Exploration and Stopping

Yuliy Sannikov*   Weijie Zhong†

October 21, 2024

Abstract

We fully characterize the possible outcomes of exploration and stopping: all state-time joint distributions achieved by stopping some martingale process with bounded variation. Utilizing this characterization, we provide a general methodology for solving an optimal exploration-stopping problem where the stopping payoff depends on state and time arbitrarily. We reveal the close relation between the pattern of exploration and time preference and apply it to study competitive exploration contests.

1 Introduction

Many economic problems involve exploration as well as a stopping decision. The payoff of the decision-maker can depend on the time of the decision, as well as the state of knowledge at that time. The extensive literature on real options emphasizes the importance of the stopping problem. For example, the classic book of Dixit and Pindick (1994) focuses on the timing of investment decisions under the assumption of exogenous information arrival. Yet, many applications involve active exploration, which often includes the choice of the type of information to acquire.

Dynamic exploration problems are complicated because the decision of what type of information to acquire can depend on the information already obtained. This paper sets off with a result which significantly simplifies these problems. Specifically, for a class of natural constraints on the rate of learning, we present a simple condition that fully characterizes the attainable joint distributions over stopping times and the state of knowledge at the stopping time, termed in shorthand the *state-time distribution*. Thus, when payoffs depend on the stopping time as well as the information available at that time, instead of solving the dynamic problem, we can simply pick the optimal state-time

*Stanford Graduate School of Business
†Stanford Graduate School of Business

distribution. Our result guarantees the existence of a dynamic exploration strategy that attains the desired distribution, as well as that no other strategy can attain a distribution outside the consideration set.

To appreciate the power of the result, consider two competitive firms that test candidate technologies in order to launch a new product. They could perform various experiments that glean different types of information. Some experiments may focus on reliability, yielding information in the form of observed potential failures. Some experiments may yield better estimates of the efficiency of the technology in the form of performance data in various circumstances. The information collected affects the success of the product, and so does the timing. Specifically, because the distribution of times that the competitor launches the product matters, the problem may be more complicated than one of exponential discounting. Each firm wants to know the attainable information given any distribution of the competitors' decision times, and our result provides exactly that.

Outline of contribution. In our general framework of exploration and stopping, we model the exploration strategy of a decision-maker (DM) as choosing a martingale process. Exploration is flexible in that any martingale process is feasible as long as it satisfies a bound on its flow variation.

Our first main result gives a complete characterization of the state-time distributions that are **embeddable**, i.e., they are the joint distributions of the stopping state and stopping time corresponding to some feasible martingale process and stopping time. We show that a state-time distribution is embeddable if and only if, in each period, a simple inequality condition holds: the *expected variation* of the stopped state plus the *variation of the expected* stopped state in the future is less than the cumulative variation bound up to the period (Theorem 1). The condition has a natural interpretation that in any period, the amount of knowledge that has been exploited plus the amount of knowledge that has been explored but not yet exploited must be less than the cumulative capacity of exploration.

Then, we consider a general exploration-stopping problem, where the DM controls the exploration strategy and stopping time to maximize the expected stopping payoff. The DM's stopping payoff depends arbitrarily on the state of the martingale and time. The embedding theory reduces the general exploration-stopping problem to a semi-static problem where the DM directly chooses the optimal embeddable state-time distribution—a simple linear program.

Our second main result provides a unified methodology for analytically solving the reduced problem. We establish a strong duality of the linear program and derive a necessary and sufficient first-order characterization of the optimal policy (Theorem 2). The constrained optimization problem is equivalent to an unconstrained dual problem where there exists a “price” (multiplier) for every period at which the DM can “buy” or “sell” information. The first-order condition states that the optimal state-time distribution concavifies (attains the upper tangent hyperplane of) a combination of the payoff function and the shadow cost/benefit of information. Then, solving the optimal exploration-stopping problem boils down to solving a single-dimensional ordinary differential equation characterizing the “prices”. In various applications, we illustrate the tremendous analytical tractability of the methodology.

Third, we derive several general implications of the optimal exploration-stopping problem. We show that a strategy with coarse support can always solve it: the support of the stopped state at each time contains a bounded number of points. When the payoff function is convex in time, we show that the optimal exploration process resembles a Poisson process that either drifts along a deterministic path or jumps into the stopping region. Conversely, when the payoff function is concave in time, the optimal exploration process necessarily involves “pure exploration” at the beginning, i.e., exploration without any immediate stopping.

Economic applications. We apply our methodology to develop tractable models for economic applications. Our first application revisits the canonical real options problem but with active and flexible exploration. We use the application as a minimal working example to explain the critical machinery of our model, establishing a connection with the recursive approach that has been almost exclusively used in the literature.

Our second application is a canonical information acquisition problem: a DM chooses a signal process with bounded informativeness to learn a binary payoff-relevant state and solves a decision-making problem upon stopping. We characterize the optimal information acquisition strategy for general discount functions. The first result reveals the connection between the risk preference toward time lotteries and the pattern of optimal exploration. Specifically, we show that the optimal exploration policy alternates between two types of strategies:

- *Pure exploration*: during a period of pure exploration, the DM’s interim belief becomes more dispersed but never sufficient to induce stopping and making the decision. A pure exploration period always ends in a region where the (adjusted) discount function is concave (indicating time-risk aversion).

- *Full exploitation*: during a period of full exploitation, the DM’s belief jumps according to a Poisson process, and the DM stops immediately upon the jump of the belief. The continuing belief remains degenerate and constant. An exploitation period typically ends in a region where the (adjusted) discount function is convex (indicating time-risk loving).

The second result reveals the connection between the discount rate's evolution and the decision's quality. We consider a setting where full exploitation is optimal and quantify the decision quality measured by the distance between the decision belief and the prior. We show that convexly decreasing decision quality over time implies an increasing discount rate. Vice versa, convexly increasing decision quality over time implies a decreasing discount rate. The decision quality is constant if and only if the discount function resembles standard exponential discounting.

Our third application models the competitive R&D example we introduced earlier. We model a continuous-time contest in which n contestants independently and privately choose their exploration strategies, i.e., each of them chooses a martingale process and a stopping time. The distance between the stopped state and the initial state represents the quality of the research. They compete in the time dimension but also value the quality dimension: the contestant who stops the first collects a reward that depends on the quality of his research. The remaining contestant gets nothing. We provide a complete characterization of all pure strategy equilibria of the contest, showing that all equilibria are symmetric and exhibit endogenous time-risk loving induced by competition: in any pure strategy equilibrium of the game, all contestants use the same Poisson exploration process, leading to an effective discount factor that is convex in time.

Related literature

The optional stopping problem in our paper resembles the canonical sequential sampling problem (see, e.g., Arrow, Blackwell, and Girshick 1949; Wald 1947) and real options problem (see, e.g., Dixit and Pindyck 1994). We merged the stopping problem with flexible active exploration, providing by far the most general solution to optimal exploration and stopping problems. Our framework fully nests Zhong 2022, D. Chen and Zhong 2024, and the majority of Hébert and Woodford 2023, each of which focuses on a specific time preference and payoff structure and makes sharply different predictions.[1] The generality of our method allows us to obtain a complete characterization of how time preference determines the optimal pattern of exploration, unifying and generalizing the existing results. A closely related but not nested paper is Georgiadis-Harris 2024, where the stopping time is *exogenous* