM has a stronger preference for frequency to precision with belief μ_1. Since $V' > 0$, the effect is first order. Therefore, ν_R is strictly preferred to ν_L at μ_1. Consider μ_2 slightly smaller than μ (in Figure 1.11-(b)). A similar analysis shows that now size of jump to ν_R is larger, and the DM has a stronger preference for precision with belief μ_2. Thus, ν_R is also strictly optimal for μ_2.

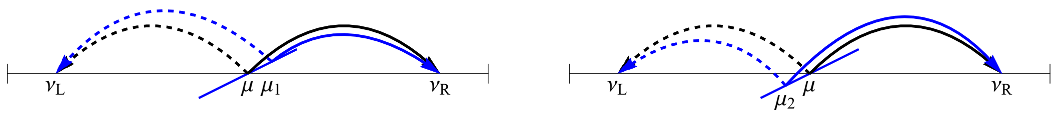


Figure 1.11: Confirmatory v.s. contradictory

In this analysis, jumping in the direction of increasing value function means the signal is confirmatory. When value function is quasi-convex, this property is equivalent to property 2 of Theorems 1.2 and 1.3. Therefore, the precision-frequency trade-off implies that the incentive for confirmatory learning is self-enforcing.

3. Poisson vs. Gaussian

Thus far, I have ignored the possibility of Gaussian signals. In fact, Gaussian signals are implicitly modeled in Equation (1.8). Consider the optimization w.r.t. Gaussian signals:

$$\begin{aligned}
&\sup_{\sigma} \sigma^2 V''(\mu) - C(-\sigma^2 H''(\mu)) \\
\Longrightarrow\ & \text{FOC} : V''(\mu) + \lambda H''(\mu) = 0 \\
\Longleftrightarrow\ & G''(\mu) = 0
\end{aligned} \tag{1.11}$$

where $\lambda = C'(-\sigma^2 H''(\mu))$ with Assumption 1.2-a or $\lambda = \frac{\rho}{c} V(\mu)$ with Assumption 1.2-b. Comparison of Equations (1.8) and (1.11) shows that Equation (1.11) is exactly the limit of Equation (1.8) when optimal posterior ν converges to prior μ. This result is intuitive since a Gaussian signal can be approximated as a Poisson signal with very low precision and high arrival rate.

The comparison of Gaussian and Poisson signals is effectively the comparison of a special imprecise Poisson signal and other Poisson signals. Therefore, this trade-off is a special case of the precision-frequency trade-off. Selecting a Gaussian signal is a corner solution when the DM wants to sacrifice almost all precision for frequency—a slightly less patient DM is willing to avoid any waiting and stop immediately, while a slightly more patient DM is willing to wait for a more precise Poisson signal. Therefore, the Gaussian signal is optimal only on the boundaries of the experimentation regions. Given this intuition, one could imagine that the Gaussian signal is generically suboptimal except for special cases where the precision-frequency trade-off is invariant. Since the preference between precision and frequency depends on the loss from delaying, the trade-off is invariant only when the DM does not discount future payoffs. This intuition is confirmed in a no-discounting special case in Section 1.7.1, as well as in the model of Hébert and Woodford (2016).

4. Continuing vs. stopping

Consider the optimal stopping time. Theorems 1.2 and 1.3 states that repeated jumps are suboptimal. I prove by showing that repeated jumps can be improved by a direct jump. Let ν be the optimal posterior for prior μ (see Figure 1.12). Then, Equation (1.8) implies that $\frac{U_0}{J_0} = \frac{U_0'}{J_0'} = \lambda(\mu)$.

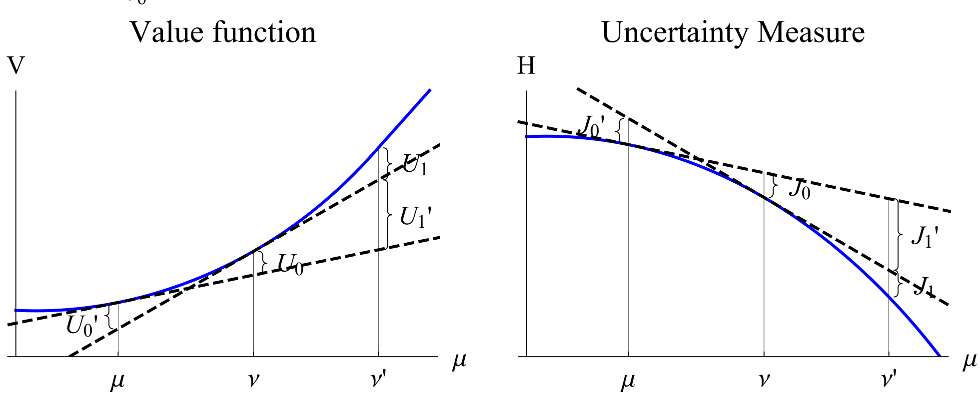


Figure 1.12: Continuing vs. stopping

Hypothetically, imagine that at ν, it is optimal to continue, and the optimal posterior is ν'. Then, $\frac{U_1}{J_1} = \lambda(\nu)$, and $\lambda(\nu) > \lambda(\mu)$ by the confirmatory evidence property and value-intensity monotonicity. I want to show that this result implies $\frac{U(\mu,\nu')}{J(\mu,\nu')} = \frac{U_1+U_1'}{J_1+J_1'} > \lambda(\mu)$, i.e., jumping to posterior ν' directly is strictly better than a two-step jump. By elementary geometry, there exists α s.t $U_1' = \alpha U_0$ and $J_1' = \alpha J_0$.[21] Therefore, the value to uncertain reduction ratio $\frac{U(\mu,\nu')}{J(\mu,\nu')} = \frac{U_1+\alpha U_0}{J_1+\alpha J_0}$ is a weighted average of $\frac{U_0}{J_0}$ and $\frac{U_1}{J_1}$, which is larger than $\lambda(\mu)$.

The intuition for the stopping rule is now clear. If we combine a two-step jump into a direct jump, the flow utility gain is a weighted sum of that of the two jumps. The flow uncertainty reduction is exactly the same weighted sum of that of the two jumps. Therefore, the net value from a direct jump is a weighted average of the net values from each jump. As a result, sequentially jumping to higher values is dominated by directly jumping to the highest value.

Remark 1.1.

The intuition behind the value-intensity monotonicity is driven purely by convexity of cost function h and is clearly independent of the formulation of the information measure. The intuition behind the optimality of a Poisson signal over a Gaussian signal is the use of the precision-frequency trade-off to compare a generic Poisson signal with an extremely imprecise Poisson signal. The result does not depend on the exact form of I. I generalize the optimality of a Poisson signal to the generic cost of information in Theorem 1.5, Section 1.7.2. I also discuss confirmatory evidence and immediate stopping properties with generic cost functions in Section 1.7.2.

The precision-frequency trade-off also does not depend on the size of the state space. I confirm this result via a general characterization of optimal strategy with more states

[21] See Figure 1.12. $\frac{U_0}{J_0} = \frac{U_0'}{J_0'} = \lambda(\mu)$ implies $\frac{U_1'}{J_1'} = \lambda(\mu)$, hence, $\frac{U_1'}{U_0} = \frac{J_1'}{J_0}$. I assume the ratio to be α.

(Theorem A.3) in Appendix A.1.3. However, the binary states assumption is crucial for proving the existence of the solution to the HJB equation. A constructive proof of the binary state case based on ODE theory is introduced in Section 1.6.2.2.

Our discussion thus far does not rely on the exact form of λ. The qualitative properties of all these trade-offs depend only on the monotonicity of λ in continuation value, which is true with both Assumptions 1.2-a and 1.2-b. Therefore, when I introduce the sketch of the proof, I discuss only Theorem 1.2, and the proof extends to Theorem 1.3.

1.6.2.2 Sketched proof of Theorem 1.2

I prove Theorem 1.2 by construction and verification. I conjecture that the optimal policy for Equation (1.4) takes the form of Theorem 1.2: a single confirmatory signal associated with an immediate action. I first construct $V(\mu)$ and $\nu(\mu)$ via three steps:

- *Step 1*. Determine μ^*. Since $\mu^* \in \arg\min V$, except for the special case where V is strictly monotonic, μ^* is essentially the unique belief at which $V'(\mu^*) = 0$, and searching for posteriors on either side of μ^* is equally good. The HJB equation implies:

$$\sup_{\nu \leqslant \mu^*} \frac{F(\nu)}{1 + \frac{\rho}{c} J(\mu^*, \nu)} = \sup_{\nu \geqslant \mu^*} \frac{F(\nu)}{1 + \frac{\rho}{c} J(\mu^*, \nu)}$$

  $V(\mu^*)$ and $\nu(\mu^*)$ are pinned down correspondingly. The special case occurs when F is strictly monotonic. Take $F' > 0$ for example. μ^* is the smallest belief that $\rho F(\mu) \leqslant \sup_{\nu \geqslant \mu} -c \frac{F(\nu) - F(\mu) - F'(\mu)(\nu - \mu)}{J(\mu, \nu)}$, and vice versa for $F' < 0$.

- *Step 2*. Solve for the value function while holding the action fixed. Let a be the optimal action for optimal posterior ν solved in step 1. Let $F_a(\mu) = E_\mu[u(a, x)]$. Now, solve for the value function given payoff $F_a(\nu)$:

$$\rho V(\mu) = \max_{\nu \geqslant \mu} -c \frac{F_a(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{H(\nu) - H(\mu) - H'(\mu)(\nu - \mu)}$$

The primitives in the objective function are all sufficiently smooth in ν. Then, the first-order condition w.r.t. ν yields a well-behaved ODE characterizing $\nu(\mu)$ with initial condition $\nu(\mu^*)$. Therefore, we can solve for the optimal policy ν and calculate value $V(\mu)$ accordingly for $\mu \geqslant \mu^*$. $V(\mu)$ and $\nu(\mu)$ for all $\mu \leqslant \mu^*$ are solved by a symmetric process.

- *Step 3*. Update the value function w.r.t. all alternative actions and smoothly paste the solved value function piece by piece. This step begins with solving the ODE defined in step 2 at μ^*. Then, I extend the value function towards $\mu = \{0, 1\}$. Whenever I reach a belief at which two actions yield the same payoff, I setup a new ODE with the new action. This process continues until the calculated value function $V(\mu)$ smoothly pastes to $F(\mu)$. This procedure generates a quasi-convex value function (minimized at μ^*).

Solving the ODE characterizing $\nu(\mu)$ directly implies monotonicity of $\nu(\mu)$ in each connected experimentation region. Now, I need to verify the optimality of the constructed strategy. The verification takes three steps, which rule out repeated jumps, contradictory evidence and Gaussian signals. The intuition for the suboptimality of these three alternative strategies is explained in Section 1.6.2. The formal proof is relegated to Appendix A.2.3.

1.7 Discussion

In this section, I discuss, in detail, the assumptions I make in the baseline model, which can be categorized into three classes.

1. *Economic assumptions:*
    - Discounting (positive ρ).
    - Informativeness measure (Assumption 1.1).
    - Convexity of cost function (Assumption 1.2).
2. *Restrictive assumptions:* Finite actions and binary states (Assumption 1.3).

3. *Technical assumptions:* Smoothness and positiveness assumptions (Assumption 1.3).

The economic assumptions are crucial for my results and deserve an in-depth discussion. To illustrate the role of discounting, in Section 1.7.1, I discuss the case with no discounting but a flow waiting cost, and show that without discounting, the trade-off between precision and frequency diminishes and the dynamics of information become irrelevant. In Section 1.7.2, I generalize Assumption 1.1 to general information measures and show that a Poisson signal almost always strictly dominates a Gaussian signal. I also explain that immediate action and confirmatory learning properties are tightly tied to Assumption 1.1. To illustrate the role of Assumption 1.2, I discuss the case where the cost function is linear in Section 1.7.3 and show that without convexity, the optimal strategy is static.

The restrictive assumptions do limit the generality of the model. However, relaxing them does not fundamentally alter the key intuition, and the methodology generalizes. The discussion of these assumptions is relegated to the appendix. In Appendix A.1.2, I relax the finite action assumption and show that the problem with a continuum of actions can be approximated well by adding actions. In Appendix A.1.3, I relax the binary state assumption. Although the constructive proof of existence no longer works with the general state space, I show that all the properties in Theorem 1.2 extend. The technical assumptions do not restrict my model in a meaningful way and are therefore not discusses.

1.7.1 Linear delay cost

As is discussed in Section 1.6.2, discounting is the key factor driving all the dynamics. With exponential discounting, the trade-off between the arrival frequency and precision of signals changes according to the continuation value. A sensible conjecture is that if we replace exponential discounting with linear discounting, i.e., the DM pays a fixed flow cost of delay, the time distribution of the utility gain and information cost no longer matters to the DM. In fact, this conjecture is correct. Consider the following problem:

$$V(\mu) = \sup_{\langle\mu_t\rangle\in\mathbb{M},\tau} E\left[F(\mu_\tau) - m\tau - \int_0^\tau C(I_t)\mathrm{d}t\right] \tag{1.12}$$

Theorem 1.4. *Given Assumptions 1.1 and 1.2, suppose $V(\mu)$ solves Equation (1.12); then:*

$$V(\mu) = \sup_{P\in\Delta^2(X),\lambda>0} E_P[F(\nu)] - \frac{m + C(\lambda)}{\lambda} E_P[H(\mu) - H(\nu)]$$

Theorem 1.4 illustrates that solving Equation (1.12) is equivalent to solving a static rational inattention problem, with $\frac{m+C(\lambda)}{\lambda}$ being the marginal cost on the information measure (see Caplin and Dean (2013) and Matějka and McKay (2014)). The optimal value function can be obtained through various learning strategies. Assuming (P^*, λ^*) to be the solution to the problem in Theorem 1.4, then *all* dynamic information acquisition strategies that eventually implement P^* (i.e., $\mu_\infty \sim P^*$) and incur flow cost λ^* achieve the same utility level $V(\mu)$.[22]

Note that in Equation (1.12), the utility depends on the decision time only through expected delay $E[\tau]$. Therefore, the previous analysis implies that all dynamic information acquisition strategies that eventually implement P^* and incur flow cost λ^* have the same expected delay. This result suggests that the cost structure specified by Assumptions 1.1 and 1.2 has the property that all learning strategies are equally *fast* on expectation, but they might differ in terms of *riskiness*. The linear delay cost case is a knife-edge case where the DM is risk neutral on the time dimension and, consequently, all learning strategies are equally good.

When the DM discounts delayed payoffs, as is assumed in the main model, she is risk loving on the time dimension; therefore, the DM prefers a riskier strategy. Intuitively,

[22] This result is stated and proved formally in Chapter 3.

the riskiest information acquisition strategy is a "greedy strategy" that maximizes the probability of early decision (at the cost of a high probability of long delays as the expected delay is fixed). The confirmatory Poisson learning strategy exactly resembles such a greedy strategy. The key property of the strategy is that all resources are used in verifying the conjectured state directly, and no intermediate step exists before a breakthrough. Alternative strategies, such as Gaussian learning and contradictory Poisson learning all involve the accumulation of substantial intermediate evidence to conclude a success. The intermediate evidence accelerates future learning and hence hedges the risk of decision time. Moreover, the decision time is further dispersed by acquiring signals with decreasing frequency.

Equation (1.12) is the dynamic learning foundation provided in Hébert and Woodford (2016) to justify Gaussian learning.[23] The analysis of Equation (1.12) suggests that a linear delay cost is a knife-edge case.

1.7.2 General information measure

Technically, Assumption 1.1 helps throughout the entire analysis. The methodology of concavifying "the gross value function" is possible only when the expected utility gain and information measure take consistent forms. However, I want to show that one key feature of the baseline model—the optimality of Poisson learning—does not depend on this assumption. Let $J(\mu,\nu)$ and $\kappa(\mu,\sigma)$ be bivariate functions. Consider the following functional equation:

$$\rho V(\mu) = \max\left\{\rho F(\mu), \sup_{p,\nu,\sigma^2} p\big(V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)\big) + \frac{1}{2}\sigma^2 V''(\mu)\right\} \tag{1.13}$$

$$\text{s.t. } pJ(\mu,\nu) + \kappa(\mu,\sigma) \leqslant c$$

[23]In Hébert and Woodford (2016), informativeness measures that are more general than Assumption 1.1 are also considered in the Appendix.

The objective function of Equation (1.13) is exactly the same as that of Equation (1.4) with Assumption 1.2-a. I assume that the DM controls a jump-diffusion belief process. The gain from information is the same as before. I assume $J(\mu,\nu)$ to be an arbitrary function that is both prior and posterior dependent. The cost of the diffusion signal is $\kappa(\mu,\sigma)$. I impose the following assumptions on $J(\mu,\nu)$ and $\kappa(\mu,\sigma)$.

Assumption 1.4.

1. $J \in C^{(4)}(0,1)^2$.
2. $\forall \mu \in (0,1)$, $J(\mu,\mu) = J'_\nu(\mu,\mu) = 0$, *and* $J''_{\nu\nu}(\mu,\mu) > 0$.
3. $\kappa(\mu,\sigma) = \frac{1}{2}\sigma^2 J''_{\nu\nu}(\mu,\mu)$.

First, J is assumed to be sufficiently smooth to eliminate technical difficulties. $J(\mu,\mu) = 0$ is the implication of "an uninformative Poisson signal is free".[24] $J'_\nu(\mu,\mu) = 0$ and $J''_{\nu\nu}(\mu,\mu) > 0$ are implications of "any informative Poisson signal is costly". Within this continuous time framework, these assumptions are imposed on J without loss of generality. The crucial assumption is the third condition: $\kappa(\mu,\sigma) = \frac{1}{2}\sigma^2 J''_{\nu\nu}(\mu,\mu)$. This assumption states that the cost functional is "continuous" in the space of the signal structures. Consider a Poisson signal (p,ν). When $\nu \to \mu$, the utility gain from learning this signal is:

$$p\big(V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)\big) = p\left(\frac{1}{2}V''(\mu)(\nu-\mu)^2 + O|\nu-\mu|^3\right)$$

Therefore, (p,ν) approximates a Gaussian signal with flow variance $p(\nu-\mu)^2$. Meanwhile, the cost of this signal is:

$$\begin{aligned} pJ(\mu,\nu) =& p\left(J(\mu,\mu) + J'_\nu(\mu,\mu)(\nu-\mu) + \frac{1}{2}J''_{\nu\nu}(\mu,\mu)(\nu-\mu)^2 + O(|\nu-\mu|^3)\right) \\ =& \frac{1}{2}p(\nu-\mu)^2 J''_{\nu\nu}(\mu,\mu) + pO(|\nu-\mu|^3) \end{aligned}$$

[24] In this setup, $J(\mu,\mu) = 0$ is WLOG. If an uninformative signal has a strictly positive cost, we can always shift the capacity constraint c to normalize $J(\mu,\mu)$ to 0.

Hence, if the cost of a Gaussian signal is consistent with the cost of imprecise Poisson signals in the limit, $\kappa(\mu,\sigma) = \frac{1}{2}\sigma^2 J''_{\nu\nu}(\mu,\mu)$.

Theorem 1.5. *Given Assumption 1.4, suppose $V \in C^{(3)}(0,1)$ solves Equation (1.13), and let $L(\mu)$ be defined by:*

$$L(\mu) = \frac{\rho}{c} J''_{\nu\nu}(\mu,\mu)^2 - \frac{2J^{(3)}_{\nu\nu\mu}(\mu,\mu)^2 + J^{(3)}_{\nu\nu\nu}(\mu,\mu)J^{(3)}_{\nu\nu\mu}(\mu,\mu)}{J''_{\nu\nu}(\mu,\mu)} + J^{(4)}_{\nu\nu\nu\mu}(\mu,\mu) + J^{(4)}_{\nu\nu\mu\mu}(\mu,\mu)$$

Then in the open region: $D = \left\{\mu \middle| V(\mu) > F(\mu) \text{ and } L(\mu) \neq 0\right\}$, the set of μ s.t.:

$$\rho V(\mu) = c\frac{V''(\mu)}{J''_{\nu\nu}(\mu,\mu)}$$

is of zero measure.

The interpretation of Theorem 1.5 is that a Poisson signal is almost always strictly superior to the diffusion signal. In the experimentation region where $L(\mu) \neq 0$, $V(\mu)$ can be achieved by a diffusion signal only at a zero measure of points. $L(\mu) = 0$ is a partial differential equation on $J(\mu,\nu)$ in the diagonal of space. Therefore, the set of points that $L(\mu) = 0$ could contain an interval only when $J(\mu,\nu)$ is a local solution to the PDE. The solution to a specific PDE is a non-generic set in the set of all functions satisfying Assumption 1.4. In this sense, for an arbitrary information measure $J(\mu,\nu)$, the optimal policy function contains a diffusion signal almost nowhere.

A trivial sufficient condition for $L(\mu) \neq 0$ is Assumption 1.1. Assumption 1.1 implies that $J^{(2)}_{\nu\nu}(\mu,\nu)$ is invariant in μ. In this case $L(\mu) = \frac{\rho}{c}J''_{\nu\nu}(\mu,\mu)^2 > 0$ for certain. The first corollary of Theorem 1.5 characterizes D when J is almost locally posterior separable. $\forall f \in C^{(1)}(0,1)^2$, define a norm $\|f(\cdot)\|_\delta = \sup_{x\in[\delta,1-\delta]}\left\{|f(x,x)|, \|\nabla f(x,x)\|_{L^2}\right\}$.

Corollary 1.5.1. *Given Assumption 1.4, suppose $V \in C^{(3)}(0,1)$ solves Equation (1.13); then, for* *any $\delta > 0$, there exists ε s.t. if $\left\| J^{(3)}_{\nu\nu\mu} \right\|_{\delta} \leqslant \varepsilon$, then in the interval $[\delta, 1-\delta]$ the set of μ s.t.:*

$$\rho V(\mu) = c \frac{V''(\mu)}{J''_{\nu\nu}(\mu, \mu)}$$

is of zero measure.

The condition in Corollary 1.5.1 states that $J''_{\nu\nu}(\mu, \nu)$ is approximately constant over μ for ν close to μ. This result verifies my analysis in Section 1.6.2.1 that the comparison of Poisson and Gaussian signals relies only on the local properties of J. Another simple sufficient condition for $L(\mu) \neq 0$ is high impatience or low learning capacity.

Corollary 1.5.2. *Given Assumption 1.4, suppose $V \in C^{(3)}(0,1)$ solves Equation (1.13). Then, for any $\delta > 0$, there exists Δ s.t. if $\frac{\rho}{c} \geqslant \Delta$, then in the interval $[\delta, 1-\delta]$, the set of μ s.t.:*

$$\rho V(\mu) = c \frac{V''(\mu)}{J''_{\nu\nu}(\mu, \mu)}$$

is of zero measure.

Corollaries 1.5.1 and 1.5.2 complement the discussion in Section 1.7.1 and illustrate the complete picture of how the DM's incentives pin down the optimal learning dynamics. First, when Assumption 1.1 holds, Theorem 1.4 implies that the cost structure does not favor any learning strategy. Any positive discount rate gives the DM incentive to choose a Poisson signal. All learning strategies, including Gaussian learning, become equally optimal only when time preference is risk neutral. Second, when Assumption 1.1 is violated by a small amount, then even though the cost structure might favor a Gaussian signal, the incentive is dominated by discounting. Third, when the cost structure provides arbitrarily strong incentive for a Gaussian signal, sufficiently high discount rate overweights the incentive.

Although *Poisson learning* is generally optimal, *immediate action* and *confirmatory evidence* are implications of Assumption 1.1. Imagine a case in which high-precision signals are relatively inexpensive (e.g., $J(\mu,\nu)$ is truncated both below and above). Then, when the prior is close to the boundary of the stopping region, seeking confirmatory evidence (with low precision and high frequency) results in very high cost, whereas seeking for a precise contradictory signal is inexpensive. Searching for a contradictory signal causes the belief to drift rapidly toward the more likely state, which effectively enables quick confirmation. Therefore, the contradictory signal becomes optimal. In fact, this example has the same intuition as the findings in Che and Mierendorff (2016). In their setup, the DM allocates limited attention to two exogenous Poisson signals, each revealing a state. When the DM is more uncertain, their model predicts that the DM acquires a confirmatory signal. However, near the stopping boundary, their model predicts a contradictory signal, as the contradictory signal approximates an infeasible confirmatory signal with low precision and high frequency.

On the other hand, consider the immediate action property. Imagine a case in which low-precision signals are inexpensive. Then, breaking a long jump into multiple short jumps may be profitable. The immediate action property is called the single experiment property (SEP) in Che and Mierendorff (2016). In their paper, SEP is also shown not to be a robust property in a generic Poisson learning model.

1.7.3 Linear flow cost

In this subsection, I study the case where the flow cost $C(I)$ is a linear function. Assumption 1.2 is replaced by the following assumption:

Assumption 1.2′ (Linear flow cost)**.** *Function h is defined by* $C(I)=\lambda I$, $\lambda>0$.

The convexity of $C(I)$ in Assumption 1.2 gives the DM incentive to smooth the acquisition of information. When $C(I)$ is a linear function, the optimal value is achieved by acquiring all the information and immediately making a decision.

Theorem 1.6. *Given Assumptions 1.1 and 1.2′, suppose $V(\mu)$ solves Equation (1.1), then:*

$$V(\mu) = \sup_{P \in \Delta^2(X)} E_P[F(\nu)] - \lambda E_P[H(\mu) - H(\nu)] \tag{1.14}$$

The intuition for this result is simple. At any instant in time, suppose that the optimal decision is to continue learning for a positive amount of time. The value is the discounted future value at the next instant of time $(t + \mathrm{d}t)$ less the flow cost of information. Now, consider moving the learning strategy at $t + \mathrm{d}t$ to the current period. Then, both the future value at $t + \mathrm{d}t$ and the cost are discounted by $\mathrm{d}t$ less. If the net utility gain from learning at $t + \mathrm{d}t$ is nonnegative, then this operation increases the current utility by reducing the waiting time.[25] If the net utility gain from learning at $t + \mathrm{d}t$ is negative, then stopping learning immediately increases current utility. This operation can always be applied recursively and strictly improves the strategy until all information is acquired at period 0.[26]

In fact, given Assumptions 1.1 and 1.2′, Equation (1.1) is a variant of the more general model in Steiner, Stewart, and Matějka (2017), which considers a varying state and repeated decision making. With linear cost function $C(I)$, no motivation for smoothing the learning behavior exists. The dynamics in Steiner, Stewart, and Matějka (2017) are a result of the intertemporal dependence of decision problems.

[25]This step utilizes Assumption 1.2′, which implies that the cost of a combined signal structure is the sum of the cost of each of them.

[26]Strictly speaking, an immediate learning strategy is not admissible because its belief path is not cadlag. However, there always exists a way to implement a signal structure in an arbitrarily short period of time, and the payoff approximates the immediate learning payoff.

1.8 Applications

1.8.1 Choice accuracy and response time

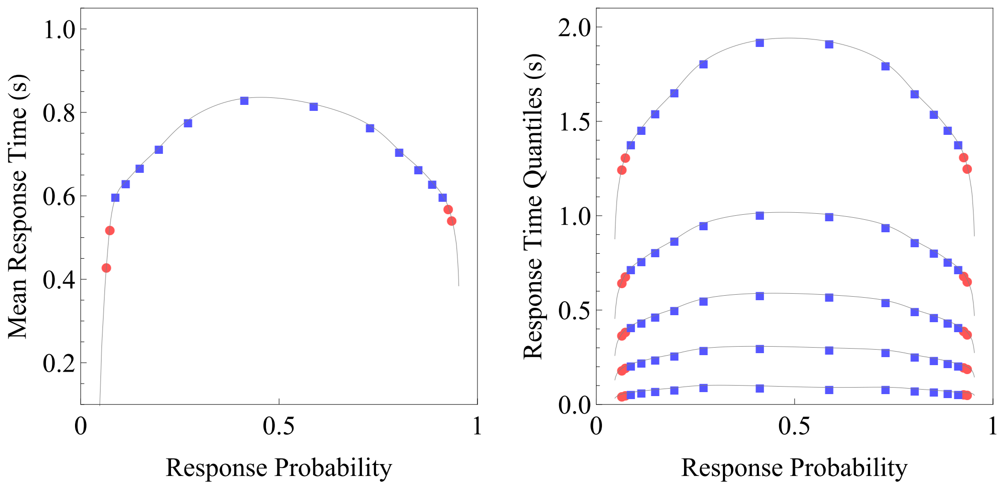
The two-choice sequential decision making problem has been extensively studied in the psychological and behavioral studies. One of the key objective is to explain the data on choice accuracy and response time from experiments. The drift-diffusion model (DDM) has been the most popular theoretical model for these decision problems, for the reason that DDM is very tractable and fits the accuracy/ response time data well. However, accounting for the *joint distribution* of choice accuracy and response time remains a challenge for DDM. In this section, I apply my model to predict a systematic feature in the data: the *crossover* of response time-accuracy relationship.

The crossover happens when the difficulty of decision problem varies: the error responses are faster than the correct responses when the task is easy; the error responses are slower than the correct responses when the task is hard (see Luce et al. (1986), Ratcliff, Van Zandt, and McKoon (1999)). First, I illustrate the crossover of time-accuracy relationship in Example 1.5.

Example 1.5. Consider the same decision problem as in Example 1.1. $F(\mu) = \max\{1-2\mu, 2\mu-1\}$ and $\rho = 1$. Assume prior belief $\mu_0 = 0.5$ and let $H_0(\mu)$ be the entropy function. Define uncertainty measure $H(\mu)$ as:

$$H(\mu) = \begin{cases} H_0(\mu) & \text{if } \mu \in [0.5, 0.65] \\ H_0(\mu) - |\mu - 0.5|^3 & \text{if } \mu < 0.5 \\ H_0(\mu) - 4|\mu - 0.65|^3 & \text{if } \mu > 0.65 \end{cases}$$

$H(\mu)$ is an asymmetric uncertainty measure, and $H(\mu)$ is slightly more concave than H_0 when $\mu < 0.5$ or $\mu > 0.65$. The different difficulty levels are modeled as different capacity constraints on $-\mathcal{L}H(\mu_t)$, the higher the capacity constraint is, the easier the decision problem is. I study the joint distribution of choice and decision time conditional on the true state being r ($\mu = 1$). Figure 1.13 depicts the *latency-probability (LP)* and *quantile-probability*



Left panel: The latency-probability function (the thin line) and the data points simulated from 8 difficulty levels. Right panel: The quantile-probability functions (the thin lines, from bottom to top: 0.1, 0.3, 0.5, 0.7, 0.9 quantile) and the data points simulated from 8 difficulty levels. The correct responses are to the right of 0.5, the errors are to the left of 0.5. Red points: the errors have shorter response times. Blue points: the errors have longer response time.

Figure 1.13: LP and QP plots

(QP) plots. The horizontal coordinates of the points to the right of $p = 0.5$ shows the choice probability of the action R (the correct choice). Each such point has a corresponding point to the left of $p = 0.5$ showing the remaining probability of the action L (the error). The vertical coordinates of all points show the response time measured by mean (in LP plot) or by quantiles (in QP plot).

The crossover of time-accuracy relationship is illustrated by the differently colored points. The red points are data points where the errors happen earlier than the correct responses (measured by both mean or quantiles). They are simulated with high capacity, thus are of higher accuracy in general. On the contrary, the blue points are data points where the errors happen later than the correct responses. They are simulated with low capacity, and of low accuracy in general. In fact, Figure 1.13 is qualitatively the same as the LP and QP plots documented in Ratcliff and Rouder (1998) and Ratcliff, Van Zandt, and McKoon (1999).

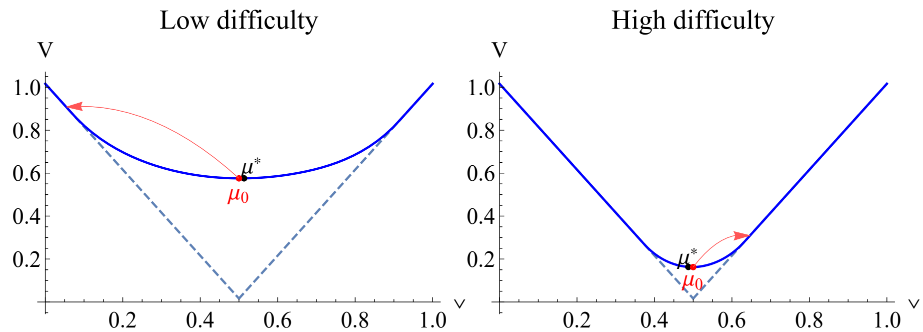


Figure 1.14: The critical beliefs of different difficulty levels

The main reason for the crossover is explained in Figure 1.14. When the capacity is low (the task difficulty is high), the optimal size of belief jump is small. By construction of $H(\mu)$, when the posterior belief is not far away from μ_0, learning the state L is more costly than learning the state R. As a result, the critical belief μ^* at which searching for both direction is indifferent is biased toward left. Since $\mu_0 > \mu^*$, the correct responses are font-loaded. Applying the same intuition, when the capacity is high, $\mu_0 < \mu^*$ and the errors are font-loaded.

Applying the idea from Example 1.5, creating a crossover of μ^* and μ_0 is necessary for creating a crossover of the response time-accuracy relationship.

Proposition 1.1. *Suppose $|A| = 2$, Assumption 1.2-a is satisfied. $H_0(\mu)$ and $F(\mu)$ are symmetric around $\mu_0 = 0.5$ and satisfy Assumption 1.3. $\forall$ partition of $\mathbb{R}^+ : \{0, c_1, \cdots, c_k, \infty\}$, there exists uncertainty measure $H(\mu)$ satisfying Assumption 1.3 such that:*

1. *When $c \in \{c_k\}$, $\mu^* = \mu_0$, and the optimal strategy at μ_0 is the sames as that with $H_0(\mu)$.*
2. *When c increases on $\mathbb{R}^+$, the sign of $\mu^* - \mu_0$ alternates on each partition.*

Proposition 1.1 states that the flexible learning model can fit an arbitrary number of crossovers of the response time-accuracy relationship at given difficulty levels. The standard DDM predicts identical decision time distribution for the correct responses and the errors (Ratcliff (1981)). To accommodate a non-trivial speed-accuracy trade-off/complementarity, DDM with varying boundary (Cisek, Puskas, and El-Murr (2009)) or DDM with random starting point and drift (Ratcliff and Rouder (1998)) are proposed, and there are a lot of debate about which variation works better. Fudenberg, Strack, and Strzalecki (2018) shows that the collapsing (expanding) boundary maps exactly to the complementarity (trade-off), and in an uncertain-difference DDM with endogenous stopping, decision boundary collapses to zero asymptotically and accuracy declines over time. These analyses suggest that DDM is able to fit the crossover, however at the cost of adding trial dependent parameters. Meanwhile, it remains to be disentangled which set of parameters in DDM are task specific and which set are subject specific. On the contrary, the flexible learning model predicts the crossovers clearly with varying only a task difficulty parameter, while keeping the task payoffs and the learning technology constant across trials.

1.8.2 Radical innovation

An important question in the study of innovation is to understand what characteristics of a firm foster innovation. The second application relates the radicality of firm's R&D and innovation to its safe option. I consider two firms: an *incumbent* (I) and an *entrant* (E). They face the identical set of risky new products. The only difference between the two firms is that the incumbent has a better existing safe product. I am interested in which firm innovates more radically in the R&D process. Intuitively, there are two competing incentives:

1. *Impatience effect*: The incumbent has an overall higher continuation value than the entrant. Therefore, by the value-precision monotonicity, the more impatient incum-

bent should prefer the frequency of signal to the precision of signal. So the impatience effect suggests that the entrant innovates more radically.

2. *Threshold effect*: The incumbent has a better outside option. Therefore, it has a higher threshold of belief for accepting a risky option. The relative value of a precise signal to an imprecise signal is higher for the incumbent. Therefore, the threshold effect suggests that the incumbent innovates more radically.

I model the problem using the following setup. There is one safe product P_s and K risky products $\{P_1, \cdots, P_K\}$. The state is $x \in \{G, B\}$. $x = G$ means the new technology is good, and the new products are better than the safe product: $\forall i, k\ u_i(P_k, G) > u_i(P_s, G)$. When $x = B$, the new technology fails, and $\forall i, k\ u_i(P_k, B) < u_i(P_s, B)$. $\forall x, k$, $u_I(P_k, x) = u_E(P_k, x)$ and $u_I(P_s) > u_E(P_s)$. The two firms share the same $H(\mu)$ function and capacity constraint c.[27] Let $\nu_i(\mu)$ be the two firms' optimal strategies. I define that a firm is looking for *more radical innovation* given belief μ *iff* $|\nu_i(\mu) - \mu| > |\nu_{-i}(\mu) - \mu|$, namely firm i is searching for a more precise Poisson signal.

Example 1.6. I calculate a simple example. There is only one risky product and $K = 1$. The incumbent's safe option pays $u_I(P_s, x) = 0.3$ and the entrant's safe option pays $u_E(P_s, x) = 0.15$. The risky option pays 1 when $x = G$ and -1 when $x = B$. H is the standard entropy function, $\rho = 1, c = 0.3$.

Figure 1.15 depicts the value functions (red curve: incumbent; blue curve: entrant). The two dashed lines are the payoffs of the corresponding safe options. Figure 1.16 depicts the policy functions (red curve: incumbent; blue curve: entrant). There is clearly a crossover of the policy functions. In the union of the two firm's experimentation regions,

[27] It is straightforward that if the cost of R&D is flexible, the incumbent invests (strictly) more as a direct implication of the *value-intensity monotonicity*. So I fix the capacity and focus on the choice of signal precision. It is not hard to extend the results to the flexible cost case.

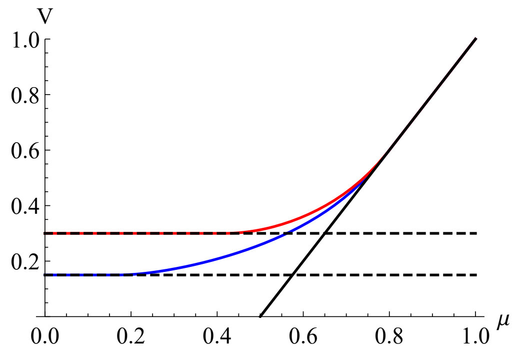


Figure 1.15: Value function

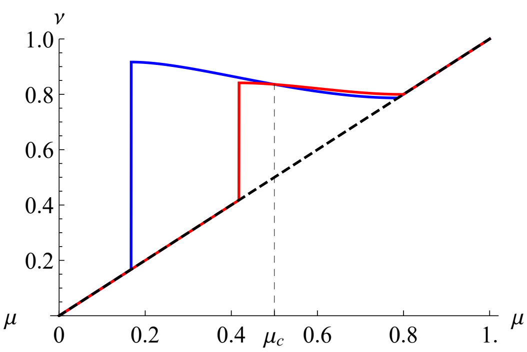


Figure 1.16: Policy function

when $\mu < \mu_c$ the entrant seeks more radical innovation, when $\mu > \mu_c$ the incumbent seeks more radical innovation.

The result of Example 1.6 can be summarized by the following proposition. Suppose $K = 1$, let E_0 be the union of the two firms' experimentation regions.

Proposition 1.2. *There exists μ_c s.t. $\forall \mu \in E_0$, $\mu > \mu_c \implies |\nu_I(\mu) - \mu| > |\nu_E(\mu) - \mu|$ and $\mu < \mu_c \implies |\nu_I(\mu) - \mu| < |\nu_E(\mu) - \mu|$. Moreover, $E_0 \bigcap (0, \mu_c) \neq \varnothing$ and $E_0 \bigcap (\mu_c, 1) \neq \varnothing$.*

Proposition 1.2 first states that there exist a threshold belief that the incumbent looks for more radical innovation if (and only if) the belief is higher than the threshold. Moreover, there exist none degenerate regions that either firm is innovating more radically than the other. Therefore, the order of radically of the two firms' innovations switches exactly once when the belief changes. Here is the intuition for the crossover. The entrant's value function is always steeper than the incumbent's, hence, the difference in the continuation value is decreasing in the belief. As a result, the impatience effect is diminishing when μ increases. On the other hand, when μ is higher, it is ex ante more likely that the risky arm will be chosen. As a result, the threshold effect outweighs the impatience effect when μ increases. Therefore, when μ increases, the incumbent is increasingly favoring a more precise signal, comparing to the entrant. Thus, there is a crossover.

Proposition 1.2 extends to multiple risky products as well. When $K > 1$, the experimentation regions are no longer simple intervals. Instead, they are unions of open intervals. In any experimentation interval where V never touches F_s, the two firms use the identical strategy (since the outside option is never triggered). So we only consider the leftmost interval in each firm's experimentation region. Let E_0 be the union of the two firms' leftmost intervals of the experimentation region.

Proposition 1.3. *There exists μ_c s.t. $\forall \mu \in E_0$, $\mu > \mu_c \implies |\nu_I(\mu) - \mu| > |\nu_E(\mu) - \mu|$ and $\mu < \mu_c \implies |\nu_I(\mu) - \mu| < |\nu_E(\mu) - \mu|$. Moreover, $E_0 \bigcap (0, \mu_c) \neq \varnothing$ and $E_0 \bigcap (\mu_c, 1) \neq \varnothing$.*

1.9 Conclusion

This chapter provides a dynamic information acquisition framework which allows fully general design of signal processes, and characterizes the optimal information acquisition strategy. My first contribution is an optimization foundation for a family of simple information generating processes: for an information acquisition problem with flexible design of information, the optimal information structure causes beliefs to follow a jump-diffusion process. Second, I characterize the optimal policy: seeking a Poisson signal whose arrival confirms the prior belief is optimal. The arrival of the signal leads to an immediate action. The absence of the signal is followed by continued learning with increasing precision and decreasing frequency.

Chapter 2

Time preference and information acquisition

2.1 Introduction

Consider a decision maker (DM) who is making a one-shot choice of action. The payoff of each action depends on an unknown state of the world. The DM can design a sequence of signal structures as her information source subject to a flow informativeness constraint. The informativeness of a signal structure is measured by a *posterior separable* measure. The DM is impatient and discounts future payoffs. Here I want to study the following question: fix a target information structure, what is the optimal learning dynamics that implements this target information structure?

In Example 2.1, I analyze this problem in a very simple toy model. In the example, I consider three simple dynamic signal structures: (i) pure accumulation of information before decision making, (ii) learning from a decisive signal arriving at a Poisson rate and (iii) learning from observing a Gaussian signal. This example suggests that different dynamic signal structures mainly differ in the induced decision time distribution. Since the form of discounting function prescribes the risk attitude on the time dimension, the discounting function (or time preference) is a key factor determining the optimal dynamic signal structure.

Example 2.1. The unknown state of the world can take two possible values $x = \{0, 1\}$. Prior belief is $\mu = 0.5$ (the probabiltiy that $x = 1$). Suppose that the target information structure is full revelation (induced posterior belief is either 0 or 1). I consider a model in continuous time. The flow information measure of belief process μ_t is assumed to be $E[-\frac{\mathrm{d}}{\mathrm{d}t}H(\mu_t)|\mathcal{F}_t]$ (the *uncertainty reduction speed*, introduced in Chapter 1), where $H(\mu) = 1 - 4(\mu - 0.5)^2$. Assume that the flow cost constraint is $c \leqslant 1$. The DM has exponential discount function e^{-t}. I assume the utility from the optimal actions associated with each state to be 1. In this example, I compare three different learning strategies:

1. *Pure accumulation:* the DM uses up all resources pushing her posterior beliefs towards

the boundary. This strategy is a continuous time extension of the suspense maximizing strategy introduced in Ely, Frankel, and Kamenica (2015). At each prior μ, the strategy is to seek a signal that induces posterior belief $\nu = 1 - \mu$ with arrival probability $p = \frac{1}{4(1-2\mu)^2}$[1]. The DM makes decision once her posterior arrives at 0 or 1. The posterior belief will either drift along one of two deterministic iso-time curve or jump between the two curves at the Poisson rate.

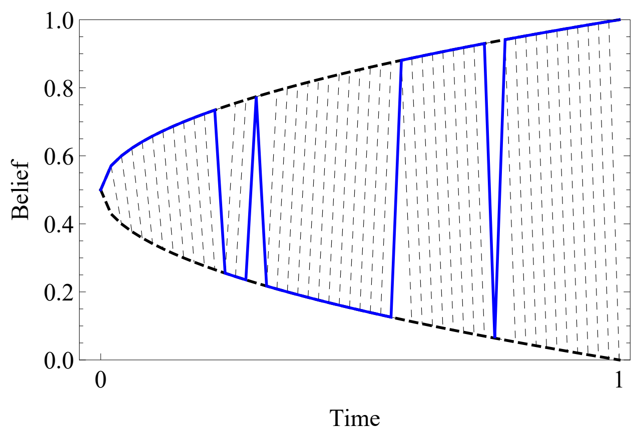


Figure 2.1: Belief trajectory

Figure 2.1 illustrates the two iso-time curves (thick dashed curves) and a possible path of belief (blue curve). By the standard property of compensated Poisson process, the DM's posterior belief drifts towards the boundary with speed $\frac{1}{4(2\mu-1)}$. Therefore, one of the belief trajectory follows the following ODE:

$$\begin{cases} \dot{\mu} = \frac{1}{4(2\mu-1)} \\ \mu(0) = 0.5 \end{cases}$$

It is easy to solve that $\mu(t) = \frac{1+\sqrt{t}}{2}$. As a result, the DM's decision time i.e. the time that belief process hits 1 is deterministic at $t = \mu^{-1}(1) = 1$. Then the expected utility from

[1]This can be calculated using the cost of Poisson signals $E[-\frac{\mathrm{d}H(\mu)}{\mathrm{d}t}] = (H(\mu) - H(\nu)p + H'(\mu)(\nu - \mu))p \leqslant c$

the pure accumulation strategy is discounted by one unit of time: $V_A = e^{-1} \approx 0.368$.

2. *Gaussian learning:* the DM observes a Gaussian signal, whose drift is the true state and variance is a control variable. By the standard property of Gaussian learning, the DM's posterior belief process follows a martingale Brownian motion. The flow variance of the posterior belief process satisfies the information cost constraint $E[-\frac{\mathrm{d}}{\mathrm{d}t}H(\mu_t)\big|\mathcal{F}_t] = -\frac{1}{2}\sigma^2 H''(\mu) \leqslant c$. Therefore, we can solve for $\sigma^2 = \frac{1}{4}$ when the constraint is binding. It is obvious that it is optimal to have the constraint binding. The value function is characterized by the following HJB:

$$V(\mu) = \frac{1}{2}\sigma^2 V''(\mu) = \frac{1}{8}V''(\mu)$$

with boundary condition $V(0) = V(1) = 1$. There is an analytical solution to the ODE:

$$V(\mu) = \frac{e^{2\sqrt{2}} + e^{4\sqrt{2}x}}{1 + e^{2\sqrt{2}}} e^{-2\sqrt{2}x}$$
$$\Longrightarrow V_G = V(0.5) \approx 0.459$$

3. *Poisson learning:* the DM learns the state perfectly at Poisson rate λ. If no information arrives, her belief stays at the prior. By the flow informativeness constraint $E[-\frac{\mathrm{d}}{\mathrm{d}t}H(\mu_t)\big|\mathcal{F}_t] = \lambda(H(\mu_t) - \frac{1}{2}H(1) - \frac{1}{2}H(0)) \leqslant c \implies \lambda = 1$. The value function is characterized by the HJB:

$$\rho V_P = \lambda(1 - V_P)$$
$$\Longrightarrow V_P = 0.5$$

Clearly:

$$V_P > V_G > V_A$$

Now we introduce the intuition why the values are ordered in this way. First, all of the three strategies induce the same expected decision time 1. This is due to the linearity of posterior separable information measure in compound experiments. The measure of a signal structure that fully reveals the state at prior 0.5 is exactly 1, and it must equal the expected sum of the total learning costs. Since in each continuing unit of time flow cost 1 is spent, expected learning time must be exactly 1. Therefore, what determines the expected decision utility is the dispersion of decision time distributions. Since exponential discount function e^{-t} is a strictly convex function, a learning strategy that creates the most dispersed decision time attains the highest expected utility. Now let us study the decision time distribution induced by the three strategies:

1. Pure accumulation: $t = 1$ with probability 1. The decision time is deterministic.

2. Gaussian learning: The decision time is the first passage time of a standard Brownian motion at either of the two absorbing barriers:

   $$\tau = \min\left\{ t \middle| \frac{1}{2} + \frac{1}{2} B_t = 0 \ or \ 1 \right\}$$

   The distribution of τ is characterized by a heat equation with two-sided boundary conditions at $x = 0, 1$. This equation has no analytical solution (solution can be characterized by series). Here I numerically simulate this process:

   Figure 2.2 depicts the evolution of the distribution of posterior beliefs over time. We can see that at any time, the distribution over posteriors is a Normal distribution cen-

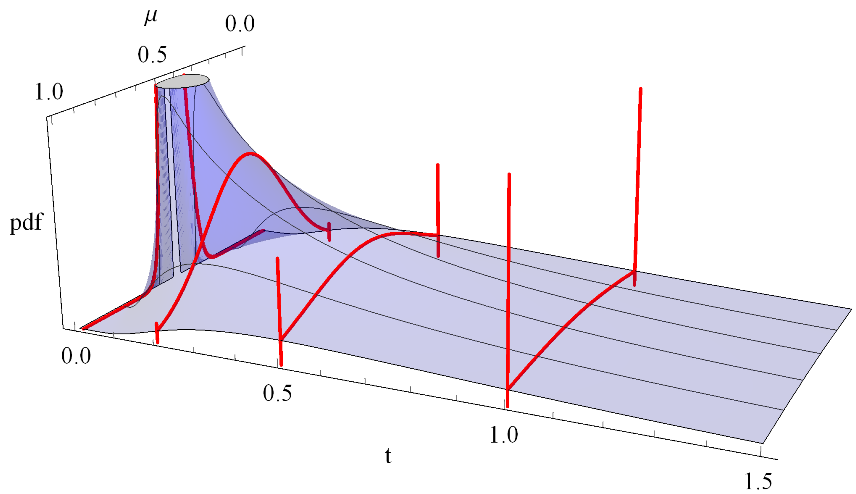


Figure 2.2: Belief distribution of Gaussian learning

sored at the two absorbing barriers [2]. The normal part is becoming flatter over time because learning leads to mean preserving spread of posterior beliefs.

3. Poisson learning: As is calculated, the Poisson signal arrives at a fixed arrival frequency $\lambda = 1$. The stopping time distribution can be calculated easily:

$$F(t) = 1 - e^{-t}$$

Evolution of posterior beliefs is shown in Figure 2.3: Figure 2.3 depicts the evolution of the distribution of posteriors over time. At any time, distribution over posteriors has three mass points at the prior and the two target posteriors. The mass on prior is decreasing over time (following an exponential distribution) and the mass on posteriors is increasing over time.

Obviously, pure accumulation is always the worst in this example since it induces deterministic decision time. By comparing Figures 2.2 and 2.3, one can easily see the dif-

[2]The distribution has point mass at $0, 1$, represented by the straight lines in Figure 2.2. The relative height represents the size of the probability mass. But the point mass part and Normal part does not share the same scale.

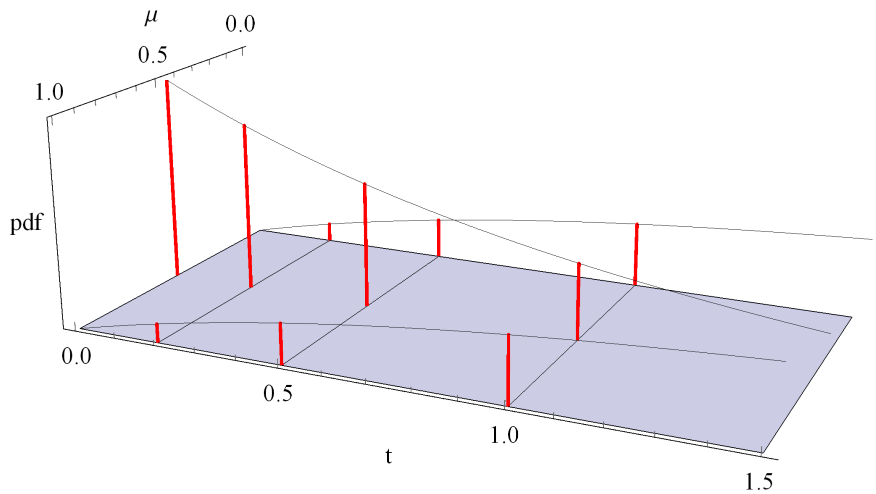


Figure 2.3: Belief distribution of Poisson learning

ference between Gaussian learning and Poisson learning: Gaussian learning accumulates some information that induces intermediate beliefs over time, while Poisson learning uses up all resources to draw conclusive signals. It seems that Poisson learning induces higher decision probability in the beginning while Gaussian learning induces higher decision probability later on (when prior becomes more dispersed). Therefore, Poisson learning has more dispersed decision time. We can verify this conjecture by plotting the PDFs and the integral of CDFs:

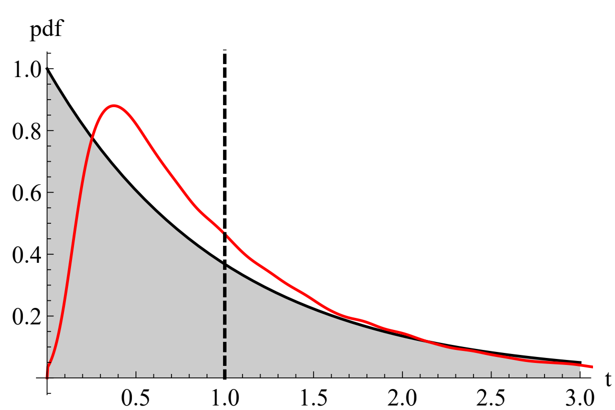


Figure 2.4: PDFs

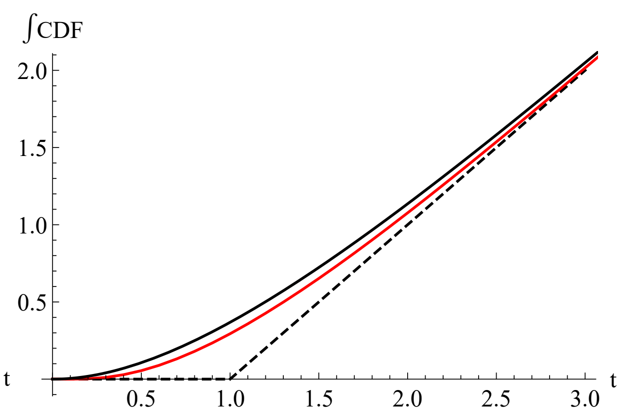


Figure 2.5: Integral of CDFs

In figure Figures 2.4 and 2.5, the black curves represent Poisson learning, the red curves represent Gaussian learning and the dashed lines represent pure accumulation. It is not hard to see from Figure 2.5 that the decision time of Poisson learning is in fact a mean-preserving spread of that of Gaussian learning. So Poisson learning dominates Gaussian learning for not only exponential discounting, but also any other convex discounting function.

In Example 2.1, I compare three kinds of dynamic learning strategies. These three strategies are chosen to be representative. First, the three strategies are simple heuristics that are very tractable. Second, these three strategies are also representative for three kinds of learning frameworks widely used in the literature:

- Pure accumulation has a flavor of the static rational inattention models. Like in Matějka and McKay (2014), decision is made once and there is no dynamics. Even in dynamic rational inattention model like Steiner, Stewart, and Matějka (2017), information is acquired in one period, and there is no smooth of information. In this example, the belief processes induced by learning has neither time dispersion nor cross-sectional dispersion when using the pure accumulation strategy.

- Gaussian learning itself is well studied in the literature, for example by Moscarini and Smith (2001), Hébert and Woodford (2016). On the other hand, Gaussian learning is one kind of symmetric drift-diffusion model (Ratcliff and McKoon (2008)). Gaussian learning captures the idea of gradual learning both over time and over beliefs.

- Poisson learning has been studied in Che and Mierendorff (2016). Poisson bandit is also used as a building block for strategic experimentation models (see a survey by Hörner and Skrzypacz (2016)). My example considers a simplest stationary Poisson stopping strategy that stochastically reveals the true state. Poisson learning is only gradual over time, but is lump sum in the belief space.

Example 2.1 suggests a key trade-off to be studied: gradual accumulation of information v.s. seeking decisive evidence. I want to study how the choice between gradual accumulation and decisive evidence seeking determines the decision time distribution. In Section 2.2, I develop an information acquisition problem that imposes no restriction on the specific form of information a decision maker can acquire. The DM can choose an arbitrary random process as signals, and she observes the signal realizations as her information. There are two constraints on the signal process. First, flow informativeness of the process is bounded. Second, the signal distribution conditional on stopping is fixed. If the DM chooses to learn gradually, then she is able to accumulate sufficient information before making any decision. After accumulating information, she can run the target experiment successfully with very high probability and achieves close to riskless decision time. On the contrary, if the DM chooses to only seek decisive signals, then the signals arrive only with low probabilities. So the corresponding decision time is riskier.

The main finding of this chapter is that among all decision time distributions induced by feasible and exhaustive[3] learning strategies, the most dispersed decision time distribution is induced by decisive Poisson learning–only decisive signals arrive as Poisson process. Meanwhile, the least dispersed time distribution is induced by pure accumulation, as I have already shown in Example 2.1.

This chapter is structured as follows. Section 2.2 setups a general discrete time information acquisition framework. Section 2.3 proves the main theorem. Section 2.4 extends the result to a continuous time model. Section 2.6 concludes.

[3] A feasible strategy is exhaustive if it is not leaving any capacity unused or acquiring unrelated information.

2.2 Setup of model

The model is in discrete time. Consider a decision maker who has a discount function ρ_t decreasing and convex (both weakly) in time t and $\lim_{t\to\infty}\sum_{s=t}^{\infty}\rho_s = 0$. There is a finite state space X and action space A. The prior belief of the unknown payoff-relevant state is $\mu \in \Delta(X)$. The DM's goal is to implement a signal structure that induces distribution $\pi \in \Delta^2(X)$ over posterior beliefs[4]. By implementing a target signal structure, I mean conditional on stopping, the signal structure in the current period must be a sufficient statistics for the target information structure. The informativeness of signal structure is measured by a posterior separable function $I(p_i, \nu_i|\mu) = \sum p_i(H(\mu) - H(\nu_i))$. In each period, the DM can acquire information for no more than c unit, i.e. $E\left[I(p_i^t, \nu_i^t|\mu^t)\right] \leqslant c$. The optimization problem is:

$$\sup_{\mathcal{S}_t,\mathcal{T}} E[\rho_{\mathcal{T}} u(\mathcal{A},\mathcal{X})] \tag{2.1}$$

$$\text{s.t.} \begin{cases} I(\mathcal{S}_t;\mathcal{X}|\mathcal{S}_{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\geqslant t}) \leqslant c \\ \mathcal{X} \to \mathcal{S}_t \to \mathcal{A} \text{ conditional on } \mathcal{T} = t \\ \mathcal{X} \to \mathcal{S}_t \to \mathbf{1}_{\mathcal{T}\geqslant t} \end{cases}$$

where $\mathcal{T} \in \Delta\mathbb{N}$ is a random stopping time. $\mathcal{S}_{\mathbf{t-1}}$ is defined as summary of past information $(\mathcal{S}_1,\ldots,\mathcal{S}_{t-1})$. $\mathcal{S}_0 \equiv c_0$ is assumed to be degenerate. The objective function in Equation (2.1) is the expected discounted utility from taking the action. The first constraint is the flow information cost constraint, it states that conditional on any history, the information cost incurred in a period is less than the constraint c. The second constraint is the target information structure constraint. It states that at each period, conditional on stopping the acquired information structure is statistically sufficient for the target ac-

[4]State and signal realization can be equivalently represented as a pair of random variables $(\mathcal{X},\mathcal{A})$.

tion profile. The remaining constraint is a natural information process constraint (or the standard measurability constraint for stopping time).

Remark 2.1. This model is restrictive in the design of information in the following sense: At any instant in time, conditional on stopping, the information acquired must be statistically sufficient for a time invariant random variable $\mathcal{A}$. Other than this restriction, the DM can freely choose her learning dynamics. The interpretation of this constraint is not easy, as it is an aggregate constraint for each period, unconditional on the history of past signals. This model does not necessarily cover Gaussian learning in general, but it does in a symmetric cases (i.e. target posterior distribution and H are symmetric around prior μ, like in Example 2.1).

The main reason for imposing this constraint is for tractability. I restrict learning dynamics in this way to abstract away from the fact that the optimal target information structure itself is changing over time, which creates time varying incentive for search direction, search precision and search intensity (highlighted in Chapter 1). In the current chapter, I want to focus on the trade-off between gradual information accumulation and decision evidence seeking.

I assume that the DM follows the suggestion of signal structure $\mathcal{A}$ in choosing the action. This is WLOG since given any signal structure, the induced optimal action itself forms a Blackwell less informative signal structure. Therefore, the original learning strategy is still statistically sufficient for the direct signal structure. So if we take the optimization of $\mathcal{A}$ also into account, then it is WLOG to assume that $\mathcal{A}$ is a direct signal. Then the optimal implementation of $\mathcal{A}$ still follows a solution to Equation (2.1). The optimization of $\mathcal{A}$ is studied in Section 2.5.1.

2.3 Solution

2.3.1 An auxiliary problem

Let $\bar{I} = I(\mathcal{A}; \mathcal{X})$ and $V^* = E[u(\mathcal{A}, \mathcal{X})]$. Consider a relaxed problem which only tracks the *average* accumulated information measure I at every time t, rather than the entire signal process conditional on all histories:

$$\sup_{p_t} \sum_{t=1}^{\infty} \rho_t (1 - P_{t-1}) p_t V^* \tag{2.2}$$
$$\text{s.t.} \begin{cases} (\bar{I} - I_t) p_t + (I_{t+1} - I_t)(1 - p_t) \leqslant c \\ P_t = P_{t-1} + (1 - P_{t-1}) p_t \\ P_0 = 0, I_1 = 0 \end{cases}$$

where $p_t \in [0,1]$ and $I_t \geqslant 0$. $1 - P_{t-1}$ is the surviving probability at period t, p_t is the conditional stopping probability. I_t is the total expected information measure of the entire path of non-stopping signals up to period t.

The constraints in the relaxed problem Equation (2.2) capture a key feature of posterior separable information measure: I_t is accumulated linearly over time and the information measure required to implement $\mathcal{S}$ is exactly the remaining information measure $\bar{I} - I_t$. It is more relaxed than Equation (2.1) in the following sense: in Equation (2.1), the flow informativeness constraint is imposed on all histories of $\mathcal{S}_{t-1}$ and $\mathbf{1}_{\mathcal{T} \leqslant t}$. However, in Equation (2.2), the first constraint is imposed only on each period unconditional on the history. In other words, the first constraint in Equation (2.2) is an average version of the flow informativeness constraint in Equation (2.1). p_t can be interpreted as the expected stopping probability and I_t's as the expected accumulated informativeness.

Lemma 2.1. *Equation* (2.1) $\leqslant$ *Equation* (2.2)

Lemma 2.1 verifies the previous intuition, that Equation (2.2) is a relaxation of Equation (2.1). Now I first solve Equation (2.2). Then we can use the auxiliary problem to provide some clue for solving the original problem Equation (2.1).

Theorem 2.1. *$p_t \equiv \frac{c}{I}$ solves Equation (2.2).*

Theorem 2.1 states that the relaxed problem Equation (2.2) has a simple solution: no information should ever be accumulated. It directly implies that $I_t \equiv 0$ and the optimal value equals $\sum_{t=0}^{\infty} \rho_t \left(1 - \frac{c}{I}\right)^{t-1} \frac{c}{I} V^*$. I prove Theorem 2.1 by approximating the convex discount function ρ_t with a finite summation of linear functions. Then for each linear discount function, I prove by backward induction that choosing $I_t \equiv 0$ is optimal.

2.3.2 Optimal learning dynamics

By Lemma 2.1 and Theorem 2.1, to solve Equation (2.1), it is sufficient to show that

$$\sum_{t=1}^{\infty} \rho_t \left(1 - \frac{c}{\overline{I}}\right)^{t-1} \frac{c}{\overline{I}} V^* \tag{2.3}$$

is attainable by a feasible strategy in Equation (2.1). Consider the following experimentation strategy: $\mathcal{A}$ is observed with probability $\frac{c}{\overline{I}}$ in each period. If $\mathcal{A}$ is successfully observed, the corresponding action is taken. If not, go to the next period and follow the same strategy. Formally, $\mathcal{S}_t$ and $\mathcal{T}$ are defined as follows. Let $s_0, c_0 \notin A$ be two distinct degenerate signals.

$$\mathcal{S}_t = \begin{cases} s_0 & \text{with probability 1 if } \mathcal{S}_{t-1} \in A \bigcup \{s_0\} \\ \mathcal{A} & \text{with probability } \frac{c}{\overline{I}} \text{ if } \mathcal{S}_{t-1} = c_0 \\ c_0 & \text{with probability } 1 - \frac{c}{\overline{I}} \text{ if } \mathcal{S}_{t-1} = c_0 \end{cases} \tag{2.4}$$

$$\mathcal{T} = t \ \text{ if } \mathcal{S}_t \in A$$

Here signal s_0 means stopping. Signal c_0 means continuation. Any signal in A indicates the action to take. Then it is not hard to verify that:

- *Objective function*:

$$
\begin{aligned}
&E[\rho_{\mathcal{T}} u(\mathcal{A}, X)] \\
=&\sum_{t=1}^{\infty} \rho_t \mathrm{P}(\mathcal{S}_t \in A) E[u(\mathcal{A}, \mathcal{X})] \\
=&\sum_{t=1}^{\infty} \rho_t \prod_{\tau=0}^{t-1} \mathrm{P}(\mathcal{S}_\tau = c_0 | \mathcal{S}_{\tau-1} = c_0) \mathrm{P}(\mathcal{S}_t \in A | \mathcal{S}_{t-1} = c_0) E[u(\mathcal{A}, \mathcal{X})] \\
=&\sum_{t=1}^{\infty} \rho_t \left(1 - \frac{c}{\bar{I}}\right)^t \frac{c}{\bar{I}} V^*
\end{aligned}
$$

- *Capacity constraint*:

$$
\begin{aligned}
&I(\mathcal{S}_t; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) \\
=&\mathbf{1}_{\mathcal{S}_{t-1}=c_0} I(\mathcal{S}_t; \mathcal{X} | \{c_0\}, 1) + \mathbf{1}_{\mathcal{S}_{t-1} \neq c_0} I(s_0; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, 0) \\
=&\mathbf{1}_{\mathcal{S}_{t-1}=c_0} (\mathrm{P}(\mathcal{S}_t \in A) I(\mathcal{A}; \mathcal{X}) + \mathrm{P}(\mathcal{S}_t = c_0) I(c_0; \mathcal{X})) + \mathbf{1}_{\mathcal{S}_{t-1} \neq c_0} \cdot 0 \\
=&\mathbf{1}_{\mathcal{S}_{t-1}=c_0} \cdot \frac{c}{\bar{I}} \cdot \bar{I} \leqslant c
\end{aligned}
$$

- *Decision time distribution*:

$$
P_t = \mathrm{P}(\mathcal{T} \leqslant t) = 1 - \left(1 - \frac{c}{\bar{I}}\right)^t \tag{2.5}
$$

I show that Equation (2.4) implements the expected utility level Equation (2.3), hence solves Equation (2.1). It is easy to see that Equation (2.4) induces expected decision time $\frac{\bar{I}}{c}$. By Lemma 2.2, which is stated below, $\frac{\bar{I}}{c}$ is the lower bound of expected decision time for all feasible strategies. In fact, the proof of Lemma 2.2 suggests that $E[\mathcal{T}] > \frac{\bar{I}}{c}$ only when there is some waste of information: either capacity constraint c is not fully used, or $\mathcal{S}_t$ contains strictly more information than $\mathcal{A}$ conditional on taking action.

Lemma 2.2. *Let $(\mathcal{S}_t, \mathcal{T})$ be a strategy that satisfies the constraints in Equation (2.1), then $E[\mathcal{T}] \geqslant \frac{\bar{I}}{c}$.*

I call an information acquisition strategy *exhaustive* if the corresponding $E[\mathcal{T}] = \frac{\bar{I}}{c}$. The decision time distribution P_t induced by strategy Equation (2.4) is a exponential distribution with parameter $\frac{c}{\bar{I}}$. Equation (2.4) being the optimal strategy, independent of choice of ρ_t implies that $\forall \rho_t$, $\forall$ information acquisition strategy $(\widetilde{\mathcal{S}}_t, \widetilde{\mathcal{T}})$:

$$
\begin{aligned}
& E[\rho_{\widetilde{\mathcal{T}}} u(\mathcal{A}; \mathcal{X})] \leqslant E[\rho_{\mathcal{T}} u(\mathcal{A}; \mathcal{X})] \\
\Longrightarrow\ & E[\rho_{\widetilde{\mathcal{T}}}] \leqslant E[\rho_{\mathcal{T}}]
\end{aligned}
$$

Since ρ_t ranges over all positive decreasing convex functions, P_t as distribution over time is second order stochastically dominated. Summarizing the analysis above, I get Theorem 2.2.

Theorem 2.2. *Equation (2.4) solves Equation (2.1). The decision time distribution of any feasible and exhaustive information acquisition strategy is a* ***meas preserve contraction*** *of P_t.*

2.3.3 Gradual learning v.s. decisive evidence

My analysis illustrates the gradual learning v.s. decisive evidence trade-off in the flexible learning environment. The trade-off is: the speed of future learning depends on how much information the DM has already possessed. Accumulating more information today speeds up future learning. So the DM is choosing between naively learning just for today or learning for the future. If all resources are invested in seeking decisive evidence, then signal arrives at a constant low probability, and the decision time distribution is dispersed. If some resources are invested in information accumulation, then learning will accelerate, at a cost of lower (or even zero) arrival rate of decisive signals in the early stage. As a result the decision time is less dispersed.

When the decision maker has convex discounting function, decisive evidence seeking is optimal. The intuition behind this result is natural. Convex discounting function means that the decision maker is risk loving towards decision time. Seeking decisive evidence is the riskiest learning strategy one can take: it payoffs quickly with high probability, but if it fails, learning is very slow in future. In practice, evidence seeking is a very natural learning strategies. A researcher tends to form a hypothesis, then seeks evidence that either confirms or contradicts the hypothesis. Usually there is a clear target of what to prove (the hypothesis), and what kind of signals (data from experiments) proves/contradicts the hypothesis. Running the research protocol itself is usually more mechanical than the designing stage. What is common in natural science is that the principal investigator(PI) designs the whole research plan. Then all experiments, data collections and computations are run by doctoral students. The PI usually has a permanent position and there is no deadline, so he can enjoy the expected payoff from this risky project design.

Two elements in my framework are crucial to this result. The first is the flexibility in the design of signal process. In contrast to my framework, if one considers a dynamic information acquisition problem with highly parametrized information process, then other kind of trade-offs tied to the parametrization constraints might have first order effects. For example, if one only allows Poisson learning or Gaussian learning, then the trade-off of gradual learning and decisive evidence is directly assumed away. As a result the choice among signal types (Che and Mierendorff (2016),Liang, Mu, and Syrgkanis (2017)) or the trade-off between intensity and information cost (Moscarini and Smith (2001)) becomes first order important. If one only allows DM to choose between to learn or not to learn in each period, then the trade-off between exploration and exploitation becomes first order. Meanwhile, in my framework, the DM can freely design the optimal signal type, and hence the corresponding decision time[5]. So the aforementioned trade-offs actually do not exist, the trade-off between gradual learning and decisive evidence becomes central to the analysis.

The second is the posterior separability assumption on information measure. Posterior separability is equivalent to the linear additivity of compound signal structures (see the discussion in Section 1.7.1). This assumption restricts the relative price between gradual learning and evidence seeking. Any amount of informativeness invested today to accumulate information transfers one-to-one to the amount of reduction of information cost tomorrow. Lemma 2.2 shows that the expected decision time is identical for all feasible and exhaustive learning strategy. As a result the trade-off between gradual learning and decisive learning translates to choice of dispersion of decision time distribution. If one assumes either sub-additivity or super-additivity in informativeness measure, then choosing different learning strategies might also change the expected decision time, which makes my key trade-off entangled with other effects.

2.4 Continuous time model

In this section, I study a continuous time version of Equation (2.1). Let $\rho_t : \mathbb{R}^* \to \mathbb{R}^*$ be a decreasing and convex discounting function. Let $F(\mu) = \sup_a E_\mu[u(a,x)]$, the expected utility from choosing the optimal action given a belief. Consider the following stochastic control problem:

$$\sup_{\langle\mu_t\rangle\in\mathbb{M},\tau} \; E[\rho_\tau E_\pi[F(\mu)]] \tag{2.6}$$
$$\text{s.t.} \begin{cases} -E\left[\frac{\mathrm{d}}{\mathrm{d}t}H(\mu_t)\middle|\mathcal{F}_t\right] \leqslant c \\ \mu_0 = \mu,\ \mu_t\big|_{\tau=t} \sim \pi \end{cases}$$

[5]The DM can affect the decision time distribution by choosing the information acquisition strategy. However, not all decision time distributions are implementable.

where τ is a stopping time measurable to the natural filtration of μ_t. The objective function of Equation (2.6) is the same as that of Equation (2.1). In the stochastic control problem, the decision maker chooses the optimal posterior belief process $\langle\mu_t\rangle$ and stopping time τ, subject to the 1) stopping time is measurable to belief process. 2) belief process is a martingale. 3) flow increase in informativeness measure is bounded by c. 4) conditional on stopping time, μ_t has distribution π.

It is not hard to see that Equation (2.6) is a continuous time extension of Equation (2.1). I take a belief based approach when formulating Equation (2.6). However, I did not formally proof how a stochastic process of posterior beliefs can be induced by a stochastic information acquisition strategy. Equation (2.6) is constructed by taking analog of Equation (2.1). Let $V^* = E_\pi[F(\mu)]$. Then

Lemma 2.3. *Equation* (2.6) $\leqslant$ *Equation* (2.7).

$$V = \sup_{p_t} \int_0^\infty \rho_t (1-P_t) p_t V^* \mathrm{d}t \tag{2.7}$$

$$\text{s.t.} \begin{cases} I_0 = 0,\ I_t \geqslant 0,\ \dot{I}_t \leqslant c - p_t(\bar{I} - I_t) \\ P_0 = 0,\ \dot{P}_t = (1-P_t)p_t \end{cases}$$

where p_t is a positive integrable function.

Theorem 2.3. *$p_t \equiv \frac{c}{\bar{I}}$ solves Equation* (2.7).

Lemma 2.3 and Theorem 2.3 are exactly the continuous time analogs of Lemma 2.1 and Theorem 2.1. Lemma 2.3 states that Equation (2.7) is a relaxed problem of Equation (2.2). Theorem 2.3 characterizes the solution of Equation (2.7): no information should ever be accumulated. $I_t \equiv 0$ and the optimal value equals $\int_0^\infty \rho_t e^{-\frac{c}{\bar{I}}t} \frac{c}{\bar{I}} V^* \mathrm{d}t$. Theorem 2.3 is proved by discretizing the continuous time problem and invoking the result of Theorem 2.1.

2.4.1 Implementation

By Lemma 2.3 and Theorem 2.3, to solve Equation (2.6), it is sufficient to show that:

$$\int_0^\infty \rho_t e^{-\frac{c}{I}t}\mathrm{d}t\frac{c}{\overline{I}}V^*$$

can be attained in Equation (2.6). Consider the following information acquisition strategy. Let ν be a random variable with distribution π and define:

$$\begin{cases} \mathrm{d}\mu_t = (\nu - \mu_t)\cdot \mathrm{d}N_t \\ \tau = t \text{ if } \mathrm{d}N_t\text{=}1 \end{cases} \tag{2.8}$$

where N_t a standard Poisson counting processes with parameter $\frac{c}{I}$ and independent to ν. $\langle\mu_t\rangle$ is by definition a stationary compound Poisson process. The jump happens when the Poisson signal arrives and belief jumps to posteriors according to distribution π. Once the jump occurs, decision is made immediately. It is easy to verify:

- *Martingale property*: We know that each compensated Poisson process $\mathrm{d}N_t - \frac{c}{I}\mathrm{d}t$ is martingale, therefore:

$$\begin{aligned} E[\mathrm{d}\mu_t|\mu_t] =& E[(\nu-\mu_t)\cdot \mathrm{d}N_t] \\ =& E_\pi[E[(\nu-\mu)\cdot \mathrm{d}N_t|\nu]] \\ =& E_\pi\Big[(\nu-\mu)\cdot E\Big[\mathrm{d}N_t - \frac{c}{\overline{I}}\mathrm{d}t\Big]\Big] + E_\pi\Big[(\nu-\mu)\cdot\frac{c}{\overline{I}}\mathrm{d}t\Big] \\ =& 0 \end{aligned}$$

  therefore, μ_t is a martingale. The second equality is the law of iterated expectation. Third equality is by $E[\nu]=\mu$ and $\mathrm{d}N_t - \frac{c}{I}\mathrm{d}t$ being martingale.

- *Capacity constraint*: If $N_t \geqslant 1$, then $E[-\frac{\mathrm{d}H(\mu_t)}{\mathrm{d}t}|\mu_t] = 0 \leqslant c$. If $N_t < 1$, then by the Ito formula for jump process:

$$\begin{aligned}
\mathrm{d}H(\mu_t) =& (H(\nu) - H(\mu)) \cdot \mathrm{d}N_t \\
\implies \quad E[-\frac{\mathrm{d}H(\mu_t)}{\mathrm{d}t}|\mu_t] =& E_\pi\left[E\left[(H(\mu_t) - H(\nu)) \cdot \frac{\mathrm{d}N_t}{\mathrm{d}t}|\nu\right]\right] \\
=& E_\pi\left[\frac{c}{\bar{I}}(H(\mu_t) - H(\nu))\right] \\
=& c
\end{aligned}$$

The second equality is the law of iterated expectation. The third equality is the martingale property of $\mathrm{d}N_t - \frac{c}{\bar{I}}\mathrm{d}t$.

- *Decision time distribution*:

$$P_t = 1 - e^{-\frac{c}{\bar{I}}t}$$

Therefore, Equation (2.8) implements utility level $\int_0^\infty \rho_t \frac{c}{\bar{I}} e^{-\frac{c}{\bar{I}}t} V^* \mathrm{d}t$.

Lemma 2.4. *Let (μ_t, τ) be a strategy that satisfies the constraints in Equation (2.6), then $E[\tau] \geqslant \frac{\bar{I}}{c}$.*

As in the discrete time case, I call an information acquisition strategy *exhaustive* if the corresponding $E[\tau] = \frac{\bar{I}}{c}$. Since Equation (2.8) is optimal independent of the choice of convex ρ_t, previous analysis implies Theorem 2.4.

Theorem 2.4. *Equation (2.8) solves Equation (2.6). The decision time distribution of any feasible and exhaustive information acquisition strategy is a mean preserving contraction of P_t.*

2.5 Discussion

2.5.1 Optimal target signal structure

In this section, I solve for the optimal target signal structure in decision problem Equation (2.6). Assume that ρ_t is differentiable. By Theorem 2.4, the optimization problem can be written as:

$$\sup_{\pi\in\Delta^2(X)} \int_0^\infty \rho_t e^{-\frac{c}{H(\mu)-E_\pi[H(\nu)]}t}\mathrm{d}t \cdot \frac{c\cdot E_\pi[F(\nu)]}{H(\mu)-E_\pi[H(\nu)]} \tag{2.9}$$
$$\text{s.t. } E_\pi[\nu]=\mu$$

Define $f(V^1,V^2)=\int_0^\infty \rho_t e^{-\frac{c}{H(\mu)-V^1}t}\mathrm{d}t\cdot\frac{c\cdot V^2}{H(\mu)-V^1}$. Then it is not hard to verify that $f(V^1,V^2)$ is differentiable[6] in V^1,V^2. Optimization problem Equation (2.9) fits in Theorem 4.2 from Chapter 4. Applying the theorem gives a necessary condition for π^* solving Equation (2.9):

$$\pi^* \in \arg\max_{\substack{\pi\in\Delta^2(X)\\ E_\pi[\nu]=\mu}} E_\pi\left[F(\nu)+\frac{\int_0^\infty(-\dot{\rho}_t)e^{-\frac{c}{H(\mu)-E_{\pi^*}[H(\nu)]}t}\frac{E_{\pi^*}[F(\nu)]}{H(\mu)-E_{\pi^*}[H(\nu)]}t\mathrm{d}t}{\int_0^\infty \rho_t e^{-\frac{c}{H(\mu)-E_{\pi^*}[H(\nu)]}t}\mathrm{d}t}\cdot H(\nu)\right]$$

Notice that the objective function is the expectation of the linear combination of two belief dependent functions. If we define:

$$g(x)=\frac{\int_0^\infty(-\dot{\rho}_t)e^{-\frac{c}{H(\mu)-x}t}\frac{t}{H(\mu)-x}\mathrm{d}t}{\int_0^\infty \rho_t e^{-\frac{c}{H(\mu)-x}t}\mathrm{d}t}$$

Then by the standard argument in Bayesian persuasion, π^* can by characterized by concavifying the *gross value function* $F+(g(E_{\pi^*}[H(\nu)])\cdot E_{\pi^*}[F(\nu)])H$. Moreover, by rem B.1, there exists π^* with support size $2|X|$ solving Equation (2.9). So I get the following characterization:

Proposition 2.1. *There exists π^* solving Equation (2.9) and $|\text{supp}(\pi^*)| \leqslant 2|X|$. Let $\lambda = g(E_{\pi^*}[H(\nu)]) \cdot E_{\pi^*}[F(\nu)]$, any maximizer π^* satisfies:*

$$\pi^* \in \arg \max_{\substack{\pi \in \Delta^2(X) \\ E_\pi[\nu]=\mu}} E_\pi[F(\nu) + \lambda \cdot H(\nu)]$$

Suppose the discounting function is a standard exponential function: $\rho_t = e^{-\rho t}$, then $g(x) = \frac{\rho}{c+\rho(H(\mu)-x)}$. Notice that the objective function:

$$V(\mu) = \int_0^\infty e^{-\left(\rho + \frac{c}{H(\mu)-E_{\pi^*}[H(\nu)]}\right)t} \frac{c \cdot E_{\pi^*}[F(\nu)]}{H(\mu) - E_{\pi^*}[H(\nu)]} \mathrm{d}t = \frac{c \cdot E_{\pi^*}[F(\nu)]}{c + \rho(H(\mu) - E_{\pi^*}[H(\nu)])}$$

Therefore, the optimality condition becomes:

$$\pi^* \in \arg \max_{\substack{\pi \in \Delta^2(X) \\ E_\pi[\nu]=\mu}} E_\pi\Big[F(\nu) + \frac{\rho}{c} V(\mu) H(\nu)\Big] \tag{2.10}$$

Equation (2.10) is very similar to the optimality condition I derived in Chapter 1, where the optimal posterior is solved from concavifying $V(\cdot) + \frac{\rho}{c}V(\mu)H(\cdot)$. The problem solved in Chapter 1 is the continuous time limit of Equation (2.1) without the restriction on constant target signal structure and with exponential discounting. In both problems, $\frac{\rho}{c}V(\mu)$ is adjusting the concavity of the gross value function. Therefore, higher continuation value corresponds to more concave gross value function and less informative signal structure. This suggests that the monotonicity in precision-frequency trade-off is extended to our model as well. In Chapter 1, the trade-off is illustrated as decrease in precision of target information structure at each decision time. In the current chapter, target information structure is forced to be constant over time. However, if I endogenize the target information structure, then at more extreme prior beliefs associated with higher decision value, less informative target information structure is optimal (and corresponding expected waiting time is shorter).

2.6 Conclusion

In this chapter, I characterize the decision time distributions that can be induced by a dynamic information acquisition strategy, and study how time preference determines the optimal form of learning dynamics. No restriction is placed on the form of information acquisition strategy, except for a time invariant target signal structure and a flow informativeness constraint. I find that all decision time distributions have the same expectation, and the maximal and minimal elements by mean-preserving spread order are exponential distribution and deterministic distribution. The result implies that when time preference is risk loving (e.g. standard or hyperbolic discounting), Poisson signal is optimal since it induces the riskiest exponential decision time distribution. When time preference is risk neutral (e.g. constant delay cost), all signal processes are equally optimal.

Chapter 3

Indirect information measure and dynamic learning

3.1 Introduction

Information plays a central role in economic activities. It affects both strategic interaction in games and single agent decision making under uncertainty. Information is often endogenously acquired by a decision maker, as opposed to being exogenously endowed. Therefore, it is important to understand how information is acquired. This boils down to a simple trade-off: the value of information and the cost of acquiring information. The value of information is often unambiguous in a single agent decision problem with uncertainty. It is measured by the increased expected utility from choosing optimal actions measurable to informative signal realizations(see Blackwell et al. (1951)). However, there has been less consensus on the proper form of information acquisition cost. One (probably most) popular measure of informativeness being used in many information acquisition models is the Entropy based mutual information and its generalizations. This approached was initiated by Sims (1998, 2003), and is applied to a wide range of problems (Matějka and McKay (2014), Steiner, Stewart, and Matějka (2017), Yang (2015a), Gentzkow and Kamenica (2014), etc.). Despite its great theoretical tractability, Entropy based models suffer from criticism on its unrealistic implications, including prior dependence, invariant likelihood ratio of action, etc.

Two approaches can be taken to build a solid foundation for studying information acquisition. One approach is to fully characterize the behavior implications associated with mutual information and its generalizations. Then we will be able to empirically test the behavior validity of these models. Caplin, Dean, and Leahy (2017) takes this approach and proposes testable axioms for the Shannon model of rational inattention and its generalizations. The other approach is to impose only minimal assumptions on the cost of information and study the robust predictions in an information acquisition problem. In this chapter I take the second robust approach and focus on a dynamic information acquisition problem: a decision maker acquires information about a payoff relevant state before choosing an action. She can choose an arbitrary random process as observed information, subject to cost on information and cost on waiting.

I accomplish two main goals. First, I characterize the “minimal assumptions” on a (static) information measure if a decision maker can choose from not only all information structures but also all sequential combinations of them to minimize expected information measure. I show that an indirect information measure is supported by expected learning cost minimization— given any general measure of information, and for any information structure (Blackwell experiment), the DM minimizes the expected total measure of a compound experiment which replicates the original information structure— if and only it satisfies three simple conditions. 1) Monotonicity: Blackwell more informative experiment has higher measure. 2) Sub-additivity: the expected total measure of a replicating compound experiment is weakly higher than the measure of the original experiment. 3) C-linearity: mixing uninformative experiment with a proportion of informative experiment has measure proportional to the mass on the informative part.

Second, I solve a dynamic information acquisition problem with those assumptions imposed on the flow information measure. I prove that solving the dynamic problem can be divided into two steps. The first step is to solve a static rational inattention problem for an optimal static information structure. The second step is to solve for the optimal dynamic implementation of the solution from the first step. The optimal information process involves direct Poisson signals: signal arrives according to a Poisson counting processes and the arrival of signal suggests the optimal action directly. When no signal arrives, posterior belief process stays at prior.

Related Literature

This chapter is closed related to two sets of works that aim at understanding the measure of information. The first tries to characterize implications (testable or non-testable) of commonly used information measures. Basic mathematical implications and characterizations for Entropy and Entropy based mutual information was provided in standard information theory text books like Cover and Thomas (2012). Matějka and McKay (2014) and Caplin and Dean (2013) study the behavior implications of rational inattention model based on Mutual information and posterior separable information measure respectively. Caplin and Dean (2015) studies the implications of rational inattention model based on general information measure. A set of full behavior characterizations for mutual information, posterior separable information cost and their generalizations are provided in Caplin, Dean, and Leahy (2017), Denti (2018), and Frankel and Kamenica (2018). Meanwhile, the second set of works seeks to build a dynamic foundation for common information measures. Morris and Strack (2017) shows that the posterior separable function can be represented as the induced cost from random sampling. Hébert and Woodford (2016) justifies a class of information cost function (including mutual information) based on a continuous-time sequential information acquisition problem. This chapter contributes to this literature by providing a new optimization foundation for posterior separability. Posterior separability is actually equivalent to additivity — both sub-additivity and sup-additivity — in the expected measure of compound experiments. I show that sub-additivity is justified by expected information cost minimization.

This chapter is also closely related to the dynamic information acquisition literature, in which the main goal is to characterize the learning dynamics. A common approach in this literature is to model information flow as a simple family of random process. The decision maker can control parameters which represents aspects of interest. Wald

(1947) first studies stopping problem with exogenous information process. Moscarini and Smith (2001) and Che and Mierendorff (2016) go further by edogenizing information process into optimization problem in Brownian motion framework and Poisson bandits framework to study dynamics of learning intensity and direction respectively. Some recent papers edogenize the random process family as well and give decision maker full flexibility in designing information. Chapter 1 studies flexible dynamic information acquisition with a posterior separable information measure and shows that confirmatory Poisson signal is optimal. Steiner, Stewart, and Matějka (2017) studies a repeated rational inattention problem with mutual information as cost. This chapter contributes by relaxing the restriction on information cost to only minimal assumptions. I show that when impatience is measure by fixed delay cost, the dynamic problem is closely related to the static rational inattention problem, and Poisson learning is robustly optimal.

The rest of this chapter is structured as follows. Section 3.2 introduces the characterization of indirect information measure based on expected information measure minimization. Section 3.3 setups a dynamic information acquisition problem and characterizes the solution.

3.2 Indirect information measure

3.2.1 *Information structure and the measure of informativeness*

In this subsection, I formally define "information" and a "measure of informativeness" in decision making problems. I extract key factors in any abstract "information" that matters in a decision making problem and characterize a well defined equivalence class that characterizes all information structures. Then, I use an "indirect information measure" characterization to derive the minimal assumptions that we should impose on an information measure.

Definition 3.1.

1. Bayesian plausible posteriors*: Let $\Delta X \subset \mathbb{R}^{|X|}$ be the belief space over X. Let $\Delta^2 X$ be the space of probability measures over ΔX. $\Pi(\mu) = \left\{\pi \in \Delta^2 X \middle| \int \nu \mathrm{d}\pi(\nu) = \mu\right\}$ is the set of Bayesian plausible posterior distributions. Let $\Gamma = \left\{(\pi, \mu) \in \Delta^2 X \times \Delta X \middle| \pi \in \Pi(\mu)\right\}$*

2. Information structure*: Let S be an arbitrary set (set of signals). Let $p \in \Delta S \times X$ be a conditional distribution over S on $x \in X$. (S, p) is an information structure. (S, p) can be equivalently represented as $\mathcal{S}$, a random variable whose realization is determined by p.*

I would like to study the "set" of all information structures as a choice set for decision maker. However, since S is an arbitrary set, the "set" of all possible S is not even a well-defined object from the perspective of set theory. Instead, I use $\Pi(\mu)$ to equivalently characterize the "set" of all information structures. $\forall$ (S, p), $\forall s \in S$, the posterior belief from observing s can be calculated according to Bayes rule. The distribution of all such posteriors forms a Bayesian plausible distribution as defined in Definition 3.1. Since different signals inducing the same posterior belief affect neither the choice of action nor the expected utility, I claim that $\Pi(\mu)$ already summarizes all possible information structures (up to the equivalence of posterior beliefs). Γ is defined as the set of all pairs (π, μ) where π represents an information structure given prior belief μ.

Definition 3.2. *An **information measure** is a mapping $I : \Gamma \to \overline{\mathbb{R}}^{+}$. I will represent $I(\pi, \mu)$ using $I(\mathcal{S}; \mathcal{X}|\mu)$ in an interchangeable way, where μ is the distribution of $\mathcal{X}$ and $\mathcal{S}$ induces belief distribution π.*

Information measure I is defined as a mapping from prior-information structure pairs in Γ to extended non-negative real numbers. The only (implicit) restriction I put on I is that different information structures that induce the same distribution of posterior π at μ have the same measure. This restriction is actually without loss of generality because the induced distribution of posterior of an information structure is always a sufficient statistics for any feasible decision rule. Suppose different information structures have different measure, then the DM is always able to choose an appropriate information structure with the lowest information measure.[1] Definition 3.2 is the same as *information cost function* defined in Caplin and Dean (2015). The only difference is that I explicitly modeled prior dependence of I: μ is an argument in I. In Caplin and Dean (2015) prior is chosen and fixed in the beginning so there is no need to explicitly specify information cost function for different priors.

From this point on, for simplicity I represent the choice set of DM with information structures $\mathcal{S}$. However, I don't differentiate two information structures that induces same distribution of posterior beliefs. By using notation $\cdot|\mathcal{S}$, I mean conditional on the posterior beliefs induced by realization of $\mathcal{S}$. The next step is to impose some restrictive assumptions on I. The restrictions I impose is about comparing measure of information structure when they satisfies some information order. So first let's formally define the information order.

Definition 3.3 (Information processing constraint)**.** *Given random variables $\mathcal{X},\mathcal{S},\mathcal{T}$ and their joint distribution $p(x,s,t)$. Let $p(t|s)$, $p(t|s,x)$ be the conditional distribution defined by Bayes rule: $p(t|s)=\frac{\int p(t,s,x)\mathrm{d}x}{\int p(t,s,x)\mathrm{d}x\mathrm{d}t}$ and $p(t|s,x)=\frac{p(t,s,x)}{\int p(t,s,x)\mathrm{d}s}$ and:*

$$p(t|s,x)=p(t|s)$$

for s,x with positive probability, then the triple $\mathcal{X},\mathcal{S},\mathcal{T}$ is defined as a Markov chain:

$$\mathcal{X}\rightarrow\mathcal{S}\rightarrow\mathcal{T}$$

[1] Discussing this issue formally leads to the problem of choosing inf from all possible $\mathcal{S}$, which is not a well defined set. I avoid dealing with this problem by making this restriction explicitly.

The information processing constraint in Definition 3.3 defines a most natural constraint in the acquisition of information: When decision time $\mathcal{T}$ is chosen based on information $\mathcal{S}$, the choice should be purely a result of information. Therefore, conditional on knowing the information, choice should not be dependent to the underlying state any more. This is the key constraint I'm going to impose in Section 3.3. The information processing constraint has several equivalent characterizations:

Proposition 3.1. *The following statements are equivalent:*

1. $\mathcal{X} \to \mathcal{S} \to \mathcal{T}$.
2. $\mathcal{X}$ *and* $\mathcal{T}$ *are independent conditional on* $\mathcal{S}$.
3. $\mathcal{S}$ *is a sufficient statistics for* $\mathcal{T}$ *w.r.t.* $\mathcal{X}$.
4. $\mathcal{S}$ *is Blackwell more informative than* $\mathcal{T}$ *about* $\mathcal{X}$.

Proposition 3.1 comes mostly from Blackwell et al. (1951) and links the information processing constraint to other well-known notions in probability theory and information theory. It is intuitive that these notions are equivalent. They essentially all characterize the fact that $\mathcal{S}$ carries more information about $\mathcal{X}$ than $\mathcal{T}$. From this point on, I use the four equivalent notions in an inter-changeable way.

Now I can define what I refer to as the minimal assumptions on the measure of information.

Assumption 3.1. $I(\mathcal{S};\mathcal{X}|\mu)$ *satisfies the following axioms:*

1. *(Monotonicity)* $\forall \mu$, *if* $\mathcal{X} \to \mathcal{S} \to \mathcal{T}$, *then:*

$$I(\mathcal{T};\mathcal{X}|\mu) \leqslant I(\mathcal{S};\mathcal{X}|\mu)$$

2. *(Sub-additivity)* $\forall \mu$, $\forall$ *information structure* $\mathcal{S}_1$ *and information structure* $\mathcal{S}_2|_{S_1}$ *whose*

distribution depends on the realization of $\mathcal{S}_1$:

$$I((\mathcal{S}_1, \mathcal{S}_2); \mathcal{X}|\mu) \leqslant I(\mathcal{S}_1; \mathcal{X}|\mu) + E[I(\mathcal{S}_2; \mathcal{X}|\mathcal{S}_1, \mu)]$$

3. *(C-linearity)* $\forall \mu$, $\forall$ *information structure* $\mathcal{S} \sim (\mu_i, p_i)$. $\forall \lambda \in [0, 1]$, *consider* $\mathcal{S}_\lambda \sim (\mu_i, \mu, \lambda p_i, 1 - \lambda)$ [2], *then:*

$$I(\mathcal{S}_\lambda; \mathcal{X}|\mu) = \lambda I(\mathcal{S}; \mathcal{S}|\mu)$$

Assumption 3.1 imposes three restrictions on the information measure I. Monotonicity states that if an information structure $\mathcal{S}$ is Blackwell more informative than (statistically sufficient for) information structure $\mathcal{T}$, then the information measure of $\mathcal{S}$ is no lower than that of $\mathcal{T}$. Sub-additivity states that if one breaks a combined information structure into the two components sequentially, then the information measure of the combined information structure is no higher than the expected total measure of the two components. C-linearity is a strengthen of sub-additivity in a special case: if a combined information structure can be decomposed into pure randomness and an informative information structure, then its information measure is exactly the expected total measure of these components.

With Assumption 3.1, my model nests some standard measures of information. Monotonicity directly states that my information measure is consistent with the Blackwell partial order of information (Blackwell et al. (1951)). My model includes the mutual information measure used in rational inattention models (Sims (2003), Matějka and McKay (2014) etc.) as a special case. Mutual information is a case where my sub-additivity

[2] $\mathcal{S}_\lambda$ is defined that with $1 - \lambda$ probability, the posterior is identical to the prior. With the remaining λ probability, the distribution of posteriors is identical to that of $\mathcal{S}$. That is to say, $\mathcal{S}_\lambda$ is obtained by mixing $\mathcal{S}$ with a constant signal by weight $(\lambda, 1 - \lambda)$.

assumption is replaced by additivity and an extra logarithm structure is imposed on the information measure. In Gentzkow and Kamenica (2014) and Chapter 1, a *posterior separable* information measure, which is more general than mutual information is used to model the cost of information. Posterior separability is equivalent to additivity (see the discussion in Chapter 1), thus a special case of sub-additivity. Generally speaking, Assumption 3.1 nests most information measures used in the recent "information design" literature, where information is modeled in a non-parametric way. However, it still excludes many interesting settings. For example, it's hard to verify whether Assumption 3.1 is satisfied in a parametric model. It also fails the prior independence, which is a very natural assumption when we think of information as objective experimentations.

3.2.2 *Information cost minimization*

Imagine that a decision maker is allowed to flexibly choose any information structure to learn. The cost of information is captured by a general measure of information as defined in Definition 3.2. Consider the information measure as the cost paid by the DM. Then if the decision maker is further allowed to choose any (sequential) combinations of a set of information structures, then she might be able to replicate a single information structure using a combination of information structures with paying a lower cost on expectation. For each single information structure, I call the minimal expected sum of information measure of any sequential replication the *Indirect information measure*. In fact, if we consider the indirect information measure as the effective measure of informativeness of information structures, then Assumption 3.1 is without loss of any generality:

Proposition 3.2. *Information measure $I^*(\mathcal{S};\mathcal{X}|\mu)$ satisfies Assumption 3.1* iff *there exists an information measure $I(\mathcal{S};\mathcal{X}|\mu)$ s.t. $\forall \mu, \mathcal{S}$:*

$$I^*(\mathcal{S};\mathcal{X}|\mu) = \inf_{(\mathcal{S}^i,N)} E\left[\sum_{i=1}^{N} I\left(\mathcal{S}^i;\mathcal{X}|\mathcal{S}^1,\ldots,\mathcal{S}^{i-1}\right)\right]$$

$$\text{s.t. } \mathcal{X} \to \left(\mathcal{S}^1, \ldots, \mathcal{S}^N\right) \to \mathcal{S}$$

Proposition 3.2 states that when a DM can choose from all sequential combinations of information structures that replicate a given information structure to minimize the expected total measure, then the effective measure for a piece of information satisfies Assumption 3.1. The intuition for Proposition 3.2 is quite simple. Consider the expected information measure as a cost of information. If a Blackwell less informative information structure has a higher measure, then it is never chosen because by choosing the more informative structure, a DM can still accomplish any decision rule feasible with the less informative structure and pays a lower cost. This implies both monotonicity and sub-additivity. C-linearity is in fact an implication of sub-additivity when adding irrelevant noise to information. On the one hand, combining noise with an information structure $\mathcal{S}$, one can create $\mathcal{S}_\lambda$, implying inequality from one direction. On the other hand, by repeatedly acquiring $\mathcal{S}_\lambda$ conditional on observing only noise, one can replicate $\mathcal{S}$. Therefore, additivity from both direction implies C-linearity.

In practice, there are many scenarios in which such minimization of expected information measure is present. If we consider information as a product provided in a competitive market, then the minimization problem in Proposition 3.2 is very natural. The price of information is the marginal cost of information. And cost minimization on the sellers' side implies that the price of information satisfies Assumption 3.1. (In a monopolistic market there might be positive markups and varying information rents so pricing might be very different, as is discussed in Zhong (2018).) Another example is information processing of a computer. Modern computer programs are designed to balance work loads from independent processes onto nodes/threads. As a result what matters is the average informational bandwidth, (as opposed to the peak bandwidth or other measures). If we consider information as data processed by a computer, then in each CPU tick time, an optimally designed algorithm will minimize expected bandwidth required to process information.

3.3 Dynamic decision problem

In this section, I study the implication of the indirect information measure in a dynamic information acquisition problem. I consider a decision maker (DM) acquiring information about the payoffs of different alternatives before making a choice. She can choose the information structure flexibly within each period, contingent on the history of signals. The cost of information acquired within a period depends on an indirect information measure, and the DM pays a constant cost of delay per period. The major finding is that this model justifies learning by acquiring Poisson type signals.

3.3.1 Model

Assume that the DM faces the following dynamic information acquisition problem:

- *Decision problem*: The time horizon $t = 0, 1, \dots, \infty$ is discrete. Length of each time interval is dt. The utility associated with action-state pair (a, x) is $u(a, x)$. The DM pays a constant cost m for delaying on period. If the DM takes action $a \in A$ at time t conditional on state being $x \in X$, then her utility gain is $u(a, x) - mt$. I assume that the utility gains from actions are bounded: $\sup_{a,x} u(a, x) < \infty$.

- *Uncertainty*: Not knowing the true state, the DM forms a prior belief $\mu \in \Delta X$ about the state. Her preference under uncertainty is expressed as von Neumann-Morgenstern expected utility. I am going to use two essentially equivalent formulations to express expected utility. 1) Given belief μ, the expected utility associated with each action $a \in A$ is $E_\mu[u(a, x)]$. 2) State and action are represented by random variables $\mathcal{X}, \mathcal{A}$. Expected utility is denoted by $E[u(\mathcal{A}, \mathcal{X})]$.

- *Information Cost*: I use the information measure I defined as in Definition 3.2 as a flow measure of information within a period, and define a time separable information cost structure. In each period, with prior belief μ, the DM pays information cost $f(I(\mathcal{S},\mathcal{X}|\mu))$ which transforms the measure of information acquired in the period into utility loss. $f:\mathbb{R}^+\to\overline{\mathbb{R}}^+$ is a non-decreasing convex function which maps to extended real values.

- *Dynamic Optimization*: The dynamic optimization problem of the DM is:

$$V(\mu)=\sup_{\mathcal{S}^t,\mathcal{A}^t,\mathcal{T}} E\left[u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right)-m\mathcal{T}-\sum_{t=0}^{\infty}f\left(I\left(\mathcal{S}^t;\mathcal{X}\middle|\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right)\right] \tag{P}$$

$$\text{s.t.}\begin{cases}\mathcal{X}\to\mathcal{S}^{t-1}\to\mathbf{1}_{\mathcal{T}\leqslant t}\\ \mathcal{X}\to\mathcal{S}^{t-1}\to\mathcal{A}^t \text{ conditional on } \mathcal{T}=t\end{cases}$$

  where $\mathcal{T}\in\Delta\mathbb{N}$, $t\in\mathbb{N}$. $\mathcal{S}^{-1}$ is defined as a degenerate random variable that induces belief same as prior belief μ of the DM (just for notational simplicity). $\mathcal{S}^{\mathbf{t-1}}$ is defined as the summary of all past information $\left(\mathcal{S}^1,\dots,\mathcal{S}^{t-1}\right)$. The DM chooses the decision time $\mathcal{T}$, the choice of action conditional on stopping $\mathcal{A}^t$ and the signal structure $\mathcal{S}^t$ subject to information cost, waiting cost and two natural constraints for information processing:

  1. The information received in last period is sufficient for stopping in current period.
  2. The information received in last period is sufficient for action in current period. [3]

In Equation (P), the DM is modeled as choosing the information process $\mathcal{S}^t$, decision time $\mathcal{T}$ and choice of action $\mathcal{A}^t$ jointly, to maximize utility gain from action profile net waiting cost and total information cost. Within each period, informativeness is measured by I and incurs cost $f(I)$. Across period, information costs are aggregated by the expected sum of f. Since the information measure is defined on information structure-prior pairs. It's

[3] Noticing that in every period, the information in current period has not been acquired yet. So decision can only be taken based on the information already acquired in the past. So the Markov chain property on information and action time/action will have information lagged by one period. This within-period timing can be defined in different ways and it doesn't affect the main results.

important to define clearly how prior is determined. In each period, information measure is evaluated conditional on the realization of past signals and choice of stopping. This is a natural setup since past information plus whether action is taken in current period is exactly what the DM "knows" in current period. Therefore this is the finest filter on which she evaluates information cost.

Let me illustrate the cost structures of dynamic information acquisition with a simple two period model: $t \in \{0,1\}$ and the DM has prior belief μ. The timing is as following: when $t = 0$, the DM first chooses whether to take an action and which action to take. Second she decides what information to acquire. When $t = 1$, DM takes action based on information acquired in period 0. First let's consider deterministic continuation decision. In period 0 no information has been acquired yet so if DM want to make a choice, her expected utility will be calculated with the prior μ: $E_\mu[u(a,\mathcal{X})]$ and there is no waiting or information cost. If DM wants to collect information before decision making, she can acquire information structure $\mathcal{S}$, now it's for sure $\mathcal{T} = 1$ and $\mathcal{X} \to \mathcal{S} \to \mathcal{A}$. Therefore she gets expected utility $E[u(\mathcal{A},\mathcal{X})]$, pays waiting cost m and information cost $f(I(\mathcal{S};\mathcal{X}|\mu))$.

The problem becomes less trivial when continuation is random: suppose DM chooses to continue with probability p (independent to states because she has no information yet about state). Only conditional on continuation, she acquires $\mathcal{S}$. Within my framework, total cost is $p \cdot f(I(\mathcal{S};\mathcal{X},\mu)) + (1-p)\cdot 0$ by calculating conditional cost on $\mathbf{1}_{\mathcal{T}\leqslant 0}$. One might think that just conditional on information but not continuation decision, the same information structure is essentially $\mathcal{S}_p$ and cost is $f\big(I\big(\mathcal{S}_p;\mathcal{X}|\mu\big)\big)$. However, this is saying that when DM is choosing information after decision making in period 0, she acquires a signal correlated to her previous choice of continuation. This piece of randomness (whether to continue) is already resolved. Since our DM can not revert time, this case is physically impossible. $f(I(\mathcal{S}_p;\mathcal{X}|\mu))$ will be the right cost if the decision of continuation is delayed to the next period.

3.3.2 Solution

In this section, I solve the dynamic information acquisition problem Equation (P) under Assumption 3.1 on the information measure. First, I characterize the optimal expected utility as a solution to a simple static information acquisition problem. Second, I provide a simple stationary strategy that implements the expected utility from choosing any information and action strategy in the equivalent static problem.

Theorem 3.1. *If I satisfies Assumption 3.1, $\forall \mu \in \Delta X$, suppose the expected utility level $V(\mu)$ solves Equation (P), then:*

$$V(\mu) = \max\left\{\sup_{a\in A} E[u(a,\mathcal{X})], \sup_{I(\mathcal{A};\mathcal{X}|\mu)\geqslant\lambda} E[u(\mathcal{A},\mathcal{X})] - \left(\frac{m}{\lambda} + \frac{f(\lambda)}{\lambda}\right) I(\mathcal{A};\mathcal{X}|\mu)\right\} \tag{3.1}$$

The first superemum is taken over a, the second superemum is take over both λ and $\mathcal{A}$.

Theorem 3.1 states that solving the optimal utility level in Equation (P) is equivalent to solving a static problem under Assumption 3.1. In the static problem, the DM pays a fixed marginal cost $\left(\frac{m}{\lambda} + \frac{f(\lambda)}{\lambda}\right)$ on each unit of information measure $I(\mathcal{A};\mathcal{X}|\mu)$. Notice that the optimal parameter λ depends on only m, f when the constraint $I(\mathcal{A};\mathcal{X}|\mu) \geqslant \lambda$ doesn't bind. There is an explicit algorithm to solve Equation (3.1):

Proposition 3.3. *If I satisfies Assumption 3.1, $V(\mu)$ solves Equation (P) if and only if it solves the following problem: Let $\lambda^* = \sup\{\lambda \in \mathbb{R}^+ | m + f(\lambda) > \lambda \cdot \partial f(\lambda)\}$ and solve for*

$$V^0(\mu) = \sup_{a\in A} E[u(a,\mathcal{X})]$$

$$V^1(\mu) = \sup_{\mathcal{A}} E[u(\mathcal{A},\mathcal{X})] - \left(\frac{m}{\lambda^*} + \frac{f(\lambda^*)}{\lambda^*}\right) I(\mathcal{A};\mathcal{X}|\mu) \tag{3.2}$$

$$V^2(\mu) = \sup_{\mathcal{A}} E[u(\mathcal{A},\mathcal{X})] - m - f(I(\mathcal{A};\mathcal{X}|\mu)) \tag{3.3}$$

Let $\mathbb{A}$ be the set of maximizers of Equation (3.2)[4]*, then*

$$V(\mu) = \begin{cases} \max\left\{V^0(\mu), V^1(\mu)\right\} & \textit{if } \sup\limits_{\mathcal{A}\in\mathbb{A}} I(\mathcal{A};\mathcal{X}|\mu) \geqslant \lambda^* \\ \max\left\{V^0(\mu), V^2(\mu)\right\} & \textit{otherwise} \end{cases}$$

Proposition 3.3 states that the value function in Equation (P) can be solved by solving three static problems. The first value V^0 is a no-information benchmark when value equals expected utility from choosing the optimal action according to only the prior. The second problem Equation (3.2) is a standard rational inattention problem with marginal cost $\frac{m}{\lambda^*} + \frac{f(\lambda^*)}{\lambda^*}$ on information measure I. The interpretation is that under Assumption 3.1, the dynamic information acquisition problem is separable in two parts. The first part is the dynamic allocation of information, keeping the aggregate information fixed. Marginal cost of increasing the aggregate information is reflected by $\frac{m}{\lambda^*} + \frac{f(\lambda^*)}{\lambda^*}$, which measures both the impatience and the smoothing incentive jointly. The second part is a static problem that optimizes the aggregate information. The third problem Equation (3.3) is a special case when there is under-smoothing. This happens only when waiting is so costly that it is optimal for decision maker to scale up information cost and wait for less than one period. Since fractional period length is not feasible, in this case decision maker solves a one-period problem.

Once the static problems Equations (3.2) and (3.3) are solved, let $\mathcal{A}$ be an optimal information structure of the static problem, then $\mathcal{A}$ can be modified to construct an optimal dynamic information structure in Equation (P).

Proposition 3.4. *If I satisfies Assumption 3.1, $\forall \mu \in \Delta X$, $\mathcal{A} \in \Delta A \times X$ and $\lambda^* < I(\mathcal{A};\mathcal{X}|\mu)$, let*

[4] If $\lambda^* = +\infty$, define $\mathbb{A} = \varnothing$. Here $\mathbb{A}$ includes both $\mathcal{A}$'s that exactly solve Equation (3.2) and sequences $\left\{\mathcal{A}^i\right\}$ that approach Equation (3.2). Given a sequence $\left\{\mathcal{A}^i\right\} \in \mathbb{A}$, $I(\mathcal{A};\mathcal{X}|\mu)$ is defined as $\limsup I(\mathcal{A}^i, \mathcal{X}|\mu)$

$(\mathcal{S}^t, \mathcal{A}^t, \mathcal{T})$ *be defined by*[5]*:*

1. $\mathcal{S}^{-1} = c_0$.

2. $\mathcal{S}^t = \begin{cases} s_0 & \text{if } \mathcal{S}^{t-1} \in A \bigcup \{s_0\} \\ \mathcal{A} \text{ with probability } \frac{\lambda^*}{I(\mathcal{A};\mathcal{X}|\mu)} & \text{if } \mathcal{S}^{t-1} = c_0 \\ c_0 \text{ with probability } 1 - \frac{\lambda^*}{I(\mathcal{A};\mathcal{X}|\mu)} & \text{if } \mathcal{S}^{t-1} = c_0 \end{cases}$

3. $\begin{cases} \mathcal{T} = t \\ \mathcal{A}^t = \mathcal{S}^{t-1} \end{cases} \quad \text{if } \mathcal{S}^{t-1} \in A.$

Then:

$$E[u(\mathcal{A},\mathcal{X})] - \left(\frac{m}{\lambda^*} + \frac{f(\lambda^*)}{\lambda^*}\right) I(\mathcal{A};\mathcal{X}|\mu) = E\left[u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right) - m\mathcal{T} - \sum_{t=0}^{\infty} f\left(I\left(\mathcal{S}^t;\mathcal{X}\middle|\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right)\right]$$

Proposition 3.4 complements Theorem 3.1 by showing that the optimal value from Equation (3.1) can be implemented using a simple stationary experimentation strategy that is feasible in Equation (P). The information structure $\mathcal{S}^t$ explicitly codes three kinds of signals: *Stop* s_0, *Wait* c_0 and *Action* in A. The first condition defines the initial information. The second condition defines the information structures in the following periods by induction: If $\mathcal{S}^{t-1} = s_0$ or $\mathcal{A}$ it means that action is already taken and information acquisition stops from now on so $\mathcal{S}^t = s_0$ and so on so forth. If $\mathcal{S}^{t-1} = c_0$ it means that do nothing and delay all decision to the current period. Conditional on continuation, $\mathcal{S}^t$ realizes as $\mathcal{A}$ with $\frac{\lambda^*}{I(\mathcal{A};\mathcal{X}|\mu)}$ probability. And in the next period the action is taken according to the realization of $\mathcal{S}^t$. With $1 - \frac{\lambda^*}{I(\mathcal{A};\mathcal{X}|\mu)}$ probability c_0 realizes and the decision is delayed to the next period. The Third condition explicitly defines $\mathcal{T}$: when action is taken in period t as indicated by $\mathcal{S}^{t-1}$, then $\mathcal{T} = t$. It's easy to verify the information processing constraints

[5] s_0 and c_0 are chosen to be distinguishable from any element in action set A.

in Equation (P) are satisfied. First, conditional on $\mathcal{S}^{t-1}$, the distribution of $\mathbf{1}_{\mathcal{T}\leqslant t}$ is degenerate. When $\mathcal{S}^{t-1}=c_0$ it's 0 and 1 otherwise. So $\mathcal{S}\to\mathcal{S}^{t-1}\to\mathbf{1}_{\mathcal{T}}\leqslant t$. Second, conditional on $\mathcal{S}^{t-1}$ and knowing $\mathcal{T}=t$, $\mathcal{A}^t$ is also degenerate. It is exactly the realization of $\mathcal{S}^{t-1}$. Therefore $\mathcal{X}\to\mathcal{S}^{t-1}\to\mathcal{A}^t$.

Sketched proof.

Here I provide a simplified proof which illustrates the main intuition for Theorem 3.1 and Proposition 3.4. Since there is no discounting on the utility gain from actions, given an action profile $\mathcal{A}^{\mathcal{T}}$, the expected utility is completely determined by 1) the aggregate distribution of actions $\mathcal{A}$. 2) the expected waiting time $E[\mathcal{T}]$. How actions are allocated over time doesn't affect the expected utility at all. Since actions are driven by information, this observation indicates that solving Equation (P) can be divide into three steps: Step 1 is to solve for the optimal distribution of information over time to minimize the information cost given any aggregate information structure and expected waiting time. Step 2 is to solve for the optimal waiting time given any fixed aggregate information structure. Step 3 is to solve for the optimal aggregate information structure and the associated action profile.

Step 1. Given any strategy $(\mathcal{S}^t,\mathcal{A}^t,\mathcal{T})$, the DM can implement the same action distribution $\mathcal{A}^{\mathcal{T}}$ and expected waiting time $E[\mathcal{T}]$ with a information process of lower cost. First, consider combining all information $\mathcal{S}=(\mathcal{S}^1,\dots,\mathcal{S}^t,\dots)$. By sub-additivity $I(\mathcal{S};\mathcal{X}|\mu)\leqslant\sum E\left[I\left(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}}\right)\right]$. Then consider averaging $I(\mathcal{S};\mathcal{X}|\mu)$ into $E[\mathcal{T}]$ periods:

$$\begin{aligned}
\frac{I(\mathcal{S};\mathcal{X}|\mu)}{E[\mathcal{T}]} &\leqslant \frac{\sum E\left[I\left(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}}\right)\right]}{E[\mathcal{T}]}\\
\Longrightarrow\ f\left(\frac{I(\mathcal{S};\mathcal{X}|\mu)}{E[\mathcal{T}]}\right) &\leqslant \frac{\sum E\left[f\left(I\left(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}}\right)\right)\right]}{E[\mathcal{T}]}\\
\Longrightarrow\ E[\mathcal{T}]f\left(\frac{I(\mathcal{A};\mathcal{X}|\mu)}{E[\mathcal{T}]}\right) &\leqslant \sum E\left[f\left(I\left(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}}\right)\right)\right]
\end{aligned}$$

The second inequality is first by monotonicity of f then by convexity of f. That is to say: there is incentive to combine small information (by sub-additivity of I) and smooth information over time (by convexity of f). The last inequality is from $I(\mathcal{S};\mathcal{X}|\mu) \leqslant I(\mathcal{A};\mathcal{X}|\mu)$ and the Blackwell monotonicity of I. Then an ideal strategy is to spend $f\left(\frac{I(\mathcal{A};\mathcal{X}|\mu)}{E[\mathcal{T}]}\right)$ on information acquisition every period.

Then I implement the aforementioned information cost using a strategy defined as in Proposition 3.4. By C-linearity, acquiring $\mathcal{A}$ with probability $\frac{1}{E[\mathcal{T}]}$ exactly has cost $f\left(\frac{I(\mathcal{A};\mathcal{X}|\mu)}{E[\mathcal{T}]}\right)$. On the other hand, taking action with probability $\frac{1}{E[\mathcal{T}]}$ in each period exactly the implements aggregate action distribution $\mathcal{A}$ and the expected waiting time $E[\mathcal{T}]$. Then it is WLOG to consider:

$$\sup_{\mathcal{A},T} E[u(\mathcal{A},\mathcal{X})] - mT - Tf\left(\frac{I(\mathcal{A};\mathcal{X}|\mu)}{T}\right)$$

where $E[\mathcal{T}]$ is replaced by T for notational simplicity.

Step 2. Maximizing over $E[\mathcal{T}]$ (or T in the simplified problem). This can be done easily by solving the first order condition w.r.t. T: $-m - f\left(\frac{I}{T}\right) + \frac{I}{T}f'\left(\frac{I}{T}\right) = 0$. Replace $\lambda = \frac{I}{T}$, we get the expression for λ: $m + f(\lambda) = \lambda f'(\lambda)$ and further simplified problem:

$$\sup_{\mathcal{A}} E[u(\mathcal{A},\mathcal{X})] - \left(\frac{m}{\lambda} + \frac{f(\lambda)}{\lambda}\right) I(\mathcal{A};\mathcal{X}|\mu)$$

The formal theorem covers general cases without smoothness assumption so f' is replaced with sub-differentials ∂f.

Step 3. I will refer to the Weierstrass theorem to show the existence of solution. See Proposition 3.5 for detailed discussion.

In the sketched proof I implicitly assumed f to be differentiable, first order condition has solution and optimal $T \geqslant 1$. The formal proof for more general cases is provided in

Appendix D.2.1

3.3.3 Existence and uniqueness

In this section, I first show a general existence result for the solution of Equations (3.2) and (3.3). Then I established its uniqueness in different dimensions. By toggling the inequalities defining the monotonicity of I, the concavity of f and the sub-additivity of I to strict inequalities, my model predicts unique belief profile, unique information cost allocation and unique strategy correspondingly.

Proposition 3.5. *If A, X are finite sets, I satisfies Assumption 3.1, then*

- *Existence: $\forall \varepsilon > 0$, let $\nabla_\varepsilon = \{\mathcal{A} | \mathrm{P}[a|x] \geqslant \varepsilon\}$, then there exists a non-empty, convex and compact set of solution $\mathbb{A}_\varepsilon$ to Equation (3.1) subject to $\mathcal{A} \in \nabla_\varepsilon$.*
  - *If $\exists \varepsilon > 0$, $\mathbb{A}_\varepsilon \bigcap \nabla^o_\varepsilon \neq \varnothing$, then $\bigcup_{\varepsilon' \leqslant \varepsilon} \mathbb{A}_{\varepsilon'}$ is the maximizer of Equation (3.1).*
  - *If $\forall \varepsilon > 0$, $\mathbb{A}_\varepsilon \bigcap \nabla^o_\varepsilon = \varnothing$, then any sequence in $\prod \mathbb{A}_\varepsilon$ approaches $V(\mu)$.*
- *Uniqueness:*
  - *If I satisfies strict-monotonicity, then posterior belief $\nu(a)$ associated with any action a is unique for all optimal $\mathcal{A}$.*
  - *If $f(\cdot)$ satisfies strict-convexity, then $\forall$ optimal strategy $\left(\mathcal{S}^t, \mathcal{A}^t, \mathcal{T}\right)$ to Equation (P), $I\left(\mathcal{S}^t; \mathcal{X} | \mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \leqslant t}\right)$ is the same.*
  - *If I satisfies strict-sub-additivity, then the solution $\left(\mathcal{S}^t, \mathcal{A}^t, \mathcal{T}\right)$ to Equation (P) is unique.*

Proposition 3.5 first states the existence of solution to Equation (3.1) and the uniqueness of different aspects of the solution. First, with Assumption 3.1, very mild extra assumptions (finite A and X) can guarantee the existence of solution to Equation (3.1) (and solution to Equation (P) as well). Second, when strictly more informative information structure has strictly larger information measure, the belief inducing each action can be uniquely pinned down by optimization. Third, when the information cost function f is strictly convex, then the optimal cost level incurred in each experimentation period is constant over time. Finally, if a combination of informative experiments has strictly larger measure than the expected summation of each component's measure, then whole dynamic strategy is uniquely pinned down.

The existence result is non-trivial in the sense that I don't impose any continuity assumption on I. However, I being an indirect information measure function actually guarantees it to be convex in an appropriate space. In Equation (3.1), the strategy space is all random variable $\mathcal{A}$. If we consider the space of all conditional distribution over A on X (Markovian transition matrices), then this is an Euclidean space and any indirect information measure I is a convex function on this space: if $\mathcal{S}$ is a linear combination of $\mathcal{S}_1$ and $\mathcal{S}_2$, then $\mathcal{S}$ can be implemented as randomly using $\mathcal{S}_1$ or $\mathcal{S}_2$ (and not knowing the choice of experiment). Therefore, monotonicity and sub-additivity guarantees $\mathcal{S}$ to have weakly lower measure than the linear combination of measures of $\mathcal{S}_1$, $\mathcal{S}_2$. Convexity of I implies both objective function to be continuous and choice set to be compact on any interior closed subset of the strategy space.

The incentive for inter-temporal smoothing of information is clearly illustrated in the proof of Proposition 3.4 and Theorem 3.1: The convexity of information cost f implies the incentive to smooth the cost over time. Sub-additivity of I implies incentive to smooth the choice of information structure over time. The incentive for choice of aggregate information structure is illustrated in the proof of existence: monotonicity and sub-additivity implies a concave objective function. Now if any of aforementioned incentives is strict, then the solution is uniquely pinned down in the corresponding aspect. First, consider the proof for convexity of I in the last part. Randomly using S_1 or S_2 (and knowing choice of experiment) carries strictly more information than $\mathcal{S}$ (which discards information about which experiment is used). Therefore, strict monotonicity implies that he objective function is strict concave (except when $\mathcal{S}_1$ and $\mathcal{S}_2$ have the same row vectors). Second, consider step 1 in the proof of Theorem 3.1. Suppose f is strictly convex, whenever information cost is not constant over time, the total cost is strictly dominated by a stationary strategy. Third, when there is strict sub-additivity, then any non-stationary experimentation strategy is dominated by the stationary one I constructed. Moreover, the objective function in Equation (3.1) is strictly concave w.r.t any $\mathcal{A}$. In this case, the whole solution is uniquely pinned down.

3.4 Conclusion

In this chapter, I explore the robust predictions we can make when the measure of signal informativeness is an indirect measure from sequential cost minimization. I first show that an indirect information measure is supported by sequential cost minimization *iff* it satisfies: 1) monotonicity in Blackwell order, 2) sub-additivity in compound experiments and 3) linearity in mixing with no information. In a sequential learning problem, if the cost of information depend on an indirect information measure and delay cost is fixed, then the optimal solution involves direct Poisson signals: arrival of signal directly suggests the optimal action, and non-arrival of signal provides no information. I also characterize the existence and uniqueness of the optimal learning dynamics.

Chapter 4

Information design possibility set

4.1 Introduction

Let X be a non-empty finite set (state space). $\Delta(X) \in \mathbb{R}^{|X|}$ is the set of all probability measures on X. Let μ denote an element in $\Delta(X)$. $\Delta^2(X)$ is the set of all probability measures (standard Borel sigma algebra) on $\Delta(X)$. Let P denote an element in Δ^2(X). Let $\{V^i\}_{i=1}^n$ be a finite set of continuous functions on $\Delta(X)$. Let $f : \mathbb{R}^n \to \mathbb{R}$ denote a continuous function. Let $D(\mu) : \Delta(X) \rightrightarrows \mathbb{R}^n$ be a closed valued correspondence.

My objective is to solve the following constrained maximization problem:

$$\sup_{P \in \Delta^2(X)} f\Big(E_P[V^1], \ldots, E_P[V^n]\Big) \tag{4.1}$$
$$\text{s.t.} \begin{cases} \Big(E_P[V^1], \ldots, E_P[V^n]\Big) \in D(\mu) \\ E_P[\nu] = \mu \end{cases}$$

Suppose $n = 1$ and $D \equiv \mathbb{R}$, then Equation (4.1) can be solved by *concavifying* $V^1(\mu)$ (Kamenica and Gentzkow (2011), Aumann, Maschler, and Stearns (1995)). And Theorem 4.8 implies that it is without loss to consider information structures with no more than $|X|$ signals. This gives tractability both analytically and computationally. However, even when $n = 2$, with a general f or a nontrivial constraint D, concavification no longer works and we might need to search over an infinite dimensional space to solve Equation (4.1).

To solve Equation (4.1), I study the set of all possible expected valuation vectors that can be implemented by designing the information structure P — the *information design possibility set*. In Section 4.3, I proved a two-step concavification method: First, the information design possibility set itself can be implemented by combining finite number of information structures that implement its extreme points. Second, each extreme point can be implemented by concavifying a linear combination of V^i's, hence involving only finite number of signals.

The general concavification method developed in this chapter can be applied to a wide range of information design problems. In Section 4.4, I first provide two applications on static information acquisition and dynamic information acquisition to show that the optimal solutions have a nice Lagrange multiplier characterization. Then I provide an application on persuading receivers with outside options to illustrate how the Lagrange characterization can simplify the optimal persuasion problem. Finally I provide an application of Lemma 4.1 on screening using information structures, to illustrate how the theory developed in this chapter reduces the dimensionality of the problem and makes the problem tractable.

4.2 Information possibility set

Notations used in this section: given a convex set C, let $\text{ext}(C)$ be the set of all extreme points of C, let $\text{ext}_k(C)$ be the set of all k-extreme points of C [1]. Let $\exp(C)$ be the set of exposed points of C. $F(C)$ is set of faces of C.

Definition 4.1. *Information possibility set $\mathcal{V}(\mu) \in \mathbb{R}^n$ is defined as:*

$$\mathcal{V}(\mu) = \left\{ \left(E_P[V^1], \dots, E_P[V^n] \right) \middle| P \in \Delta^2(X), E_P[\nu] = \mu \right\}$$

Lemma 4.1. *$\forall \mu$, $\mathcal{V}(\mu)$ is a compact and convex set. $\forall v \in \text{ext}_k(\mathcal{V}(\mu))$, there exists $P \in \Delta^2(X)$ such that:*

$$\begin{cases} v = \left(E_P[V^1], \dots, E_P[V^n] \right) \\ |\text{supp}(P)| \leqslant (k+1)|X| \end{cases}$$

Proof. First of all, we prove that $\mathcal{V}(\mu)$ is compact and convex.

[1] $\text{ext}(C) = \text{ext}_0(C)$ and $C = \bigcup_{k \leqslant n} \text{ext}_k(C)$.

- Boundedness: $\forall P \in \Delta^2(X)$, $\min_{\mu\in\Delta(X)} V^i(\mu) \leqslant E_P[V^i] \leqslant \max_{\mu\in\Delta(X)} V^i(\mu)$. Therefore, $\forall v \in \mathcal{V}(\mu)$, it is bounded from 0 by $\max_{\mu\in\Delta(X),i}|V^i(\mu)|$ by sup norm. So $\mathcal{V}(\mu)$ is a bounded set.

- Convexity: $\forall v_1, v_2 \in \mathcal{V}(\mu)$, there exists $P_1, P_2 \in \mathcal{D}^2(X)$ s.t. $v_i = \left(E_{P_i}[V^1], \ldots, E_{P_i}[V^n]\right)$. Since $\Delta^2(X)$ is a linear space and expectation operator is linear functional, $\forall \beta \in [0,1]$, $P_\beta = \beta P_1 + (1-\beta)P_2 \in \Delta^2(X)$ and:

$$\begin{aligned} v_\beta =&\left(E_{P_\beta}[V^1], \ldots, E_{P_\beta}[V^n]\right) \\ =&\beta\left(E_{P_1}[V^1], \ldots, E_{P_1}[V^n]\right) + (1-\beta)\left(E_{P_2}[V^1], \ldots, E_{P_2}[V^n]\right) \\ =&\beta v_1 + (1-\beta)v_2 \end{aligned}$$

  Therefore, $\beta v_1 + (1-\beta)v_2 \in \mathcal{V}(\mu)$ so $\mathcal{V}(\mu)$ is a convex set.

- Closeness: $\Delta(X)$ is a finite dimensional simplex. If we consider the Prokhorov metric on $\Delta^2(X)$, then $\Delta^2(X)$ is a complete and separable space (Theorem 6.8 of Billingsley (2013)). Now since $\Delta(X)$ is compact, by Theorem 4.9, $\Delta^2(X)$ is a compact, complete and separable space with the Prokhorov metric. Prokhorov metric induces a topology equivalent to weak* topology(by Theorem 6.8 of Billingsley (2013)). So $\forall v_k \in \mathcal{V}(\mu)$, if $v_k \to v$, then consider the sequence P_k such that $v_k = E_{P_k}\left[(V^i)\right]$. By compactness of $\Delta^2(X)$, pick a subsequence $P_k \xrightarrow{w-*} P$. Then $\forall V^i$, since V^i is continuous, $E_{P_k}[V^i] \to E_P[V^i]$. So $v \in \mathcal{V}(\mu)$ and $\mathcal{V}(\mu)$ is a closed set.

- Compactness: $\mathcal{V}(\mu)$ is a finite dimensional bounded and closed set, so it is compact.

$\forall v \in \text{ext}_k(\mathcal{V}(\mu))$, v is an interior point of a k-dimensional face F of $\mathcal{V}(\mu)$. Then by Theorem 4.7, $v \in \text{conv}(\text{ext}(F))$. By Theorem 4.8, there exists $\{v_j\}_{j=1}^{k+1} \subset \text{ext}(F)$ and $\sum \pi_j = 1$ s.t. $\sum \pi_j v_j = v$. By Lemma 4.4, $\{v_j\} \subset \text{ext}(\mathcal{V}(\mu))$. The next step is to prove that $\forall j$, there exists $P_j \in \Delta^2(X)$ s.t. $v_j = \left(E_P[V^1], \ldots, E_P[V^n]\right)$ and $|\text{supp}(P_j)| \leqslant |X|$.

Lemma 4.2. *$\forall \mu$, $\forall v \in \exp(\mathcal{V}(\mu))$, $\exists P \in \Delta^2(X)$ and $|\mathrm{supp}(P)| \leqslant |X|$ s.t. $v = E_P\big[(V^i)\big]$.*

Proof. By definition of exposed points, there exists a linear function $l \in L(\mathbb{R}^n)$ s.t.

$$l(v) > l(v') \ \forall v' \in \mathcal{V}(x), v' \neq v$$

In finite dimensional space, a linear function $l(v)$ can be equivalently written as $\sum \lambda_i v_i + c$. Consider the following maximization problem:

$$\sup_{P \in \Delta^2(X)} E_P\Big[\sum \lambda_i V^i + c\Big] \tag{4.2}$$
$$\text{s.t. } E_P[\nu] = \mu$$

By Theorem 4.8, Equation (4.2) can be solved by convexifying the graph of $\sum \lambda_i V^i(\mu) + c$. The maximum is achieved by a P s.t. $|\mathrm{supp}(P)| \leqslant |X|$. Of course $E_P[(V^i)] \in \mathcal{V}(\mu)$. Then by definition of l, $l(v) \geqslant E_P\big[\sum \lambda_i V^i + c\big]$. On the other hand, there exists $P' \in \Delta^2(X)$ s.t. $v = E_{P'}\big[(V^i)\big]$, by optimality of P, $l(v) \leqslant E_P\big[\sum \lambda_i V^i + c\big]$. Therefore, since v is the unique element in $\mathcal{V}(\mu)$ achieving $l(v)$, we have $E_P[(V^i)] = v$ and $|\mathrm{supp}(P)| \leqslant |X|$. ■

$\forall v^j \in \mathrm{ext}(\mathcal{V}(\mu))$, by Theorem 4.6, there exists $\{v^{jl}\}_{l=1}^{\infty} \subset \exp(\mathcal{V}(\mu))$ and $\lim_{l\to\infty} v^{jl} = v^j$. By Lemma 4.2, there exists $P^{jl} \in \Delta^2(X)$ s.t. $\big|\mathrm{supp}(P^{jl})\big| \leqslant |X|$ and $v^{jl} = E_{P^{jl}}\big[(V^i)\big]$. Now each P^{jl} can be represented as $\left(p_t^{jl}, \mu_t^{jl}\right)_{t=1}^{|X|} \in \mathbb{R}^{2|X|}$, where:

$$\begin{cases} \sum_t p_t^{jl} = 1 \\ \sum_t p_t^{jl} \mu_t^{jl} = \mu \\ \sum_t p_t^{jl} V^i(\mu_t^{jl}) = v_i^{jl} \ \forall i \end{cases}$$

Since $\left(p_t^{jl}, \mu_t^{jl}\right)$ is in finite dimensional vector space, there exists a subsequence converging to $\left(p_t^j, \mu_t^j\right)$ when $l \to \infty$. Therefore, since V^i is each continuous, it is easy to verify that:

$$\begin{cases} \sum_t p_t^j = 1 \\ \sum_t p_t^j \mu_t^j = \mu \\ \sum_t p_t^j V^i(\mu_t^j) = v_i^j \ \forall i \end{cases}$$

Therefore, v^j is implemented by $P^j \in \Delta^2(X)$ and $\left|\text{supp}(P^j)\right| \leqslant |X|$. So $P = \sum \pi_j P^j \in \Delta^2(X)$ and $|\text{supp}(P)| \leqslant (k+1) \cdot |X|$. By linearity of expectation operator, $E_P\left[(V^i)\right] = \sum \pi_j E_{P^j}\left[(V^i)\right] = \sum \pi_j v^j = v$. ■

Lemma 4.3. *Correspondence $\mathcal{V} : \Delta(X) \rightrightarrows \mathbb{R}^n$ is continuous.* $\text{Gr}(\mathcal{V})$ *is convex and compact.*

Proof.

- Boundedness: $\Delta(X)$ is a bounded set. $\forall \mu \in \Delta(X)$, $\mathcal{V}$ is uniformly bounded by radious $\max_{\mu \in \Delta(X), i}\left|V^i(\mu)\right|$ by sup norm. So $\text{Gr}(\mathcal{V})$ is bounded.

- Convexity: $\forall (\mu_1, v_1), (\mu_2, v_2) \in \text{Gr}(\mathcal{V})$. $\forall \alpha \in [0,1]$. Since $\Delta(X)$ is convex, $\mu_\alpha = \alpha\mu_1 + (1-\alpha)\mu_2 \in \Delta(X)$. Now we prove that $v_\alpha = \alpha v_1 + (1-\alpha) v_2 \in \mathcal{V}(\mu_\alpha)$. By definition, there exists $P_1, P_2 \in \Delta^2(X)$ s.t. $E_{P_1}\left[(V^i)\right] = v_1$, $E_{P_1}[\nu] = \mu_1$ and $E_{P_2}\left[(v^i)\right] = v_2$, $E_{P_2}[\nu] = \mu_2$. Define $P_\alpha = \alpha P_1 + (1-\alpha) P_2$, then by linearity of the expectation operator, $E_{P_\alpha}[\nu] = \alpha E_{P_1}[\nu] + (1-\alpha) E_{P_2}[\nu] = \mu_\alpha$. $E_{P_\alpha}\left[(V^i)\right] = \alpha E_{P_1}[(V^i)] + (1-\alpha) E_{P_2}[(V^i)] = v_\alpha$. Therefore, $v_\alpha \in \mathcal{V}(\mu_\alpha)$. So $(\mu_\alpha, v_\alpha) \in \text{Gr}(\mathcal{V})$.

- Closedness: $\forall \left\{(\mu_j, v_j)\right\} \subset \text{Gr}(\mathcal{V})$, suppose $\mu_j \to \mu$, $v_j \to v$. Want to show that $\mu \in \Delta(X)$ and $v \in \mathcal{V}(\mu)$. First of all, since $\Delta(X)$ is complete, $\mu \in \Delta(X)$. Now by

Lemma 4.1, there exists (p_j, ν_j) such that:

$$\begin{cases} \sum_{k=1}^{(n+1)|X|} p_j^k = 1 \\ \sum_{k=1}^{(n+1)|X|} p_j^k \nu_j^k = \mu_j \\ \sum_{k=1}^{(n+1)|X|} p_j^k V^i(\nu_j^k) = v_j^i \end{cases}$$

Now since $p_j \in \Delta((n+1)|X|)$ and $\nu_j \in \Delta(X)$ are both compact spaces. Consider stadard Euclidean metric on product space $\Delta((n+1)|X|) \times \Delta(X)^{(n+1)|X|}$, it is also compact. Therefore there exists convergincing subsequence $p_j \to p$ and $\nu_j^k \to \nu^k$. Then

$$\begin{cases} \sum_{k=1}^{(n+1)|X|} p^k = \lim_{j\to\infty} \sum_{k=1}^{(n+1)|X|} p_j^k = 1 \\ \sum_{k=1}^{(n+1)|X|} p^k \nu^k = \lim_{j\to\infty} \sum_{k=1}^{(n+1)|X|} p_j^k \nu_j^k = \lim_{j\to\infty} \mu_j = \mu \\ \sum_{k=1}^{(n+1)|X|} p^k V^i(\nu^k) = \lim_{j\to\infty} \sum_{k=1}^{(n+1)|X|} p_j^k V^i(\nu_j^k) = \lim_{j\to\infty} v_j^i = v^i \end{cases}$$

Therefore, (p, ν) implements v at μ. So $v \in \mathcal{V}(\mu)$.

- Compactness: Since $\mathrm{Gr}(\mathcal{V})$ is closed and bounded, it is compact.

- Continuity: Since $\mathrm{Gr}(\mathcal{V})$ is compact, $\mathcal{V}(\mu)$ is upper hemicontinuous. Now we only need to show lower hemicontinuity. $\forall (\mu_m) \subset \Delta(X)$, $\mu_m \to \mu \in \Delta(X)$. $\forall v \in \mathcal{V}(\mu)$. By Lemma 4.1, v is impelemnted by (p, ν) with support size $(n+1)|X|$. There exists a stochastic matrix q_{jk} such that:

$$\begin{cases} \nu_j = \frac{1}{\sum_k \mu_k q_{jk}} (\mu_1 q_{j1}, \ldots, \mu_{-1} q_{j,-1}) \\ p_j = \sum_k \mu_k q_{jk} \end{cases}$$

$$\Longrightarrow \begin{cases} \dfrac{\partial p_j}{\mu_k} = q_{jk} \\ \dfrac{\partial \nu_{jl}}{\partial \mu_k} = \begin{cases} \dfrac{p_j q_{jl} - \mu_l q_{jl}^2}{p_j^2} & \text{when } k = l \\ \dfrac{-\mu_l q_{jl} q_{jk}}{p_j^2} & \text{when } k \neq l \end{cases} \end{cases}$$

Therefore, since each $p_j > 0$, when μ_m is sufficiently close to μ, corresponding (p_m, ν_m) will be bounded from (p, ν) by $|\mu - \mu_m|$. By continuity of V^i, $v_m = \left(\sum p_m V^i(\nu_m)\right) \to \left(\sum p V^i(\nu)\right) = v$. Therefore, $\mathcal{V}(\mu)$ is both upper hemicontinuous and lower hemicontinuous.

■

4.3 Main theorem

4.3.1 Existence and finite support

Theorem 4.1. *Let X be finite and non-empty, $\{V^i\}_{i=1}^n \subset C\Delta(X)$, $f \in C\mathbb{R}^n$. $\forall \mu \in \Delta(X)$, suppose $\mathcal{V}(\mu) \bigcap D(\mu) \neq \emptyset$, then there exists $P^* \in \Delta^2(X)$ solving Equation (4.1) and $|\mathrm{supp}(P^*)| \leqslant (n+1) \cdot |X|$.*

Proof. By definition of $\mathcal{V}(\mu)$, Equation (4.1) is equivalent to the following problem:

$$\sup_{v \in D \bigcap \mathcal{V}(\mu)} f(v) \tag{4.3}$$

By Lemma 4.1, $\mathcal{V}(\mu)$ is a compact set. Then $\mathcal{V}(\mu) \bigcap D(\mu)$ is compact and non-empty. By Weierstrass's theorem, there exists $v^* \in \mathcal{V}(\mu) \bigcap D(\mu)$ solving Equation (4.3). Then by Lemma 4.1, there exists $P^* \in \Delta^2(X)$ s.t. $v^* = \left(E_{P*}[V^1], \ldots, E_{P*}[V^n]\right)$ and $|\mathrm{supp}(P^*)| \leqslant (n+1) \cdot |X|$. Therefore, P^* solves Equation (4.1). ■

4.3.2 Necessary condition for the maximizer

Theorem 4.2. *Let X be finite and non-empty, $\{V^i\}_{i=1}^n \subset C\Delta(X)$, $f : \mathbb{R}^n \to \mathbb{R}$ is differentiable. Let $D \equiv \mathbb{R}^n$. Then a necessary condition for P^* solving Equation (4.1) is:*

$$P^* \in \arg \max_{\substack{P \in \Delta^2(X) \\ E_P[v]=\mu}} \nabla f(E_{P*}[V^1], \ldots, E_{P*}[V^n]) \cdot \Big(E_P[V^1], \ldots, E_P[V^n]\Big) \tag{4.4}$$

Proof. Solving Equation (4.1) is equivalent to solving Equation (4.3). Suppose by contradiction that Equation (4.4) is violated at optimal P^*. Then it is equivalently saying that there exists $v \in \mathcal{V}(\mu)$ such that:

$$\nabla f(v^*) \cdot v^* < \nabla f(v^*) \cdot v$$

By Lemma 4.1, $\mathcal{V}(\mu)$ is a convex set. Therefore $v_\alpha = (1-\alpha)v^* + \alpha v \in \mathcal{V}(\mu)$. Consider $h(\alpha) = f(v_\alpha)$. Then $h'(0) = \nabla f(v^*) \cdot (v - v^*) > 0$. So there exists $\alpha' > 0$ s.t. $h(\alpha') > h(0)$. Then $f(v^*) < f(v_{\alpha'})$. Contradicting optimality of v^*. ■

Theorem 4.3. *Let X be finite and non-empty, $\{V^i\}_{i=1}^{n+m} \subset C\Delta(X)$, $f : \mathbb{R}^{n+m} \to \mathbb{R}$ is constant in the last m arguments. Let $D \equiv \{v | v^i \geqslant 0 \; \forall i > n\}$. Then there exists P^* solving Equation (4.1) and $\lambda \in B_{m+n}$ such that:*

$$P^* \in \arg \max_{\substack{P \in \Delta^2(X) \\ E_P[v]=\mu}} E_P\Big[\sum \lambda^i V^i\Big]$$

Proof. $\forall P^*$ solving Equation (4.1), let v^* be corresponding value. Define:

$$v_\alpha = v^* + \alpha(\underbrace{0,\ldots,0}_{n},\underbrace{1,\ldots,1}_{m})$$

Then by definition $f(v_\alpha) = f(v^*)$. $v_0 = v^* \in \mathcal{V}(\mu)$. Since $\mathcal{V}(\mu)$ is bounded, for large enough α, $v_\alpha \notin \mathcal{V}(\mu)$. Then since $\mathcal{V}(\mu)$ is compact, there exists α s.t. $v_\alpha \in \partial\mathcal{V}(\mu)$. Since $\mathcal{V}(\mu)$ is convex, there exists $l \in L(\mathbb{R}^{m+n})$ s.t. $v_\alpha \in \arg\max_{v\in\mathcal{V}(\mu)} l(v)$. Let $l = \sum \lambda^i v^i$, then:

$$v_\alpha \in \arg\max_{v\in\mathcal{V}(\mu)} \sum \lambda^i v^i$$

Let P_α be the corresponding information structure implementing v_α (existence of P_α guaranteed by Lemma 4.1). Then

$$P_\alpha \in \arg\max_{\substack{P\in\Delta^2(X)\\ E_P[\nu]=\mu}} E_P\left[\sum \lambda^i V^i\right]$$

Since $f(v_\alpha) = f(v^*)$, P_α solves Equation (4.1) as well. ■

4.3.3 Convex optimization

Theorem 4.4. *Let X be finite and non-empty, $\{V^i\}_{i=1}^n \subset C\Delta(X)$, $D \equiv \{v|g(v) \geqslant 0\}$. If both f and g are quasi-concave and continuous, then there exists P^* solving Equation (4.1), $v^* = (E_P[V^i])$ and $\lambda \in B_n$ such that:*

$$\begin{cases} P^* \in \arg\max\limits_{\substack{P\in\Delta^2(X)\\ E_P[\nu]=\mu}} E_P\left[\sum \lambda^i V^i\right] \\ v^* \in \arg\min\limits_{f(v)\geqslant f(v^*),v\in D} \lambda\cdot v \end{cases}$$

Proof. First, by Theorem 4.1, P^* solving Equation (4.1) exists. Then by optimality of P^*:

$$\mathcal{V}(\mu) \bigcap \{v | v \in D, f(v) > f(v^*)\} = \varnothing$$

Since f and g are quasi-convex, $\{v | v \in D, f(v) > f(v^*)\}$ is a convex set. Then by separating hyperplane theorem, there exists c and λ s.t. $\forall v \in \mathcal{V}(\mu)$, $v' \in D$ and $f(v') > f(v^*)$:

$$\lambda \cdot v \leqslant c \text{ and } \lambda \cdot v' > c$$

By continuity of f and g, $v^* \in \text{cl}(\{v | v \in D, f(v) > f(v^*)\})$. So $\lambda \cdot v^* = c$. Then it is easy to verify that λ satisfies the conditions in Theorem 4.4. ■

Corollary 4.4.1. *Let X be finite and non-empty, $\{V^i\}_{i=1}^n \subset C\Delta(X)$, $f : \mathbb{R}^n \to \mathbb{R}$ is quasi-concave. Let $D \equiv \{v | g(v) \geqslant 0\}$, g is quasi-concave. If f and g are both differentiable, then there exists P^* solving Equation (4.1), $v^* = (E_P[V^i])$ and $\gamma, \eta \geqslant 0$ such that:*

$$P^* \in \arg \max_{\substack{P \in \Delta^2(X) \\ E_P[v] = \mu}} (\eta \nabla f(v^*) + \gamma \cdot \text{J}g(v^*)) \cdot \left(E_P[V^1], \cdots, E_P[V^n]\right)$$

Proof. By Theorem 4.4:

$$v^* \in \arg \min_{f(v) \geqslant f(v^*), v \in D} \lambda \cdot v \tag{4.5}$$

It is easy to verify that Equation (4.5) as a dual problem is a convex optimization problem. Since both f and g are differentiable, by Kuhn-Tucker condition, there exists $\gamma, \eta \geqslant 0$ such that:

$$\lambda - \eta \cdot \nabla f(v^*) - \gamma \cdot \mathrm{J}g(v^*) = 0$$

Then by definition of λ:

$$P^* \in \arg \max_{\substack{P \in \Delta^2(X) \\ E_P[\nu]=\mu}} (\eta \nabla f(v^*) + \gamma \cdot \mathrm{J}g(v^*)) \cdot \left(E_P[V^1], \cdots, E_P[V^n]\right)$$

■

4.3.4 Maximum theorem

Theorem 4.5. *Let X be finite and non-empty, $\{V^i\}_{i=1}^n \subset C\Delta(X)$, $f \in C\mathbb{R}^n$. Suppose $D(\mu)$ is a continuous correspondence and $\forall \mu \in \Delta(X)$, $\mathcal{V}(\mu) \bigcap D(\mu) \neq \varnothing$. Let $\kappa(\mu)$ be the maximum of Equation (4.1) and $\mathcal{P}(\mu)$ be the maximizer of Equation (4.1), then $\kappa(\mu)$ is continuous and $\mathcal{P}(\mu)$ is compact-valued and upper hemicontinuous*[2].

Proof. Theorem 4.5 is an application of the maximum theorem. Since by Lemma 4.3 $\mathcal{V}(\mu)$ and $D(\mu)$ are both continuous, $\mathcal{V}(\mu) \bigcap D(\mu)$ is non-empty, compact valued and continuous. Equation (4.1) is equivalent to maximizing $f(v)$ on $\mathcal{V}(\mu) \bigcap D(\mu)$. Therefore, by maximum theorem, $\kappa(\mu)$ is continuous and the argmax correspondence $V^*(\mu)$ is comapct-valued and upper hemicontinuous.

Now we show that $\mathcal{P}(\mu)$ is compact valued and upper hemicontinuous.

- compactness: (sequential comapctness will be sufficient) $\forall \{P_m\} \subset \mathcal{P}(\mu)$, consider $v_m = E_{P_m}\left[(V^i)\right]$. Then $v_m \in V^*(\mu)$, so there exists subsequence (without loss assume

[2] with respect to Prokhorov metric.

to be $v)m$ itself) $v_m \to v \in V^*(\mu)$. Then since $\Delta^2(X)$ is compact by Theorem 4.9, there exists subsequence $P_m \xrightarrow{w-*} P \in \Delta^2(X)$. Then $E_P[(V^i)] = \lim E_{P_m}[(V^i)] = \lim v_m = v \in V^*(\mu)$. So $P \in \mathcal{P}(\mu)$.

- upper hemicontinuity: $\forall \mu_m \to \mu$, $P_m \xrightarrow{w-*} P$ and $P_m \in \mathcal{P}(\mu_m)$. Then $v_m = E_{P_m}[(V^i)] \in V^*(\mu_m)$. By definition of w-* convergence, $v_m \to v = E_P[(V^i)]$. By upper hemicontinuity of $V^*(\mu)$, $v \in V^*(\mu)$. Therefore, $P \in \mathcal{P}(\mu)$.

■

4.4 Applications

4.4.1 *Costly Information acquisition*

A direct application of Theorem 4.1 is costly information acquisition problems. Consider a variant of the rational inattention model. Decision utility at each belief is $F(\mu) = \max_a E_\mu[u(a,x)]$. The information measure of any experiment P is $I(P|\mu) = E_P[H(\mu) - H(\nu)]$ where H is the standard entropy function. Assume that the cost of experiments is convex in their measure, the decision problem can be written as:

$$\sup_{\substack{P \in D^2(X) \\ E_P[\nu]=\mu}} E_P[F(\nu)] - f(E_P[H(\mu) - H(\nu)]) \tag{4.6}$$

In a standard rational inattention problem, f is linear. Then standard concavification method suggests that optimal experiment involves signals no more than $|X|$. The reason why we want to deviate from a linear f is that standard RI has two kind of debatable predictions: 1) prior invariant choice of optimal posteriors (see Caplin and Dean (2013)). 2) no dynamics if we allow repeated experiments (see Steiner, Stewart, and Matějka (2017)). However, when f is more general, say convex, we knew little about how to solve Equation (4.6). Theorem 4.2 becomes useful.

Proposition 4.1. *There exists P^* solving Equation (4.6), $|\text{supp}(P^*)| = 2|X|$. Moreover, if f is differentiable, P^* solves:*

$$P^* \in \arg \max_{\substack{P \in \Delta^2(X) \\ E_P[\nu]=\mu}} E_P\big[F(\nu) - f'(E_{P*}[H(\mu) - H(\nu)]) \cdot H(\nu)\big]$$

4.4.2 Dynamic information design

Consider the following Bellman equation:

$$V(\mu) = \max\left\{F(\mu), \sup_{P \in \Delta^2(X)} e^{-\rho dt} E_P[V(\nu)] - f(E_P[H(\mu) - H(\nu)])\right\} \tag{4.7}$$

$$\text{s.t.} \begin{cases} E_P[\nu] = \mu \\ E_P[H(\mu) - H(\nu)] \leqslant C \end{cases}$$

Proposition 4.2. *If $F, H \in C\Delta(X)$, $f \in C\mathbb{R}$. $F(x), f(x), C \geqslant 0$. Then there exists unique $V \in C\Delta(X)$ solving Equation (4.7).*

Proof. Let $\mathcal{Z} = \{V \in C\Delta(X) | F \leqslant V \leqslant \text{co}(F)\}$. We define operator:

$$T(V)(\mu) = \max\left\{F(\mu), \sup_{P \in \Delta^2(X)} e^{-\rho dt} E_P[V(\nu)] - f(E_P[H(\mu) - H(\nu)])\right\} \tag{4.8}$$

$$\text{s.t.} \begin{cases} E_P[\nu] = \mu \\ E_P[H(\mu) - H(\nu)] \leqslant C \end{cases}$$

By Theorem 4.1, the max operator is well defined. When $P = \delta_\mu$, $E_P[\nu] = \mu$ and $E_P[H(\mu) - H(\nu)] = 0$ so the sup operator is also well defined. Now we prove that T is a contraction mapping on $(\mathcal{Z}, L_\infty)$.

- $T(\mathcal{Z}) \subset \mathcal{Z}$: First of all, given the outter max operator in Equation (4.8), $T(V)(\mu) \geqslant$

$F(\mu)$. Then $\forall P \in \Delta^2(X)$ such that $E_P[\nu] = \mu$ and $E_P[H(\mu) - H(\nu)] \leqslant C$:

$$\begin{aligned}
&e^{-\rho dt} E_P[V(\nu)] - f(E_P[H(\mu) - H(\nu)]) \\
\leqslant & e^{-\rho dt} E_P[V(\nu)] \\
\leqslant & E_P[\mathrm{co}(F)(\nu)] \\
= & \mathrm{co}(F)(\mu)
\end{aligned}$$

First inequality is from f being non-negative, second inqeuality is from V being non-negative, $e^{-\rho dt} < 1$ and $V \leqslant \mathrm{co}(F)$. Last equality is from $\mathrm{co}(F)$ being linear. Last step is to show $T(\mathcal{Z})(\mu) \in C\Delta(X)$. This is directly implied by Theorem 4.5.

- $T(V)$ is monotonic: Suppose $U(\mu) \geqslant 0$ and $U + V \in \mathcal{Z}$ If $T(V)(\mu) = F(\mu)$, then by construction $T(V + U) \geqslant F(\mu) = T(V)(\mu)$. If $T(V)(\mu) > F(\mu)$, let P be solution to Equation (4.8) at μ for V:

$$\begin{aligned}
T(V + U)(\mu) \geqslant & e^{-\rho dt} E_P[V(\nu) + U(\nu)] - f(E_P[H(\mu) - H(\nu)]) \\
= & T(V)(\mu) + e^{-\rho dt} E_P[U(\nu)] \\
\geqslant & T(V)(\mu)
\end{aligned}$$

  And constraints $E_P[H(\mu) - H(\nu)] \leqslant C$ and $E_P[\nu] = \mu$ are independent of choice of V so still satisfied.

- $T(V)$ is contraction. We claim that $T(V + \alpha)(\mu) \leqslant T(V)(\mu) + e^{-\rho dt}\alpha$. Suppose not true at μ. Obviously $T(V + \alpha)(\mu) > F(\mu)$. Then let P be the solution of Equation (4.8) at μ for $V + \alpha$.

$$T(V)(\mu) \geqslant e^{-\rho dt} E_P[V(\nu)] - f(E_P[H(\mu) - H(\nu)])$$

$$
\begin{aligned}
&= e^{-\rho dt} E_P[V(\nu) + \alpha] - f(E_P[H(\mu) - H(\nu)]) - e^{-\rho dt}\alpha \\
&= T(V + \alpha)(\mu) - e^{-\rho dt}\alpha \\
&> T(V)(\mu)
\end{aligned}
$$

Similar to last part, constraints $E_P[H(\mu) - H(\nu)] \leqslant C$ and $E_P[\nu] = \mu$ are still satisfied. Contradiction.

Therefore, by Blackwell condition, $T(V)$ is a contraction mapping on $\mathcal{Z}$. There exists a unique solution $V \in \mathcal{Z}$ solving the fixed point problem $T(V) = V$. ∎

4.4.3 Persuade voters with outside options

Consider a politician who can strategically design a public signal to voters to influence their voting behavior (the setup in Alonso and Câmara (2016)).

Voting game: There are $n \geqslant 1$ voters who chooses from a binary policy set $A = \{a_0, a_1\}$. There are two states $X = \{x_0, x_1\}$. Each voter gets Bernoulli utility $u_i(a, x)$ from voting for the policy a. Assume that a_1 is unanimously preferred to a_0 when x_1 is the true state and vice versa. The politician has state independent utility over policies and prefers a_1 strictly to a_0. I assume that a_0 is a default policy. For a_1 to be proved, the politician needs more than m ($m \leqslant n$) voters to voter for a_1. The politician can design a signal structure to influence voters' decisions. Equivalently, I assume that the politician chooses a distribution over posterior beliefs $P \in \Delta^2(X)$.

Outside option: Different from Alonso and Câmara (2016), where number of potential voters is fixed, I assume that each voter has opportunity cost c_i of participating in the voting game. Therefore, to approve the new policy, the politician should first attract at least m voters to the game and then persuade them to vote for a_1.

To simplify notation, I write all voter's utility as functions of belief $F_i(\mu)$. Let $\bar{\mu}_i$ be the threshold belief for each voter to vote for a_1 The politician's optimization problem can be written as:

$$\sup_{i_1,\ldots,i_k,P} E_P\left[\mathbf{1}_{\#(\mu \geqslant \bar{\mu}_{i_j}) \geqslant m}\right] \tag{4.9}$$
$$\text{s.t.} \begin{cases} E_P[F_{i_j}] \geqslant c_{i_j} \\ E_P[\nu] = \mu \end{cases}$$

Notice that in Equation (4.9), the politician doesn't necessarily need to exclude voters outside of $\{i_1, \ldots, i_k\}$, so the maximum from Equation (4.9) must be weakly larger than the politician's optimal utility. On the other hand, for any strategy in Equation (4.9), potentially including more voters to the voting game can only make the politician better off. So Equation (4.9) exactly characterizes the politician's optimization problem.

For any voter, except for $\bar{\mu}_i$, there is another critical belief $\tilde{\mu}_i$:

$$\frac{\tilde{\mu}_i - \mu}{\tilde{\mu}_i} F_i(0) + \frac{\mu}{\tilde{\mu}_i} F_i(\tilde{\mu}_i) = c_i$$

Suppose voter observes information structure inducing posterior belief 0 and $\tilde{\mu}_i$, then the voter is exactly indifferent between paying the opportunity cost and entering the voting game and not.

Proposition 4.3. *Let μ^* be the smallest belief s.t. $\#\{i|\bar{\mu}_i \geqslant \mu^*\} \geqslant m$ and $\#\{i|\tilde{\mu}_i \geqslant \mu^*\} \geqslant m$, then the optimal strategy for Equation (4.9) is:*

$$\begin{cases} P(0) = \frac{\mu^* - \mu}{\mu^*} \\ P(\mu^*) = \frac{\mu}{\mu^*} \end{cases}$$

and $\{i_1, \ldots, i_k\} = \{i|\min\{\tilde{\mu}_i, \bar{\mu}_i\} \geqslant \mu^*\}$.

Proposition 4.3 states that when voters must pay opportunity cost to enter the voting game, then there are potentially two pivotal voters. One is the one who's most difficult to persuade to adopt a_1, and the other is the one who's most difficult to attract to the voting game. Both *difficulty* levels are measured by the location of the critical beliefs.

Proof. The key step of proving Proposition 4.3 is to apply Corollary 4.4.1 to Equation (4.9). Notice that the objective function is Equation (4.9) is in fact an indicator function with some threshold belief level (say μ', which is the lowest belief to persuade at least m voters to vote for a_1). So Corollary 4.4.1 is directly applicable to Equation (4.9), and the objective function is in the form of:

$$\sum \lambda_{i_j} \max\left\{0, \mu - \bar{\mu}_{i_j}\right\} + \mathbf{1}_{\mu \geqslant \mu'} \tag{4.10}$$

It is easy to see that Equation (4.10) is a convex function on $\mu \in [0, \mu']$ and a linear function on $\mu \in [\mu', 1]$ (there is no point to include voters who will never vote for a_1.). So optimal persuasion strategy must induce either belief 0 or interior belief $\nu > \mu'$. Of course since at least m voters are included and persuaded, $\nu \geqslant \mu^*$. On the other hand, it is easy to verify that the strategy define by μ^* induces at least m voters to participate, so μ^* is optimal. ■

4.4.4 *Screening with information*

Consider a problem of Bayesian persuasion with unknown receiver types. Let Θ be the set of receiver types, X be the finite set of states and A be the set of actions. $\forall \theta \in \Theta$, decision utility at each belief is $F_\theta(\mu) = \max_a E_\mu[u(a, x, \theta)]$. Sender's utility at each belief given receiver type θ is $V_\theta(\mu)$. Assume that the type distribution is $\pi(\theta) \in \Delta(\Theta)$. The sender can screen the receivers by providing a menu of information structures. Then by revelation principle, sender's optimization problem is:

$$\sup_{P_\theta \in \Theta \times \Delta^2(X)} \int E_{P_\theta}[V_\theta] \pi(\theta) \mathrm{d}\theta \tag{4.11}$$

$$\text{s.t.} \begin{cases} E_{P_\theta}[F_\theta] \geqslant E_{P_{\theta'}}[F_\theta] \; \forall \theta, \theta' \in \Theta \\ E_{P_\theta}[\nu] = \mu \; \forall \theta \in \Theta \end{cases}$$

When Θ and A are both infinite, solving Equation (4.11) is difficult due to the dimensionality of strategy space. When A is finite, it is WLOG to restrict the sender to use direct message which suggests the actions being played conditional on the state. Then Equation (4.11) reduces to a screening problem with finite dimensional strategy function (plus a few more obedience constraints). In the remaining case where Θ is finite but A is infinite, it is still unclear whether it is WLOG to consider only finite dimensional screening mechanisms.

Now consider the finite Θ case. Suppose $\Theta = \{1, \ldots, N\}$. Define:

$$\begin{cases} \mathcal{V}_i(\mu) = \left\{ E_P[V_i], E_P[F_1], \ldots, E_P[F_N] \middle| P \in \Delta^2(X), E_P[\nu] = \mu \right\} \\ D(\mu) = \left\{ v \in \mathbb{R}^{(N\times(N+1))} \middle| v_i^{i+1} \geqslant v_j^{i+1} \forall i, j \right\} \end{cases}$$

Then Equation (4.11) is equivalent to the following problem:

$$\sup_{v \in D(\mu) \bigcap \times_{i=1}^N \mathcal{V}_i(\mu)} \pi_i v_i^1 \tag{4.12}$$

By Lemma 4.1, each $\mathcal{V}_i(\mu)$ is compact set. Therefore, $D(\mu) \bigcap \times \mathcal{V}_i(\mu)$ is compact. It is easy to see that $D(\mu) \bigcap \times \mathcal{V}_i(\mu)$ is non-empty. By Weierstrass's theorem, there exists v^* solving Equation (4.12). Then by Lemma 4.1, there exists $P_i^* \in \Delta^2(X)$ s.t. $v_i^* = \left(E_{P_i^*}[V_i], E_{P_i^*}[F_1], \ldots, E_{P_i^*}[F_N] \right)$ and $\left|\text{supp}(P_i^*)\right| \leqslant (N+2) \cdot |X|$. Therefore, $\left(P_1^*, \ldots, P_N^*\right)$ solves Equation (4.11) and we get the following proposition:

Proposition 4.4. *If Θ is finite, then $\forall \mu \in \Delta(X)$, there exists $\left(P_1^*, \ldots, P_N^*\right) \in \Delta^2(X)^N$ solving Equation (4.11) and each $\left|\text{supp}(P_i^*)\right| \leqslant (N+2) \cdot |X|$.*

Proposition 4.4 states that it is WLOG to consider only mechanisms with finite support when solving Equation (4.11). Therefore, it is sufficient to maximize over $N(N+2) \cdot |X|$ posterior beliefs and $N(N+2) \cdot |X|$ corresponding probabilities to solve constrained optimization problem Equation (4.11), which is a computationally tractable problem.

4.5 Conclusion

In this chapter, I study the set of all possible combinations of expected valuations that can be implemented by designing information. I show that the set can be implemented only using information structures with finite realizations, and all extreme points of the set can be characterized using a concavification characterization. I developed a Lagrange method in the information design setup, and applied the results to various applications including static and dynamic information acquisition, persuasion of receivers with outside options and screening using information.

4.6 Theorems used in proof

Here I list the key theorems used for my proof. Theorem 4.6 is Straszewicz's theorem (Straszewicz (1935), see Theorem 18.6 of Rockafellar (1969)). Theorem 4.7 is Krein-Milman theorem(see Theorem 3.23 of Rudin (1991)). Theorem 4.8 is Carathéodory's theorem (Carathéodory (1907)).Theorem 4.9 is Prokhorov's theorem (see Theorem 5.1 of Billingsley (2013))

Theorem 4.6. *Let $C \in \mathbb{R}^n$ be a closed convex set,* $\mathrm{cl}(\exp(C)) = \mathrm{ext}(C)$.

Theorem 4.7. *Let $C \in \mathbb{R}^n$ be a compact and convex set,* $C = \mathrm{conv}(\mathrm{ext}(C))$.

Theorem 4.8. *Let $C \in \mathbb{R}^n$, if $x \in \mathrm{conv}(C)$ then $x \in \mathrm{conv}(R)$ for $R \subset C$, $|R| \leqslant n+1$.*

Theorem 4.9. *A tight set Π of probability measures on Borel sets of metric topological space $\mathcal{X}$ is relative compace in weak-* topology.*

Lemma 4.4. *Let C be a convex set in $\mathbb{R}^n$. Then $\forall F \in F(C)$, $\mathrm{ext}(F) \subset \mathrm{ext}(C)$.*

Proof. $\forall x \in \mathrm{ext}(F)$ there exists affine f defining face F. $\forall y, z \in C$. Suppose $y \in F$, then $f(x) = f(y)$. If there exists $\alpha \in (0,1)$ s.t. $\alpha y + (1-\alpha)z = x$, then $\alpha f(y) + (1-\alpha) f(z) = f(x) \implies f(z) = f(x) = f(y)$ so $z \in F$. Since $x \in \mathrm{ext}(F)$, $x \in \{y, z\}$. Suppose $y \notin F$, then $f(x) = \alpha f(y) + (1-\alpha) f(z) < f(x)$ by definition of f, contradiction. To sum up, $x \in \mathrm{ext}(C)$. ■

Bibliography

Alonso, Ricardo and Odilon Câmara (Nov. 2016). "Persuading Voters." In: *American Economic Review* 106.11, pp. 3590–3605. DOI: 10.1257/aer.20140737. URL: http://www.aeaweb.org/articles?id=10.1257/aer.20140737.

Arrow, Kenneth J, David Blackwell, and Meyer A Girshick (1949). "Bayes and minimax solutions of sequential decision problems." In: *Econometrica, Journal of the Econometric Society*, pp. 213–244.

Aumann, Robert J, Michael Maschler, and Richard E Stearns (1995). *Repeated games with incomplete information*. MIT press.

Bellman, Richard (1957). *Dynamic programming*. Princeton University Press.

Bergemann, Dirk and Stephen Morris (2017). "Information design: A unified perspective."

Bergemann, Dirk and Juuso Välimäki (1996). "Learning and strategic pricing." In: *Econometrica: Journal of the Econometric Society*, pp. 1125–1149.

Billingsley, Patrick (2013). *Convergence of probability measures*. John Wiley & Sons.

Blackwell, David et al. (1951). "Comparison of experiments." In: *Proceedings of the second Berkeley symposium on mathematical statistics and probability*. Vol. 1, pp. 93–102.

Boel, René and Michael Kohlmann (1980). "Semimartingale models of stochastic optimal control, with applications to double martingales." In: *SIAM Journal on Control and Optimization* 18.5, pp. 511–533.

Bolton, Patrick and Christopher Harris (1999). "Strategic experimentation." In: *Econometrica* 67.2, pp. 349–374.

Callander, Steven (2011). "Searching and learning by trial and error." In: *American Economic Review* 101.6, pp. 2277–2308.

Caplin, Andrew and Mark Dean (Aug. 2013). *Behavioral Implications of Rational Inattention with Shannon Entropy*. Working Paper 19318. National Bureau of Economic Research. DOI: 10.3386/w19318. URL: http://www.nber.org/papers/w19318.

— (2015). "Revealed preference, rational inattention, and costly information acquisition." In: *The American Economic Review* 105.7, pp. 2183–2203.

Caplin, Andrew, Mark Dean, and John Leahy (Aug. 2017). *Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy*. Working Paper 23652. National Bureau of Economic Research. DOI: 10.3386/w23652. URL: http://www.nber.org/papers/w23652.

Carathéodory, Constantin (1907). "Über den Variabilitätsbereich der Koeffizienten von Potenzreihen, die gegebene Werte nicht annehmen." In: *Mathematische Annalen* 64.1, pp. 95–115.

Che, Yeon-Koo and Konrad Mierendorff (2016). "Optimal sequential decision with limited attention."

Cisek, Paul, Geneviève Aude Puskas, and Stephany El-Murr (2009). "Decisions in changing conditions: the urgency-gating model." In: *Journal of Neuroscience* 29.37, pp. 11560–11571.

Clark, Aubrey (2016). *Contracts for information acquisition*. unpublished manuscript.

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

Crandall, Michael G, Hitoshi Ishii, and Pierre-Louis Lions (1992). "User's guide to viscosity solutions of second order partial differential equations." In: *Bulletin of the American mathematical society* 27.1, pp. 1–67.

Davis, Mark HA (1979). "Martingale methods in stochastic control." In: *Stochastic Control Theory and Stochastic Differential Systems*. Springer, pp. 85–117.

Dellacherie, Claude and P-A Meyer (1979). *Probabilities and potential*. Vol. 29. Elsevier.

Denti, Tommaso (2015). *Unrestricted information acquisition*. Tech. rep.

— (2018). *Posterior-separable cost of information*. Tech. rep. working paper.

Doval, Laura (2018). "Whether or not to open Pandora's box." In: *Journal of Economic Theory* 175, pp. 127–158.

Ely, Jeffrey (2017). "Beeps." In: *American Economic Review* 107.1, pp. 31–53.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica (2015). "Suspense and surprise." In: *Journal of Political Economy* 123.1, pp. 215–260.

FDA, US (2018). *Guidance for industry: Adaptive Designs for Clinical Trials of Drugs and Biologics*. Tech. rep.

Fleming, Wendell H (1969). "Optimal continuous-parameter stochastic control." In: *SIAM review* 11.4, pp. 470–509.

Frankel, Alexander and Emir Kamenica (2018). "Quantifying information and uncertainty."

Fudenberg, Drew, Philipp Strack, and Tomasz Strzalecki (2018). "Speed, accuracy, and the optimal timing of choices." In: *American Economic Review* 108.12, pp. 3651–84.

Gentzkow, Matthew and Emir Kamenica (2014). "Costly persuasion." In: *American Economic Review* 104.5, pp. 457–62.

Gittins, John (1974). "A dynamic allocation index for the sequential design of experiments." In: *Progress in statistics*, pp. 241–266.

Hanson, Floyd B (2007). *Applied stochastic processes and control for Jump-diffusions: modeling, analysis, and computation*. Vol. 13. Siam.

Hébert, Benjamin and Michael Woodford (2016). "Rational Inattention with Sequential Information Sampling."

Hörner, Johannes and Andrzej Skrzypacz (2016). "Learning, experimentation and information design."

Kamenica, Emir and Matthew Gentzkow (2011). "Bayesian persuasion." In: *American Economic Review* 101.6, pp. 2590–2615.

Ke, T Tony and J Miguel Villas-Boas (2016). "Optimal Learning before Choice."

Keller, Godfrey and Sven Rady (2010). "Strategic experimentation with Poisson bandits." In: *Theoretical Economics* 5.2, pp. 275–311.

Keller, Godfrey, Sven Rady, and Martin Cripps (2005). "Strategic experimentation with exponential bandits." In: *Econometrica* 73.1, pp. 39–68.

Klabjan, Diego, Wojciech Olszewski, and Asher Wolinsky (2014). "Attributes." In: *Games and Economic Behavior* 88, pp. 190–206.

Liang, Annie, Xiaosheng Mu, and Vasilis Syrgkanis (2017). "Optimal Learning from Multiple Information Sources."

Lowther, George (2009). *Cadlag Modifications*. https://almostsure.wordpress.com/2009/12/18/cadlag-modifications/. Blog.

Luce, R Duncan et al. (1986). *Response times: Their role in inferring elementary mental organization*. 8. Oxford University Press on Demand.

Matějka, Filip (2015). "Rationally inattentive seller: Sales and discrete pricing." In: *The Review of Economic Studies* 83.3, pp. 1125–1155.

Matějka, Filip and Alisdair McKay (2012). "Simple market equilibria with rationally inattentive consumers." In: *American Economic Review* 102.3, pp. 24–29.

— (2014). "Rational inattention to discrete choices: A new foundation for the multinomial logit model." In: *The American Economic Review* 105.1, pp. 272–298.

Mathevet, Laurent, Jacopo Perego, and Ina Taneva (2017). "On information design in games."

Matyskova, Ludmila (2018). *Bayesian persuasion with costly information acquisition*. unpublished manuscript.

Mayskaya, Tatiana (2016). "Dynamic Choice of Information Sources."

Morris, Stephen and Philipp Strack (June 2017). *The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs*. Working Paper. URL: https://ssrn.com/abstract=2991567.

Moscarini, Giuseppe and Lones Smith (2001). "The optimal level of experimentation." In: *Econometrica* 69.6, pp. 1629–1644.

Rappoport, Daniel and Valentin Somma (2017). *Incentivizing Information Design*. unpublished manuscript.

Ratcliff, Roger (1981). "A theory of order relations in perceptual matching." In: *Psychological Review* 88.6, p. 552.

Ratcliff, Roger and Gail McKoon (2008). "The diffusion decision model: theory and data for two-choice decision tasks." In: *Neural computation* 20.4, pp. 873–922.

Ratcliff, Roger and Jeffrey N Rouder (1998). "Modeling response times for two-choice decisions." In: *Psychological Science* 9.5, pp. 347–356.

Ratcliff, Roger, Trisha Van Zandt, and Gail McKoon (1999). "Connectionist and diffusion models of reaction time." In: *Psychological review* 106.2, p. 261.

Ravid, Doron (2018). "Bargaining with rational inattention."

Rockafellar, RT (1969). *Convex analysis*. Princeton university press.

Rudin, Walter (1991). "Functional analysis 2nd ed." In: *International Series in Pure and Applied Mathematics. McGraw-Hill, Inc., New York*.

Shannon, Claude Elwood (1948). "A mathematical theory of communication." In: *Bell system technical journal* 27.3, pp. 379–423.

Sims, Christopher A (1998). "Stickiness." In: *Carnegie-Rochester Conference Series on Public Policy*. Vol. 49. Elsevier, pp. 317–356.

— (2003). "Implications of rational inattention." In: *Journal of monetary Economics* 50.3, pp. 665–690.

Steiner, Jakub, Colin Stewart, and Filip Matějka (2017). "Rational Inattention Dynamics: Inertia and Delay in Decision-Making." In: *Econometrica* 85.2, pp. 521–553.

Straszewicz, Stefan (1935). "Über exponierte punkte abgeschlossener punktmengen." In: *Fundamenta Mathematicae* 24, pp. 139–143.

Striebel, Charlotte (1984). "Martingale conditions for the optimal control of continuous time stochastic systems." In: *Stochastic processes and their applications* 18.2, pp. 329–347.

Tao, Terence (2016). *Analysis II*. Springer.

Wald, Abraham (1947). "Foundations of a general theory of sequential decision functions." In: *Econometrica, Journal of the Econometric Society*, pp. 279–313.

Weber, Richard et al. (1992). "On the Gittins index for multiarmed bandits." In: *The Annals of Applied Probability* 2.4, pp. 1024–1033.

Weitzman, Martin L (1979). "Optimal search for the best alternative." In: *Econometrica: Journal of the Econometric Society*, pp. 641–654.

Yang, Ming (2015a). "Coordination with flexible information acquisition." In: *Journal of Economic Theory* 158, pp. 721–738.

— (2015b). “Optimality of debt under flexible information acquisition.”

Zhong, Weijie (Sept. 2018). “Selling Information.” In: *arXiv e-prints*. arXiv: 1809.06770 [econ.TH].

Appendix A

Appendix for Chapter 1

A.1 Further discussions

In Appendix A.1, I first discuss the convergence of discrete-time optimal policy in Appendix A.1.1. It is shown that the discrete-time optimal policy's support as a correspondence of prior belief converges to that of the continuous-time optimal policy. Then I complete the discussion in Section 1.7 by generalizing each of the restrictive assumptions. Appendix A.1.2 generalizes the finite actions assumption and shows that the solution of a problem with infinite actions can be approximated by solutions to a series of problems with increasing number of actions. Appendix A.1.3 generalizes the binary states assumption in Assumption 1.3 and shows that the properties of optimal policy in Theorem 1.2 all extend in a problem with general finite state space. The proofs of theorems stated in this section are relegated to Appendix B.5.

A.1.1 Convergence of policy

By Theorems 1.2 and 1.3, the optimal policy solving Equation (1.4) is essentially unique in the jump-diffusion class. However, Theorem 1.1 does not rule out other possible optimal policies for the original stochastic control problem Equation (1.1). To get behavior predictions from my model, additional refinement of optimal policy of Equation (1.1) is necessary. In this discussion, I show that the discrete-time optimal policy of Equation (1.6) converges to the solutions defined in Theorems 1.2 and 1.3. I define a modified version of Lévy distance that characterizes the difference between two policy correspondences:

Definition A.1 (Lévy metric)**.** *Let F, G: $[0,1] \to 2^{[0,1]}$ be two correspondences. The Lévy metric $d_{\mathscr{L}}(F,G)$ is defined as:*

$$d_{\mathscr{L}}(F,G) := \inf\left\{\varepsilon > 0 \middle| \inf_{|y-x|\leqslant\varepsilon} d_H(F(x),G(y)) \leqslant \varepsilon, \forall x \in [0,1]\right\}$$

where d_H is the standard Hausdorff metric on $\mathbb{R}$.

$d_{\mathscr{L}}(F,G) = a$ means that $\forall \mu \in [0,1], \forall y \in F(\mu)$, there exists some μ' in a-neighbourhood of μ such that y is in the a-neighbourhood of $G(\mu')$. When G is continuous at μ, and a is sufficiently small, it simply states that the images of F and G at μ are close to each other (measured by d_H). If $d_{\mathscr{L}}(F,G) = 0$ then F and G are identical.

Theorem A.1 (Convergence of policy)**.** *Given either Assumptions 1.1, 1.2-a and 1.3 or Assumptions 1.1, 1.2-b and 1.3, let $\nu(\mu)$ be the policy correspondence solving Equation (1.4). Let $N(\mu) = \{\mu\} \bigcup \nu(\mu)$. Let $N_{dt}(\mu)$ be the support of optimal posteriors solving Equation (1.6). Then:*

$$\lim_{dt\to 0} d_{\mathscr{L}}(N, N_{dt}) = 0$$

Theorem A.1 states that the graph of policy function of discrete-time problem Equation (1.5′) converges to the graph of the continuous solution defined in Theorems 1.2 and 1.3. The convergence is illustrated in Figure A.1. I calculate the discrete-time policy function using parameters in Example 1.2. The red, blue and green lines represent the set of optimal posteriors as functions of prior when $V_{dt} > F$ with $dt = 10^{-5}, 10^{-3}$ and 10^{-2}. As is shown in the figure, when $dt \to 0$, one of optimal posterior is converging to the prior, and the other optimal posterior is converging to the continuous time solution. The posterior converging to prior captures a drift term and the other posterior captures a Poisson jump in the limit.

A.1.2 Infinite action space

In this section, I extend my model to accommodate infinite actions (or even continuum of actions) in the underlying decision problem, i.e. $|A| = \infty$. Mathematically, the difference is that the value from immediate action $F(\mu) = \sup_{a\in A} E[u(a,x)]$ is no-longer a piecewise linear function. There are several technical problems arising from a continuum of actions. For example whether the supremum is indeed achieved and whether F has

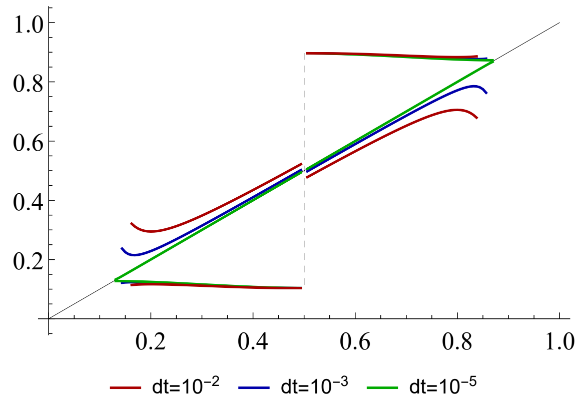


Figure A.1: Convergence of policy function

bounded subdifferentials. I impose the following assumption to rule out these technical issues:

Assumption A.1. $F(\mu) = \max_{a \in A} E[u(a, x)]$ *has bounded subdifferentials.*

Assumption A.1 rules out two cases. The first case is that the supremum is not achievable. The second case is that some optimal action being infinitely risky: the optimal action with belief approaching $x = 0$ has utility approaching $-\infty$ at state 1 (and similar case with states swapped). A sufficient condition for Assumption A.1 is:

Assumption A.1′. *A is a compact set.* $\forall x \in X$, $u(a, x) \in C(A) \bigcap \mathrm{TB}(A)$.

It is useful to notice that the proof of Theorem 1.1 does not rely on the fact that $F(\mu)$is piecewise linear. Actually the only necessary properties of $F(\mu)$ are boundedness and continuity in Lemma 1.2, which prove the existence of solution to discrete time functional equation Equation (B.1). Therefore Assumption A.1 guarantees that Lemma 1.2 and Lemma B.8 still hold when there is a continuum of actions. With Assumption A.1, the problem with continuum of actions can be approximated well by a sequence of problems with discrete actions. I first define the following notation: $\forall F$ satisfying Assumption A.1, $\mathcal{V}_{dt}(F)$ is the unique solution of Equation (1.6) and $\mathcal{V}(F) = \lim_{dt \to 0} \mathcal{V}_{dt}(F)$[1].

[1]The existence of limit is guaranteed by monotonic convergence theorem.

Lemma A.1. *Given Assumption A and Assumptions 1.2 and A.1, $\mathcal{V}$ is a Lipschitz continuous functional under L_∞ norm.*

Lemma A.1 implies that a problem with continuum of actions can be approximated well by a sequence of problems with discrete actions in the sense of value function convergence. Next, I push the convergence criteria further to the convergence of policy function.

Theorem A.2. *Given Assumptions 1.1, 1.2-a, 1.3 and A.1, let $\{F_n\}$ be a set of piecewise linear functions on [0,1] satisfying:*

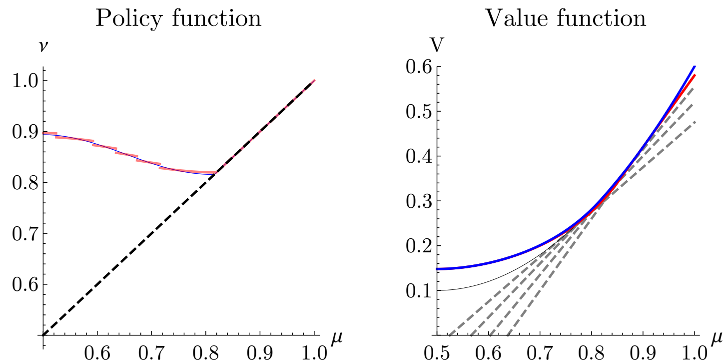
1. $\|F_n - F\|_\infty \to 0$;
2. $\forall \mu \in [0,1]$, $\lim F_n'(\mu) = F'(\mu)$.

Then $|\mathcal{V}(F) - \mathcal{V}(F_n)| \to 0$ and:

1. *$\mathcal{V}(F)$ solves Equation (1.4).*
2. *$\forall \mu$ s.t. $V(\mu) > F(\mu)$, if each ν_n is maximizer of $\mathcal{V}(F_n)$ and $\nu = \lim_{n\to\infty} \nu_n$ exists, then ν is the optimal posterior in Equation (1.4) at μ.*

Theorem A.2 states that to solve the problem with a continuum of actions, one can simply use both value function and policy function from problems with finite actions to approximate. As long as the immediate action values F_n converge uniformly in value and pointwise in first derivative, the optimal value functions have a uniform limit. The limit solves Equation (1.4) and the optimal policy function is the pointwise limit of policy functions from the finite action problems.

Figure A.2 illustrates this approximation process. On both panels, only $\mu \in [0.5, 1]$ is plotted (policy and value on $[0, 0.5]$ are symmetric). On the right panel, the thin black curve shows a smooth $F(\mu)$ associated with continuum of actions. Since optimal policy only utilizes a subset of actions, I approximate the smooth function only locally as the upper envelope of dashed lines (each represents one action). The optimal value function with continuous actions is the blue curve and the discrete action approximation is the red



Left panel shows the optimal policy function of discrete actions (red) and continuous actions (blue). The dashed line is $\nu = \mu$. Right panel shows the optimal value function. The thin black line is value from immediate action $F(\mu)$, the dashed lines are discrete approximations of the continuous function F.

Figure A.2: Approximation of a continuum of actions

curve. The left panel shows the approximation of policy function. The blue smooth curve is the optimal policy of the continuous action problem and the red curve with breaks is the optimal policy of the discrete action problem.

To approximate a smooth $F(\mu)$, one can simply add more and more actions to the finite action problem and use F's supporting hyper planes to approximate it. Then the optimal policy functions have more and more breaks as optimal policys involve more frequent jumps among actions. In the limit, as number of breaks grows to infinity, the size of breaks shrinks to zero and approaches a continuous policy function.

A.1.3 General state space

In this section, I extend the size of state space. The constructive proof for Theorems 1.2 and 1.3 relies on the ODE theory to guarantee existence of solution. With a larger state space, construction of value function relies on existence of PDE. There is no general theory ensuring existence of solution.[2] Nevertheless, the verification part still works. In fact, the

[2]The maximization problem can be translated into a PDE system. What is problematic is the boundary conditions. In fact, to solve for $V(\mu)$ searching over one action, I need to use the value function at regions where DM is indifferent between two actions as a boundary condition. That boundary condition is

discussion in Section 1.6.2 seems to extend to higher dimensional spaces in a natural way. I formalize a partial characterization theorem in the section.

Let $n = |X|$. Consider value function $V(\mu)$ on $\Delta(X)$. Let $V(\mu) \in C\Delta(X)$ and $C^{(2)}$ smooth when $V(\mu) > F(\mu)$. Consider the following HJB equation:

$$\rho V(\mu) = \max\left\{\rho F(\mu), \max_{\nu,p,\sigma} p(V(\nu)-V(\mu)-\nabla V(\mu)\cdot(\nu-\mu))+\sigma^T \mathrm{H}V(\mu)\sigma\right\} \tag{A.1}$$
$$s.t.\ -p(H(\nu)-H(\mu)-\nabla H(\mu)\cdot(\nu-\mu))-\sigma^T \mathrm{H}H(\mu)\sigma \leqslant c$$

where $\nu \in \Delta(\mathrm{supp}(\mu))$, $p \in \Delta I$ and $\sigma \in \mathbb{R}^{|\mathrm{supp}(\mu)|}$. Equation (A.1) comes from applying Assumption 1.2-a and smoothness condition to Equation (1.4). [3] I only discuss Assumption 1.2-a because the intuition is the same and similar proof methodology can be applied to Assumption 1.2-b to show an analog result.

Theorem A.3. *Let $E = \{\mu \in \Delta(X) | V(\mu) > F(\mu)\}$ be the experimentation region. Suppose there exists $C^{(2)}$ smooth $V(\mu)$ on E solving Equation (A.1), then $\exists$ policy function $\nu : E \mapsto \Delta(X)$ s.t.*

$$\rho V(\mu) = -c\frac{F(\nu(\mu)) - V(\mu) - \nabla V(\mu)(\nu(\mu) - \mu)}{H(\nu(\mu)) - H(\mu) - \nabla H(\mu)(\nu(\mu) - \mu)}$$

and ν satisfies the following properties:

1. *Poisson learning: $\rho V(\mu) \geqslant \sup_{\sigma} -c\frac{\sigma^T \mathrm{H}V(\mu)\sigma}{\sigma^T \mathrm{H}H(\mu)\sigma}$.*
2. *Direction: $D_{\nu(\mu)-\mu}V(\mu) \geqslant 0$.*
3. *Precision: $D_{\mu-\nu(\mu)}\nu(\mu) \cdot \mathrm{H}H(\nu)(\nu - \mu) \leqslant 0$.*
4. *Stopping time: $\nu(\mu) \in E^C$.*

There exists a nowhere dense set K s.t. strict inequality holds on $E \backslash K$ in property 1,3 and 4.

unknown, in contrast to the one dimensional analog $V(\mu^*)$ which can be easily calculated.

[3] $\mathrm{H}H(\mu)$ is defined on boundary where $V(\mu) = F(\mu)$ as continuous extension of interior Hessian's by Kirszbraun theorem.

Theorem A.3 states that if a solution $V(\mu)$ to Equation (A.1) exists, then $V(\mu)$ can be solved with only Poisson signals. The four properties are extensions to the four properties in Theorem 1.2 respectively. Property 1 and 4 are exactly the suboptimality of Gaussian signal and the immediate action property. Property 2 and 3 are weaker than the corresponding properties in Theorem 1.2. Property 2 is the extension to the confirmatory signal property. It states that optimal direction of jump is in the myopic direction that value function increases. Property 3 is the extension to the increasing precision property. $D_{\mu-\nu}\nu(\mu)$ is the direction ν is moving when μ is moving against ν. $\mathrm{H}H(\nu)(\nu-\mu)$ is the direction $(\nu-\mu)$ distorted by a negative definite matrix $\mathrm{H}H(\nu)$. In a special case when $H(\mu) = \|\mu-\mu_0\|_2^2$, $\mathrm{H}H(\nu)(\nu-\mu)$ is in the same direction as $(\mu-\nu)$, which implies (together with property 3) that the distance between μ and ν is increasing when μ is drifting against ν. In a generic case, this property does not directly predict how $\|\nu-\mu\|$ changes.

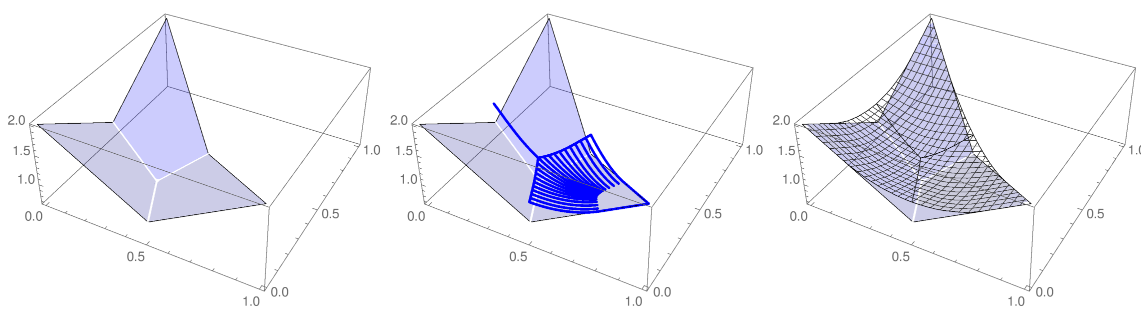
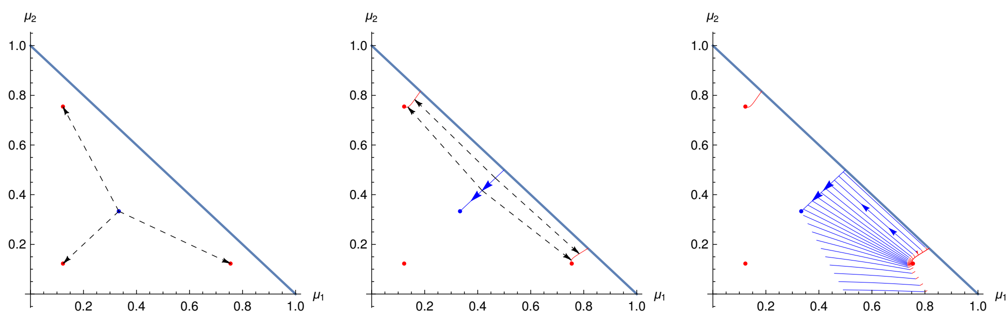


Figure A.3: Value function with 3 states

Figures A.3 and A.4 illustrate Theorem A.3 in a numerical example. There are three states and three actions. Belief space is a two-dimensional simplex. $F(\mu)$ is assumed to be a centrally symmetric function on belief space (Figure A.3-(a)). Value function $V(\mu)$ is the meshed manifold in Figure A.3-(c). Each blue curve in Figure A.3-(b) shows a drifting path of posterior beliefs. Take a prior in lower right region. The optimal policy is to search for one posterior (red points in lower right corner of Figure A.4-(c)), and posterior belief conditional on receiving no signal drifts along the curve in arrowed direction as in Figure A.4-(c). Once belief reaches the boundary, optimal policy becomes searching for two posteriors in a balanced way and posterior drifts towards center of belief space (see Figure A.4-(b), arrowed blue curve is belief trajectory and dashed arrows points to optimal posterior). Finally, if belief reaches center, optimal policy is to search for three posteriors in a balanced way (Figure A.4-(a)).



Dashed arrows start from priors and point to optimal posteriors. Blue arrows represents drift of posrtior beliefs conditional on no signal arrival. Left panel shows a point at which a balanced search over three posteriors is optimal. MIddle panel shows a curve along which searching over two posteriors is optimal. Right panel shows curves along witch searching over one unique posterior is optimal.

Figure A.4: Policy function with 3 states

A.1.4 Discrete-time information acquisition

In this section, I introduce a general discrete-time information acquisition problem. In the general problem, information is explicitly modeled as state-dependent signal process, and the cost of information is defined using a *posterior separable* function. I show that the discrete-time auxiliary problem Equation (1.5) introduced in Section 1.5.1 is a reduced form of the general problem. In Appendix A.1.4.1, I axiomatize posterior separability.

Decision problem: Time is discrete $t \in \mathbb{N}$. Period length is $dt > 0$. The other primitives (A, X, u, μ, ρ) are the same as in Section 1.3. The Bernoulli utility of action-state pair (a, x) in period t is $e^{-\rho dt \cdot t} u(a, x)$.

Strategy: a strategy is a triplet $(\mathcal{S}^t, \tau, \mathcal{A}^t)$. $\mathcal{S}^t$ is a random process correlated with the state, called an *information structure*. The realization of $\mathcal{S}^t$ is called a *signal history*.

The signal history up to period t is denoted by $\mathcal{S}^{\mathbf{t}}$. Each $\mathcal{S}^t$ specifies the signal structure acquired in period t conditional on all histories up to period t.[4] τ is a random variable whose realization is in $\mathbb{N}$. τ specifies a random decision time. The action choice $\mathcal{A}^t$ is a random process whose realization is in A. Each $\mathcal{A}^t$ specifies the joint distribution of action choice and state conditional on making decision in period t. Let the marginal distribution of the state be denoted by random variable $\mathcal{X}$.

Cost of information: Define $C_{dt}(I) = C\left(\frac{I}{dt}\right)dt$. The per-period cost of information is $C_{dt}(I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\tau\leqslant t}))$,[5] where the measure of signal informativeness I is defined as:

Assumption A. $I(\mathcal{S};\mathcal{X}|\mu) = E_s[H(\mu) - H(\nu(\cdot|s))]$, *where ν is the posterior belief about x according to Bayes rule.*

It is not difficult to see that $I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\tau\leqslant t})$ is exactly the finite difference formulation of $-\mathcal{L}_t H(\mu_t)\mathrm{d}t$. Assumption A is called *(uniform) posterior separability* in the literature. If H is the standard entropy function, then I is the mutual information between signal $\mathcal{S}^t$ and unknown state $\mathcal{X}$ (conditional on history).

Dynamic optimization: The dynamic optimization problem of the DM is:

$$V_{dt}(\mu) = \sup_{\mathcal{S}^t,\tau,\mathcal{A}^t} E\left[e^{-\rho dt\cdot\tau}u(\mathcal{A}^\tau,\mathcal{X}) - \sum_{t=0}^{\infty} e^{-\rho dt\cdot t}C_{dt}\left(I\left(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\tau\leqslant t}\right)\right)\right] \tag{1.5$'$}$$
$$\text{s.t.} \begin{cases} \mathcal{X} \to \mathcal{S}^{t-1} \to \mathbf{1}_{\tau\leqslant t} \\ \mathcal{X} \to \mathcal{S}^{t-1} \to \mathcal{A}^t \text{ conditional on } \tau = t \end{cases}$$

The two constraints in Equation (1.5′) are called the *information processing constraints*. Notation $\mathcal{X} \to \mathcal{S} \to \mathcal{T}$ means $\mathcal{X} \perp\!\!\!\perp \mathcal{T}|\mathcal{S}$. The first constraint states that signal history prior

[4] $\mathcal{S}^{-1}$ is defined as a degenerate random variable that induces belief same as prior belief μ for notation simplicity.

[5] $\mathbf{1}_{\tau\leqslant t}$ is an indicator whether learning is already stopped up to current period, which is known to the DM. So $(\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t})$ summarizes all knowledge of the DM.

to action time is sufficient for action time. The second constraint states that signal history prior to period t is sufficient for action at time t.[6] They are extensions to the standard measurability requirement, allowing randomness unrelated to unknown state to be added.

Equation (1.5′) is more general than Equation (1.5) in that it explicitly models the fully flexible choice of information. Take any strategy in Equation (1.5), if we consider belief as direct signal, then it resembles a special kind of strategy which is feasible in Equation (1.5′). These special strategies involve no *irrelevant randomness* and *unused information*, which are permitted in Equation (1.5′). In fact, Equation (1.5′) is more general than Equation (1.5) *only* in permitting irrelevant randomness and unused information. It is quite intuitive that allowing those more general strategies doesn't improve utility at all. In fact, it is proved in Lemmas B.4 and B.5 that V_{dt} defined by Equation (1.5′) is identical to that defined by Equation (1.5), for which reason I do not differentiate the notation.

Given the discussion above, Equation (1.5′) serves as a formal justification for using a belief based approached to model dynamic information acquisition. Moreover, it also relates Assumption 1.1 to posterior separable function — a measure for information widely used in rational inattention problems. In addition to existing attempts to axiomatize or microfound Assumption A, I provide a different axiomitization based on sequential information decomposition in Appendix A.1.4.1.

A.1.4.1 Axiom for Assumption A

Theorem A.4. *$I(\mathcal{S};\mathcal{X}|\mu)$ is a non-negative function of information structure and prior belief. I satisfies Assumption A if and only if the following axiom holds:*

Axiom: *$\forall\mu$, $\forall$ information structure $\mathcal{S}_1$ and information structure $\mathcal{S}_2|_{S_1}$ whose distribution de-*

[6] Notice that in every period, the information in current period has not been acquired yet. So decision can only be taken based on the information already acquired in the past. As a result in the information processing constraints information is advanced by one period. This within period timing issue does not make a difference when going to continuous-time limit.

pends on realization of $\mathcal{S}_1$:

$$I((\mathcal{S}_1, \mathcal{S}_2); \mathcal{X}|\mu) = I(\mathcal{S}_1; \mathcal{X}|\mu) + E[I(\mathcal{S}_2; \mathcal{X}|\mathcal{S}_1, \mu)]$$

Theorem A.4 states that the *chain rule* (the name for a key property of mutual information in Cover and Thomas (2012)) is not only a necessary condition but also a sufficient condition for posterior separability. Given any experiment, we can divide it into multiple stages of “smaller“ experiments. This axiom requires that the total informativeness of this sequence of small experiments is ”path-independent“: it always equals to the informativeness of the compound experiment.

A.2 Omitted proofs

A.2.1 Roadmap for proofs

Figure A.5: Roadmap for proofs

- Theorem B.1, P186
  - Corollary of Theorem 4.4, Chapter 4
- Theorem 1.1, P26
  - Lemma 1.2, P30
    - Lemma B.4, P202 — Lemma B.1, P185
    - Lemma B.5, P205
    - Lemma B.6, P207
    - Theorem B.1, P186
  - Lemma 1.1, P30
    - Lemma 1.2, P30
    - Lemma B.3, P194
    - Lemma B.8, P210
      - Lemma B.3, P194
      - Lemma 1.2, P30
  - Lemma 1.3, P32
    - Lemma B.9, P214
    - Lemma B.10, P218
    - Lemma B.7, P209
- Theorem 1.2, P33
  - Lemma A.2, P177
  - Lemma A.3, P180 — Lemma B.16, P226
    - Lemma B.15, P224
    - Lemma B.17, P231
  - Lemma A.3′, P233 — Lemma B.16′, P234
    - Lemma B.15, P224
    - Lemma B.17′, P234
  - Lemma B.18, P235
  - Lemma B.18′, P239
  - Lemmas B.11, B.12, B.13 and B.14, P221
- Theorem 1.3, P35
  - Lemma B.20, P250
  - Lemma B.21, P254 — Lemma B.22, P257
  - Lemma B.21′, P260 — Lemma B.22′, P261
  - Lemma B.23, P261
  - Lemma B.23′, P265
  - Lemmas B.11, B.12 and B.13, P221
  - Lemma B.19, P242 — Lemmas B.11 and B.12, P219
- Theorem 1.4, P52
- Theorem 1.5, P55
- Theorem 1.6, P58
- Theorem A.1, P140
  - Lemma B.25, P287
- Theorem A.2, P142
  - Lemma A.1, P142 — Lemma 1.3, P32
  - Lemma B.27, P297 — Lemma B.26, P296 — Lemma A.2, P177
- Theorem A.3, P144
- Theorem A.4, P148

Figure A.5 illustrates the roadmap for proofs in Chapter 1. Each node in the figure displays a theorem/lemma's name and its page number. Proof of each node depends (indirectly) on all nodes linked (indirectly) to it on the right. From top to bottom, the nodes are ordered by order of proofs: each node only depends on nodes on the right of it or above it. So it is clear that there is no circular argument. Dependent nodes that have been proved earlier are boxed by dashed lines. From left to right, the nodes are ordered by importance. Lemmas in the first layer are conceptually important and are directly supporting the proof for theorems. Lemmas in the second layer or above are more technical lemmas.

A.2.2 Proof of Theorem 1.1

The general road map for proving Theorem 1.1 is introduced in Section 1.5.3. The proof relies on three lemmas. Lemma 1.1 proves that the value function V_{dt} of discrete-time optimization problem Equation (1.5′) converges to the value function V of continuous-time optimization problem Equation (1.1) as $dt \to 0$. Lemma 1.3 proves that the solution of discrete time Bellman Equation (1.6) converges to the solution of continuous time HJB Equation (1.4) as $dt \to 0$. Lemma 1.2 proves that V_{dt} is also the solution of Bellman Equation (1.6). Therefore, V is the solution of HJB Equation (1.4).

Among the three lemmas, Lemmas 1.1 and 1.2 are quite standard, and the proofs are mostly variations of standard arguments. In Appendices A.2.2.1 and A.2.2.2, I discuss only the main proof ideas and some non-standard details and relegate the standard parts and purely technical details to Appendix B.2.1.

Lemma 1.3 is the key lemma for Theorem 1.1, as it provides an important link between discrete time Bellman and continuous time HJB. Proof of Lemma 1.3 is provided in details in Appendix A.2.2.3. The discussion also formalizes the definition of HJB Equation (1.4) by clarifying the notion of viscosity solution I am using.

A.2.2.1 Proof of Lemma 1.1

Remark A.1. The proof of Lemma 1.1 uses Lemma 1.2 for some minor technical arguments. However the main proof idea does not conceptually depend on Lemma 1.2. So I show the proof of Lemma 1.1 first.

Proof. As already stated in Section 1.5.1, it is sufficient to show that the order of limits can be switched:

$$\sup_{\langle\mu_t\rangle,\tau}\lim_{dt\to 0} W_{dt}(\mu_t,\tau) = \lim_{dt\to 0}\sup_{\langle\mu_t\rangle,\tau} W_{dt}(\mu_t,\tau) \tag{A.2}$$

Here $W_{dt}(\mu_t,\tau)$ is defined in Section 1.5.1 as the *discretized* payoff of continuous time strategy $\langle\mu_t\rangle,\tau$. The inner limit of LHS in Equation (A.2) is then by definition the payoff of strategy $\langle\mu_t\rangle,\tau$ in the continuous time problem Equation (1.1). So the LHS is $V(\mu)$. The inner limit of RHS is $V_{dt}(\mu)$ (as the problem optimizing W_{dt} is a discrete time problem equivalent to Equation (1.5′), formally shown in Lemma B.5, a dependence lemma for Lemma 1.2). So RHS is $\lim V_{dt}$ (a technical lemma Lemma B.8 guarantees existence of such limit).

I prove by showing inequality in two directions. The direction $V(\mu) \leqslant \lim V_{dt}(\mu)$ is trivial since $W_{dt}(\mu_t,\tau) \leqslant V_{dt}(\mu)$ for all $\langle\mu_t\rangle,\tau,dt$. The key is to prove the other direction $V(\mu) \geqslant \lim V_{dt}(\mu)$. I prove this claim by showing that $\forall dt>0$, there exists a continuous time strategy that achieves a payoff in Equation (1.1) no less than $V_{dt}(\mu)$.

Given time period dt, by Lemma 1.2 there exists discrete time optimal solution μ_t^* and τ^*, where $\mu_{t+1}^*|\mathcal{F}_t$ has support size N. The goal is to construct an admissible continuous-time belief process $\langle\mu_t\rangle$, which satisfies two properties: 1) at each discrete time idt, μ_t has exactly the same distribution as μ_i^*, 2) within each dt period, uncertainty reduction speed of μ_t is exactly $E[H(\mu_i^*)-H(\mu_{i+1}^*)|\mathcal{F}_i]/dt$. Such $\langle\mu_t\rangle$ with stopping time τ^*

achieves higher payoff than $V_{dt}(\mu)$. Now this construction can be done by a technique introduced in Lemma B.3. $\forall i$ and conditional on $\mathcal{F}_i$, apply Lemma B.3 to the distribution of μ^*_{i+1} to smooth it on $[idt,(i+1)dt]$. Lemma B.3 states that there exists a continuous-time martingale $\langle\widetilde{\mu}_t\rangle$ (with a corresponding probability space) satisfying: $\forall s,t\in[0,1]$, $s>t$: $E[H(\mu_t)-H(\mu_s)|\mathcal{F}_t]=(s-t)E[H(\mu^*_i)-H(\mu^*_{i+1})|\mathcal{F}_i]$. For $t\in[idt,(i+1)dt]$, define $\mu_t|\mathcal{F}_{idt}=\widetilde{\mu}_{\frac{t-idt}{dt}}|\mathcal{F}_i$. Therefore, $\forall t\in[idt,(i+1)dt)$:

$$\begin{aligned}
-\mathcal{L}_tH(\mu_t)=&\lim_{s\to t^+}E\left[\frac{H(\mu_t)-H(\mu_s)}{s-t}\middle|\mathcal{F}_t\right]\\
=&\lim_{s\to t^+}\frac{(s-t)E\big[H(\mu^*_i)-H(\mu^*_{i+1})\big|\mathcal{F}_i\big]}{s-t}\\
=&H(\mu^*_i)-\sum p_i^jH(\mu^{*j}_{i+1})
\end{aligned}$$

Let $\tau=\tau^*dt$. It is easy to see that by construction τ is measurable to the natural filtration of μ_t. Therefore:

$$\begin{aligned}
V(\mu)\geqslant&E\left[e^{-\rho\tau}F(\mu_\tau)-\int_0^\tau e^{-\rho t}C(I_t)\mathrm{d}t\right]\\
=&E\left[e^{-\rho dt\cdot\tau^*}F(\mu_{\tau^*})-\sum_{t=0}^{\tau^*-1}C\left(\frac{H(\mu^*_i)-\sum p_i^jH(\mu^{*j}_{i+1})}{dt}\right)e^{-\rho dt\cdot t}\cdot\frac{1-e^{-\rho dt}}{\rho}\right]\\
\geqslant&E\left[e^{-\rho dt\cdot\tau^*}F(\mu_{\tau^*})-\sum_{t=0}^{\tau^*-1}C\left(\frac{H(\mu^*_i)-\sum p_i^jH(\mu^{*j}_{i+1})}{dt}\right)e^{-\rho dt\cdot t}\cdot dt\right]\\
=&E\left[e^{-\rho dt\cdot\tau^*}F(\mu_{\tau^*})-\sum_{t=0}^{\tau^*-1}C_{dt}\Big(H(\mu^*_i)-\sum p_i^jH(\mu^{*j}_{i+1})\Big)e^{-\rho dt\cdot t}\right]=V_{dt}(\mu)
\end{aligned}$$

Second inequality is from $1-e^{-x}\leqslant x$. Therefore, $V(\mu)\geqslant\lim V_{dt}(\mu)$. ■

Remark A.2 (Non-integrable $\langle\mu_t\rangle$). In fact, the integrability requirement introduced in Equation (1.1) (defined as existence of $\lim W_{dt}$ in Section 1.5.1) is not necessary for my analysis of Theorem 1.1. Suppose now I extend the set of admissible belief profiles $\mathbb{M}$ to satisfy only the first two conditions: cadlag path, martingale property and initial value $\mu_0 = \mu$. Then the limit of finite Riemann sum $W_{dt}(\mu_t, \tau)$ might not exist (although each finite Riemann sum is always well defined). Whenever this is the case, I define the payoff of strategy $\langle \mu_t \rangle, \tau$ as:

$$E\left[e^{-\rho\tau}F(\mu_\tau) - \int_0^\tau e^{-\rho t}C(-\mathcal{L}_t H(\mu_t))\mathrm{d}t\right] \triangleq \limsup_{dt\to 0} W_{dt}(\mu_t, \tau) \tag{A.3}$$

Since $W_{dt}(\mu_t, \tau)$ is bounded above by $\max F$, Equation (A.3) is always well defined. Equation (A.3) is the essential upper-bound of payoff of an ill-behaved strategy, and when $\langle \mu_t \rangle$ is integrable it is consistent with the original definition of V. Obviously, such extension of admissible strategy set weakly increases the value of $V(\mu)$. Here I call the extended value function $\widehat{V}(\mu) = \sup_{\langle \mu_t \rangle, \tau} \limsup_{dt\to 0} W_{dt}(\mu_t, \tau)$.

In the proof of Theorem 1.1, Lemmas 1.2 and 1.3 are not affected at all since they are about the discrete-time problem and corresponding value function V_{dt}. If Lemma 1.1 can be extended to $\widehat{V}(\mu) = \lim_{dt\to 0} V_{dt}$, then Theorem 1.1 still holds with V replaced with $\widehat{V}$. This extension is quite trivial by observing $\forall \langle \mu_t \rangle, \tau, dt,\ W_{dt}(\mu_t, \tau) \leqslant V_{dt}(\mu) \implies \limsup W_{dt}(\mu_t, \tau) \leqslant \lim V_{dt}(\mu) \implies \widehat{V}(\mu) = \limsup \leqslant \lim V_{dt}(\mu)$.

To sum up, if we extend the admissible strategy set, and relax the definition of the objective function to its essential upper-bound, a solution to HJB Equation (1.4) still achieves the value function. Therefore, it is WLOO to eliminate all those ill-behaved strategies from the admissible control set.

A.2.2.2 *Proof of Lemma 1.2*

Remark A.3. The proof presented here is stronger than the statement of Lemma 1.2 in Section 1.5.2. It proves that the Bellman Equation (1.6) characterizes both Equations (1.5) and (1.5′) (while Lemma 1.2 only states that Equation (1.5) is characterized by Equation (1.6)). The first step of the proof shows that V_{dt} defined by Equations (1.5) and (1.5′)

are identical (Lemmas B.4 and B.5), and can be rewritten as a recursive problem (Lemma B.6). To proof the Lemma 1.2 exactly stated in Section 1.5.2, one can simply skip Lemmas B.4 and B.5 and start with Lemma B.6, noticing that Equation (B.9) is simply rewriting Equation (1.5).

Proof. The proof of Lemma 1.2 is mostly the standard theory of discrete-time dynamic programming with a few tweaks. The proof involves 4 steps:

Step 1. Rewrite the sequential problem into the recursive problem. The technical details of the rewriting of problem is shown in Lemmas B.4, B.5 and B.6. The only non-standard analysis is to show that in Equation (1.5′), $\mathcal{S}_t$ may contain unused information/ randomness which can be discarded without loss of utility. Then the sequential problem without any redundant information can be represented in the belief space and easily written as a recursive problem.

Step 2. Verify the standard transversality condition. This is trivial as the payoff is bounded by $\max F$ and discounted exponentially.

Step 3. Verify the Blackwell contract mapping condition. The contraction parameter in Equation (1.6) is trivially the discount factor $e^{-\rho dt}$. The non-standard analysis is to show that the optimization operation is into the domain $C(\Delta X)$. To show this I invoke a maximum theorem in information design problems (Theorem 4.5 of Chapter 4, it shows the existence of maximum as well).

Step 4. With steps 1-3, I invoke the standard contract-mapping fixed point theorem and show that value function V_{dt} is the unique solution to Equation (1.6). The final bits show that I can restrict the optimal strategy of Equation (1.6) to have support size N. This part is proved using a generalized concavification result: Notice that the objective function in Equation (1.6) is not in the standard "expected valuation" form as in the literature of information design (see Kamenica and Gentzkow (2011)). Instead, there is an extra $C_{dt}(\cdot)$

term. However, intuitively this problem can be handle using a Lagrange method and take the term inside $C_{dt}(\cdot)$ to combine it with $E[V]$ linearly. This intuition is formalized by Theorem B.1, which is a corollary of a more general result in Chapter 4. ■

A.2.2.3 Proof of Lemma 1.3

Before going to the proof of Lemma 1.3, I first formally rewrite the problem to accommodate viscosity solutions (see Crandall, Ishii, and Lions (1992)). First define a space of functions on $\Delta(X)$:

$$\mathcal{L} = \left\{ V : \Delta(X) \mapsto \mathbb{R}^+ \middle| \forall \mu \in \Delta X, \mu' \in \Delta(\mathrm{supp}(\mu)), \lim\sup_{\mu' \to \mu} \frac{|V(\mu') - V(\mu)|}{\|\mu' - \mu\|} \in \mathbb{R} \right\}$$

where $\|\cdot\|$ is Euclidean norm on ΔX. By definition, $\mathcal{L}$ is the set of pointwise Lipschitz functions on $\Delta(X)$. Two technical lemmas Lemmas B.8 and B.9 guarantee that $\lim V_{dt}$ is well defined, and there exists $\overline{V} \in \mathcal{L}$ which is the uniform limit of V_{dt}. Now I show that $\overline{V}$ coincides with the solution of the HJB equation. Consider the following HJB equation defined on $\mathcal{L}$:

$$\begin{aligned}\rho V(\mu) = \max \Bigg\{ \rho F(\mu), \sup_{\substack{\nu_i \in \Delta(\mathrm{supp}(\mu)), \\ p_i \in \mathbb{R}^+, \\ \widehat{\sigma} \in \mathbb{R}^{|\mathrm{supp}(\mu)|}}} & \sum p_i (V(\nu_i) - V(\mu)) - DV\Big(\mu, \sum p_i \nu_i - \mu\Big) \Big\| \Big(\sum p_i \nu_i - \mu\Big) \Big\| \\ & + \frac{1}{2} \|\widehat{\sigma}\|^2 D^2 V(\mu, \widehat{\sigma}) \qquad \text{(A.4)} \\ & - C\Big(- \sum p_i (H(\nu_i) - H(\mu) - \nabla H(\mu) \cdot (\nu_i - \mu)) - \frac{1}{2} \widehat{\sigma}^T \cdot \mathrm{H}H(\mu) \cdot \widehat{\sigma} \Big) \Bigg\}\end{aligned}$$

∇ and H denote gradient and Hessian operator (well-defined on all interior points). Since $\overline{V}$ is not necessarily differentiable, I use operator D and D^2 to replace the Jacobian and Hessian operators on $\overline{V}$. D and D^2 are defined as follows. $\forall y \in B^{|\mathrm{supp}(x)|-1}$ (Unit ball in $|\mathrm{supp}(x)| - 1$ dimensional space):

Definition A.2 (General differentials)**.** *$\forall f \in \mathcal{L}$:*

$$\begin{cases} Df(x,y) = \liminf_{\delta\to 0} \frac{f(x)-f(x-\delta y)}{\delta\|y\|} \\ D^2 f(x,y) = \limsup_{\delta\to 0} 2\frac{f(x+\delta y)-f(x)-\delta\cdot Df(x,y)\cdot\|y\|}{\delta\|y\|^2} \end{cases}$$

Notice that if $f \in C^{(1)}(\Delta X)$, then $Df(x,y) = \frac{\nabla f(x)\cdot y}{\|y\|}$. If $f \in C^{(2)}(\Delta X)$ then $D^2 f(x,y) = \frac{y^T\cdot \mathrm{H}f(x)\cdot y}{\|y\|^2}$. It is not hard to verify that for $C^{(1)}$ smooth value function $V(\mu)$, Equation (A.4) is equivalent to Equation (1.4).

Proof.

Consider Lemma 1.3 by replacing Equation (1.4) with Equation (A.4). If the statement is proved with Equation (A.4), then since $\overline{V} = V$ is $C^{(1)}$ smooth, $\overline{V}$ is smooth and Equation (1.4) automatically holds. I prove by induction on dimensionality of $\mathrm{supp}(\mu)$. First of all, Lemma 1.3 is trivially true when $\mu = \delta_x$ since $V(\mu) = \overline{V}(\mu) = F(\mu)$ when the state is deterministic. Now it is sufficient to prove $\overline{V} = V$ on interior of ΔX conditional on $\overline{V} = V$ being true on $\partial\Delta X$ (boundary of ΔX).

The proof takes three steps. Before going to the details, I introduce the steps briefly. The first step is to show that $\overline{V}$ is unimprovable in HJB Equation (A.4). The proof is quite standard as any continuous-time strategy that improves $\overline{V}$ can be approximated by a discrete-time strategy. The second step shows $\overline{V} \geqslant V$. Proof is by a standard contradiction argument. If $\overline{V} < V$, then there exists a belief s.t. the same strategy implements strictly higher HJB with $\overline{V}$, which violates unimprovability. The last and most difficult step is to show that $V \geqslant \overline{V}$.

Unimprovability: First I show that $\overline{V}$ is unimprovable in Equation (A.4). Suppose for the sake of contradiction that $\overline{V}$ is improvable at interior μ, then there exists $p_i, \nu_i, \widehat{\sigma}, I$

such that:

$$\rho\overline{V}(\mu) < \sum p_i\big(\overline{V}(\nu_i) - \overline{V}(\mu)\big) - D\overline{V}(\mu, \mu - \sum p_i\nu_i)\Big\|\sum p_i\nu_i - \mu\Big\| + \sum D^2\overline{V}(\mu, \widehat{\sigma}_j)\|\widehat{\sigma}\|_j^2 - C(I)$$

$$\text{where } I = -\sum p_i(H(\nu_i) - H(\mu) - \nabla H(\mu)\cdot(\nu_i - \mu)) - \sum \widehat{\sigma}_j^T \mathrm{H}H(\mu)\widehat{\sigma}_j$$

Then if we compare the following two ratios:

$$\frac{\sum p_i\big(\overline{V}(\nu_i) - \overline{V}(\mu)\big) - D\overline{V}(\mu, \sum p_i\nu_i - \mu)\|\sum p_i\nu_i - \mu\|}{-\sum p_i(H(\nu_i) - H(\mu) - \nabla H(\mu)\cdot(\nu_i - \mu))}; \ \frac{D^2\overline{V}(\mu,\widehat{\sigma})\|\widehat{\sigma}\|^2}{-\widehat{\sigma}^T\mathrm{H}H(\mu)\widehat{\sigma}}$$

At least one of them must be larger than $\frac{\rho\overline{V}(\mu)+C(I)}{I}$.

- *Case 1:*

$$\frac{\sum p_i\big(\overline{V}(\nu_i) - \overline{V}(\mu)\big) - D\overline{V}(\mu, \sum p_i\nu_i - \mu)\|\sum p_i\nu_i - \mu\|}{-\sum p_i(H(\nu_i) - H(\mu) - \nabla H(\mu)\cdot(\nu_i - \mu))} > \frac{\rho}{I}\overline{V}(\mu) + \frac{C(I)}{I}$$

By Definition A.2, there exists $\delta, \varepsilon > 0$ s.t. :

$$\frac{\sum p_i\big(\overline{V}(\nu_i) - \overline{V}(\mu)\big) - \frac{\overline{V}(\mu) - \overline{V}(\mu - \delta(\sum p_i\nu_i - \mu))}{\delta}}{\sum p_i(H(\mu) - H(\nu_i)) + \frac{H(\mu) - H(\mu - \delta(\sum p_i\nu_i - \mu))}{\delta}} \geqslant \frac{\rho}{I}\overline{V}(\mu) + \frac{C(I)}{I} + \varepsilon \tag{A.5}$$

where δ is sufficiently small that $\mu_0 = \mu - \delta(\sum p_i\nu_i - \mu) \in \Delta X^{\mathrm{o}}$. Then by construction, if we assume:

$$\begin{cases} p_0' = \frac{1}{1+\delta} \\ p_i' = \frac{\delta}{1+\delta}p_i \end{cases}$$

Then (p_i', ν_i') is Bayesian plausible:

$$\begin{cases} \sum p_i' = 1 \\ \sum p_i' \nu_i = \mu \end{cases}$$

where 0 is also included in indices i's. Replacing terms in Equation (A.5) and let $I(\nu_i|\mu) = H(\mu) - \sum p_i' H(\nu_i)$:

$$\begin{aligned} &\frac{\sum p_i' \overline{V}(\nu_i) - \overline{V}(\mu)}{-\sum p_i' H(\nu_i) + H(\mu)} \geqslant \frac{\rho}{I}\overline{V}(\mu) + \frac{C(I)}{I} + \varepsilon \\ \Longrightarrow &\sum p_i' \overline{V}(\nu_i) - \frac{I(\nu_i|\mu)}{I} C(I) \geqslant \left(1 + \rho \frac{I(\nu_i|\mu)}{I}\right) \overline{V}(\mu) + \varepsilon I(\nu_i|\mu) \end{aligned} \tag{A.6}$$

It is easy to verify that $I(\nu_i|\mu)$ is continuous in δ and it is zero when $\delta = 0$. So δ can be chosen sufficiently small that

$$e^{\rho \frac{I(\nu_i|\mu)}{I}} - \left(1 + \rho \frac{I(\nu_i|\mu)}{I}\right) = \sum_{k=1}^{\infty} \frac{1}{(k+1)!} \left(\frac{\rho}{I}\right)^{k+1} I(\nu_i|\mu)^k \cdot I(\nu_i|\mu) \leqslant \frac{\varepsilon I(\nu_i|\mu)}{4 \sup F} \tag{A.7}$$

The equality is from Taylor expansion of exponential function. Plug Equation (A.7) into Equation (A.6):

$$\begin{aligned} &\sum p_i' \overline{V}(\nu_i) - \frac{I(\nu_i|\mu)}{I} C(I) \geqslant e^{\rho \frac{I(\nu_i|\mu)}{I}} \overline{V}(\mu) + \frac{\varepsilon}{4} I(\nu_i|\mu) \\ \Longrightarrow\ & e^{-\rho \frac{I(\nu_i|\mu)}{I}} \left(\sum p_i' \overline{V}(\nu_i)\right) - \frac{I(\nu_i|\mu)}{I} C(I) \\ &\geqslant \overline{V}(\mu) + e^{-\rho \frac{I(\nu_i|\mu)}{I}} \frac{\varepsilon I(\nu_i|\mu)}{4} - \left(1 - e^{-\rho \frac{I(\nu_i|\mu)}{I}}\right) \frac{I(\nu_i|\mu)}{I} C(I) \end{aligned} \tag{A.8}$$

Noticing that $\left(1 - e^{-\rho \frac{I(\nu_i|\mu)}{I}}\right) I(\nu_i|\mu)$ is a second order small term. Then we can pick δ

such that Equation (A.8) implies:

$$e^{-\rho\frac{I(\nu_i|\mu)}{I}}\left(\sum p_i'\overline{V}(\nu_i)\right)-\frac{I(\nu_i|\mu)}{I}C(I)\geqslant\overline{V}(\mu)+\frac{\varepsilon}{8}I(\nu_i|\mu)$$

From now on, we fix ε and δ. Pick $dt=\frac{I(\nu_i|\mu)}{I}$, $dt_m=\frac{dt}{m}$. By uniform convergence, there exists N s.t. $\forall m\geqslant N$:

$$\begin{aligned}&e^{-\rho dt}\left(\sum p_i'V_{dt_m}(\nu_i)\right)-dt\cdot C\left(\frac{I(\nu_i|\mu)/m}{dt_m}\right)>V_{dt_m}(\mu)\\ \Longrightarrow\ &e^{-\rho m dt_m}\left(\sum p_i'V_{dt_m}(\nu_i)\right)-\sum_{\tau=0}^{m-1}e^{-\rho\tau dt_m}C_{dt_m}\left(\frac{I(\nu_i|\mu)}{m}\right)>V_{dt_m}(\mu)\end{aligned}$$

That is to say we find a feasible experiment, whose cost can be spread into m periods (the split of experiment is done by applying Lemma B.3). This experiment strictly dominates the optimal experiment at μ for discrete time problem with dt_m. Contradiction. Therefore, $\overline{V}$ must be unimprovable at μ.

- *Case 2*:

$$\frac{D^2\overline{V}(\mu,\widehat{\sigma})\|\widehat{\sigma}\|^2}{-\widehat{\sigma}^T\mathrm{H}H(\mu)\widehat{\sigma}}>\frac{\rho}{I}\overline{V}(\mu)+\frac{C(I)}{I}$$

Then by the definition of operator D^2 in Definition A.2, there exists $\widehat{\sigma}$, $\delta,\varepsilon>0$ s.t.:

$$\frac{\overline{V}(\mu+\delta\widehat{\sigma})-\overline{V}(\mu)-\delta D\overline{V}(\mu,\widehat{\sigma})\|\widehat{\sigma}\|}{-H(\mu+\delta\widehat{\sigma})+H(\mu)+\delta\nabla H(\mu)\cdot\widehat{\sigma}}\geqslant\frac{\rho}{I}\overline{V}(\mu)+\frac{C(I)}{I}+2\varepsilon$$

Then by the definition of operator D in Definition A.2, there exists δ' s.t.:

$$\frac{\overline{V}(\mu+\delta\widehat{\sigma})-\overline{V}(\mu)-\delta\frac{\overline{V}(\mu)-\overline{V}(\mu-\delta'\widehat{\sigma})}{\delta'}}{-H(\mu+\delta\widehat{\sigma})+H(\mu)+\delta\frac{H(\mu)-H(\mu-\delta'\widehat{\sigma})}{\delta'}}\geqslant\frac{\rho}{I}\overline{V}(\mu)+\frac{C(I)}{I}+\varepsilon$$

Let $\mu_1 = \mu - \delta'\widehat{\delta}$ and $\mu_2 = \mu + \delta\widehat{\sigma}$, $p_1 = \frac{\delta'}{\delta+\delta'}$, $p_2 = \frac{\delta}{\delta+\delta'}$, then:

$$\sum p_i \overline{V}(\nu_i) \geqslant \left(1 + \rho\frac{I(\nu_i|\mu)}{I}\right)\overline{V}(\mu) + \frac{I(\nu_i|\mu)}{I}C(I) + \varepsilon I(\nu_i|\mu) \tag{A.9}$$

Noticing that Equation (A.9) is exactly the same as Equation (A.6) in Case 1. Then using same argument, This case is also ruled out.

Equality: I show that $\forall$ smooth function V solving Equation (A.4), $\overline{V} = V$. Notice that this automatically proves the uniqueness of solution of Equation (A.4). I prove inequality from both directions for $\mu \in \Delta(X)^{o}$:

- $\overline{V}(\mu) \geqslant V(\mu)$: Suppose not, then consider $U(\mu) = \overline{V}(\mu) - V(\mu)$. Since both V and $\overline{V}$ are continuous, U is continuous. Therefore $\arg\min U$ is non empty and $\min U < 0$ according to our assumption. Choose $\mu \in \arg\min U$ ($\mu \in \Delta X^o$ since $V = \overline{V}$ on boundary). Since $\overline{V}(\mu) \geqslant F(\mu)$, $V(\mu) > F(\mu)$. Let $(p_i, \nu_i, \hat{\sigma})$ be a strategy solving $V(\mu)$:

$$\begin{aligned}\rho V(\mu) = &\sum p_i(V(\nu_i) - V(\mu)) - DV\left(\mu, \sum p_i\nu_i - \mu\right)\left\|\sum p_i(\nu_i - \mu)\right\| \\ &+ \frac{1}{2}D^2V(\mu,\widehat{\sigma})\|\widehat{\sigma}\|^2 \\ &- C\left(-\sum p_i(H(\nu_i) - H(\mu) - \nabla H(\mu)(\nu_i - \mu)) - \frac{1}{2}\widehat{\sigma}^T \mathrm{H}H(\mu)\widehat{\sigma}\right)\end{aligned} \tag{A.10}$$

Now compare $D\overline{V}$ and DV:

$$\begin{aligned}&\frac{\overline{V}(\mu) - \overline{V}(\mu')}{\|\mu - \mu'\|} = \frac{V(\mu) - V(\mu') + U(\mu) - U(\mu')}{\|\mu - \mu'\|} \leqslant \frac{V(\mu) - V(\mu')}{\|\mu - \mu'\|} \\ \Longrightarrow\ &\liminf \frac{\overline{V}(\mu) - \overline{V}(\mu')}{\|\mu - \mu'\|} \leqslant \lim \frac{V(\mu) - V(\mu')}{\|\mu - \mu'\|} \\ \Longrightarrow\ &D\overline{V}(\mu, \mu' - \mu)\|\mu' - \mu\| \leqslant \nabla V(\mu) \cdot (\mu' - \mu)\end{aligned}$$

Compare $D^2\overline{V}$ and D^2V:

$$\frac{\overline{V}(\mu')-\overline{V}(\mu)-D\overline{V}(\mu,\mu'-\mu)\|\mu'-\mu\|}{\|\mu'-\mu\|^2}\geqslant\frac{V(\mu')-V(\mu)-\nabla V(\mu)\cdot(\mu'-\mu)+U(\mu')-U(\mu)}{\|\mu-\mu'\|^2}$$
$$\implies D^2\overline{V}(\mu,\widehat{\sigma})\geqslant D^2V(\mu,\widehat{\sigma})$$

Therefore Equation (A.10) implies:

$$\begin{aligned}\rho V(\mu)\leqslant&\sum p_i\big(\overline{V}(\nu_i)-\overline{V}(\mu)-(U(\nu_i)-U(\mu))\big)\\&-D\overline{V}(\mu,\sum\nu_i-\mu)\Big\|\sum\nu_i-\mu\Big\|+\frac{1}{2}D^2\overline{V}(\mu,\widehat{\sigma})\|\widehat{\sigma}\|^2\\&-C\Big(-\sum p_i(H(\nu_i)-H(\mu)+\nabla H(\mu)(\nu_i-\mu))-\frac{1}{2}\widehat{\sigma}^T\mathrm{H}H(\mu)\widehat{\sigma}\Big)\\\leqslant&\rho\overline{V}(\mu)\end{aligned}$$

The first inequality comes from replacing DV and D^2V with $D\overline{V}$ and $D^2\overline{V}$. The second inequality comes from $U(\nu_i)-U(\mu)\geqslant 0$ and unimprovability of $\overline{V}$. Contradiction.

- $V(\mu)\geqslant\overline{V}(\mu)$: I prove by showing that $\forall dt>0$, $V\geqslant V_{dt}$. Suppose not, then there exists μ',dt s.t. $V_{dt}(\mu')>V(\mu')$. Let $dt_n=\frac{dt}{2^n}$. Since V_{dt_n} is increasing in n, there exists $\varepsilon>0$ s.t. $V_{dt_n}(\mu')-V(\mu')\geqslant\varepsilon\ \forall n\in\mathbb{N}$. Now consider $U_n=V-V_{dt_n}$. U_n is continuous by Lemma 1.2 and $U_n(\mu')\leqslant-\varepsilon$. Pick $\mu^n\in\arg\min U_n$. Since $\Delta(X)$ is compact, there exists a converging sequence $\lim\mu^n=\mu$. By assumption, $U_n(\mu^n)\leqslant-\varepsilon$, therefore since $U(\mu)=\lim U_n(\mu^n)\leqslant-\varepsilon$, μ must be in interior of $\Delta(X)$. So without loss, μ^n can be picked that $\mu^n\in\Delta(X)^O$. Now consider the optimal strategy of discrete time problem:

$$\begin{cases}V_{dt_n}(\mu^n)=e^{-\rho dt_n}\sum p_i^nV_{dt_n}(\nu_i^n)-dt_nC(I_n)\\\sum p_i^n(H(\mu^n)-H(\nu_i^n))=I_ndt_n\\\sum p_i^n\nu_i^n=\mu^n;\ \sum p_i^n=1\end{cases}$$

By definition of $U_n(\mu)$:

$$\begin{aligned}
\sum p_i^n(V(\nu_i^n) - V(\mu^n)) &= \sum p_i^n \left(V_{dt_n}(\nu_i^n) - V_{dt_n}(\mu^n) - U_n(\mu^n) + U(\nu_i^n)\right) \\
&\geqslant \sum p_i^n \left(V_{dt_n}(\nu_i^n) - V_{dt_n}(\mu^n)\right) \\
&= \left(e^{\rho dt_n} - 1\right) V_{dt_n}(\mu^n) + e^{\rho dt_n} dt_n C(I_n) \\
&\geqslant \rho dt_n V_{dt_n}(\mu^n) + e^{\rho dt_n} dt_n C(I_n) \\
&\geqslant \rho dt_n \varepsilon + \rho dt_n V(\mu^n) + e^{\rho dt_n} dt_n C(I_n) \\
\implies \rho V(\mu_n) &\leqslant -\rho\varepsilon + \sum \frac{p_i^n}{dt_n}(V(\nu_i^n) - V(\mu^n)) - e^{\rho dt_n} C(I_n) \\
\implies \rho V(\mu_n) &\leqslant -\rho\varepsilon + \sum \frac{p_i^n}{dt_n}(V(\nu_i^n) - V(\mu^n)) - C(I_n) \qquad \text{(A.11)}
\end{aligned}$$

The first equality is by the definition of U_n. The first inequality is from $\mu^n \in \arg\min U_n$. The second inequality is from $e^x - 1 \geqslant x$. The third inequality is from $U_n(\mu^n) \leqslant -\varepsilon$. Now since the number of posteriors ν_i^n is no more than $2|X|$, we can take a subsequence of n such that all $\lim \nu_i^n = \nu_i$. Partition ν_i^n into two kinds: $\lim \nu_i^n = \nu_i \neq \mu$, $\lim \nu_j^n = \mu$. Since V is unimprovable, $\forall c, \widehat{\sigma}$ we have $D^2V(\mu, \widehat{\sigma})\|\widehat{\sigma}\|^2 \leqslant -\widehat{\sigma}^T \mathrm{H}H(\mu)\widehat{\sigma}\left(\frac{\rho}{I}V(\mu) + \frac{C(I)}{I}\right)$. Since $V \in C^{(1)}$, $H \in C^{(2)}$, $\forall \eta$, there exists δ s.t. $\forall |\mu' - \mu| \leqslant \delta$:

$$\begin{aligned}
&\begin{cases}
\|\mathrm{H}H(\mu) - \mathrm{H}H(\mu')\| \leqslant \eta \\
|V(\mu) - V(\mu')| \leqslant \eta
\end{cases} \\
\implies D^2V(\mu', \widehat{\sigma}) &\leqslant \left(\frac{\rho}{I}V(\mu') + \frac{C(I)}{I}\right)\left(-\frac{\widehat{\sigma}^T \mathrm{H}H(\mu')\widehat{\sigma}}{\|\widehat{\sigma}\|^2}\right) \\
&\leqslant \left(\frac{\rho}{I}V(\mu) + \frac{C(I)}{I}\right)\left(-\frac{\widehat{\sigma}^T \mathrm{H}H(\mu)\widehat{\sigma}}{\|\widehat{\sigma}\|^2}\right) + \left(\frac{\rho}{I}\sup F + \frac{C(I)}{I}\right)\eta + \frac{\rho}{I}\eta\|\mathrm{H}H(\mu)\|
\end{aligned}$$

If we pick η and δ properly:

$$D^2V(\mu', \widehat{\sigma}) \leqslant \left(\frac{\rho}{I}V(\mu) + \frac{C(I)}{I}\right)\left(-\frac{\widehat{\sigma}^T \mathrm{H}H(\mu)\widehat{\sigma}}{\|\widehat{\sigma}\|^2}\right) + \frac{1+C(I)}{I}\eta$$

Then there exists N s.t. $\forall n \geqslant N$, $\left|\nu_j^n - \mu\right| < \delta$, $|\mu^n - \mu| < \delta$. Now I want to do a second-order approximation of $V(\nu_j^n) - V(\mu^n) - \nabla V(\mu^n)(\nu_j^n - \mu^n)$. To apply Taylor expansion to a not necessarily twice differentiable function V, I invoke a technical Lemma B.10 to $g(\alpha) = V(\alpha\nu_j^n + (1-\alpha)\mu^n)$:

$$\begin{aligned}
&V(\nu_j^n) - V(\mu^n) - \nabla V(\mu^n)(\nu_j^n - \mu^n) = g(1) - g(0) - g'(0) \\
\leqslant&\frac{1}{2}\sup_{\alpha\in(0,1)} D^2 g(\alpha, 1) = \sup_{\alpha\in(0,1)} \limsup_{d\to 0} \frac{g(\alpha+d) - g(\alpha) - g'(\alpha)d}{d^2} \\
=&\sup_{\xi\in(\mu^n,\nu_j^n)} \limsup_{d\to 0} \frac{V(\xi + d(\nu_j^n) - \mu^n) - V(\xi) - dJV(\xi)(\nu_j^n - \mu^n)}{d^2} \\
\leqslant&\frac{1}{2}\sup_{|\xi-\mu|\leqslant\delta} D^2V(\xi, \nu_j^n - \mu^n)\left\|\nu_j^n - \mu^n\right\|^2 \\
\leqslant&-\frac{1}{2}\left(\frac{\rho}{I}V(\mu) + \frac{C(I)}{I}\right)(\nu_j^n - \mu^n)^T \mathrm{H}H(\mu)(\nu_j^n - \mu^n) + \frac{1+C(I)}{2I}\eta\left\|\nu_j^n - \mu^n\right\|^2
\end{aligned} \tag{A.12}$$

Therefore, by applying Equation (A.12):

$$\begin{aligned}
&\sum p_{i,j}^n\left(V(\nu_{i,j}^n) - V(\mu^n)\right) \\
=&\sum p_i^n(V(\nu_i^n) - V(\mu^n) - \nabla V(\mu^n)(\nu_i^n - \mu^n)) + \sum p_j^n\left(V(\nu_j^n) - V(\mu^n) - \nabla V(\mu^n)\left(\nu_j^n - \mu^n\right)\right) \\
\leqslant&\sum p_i^n(V(\nu_i^n) - V(\mu^n) - \nabla V(\mu^n)(\nu_i^n - \mu^n)) \\
&-\frac{1}{2}\left(\frac{\rho}{I}V(\mu) + \frac{C(I)}{I}\right)\sum p_j^n\left(\nu_j^n - \mu^n)^T \mathrm{H}H(\mu)(\nu_j^n - \mu^n)\right) + \frac{1+C(I)}{2I}\eta\sum p_j^n\left\|\nu_j^n - \mu^n\right\|^2
\end{aligned} \tag{A.13}$$

Notice that Equations (A.12) and (A.13) are true uniform to I, so we can replace I with

I_n and Equation (A.13) is still true. Now let $\bar{p}_i^n = \frac{p_i^n}{dt_n}$,
$-\hat{\sigma}_n^T \mathrm{H}H(\mu^n)\hat{\sigma}^n dt_n = \sum p_j^n \Big(H(\mu^n) - H(\nu_j^n) + \nabla H(\mu)(\nu_j^n - \mu^n) \Big)$, we have:

$$\sum \bar{p}_i^n \big(H(\mu^n) - H(\nu_i^n) + H'(\mu^n)(\nu_i^n - \mu^n) \big) - \hat{\sigma}_n^T \mathrm{H}H(\mu^n)\hat{\sigma}_n = I_n \tag{A.14}$$

$\big(\bar{p}_i^n, \nu_i^n, \hat{\sigma}_n\big)$ is a feasible experiment for Equation (A.4). Therefore, by optimality of V at μ^n, we have

$$\begin{cases} \sum \bar{p}_i^n (V(\nu_i^n) - V(\mu^n) - \nabla V(\mu^n)(\nu_i^n - \mu^n)) \leqslant \Big(I_n + \hat{\sigma}^{nT} \mathrm{H}H(\mu^n)\hat{\sigma}^n \Big) \Big(\frac{\rho}{I_n} V(\mu^n) + \frac{C(I_n)}{I_n} \Big) \\ D^2 V(\mu^n, \hat{\sigma}^n) \leqslant -\frac{\hat{\sigma}^{nT} \mathrm{H}H(\mu)\hat{\sigma}^n}{\|\hat{\sigma}^n\|^2} \Big(\frac{\rho}{I_n} \overline{V}(\mu^n) + \frac{C(I_n)}{I_n} \Big) \end{cases} \tag{A.15}$$

Then we study term $\sum p_j^n (\nu_j^n - \mu^n)^2$. Apply Lemma B.10 to $g(\alpha) = H(\alpha \nu_j^n + (1-\alpha)\mu^n)$:

$$\begin{aligned} &\sum p_j^n \Big(H(\mu^n) - H(\nu_j^n) + \nabla H(\mu^n)(\nu_j^n - \mu^n) \Big) \\ \geqslant &\frac{1}{2} \inf_{\xi_j^n \in [\mu^n, \nu_n^j]} \sum p_j^n \Big(-(\nu_j^n - \mu^n)^T \mathrm{H}H(\xi_j^n)(\nu_j^n - \mu^n) \Big) \\ \geqslant &-\frac{1}{2} \sum p_j^n ((\nu_j^n - \mu^n)^T \mathrm{H}H(\mu)(\nu_j^n - \mu^n)) - \frac{1}{2} \eta \sum p_j^n \Big\| \nu_j^n - \mu^n \Big\|^2 \end{aligned} \tag{A.16}$$

Therefore, to sum up:

$$\begin{aligned} \sum \frac{p_{i,j}^n}{dt_n} \Big(V(\nu_{i,j}^n) - V(\mu^n) \Big) \leqslant & \sum \bar{p}_i^n (V(\nu_i^n) - V(\mu^n) - \nabla V(\mu^n)(\nu_i^n - \mu^n)) \\ & + \frac{1}{2} \sum \frac{p_j^n}{dt_n} \Big(-(\nu_j^n - \mu^n)^T \mathrm{H}H(\mu)(\nu_j^n - \mu^n) \Big(\frac{\rho}{I_n} V(\mu) + \frac{C(I_n)}{I_n} \Big) \Big) \\ & + \sum \frac{p_j^n}{dt_n} \Big(\frac{1 + C(I_n)}{2I_n} \eta \Big\| \nu_j^n - \mu^n \Big\|^2 \Big) \\ \leqslant & \Big(I_n + \hat{\sigma}^{nT} \mathrm{H}H(\mu^n)\hat{\sigma}^n \Big) \Big(\frac{\rho}{I_n} V(\mu^n) + \frac{C(I_n)}{I_n} \Big) \end{aligned}$$

$$
\begin{aligned}
&+\Bigg(\sum \frac{p_j^n}{dt_n}(H(\mu^n)-H(\nu_j^n)+\nabla H(\mu^n)(\nu_j^n-\mu^n)) \\
&+\frac{1}{dt_n}\frac{\eta}{2}\sum p_j^n \left\|\nu_j^n-\mu^n\right\|^2\Bigg)\left(\frac{\rho}{I_n}V(\mu)+\frac{C(I_n)}{I_n}\right) \\
&+\frac{1}{dt_n}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\frac{1+C(I_n)}{2I_n}\eta \\
=&\left(I_n+\widehat{\sigma}^{nT}\mathrm{H}H(\mu^n)\widehat{\sigma}^n\right)\left(\frac{\rho}{I_n}V(\mu^n)+\frac{C(I_n)}{I_n}\right) \\
&+\left(-\widehat{\sigma}^{nT}\mathrm{H}H(\mu^n)\widehat{\sigma}^n+\frac{1}{dt_n}\frac{\eta}{2}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\right)\left(\frac{\rho}{I_n}V(\mu)+\frac{C(I_n)}{I_n}\right) \\
&+\frac{1}{dt_n}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\frac{1+C(I_n)}{2I_n}\eta \\
\leqslant&\rho V(\mu^n)+C(I_n)+\frac{1}{dt_n}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\left(\frac{1+\rho V(\mu)+2C(I_n)}{2I_n}\right)\eta+\rho\eta
\end{aligned}
$$

The first inequality is Equation (A.13). The second inequality comes from Equation (A.15) and Equation (A.16). The next equality comes from definition of $\widehat{\sigma}_n^2$. The last inequality comes from canceling out terms and $-\widehat{\sigma}^{nT}\mathrm{H}H(\mu^n)\widehat{\sigma}^n \leqslant I_n$ (Notice the difference between $V(\mu)$ and $V(\mu^n)$). Then by plug into Equation (A.11):

$$
\rho V(\mu^n)\leqslant -\rho\varepsilon+\rho V(\mu^n)+\frac{1}{dt_n}\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\left(\frac{1+\rho V(\mu)+2C(I_n)}{2I_n}\right)\eta+\rho\eta
$$

Moreover:

$$
\begin{aligned}
&\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\inf_\sigma\frac{\left|\sigma^T\mathrm{H}H(\mu)\sigma\right|}{\|\sigma\|^2} \\
&\leqslant\sum p_j^n(\nu_j^n-\mu^n)\mathrm{H}H(\mu)(\mu^n-\mu^n)\leqslant I_n dt_n+\eta\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2 \\
\Longrightarrow&\sum p_j^n\left\|\nu_j^n-\mu^n\right\|^2\leqslant\frac{I_n dt_n}{\inf_\sigma\frac{\left|\sigma^T\mathrm{H}H(\mu)\sigma\right|}{\|\sigma\|^2}-\eta} \\
\Longrightarrow&\rho\varepsilon\leqslant\frac{1}{2}(1+\rho V(\mu)+2C(I_n))\frac{\eta}{\inf_\sigma\frac{\left|\sigma^T\mathrm{H}H(\mu)\sigma\right|}{\|\sigma\|^2}-\eta}+\rho\eta
\end{aligned}
$$

By Lemma B.7, $C(I_n)$ is uniformly bounded above. Since H is strictly concave $\inf_\sigma \frac{\left|\sigma^T \mathrm{H}H(\mu)\sigma\right|}{\|\sigma\|^2}$ is positive. The inequality holds when η is chosen smaller than $\inf_\sigma \frac{\left|\sigma^T \mathrm{H}H(\mu)\sigma\right|}{\|\sigma\|^2}$. By taking $\eta \to 0$, the LHS is eventually larger than the RHS. Contradiction. Therefore:

$$V(\mu) = \limsup_{dt \to 0} V_{dt}(\mu) = \overline{V}(\mu)$$

■

A.2.3 Proof of Theorem 1.2

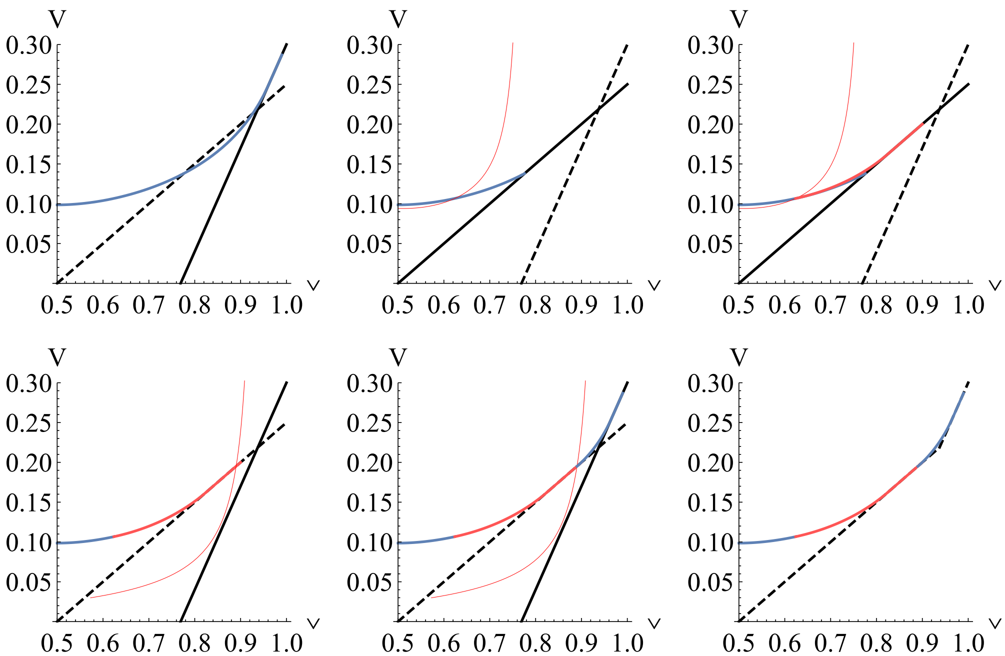
Proof. I prove Theorem 1.2 by guess and verification. To simplify notation, I define a flow version of information measure:

$$J(\mu, \nu) = H(\mu) - H(\nu) + H'(\mu)(\nu - \mu)$$

Then total flow information cost is $p \cdot J(\mu, \nu)$. Let $F_m = E_\mu[u(a_m, x)]$ and reorder a_m s.t. F'_m is increasing in m. Let $\underline{\mu}_k$ be each kink points of F: $F(\mu) = F_k(\mu) \iff \mu \in \left[\underline{\mu}_{k-1}, \underline{\mu}_k\right]$. $\overline{m}$ is the smallest index s.t. $F'_{\overline{m}} \geqslant 0$.

Algorithm:

In this part, I introduce the algorithm for constructing $V(\mu)$ and $\nu(\mu)$. I only discuss the case $\mu \geqslant \mu^*$. The remaining case $\mu \leqslant \mu^*$ follows by a symmetric method. The main steps are illustrated in Figure A.6. The first step is to find critical the belief μ^* at which two sided stationary Poisson signal is optimal (μ^*=0.5 in a symmetric problem). Then value function is solved by searching over optimal posterior beliefs, given choosing an action (say a_m). Then the remaining actions are added one by one to consideration. And value function is updated when each additional action is added. Finally, after all actions have been considered, I complete the construction of value function.



The two black (dashed and solid) lines are $F_{m-1}(\mu)$, $F_m(\mu)$.
The blue line is optimal value function from taking immediate action m.
The red line is optimal value function from taking immediate action $m-1$.

Figure A.6: Construction of optimal value function.

- *Step 1*: Define:

$$\overline{V}^+(\mu) = \max_{\nu \geqslant \mu} \frac{F_m(\nu)}{1 + \frac{\rho}{c} J(\mu, \nu)}$$
$$\overline{V}^-(\mu) = \max_{\nu \leqslant \mu} \frac{F_m(\nu)}{1 + \frac{\rho}{c} J(\mu, \nu)}$$

In Lemma A.2 I analyze the technical details of $\overline{V}^+$ and $\overline{V}^-$. The main property is that: $\overline{V}^+$ is increasing and $\overline{V}^-$ is decreasing. There exists $\mu^* \in [0,1]$ s.t. $\overline{V}^+(\mu) \geqslant \overline{V}^-(\mu)$ when $\mu \geqslant \mu^*$ and $\overline{V}^-(\mu) \leqslant \overline{V}^-(\mu)$ when $\mu \leqslant \mu^*$. Define $\overline{V}(\mu) = \max\{\overline{V}^+(\mu), \overline{V}^-(\mu)\}$.

- *Step 2:* I construct the first piece of $V(\mu)$ to the right of μ^*. There are three possible cases of μ^* to be discussed (I omitted $\mu^* = 1$ by symmetry).

Case 1: Suppose $\mu^* \in (0,1)$ and $\overline{V}(\mu^*) > F(\mu^*)$. Then, there exists m and $\nu(\mu^*) \in (\mu^*, 1)$

s.t.

$$\overline{V}(\mu^*) = \frac{F_m(\nu(\mu^*))}{1 + \frac{\rho}{c} J(\mu^*, \nu(\mu^*))}$$

Initial condition $\left(\mu_0 = \mu^*, V_0 = \overline{V}(\mu^*), V_0' = 0\right)$ satisfies Lemma A.3, which states that there exists $V_m(\mu)$ solving:

$$V_m(\mu) = \max_{\nu \geqslant \mu} \frac{c}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)}$$

This refers to Figure A.6-1. Define

$$V_{\mu^*}(\mu) = \begin{cases} F(\mu) \text{ if } \mu \leqslant \mu^* \\ V_m(\mu) \text{ if } \mu \geqslant \mu^* \end{cases}$$

Be Lemma A.3, when $V_{\mu^*}(\mu) > F(\mu)$, V_{μ^*} is smoothly increasing and optimal $\nu(\mu)$ is smoothly decreasing.

Now update $V_{\mu^*}(\mu)$ with respect to more actions (in the order of decreasing index m). First consider F_{m-1} and let $\hat{\mu}_m$ be the smallest $\mu \geqslant \mu^*$ such that:

$$V_{\mu^*}(\widehat{\mu}_m) = \max_{\nu \geqslant \widehat{\mu}_m} \frac{c}{\rho} \frac{F_{m-1}(\nu) - V_{\mu^*}(\widehat{\mu}_m) - V'_{\mu^*}(\widehat{\mu}_m)(\nu - \widehat{\mu}_m)}{J(\widehat{\mu}_m, \nu)} \tag{A.17}$$

At $\widehat{\mu}_m$, searching posterior on F_{m-1} first dominates searching posterior on F_m[7]. This step refers to Figure A.6-2. $\widehat{\mu}_m$ is the smallest intersection point of blue curve ($V_{\mu^*}(\mu)$, LHS of Equation (A.17)) and thin red curve (RHS of Equation (A.17)). If $V_m(\widehat{\mu}_m) > F_{m-1}(\widehat{\mu}_m)$ then solve for V_{m-1} with initial condition $\mu_0 = \hat{\mu}_m, V_0 = V_m(\hat{\mu}_m), V_0' = V_m'(\hat{\mu}_m)$

[7]Existence is guaranteed by smoothness of V_{μ^*} and J. Noticing that $V_m(\hat{\mu}_m) \geqslant F_{m-1}(\hat{\mu}_m)$. Otherwise, there will be a $\widehat{\mu}_m' < \widehat{\mu}_m$ s.t. $V_m(\widehat{\mu}_m') = F_{m-1}(\widehat{\mu}_m')$ and it is easy to verify that V_m is weakly larger than the maximum. So there is an even smaller $\widehat{\mu}_m$, contradiction.

according to Lemma A.3 and redefine $V_{\mu*}(\mu) = V_{m-1}(\mu)$ when $\mu \geqslant \widehat{\mu}_m$. Otherwise skip to looking for $\widehat{\mu}_{m-1}$. If $m-1 > \overline{m}$, continue this procedure by looking for $\hat{\mu}_{m-1}$ and update $V_{\mu^*}|_{\mu \geqslant \widehat{\mu}_{m-1}}$ with corresponding V_{m-2} ...until $m = \overline{m}$ (No action with the slope of F'_m being negative is considered). This refers to Figure A.6-3. Now suppose $V_{\overline{m}}$ first hits $F(\mu)$ at some point μ^{**} ($\mu^{**} > \mu^*$ since $V_m(\mu^*) > F(\mu^*)$). V_{μ^*} is a (piecewise) smooth function on $[\mu^*, \mu^{**}]$ such that:

$$V_{\mu^*}(\mu) = \begin{cases} F(\mu) \text{ if } \mu \leqslant \mu^* \text{ or } \mu \geqslant \mu^{**} \\ V_k(\mu) \text{ if } \mu \in [\widehat{\mu}_k, \widehat{\mu}_{k-1}]^{8} \end{cases}$$

By construction, optimal posterior $\nu_{\mu^*}(\mu)$ is smoothly decreasing on each $(\widehat{\mu}_{k+1}, \widehat{\mu}_k)$ and jumps down at each $\widehat{\mu}_k$[9]. Notice that it is not yet proved that this order of value function updating is WLOO. It is possible that optimal policy function is non-monotonic. This is taken care of by Lemma B.18, which proves the order of updating being WLOO. I relegate the proof of Lemma B.18 to supplemental materials to conserve space, but it uses exactly the techniques of the verification step 2. Now I can claim that $\forall \mu \in [\mu^*, \mu^{**}]$:

$$V_{\mu^*}(\mu) = \max_{\nu \geqslant \mu, k} \frac{c}{\rho} \frac{F_k(\nu) - V_{\mu^*}(\mu) - V'_{\mu^*}(\mu)(\nu - \mu)}{J(\mu, \nu)} \tag{A.18}$$

Case 2: Suppose $\mu^* \in (0,1)$ but $\overline{V}(\mu^*) = F(\mu^*)$, let $\mu^{**} = \inf\{\mu \geqslant \mu^* | \overline{V}(\mu) > F(\mu)\}$.

Case 3: Suppose $\mu^* = 0$, then $F'(0) \geqslant 0$ (by Lemma A.2). Consider

$$\widetilde{V}(\mu) = \max_{\nu \geqslant \mu, k} \frac{c}{\rho} \frac{F_k(\nu) - F_1(\mu) - F'_1(\nu - \mu)}{J(\mu, \nu)}$$

[8] Define $\widehat{\mu}_{m+1} = \mu^*$ and $\widehat{\mu}_{\overline{m}} = \mu^{**}$ for consistency.

[9] Since F_{k-1} always crosses F_k from above, when indifference between choosing F_{k-1} and F_k, the posterior corresponding to F_{k-1} must be smaller.

Define, $\mu^{**} = \inf\left\{\mu|\widetilde{V}(\mu) > F_1(\mu)\right\}$. By Assumption 1.3, $\lim_{\mu\to 0}|H'(\mu)| = \infty$, then there exists δ s.t. $\forall \mu < \delta$, $\forall \nu \geqslant \underline{\mu}_2$, $\frac{\sup F}{J(\mu,\nu)} \leqslant \inf F$. Therefore $\mu^{**} \geqslant \delta > 0$. This step refers to Figure A.6-4.

- *Step 3:* Solve for V to the right of μ^{**}. For all $\mu^{\diamond} \geqslant \mu^{**}$ such that:

$$F(\mu^{\diamond}) = \max_{\nu\geqslant\mu,k} \frac{c}{\rho}\frac{F_k(\nu) - F(\mu^{\diamond}) - F'^{-}(\mu^{\diamond})(\nu-\mu^{\diamond})}{J(\mu^{\diamond},\nu)} \tag{A.19}$$

Let m be the index of optimal action. Solve for V_m with initial condition $\mu_0 = \mu^{\diamond}, V_0 = F(\mu^{\diamond}), V_0' = F'^{-}(\mu^{\diamond})$. [10] Then take same steps in *Step 2* and solve for $\widehat{\mu}_k$ and V_{k-1} sequentially until V_{m_0} first hits F. This step refers to Figure A.6-4,5. Now suppose V_{m_0} first hits $F(\mu)$ at some point $\mu^{\diamond\diamond}$ (can potentially be μ), define:

$$V_{\mu^{\diamond}}(\mu) = \begin{cases} F(\mu) & \text{if } \mu < \mu^{\diamond} \text{ or } \mu > \mu^{\diamond\diamond} \\ V_k(\mu) & \text{if } \mu \in [\widehat{\mu}_{k+1}, \widehat{\mu}_k]^{11} \end{cases}$$

By Lemma A.3, V_μ is piecewise smooth are pasted smoothly. So V_μ is a smooth function on $[\mu, \mu'']$. Optimal posterior $\nu_{\mu^{\diamond}}(\mu)$ is smoothly decreasing on each $(\widehat{\mu}_{k+1}, \widehat{\mu}_k)$ and jumps down at each $\widehat{\mu}_k$. By Lemma B.18 and our construction, $\forall \mu \in [\mu^{\diamond}, \mu^{\diamond\diamond}]$:

$$V_\mu(\mu) = \max_{\nu\geqslant\mu^{\diamond},k} \frac{c}{\rho}\frac{F_k(\nu) - V_{\mu^{\diamond}}(\mu) - V'_{\mu^{\diamond}}(\mu)(\nu-\mu)}{J(\mu,\nu)} \tag{A.18}$$

Let Ω be the set of all such $\mu^{\diamond}$'s.

[10] By definition of μ^{**}, μ_0 is bounded away from $\{0,1\}$ and Equation (A.19) implies conditions in Lemma A.3 are satisfied.

[11] Define $\widehat{\mu}_{m+1} = \mu^{\diamond}$ and $\widehat{\mu}_{m_0} = \mu^{\diamond\diamond}$ for consistency.

- *Step 4:* Define:

$$V(\mu) = \begin{cases} V_{\mu^*}(\mu) & \text{if } \mu \in [\mu^*, \mu^{**}] \\ \sup_{\mu^\diamond \in \Omega} \left\{V_{\mu^\diamond}(\mu)\right\} & \text{if } \mu \geqslant \mu^{**} \end{cases} \tag{A.20}$$

In the algorithm, I only discussed the case $\mu^* < 1$ and constructed the value function on the right of μ^*. On the left of μ^*, V can be defined using a totally symmetric argument by referring to Lemma A.3′ and Lemma B.18′.

Smoothness:

I need to verify that $V(\mu)$ that defined as Equation (A.20) is a $C^{(1)}$ smooth function on $[0,1]$. This claim is purely for technical use (for example, the validity of using V' and V''). I relegate this technical proof to Appendix B.2.1 in Lemmas B.11, B.12, B.13 and B.14. In addition, it is shown in Appendix B.2.1 that there exists a set of μ_0 such that on each interval when $V(\mu) > F(\mu)$, $V(\mu)$ is defined as one V_{μ_0}.

Unimprovability:

Finally, I prove unimprovability of $V(\mu)$.

- *Step 1*: I first show that $V(\mu)$ solves the following problem:

$$V(\mu) = \max\left\{F(\mu), \max_{\nu,m} \frac{c}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{J(\mu,\nu)}\right\} \tag{P-C}$$
$$\begin{cases} \nu \geqslant \mu \text{ when } \mu \geqslant \mu^* \\ \nu \leqslant \mu \text{ when } \mu \leqslant \mu^* \end{cases}$$

Equation (P-C) is the maximization problem over all confirmatory evidence seeking with immediate decision making upon arrival of signals. Equation (P-C) is implied by Equation (A.18) for $\mu \in E$. So it is sufficient to prove Equation (P-C) for $\mu \in E^C$. Suppose for the sake of contradictoin that there exists $\mu \geqslant \mu^*$ s.t. Equation (P-C) is violated. Let

$F(\mu) = F_k(\mu)$. Then it is equivalently stating that:

$$U(\mu) = \max_{\nu, k' > k} \frac{c}{\rho} \frac{F'_k(\nu) - F_k(\mu) - F'_k(\nu - \mu)}{J(\mu, \nu)} > F_k(\mu)$$

Consider $\underline{\mu}_k$ (the intercection of F_k and F_{k-1}). By Lemma B.11, there exists I_k s.t. $\underline{\mu}_k \in I_k$. At $b_k = \sup I_k$, $U(b_k) \leqslant F_k(b_k)$. Therefore, since $U(\mu)$ is continuous, by intermediate value theorem there exists largest μ' between $\underline{\mu}_k$ and μ s.t .$U(\mu') = F_k(\mu')$. Then Equation (A.19) is satisfied at μ' so consider $V_{\mu'}$. Sicne $V_{\mu'}(\mu) \leqslant V(\mu) = F_k(\mu)$, there exists $\mu'' \in (\mu', \mu)$ s.t. $V_{\mu'}(\mu'') \leqslant F_k(\mu'')$ and $V'_{\mu'}(\mu'') \leqslant F_k(\mu'')$. Therefore $U(\mu'') > F_k(\mu'')$ implies $V_{\mu'}(\mu'') > F_k(\mu'')$, contradiction. Apply a symmetric argument to $\mu \leqslant \mu^*$, I prove Equation (P-C).

- *Step 2*: I show that $V(\mu)$ solves the following problem:

$$V(\mu) = \max \left\{ F(\mu), \max_{\nu} \frac{c}{\rho} \frac{V(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)} \right\} \qquad \text{(P-D)}$$
$$\begin{cases} \nu \geqslant \mu \text{ when } \mu \geqslant \mu^* \\ \nu \leqslant \mu \text{ when } \mu \leqslant \mu^* \end{cases}$$

Equation (P-D) is the maximization problem over all confirmatory learning strategies. It has less constraint than Equation (P-C): when a signal arrives and posterior belief ν is realized, the DM is allowed to continue experimentation instead of being forced to take an action.

I only show the case $\mu \geqslant \mu^*$ and a totally symmetric argument applies to $\mu \leqslant \mu^*$. Suppose Equation (P-C) is violated at μ, then there exists ν' such that:

$$V(\mu) = \max_{\nu \geqslant \mu, m} \frac{c}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)} < \frac{c}{\rho} \frac{V(\nu') - V(\mu) - V'(\mu)(\nu' - \mu)}{J(\mu, \nu')} \qquad \text{(A.21)}$$

Let $\widetilde{V} = V(\mu)$. Suppose the maximizer is ν, m. Optimality implies first order conditions Equation (A.27) and Equation (A.26):

$$\begin{cases} F'_m + \frac{\rho}{c}\widetilde{V}H'(\nu) = V'(\mu) + \frac{\rho}{c}\widetilde{V}H'(\mu) \\ \left(F_m(\nu) + \frac{\rho}{c}\widetilde{V}H(\nu)\right) - \left(V(\mu) + \frac{\rho}{c}\widetilde{V}H(\mu)\right) = \left(V'(\mu) + \frac{\rho}{c}V(\mu)H'(\mu)(\nu - \mu)\right) \end{cases}$$

We define $L(V,\lambda,\mu)(\nu)$ and $G(V,\lambda)(\mu)$ as:

$$\begin{cases} L(V,\lambda,\mu)(\nu) = (V(\mu) + \lambda H(\mu)) + (V'(\mu) + \lambda H'(\mu))(\nu - \mu) \\ G(V,\lambda)(\nu) = V(\nu) + \lambda H(\nu) \end{cases} \tag{A.22}$$

Then L is a linear function of ν and $G(F_m, \frac{\rho}{c}\widetilde{V})(\nu)$ is a strictly concave smooth function of ν. Consider:

$$L\left(V, \frac{\rho}{c}\widetilde{V}, \mu\right)(\nu) - G\left(F_m, \frac{\rho}{c}\widetilde{V}\right)(\nu)$$

Equation (A.27) implies that it attains minimum 0 at ν. For any m' other than m,

$$L\left(V, \frac{\rho}{c}\widetilde{V}, \mu\right)(\nu) - G\left(F_{m'}, \frac{\rho}{c}\widetilde{V}\right)(\nu)$$

is convex and weakly larger than zero. However by Equation (A.21):

$$L\left(V, \frac{\rho}{c}\widetilde{V}, \mu\right)(\nu') - G\left(V, \frac{\rho}{c}\widetilde{V}\right)(\nu') = -\left(V(\nu') - V(\mu) - V'(\mu)(\nu' - \mu) - \frac{\rho}{c}\widetilde{V}J(\mu, \nu')\right) < 0$$

Therefore $L\left(V, \frac{\rho}{c}\widetilde{V}, \mu\right)(\nu) - G\left(V, \frac{\rho}{c}\widetilde{V}\right)(\nu)$ has strictly negative minimum. Suppose it's minimized at $\widetilde{\mu}$ ($\widetilde{\mu} > \mu$ since $L(V,\lambda,\mu)(\mu) \equiv G(V,\lambda)(\mu)$). Then FOC is a necessary condition:

$$V'(\mu)+\frac{\rho}{c}\widetilde{V}H'(\mu)=V'(\widetilde{\mu})+\frac{\rho}{c}\widetilde{V}H'(\widetilde{\mu})$$

Consider:

$$\begin{aligned}
&L\Big(V,\frac{\rho}{c}\widetilde{V},\widetilde{\mu}\Big)(\nu(\widetilde{\mu}))-G\Big(F_m,\frac{\rho}{c}\widetilde{V}\Big)(\widetilde{\nu})\\
=&L\Big(V,\frac{\rho}{c}\widetilde{V},\mu\Big)(\nu(\widetilde{\mu}))-G\Big(F_m,\frac{\rho}{c}\widetilde{V}\Big)(\nu(\widetilde{\mu}))\\
&+V(\widetilde{\mu})-V(\mu)+\frac{\rho}{c}\widetilde{V}(H(\widetilde{\mu})-H(\mu))-\Big(V'(\mu)+\frac{\rho}{c}\widetilde{V}H'(\mu)\Big)(\widetilde{\mu}-\mu)\\
\geqslant&V(\widetilde{\mu})-V(\mu)+\frac{\rho}{c}\widetilde{V}(H(\widetilde{\mu})-H(\mu))-\Big(V'(\mu)+\frac{\rho}{c}\widetilde{V}H'(\mu)\Big)(\widetilde{\mu}-\mu)\\
=&G\Big(V,\frac{\rho}{c}\widetilde{V}\Big)(\widetilde{\mu})-L\Big(V,\frac{\rho}{c}\widetilde{V},\mu\Big)(\widetilde{\mu})>0
\end{aligned}$$

In the first equality I used Equation (A.27) at $\widetilde{\mu}$. In first inequality I used suboptimality of $\widetilde{\mu}$ at μ. However for m' and $\nu(\widetilde{\mu})$ being optimizer at $\widetilde{\mu}$:

$$\begin{aligned}
0=&L\Big(V,\frac{\rho}{c}V(\widetilde{\mu}),\widetilde{\mu}\Big)(\nu(\widetilde{\mu}))-G\Big(F_{m'},\frac{\rho}{c}V(\widetilde{\mu})\Big)(\nu(\widetilde{\mu}))\\
=&L\Big(V,\frac{\rho}{c}\widetilde{V},\widetilde{\mu}\Big)(\nu(\widetilde{\mu}))-G\Big(F_{m'},\frac{\rho}{c}\widetilde{V}\Big)(\nu(\widetilde{\mu}))\\
&+\frac{\rho}{c}(V(\widetilde{\mu})-\widetilde{V})(H(\widetilde{\mu})-H(\nu(\widetilde{\mu}))+H'(\widetilde{\mu})(\nu(\widetilde{\mu})-\widetilde{\mu}))\\
>&\frac{\rho}{c}(V(\widetilde{\mu})-\widetilde{V})J(\widetilde{\mu},\nu(\widetilde{\mu}))
\end{aligned}$$

Contradiction. Therefore, I proved Equation (P-D).

- *Step 3*: I show that V satisfies Equation (A.4), which is less restrictive than Equation (P-D) by allowing 1) diffusion experiments. 2) evidence seeking of all possible posteriors instead of just confirmatory evidence.

  First, since V is smoothly increasing and has a piecewise differentiable optimizer ν, envelope theorem implies:

$$\begin{aligned}
V'(\mu) =& \frac{c}{\rho}\frac{-V''(\mu)(\nu-\mu)}{J(\mu,\nu)} + V(\mu)\frac{-H''(\mu)(\nu-\mu)}{J(\mu,\nu)} \\
=& -\frac{c}{\rho}\frac{\nu-\mu}{J(\mu,\nu)}\Big(V''(\mu)+\frac{\rho}{c}V(\mu)H''(\mu)\Big) > 0 \\
\implies & V''(\mu)+\frac{\rho}{c}V(\mu)H''(\mu) < 0
\end{aligned}$$

Therefore, allocating to diffusion experiment is strictly suboptimal. Moreover, consider:

$$\begin{aligned}
V^{-}(\mu) =& \max_{\nu\leqslant\mu}\frac{c}{\rho}\frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)} \\
\implies V^{-\prime}(\mu) =& -\frac{c}{\rho}\frac{\nu-\mu}{J(\mu,\nu)}\Big(V''(\mu)+\frac{\rho}{c}V^{-}H''(\mu)\Big)
\end{aligned}$$

$V^{-}(\mu)$ is by definition the utiltiy gain from searching contradictive evidence, given value function $V(\mu)$. By definition of μ^*, $V^{-}(\mu^*) = V(\mu^*)$ and whenever $V(\mu) = V^{-}(\mu)$ $V^{-\prime}(\mu) < 0$. Therefore, $V^{-}(\mu)$ can never cross $V(\mu)$ from below — $V^{-}(\mu)$ is lower than $V(\mu)$ and $V(\mu)$ is unimprovable with contraditive evidence. That is to say:

$$\begin{aligned}
\rho V(\mu) =& \max\Big\{\rho F(\mu), \max_{\nu,p,\sigma} p(V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)) + \frac{1}{2}V''(\mu)\sigma^2\Big\} \\
&\text{s.t. } pJ(\mu,\nu)+\frac{1}{2}H''(\mu)\sigma^2 \leqslant c
\end{aligned}$$

To sum up, I construct a policy function $\nu(\mu)$ and value function $V(\mu)$ solving Equation (A.4). Now consider the four properties in Theorem 1.2. First, by my construction algorithm, in the case $\mu^* \in \{0,1\}$, I can replace μ^* with $\mu^{**} \in (0,1)$. Therefore WLOG $\mu^* \in (0,1)$. Second, $E = \big\{\mu \in [0,1] \big| V(\mu) > F(\mu)\big\}$ is a union of disjoint open intervals $E = \bigcup I_m$. By my construction, $V(\mu) = V_{\mu^m}(\mu)|_{\mu\in I_m}$. On each I_m, $\nu_{\mu^m}(\mu)$ is sctrictly decreasing and jumps down at finite $\widehat{\mu}_k$'s. Finally, uniqueness argument in Lemma A.3

implies that ν is uniquely determined by FOC. Therefore, except for those discountinous points of ν, ν is uniquely defined. Number of such discontinuous points is countable, thus of zero measure. ■

Lemma A.2. *Define $\overline{V}^+$ and $\overline{V}^-$:*

$$\overline{V}^+(\mu) = \max_{\nu \geqslant \mu, m} \frac{F_m(\nu)}{1 + \frac{\rho}{c} J(\mu, \nu)}$$
$$\overline{V}^-(\mu) = \max_{\nu \leqslant \mu, m} \frac{F_m(\nu)}{1 + \frac{\rho}{c} J(\mu, \nu)}$$

There exists $\mu^* \in [0, 1]$ s.t. $\overline{V}^+(\mu) \geqslant \overline{V}^-(\mu)\ \forall \mu \geqslant \mu^*$; $\overline{V}^+(\mu) \leqslant \overline{V}^-(\mu)\ \forall \mu \leqslant \mu^*$.

Proof. I define function U_m^+ and U_m^- as follows:

$$U_m^+(\mu) = \max_{\nu \geqslant \mu} \frac{F_m(\nu)}{1 + \frac{\rho}{c} J(\mu, \nu)}$$
$$U_m^-(\mu) = \max_{\nu \leqslant \mu} \frac{F_m(\nu)}{1 + \frac{\rho}{c} J(\mu, \nu)}$$

First of all, I solve U_m^+, U_m^- on interior $\mu \in (0, 1)$. Since $F_m(\mu)$ is a linear function, $J(\mu, \nu) \geqslant 0$ is smooth, the objective function is a continuous function on compact domain. Therefore both maximization operators are well defined. Existence is already guaranteed, therefore I can refer to first order condition to characterize the maximizer:

$$\text{FOC} : F_m'\Big(1 + \frac{\rho}{c} J(\mu, \nu)\Big) + F_m(\nu)\frac{\rho}{c}\big(H'(\nu) - H'(\mu)\big) = 0 \tag{A.23}$$
$$\text{SOC} : \frac{\rho}{c} F_m'\big(H'(\nu) - H'(\mu)\big) \tag{A.24}$$

First discuss solving for $\nu \geqslant \mu$. Since $(1 + \frac{\rho}{c} J) > 0$, $H'' < 0$, $H'(\nu) - H'(\mu) \leqslant 0$ and inequality is strict when $\nu > \mu$. Therefore, if $F_m' < 0$, FOC being held will imply SOC being strictly positive at $\nu > \mu$. So $\forall F_m' < 0$, optimal ν is a corner solution. Moreover:

$$\frac{F_m(\mu)}{1+\frac{\rho}{c}J(\mu,\mu)} = F_m(\mu) > F_m(1) > \frac{F_m(1)}{1+\frac{\rho}{c}J(\mu,1)}$$

So $U_m^+(\mu) = F_m(\mu)$. If $F_m' = 0$, then $\forall \nu > \mu$:

$$\frac{F_m(\mu)}{1+\frac{\rho}{c}J(\mu,\mu)} = F_m(\mu) = F_m(\nu) \geqslant \frac{F_m(\nu)}{1+\frac{\rho}{c}J(\mu,\nu)}$$

Therefore $\forall F_m' \leqslant 0$, $U_m^+(\mu) = F_m(\mu)$. Then consider the case $F_m' > 0$. It can be easily verified that SOC is strictly negative when FOC holds and $\nu > \mu$. Therefore solution of FOC characterizes maximizer. Consider:

$$\begin{aligned}
&\lim_{\nu\to\mu} F_m'(1+\frac{\rho}{c}J(\mu,\nu)) + F_m(\nu)\frac{\rho}{c}(H'(\nu)-H'(\mu)) = F_m' > 0\\
&\lim_{\nu\to 1} F_m'(1+\frac{\rho}{c}J(\mu,\nu)) + F_m(\nu)\frac{\rho}{c}(H'(\nu)-H'(\mu)) = -\infty
\end{aligned}$$

Therefore be intermediate value theorem a unique solution $\nu \in (\mu, 1)$ exists by solving FOC. Since FOC is a smooth function of μ, ν and SOC is strictly negative, implicit function theorem implies ν being a smooth function of μ. This is sufficient to apply envelope theorem:

$$\frac{d}{d\mu}U_m^+(\mu) = \frac{F_m(\nu)(-H''(\mu)(\nu-\mu))}{\left(1+\frac{\rho}{c}J(\mu,\nu)\right)^2} > 0$$

Moreover, Equation (A.23) is strictly positive when $\nu = \mu$. This implies $U_m^+(\mu) > F_m(\mu)$ when $F_m' > 0$.

New consider limit of U_m^+ when $\mu \to 0, 1$. When $\mu \to 1$, $U_m^+(\mu) \leqslant \max_{\nu \geqslant \mu} F_m(\nu) = F(1)$.

When $\mu \to 0$, consider FOC Equation (A.23):

$$\begin{aligned}
&\lim_{\mu\to 0} F'_m\Big(1+\frac{\rho}{c}J(\mu,\nu)\Big)+F_m(\nu)\frac{\rho}{c}\big(H'(\nu)-H'(\mu)\big)\\
=&\lim_{\mu\to 0} F'_m\Big(1+\frac{\rho}{c}J(\nu,\mu)\Big)+F_m(\mu)\frac{\rho}{c}\big(H'(\nu)-H'(\mu)\big)\\
=&F'_m\Big(1+\frac{\rho}{c}J(\nu,0)\Big)+\lim_{\mu\to 0} F_m(\mu)\frac{\rho}{c}\big(H'(\nu)-H'(\mu)\big)=-\infty
\end{aligned}$$

Therefore, when $\mu \to 0$, optimal $\nu \to 0$. Therefore $\frac{F_m(\nu)}{1+\frac{\rho}{c}J(\mu,\nu)} \leqslant F_m(\nu) \to F_m(0)$. To conclude, $U^+_m(\mu) = F_m(\mu)$ when $\mu = 0, 1$. Let $\overline{m}$ be the first $F'_{\overline{m}} > 0$ (not necessarily exists). Let:

$$U^+(\mu) = \max_{m \geqslant \overline{m}} U^+_m(\mu)$$

Then $U^+(\mu)$ is a strictly increasing function when $\overline{m}$ exists. Symmetrically I can define $\underline{m}$ to be last $F'_m < 0$ and:

$$U^-(\mu) = \max_{m \leqslant \underline{m}} U^-_m(\mu)$$

There are three cases:

- *Case 1*: when F is not monotonic, then both U^+ and U^- exists. Moreover, $F(0) > F_{\overline{m}}(0)$ and $F(1) > F_{\underline{m}}(1)$. Therefore, $U^+(0) < U^-(0)$ and $U^+(1) > U^-(1)$. There must exists unique $\mu^* \in (0,1)$ s.t. $U^+(\mu^*) = U^-(\mu^*)$.
- *Case 2*: when $F' \geqslant 0$, then define $\mu^* = 0$.
- *Case 3*: when $F' \leqslant 0$, then define $\mu^* = 1$.

Finally, define $\overline{V}$:

$$\overline{V}^+(\mu) = \max\{F(\mu), U^+(\mu)\}$$

$$\overline{V}^{-}(\mu) = \max\{F(\mu), U^{-}(\mu)\}$$
$$\overline{V}(\mu) = \max\{\overline{V}^{+}(\mu), \overline{V}^{-}(\mu)\}$$

Given our construction, μ^* always exists and satisfies the conditions in Lemma A.2.

∎

Lemma A.3. *Assume $\mu_0 \geqslant \mu^*$, $F_m' \geqslant 0$, $V_0, V_0' \geqslant 0$ satisfies:*

$$\begin{cases} \overline{V}(\mu_0) \geqslant V_0 \geqslant F_m(\mu_0) \\ V_0 = \max\limits_{\nu \geqslant \mu_0} \dfrac{c}{\rho} \dfrac{F_m(\nu) - V_0 - V_0'(\nu - \mu_0)}{J(\mu_0, \nu)} \end{cases}$$

Then there exists a $C^{(1)}$ smooth and strictly increasing $V(\mu)$ defined on $[\mu_0, 1]$ satisfying

$$V(\mu) = \max_{\nu \geqslant \mu} \frac{c}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)} \tag{A.25}$$

and initial condition $V(\mu_0) = V_0, V'(\mu_0) = V_0'$. Maximizer $\nu(\mu)$ is $C^{(1)}$ and strictly decreasing on $\{\mu | V(\mu) > F_m(\mu)\}$.

Proof. I start from deriving the FOC and SOC for Equation (A.25):

$$\begin{aligned} &\text{FOC: } \frac{F_m' - V'(\mu)}{J(\mu, \nu)} + \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)^2} \big(H'(\nu) - H'(\mu)\big) = 0 \\ &\text{SOC: } \frac{H'(\nu) - H'(\mu)}{J(\mu, \nu)} \left(\frac{F_m' - V'(\mu)}{J(\mu, \nu)} + \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)^2} \big(H'(\nu) - H'(\mu)\big) \right) \\ &\qquad + \frac{H''(\nu)}{J(\mu, \nu)} \big(F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)\big) \leqslant 0 \end{aligned}$$

If feasibility is imposed:

$$V(\mu) = \frac{c}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)} \tag{A.26}$$

FOC and SOC reduces to:

$$\text{FOC: } F'_m - V'(\mu) + \frac{\rho}{c}V(\mu)(H'(\nu) - H'(\mu)) = 0 \tag{A.27}$$

$$\text{SOC: } \frac{\rho}{c}H''(\nu)V(\mu) \leqslant 0 \tag{A.28}$$

Let us proceed as follows. I use FOC and feasiblity to derive an ODE system with intial value defined by V_0, V'_0. Then I prove that the solution V must be strictly positive. Therefore, SOC is strict at the point where FOC is satisfied, the solution must be a local maximizer. Moreover, since $H'(\nu) - H'(\mu) < 0$, when FOC is negative, SOC must be strictly negative, then FOC can cross zero only from above and hence the solution to FOC is unique. Therefore the solution I get from the ODE system is the maximizer in Equation (A.25).

$$\begin{aligned}
&\begin{cases}
\text{Equation (A.26)} \implies V(\mu) = \dfrac{F_m(\nu) - V'(\mu)(\nu - \mu)}{1 + \frac{\rho}{c}J(\mu, \nu)} \\
\text{Equation (A.27)} \implies V'(\mu) = F'_m + \dfrac{\rho}{c}V(\mu)(H'(\nu) - H'(\mu))
\end{cases} \\
\implies &\begin{cases}
V(\mu) = \dfrac{F_m(\mu)}{1 - \frac{\rho}{c}J(\nu, \mu)} \\
V'(\mu) = F'_m + \dfrac{\frac{\rho}{c}F_m(\mu)(H'(\nu) - H'(\mu))}{1 - \frac{\rho}{c}J(\nu, \mu)}
\end{cases}
\end{aligned} \tag{A.29}$$

Consistency of Equation (A.29) implies that $\nu = \nu(\mu)$ is characterized by the following ODE:

$$\frac{\partial}{\partial \mu}\frac{F_m(\mu)}{1 - \frac{\rho}{c}J(\nu, \mu)} + \frac{\partial}{\partial \nu}\frac{F_m(\mu)}{1 - \frac{\rho}{c}J(\nu, \mu)}\dot{\nu} = F'_m + \frac{\frac{\rho}{c}F_m(\mu)(H'(\nu) - H'(\mu))}{1 - \frac{\rho}{c}J(\nu, \mu)} \tag{A.30}$$

Simplifying Equation (A.30):

$$
\begin{aligned}
&\frac{F_m'}{1-\frac{\rho}{c}J(\nu,\mu)} + \frac{\frac{\rho}{c}F_m(\mu)(H'(\nu)-H'(\mu))}{\left(1-\frac{\rho}{c}J(\nu,\mu)\right)^2} + \frac{\frac{\rho}{c}F_m(\mu)H''(\nu)(\mu-\nu)}{\left(1-\frac{\rho}{c}J(\nu,\mu)\right)^2}\dot{\nu} \\
=&\frac{F_m' + \frac{\rho}{c}(-F_m'J(\nu,\mu)+F_m(\mu)(H'(\nu)-H'(\mu)))}{1-\frac{\rho}{c}J(\nu,\mu)} \\
\implies & F_m(\mu)(H'(\nu)-H'(\mu)) + F_m(\mu)H''(\nu)(\mu-\nu)\dot{\nu} \\
& = (-F_m'J(\nu,\mu)+F_m(\mu)(H'(\nu)-H'(\mu)))(1-\frac{\rho}{c}J(\nu,\mu)) \\
\implies & F_m(\mu)H''(\nu)(\mu-\nu)\dot{\nu} = -F_m'J(\nu,\mu)(1-\frac{\rho}{c}J(\nu,\mu)) - \frac{\rho}{c}J(\nu,\mu)F_m(\mu)(H'(\nu)-H'(\mu)) \\
\implies & \dot{\nu} = J(\nu,\mu)\frac{F_m'\left(1-\frac{\rho}{c}J(\nu,\mu)\right)+\frac{\rho}{c}F_m(\mu)(H'(\nu)-H'(\mu))}{F_m(\mu)H''(\nu)(\nu-\mu)}
\end{aligned}
$$

Since I want to solve for V_0 on $[\mu_0, 1]$, I solve for ν_0 at μ_0 as the initial condition of ODE for ν. The technical details proving the existence of solution to the ODE is relegated to Lemma B.16, which checks standard conditions and invokes the Picard-Lindelof theorem. Lemma B.16 requires an inequality condition and I show it here:

The FOC characterizing ν is Equation (A.29):

$$
\begin{aligned}
& (F_m' - V_0')\Big(1-\frac{\rho}{c}J(\nu_0,\mu_0)\Big) + \frac{\rho}{c}F_m(\mu_0)\big(H'(\nu_0)-H'(\mu_0)\big) = 0 \\
\iff & F_m'\Big(1+\frac{\rho}{c}J(\mu_0,\nu_0)\Big) + \frac{\rho}{c}F_m(\nu_0)(H'(\nu_0)-H'(\mu)) = V_0'\Big(1-\frac{\rho}{c}J(\nu_0,\mu_0)\Big) \\
\iff & F_m(\mu_0)\Big(F_m'\Big(1+\frac{\rho}{c}J(\mu_0,\nu_0)\Big) + \frac{\rho}{c}F_m(\nu_0)(H'(\nu_0)-H'(\mu))\Big) = V_0'F_m(\mu_0)\Big(1-\frac{\rho}{c}J(\nu_0,\mu_0)\Big)
\end{aligned}
$$

Since $V_0 = \frac{F_m(\mu_0)}{1-\frac{\rho}{c}J(\nu_0,\mu_0)} \geqslant 0$, LHS is weakly positive. This satisifes the condition in Lemma B.16. Then Lemma B.16 guarantees existence of unique $\nu(\mu)$, and $\nu(\mu)$ is continuously decreasing from μ_0 until it hits $\nu = \mu$. Suppose $\nu(\mu)$ hits $\nu = \mu$ at $\bar{\mu}_m < 1$, define $V(\mu)$ as following:

$$V(\mu) = \begin{cases} \dfrac{F_m(\mu)}{1 - \frac{\rho}{c} J(\nu(\mu), \mu)} & \text{if } \mu \in [\mu_0, \bar{\mu}_m) \\ F_m(\mu) & \text{if } \mu \in [\bar{\mu}_m, 1] \end{cases}$$

Then I prove the properies of V:

1. V is by construction smooth except for at $\bar{\mu}$. When $\mu \to \bar{\mu}_m$, $\nu(\mu) \to \mu$. Therefore $J(\nu, \mu) \to 0$. This implies $V(\mu) \to F_m(\mu)$. So V is continuous.

2. By Equation (A.29), when $\mu \in [\mu_0, \bar{\mu}_m)$:

$$V'(\mu) = F'_m + \frac{F_m(\mu)(H'(\nu(\mu)) - H'(\mu))}{\frac{c}{\rho} - J(\nu(\mu), \mu)}$$

   When $\mu \to \bar{\mu}_m$, $H'(\nu(\mu)) - H'(\mu) \to 0$, $J(\nu(\mu), \mu) \to 0$. Thus $V'(\mu) \to F'_m$. So $V' \in C[\mu_0, 1] \implies V \in C^{(1)}[\mu_0, 1]$.

3. Rewrite Equation (A.29) on $[\mu_0, 1]$:

$$V'(\mu) = \frac{F'_m\big(1 + \frac{\rho}{c} J(\mu, \nu)\big) + F_m(\nu)(H'(\nu) - H'(\mu))}{1 - \frac{\rho}{c} J(\nu, \mu)} \tag{A.31}$$

   According to proof of Lemma B.16, $V'(\mu) > 0\ \forall \mu \in (\mu_0, 1]$. Moreover since $V_0 \geqslant 0$, $\forall \mu \in (\mu_0, 1]\ V(\mu) > 0$.

■

Appendix B

Supplemental materials for Chapter 1

B.1 Proofs in Section 1.5

This section contains formal proofs for theorems and lemmas in Section 1.5.

B.1.1 Useful lemmas

I first establish a useful Lemma B.1. Lemma B.1 is an analog to three key theorems on mutual information proved in Cover and Thomas (2012), generalizing the log-sum structure in mutual information to any function while keeping the key posterior separabiltiy.

B.1.1.1 Information theory resutls

Lemma B.1. *Information measure $I(\mathcal{S};\mathcal{X}|\mu)$ satisfies the following properties:*

1. *Markov property: If $\mathcal{X}\rightarrow\mathcal{S}\rightarrow\mathcal{T}$, then $I(\mathcal{T};\mathcal{X}|\mathcal{S})=0$.*
2. *Linear additivity: $I(\mathcal{S},\mathcal{T};\mathcal{X}|\mu)=I(\mathcal{S};\mathcal{X}|\mu)+E[I(\mathcal{T};\mathcal{X}|\mathcal{S},\mu)]$.*
3. *Information processing inequality: If $\mathcal{X}\rightarrow\mathcal{S}\rightarrow\mathcal{T}$, then $I(\mathcal{T};\mathcal{X}|\mu)\leqslant I(\mathcal{S};\mathcal{X}|\mu)$.*

Proof.

1. *Markov property*: Suppose the signal realization of $\mathcal{S},\mathcal{T}$ are denoted by s,t. Then:

$$\begin{aligned}
I(\mathcal{T};\mathcal{X}|\mathcal{S}) &= E_s[H(\mu(x|s))-E_t[H(\mu(x|t,s))|s]]\\
&= E_s[H(\mu(x|s))-E_t[H(\mu(x|s))|s]]\\
&= 0
\end{aligned}$$

   First equality is by definition of I. Second equality is by $\mathcal{T}\perp\mathcal{X}\big|\mathcal{S}$, then conditional on s,t will not shift belief of $\mathcal{X}$ at all.

2. *Chain rule*: Suppose the signal realization of $\mathcal{S},\mathcal{T}$ are denoted by s,t. Then:

$$\begin{aligned}
I(\mathcal{S},\mathcal{T};\mathcal{X}|\mu) &= E_{s,t}[H(\mu)-H(\mu(x|s,t))]\\
&= E_{s,t}[H(\mu)-H(\mu(x|s))+(H(\mu(x|s))-H(\mu(x|s,t)))]
\end{aligned}$$

$$= (H(\mu) - E_s[H(\mu(x|s))]) + (E_s[H(\mu(x|s)) - E_t[H(\mu(x|s,t))|s]])$$
$$= I(\mathcal{S};\mathcal{X}|\mu) + E[I(\mathcal{T};\mathcal{X}|\mathcal{S},\mu)]$$

First equality is by definition. Second equality is trivial. Third equality is by chain rule of conditional expectation.

3. *Information processing inequality*:

$$\begin{aligned} I(\mathcal{S};\mathcal{X}|\mu) &= I(\mathcal{S},\mathcal{T};\mathcal{X}|\mu) - I(\mathcal{T};\mathcal{X}|\mathcal{S},\mu) \\ &= I(\mathcal{S},\mathcal{T};\mathcal{X}|\mu) \\ &= I(\mathcal{T};\mathcal{X}|\mu) + I(\mathcal{S};\mathcal{X}|\mathcal{T},\mu) \\ &\geqslant I(\mathcal{T};\mathcal{X}|\mu) \end{aligned}$$

First and third equalities are from chain rule. Second equality is from Markov property.

∎

B.1.1.2 Concavification theorem

Theorem B.1 (Concavification). *Let X be a finite state space, $V \in C(\Delta X)$, $\mu \in \Delta X$. $H \in C(\Delta X)$ is non-negative. $f : \mathbb{R}^+ \mapsto \overline{\mathbb{R}}^+$ continuous, increasing and convex. Then there exists P s.t. $|\text{supp}(P)| \leqslant 2|X|$ solving:*

$$\sup_{P\in\Delta^2 X} E_P[V(\nu)] - f(H(\mu) - E_P[H(\nu)]) \tag{B.1}$$
$$s.t.\ E_P[\nu] = \mu$$

Let $I^* = H(\mu) - E_P[H(\nu)]$, there exists $\lambda \in \mathrm{d}f(I^*)$ such that:

$$\text{co}(V + \lambda H)(\mu) = E_P[(V + \lambda H)(\mu)]$$

Proof. Theorem B.1 is a corollary of Lemma 4.1 and Theorem 4.4 of Chapter 4.

Support size: since objective function is monototic in $(E_P[V], E_P[H])$, optimal solution must be on the boudary of set $\{(E_P[V], E_P[H])|E_P[\nu] = \mu\}$. Lemma 4.1 implies that there exists P solving Equation (B.1) and $|\text{supp}(P)| \leqslant 2|X|$.

Concavification: Suppose $f(I) = \infty \iff I > \bar{I}$. Since $v - f(H(\mu) - h)$ is a concave function in (v, h), and $H(\mu) - h \leqslant \bar{I}$ is a linear constraint, we can apply Theorem 4.4 of Chapter 4: let V^* be maximum of Equation (B.1), there exists λ s.t.

$$\begin{cases} P \in \arg \max\limits_{\substack{P \in \Delta^2(x) \\ E_P[\nu] = \mu}} E_P[\lambda_1 V - \lambda_2 H] \\ (E_P[V], I^*) \in \arg \min\limits_{\substack{I \leqslant \bar{I} \\ v - f(I) > V^*}} \lambda_1 v - \lambda_2 I \end{cases}$$

Then by Kuhn-Tucker condition (generalized to subgradients), there exists $\eta, \gamma \geqslant 0$ such that:

$$\begin{cases} \lambda_1 = \eta \\ \lambda_2 \in -\eta \partial f(I^*) - \gamma \end{cases}$$

If $\eta = 0$, then $\gamma > 0$ and P maximizes $E_P[\gamma H]$, then optimal P is uninformative and $I^* = 0$, contradiction. So $\eta > 0$. If $\gamma = 0$, then we can normalize $(\lambda_1, -\lambda_2)$ to $(1, \lambda)$ and $\lambda \in \partial f(I^*)$. If $\gamma > 0$, the complementary slackness condition implies that $I^* = \bar{I}$ and $\lambda^2/\eta \in \partial f(I^*)$. So we can also normalize $(\lambda_1, -\lambda_2)$ to $(1, \lambda)$ and $\lambda \in \partial f(I^*)$. ∎

B.1.1.3 Decomposition of information

In this section, I prove two important lemmas. Lemmas B.2 and B.3 shows that any static information structure can be decomposed into a continuous time belief process on unit time interval such that the flow reduction of informativeness is constant.

Lemma B.2. *$\forall \mu \in \Delta(X)$, $\forall \pi \in \Delta^2(X)$, $\int \pi(\nu)\mathrm{d}\nu = \mu$ and $|\pi|$ is finite. Then there exists probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and stochastic process $\langle \mu_t \rangle$ s.t.:*

1. *$\langle \mu_t \rangle$ is a Markovian martingale.*
2. *$\mu_0 = \mu$, $\mu_1 \sim \pi$.*
3. *$\forall t_1, t_2 \in [0,1]$ and $t_1 < t_2$, $E[H(\mu_{t_1}) - H(\mu_{t_2})|\mathcal{F}_{t_1}] = (t_2 - t_1)E[H(\mu_0) - H(\mu_1)]$.*

Proof. Define $\pi(\nu)$ as $(\nu_i, \pi_i)_{i=1}^N$. Let $M = \sum \pi_i H(\nu_i) - H(\mu)$. Consider the space $\Delta(N)$. Let $x_i = (0, \cdots, 0, 1, 0, \cdots, 0)$, which is an N dimensional vector with only ith element being 1. Then $\{x_i\}_{i=1}^N$ is an orthogonal normal base of $\Delta(N)$. Let $x_\mu = (\pi_1, \cdots, \pi_N)$. Then it is easy to see that $x_\mu \in \Delta(N)$. $\forall \lambda_i \in [0,1]$, let $x_{i,\lambda_i} = \lambda_i x_\mu + (1-\lambda_i)x_i \in \Delta(N)$.

Define map $f : \Delta(N) \to \Delta(X)$ as $f(x) = \sum_{j=1}^N x^j \nu_j$ (f is a linear map). Then consider

$$Q_i(\lambda_i) = \sum x^j_{i,\lambda_i}(H(f(x_{i,\lambda_i})) - H(\nu_j)) \tag{B.2}$$

Now consider properties of Q_i. First of all, since H is continuous and f is linear, $Q_i(\lambda_i)$ is continuous in λ_i. Second, suppose $\lambda_i' > \lambda_i$ and $\lambda_{i,\alpha} = \alpha\lambda_i + (1-\alpha)\lambda_i'$, consider:

$$\begin{aligned}
&Q_i(\lambda_{i,\alpha}) - \alpha Q_i(\lambda_i') - (1-\alpha)Q_i(\lambda_i)\\
=&H(f(x_{i,\lambda_{i,\alpha}})) - \alpha H(f(x_{i,\lambda_i})) - (1-\alpha)H(f(x_{i,\lambda_i'}))\\
&-\sum\left(x^j_{i,\lambda_{i,\alpha}} - \alpha x^j_{i,\lambda_i} - (1-\alpha)x^j_{i,\lambda_i'}\right)H(\nu_j)\\
=&H(f(x_{i,\lambda_{i,\alpha}})) - \alpha H(f(x_{i,\lambda_i})) - (1-\alpha)H(f(x_{i,\lambda_i'}))\\
=&H(f(\alpha x^j_{i,\lambda_i} + (1-\alpha)x^j_{i,\lambda_i'})) - \alpha H(f(x_{i,\lambda_i})) - (1-\alpha)H(f(x_{i,\lambda_i'}))\\
=&H(\alpha f(x^j_{i,\lambda_i}) + (1-\alpha)f(x^j_{i,\lambda_i'})) - \alpha H(f(x_{i,\lambda_i})) - (1-\alpha)H(f(x_{i,\lambda_i'}))\\
\geqslant&0
\end{aligned}$$

The first equality is by definition of Q_i. The second and third equalities is from linearity of x_{i,λ_i} in λ_i. The forth equality is from linearity of f. The last equality is from concavity of H. Hence, $Q_i(\lambda_i)$ is concave in λ_i. It is easy to verify that when $\lambda_i = 0$, $x_{i,\lambda_i} = x_i$ and $f(x_{i,\lambda_i}) = \nu_i$ so $Q_i(0) = 0$. When $\lambda_i = 1$, $x_{i,\lambda_i} = x_\mu$ and $f(x_{i,\lambda_i}) = \mu$ so $Q_i(1) = \sum \pi_j (H(\mu) - H(\nu_j)) = M$. Since Q_i is concave, the only possibility is that Q_i is first increasing then decreasing. Since Q_i is a continuous function, $\forall t \in [0,1]$, there exists λ_i in increasing region of Q_i s.t.:

$$Q_i(\lambda_i(t)) = (1-t)M$$

Since $(1-t)M$ is strictly decreasing M, $\lambda_i(t)$ is strictly decreasing in M. When $t \in (0,1]$, $\lambda_i(t) \in [0,1)$. Let $f(x_{i,\lambda_i(t)}) = \mu_i(t)$. Define:

$$\pi_i(t) = \frac{\frac{\pi_i}{1-\lambda_i(t)}}{\sum_j \frac{\pi_j}{1-\lambda_j(t)}}$$

It is easy to verify that:

$$\begin{aligned}
\sum \pi_i(t)\mu_i(t) &= f\left(\sum \pi_i(t) x_{i,\lambda_i(t)}\right) \\
&= f\left(\sum \pi_i(t)(\lambda_i(t)x_\mu + (1-\lambda_i(t))x_i)\right) \\
&= f\left(\frac{\sum \frac{\pi_i \lambda_i x_\mu}{1-\lambda_i(t)} + \sum \pi_i x_i}{\sum \frac{\pi_i}{1-\lambda_i}}\right) \\
&= f\left(\frac{\sum \frac{\pi_i \lambda_i}{1-\lambda_i} + 1}{\sum \frac{\pi_i}{1-\lambda_i}} x_\mu\right) \\
&= f(x_\mu) = \mu
\end{aligned}$$

Now for any $t, t' \in (0,1]$, and $t' > t$, define:

$$\pi_j(t'|\mu_i(t)) = \begin{cases} \dfrac{\lambda_i(t)-\lambda_i(t')}{1-\lambda_i(t')}\pi_j(t') & \text{if } i \neq j \\ \dfrac{\lambda_i(t)-\lambda_i(t')}{1-\lambda_i(t')}\pi_i(t') + \dfrac{1-\lambda_i(t)}{1-\lambda_i(t')} & \text{if } i = j \end{cases}$$

It is easy to verify that:

$$\begin{aligned}
\sum_j \pi_j(t'|\mu_i(t))\mu_j(t') =& f\left(\frac{1-\lambda_i(t)}{1-\lambda_i(t')}x_{i,\lambda_i(t')} + \frac{\lambda_i(t)-\lambda_i(t')}{1-\lambda_i(t')}\sum \pi_j(t')x_{j,\lambda_j(t')}\right) \\
=& f\left(\frac{1-\lambda_i(t)}{1-\lambda_i(t')}x_{i,\lambda_i(t')} + \frac{\lambda_i(t)-\lambda_i(t')}{1-\lambda_i(t')}x_\mu\right) \\
=& f\left(\frac{1-\lambda_i(t)}{1-\lambda_i(t')}\left(\lambda_i(t')x_\mu + (1-\lambda_i(t'))x_i\right) + \frac{\lambda_i(t)-\lambda_i(t')}{1-\lambda_i(t')}x_\mu\right) \\
=& f\left(\lambda_i(t)x_\mu + (1-\lambda_i(t))x_i\right) \\
=& \mu_i(t)
\end{aligned} \tag{B.3}$$

and:

$$\begin{aligned}
\sum_j \pi_j(t)\pi_i(t'|\mu_j(t)) =& \frac{1}{\sum \frac{\pi_j}{1-\lambda_j(t)}}\left(\frac{1-\lambda_i(t)}{1-\lambda_i(t')}\frac{\pi_i}{1-\lambda_i(t)} + \sum_j \frac{\lambda_j(t)-\lambda_j(t')}{1-\lambda_j(t')}\frac{\pi_j}{1-\lambda_j(t)}\pi_i(t')\right) \\
=& \frac{1}{\sum \frac{\pi_j}{1-\lambda_j(t)}}\left(\frac{\pi_i}{1-\lambda_i(t')} + \left(\sum \frac{\pi_j}{1-\lambda_j(t)} - \sum \frac{\pi_j}{1-\lambda_j(t')}\right)\pi_i(t')\right) \\
=& \frac{\frac{\pi_i}{1-\lambda_i(t')}}{\sum \frac{\pi_j}{1-\lambda_j(t)}} + \pi_i(t') - \frac{\frac{\pi_i}{1-\lambda_i(t')}}{\sum \frac{\pi_j}{1-\lambda_j(t)}} \\
=& \pi_i(t')
\end{aligned} \tag{B.4}$$

Now we verify the dynamic consistency of π_i. $\forall r > s > t$:

$$\begin{aligned}
&\sum_j \pi_j(s|\mu_i(t))\pi_k(r|\mu_j(s)) \\
=&\sum_j \frac{\lambda_i(t)-\lambda_i(s)}{1-\lambda_i(s)}\pi_j(s)\pi_k(r|\mu_j(s)) + \frac{1-\lambda_i(t)}{1-\lambda_i(s)}\pi_k(r|\mu_i(s)) \\
=&\frac{\lambda_i(t)-\lambda_i(s)}{1-\lambda_i(s)}\pi_k(r) + \frac{1-\lambda_i(t)}{1-\lambda_i(s)}\frac{\lambda_i(s)-\lambda_i(r)}{1-\lambda_i(r)}\pi_k(r) + \mathbf{1}_{k=i}\frac{1-\lambda_i(t)}{1-\lambda_i(s)}\frac{1-\lambda_i(s)}{1-\lambda_i(r)} \\
=&\frac{\lambda_i(t)-\lambda_i(r)}{1-\lambda_i(r)}\pi_k(r) + \mathbf{1}_{k=i}\frac{1-\lambda_i(t)}{1-\lambda_i(r)} \\
=&\pi_k(r|\mu_i(t))
\end{aligned} \tag{B.5}$$

The second equality is from Equation (B.4). Now define the stochastic process $\langle\mu_t\rangle$. First, I complete the definition of $\mu_i(t)$ and $\pi_i(t)$. Let $\mu(0) = \mu$, $\pi_i(t|\mu(0)) = \pi_i(t)$. $\forall t > 1$, define $\mu_i(t) = \nu_i$, $\pi_j(t|\mu_i(1)) = \mathbf{1}_{i=j}$. Define $S_i = \{\mu_i(t)|t \in (0,1]\}$. Since ν_i are distinct, S_i are disjoint sets. Since $\lambda_i(t)$ is strictly decreasing, $\mu_i(t)$ is a one-to-one map from $(0,1]$ to S_i. Let $S = (\bigcup S_i)\bigcup\{\mu\}$. Define: $\tau : S \to [0,1]$:

$$\tau(\nu) = \begin{cases} \mu_i(t)^{-1}(\nu) & \text{if } \mu \in S_i \\ 0 & \text{if } \nu = \mu \end{cases}$$

Now we can define the Markov transition kernel of $\langle\mu_t\rangle$: $\forall x, y \in S$, $t \in \mathbb{R}^+$,

$$P_t(x,y) = \sum_i \mathbf{1}_{y=\mu_i(\tau(x)+t)}\pi_i(\tau(x)+t|x)$$

We verify the Chapman-Kolmogorov equation: $\forall x, z \in S$, $t, s \in \mathbb{R}^+$:

- If $\tau(x) + t + s \leqslant 1$, then:

$$\int P_t(x,y)P_s(y,z)\mathrm{d}y = \sum_{i,j} \mathbf{1}_{z=\mu_i(\tau(\mu_j(\tau(x)+t))+s)}\pi_j(\tau(x)+t|x)\pi_i\big(\tau(\mu_j(\tau(x)+t))+s|\mu_j(\tau(x)+t)\big)$$

$$
\begin{aligned}
&= \sum_{i,j} \mathbf{1}_{z=\mu_i(\tau(x)+t+s)} \pi_j(\tau(x)+t|x)\pi_i(\tau(x)+t+s|\mu_j(\tau(x)+t)) \\
&= \sum_{i} \mathbf{1}_{z=\mu_i(\tau(x)+s+t)} \pi_i(\tau(x)+t+s|x) \\
&= P_{t+s}(x,z)
\end{aligned}
$$

The second equality is from definition of τ. The third equality is from Equation (B.5).

- If $\tau(x)+t=1$, then:

$$
\begin{aligned}
\int P_t(x,y)P_s(y,z)\mathrm{d}y &= \sum_{i,j} \mathbf{1}_{z=\mu_i(\tau(\mu_j(\tau(x)+t))+s)} \pi_j(\tau(x)+t|x)\pi_i\big(\tau(\mu_j(\tau(x)+t))+s|\mu_j(\tau(x)+t)\big) \\
&= \sum_{i,j} \mathbf{1}_{z=\nu_i} \pi_j(1|x)\pi_i(1+s|\nu_j) \\
&= \sum_{i} \mathbf{1}_{z=\nu_i} \pi_i(1|x) \\
&= \sum_{i} \mathbf{1}_{z=\mu_i(\tau(x)+t+s)} \pi_i(\tau(x)+t+s|x) \\
&= P_{t+s}(x,z)
\end{aligned}
$$

- If $\tau(x)=1$, then the C-K equation is trivially satisfied:

$$
\int P_t(x,y)P_s(y,z)\mathrm{d}y = \sum_{i,j} \mathbf{1}_{z=\nu_i}\mathbf{1}_{\nu_i=\nu_j}\mathbf{1}_{\nu_j=x} = \mathbf{1}_{z=x} = P_{t+s}(x,z)
$$

- Now for any general case $\tau(x)<1$ and $\tau(x)+t+s>1$, we can add 1, and apply the C-K equation in the last two cases jointly to establish the C-K equation in the general case.

Since we verified the C-K equation for Markov transition kernel $P_t(\cdot,\cdot)$, it is easy to see that $\forall t_1, \ldots, t_n$, the measure given by:

$$P_{t_1,\ldots,t_n}(x_1, \ldots, x_n) = P_{t_1}(\mu, x_1) \prod P_{t_{i+1}-t_i}(x_i, x_{i+1})$$

satisfies the conditions in Daniell-Kolmogorov theorem (see Dellacherie and Meyer (1979)). Then, a simple corollary of Daniell-Kolmogorov theorem states that there exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and stochastic process $\langle \mu_t \rangle$ such that any finite dimensional marginal distribution of $\langle \mu_t \rangle$ is given by P. Now Equation (B.3) implies $\langle \mu_t \rangle$ is a martingale and the C-K equation implies that $\langle \mu_t \rangle$ is Markovian.

What remains to be verified is the third property of Lemma B.3. $\forall t_1, t_2 \in [0,1]$ and $t_1 < t_2$, $\forall \mu_{t_1} \in \{\mu_i(t_1)\}$,

$$\begin{aligned}
E[H(\mu_{t_1}) - H(\mu_{t_2})|\mathcal{F}_{t_1}] =& H(\mu_{t_1}) - E[H(\mu_{t_2})|\mu_{t_1}] \\
=& H(\mu_{t_1}) - \int P_{t_2-t_1}(\mu_{t_1}, \mu_{t_2}) H(\mu_{t_2}) \mathrm{d}y \\
=& H(\mu_{t_1}) - \sum_i \pi_i(\tau(\mu_{t_1}) + t_2 - t_1|\mu_{t_1}) H(\mu_i(\tau(\mu_{t_1}) + t_2 - t_1)) \\
=& H(\mu_{t_1}) - \sum_i \pi_i(t_2|\mu_{t_1}) H(\mu_i(t_2)) \\
=& H(\mu_{t_1}) - \sum_i \pi_i(t_2|\mu_{t_1}) \left(\sum_j \pi_j(1|\mu_i(t_2)) H(\nu_j) \right) \\
& - \sum_i \pi_i(t_2|\mu_{t_1}) \left(H(\mu_i(t_2)) - \sum_j \pi_j(1|\mu_i(t_2)) H(\nu_j) \right) \\
=& \left(H(\mu_{t_1}) - \sum_j \pi_j(1|\mu_{t_1}) H(\nu_j) \right) - \sum_i \pi_i(t_2|\mu_{t_1}) Q_i(\lambda_i(t_2)) \\
=& (1 - t_1)M - \sum_i \pi_i(t_2|\mu_{t_1})(1 - t_2)M \\
=& (t_2 - t_1)(H(\mu_0) - E[H(\mu_1)])
\end{aligned}$$

■

Lemma B.3. *$\forall \mu \in \Delta(X)$, $\pi \in \Delta^2(X)$ and $\int \pi(\nu)\mathrm{d}\nu = \mu$. Then there exists probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and stochastic process $\langle \mu_t \rangle$ s.t.:*

1. *$\langle \mu_t \rangle$ is a martingale.*
2. *$\mu_0 = \mu$, $\mu_1 \sim \pi$.*
3. *$\forall t_1, t_2 \in [0,1]$ and $t_1 < t_2$, $E[H(\mu_{t_1}) - H(\mu_{t_2}) | \mathcal{F}_{t_1}] = (t_2 - t_1)E[H(\mu_0) - H(\mu_1)]$.*

Proof. If $|\mathrm{Supp}(\pi)|$ is finite, the Lemma B.3 is identical to Lemma B.2 and the proof is done. Now I discuss the general case where $\mathrm{Supp}(\pi)$ is an infinite set. Let $M = \int \pi(\nu)(H(\mu) - H(\nu))\mathrm{d}\nu$.

Step 1. Discretizing $\Delta(X)$. Since $H(\mu)$ is a continuous function on $\Delta(X)$, by Heine-Cantor theorem $H(\mu)$ is uniformly continuous. Pick $\varepsilon_k = \frac{M}{2^k}$ and let δ_k be corresponding continuity parameter for ε_k. Discretize $\Delta(X)$ into a set of nested grids with grid size $d_k \leqslant \delta_k$. Let D^k_ν be each d_k-cube containing μ. Then $\forall \mu \in \Delta(X)$, $\forall \pi' \in \Delta(D^k_\mu)$:

$$\int \pi'(\nu)(H(\mu) - H(\nu)) \leqslant \varepsilon_k$$

Step 2. Index all d_1-cubes as $\left\{D^1_{i_1}\right\}_{i_1 \in I_1}$. $\forall i_1 \in I_1$, define:

$$\begin{cases} \mu^1_{i_1} = \dfrac{\int_{\nu \in D^1_{i_1}} \nu \pi(\nu)\mathrm{d}\nu}{\int_{\nu \in D^1_{i_1}} \pi(\nu)\mathrm{d}\nu} \\ \pi^1_{i_1}(\nu) = \dfrac{\mathbf{1}_{\nu \in D^1_{i_1}} \pi(\nu)}{\int_{\nu \in D^1_{i_1}} \pi(\nu)\mathrm{d}\nu} \\ q^1_{i_1} = \displaystyle\int_{\nu \in D^1_{i_1}} \pi(\nu)\mathrm{d}\nu \end{cases} \quad \text{and} \quad \begin{cases} M^1_{i_1} = \displaystyle\int \pi^1_{i_1}(\nu)(H(\mu^1_{i_1}) - H(\nu))\mathrm{d}\nu \\ M^0 = \displaystyle\sum_{i_1} q^1_{i_1}\left(H(\mu) - H(\mu^1_{i_1})\right) \end{cases}$$

It is easy to verify that:

$$
\begin{cases}
\sum q_{i_1}^1 \mu_{i_1}^1 = \mu \\
\int \pi_{i_1}^1 \nu \mathrm{d}\nu = \mu_{i_1}^1 \\
\sum q_{i_1}^1 \pi_{i_1}^1(\nu) = \pi(\nu) \\
M_{i_1}^1 \leqslant \varepsilon_1 \\
M^1 + \sum q_{i_1}^1 M_{i_1}^1 = M
\end{cases}
$$

Now consider distribution $(q_{i_1}^1, \mu_{i_1}^1)$. Let $x_{i_1} = (0, \ldots, 0, 1, 0, \ldots, 0)$ where only i_1th element is 1. Then $\{x_{i_1}\}$ is an orthogonal normal base of $\Delta(I_1)$. Let $x_\mu = (q_1^1, \ldots, q_{I_1}^1)$. Then it is easy to see than $x_\mu \in \Delta(N)$. $\forall \lambda \in [0,1]$, define $x_{i_1,\lambda} = \lambda x_\mu + (1-\lambda) x_{i_1}$.

Define linear map $f : \Delta(I_1) \to \Delta(X)$ as $f(x) = \sum_{i=1}^{I_1} x^{i_1} \mu_{i_1}^1$ (x^{i_1} is i_1th coordinate of vector x). Then consider:

$$
Q_{i_1}(\lambda) = \sum_{j=1}^{I_1} x_{i_1,\lambda}^j \left(H(f(x_{i_1,\lambda})) - H(\mu_j^1) + M_j^1\right)
$$

Now consider properties of Q_{i_1}. First of all, since H is continuous and f is linear, $Q_{i_1}(\lambda)$ is continuous in λ. Second, $Q_{i_1}(0) = M_{i_1}^1$ and $Q_{i_1}(1) = M$. Since $M > \varepsilon_1 \geqslant M_{i_1}^1$, by intermediate value theorem there exists λ_{i_1} s.t. $Q_{i_1}(\lambda_{i_1}) = \varepsilon_1$. Now define:

$$
\begin{cases}
\widetilde{\mu}_{i_1}^1 = f(x_{i_1,\lambda_{i_1}}) \\
\widetilde{q}_{i_1}^1 = \dfrac{\frac{q_{i_1}^1}{1-\lambda_{i_1}}}{\sum_{j=1}^{I_1} \frac{\pi_j^1}{1-\lambda_j}} \\
\widetilde{\pi}_{i_1}^1(\nu) = \sum_{j=1}^{I_1} x_{i_1,\lambda_{i_1}}^j \pi_j^1(\nu)
\end{cases}
\quad \text{and} \quad
\begin{cases}
\widetilde{M}_{i_1}^1 = \int \widetilde{\pi}_{i_1}^1(\nu)\left(H(\widetilde{\mu}_{i_1}^1) - H(\nu)\right)\mathrm{d}\nu \\
\widetilde{M}^0 = \sum_{I_1} \widetilde{q}_{i_1}^1 \left(H(\mu) - H(\widetilde{\mu}_{i_1}^1)\right)
\end{cases}
$$

It is easy to verify that:

$$\begin{cases} \sum \tilde{q}^1_{i_1} \tilde{\mu}^1_{i_1} = \mu \\ \int \tilde{\pi}^1_{i_1} \nu \mathrm{d}\nu = \tilde{\mu}^1_{i_1} \\ \sum \tilde{q}^1_{i_1} \tilde{\pi}^1_{i_1}(\nu) = \pi(\nu) \\ \widetilde{M}^1_{i_1} \equiv \frac{M}{2}, \ \widetilde{M}^0 = \frac{M}{2} \end{cases}$$

Step 3. Recursively apply *step 2*. Suppose I have defined a discrete time stochastic process for $i \in \{0, \ldots, k\}$ satisfying $\mu^0 = \mu$, $\mu^k \sim \pi$ and:

$$\begin{cases} \left|\mathrm{Supp}(\mu^i)\middle|\mathcal{F}_i\right| \leqslant I_i < \infty \ \forall i < k \\ E\left[\mu^i\middle|\mathcal{F}_j\right] = \mu^i \ \forall j < i \\ E\left[H(\mu^j) - H(\mu^i)\middle|\mathcal{F}_j\right] = \sum_{l=j}^{i-1} \frac{M}{2^{l+1}} \\ E\left[H(\mu^k) - H(\mu^i)\middle|\mathcal{F}_i\right] = \frac{M}{2^i} \end{cases} \tag{B.6}$$

Noticing that I have verified that $\left(\mu, (\tilde{q}^i_{i_1}, \tilde{\mu}^1_{i_1}), \tilde{\pi}^1_{i_1}\right)$ we find in step one satisfies this condition for $k = 1$. Now we prove that we can construct a discrete time stochastic process with $k + 1$. Define a new process $\langle \mu^i \rangle$ exactly as in the assumption for $i < k$. Now for any sample path in $\mathcal{F}_{k-1}$, applying the procedure in *step 2* to prior μ^{k-1} and distribution of

(μ^k). Then I get $(\tilde{q}^k_{i_k}, \tilde{\mu}^k_{i_k}, \tilde{\pi}^{k_{i_k}})$ satisfying:

$$\begin{cases} \sum \tilde{q}^k_{i_k} \tilde{\mu}^k_{i_k} = \mu^{k-1} \\ \int \tilde{\pi}^k_{i_k}(\nu) = \tilde{\mu}^k_{i_k} \\ \sum \tilde{q}^k_{i_k} \tilde{\pi}^k_{i_k}(\nu) \sim \mu^k | \mathcal{F}_{k-1} \\ \int \tilde{\pi}^k_{i_k}(\nu)(H(\tilde{\mu}^k_{i_k}) - H(\nu)) \mathrm{d}\nu \equiv \frac{M}{2^k} = \sum_{I_K} \tilde{q}^k_{i_k} \Big(H(\mu^{k-1}) - H(\tilde{\mu}^k_{i_k}) \Big), \end{cases} \tag{B.7}$$

In the new process, let $\mu^k | \mathcal{F}_{k-1} \sim (\tilde{q}^k_{i_k}, \tilde{\mu}^k_{i_k})$ and $\mu^{k+1} | \mathcal{F}_{k-1}, \tilde{\mu}^k_{i_k} \sim \tilde{\pi}^k_{i_k}$. Now let us verify Equation (B.6). The first condition is trivially satisfied for $i < k$. If $i = k$, by construction the support size of $\tilde{\mu}^k_{i_i}$ is finite. The second condition is true for $i = k, k+1$ given Equation (B.7)'s first two properties. The third and forth condition are implied the last condition of Equation (B.7). μ^0 is still μ and $\mu^{k+1} \sim \pi$ by third property of Equation (B.7).

Hence, for any positive K, a $\langle \mu^i \rangle_{i \leqslant K}$ is well defined. And by construction, for any $K_1 < K_2$, the two processes have exactly same path distribution for $i < K_1$. So except if I need to explicitly use the distribution of μ^K, otherwise I refer to $\langle \mu^i \rangle$ as an infinite process.

Step 4. Extension to continuous time. Let $T_k = 1 - \frac{1}{2^k}$. The main idea is to define finite dimensional joint distribution at T_k's according to $\langle \mu^k \rangle$. Then within each interval $[T_k, T_{k+1}]$, the process is defined using Lemma B.2. For any sequence of $\mu_0, \dots, \mu_{T_k}$, define:

$$P(\mu_{T_k} = \mu^k | \mu_0, \dots \mu_{T_{k-1}}) = \mathbb{P}(\mu^k \Big| \mu^l_{i_l} = \mu_{T_l}, \ \forall l < k)$$

Now for any $t_1, \dots, t_k$ and $\mu_{t_1}, \dots, \mu_{t_k}$, I define the joint distribution of the sample path. First assume $t_k < 1$. Then there exists a unique sequence of:

$$0 \cdots \underbrace{T_{l_1}, t_1 \cdots t_{m_1}, T_{l_1+1}}_{\text{Interval 1}} \cdots \underbrace{T_{l_2}, t_{m_1+1} \cdots t_{m_2}, T_{l_2+1}}_{\text{Interval 2}} \cdots \underbrace{T_{l_n}, t_{m_{n-1}+1} \cdots t_k}_{\text{Interval n}}$$

Noticing that t_{m_s+1} can be same as $T_{l_{s+1}}$ and t_{m_s} can be same as T_{l_s+1}. Now for any sequence of $(\mu_{T_{l_s}})$, apply Lemma B.2 to prior $\mu_{T_{l_s}}$ and distribution $P(\mu_{T_{l_s+1}}|\mu_0,\dots,\mu_{T_{l_s}})$. Lemma B.2 implies that there exists a space $(\Omega,\mathcal{F},\mathbb{P})$ and $\langle\widetilde{\mu}_t\rangle$ s.t. $\widetilde{\mu}_0=\mu_{T_{l_s}}$ and $\widetilde{\mu}_1\sim\mathbb{P}(\mu_{T_{l_s+1}})$ (the dependence of all terms on $(\mu_{T_{l_s}})$ is omitted for notational simplicity). Define:

$$
\begin{aligned}
&P\Big(\mu_{t_{m_{s-1}+1}},\dots,\mu_{t_{m_s}},\mu_{T_{l_s+1}}\Big|\mu_0,\dots,\mu_{T_{l_s}}\Big)\\
=&\mathbb{P}\Big(\widetilde{\mu}_{2^{l_s}\cdot(t_{m_{s-1}+1}-T_{l_s})}=\mu_{t_{m_{s-1}+1}},\dots,\widetilde{\mu}_{2^{l_s}\cdot(t_{m_s}-T_{l_s})}=\mu_{t_{m_s}},\widetilde{\mu}_1=\mu_{T_{l_s+1}}\Big|\mu_0,\dots,\mu_{T_{l_s}}\Big)
\end{aligned}
$$

Now we can define the finite joint distribution of μ_{t_k}:

$$
\begin{aligned}
&P\big(\mu_0,\dots,\mu_{t_k}\big)\\
=&\int\prod_{s=1}^{n-1}\left[\left(\prod_{j=l_s+1}^{l_{s+1}}P(\mu_{T_j}|\mu_0,\dots,\mu_{T_{j-1}})\right)\cdot P\Big(\mu_{t_{m_s}},\dots,\mu_{t_{m_{s+1}}},\mu_{T_{l_{s+1}+1}}\Big|\mu_0,\dots\mu_{T_{l_{s+1}}}\Big)\right]\mathrm{d}\mu_{T_1},\dots\mu_{T_{l_n+1}}
\end{aligned}
$$

Noticing that by definition of $\langle\mu_k\rangle$, each $P(\mu_{T_j}|\mu_0,\dots,\mu_{T_{j-1}})$ is a probability measure. By definition of $\langle\widetilde{\mu}_t\rangle$, each $P\Big(\mu_{t_{m_s}},\dots,\mu_{t_{m_{s+1}}},\mu_{T_{l_{s+1}+1}}\Big|\mu_0,\dots\mu_{T_{l_{s+1}}}\Big)$ is a joint probability measure. Therefore, $P(\mu_0,\dots,\mu_{t_k})$ is a joint probability measure.

Now consider the case $t_k=1$. Since $t_{k-1}<1$, there must be some finite $T_l>t_{k-1}$. There exists a unique sequence of:

$$
0\cdots\underbrace{T_{l_1},t_1\cdots t_{m_1},T_{l_1+1}}_{\text{Interval 1}}\cdots\underbrace{T_{l_2},t_{m_1+1}\cdots t_{m_2},T_{l_2+1}}_{\text{Interval 2}}\cdots,t_{k-1},T_{l_n+1}\cdots t_k
$$

Pick $K=l_n+2$, for any given sequence of $\mu_0,\dots,\mu_{T_{l_n+1}}$, $\forall S\subset\Delta(X)$ define:

$$
P\Big(S\Big|\mu_0,\dots,\mu_{T_{l_n+1}}\Big)=\mathbb{P}\Big(\mu^K\in S\Big|\mu^l_{i_l}=\mu_{T_l},\forall l<K\Big)
$$

Now we can define the finite joint distribution of μ_{t_k}:

$$\begin{aligned}
&P(\mu_0, \ldots, \mu_{t_k}) \\
=&\int \prod_{s=1}^{n-1} \left[\left(\prod_{j=l_s+1}^{l_{s+1}} P(\mu_{T_j} | \mu_0, \ldots, \mu_{T_{j-1}}) \right) \cdot P\left(\mu_{t_{m_s}}, \ldots, \mu_{t_{m_{s+1}}}, \mu_{T_{l_{s+1}+1}} \middle| \mu_0, \ldots \mu_{T_{l_{s+1}}} \right) \right] \\
&\cdot P\left(S \middle| \mu_0, \ldots, \mu_{T_{l_n+1}}\right) \mathrm{d}\mu_{T_1}, \ldots \mu_{T_{l_n+1}}
\end{aligned}$$

Same as the previous case, each $P(\mu_{T_j} | \mu_0, \ldots, \mu_{T_{j-1}})$ and $P\left(\mu_{t_{m_s}}, \ldots, \mu_{t_{m_{s+1}}}, \mu_{T_{l_{s+1}+1}} \middle| \mu_0, \ldots \mu_{T_{l_{s+1}}} \right)$ are joint probability measures. Moreover, $P(S|\mu_0, \ldots, \mu_{T_{K-1}})$ is a probability measure. Importantly, by definition of $\langle \mu^k \rangle$, for $K_1 < K$, P is defined consistently:

$$P(S|\mu_0, \ldots, \mu_{K_1-1}) = \int P(S|\mu_0, \ldots, \mu_{K-1}) \mathrm{d}\mu_{K_1}, \ldots, \mu_{K-1}$$

Therefore, $P(\mu_0, \ldots, \mu_{t_k})$ is a joint probability measure. To sum up, we defined a finite dimensional joint probability measure satisfying the conditions in Daniell-Kolmogorov theorem. Hence, there exists probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\langle \mu_t \rangle_{t \in [0,1]}$ satisfying all finite distributions.

Step 5. Verify that $\langle \mu_t \rangle$ satisfies Lemma B.3. $\mu_0 = \mu$ is true by construction. $\forall S \subset \Delta(X)$:

$$\mathbb{P}(S) = \mathbb{P}(\mu^1 \in S) = \pi(S)$$

So $\mu_1 \sim \pi$. Now I verify property 1 and property 3. $\forall t_1, t_2 \in [0,1]$ and $t_2 > t_1$,

- *Case 1*. If $\exists k$ s.t. $T_k \leqslant t_1 < t_2 \leqslant T_{k+1}$, then by construction of $\langle \mu_t \rangle$ in $[T_j, T_{j+1}]$, μ_t is Markovian and $E[\mu_{t_2} | \mathcal{F}_{t_1}] = E[\mu_{t_2} | \mu_{t_1}] = \mu_{t_1}$.

$$\begin{aligned}
E[H(\mu_{t_1}) - H(\mu_{t_2}) | \mu_{t_1}] =& 2^k (t_2 - t_1) E[H(\mu_{t_k}) - H(\mu_{t_{k+1}}) | \mathcal{F}_{t_k}] \\
=& (t_2 - t_1) 2^k \cdot \frac{M}{2^k}
\end{aligned}$$

$$=(t_2 - t_1)M$$

- *Case 2*. There exists a unique sequence of:

$$0 \cdots T_j, t_1, T_{j+1} \cdots T_k, t_2, T_{k+1}$$

By construction, for all path on $[0, t_1]$,

$$\begin{aligned}
E\big[\mu_{t_2}|\mathcal{F}_{t_1}\big] =&E\Big[E\big[\mu_{t_2}|\mathcal{F}_{T_k}\big]\Big|\mathcal{F}_{t_1}\Big]\\
=&E\Big[\mu_{T_k}\Big|\mathcal{F}_{t_1}\Big]\\
&\vdots\\
=&E\Big[\mu_{T_{j+1}}|\mathcal{F}_{t_1}\Big]\\
=&\mu_{t_1}
\end{aligned}$$

and

$$\begin{aligned}
&H(\mu_{t_1}) - E\big[H(\mu_{t_2})|\mathcal{F}_{t_1}\big]\\
=&H(\mu_{t_1}) - E\Big[H(\mu_{T_{j+1}})|\mathcal{F}_{t_1}\Big] + E\Big[H(\mu_{T_{j+1}})|\mathcal{F}_{t_1}\Big] - E\big[H(\mu_{t_2})|\mathcal{F}_{t_1}\big]\\
=&2^j(T_{j+1} - t_1) \cdot \frac{M}{2^j} + E\Big[E\Big[H(\mu_{T_{j+1}}) - H(\mu_{t_2})|\mathcal{F}_{T_{j+1}}\Big]\Big|\mathcal{F}_{t_1}\Big]\\
=&(T_{j+1} - t_1)M + E\Big[H(\mu_{T_{j+1}}) - E\Big[H(\mu_{T_{j+2}})|\mathcal{F}_{T_{j+1}}\Big] + E\Big[H(\mu_{T_{j+2}}) - H(\mu_{t_2})|\mathcal{F}_{T_{j+1}}\Big]\Big|\mathcal{F}_{t_1}\Big]\\
=&(T_{j+1} - t_1)M + E\left[\frac{M}{2^{j+2}} + E\Big[H(\mu_{T_{j+2}}) - H(\mu_{t_2})|\mathcal{F}_{T_{j+1}}\Big]\Big|\mathcal{F}_{t_1}\right]\\
&\vdots\\
=&(1 - \frac{1}{2^{j+1}} - t_1)M + \sum_{j+1}^{k-1} \frac{M}{2^{l+1}} + E\Big[E\big[H(\mu_{T_k}) - H(\mu_{t_2})|\mathcal{F}_{T_k}\big]\Big|\mathcal{F}_{t_1}\Big]
\end{aligned}$$

$$=(1-\frac{1}{2^{j+1}}-t_1)M+\sum_{j+1}^{k-1}\frac{M}{2^{l+1}}+\left(t_2-1+\frac{1}{2^k}\right)M$$

$$=(t_2-t_1)M$$

Case 3. $t_2 = 1$. Then there exists k s.t $T_k \leqslant t_1 < T_{k+1}$. By construction, for all path on $[0, t_1]$:

$$\begin{aligned}
E[\mu_1|\mathcal{F}_{t_1}] =& E\big[E\big[\mu_t|\mathcal{F}_{T_{k+1}}\big]|\mathcal{F}_{t_1}\big] \\
=& E\big[\mu_{T_{k+1}}|\mathcal{F}_{t_1}\big] \\
=& \mu_{t_1}
\end{aligned}$$

and

$$\begin{aligned}
& H(\mu_{t_1}) - E[H(\mu_1)|\mathcal{F}_{t_1}] \\
=& H(\mu_{t_1}) - E\big[H(\mu_{T_{k+1}})|\mathcal{F}_{t_1}\big] + E\big[E\big[H(\mu_{t_{K+1}}) - H(\mu_1)|\mathcal{F}_{T_{k+1}}\big]|\mathcal{F}_{t_1}\big] \\
=& \left(1-\frac{1}{2^{k+1}}-t_1\right)M+\frac{M}{2^{k+1}} \\
=& (t_2-t_1)M
\end{aligned}$$

■

B.1.2 Proof of Lemma 1.2

Proof. I break the proof of Lemma 1.2 into three lemmas. Lemma B.4 shows that solving Equation (1.5′) is equivalent to solving Equation (B.8), which reduces the signal structure to be nested, and containing only action as direct signals and continuation signals. Then Lemma B.5 shows that solving Equation (B.8) is equivalent to solving Equation (B.9), which transforms signal process formulation to conditional distribution formulation. Then

Lemma B.6 shows that solving functional equation Equation (B.10) is equivalent to solving sequential problem Equation (B.9) using the standard methodology. Finally, we apply Theorem B.1 to Equation (B.10) to further reduce the dimensionality of strategy space to Equation (1.6). ∎

Lemma B.4 (Reduction of redundency). $\left(\mathcal{S}^t, \mathcal{T}, \mathcal{A}^{\mathcal{T}}\right)$ *solves* Equation (1.5′) *if and only if there exists* $\left(\tilde{\mathcal{S}}^t, \mathcal{T}, \mathcal{A}^{\mathcal{T}}\right)$ *solving :*

$$\sup_{\mathcal{S}^t, \mathcal{T}, \mathcal{A}^{\mathcal{T}}} \sum_{t=0}^{\infty} e^{-\rho dt \cdot t} \Big(\mathbf{P}[\mathcal{T} = t]\big(E\big[u(\mathcal{A}^t, \mathcal{X})|\mathcal{T} = t\big]\big) - \mathbf{P}[\mathcal{T} > t]E\Big[C_{dt}\Big(I\Big(\tilde{\mathcal{S}}^t; \mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}}\Big)\Big)|\mathcal{T} > t\Big]\Big)$$

$$\text{s.t. } \tilde{\mathcal{S}}^t = \begin{cases} s_0 & \text{when } \mathcal{T} < t+1 \\ \mathcal{A}^{t+1} & \text{when } \mathcal{T} = t+1 \\ \mathcal{S}^t & \text{when } \mathcal{T} > t+1 \end{cases} \tag{B.8}$$

What's more, the optimal utility level is same in Equation (1.5′) *and* Equation (B.8).

Proof. Suppose $(\mathcal{S}^t, \mathcal{T}, \mathcal{A}^t)$ is a feasible strategy to Equation (1.5′). I first show that it is WLOO that the DM discards all information after taking an action: take given $\mathcal{T}$ and $\mathcal{A}^t$, let s_0 be a degenerate signal, define signal process $\hat{\mathcal{S}}^t$ as:

$$\hat{\mathcal{S}}^t = \begin{cases} \mathcal{S}^t & \text{when } \mathcal{T} \geqslant t+1 \\ s_0 & \text{when } \mathcal{T} \leqslant t \end{cases}$$

By definition, $\hat{\mathcal{S}}^{\mathbf{t}} = \mathcal{S}^{\mathbf{t}}$ conditional on $\mathcal{T} \geqslant t+1$. Therefore:

$$I\Big(\hat{\mathcal{S}}^t; \mathcal{X}|\hat{\mathcal{S}}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \leqslant t}\Big) = \begin{cases} I\big(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \leqslant t}\big) & \text{when } \mathcal{T} \geqslant t+1 \\ 0 & \text{when } \mathcal{T} \leqslant t \end{cases}$$

$$\mathcal{X} \to \hat{\mathcal{S}}^t \to \mathcal{A}^{t+1} \text{ conditional on } \mathcal{T} = t$$

1. By definition, when $\mathcal{T} \geqslant t+1$, $\widehat{\mathcal{S}}^t = \mathcal{S}^t$. So conditional on $\mathcal{T} = t+1$, $\mathcal{X} \to \mathcal{S}^t \to \mathcal{A}^{t+1}$ implies $\mathcal{X} \to \widehat{\mathcal{S}}^t \to \mathcal{A}^{t+1}$.

2. When $\widehat{\mathcal{S}}^{t-1} = s_0$, $\mathbf{1}_{\mathcal{T}\leqslant t} = 1$. When $\widehat{\mathcal{S}}^{t-1} \neq s_0$:

$$\begin{aligned}
&\text{Prob}\left(\mathcal{T} > t|\mathcal{S}^{t-1}\right) = \text{Prob}\left(\mathcal{T} > t|\mathcal{S}^{t-1}, \mathcal{X}, \mathcal{T} \geqslant t\right)\text{Prob}\left(\mathcal{T} \geqslant t|\mathcal{S}^{t-1}, \mathcal{X}\right) \\
&= \text{Prob}\left(\mathcal{T} > t|\mathcal{S}^{t-1}, \mathcal{X}, \mathcal{T} \geqslant t\right) \\
\Longrightarrow\ &\text{Prob}\left(\mathcal{T} > t|\widehat{\mathcal{S}}^{t-1}, \mathcal{X}\right) = \text{Prob}\left(\mathcal{T} > t|\widehat{\mathcal{S}}^{t-1}\right)
\end{aligned}$$

which is independent to realization of $\mathcal{X}$. So $\mathcal{X} \to \widehat{\mathcal{S}}^{t-1} \to \mathbf{1}_{\mathcal{T}\leqslant t}$. The first equality is by the law of total probability (conditional on $\mathcal{T} \geqslant t$), $\mathcal{X} \to \mathcal{S}^{t-1} \to \mathbf{1}_{\mathcal{T}\leqslant t}$ and when $\widehat{\mathcal{S}}^{t-1} \neq s_0$, $\text{Prob}(\mathcal{T} = t) = 0$. The second equality is by when $\widehat{\mathcal{S}}^{t-1} \neq s_0$, $\text{Prob}(\mathcal{T} \geqslant t) = 1$.

3. Total information cost:

$$\begin{aligned}
&E\left[\sum_{t=0}^{\infty} e^{-\rho dt\cdot t} C_{dt}\left(I\left(\widehat{\mathcal{S}}^t; \mathcal{X}|\widehat{\mathcal{S}}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}\right)\right)\right] = E\left[\sum_{t=0}^{\mathcal{T}-1} e^{-\rho dt\cdot t} C_{dt}\left(I\left(\widehat{\mathcal{S}}^t; \mathcal{X}|\widehat{\mathcal{S}}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}\right)\right)\right] \\
=&E\left[\sum_{t=0}^{\mathcal{T}-1} e^{-\rho dt\cdot t} C_{dt}(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}))\right] \leqslant E\left[\sum_{t=0}^{\infty} e^{-\rho dt\cdot t} C_{dt}(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}))\right]
\end{aligned}$$

The first equality is by $\widehat{\mathcal{S}}^t$ being degenerate when $t \geqslant \mathcal{T}$. The second equality is from $\widehat{\mathcal{S}}^t = \mathcal{S}^t$ when $\mathcal{T} > t$. Therefore, $\left(\widehat{\mathcal{S}}^t, \mathcal{A}^t, \mathcal{T}\right)$ is a feasible strategy dominating $(\mathcal{S}^t, \mathcal{A}^t, \mathcal{T})$. Now we define $\widetilde{\mathcal{S}}^t$:

$$\widetilde{\mathcal{S}}^t = \begin{cases} s_0 & \text{when } \mathcal{T} < t+1 \\ \mathcal{A}^{t+1} & \text{when } \mathcal{T} = t+1 \\ \widehat{\mathcal{S}}^t & \text{when } \mathcal{T} > t+1 \end{cases}$$

Initial information $\tilde{\mathcal{S}}^{-1}$ is defined as a degenerate(uninformative) signal and induced belief is the prior. $\tilde{\mathcal{S}}^t$ replaces the signal defined in $\hat{\mathcal{S}}^t$ by a direct signal that suggests the corresponding action profile in next period when $\mathcal{T} = t+1$. Verify that the $\tilde{\mathcal{S}}^t$ satisfies the information processing constraints in Equation (1.5′) and improves utility:

1. When $\tilde{\mathcal{S}}^{t-1} \in \{s_0\} \bigcup A$, it's for sure that $\mathcal{T} \leqslant t$. Otherwise, $\mathcal{T} > t$. Therefore $\mathbf{1}_{\mathcal{T}\leqslant t}$ is a direct garbling of $\tilde{\mathcal{S}}^{t-1}$. So $\mathcal{X} \to \tilde{\mathcal{S}}^{t-1} \to \mathbf{1}_{\mathcal{T}\leqslant t}$.
2. When $\mathcal{T} = t$, $\mathcal{A}^t = \tilde{\mathcal{S}}^{t-1}$. Therefore $\mathcal{S} \to \hat{\mathcal{S}}^{t-1} \to \mathcal{A}^t$ implies $\mathcal{X} \to \tilde{\mathcal{S}}^{t-1} \to \mathcal{A}^t$ conditional on $\mathcal{T} = t$.
3. Information measure associated with $\left(\tilde{\mathcal{S}}^t, \mathcal{A}^t, \mathcal{T}\right)$ when $\mathcal{T} > t$:

$$
\begin{aligned}
&I\left(\tilde{\mathcal{S}}^t; \mathcal{X} \middle| \tilde{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > t\right)\\
=&\mathbf{1}_{\mathcal{T}=t+1} I\left(\mathcal{A}^{t+1}; \mathcal{X} \middle| \tilde{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} = t+1\right) + \mathbf{1}_{\mathcal{T}>t+1} I\left(\tilde{\mathcal{S}}^t; \mathcal{X} \middle| \tilde{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > t+1\right)\\
=&\mathbf{1}_{\mathcal{T}=t+1} I\left(\mathcal{A}^{t+1}; \mathcal{X} \middle| \hat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} = t+1\right) + \mathbf{1}_{\mathcal{T}>t+1} I\left(\hat{\mathcal{S}}^t; \mathcal{X} \middle| \hat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > t+1\right)\\
\leqslant&\mathbf{1}_{\mathcal{T}=t+1} I\left(\hat{\mathcal{S}}^t; \mathcal{X} \middle| \hat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} = t+1\right) + \mathbf{1}_{\mathcal{T}>t+1} I\left(\hat{\mathcal{S}}^t; \mathcal{X} \middle| \hat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > t+1\right)\\
=&I\left(\hat{\mathcal{S}}^t; \mathcal{X} \middle| \hat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > 1\right)
\end{aligned}
$$

First equality is simply rewriting two possible cases of $\mathcal{T}$. Second equality is from definition of $\tilde{\mathcal{S}}^t$ when $\mathcal{T} > t+1$. The inequality is from $\mathcal{X} \to \hat{\mathcal{S}}^t \to \mathcal{A}^{t+1}$ conditional on $\mathcal{T} = t+1$. Therefore, $\left(\tilde{\mathcal{S}}^t, \mathcal{T}, \mathcal{A}^t\right)$ dominates the original solution in Equation (1.5′) by achieving same action profile at lower costs. $\left(\tilde{\mathcal{S}}^t, \mathcal{T}, \mathcal{A}^t\right)$ is a feasible solution to Equation (B.8). Therefore solving Equation (B.8) yields a weakly higher utility than Equation (1.5′). What remains to be proved is that any $\left(\tilde{\mathcal{S}}^t, \mathcal{T}, \mathcal{A}^t\right)$ feasible in Equation (B.8) can be dominated by some strategy feasible in Equation (1.5′). It's not hard to see that the strategy is feasible in Equation (1.5′). Finally we show that the two formulation gives same utility:

$$
\begin{aligned}
&E\left[e^{-\rho dt\cdot\mathcal{T}}E[u(\mathcal{A}^{\mathcal{T}},\mathcal{X})]-\sum_{t=0}^{\infty}e^{-\rho dt\cdot t}C_{dt}\left(I\left(\widetilde{\mathcal{S}}^t;\mathcal{X}|\widetilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right)\right]\\
=&\sum_{t=0}^{\infty}e^{-\rho dt\cdot t}\left(\mathbf{P}[\mathcal{T}=t]\left(E\left[u(\mathcal{A}^t,\mathcal{X})|\mathcal{T}=t\right]\right)-E\left[C_{dt}\left(I\left(\widetilde{\mathcal{S}}^t;\mathcal{X}|\widetilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right)\right]\right)\\
=&\sum_{t=0}^{\infty}e^{-\rho dt\cdot t}\left(\mathbf{P}[\mathcal{T}=t]\left(E\left[u(\mathcal{A}^t,\mathcal{X})|\mathcal{T}=t\right]\right)-\mathbf{P}[\mathcal{T}>t]E\left[C_{dt}\left(I\left(\widetilde{\mathcal{S}}^t;\mathcal{X}|\widetilde{\mathcal{S}}^{\mathbf{t-1}}\right)\right)\Big|\mathcal{T}>t\right]\right)
\end{aligned}
$$

First equality is from rewriting the utility part conditional on decision time $\mathcal{T}=t$. Second equality is from rewriting the information cost part conditional on decision time $\mathcal{T}\leqslant t$ and $\mathcal{T}>t$. Therefore, Equation (1.5′) is equivalent to Equation (B.8). ∎

Lemma B.5 (Tranformation of space)**.** *With Assumption A satisfied, $(\mathcal{S}^t,\mathcal{T},\mathcal{A}^{\mathcal{T}})$ solves Equation (1.5′) if and only if there exists $p^t(\mu^{t+1}|\mu^t):\Delta X\mapsto\Delta^2 X$ and $q_s^t(\mu^t):\Delta X\mapsto[0,1]$ solving:*

$$
\begin{aligned}
\sup_{(p^t,q_s^t)}\ &\sum_{t=0}^{\infty}e^{-\rho dt\cdot t}\int_{\Delta X}\Bigg[\left(\max_a\sum_j u(a,x_j)\cdot\mu_j\right)q_s^t(\mu^t)\\
&-C_{dt}\left(H(\mu^t)-\int_{\Delta X}H(\mu^{t+1})p^t(\mu^{t+1}|\mu^t)\mathrm{d}\mu^{t+1}\right)(1-q_s^t(\mu^t))\Bigg]\\
&\left(\int_{\Delta X^{t-1}}\prod_{\tau=0}^{t-1}p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))\mathrm{d}\mu^1\ldots\mu^{t-1}\right)\mathrm{d}\mu^t\\
\text{s.t. }&\int_{\Delta X}\mu p^t(\mu|\mu^t)\mathrm{d}\mu=\mu^t
\end{aligned}
\tag{B.9}
$$

What's more, the optimal utility level is same in Equation (1.5′) and Equation (B.9).

Proof. Let $p^t(\cdot|\mu^t)$ be the distribution of posteriors generated by $\widetilde{\mathcal{S}}^t|_{\mathcal{T}>t,\widetilde{\mathcal{S}}^{t-1}=\widetilde{S}^{t-1}}$, where μ^t is posterior belief associated with signal $\widetilde{S}^{\mathbf{t-1}}$. Let $q_s^t(\mu^t)=\mathbf{P}\left[\mathcal{T}=t|\widetilde{\mathcal{S}}^{\mathbf{t-1}}=\widetilde{S}^{\mathbf{t-1}},\mathcal{T}\geqslant t\right]$. Now we can explicitly represent the distribution of $\widetilde{\mathcal{S}},\mathcal{T},\mathcal{A}$ with the conditional distributions. First, $\mathbf{P}[\mathcal{T}=t]$ and $\mathbf{P}[\mathcal{T}>t]$ can be calculated by integrating $q_s^t(\mu^t)$:

$$\begin{aligned}
\mathbf{P}[\mathcal{T}=t] =& E\Big[\mathbf{P}[\mathcal{T}=t|\widetilde{\mathcal{S}}^{-1}]\Big] \\
=& E\Big[\mathbf{P}\Big[\mathcal{T}=t|\widetilde{\mathcal{S}}^{-1},\mathcal{T}>0\Big]\mathbf{P}\Big[\mathcal{T}>0|\widetilde{\mathcal{S}}^{-1}\Big]\Big] \\
=& (1-q_s^0(\mu^0))\mathbf{P}\big[\mathcal{T}=t|\mathcal{T}>0\big] \\
=& (1-q_s^0(\mu^0))E\Big[\mathbf{P}\Big[\mathcal{T}=t|\mathcal{T}>0,\widetilde{\mathcal{S}}^0\Big]\Big] \\
=& (1-q_s^0(\mu^0))\int \mathbf{P}\Big[\mathcal{T}=t|\mathcal{T}\geqslant 1,\widetilde{\mathcal{S}}^0\Big]p^0(\mu^1|\mu^0)\mathrm{d}\mu^1 \\
=& (1-q_s^0(\mu^0))\int \mathbf{P}\Big[\mathcal{T}=t|\mathcal{T}>1,\widetilde{\mathcal{S}}^0\Big]\mathbf{P}\Big[\mathcal{T}>1|\mathcal{T}\geqslant 1,\widetilde{\mathcal{S}}^0\Big]p^0(\mu^1|\mu^0)\mathrm{d}\mu^1 \\
=& (1-q_s^0(\mu^0))\int E\Big[\mathbf{P}\Big[\mathcal{T}=t|\mathcal{T}>1,\widetilde{\mathcal{S}}^1\Big]|\mu^1\Big](1-q_s^1(\mu^1))p^0(\mu^1|\mu^0)\mathrm{d}\mu^1 \\
=& \cdots \\
=& \int\prod_{\tau=0}^{t-1}p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))q_s^t(\mu^t)\mathrm{d}\mu^1\ldots\mu^t
\end{aligned}$$

Similarly, we can get:

$$\mathbf{P}[\mathcal{T}>t]=\int\prod_{\tau=0}^{t-1}p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))(1-q_s^t(\mu^t))\mathrm{d}\mu^1\ldots\mu^t$$

Then we can calculate the joint distribution of $\mathcal{T}$ and μ^t:

$$\begin{cases}
\mathbf{P}\big[\mathcal{T}=t,\mu^t=\nu\big]=\int\prod_{\tau=0}^{t-1}p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))q_s^t(\mu^t)\mathrm{d}\mu^1\ldots\mu^{t-1} \\
\mathbf{P}\big[\mathcal{T}>t,\mu^t=\nu\big]=\int\prod_{\tau=0}^{t-1}p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))(1-q_s^t(\mu^t))\mathrm{d}\mu^1\ldots\mu^{t-1}
\end{cases}$$

Therefore:

$$\begin{cases} \mathcal{A}^t\big|_{\mathcal{T}=t} \sim \dfrac{\int \prod_{\tau=0}^{t-1} p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))\mathrm{d}\mu^1 \ldots \mu^{t-1} q_s^t(\mu^t)}{\int \prod_{\tau=0}^{t-1} p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))q_s^t(\mu^t)\mathrm{d}\mu^1 \ldots \mu^t} \\ \widetilde{\mathcal{S}}^t\big|_{\mathcal{T}>t} \sim \dfrac{\int \prod_{\tau=0}^{t-1} p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))\mathrm{d}\mu^1 \ldots \mu^{t-1}(1-q_s^t(\mu^t))}{\int \prod_{\tau=0}^{t-1} p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau)(1-q_s^t(\mu^t))\mathrm{d}\mu^1 \ldots \mu^t} \end{cases}$$

This implies:

$$\begin{aligned} &\mathbf{P}[\mathcal{T}=t]E\Big[u(\mathcal{A}^t,\mathcal{X})\big|\widetilde{\mathcal{S}}^{\mathbf{t-1}},\mathcal{T}=t\Big] \\ =&\int_{\Delta X} \max_a \sum_j u(a,x_j)\mu_j^t \int_{\Delta X^{t-1}} \prod_{\tau=0}^{t-1} p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))q_s^t(\mu^t)\mathrm{d}\mu^1 \ldots \mu^{t-1}\mathrm{d}\mu^t \\ &\mathbf{P}[\mathcal{T}>t]E\Big[C_{dt}\Big(I\Big(\widetilde{\mathcal{S}}^t;\mathcal{X}\big|\widetilde{\mathcal{S}}^{\mathbf{t-1}}\Big)\Big)\big|\mathcal{T}>t\Big] \\ =&\int_{\Delta X} C_{dt}\left(H(\mu^t)-\int_{\Delta X} H(\mu^{t+1})p^t(\mu^{t+1}|\mu^t)\mathrm{d}\mu^{t+1}\right) \\ &\times \int \prod_{\tau=0}^{t-1} p^\tau(\mu^{\tau+1}|\mu^\tau)(1-q_s^\tau(\mu^\tau))(1-q_s^t(\mu^t))\mathrm{d}\mu^1 \ldots \mu^{t-1}\mathrm{d}\mu^t \end{aligned}$$

To sum up, we showed that starting from $\widetilde{\mathcal{S}},\mathcal{T},\mathcal{A}$ solving Equation (B.8), we can construct p^t, q_s^t such that the value of Equation (B.8) is achieved in Equation (B.9). Next, we start from (p^t, q^t) solving Equation (B.9). We can easily define $\mathcal{T}$: $\mathcal{T}\big|_{\mathcal{T}\geqslant t,\mu^t} \sim B(q_s^t(\mu^t))$ conditionally independent across all t,μ^t. $\widetilde{\mathcal{S}}^t\big|_{\mathcal{T}>t,\mu^t} \sim p^t(\cdot|\mu^t)$, $\mathcal{A}^t\big|_{\mathcal{T}=t,\mu^t} = \arg\max \sum u(a,x_j)\mu_j^t$. Therefore, the previous calculation shows that the value of Equation (B.9) is also achieved in Equation (B.8). Combining with the previous result, we conclude that Equation (B.8) and Equation (B.9) are equivalent in the sense that $\left(\widetilde{\mathcal{S}},\mathcal{T},\mathcal{A}\right)$ solves Equation (B.8) if and only if the corresponding (p^t, q_s^t) solves Equation (B.9). ■

Lemma B.6 (Recursive representation). *$V_{dt}(\mu)$ is the optimal utility level solving Equation* (B.9) *given initial belief μ if and only if $V_{dt}(\mu)$ satisfies the following functional equation:*

$$V_{dt}(\mu) = \max\left\{\max_a E[u(a,x)|\mu],\ \sup_{p\in\Delta^2 X} e^{-\rho dt}\int_{\Delta X} V_{dt}(\mu)p(\mu)\mathrm{d}\mu - C_{dt}\left(H(\mu) - \int_{\Delta X} H(\nu)p(\nu)\mathrm{d}\nu\right)\right\} \tag{B.10}$$

$$s.t. \int_{\Delta X} \nu p(\nu)\mathrm{d}\nu = \mu$$

Proof. We first derive the recursive representation of Equation (B.9). Consider the following functional equation:

$$\begin{aligned} V_{dt}(\mu) = \sup_{q_s(\mu),p(\cdot|\mu)} & q_s(\mu)\Big(\max_a \sum u(a,x_j)\mu_j\Big) \\ & + (1-q_s(\mu))\left[\int_{\Delta X} V_{dt}(\nu)p(\nu|\mu)\mathrm{d}\nu - C_{dt}\left(H(\mu) - \int_{\Delta X} H(\nu)p(\nu|\mu)\mathrm{d}\nu\right)\right] \\ \text{s.t. } & \int_{\Delta X} \nu p(\nu|\mu)\mathrm{d}\nu = \mu \end{aligned}$$

Since RHS is linear in $q_s(\mu)$, it will be WLOG that we only consider boundary solution $q_s(\mu) \in \{0,1\}$. Therefore, it is be exactly the same as Equation (B.10).

Now consider the equivalence between the sequential problem and the recursive problem. By assumption $E[u(a,x)|\mu]$ is bounded above by $\max_{a,x} u(a,x)$. Therefore, $e^{-\rho dt\cdot t}E[u(a,x)|\mu]$ is uniformly (for all choice of μ, a) converging to zero when $t \to \infty$. Then $V_{dt}(\mu)$ is the solution of Equation (B.9) by the standard theory of dynamic programming. ■

B.1.3 Convergence

I first prove two useful lemmas. Lemma B.7 shows that optimal strategy has informativeness of signal in each period of same order of dt. Lemma B.8 shows that there exists a unique limit of V_{dt} in L_∞ norm.

B.1.3.1 Bounded flow cost

Lemma B.7 (Bounded flow cost). *With Assumption 1.2 satisfied, there* $\exists \Delta \in \mathbb{R}^+$ *s.t.* $I_{dt}^*(\mu) \leqslant \Delta dt$. $\forall \mu, dt$. *Where* $I_{dt}^*(\mu) = \sum p_i(H(\mu) - H(\nu_i))$ *for optimal* (p_i, ν_i) *in Equation (1.6)*

Proof. $\forall$ (p_i, ν_i) which solves Equation (1.6), assume the value is $V_{dt}(\mu)$ and $I_{dt}^*(\mu) = \sum p_i(H(\mu) - H(\nu_i))$. Now for $I < I_{dt}^*$, consider a different information acquisition strategy:

- At prior μ, use the following information structure:

$$\begin{cases} \mu_i' = \nu_i & \text{with probability } \frac{I}{I_{dt}^*} p_i \\ \mu_0' = \mu & \text{with probability } 1 - \frac{I}{I_{dt}^*} \end{cases}$$

  This information structure mixes uninformative signal into (p_i, ν_i) with probability $1 - \frac{I}{I_{dt}^*} > 0$. It is Bayes plausible by definition.
- At any posterior other than μ, follow the original strategy.

Now let's calculate the expected utility of this strategy. The utility gain from experimentation is:

$$\begin{aligned}
V'(\mu) =& e^{-\rho dt} \frac{I}{I_{dt}^*} \sum p_i V(\nu_i) + e^{-\rho dt} \left(1 - \frac{I}{I_{dt}^*}\right) V'(\mu) \\
=& \left(\sum p_i V(\nu_i)\right) \cdot \frac{e^{-\rho dt} \frac{I}{I_{dt}^*}}{1 - e^{-\rho dt}\left(1 - \frac{I}{I_{dt}^*}\right)} \\
\implies \sum p_i V(\nu_i) - e^{\rho dt} V'(\mu) =& \left(\sum p_i V(\nu_i)\right) \cdot \frac{\left(1 - e^{-\rho dt}\right)\left(1 - \frac{I}{I_{dt}^*}\right)}{1 - e^{-\rho dt}\left(1 - \frac{I}{I_{dt}^*}\right)} \\
\leqslant& \max v \cdot \frac{\left(1 - e^{-\rho dt}\right)\left(1 - \frac{I}{I_{dt}^*}\right)}{1 - e^{-\rho dt}\left(1 - \frac{I}{I_{dt}^*}\right)} \\
\leqslant& \max v \cdot 2\rho dt \left(\frac{I_{dt}^*}{I} - 1\right)
\end{aligned}$$

The first inequality is from bounding V with $\max v$. The second inequality is from $1 - e^{-\rho dt} < 2\rho dt$, $e^{-\rho dt} < 1$ and $1 - \frac{I}{I^*_{dt}} > 0$. On the other hand, the discounted total cost of this strategy is:

$$\begin{aligned} Cost'(\mu) =& C_{dt}(I) + e^{-\rho dt}\left(1 - \frac{I}{I^*_{dt}}\right) Cost'(\mu) = \frac{C_{dt}(I)}{1 - e^{-\rho dt}\left(1 - \frac{I}{I^*_{dt}}\right)} \\ \implies \quad C_{dt}(I^*_{dt}(\mu)) - Cost'(\mu) =& C_{dt}(I^*_{dt}) - \frac{C_{dt}(I)}{1 - e^{-\rho dt}\left(1 - \frac{I}{I^*_{dt}}\right)} \\ \geqslant& C_{dt}(I^*_{dt}) - \frac{C_{dt}(I) I^*_{dt}}{I} \end{aligned}$$

The inequality is from $e^{-\rho dt} < 1$ and $1 - \frac{I}{I^*_{dt}} > 0$. Therefore, by optimality of (p_i, ν_i) at μ, the new strategy I defined should not improve the expected utility:

$$\begin{aligned} & e^{-\rho dt}\left(\sum p_i V(\nu_i)\right) - V'(\mu) - \left(C_{dt}(I^*_{dt}) - Cost'(\mu)\right) \geqslant 0 \\ \implies & \sum p_i V(\nu_i) - e^{\rho dt} V'(\mu) - \left(C_{dt}(I^*_{dt}) - Cost'(\mu)\right) \geqslant 0 \\ \implies & \max v \cdot 2\rho dt \left(\frac{I^*_{dt}}{I} - 1\right) \geqslant C_{dt}(I^*_{dt}) - \frac{C_{dt}(I) I^*_{dt}}{I} \\ \implies & \max v \cdot 2\rho \left(1 - \frac{I}{I^*_{dt}}\right) \geqslant \frac{I}{I^*_{dt}} C\left(\frac{I^*_{dt}}{dt}\right) - C\left(\frac{I^*_{dt}}{dt} \cdot \frac{I}{I^*_{dt}}\right) \end{aligned} \tag{B.11}$$

By Assumption 1.2, $\exists\ \Delta$ s.t. $\forall \frac{I^*_{dt}}{dt} \geqslant \Delta$, there exists $\alpha \in (0, 1)$ s.t. $\alpha C\left(\frac{I^*_{dt}}{dt}\right) - C\left(\alpha \frac{I^*_{dt}}{dt}\right) > 2\rho \max v$. Let $I = \alpha I^*_{dt}$, then $I < I^*_{dt}$ and Equation (B.11) is violated. By contradiction, $I^*_{dt} \leqslant \Delta dt$. ■

B.1.3.2 Convergence of V_{dt}

Lemma B.8. *With Assumption A and Assumption 1.2 satisfied. Let* $\overline{V}(\mu) = \limsup_{dt \to 0} V_{dt}(\mu)$. *Then* $\lim_{dt \to 0} \left\| V_{dt}(\mu) - \overline{V}(\mu) \right\|_\infty = 0$.

Proof. We break down the proof of Lemma Lemma B.8 into three steps:

- *Step 1*: Prove that if $V_{dt} = \limsup_{n\to\infty} V_{\frac{dt}{2^n}}$, then $\left\| V_{dt} - V_{\frac{dt}{2^n}} \right\| \to 0$. First $V_{\frac{dt}{2^n}}$ is an increasing sequence, because every experimentation strategy associated with $\frac{dt}{2^n}$ can be replicated in a problem with $\frac{dt}{2^{n+1}}$: the DM can always split the experiment into two stages with equal cost in two periods and get an identical distribution of posterior beliefs at the end of second period (Lemma B.3). Moreover, $V_{\frac{dt}{2^n}}$ is always bounded above by fully informed utility. Then existence of $V_{dt} = \lim V_{\frac{dt}{2^n}}$ is guaranteed by monotonic convergence theorem.

Now let's prove the convergence is uniform in sup norm, i.e. $V_{\frac{dt}{2^n}}$ is a Cauchy sequence under sup norm. $\forall m > n$, $\forall \mu_0$, consider the problem with $\frac{dt}{2^m}$, consider the optimal experimentation $(p_i(\mu), \nu_i(\mu))$ and associated action rule A_T, information measure I_T, the expected utility is:

$$\begin{aligned} V_{\frac{dt}{2^m}}(\mu_0) &= \sum e^{-\rho T \frac{dt}{2^m}} E_{\mu_0}\Big[u(A_T, X) - C_{\frac{dt}{2^m}}(I_T) \Big]. \\ &= \sum e^{-\rho T \frac{dt}{2^n}} \sum_{\tau=0}^{2^{m-n}-1} e^{-\rho\tau \frac{dt}{2^m}} E_{\mu_0}\Big[u(A_{2^{m-n}T+\tau}, X) - C_{\frac{dt}{2^m}}(I_{2^{m-n}T+\tau}) \Big] \end{aligned} \tag{B.12}$$

The second equality is get by rewriting $T = 2^{m-n}T' + \tau$. Then take summation first over τ then over T' (and relabel T' to be T).

Now we construct an experimentation strategy for problem with $\frac{dt}{2^n}$. We combine all experiments between $2^{m-n}T$ and $2^{m-n}(T+1)$, and get the joint distribution of posteriors. We use this as the signal structure in each period T. Given this construction, at the end of each $2^{m-n}T$, the posterior distribution will be exactly same as that using original experiment. Then we assign same action as before to each posterior. By construction this action profile satisfies Markov property of information (i.e. signal realization is a sufficient statistics for action). Therefore if we let $U(\mu_0)$ be the discounted expected utility associated with the aforementioned strategy at μ_0:

$$
\begin{aligned}
& V_{\frac{dt}{2^n}}(\mu_0) \\
\geqslant & U(\mu_0) \\
= & \sum e^{-\rho T \frac{dt}{2^n}} \left(\sum_{\tau=0}^{2^{m-n}-1} e^{-\rho \frac{dt}{2^n}} E_{\mu_0}[u(A_{2^{m-n}T+\tau}, X)] - E_{\mu_0}\left[C_{\frac{dt}{2^n}} \left(\sum_{\tau=0}^{2^{m-n}-1} E_{\mu_{2^{m-n}T}}[I_{2^{m-n}T+\tau}] \right) \right] \right) \\
\geqslant & \sum e^{-\rho T \frac{dt}{2^n}} \left(\sum_{\tau=0}^{2^{m-n}-1} e^{-\rho \frac{dt}{2^n}} E_{\mu_0}[u(A_{2^{m-n}T+\tau}, X)] - E_{\mu_0}\left[C_{\frac{dt}{2^n}} \left(\sum_{\tau=0}^{2^{m-n}-1} I_{2^{m-n}T+\tau} \right) \right] \right) \\
\geqslant & \sum e^{-\rho T \frac{dt}{2^n}} \left(\sum_{\tau=0}^{2^{m-n}-1} e^{-\rho \frac{dt}{2^n}} E_{\mu_0}[u(A_{2^{m-n}T+\tau}, X)] - E_{\mu_0}\left[\sum_{\tau=0}^{2^{m-n}-1} \frac{1}{2^{m-n}} C_{\frac{dt}{2^n}} \left(2^{m-n} \cdot I_{2^{m-n}T+\tau} \right) \right] \right) \\
= & \sum e^{-\rho T \frac{dt}{2^n}} \left(\sum_{\tau=0}^{2^{m-n}-1} e^{-\rho \frac{dt}{2^n}} E_{\mu_0}[u(A_{2^{m-n}T+\tau}, X)] - E_{\mu_0}\left[\sum_{\tau=0}^{2^{m-n}-1} C_{\frac{dt}{2^m}} \left(I_{2^{m-n}T+\tau} \right) \right] \right) \qquad \text{(B.13)} \\
= & e^{-\rho \frac{dt}{2^n}} \sum e^{-\rho T \frac{dt}{2^n}} \sum_{\tau=0}^{2^{m-n}-1} E_{\mu_0}[u(A_{2^{m-n}T+\tau}, X)] - \sum e^{-\rho T \frac{dt}{2^n}} \sum_{\tau=0}^{2^{m-n}-1} E_{\mu_0} C_{\frac{dt}{2^m}} \left(I_{2^{m-n}T+\tau} \right) \\
> & e^{-\rho \frac{dt}{2^n}} \sum e^{-\rho T \frac{dt}{2^n}} \sum_{\tau=0}^{2^{m-n}-1} e^{-\rho \tau \frac{dt}{2^m}} E_{\mu_0}[u(A_{2^{m-n}T+\tau}, X)] \\
& - e^{\rho \frac{dt}{2^n}} \sum e^{-\rho T \frac{dt}{2^n}} \sum_{\tau=0}^{2^{m-n}-1} e^{-\rho \tau \frac{dt}{2^m}} E_{\mu_0}\left[C_{\frac{dt}{2^m}} \left(I_{2^{m-n}T+\tau} \right) \right] \\
\geqslant & V_{\frac{dt}{2^m}}(\mu_0) - \left(1 - e^{-\rho \frac{dt}{2^n}} \right) \max v - \left(e^{\rho \frac{dt}{2^n}} - 1 \right) \max v \\
= & V_{\frac{dt}{2^m}}(\mu_0) - \left(e^{\rho \frac{dt}{2^n}} - e^{-\rho \frac{dt}{2^n}} \right) \max v \qquad \text{(B.14)}
\end{aligned}
$$

Where $\max v$ is an upper bounded of total utility from action. The second and third inequalities are from concavity of f. Equation (B.13) is obtained by definition of $C_{dt}(\cdot) = dt \cdot C(\frac{\cdot}{dt})$. Noticing that Equation (B.13) is different from Equation (B.12) by only one term: the discounting term in inner summation ($e^{-\rho \frac{dt}{2^m}}$ instead of $e^{-\rho \frac{dt}{2^n}}$). This characterizes the experiment design in problem $\frac{dt}{2^n}$. In each period T, actions are all postponed to the end of period. Therefore they are discounted by at most $\frac{dt}{2^n}$, which is period length and costs are shifted to the beginning of each period. The next inequality is from $e^{-\rho\frac{dt}{2^m}} < 1$ and $m > n$. By Lemma 1.2, both utility gain and information cost are uniformly bounded by $\max v$, then $\left\|V_{\frac{dt}{2^n}} - V_{\frac{dt}{2^m}}\right\| \leqslant \max v\left(e^{\rho\frac{dt}{2^n}} - e^{-\rho\frac{dt}{2^n}}\right) \to 0$ when $n \to 0$.

- *Step 2*: Prove that $\forall dt > 0$, V_{dt} are identical, WLOG we can call it $\overline{V}(\mu)$. $\forall dt, dt' > 0$, $\forall n$, consider $V_{\frac{dt}{2^n}}$. Pick m large enough that there exists N s.t. $\frac{dt}{2^{n+1}} \leqslant N\frac{dt'}{2^m} \leqslant \frac{dt}{2^n} \leqslant (N+1)\frac{dt'}{2^m}$. Consider optimal experimentation and action associated with $\frac{dt}{2^n}$, we construct experimentation strategy for problem with $\frac{dt'}{2^m}$. For each time period T in the original problem, split the experiment in period T into $N+1$ periods and take any action at the end of $N+1$th period (apply Lemma B.3 recursively). In the new experiment strategy, the effective period length will increase from $\frac{dt}{2^n}$ to $(N+1)\frac{dt'}{2^m}$. First, suppose the information measure incurred in any period is I in problem with $\frac{dt}{2^n}$. Then per-period information measure from the aforementioned strategy is $\frac{I}{N+1} \leqslant 2^{n-m}\frac{dt'}{dt}I$. This leads to per-period cost $\frac{dt'}{2^m}\cdot C\left(\frac{I\cdot 2^m}{(N+1)dt'}\right) \leqslant \frac{1}{N}\frac{dt}{2^n}\cdot C\left(\frac{2^n\cdot I}{dt}\right)$. Therefore, the total cost from experimentation will increase by no more than $\frac{N+1}{N}$ times and that will be bounded by $\frac{1}{N}\max v$. Second, since induced posterior distribution and action distribution are still the same, Markov property still holds. Finally:

$$
\begin{aligned}
V_{\frac{dt'}{2^m}}(\mu_0) - V_{\frac{dt}{2^n}}(\mu_0) \geqslant & \sum e^{-\rho T(N+1)\frac{dt'}{2^m}} E_{\mu_0}[u(A_T, X)] - \sum e^{\rho T\frac{dt}{2^n}} E_{\mu_0}[u(A_T, X)] - \frac{1}{N}\max v \\
= & -\sum\left(e^{-\rho T\frac{dt}{2^n}} - e^{-\rho T(N+1)\frac{dt'}{2^m}}\right) E_{\mu_0}[u(A_T, X)] - \frac{1}{N}\max v \\
\geqslant & -\max v\left|\sum e^{-\rho T\frac{dt}{2^n}} - \sum e^{-\rho T(N+1)\frac{dt'}{2^m}}\right| - \frac{1}{N}\max v \\
= & -\max v\frac{e^{-\rho\frac{dt}{2^n}} - e^{-\rho(N+1)\frac{dt'}{2^m}}}{\left(1 - e^{-\rho\frac{dt}{2^n}}\right)\left(1 - e^{-\rho(N+1)\frac{dt'}{2^m}}\right)} - \frac{1}{N}\max v \\
\geqslant & -\max v\frac{e^{-\rho N\frac{dt'}{2^m}} - e^{-\rho(N+1)\frac{dt'}{2^m}}}{(1 - e^{-\rho\frac{dt}{2^n}})^2} - \frac{1}{N}\max v
\end{aligned}
$$

$$= -\max v \frac{e^{-\rho N \frac{dt'}{2^m}}}{(1-e^{-\rho \frac{dt}{2^n}})^2}(e^{\rho \frac{dt'}{2^m}} - 1) - \frac{1}{N}\max v$$
$$\geqslant -\max v \left(\frac{e^{-\rho \frac{dt}{2^{n+1}}}}{(1-e^{-\rho \frac{dt}{2^n}})^2}(e^{\rho \frac{dt'}{2^m}} - 1) + \frac{dt'}{dt \cdot 2^{m-n-1}} \right)$$

First inequality is from suboptimality of the constructed experiment and bound of cost difference. Second inequality is from $e^{-\rho T \frac{dt}{2^n}} \geqslant e^{-\rho T(N+1)\frac{dt'}{2^m}}$. Third inequality is from $\frac{dt}{2^n} \geqslant N\frac{dt'}{2^m}$. Last inequality is from $N\frac{dt'}{2^m} \geqslant \frac{dt}{2^{n+1}}$. Take $m \to \infty$ on both side, we have $V_{dt'}(\mu_0) \geqslant V_{\frac{dt}{2^n}}(\mu_0)$. Then take $n \to 0$ on both side $V_{dt'}(\mu_0) \geqslant V_{dt}(\mu_0)$. Since this holds for arbitrary dt, dt' and μ_0, we conclude that $V_{dt} = V_{dt'}$.

- *Step 3*: $\left\|V_{dt} - \overline{V}\right\| \to 0$ when $dt \to 0$. Fix any $dt > 0$, then $\forall \varepsilon > 0$, there exists N s.t. $\forall n \geqslant N$, $\left\|V_{\frac{dt}{2^n}} - \overline{V}\right\| < \frac{\varepsilon}{2}$. Then given the proof in last part, for any $dt' < \frac{dt}{2^n}$, suppose there exists N s.t. $\frac{dt}{2^{n-1}} \leqslant Ndt' \leqslant \frac{dt}{2^n} \leqslant (N+1)dt'$, then the difference between $V_{\frac{dt}{2^n}}$ and $V_{dt'}$ will be bounded by:

$$\max v \left(\frac{e^{-\rho \frac{dt}{2^{n+1}}}}{(1-e^{-\rho \frac{dt}{2^n}})}(e^{\rho dt'} - 1) + \frac{2^{n+1}}{dt}dt' \right)$$

Actually such $N = \left[\frac{dt}{2^n dt'}\right]$ exists for any $dt' \leqslant \frac{dt}{2^n}$. Thus there exists δ s.t. $\forall dt' < \delta$, $\left\|V_{dt'} - V_{\frac{dt}{2^n}}\right\| < \frac{\varepsilon}{2}$, then $\left\|V_{dt'} - \overline{V}\right\| < \varepsilon$.

■

B.1.3.3 Lemmas for Lemma 1.3

Lemma B.9. *With Assumption A and Assumption 1.2 satisfied. Let* $\overline{V}(\mu) = \lim_{dt \to 0} V_{dt}(\mu)$*. Then* $\overline{V} \in \mathcal{L}$ *(pointwise Lipschitz function).*

Proof. We prove by induction on dimensionality of μ. When $\mu = \delta_x$, $\text{supp}(\mu)$ is singleton. So Lemma B.9 is trivially satisifed. Now it is sufficient to prove that $\overline{V}$ is pointwise

Lipschitz at any interior μ.

First, since $\overline{V}$ is the uniform limit of continuous V_{dt}, $\overline{V}$ is continuous. $\forall \mu \in \Delta X^o$, suppose by contradiction $\overline{V}$ is not pointwise Lipschitz. Then $\exists\ \mu_n \to \mu$, $\frac{|\overline{V}(\mu_n)-\overline{V}(\mu)|}{\|\mu_n-\mu\|} \geqslant n$. There are two possibilities:

- $\frac{\overline{V}(\mu_n)-\overline{V}(\mu)}{\|\mu_n-\mu\|} \geqslant n$. Now let ν_n be a point in $\partial\Delta X$ s.t. μ_n, μ, ν_n are three ordered points on a straight line. Let p_n, q_n be such that $p_n + q_n = 1$, $p_n\mu_n + q_n\nu_n = \mu$. Pick any I s.t. $C(I) < \infty$ We have:

$$I\frac{\overline{V}(\nu_n)-\overline{V}(\mu)+\frac{\overline{V}(\mu_n)-\overline{V}(\mu)}{\|\mu_n-\mu\|}\|\nu_n-\mu\|}{H(\mu)-H(\nu_n)-\frac{H(\mu_n)-H(\mu)}{\|\mu_n-\mu\|}\|\nu_n-\mu\|} \geqslant I\frac{\overline{V}(\nu_n)-\overline{V}(\mu)+n\|\nu_n-\mu\|}{H(\mu)-H(\nu_n)-\frac{H(\mu_n)-H(\mu)}{\|\mu_n-\mu\|}\|\nu_n-\mu\|}$$

Noticing that the only difference between LHS and RHS is that $\frac{\overline{V}(\mu_n)-\overline{V}(\mu)}{\|\mu_n-\mu\|}$ is replaced with n on RHS. Since the nominator is bounded, μ being interior suggesting $\|\nu_n - \mu\|$ is strictly positive in the limit. Take $n \to \infty$ on RHS, we observe that RHS goes to infinity. Therefore, there exists N s.t. $\forall n \geqslant N$, RHS is larger than $3\rho \sup F + 2C(I)$.

$$\begin{aligned}
&I\frac{\overline{V}(\nu_n)-\overline{V}(\mu)+\frac{\overline{V}(\mu_n)-\overline{V}(\mu)}{\|\mu_n-\mu\|}\|\nu_n-\mu\|}{H(\mu)-H(\nu_n)-\frac{H(\mu_n)-H(\mu)}{\|\mu_n-\mu\|}\|\nu_n-\mu\|} \geqslant 3\rho\sup F + 2C(I)\\
\Longrightarrow\ &\frac{(\|\mu_n-\mu\|)\overline{V}(\nu_n)+\|\nu_n-\mu\|\overline{V}(\mu_n)-(\|\mu_n-\mu\|+\|\nu_n-\mu\|)\overline{V}(\mu)}{-\|\mu_n-\mu\|H(\mu_n)-\|\nu_n-\mu\|H(\mu_n)+(\|\mu_n-\mu\|+\|\nu_n-\mu\|)H(\mu)} \geqslant \frac{3\rho}{I}\sup F + \frac{2C(I)}{I}\\
\Longrightarrow\ &\frac{p_n\overline{V}(\mu_n)+q_n\overline{V}(\nu_n)-\overline{V}(\mu)}{-p_nH(\mu_n)-q_nH(\nu_n)+H(\mu)} \geqslant \frac{3\rho}{I}\sup F + \frac{2C(I)}{I}\\
\Longrightarrow\ &\frac{p_n\overline{V}(\mu_n)+q_n\overline{V}(\nu_n)-\overline{V}(\mu)}{I(\mu_n,\nu_n|\mu)} \geqslant \frac{3\rho}{I}\sup F + \frac{2C(I)}{I}\\
\Longrightarrow\ &p_n\overline{V}(\mu_n)+q_n\overline{V}(\nu_n)-V(\mu) \geqslant \frac{3\rho}{I}\sup FI(\mu_n,\nu_n|\mu)+2C(I)\frac{I(\mu_n,\nu_n|\mu)}{I}\\
\Longrightarrow\ &p_n\overline{V}(\mu_n)+q_n\overline{V}(\nu_n)-2C(I)\frac{I(\mu_n,\nu_n|\mu)}{I} \geqslant \overline{V}(\mu)\left(1+2\frac{\rho}{I}I(\mu_n,\nu_n|\mu)\right)+\sup F\frac{\rho}{I}I(\mu_n,\nu_n|\mu)\\
\Longrightarrow\ &p_n\overline{V}(\mu_n)+q_n\overline{V}(\nu_n)-2C(I)\frac{I(\mu_n,\nu_n|\mu)}{I} \geqslant \overline{V}(\mu)e^{\frac{\rho}{I}I(\mu_n,\nu_n|\mu)}+\sup F\frac{\rho}{I}I(\mu_n,\nu_n|\mu)
\end{aligned}$$

Last inequality comes from $\forall x > 0,\ 1 + 2x > e^x$. Now we have:

$$
\begin{aligned}
& e^{-\rho \frac{I(\mu_n,\nu_n|\mu)}{I}} \left(p_n \overline{V}(\mu_n) + q_n \overline{V}(\nu_n) \right) - 2e^{-\rho \frac{I(\mu_n,\nu_n|\mu)}{I}} C(I) \frac{I(\mu_n,\nu_n|\mu)}{I} \\
\geqslant & \overline{V}(\mu) + e^{-\rho \frac{I(\mu_n,\nu_n|\mu)}{I}} \sup F \frac{\rho}{I} I(\mu_n,\nu_n|\mu)
\end{aligned}
$$

Since μ_n are converging to μ, $\lim_{n\to\infty} I(\mu_n,\nu_n|\mu) = 0$. Then we can pick N sufficiently large that $\forall n \geqslant N$:

$$
e^{-\rho \frac{I(\mu_n,\nu_n|\mu)}{I}} \left(p_n \overline{V} \mu_n + q_n \overline{V}(\nu_n) \right) - \frac{I(\mu_n,\nu_n|\mu)}{I} C(I) \geqslant \overline{V}(\mu) + \frac{\rho I(\mu_n,\nu_n|\mu)}{2I} \sup F
$$

From now on, we keep n fixed. Then we pick $dt = \frac{I(\mu_n,\nu_n|\mu)}{I}$ and $dt_m = \frac{dt}{2^m}$. m is chosen sufficiently large that $\left|\overline{V} - V_{dt_m}\right| e^{\frac{\rho}{I} I(\mu_n,\nu_n|\mu)} < \frac{\rho I(\mu_n,\nu_n|\mu)}{8c} \sup F$, then:

$$
e^{-\rho \frac{I(\mu_n,\nu_n|\mu)}{I}} \left(p_n V_{dt_m}(\mu_n) + q_n V_{dt_m}(\nu_n) \right) - dt C\left(\frac{I(\mu_n,\nu_n|\mu)}{dt} \right) \geqslant V_{dt_m}(\mu) + \frac{\rho dt}{4} \sup F
$$

We consider an experimentation strategy that divides information measure $I(\mu_n,\nu_n|\mu)$ into 2^m periods uniformly, and wait until the end of the 2^m periods to take action:

$$
\begin{aligned}
& e^{-\rho dt} \left(p_n V_{dt_m}(\mu_n) + q_n V_{dt_m}(\nu_n) \right) - \sum_{\tau=0}^{2^m-1} e^{-\rho \tau dt_m} dt_m \cdot C\left(\frac{I(\mu_n,\nu_n|\mu)/2^m}{dt_m} \right) \\
> & e^{-\rho dt} \left(p_n V_{dt_m}(\mu_n) + q_n V_{dt_m}(\nu_n) \right) - \sum_{\tau=0}^{2^m-1} e^{-\rho dt} dt_m \cdot C\left(\frac{I(\mu_n,\nu_n|\mu)/2^m}{dt_m} \right) \\
= & e^{-\rho dt} \left(p_n V_{dt_m}(\mu_n) + q_n V_{dt_m}(\nu_n) - dt \cdot C\left(\frac{I(\mu_n,\nu_n|\mu)}{dt} \right) \right) \\
\geqslant & V_{dt_m}(\mu) + \frac{\rho dt}{4} \sup F
\end{aligned}
$$

First line is expected utility from taking the aforementioned experiment at μ. Second line is replacing all discounting in cost with a term larger than 1. Taking m sufficiently large, last line will be strictly larger than $V_{dt_m}(\mu)$. Thus this experiment dominates optimal value of dt_m problem at μ. Contradiction.

- $\frac{\overline{V}(\mu_n)-\overline{V}(\mu)}{\|\mu_n-\mu\|} \leqslant -n$. Then pick $\nu_n \in \partial\Delta X$ s.t. μ, μ_n, ν_n are three ordered points on a straight line. Let p_n, q_n be such that $p_n + q_n = 1$, $p_n\mu + q_n\nu_n = \mu_n$. Pick any I s.t. $C(I) < \infty$. We have:

$$I\frac{\overline{V}(\nu_n) - \overline{V}(\mu_n) + \frac{\overline{V}(\mu)-\overline{V}(\mu_n)}{\|\mu_n-\mu\|}\|\nu_n - \mu_n\|}{H(\mu_n) - H(\nu_n) - \frac{H(\mu)-H(\mu_n)}{\|\mu_n-\mu\|}\|\nu_n - \mu_n\|} \geqslant I\frac{\overline{V}(\nu_n) - \overline{V}(\mu_n) + n\|\nu_n - \mu_n\|}{H(\mu_n) - H(\nu_n) - \frac{H(\mu)-H(\mu_n)}{\|\mu_n-\mu\|}\|\nu_n - \mu_n\|}$$

Take $n \to \infty$ on RHS, we observe that RHS goes to infinity. Therefore, there exists N s.t. $\forall n \geqslant N$, RHS is larger than $3\rho \sup F + 2C(I)$.

$$\begin{aligned}\implies p_n\overline{V}(\mu) + q_n\overline{V}(\nu_n) - 2C(I)\frac{I(\mu, \nu_n|\mu_n)}{I} &\geqslant \overline{V}(\mu_n) + 3\frac{\rho I(\mu, \nu_n|\mu_n)}{I}\sup F \\ &\geqslant e^{\rho\frac{I(\mu,\nu_n|\mu_n)}{I}}\overline{V}(\mu_n) + \frac{\rho I(\mu, \nu_n|\mu_n)}{I}\sup F\end{aligned}$$

Similar to last part, N can be chosen sufficiently large that:

$$e^{-\rho\frac{I(\mu,\nu_n|\mu_n)}{I}}\left(p_n\overline{V}(\mu) + q_n\overline{V}(\nu_n)\right) - \frac{I(\mu, \nu_n|\mu_n)}{I}C(I) \geqslant \overline{V}(\mu_n) + \frac{\rho I(\mu, \nu_n|\mu_n)}{I}\sup F$$

Then pick $dt = \frac{I(\mu,\nu_n|\mu_n)}{I}$ and $dt_m = \frac{dt}{2^m}$. m can be chosen sufficiently large that:

$$e^{-\rho dt}\left(p_n V_{dt_m}(\mu) + q_n V_{dt_m}(\nu_n)\right) - dtC(I) \geqslant V_{dt_m}(\mu_n) + \frac{\rho dt}{2}\sup F$$

We consider a similar experimentation strategy as before that divides experiment uniformly:

$$e^{-\rho dt}\left(p_n V_{dt_m}(\mu) + q_n V_{dt_m}(\nu_n)\right) - \sum_{\tau=0}^{2^m-1} e^{-\rho\tau dt_m}dt_m \cdot C\left(\frac{I(\mu, \nu_n|\mu_n)/2^m}{dt_m}\right)$$

$$\geqslant V_{dt_m}(\mu_n) + \frac{\rho dt}{4} \sup F$$

RHS is strictly larger than $V_{dt_m}(\mu_n)$. This experiment dominates optimal experiment of dt_m problem at μ_n. Contradiction.

■

Lemma B.10. *$\forall f(x)$ differentiable on (a, b). $\forall x, y \in (a, b)$,*

$$\frac{1}{2} \inf_{z \in (x,y)} D^2 f(z, y)|y - x|^2 \leqslant f(y) - f(x) - f'(x)(y - x) \leqslant \frac{1}{2} \sup_{z \in (x,y)} D^2 f(z, y)|y - x|^2$$

Proof.

- First inequality: let $\underline{D} = \inf_{z \in (x,y)} D^2 f(z, y)$. Suppose by contradiction the statement is not true, then there exists ε s.t. $\frac{\underline{D}-\varepsilon}{2}|y - x|^2 > f(y) - f(x) - f'(x)(y - x)$. Let $h(w) = f(w) - f(x) - f'(x)(w - x) - \frac{\underline{D}-\varepsilon}{2}(w - x)^2$. Then $h(x) = 0$, $h'(x) = 0$ and $h(y) < 0$. Now consider $\max_z h(z) - \frac{h(y)}{y-x}(z - x)$. By continuity of h, maximizer z^* exists in $[x, y]$. FOC implies $h'(z^*) = \frac{h(y)}{y-x}$ so $z^* \neq x$. The objective function is 0 at both x, y so $z^* \neq y$. Then optimality of z^* implies $\forall dz$ sufficiently small:

$$h(z^* + dz) - \frac{h(y)}{y - x}(z^* + dz - x) \leqslant h(z^*) - \frac{h(y)}{y - x}(z^* - x)$$
$$\Longrightarrow f(z^* + dz) - f(z^*) - f'(x)dz - \frac{\underline{D} - \varepsilon}{2}(2z^* - 2x + dz)dz$$
$$\leqslant dz(f'(z^*) - f'(x) - (\underline{D} - \varepsilon)(z^* - x))$$
$$\Longrightarrow \frac{f(z^* + dz) - f(z^*) - f'(z^*)dz}{dz^2} \leqslant \frac{\underline{D} - \varepsilon}{2}$$
$$\Longrightarrow D^2 f(z^*, y) < \underline{D}$$

  Contradiction.

- Second inequality: let $\overline{D} = \sup_{z\in(x,y)} D^2(z,y)$. Suppose by contradiction the statement is not true, then there exists ε s.t. $\frac{\overline{D}+\varepsilon}{2}|y-x|^2 < f(y) - f(x) - f'(x)(y-z)$. Let $h(w) = f(w) - f(x) - f'(x)(w-x) - \frac{\overline{D}+\varepsilon}{2}(w-x)^2$. Then $h(x) = 0, h'(x) = 0$ and $h(y) > 0$. Now consider $\min_z h(z) - \frac{h(y)}{y-x}(z-x)$. By continuity of h, minimizer z^* exists in $[x,y]$. FOC implies $h'(z^*) = \frac{h(y)}{y-z}$ so $z^* \neq x$. Then optimality of z^* implies $\forall dz$ sufficiently small:

$$
\begin{aligned}
&h(z^*+dz) - \frac{h(y)}{y-x}(z^*+dz-x) \geqslant h(z^*) - \frac{h(y)}{y-x}(z^*-x)\\
\implies &f(z^*+dz) - f(z^*) - f'(x)dz - \frac{\overline{D}+\varepsilon}{2}(2z^*-2x+dz)dz\\
&\geqslant dz(f'(z^*) - f'(x) - (\overline{D}+\varepsilon)(z^*-x))\\
\implies &\frac{f(z^*+dz) - f(z^*) - f'(z^*)dz}{dz^2} \geqslant \frac{\overline{D}+\varepsilon}{2}\\
\implies &D^2 f(z^*,y) > \overline{D}
\end{aligned}
$$

Contradiction.

■

B.2 Proofs in Section 1.6

B.2.1 *Proof and lemmas of Theorem 1.2*

Proof of smoothness in Theorem 1.2

I first show that there exists a set of μ_0 such that on each interval when $V(\mu) > F(\mu)$, $V(\mu)$ is defined a V_{μ_0}. Then I utilize this result to show that V is $C^{(1)}$ smooth on $[0,1]$.

Proof. This is true when $\mu \leqslant \mu^{**}$ by definition of V_{μ^*}. So I prove this for $\mu > \mu^{**}$. First prove some useful lemmas:

Lemma B.11. *$\forall k$, there exists $\mu_0 \in \Omega$ s.t. $V_{\mu_0}(\underline{\mu}_k) > F(\underline{\mu}_k)$.*

Proof. Suppose $F = F_{k-1}$ at μ^{**}. Equation (A.18) implies $\underline{\mu}_k > \mu^{**} > \underline{\mu}_{k-1}$. Consider:

$$U_k(\mu) = \max_{\nu \geqslant \mu} \frac{c}{\rho} \frac{F(\nu) - F(\mu) - F'^{-}(\mu)(\nu - \mu)}{J(\mu, \nu)}$$

U_k is continuous by maximum theorem on $[\mu^{**}, \underline{\mu}_k)$. Since $U_k(\mu^{**}) = F(\mu^{**})$, $\lim_{\mu \to \underline{\mu}_k} U_k(\mu) = +\infty$, there exists μ_0 s.t. $U_k(\mu_0) = F(\mu_0)$ and $U_k(\mu) > F(\mu)\ \forall \mu \in (\mu_0, \underline{\mu}_k)$. Now consider $V_{\mu_0}(\mu)$. I claim that $V_{\mu_0}(\mu) > F(\mu)\ \forall \mu \in (\mu_0, \underline{\mu_k})$. Suppose not, then by intermediate value theorem, there exists μ' s.t. $V_{\mu_0}(\mu') \leqslant F(\mu)$ and $V'_{\mu_0}(\mu') \leqslant F(\mu)$. However, this implies

$$V_{\mu_0}(\mu') = \max_{\nu \geqslant \mu} \frac{c}{\rho} \frac{F(\nu) - V_{\mu_0}(\mu') - V'_{\mu_0}(\nu - \mu')}{J(\mu', \nu)} \geqslant U_k(\mu') > F(\mu')$$

Contradiction. Now assume V_{μ_0} hits F at μ'_0. Then $U_{k+1}(\mu'_0) \leqslant 0$ and $lim_{\mu \to \underline{\mu}_{k+1}}(\mu) = +\infty$, so we can find $V_{\mu_1}(\underline{\mu}_{k+1}) > F(\underline{\mu}_{k+1})$. By induction on k, Lemma B.11 is true. ■

Lemma B.12. *$\forall \mu_0 \leqslant \mu_1 \in \Omega$, let $I_i = \{\mu | V_{\nu_i}(\mu) > F(\mu)\}$. Then either $I_0 \bigcap I_1 = \varnothing$, or $I_1 \subset I_0$ and $V_{\mu_0} \geqslant V_{\mu_1}$.*

Proof. The only possible contradiction of Lemma B.12 is that $\exists \mu' \in I_0 \bigcap I_1$ s.t. $V_{\mu_1}(\mu') > V_{\mu_0}(\mu')$. Since at μ_1, $V_{\mu_0}(\mu_1) > V_{\mu_1}(\mu_1) = F(\mu_1)$, by intermediate value theorem, there exists $\xi \in (\mu_1, \mu')$ s.t. $V_{\mu_1}(\xi) > V_{\mu_0}(\xi)$ and $V'(\mu_1)(\xi) > V'(\mu_0)(\xi)$. Since $\xi \in I_1$, there exists ν, m solving Equation (A.18) for $V_{\mu_1}(\xi)$:

$$\begin{aligned} V_{\mu_0}(\xi) \geqslant & \frac{c}{\rho} \frac{F_m(\nu) - V_{\mu_0}(\xi) - V'_{\mu_0}(\xi)(\nu - \xi)}{J(\xi, \nu)} \\ > & \frac{c}{\rho} \frac{F_m(\nu) - V_{\mu_1}(\xi) - V'_{\mu_1}(\xi)(\nu - \xi)}{J(\xi, \nu)} = V_{\mu_1}(\xi) > V_{\mu_0}(\xi) \end{aligned}$$

Contradiction. So Lemma B.12 is true. ■

Lemma B.13. $\mathbb{V} = \{\max_{i=1}^{n}\{V_{\nu_i}\}\}_{\nu_i \in \Omega, n \in \mathbb{N}}$ *is totally bounded and equi-continuous on* $[\mu^{**}, 1]$.

Proof. $V_{\mu^\diamond}$ is bounded above by $\sup F(\mu)$ and below by $\inf F(\mu)$. Consider $V'_{\mu^\diamond}$. When $V_{\mu^\diamond}(\mu) = F(\mu)$, obviously direvative is bounded by $\max|F'|$. When $V_{\mu^\diamond}(\mu) > F(\mu)$. Suppose $V'_{\mu^\diamond}(\mu) > \max|F'|$, then $F(\nu) - V_{\mu^\diamond}(\mu) - V'_{\mu^\diamond}(\mu)(\nu-\mu) < F(\nu) - F(\mu) - F'(\nu)(\nu-\mu) \leqslant 0$, contradiction. By Lemma A.3, $V'_{\mu^\diamond} \geqslant 0$. So $V'_{\mu^\diamond}$ are uniformly bounded in $[0, \max|F'|]$.

Now consider $\forall n$, $\forall \nu_i \in \Omega$, $V_{\nu_i} \in [\inf F, \sup F] \implies \max_i\{V_{\nu_i}\} \in [\inf F, \sup F]$. By Lemma B.12, $\max\{V_{\nu_i}\}$ is piecewisely defined as V_{ν_i} on finite disjoint intervals. So its derivative is piecewisely defined as V'_{ν_i}, therefore bounded in $[0, \max|F'|]$. Therefore $\mathbb{V}$ is totally bounded and equi-continuous on $[\mu^{**}, 1]$. ∎

Lemma B.14. *There exists Δ s.t. $\forall \nu_i \in \Omega$, on $\{\mu | V_{\nu_i}(\mu) > F(\mu)\}$, $V'(\mu)$ has Lipschitz parameter Δ.*

Proof. $\forall \mu \in (\widehat{\mu}_{k+1}, \widehat{\mu}_k)$, ν is smooth in μ and $V'_{\nu_i} > 0$, by envelope theorem:

$$V'_{\nu_i}(\mu) = -\frac{c}{\rho}\frac{\nu-\mu}{J(\mu,\nu)}\Big(V''_{\nu_i}(\mu) + \frac{\rho}{c}V(\mu)H''(\mu)\Big) > 0$$
$$\implies V''_{\nu_i}(\mu) + \frac{\rho}{c}V_{\nu_i}(\mu)H''(\mu) < 0$$

$V_{\nu_i}(\mu)$ is bounded by $\sup F$. It is easy to see that $\sup \Omega < \underline{\mu}_n$ (where n is the largest index). By Lemma B.11, there is $\mu_0 \in \Omega$ s.t. $V_{\mu_0}(\underline{\mu}_n) > F(\underline{\mu}_n)$. By Lemma B.12, $\sup \Omega = \sup\{\mu | V_{\mu_0}(\mu) > F(\mu)\} < \nu(\mu_0) < 1$. Therefore, μ is bounded away from 1. Then by Assumption 1.3, $-H''(\mu)$ is bounded above. Therefore, Δ exists for all such μ.

Then consider $\mu = \widehat{\mu}_k$, since V''_{ν_i} is bounded on both side by Δ, $V''_{\nu_i}(\mu) \leqslant \Delta$. Therefore at μ V'_{ν_i} has Lipschitz parameter Δ by Kirszbraun theorem. ∎

- *Step 1*: prove $V \in C[\mu^{**}, 1]$.

  Sort all rational numbers in $[\mu^{**}, 1]$ as $\{r_n\}$. $\forall N$, there exists $\mu_{n,M} \in \Omega$ s.t. $V(r_n) -$

$V_{\mu_{n,M}}(r_n) \leqslant \frac{1}{N}$. Let $V_N = \max_n\{V_{\mu_{n,N}}\}$, then $\{V_N\} \subset \mathbb{V}$ and V_N converges to V pointwisely on $\{r_n\}$. Let $\widehat{V} = \lim V_n$, by Lemma B.13, $\widehat{V} \in C[\mu^{**}, 1]$. By definition $\widehat{V} \leqslant V$. Suppose $\widehat{V}(\mu) < V(\mu)$, then there exists $V_{\mu_0}(\mu) > \widehat{V}(\mu)$. Since both V_{μ_0} and $\widehat{V}$ are continuous, $V_{\mu_0} > \widehat{V}$ on an open interval, containing some r_n. Contradiction. So $\widehat{V} = V \in C[\mu^{**}, 1]$. Let $\{\mu \geqslant \mu^{**} | V(\mu) > F(\mu)\} = \bigcup I_m$ where I_m are disjoint open intervals.

- *Step 2*: prove $\forall I_m$, exists $\mu_n \in \Omega$ s.t. $V(\mu) = \lim V_{\mu_n}(\mu)$ and $V'(\mu) = \lim V'_{\mu_n}(\mu)$ on I_m.

  Pick any $\mu \in I_m$. Let $\Theta(\mu) = \{\mu^\diamond \in \Omega | V_{\mu^\diamond}(\mu) > F(\mu)\}$. Then by definition of $V(\mu)$, $\Theta(\mu)$ is non-empty. Let $\widetilde{V} = \sup_{\mu^\diamond \in \Theta(\mu)} V_{\mu^\diamond}$. $\forall N$, there exists $\mu_{n,M} \in \Theta(\mu)$ s.t. $\widetilde{V}(r_n) - V_{\mu_{n,N}}(r_n) \leqslant \frac{1}{N}$. Since $V_{\mu_{n,N}}(\mu) > F(\mu)$, by Lemma B.12, there exists $V_{\mu_N} = \max\{V_{\mu_{n,N}}\}$. Therefore, $\lim V_{\mu_N} = \widetilde{V}$ on $\{r_n\}$. By Lemma B.13 $\widetilde{V} = \lim V_{\mu_N} \in C[\mu^{**}, 1]$. Now suppose $V(\mu) > \widetilde{V}(\mu)$, then there exists $V_{\mu^\diamond}(\mu) > V_{\mu_n}(\mu) > F(\mu)$. Then $\mu^\diamond \in \Theta(\mu)$ by Lemma B.12, contradiction. Therefore, $\lim V_{\mu_n} = V$ on I_m.

  Let $I_m = (a_m, b_m)$. Now consider $\{V'_{\mu_n}\}$. $V'_{\mu_n}(a_m) = F'(a_m)$. Lemma B.14 implies that V'_{μ_n} are totally bounded and equi-continuous on I_m. Therefore, there exists subsequence V'_{μ_n} being Cauchy w.r.t. sup norm on $[a_m, b_m]$. So V as limit of V_{μ_n} is differentiable on $[a_m, b_m]$ and $V' = \lim V'_{\mu_n}$.[1]

- *Step 3*: prove $\forall I_m$, exists $\mu^m \in \Omega$ s.t. $V(\mu) = V_{\mu^m}$ on I_m.

  Let $\mu^m = \inf I_m$. By step 2, it is easy to verify that $\mu_n \to \mu^m$. Then since Equation (A.19) is continuous in μ, it is satisfied at μ^m and $\mu^m \in \Omega$. Since both V_{μ_n} and V'_{μ_n} converges on I_m, Equation (A.18) is satisfied for V on I_m. Let $F(\mu^m) = F_k(\mu^m)$.

  As an intermediate step, I first prove that Equation (A.18) is solved for $k' > k$ in a non-degenerate neighbour of μ^m. Take any $\mu' > \mu^m$ s.t. $V(\mu') > F(\mu')$, since $V(\mu^m) = F_k(\mu^m)$, there exists $\mu^\star \in (\mu^m, \mu')$ and $\varepsilon > 0$ s.t $\forall \mu \in (\mu^m, \mu^\star)$ $V(\mu) - F_k(\mu) < V(\mu') - F_k(\mu') - \varepsilon$.

[1] This result is ex. 14.2.7 from Tao (2016).

I claim that Equation (A.18) is solved at all $\mu \in (\mu^m, \mu^\star)$ with $k' > k$. Suppose not, then for n sufficiently large:

$$\begin{aligned}
V_{\mu_n}(\mu) =& \frac{c}{\rho}\frac{F_k(\nu) - V_{\mu_n}(\mu) - V'_{\mu_n}(\mu)(\nu-\mu)}{J(\mu,\nu)} \\
\leqslant& \frac{c}{\rho}\frac{F_k(\nu) - F_k(\mu) - V'_{\mu_n}(\mu)(\nu-\mu)}{J(\mu,\nu)} \\
=& (F'_k - V'_{\mu_n}(\mu))\frac{\nu-\mu}{J(\mu,\nu)}
\end{aligned}$$

Therefore $F'_k \geqslant V'_{\mu_n}(\mu)$. By construction of V_{μ_n} at any $\mu'' \geqslant \mu$ Equation (A.18) is solved with k, therefore $F'_k \geqslant V_{\mu_n}(\mu'')$ holds for all $\mu'' \geqslant \mu$. This implies $\forall \mu'' \geqslant \mu$, $V_{\mu_n}(\mu'') - F_k(\mu'') \leqslant V_{\mu_n}(\mu) - F_k(\mu) < V(\mu') - F_k(\mu') - \varepsilon$. Take $n \to \infty$ and $\mu'' = \mu'$, contradiction. Therefore, Equation (A.18) is solved at all $\mu \in (\mu^m, \mu^\star)$ for $V(\mu)$ with $k' > k$.

Now consider $V_{\mu^m}(\mu)$. By my construction, suppose V_{μ^m} is updated up to action $k+1$. I claim that $V_{\mu^m} = V$ when $\mu \in [\mu^m, \mu^\star)$. Suppose not true, then there exists μ at which $V_{\mu^m}(\mu) < V(\mu)$, $V'_{\mu^m}(\mu) < V'(\mu)$. It is easy to verify that Equation (A.18) is violated at $V_{\mu_m}(\mu)$. Therefore, if $V_{\mu^m} \neq V$, it must happen in $(\mu^\star, b_m)$. Again we can find $\mu \in (\mu^\star, b_m)$ s.t. $V_{\mu^m}(\mu) < V(\mu)$, $V'_{\mu^m}(\mu) < V'(\mu)$, which is not possible. So $V(\mu) = V_{\mu^m}(\mu)$ on I_m.

To sum up, V can be represented as:

$$V(\mu) = \begin{cases} V_{\mu^*}(\mu) & \text{if } \mu \in [\mu^*, \mu^{**}] \\ V_{\mu^m}(\mu) & \text{if } \mu \in I^m \\ F(\mu) & \text{otherwise} \end{cases}$$

Now I prove smoothness of $V(\mu)$ on $[\mu^*, 1]$. By Lemma B.14, $\forall \mu \in I_m$:

$$F'(a_m) - \Delta|\mu - a_m| \leqslant V'(\mu) \leqslant F'(a_m) + \Delta|\mu - a_m|$$

$$F'(b_m) + \Delta|\mu - b_m| \geqslant V'(\mu) \geqslant F'(b_m) - \Delta|\mu - b_m|$$

Therefore $|V'(\mu) - F'(\mu)|$ is bounded by $\Delta|I_n|$. Define:

$$V_n(\mu) = \begin{cases} V_{\mu^m}(\mu) & \text{when } \mu \in I_m,\, m \leqslant n \\ F(\mu) & \text{otherwise} \end{cases}$$

Then $V_n(\mu) \to V(\mu)$. By Lemma B.11, we can without loss assume first n V_{μ^m} have I_m covering $\underline{\mu}_m$. Fix n, $\forall \mu$, $\forall m \geqslant n$, if $\mu \in I_m$ and $m \leqslant n$ or $\mu \notin \bigcup I_m$, then $V_n'(\mu) = V_m'(\mu)$, else if $\mu \in I_m$, $m > n$, then $|V_n'(\mu) - F'(\mu)|$ and $|V_m'(\mu) - F'(\mu)|$ are all bounded by $\Delta|I_m|$. Therefore, $V_n'(\mu)$ is a Cauchy sequence. Then $V_n'(\mu) \to V'(\mu)$ pointwise. Since each V_n' is continuous, V is a smooth function on $[0, 1]$ and $V' = F'$ when $V = F$. ∎

Other lemmas for Theorem 1.2

Lemma B.15. *$\forall \delta, \eta > 0$, $\forall \mu, \nu$ s.t. $\mu, \nu \in (\delta, 1 - \delta)$, $|F_m(\mu)| > \eta$,*

$$L(\mu, \nu) = J(\nu, \mu) \frac{F_m'\big(1 + \frac{\rho}{c} J(\mu, \nu)\big) + \frac{\rho}{c} F_m(\nu)(H'(\nu) - H'(\mu))}{(\nu - \mu) F_m(\mu) H''(\nu)}$$

$L(\mu, \nu)$ is uniformly Lipschtiz continuous in ν and continuous in μ.

Proof. There exists $\sigma, \Delta > 0$ s.t. $\forall \mu \in (\delta, 1 - \delta)$

$$\begin{cases} \Delta \geqslant |F_m(\mu)| \geqslant \eta \\ \Delta \geqslant \big|H''(\mu)\big| \geqslant \varepsilon \\ \Delta \geqslant \big|H'(\mu)\big|, |H(\mu)|, \big|F_m'\big| \end{cases}$$

Since $[\delta, 1 - \delta]$ is compact, H'' is Lipschitz continuous on $[\delta, 1 - \delta]$ with Lipschitz parameter Δ. Then:

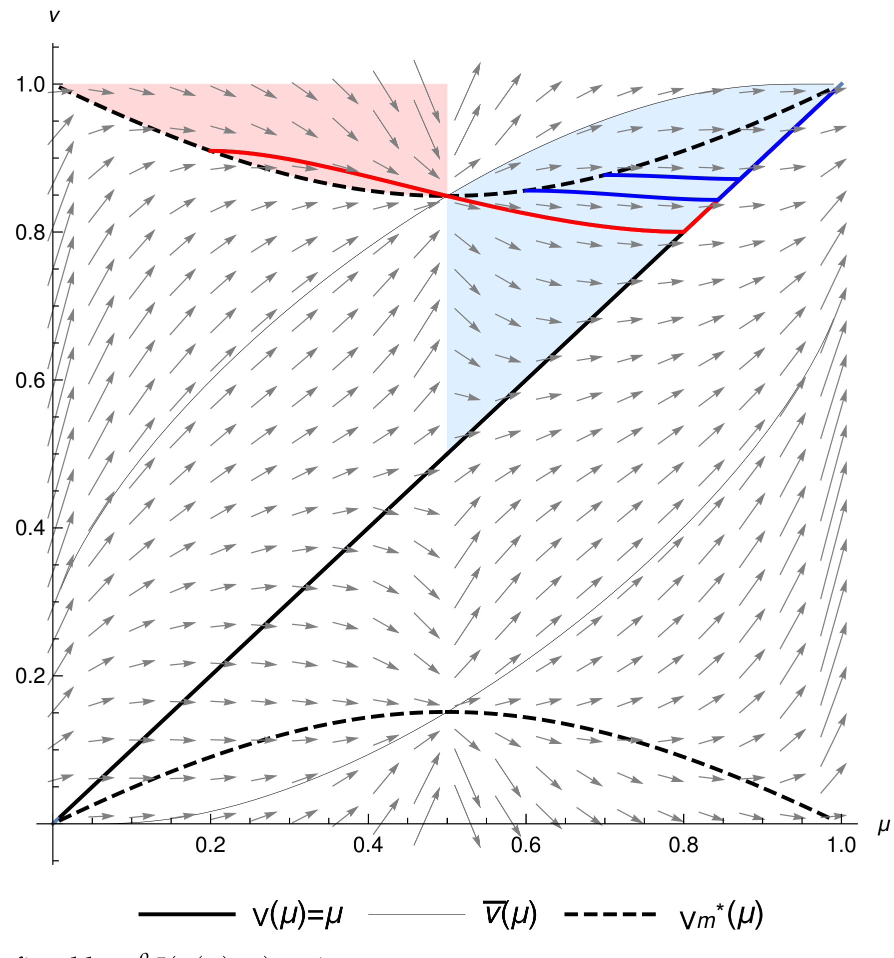
$$
\begin{aligned}
&\left|L(\mu,\nu)-L(\mu,\nu')\right| \\
\leqslant&\left|\frac{J(\nu,\mu)}{(\nu-\mu)F_m(\mu)}\right| \\
&\cdot\left|\frac{F_m'\big(1+\frac{\rho}{c}J(\mu,\nu)\big)+\frac{\rho}{c}F_m(\nu)(H'(\nu)-H'(\mu))}{H''(\nu)}-\frac{F_m'\big(1+\frac{\rho}{c}J(\mu,\nu')\big)+\frac{\rho}{c}F_m(\nu')(H'(\nu')-H'(\mu))}{H''(\nu')}\right| \\
&+\left|\frac{J(\nu,\mu)}{\nu-\mu}-\frac{J(\nu',\mu)}{\nu'-\mu}\right|\left|\frac{F_m'\big(1+\frac{\rho}{c}J(\mu,\nu')\big)+\frac{\rho}{c}F_m(\nu')(H'(\nu')-H'(\mu))}{H''(\nu')}\right| \\
\leqslant&\frac{2\Delta}{\eta\varepsilon}\left|\begin{matrix}F_m'\big(1+\frac{\rho}{c}J(\mu,\nu)\big)+\frac{\rho}{c}F_m(\nu)(H'(\nu)-H'(\mu))\\-F_m'\big(1+\frac{\rho}{c}J(\mu,\nu')\big)+\frac{\rho}{c}F_m(\nu')(H'(\nu')-H'(\mu))\end{matrix}\right| \\
&+\frac{2\Delta}{\eta}\left|\frac{H''(\nu')-H''(\mu)}{H''(\nu')H''(\mu)}\right|\left|F_m'\left(1+\frac{\rho}{c}J(\mu,\nu')\right)+\frac{\rho}{c}F_m(\nu')(H'(\nu')-H'(\mu))\right| \\
&+\left|\frac{J(\nu,\mu)}{\nu-\mu}-\frac{J(\nu',\mu)}{\nu'-\mu}\right|\left|\frac{F_m'\big(1+\frac{\rho}{c}J(\mu,\nu')\big)+\frac{\rho}{c}F_m(\nu')(H'(\nu')-H'(\mu))}{H''(\nu')}\right| \\
\leqslant&\frac{2\Delta}{\eta\varepsilon}\left|\begin{matrix}F_m'\big(1+\frac{\rho}{c}J(\mu,\nu)\big)+\frac{\rho}{c}F_m(\nu)(H'(\nu)-H'(\mu))\\-F_m'\big(1+\frac{\rho}{c}J(\mu,\nu')\big)+\frac{\rho}{c}F_m(\nu')(H'(\nu')-H'(\mu))\end{matrix}\right| \\
&+\frac{2\Delta}{\eta}\left|\frac{H''(\nu')-H''(\mu)}{H''(\nu')H''(\mu)}\right|\left(\Delta+\frac{\rho}{c}5\Delta^2\right)+\left|\frac{J(\nu,\mu)}{\nu-\mu}-\frac{J(\nu',\mu)}{\nu'-\mu}\right|\frac{\Delta+\frac{\rho}{c}5\Delta^2}{\varepsilon} \\
\leqslant&\frac{2\Delta}{\eta\varepsilon}\left|F_m'\left(\frac{\rho}{c}H'(\nu)\right)+\frac{\rho}{c}F_m(\mu)H''(\tilde{\nu})\right|\left|\nu-\nu'\right|+\frac{2\Delta^2+10\frac{\rho}{c}\Delta^3}{\eta\varepsilon^2}\Delta\left|\nu'-\nu\right| \\
+&\frac{\Delta+\frac{\rho}{c}5\Delta^2}{\varepsilon}\left|-H''(\tilde{\nu}-\frac{J(\tilde{\nu},\mu)}{(\tilde{\nu}-\mu)^2})\right|\left|\nu'-\mu\right| \\
\leqslant&\frac{2\Delta}{\eta\varepsilon}\left|2\frac{\rho}{c}\Delta^2\right|\left|\nu-\nu'\right|+\frac{2\Delta^2+10\frac{\rho}{c}\Delta^3}{\eta\varepsilon^2}\Delta\left|\nu'-\nu\right|+\frac{\Delta+\frac{\rho}{c}5\Delta^2}{\varepsilon}\left|-H''(\tilde{\nu})+\frac{1}{2}H''(\tilde{\tilde{\nu}})\right|\left|\nu'-\mu\right| \\
\leqslant&\left(\frac{4\frac{\rho}{c}\Delta^3}{\eta\varepsilon}+\frac{2\Delta^3+10\frac{\rho}{c}\Delta^4}{\eta\varepsilon^2}+\frac{2\Delta^2+10\frac{\rho}{c}\Delta^3}{\varepsilon}\right)\left|\nu'-\nu\right|
\end{aligned}
$$

Therefore, $L(\mu,\nu)$ is uniformly Lipschitz continuous in ν. It is easy to see that $L(\mu,\nu)$ is continuous in μ when μ is bounded away from ν. Now we only need to consider $\mu \to \nu$:

$$
\left|\frac{L(\mu,\nu)}{\nu-\mu}\right|=\left|\frac{(\mu-\nu)H''(\tilde{\nu})\frac{\rho}{c}F_m(\tilde{\nu}')(\nu-\mu)}{\frac{\rho}{c}(\nu-\mu)^2F_m(\mu)H''(\nu)}\right|\leqslant\frac{\Delta^2}{\eta}
$$

Therefore, $L(\nu, \mu)$ is uniformly Lipschitz continuous in ν and continuous in μ.

■



$\bar{\nu}(\mu)$ is defined by: $\frac{\rho}{c}J(\bar{\nu}(\mu), \mu) = 1$.
$\nu_m^*(\mu)$ is defined by: $F_m'\left(1 + \frac{\rho}{c}J(\mu, \nu_m^*(\mu))\right) + \frac{\rho}{c}F_m(\nu_m^*(\mu))(H'(\nu^*(\mu)) - H'(\mu)) = 0$.
The red line and blue lines are solution path of ODE $\dot{\mu} = L(\mu, \nu)$ with initial value satisfying Lemma B.16.

Figure B.1: Phase diagram of $(\dot{\mu}, \dot{\nu})$.

Lemma B.16. *Assume* $\mu_0 \in [\mu^*, 1)$, $F_m(\mu_0) \neq 0$, $F_m' \geqslant 0$, $\nu_0 \in [\mu_0, 1)$ *satisfies:*

$$F_m(\mu_0)\left(F_m'\left(1 + \frac{\rho}{c}J(\mu_0, \nu_0)\right) + \frac{\rho}{c}F_m(\nu_0)(H'(\nu_0) - H'(\mu_0))\right) \geqslant 0$$

Then there is a continuous function ν on $[\mu_0, 1]$ satisfying initial condition $\nu(\mu_0) = \nu_0$. On $\{\mu|\nu(\mu) > \mu\}$, ν is differentiable, strictly decreasing and satisfis ODE:

$$\dot{\nu} = J(\nu, \mu)\frac{F'_m\big(1 + \frac{\rho}{c}J(\mu, \nu)\big) + \frac{\rho}{c}F_m(\nu)(H'(\nu) - H'(\mu))}{(\nu - \mu)F_m(\mu)H''(\nu)}$$

Proof. Before we proceed to solving the ODE, we characterize the dynamics of (μ, ν) on $[0, 1]^2$. Figure B.1 shows the phase diagram of $\dot{\mu}, \dot{\nu}$ on $[0, 1]^2$ and some important functions that determines the dynamics of (μ, ν). The horizongtal axis is μ and vertical axis is ν. The black line is $\nu = \mu$. The two thin black lines characterizes $\bar{\nu}(\mu)$ as the solutions to:

$$1 - \frac{\rho}{c}J(\bar{\nu}(\mu), \mu) = 0$$

The two dashed black lines characterizes $\nu^*(\mu)$ as the two solutions to:

$$F'_m\left(1 + \frac{\rho}{c}J(\mu, \nu^*(\mu))\right) + \frac{\rho}{c}F_m(\nu^*(\mu))(H'(\nu^*(\mu)) - H'(\mu)) = 0$$

Since we are discussing the case $\nu \geqslant \mu$, we only focus on the upper left half of the graph:

- $F(\mu_0) < 0$. This corresponds to the left half of the graph.

$$\begin{aligned} &F'_m\left(1 + \frac{\rho}{c}J(\mu_0, \nu_0)\right) + \frac{\rho}{c}F_m(\nu_0)(H'(\nu_0) - H'(\mu_0)) \leqslant 0 \\ \Longrightarrow\ & \nu_0 \geqslant \nu^*(\mu_0) \end{aligned}$$

  Therefore our initial condition means (μ_0, ν_0) lies in the red region. $\dot{\nu} = 0$ when $\nu(\mu) = \nu^*$. otherwise $\dot{\nu} < 0$. When $F(\mu)$ is close to 0, $\dot{\nu}$ goes to negative infinity if $\nu > \nu^*(\mu)$. So the dynamics of ν in this region must have ν strictly decreasing and reaches ν^* when $F(\mu) = 0$. Intuitively, ν will never reach the region $\nu > \nu_0$. Then uniform Lipschtiz continuity of $L(\mu, \nu)$ on $\nu \in [\mu, \nu_0]$, for $\mu \in [\mu_0, F^{-1}(-\eta)]$ will be enough to guarantee existence of solution.

- $F(\mu_0) > 0$. This corresponds to the right half of the graph.

$$\begin{aligned}&F'_m\Big(1+\frac{\rho}{c}J(\mu_0,\nu_0)\Big)+\frac{\rho}{c}F_m(\nu_0)(H'(\nu_0)-H'(\mu_0))\geqslant 0\\ \implies &\nu_0\leqslant \nu^*(\mu_0)\end{aligned}$$

Our intial condition will lie below the dashed line in blue region. $L(\mu,\nu) < 0$ in this region and $L(\mu,\nu^*) = 0$. So the dynamics of ν in this region must have ν strictly decreasing until it reaches $\nu = \mu$. Then uniform Lipschtiz continuity of $L(\mu,\nu)$ on $\nu \in [\mu,\nu_0]$ for $\mu \in [\mu_0, 1]$ will be sufficient ot guanrantee existence of solution.

Then we characterize formally the solution of ODEs:

- $F_m(\mu_0) > 0$. Our conjecture is that solution ν will be no larger than ν_0 within the region: $\mu \in [\mu_0,\nu_0]$, $\nu \in [\mu_0,\nu_0]$. Therefore, we modify $L(\mu,\nu)$ to define $\tilde{L}(\mu,\nu)$ on the whole space:

$$\tilde{L}(\mu,\nu) = L(\max\{\min\{\mu,\nu_0\},\mu_0\},\max\{\min\{\nu,\nu_0\},\mu_0\})$$

It's not hard to see that $\tilde{L}$ is uniformly Lipschtiz continuous w.r.t $\nu \in \mathbb{R}$ for $\mu \in [0,1]$ and continuous in $\mu \in [0,1]$. We can apply Picard-Lindelof to solve for ODE $\dot{\tilde{\nu}} = \tilde{L}(\mu,\tilde{\nu})$ on the space with initial condition $\tilde{\nu}(\mu_0) = \nu_0$.

  - Consider $\tilde{\nu}$ on $[\mu_0, 1]$, it starts at $\nu_0 > \mu_0$. It first reaches $\nu = \mu$ at $\overline{\mu} \in (\mu_0, 1]$ (we define it to be 1 when it doesn't exist). Then for $\mu \in (\mu_0,\overline{\mu})$, we must have $L(\mu,\tilde{\nu}) <$ 0. Suppose not, then there exists $\tilde{\nu}(\mu) \geqslant \nu^*_m(\mu) > \nu_0$. We pick a smallest μ such that this is true. Then this μ must be strictly larger than μ_0 because $L(\mu_{,0},\nu_0) = 0 <$

$\dot{\nu}_m^*(\mu_0)$. Then at μ, $\dot{\tilde{v}}(\mu) = 0$ but $\dot{\nu}_m^*(\mu) > 0$. It's impossible that $\tilde{v}$ crosses ν_m^* from below. Contradiction. Then $\dot{\tilde{\nu}} < 0$ until it hits $\nu = \mu$.

- $\bar{\mu} < \nu_0$. Suppose $\bar{\mu} \geqslant \mu_0$, since $\tilde{\nu} < 0$ on $(\mu_0, \bar{\mu})$, $\tilde{v}(\bar{\mu}) < \nu_0$. Contradiction. Therefore, $\tilde{\nu}$ on $[\mu_0, \bar{\mu}]$ will be with region $[\mu_0, \nu_0]$.

In the region $[\mu_0, \bar{\mu}] \times [\mu_0, \nu_0]$, $\tilde{L}$ coincides L. Therefore, $\tilde{v}$ is a solution to original ODE Equation (A.30). We define ν:

$$\nu(\mu) = \begin{cases} \tilde{v}(\mu) & \text{if } \mu \in [\mu_0, \bar{\mu}] \\ \mu & \text{if } \mu \in [\bar{\mu}, 1] \end{cases}$$

It's easy to verify that ν satisfies Lemma B.16. The blue line on Figure B.1 illustrates a solution in this case.

- $F_m(\mu_0) < 0$. Define $\mu^0 = F^{-1}(0)$, our conjecture is that solution ν will be decreasing on $[\mu_0, \mu^0)$. $\forall \eta > 0$, define $\mu^\eta = F_{-1}(-\eta)$, we modify $L(\mu, \nu)$ to define $\tilde{L}(\mu, \nu)$ on the whole space:

$$\tilde{L}(\mu, \nu) = L(\max(\min(\mu, \mu^\eta), \mu_0), \max\{\min\{\nu, \nu_0\}, \nu_m^*(\mu)\})$$

It's not hard to see that $\tilde{L}$ is uniformly Lipschtiz continuous w.r.t. $\nu \in \mathbb{R}$ for $\mu \in [0, 1]$ and continuous in $\mu \in [0, 1]$. We can apply Picard-Lindelof to solve for ODE $\dot{\tilde{\nu}} = \tilde{L}(\mu, \tilde{v})$ on the space with initial condition $\tilde{v}(\mu_0) = \nu_0$. $\tilde{v}$ will be strictly decreasing on $(\mu_0, \mu^\eta]$. Because when $\tilde{v}$ first touches ν_m^* is must crosses from below and this is not possible. Then, when $\mu \in [\mu_0, \mu^\eta]$, we have $L(\mu, \tilde{v}) = \tilde{L}(\mu, \tilde{v})$. Therefore $\tilde{v}$ is a solution to original ODE Equation (A.30).

Then we extend $\tilde{\nu}$ to $[\mu_0, \mu^0)$ by taking $\eta \to 0$ and define:

$$\nu(\mu) = \begin{cases} \tilde{\nu}(\mu) & \text{if } \mu \in [\mu_0, \mu^0) \\ \overline{\lim}_{\mu \to F^{-1}(0)} \tilde{\nu}(\mu) & \text{if } \mu = F^{-1}(0) \end{cases}$$

First since $\tilde{\nu}$ is decreasing, the sup limit will actually be the limit and $\nu \in C[\mu_0, \mu^0]$. Then we show that this extension is left differentiable at μ^0. Consider:

$$V(\mu) = \frac{F_m(\mu)}{1 - \frac{\rho}{c} J(\nu(\mu), \mu)}$$

By Equation (A.31), we know that on $[\mu_0, \mu^0)$ sign of V' is determined by sign of $1 - \frac{\rho}{c} J(\nu(\mu), \mu)$. At initial value, $V_0 \geqslant 0 \implies 1 - \frac{\rho}{c} J(\nu_0, \mu_0) > 0$. On the other hand, $V(\mu)$ will be bounded above by $\overline{V}$. So $1 - \frac{\rho}{c} J(\nu(\mu), \mu)$ as a continuous function of μ has to stay above 0. Therefore $V'(\mu) > 0$ on $[\mu_0, \mu^0)$. By monotonic convergence, there exists $\lim_{\mu \to \mu^0 -} V(\mu)$. Define it as $V(\mu^0)$. We define:

$$\dot{\nu}(\mu^0) = \frac{\frac{F_m'}{V(\mu_0)} + \frac{\rho}{c}\big(H'(\nu(\mu^0)) - H'(\mu^0)\big)}{\frac{\rho}{c} H''(\nu(\mu^0))(\nu(\mu^0) - \mu^0)}$$

Now we show that $\dot{\nu}(\mu^0) = \lim_{\mu \to \mu^0} \frac{\nu(\mu) - \nu(\mu^0)}{\mu - \mu^0}$. Suppose not, there exists $\varepsilon > 0$, $\mu_n \to \mu^0$ s.t. $\left|\dot{\nu}(\mu^0) - \frac{\nu(\mu_n) - \nu(\mu^0)}{\mu_n - \mu^0}\right| > \varepsilon$. Suppose $\nu(\mu_n) > \nu(\mu^0) + \big(\dot{\nu}(\mu^0) - \varepsilon\big)(\mu_n - \mu^0)$:

$$\begin{aligned} V(\mu_n) <& \frac{F_m(\mu)}{1 - \frac{\rho}{c} J(\nu^0 + (\dot{\nu}(\mu^0) - \varepsilon)(\mu_n - \mu^0), \mu_n)} \\ \implies \lim_{n \to \infty} V(\mu_n) \leqslant& \frac{F_m'}{\frac{\rho}{c}(-H'(\nu(\mu^0)) + H'(\mu^0) + H''(\nu(\mu^0))(\nu(\mu^0) - \mu^0)(\dot{\nu}(\mu^0) - \varepsilon))} \\ <& \frac{F_m'}{\frac{\rho}{c}(-H'(\nu(\mu^0)) + H'(\mu^0) + H''(\nu(\mu^0))(\nu(\mu^0) - \mu^0)\dot{\nu}(\mu^0))} \\ =& V(\mu^0) \end{aligned}$$

First strict inequality is from $1 - \frac{\rho}{c}J(\nu, \mu)$ strictly increasing in ν. When $F_m(\mu) < 0$, $\frac{F_m(\mu)}{1-\frac{\rho}{c}J(\nu,\mu)}$ will be decreasing in ν. Second inequality is by taking limit of lower bounded of $V(\mu_n)$ with L'Hospital rule. Third strict inequality is from $\varepsilon > 0$, $H'' < 0$. Last equality is from definition of $\dot{\nu}(\mu^0)$. We get contradiction. Similarly, we cna rule out $\nu(\mu_n) < \nu(\mu^0) + (\dot{\nu}(\mu^0) + \varepsilon)(\mu_n - \mu^0)$. Therefore, we extended ν to $[\mu_0, \mu^0]$ such that it's differentiable on $[\mu_0, \mu^0]$ and smooth on (μ_0, μ^0).

Let $\mu_0 = \mu^0$, $\nu_0 = \nu(\mu^0)$, $\nu_0' = \dot{\nu}(\mu^0)$, then $\nu_0 > \mu_0$ and

$$\begin{cases} 1 - \frac{\rho}{c}J(\nu_0, \mu_0) = 0 \\ 0 < \frac{F_m'}{\frac{\rho}{c}(H'(\mu_0) - H'(\nu_0) + H''(\nu_0)(\nu_0 - \mu_0)\nu_0')} = V(\mu_0) \leqslant \overline{V}(\mu_0) \end{cases}$$

Then by Lemma B.17, we can solve for $\nu(\mu)$ on $[\mu^0, 1]$ satisfying the conditions in Lemma B.17. Moreover, $\dot{\nu}(\mu^0) = \nu_0$, then ν is differentiable at μ^0. For any other points in $\{\mu|\nu(\mu) > \mu\}$, ν is $C^{(1)}$ smooth. Since $\nu_0' < 0$, then the solved ν will be strictly decreasing.

■

Lemma B.17. *Assume $F_m(\mu_0) = 0$, $F_m' > 0$, $\nu_0 \in [\mu_0, 1)$, ν_0' satisfies*

$$\begin{cases} 1 - \frac{\rho}{c}J(\nu_0, \mu_0) = 0 \\ 0 < \frac{F_m'}{\frac{\rho}{c}\left(H'(\mu_0) - H'(\nu_0) + H''(\nu_0)(\nu_0 - \mu_0)\nu_0'\right)} \leqslant \overline{V}(\mu_0) \end{cases}$$

Then there is a continuous function ν on $[\mu_0, 1]$ satifying initial condition $\nu(\mu_0) = \nu_0$, $\dot{\nu}(\mu_0) = \nu_0'$. On $\{\mu|\nu(\mu) > \mu\}$, ν is differentiable, strictly decreasing and satisfies ODE:

$$\dot{\nu} = J(\nu, \mu)\frac{F_m'\left(1 + \frac{\rho}{c}J(\mu, \nu)\right) + \frac{\rho}{c}F_m(\nu)(H'(\nu) - H'(\mu))}{(\nu - \mu)F_m(\mu)H''(\nu)}$$

Proof. $\forall\ \mu_1 \in (\mu_0, 1)$, $\forall \nu_1 \in [\mu_1, \nu_m^*(\mu_1))$, we consider the solution of ODE with initial condition (μ_0, ν_0). $\forall \eta > 0$, define $\mu^\eta = F^{-1}(\eta)$. Then like the proof of Lemma B.16, we can solve for a smooth ν on $[\mu^\eta, \bar{\mu}]$. ν will be strictly decreasing below ν_m^* and strictly increasing over ν_m^*. Consider the slope of $\bar{\nu}$:

$$\dot{\bar{\nu}} = \frac{H'(\bar{\nu}) - H'(\mu)}{H''(\bar{\nu})(\bar{\nu} - \mu)} = L(\mu, \bar{\nu})$$

$\bar{\nu}$ itself satisfies ODE Equation (A.30), then uniqness of solution to ODE implies $\nu < \bar{\nu}$ $\forall \mu \in [\mu^\eta, \bar{\mu}]$. So solution must lies in the blue region in Figure B.1. Let

$$V(\mu) = \frac{F_m(\mu)}{1 - \frac{\rho}{c} J(\nu(\mu), \mu)}$$

When $\nu_1 \to \bar{\nu}(\mu_1)$, $1 - \frac{\rho}{c} J(\nu(\mu), \mu) \to 0$. Thus $V(\mu) \to \infty$. On the other hand, when $\mu_1 \to \mu_0, \nu_1 = \mu_1$, $V(\mu) \to F_m(\mu_0) = 0$. Define

$$V_0 = \frac{F_m'}{\frac{\rho}{c}\left(H'(\mu_0) - H'(\nu_0) + H''(\nu_0)(\nu_0 - \mu_0)\nu_0'\right)}$$

I want to show that there exists there exists μ_1, ν_1 s.t. $V(\mu) \to V_0$ when $\mu \to \mu_0$.

Index $V(\mu^0)$ by initial value (μ_1, ν_1): $V_0(\mu_1, \nu_1)$. I claim that $V_0(\mu_1, \nu_1)$ is continuous in (μ_1, ν_1). Suppose not, then there exists $\lim_{\mu_1^n, \nu_1^n \to \mu_1, \nu_1} V_0(\mu_1^n, \nu_1^n) \neq V_0(\mu_1, \nu_1)$. On the other hand, index $V(\mu^\eta)$ by initial value (μ_1, ν_1): $V_\eta(\mu_1, \nu_1)$, then continuous dependence of ODE guanrantees that $\lim_{\mu_1^n, \nu_1^n \to \mu_1, \nu_1} V_\eta(\mu_1^n, \nu_1^n) = V_\eta(\mu_1, \nu_1)$. Therefore, $\forall N$, there exists η s.t.

$$\frac{\left|\lim_{\mu_1^n, \nu_1^n \to \mu_1, \nu_1} V_0(\mu_1^n, \nu_1^n) - V_0(\mu_1, \nu_1)\right|}{|\mu^0 - \mu^\eta|} > 3N$$

Then by continuity, we can have η sufficiently small that:

$$\frac{\left|\lim_{\mu_1^n,\nu_1^n\to\mu_1,\nu_1} V_0(\mu_1^n,\nu_1^n) - V_\eta(\mu_1,\nu_1)\right|}{|\mu^0-\mu^\eta|} > 2N$$

Then we can have n sufficiently large that:

$$\frac{\left|V_0(\mu_1^n,\nu_1^n) - V_\eta(\mu_1^n,\nu_1^n)\right|}{\left|\mu_0-\mu_\eta\right|} > N$$

Then there must exists $\tilde{\mu}_N$ s.t. $|V'(\tilde{\mu}_N)| > N$. On the other hand, $|V'|$ must be bounded because:

$$V(\mu) = \frac{F_m(\nu) - V'(\mu)(\nu-\mu)}{1+\frac{\rho}{c}J(\mu,\nu)}$$

When V' going to positive infinity, $V(\mu)$ will go to $F_m(\mu)$. When V' going to negative infinity, $V(\mu)$ will go to positive infinity. Both cases are impossible. Therefore, $V_0(\mu_1,\nu_1)$ will be a continuous function on initial value. There exists μ_1,ν_1 such that $\lim_{\eta\to 0} V(\mu^\eta) = V_0$. Apply L'Hospital rule to $V(\mu) = \frac{F_m(\mu)}{1-\frac{\rho}{c}J(\nu(\mu),\mu)}$, we ge that:

$$V_0 = \frac{F_m'}{\frac{\rho}{c}\left(H'(\mu_0) - H'(\nu_0) + H''(\nu_0)\lim_{\mu\to\mu_0}\nu'(\mu)\right)} \implies \lim_{\mu\to\mu_0}\nu'(\mu) = \nu_0'$$

Smoothly extend $\nu(\mu)$ to μ_0. Therefore, $\nu(\mu)$ associated with initial value (μ_1,ν_1) satisfies $\dot{\nu}(\mu^0) = \nu_0'$. Since $\bar{\nu}$ satisfies $\frac{F_m'}{\frac{\rho}{c}(H'(\mu_0)-H'(\bar{\nu}(\mu_0)))} = \overline{V}(\mu_0)$, the assumption in Lemma B.17 implies $\nu_0' \leqslant 0$. ∎

Lemma A.3′. *Assume $\mu_0 \leqslant \mu^*$, $F_m' \leqslant 0$, V_0, V_0' satisfies:*

$$\begin{cases} \overline{V}(\mu_0) \geqslant V_0 \geqslant F_m(\mu_0) \\ V_0 = \max_{\nu \leqslant \mu_0} \frac{c}{\rho} \frac{F_m(\nu) - V_0 - V_0'(\nu - \mu_0)}{J(\mu_0, \nu)} \end{cases}$$

Then there exists a $C^{(1)}$ smooth and strictly decreasing $V(\mu)$ defined on $[0, \mu_0]$ satisfying

$$V(\mu) = \max_{\nu \leqslant \mu} \frac{c}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)} \tag{A.25′}$$

and initial condition $V(\mu_0) = V_0, V'(\mu_0) = V_0'$.

Lemma B.16′. *Assume $\mu_0 \in (0, \mu^*]$, $F_m(\mu_0) \neq 0$, $F_m' \leqslant 0$, $\nu_0 \in (0, \mu_0]$ satisfies:*

$$F_m(\mu_0)\Big(-F_m'\Big(1 + \frac{\rho}{c} J(\mu_0, \nu_0)\Big) + \frac{\rho}{c} F_m(\nu_0)(H'(\mu_0) - H'(\nu_0))\Big) \geqslant 0$$

Then $\exists \nu \in C[0, \mu_0]$ satisfying initial condition $\nu(\mu_0) = \nu_0$. On $\{\mu | \nu(\mu) > \nu\}$, ν is differentiable, strictly decreasing and satifies ODE:

$$\nu = J(\nu, \mu) \frac{F_m'\big(1 + \frac{\rho}{c} J(\mu, \nu)\big) + \frac{\rho}{c} F_m(\nu)(H'(\nu) - H'(\mu))}{(\nu - \mu) F_m(\mu) H''(\nu)}$$

Lemma B.17′. *Assume $F_m(\mu_0) = 0$, $F_m' < 0$, $\nu_0 \in (0, \mu_0]$, ν_0' satisfies*

$$\begin{cases} 1 - \frac{\rho}{c} J(\nu_0, \mu_0) = 0 \\ 0 > \frac{\rho}{c}(H'(\mu_0) - H'(\nu_0) + J''(\nu_0)(\nu_0 - \mu_0)\nu_0) \geqslant \frac{\overline{V}(\mu_0)}{F_m'} \end{cases}$$

Then $\exists\ \nu \in C[0, \mu_0]$ satifying initial condition $\nu(\mu_0) = \nu_0$, $\dot{\nu}(\mu_0) = \nu_0'$. On $\{\mu | \nu(\mu) > \mu\}$, ν is *differentiable, strictly decreasing and satisfies ODE:*

$$\nu = J(\nu,\mu)\frac{F_m'\left(1+\frac{\rho}{c}J(\mu,\nu)\right)+\frac{\rho}{c}F_m(\nu)(H'(\nu)-H'(\mu))}{(\nu-\mu)F_m(\mu)H''(\nu)}$$

Lemma B.18. *Suppose at μ_0, V_0, V_0', $k \geqslant 1$ satisfies:*

$$\begin{cases} V_0 = \max\limits_{\nu\geqslant\mu_0} \dfrac{c}{\rho}\dfrac{F_{m-k}(\nu)-V_0-V_0'(\nu-\mu_0)}{J(\mu_0,\nu)} \geqslant \max\limits_{\nu\geqslant\mu_0} \dfrac{c}{\rho}\dfrac{F_m(\nu)-V_0-V_0'(\nu-\mu_0)}{J(\mu_0,\nu)} \\ \overline{V}(\mu_0) \geqslant V_0 \geqslant F_{m-k}(\mu_0) \end{cases}$$

V_{m-k} is the solution as defined in Lemma A.3 with initial condition V_0, V_0', then $\forall \mu \in [\mu_0, \nu(\mu_0)]$:

$$V_{m-k}(\mu) \geqslant \max_{\nu\geqslant\mu, m'\in[m-k,m]} \frac{c}{\rho}\frac{F_{m'}(\nu)-V_{m-k}(\mu)-V_{m-k}'(\mu)(\nu-\mu)}{J(\mu,\nu)}$$

Proof. I first claim that:

$$V_0 \geqslant \max_{\nu\in\left[\mu_0,\underline{\mu}_m\right]} \frac{c}{\rho}\frac{V_{m-k}(\nu)-V_0-V_0'(\nu-\mu_0)}{J(\mu_0,\nu)}$$

Suppose not, then there exists μ' s.t.

$$V_0 < \frac{c}{\rho}\frac{V_{m-k}(\mu')-V_0-V_0'(\mu'-\mu_0)}{J(\mu_0,\mu')} \tag{B.15}$$

By definition of V_0, we must have $V_{m-k}(\mu') > F_{m-k}(\mu')$. The inequality is trivial because if $F_{m-k}(\mu') = V_{m-k}(\mu')$, then choosing μ' will be suboptimal. Therefore $\nu(\mu') > \mu'$. Optimality implies Equation (A.27) and Equation (A.26) at $\mu = \mu_0$:

$$\begin{cases} F_{m-k}' + \dfrac{\rho}{c}V_0H'(\nu(\mu)) = V_0' + \dfrac{\rho}{c}V_0H'(\mu) \\ \left(F_{m-k}(\nu(\mu)) + \dfrac{\rho}{c}V_0H(\nu(\mu))\right) - \left(V_0 + \dfrac{\rho}{c}V_0H(\mu)\right) = \left(V_0' + \dfrac{\rho}{c}V_0H'(\mu)\right)(\nu(\mu)-\mu) \end{cases}$$

We define $L(V, \lambda, \mu)(\mu')$ as a linear function of μ':

$$L(V, \lambda, \mu)(\mu') = (V(\mu) + \lambda H(\mu)) + (V'(\mu) + \lambda H'(\mu))(\mu' - \mu) \tag{B.16}$$

Define $G(V, \lambda)(\mu)$ as a function of μ:

$$G(V, \lambda)(\mu) = V(\mu) + \lambda H(\mu) \tag{B.17}$$

Then $G(F_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0))(\mu')$ is a concave function of μ'. Consider:

$$L\Big(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0), \mu_0\Big)(\mu') - G\Big(F_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0)\Big)(\mu')$$

This is a convex function and have unique minimum. Therefore, the minimum will be determined by FOC. Simple calculation shows that it is minimized at $\nu(\mu_0)$ and the minimal value is 0.

$$\text{FOC} : V'_{m-k}(\mu_0) + \frac{\rho}{c} V_{m-k}(\mu_0) H'(\mu_0) = F'_{m-k} + \frac{\rho}{c} V_{m-k}(\mu_0) H'(\mu')$$

It's easy to see that this equation is identical to the FOC for $\nu(\mu_0)$. Now consider:

$$\begin{aligned}
&L\Big(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0), \mu_0\Big)(\mu') - G\Big(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0)\Big)(\mu') \\
=&\Big(V_{m-k}(\mu_0) + \frac{\rho}{c} V_{m-k}(\mu_0) H(\mu_0)\Big) + \Big(V'_{m-k}(\mu_0) + \frac{\rho}{c} V_{m-k}(\mu_0) H'(\mu_0)\Big)(\mu' - \mu_0) \\
&- \Big(V_{m-k}(\mu') + \frac{\rho}{c} V_{m-k}(\mu_0) H(\mu')\Big) \\
=&-\Big(V_{m-k}(\mu') - V_{m-k}(\mu_0) - V'_{m-k}(\mu_0)(\mu' - \mu_0) - \frac{\rho}{c} V_{m-k}(\mu_0) J(\mu_0, \mu')\Big) < 0
\end{aligned}$$

The last inequality is from rewriting Equation (B.15). Therefrore, $L(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0), \mu_0)(\mu') - G(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0))(\mu')$ will have minimum strictly negative. Suppose it's minimized at

μ' (Since $L(\mu_0) - G(\mu_0) = 0$, μ' must be bounded away from μ_0). Then FOC implies:

$$V'_{m-k}(\mu_0) + \frac{\rho}{c} V_{m-k}(\mu_0) H'(\mu_0) = V'_{m-k}(\mu') + \frac{\rho}{c} V_{m-k}(\mu_0) H(\mu')$$

Consider:

$$\begin{aligned}
&L\Big(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0), \mu'\Big)(\nu(\mu')) - G\Big(F_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0)\Big)(\nu(\mu')) \\
=&L\Big(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0), \mu_0\Big)(\nu(\mu')) - G\Big(F_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0)\Big)(\nu(\mu')) \\
&+ V_{m-k}(\mu') - V_{m-k}(\mu_0) + \frac{\rho}{c} V_{m-k}(\mu_0)(H(\mu') - H(\mu_0)) - (V'_{m-k}(\mu_0) + \frac{\rho}{c} H'(\mu_0))(\mu' - \mu_0) \\
\geqslant& V_{m-k}(\mu') - V_{m-k}(\mu_0) + \frac{\rho}{c} V_{m-k}(\mu_0)(H(\mu') - H(\mu_0)) - (V'_{m-k}(\mu_0) + \frac{\rho}{c} H'(\mu_0))(\mu' - \mu_0) \\
=&G\Big(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0)\Big)(\mu') - L\Big(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0), \mu_0\Big)(\mu') > 0
\end{aligned}$$

In the first equality we used FOC. In the first inequality we used suboptimality of $\nu(\mu')$ at μ_0. However:

$$\begin{aligned}
0 =&L\Big(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu'), \mu'\Big)(\nu(\mu')) - G\Big(F_{m-k}, \frac{\rho}{c} V_{m-k}(\mu')\Big)(\nu(\mu')) \\
=&L\Big(V_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0), \mu'\Big)(\nu(\mu')) - G\Big(F_{m-k}, \frac{\rho}{c} V_{m-k}(\mu_0)\Big)(\nu(\mu')) \\
&+ \frac{\rho}{c}(V_{m-k}(\mu') - V_{m-k}(\mu))\big(H(\mu') - H(\nu(\mu')) + H'(\mu')(\nu(\mu') - \mu')\big) \\
>&\frac{\rho}{c}(V_{m-k}(\mu') - V_{m-k}(\mu)) J(\mu', \nu(\mu')) > 0
\end{aligned}$$

Contradiction.

Now we show Lemma B.18. Suppose that it is not true, then there exists $\mu' \in (\mu_0, \nu(\mu_0))$ and $\mu'' \geqslant \underline{\mu}_{m'}$ s.t.:

$$V_{m-k}(\mu') < \frac{c}{\rho} \frac{F_{m'}(\mu'') - V_{m-k}(\mu') - V'_{m-k}(\mu')(\mu'' - \mu')}{J(\mu', \mu'')}$$

Then by definition:

$$\begin{aligned}
0 \leqslant & L\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu_0), \mu_0\Big)(\mu'') - G\Big(F_{m'}, \frac{\rho}{c}V_{m-k}(\mu_0)\Big)(\mu'') \\
= & L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu_0), \nu(\mu_0)\Big)(\mu'') - G\Big(F_{m'}, \frac{\rho}{c}V_{m-k}(\mu_0)\Big)(\mu'') \\
0 \leqslant & L\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu_0), \mu_0\Big)(\mu') - G\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu_0)\Big)(\mu') \\
= & L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu_0), \nu(\mu_0)\Big)(\mu') - G\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu_0)\Big)(\mu') \\
\implies & L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \nu(\mu_0)\Big)(\mu'') - G\Big(F_{m'}, \frac{\rho}{c}V_{m-k}(\mu')\Big)(\mu'') \\
= & L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu_0), \nu(\mu_0)\Big)(\mu'') - G\Big(F_{m'}, \frac{\rho}{c}V_{m-k}(\mu_0)\Big)(\mu'') \\
& + \frac{\rho}{c}\big(V_{m-k}(\mu') - V_{m-k}(\mu_0)\big)J(\mu_0, \mu'') \\
> & 0 \\
& L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \nu(\mu_0)\Big)(\mu') - G\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu')\Big)(\mu') \\
= & L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu_0), \nu(\mu_0)\Big)(\mu') - G\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu_0)\Big)(\mu') \\
& + \frac{\rho}{c}\big(V_{m-k}(\mu') - V_{m-k}(\mu_0)\big)J(\mu_0, \mu') \\
> & 0
\end{aligned}$$

Now we consider $L\big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \mu'\big)(\cdot)$:

$$\begin{aligned}
& \begin{cases}
L\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \mu'\Big)(\mu') = G\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu')\Big)(\mu') \\
L\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \mu'\Big)(\nu(\mu_0)) \geqslant G\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu')\Big)(\nu(\mu_0)) \\
L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \nu(\mu_0)\Big)(\mu') > G\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu')\Big)(\mu') \\
L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \nu(\mu_0)\Big)(\nu(\mu_0)) = G\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu')\Big)(\nu(\mu_0))
\end{cases} \\
\implies & \begin{cases}
L\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \mu'\Big)(\nu(\mu_0)) \geqslant L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \nu(\mu_0)\Big)(\nu(\mu_0)) \\
L\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \mu'\Big)(\mu') < L\Big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \nu(\mu_0)\Big)(\mu')
\end{cases}
\end{aligned}$$

The two equalities are directly from definition of L and G. First inequality is from suboptimality, second inequality is from previous calculation. Therefore $L\big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \mu'\big)(\cdot)$ is lower at μ' and $L\big(F_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \nu(\mu_0)\big)(\cdot)$ is lower at $\nu(\mu_0)$. Since both of them are linear fuinctions, then $L\big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \mu'\big)(\cdot)$ must be higher at any $\mu'' > \nu(\mu_0)$. Therefore, this implies:

$$L\Big(V_{m-k}, \frac{\rho}{c}V_{m-k}(\mu'), \mu'\Big)(\mu'') > G\Big(F_{m'}, \frac{\rho}{c}V_{m-k}(\mu')\Big)(\mu'')$$

Contradicting that μ'' is superior than $\nu(\mu')$. ■

Lemma B.18′. *Suppose at μ_0, V_0, V_0', $k \geqslant 1$ satisfies:*

$$\begin{cases} V_0 = \max\limits_{\nu \leqslant \mu_0} \dfrac{c}{\rho} \dfrac{F_{m-k}(\nu) - V_0 - V_0'(\nu - \mu_0)}{J(\mu_0, \nu)} \leqslant \max\limits_{\nu \geqslant \mu_0} \dfrac{c}{\rho} \dfrac{F_m(\nu) - V_0 - V_0'(\nu - \mu_0)}{J(\mu_0, \nu)} \\ \overline{V}(\mu_0) \geqslant V_0 \geqslant F_{m+k}(\mu_0) \end{cases}$$

V_{m+k} is the solution as defined in Lemma A.3 with initial condition V_0, V_0', then $\forall \mu \in [\nu(\mu_0), \mu_0]$:

$$V_{m+k}(\mu) \geqslant \max_{\nu \leqslant \mu, m' \in [m, m+k]} \frac{c}{\rho} \frac{F_{m'}(\nu) - V_{m-k}(\mu) - V_{m-k}'(\mu)(\nu - \mu)}{J(\mu, \nu)}$$

B.2.2 *Proof of Theorem 1.3*

Proof. In this part, we introduce the algorithm to construct $V(\mu)$ and $\nu(\mu)$. We only discuss the case $\mu \geqslant \mu^*$ and the case $\mu \leqslant \mu^*$ will follow by a symmetric method.

Algorithm:

- *Step 1*: Define:

$$\overline{V}^+(\mu) = \max_{\nu \geqslant \mu} \frac{F_m(\nu) - \frac{C(I)}{I} J(\mu, \nu)}{1 + \frac{\rho}{I} J(\mu, \nu)}$$

$$\overline{V}^-(\mu) = \max_{\nu \leqslant \mu} \frac{F_m(\nu) - \frac{C(I)}{I} J(\mu,\nu)}{1 + \frac{\rho}{I} J(\mu,\nu)}$$

$\overline{V}^+$ is increasing and $\overline{V}^-$ is decreasing. There exists $\mu^* \in [0,1]$ s.t. $\overline{V}^+(\mu) \geqslant \overline{V}^-(\mu)$ when $\mu \geqslant \mu^*$ and $\overline{V}^-(\mu) \leqslant \overline{V}^-(\mu)$ when $\mu \leqslant \mu^*$ (See Lemma B.20). Define $\overline{V}(\mu) = \max\{\overline{V}^+(\mu), \overline{V}^-(\mu)\}$.

- *Step 2*: I construct the first piece of $V(\mu)$ to the right of μ^*. By Lemma B.20, there are three possible cases of μ^* to discuss ($\mu^* = 1$ is omitted by symmetry).

  Case 1: Suppose $\mu^* \in (0,1)$ and $\overline{V}(\mu^*) > F(\mu^*)$. Then there exists $(m, \nu(\mu^*) > \mu^*, I)$ s.t.

$$\overline{V}(\mu^*) = \frac{F_m(\nu(\mu^*)) - \frac{C(I)}{I} J(\mu^*, \nu(\mu^*))}{1 + \frac{\rho}{I} J(\mu^*, \nu(\mu^*))}$$

  With initial condition $(\mu_0 = \mu^*, V_0 = \overline{V}(\mu^*), V_0' = 0)$, we solve for $V_m(\mu)$ on $[\mu^*, 1]$ as defined by Lemma B.22. Define

$$V_{\mu*}(\mu) = \begin{cases} F(\mu) \text{ if } \mu \leqslant \mu^* \\ V_m(\mu) \text{ if } \mu \geqslant \mu^* \end{cases}$$

  Be Lemma B.22, when $V_{\mu*}(\mu) > F(\mu)$, $V_{\mu*}$ is smoothly increasing and optimal $\nu(\mu)$ is smoothly decreasing.

  Now update $V_{\mu*}(\mu)$ with respect to more actions. Let $\hat{\mu}_m$ be the smallest $\mu \geqslant \mu^*$ that:

$$V_m(\mu) = \max_{\nu \geqslant \mu, I} \frac{I}{\rho} \frac{F_{m-1}(\nu) - V_m(\mu) - V_m'(\mu)(\nu - \mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho}$$

  If $V_m(\widehat{\mu}_m) > F_{m-1}(\widehat{\mu}_m)$ we solve for V_{m-1} with initial condition $\mu_0 = \hat{\mu}_m, V_0 = V_m(\hat{\mu}_m), V_0' = V_m'(\hat{\mu}_m)$. Then redefine $V_{\mu*}(\mu)|_{\mu \geqslant \widehat{\mu}_m} = V_{m-1}(\mu)$. Otherwise skip to looking for $\widehat{\mu}_{m-1}$. If $m - 1 > \overline{m}$, we continue this procedure by looking for $\hat{\mu}_{m-1}$ until $m = \overline{m}$. Now suppose

$V_{\overline{m}}$ first hits $F(\mu)$ at $\mu^{**} > \mu^*$. V_{μ^*} is a smooth function on $[\mu^*, \mu^{**}]$ such that:

$$V_{\mu^*}(\mu) = \begin{cases} F(\mu) \text{ if } \mu \leqslant \mu^* \text{ or } \mu \geqslant \mu^{**} \\ V_k(\mu) \text{ if } \mu \in [\widehat{\mu}_k, \widehat{\mu}_{k-1}]^2 \end{cases}$$

By construction, optimal posterior $\nu_{\mu^*}(\mu)$ is smoothly decreasing on each $(\widehat{\mu}_{k+1}, \widehat{\mu}_k)$ and jumps down at each $\widehat{\mu}_k$. By Lemma B.23 and our construction, $\forall \mu \in [\mu^*, \mu^{**}]$:

$$V_{\mu^*}(\mu) = \max_{\nu \geqslant \mu, k, I} \frac{I}{\rho} \frac{F_k(\nu) - V_{\mu^*}(\mu) - V'_{\mu^*}(\mu)(\nu - \mu)}{J(\mu, \nu)} - \frac{C(I)}{\rho} \tag{B.18}$$

Case 2: Suppose $\mu^* \in (0,1)$ but $\overline{V}(\mu^*) = F(\mu^*)$, let $\mu^{**} = \inf\{\mu | \overline{V}(\mu) > F(\mu)\}$.

Case 3: Suppose $\mu^* = 0$, consider:

$$\widetilde{V}(\mu) = \max_{\nu \geqslant \mu, k, I} \frac{I}{\rho} \frac{F_k(\nu) - F_1(\mu) - F'_1(\nu - \mu)}{J(\mu, \nu)} - \frac{C(I)}{\rho}$$

Define $\mu^{**} = \inf\left\{\mu \middle| \widetilde{V}(\mu) > F_1(\mu)\right\} > 0$.

- *Step 3:* Solve for V to the right of μ^{**}. For all $\mu^\diamond \geqslant \mu^{**}$ such that:

$$F(\mu^\diamond) = \max_{\nu \geqslant \mu, k} \frac{I}{\rho} \frac{F_k(\nu) - F(\mu^\diamond) - F'^{-}(\mu^\diamond)(\nu - \mu^\diamond)}{J(\mu^\diamond, \nu)} - \frac{C(I)}{\rho} \tag{B.19}$$

Let m be the index of optimal action. Solve for V_m with initial condition $\mu_0 = \mu^\diamond, V_0 = F(\mu^\diamond), V'_0 = F'^{-}(\mu^\diamond)$. Then take same steps in *Step 3* and solve for $\widehat{\mu}_k$ and V_{k-1} sequentially until V_{m_0} first hits F. This step refers to Figure A.6-4,5. Now suppose V_{m_0} first hits

[1] Define $\widehat{\mu}_{m+1} = \mu^*$ and $\widehat{\mu}_{\overline{m}} = \mu^{**}$ for consistency.

$F(\mu)$ at some point $\mu^{\diamond\diamond}$ (can potentially be μ), define:

$$V_{\mu^\diamond}(\mu) = \begin{cases} F(\mu) & \text{if } \mu < \mu^\diamond \text{ or } \mu > \mu^{\diamond\diamond} \\ V_k(\mu) & \text{if } \mu \in [\widehat{\mu}_{k+1}, \widehat{\mu}_k]^3 \end{cases}$$

By Lemma B.21, V_μ is piecewise smooth are pasted smoothly. So V_μ is a smooth function on $[\mu, \mu'']$. Optimal posterior $\nu_{\mu^\diamond}(\mu)$ is smoothly decreasing on each $(\widehat{\mu}_{k+1}, \widehat{\mu}_k)$ and jumps down at each $\widehat{\mu}_k$. By Lemma B.23 and our construction, $\forall \mu \in [\mu^\diamond, \mu^{\diamond\diamond}]$:

$$V_\mu(\mu) = \max_{\nu \geqslant \mu^\diamond, k} \frac{I}{\rho} \frac{F_k(\nu) - V_{\mu^\diamond}(\mu) - V'_{\mu^\diamond}(\mu)(\nu - \mu)}{J(\mu, \nu)} - \frac{C(I)}{\rho} \tag{B.18}$$

Let Ω be the set of all such $\mu^\diamond$'s.

- *Step 4:* Define:

$$V(\mu) = \begin{cases} V_{\mu^*}(\mu) & \text{if } \mu \in [\mu^*, \mu^{**}] \\ \sup_{\mu^\diamond \in \Omega} \left\{ V_{\mu^\diamond}(\mu) \right\} & \text{if } \mu \geqslant \mu^{**} \end{cases}$$

Smoothness:

I want to show that $V(\mu)$ is piecewisely defined as V_{μ_0}'s. This is true when $\mu \leqslant \mu^{**}$ by definition of V_{μ^*}. So I prove this for $\mu > \mu^{**}$. First it is easy to verify that Lemmas B.11, B.12 and B.13 still hold. The original proof directly applies by replacing Equation (A.18) with Equation (B.18) and Lemma A.3 with Lemma B.21.

Lemma B.19. *There exists Δ s.t. $\forall \mu_i \in \Omega$, on $\left\{\mu | V_{\mu_i}(\mu) > F(\mu)\right\}$, $V'(\mu)$ has Lipschitz parameter Δ.*

[3] Define $\widehat{\mu}_{m+1} = \mu^\diamond$ and $\widehat{\mu}_{m_0} = \mu^{\diamond\diamond}$ for consistency.

Proof. $\forall \mu \in (\widehat{\mu}_{k+1}, \widehat{\mu}_k)$, ν is smooth in μ and $V'_{\mu_i} > 0$, by envelope theorem:

$$\begin{aligned} V'_{\mu_i}(\mu) &= -\frac{I}{\rho}\frac{\nu-\mu}{J(\mu,\nu)}\Big(V''_{\mu_i}(\mu) + C'(I)H''(\mu)\Big) > 0 \\ \implies & V''_{\mu_i}(\mu) + C'(I)H''(\mu) < 0 \end{aligned}$$

$C'(I)$ is bounded since $C(I)$ is bounded by $\sup F$. It is easy to see that $\sup \Omega < \underline{\mu}_n$ (where n is the largest index). By Lemma B.11, there is $\mu_0 \in \Omega$ s.t. $V_{\mu_0}(\underline{\mu}_n) > F(\underline{\mu}_n)$. By Lemma B.12, $\sup \Omega = \sup\{\mu | V_{\mu_0}(\mu) > F(\mu)\} < \nu(\mu_0) < 1$. Therefore, μ is bounded away from 1. Then by Assumption 1.3, $-H''(\mu)$ is bounded above. Therefore, Δ exists for all such μ.

Then consider $\mu = \widehat{\mu}_k$, since V''_{μ_i} is bounded on both side by Δ, $V''_{\mu_i}(\mu) \leqslant \Delta$. Therefore at μ V'_{μ_i} has Lipschitz parameter Δ by Kirszbraun theorem. ■

- *Step 1*: $V \in C[\mu^{**}, 1]$. Sort all rational numbers in $[\mu^{**}, 1]$ as $\{r_n\}$. $\forall N$, there exists $\mu_{n,M} \in \Omega$ s.t. $V(r_n) - V_{\mu_{n,M}}(r_n) \leqslant \frac{1}{N}$. Let $V_N = \max_n\{V_{\mu_{n,N}}\}$, then $\{V_N\} \subset \mathbb{V}$ and V_N converges to V pointwisely on $\{r_n\}$. Let $\widehat{V} = \lim V_n$, by Lemma B.13, $\widehat{V} \in C[\mu^{**}, 1]$. By definition $\widehat{V} \leqslant V$. Suppose $\widehat{V}(\mu) < V(\mu)$, then there exists $V_{\mu_0}(\mu) > \widehat{V}(\mu)$. Since both V_{μ_0} and $\widehat{V}$ are continuous, $V_{\mu_0} > \widehat{V}$ on an open interval, containing some r_n. Contradiction. So $\widehat{V} = V \in C[\mu^{**}, 1]$. Let $\{\mu \geqslant \mu^{**} | V(\mu) > F(\mu)\} = \bigcup I_m$ where I_m are disjoint open intervals.

- *Step 2*: $\forall I_m$, exists $\mu_n \in \Omega$ s.t. $V(\mu) = \lim V_{\mu_n}(\mu)$ and $V'(\mu) = \lim V'_{\mu_n}(\mu)$ on I_m. Pick any $\mu \in I_m$. Let $\Theta(\mu) = \{\mu^\diamond \in \Omega | V_{\mu^\diamond}(\mu) > F(\mu)\}$. Then by definition of $V(\mu)$, $\Theta(\mu)$ is non-empty. Let $\widetilde{V} = \sup_{\mu^\diamond \in \Theta(\mu)} V_{\mu^\diamond}$. $\forall N$, there exists $\mu_{n,M} \in \Theta(\mu)$ s.t. $\widetilde{V}(r_n) - V_{\mu_{n,N}}(r_n) \leqslant \frac{1}{N}$. Since $V_{\mu_{n,N}}(\mu) > F(\mu)$, by Lemma B.12, there exists $V_{\mu_N} = \max\{V_{\mu_{n,N}}\}$. Therefore, $\lim V_{\mu_N} = \widetilde{V}$ on $\{r_n\}$. By Lemma B.13 $\widetilde{V} = \lim V_{\mu_N} \in C[\mu^{**}, 1]$. Now suppose $V(\mu) > \widetilde{V}(\mu)$, then there exists $V_{\mu^\diamond}(\mu) > V_{\mu_n}(\mu) > F(\mu)$. Then $\mu^\diamond \in \Theta(\mu)$ by Lemma B.12, contradiction. Therefore, $\lim V_{\mu_n} = V$ on I_m.

Let $I_m = (a_m, b_m)$. Now consider $\{V'_{\mu_n}\}$. $V'_{\mu_n}(a_m) = F'(a_m)$. Lemma B.19 implies that V'_{μ_n} are totally bounded and equi-continuous on I_m. Therefore, there exists subsequence V'_{μ_n} being Cauchy w.r.t. sup norm on $[a_m, b_m]$. So V as limit of V_{μ_n} is differentiable on $[a_m, b_m]$ and $V' = \lim V'_{\mu_n}$.

- *Step 3* $\forall I_m$, exists $\mu^m \in \Omega$ s.t. $V(\mu) = V_{\mu^m}$ on I_m. Let $\mu^m = \inf I_m$. By step 2, it is easy to verify that $\mu_n \to \mu^m$. Then since Equation (B.19) is continuous in μ, it is satisfied at μ^m and $\mu^m \in \Omega$. Since both V_{μ_n} and V'_{μ_n} converges on I_m, Equation (B.18) is satisfied for V on I_m. Let $F(\mu^m) = F_k(\mu^m)$.

  As an intermediate step, I first prove that Equation (B.18) is solved for $k' > k$ in a non-degenerate neighbour of μ^m. Take any $\mu' > \mu^m$ s.t. $V(\mu') > F(\mu')$, since $V(\mu^m) = F_k(\mu^m)$, there exists $\mu^\star \in (\mu^m, \mu')$ and $\varepsilon > 0$ s.t $\forall \mu \in (\mu^m, \mu^\star)$ $V(\mu) - F_k(\mu) < V(\mu') - F_k(\mu') - \varepsilon$. I claim that Equation (B.18) is solved at all $\mu \in (\mu^m, \mu^\star)$ with $k' > k$ and I. Suppose not, then for n sufficiently large:

$$
\begin{aligned}
V_{\mu_n}(\mu) =& \frac{I}{\rho} \frac{F_k(\nu) - V_{\mu_n}(\mu) - V'_{\mu_n}(\mu)(\nu - \mu)}{J(\mu, \nu)} - \frac{C(I)}{\rho} \\
\leqslant& \frac{I}{\rho} \frac{F_k(\nu) - F_k(\mu) - V'_{\mu_n}(\mu)(\nu - \mu)}{J(\mu, \nu)} \\
=& \left(F'_k - V'_{\mu_n}(\mu) \right) \frac{\nu - \mu}{J(\mu, \nu)}
\end{aligned}
$$

  Therefore $F'_k \geqslant V'_{\mu_n}(\mu)$. By construction of V_{μ_n} at any $\mu'' \geqslant \mu$ Equation (B.18) is solved with k, therefore $F'_k \geqslant V_{\mu_n}(\mu'')$ holds for all $\mu'' \geqslant \mu$. This implies $\forall \mu'' \geqslant \mu$, $V_{\mu_n}(\mu'') - F_k(\mu'') \leqslant V_{\mu_n}(\mu) - F_k(\mu) < V(\mu') - F_k(\mu') - \varepsilon$. Take $n \to \infty$ and $\mu'' = \mu'$, contradiction. Therefore, Equation (B.18) is solved at all $\mu \in (\mu^m, \mu^\star)$ for $V(\mu)$ with $k' > k$.

  Now consider $V_{\mu^m}(\mu)$. By my construction, suppose V_{μ^m} is updated up to action $k + 1$. I claim that $V_{\mu^m} = V$ when $\mu \in [\mu^m, \mu^\star)$. Suppose not true, then there exists μ at which $V_{\mu^m}(\mu) < V(\mu)$, $V'_{\mu^m}(\mu) < V'(\mu)$. It is easy to verify that Equation (B.18) is violated at

$V_{\mu_m}(\mu)$. Therefore, if $V_{\mu^m} \neq V$, it must happen in $(\mu^\star, b_m)$. Again we can find $\mu \in (\mu^\star, b_m)$ s.t. $V_{\mu^m}(\mu) < V(\mu)$, $V'_{\mu^m}(\mu) < V'(\mu)$, which is not possible. So $V(\mu) = V_{\mu^m}(\mu)$ on I_m.

To sum up, V can be represented as:

$$V(\mu) = \begin{cases} V_{\mu^*}(\mu) & \text{if } \mu \in [\mu^*, \mu^{**}] \\ V_{\mu^m}(\mu) & \text{if } \mu \in I^m \\ F(\mu) & \text{otherwise} \end{cases}$$

Now I prove smoothness of $V(\mu)$ on $[\mu^*, 1]$. By Lemma B.19 $|V'(\mu) - F'(\mu)|$ is bounded by $\Delta|I_n|$. Define:

$$V_n(\mu) = \begin{cases} V_{\mu^m}(\mu) & \text{when } \mu \in I_m,\, m \leqslant n \\ F(\mu) & \text{otherwise} \end{cases}$$

Then $V_n(\mu) \to V(\mu)$. By Lemma B.11, we can without loss assume first n V_{μ^m} have I_m covering $\underline{\mu}_m$. Fix n, $\forall \mu$, $\forall m \geqslant n$, if $\mu \in I_m$ and $m \leqslant n$ or $\mu \notin \bigcup I_m$, then $V'_n(\mu) = V'_m(\mu)$, else if $\mu \in I_m$, $m > n$, then $|V'_n(\mu) - F'(\mu)|$ and $|V'_m(\mu) - F'(\mu)|$ are all bounded by $\Delta|I_m|$. Therefore, $V'_n(\mu)$ is a Cauchy sequence. Then $V'_n(\mu) \to V'(\mu)$ pointwise. Since each V'_n is continuous, V is a smooth function on $[0, 1]$ and $V' = F'$ when $V = F$.

Unimprovability

Finally, I prove unimprovability of $V(\mu)$.

- *Step 1*: We first show that $V(\mu)$ solves the following problem:

$$V(\mu) = \max\left\{F(\mu), \max_{\nu,m,I} \frac{I}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)} - \frac{C(I)}{\rho}\right\} \tag{P-C1}$$
$$\begin{cases} \nu \geqslant \mu \text{ when } \mu \geqslant \mu^* \\ \nu \leqslant \mu \text{ when } \mu \leqslant \mu^* \end{cases}$$

We still focus on the case $\mu \geqslant \mu^*$. For the case $\mu \leqslant \mu^*$, a totally symmetric argument applies by referring to Lemma B.23′. Equation (P-C1) is implied by Equation (B.18) for $\mu \in E$. So it is sufficient to prove Equation (P-C1) for $\mu \in E^C$. Suppose there exists $\mu \geqslant \mu^*$ s.t. Equation (P-C) is violated. Let $F(\mu) = F_k(\mu)$. Then without loss we can assume that:

$$U(\mu) = \max_{\nu,k'>k,I} \frac{I}{\rho} \frac{F_k'(\nu) - F_k(\mu) - F_k'(\nu - \mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho} > F_k(\mu)$$

By Lemma B.11, there exists I_k s.t. $\underline{\mu}_k \in I_k$. At $b_k = \sup I_k$, $U(b_k) \leqslant F_k(b_k)$. Therefore, since $U(\mu)$ is continuous there exists largest $\mu' < \mu$ s.t .$U(\mu') = F_k(\mu')$. Then Equation (B.19) is satisfied at μ' so consider $V_{\mu'}$. Sicne $V_{\mu'}(\mu) \leqslant V(\mu) = F_k(\mu)$, there exists $\mu'' \in (\mu', \mu)$ s.t. $V_{\mu'}(\mu'') \leqslant F_k(\mu)$ and $V'_{\mu'}(\mu'') \leqslant F_k(\mu)$. Therefore $U(\mu'') > F_k(\mu'')$ implies $V_{\mu'}(\mu'') > F_k(\mu'')$, contradiction. Apply a symmetric argument to $\mu \leqslant \mu^*$, I proved Equation (P-C).

- *Step 2*: Then we show that $V(\mu)$ solves the following problem:

$$V(\mu) = \max\left\{F(\mu), \max_{\nu,I} \frac{I}{\rho} \frac{V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho}\right\} \tag{P-D1}$$

$$\begin{cases} \nu \geqslant \mu \text{ when } \mu \geqslant \mu^* \\ \nu \leqslant \mu \text{ when } \mu \leqslant \mu^* \end{cases}$$

Suppose not, then there exists:

$$\begin{aligned} \widetilde{V} &= \max_{\nu \geqslant \mu, I} \frac{I}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho} \\ &\leqslant V(\mu) < \frac{I''}{\rho} \frac{V(\mu'') - V(\mu) - V'(\mu)(\mu''-\mu)}{J(\mu,\mu'')} - \frac{C(I'')}{\rho} \end{aligned}$$

Suppose the optimizer is ν, m, I. Optimality implies Equation (B.22):

$$\frac{V(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)} = C'(I')$$

Together with Equation (B.20), we have $I'C'(I') = \rho\widetilde{V} + C(I')$. Then combine with Equation (B.21), we get:

$$\begin{cases} F_m' + C'(I)H'(\nu) = V'(\mu) + C'(I')H'(\mu) \\ (F_m(\nu) + C'(I')H(\nu)) - (V(\mu) + C'(I')H(\mu)) = (V'(\mu) + C'(I')H'(\mu))(\nu - \mu) \end{cases}$$

We define L and G as in Theorem 1.2. Then L will be linear and $G(F_m, C'(I'))(\nu)$ will be a concave function of ν. Consider:

$$L(V, C'(I'), \mu)(\nu) - G(F_m, C'(I'))$$

FOC implies that it will be convex and attains minimum 0 at ν. For any m' other than m,

$$L(V, C'(I'))(\nu) - G(F_{m'}, C'(I'))(\nu)$$

will be convex and weakly larger than zero. However:

$$\begin{aligned} & L(V, C'(I'), \mu)(\mu'') - G(V, C'(I'))(\mu'') \\ = & -(V(\mu'') - V(\mu) - V'(\mu)(\mu'' - \mu) - C'(I')J(\mu, \mu'')) \\ < & 0 \end{aligned}$$

The inequality is from definition of I':

$$\begin{aligned} & I'C'(I') - C(I') < I''C'(I'') - C(I'') \\ \implies & C'(I') < C'(I'') \\ \implies & C'(I') < \frac{V(\mu'') - V(\mu) - V'(\mu)(\mu'' - \mu)}{J(\mu, \mu'')} \end{aligned}$$

Therefore, $L(V, C'(I'), \mu)(\cdot) - G(V, C'(I'))(\cdot)$ will have a strictly negative minimum. Suppose it's minimized at $\widetilde{\mu}$, Then FOC implies:

$$V'(\mu) + C'(I')H'(\mu) = V'(\widetilde{\mu}) + C'(I')H'(\widetilde{\mu})$$

Consider:

$$\begin{aligned} & L\big(V, C'(I'), \widetilde{\mu}\big)(\nu(\widetilde{\mu})) - G\big(F_m, C'(I')\big)(\widetilde{\nu}) \\ = & L\big(V, C'(I'), \mu\big)(\nu(\widetilde{\mu})) - G\big(F_m, C'(I')\big)(\nu(\widetilde{\mu})) \\ & + V(\widetilde{\mu}) - V(\mu) + C'(I')(H(\widetilde{\mu}) - H(\mu)) - \big(V'(\mu) + C'(I')H(\mu)\big)(\widetilde{\mu} - \mu) \\ \geqslant & V(\widetilde{\mu}) - V(\mu) + C'(I')(H(\widetilde{\mu}) - H(\mu)) - \big(V'(\mu) + C'(I')H'(\mu)\big)(\widetilde{\mu} - \mu) \\ = & G\big(V, C'(I')\big)(\widetilde{\mu}) - L\big(V, C'(I'), \mu\big)(\widetilde{\mu}) \\ > & 0 \end{aligned}$$

Let $m', \nu(\widetilde{\mu}), \widetilde{I}$ be maximizer at $\widetilde{\mu}$, $\widetilde{I}C'(\widetilde{I}) = \rho V(\widetilde{\mu}) + C(\widetilde{I})$:

$$\begin{aligned} 0 = & L(V, C'(\widetilde{I}), \widetilde{\mu})(\nu(\widetilde{\mu})) - G(F_{m'}, C'(\widetilde{I}))(\nu(\widetilde{\mu})) \\ = & L(V, C'(I'), \widetilde{\mu})(\nu(\widetilde{\mu})) - G(F_{m'}, C'(I'))(\nu(\widetilde{\mu})) \\ & + (C'(\widetilde{I}) - C'(I'))J(\widetilde{\mu}, \nu(\widetilde{\mu})) \\ > & (C'(\widetilde{I}) - C'(I'))J(\widetilde{\mu}, \nu(\widetilde{\mu})) \end{aligned}$$

Since $\widetilde{\mu} > \mu$, we have $C'(\widetilde{I}) - C'(I') > 0$. Contradiction. Therefore we proved Equation (P-D1).

- *Step 3*: We show that V satisfies Equation (1.4). First, since V is smooth, envelope theorem implies:

$$\begin{aligned} V'(\mu) &= -\frac{I}{\rho}\frac{\nu-\mu}{J(\mu,\nu)}(V''(\mu)+C'(I)H''(\mu)) \\ &> 0 \\ \implies & V''(\mu)+C'(I)H''(\mu) < 0 \end{aligned}$$

Therefore, allocating to diffusion experiment will always be suboptimal. What's more, consider:

$$\begin{aligned} V^{-}(\mu) &= \max_{\nu\leqslant\mu,I}\frac{I}{\rho}\frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho} \\ \implies V'^{-}(\mu) &= -\frac{I}{\rho}\frac{\nu-\mu}{J(\mu,\nu)}(V''(\mu)+C'(I)H''(\mu)) \end{aligned}$$

$V^{-}(\mu^*) = V(\mu^*)$ and whenever $V(\mu) = V^{-}(\mu)$, we will have $V^{-\prime}(\mu) < 0$. Therefore, $V^{-}(\mu)$ can never cross from below, that is to say:

$$\begin{aligned} \rho V(\mu) = \max\Big\{\rho F(\mu), \max_{\nu,p,\sigma,I} p(V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)) + V''(\mu)\sigma^2 - C(I)\Big\} \\ \text{s.t. } pJ(\mu,\nu) + H''(\mu)\sigma^2 = I \end{aligned}$$

■

Lemma B.20. *Define $\overline{V}^{+}$ and $\overline{V}^{-}$:*

$$\overline{V}^{+}(\mu) = \max_{\nu\geqslant\mu,m,I}\frac{IF_m(\nu)-C(I)J(\mu,\nu)}{I+\rho J(\mu,\nu)}$$

$$\overline{V}^{-}(\mu) = \max_{\nu \leqslant \mu, m, I} \frac{IF_m(\nu) - C(I)J(\mu,\nu)}{I + \rho J(\mu,\nu)}$$

There exists $\mu^* \in [0,1]$ *s.t.* $\overline{V}^{+}(\mu) \geqslant \overline{V}^{-}(\mu)\ \forall \mu \geqslant \mu^*$; $\overline{V}^{+}(\mu) \leqslant \overline{V}^{-}(\mu)\ \forall \mu \leqslant \mu^*$. *Moreover* $\overline{V}(\mu^*) > 0$.

Proof. We define function U_m^+ and U_m^- as following:

$$\overline{U}_m^{+}(\mu) = \max_{\nu \geqslant \mu, I} \frac{IF_m(\nu) - C(I)J(\mu,\nu)}{I + \rho J(\mu,\nu)}$$
$$\overline{U}_m^{-}(\mu) = \max_{\nu \leqslant \mu, I} \frac{IF_m(\nu) - C(I)J(\mu,\nu)}{I + \rho J(\mu,\nu)}$$

Since $C(I)$, $F_m(\mu)$ and $J(\mu,\nu)$ are all smooth functions, the objective function will be smooth. First consider FOCs and SOCs:

$$\text{FOC-}\nu: F_m'\Big(1 + \frac{\rho}{I}J(\mu,\nu)\Big) - \Big(\frac{C(I)}{I} + \frac{\rho}{I}F_m(\nu)\Big)(H'(\mu) - H'(\nu)) = 0$$
$$\text{FOC-}I: \rho F_m(\nu) + C(I) - C'(I)(I + \rho J(\mu,\nu)) = 0$$
$$\text{SOC}: \quad H = \begin{bmatrix} I(\rho F_m(\nu) + C(I))(I + \rho J(\mu,\nu))H''(\nu) & 0 \\ 0 & -J(\mu,\nu)(I + \rho J(\mu,\nu))^2 C''(I) \end{bmatrix}$$

Noticing that SOC is evaluated at the pairs (ν, I) at which FOC holds.

Remark B.1. Details of calculation of second derivatives:

- $H_{\nu,\nu}$:

$$\begin{aligned}
&\frac{\partial^2}{\partial \nu^2}\frac{IF_m(\nu) - C(I)J(\mu,\nu)}{I + \rho J(\mu,\nu)} \\
=&\frac{1}{(I + \rho J(\mu,\nu))^3}\Big[2\rho^2(IF_m(\nu) - C(I)J(\mu,\nu))(H'(\mu) - H'(\nu))^2 \\
&- 2\rho(I + \rho J(\mu,\nu))(H'(\mu) - H'(\nu))\big(IF_m' - C(I)(H'(\mu) - H'(\nu))\big)
\end{aligned}$$

$$
\begin{aligned}
&+\rho(I+\rho J(\mu,\nu))(IF_m(\nu)-C(I)J(\mu,\nu))H''(\nu)\\
&+(I+\rho J(\mu,\nu))^2C(I)H''(\nu)\Big]
\end{aligned}
$$

$$
\begin{aligned}
\text{FOC-}\nu \implies & F_m' = \frac{(C(I)+\rho F_m(\nu))(H'(\mu)-H'(\nu))}{I+\rho J(\mu,\nu)}\\
\implies & \frac{\partial^2}{\partial \nu^2}\frac{IF_m(\nu)-C(I)J(\mu,\nu)}{I+\rho J(\mu,\nu)}\\
=& \frac{1}{(I+\rho J(\mu,\nu))^3}\Big[2\rho^2(IF_m(\nu)-C(I)J(\mu,\nu))(H'(\mu)-H'(\nu))^2\\
&+2\rho(I+\rho J(\mu,\nu))C(I)(H'(\mu)-H'(\nu))^2\\
&-2\rho(C(I)+\rho F_m(\nu))(H'(\mu)-H'(\nu))^2\\
&+\rho(I+\rho J(\mu,\nu))(IF_m(\nu)-C(I)J(\mu,\nu))H''(\nu)\\
&+(I+\rho J(\mu,\nu))^2C(I)H''(\nu)\Big]\\
=& \frac{1}{(I+\rho J(\mu,\nu))^3}\Big[\rho(I+\rho J(\mu,\nu))(IF_m(\nu)-C(I)J(\mu,\nu))H''(\nu)\\
&+(I+\rho J(\mu,\nu))^2C(I)H''(\nu)\Big]\\
=&(I+\rho J(\mu,\nu))H''(\nu)(\rho IF_m(\nu)-\rho C(I)J(\mu,\nu)+IC(I)+\rho C(I)J(\mu,\nu))\\
=& \frac{I(\rho F_m(\nu)+C(I))(I+\rho J(\mu,\nu))H''(\nu)}{(I+\rho J(\mu,\nu))^3}
\end{aligned}
$$

- $H_{I,I}$:

$$
\begin{aligned}
& \frac{\partial^2}{\partial I^2}\frac{IF_m(\nu)-C(I)J(\mu,\nu)}{I+\rho J(\mu,\nu)}\\
=& \frac{1}{(I+\rho J(\mu,\nu))^3}\Big[2(IF_m(\nu)-C(I)J(\mu,\nu))\\
&-2(I+\rho J(\mu,\nu))(F_m(\nu)-C'(I)J(\mu,\nu))\\
&-J(\mu,\nu)(I+\rho J(\mu,\nu))^2C''(I)\Big]
\end{aligned}
$$

$$
\text{FOC-}I \implies IF_m(\nu)-C(I)J(\mu,\nu) = (I+\rho J(\mu,\nu))(F_m(\nu)-C'(I)J(\mu,\nu))
$$

$$
\begin{aligned}
\implies & \frac{\partial^2}{\partial I^2}\frac{IF_m(\nu)-C(I)J(\mu,\nu)}{I+\rho J(\mu,\nu)} \\
=&\frac{1}{(I+\rho J(\mu,\nu))^3}\Big[2(I+\rho J(\mu,\nu))(F_m(\nu)-C'(I)J(\mu,\nu)) \\
&-2(I+\rho J(\mu,\nu))(F_m(\nu)-C'(I)J(\mu,\nu)) \\
&-J(\mu,\nu)(I+\rho J(\mu,\nu))^2C''(I)\Big] \\
=&\frac{-J(\mu,\nu)(I+\rho J(\mu,\nu))^2C''(I)}{(I+\rho J(\mu,\nu))^3}
\end{aligned}
$$

- $H_{\nu,I}$:

$$
\begin{aligned}
&\frac{\partial^2}{\partial I\partial \nu}\frac{IF_m(\nu)-C(I)J(\mu,\nu)}{I+\rho J(\mu,\nu)} \\
=&\frac{1}{(I+\rho J(\mu,\nu))^3}\Big[2\rho(IF_m(\nu)-C(I)J(\mu,\nu))(H'(\mu)-H'(\nu)) \\
&-\rho(I+\rho J(\mu,\nu))(F_m(\nu)-C'(I)J(\mu,\nu))(H'(\mu)-H'(\nu)) \\
&-(I+\rho J(\mu,\nu))(IF_m'-C(I)(H'(\mu-H'(\nu)))) \\
&+(I+\rho J(\mu,\nu))^2\big(F_m'-C'(I)(H'(\mu)-H'(\nu))\big)\Big] \\
=&\frac{1}{(I+\rho J(\mu,\nu))^3}\Big[2\rho(IF_m(\nu)-C(I)J(\mu,\nu))(H'(\mu)-H'(\nu)) \\
&-\rho(IF_m(\nu)-C(I)J(\mu,\nu))(H'(\mu)-H'(\nu)) \\
&-(I+\rho J(\mu,\nu))(I\frac{(C(I)+\rho F_m(\nu))(H'(\mu)-H'(\nu))}{I+\rho J(\mu,\nu)}C(I)(H'(\mu-H'(\nu)))) \\
&+(I+\rho J(\mu,\nu))^2\left(\frac{(C(I)+\rho F_m(\nu))(H'(\mu)-H'(\nu))}{I+\rho J(\mu,\nu)}-C'(I)(H'(\mu)-H'(\nu))\right)\Big] \\
=&\frac{H'(\mu)-H'(\nu)}{(I+\rho J(\mu,\nu))^3}(\rho IF_m(\nu)-\rho C(I)J(\mu,\nu)-I(C(I)+\rho F_m(\nu)) \\
&+(I+\rho J(\mu,\nu))C(I)+(I+\rho J(\mu,\nu)(C(I)+\rho F_m(\nu)))-(I+\rho J(\mu,\nu))^2C'(I)\big) \\
=&0
\end{aligned}
$$

The only term we don't know its sign is

$$\rho F_m(\nu) + C(I) = \frac{I + \rho J(\mu, \nu)}{H'(\mu) - H'(\nu)} F_m'$$

Therefore, H will be ND if $\nu > \mu$ and $F_m' > 0$, or $\nu < \mu$ and $F_m' < 0$. In these cases, FOC uniquely characterizes the maximum. Suppose $\nu > \mu$ and $F_m' < 0$ or $\nu < \mu$ and $F_m' > 0$, the H will never be ND, and choice of ν will be on boundary. What's more, simple calculation shows that choosing $\nu = \mu$ will dominate choosing $\nu = 0, 1$. Therefore:

$$\overline{U}_m^+(\mu) = F_m(\mu) \text{ when } F_m' < 0$$
$$\overline{U}_m^-(\mu) = F_m(\mu) \text{ when } F_m' > 0$$

When $F_m' > 0$, envelope condition implies:

$$\frac{d}{d\mu}\overline{U}_m^+(\mu) = \frac{-H''(\mu)(\nu - \mu)\big(C(I) + \frac{\rho}{I}F_m(\nu)\big)}{\big(1 + \frac{\rho}{I}J(\mu, \nu)\big)^2} > 0$$

Similarly, when $F_m' < 0$, envelope condition implies:

$$\frac{d}{d\mu}\overline{U}_m^-(\mu) = \frac{-H''(\mu)(\nu - \mu)\big(C(I) + \frac{\rho}{I}F_m(\nu)\big)}{\big(1 + \frac{\rho}{I}J(\mu, \nu)\big)^2} < 0$$

Therefore, $\overline{U}_m^+$ and $\overline{U}_m^-$ have exactly the same properties as in Lemma A.2, the rest of proofs simply follow Lemma A.2. What's more, we define ν_m^* and I_m^* as the maximizer in this problem.

Now I prove that $\overline{V}(\mu^*) > 0$. We know that $\overline{V}(\mu^*)$ solves:

$$\overline{V}(\mu^*) = \max_{\nu \geqslant \mu^*, I} \frac{F(\nu) - \frac{C(I)}{I}J(\mu^*, \nu)}{1 + \frac{\rho}{I}J(\mu^*, \nu)} = \max_{\nu \leqslant \mu^*, I} \frac{F(\nu) - \frac{C(I)}{I}J(\mu^*, \nu)}{1 + \frac{\rho}{I}J(\mu^*, \nu)}$$

Consider the following term:

$$\underline{V} = \max_{\mu_i, p_i, I} \frac{p_1 F(\mu_1) + p_2 F(\mu_2) - \frac{C(I)}{I} I(\mu_i|\mu^*)}{1 + \frac{\rho}{I} I(\mu_i|\mu^*)}$$

Suppose μ_i, p_i, I solves $\underline{V}$. Then:

$$\frac{\rho}{I}\underline{V} + \frac{C(I)}{I} = \frac{p_1 F(\mu_1) + p_2 F(\mu_2) - \underline{V}}{-p_1 H(\mu_1) - p_2 H(\mu_2) + H(\mu^*)}$$

I want to claim that $\underline{V} \leqslant \overline{V}(\mu^*)$. Suppose not, then:

$$\begin{aligned}\frac{p_1 F(\mu_1) + p_2 F(\mu_2) - \overline{V}(\mu^*)}{-p_1 H(\mu_1) - p_2 H(\mu_2) + H(\mu^*)} &> \frac{p_1 F(\mu_1) + p_2 F(\mu_2) - \underline{V}}{-p_1 H(\mu_1) - p_2 H(\mu_2) + H(\mu^*)} \\ &\geqslant \frac{\rho}{I}\overline{V}(\mu^*) + \frac{C(I)}{I}\end{aligned}$$

Then at least one of the following:

$$\frac{F(\mu_1) - \overline{V}(\mu^*)}{-H(\mu_1) + H(\mu^*) + H'(\mu^*)(\mu_1 - \mu^*)}; \frac{F(\mu_2) - \overline{V}(\mu^*)}{-H(\mu_2) + H(\mu^*) + H'(\mu^*)(\mu_2 - \mu^*)}$$

is larger than $\frac{\rho}{I}\overline{V}(\mu^*) + \frac{C(I)}{I}$. Suppose the fisrt term does, then:

$$\rho\overline{V}(\mu^*) < I\frac{F(\mu_1) - \overline{V}(\mu^*)}{J(\mu^*, \mu_1)} - C(I)$$

Contradicting optimality of $\overline{V}(\mu^*)$. Same argument applies to the second term. So $\overline{V}(\mu^*) \geqslant \underline{V}$. However:

$$\lim_{c\to 0} p_1 F(\mu_1) + p_2 F(\mu_2) - \frac{C(I)}{I} I(\mu_i|\mu^*) = p_1 F(\mu_1) + p_2 F(\mu_2) - C'(0) I(\mu_i|\mu) > 0$$

Therefore, $\overline{V}(\mu^*) \geqslant \underline{V} > 0$. ■

Lemma B.21. *Assume $\mu_0 \geqslant \mu^*$, $F_m' \geqslant 0$, V_0, V_0' satisfies:*

$$\begin{cases} \overline{V}(\mu_0) \geqslant V_0 > F_m(\mu_0) \\ V_0 = \max\limits_{\nu \geqslant \mu_0, I} \dfrac{I}{\rho} \dfrac{F_m(\nu) - V_0 - V_0'(\nu - \mu)}{J(\mu,\nu)} - \dfrac{C(I)}{\rho} \end{cases}$$

Then there exists a $C^{(1)}$ smooth and strictly increasing $V(\mu)$ defined on $[\mu_0, 1]$ satisfying:

$$V(\mu) = \max_{\nu \geqslant \mu, I} \frac{I}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho} \tag{A.25-c}$$

and initial condition $V(\mu_0) = V_0$, $V'(\mu_0) = V_0'$.

Proof. We start from deriving FOC and SOC for Equation (A.25-c):

$$\text{FOC-}\nu\text{: } \frac{I}{\rho}\left(\frac{F_m' - V'(\mu)}{J(\mu,\nu)} + \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{J(\mu,\nu)^2}(H'(\nu) - H'(\mu))\right) = 0$$

$$\text{FOC-}I\text{: } \frac{1}{\rho}\left(\frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{J(\mu,\nu)} - C'(I)\right) = 0$$

$$\text{SOC:} \quad H = \begin{bmatrix} \frac{-2(H'(\mu)-H'(\nu))(\text{FOC-}\nu)}{J(\mu,\nu)} + \frac{I}{\rho}\frac{(F_m(\nu)-V(\mu)-V'(\mu)(\nu-\mu))H''(\nu)}{J(\mu,\nu)^2} & \frac{1}{I}\text{FOC-}\nu \\ \frac{1}{I}\text{FOC-}\nu & -\frac{C''(I)}{\rho} \end{bmatrix}$$

Noticing that $H_{I,I} < 0$, therefore I satisfying FOC will be unique given μ, ν. On the other hand, FOC-ν is independent of I. $H_{\nu,\nu} < 0$ when FOC-$\nu \geqslant 0$. Therefore, solution of F)C-ν will be unique. When FOCs are satisfied, H is strictly ND, then the solution of FOCs are going to be maximizer. Therefore, FOC-ν and FOC-I uniquely characterize optimal choice of ν, I. Now we impose feasibility:

$$V(\mu) = \frac{I}{\rho} \frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho} \tag{B.20}$$

FOCs reduces to:

$$\text{FOC-}\nu\text{:}(F_m' - V'(\mu)) + \frac{\rho V(\mu) + C(I)}{I}(H'(\nu) - H'(\mu)) = 0 \tag{B.21}$$

$$\text{FOC-}I\text{:}\, IC'(I) = \rho V(\mu) + C(I) \tag{B.22}$$

Differentiate FOC-I, we get:

$$\begin{cases} V(\mu) = \dfrac{IC'(I) - C(I)}{\rho} \\ V'(\mu) = \dfrac{IC''(I)}{\rho}\dot{I} \end{cases} \tag{B.23}$$

Plug Equation (B.23) into Equation (B.21) and Equation (B.20):

$$\begin{cases} \dot{I} = \dfrac{\rho}{IC''(I)}\left(F_m' + C'(I)(H'(\nu) - H'(\mu))\right) \\ J(\nu, \mu) = \dfrac{1}{\rho}\left(I - \dfrac{C(I) + \rho F_m(\mu)}{C'(I)}\right) \end{cases} \tag{B.24}$$

We obtained an equation system with one ODE of $(c, \dot{I})$ and one regular equation for ν. Since $J(\nu, \mu)$ is strictly monotonic for $\nu \geqslant \mu$, we can also define an implicit inverse function M to eliminate ν in the equation.

$$J(M(y, \mu), \mu) = y$$

Therefore we get an ODE:

$$\dot{I} = \frac{\rho}{IC''(I)}\left(F_m' + C'(I)\left(H'\left(M\left(\frac{1}{\rho}\left(I - \frac{C(I) + \rho F_m(\mu)}{C'(I)}\right)\right), \mu\right) - H'(\mu)\right)\right) \tag{B.25}$$

We define $\underline{I}_m(\mu_0)C'(\underline{I}_m(\mu_0)) - C(\underline{I}_m(\mu_0)) = \rho F_m(\mu)$ when this equation has solution and $\underline{I}_m(\mu) = 0$ when $\rho F_m(\mu)$ is so small that this equation has no solution. Since $F_m(\mu)$ is increasing in μ, $\underline{I}_m(\mu)$ is increasing and strictly increasing when $\underline{I}_m(\mu) > 0$. We consider the initial conditions:

$$\begin{aligned} F_m(\mu_0) < V_0 = \frac{I_0 C'(I_0) - C(I_0)}{\rho} \leqslant \overline{V}(\mu_0) \\ \implies \underline{I}_m(\mu_0) < I_0 \leqslant I_m^*(\mu_0) \end{aligned}$$

Then Lemma B.22 guaranteed the existence of an increasing function $I(\mu)$ on $[\mu_0, 1]$. ■

Lemma B.22. *Define M as $J(M(y,\mu),\mu) = y$. Assume $\mu_0 \in [\mu^*, 1)$, I_0 satisfies:*

$$\underline{I}_m(\mu_0) < I_0 \leqslant I_m^*(\mu_0)$$

Then there exists a $C^{(1)}$ and strictly increasing I on $[\mu_0, 1]$ satisfying initial condition $I(\mu_0) = I_0$. On $\{\mu | I(\mu) > \underline{I}_m(\mu)\}$, I solves:

$$\dot{I} = \frac{\rho}{IC''(I)}\left(F_m' + C'(I)\left(H'\left(M\left(\frac{1}{\rho}\left(I - \frac{C(I) + \rho F_m(\mu)}{C'(I)}\right)\right), \mu\right) - H'(\mu)\right)\right) \tag{B.25}$$

Proof. We first characterize some useful properties of the ODE. We denote the ODE by $\dot{I} = R(\mu, I)$.

- *Domain*: By definition of $\underline{I}_m(\mu)$, $\forall \mu \in (0,1)$

$$\underline{I}_m(\mu) - \frac{C(\underline{I}_m(\mu)) + \rho F_m(\mu)}{C'(\underline{I}_m(\mu))} = 0$$

  Since $\underline{I}_m \geqslant 0$, then $C(\underline{I}_m(\mu)) + \rho F_m(\mu) \geqslant 0$. Therefore at $I = \underline{I}_m(\mu)$:

$$\frac{\partial}{\partial I}\left(I - \frac{C(I) + \rho F_m(\mu)}{C'(I)}\right) = \frac{C(I) + \rho F_m(\mu)}{C'(I)^2} C''(I) > 0$$

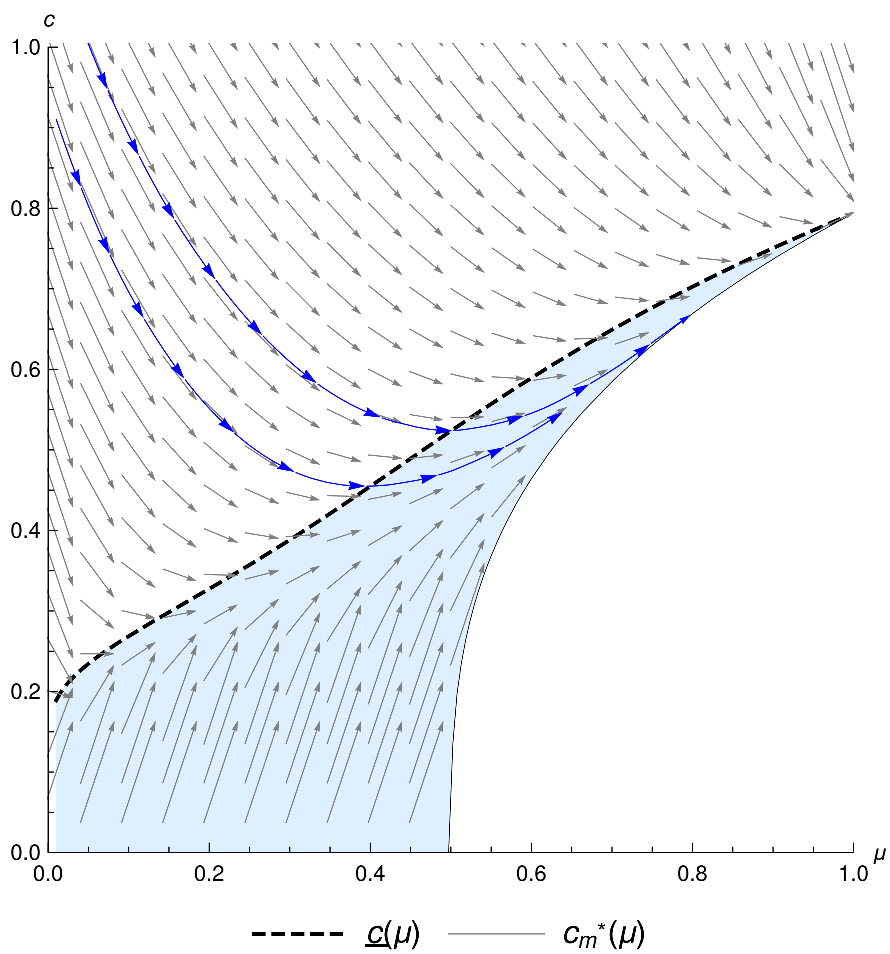


Figure B.2: Phase diagram of $\left(\dot{\mu}, \dot{I}\right)$.

Therefore, $\forall I \geqslant \underline{I}_m(\mu)$, $I - \frac{C(I)+\rho F_m(\mu)}{C'(I)} \geqslant 0$. Strictly inequality holds when $I > \underline{I}_m(\mu)$. On the other hand, when $I < \underline{I}_m(\mu)$, if $F_m(\mu) \geqslant 0$, then $I - \frac{C(I)+\rho F_m(\mu)}{C'(I)} < 0$. Else if $F_m(\mu) \leqslant 0$, then $\underline{I}_m(\mu) = 0$. Since M only applies to non-negative reals, we know that the ODE is only well defined in the region: $\{I | I \geqslant \underline{I}_m(\mu)\}$.

- *Continuity*: It is not hard to verify that the ODE is well behaved (satisfying Picard-Lindelof) when μ is strictly bounded away from $\{0, 1\}$, I is uniformly bounded away from $\underline{I}_m(\mu)$. One just need to calculate $M_y(y, \mu)$ by implied function theorem:

$$\frac{\partial}{\partial y} M_y(y, \mu) = -\frac{1}{H''(M(y,\mu))(M(y,\mu) - \mu)}$$

$M(y, \mu) = \mu$ implies $J(\nu, \mu) = 0$, implies $\frac{1}{\rho}\left(I - \frac{C(I)+\rho F_m(\mu)}{C'(I)}\right) = 0$. Since I is uniformly bounded away from $\underline{I}_m(\mu)$, then $M(y, \mu) - \mu$ is uniformly bounded away from 0.

- *Monotonicity*: When $I = I^*_m(\mu)$, $\dot{I} = 0$. This can be shown by considering FOC on I^*_m:

$$\begin{cases} F'_m - C'(I)(H'(\mu) - H'(\nu)) = 0 \\ (I + \rho J(\mu, \nu))C'(I) = C(I) + \rho F_m(\nu) \end{cases}$$

$$\Longrightarrow (I - \rho J(\nu, \mu))C'(I) = C(I) + \rho F_m(\mu) + \rho F'_m(\nu - \mu) + C'(I)(H'(\nu) - H'(\mu))(\nu - \mu)$$

$$\Longrightarrow (I - \rho J(\nu, \mu))C'(I) = C(I) + \rho F_m(\mu)$$

$$\Longrightarrow F'_m + C'(I)\left(H'\left(M\left(\frac{1}{\rho}\left(I - \frac{C(I) + \rho F_m(\mu)}{C'(I)}\right), \mu\right)\right) - H'(\mu)\right) = 0$$

$$\Longrightarrow \dot{I} = R(\mu, I) = 0$$

Then we consider the monotonicity of $R(\mu, I)$:

$$\frac{\partial}{\partial I} R(\mu, I) = C''(I)\big(H'(M) - H'(\mu)\big) + C'(I)\frac{H''(M)}{H''(M)(\mu - M)}\frac{1}{\rho}\frac{C(I) + \rho F_m(\mu)}{C'(I)^2} C''(I) < 0$$

Therefore, $R(\mu, I)$ will be positive in $\{\underline{I}_m(\mu) < I \leqslant I_m^*(\mu)\}$. This refers to the blue region in Figure B.2.

$\forall\delta > 0$, we consider solving the ODE $\dot{I} = R(\mu, I)$ in region: $\mu \in [\delta, 1-\delta], I \in [\underline{I}_m(\mu) + \delta, I_m^*(\mu)]$. The initial condition (μ_0, I_0) is in the blue region of Figure B.2. (When H' is finite, we can take $\mu \in [0, 1]$.) Picard-Lindelof guarantees a unique solution satisfying the ODE in the region. What's more, it's straight forward that the solution $I(\mu)$ will be increasing. A solution is a blue line with arrows in Figure B.2. A solution $I(\mu)$ will lie between $\underline{I}_m(\mu)$ and $I_m^*(\mu)$ until it hits the boundary of region.

Now we can take $\delta \to 0$ and extend $I(\mu)$ towards the boundary. Since the end point of $I(\mu)$ has both μ, I monotonically increasing, there is a limit $\bar{I}, \bar{\mu}$ with $\underline{I}_m(\bar{\mu}) = \bar{I}$. Then since $R(\mu, I)$ has a limit $\frac{\rho F_m'}{\bar{I}h''(\bar{I})}$, we actually have $\lim_{\mu\to\bar{\mu}} V'(\mu) = F_m'$ by Equation (B.23). So the resulting $V(\mu)$ calculated from

$$V(\mu) = \frac{I(\mu)C'(I(\mu)) - C(I(\mu))}{\rho}$$

will be smooth on $[\mu_0, 1]$. ■

Lemma B.21′. *Assume $\mu_0 \leqslant \mu^*$, $F_m' \geqslant 0$, V_0, V_0' satisfies:*

$$\begin{cases} \overline{V}(\mu_0) \geqslant V_0 > F_m(\mu_0) \\ V_0 = \max\limits_{\nu\leqslant\mu_0, I} \frac{I}{\rho}\frac{F_m(\nu) - V_0 - V_0'(\nu-\mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho} \end{cases}$$

Then there exists a $C^{(1)}$ smooth and strictly decreasing $V(\mu)$ defined on $[0, \mu_0]$ satisfying:

$$V(\mu) = \max_{\nu\leqslant\mu, I} \frac{I}{\rho}\frac{F_m(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho} \qquad \text{(A.25-I')}$$

and initial condition $V(\mu_0) = V_0$, $V'(\mu_0) = V_0'$.

Lemma B.22′. *Define M as $J(M(y,\mu),\mu) = y$. Assume $\mu_0 \in (0,\mu^*]$, I_0 satisfies:*

$$\underline{I}_m(\mu_0) < I_0 \leqslant I_m^*(\mu_0)$$

Then there exists a $C^{(1)}$ and strictly decreasing I on $[0,\mu_0]$ satisfying initial condition $c(\mu_0) = I_0$. On $\{\mu | I(\mu) > \underline{I}_m(\mu)\}$, I solves:

$$\dot{I} = \frac{\rho}{IC''(I)}\left(F_m' + C'(I)\left(H'\left(M\left(\frac{1}{\rho}\left(I - \frac{C(I)+\rho F_m(\mu)}{C'(I)}\right)\right),\mu\right) - H'(\mu)\right)\right) \tag{B.25′}$$

Lemma B.23. *Suppose at μ_0, V_0, V_0', $k \geqslant 1$ satisfies:*

$$\begin{cases} V_0 = \max\limits_{\nu \geqslant \mu_0, I} \dfrac{I}{\rho}\dfrac{F_{m-k}(\nu) - V_0 - V_0'(\nu-\mu_0)}{J(\mu_0,\nu)} - \dfrac{C(I)}{\rho} \geqslant \max\limits_{\nu \geqslant \mu_0, I} \dfrac{I}{\rho}\dfrac{F_m(\nu) - V_0 - V_0'(\nu-\mu_0)}{J(\mu_0,\nu)} - \dfrac{C(I)}{\rho} \\ \overline{V}(\mu_0) \geqslant V_0 \geqslant F_{m-k}(\mu_0) \end{cases}$$

V_{m-k} is the solution as defined in Lemma B.22 with initial condition μ_0, V_0, V_0', then $\forall \mu \in [\mu_0, \nu(\mu_0)]$:

$$V_{m-k}(\mu) \geqslant \max_{\nu \geqslant \mu, m' \in [m-k,m], I} \frac{I}{\rho}\frac{F_{m'}(\nu) - V_{m-k}(\mu) - V_{m-k}'(\mu)(\nu-\mu)}{J(\mu,\nu)} - \frac{C(I)}{\rho}$$

Proof. We first show that:

$$V_0 \geqslant \max_{\nu \in [\mu_0, \overline{\mu}_m], I} \frac{I}{\rho}\frac{V_{m-k}(\nu) - V_0 - V_0'(\nu-\mu_0)}{J(\mu_0,\nu)} - \frac{C(I)}{\rho}$$

Suppose not, then there exists ν, I' s.t.

$$\begin{cases} V_0 < \dfrac{I'}{\rho}\dfrac{V_{m-k}(\nu) - V_0 - V_0'(\nu-\mu_0)}{J(\mu_0,\nu)} - \dfrac{C(I')}{\rho} \\ \dfrac{V_{m-k}(\nu) - V_0 - V_0'(\nu-\mu_0)}{J(\mu_0,\nu)} = C'(I') \end{cases} \tag{B.26}$$

Let $I_0 C'(I_0) = \rho V_0 + C(I_0)$, then optimality implies Equation (B.20) and Equation (B.21) at $\mu = \mu_0$:

$$\begin{cases} F'_{m-k} + C'(I_0)H'(\nu(\mu)) = V'_0 + C'(I_0)H'(\mu) \\ \big(F_{m-k}(\nu(\mu)) + C'(I_0)H(\nu(\mu))\big) - \big(V_0 + C'(I_0)H(\mu)\big) = \big(V'_0 + C'(I_0)H'(\mu)\big)(\nu(\mu) - \mu) \end{cases}$$

We define $L(V, \lambda, \mu)(\nu)$ and $G(V, \lambda)(\mu)$ as Equation (B.16), Equation (B.17). Consider:

$$L\big(V_{m-k}, C'(I_0), \mu_0\big)(\nu) - G\big(F_{m-k}, C'(I_0)\big)(\nu)$$

L is a linear function and G is a concave function. Therefore this is a convex function and have unique minimum determined by FOC. Simple calculation shows that it is minimized at $\nu(\mu_0)$ and the minimal value is 0. Now consider

$$\begin{aligned} & L\big(V_{m-k}, C'(I_0), \mu_0\big)(\nu) - G\big(V_{m-k}, C'(I_0)\big)(\nu) \\ = & -\left(V_{m-k}(\nu) - V_{m-k}(\mu_0) - V'_{m-k}(\mu_0)(\nu - \mu_0) - C'(I_0)J(\mu_0, \nu)\right) \\ < & 0 \end{aligned}$$

The inequality is from Equation (B.26) and definition of I_0:

$$\begin{aligned} & I_0 C'(I_0) - C(I_0) < I'C'(I') - C(I') \\ \Longrightarrow & C'(I_0) < C'(I') \\ \Longrightarrow & C'(I_0) < \frac{V_{m-k}(\nu) - V_0 - V'_0(\nu - \mu_0)}{J(\mu_0, \nu)} \end{aligned}$$

Therefore $L(V_{m-k}, C'(I_0), \mu_0)(\nu) - G(V_{m-k}, C'(I_0))(\nu)$ will be strictly negative at ν and will have minimum strictly negative. Suppose it's minimized at μ'' ($\mu'' > \mu_0$), then FOC implies:

$$V'_{m-k}(\mu_0) + C'(I_0)H'(\mu_0) = V'_{m-k}(\mu'') + C'(I_0)H(\mu'')$$

Let $I''C'(I'') = \rho V_{m-k}(\mu'') + C(I'')$, then we have $I'' > I_0$ and $C'(I'') > C'(I_0)$. Consider:

$$\begin{aligned}
&L(V_{m-k}, C'(I_0), \mu'')(\nu(\mu'')) - G(F_{m-k}, C'(I_0))(\nu(\mu'')) \\
=&L(V_{m-k}, C'(I_0), \mu_0)(\nu(\mu'')) - G(F_{m-k}, C'(I_0))(\nu(\mu'')) \\
&+ V_{m-k}(\mu'') - V_{m-k}(\mu_0) + C'(I_0)(H(\mu'') - H(\mu_0)) - (V'(\mu_0) + C'(I_0))(\mu'' - \mu_0) \\
\geqslant& V_{m-k}(\mu'') - V_{m-k}(\mu_0) + C'(I_0)(H(\mu'') - H(\mu_0)) - (V'(\mu_0) + C'(I_0))(\mu'' - \mu_0) \\
=&G(V_{m-k}, C'(I_0))(\mu'') - L(V_{m-k}, C'(I_0), \mu_0)(\mu'') > 0
\end{aligned}$$

However:

$$\begin{aligned}
0 =&L(V_{m-k}, C'(I''), \mu'')(\nu(\mu'')) - G(F_{m-k}, C'(I''))(\nu(\mu'')) \\
=&L(V_{m-k}, C'(I_0), \mu'')(\nu(\mu'')) - G(F_{m-k}, C'(I_0))(\nu(\mu'')) \\
&+ (C'(\mu'') - C'(I_0))(H(\mu'') - H(\nu(\mu'')) + H'(\mu'')(\nu(\mu'') - \mu'')) \\
>&(C'(I'') - C'(I_0))J(\mu'', \nu(\mu'')) > 0
\end{aligned}$$

Contradiction.

Now we show Lemma B.23. Suppose it's not true, then there exists $\nu \in (\mu_0, \nu(\mu_0))$, $\mu'' \geqslant \underline{\mu}_m$, and I'' s.t.

$$\begin{cases}
V_{m-k}(\nu) < \dfrac{I''}{\rho} \dfrac{F_{m'}(\mu'') - V'_{m-k}(\nu) - V'_{m-k}(\nu)(\mu'' - \nu)}{J(\nu, \mu'')} - \dfrac{C(I'')}{\rho} \\
\dfrac{F_{m'}(\mu'') - V'_{m-k}(\nu) - V'_{m-k}(\nu)(\mu'' - \nu)}{J(\nu, \mu'')} = C'(I'')
\end{cases}$$

If we let $I'C'(I') = \rho V(\nu) + C(I')$, then $I' > I_0$ and $C'(I') > C'(I_0)$. By definition:

$$
\begin{aligned}
0 \leqslant & L(V_{m-k}, C'(I_0), \mu_0)(\mu'') - G(F_{m'}, C'(I_0))(\mu'') \\
= & L(F_{m-k}, C'(I_0), \nu(\mu_0))(\mu'') - G(F_{m'}, C'(I_0))(\mu'') \\
0 \leqslant & L(V_{m-k}, C'(I_0), \mu_0)(\nu) - G(F_{m'}, C'(I_0))(\nu) \\
= & L(F_{m-k}, C'(I_0), \nu(\mu_0))(\nu) - G(F_{m'}, C'(I_0))(\nu) \\
\implies & L(F_{m-k}, C'(I'), \nu(\mu_0))(\mu'') - G(F_{m'}, C'(I'))(\mu'') \\
= & L(F_{m-k}, C'(I_0), \nu(\mu_0))(\mu'') - G(F_{m'}, C'(I_0))(\mu'') \\
& + (C'(I') - C'(I_0))J(\mu_0, \mu'') \\
> & 0 \\
& L(F_{m-k}, C'(I'), \nu(\mu_0))(\mu'') - G(F_{m'}, C'(I'))(\mu'') \\
= & L(F_{m-k}, C'(I_0), \nu(\mu_0))(\nu) - G(F_{m'}, C'(I_0))(\nu) \\
& + (C'(I') - C'(I_0))J(\mu_0, \nu) > 0
\end{aligned}
$$

No we consider $L(V_{m-k}, C'(I'), \nu)(\cdot)$ and $L(F_{m-k}, C'(I'), \nu(\mu_0))(\cdot)$:

$$
\begin{aligned}
& \begin{cases}
L(V_{m-k}, C'(I'), \nu)(\nu) = G(V_{m-k}, C'(I'))(\nu) \\
L(v_{m-k}, C'(I'), \nu)(\nu(\mu_0)) \geqslant G(V_{m-k}, C'(I'))(\nu(\mu_0)) \\
L(F_{m-k}, C'(I'), \nu(\mu_0))(\nu) > G(V_{m-k}, C'(I'))(\nu) \\
L(F_{m-k}, C'(I'), \nu(\mu_0))(\nu(\mu_0)) = G(V_{m-k}, C'(I'))(\nu(\mu_0))
\end{cases} \\
\implies & \begin{cases}
L(V_{m-k}, C'(I'), \nu)(\nu(\mu_0)) \geqslant L(F_{m-k}, C'(I'), \nu(\mu_0))(\nu(\mu_0)) \\
L(V_{m-k}, C'(I'), \nu)(\nu) < L(F_{m-k}, C'(I'), \nu(\mu_0))(\nu)
\end{cases}
\end{aligned}
$$

Since both functions are linear: $\frac{\mathrm{d}}{\mathrm{d}\mu} L(V_{m-k}, C'(I'), \nu)(\mu) > \frac{\mathrm{d}}{\mathrm{d}\mu} L(F_{m-k}, C'(I')\nu(\mu_0))(\mu)$, then

$L(V_{m-k}, C'(I'), \nu)(\cdot)$ must be larger than $L(F_{m-k}, C'(I'), \nu(\mu_0))(\cdot)$ at any $\mu'' \geqslant \nu(\mu_0)$. This implies:

$$L(V_{m-k}, C'(I'), \nu)(\mu'') > G(F_{m'}, C'(I'))(\mu'')$$

Contradicting the assumption. ∎

Lemma B.23′. *Suppose at μ_0, V_0, V_0', $k \geqslant 1$ satisfies:*

$$\begin{cases} V_0 = \max_{\nu \leqslant \mu_0, I} \frac{I}{\rho} \frac{F_{m+k}(\nu) - V_0 - V_0'(\nu - \mu)}{J(\mu_0, \nu)} - \frac{C(I)}{\rho} \geqslant \max_{\nu \leqslant \mu_0, I} \frac{I}{\rho} \frac{F_m(\nu) - V_0 - V_0'(\nu - \mu_0)}{J(\mu_0, \nu)} - \frac{C(I)}{\rho} \\ \overline{V}(\mu_0) \geqslant V_0 \geqslant F_{m+k}(\mu_0) \end{cases}$$

V_{m+k} is the solution as defined in Lemma B.22 with initial condition μ_0, V_0, V_0', then $\forall \mu \in [\nu(\mu_0), \mu_0]$:

$$V_{m+k}(\mu) \geqslant \max_{\nu \leqslant \mu, \nu \in [m, m+k], I} \frac{I}{\rho} \frac{F_{m'}(\nu) - V_{m-k}(\mu) - V_{m-k}'(\mu)(\nu - \mu)}{J(\mu, \nu)} - \frac{C(I)}{\rho}$$

B.3 Proofs in Section 1.7

B.3.1 Linear delay cost

B.3.1.1 Proof of Theorem 1.4

Proof. $\forall \langle \mu \rangle \in \mathbb{M}, \tau$,

$$\begin{aligned} E\left[\int_0^\tau C(I_t)\mathrm{d}t\right] \geqslant & C\left(\frac{E\left[\int_0^\tau I_t \mathrm{d}t\right]}{E\left[\int_0^\tau \mathrm{d}t\right]}\right) E\left[\int_0^\tau \mathrm{d}t\right] = C\left(\frac{E\left[\int_0^\tau -E\left[\frac{\mathrm{d}H(\mu_t)}{\mathrm{d}t} | \mathcal{F}_t\right]\mathrm{d}t\right]}{E[\tau]}\right) E[\tau] \\ = & C\left(\frac{E\left[-\int_0^\tau \frac{\mathrm{d}H(\mu_t)}{\mathrm{d}t}\mathrm{d}t\right]}{E[\tau]}\right) E[\tau] = C\left(\frac{E[H(\mu) - H(\mu_\tau)]}{E[\tau]}\right) E[\tau] \end{aligned}$$

First inequality is by Jensen's inequality. First equality is by definition of I_t. Second inequality is by iterated law of expectation. Last equality is straight forward. Since $\langle\mu_t\rangle \in \mathbb{M}$, $E[\mu_\tau] = \mu$. Let $\mu_\tau \sim P$ and $\lambda = E[H(\mu) - H(\mu_\tau)]/E[\tau]$, then:

$$\begin{aligned} &E\left[F(\mu_\tau) - m\tau - \int_0^\tau C(I_t)\mathrm{d}t\right] \leqslant E_P[F(\nu)] - \frac{E_P[H(\mu) - H(\nu)]}{\lambda}(m + C(\lambda)) \\ \Longrightarrow\ & V(\mu) \leqslant \sup_{P\in\Delta^2(X),\lambda>0} E_P[F(\nu)] - \frac{m + C(\lambda)}{\lambda}E_P[H(\mu) - H(\nu)] \end{aligned}$$

On the other hand, $\forall P \in \Delta^2(X), \lambda > 0$, let $\langle\mu_t\rangle$ be a compound Poisson process which realizes according to P with Poisson rate $\frac{\lambda}{E_p[H(\mu)-H(\nu)]}$, τ is jump time of $\langle\mu_t\rangle$. Then it is easy to verify that:

$$E\left[F(\mu_\tau) - m\tau - \int_0^\tau C(I_t)\mathrm{d}t\right] = E_P[F(\nu)] - \frac{m + C(\lambda)}{\lambda}E_p[H(\mu) - H(\nu)]$$

■

B.3.2 General information measure

B.3.2.1 Proof of Theorem 1.5

Proof. Consider Equation (1.13), it's sasy to see that both the inner maximization problem and the constraint are linear in p_i and σ^2. Therefore, Equation (1.13) can be written equivalently as choosing either one posterior or a diffusion experiment:

$$\rho V(\mu) = \max\left\{\rho F(\mu), \sup_\nu \frac{c(V(\nu) - V(\mu) - V'(\mu)(\nu - \mu))}{J(\mu,\nu)}, \frac{cV''(\mu)}{J''_{\nu\nu}(\mu,\mu)}\right\}$$

Now suppose $\mu \in D$ and $\rho V(\mu) = c\frac{V''(\mu)}{J''_{\nu\nu}(\mu,\mu)}$. This is saying, the maximization problem:

$$\sup_\nu \frac{c(V(\nu) - V(\mu) - V'(\mu)(\nu - \mu))}{J(\mu,\nu)}$$

will be solved for $\nu \to \mu$. Therefore, consider the FOC:

$$\text{FOC: } \frac{V'(\nu)-V'(\mu)}{J(\mu,\nu)} - \frac{J'_\nu(\mu,\nu)}{J(\mu,\nu)^2}\big(V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)\big)$$

It must be $\leqslant 0$ when $\nu \to \mu^+$ and $\geqslant 0$ when $\nu \to \mu^-$. Otherwise, the diffusion experiment will be locally dominated by some Poisson experiment. When $\nu \to \mu$, $J(\mu,\nu) \to 0$, $V'(\nu) \to V'(\mu)$, $V(\nu)-V(\mu)-V'(\mu)(\nu-\mu) \to 0$. Therefore, we can apply L'Hospital's rule:

$$\begin{aligned}
\lim_{\nu\to\mu} \text{FOC} =& \frac{\lim_{\nu\to\mu}\left(V''(\nu) - J''_{\nu\nu}(\mu,\nu)\frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)} - J'_\nu(\mu,\nu)\cdot FOC\right)}{\lim_{\nu\to\mu} J'_\nu(\mu,\nu)} \\
=& \frac{1}{2}\frac{\lim_{\nu\to\mu}\left(V''(\nu) - J''_{\nu\nu}(\mu,\nu)\frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}\right)}{\lim_{\nu\to\mu} J'_\nu(\mu,\nu)} \\
=& \frac{1}{2}\frac{\lim_{\nu\to\mu}\left(V^{(3)}(\nu) - J^{(3)}_{\nu\nu\nu}(\mu,\nu)\frac{V(\nu)-V(\mu)(\nu-\mu)}{J(\mu,\nu)} - J''_{\nu\nu}(\mu,\nu)\cdot \text{FOC}\right)}{\lim_{\nu\to\mu} J''_{\nu\nu}(\mu,\nu)} \\
=& \frac{1}{3}\frac{V^{(3)}(\mu) - J^{(3)}_{\nu\nu\nu}(\mu,\mu)\frac{V''(\mu)}{J''_{\nu\nu}(\mu,\mu)}}{J''_{\nu\nu}(\mu,\mu)}
\end{aligned} \tag{B.27}$$

Now consider $V(\mu) - \frac{c}{\rho}\frac{V''(\mu)}{J''_{\nu\nu}(\mu,\mu)}$. By assumption, it's non-negative and achieves 0 at μ. Therefore it is locally minimized at μ:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\mu}\left(V(\mu) - \frac{c}{\rho}\frac{V''(\mu)}{J''_{\nu\nu}(\mu,\mu)}\right) = 0 \\
\Longrightarrow\ & \frac{\rho}{c}V'(\mu) - \frac{V^{(3)}(\mu)}{J''_{\nu\nu}(\mu,\mu)} + \frac{V''(\mu)}{J''_{\nu\nu}(\mu,\mu)^2}\left(J^{(3)}_{\nu\nu\nu}(\mu,\mu) + J^{(3)}_{\nu\nu\mu}(\mu,\mu)\right) = 0 \\
\Longrightarrow\ & \frac{V^{(3)}(\mu) - J^{(3)}_{\nu\nu\nu}(\mu,\mu)\frac{V''(\mu)}{J''_{\nu\nu}(\mu,\mu)}}{J''_{\nu\nu}(\mu,\mu)} = \frac{\rho}{c}V'(\mu) + V''(\mu)\frac{J^{(3)}_{\nu\nu\mu}(\mu,\mu)}{J''_{\nu\nu}(\mu,\mu)^2} \\
\Longrightarrow\ & \frac{V^{(3)}(\mu) - J^{(3)}_{\nu\nu\nu}(\mu,\mu)\frac{V''(\mu)}{J''_{\nu\nu}(\mu,\mu)}}{J''_{\nu\nu}(\mu,\mu)} = \frac{\rho}{c}V'(\mu) + \frac{\rho}{c}V(\mu)\frac{J^{(3)}_{\nu\nu\mu}(\mu,\mu)}{J''_{\nu\nu}(\mu,\mu)}
\end{aligned} \tag{B.28}$$

By smoothness of V and J, for FOC to be non-positive when $\nu \to \mu^+$ and non-negative when $\nu \to \mu^-$. So Equations (B.27) and (B.28) implies:

$$V'(\mu)J''_{\nu\nu}(\mu,\mu) + V(\mu)J^{(3)}_{\nu\nu\mu}(\mu,\mu) = 0$$

Now suppose there exists $\mu_n \to \mu$ s.t. $\rho V(\mu_n) = c\frac{V''(\mu_n)}{J''_{\nu\nu}(\mu_n,\mu_n)}$, we have:

$$V'(\mu_n)J''_{\nu\nu}(\mu_n,\mu_n) + V(\mu_n)J^{(3)}_{\nu\nu\mu}(\mu_n,\mu_n) = 0$$

By differentiability of the whole term, we have:

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}\mu}\Big(V'(\mu)J''_{\nu\nu}(\mu,\mu) + V(\mu)J^{(3)}_{\nu\nu\mu}(\mu,\mu)\Big) = 0 \\
\Longrightarrow\ & V''(\mu)J''_{\nu\nu}(\mu,\mu) + V'(\mu)\Big(2J^{(3)}_{\nu\nu\mu}(\mu,\mu) + J^{(3)}_{\nu\nu\nu}(\mu,\mu)\Big) + V(\mu)\Big(J^{(4)}_{\nu\nu\nu\mu}(\mu,\mu) + J^{(4)}_{\nu\nu\mu\mu}(\mu,\mu)\Big) = 0 \\
\Longrightarrow\ & \frac{\rho}{c}V(\mu)J''_{\nu\nu}(\mu,\mu)^2 - \frac{V(\mu)}{J''_{\nu\nu}(\mu,\mu)}\Big(2J^{(3)}_{\nu\nu\mu}(\mu,\mu)^2 + J^{(3)}_{\nu\nu\nu}(\mu,\mu)J^{(3)}_{\nu\nu\mu}(\mu,\mu)\Big) \\
& + V(\mu)\Big(J^{(4)}_{\nu\nu\nu\mu}(\mu,\mu) + J^{(4)}_{\nu\nu\mu\mu}(\mu,\mu)\Big) = 0 \\
\Longrightarrow\ & \frac{\rho}{c}J''_{\nu\nu}(\mu,\mu)^2 - \frac{2J^{(3)}_{\nu\nu\mu}(\mu,\mu)^2 + J^{(3)}_{\nu\nu\nu}(\mu,\mu)J^{(3)}_{\nu\nu\mu}(\mu,\mu)}{J''_{\nu\nu}(\mu,\mu)} + J^{(4)}_{\nu\nu\nu\mu}(\mu,\mu) + J^{(4)}_{\nu\nu\mu\mu}(\mu,\mu) = 0
\end{aligned}$$

By assumption, $\mu \in D$, therefore the differential equation must not be satisfied. This implies that there doesn't exist such $\mu_n \to \mu$. So the set:

$$\left\{\mu \in D \middle| \rho V(\mu) = c\frac{V''(\mu)}{J''_{\nu\nu}(\mu,\mu)}\right\}$$

is a closed set (closed w.r.t. D) containing no limiting point. That is to say, within any compact subset of D, this set is finite. This set is a Borel set, thus Lebesgue measurable. By definition of Lebesgue measure, the measure of a set can be approximated by compact subsets from below. Therefore, this set has zero-measure. ■

B.3.2.2 Construction of a special cost function

Take any general cost structure $J(\mu,\nu)$ and $\kappa(\mu,\sigma)$ that satisfies Assumption 1.4. In the section, I introduce the method to construct a cost structure such that (i) the cost of Gaussian learning is $\kappa(\mu,\sigma)$ and (ii) the DM is exactly indifferent between using Guassian learning and Poisson learning.

Step 1. Let $g(\mu) = J''_{\nu\nu}(\mu,\mu)$ (then $\kappa(\mu,\sigma) = \frac{1}{2}g(\mu)\sigma^2$). Restrict the DM to using only Gaussian learning, then Equation (1.13) becomes:

$$\rho V(\mu) = \max\left\{\rho F(\mu), \frac{c}{g(\mu)}V''(\mu)\right\} \tag{B.29}$$

Equation (B.29) can be solved by solving the following ODE and applying smooth pasting:

$$V(\mu) = \frac{c}{\rho}\frac{V''(\mu)}{g(\mu)} \tag{B.30}$$

Change parameter and let $v(\mu) = \frac{\mathrm{d}}{\mathrm{d}\mu}\log(V(\mu))$, then $v(\mu)$ satisfies the following ODE:

$$v'(\mu) + v(\mu)^2 = \frac{\rho}{c}g(\mu) \tag{B.31}$$

By Assumption 1.4, $g(\mu)$ is a smooth function on $(0,1)$. Therefore it is easy to verify that on any closed sub-interval of $(0,1)$, Picard-Lindelöf is satisfied that there exist unique solution to Equation (B.31) given initial condition. Let $v(\mu, C_1)$ be the solution indexed by free parameter C_1, then $V(\mu) = C_2 e^{\int_0^\mu v(C_1,\nu)\mathrm{d}\nu}$. The two free parameters (C_1, C_2) can be pinned down by two smooth pasting conditions:

$$\begin{cases} C_2 e^{\int_0^{\mu_1} v(C_1,\nu)\mathrm{d}\nu} = F(\mu_1) \\ C_2 e^{\int_0^{\mu_1} v(C_1,\nu)\mathrm{d}\nu} v(C_1,\mu_1) = F'(\mu_1) \\ C_2 e^{\int_0^{\mu_2} v(C_1,\nu)\mathrm{d}\nu} = F(\mu_2) \\ C_2 e^{\int_0^{\mu_1} v(C_1,\nu)\mathrm{d}\nu} v(C_1,\mu_2) = F'(\mu_2) \end{cases}$$

Notice that smooth pasting need to be checked for at most $C^2_{|A|}$ pairs of actions, index all solutions by V_i. Then $V(\mu) = \max\{V_i(\mu)\}$ solves Equation (B.29). Let $E = \{\mu | V(\mu) > F(\mu)\}$.

Step 2. $\forall \mu \in E$, define $J_0(\mu,\nu)$ as:

$$J_0(\mu,\nu) = \frac{c}{\rho}\frac{V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{V(\mu)}$$

It is easy to verify that $J_0 > 0$. Now let us verify the solution of Equation (1.13):

$$\rho V(\mu) = \max\left\{\rho F(\mu), \sup_{p,\nu,\sigma^2} p\big(V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)\big) + \frac{1}{2}\sigma^2 V''(\mu)\right\}$$
$$\text{s.t. } pJ_0(\mu,\nu) + \frac{1}{2}J_{0\nu\nu}(\mu,\mu)\sigma^2 \leqslant c$$

First of all, by definition, $V(\mu)$, $p \equiv 0$ and $\sigma^2 = \frac{c}{J_{0\nu\nu}(\mu,\mu)}$ is feasible and satisfies the equality condition. Now $\forall \nu \in [0,1]$:

$$c\frac{V(\nu) - V(\mu) - V'(\mu)(\nu-\mu)}{J_0(\mu,\nu)} = \rho V(\mu)$$

Therefore, any Poisson learning strategy is as good as the Gaussian learning strategy. By definition of J_0, $\frac{1}{2}J_{0\nu\nu}(\mu,\mu)\sigma^2 = \frac{1}{2}g(\mu)\sigma^2 = \kappa(\mu,\sigma)$.

Step 3. Smooth extension of J_0. So far, J_0 is only defined on E. J_0 can be extended smoothly onto $[0,1]$ satisfying $\forall \mu \in E^C$:

$$\begin{cases} J_{0\nu\nu}(\mu,\mu) = g(\mu) \\ J_0(\mu,\nu) \geqslant \dfrac{c}{\rho}\dfrac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{V(\mu)}, \forall \nu \end{cases}$$

The extension is intuitively simple but quite technical, which is ommited here.

Such a $J_0(\mu,\nu)$ is uniquely defined for $\mu \in E$. For $\mu \in E^C$, there can be many degrees of freedom because Gaussian learning is anyway strictly dominated by stopping. So it is sufficienly to make Poisson learning also dominated. Now, suppose $J(\mu,\nu)$ and $\kappa(\mu,\sigma)$ are such that Gaussian learning is weakly optimal. Then $J(\mu,\nu)$ must be pointwise weakly higher than $J_0(\mu,\nu)$, $\forall \mu \in E$. On the other hand, since $J_{\nu\nu}(\mu,\mu) = J_{0\nu\nu}(\mu,\mu)$, this implies $J_{\nu^3}(\mu,\mu) = J_{0\nu^3}(\mu,\mu)$. That is to say, assuming Gaussian learning being weakly optimal is imposing an additional third derivative constraint on $J(\mu,\nu)$ on the constriants in Assumption 1.4, making the set of cost functions non-generic.

B.3.3 *Linear cost function*

B.3.3.1 *Proof of Theorem 1.6*

Proof. I first prove a result in discrete time. Take any information acquisition strategy $(\mathcal{S}^t, \mathcal{A}^t, \mathcal{T})$ that satisfies the constraints in Equation (1.5′). The achieved expected utility will be:

$$E\left[e^{-\rho dt\cdot\mathcal{T}}u(\mathcal{A}^{\mathcal{T}},\mathcal{X}) - \sum_{t=0}^{\infty} e^{-\rho dt\cdot t}\lambda I\left(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right]$$

We can separate the utility gain part and information cost part. Utility gain part is:

$$E\left[e^{-\rho dt\cdot\mathcal{T}}u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right)\right]$$

It's easy to see that this is determined only by action time $\mathcal{T}$ and action process $\mathcal{A}^{\mathcal{T}}$. Let $\tilde{\mathcal{S}}^{t-1} = (\mathbf{1}_{\mathcal{T}=t}, \mathcal{A}^t|_{\mathcal{T}=t})$. Then by information processing constraint in Equation (1.5′), we have:

$$\begin{aligned}
\text{Prob}\left(\tilde{\mathcal{S}}^{\mathbf{t}}|\mathcal{S}^{\mathbf{t}},\mathcal{X}\right) =& \text{Prob}(\mathbf{1}_{\mathcal{T}=\mathbf{t+1}},\mathcal{A}^{\mathbf{t+1}}|\mathcal{S}^{\mathbf{t}},\mathcal{X}) = \text{Prob}(\mathbf{1}_{\mathcal{T}\leqslant\mathbf{t+1}},\mathcal{A}^{\mathbf{t+1}}|\mathcal{S}^{\mathbf{t}},\mathcal{X}) \\
=& \text{Prob}(\mathbf{1}_{\mathcal{T}\leqslant\mathbf{t+1}},\mathcal{A}^{\mathbf{t+1}}|\mathcal{S}^{\mathbf{t}}) = \text{Prob}\left(\tilde{\mathcal{S}}^{\mathbf{t}}|\mathcal{S}^{\mathbf{t}}\right) \\
\Longrightarrow& \mathcal{X} \to \mathcal{S}^{\mathbf{t}} \to \tilde{\mathcal{S}}^{\mathbf{t}}
\end{aligned}$$

Therefore:

$$\begin{aligned}
&\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}\lambda E\left[I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t})\right] \\
=&\lambda\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}E\left[I(\mathcal{S}^{\mathbf{t}},\mathbf{1}_{\mathcal{T}\leqslant t};\mathcal{X}) - I(\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t};\mathcal{X})\right] \\
=&\lambda\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}E\left[I(\mathcal{S}^{\mathbf{t}};\mathcal{X}) + I(\mathbf{1}_{\mathcal{T}\leqslant t};\mathcal{X}|\mathcal{S}^{\mathbf{t}}) - I(\mathcal{S}^{\mathbf{t-1}};\mathcal{X}) - I(\mathbf{1}_{\mathcal{T}\leqslant t};\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})\right] \\
=&\lambda\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}E\left[I(\mathcal{S}^{\mathbf{t}};\mathcal{X}) - I(\mathcal{S}^{\mathbf{t-1}};\mathcal{X})\right] \\
=&\lambda\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}E\left[I\left(\tilde{\mathcal{S}}^{\mathbf{t}};\mathcal{X}\right) + I\left(\mathcal{S}^{\mathbf{t}};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t}}\right) - I\left(\tilde{\mathcal{S}}^{\mathbf{t}};\mathcal{X}|\mathcal{S}^{\mathbf{t}}\right)\right] \\
&-\lambda\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}E\left[I\left(\tilde{\mathcal{S}}^{\mathbf{t-1}};\mathcal{X}\right) + I\left(\mathcal{S}^{\mathbf{t-1}};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}}\right) - I\left(\tilde{\mathcal{S}}^{\mathbf{t-1}};\mathcal{X}|\mathcal{S}^{\mathbf{t-1}}\right)\right] \\
=&\lambda\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}E\left[I\left(\tilde{\mathcal{S}}^{\mathbf{t}};\mathcal{X}\right) + I\left(\mathcal{S}^{\mathbf{t}};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t}}\right)\right] \\
&-\lambda\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}E\left[I\left(\tilde{\mathcal{S}}^{\mathbf{t-1}};\mathcal{X}\right) + I\left(\mathcal{S}^{\mathbf{t-1}};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}}\right)\right] \\
=&\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}\lambda E\left[I\left(\tilde{\mathcal{S}}^t;\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right] + \lambda\sum e^{-\rho dt\cdot t}\left(1-e^{-\rho dt}\right)E\left[I\left(\mathcal{S}^{\mathbf{t}};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t}}\right)\right] \\
\geqslant&\sum_{t=0}^{\infty} e^{-\rho dt\cdot t}\lambda E\left[I\left(\tilde{\mathcal{S}}^t;\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right]
\end{aligned}$$

Therefore, by replacing signal process $\mathcal{S}^t$ with $\tilde{\mathcal{S}}^t$, the DM can achieve the same utility gain and pay a weakly lower information cost. Now consider:

$$
\begin{aligned}
&E\left[e^{-\rho dt\cdot\mathcal{T}}u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right)\right]-\lambda\sum_{t=0}^{\infty}e^{-\rho dt\cdot t}E\left[I\left(\tilde{\mathcal{S}}^t;\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right]\\
=&\mathrm{Prob}(\mathcal{T}=0)E\left[u\left(\mathcal{A}^0,\mathcal{X}\right)\right]\\
&+\mathrm{Prob}(\mathcal{T}\geqslant 1)E\left[e^{-\rho dt\cdot\mathcal{T}}u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right)|\mathcal{T}\geqslant 1\right]\\
&-\lambda I\left(\tilde{\mathcal{S}}^0;\mathcal{X}|\mu\right)-\lambda\sum_{t=1}^{\infty}e^{-\rho dt\cdot t}E\left[I\left(\tilde{\mathcal{S}}^t;\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right]\\
=&\mathrm{Prob}(\mathcal{T}=0)E\left[u\left(\mathcal{A}^0,\mathcal{X}\right)\right]\\
&+\mathrm{Prob}(\mathcal{T}=1)E\left[e^{-\rho dt}u\left(\mathcal{A}^1,\mathcal{X}\right)|\mathcal{T}=1\right]-\lambda I\left(\tilde{\mathcal{S}}^0;\mathcal{X}|\mu,\mathcal{T}>0\right)\\
&+\mathrm{Prob}(\mathcal{T}\geqslant 2)E\left[e^{-\rho dt\cdot\mathcal{T}}u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right)|\mathcal{T}\geqslant 2\right]-\lambda\sum_{t=1}^{\infty}e^{-\rho dt\cdot t}E\left[I\left(\tilde{\mathcal{S}}^t;\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right]
\end{aligned}
$$

Suppose the term:

$$
\mathrm{Prob}(\mathcal{T}\geqslant 2)E\left[e^{-\rho dt\cdot\mathcal{T}}u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right)|\mathcal{T}\geqslant 2\right]-\lambda\sum_{t=1}^{\infty}e^{-\rho dt\cdot t}E\left[I\left(\tilde{\mathcal{S}}^t;\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right] \tag{B.32}
$$

is negative, then discard all actions and information after first period will give the DM higher expected utility. This information and action process satisfies this theorem. Therefore, WLOG we assume Equation (B.32), as well as all continuation payoffs are non-negative. Then:

$$
\begin{aligned}
&E\left[e^{-\rho dt\cdot\mathcal{T}}u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right)\right]-\lambda\sum_{t=0}^{\infty}e^{-\rho dt\cdot t}E\left[I\left(\tilde{\mathcal{S}}^t;\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right]\\
=&\mathrm{Prob}(\mathcal{T}=0)E\left[u\left(\mathcal{A}^0,\mathcal{X}\right)\right]\\
&+\mathrm{Prob}(\mathcal{T}=1)\left(E\left[e^{-\rho dt}u\left(\mathcal{A}^1,\mathcal{X}\right)|\mathcal{T}=1\right]-\lambda I\left(\tilde{\mathcal{S}}^0;\mathcal{X}|\mu,\mathcal{T}>0\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&\quad + \text{Prob}(\mathcal{T} \geqslant 2)E\Big[e^{-\rho dt\cdot\mathcal{T}}u\Big(\mathcal{A}^{\mathcal{T}},\mathcal{X}\Big)|\mathcal{T} \geqslant 2\Big] - \lambda\sum_{t=1}^{\infty}e^{-\rho dt\cdot t}E\Big[I\Big(\tilde{\mathcal{S}}^{t};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\Big)\Big]\\
\leqslant&\text{Prob}(\mathcal{T} = 0)E\Big[u\Big(\mathcal{A}^{0},\mathcal{X}\Big)\Big]\\
&\quad + \text{Prob}(\mathcal{T} = 1)\Big(E\Big[e^{-\rho dt}u\Big(\mathcal{A}^{1},\mathcal{X}\Big)|\mathcal{T} = 1\Big] - \lambda I\Big(\tilde{\mathcal{S}}^{0};\mathcal{X}|\mu,\mathcal{T} > 0\Big)\Big)\\
&\quad + \text{Prob}(\mathcal{T} \geqslant 2)E\Big[e^{-\rho dt\cdot(\mathcal{T}-1)}u\Big(\mathcal{A}^{\mathcal{T}},\mathcal{X}\Big)|\mathcal{T} \geqslant 2\Big] - \lambda\sum_{t=1}^{\infty}e^{-\rho dt\cdot(t-1)}E\Big[I\Big(\tilde{\mathcal{S}}^{t};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\Big)\Big]\\
=&\text{Prob}(\mathcal{T} = 0)E\Big[u\Big(\mathcal{A}^{0},\mathcal{X}\Big)\Big]\\
&\quad + \text{Prob}(\mathcal{T} = 1)\Big(E\Big[e^{-\rho dt}u\Big(\mathcal{A}^{1},\mathcal{X}\Big)|\mathcal{T} = 1\Big] - \lambda I\Big(\tilde{\mathcal{S}}^{0};\mathcal{X}|\mu,\mathcal{T} > 0\Big)\Big)\\
&\quad + \text{Prob}(\mathcal{T} = 2)\Big(E\Big[e^{-\rho dt}u\Big(\mathcal{A}^{2},\mathcal{X}\Big)|\mathcal{T} = 2\Big] - \lambda I\Big(\tilde{\mathcal{S}}^{1};\mathcal{X}|\tilde{\mathcal{S}}^{0},\mathcal{T} > 1\Big)\Big)\\
&\quad + \text{Prob}(\mathcal{T} \geqslant 3)E\Big[e^{-\rho dt\cdot(\mathcal{T}-1)}u\Big(\mathcal{A}^{\mathcal{T}},\mathcal{X}\Big)|\mathcal{T} \geqslant 3\Big] - \lambda\sum_{t=2}^{\infty}e^{-\rho dt\cdot(t-1)}E\Big[I\Big(\tilde{\mathcal{S}}^{t};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\Big)\Big]\\
\leqslant&\text{Prob}(\mathcal{T} = 0)E\Big[u\Big(\mathcal{A}^{0},\mathcal{X}\Big)\Big]\\
&\quad + \text{Prob}(\mathcal{T} = 1)\Big(E\Big[e^{-\rho dt}u\Big(\mathcal{A}^{1},\mathcal{X}\Big)|\mathcal{T} = 1\Big] - \lambda I\Big(\tilde{\mathcal{S}}^{0};\mathcal{X}|\mu,\mathcal{T} > 0\Big)\Big)\\
&\quad + \text{Prob}(\mathcal{T} = 2)\Big(E\Big[e^{-\rho dt}u\Big(\mathcal{A}^{2},\mathcal{X}\Big)|\mathcal{T} = 2\Big] - \lambda I\Big(\tilde{\mathcal{S}}^{1};\mathcal{X}|\tilde{\mathcal{S}}^{0},\mathcal{T} > 1\Big)\Big)\\
&\quad + \text{Prob}(\mathcal{T} \geqslant 3)E\Big[e^{-\rho dt\cdot(\mathcal{T}-2)}u\Big(\mathcal{A}^{\mathcal{T}},\mathcal{X}\Big)|\mathcal{T} \geqslant 3\Big] - \lambda\sum_{t=2}^{\infty}e^{-\rho dt\cdot(t-2)}E\Big[I\Big(\tilde{\mathcal{S}}^{t};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\Big)\Big]\\
&\ \vdots\\
\leqslant&\text{Prob}(\mathcal{T} = 0)E\Big[u\Big(\mathcal{A}^{0},\mathcal{X}\Big)\Big]\\
&\quad + \text{Prob}(\mathcal{T} \geqslant 1)\left(E\Big[e^{-\rho dt}u\Big(\mathcal{A}^{\mathcal{T}};\mathcal{X}\Big)|\mathcal{T} \geqslant 1\Big] - \lambda\sum_{t=0}^{\infty}E\Big[I\Big(\tilde{\mathcal{S}}^{t};\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t},\mathcal{T} \geqslant 1\Big)\Big]\right)\\
=&\text{Prob}(\mathcal{T} = 0)E\Big[u\Big(\mathcal{A}^{0},\mathcal{X}\Big)\Big]\\
&\quad + \text{Prob}(\mathcal{T} \geqslant 1)\Big(E\Big[e^{-\rho dt}u\Big(\mathcal{A}^{\mathcal{T}};\mathcal{X}\Big)|\mathcal{T} \geqslant 1\Big] - \lambda\lim_{t\to\infty}I\Big(\tilde{\mathcal{S}}^{\mathbf{t}};\mathcal{X}|\mathcal{T} \geqslant 1\Big)\Big)\\
=&\text{Prob}(\mathcal{T} = 0)E\Big[u\Big(\mathcal{A}^{0},\mathcal{X}\Big)\Big]\\
&\quad + \text{Prob}(\mathcal{T} \geqslant 1)\Big(E\Big[e^{-\rho dt}u\Big(\mathcal{A}^{\mathcal{T}};\mathcal{X}\Big)|\mathcal{T} \geqslant 1\Big] - \lambda I\Big(\mathcal{A}^{\mathcal{T}};\mathcal{X}|\mathcal{T} \geqslant 1\Big)\Big)
\end{aligned}
$$

Therefore:

$$
\begin{aligned}
&E\left[e^{-\rho dt\cdot\mathcal{T}}u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right)\right]-\lambda\sum_{t=0}^{\infty}e^{-\rho dt\cdot t}E\left[I\left(\widetilde{\mathcal{S}}^{t};\mathcal{X}\middle|\widetilde{\mathcal{S}}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t}\right)\right]\\
\leqslant&\mathrm{P}[\mathcal{T}=0]F(\mu)+(1-\mathrm{P}[\mathcal{T}=0])\left(E\left[e^{-\rho dt}u\left(\mathcal{A}^{\mathcal{T}},\mathcal{X}\right)\right]-\lambda I\left(\mathcal{A}^{\mathcal{T}};\mathcal{X}\middle|\mu\right)\right)\\
\leqslant&\max\left\{F(\mu),\sup_{\mathcal{A}}E\left[e^{-\rho dt}u(\mathcal{A},\mathcal{X})\right]-\lambda I(\mathcal{A};\mathcal{X})\right\}
\end{aligned}
$$

Therefore, we showed that any dynamic information acquisition strategy solving Equation (1.5′) will have weakly lower expected utility level than a static information acquisition strategy solving Equation (1.14).

Now let us consider the continuous time problem. It is clear by Lemma B.5 that any discretization of Equation (1.1) can be implemented by Equation (1.5′). Hence,

$$
\begin{aligned}
V(\mu)\leqslant\lim V_{dt}(\mu)\leqslant\overline{\lim}\ &\max\left\{F(\mu),\sup_{\mathcal{A}}E\left[e^{-\rho dt}u(\mathcal{A},\mathcal{X})\right]-\lambda I(\mathcal{A};\mathcal{X})\right\}\\
=&\sup_{\mathcal{A}}E[u(\mathcal{A},\mathcal{X})]-\lambda I(\mathcal{A};\mathcal{X})\\
=&\sup_{P\in\Delta^2(X)}E_P[F(\nu)-\lambda(H(\mu)-H(\nu))]
\end{aligned}
$$

On the other hand, take any P and $dt>0$, by Lemma B.3, there exists $\langle\mu_t\rangle\in\mathbb{M}$ such that:

$$
\begin{aligned}
&E\left[e^{-\rho dt}F(\mu_{dt})-\int_0^{dt}e^{-\rho t}\lambda\frac{E_P[H(\mu)-H(\nu)]}{dt}\mathrm{d}t\right]\\
=&E_P[e^{-\rho dt}F(\nu)]-\frac{1-e^{-\rho dt}}{\rho dt}\lambda E_P[H(\mu)-H(\nu)]\\
\Longrightarrow V(\mu)\geqslant&\sup_{P\in\Delta^2(X)}\lim_{dt\to0}E_P[e^{-\rho dt}F(\nu)]-\frac{1-e^{-\rho dt}}{\rho dt}\lambda E_P[H(\mu)-H(\nu)]\\
\Longrightarrow V(\mu)\geqslant&\sup_{P\in\Delta^2(X)}E_P[F(\nu)-\lambda(H(\mu)-H(\nu))]
\end{aligned}
$$

Combining the two inequalities, Equation (1.14) is proved. ■

B.4 Proofs in Section 1.8

B.4.1 Choice accuracy and response time: proof of Proposition 1.1

Proof. Since both $H_0(\mu)$ and $F(\mu)$ are symmetric functions around $\mu_0 = 0.5$, by symmetry and quasi-concavity of value function (Theorem 1.2), $\forall c_k$, $\mu^* = \mu_0$. Let the expected utility of the action favoring beliefs > 0.5 be $F_r(\mu)$, and the utility of the other action $F_l(\mu)$. $\forall c_k$, by the proof of Lemma A.2, there exists unique ν^l_k and $\nu^r_k = 1 - \mu^r_k$ s.t.

$$\begin{cases} \nu^r_k \in \arg\max\limits_{\nu \geqslant \mu} \dfrac{F_r(\nu)}{1 + \frac{\rho}{c_k} J_0(\mu_0, \nu)} \\ \nu^l_k \in \arg\max\limits_{\nu \leqslant \mu} \dfrac{F_l(\nu)}{1 + \frac{\rho}{c_k} J_0(\mu_0, \nu)} \end{cases}$$

Where $J_0(\mu, \nu) = H_0(\mu) - H_0(\nu) + H_0'(\mu)(\nu - \mu)$. Now I determine the location of $\{\nu^l_k, \nu^r_k\}$ by studying the following cross derivative:

$$\begin{aligned} &\frac{\mathrm{d}^2}{\mathrm{d}c, \mathrm{d}\nu} \frac{F_r(\nu)}{1 + \frac{\rho}{c} J_0(\mu, \nu)} \bigg|_{\nu = \nu^{r*}, \mu = \mu_0} \\ =& \frac{\rho^2}{c^3} \frac{F_r(\nu)(H_0'(\mu) - H_0'(\nu))(F_r(\mu)(H_0'(\mu) - H_0'(\nu)) + F_r' J_0(\nu, \mu))}{F_r'\left(1 + \frac{\rho}{c} J_0(\mu, \nu)\right)} \Bigg|_{\nu = \nu^{r*}, \mu = \mu_0} > 0 \end{aligned}$$

The equality is by plug the FOC determining ν^* into the cross derivative. The inequality follows by $H_0(\mu)$ being strictly concave, $F_r' > 0$ and $F_r(\mu_0) = F(\mu_0) > 0$. Since the cross derivative w.r.t. ν and c is strictly positive at ν^{r*}, the standard comparative statics analysis suggests that the optimal belief ν^{r*} is strictly increasing in parameter c. A completely symmetric argument applies to ν^{l*} and ν^{l*} is strictly decreasing in parameter c. Therefore, all the $\{\nu_k^r, \nu_k^l\}$ are ordered on $[0,1]$:

$$0, \nu_k^l, \cdots, \nu_1^l, \mu_0, \nu_1^r, \cdots, \nu_k^r, 1$$

Moreover, $\forall c \in (c_i, c_{i+1})$, $\nu^{r*}(c) \in (\nu_i^r, \nu_{i+1}^r)$ and $\nu^{l*}(c) \in (\nu_{i+1}^l, \nu_i^l)$. Now assume that the goal is to make the sign of $\mu^* - \mu_0$ strictly positive (negative) when $c \in (c_{2i}, c_{2i+1})$ ($c \in (c_{2i-1}, c_{2i})$). To achieve this goal, define $H(\mu)$ based on $H_0(\mu)$. Let $M_{a,b}(\mu)$ be a function on $\mathbb{R}$ with the following properties:

- Parameter $a, b \in \mathbb{R}$ and $a < b$.
- $\forall a, b, \mu$, $M_{a,b}(\mu) < 0$ if $\mu \in (a,b)$ and $M_{a,b} = 0$ otherwise.
- $\forall a, b$, $M_{a,b}(\mu)$ is $C^{(2)}$ smooth on $\mathbb{R}$ and $|M''_{a,b}(\mu)|$ is bounded by 1.

The choice of function M can be quite arbitrary. For example, it is not hard to verify that:

$$M_{a,b}(\mu) = -\mathbf{1}_{a<\mu<b} \frac{(b-a)^4}{256 e^{-\frac{16}{(b-a)^2}}} e^{-\left(\frac{1}{\mu-a} + \frac{1}{b-\mu}\right)^2}$$

satisfies these properties. Define $\nu_{k+1}^r = \frac{1+\nu_k^r}{2}$ and $\nu_{k+1}^l = \frac{\nu_k^l}{2}$. Since $H_0(\mu)$ satisfies Assumption 1.2-a, there exists ε s.t. $\forall \mu \in \left[\nu_{k+1}^l, \nu_{k+1}^r\right]$, $H''(\mu) \leqslant -2\varepsilon$. Now define $H(\mu)$:

$$H(\mu) = H_0(\mu) + \sum M_{\nu_{2i}^r, \nu_{2i+1}^r}(\mu) + \sum M_{\nu_{2i}^l, \nu_{2i-1}^l}(\mu)$$

I verify that $H(\mu)$ satisfies the conditions in Proposition 1.1. It is easy to verify that $J(\mu_0, \nu) = J(\mu_0, \nu)$ when $\nu \notin \bigcup(\nu_{2i}^r, \nu_{2i+1}^r)$ or $\bigcup(\nu_{2i}^l), \nu_{2i-1}^l$. $J(\mu_0, \nu) > J_0(\mu_0, \nu)$ otherwise. First, when $c \in \{c_k\}$:

$$\sup_{\nu \geqslant \mu} \frac{F_r(\nu)}{1 + \frac{\varrho}{c} J(\mu_0, \nu)} \geqslant \sup_{\nu \geqslant \mu} \frac{F_r(\nu)}{1 + \frac{\varrho}{c} J_0(\mu_0, \nu)} \geqslant \frac{F_r(\nu_i^r)}{1 + \frac{\varrho}{c_i} J(\mu_0, \nu_i^r)} = \frac{F_r(\nu_i^r)}{1 + \frac{\varrho}{c_i} J_0(\mu_0, \nu_i^r)} \geqslant \sup_{\nu \geqslant \mu} \frac{F_r(\nu)}{1 + \frac{\varrho}{c} J(\mu_0, \nu)}$$

$$\sup_{\nu \leqslant \mu} \frac{F_l(\nu)}{1+\frac{\rho}{c}J(\mu_0,\nu)} \geqslant \sup_{\nu \leqslant \mu} \frac{F_l(\nu)}{1+\frac{\rho}{c}J_0(\mu_0,\nu)} \geqslant \frac{F_l(\nu_i^l)}{1+\frac{\rho}{c_i}J(\mu_0,\nu_i^l)} = \frac{F_l(\nu_i^l)}{1+\frac{\rho}{c_i}J_0(\mu_0,\nu_i^l)} \geqslant \sup_{\nu \leqslant \mu} \frac{F_l(\nu)}{1+\frac{\rho}{c}J(\mu_0,\nu)}$$
$$\implies \frac{F_r(\nu_i^r)}{1+\frac{\rho}{c_i}J(\mu_0,\nu_i^r)} = \frac{F_l(\nu_i^l)}{1+\frac{\rho}{c_i}J(\mu_0,\nu_i^l)} = \sup_{\nu \geqslant \mu} \frac{F_r(\nu)}{1+\frac{\rho}{c}J_0(\mu_0,\nu)} = \sup_{\nu \leqslant \mu} \frac{F_l(\nu)}{1+\frac{\rho}{c}J_0(\mu_0,\nu)}$$

Therefore, $\mu^* = \mu_0$ and the optimal strategy at μ_0 is the same as that with $H_0(\mu)$.

Now I prove that the sign of $\mu^* - \mu_0$ strictly positive (negative) when $c \in (c_{2i}, c_{2i+1})$ ($c \in (c_{2i-1}, c_{2i})$). For the first case $c \in (c_{2i}, c_{2i+1})$, since $J(\mu_0, \nu^{r*}(c)) > J_0(\mu, \nu^{r*}(c))$ and $J(\mu_0, \nu^{l*}(c)) = J_0(\mu_0, \nu^{l*}(c))$,

$$\overline{V}^+(\mu_0) = \max_{\nu \geqslant \mu} \frac{F_r(\nu)}{1+\frac{\rho}{c}J(\mu_0,\nu)} < \max_{\nu \geqslant \mu} \frac{F_r(\nu)}{1+\frac{\rho}{c}J_0(\mu_0,\nu)}$$
$$\overline{V}^-(\mu_0) = \max_{\nu \leqslant \mu} \frac{F_l(\nu)}{1+\frac{\rho}{c}J(\mu_0,\nu)} = \max_{\nu \leqslant \mu} \frac{F_r(\nu)}{1+\frac{\rho}{c}J_0(\mu_0,\nu)}$$

The first strict inequality is from uniqueness of optimal $\nu^{r*}(c)$, $J(\mu_0, \nu^{r*}(c)) > J_0(\mu, \nu^{r*}(c))$ and continuity of the objective function in ν. Therefore, $\overline{V}^+(\mu_0) < \overline{V}^-(\mu_0)$. Since $\overline{V}^+(\mu)$ is increasing in μ and $\overline{V}^-(\mu)$ is decreasing in μ, their crossing point $\mu^* > \mu_0$. For the other case $c \in (c_{2i-1}, c_{2i})$, $J(\mu_0, \nu^{r*}(c)) = J_0(\mu, \nu^{r*}(c))$ and $J(\mu_0, \nu^{l*}(c)) < J_0(\mu_0, \nu^{l*}(c))$. Therefore,

$$\overline{V}^+(\mu_0) = \max_{\nu \geqslant \mu} \frac{F_r(\nu)}{1+\frac{\rho}{c}J(\mu_0,\nu)} = \max_{\nu \geqslant \mu} \frac{F_r(\nu)}{1+\frac{\rho}{c}J_0(\mu_0,\nu)}$$
$$\overline{V}^-(\mu_0) = \max_{\nu \leqslant \mu} \frac{F_l(\nu)}{1+\frac{\rho}{c}J(\mu_0,\nu)} < \max_{\nu \leqslant \mu} \frac{F_r(\nu)}{1+\frac{\rho}{c}J_0(\mu_0,\nu)}$$

Therefore, $\overline{V}^+(\mu_0) > \overline{V}^-(\mu_0)$. Since $\overline{V}^+(\mu)$ is increasing in μ and $\overline{V}^-(\mu)$ is decreasing in μ, their crossing point $\mu^* < \mu_0$. ■

B.4.2 *Radical innovation: proof of Propositions 1.2 and 1.3*

Proof. Consider the solution to the problem of firm L, where the payoff to riskless arm is $u_L(P_s)$. By Theorem 1.2, the policy function $\nu_L(\nu)$ is a strictly decreasing function on experimentation region E_L. $\nu_L(\mu)$ is piecewise smooth. Each discrete point of $\nu_L(\mu)$ corresponds to a critical point where the DM is indifferent between confirming two beliefs associated with two different actions. Now I call the set of those critical beliefs $\{\mu_j\}_{j=1}^J$, where μ_J is the smallest and μ_1 is the largest. I first prove the following useful lemma.

Lemma B.24. *At each μ_j, let $\underline{\nu}_L^j < \bar{\nu}_L^j$ be the smallest and largest optimal posterior beliefs for firm L. Then either $\nu_S(\mu_j) < \underline{\nu}_L^j$ or $\nu_S(\mu_j) > \bar{\nu}_L^j$.*

Proof. Define $L(V,\lambda,\mu)(\nu)$ and $G(V,\lambda,\mu)(*)\nu$ and $G(V,\lambda)(\mu)$ as Equation (A.22). Consider:

$$L\left(V_L, \frac{\rho}{c}V_L(\mu_j), \mu_j\right)(\nu) - G\left(F, \frac{\rho}{c}V_L(\mu_j)\right)(\nu) \tag{B.33}$$

Optimality condition Equations (A.26) and (A.27) implies that it attains minimum 0 at both $\underline{\nu}_L^j$ and $\bar{\nu}_L^j$ (and at no other beliefs outside of the range $(\underline{\nu}_L^j), \bar{\nu}_L^j)$. Now consider:

$$L\left(V_S, \frac{\rho}{c}V_L(\mu_j), \mu_j\right)(\nu) - G\left(F, \frac{\rho}{c}V_L(\mu_j)\right)(\nu)$$

Since in L's experimentation region $V_L > V_S$, the term is strictly positive for $\nu > \mu_j$. Now the optimality condition implies that

$$L\left(V_S, \frac{\rho}{c}V_S(\mu_j), \mu_j\right)(\nu) - G\left(F, \frac{\rho}{c}V_S(\mu_j)\right)(\nu) \tag{B.34}$$

attains minimum 0 at $\nu_S(\mu_j)$. Notice that Equation (B.34) is equivalent to:

$$\begin{aligned}
&L\left(V_S, \frac{\rho}{c}V_L(\mu_j), \mu_j\right)(\nu) - G\left(F, \frac{\rho}{c}V_L(\mu_j)\right)(\nu) \\
=&L\left(V_L, \frac{\rho}{c}V_L(\mu_j), \mu_j\right)(\nu) - G\left(F, \frac{\rho}{c}V_L(\mu_j)\right)(\nu) \\
+&\left(V_S(\mu_j) - V_L(\mu_j) + (V_S'(\mu_j) - V_L'(\mu_j))(\nu - \mu_j)\right)
\end{aligned}$$

$$+\frac{\rho}{c}(V_S(\mu_j) - V_L(\mu_j))\big(H(\mu) - H(\nu) + H'(\mu)(\nu - \mu)\big)$$

Notice that the second term is linear in ν and the third term is concave in μ. Since Equation (B.33) is minimized at $\underline{\nu}_L^j$ and $\bar{\nu}_L^j$, the two points share the same supporting hyperplane. Now Equation (B.34) equals Equation (B.33) plus a strictly concave term. As a result, Equation (B.34) has positive first derivative at $\underline{\nu}_L^j$ and negative first derivative at $\overline{V}_L^j$. Therefore, the posterior belief that minimizes Equation (B.34) is either strictly less than $\underline{\nu}_L^j$ or strictly larger than $\overline{\nu_L^j}$. ∎

Step 1. I prove that in the region $\mu \geqslant \mu_1$, there exists critical belief μ_c that the satisfies the property of Proposition 1.3. By Lemma B.24, there are two possible cases.

The first case is that $\nu_S(\mu_1) < \underline{\nu}_L^1$. Now if at belief μ_1, firm S's optimal posterior is already associated with a less risky action, then since by definition of μ_1 firm L doesn't use any action less risky than that associated with $\underline{\mu}_L^1$ at all. So $\mu_c = \mu_1$. If otherwise firm S's and firm L's optimal beliefs are associated with the same action, then by the previous analysis, $\nu_S(\mu_j) < \underline{\nu}_L^j$. Now I prove that $\nu_S(\mu) < \nu_L(\mu)$ for all $\mu \geqslant \mu_j$. This can be easily seen from the phase diagram Figure B.1. Since the two firms are using the same action, their optimal belief is characterized by the same set of ODEs (except for different in initial value). Since we know that $V_L > V_S$, then the policy function ν_L must touch the diagonal line later than ν_S. By Picard-Lindelof solution to ODE doesn't cross, $\nu_L(\mu) > \nu_S(\mu)$ for $\mu \geqslant \mu_0$ and $\nu_L(\mu) < \nu_S(\mu)$ for $\mu \leqslant \mu_0$ (μ_0 is the critical belief the action giving zero expected payoff). Giving this single crossing property, since $\nu_L(\mu_1^+) > \nu_S(\mu_1^+)$, $\nu_L(\mu) > \nu_S(\mu)$ for all $\mu \geqslant \mu_1$.

The second case is that $\nu_S(\mu_1) > \bar{\nu}_L^1$. Now for some $\mu > \mu_1$, $\nu_S(\mu)$ stays above $\nu_L(\mu)$ whenever it corresponds to a more risky action. However, the analysis in the first case shows that when firm S switches action, it either jumps to a strictly less risky action, or stays at the same action as firm L for some beliefs (but $\nu_L(\mu)$ and $\nu_S(\mu)$ crosses once). In either cases, the single crossing property holds. So there exists such critical belief μ_c. Notice that the analysis in this region is already sufficient to prove Proposition 1.2.

Step 2. I prove Proposition 1.3 by induction. I prove the following statement that *if* for μ_j, there are two possibles cases: $\nu_S(\mu_j) < \underline{\nu}_L^j$ and $\nu_S(\mu) < \nu_L(\mu)\ \forall \mu \geqslant \mu_j$; or $\nu_S(\mu_j) > \bar{\nu}_L^j$ and there exists $\mu_c > \mu_j$, *then* the same statement is true for μ_{j+1}.
If $\nu_S(\mu_{j+1}) < \underline{\nu}_L^{j+1}$, then the argument is simple. Case 1 is that firm S has already switched to a less risky action, then before the firm H switches, $\nu_L(\mu) > \nu_S(\mu)$ for sure, up to μ_j. Then $\nu_L(\mu) > \nu_S(\mu)$ for $\mu \geqslant \mu_j$ as well by assumption in induction. Case 2 is that firm S is using the same action as firm L. The by the argument in step 1, before either firm switches to a less risky action, $\nu_L(\mu) > \nu_S(\mu)$. Suppose by contradiction that firm L first switches to a less risky action at μ_j, then by Lemma B.24, $\nu_S(\mu_j^-) > \nu_L(\mu_j^-)$, contradiction. Therefore, to sum up $\nu_S(\mu) < \mu_L(\mu)\ \forall \mu > \underline{\nu}_L^{j+1}$.
If $\nu_S(\mu_{j+1}) > \bar{\nu}_L^{j+1}$, the we only need to discuss that firm S ever uses the same action as firm (because otherwise either single crossing happens and we are in the case $\nu_S(\mu_j) < \underline{\nu}_L^j$, then the induction assumptions shows $\nu_S(\mu) < \nu_L(\mu)$ for all $\mu \geqslant \mu_j$; or crossing doesn't happen, then the induction assumptions shows that single crossing happens for $\mu_c > \mu_j$). In this case, the analysis in step 1 shows that $\nu_L(\mu)$ and $\nu_S(\mu)$ crosses at most once, and afterwards, the induction assumptions shows $\nu_S(\mu) < \nu_L(\mu)$ for all $\mu \geqslant \mu_j$. To sum up, I prove that ν_L and ν_S crosses at most once in firm L's experimentation region. Notice that $V_L > V_S$, therefore, there exists μ in L's experimentation region where $V_S(\mu) = F(\mu)$ already. Obviously for such belief $\nu_L(\mu) > \nu_S(\mu)$. On the other hand, on the left end of L's experimentation region:

$$\frac{F(\nu) - V_L(\mu)}{H(\mu) - H(\nu) + H'(\mu)(\nu - \mu)} < \frac{F(\nu) - V_S(\mu)}{H(\mu) - H(\nu) + H'(\mu)(\nu - \mu)} \leqslant V_S(\mu)$$

So since $V_L(\mu) > V_S(\mu)$ it must be that $V_S'(\mu) > 0$. Therefore, there exist μ in S's experimentation region where $V_L(\mu) = F(\mu)$. This proves that $E_0 \bigcap (0, \mu_c) \neq \varnothing$ and $E_0 \bigcap (\mu_c, 1) \neq \varnothing$. ■

B.5 Proofs in Appendix A.1

B.5.1 Convergence of policy

B.5.1.1 Proof of Theorem A.1

Proof. The original statement in Theorem A.1 is equivalent to: $\forall \varepsilon > 0$, there exists δ s.t. $\forall dt \leqslant \delta$, $\forall \mu \in [0,1]$, there exists $|\mu' - \mu| \leqslant \varepsilon$ and any optimal posterior induced in discrete time problem with period length dt will be within either $[\mu' - \varepsilon, \mu' + \varepsilon]$ or $[\nu(\mu') - \varepsilon, \nu(\mu') + \varepsilon]$.Now pick any $\varepsilon > 0$, let's discuss two cases separately:

Case 1: $\mu \in [0,1] \backslash E$. I first prove the case with Assumption A and Assumptions 1.2-a and 1.3. I will show that for any dt, any informative experiment is suboptimal. Suppose not, and there exists $\nu_i \neq \mu$ s.t.:

$$e^{-\rho dt} \sum p_i V_{dt}(\nu_i) \geqslant V_{dt}(\mu)$$

and

$$\begin{cases} \sum p_i \nu_i = \mu \\ H(\mu) - \sum p_i H(\nu_i) \leqslant cdt \end{cases}$$

Now consider a problem with $\frac{dt}{2}$. Consider the following strategy: mix experiment p_i, ν_i and prior with probability $\frac{1}{2}$. Then obviously Bayes plausibility and capacity constraint are satisfied. The expected utility from this strategy is:

$$V_{\frac{dt}{2}}(\mu) \geqslant \sum_{t=1} e^{-\rho \frac{dt}{2} \cdot t} \sum p_i V_{\frac{dt}{2}}(\nu_i) \cdot \frac{1}{2^t} = \frac{1}{2 - e^{-\frac{\rho dt}{2}}} \sum p_i V_{\frac{dt}{2}}(\nu_i)$$

$$
\begin{aligned}
&> e^{-\rho dt} \sum p_i V_{\frac{dt}{2}}(\nu_i) \\
&\geqslant e^{-\rho dt} \sum p_i V_{dt}(\nu_i) \\
&\geqslant F(\mu)
\end{aligned}
$$

First inequality is from optimality of $V_{\frac{dt}{2}}$. Second inequality is from $\frac{1}{2-x} > x^2$ for $x \in (0,1)$. Third inequality is from $V_{\frac{dt}{2}} \geqslant V_{dt}$. Last inequality is from assumption. Therefore $F(\mu) = V(\mu) \geqslant V_{\frac{dt}{2}} > F(\mu)$. Contradiction. So for $\mu \in [0,1]$, $N_{dt}(\mu) = \{\mu\}$ for any $dt > 0$. Noticing that this satisfies Theorem A.1 independent of choice of dt and ε.

Then consider the case with Assumption A and Assumptions 1.2-b and 1.3. Suppose not true, and there exists $\nu_i \neq \mu$ s.t.:

$$
e^{-\rho dt} \sum p_i V_{dt}(\nu_i) - dt \cdot C\left(\frac{I}{dt}\right) \geqslant V_{dt}(\mu)
$$

and

$$
\begin{cases}
\sum p_i \nu_i = \mu \\
H(\mu) - \sum p_i H(\nu_i) = I
\end{cases}
$$

Now consider a problem with $\frac{dt}{2}$. Consider the following strategy: mix experiment p_i, ν_i and prior with probability $\frac{1}{2}$. Then obviously Bayes plausibility and capacity constraint are satisfied. The expected utility from this strategy is:

$$
\begin{aligned}
& V_{\frac{dt}{2}}(\mu) \\
\geqslant & \sum_{t=1} e^{-\rho \frac{dt}{2} \cdot t} \sum p_i V_{\frac{dt}{2}}(\nu_i) \cdot \frac{1}{2^t} - \sum_{t=0} e^{-\rho \frac{dt}{2} t} \frac{1}{2^t} \frac{dt}{2} C\left(\frac{I}{dt}\right) \\
= & \frac{e^{\rho dt}}{2 - e^{-\frac{\rho dt}{2}}} \left(e^{-\rho dt} \sum p_i V_{\frac{dt}{2}}(\nu_i) - e^{-\frac{3\rho dt}{2}} dt C\left(\frac{I}{dt}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&> e^{-\rho dt} \sum p_i V_{\frac{dt}{2}}(\nu_i) - dt C\left(\frac{I}{dt}\right) \\
&\geqslant e^{-\rho dt} \sum p_i V_{dt}(\nu_i) - dt C\left(\frac{I}{dt}\right) \\
&\geqslant F(\mu)
\end{aligned}
$$

First inequality is from optimality of $V_{\frac{dt}{2}}$. Second inequality is from $\frac{1}{2-x} > x^2$ for $x \in (0,1)$. Third inequality is from $V_{\frac{dt}{2}} \geqslant V_{dt}$. Last inequality is from assumption. Therefore $F(\mu) = V(\mu) \geqslant V_{\frac{dt}{2}} > F(\mu)$. Contradiction. So for $\mu \in [0,1]$, $N_{dt}(\mu) = \{\mu\}$ for any $dt > 0$. Noticing that this satisfies Theorem A.1 independent of choice of dt and ε.

Case 2: $\mu \in E$. Suppose Theorem A.1 is not true. Then there exists ε s.t. $\forall dt$, there exists $\mu_{dt} \in E$ s.t. $\exists \nu_{dt} \in N_{dt}(\mu_{dt})$ and $\forall \mu \in B_\varepsilon(\mu_{dt})$, $d_H(\nu_{dt}, N(\mu)) > \varepsilon$. Now pick $dt_n = 2^{-n} \to 0$. Since (μ_{dt_n}, ν_{dt_n}) is an infinite sequence in compact space $[0,1]^2$, we can WLOG assume $\left(\mu_{dt_n}, \nu_{dt_n}\right) \to (\mu, \nu)$. $\forall \mu' \in B_{\frac{\varepsilon}{2}}(\mu)$, there exists N sufficiently large that $\forall n \geqslant N$, $\mu' \in B_\varepsilon(\mu_{dt_n})$ and $d_H\left(\nu_{dt_n}, N(\mu')\right) > \varepsilon$. N can be picked sufficiently large that $\left|\nu_{dt_n} - \nu\right| < \frac{\varepsilon}{2}$. Therefore $d_H(\nu, N(\mu')) > \frac{\varepsilon}{2}$. To sum up, we find a converging sequence $\left(\mu_{dt_n}, \nu_{dt_n}\right)$ to (μ, ν), which is bounded away by $\frac{\varepsilon}{2}$ from the graph of $N(\cdot)$.

Let $\widetilde{\nu}$ be the non-empty set of optimal posteriors (including μ itself) at μ solving Equation (1.4). Let

$$
\lambda = \begin{cases} \dfrac{\rho}{c} V(\mu) & \text{with Assumption 1.2-a} \\ C'(I(\mu)) & \text{with Assumption 1.2-b} \end{cases}
$$

Consider:

$$
G(\cdot) = V(\cdot) + \lambda H(\cdot)
$$

Then optimality condition implies that:

$$\begin{cases} G(\nu) = G(\mu) + G'(\mu)(\nu - \mu) & \forall \nu \in \tilde{\nu} \\ G(\nu) < G(\mu) + G'(\mu)(\nu - \mu) & \text{otherwise} \end{cases} \tag{B.35}$$

By Theorem B.1, $\forall dt$, there exists λ_{dt_n} s.t. Equation (1.6) is solved by concavifying $G_n = V_{dt_n} + \lambda_{dt_n} H$ at μ_{dt_n}.

Obviously λ_{dt_n} is non-negative. Suppose it diverges to $+\infty$. Then consider function $G_{dt_n}(\cdot) = V_{dt_n}(\cdot) + \lambda_{dt_n} H(\cdot)$. Let $\nu_{dt_n,1}$ and $\nu_{dt_n,2}$ be two optimal posterior. Let $\nu'_{dt_n} = \frac{1}{2}(\nu_{dt_n,1} + \nu_{dt_n,2})$. By $|\nu_{dt_n} - \mu_{dt_n}| > \frac{\varepsilon}{2}$, we know that $|\nu_{dt_n,1} - \nu_{dt_n,2}| > \frac{\varepsilon}{2}$ by Lemma B.25.

$$\begin{aligned}
&G_{dt_n}(\nu'_{dt_n}) - \frac{1}{2}\left(G_{dt_n}(\nu_{dt_n,1}) + G_{dt_n}(\nu_{dt_n,2})\right) \\
=&V_{dt_n}(\nu'_{dt_n}) - \frac{1}{2}\left(V_{dt_n}(\mu_{dt_n,1}) + V_{dt_n}(\nu_{dt_n,1})\right) \\
&+ \lambda_{dt_n}\left(H(\nu'_{dt_n}) - \frac{1}{2}\left(H(\nu_{dt_n,1}) + H(\nu_{dt_n,2})\right)\right) \\
\leqslant &\sup F + \lambda_{dt_n}\frac{1}{8}\sup H''\left(\nu_{dt_n,1} - \nu_{dt_n,2}\right)^2 \to -\infty
\end{aligned}$$

So for n sufficiently large, μ'_{dt_n} will be higher than the connected straight line of μ_{dt_n} and ν_{dt_n} on G_{dt_n}. Contradicting optimality of ν_{dt_n}. So λ_{dt_n} is a bounded sequence.

Suppose there exists convergent subsequence $\lim \lambda_{dt_n} < \lambda$. If $\mu \neq \mu^*$ then pick μ' such that μ' is in the same interval as μ in E and $\lim \lambda_{dt_n} < \lambda(\mu') < \lambda$, let $\lambda' = \lambda(\mu')$. If $\mu = \mu^*$ then let $\lambda(\mu^*) > \lambda' > \lim \lambda_{dt_n}$. Now consider concavifying $V + \lambda' H$. By monotonicity in Theorem 1.2 and definition of μ^*, we know that optimal posteriors are bounded away from μ. Moreover, $V(\mu) + \lambda' H(\mu) < \text{cov}(V + \lambda' H)(\mu)$. Pick $\varepsilon > 0$ sufficiently small such that optimal posteriors of $V + \lambda' H$ are bounded away from μ by ε and $V(\mu) + \lambda' H(\mu) + \varepsilon < \text{cov}(V + \lambda' H)(\mu)$. Let $\nu_1 < \mu < \nu_2$ be two optimal posterios for $V + \lambda' H$ closest to μ. By continuity, there exists δ s.t $\forall |\mu'' - \mu| < \delta$, $V(\mu'') + \lambda' H(\mu'') + \frac{\varepsilon}{2} < \text{cov}(V + \lambda' H)(\mu'')$. Pick

dt_n s.t. $\|V_{dt_n} - V\| < \frac{\varepsilon}{8}$, then

$$\begin{aligned}
&V_{dt_n}(\mu'') + \lambda' H(\mu'') \\
<&V(\mu'') + \lambda' H(\mu'') + \frac{\varepsilon}{8} \\
<&\mathrm{cov}(V + \lambda' H)(\mu'') - \frac{3\varepsilon}{8} \\
\leqslant&\mathrm{cov}(V_{dt_n} + \lambda' H)(\mu'') - \frac{\varepsilon}{4}
\end{aligned}$$

The last inequality comes from the fact that any convex combination of points on $V + \lambda' H$ is less than $\frac{\varepsilon}{8}$ higher than convex combination of those points on $V_{dt_n} + \lambda' H$, therefore less than $\mathrm{cov}\big(V_{dt_n} + \lambda' H\big) + \frac{\varepsilon}{8}$. Therefore, we showed that any point μ'' within δ ball of μ can't be on supporting hyperplane of $V_{dt_n} + \lambda' H$. So any optimal posterior of $V_{dt_n} + \lambda' H$ is bounded away from μ by δ. Pick N sufficently large than $\forall n \geqslant N$, $\left|\mu_{dt_n} - \mu\right| < \frac{\delta}{2}$. Then, optimal posterior of $V_{dt_n} + \lambda' H$ is bounded away from μ_{dt_n} by $\frac{\delta}{2}$. By definition of λ', N can be picked sufficently large that $\forall n \geqslant N$ $\lambda_{dt_n} \leqslant \lambda'$. Therefore, by Lemma B.25, optimal posteriors are even further from μ_{dt_n}. To sum up, we found N s.t. $\forall n \geqslant N$, the optimal posteriors from concavifying $V_{dt_n} + \lambda_{dt_n} H$ are bounded away from μ_{dt_n} by $\frac{\delta}{2}$. The experimentation cost of any such information structure is:

$$\begin{aligned}
&\sum p_i\big(H(\mu_{dt_n}) - H(\nu_{dt_n,i})\big) \\
=&\sum p_i\left(H'(\mu_{dt_n})\big(\mu_{dt_n} - \nu_{dt_n,i}\big) - \frac{1}{2}H''(\xi_i)\big(\nu_{dt_n,i} - \mu_{dt_n}\big)^2\right) \\
\geqslant& -\sup H'' \frac{\delta^2}{4} > 0
\end{aligned}$$

Therefore, for sufficently large n, experimentation cost will eventually exceed cdt_n. Contradiction.

Suppose there exists subsequence $\lim \lambda_{dt_n} = \lambda' \geqslant \lambda$. By definition, there exists linear function $L_n(\mu)$ s.t.

$$\begin{cases} G_n(\nu_{dt_n}) = L_n(\nu_{dt_n}) \\ G_n(\mu_{dt_n}) = L_n(\mu_{dt_n}) \\ G_n(\nu) \leqslant L_n(\nu) \end{cases}$$

Since $\lambda_{dt_n} \to \lambda'$, G_n is bounded at μ_{dt_n} and ν_{dt_n}. Therefore, L_n has bounded slope and constant term. It's easy to see that L_n will converge uniformly to linear function L_∞ on belief space ΔX. Moreover, $\forall \nu \in \Delta X$:

$$\begin{aligned} G_n(\nu) = V_{dt_n}(\nu) + \lambda_{dt_n} H(\nu) \to V(\nu) + \lambda' H(\nu) = \widetilde{G}(\nu) \leqslant L_\infty(\nu) \\ G_n(\mu_{dt_n}) = V_{dt_n}(\mu_{dt_n}) + \lambda_{dt_n} H(\mu_{dt_n}) \to \widetilde{G}(\mu) = L_\infty(\mu) \\ G_n(\nu_{dt_n}) = V_{dt_n}(\nu_{dt_n}) + \lambda_{dt_n} H(\nu_{dt_n}) \to \widetilde{G}(\nu) = L_\infty(\nu) \end{aligned} \tag{B.36}$$

Second and third convergence comes from V_{dt_n} uniformly convergent and V continuous. $\widetilde{G}(\mu) = G(\mu) + (\lambda' - \lambda)H(\mu)$. Equation (B.36) implies that L_∞ is a supporting hyperplane of graph of $\widetilde{G}$, tangents $\widetilde{G}$ at μ and ν. Since $\widetilde{G}$ is a smooth function, we know that $\widetilde{G}'(\mu) = \frac{\widetilde{G}(\nu) - \widetilde{G}(\mu)}{\nu - \mu}$. On the other hand, Equation (B.35) implies that:

$$\begin{aligned} & G(\nu) < G(\mu) + G'(\mu)(\nu - \mu) \\ \Longrightarrow & \left(\widetilde{G}(\nu) - \widetilde{G}(\mu)\right) - (\lambda' - \lambda)(H(\nu) - H(\mu)) < \left(\widetilde{G}'(\mu) - (\lambda' - \lambda)H'(\mu)\right)(\nu - \mu) \\ \Longrightarrow & \widetilde{G}(\nu) - \widetilde{G}(\mu) < \widetilde{G}'(\mu)(\nu - \mu) \end{aligned}$$

Contradiction. Last inequality is from concavity of $H : H(\nu) - H(\mu) < H'(\mu)(\nu - \mu)$. Therefore Theorem A.1 is true.

■

Lemma B.25. *Let X be closed interval in $\mathbb{R}$. Let V be a continuous function on X, H be a concave function on X. Let $E_\lambda = \{x \in X | cov(V + \lambda H)(x) > V(x) + \lambda H(x)\}$. Then $\{E_\lambda\}$ are ordered monotonically as λ by set inclusion: if $\lambda \geqslant \lambda'$, then $\forall$ interval I in E_λ, there exists interval I' in $E_{\lambda'}$ s.t. $I \subset I'$.*

Proof. $\forall \lambda$, take any $I \in E_\lambda$. Let $I = [x, y]$. Define:

$$L(z) = V(x) + \lambda H(x) + \frac{V(y) - V(x) + \lambda H(y) - \lambda H(x)}{y - x}(z - x)$$

Then $\forall z \in X$:

$$\begin{cases} L(x) = V(x) + \lambda H(x) \\ L(y) = V(y) + \lambda H(y) \\ L(z) > V(z) + \lambda H(z) & \text{if } z \in (x, y) \\ L(z) \geqslant V(z) + \lambda H(z) & \text{if } z < x \text{ or } z > y \end{cases}$$

Now take any $\lambda' < \lambda$ and consider $V + \lambda' H$. Let:

$$\begin{aligned} \widetilde{L}(z) &= V(x) + \lambda' H(x) + \frac{V(y) - V(x) + \lambda' H(y) - \lambda' H(x)}{y - x}(z - x) \\ &= L(z) + (\lambda' - \lambda)\left(H(x) + \frac{H(y) - H(x)}{y - x}(z - x)\right) \\ &\begin{cases} \geqslant L(z) + (\lambda' - \lambda)H(z) & \text{if } z \in [x, y] \\ \leqslant L(z) + (\lambda' - \lambda)H(z) & \text{if } z \notin [x, y] \end{cases} \\ \Longrightarrow &\begin{cases} \widetilde{L}(x) = V(x) + \lambda' H(x) \\ \widetilde{L}(y) = V(y) + \lambda' H(y) \\ \widetilde{L}(z) > V(z) + \lambda' H(z) \text{ if } z \in (x, y) \end{cases} \end{aligned}$$

Therefore, $\forall z \in (x, y)$, $\text{cov}(V + \lambda' H)(z) > V(z) + \lambda' H(z)$. So there exists interval $I' \in E_{\lambda'}$

s.t. $I \subset I'$. ■

B.5.2 Continuum of actions

B.5.2.1 Proof of Lemma A.1

Proof. We prove with two steps:

Step 1: We first show that if we let $\mathcal{V}_{dt}(F)$ be the solution to Equation (1.6), then $\mathcal{V}_{dt}$ is Lipschitz continuous in F under L_∞ norm. $\forall F_1, F_2$ convex and with bounded subdifferentials, consider $\overline{F} = \max\{F_1, F_2\}$, $\underline{F} = \min\{F_1, F_2\}$. Then by properties of convex functions, $\overline{F}, \underline{F}$ are convex. $\partial\underline{F}(\mu), \partial\overline{F}(\mu) \subset \partial F_1(\mu) \bigcup \partial F_2(\mu)$. Therefore $\overline{F}$ and $\underline{F}$ are both within the domain of convex and bounded subdifferential functions with the following quantitative property:

$$\begin{cases} \overline{F} \geqslant F_1, F_2 \geqslant \underline{F} \\ \left|\overline{F} - \underline{F}\right| = |F_1 - F_2| \end{cases}$$

It's not hard to see that $\mathcal{V}$ is monotonicaly increasing in F. Therefore, we have:

$$\mathcal{V}_{dt}(\underline{F}) \leqslant \mathcal{V}_{dt}(F_1), \mathcal{V}_{dt}(F_2) \leqslant \mathcal{V}_{dt}(\overline{F})$$

Now let (p_i, μ_i) be the policy solving $\mathcal{V}_{dt}(\overline{F})$. Let $\overline{V}_{dt} = \mathcal{V}_{dt}(\overline{F}), \underline{V}_{dt} = \mathcal{V}_{dt}(\underline{F})$. Let C be total expected cost associate with this strategy. Then consider:

$$\begin{aligned}
\underline{V}_{dt}(\mu) \geqslant & \mathbf{1}_{\overline{V}_{dt}(\mu) \leqslant \overline{F}(\mu)} \underline{F}(\mu) + \mathbf{1}_{\overline{V}_{dt}(\mu) > \overline{F}(\mu)} e^{-\rho dt} \sum p_i^1(\mu) \underline{V}_{dt}(\mu_i^1) - C \\
\geqslant & \mathbf{1}_{\overline{V}_{dt}(\mu) \leqslant \overline{F}(\mu)} \underline{F}(\mu) + \mathbf{1}_{\overline{V}_{dt}(\mu) > \overline{F}(\mu)} e^{-\rho dt} \sum p_i^1(\mu) \mathbf{1}_{\overline{V}_{dt}(\mu_i^1) \leqslant \overline{F}(\mu_i^1)} \underline{F}(\mu_i^1) - C \\
& + \mathbf{1}_{\overline{V}_{dt}(\mu) > \overline{F}(\mu)} e^{-2\rho dt} \sum p_i^1(\mu) \mathbf{1}_{\overline{V}_{dt}(\mu_i^1) > \overline{F}(\mu_i^1)} \sum p_i^2(\mu_i^1) \underline{V}_{dt}(\mu_i^2) - C \\
\geqslant & \cdots
\end{aligned}$$

$$
\begin{aligned}
=&\sum_t e^{-\rho t\cdot dt}\sum_{i^1,\dots,i^{t-1}}\prod p_i^\tau(\mu_i^{\tau-1})\mathbf{1}_{\overline{V}_{dt}(\mu_i^\tau)>\overline{F}(\mu_i^\tau)}\sum p_i^t(\mu_i^{t-1})\mathbf{1}_{\overline{V}_{dt}(\mu_i^t)\leqslant\overline{F}(\mu_i^t)}\underline{F}(\mu_i^t)-C\\
\geqslant&\sum_t e^{-\rho t\cdot dt}\sum_{i^1,\dots,i^{t-1}}\prod p_i^\tau(\mu_i^{\tau-1})\mathbf{1}_{\overline{V}_{dt}(\mu_i^\tau)>\overline{F}(\mu_i^\tau)}\sum p_i^t(\mu_i^{t-1})\mathbf{1}_{\overline{V}_{dt}(\mu_i^t)\leqslant\overline{F}(\mu_i^t)}\overline{F}(\mu_i^t)-C\\
&-\sum_t\sum_{i^1,\dots,i^{t-1}}\prod p_i^\tau(\mu_i^{\tau-1})\mathbf{1}_{\overline{V}_{dt}(\mu_i^\tau)>\overline{F}(\mu_i^\tau)}\sum p_i^t(\mu_i^{t-1})\mathbf{1}_{\overline{V}_{dt}(\mu_i^t)\leqslant\overline{F}(\mu_i^t)}\left|\overline{F}-\underline{F}\right|\\
=&\overline{V}_{dt}(\mu)-\left|\overline{F}-\underline{F}\right|
\end{aligned}
$$

Therefore, $\left|\overline{V}_{dt}-\underline{V}_{dt}\right|\leqslant\left|\overline{F}-\underline{F}\right| \implies |\mathcal{V}_{dt}(F_1)-\mathcal{V}_{dt}(F_2)|\leqslant|F_1-F_2|$. $\mathcal{V}_{dt}(F)$ has Lipschitz parameter 1.

Step 2: $\forall F_1,F_2,\forall\varepsilon>0$, by Lemma 1.3, there exists dt s.t. $|\mathcal{V}(F_i)-\mathcal{V}_{dt}(F_i)|\leqslant\varepsilon|F_1-F_2|$. Therefore:

$$
\begin{aligned}
|\mathcal{V}(F_1)-\mathcal{V}(F_2)|\leqslant&|\mathcal{V}(F_1)-\mathcal{V}_{dt}(F_1)|+|\mathcal{V}(F_2)-\mathcal{V}_{dt}(F_2)|+|\mathcal{V}_{dt}(F_1)-\mathcal{V}_{dt}(F_2)|\\
\leqslant&(1+2\varepsilon)|F_1-F_2|
\end{aligned}
$$

Take $\varepsilon\to 0$, since LHS is not a function of ε, we conclude that $\mathcal{V}(F)$ is Lipschitz continuous in F with Lipschitz parameter 1. ∎

B.5.2.2 *Proof of Theorem A.2*

Proof. We prove the three main results in following steps:

- *Lipschitz continuity*. By Lemma A.1, we directly get Lipschitz continuity of operator $\mathcal{V}$ on $\{F_n,F\}$ and the Lipschitz parameter being 1.

- *Convergence of derivatives*. Let $V_n=\mathcal{V}(F_n)$, $V=\mathcal{V}(F)$, we show that $\forall\mu$ s.t. $V(\mu)>F(\mu)$, $V'(\mu)=\lim V_n'(\mu)$. Since $V(\mu)>F(\mu)$, by continuity strict inequality holds in an closed interval $[\mu_1,\mu_2]$ around μ. Then by Lemma B.27, $\lim_{n\to\infty}V_n'(\mu')$ exists $\forall\mu'\in[\mu_1,\mu_2]$. Now consider function $V_n'(\mu)$. Since $V_n''(\mu)$ is uniformly bounded for all n, $V_n'(\mu)$ are uniformly Lipschitz continuous, thus equicontinuous and totally bounded. Therefore by lemma Arzela-Ascolli, V_n' converges uniformly to $\lim V_n'$. By convergence theorem of derivatives, $V' = \lim V_n'$ on $[\mu_1, \mu_2]$. Therefore, $V'(\mu) = \lim_{n\to\infty} V_n'(\mu)$.

- *Feasibility*. For μ s.t. $V(\mu) = F(\mu)$, feasibility is trivial. Now we discuss the case $V(\mu) > F(\mu)$. We only prove for $\mu > \mu^*$ and $\mu = \mu^*$, the case $\mu < \mu^*$ follows by symmetry. If $\mu > \mu^*$, there exists N s.t. $\forall n \geqslant N$, $\mu > \mu_n^*$. N can be picked large enough that in a closed interval around μ, $V_n(\mu) > F_n(\mu)$. Therefore, there exists ν_n being maximizer for $V_n(\mu)$ bounded away from μ and satisfying:

$$V_n(\mu) = \frac{c}{\rho}\frac{F_n(\nu_n) - V_n(\mu) - V_n'(\mu)(\nu_n - \mu)}{J(\mu, \nu_n)}$$

Pick a converging subsequence $\nu_n \to \nu$:

$$\begin{aligned}&\frac{c}{\rho}\frac{F(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu_n)}\\ =&\lim_{n\to\infty}\frac{c}{\rho}\frac{F_n(\nu_n) - V_n(\nu) - V_n'(\nu)(\nu_n - \mu)}{J(\mu, \nu_n)}\\ =&\lim_{n\to\infty} V_n(\mu)\\ =&V(\mu)\end{aligned}$$

Therefore $V(\mu)$ is feasible in Equation (A.4).

Suppose $\mu = \mu^*$. Then there exists a subsequence of μ_n^* converging from one side of μ^*. Suppose they are converging from left. Then $\mu \geqslant \mu_n^*$. Previous proof still works. Essentially, what we showed is that the limit of strategy in discrete action problem achieves $V(\mu)$ in the continuous action limit.

- *Unimprovability*. First, when $\mu \in \{0, 1\}$, information provides no value but discounting is costly, therefore $V(\mu)$ is unimprovable. We now show unimprovability on $(0, 1)$ by adding more feasible information acquisition strategies in several steps.

– *Step 1*. Poisson experiments at $V(\mu) > F(\mu)$. In this step, we show that $\forall \mu \geqslant \mu^*$ and $V(\mu) > F(\mu)$:

$$\rho V(\mu) = \max_{\nu \geqslant \mu} c \frac{F(\nu) - V(\mu) - V'(\mu)(\nu - \mu)}{J(\mu, \nu)}$$

Suppose not true, then there exists ν s.t.:

$$\begin{aligned} \lim_{n\to\infty} \rho V_n(\mu) =& \rho V(\mu) \\ <& c \frac{F(\nu) - V(\mu) - V'(\nu)(\nu - \mu)}{J(\mu, \nu)} \\ =& \lim_{n\to\infty} c \frac{F_n(\nu) - V_n(\mu) - V_n'(\mu)(\nu - \mu)}{J(\mu, \nu)} \\ \leqslant& \lim_{n\to\infty} \rho V_n(\mu) \end{aligned}$$

Second line is by the contradictory assumption. Third line is by convergence of F_n by assumption, convergence of V_n by Lemma A.1 and convergence of V_n' by Lemma B.27. Last inequality is by suboptimality of ν.

Similarly, for the case $\mu \leqslant \mu^*$, we can apply a symmetric argument to prove.

– *Step 2*. Poisson experiments at $V(\mu) = F(\mu)$. In this step, we show that $\forall \mu \geqslant \mu^*$ and $V(\mu) = F(\mu)$ (The symmetric case $\mu \leqslant \mu^*$ is ommited).

First of all, we show that V is differentiable at μ and $V'(\mu) = F'(\mu)$. Suppose not, then since $V(\mu) = F(\mu)$ and $V \geqslant F$, we know that $V - F$ is locally minimized at μ. Therefore $DV_+(\mu) > DV_-(\mu)$. By Definition A.2, there exists $\varepsilon > 0$, $\mu_1^n \nearrow \mu$ and $\mu_2^n \searrow \mu$ s.t. $\frac{V(\mu_2^n) - V(\mu)}{\mu_2^n - \mu} \geqslant \varepsilon + \frac{V(\mu) - V(\mu_1^n)}{\mu - \mu_1^n}$. Let $\delta_1^n = \mu - \mu_1^n, \delta_2^n = \mu_2^n - \mu$, this implies:

$$\frac{\mu - \mu_1^n}{\mu_2^n - \mu_1^n}(V(\mu_2^n) - V(\mu)) + \frac{\mu_2^n - \mu}{\mu_2^n - \mu_1^n}(V(\mu_1^n) - V(\mu)) \geqslant \varepsilon \frac{(\mu_2^n - \mu)(\mu - \mu_1^n)}{\mu_2^n - \mu_1^n}$$

$$\implies \frac{\mu-\mu_1^n}{\mu_2^n-\mu_1^n}V(\mu_2^n)+\frac{\mu_2^n-\mu}{\mu_2^n-\mu_1^n}V(\mu_1^n)\geqslant V(\mu)+\varepsilon\cdot\min\{\delta_1^n,\delta_2^n\}$$

On the other hand:

$$\begin{aligned}
&\frac{\mu-\mu_1^n}{\mu_2^n-\mu_1^n}(H(\mu)-H(\mu_2^n))+\frac{\mu_2^n-\mu}{\mu_2^n-\mu_1^n}(H(\mu)-H(\mu_1^n))\\
=&\frac{\mu-\mu_1^n}{\mu_2^n-\mu_1^n}\left(H'(\mu)(\mu-\mu_2^n)+\frac{1}{2}H''(\xi_2^n)(\mu-\mu_2^n)\right)\\
&+\frac{\mu_2^n-\mu}{\mu_2^n-\mu_1^n}\left(H'(\mu)(\mu-\mu_1^n)+\frac{1}{2}H''(\xi_1^n)(\mu-\mu_1^n)^2\right)\\
=&\frac{1}{2}\frac{(\mu_2^n-\mu)(\mu-\mu_1^n)}{\mu_2^n-\mu_1^n}\left(H''(\xi_2^n)(\mu_2^n-\mu)+H''(\xi_1^n)(\mu-\mu_1^n)\right)
\end{aligned}$$

ξ_1^n and ξ_2^n are determined by applying intermediate value theorem on H'. Now we can choose N s.t. $\forall n\geqslant N$, $\max_{\mu'\in[\mu_1^n,\mu_2^n]}\{H''(\mu')\}\leqslant 2H''(\mu)$. Therefore:

$$\begin{aligned}
&\frac{\mu-\mu_1^n}{\mu_2^n-\mu_1^n}(H(\mu)-H(\mu_2^n))+\frac{\mu_2^n-\mu}{\mu_2^n-\mu_1^n}(H(\mu)-H(\mu_1^n))\\
\leqslant&H''(\mu)(\mu_2^n-\mu)(\mu-\mu_1^n)\\
=&H''(\mu)\delta_1^n\delta_2^n
\end{aligned}$$

Now we consider a stationary experiment at μ that takes any experiment with posteriors (μ_1^n,μ_2^n) with flow probability $\frac{c}{H''(\mu)\delta_1^n\delta_2^n}$. Then by definition the flow cost of this information acquisition strategy is less than c, thus is feasible. The expected utility is:

$$\begin{aligned}
\widetilde{V}(\mu)=&\frac{c}{\rho}\frac{\frac{\mu-\mu_1^n}{\mu_2^n-\mu_1^n}V(\mu_2^n)+\frac{\mu_2^n-\mu}{\mu_2^n-\mu_1^n}V(\mu_1^n)-\widetilde{V}(\mu)}{\frac{\mu-\mu_1^n}{\mu_2^n-\mu_1^n}\left(H(\mu)-H(\mu_2^n)\right)+\frac{\mu_2^n-\mu}{\mu_2^n-\mu_1^n}\left(H(\mu)-H(\mu_1^n)\right)}\\
\geqslant&\frac{V(\mu)-\widetilde{V}(\mu)+\varepsilon\min\{\delta_1^n,\delta_2^n\}}{H''(\mu)\delta_1^n\delta_2^n}
\end{aligned}$$

$$\begin{aligned}
\implies \widetilde{V}(\mu) \geqslant & \frac{V(\mu) + \varepsilon \min\{\delta_1^n, \delta_2^n\}}{1 + \frac{\rho}{c} H''(\mu)\delta_1^n\delta_2^n} \\
= & V(\mu) + \frac{\varepsilon \min\{\delta_1^n, \delta_2^n\} - \frac{\rho}{c} H''(\mu)\delta_1^n\delta_2^n}{1 + \frac{\rho}{c} H''(\mu)\delta_1^n\delta_2^n} \\
= & V(\mu) + \min\{\delta_1^n, \delta_2^n\} \frac{\varepsilon - H''(\mu)\max\{\delta_1^n, \delta_2^n\}}{1 + \frac{\rho}{c} H''(\mu)\delta_1^n\delta_2^n}
\end{aligned}$$

n can be pick large enough that $\varepsilon - H''(\mu)\max\{\delta_1^n, \delta_2^n\}$ is positive. Therefore $\widetilde{V}(\mu) > V(\mu)$. Now fix n and define:

$$\begin{aligned}
\widetilde{V}_m(\mu) = & \frac{c}{\rho} \frac{\frac{\mu - \mu_1^n}{\mu_2^n - \mu_1^n} V_m(\mu_2^n) + \frac{\mu_2^n - \mu}{\mu_2^n - \mu_1^n} V_m(\mu_1^n) - \widetilde{V}_m(\mu)}{\frac{\mu - \mu_1^n}{\mu_2^n - \mu_1^n}\left(H(\mu) - H(\mu_2^n)\right) + \frac{\mu_2^n - \mu}{\mu_2^n - \mu_1^n}\left(H(\mu) - H(\mu_1^n)\right)} \\
\implies \lim_{m\to\infty} \widetilde{V}_m(\mu) = & \widetilde{V}(\mu) > \lim_{m\to\infty} V_m(\mu)
\end{aligned}$$

There exists m large enough that $\widetilde{V}_m(\mu) > V_m(\mu)$, violating optimality of V_m. Contradiction. Therefore, we showed that $V'(\mu) = F'(\mu)$.

Next we show unimperovability. Suppose not, then $\exists \nu$ s.t.:

$$F(\mu) < \frac{c}{\rho} \frac{F(\nu) - F(\mu) - F'(\mu)(\nu - \mu)}{J(\mu, \nu)}$$

By continuity of V, $\exists \varepsilon$ and a neighbourhood $\mu \in O$, $\forall \mu' \in O$:

$$V(\mu') + \varepsilon \leqslant \frac{c}{\rho} \frac{F(\nu) - V(\mu') - F'(\mu)(\nu - \mu')}{J(\mu', \nu)}$$

By uniform convergence of F_n and V_n, there exists $\varepsilon > 0$ and N s.t. $\forall n \geqslant N$:

$$\begin{aligned}
& V_n(\mu') + \frac{\varepsilon}{2} \leqslant \frac{c}{\rho} \frac{F_n(\nu) - V_n(\mu') - F'(\mu)(\nu - \mu')}{J(\mu', \nu)} \\
\implies & \frac{c}{\rho} \frac{F_n(\nu) - V_n(\mu') - V_n'(\mu')(\nu - \mu')}{J(\mu', \nu)} + \frac{\varepsilon}{2} \leqslant \frac{c}{\rho} \frac{F_n(\nu) - V_n(\mu') - F'(\mu)(\nu - \mu')}{J(\mu', \nu)}
\end{aligned}$$

$$\implies V_n'(\mu') \geqslant F'(\mu) + \frac{\rho\varepsilon}{2c}\frac{J(\mu',\nu)}{\nu-\mu'}$$

In an interval around μ, $V_n'(\mu') - F'(\mu) \geqslant \frac{\rho\varepsilon}{2c}\frac{J(\mu',\nu)}{\nu-\mu'}$, which is a positive number independent of n and uniformly bounded away from 0 for all μ'. Then it's impossible that $V'(\mu) = F'(\mu)$. Contradiction.

What's more, since V' is Lipschitz continuous at any $V(\mu) > F(\mu)$, it can be extended smoothly to the boundary. Since $V' = F'$ at $V(\mu) = F(\mu)$, then the limit of this smooth extension has $\lim V'(\mu) = F'(\mu)$. Therefore V is $C^{(1)}$ smooth on $[0,1]$.

– *Step 3*. Repeated experiments and contradictory experiments. With the convergence result we have on hand, we can apply similar proof by contradiction method in step 1 and 2 to rule out these two cases. We omitted the proofs here. Therefore:

$$V(\mu) = \max\left\{F(\mu), \max_{\nu}\frac{c}{\rho}\frac{V(\nu)-V(\mu)-V'(\mu)(\nu-\mu)}{J(\mu,\nu)}\right\}$$

– *Step 4*. Diffusion experiments. Suppose at μ, diffusion experiment is strictly optimal:

$$V(\mu) < -\frac{c}{\rho}\frac{D^2V(\mu)}{H''(\mu)}$$

Then by Definition A.2, there exists ε, δ_1 s.t.:

$$V(\mu) + \varepsilon \leqslant \frac{c}{\rho}\frac{V(\mu+\delta_1)-V(\mu)-V'(\mu)\delta_1}{H(\mu)-H(\mu+\delta_1)+H'(\mu)\delta_1}$$

Then by definition of derivative, there exists δ_2 s.t.:

$$V(\mu) + \frac{\varepsilon}{2} \leqslant \frac{c}{\rho}\frac{\frac{\delta_2}{\delta_1+\delta_2}(V(\mu+\delta_1)-V(\mu)) + \frac{\delta_2}{\delta_1+\delta_2}(V(\mu-\delta_2)-V(\mu))}{\frac{\delta_2}{\delta_1+\delta_2}(H(\mu)-H(\mu+\delta_1)) + \frac{\delta_2}{\delta_1+\delta_2}(H(\mu)-H(\mu-\delta_2))}$$

By convergence of V_n, there exists n s.t.:

$$
\begin{aligned}
&V_n(\mu)+\frac{\varepsilon}{4}\leqslant\frac{c}{\rho}\frac{\frac{\delta_2}{\delta_1+\delta_2}(V_n(\mu+\delta_1)-V_n(\mu))+\frac{\delta_2}{\delta_1+\delta_2}(V_n(\mu-\delta_2)-V_n(\mu))}{\frac{\delta_2}{\delta_1+\delta_2}(H(\mu)-H(\mu+\delta_1))+\frac{\delta_2}{\delta_1+\delta_2}(H(\mu)-H(\mu-\delta_2))}\\
\implies&\frac{\delta_2}{\delta_1+\delta_2}V_n(\mu+\delta_1)+\frac{\delta_1}{\delta_1+\delta_2}V_n(\mu-\delta_2)\\
&\geqslant V_n(\mu)\left(1+\frac{\rho}{c}\left(H(\mu)-\frac{\delta_2}{\delta_1+\delta_2}H(\mu+\delta_1)-\frac{\delta_1}{\delta_1+\delta_2}H(\mu-\delta_2)\right)\right)\\
&\quad+\frac{\rho}{c}\left(H(\mu)-\frac{\delta_2}{\delta_1+\delta_2}H(\mu+\delta_1)-\frac{\delta_1}{\delta_1+\delta_2}H(\mu-\delta_2)\right)\frac{\varepsilon}{4}
\end{aligned}
$$

If we consider the experiment with posterior beliefs $\mu+\delta_1$, $\mu-\delta_2$ at μ. Taking this experiment at μ with flow probability:

$$
\frac{c}{H(\mu)-\frac{\delta_2}{\delta_1+\delta_2}H(\mu+\delta_1)-\frac{\delta_1}{\delta_1+\delta_2}H(\mu-\delta_2)}
$$

Then the flow cost constraint will be satisfied and the utility gain is:

$$
\begin{aligned}
\widetilde{V}_n(\mu)=&\frac{\frac{\delta_2}{\delta_1+\delta_2}V_n(\mu+\delta_1)+\frac{\delta_1}{\delta_1+\delta_2}V_n(\mu-\delta_2)}{1+\frac{\rho}{c}\left(H(\mu)-\frac{\delta_2}{\delta_1+\delta_2}H(\mu+\delta_1)-\frac{\delta_1}{\delta_1+\delta_2}H(\mu-\delta_2)\right)}\\
&\geqslant V_n(\mu)+\frac{\frac{\rho}{c}\left(H(\mu)-\frac{\delta_2}{\delta_1+\delta_2}H(\mu+\delta_1)-\frac{\delta_1}{\delta_1+\delta_2}H(\mu-\delta_2)\right)}{1+\frac{\rho}{c}\left(H(\mu)-\frac{\delta_2}{\delta_1+\delta_2}H(\mu+\delta_1)-\frac{\delta_1}{\delta_1+\delta_2}H(\mu-\delta_2)\right)}\frac{\varepsilon}{4}\\
&>V_n(\mu)
\end{aligned}
$$

Contradiction.

To sum up, we proved that $V(\mu)$ solves Equation (A.4).

∎

Lemma B.26 (Convergence of μ^*). *Suppose Assumption A and Assumptions 1.3, 1.2-a and A.1 are satisfied. Let F_n be piecewise linear function on [0,1] satisfying:*

1. $|F_n - F| \to 0$;
2. $\forall \mu \in [0,1]$, $\lim F_n'(\mu) = F'(\mu)$.

Let μ_n^* be as defined in Lemma A.2 associated with F_n. Suppose $\mu^* = \lim \mu_n^*$. Then,

1. $\forall \mu > \mu^*, \exists N$ *s.t.* $\forall n \geqslant N, \nu_n(\mu) \geqslant \mu$.
2. $\forall \mu < \mu^*, \exists N$ *s.t.* $\forall n \geqslant N, \nu_n(\mu) \leqslant \mu$.

Proof. $\forall \mu > \mu^*$, by definition $\lim \mu_n^* = \mu^*$, there exists N s.t. $\forall n \geqslant N$: $|\mu_n^* - \mu^*| < |\mu - \mu^*|$. Therefore $\mu > \mu_n^*$ and thus $\nu_n(\mu) \geqslant \mu$. Same argument applies to $\mu < \mu^*$. ■

Lemma B.27. *Suppose Assumption A and Assumptions 1.3, 1.2-a and A.1 are satisfied. Let F_n be piecewise linear function on [0,1] satisfying:*

1. $|F_n - F| \to 0$;
2. $\forall \mu \in [0,1]$, $\lim F_n'(\mu) = F'(\mu)$.

Define $V_n = \mathcal{V}(F_n)$ and $V = \mathcal{V}(F)$. Then: $\forall \mu \in [0,1]$ s.t. $V(\mu) > F(\mu)$, $\exists \lim V_n'(\mu)$.

Proof. With Lemma B.26, we can define $\mu^* \in [0,1]$ (we pick an arbitrary limiting point when there are multiple ones). First by assumption $\lim F_n'(\mu) = F'(\mu)$, and $V_n' = F_m'$ on the boundary by construction in Theorem 1.2, the statement is automatically true for $\mu \in \{0,1\}$. We discuss three possible cases for different $\mu \in (0,1)$ separately.

- *Case 1*: $\mu > \mu^*$. If $V(\mu) > F(\mu)$, then by convergence in L_∞ norm, there exists N and neighbourhood $\mu \in O$ s.t. $\forall n \geqslant N$, $\mu' \in O$, $V_n(\mu') > F_n(\mu')$. We know that by no-repeated-experimentation property of solution $\nu_n(\mu)$ to problem with F_n, $\nu_n(\mu) > \sup O$. Now consider $V_n'(\mu)$. Suppose $V_n'(\mu)$ have unlimited limiting point. Then exists subsequence $\lim V_n'(\mu) = \infty$ or $-\infty$. If $\lim V_n'(\mu) = \infty$, consider $\nu = 0$, else if $\lim V_n'(\mu) = -\infty$, consider $\nu = 1$:

$$V(\mu) = \lim_{n\to\infty} V_n(\mu)$$

$$\begin{aligned}
&\geqslant \lim_{n\to\infty} \frac{c}{\rho} \frac{F_n(\nu) - V_n(\mu) - V_n'(\mu)(\nu-\mu)}{J(\mu,\nu)} \\
&= \frac{c}{\rho} \frac{F(\nu) - V(\mu)}{J(\mu,\nu)} - \frac{c}{\rho} \lim_{n\to\infty} V_n'(\mu) \frac{\nu-\mu}{J(\mu,\nu)} \\
&= +\infty
\end{aligned}$$

Contradiction. Therefore we know that $V_n'(\mu)$ must have finite limiting points. Now suppose $V_n'(\mu)$ doesn't converge, then there exists two subsequences $\lim V_n'(\mu) = V_1'$ and $\lim V_m'(\mu) = V_2'$, $V_1' \neq V_2' \in \mathbb{R}$. Suppose $V_1' > V_2'$. Now take a converging subsequence of optimal policy at μ $\nu_{n_k} \to \nu^1$. By previous result $\nu^1 \geqslant \sup O$. Therefore ν^1 will be bounded away from μ. Consider:

$$\begin{aligned}
V(\mu) &= \lim_{k\to\infty} V_{n_k}(\mu) \\
&\geqslant \lim_{k\to\infty} \frac{c}{\rho} \frac{F_{m_k}(\nu^1) - V_{m_k}(\mu) - V_{m_k}'(\mu)(\nu^1-\mu)}{J(\mu,\nu^1)} \\
&= \frac{c}{\rho} \frac{F(\nu^1) - V(\mu) - V_2'(\nu^1-\mu)}{J(\mu,\nu^1)} \\
&= \lim_{k\to\infty} \frac{F_{n_k}(\nu_{n_k}) - V_{n_k}(\mu) - V_{n_K}'(\mu)(\nu_{n_k}-\mu)}{J(\mu,\nu_{n_k})} + \frac{(V_1'-V_2')(\nu^1-\mu)}{J(\mu,\nu^1)} \\
&> V(\mu)
\end{aligned}$$

Contradiction. Therefore, limit point of $V_n'(\mu)$ must be unique. Such limit point exists since V_n' are uniformly bounded. To sum up, there exists $\lim V_n'(\mu)$.

- *Case 2*: $\mu = \mu^*$. Since $V(\mu^*) > F(\mu^*)$. This implies that $\exists N$ s.t. $\forall n \geqslant N$, $V_n(\mu^*) > F_n(\mu^*)$. In this case, by Lemma A.2, μ_n^* are unique. Since μ_n^* is the unique intersection of U^{n+} and U^{n-} (Definition of U^{n+}, U^{n-1} are as in Lemma A.2, n is index), we can first establish convergence of μ^* through convergence of U^{n+} and U^{n-1}. By definition:

$$U^+(\mu) = \max_{\mu' \geqslant \mu, m \geqslant \underline{\mu}} \frac{F_m(\mu')}{1 + \frac{\rho}{c} J(\mu, \mu')}$$

Therefore, suppose the maximizer for index n is ν_n, m_n, then for index n':

$$\begin{aligned} U^{n'+}(\mu) \geqslant & \frac{F_{n'}(\nu_n)}{1 + \frac{\rho}{c} J(\mu, \nu_n)} \\ \geqslant & U^{n+}(\mu) + \frac{F_n(\nu_n) - F_{n'}(\nu_n)}{1 + \frac{\rho}{c} J(\mu, \nu_n)} \\ \geqslant & U^{n+}(\mu) - |F_n - F_{n'}| \end{aligned}$$

Since n and n' are totally symmetric, we actually showed that the functional map from F_n to U^{n+} is Lipschitz continuous in F_n with Lipschitz parameter 1. Symmetric argument shows that same property for U^{n-}. Since by assumption F_n is uniformly converging, we can conclude that U^{n+} and U^{n-} are Cauchy sequence with L_∞ norm. Therefore converging. Then $U^{n+} - U^{n-}$ uniformly converges and their roots will be UHC when $n \to \infty$. To show convergence of μ_n^*, it's sufficient to show that such limit is unique. This is not hard to see by applying envelope theory to U^{n+} and U^{n-}: $\frac{\mathrm{d}}{\mathrm{d}\mu} U^{n+}(\mu) = -\frac{\rho}{c} \frac{F(\nu_n) H''(\mu)(\nu_n - \mu)}{J(\mu, \nu_n)^2}$. Therefore $U^{n+} - U^{n-1}$ will have slope bounded below from zero, therefore the limit will also be strictly increasing. So μ^* is unique.

Since $\mu_n^* \to \mu$, and $V_n''(\mu)$ are all bounded from above:

$$\begin{aligned} V_n'(\mu^*) = & V_n'(\mu_n^*) + V_n''(\xi_n)(\mu^* - \mu_n^*) \\ = & V_n''(\xi_n)(\mu^* - \mu_n^*) \to 0 \end{aligned}$$

- *Case 3*: $\mu < \mu^*$. We can apply exactly the symmetric proof of case 1.

■

B.5.3 General state space

B.5.3.1 Proof of Theorem A.3

Proof. $\forall \mu \in E$, consider $X = \text{supp}(\mu)$ (This is without loss since we can always focus on only the support of μ). Let (p, ν, Σ) be optimal policy at μ.

Step 1. Derive optimality condition. Suppose $p \neq 0$:

$$\rho V(\mu) = -c\frac{V(\nu) - V(\mu) - \nabla V(\mu)(\nu - \mu)}{H(\nu) - H(\mu) - \nabla H(\mu)(\nu - \mu)} \tag{B.37}$$

Now let $p = -\frac{c}{H(\nu)-H(\mu)-\nabla H(\mu)(\nu-\mu)}$. As an analog to Equation (1.8), first order condition implies:

$$\begin{aligned}
&\text{FOC} - \nu : \nabla V(\nu) - \nabla V(\mu) + \lambda(\nabla H(\nu) - \nabla H(\mu)) = 0 \\
&\text{FOC} - p : V(\nu) - V(\mu) - \nabla V(\mu)(\nu - \mu) + \lambda((H(\nu) - H(\mu) - \nabla H(\mu)(\nu - \mu)) = 0 \\
&\xRightarrow{G=V+\lambda H} \begin{cases} \nabla G(\nu) = \nabla G(\mu) \\ G(\nu) - G(\mu) - \nabla G(\mu)(\nu - \mu) = 0 \end{cases}
\end{aligned} \tag{B.38}$$

Feasibility condition Equation (B.37) implies $\lambda = \frac{\rho}{c}V(\mu)$.Moreover, optimality implies $\forall \nu' \in \Delta(X)$:

$$\begin{aligned}
&\rho V(\mu) \geqslant -c\frac{V(\nu') - V(\mu) - \nabla V(\mu)(\nu' - \mu)}{H(\nu') - H(\mu) - \nabla H(\mu)(\nu' - \mu)} \\
\Longrightarrow\ & G(\nu') - G(\mu) - \nabla G(\mu)(\nu' - \mu) \leqslant 0
\end{aligned} \tag{B.39}$$

Suppose $p = 0$, then $\Sigma \neq 0$. Pick any non-zero row σ, then feasibility condition of

Equation (A.1) implies:

$$\rho V(\mu) = -c\frac{\sigma^T \mathrm{H}V(\mu)\sigma}{\sigma^T \mathrm{H}H(\mu)\sigma}$$

Optimality condition also implies Equation (B.39).

Step 2. Prove $V(\nu) > V(\mu)$. Suppose by contradiction that $V(\nu) \leqslant V(\mu)$. Consider $V(\mu_\alpha)$ where $\mu_\alpha = \alpha\nu + (1-\alpha)\mu$, $\alpha \in (0,1)$. Since $\Delta(X)$ is convex, $\mu_\alpha \in \Delta(X)$. Now by Equation (B.39), $G(\mu_\alpha) \leqslant G(\mu) + \nabla G(\mu)(\mu_\alpha - \mu)$. For α sufficiently small, $\mu_\alpha \in E$. $\forall \lambda' < \lambda$, let $G' = V + \lambda' H$. Then since H is strictly concave, G' is more convex that G, therefore

$$\begin{aligned}
&\begin{cases}
G'(\mu_\alpha) - G'(\mu) - \nabla G'(\mu)(\mu_\alpha - \mu) < 0 \\
G'(\mu_\alpha) - G'(\nu) - \nabla G'(\nu)(\mu_\alpha - \nu) < 0
\end{cases} \\
\Longrightarrow\ & G'(\mu_\alpha) + \nabla G'(\mu_\alpha)(\mu - \mu_\alpha) < G'(\mu) \\
\text{or}\ & G'(\mu_\alpha) + \nabla G'(\mu_\alpha)(\nu - \mu_\alpha) < G'(\nu)
\end{aligned}$$

So optimality condition is not satisfied at μ_α. Suppose $V(\mu_\alpha)$ is achieved with non-zero p_i, Then λ characterizing FOC at μ_α must be strictly larger than λ. Therefore $V(\mu_\alpha) > V(\mu)$. Suppose $V(\mu_\alpha)$ is achieved with zero p_i. Then $V(\mu_\alpha) \leqslant V(\mu)$ again implies Equation (B.39) violated. So $V(\mu_\alpha) > V(\mu)$. This implies

$$\begin{aligned}
&\frac{\mathrm{d}}{d_\alpha} V(\mu_\alpha) \geqslant 0 \\
\Longleftrightarrow\ & \nabla V(\mu)(\nu - \mu) \geqslant 0 \\
\Longrightarrow\ & V(\nu) - V(\mu) - \nabla V(\mu)(\nu - \mu) \leqslant 0
\end{aligned}$$

Contradicting Equation (B.37).

Step 3. Prove $V(\nu) = F(\nu)$. Suppose by contradiction that $V(\nu) > F(\nu)$. By the analysis in step 2, let $\lambda = \frac{\rho V(\mu)}{c}$ and $G = V + \lambda H$. Let $\lambda' = \frac{\rho V(\nu)}{c}$ and $G' = V + \lambda' H$. Then $\forall \nu' \in \Delta(X)$, $\nu' \neq \nu$:

$$\begin{aligned} G(\nu') \leqslant & G(\nu) + \nabla G(\nu)(\nu' - \nu) \\ \implies G'(\nu') = & G(\nu') + (\lambda' - \lambda)H(\nu') \\ \leqslant & G(\nu) + \nabla G(\nu)(\nu' - \nu) + (\lambda' - \lambda)H(\nu') \\ < & G(\nu) + \nabla G(\nu)(\nu' - \nu) + (\lambda' - \lambda)H(\nu) + \nabla H(\nu)(\nu' - \nu) \\ = & G'(\nu) + \nabla G'(\nu)(\nu' - \nu) \end{aligned}$$

On the other hand, $\forall \nu'$, $G(\nu') \leqslant G(\nu) + \nabla G(\nu)(\nu' - \nu)$ implies $\mathrm{H}G(\nu)$ being negative semi-definite. Then $\forall \sigma$, $\sigma^T \mathrm{H}G(\nu)\sigma \leqslant 0$. Therefore, $\forall \sigma$, $\sigma^T \mathrm{H}G(\nu)\sigma + (\lambda' - \lambda)\sigma^T \mathrm{H}H(\nu)\sigma < 0 \implies \frac{\rho}{c}V(\nu) < -\frac{\sigma^T \mathrm{H}V(\nu)\sigma}{\sigma^T \mathrm{H}H(\nu)\sigma}$. Contradicting $V(\nu)$ being sovled in Equation (A.1).

Step 4. Prove that the set of μ at which $\frac{\rho}{c}V(\mu) = -\frac{\sigma^T \mathrm{H}V(\mu)\sigma}{\sigma^T \mathrm{H}H(\mu)\sigma}$ is no where dense. Suppose by contradiction that there exists an open ball $O \subset E$ on which $\forall \mu$, $\frac{\rho}{c}V(\mu) = \max_\sigma -\frac{\sigma^T \mathrm{H}V(\mu)\sigma}{\sigma^T \mathrm{H}H(\mu)\sigma}$. Let $\underline{O}$ be a non-degenerate closed ball contained in O. Since V is continuous on V, there exists $\mu^* \in \arg\min_{\mu \in \underline{O}} V(\mu)$. $\forall \mu \in \underline{O}$, by definition $\mathrm{H}V(\mu) + \frac{\rho V(\mu)}{c}\mathrm{H}H(\mu)$ is negative semi-definite. Therefore, $\mathrm{H}V(\mu) + \frac{\rho V(\mu^*)}{c}\mathrm{H}H(\mu)$ is negative semi-definite. Now consider $G(\mu) = V(\mu) + \frac{\rho}{c}V(\mu^*)H(\mu)$ on $\underline{O}$. $G(\mu)$ has pointwise negative semi-definite Hessian. So $G(\mu)$ is a convex function. On the other hand, optimality of Gaussian signal at μ^* implies $G(\mu)$ to be concave. Therefore $G(\mu)$ is linear on $\underline{O}$. So $V(\mu) = L(\mu) - \frac{\rho}{c}V(\mu^*)H(\mu)$ on $\underline{O}$, where $L(\mu)$ is a linear function.

Now I show that $V(\mu)$ is a constant on $\underline{O}$. Suppose not, $V(\mu) > V(\mu^*)$. Then $V(\cdot) + \frac{\rho V(\mu)}{c}H(\cdot) = L(\cdot) + \frac{\rho}{c}(V(\mu) - V(\mu^*))H(\cdot)$ has negative-definit Hessian at μ. So there exists no σ s.t. $\sigma^T \mathrm{H}V(\mu)\sigma + \frac{\rho V(\mu)}{c}\sigma^T \mathrm{H}H(\mu)\sigma = 0$. Contradiction. However, $V(\mu)$ being a constant on $\underline{O}$ implies $\mathrm{H}V(\mu) \equiv 0$ on $\underline{O}$, contradiction.

Step 5. Prove that $\forall \mu \in E$, exists $\nu \in E^C$ satisfying Equation (B.37). Suppose $p > 0$, then as discussed in step 1, proof is done. Now suppose $p = 0$. Then by step 4, there is a converging sequence of $\mu_n \to \mu$ and ν_n satisfying Equation (B.37) for each μ_n. By step 3, $\nu_n \in E^C$ so ν_n are bounded away from μ_n by positive distance. Since $\nu_n \in E^C$ and E^C is closed subset of $\Delta(X)$, there exists converging subsequence $\nu_n \to \nu \in E^C$. Therefore, by smoothness of V and H,

$$\begin{aligned}
V(\mu) = \lim_{n\to\infty} V(\mu_n) = \lim_{n\to\infty} -c\frac{V(\nu_n) - V(\mu_n) - \nabla V(\mu_n)(\nu_n - \mu_n)}{H(\nu_n) - H(\mu_n) - \nabla H(\mu_n)(\nu_n - \mu_n)} \\
= -c\frac{V(\nu) - V(\mu) - \nabla V(\mu)(\nu - \mu)}{H(\nu) - H(\mu) - \nabla H(\mu)(\nu - \mu)}
\end{aligned}$$

Step 6. Prove the strict inequality. Define $K = \left\{\mu \middle| \rho V(\mu) = \sup_\sigma -c\frac{\sigma^T \mathrm{H}V(\mu)\sigma}{\sigma^T \mathrm{H}H(\mu)\sigma}\right\}$. Then by *step 4*, K is a nowhere dense set and the inequality in property 4 is satisfied by construction. Now I prove property 1 on $E\backslash K$:

$$\begin{aligned}
D_{\nu(\mu)-\mu}V(\mu) =& (\nu(\mu) - \mu)^T \cdot \nabla V(\mu) = (\nu(\mu) - \mu)^T \cdot \frac{\partial}{\partial \mu}\left(-\frac{c}{\rho}\frac{F(\nu(\mu)) - V(\mu) - \nabla V(\mu)(\nu(\mu) - \mu)}{H(\nu(\mu)) - H(\mu) - \nabla H(\mu)(\nu(\mu) - \mu)}\right) \\
=& (\nu(\mu) - \mu)^T\left(-\frac{c}{\rho}\frac{-\mathrm{H}V(\mu)(\nu(\mu) - \mu) + \frac{\rho}{c}V(\mu)(-\mathrm{H}H(\mu)(\nu(\mu) - \mu))}{H(\nu(\mu)) - H(\mu) - \nabla H(\mu)(\nu(\mu) - \mu)}\right) > 0
\end{aligned}$$

Now I prove property 3 on $E\backslash K$: Define $J(\mu, \nu) = H(\mu) - H(\nu) + \nabla H(\mu)(\nu - \mu)$. Then Equations (B.37) and (B.38) implies

$$\begin{aligned}
&\begin{cases}
V(\mu) = \dfrac{F(\nu) - (\nu - \mu)^T \cdot \nabla V(\mu)}{1 + \frac{\rho}{c}J(\mu, \nu)} \\
\nabla V(\mu) = \left((\nabla H(\nu) - \nabla H(\mu))(\nu - \mu)^T + (1 + \dfrac{\rho}{c}J(\mu, \nu))I\right)^{-1} \\
\qquad\qquad \cdot \left(F(\nu)(\nabla H(\nu) - \nabla H(\mu)) + (1 + \dfrac{\rho}{c}J(\mu, \nu))\nabla F\right)
\end{cases} \\
\Longrightarrow &\begin{cases}
V(\mu) = \dfrac{F(\mu)}{1 - \frac{\rho}{c}J(\nu, \mu)} \\
\nabla V(\mu) = \nabla F + \dfrac{\frac{\rho}{c}F(\mu)(\nabla H(\nu) - \nabla H(\mu))}{1 - \frac{\rho}{c}J(\nu, \mu)}
\end{cases}
\end{aligned}$$

Then $\nu = \nu(\mu)$ satisfies the following PDE $\forall \alpha$:

$$\begin{aligned}
&\alpha^T \cdot \frac{\partial}{\partial \mu} \frac{F(\mu)}{1 - \frac{\rho}{c} J(\nu, \mu)} + D_\alpha \nu \frac{\partial}{\partial \nu} \frac{F(\mu)}{1 - \frac{\rho}{c} J(\nu, \mu)} = \alpha^T \cdot \left(\nabla F + \frac{\frac{\rho}{c} F(\mu)(\nabla H(\nu) - \nabla H(\mu))}{1 - \frac{\rho}{c} J(\nu, \mu)} \right) \\
\implies & F(\mu) D_\alpha \nu \cdot \mathrm{H}H(\nu)(\nu - \mu) = J(\nu, \mu) \Big(\alpha^T \nabla F \Big(1 - \frac{\rho}{c} J(\nu, \mu) \Big) + \frac{\rho}{c} F(\mu) \alpha^T (\nabla H(\nu) - \nabla H(\mu)) \Big) \\
\implies & D_\alpha \nu \cdot \mathrm{H}H(\nu)(\nu - \mu) = \frac{J(\nu, \mu)}{F(\mu)\big(1 - \frac{\rho}{c} J(\nu, \mu)\big)} D_\alpha V(\mu) \\
\implies & D_{\mu - \nu} \nu \cdot \mathrm{H}H(\nu)(\nu - \mu) = \frac{J(\nu, \mu)}{F(\mu)\big(1 - \frac{\rho}{c} J(\nu, \mu)\big)} \big(-D_{\nu - \mu} V(\mu) \big) < 0
\end{aligned}$$

The inequality comes from $V(\mu) > 0$ and $D_{\nu-\mu} V(\mu) > 0$. ■

B.5.4 *Axiom for posterior separability*

B.5.4.1 *Proof of Theorem A.4*

Proof. Let $\mathcal{S}_0$ be a fully revealing information structure i.e. with any prior belief μ, each signal induces posterior belief δ_{x_i} with probability $\mu(x_i)$. $\forall \mu \in \Delta X$, define:

$$H(\mu) = I(\mathcal{S}_0; \mathcal{X} | \mu)$$

$\forall \mathcal{S}$ which induces ν with probability $h(\nu)$ with prior μ:

$$\begin{aligned}
I(\mathcal{S}_0; \mathcal{X} | \mu) =& I(\mathcal{S}; \mathcal{X} | \mu) + E[I(\mathcal{S}_0; \mathcal{X} | \mathcal{S}, \mu)] \\
=& I(\mathcal{S}; \mathcal{X} | \mu) + \int I(\mathcal{S}_0; \mathcal{X} | \nu) h(\nu) \mathrm{d}\nu \\
=& H(\mathcal{S}; \mathcal{X} | \mu) + E_h[I(\mathcal{S}_0; \mathcal{X} | \nu)] \\
\implies \quad I(\mathcal{S}; \mathcal{X} | \mu) =& H(\mu) - E_h[H(\nu)]
\end{aligned}$$

Moreover, $H(E_h[\nu]) - E_h[H(\nu)] \geqslant 0$ for all distribution h implies that H is a concave function on ΔX. ■

Appendix C

Appendix for Chapter 2

C.1 Omitted proofs

C.1.1 Proof of Lemma 2.1

Proof.

Step 1. Value from solving Equation (2.1) is no larger than value from solving Equation (C.1):

$$\sup_{\mathcal{S}_t,\mathcal{T}} E[\rho_{\mathcal{T}} u(\mathcal{A},\mathcal{X})] \tag{C.1}$$
$$\text{s.t.} \begin{cases} E\big[I(\mathcal{S}_t;\mathcal{X}|\mathcal{S}_{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\geqslant t})\big|\mathcal{T}\geqslant t\big] \leqslant c \\ \mathcal{X}\to\mathcal{S}_t\to\mathcal{A} \text{ conditional on } \mathcal{T}=t \\ \mathcal{X}\to\mathcal{S}_t\to\mathbf{1}_{\mathcal{T}\geqslant t} \end{cases}$$

Equation (C.1) is more relaxed than Equation (2.1) in the first constraint. In Equation (2.1), the flow cost constraint is imposed on each prior induced by previous information and decision choice. Equation (C.1) only requires the average cost conditional on not having stopped yet being bounded by c:

$$\begin{aligned} & I(\mathcal{S}_t;\mathcal{X}|\mathcal{S}_{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\geqslant t}) \leqslant c \\ \Longrightarrow\ & E\big[I(\mathcal{S}_t;\mathcal{X}|\mathcal{S}_{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\geqslant t})\big|\mathcal{T}\geqslant t\big] \leqslant E[c|\mathcal{T}\geqslant t] = c \end{aligned}$$

Therefore, any feasible strategy for Equation (2.1) is feasible for Equation (C.1). So Equation (C.1) is a more relaxed problem than Equation (2.1).

Step 2. Value from solving Equation (C.1) is no larger than value from solving Equation (2.2). $\forall\ (\mathcal{S}_t, \mathcal{T})$ satisfying constraints in Equation (C.1), define:

$$\begin{cases} I_t = I(\mathcal{S}_{\mathbf{t-1}}; \mathcal{X} | \mathcal{T} \geqslant t) \\ p_t = \mathrm{P}(\mathcal{T} = t | \mathcal{T} \geqslant t) \\ P_t = \mathrm{P}(\mathcal{T} \leqslant t) \end{cases}$$

Want to show that (I_t, p_t) is feasible and implements same utility in Equation (2.2) as $(\mathcal{S}_t, \mathcal{T})$ in Equation (C.1). First, consider the objective function:

$$\begin{aligned} & E[\rho_{\mathcal{T}} u(\mathcal{A}; \mathcal{X})] \\ =& \sum_{t=0}^{\infty} \mathrm{P}(\mathcal{T} = t) \rho_t E[u(\mathcal{A}; \mathcal{X}) | \mathcal{T} = t] \\ =& \sum_{t=0}^{\infty} \mathrm{P}(\mathcal{T} = t | \mathcal{T} \geqslant t) \mathrm{P}(\mathcal{T} \geqslant t) \rho_t V^* \\ =& \sum_{t=0}^{\infty} \rho_t (1 - P_{t-1}) p_t V^* \end{aligned}$$

Second, consider feasibility constraint:

$$\begin{aligned} c \geqslant & E\big[I(\mathcal{S}_t; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) | \mathcal{T} \geqslant t\big] \\ =& \mathrm{P}(\mathcal{T} = t | \mathcal{T} \geqslant t) E[I(\mathcal{S}_t; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) | \mathcal{T} = t] \\ &+ \mathrm{P}(\mathcal{T} > t | \mathcal{T} \geqslant t) E[I(\mathcal{S}_t; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) | \mathcal{T} > t)] \\ =& p_t (I(\mathcal{S}_{\mathbf{t}}, \mathbf{1}_{\mathcal{T} \geqslant t}; \mathcal{X} | \mathcal{T} = t) - I(\mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}; \mathcal{X} | \mathcal{T} = t)) \\ &+ (1 - p_t)(I(\mathcal{S}_{\mathbf{t}}, \mathbf{1}_{\mathcal{T} \geqslant t}; \mathcal{X} | \mathcal{T} > t) - I(\mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}; \mathcal{X} | \mathcal{T} > t)) \\ =& p_t I(\mathcal{S}_{\mathbf{t}}; \mathcal{X} | \mathcal{T} = t) + (1 - p_t) I(\mathcal{S}_{\mathbf{t}}; \mathcal{X} | \mathcal{T} > t) \\ &- (p_t I(\mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}; \mathcal{X} | \mathcal{T} = t) + (1 - p_t) I(\mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}; \mathcal{X} | \mathcal{T} > t)) \\ \geqslant & p_t I(\mathcal{A}; \mathcal{X}) + (1 - p_t) I_{t+1} - I(\mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}; \mathcal{X} | \mathcal{T} \geqslant t) \end{aligned}$$

$$=p_t\bar{I} + (1-p_t)I_{t+1} - I_t$$

First inequality is feasibility constraint. Fist equality is law of iterated expectation. Second equality is chain rule from posterior separability. Third equality is rewriting terms. Noticing that condition on $\mathcal{T} = t+1$, $\mathbf{1}_{\mathcal{T}\leqslant t}$ is degenerate. Second inequality is from information processing inequality and applying chain rule again. Last equality is by definition. It is easy to verify by law of total probability that:

$$\begin{aligned}
P_t =& \mathrm{P}(\mathcal{T} \leqslant t) = \mathrm{P}(\mathcal{T} = t) + \mathrm{P}(\mathcal{T} \leqslant t-1) \\
=& \mathrm{P}(\mathcal{T} \geqslant t)\mathrm{P}(\mathcal{T} = t|\mathcal{T} \geqslant t) + \mathrm{P}(\mathcal{T} \leqslant t-1) \\
=& (1-P_{t-1})p_t + P_{t-1}
\end{aligned}$$

Then we verify initial conditions:

$$\begin{cases} I_1 = I(\mathcal{S}_0; \mathcal{X}|\mathcal{T} \geqslant 1) = 0 \\ P_0 = \mathrm{P}(\mathcal{T} \leqslant 0) = 0 \end{cases}$$

■

C.1.2 Proof of Theorem 2.1

Proof. First, assume $\rho_t = \max\{0, \frac{T-t}{T}\}$. We show that the statement in Theorem 2.1 is correct with the assumed ρ_t. Since $\rho_t = 0$ when $t \geqslant T$, Equation (2.2) is finite horizon. So we can apply backward induction. Define:

$$V_t(I) = \sup_{p_\tau} \sum_{\tau=t}^{T} \frac{T-\tau}{T}(1-P_{\tau-1})p_\tau V^*$$

$$\text{s.t.} \begin{cases} (\bar{I} - I_\tau)p_\tau + (I_{\tau+1} - I_\tau)(1 - p_\tau) \leqslant c \\ P_\tau = P_{\tau-1} + (1 - P_{\tau-1})p_\tau \\ P_{t-1} = 0, I_t = I \end{cases}$$

Then V_t solves functional equation:

$$V_t(I) = \sup_p \frac{T-t}{T} pV^* + (1-p)V_{t+1}(I') \tag{C.2}$$
$$\text{s.t. } (\bar{I} - I)p + (I' - I)(1-p) \leqslant c$$

I conjecture that for $I \geqslant 0$:

$$\widetilde{V}_t(I) = \begin{cases} \frac{T-t}{T}\frac{c+I}{\bar{I}}V^* + \left(1 - \frac{c+I}{\bar{I}}\right)\sum_{\tau=t+1}^{T} \frac{T-\tau}{T}V^*\frac{c}{\bar{I}}\left(1 - \frac{c}{\bar{I}}\right)^{\tau-t-1} & \text{when } \frac{c+I}{\bar{I}} < 1 \\ \frac{T-t}{T}V^* & \text{when } \frac{c+I}{\bar{I}} \geqslant 1 \end{cases} \tag{C.3}$$

solves Equation (C.2). This is clearly true for $t = T-1$. Since when $t = T-1$, $V_{t+1} \equiv 0$ so there is no utility gain from accumulating I. Now we prove the conjecture by backward induction on t. Suppose the conjecture is true for t. Consider solving V_{t-1} from Equation (C.2).

- *Case 1*: $\bar{I} \leqslant c + I$. Then choosing $p = 1$ gives utility $\frac{T-\tau}{T}V^*$ immediately, thus optimal and $V_t(I) = \frac{T-t}{T}V^* = \widetilde{V}_t(I)$.

- *Case 1*: $\bar{I} > c + I$. Consider the one-step optimization problem choosing I':

$$V_t(I) = \sup_{I' \geqslant 0} \frac{T-t}{T}\frac{c+I-I'}{\bar{I}-I'}V^* + \frac{\bar{I}-I-c}{\bar{I}-I'}\widetilde{V}_{t+1}(I')$$

When $I' \leqslant \bar{I} - c$, the objective function is:

$$\frac{T-t}{T}\frac{c+I-I'}{\bar{I}-I'}V^* + \frac{\bar{I}-I-c}{\bar{I}-I'}\frac{T-t-1}{T}V^*$$
$$\Longrightarrow \text{FOC}: \ -\frac{1}{T}\frac{\bar{I}-I-c}{(\bar{I}-I')^2}V^* \leqslant 0$$

When $I' > \bar{I} - c$, the objective function is:

$$\frac{T-t}{T}\frac{c+I-I'}{\bar{I}-I'}V^* + \frac{\bar{I}-I-c}{\bar{I}-I'}\left(\frac{T-t-1}{T}\frac{c+I'}{\bar{I}}V^* + \left(1-\frac{c+I'}{\bar{I}}\right)\sum_{\tau=t+2}^{T}\frac{T-\tau}{T}V^*\frac{c}{\bar{I}}\left(1-\frac{c}{\bar{I}}\right)^{\tau-t-2}\right)$$
$$\Longrightarrow \text{FOC}: \ -\left(1-\frac{c}{\bar{I}}\right)^{T-t-1}\frac{\bar{I}-I-c}{T(\bar{I}-I')^2}V^* < 0$$

To sum up, decreasing I' is always utility improving. So optimal $I' = 0$ and optimal solution of Equation (C.2) is

$$\begin{aligned}
V_t(I) =& \frac{T-t}{T}\frac{c+I}{\bar{I}}V^* + \left(1-\frac{c+I}{\bar{I}}\right)\left(\frac{T-t-1}{T}\frac{c}{\bar{I}}V^* + \left(1-\frac{c}{\bar{I}}\right)\sum_{\tau=t+2}^{T}\frac{T-\tau}{T}V^*\frac{c}{\bar{I}}\left(1-\frac{c}{\bar{I}}\right)^{\tau-t-2}\right)\\
=& \frac{T-t}{T}\frac{c+I}{\bar{I}}V^* + \left(1-\frac{c+I}{\bar{I}}\right)\sum_{\tau=t+1}^{T}\frac{T-\tau}{T}V^*\frac{c}{\bar{I}}\left(1-\frac{c}{\bar{I}}\right)^{\tau-t-1}\\
=& \widetilde{V}_t(I)
\end{aligned}$$

Therefore, $\widetilde{V}_t(I)$ solves Equation (C.2). So with ρ_t defined by $\max\{0, \frac{T-t}{T}\}$, Equation (2.2) is solved by strategy $I_t \equiv 0$ (i.e. $p_t = \frac{c}{\bar{I}}$) and optimal utility is $\widetilde{V}_1(0)$.

Now, consider a general convex ρ_t. We want to show that $p_t = \frac{c}{\bar{I}}$ is still optimal strategy for Equation (2.2). By definition $\lim_{t\to\infty}\sum_{\tau\geqslant t}\rho_\tau = 0$, so $\forall\varepsilon$ there exists T s.t $\sum_{t\geqslant T}\rho_t < \varepsilon$. Pick T to be an even number. Now define ρ^t_τ recursively:

- $\rho^T_\tau = \max\{\rho_T + (\tau - T)(\rho_T - \rho_{T-1}), 0\}$. Define $\hat{\rho}^T_\tau = \rho_\tau - \rho^T_\tau$ when $\tau \leqslant T$ and $\hat{\rho}^T_\tau = 0$ otherwise. It is not hard to verify that $\hat{\rho}^T_\tau$ is convex in τ and $\hat{\rho}^T_\tau = 0\ \forall\tau \geqslant T-1$.
- $\rho^{T-2}_\tau = \max\{\hat{\rho}^T_{T-2} + (\tau - T + 2)(\hat{\rho}^T_{T-2} - \hat{\rho}^T_{T-3}), 0\}$. Define $\hat{\rho}^{T-2}_\tau = \hat{\rho}^T_\tau - \rho^{T-2}_\tau$. It is not

hard to verify that $\widehat{\rho}_\tau^{T-2}$ is convex in τ and $\widehat{\rho}_\tau^{T-2} = 0\ \forall \tau \geqslant T-3$.

- $\cdots$

- $\rho_\tau^{T-2k} = \max\left\{\widehat{\rho}_{T-2k}^{T-2k+2} + (\tau - T + 2k)\left(\widehat{\rho}_{T-2k}^{T-2k+2} - \widehat{\rho}_{T-2k-1}^{T-2k+2}\right), 0\right\}$. Define $\widehat{\rho}_\tau^{T-2k} = \widehat{\rho}_\tau^{T-2k+2} - \rho_\tau^{T-2k}$.

$\forall p'_t$ satisfying constraints in Equation (2.2) and corresponding P'_t:

$$\begin{aligned}
&\sum_{t=1}^{\infty} \rho_t (1 - P'_{t-1}) p'_t V^* \\
&= \sum_{t=1}^{\infty} \left(\sum_{k=1}^{[T/2]} \rho_t^{T-2k} + \left(\rho_t - \sum_{k=1}^{[T/2]} \rho_t^{T-2k} \right) \right) (1 - P'_{t-1}) p'_t V^* \\
&< \sum_{k=1}^{[T/2]} \sum_{t=1}^{\infty} \rho_t^{T-2k} (1 - P'_{t-1}) p'_t V^* + \varepsilon V^* \\
&\leqslant \sum_{k=1}^{[T/2]} \sum_{t=1}^{\infty} \rho_t^{T-2k} (1 - P_{t-1}) p_t V^* + \varepsilon V^* \\
&\leqslant \sum_{t=1}^{\infty} \rho_t (1 - P_{t-1}) p_t V^* + \varepsilon V^*
\end{aligned}$$

First inequality is from $\sum_{t \geqslant T} \rho_t < \varepsilon$. Second inequality is from optimality of p_t in last part. Last inequality is from $\sum_{t \geqslant T} \rho_t \geqslant 0$. Therefore, by taking $\varepsilon \to 0$, we showed that:

$$\sum_{t=1}^{\infty} \rho_t (1 - P'_{t-1}) p'_t V^* \leqslant \sum_{t=1}^{\infty} \rho_t (1 - P_{t-1}) p_t V^*$$

■

C.1.3 Proof of Lemma 2.2

Proof. First of all, redefine $\widetilde{\mathcal{S}}_t$ s.t.

$$\widetilde{\mathcal{S}}_t = \begin{cases} \mathcal{S}_t & \text{conditional on } \mathcal{T} \geqslant t \\ s_0 & \text{conditional on } \mathcal{T} < t \end{cases}$$

where equality is defined as signal distribution conditional on $\mathcal{X}$ and $\mathcal{T}$ being identical. It is not hard to verify that $\widetilde{\mathcal{S}}_t, \mathcal{T}$ still satisfies constraints in Equation (2.1):

- If $\mathcal{T} < t$, $I(\widetilde{\mathcal{S}}_t; \mathcal{X} | \widetilde{\mathcal{S}}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) = 0$ since $\widetilde{\mathcal{S}}_t$ is degenerate. If $\mathcal{T} \geqslant t$, then $\mathcal{T} \geqslant t-1$ so $I(\widetilde{\mathcal{S}}_t; \mathcal{X} | \widetilde{\mathcal{S}}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) = I(\mathcal{S}_t; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) \leqslant c$.
- Conditional on $\mathcal{T} = t$, $\widetilde{\mathcal{S}}_t = \mathcal{S}_t$ so $\mathcal{X} \to \widetilde{\mathcal{S}}_t \to \mathcal{A}$.
- If $\widetilde{\mathcal{S}}_t = s_0$, then $\mathcal{T} < t$ for sure, so $\mathbf{1}_{\mathcal{T} \geqslant t}$ is independent to $\mathcal{X}$. If $\widetilde{\mathcal{S}}_t \neq s_0$, then $\mathcal{T} \geqslant t$ for sure, so $\mathbf{1}_{\mathcal{T} \geqslant t}$ is independent to $\mathcal{X}$.

So replacing $\mathcal{S}$ with $\widetilde{\mathcal{S}}$ we still get a feasible strategy and induced decision time distribution $\mathcal{T}$ is unchanged. From now on, we assume WLOG that $\mathcal{S}_t \equiv s_0$ when $\mathcal{T} < t$. I only discuss the case $E[\mathcal{T}] < \infty$. If $E[\mathcal{T}] = \infty$ then Lemma 2.2 is automatically true.

$$\begin{aligned}
E[\mathcal{T}] =& \sum_{t=1}^{\infty} \mathrm{P}(\mathcal{T} = t) \cdot t = \sum_{t=1}^{\infty} \mathrm{P}(\mathcal{T} = t) \sum_{\tau=1}^{t} 1 = \sum_{\tau=1}^{\infty} \sum_{t=\tau}^{\infty} \mathrm{P}(\mathcal{T} = t) \\
=& \sum_{t=1}^{\infty} \mathrm{P}(\mathcal{T} \geqslant t) = \frac{1}{c} \sum_{t=0}^{\infty} \mathrm{P}(\mathcal{T} \geqslant t) \cdot c \\
\geqslant& \frac{1}{c} \sum_{t=1}^{\infty} \mathrm{P}(\mathcal{T} \geqslant t) E\big[I(\mathcal{S}_t; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) \big| \mathcal{T} \geqslant t\big] \\
=& \frac{1}{c} \sum_{t=1}^{\infty} \Big(\mathrm{P}(\mathcal{T} \geqslant t) E\big[I(\mathcal{S}_t; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) \big| \mathcal{T} \geqslant t\big] + \mathrm{P}(\mathcal{T} < t) E\big[I(\mathcal{S}_t; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t}) \big| \mathcal{T} < t\big]\Big) \\
=& \frac{1}{c} \sum_{t=1}^{\infty} E[I(\mathcal{S}_t; \mathcal{X} | \mathcal{S}_{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \geqslant t})] = \frac{1}{c} \sum_{t=1}^{\infty} (I(\mathcal{S}_{\mathbf{t}}; \mathcal{X}) - I(\mathcal{S}_{\mathbf{t-1}}; \mathcal{X}))
\end{aligned}$$

$$
\begin{aligned}
&=\frac{1}{c}\left(\lim_{t\to\infty} I(\mathcal{S}_{\mathbf{t}};\mathcal{X})+\lim_{\tau\to\infty}\sum_{t=\tau}^{\infty} I(\mathcal{S}_{\mathbf{t}};\mathcal{X})-I(\mathcal{S}_{\mathbf{t-1}};\mathcal{X})\right)\\
&\geqslant\frac{1}{c}\lim_{\tau\to\infty}\sum_{t=1}^{\tau}\mathrm{P}(\mathcal{T}=t)E\big[I(\mathcal{S}_{\mathbf{t}};\mathcal{X}|\mathcal{T}=t)\big|\mathcal{T}=t\big]\\
&\geqslant\frac{1}{c}\lim_{\tau\to\infty}\sum_{t=1}^{\tau}\mathrm{P}(\mathcal{T}=t)I(\mathcal{A};\mathcal{X})\\
&=\frac{I(\mathcal{A};\mathcal{X})}{c}
\end{aligned}
$$

Third line is from flow informativeness constraint. Forth line is from $\mathcal{S}_t\big|_{\mathcal{T}<t}\equiv s_0$. Fifth and sixth line is from chain rule of posterior separable information measure. Seventh line is from information process inequality and law of interated expectation. Second last line is from information processing constraint. ■

C.1.4 Proof of Lemma 2.3

Proof. Take any strategy $(\mu_t,\mathcal{T})$ feasible in Equation (2.6). Define

$$
\begin{cases}
P_t=\mathrm{P}(\mathcal{T}\leqslant t)\\
I_t=E\big[H(\mu)-H(\mu_t)\big|\mathcal{T}>t\big]
\end{cases}
\tag{C.4}
$$

Now we prove that Equation (C.4) is a feasible strategy in Equation (2.7) and implements same value. First, since H is concave, then $I_t\geqslant 0$. Since $\mu_0=\mu$, $I_0=0$. Since $\mu_t|_{\mathcal{T}=t}=\pi$ and $\mu_0=\mu$, then $P_0=0$. Now we verify $\dot{I}_t\leqslant c-p_t(\bar{I}-I_t)$

$$
\begin{aligned}
E[H(\mu_{t+dt})|\mathcal{T}>t]=&-H(\mu)-E[H(\mu_{t+dt})|\mathcal{T}>t+dt]\mathrm{P}(\mathcal{T}>t+dt|\mathcal{T}>t)\\
&-E[H(\mu_{t+dt})|\mathcal{T}\in(t,t+dt]]\mathrm{P}(\mathcal{T}\in(t+dt]|\mathcal{T}>t)\\
\implies\quad I_{t+dt}=&H(\mu)-\frac{1-P_t}{1-P_{t+dt}}\left(E[H(\mu_{t+dt})|\mathcal{T}>t]-\frac{P_{t+dt}-P_t}{1-P_t}E[H(\mu_{t+dt})|\mathcal{T}\in(t,t+dt]]\right)\\
\implies\quad I_{t+dt}-I_t=&E[H(\mu_t)|\mathcal{T}>t]
\end{aligned}
$$

$$
\begin{aligned}
&-\frac{1-P_t}{1-P_{t+dt}}\left(E[H(\mu_{t+dt})|\mathcal{T}>t]-\frac{P_{t+dt}-P_t}{1-P_t}E[H(\mu_{t+dt})|\mathcal{T}\in(t,t+dt]]\right)\\
=&\frac{1-P_{t+dt}}{1-P_t}(E[H(\mu_t)-H(\mu_{t+dt})|\mathcal{T}>t])\\
-&\frac{1-P_{t+dt}}{1-P_t}\frac{P_{t+dt}-P_t}{1-P_t}((E[H(\mu)-H(\mu_{t+dt})|\mathcal{T}\in(t,t+dt]])-E[H(\mu)-H(\mu_t)|\mathcal{T}>t])\\
-&\left(\frac{P_{t+dt}-P_t}{1-P_t}\right)^2E[H(\mu_t)|\mathcal{T}>t]\\
\implies \mathrm{d}I_t=&E[-\mathrm{d}H(\mu_t)|\mathcal{T}>t]-\frac{\mathrm{d}P_t}{1-P_t}(E_\pi[H(\mu)-H(\nu)]-I_t)\\
\implies \dot{I}_t\leqslant& c-\frac{\dot{P}_t}{1-P_t}(\bar{I}-I_t)
\end{aligned}
$$

First equality is law of iterated expectation. Second, third and forth equalities are rearranging terms. Fifth equality is from taking $dt\to 0$. Inequality is from $E[\mathrm{d}H(\mu_t)|\mu_t]\leqslant d\mathrm{d}t$.

Finally, deifne $p_t=\frac{\dot{P}_t}{1-P_t}$. Then

$$
E[\rho_{\mathcal{T}}]=\int_0^\infty \rho_t\mathrm{d}P_t=\int_0^\infty \rho_t(1-P_t)p_t\mathrm{d}t
$$

To sum up, for any feasible strategy in Equation (2.6), there exists an feasible strategy in Equation (2.7) attaining same value. So the statement in Lemma 2.3 is true. ■

C.1.5 proof of Theorem 2.3

Proof. It is easy to verify that $p_t\equiv\frac{c}{\bar{I}}$ is feasible in Equation (2.7) and the objective function is exactly $\int_0^\infty \rho_t e^{-\frac{c}{\bar{I}}t}\frac{c}{\bar{I}}\mathrm{d}t$. Therefore, it is sufficient to show that $V\leqslant\int_0^\infty \rho_t e^{-\frac{c}{\bar{I}}t}\frac{c}{\bar{I}}\mathrm{d}t$. Pick any p_t satisfying constraints in Equation (2.7). Now since p_t and ρ_t are integrable, $\forall\varepsilon>0$, there exists T s.t.

$$
\int_0^\infty \rho_t(1-P_t)p_t\mathrm{d}t\leqslant\int_0^T \rho_t(1-P_t)p_t\mathrm{d}t+\varepsilon
$$

Then there exists $dt > 0$ s.t.:

$$\begin{aligned}
\int_0^T \rho_t(1-P_t)p_t\mathrm{d}t &\leqslant \sum_{k=1}^{[T/dt]} \rho_{kdt} \int_{kdt}^{(k+1)dt} (1-P_\tau)p_\tau \mathrm{d}\tau + \varepsilon \\
&= \sum_{k=1}^{[T/dt]} \rho_{kdt} \int_{kdt}^{(k+1)dt} e^{-\int_0^\tau p_s \mathrm{d}s} p_\tau \mathrm{d}\tau + \varepsilon \\
&= \sum_{k=1}^{[T/dt]} \rho_{kdt} \left(e^{-\int_0^{kdt} p_\tau \mathrm{d}\tau} - e^{-\int_0^{(k+1)dt} p_\tau \mathrm{d}\tau} \right) + \varepsilon \\
&= \sum_{k=1}^{[T/dt]} \rho_{kdt} e^{-\int_0^{kdt} p_\tau \mathrm{d}\tau} \left(1 - e^{-\int_{kdt}^{(k+1)dt} p_\tau \mathrm{d}\tau} \right) + \varepsilon \\
&= \sum_{k=1}^{[T/dt]} \rho_{kdt} P_{kdt} \left(1 - e^{-\int_{kdt}^{(k+1)dt} p_\tau \mathrm{d}\tau} \right) + \varepsilon
\end{aligned}$$

Now consider the following sequence:

$$\begin{cases}
\widehat{\rho}_k = \rho_{k \cdot dt} \\
\widehat{p}_k = 1 - e^{-\int_{kdt}^{(k+1)dt} p_\tau \mathrm{d}\tau} \\
\widehat{P}_{k-1} = P_{kdt} \\
\widehat{I}_k = I_{kdt} \\
\widehat{c} = cdt
\end{cases}$$

We verify that:

$$\begin{cases}
(\bar{I} - \widehat{I}_k)\widehat{p}_k + (\widehat{I}_{k+1} - \widehat{I}_k)(1-\widehat{p}_k) \leqslant \widehat{c} \\
\widehat{P}_k = \widehat{P}_{k-1} + (1-\widehat{P}_{k-1})\widehat{p}_k
\end{cases}$$

- Solve ODE defining P_t, we get $P_t = 1 - e^{-\int_0^t p_\tau \mathrm{d}\tau}$. Apply this to calculate $\widehat{P}_k - \widehat{P}_{k-1} = P_{(k+1)dt} - P_{kdt} = (1-P_{kdt})\left(1 - e^{-\int_{kdt}^{(k+1)dt} p_\tau \mathrm{d}\tau}\right) = (1-\widehat{P}_{k-1})\widehat{p}_k$.

- Solve ODE defining I_t, we get:

$$
\begin{aligned}
I_t =& \int_0^t e^{\int_\tau^t p_s \mathrm{d}s}(c - \bar{I}p_\tau)\mathrm{d}\tau \\
\implies \quad I_{(k+1)dt} - I_{kdt} =& \int_0^{(k+1)dt} e^{\int_\tau^{(k+1)dt} p_s \mathrm{d}s}(c - \bar{I}p_\tau)\mathrm{d}\tau - \int_0^{kdt} e^{\int_\tau^{kdt} p_s \mathrm{d}s}(c - \bar{I}p_\tau)\mathrm{d}\tau \\
=& \int_{kdt}^{(k+1)dt} e^{\int_\tau^{kdt} p_s \mathrm{d}s}(c - \bar{I}p_\tau)\mathrm{d}\tau + \left(e^{\int_{kdt}^{(k+1)dt} p_s \mathrm{d}s} - 1\right) \int_0^{(k+1)dt} e^{\int_\tau^t p_s \mathrm{d}s}(c - \bar{I}p_\tau)\mathrm{d}\tau \\
=& \left(e^{\int_{kdt}^{(k+1)dt} p_s \mathrm{d}s} - 1\right) I_{kdt} + e^{\int_{kdt}^{(k+1)dt} p_s \mathrm{d}s} \int_{kdt}^{(k+1)dt} e^{\int_\tau^{kdt} p_s \mathrm{d}s}(c - \bar{I}p_\tau)\mathrm{d}\tau \\
\implies \quad \Delta\widehat{I}_k(1 - \widehat{p}_k) =& \left(1 - e^{-\int_{kdt}^{(k+1)dt} p_s \mathrm{d}s}\right)\widehat{I}_k + c\int_{kdt}^{(k+1)dt} e^{\int_\tau^{kdt} p_s \mathrm{d}s}\mathrm{d}\tau - \bar{I}\int_{kdt}^{(k+1)dt} e^{\int_\tau^{kdt} p_s \mathrm{d}s} p_\tau \mathrm{d}\tau \\
=& \widehat{p}_k\left(\widehat{I}_k - \bar{I}\right) + c\int_{kdt}^{(k+1)dt} e^{\int_\tau^{kdt} p_s \mathrm{d}s}\mathrm{d}\tau - \bar{I}\left(\int_{kdt}^{(k+1)dt} e^{\int_\tau^{kdt} p_s \mathrm{d}s} p_\tau \mathrm{d}\tau - \widehat{p}_k\right)
\end{aligned}
$$

First, since when $\tau \in [kdt, (k+1)dt]$, $\int_\tau^{kdt} p_s \mathrm{d}s \leqslant 0$, $\int_{kdt}^{(k+1)dt} e^{\int_\tau^{kdt} p_s \mathrm{d}s}\mathrm{d}\tau \leqslant dt$. Then we consider

$$
\begin{aligned}
& \int_{kdt}^{(k+1)dt} e^{\int_\tau^{kdt} p_s \mathrm{d}s} p_\tau \mathrm{d}\tau - \widehat{p}_k \\
=& \int_{kdt}^{(k+1)dt} e^{\int_\tau^{kdt} p_s \mathrm{d}s} p_\tau \mathrm{d}\tau - 1 + e^{-\int_{kdt}^{(k+1)dt} p_s \mathrm{d}s}
\end{aligned}
$$

Let

$$
\begin{aligned}
H(t,t') =& \int_t^{t'} e^{\int_\tau^t p_s \mathrm{d}s} p_\tau \mathrm{d}\tau - 1 + e^{-\int_t^{t'} p_s \mathrm{d}s} \\
\implies \quad \frac{\partial H(t,t')}{\partial t'} =& 0 \& H(t,t) = 0 \\
\implies \quad H(t,t') \equiv& 0
\end{aligned}
$$

Therefore, to sum up:

$$
\Delta\widehat{I}_k(1 - \widehat{p}_k) + \widehat{p}_k(\bar{I} - \widehat{I}_k) \leqslant \widehat{c}
$$

We have checked that $\widehat{p}_k, \widehat{P}_k, \widehat{I}_k$ is feasible in problem Equation (2.2) with parameter $\widehat{\rho}_k$ and $\widehat{c}$. Then by Theorem 2.1:

$$\begin{aligned}
\sum_{k=1}^{[T/dt]} \widehat{\rho}_k \left(1 - \widehat{P}_{k-1}\right) \widehat{p}_k &\leqslant \sum_{k=1}^{\infty} \widehat{\rho}_t \left(\frac{\bar{I} - \widehat{c}}{\bar{I}}\right)^{k-1} \frac{\widehat{c}}{\bar{I}} \\
\implies \int_0^{\infty} \rho_t (1 - P_t) p_t \mathrm{d}t &\leqslant \sum_{k=1}^{\infty} \widehat{\rho}_t \left(\frac{\bar{I} - \widehat{c}}{\bar{I}}\right)^{k-1} \frac{\widehat{c}}{\bar{I}} + 2\varepsilon \\
&= \sum_{k=1}^{\infty} \rho_{kdt} \left(1 - \frac{c}{\bar{I}} dt\right)^k \frac{c}{\bar{I}} dt + \varepsilon
\end{aligned}$$

Since $\log(1-x) \leqslant -x$, $\left(1 - \frac{cdt}{\bar{I}}\right)^k \leqslant e^{-\frac{c}{\bar{I}}kdt}$ Then:

$$\int_0^{\infty} \rho_t (1 - P_t) p_t \mathrm{d}t \leqslant \sum_{k=1}^{\infty} \rho_{kdt} e^{-\frac{c}{\bar{I}}kdt} \frac{c}{\bar{I}} dt + 2\varepsilon$$

On the other hand, since $\rho_t e^{-\frac{c}{\bar{I}}t}$ is integrable, there exists dt sufficiently small that:

$$\begin{aligned}
\sum_{k=1}^{\infty} \rho_{kdt} e^{-\frac{c}{\bar{I}}kdt} \frac{c}{\bar{I}} dt &\leqslant \int_{t=0}^{\infty} \rho_t e^{-\frac{c}{\bar{I}}t} \mathrm{d}t + \varepsilon \\
\implies \int_0^{\infty} \rho_t (1 - P_t) p_t \mathrm{d}t &\leqslant \int_0^{\infty} \rho_t e^{-\frac{c}{\bar{I}}t} \mathrm{d}t + 3\varepsilon
\end{aligned}$$

Let $\varepsilon \to 0$, then we showed that:

$$V \leqslant \int_0^{\infty} \rho_t e^{-\frac{c}{\bar{I}}t} \mathrm{d}t$$

■

C.1.6 Proof of Lemma 2.4

Proof. Similar to discussion in proof of Lemma 2.2, I only prove for $E[\mathcal{T}] < \infty$. Let:

$$\begin{cases} P_t = \mathrm{P}(\mathcal{T} \leqslant t) \\ I_t = E[H(\mu) - H(\mu_t)|\mathcal{T} > t] \end{cases}$$

Then be proof of Lemma 2.3:

$$\mathrm{d}I_t = E[-\mathrm{d}H(\mu_t)|\mathcal{T} > t] - \frac{\mathrm{d}P_t}{1 - P_t}\left(\bar{I} - I_t\right) \tag{C.5}$$

Consider $E[\mathcal{T}]$:

$$\begin{aligned} E[\mathcal{T}] &= \frac{1}{c}\int_0^\infty (1 - P_t)c\mathrm{d}t \\ &\geqslant \frac{1}{c}\int_0^\infty (1 - P_t)E[-\mathrm{d}H(\mu_t)|\mathcal{T} > t] \\ &= \frac{1}{c}\left(\int_0^\infty (1 - P_t)\mathrm{d}I_t + \int_0^\infty (\bar{I} - I_t)\mathrm{d}P_t\right) \\ &= \frac{\bar{I}}{c} + \int_0^\infty ((1 - P_t)\mathrm{d}I_t + I_t\mathrm{d}(1 - P_t)) \\ &= \frac{\bar{I}}{c} + I_t(1 - P_t)|_0^\infty \\ &= \frac{\bar{I}}{c} \end{aligned}$$

Inequality is flow informativeness constraint. Second equality is by Equation (C.5). Forth equality is by intergral by part. ■

Appendix D

Appendix for Chapter 3

D.1 Proof in Section 1.3

D.1.1 Proof of Proposition 3.2

Proof. (Necessity) First suppose $I^*(\mathcal{S};\mathcal{X}|\mu)$ satisfies Assumption 3.1. Then choose I^* itself as I. $\forall\mu$ and $\mathcal{S}$. $\forall\mathcal{X}\to\left(\mathcal{S}^1,\ldots,\mathcal{S}^N\right)\to\mathcal{S}$:

$$\begin{aligned} E\left[\sum_{i=1}^{N} I^*\left(\mathcal{S}^i;\mathcal{X}|\mathcal{S}^1,\ldots,\mathcal{S}^{i-1}\right)\right] &\geqslant I^*\left(\left(\mathcal{S}^1,\ldots\mathcal{S}^N\right);\mathcal{X}|\mu\right) \\ &\geqslant I^*(\mathcal{S};\mathcal{X}|\mu) \\ \Longrightarrow \inf_{\left(\mathcal{S}^i,N\right)} E\left[\sum_{i=1}^{N} I^*\left(\mathcal{S}^i;\mathcal{X}|\mathcal{S}^1,\ldots,\mathcal{S}^{i-1}\right)\right] &\geqslant I^*(\mathcal{S};\mathcal{X}|\mu) \end{aligned}$$

First inequality is from sub-additivity. Second inequality is from monotonicity. On the other hand, let $\mathcal{S}^1=\mathcal{S}$, $N=1$, then

$$\begin{aligned} E\left[\sum_{i=1}^{N} I^*\left(\mathcal{S}^i;\mathcal{X}|\mathcal{S}^1,\ldots,\mathcal{S}^{i-1}\right)\right] &= I^*(\mathcal{S};\mathcal{X}|\mu) \\ \Longrightarrow \inf_{\left(\mathcal{S}^i,N\right)} E\left[\sum_{i=1}^{N} I^*\left(\mathcal{S}^i;\mathcal{X}|\mathcal{S}^1,\ldots,\mathcal{S}^{i-1}\right)\right] &\leqslant I^*(\mathcal{S};\mathcal{X}|\mu) \end{aligned}$$

Combining the two direction of inequality:

$$\inf_{\left(\mathcal{S}^i,N\right)} E\left[\sum_{i=1}^{N} I^*\left(\mathcal{S}^i;\mathcal{X}|\mathcal{S}^1,\ldots,\mathcal{S}^{i-1}\right)\right] = I^*(\mathcal{S};\mathcal{X}|\mu)$$

(Sufficiency) On the other hand, suppose given $I(\mathcal{S};\mathcal{X}|\mu)$,

$$\begin{aligned} I^*(\mathcal{S};\mathcal{X}|\mu) = \inf_{\left(\mathcal{S}^i,N\right)} & E\left[\sum_{i=1}^{N} I\left(\mathcal{S}^i;\mathcal{X}|\mathcal{S}^1,\ldots,\mathcal{S}^{i-1}\right)\right] \\ &\text{s.t. } \mathcal{X}\to\left(\mathcal{S}^1,\ldots,\mathcal{S}^N\right)\to\mathcal{S} \end{aligned}$$

Then

0. *Uninformative signal*: First it's not hard to observe that acquiring no informaiton is sufficient for an uninformative signal $\mathcal{S}$. Therefore if choose $N = 0$ we have, $0 \geqslant I^*(\mathcal{S};\mathcal{X}|\mu)$. Then:

$$I^*(\mathcal{S};\mathcal{X}|\mu) = 0$$

1. *Monotonicity*: $\forall \left(\mathcal{S}^i\right)$ s.t. $\mathcal{X} \to \left(\mathcal{S}^1,\ldots,\mathcal{S}^N\right) \to \mathcal{S}$. Since $\mathcal{X} \to \mathcal{S} \to \mathcal{T}$, we have $\mathcal{X} \to \left(\mathcal{S}^1,\ldots,\mathcal{S}^N\right) \to \mathcal{T}$. Therefore:

$$\begin{aligned}
&\inf_{(\mathcal{T}^i,N)} E\left[\sum_{i=1}^{N} I\left(\mathcal{T}^i;\mathcal{X}|\mathcal{T}^1,\ldots,\mathcal{T}^{i-1}\right)\right] \\
&\text{s.t. } \mathcal{X} \to \left(\mathcal{T}^1,\ldots,\mathcal{T}^N\right) \to \mathcal{T} \\
\leqslant &E\left[\sum_{i=1}^{N} I\left(\mathcal{S}^i;\mathcal{X}|\mathcal{S}^1,\ldots,\mathcal{S}^{i-1}\right)\right] \\
\Longrightarrow \quad &\inf_{(\mathcal{T}^i,N)} E\left[\sum_{i=1}^{N} I\left(\mathcal{T}^i;\mathcal{T}|\mathcal{T}^1,\ldots,\mathcal{T}^{i-1}\right)\right] \\
&\text{s.t. } \mathcal{X} \to \left(\mathcal{T}^1,\ldots,\mathcal{T}^N\right) \to \mathcal{T} \\
\leqslant &\inf_{(\mathcal{S}^i,N)} E\left[\sum_{i=1}^{N} I\left(\mathcal{S}^i;\mathcal{X}|\mathcal{S}^1,\ldots,\mathcal{S}^{i-1}\right)\right] \\
&\text{s.t. } \mathcal{X} \to \left(\mathcal{S}^1,\ldots,\mathcal{S}^N\right) \to \mathcal{S} \\
\Longrightarrow \quad &I^*(\mathcal{T};\mathcal{X}|\mu) \leqslant I^*(\mathcal{S};\mathcal{X}|\mu)
\end{aligned}$$

First inequality comes from that factor that $\left(\mathcal{S}^i\right)$ serves as one feasible group of $\left(\mathcal{T}^i\right)$ in the minimization. Second inequality comes from taking inf on RHS. Final inequality comes from definition of I^*.

2. *Sub-additivity*: Suppose $\mathcal{S} = (\mathcal{S}_1,\mathcal{S}_2)$. $\forall \left(\mathcal{S}_1^1,\ldots,\mathcal{S}_1^{N_1}\right)$ s.t. $\mathcal{X} \to \left(\mathcal{S}_1^1,\ldots,\mathcal{S}_1^{N_1}\right) \to \mathcal{S}_1$. $\forall \left(\mathcal{S}_2^1,\ldots,\mathcal{S}_2^{N_2}\right)$ conditinal on $\mathcal{S}_1$ s.t. $\forall$ realization of $\mathcal{S}_1$, $\mathcal{X} \to \left(\mathcal{S}_2^1,\ldots,\mathcal{S}_2^{N_2}\right) \to \mathcal{S}_2$.

Therefore:

$$\begin{aligned}
&\mathcal{X} \to \left(\mathcal{S}_1^1, \ldots, \mathcal{S}_1^{N_1}, \mathcal{S}_2^1, \ldots, \mathcal{S}_2^{N_2}\right) \to (\mathcal{S}_1, \mathcal{S}_2) \to \mathcal{S} \\
\implies & I^*(\mathcal{S}; \mathcal{X}|\mu) \leqslant E\left[\sum_{i=1}^{N_1} I(\mathcal{S}_1^i; \mathcal{S}|\mathcal{S}_1^1, \ldots, \mathcal{S}_1^{i-1})\right] + E\left[\sum_{i=1}^{N_2} I(\mathcal{S}_2^i; \mathcal{X}|\mathcal{S}_1, \mathcal{S}_2^1, \ldots, \mathcal{S}_2^{i-1})\right] \\
\implies & I^*(\mathcal{S}; \mathcal{X}|\mu) \leqslant \inf E\left[\sum_{i=1}^{N_1} I(\mathcal{S}_1^i; \mathcal{S}|\mathcal{S}_1^1, \ldots, \mathcal{S}_1^{i-1})\right] + \inf E\left[\sum_{i=1}^{N_2} I(\mathcal{S}_2^i; \mathcal{X}|\mathcal{S}_1, \mathcal{S}_2^1, \ldots, \mathcal{S}_2^{i-1})\right] \\
\implies & I^*(\mathcal{S}; \mathcal{X}|\mu) \leqslant I^*(\mathcal{S}_1; \mathcal{X}|\mu) + E[I^*(\mathcal{S}_2; \mathcal{X}|\mathcal{S}_1, \mu)]
\end{aligned}$$

3. *C-linearity*: $\forall \mathcal{S}$, consider $\mathcal{S}^1 = (\{0,1\}, \lambda, 1-\lambda)$ being an uninformative binary signal. $\mathcal{S}^2 = \mathcal{S}$ when $\mathcal{S}^1 = 0$ and constant when $\mathcal{S}^1 = 1$. Therefore $\left(\mathcal{S}^1, \mathcal{S}^2\right) = \mathcal{S}_\lambda$. By sub-additivity:

$$I^*(\mathcal{S}_\lambda; \mathcal{X}|\mu) \leqslant \lambda I^*(\mathcal{S}; \mathcal{X}|\mu)$$

On the other hand, consider $\mathcal{S}^1$ conditional on $\mathcal{S}_\lambda$. If $\mathcal{S}_\lambda$ induces $\nu \neq \mu$, then $\mathcal{S}^1$ is uninformative. Otherwise $\mathcal{S}^1 = \mathcal{S}$. Then $\left(\mathcal{S}_\lambda, \mathcal{S}^1\right) = \mathcal{S}$, by sub-additivity:

$$\begin{aligned}
& I^*(\mathcal{S}; \mathcal{X}|\mu) \leqslant I^*(\mathcal{S}_\lambda; \mathcal{X}|\mu) + (1-\lambda) I^*(\mathcal{S}; \mathcal{X}|\mu) \\
\implies & \lambda I^*(\mathcal{S}; \mathcal{X}|\mu) \leqslant I^*(\mathcal{S}_\lambda; \mathcal{X}|\mu)
\end{aligned}$$

To sum up, $\lambda I^*(\mathcal{S}; \mathcal{X}|\mu) = I^*(\mathcal{S}_\lambda; \mathcal{X}|\mu)$.

■

D.2 Proof in Section 3.3

D.2.1 Proof of Theorem 3.1

Proof. Let $V(\mu)$ be expected utility in Equation (P). Then by assumption $V(\mu) \geqslant 0$. Suppose $V(\mu) = 0$, then Theorem 3.1 is straight forward. $V(\mu)$ is achieved by choosing doing nothing and acquiring no information. From now on, we assume $V(\mu) > 0$. Pick any $\varepsilon < V(\mu)$, we want to show that there exists $\mathcal{A}, T$ s.t.:

$$V(\mu) - \varepsilon \leqslant E[u(\mathcal{A}, \mathcal{X})] - mT - Tf\left(\frac{I(\mathcal{A}; \mathcal{X}|\preceq)}{T}\right)$$

Suppose $\left(\mathcal{S}^t, \mathcal{A}^t, \mathcal{T}\right)$ solves Equation (P) approches $V(\mu)$ up to $\frac{\varepsilon}{2}$:

$$V(\mu) - \frac{\varepsilon}{2} \leqslant E\left[u\left(\mathcal{A}^{\mathcal{T}}, \mathcal{X}\right) - m\mathcal{T} - \sum_{t=0}^{\infty} f\left(I\left(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}\right)\right)\right]$$

$$\text{where} \begin{cases} \mathcal{X} \to \mathcal{S}^{t-1} \to \mathbf{1}_{\mathcal{T}\leqslant t} \\ \mathcal{X} \to \mathcal{S}^{t-1} \to \mathcal{A}^t \text{ conditional on } \mathcal{T} = t \end{cases}$$

Lemma D.1 shows that we can assume that the signal structure WLOG takes the following form:

$$\mathcal{S}^t = \begin{cases} s_0 & \text{when } \mathcal{T} \leqslant t \\ \mathcal{A}^{t+1} & \text{when } \mathcal{T} = t+1 \end{cases}$$

Therefore, $\mathcal{A}^{t+1}$,$\mathbf{1}_{\mathcal{T}\leqslant t}$ and $\mathbf{1}_{\mathcal{T}=t+1}$ are all explicitly signal realizations included in $\mathcal{S}^t$. We discuss two cases separately:

Case 1. $E[\mathcal{T}] \geqslant 1$: Consider $\mathcal{S}^{\mathbf{T}} = \left(\mathcal{S}^0, \mathcal{S}^1, \cdots, \mathcal{S}^T\right)$ as a combined information structure of all signals in first T periods. By sub-additivity in Assumption 3.1:

$$
\begin{aligned}
&E\left[\sum_{t=0}^{\infty} I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}},\mathbf{1}_{\mathcal{T}\leqslant t})\right]\\
=&E\left[\sum_{t=0}^{\infty} I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})\right]\\
=&I\left(\mathcal{S}^0;\mathcal{X}|\mu\right)+E\left[\sum_{t=1}^{\infty} I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})\right]\\
=&I\left(\mathcal{S}^0;\mathcal{X}|\mu\right)+E\left[I\left(\mathcal{S}^1;\mathcal{X}|\mathcal{S}^0\right)\right]+E\left[\sum_{t=2}^{\infty} I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})\right]\\
\geqslant&I(\mathcal{S}^{\mathbf{1}};\mathcal{X}|\mu)+E\left[\sum_{t=2}^{\infty} I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})\right]\\
\geqslant&\dots\\
\geqslant&I\left(\mathcal{S}^{\mathbf{T}};\mathcal{X}|\mu\right)+E\left[\sum_{t=T+1}^{\infty} I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})\right]\\
\Longrightarrow&\sum_{t=0}^{\infty} E[f(I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}}))]\geqslant I\left(\mathcal{S}^{\mathbf{T}};\mathcal{X}|\mu\right)\ \forall T
\end{aligned}
\tag{D.1}
$$

Now consider:

$$
\begin{aligned}
&\sum_{t=0}^{\infty} E[I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})]\\
=&\sum_{t=0}^{\infty}(\mathrm{Prob}(\mathcal{T}\leqslant t)E\left[I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})|\mathcal{T}\leqslant t\right]+\mathrm{Prob}(\mathcal{T}>t)E\left[I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})|\mathcal{T}>t\right])\\
=&\sum_{t=0}^{\infty}(\mathrm{Prob}(\mathcal{T}\leqslant t)E\left[I(s_0;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})|\mathcal{T}\leqslant t\right]+\mathrm{Prob}(\mathcal{T}>t)E\left[I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})|\mathcal{T}>t\right])\\
=&\sum_{t=0}^{\infty}\mathrm{Prob}(\mathcal{T}>t)E\left[I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t-1}})|\mathcal{T}>t\right]
\end{aligned}
\tag{D.2}
$$

Since $V(\mu) - \varepsilon > 0$, then:

$$
\begin{aligned}
& m \cdot E[\mathcal{T}] \leqslant \max v \\
\implies & \operatorname{Prob}(\mathcal{T} > T) \cdot T \cdot m \leqslant \max v \\
\implies & \operatorname{Prob}(\mathcal{T} > T) \leqslant \frac{\max v}{mT} \\
\implies & \operatorname{Prob}(\mathcal{T} \leqslant T) E\Big[u\Big(\mathcal{A}^{\mathcal{T}}, \mathcal{X}\Big) | \mathcal{T} \leqslant t\Big] \geqslant E\Big[u\Big(\mathcal{A}^{\mathcal{T}}, \mathcal{X}\Big)\Big] - \operatorname{Prob}(\mathcal{T} > T) \cdot \max v \\
\implies & \operatorname{Prob}(\mathcal{T} \leqslant T) E\Big[u\Big(\mathcal{A}^{\mathcal{T}}, \mathcal{X}\Big) | \mathcal{T} \leqslant t\Big] \geqslant E\Big[u\Big(\mathcal{A}^{\mathcal{T}}, \mathcal{X}\Big)\Big] - \frac{\max v^2}{mT}
\end{aligned}
$$

Choose $T + 1 > \frac{m\varepsilon}{\max v^2}$. Now combine Equations (D.1) and (D.2), and $\sum_{t=0}^{\infty} \operatorname{Prob}(\mathcal{T} > t) = E[\mathcal{T}]$, then we have:

$$
\begin{aligned}
& I\Big(\mathcal{S}^{\mathbf{T}}; \mathcal{X} | \mu\Big) \leqslant \sum_{t=0}^{\infty} \operatorname{Prob}(\mathcal{T} > t) E\big[I\big(\mathcal{S}^t; \mathcal{X} | \mathcal{S}^{\mathbf{t-1}}\big) | \mathcal{T} > t\big] \\
\implies & \frac{I(\mathcal{S}^{\mathbf{T}}; \mathcal{X} | \mu)}{E[\mathcal{T}]} \leqslant \sum_{t=0}^{\infty} \frac{\operatorname{Prob}(\mathcal{T} > t)}{\sum \operatorname{Prob}(\mathcal{T} > \tau)} E\big[I\big(\mathcal{S}^t; \mathcal{X} | \mathcal{S}^{\mathbf{t-1}}\big) | \mathcal{T} > t\big] \\
\implies & f\left(\frac{I(\mathcal{S}^{\mathbf{T}}; \mathcal{X} | \mu)}{E[\mathcal{T}]}\right) \leqslant \sum_{t=0}^{\infty} \frac{\operatorname{Prob}(\mathcal{T} \geqslant t)}{\sum \operatorname{Prob}(\mathcal{T} > \tau)} f\big(E\big[I\big(\mathcal{S}^t; \mathcal{X} | \mathcal{S}^{\mathbf{t-1}}\big) | \mathcal{T} > t\big]\big) \\
\implies & E[\mathcal{T}] f\left(\frac{I(\mathcal{S}^{\mathbf{T}}; \mathcal{X} | \mu)}{E[\mathcal{T}]}\right) \leqslant \sum_{t=0}^{\infty} \operatorname{Prob}(\mathcal{T} > t) E\big[f(I(\mathcal{S}^t; \mathcal{X} | \mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} > t}))\big] \\
\implies & E[\mathcal{T}] f\left(\frac{I(\mathcal{S}^{\mathbf{T}}; \mathcal{X} | \mu)}{E[\mathcal{T}]}\right) \leqslant E\left[\sum_{t=0}^{\infty} f(I(\mathcal{S}^t; \mathcal{X} | \mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \leqslant t}))\right]
\end{aligned}
$$

Consider $\mathcal{A}^{\mathbf{T+1}} = \big(\mathcal{A}^0, \mathcal{A}^2, \cdots, \mathcal{A}^{T+1}\big)$ as a random variable which summarizes realizations of all $\mathcal{A}^t$. Since $\mathcal{A}^{t+1}$ are directly included in $\mathcal{S}^t$, we have $\mathcal{X} \to \mathcal{S}^{\mathbf{T}} \to \mathcal{A}^{\mathbf{T+1}}$. Therefore, by Assumption 3.1:

$$
I\Big(\mathcal{A}^{\mathbf{T}+1}; \mathcal{X} | \mu\Big) \leqslant I\Big(\mathcal{S}^{\mathbf{T}}; \mathcal{X} | \mu\Big)
$$

$$\Longrightarrow E[\mathcal{T}]f\left(\frac{I(\mathcal{A}^{\mathbf{T}+\mathbf{1}};\mathcal{X}|\mu)}{E[\mathcal{T}]}\right) \leqslant E\left[\sum_{t=0}^{\infty} f(I(\mathcal{S}^t;\mathcal{X}|\mathcal{S}^{\mathbf{t}-\mathbf{1}}))\right]$$

That's to say, if we can implement $\mathcal{A}^{\mathbf{T}+\mathbf{1}}$ with expected waiting time $E[\mathcal{T}]$ and information cost $E[\mathcal{T}]f\left(\frac{I(\mathcal{A}^{\mathbf{T}+\mathbf{1}};\mathcal{X}|\mu)}{E[\mathcal{T}]}\right)$, then utility level will be weakly higher than V. We define the new strategy as follows:

1. In each period, acquire a combined information structure by mixing $\mathcal{A}^{\mathbf{T}+\mathbf{1}}$ with probability $\frac{1}{E[\mathcal{T}]}$ and uninformative signal structure with probability $1-\frac{1}{E[\mathcal{T}]}$.
2. Following arrival of signal $\mathcal{A}^{\mathbf{T}+\mathbf{1}}$, choosing the corresponding action.
3. If no informative signal arrive, do nothing and go to next period.

It's not hard to see that in this strategy, action and signal are identical thus the three information processing constraint are naturally satisfied. In each period, the probabiltiy of decision making is $\frac{1}{E[\mathcal{T}]}$ and the distribution of actions is $\mathcal{A}^{\mathbf{T}+\mathbf{1}}$. Therefore, totally utility gain is:

$$\sum_{t=0}^{\infty}\left(1-\frac{1}{E[\mathcal{T}]}\right)^t \frac{1}{E[\mathcal{T}]}E[u(\mathcal{A}^{\mathbf{T}+\mathbf{1}},\mathcal{X})] = E\left[u(\mathcal{A}^{\mathbf{T}+\mathbf{1}},\mathcal{X})\right]$$

Expected waiting time is:

$$\sum_{t=0}^{\infty}\left(1-\frac{1}{E[\mathcal{T}]}\right)^t \frac{1}{E[\mathcal{T}]}\cdot t = E[\mathcal{T}]$$

Expected experimentation cost is:

$$\sum_{t=0}^{\infty}\left(1-\frac{1}{E[\mathcal{T}]}\right)^t f\left(\frac{I(\mathcal{A}^{\mathbf{T}+\mathbf{1}};\mathcal{X}|\mu)}{E[\mathcal{T}]}\right) = E[\mathcal{T}]f\left(\frac{I(\mathcal{A}^{\mathbf{T}+\mathbf{1}};\mathcal{X}|\mu)}{E[\mathcal{T}]}\right)$$

Therefore, we find a strategy which is no worse than original strategy than $\frac{\varepsilon}{2}$. Then:

$$\begin{aligned}
V(\mu) \leqslant & E[u(\mathcal{A}^{\mathbf{T+1}}, \mathcal{X})] - mE[\mathcal{T}] - E[\mathcal{T}] f\left(\frac{I(\mathcal{A}^{\mathbf{T+1}}; \mathcal{X}|\mu)}{E[\mathcal{T}]}\right) + \varepsilon \\
\leqslant & \sup_{\mathcal{A},T} E[u(\mathcal{A}, \mathcal{X})] - mT - Tf\left(\frac{I(\mathcal{A}; \mathcal{X}|\preceq)}{T}\right) + \varepsilon \ \forall \varepsilon \\
\implies V(\mu) \leqslant & \sup_{\mathcal{A},T} E[u(\mathcal{A}, \mathcal{X})] - mT - Tf\left(\frac{I(\mathcal{A}; \mathcal{X}|\preceq)}{T}\right)
\end{aligned}$$

Therefore, we proved Theorem 3.1 when $E[\mathcal{T}] \geqslant 1$.

Case 2. $E[\mathcal{T}] < 1$: Since $\mathcal{T} \in \mathbb{N}$, $E[\mathcal{T}] < 1$ means $P(\mathcal{T} = 0) > 0$. When $\mathcal{T} = 0$, no informatin is acquired yet and decision making is based on prior. Therefore:

$$\begin{aligned}
& E\left[u(\mathcal{A}^{\mathcal{T}}, \mathcal{X}) - m\mathcal{T} - \sum_{t=0}^{\mathcal{T}} f\big(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{\mathbf{t-1}})\big)\right] \\
= & \mathrm{Prob}(\mathcal{T} = 0) E\Big[u(\mathcal{A}^0, \mathcal{X})|\mathcal{T} = 0\Big] \\
& + \mathrm{Prob}(\mathcal{T} \geqslant 1) E\left[u(\mathcal{A}^{\mathcal{T}}, \mathcal{X}) - m\mathcal{T} - \sum_{t=1}^{\mathcal{T}} f\big(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{\mathbf{t-1}})\big)\middle|\mathcal{T} \geqslant 1\right] \\
\leqslant & \mathrm{Prob}(\mathcal{T} = 0) \max_a E_\mu[u(a, \mathcal{X})] \\
& + \mathrm{Prob}(\mathcal{T} \geqslant 1) E\left[u(\mathcal{A}^{\mathcal{T}}, \mathcal{X}) - m\mathcal{T} - \sum_{t=1}^{\mathcal{T}} f\big(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{\mathbf{t-1}})\big)\middle|\mathcal{T} \geqslant 1\right]
\end{aligned}$$

First equality is from law of iterated expectation. Inequality is from when $\mathcal{T} = 0$, choice of $\mathcal{A}^0$ is not necessarily optimal. Suppose:

$$\begin{aligned}
& \max_a E_\mu[u(a, \mathcal{X})] \geqslant E\left[u(\mathcal{A}^{\mathcal{T}}, \mathcal{X}) - m\mathcal{T} - \sum_{t=1}^{\mathcal{T}} f\Big(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{t-1})\Big)\middle|\mathcal{T} \geqslant 1\right] \\
\implies & E\left[u(\mathcal{A}^{\mathcal{T}}, \mathcal{X}) - m\mathcal{T} - \sum_{t=0}^{\mathcal{T}} f\Big(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{t-1})\Big)\right] \leqslant \max_a E_\mu[u(a, \mathcal{X})]
\end{aligned}$$

Then strategy $(\mathcal{S}^t, \mathcal{A}^t, \mathcal{T})$ is dominated by acquiring no information and this already proves Theorem 3.1. Suppose on the other hand:

$$\begin{aligned}
&\max_a E_\mu[u(a, \mathcal{X})] < E\left[u(\mathcal{A}^\mathcal{T}, \mathcal{X}) - m\mathcal{T} - \sum_{t=1}^{\mathcal{T}} f\Big(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{t-1})\Big)\middle|\mathcal{T} \geqslant 1\right] \\
\implies &E\left[u(\mathcal{A}^\mathcal{T}, \mathcal{X}) - m\mathcal{T} - \sum_{t=0}^{\mathcal{T}} f\Big(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{t-1})\Big)\right] \\
&< E\left[u(\mathcal{A}^\mathcal{T}, \mathcal{X}) - m\mathcal{T} - \sum_{t=1}^{\mathcal{T}} f\Big(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{t-1})\Big)\middle|\mathcal{T} \geqslant 1\right]
\end{aligned}$$

Then we define strategy $\mathcal{S}_1^t, \mathcal{A}_1^t, \mathcal{T}_1$ where: $(\mathcal{S}_1^t, \mathcal{A}_1^t, \mathcal{T}_1) = (\mathcal{S}^t, \mathcal{A}^t, \mathcal{T} \mid \mathcal{T} \geqslant 1)$. Then it's straight foward that:

$$\begin{aligned}
&E\left[u(\mathcal{A}_1^{\mathcal{T}_1}, \mathcal{X}) - m\mathcal{T}_1 - \sum_{t=0}^{\mathcal{T}_1} f\Big(I(\mathcal{S}_1^t; \mathcal{X}|\mathcal{S}_1^{t-1})\Big)\right] \\
=&E\left[u(\mathcal{A}^\mathcal{T}, \mathcal{X}) - m\mathcal{T} - \sum_{t=1}^{\mathcal{T}} f\Big(I(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{t-1})\Big)\middle|\mathcal{T} \geqslant 1\right]
\end{aligned}$$

We only need to verify the information processing constraints.

- When $\mathcal{T}_1 \leqslant t$: $\mathcal{S}_1^t = s_0$
- When $\mathcal{T}_1 = t + 1$: $\mathcal{S}_1^t = \mathcal{S}^t = \mathcal{A}^{t+1} = \mathcal{A}_1^{t+1}$.
- $\mathcal{T}_1 = 0$ happen with zero probability.

However, in this case $E[\mathcal{T} \geqslant 1]$. Therefore this goes back to case one. To sum up, we showed that:

$$V(\mu) \leqslant \max\left\{\sup_{\mathcal{A}} E[u(\mathcal{A}, \mathcal{X})],\ \sup_{\mathcal{A}, T \geqslant 1} E[u(\mathcal{A}, \mathcal{X})] - mT - Tf\left(\frac{I(\mathcal{A}; \mathcal{X}|\preceq)}{T}\right)\right\}$$

On the other hand the inequality of the other hand is straight foward, any strategy achieve the RHS can be achieved in origianl problem Equation (P). Therefore:

$$V(\mu) = \max\left\{\sup_{\mathcal{A}} E[u(\mathcal{A},\mathcal{X})],\ \sup_{\mathcal{A},T\geqslant 1} E[u(\mathcal{A},\mathcal{X})] - mT - Tf\left(\frac{I(\mathcal{A};\mathcal{X}|\preceq)}{T}\right)\right\} \tag{D.3}$$

Finally, we consider solving optimal T in Equation (3.1). Fix I, consider:

$$\inf_{T\geqslant 1}\left(mT + Tf\left(\frac{I}{T}\right)\right)$$

I first show that the objective function is quasi-convex. mT is already linear, so it's sufficient to show quasi-convexity of $T\left(\frac{I}{T}\right)$. By transforming argument, it's not hard to see that it's equivalent to show quasi-convexity of $\frac{f(I)}{I}$ w.r.t. I. Now consider $I_1 < I_2$ and $\lambda \in (0,1)$. Suppose by contradiction:

$$\begin{aligned}
&\frac{f(I_1)}{I_1} = \frac{f(I_2)}{I_2} < \frac{f(\lambda I_1(1-\lambda)I_2)}{\lambda I_1 + (1-\lambda)I_2} \\
\implies &\frac{\lambda f(I_1) + (1-\lambda)f(I_2)}{\lambda I_1 + (1-\lambda)I_2} < \frac{f(\lambda I_1(1-\lambda)I_2)}{\lambda I_1 + (1-\lambda)I_2}
\end{aligned}$$

contradicting convexity of $f(I)$. Therefore, $mT + T\left(\frac{I}{T}\right)$ is quasi-convex in T. Since f is convex, it always has one-side derivatives well defined. So an necessary condition for T solving the problem will be:

$$\begin{aligned}
&m + f\left(\frac{I}{T}\right) - \frac{I}{T}f'_{+}\left(\frac{I}{T}\right) \leqslant 0 \leqslant m + f\left(\frac{I}{T}\right) - \frac{I}{T}f'_{-}\left(\frac{I}{T}\right) \\
\overset{\lambda=\frac{I}{T}}{\iff} &m + f(\lambda) - \lambda f'_{+}(\lambda) \leqslant 0 \leqslant m + f(\lambda) - \lambda f'_{-}(\lambda) \\
\iff &\frac{m + f(\lambda)}{\lambda} \in \partial f(\lambda)
\end{aligned}$$

What's more, since f is convex, the correspondence $f(\lambda) - \lambda f'(\lambda)$ is increasing (in set order). Therefore, the set of λ such that $\frac{m+f(\lambda)}{\lambda} \in \partial f(\lambda)$ must be an connected interval. Therefore, $\frac{m+f(\lambda)}{\lambda} \in \partial f(\lambda)$ is actually also sufficient for minimizing $mT + Tf(\frac{I}{T})$.

Case 1.: $\{\lambda | m + f(\lambda) \in \lambda \partial f(\lambda)\} \neq \varnothing$: Since f is convex, ∂f is a continuous correspondence, therefore the set is closed. Pick the smallest λ:

$$\begin{aligned} mT + Tf\left(\frac{I}{T}\right) =& m\frac{I}{\lambda} + \frac{I}{\lambda}f(\lambda) \\ =& \left(\frac{m}{\lambda} + \frac{f(\lambda)}{\lambda}\right) I \end{aligned}$$

Therefore, the total cost paid can be summarized by:

$$\left(\frac{m}{\lambda} + \frac{f(\lambda)}{\lambda}\right) I(\mathcal{A}; \mathcal{X} | \mu)$$

Finally, the constraint $T \geqslant 1$ can be replaced by:

$$\begin{aligned} \frac{I(\mathcal{A}; \mathcal{X} | \mu)}{\lambda} \geqslant 1 \\ \iff I(\mathcal{A}; \mathcal{X} | \mu) \geqslant \lambda \end{aligned}$$

Theorem 3.1 is proved.

Case 2.: $m + f(\lambda) - \lambda \partial f(\lambda) > 0 \ \forall \lambda$. That is to say:

$$mT + Tf\left(\frac{I}{T}\right)$$

is strictly increasing in $T \ \forall I$. Therefore, independent of choice I, choosing smaller T will yield higher utility. T will eventually be smaller than 1. So we can rule out this case.

Case 3.: $m + f(\lambda) - \lambda\partial f(\lambda) < 0\ \forall\lambda$. That is to say:

$$mT + Tf\left(\frac{I}{T}\right)$$

is strictly decreasing in $T\ \forall I$. However this is not possible since:

$$\lim_{T\to\infty} mT + Tf\left(\frac{I}{T}\right) = +\infty$$

To sum up, if $\{\lambda | m + f(\lambda) \in \lambda\partial f(\lambda)\} = \varnothing$, then we define $\lambda = \infty$. Then the constraint for second term in Equation (D.3) can never be satisfied and $V(\mu) = \sup_a E[u(a, \mathcal{X})]$. ■

Lemma D.1 (Reduction of redundency). *$\left(\mathcal{S}^t, \mathcal{T}, \mathcal{A}^{\mathcal{T}}\right)$ solves Equation (P) if and only if there exists $\left(\tilde{\mathcal{S}}^T, \mathcal{T}, \mathcal{A}^{\mathcal{T}}\right)$ solving :*

$$\begin{aligned}
\sup_{\mathcal{S}^t, \mathcal{T}, \mathcal{A}^{\mathcal{T}}} \sum_{t=0}^{\infty} \Bigg(&\mathbf{P}[\mathcal{T} = t]\left(E\left[u(\mathcal{A}^t, \mathcal{X}) | \mathcal{T} = t\right] - m \cdot t\right) \\
&- \mathbf{P}[\mathcal{T} > t] E\left[f\left(I\left(\tilde{\mathcal{S}}^t; \mathcal{X} | \tilde{\mathcal{S}}^{\mathbf{t-1}}\right)\right) | \mathcal{T} > t\right]\Bigg)
\end{aligned} \tag{D.4}$$

$$s.t.\ \tilde{\mathcal{S}}^t = \begin{cases} s_0 & \text{when } \mathcal{T} < t+1 \\ \mathcal{A}^{t+1} & \text{when } \mathcal{T} = t+1 \\ \mathcal{S}^t & \text{when } \mathcal{T} > t+1 \end{cases}$$

What's more, the optimal utility level is same in Equation (P) and Equation (D.4).

Proof. Suppose $\left(\mathcal{S}^t, \mathcal{T}, \mathcal{A}^t\right)$ is a feasible strategy to Equation (P). Let first show that it's WLOG that the DM can discard all information after taking an action: take given $\mathcal{T}$ and

$\mathcal{A}^t$, take s_0 as a given degenerate signal, define $\hat{\mathcal{S}}^t$ as:

$$\hat{\mathcal{S}}^t = \begin{cases} \mathcal{S}^t & \text{when } \mathcal{T} \geqslant t+1 \\ s_0 & \text{when } \mathcal{T} \leqslant t \end{cases}$$

By definition, $\hat{\mathcal{S}}^{\mathbf{t}} = \mathcal{S}^{\mathbf{t}}$ conditional on $\mathcal{T} \geqslant t+1$. Therefore:

$$I\left(\hat{\mathcal{S}}^t; \mathcal{X} | \hat{\mathcal{S}}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \leqslant t}\right) = \begin{cases} I\left(\mathcal{S}^t; \mathcal{X} | \mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T} \leqslant t}\right) & \text{when } \mathcal{T} \leqslant t \\ 0 & \text{when } \mathcal{T} \geqslant t+1 \end{cases}$$

$$\mathcal{X} \to \hat{\mathcal{S}}^t \to \mathcal{A}^{t+1} \text{ conditional on } \mathcal{T} = t$$

To show that the first information processing constraint is satisfied, we discuss the case $\hat{\mathcal{S}} = s_0$ and $\hat{\mathcal{S}} \neq s_0$ separately:

- When $\hat{\mathcal{S}}^{t-1} = s_0$, $\mathcal{T} \leqslant t-1$. Therefore:

$$\text{Prob}\left(\mathcal{T} > t \middle| \hat{\mathcal{S}}^{t-1} = s_0, \mathcal{X}\right) = 0$$
$$\text{Prob}\left(\mathcal{T} \leqslant t \middle| \hat{\mathcal{S}}^{t-1} = s_0, \mathcal{X}\right) = 1$$

  which is independent of realization of $\mathcal{X}$.

- When $\hat{\mathcal{S}}^{t-1} \neq s_0$, $\mathcal{T} \geqslant t$. Then by law of total probability:

$$\begin{aligned} &\text{Prob}\left(\mathcal{T} > t \middle| \mathcal{S}^{t-1}\right) \\ =&\text{Prob}\left(\mathcal{T} > t \middle| \mathcal{S}^{t-1}, \mathcal{X}\right) \\ =&\text{Prob}\left(\mathcal{T} > t \middle| \mathcal{S}^{t-1}, \mathcal{X}, \mathcal{T} \geqslant t\right) \text{Prob}(\mathcal{T} \geqslant t | \mathcal{S}^{t-1}, \mathcal{X}) \end{aligned}$$

$$
\begin{aligned}
&+\operatorname{Prob}\Big(\mathcal{T} > t\big|\mathcal{S}^{t-1}, \mathcal{X}, \mathcal{T} < t\Big)\operatorname{Prob}(\mathcal{T} < t|\mathcal{S}^{t-1}, \mathcal{X}) \\
&=\operatorname{Prob}\Big(\mathcal{T} > t\big|\mathcal{S}^{t-1}, \mathcal{X}, \mathcal{T} \geqslant t\Big)\operatorname{Prob}(\mathcal{T} \geqslant t|\mathcal{S}^{t-1}, \mathcal{X}) \\
\implies &\operatorname{Prob}\Big(\mathcal{T} > t\big|\widehat{\mathcal{S}}^{t-1}, \mathcal{X}\Big) \\
&=\frac{\operatorname{Prob}\Big(\mathcal{T} > t|\widehat{\mathcal{S}}^{t-1}\Big)}{\operatorname{Prob}\Big(\mathcal{T} \geqslant t|\widehat{\mathcal{S}}^{t-1}, \mathcal{X}\Big)} \\
&=\operatorname{Prob}\Big(\mathcal{T} > t|\widehat{\mathcal{S}}^{t-1}\Big)
\end{aligned}
$$

which is independent of realization of $\mathcal{X}$.

Therefore, we proved that:

$$\mathcal{X} \to \widehat{\mathcal{S}}^{t-1} \to \mathbf{1}_{\mathcal{T} \leqslant t}$$

Therefore $\left(\widehat{\mathcal{S}}^t, \mathcal{A}^t, \mathcal{T}\right)$ is feasible and :

$$
\begin{aligned}
&E\left[\sum_{t=0}^{\infty} f\Big(I\Big(\widehat{\mathcal{S}}^t; \mathcal{X}\big|\widehat{\mathcal{S}}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}\Big)\Big)\right] \\
=&E\left[\sum_{t=0}^{\mathcal{T}-1} f\Big(I\Big(\widehat{\mathcal{S}}^t; \mathcal{X}\big|\widehat{\mathcal{S}}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}\Big)\Big)\right] \\
=&E\left[\sum_{t=0}^{\mathcal{T}-1} f\big(I\big(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}\big)\big)\right] \\
\leqslant&E\left[\sum_{t=0}^{\infty} f\big(I\big(\mathcal{S}^t; \mathcal{X}|\mathcal{S}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}\big)\big)\right]
\end{aligned}
$$

Therefore, $\left(\widehat{\mathcal{S}}^t, \mathcal{A}^t, \mathcal{T}\right)$ is a feasible strategy dominating $\left(\mathcal{S}^t, \mathcal{A}^t, \mathcal{T}\right)$. Now we define $\widetilde{\mathcal{S}}^t$:

$$\widetilde{\mathcal{S}}^t = \begin{cases} s_0 & \text{when } \mathcal{T} < t+1 \\ \mathcal{A}^{t+1} & \text{when } \mathcal{T} = t+1 \\ \widehat{\mathcal{S}}^t & \text{when } \mathcal{T} > t+1 \end{cases}$$

Initial information $\widetilde{\mathcal{S}}^{-1}$ is defined as a degenerate(uninformative) signal and induced belief is the prior. Verify the properties of $\widetilde{\mathcal{S}}^t$:

1. When $\widetilde{\mathcal{S}}^{t-1} \in \{s_0\} \cup A$, it's for sure that $\mathcal{T} \leqslant t$. Otherwise, $\mathcal{T} > t$. Therefore $\mathbf{1}_{\mathcal{T} \leqslant t}$ is a direct garbling of $\widetilde{\mathcal{S}}^{t-1}$. So we must have $\mathcal{X} \to \widetilde{\mathcal{S}}^{t-1} \to \mathbf{1}_{\mathcal{T} \leqslant t}$.
2. When $\mathcal{T} = t$, $\mathcal{A}^t = \widetilde{\mathcal{S}}^{t-1}$. Therefore $\mathcal{X} \to \widetilde{\mathcal{S}}^{t-1} \to \mathcal{A}^t$ conditional on $\mathcal{T} = t$.
3. Information measure associated with $\left(\widetilde{\mathcal{S}}^t, \mathcal{A}^t, \mathcal{T}\right)$ when $\mathcal{T} > t$:

$$\begin{aligned}
& I\left(\widetilde{\mathcal{S}}^t; \mathcal{X} \middle| \widetilde{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > t\right) \\
=& \mathbf{1}_{\mathcal{T}=t+1} I\left(\mathcal{A}^{t+1}; \mathcal{X} \middle| \widetilde{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} = t+1\right) \\
& + \mathbf{1}_{\mathcal{T}>t+1} I\left(\widetilde{\mathcal{S}}^t; \mathcal{X} \middle| \widetilde{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > t+1\right) \\
=& \mathbf{1}_{\mathcal{T}=t+1} I\left(\mathcal{A}^{t+1}; \mathcal{X} \middle| \widehat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} = t+1\right) \\
& + \mathbf{1}_{\mathcal{T}>t+1} I\left(\widehat{\mathcal{S}}^t; \mathcal{X} \middle| \widehat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > t+1\right) \\
\leqslant& \mathbf{1}_{\mathcal{T}=t+1} I\left(\widehat{\mathcal{S}}^t; \mathcal{X} \middle| \widehat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} = t+1\right) \\
& + \mathbf{1}_{\mathcal{T}>t+1} I\left(\widehat{\mathcal{S}}^t; \mathcal{X} \middle| \widehat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > t+1\right) \\
=& I\left(\widehat{\mathcal{S}}^t; \mathcal{X} \middle| \widehat{\mathcal{S}}^{\mathbf{t-1}}, \mathcal{T} > 1\right)
\end{aligned}$$

First inequality is simply rewriting two possible cases of $\mathcal{T}$. Second equality is from definition of $\widetilde{\mathcal{S}}^t$ when $\mathcal{T} > t+1$. First inequality is from $\mathcal{X} \to \widehat{\mathcal{S}}^t \to \mathcal{A}^{t+1}$ conditional on

$\mathcal{T} = t+1$. Therefore, $\left(\tilde{\mathcal{S}}^t, \mathcal{T}, \mathcal{A}^t\right)$ dominates the original solution in Equation (P) by achieving same action profile but lower costs. $\left(\tilde{\mathcal{S}}^t, \mathcal{T}, \mathcal{A}^t\right)$ is a feasible solution to Equation (D.4). Therefore solving Equation (D.4) yields a weakly higher utility than Equation (P). What remains to be proved is that any $\left(\tilde{\mathcal{S}}^t, \mathcal{T}, \mathcal{A}^t\right)$ feasible in Equation (D.4) can be dominated by some strategy feasible in Equation (P). It's not hard to see that it's feasible in Equation (P). Finally we show that the two formulation gives same utility:

$$\begin{aligned}
&E\left[E[u(\mathcal{A}^{\mathcal{T}}, \mathcal{X})] - m\cdot\mathcal{T} - \sum_{t=0}^{\infty} e^{-\rho dt\cdot t} f\left(I\left(\tilde{\mathcal{S}}^t; \mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}\right)\right)\right]\\
=&\sum_{t=0}^{\infty}\left(\mathbf{P}[\mathcal{T}=t]\left(E\left[u(\mathcal{A}^t,\mathcal{X})|\mathcal{T}=t\right] - m\cdot t\right) - E\left[f\left(I\left(\tilde{\mathcal{S}}^t;\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}}, \mathbf{1}_{\mathcal{T}\leqslant t}\right)\right)\right]\right)\\
=&\sum_{t=0}^{\infty}\left(\mathbf{P}[\mathcal{T}=t]\left(E\left[u(\mathcal{A}^t,\mathcal{X})|\mathcal{T}=t\right] - m\cdot t\right)\mathbf{P}[\mathcal{T}>t]E\left[f\left(I\left(\tilde{\mathcal{S}}^t;\mathcal{X}|\tilde{\mathcal{S}}^{\mathbf{t-1}}\right)\right)\middle|\mathcal{T}>t\right]\right)
\end{aligned}$$

Therefore, Equation (P) is equivalent to Equation (D.4). ■

D.2.2 *Proof of Proposition 3.3*

Proof. The outer maximization of Equation (3.1) is trivial. We focus on solving:

$$\overline{V}(\mu) = \sup_{I(\mathcal{A},\mathcal{X}|\mu)\geqslant\lambda} E[u(\mathcal{A},\mathcal{X})] - \left(\frac{m}{\lambda} + \frac{f(\lambda)}{\lambda}\right) I(\mathcal{A};\mathcal{X}|\mu) \tag{D.5}$$

Case 1. $\lambda^* < \infty$. By definition of λ^*, we know that

$$\lambda^* = \inf \arg\min_{\lambda}\left(\frac{m}{\lambda} + \frac{f(\lambda)}{\lambda}\right)$$

Let

$$g(I) = \left(\frac{m + f(\min\{I,\lambda^*\})}{\min\{I,\lambda^*\}}\right) I$$

Then $\frac{m+f(\lambda^*)}{\lambda^*} I \leqslant g(I) \leqslant m + f(I)$ and $g(I)$ is a convex function on $[0, \infty)$. Equation (D.5) can be rewritten as:

$$\overline{V}(\mu) = \sup_{\mathcal{A}} E[u(\mathcal{A}, \mathcal{X})] - g(I(\mathcal{A}; \mathcal{X}|\mu)) \tag{D.6}$$

Therefore by definition:

$$V^2(\mu) \leqslant \overline{V}(\mu) \leqslant V^1(\mu)$$

Now it is sufficient to show that if $\sup_{\mathcal{A}\in\mathbb{A}} I(\mathcal{A}; \mathcal{X}|\mu) \geqslant \lambda^*$ then $\overline{V}(\mu) \geqslant V^1(\mu)$, otherwise $\overline{V}(\mu) \leqslant V^2(\mu)$. First of all, suppose $\sup_{\mathcal{A}\in\mathbb{A}} I(\mathcal{A}; \mathcal{X}|\mu) \geqslant \lambda^*$, then by definition of $\sup_{\mathcal{A}\in\mathbb{A}} I(\mathcal{A}; \mathcal{X}|\mu)$ there exists $\left\{\mathcal{A}_j^i\right\}$ s.t:

$$\begin{aligned}
&E[u(\mathcal{A}_j^i, \mathcal{X})] - \left(\frac{m}{\lambda^*} + \frac{f(\lambda^*)}{\lambda^*}\right) I\left(\mathcal{A}_j^i, \mathcal{X}|\mu\right) \geqslant V^1(\mu) - \frac{1}{i} \\
&I(\mathcal{A}_j^i; \mathcal{X}|\mu) \geqslant \lambda^* - \frac{1}{j} - \frac{1}{i} \\
\Longrightarrow &\begin{cases} E[u(\mathcal{A}_i^i, \mathcal{X})] - \dfrac{m+f(\lambda^*)}{\lambda^*} I\left(\mathcal{A}_i^i; \mathcal{X}|\mu\right) \to V^1(\mu) \\ I(\mathcal{A}_i^i; \mathcal{X}|\mu) \to \lambda^* \end{cases} \\
\Longrightarrow &V^2(\mu) \geqslant E[u(\mathcal{A}_i^i, \mathcal{X})] - m - f\left(I\left(\mathcal{A}_i^i; \mathcal{X}\middle|\mu\right)\right) \\
&= E[u(\mathcal{A}_i^i, \mathcal{X})] - \frac{m+f(\lambda^*)}{\lambda^*} I\left(\mathcal{A}_i^i; \mathcal{X}|\mu\right) + \left(\frac{m + f\left(I\left(\mathcal{A}_i^i; \mathcal{X}\middle|\mu\right)\right)}{I\left(\mathcal{A}_i^i; \mathcal{X}\middle|\mu\right)} - \frac{m+f(\lambda^*)}{\lambda^*}\right) I\left(\mathcal{A}_i^i; \mathcal{X}\middle|\mu\right) \\
&\to V^1(\mu) \\
\Longrightarrow &V^2(\mu) = \overline{V}(\mu)
\end{aligned}$$

Now suppose $\sup_{\mathcal{A}\in\mathbb{A}} I(\mathcal{A}; \mathcal{X}|\mu) < \lambda^*$. Assume by contradiction that $\overline{V}(\mu) > V^2(\mu)$. Then I first claim that $\forall \mathcal{A}^i$ solving Equation (D.6), $\limsup I(\mathcal{A}^i; \mathcal{X}|\mu) \leqslant \lambda^*$. If this claim is true, then there is immediately a contradiction:

$$\begin{cases} \lim I(\mathcal{A}^i;\mathcal{X}|\mu) = \lambda^* \\ \lim E[u(\mathcal{A}^i,\mathcal{X})] - g\Big(I\Big(\mathcal{A}^i;\mathcal{X}|\mu\Big)\Big) = \overline{V}(\mu) \end{cases}$$
$$\implies \lim E[u(\mathcal{A}^i,\mathcal{X})] - g(\lambda^*) = \overline{V}(\mu)$$
$$\implies \lim E[u(\mathcal{A}^i,\mathcal{X})] - f\Big(I(\mathcal{A}^i;\mathcal{X}|\mu)\Big) = \overline{V}(\mu) > V^2(\mu)$$

Suppose the claim is not true, then $\overline{V}(\mu) < V^1(\mu)$ and there exists:

$$\begin{cases} \lim I(\mathcal{A}^i_1;\mathcal{X}|\mu) = \lambda' > \lambda^* \\ \lim E[u(\mathcal{A}^i_1,\mathcal{X})] - \dfrac{m+f(\lambda^*)}{\lambda^*} I\Big(\mathcal{A}^i_1;\mathcal{X}|\mu\Big) = \overline{V}(\mu) \end{cases}$$
$$\begin{cases} \lim I(\mathcal{A}^i_2;\mathcal{X}|\mu) = \lambda'' < \lambda^* \\ \lim E[u(\mathcal{A}^i_2,\mathcal{X})] - \dfrac{m+f(\lambda^*)}{\lambda^*} I\Big(\mathcal{A}^i_2;\mathcal{X}|\mu\Big) = V^1(\mu) \end{cases}$$

$\forall \alpha \in [0,1]$ consider compound experiment: $\mathcal{S}^0$ is an unrelated random draw with outcome 1 with probability $1-\alpha$ and 2 with α. Conditional on 1, do experiment $\mathcal{A}^i_1$ and follow recommendation. Otherwise do $\mathcal{A}^i_2$ and follow recommendation. Call this information structure $\mathcal{A}^i_\alpha$. Then Assumption 3.1 implies:

$$I(\mathcal{A}^i_\alpha;\mathcal{X}|\mu) \leqslant (1-\alpha) I(\mathcal{A}^i_1;\mathcal{X}|\mu) + \alpha I(\mathcal{A}^i_2;\mathcal{X}|\mu)$$

Since $l' > \lambda^* > \lambda''$, WLOG we can assume $I(\mathcal{A}^i_j;\mathcal{X}|\mu)$ is bounded within λ',λ'' by ε and $2\varepsilon < \lambda' - \lambda^*$. Now consider the utlity of strategy $\mathcal{A}^i_\alpha$ in Equation (D.6). Suppose $I(\mathcal{A}^i_\alpha;\mathcal{X}|\mu) < \lambda^*$ for all $\alpha > 0$, then:

$$\lim_{\alpha \to 0} E\Big[u(\mathcal{A}^i_\alpha;\mathcal{X})\Big] - g\Big(I(\mathcal{A}^i_\alpha;\mathcal{X}|\mu)\Big)$$

$$\begin{aligned}
&\geqslant E\Big[u\Big(\mathcal{A}_1^i\Big);\mathcal{X}\Big] - g(\lambda^*) \\
&\geqslant \overline{V}(\mu) + \big(g(\lambda' - \varepsilon) - g(\lambda^*)\big) - \frac{1}{i}
\end{aligned}$$

Since g is a strictly increasing function with $\lambda > \lambda^*$, given any $\delta < g(\lambda' - \varepsilon) - g(\lambda^*)$, there exists α^i s.t.

$$E\Big[u(\mathcal{A}_{\alpha^i}^i;\mathcal{X})\Big] - g\Big(I(\mathcal{A}_{\alpha^i}^i;\mathcal{X}|\mu)\Big) \geqslant \overline{V}(\mu) - \frac{1}{i} + \delta$$

Suppose there exists α^i s.t. $I(\mathcal{A}_{\alpha^i}^i;\mathcal{X}|\mu) = \lambda^*$, then:

$$\begin{aligned}
&E\Big[u(\mathcal{A}_{\alpha^i}^i;\mathcal{X})\Big] - g\Big(I(\mathcal{A}_{\alpha^i}^i;\mathcal{X}|\mu)\Big) \\
=&E\Big[u(\mathcal{A}_{\alpha^i}^i;\mathcal{X})\Big] - \frac{m + f(\lambda^*)}{\lambda^*} I(\mathcal{A}_{\alpha^i}^i;\mathcal{X}|\mu) \\
\geqslant&\overline{V}(\mu) + \alpha(V^1(\mu) - \overline{V}(\mu)) - \frac{1}{i} \\
&+ \frac{m + f(\lambda^*)}{\lambda^*}\Big((1-\alpha)I(\mathcal{A}_1^i;\mathcal{X}|\mu) + \alpha I(\mathcal{A}_2^i;\mathcal{X}|\mu) - I(\mathcal{A}_{\alpha^i}^i;\mathcal{X}|\mu)\Big) \\
\geqslant&\max\begin{cases} \overline{V}(\mu) + \alpha(V^1(\mu) - \overline{V}(\mu)) - \frac{1}{i} \\ \overline{V}(\mu) - \frac{1}{i} + \frac{m + f(\lambda^*)}{\lambda^*}\big((1-\alpha)(\lambda' - \varepsilon) + \alpha(\lambda'' - \varepsilon) - \lambda^*\big) \end{cases} \\
=&\overline{V}(\mu) - \frac{1}{i} + \max\Big\{\alpha(V^1(\mu) - \overline{V}(\mu)), \frac{m + f(\lambda^*)}{\lambda^*}\big(\lambda' - \alpha(\lambda' - \lambda'') - \varepsilon - \lambda^*\big)\Big\}
\end{aligned}$$

The maximum is independent to i and strictly positive for any α. Therefore:

$$\lim_{i\to\infty} E\Big[u(\mathcal{A}_{\alpha^i}^i;\mathcal{X})\Big] - g\Big(I(\mathcal{A}_{\alpha^i}^i;\mathcal{X}|\mu)\Big) > \overline{V}(\mu)$$

Contradicting optimality of $\overline{V}(\mu)$. To sum up, I show that when $\sup_{\mathcal{A}\in\mathbb{A}} I(\mathcal{A};\mathcal{X}|\mu) < \lambda^*$,

$\overline{V}(\mu) = V^2(\mu)$. Therefore:

$$\overline{V}(\mu) = \begin{cases} V^1(\mu) & \text{if } \sup_{\mathcal{A}\in\mathbb{A}} I(\mathcal{A};\mathcal{X}|\mu) \geqslant \lambda^* \\ V^2(\mu) & \text{if } \sup_{\mathcal{A}\in\mathbb{A}} I(\mathcal{A};\mathcal{X}|\mu) < \lambda^* \end{cases}$$

Case 2. $\lambda^* = +\infty$. By definition of λ^*, $\left(\frac{m}{\lambda} + \frac{f(\lambda)}{\lambda}\right)$ is strictly decreasing in λ. $\forall \mathcal{A}, \lambda$ being feasible in Equation (D.5), it can be improved by replacing λ with $I(\mathcal{A};\mathcal{X}|\mu)$ (feasibility is still satisfied). Therefore, it is without loss of optimality to assume constraint binding and Equation (D.5) becomes:

$$\sup_{\mathcal{A}} E[u(\mathcal{A},\mathcal{X})] - m - f(I(\mathcal{A};\mathcal{X}|\mu))$$

which is exactly Equation (3.3). ∎

D.2.3 *Proof of Proposition 3.5*

Proof. *Existence:* Equations (3.2) and (3.3) can be solved prior by prior. Therefore, I sometimes don't explicitly include prior any more in this proof. It's not hard to see that it's sufficient to prove existence of solution to:

$$\sup_{\mathcal{A}} E[u(\mathcal{A},\mathcal{X}) - f(I(\mathcal{A};\mathcal{X}|\mu))] \tag{D.7}$$

where $\mathcal{A} \in \Delta A \times X$ and f is convex. Equation (D.7) can be modified to accommodate Equation (3.2) by setting f to be a linear function. This can be WLOG directly modeled by changing the information measure I. Equation (D.7) is different from Equation (3.3) by only a constant. Therefore, it is sufficient to show existence of solution to Equation (D.7) under Assumption 3.1.

Next let's explicitly model the set of all feasible $\mathcal{A}$'s as Markovian transition matrices:

$\Delta A \times X \in \mathbb{R}^{(|A|-1)\times|X|}$. Let's call this set Λ and any conditional distribution $p(a|x) \in \Lambda$. We define $\widetilde{I} : \Lambda \to \mathbb{R}^+$:

$$\widetilde{I}(p(\cdot|\cdot)) = I(\gamma)$$

where $\gamma = (\pi, \mu) \in \Gamma$ and π is defined by distribution of posteriors induced by p:

$$\begin{cases} \mu_s(x) = \frac{p(s|x)\mu(\mu)}{\sum_y p(s|y)\mu(\mu)} \\ \pi(\mu_s) = \sum_y p(s|y)\mu(y) \end{cases}$$

Our original problem Equation (D.7) can be written as:

$$\sup_{p\in\Lambda} \sum_{a,x} p(a|x)\mu(x)u(a,x) - f\left(\widetilde{I}(p)\right)$$

To prove Proposition 3.5, it is sufficient to show the convexity of $\widetilde{I}$. If $\widetilde{I}$ is convex, the objective function is continuous in p on the interior of Λ and any space Λ_ε is compact (a closed and bounded set in Euclidean space). Now let's study the convexity of $\widetilde{I}$. Consider $\forall\ p_1, p_2 \in \Lambda$. Let $p = \lambda p_1 + (1-\lambda)p_2$. It's not hard to verify that $p \in \Lambda$ as well. Want to show:

$$\widetilde{I}(p) \leqslant \lambda\widetilde{I}(p_1) + (1-\lambda)\widetilde{I}(p_2)$$

Now define p' on $A \times \{1,2\} = \{a_1, a_2, \dots\}$ with twice number of signals than A. Let $\lambda_1 = \lambda, \lambda_2 = 1-\lambda, \forall a, x$

$$p'(a_i|x) = \lambda_i p_i(a|x)$$

Then p' will be Blackwell more informative than p:

$$\begin{bmatrix} 1\,1 & 0\,0 & \cdots & 0\,0 \\ 0\,0 & 1\,1 & \cdots & 0\,0 \\ \vdots & \vdots & \ddots & \vdots \\ 0\,0 & 0\,0 & \cdots & 1\,1 \end{bmatrix} \cdot p' = p \tag{D.8}$$

On the other hand, p' can be written as combination of p_1 and p_2. Let $\mathcal{S}_0$ be randomly $\{1,2\}$ with probability λ_1, λ_2. Let $(\mathcal{S}_1|1,\mu) \sim p_1$ and $(\mathcal{S}_1|2,\mu) \sim p_2$. Then it's easy to see that $(\mathcal{S}_0, \mathcal{S}_1, \mu) \sim p'$. Therefore:

$$\begin{aligned}
\widetilde{I}(p) \leqslant & \widetilde{I}(p') \\
= & I(\mathcal{S}_0, \mathcal{S}_1; \mathcal{X}|\mu) \\
\leqslant & I(\mathcal{S}_0; \mathcal{X}|\mu) + \lambda_1 I(\mathcal{S}_1|1; \mathcal{X}|\mu) + \lambda_2 I(\mathcal{S}_1|2; \mathcal{X}|\mu) \\
= & \lambda \widetilde{I}(p_1) + (1-\lambda)\widetilde{I}(p_2)
\end{aligned}$$

First inequality is from monotonicity, second inequality is from sub-additivity. Therefore $\widetilde{I}$ is a convex (and continuous) function. It's easy to see that Λ is a compact set. So we can apply Weierstrass theorem to conclude existence of solution.

Now suppose p_1,p_2 are two distinct maximizer. Consider $p = \alpha p_1 + (1-\alpha)p_2$. By convexity of $\widetilde{I}$ and f:

$$\begin{aligned}
E_\mu[u(a,x)p(a,x)] = & \alpha E_\mu[u(a,x)p_1(a,x)] + (1-\alpha)E_\mu[u(a,x)p_2(a,x)] \\
f\left(\widetilde{I}(p)\right) \leqslant & \alpha f\left(\widetilde{I}(p_1)\right) + (1-\alpha) f\left(\widetilde{I}(p_2)\right)
\end{aligned}$$

Therefore p weakly dominates p_1 and p_2 and $p \in \mathbb{A}$. $\mathbb{A}$ is convex.

Uniqueness: Now suppose I also satisfies strict-monotonicity. Then consider proof in last section. First, let $p_1 \neq p_2$. Suppose equality $\widetilde{I}(p) = \widetilde{I}(p')$ holds, then strict-monotonicity implies that p is Blackwell sufficient for p':

$$M \cdot p = p'$$

Where M is a stochastic matrix. Consider the following operation: If $p'_1 \sim p'_2$, then proof is done. Otherwise, first remove replication of p' (when two rows of p' are mutiplications of each other, then add them up) and get $\widetilde{p}'$. Since $p'_1 \nsim p'_2$, we can assume $\widetilde{p}'_1 = p'_1, \widetilde{p}'_2 = p'_2$. Define $\widehat{p}_1 = p'_1 + p'_2$ and $\widehat{p}_i = \widetilde{p}'_{i+1}$. By definition $\widetilde{p}'$ Blackwell dominates $\widehat{p}$. On the other hand, $\widehat{p}$ Blackwell dominates p, so dominates p', and $\widetilde{p}'$. By Lemma D.2, $\widetilde{p}'$ and $\widehat{p}$ are identical up to permutation. Then p'_1 must equal to some $\widehat{p}_i$.

- *Case 1*. If $i = 1$, then $p'_1 + p'_2$ is a multiplication of $\widetilde{p}'^1_1$, which is a multiplication of $\widetilde{p}'_1$. This means p'_1 and p'_2 are replication, contradiction.

- *Case 2*. If $i > 1$, then $\widetilde{p}'_1$ is a multiplication of $\widehat{p}_i$, which is a multiplication of $\widetilde{p}'_{i+1}$. Contradicting definition of $\widetilde{p}'$.

Therefore, p'_1 and p'_2 are replications. Now permute p' and apply the same analysis on all p'_{2i-1}, p'_{2i}. We can conclude that any row of p_1 is a replication of that of p_2. To sum up, a necessary condition for $\widetilde{I}(p) = \alpha\widetilde{I}(p_1) + (1-\alpha)\widetilde{I}(p_2)$ is that each row in p_1 and p_2 induces same posterior belief ν.

Now consider $\mathbb{A}$ being set of solutions to Equation (3.1). Suppose by contradiction there exists $\mathcal{A}_1$ and $\mathcal{A}_1$ and a such that they induces different posterior with realization a. Let p_1, p_2 be corresponding stochastic matrices, consider any $\mathcal{A} \sim \alpha p_1 + (1-\alpha)p_2$. By previous proof, $I(\mathcal{A}; \mathcal{X}|\mu) < \alpha\widetilde{I}(p_1) + (1-\alpha)\widetilde{I}(p_2)$. In first part, we show that $\mathbb{A}$ is convex, so $\mathcal{A}$ is feasible. This contradicts unimprovability.

To sum up, solutions to Equation (D.7) always have the same support. Of course if $\mathcal{A}$

is uninformative, then it induces prior μ. In both case, support of posteriors is uniquely determined.

■

Lemma D.2 (Blackwell equivalence). *Let P and P' be two stochastic matrices. P has no replication of rows. Suppose there exists stachatic matrices $M_{PP'}$ and $M_{P'P}$ s.t.:*

$$P' = M_{P'P} \cdot P$$

$$P = M_{PP'} \cdot P'$$

Then $M_{PP'}$ and $M_{P'P}$ are permutation matrices.

Proof. Let $P_i = (p_{i1}, p_{i2}, \ldots)$ be ith row of P. Suppose P_i can not be represented as positive combination of P_{-i}'s. Then by construction $P_i = M_{PP'i} \cdot M_{P'P} \cdot P$, we have:

$$M_{PP'i} \cdot M_{P'P} = (\underbrace{0, \cdots, 0}_{i-1}, 1, 0, \cdots, 0)$$

Then by non-negativity of stochastic matrices, suppose $M_{PP'ij} > 0$, then $M_{P'Pj}$ are all 0 except $M_{P'Pji}$. Then for all such rows j, we have $M_{PP'j}$ be a vector with only ith column being non-zero. However this suggests they are replicated rows. So the only possibility is that j s.t. $M_{PP'ij} > 0$ is unique. And

$$M_{PP'ij} \times M_{P'Pji} = 1$$

Since stochastic matrices have elements no larger than 1, it must be $M_{PP'ij} = M_{P'Pji} = 1$. This is equivalently saying $P'_j = P_i$. Since permutation of rows of P' doesn't affect our statement, let's assume $P'_i = P_i$ afterwards for simplicity.

So far we showed that if P_i is not a positive combinations of P_{-i}'s, then $P'_i = P_i$. We do the folloing transformation: $\widetilde{P}, \widetilde{P}'$ are P, P' removing ith row. $\widetilde{M}_{PP'}, \widetilde{M}_{P'P}$ are $M_{PP'}, M_{P'P}$ removing ith row and column. It's easy to verify that we still have:

$$\widetilde{P}' = \widetilde{M}_{P'P} \cdot \widetilde{P}$$
$$\widetilde{P} = \widetilde{M}_{PP'} \cdot \widetilde{P}'$$

and $\widetilde{M}_{PP'}, \widetilde{M}_{P'P}$ still being stochastic matrices since previous argument shows $M_{PP'ii}$ and $M_{P'Pii}$ being the only non-zero element in their rows. Since they are both 1, they must also be only non-zero element in their columns. So removing them doesn't affect the matrices being stochatic matrices.

Now we can repeat this process iteratively until any row $\widetilde{P}_i$ will be a positive combination of $\widetilde{P}_{-i}$. If $\widetilde{P}$ has one unique row, then the proof is done. We essentially showed that $P = P'$ (up to permutation of rows). Therefore we only need to exclude the possibility of $\widetilde{P}$ having more than one rows.

Suppose $\widetilde{P}$ has n rows. Then $\widetilde{P}_1$ is a posistive combination of $\widetilde{P}_{-i}$'s:

$$\widetilde{P}_1 = \sum_{i=2}^{n} a_i^1 \widetilde{P}_i$$

and $\widetilde{P}_2$ is a positive combination of $\widetilde{P}_{-i}$'s:

$$\begin{aligned}
\widetilde{P}_2 =& \sum_{i \neq 2}^{n} a_i^2 \widetilde{P}_i \\
=& a_1^2 \widetilde{P}_1 + \sum_{i>2}^{n} a_i^2 \widetilde{P}_i \\
=& a_1^2 a_2^1 \widetilde{P}_2 + \sum_{i>2}^{n} \left(a_i^2 + a_1^2 a_i^1 \right) \widetilde{P}_i
\end{aligned}$$

Since all rows in $\widetilde{P}$ are non-negative (and strictly positive in some elements). This is possible only in two cases:

- *Case 1.* $a_1^2 a_2^1 = 1$ and $\sum_{i>2}(a_i^2 + a_1^2 a_i^1) = 0$. This implies $\widetilde{P}_1 = a_2^1 \widetilde{P}_2$. Contradicting non-replication.

- *Case 2.* $a_1^2 a_2^1 < 1$. Then $\widetilde{P}_2$ is a positive combination of $\widetilde{P}_{i>2}$. Of course $\widetilde{P}_1$ is also a positive combination of $\widetilde{P}_{i>2}$.

Now by induction suppose $\widetilde{P}_1, \ldots, \widetilde{P}_i$ are positive combinations of $\widetilde{P}_{j>i}$. Then:

$$\begin{aligned}
\widetilde{P}_{i+1} &= \sum_{j=1}^{i} a_j^{i+1} \widetilde{P}_j + \sum_{j=i+1}^{n} a_j^{i+2} \widetilde{P}_j \\
&= \sum_{k=i}^{n} \left(\sum_{j=1}^{i} a_j^{i+1} a_k^j \right) \widetilde{P}_k + \sum_{j=i+2}^{n} \widetilde{P}_j \\
&= \sum_{j=1}^{i} a_j^{i+1} a_{i+1}^j \widetilde{P}_{i+1} + \sum_{k=i+2}^{n} \left(\sum_{j=1}^{i} a_j^{i+1} a_k^j + a_j^{i+1} \right) \widetilde{P}_j
\end{aligned}$$

Similar to previous analysis, non-replication implies $\sum_{j=1}^{i} a_j^{i+1} < 1$ and $\widetilde{P}_{i+1}$ is a positive combination of $\widetilde{P}_{j>i+1}$. Then by replacing $\widetilde{P}_{i+1}$ in combination of all $\widetilde{P}_{j\leqslant i}$, we can conclude that $\widetilde{P}_1, \ldots, \widetilde{P}_{i+1}$ are all positive combinations of $\widetilde{P}_{j>i+1}$. Finally, by induction we have all $\widetilde{P}_{i<n}$ being positive combination of $\widetilde{P}_n$. However, this contradicts non-replication. To sum up, we proved by contradiction that $\widetilde{P}$ has one unique row. Therefore, P must be identical to P' up to permutations. ■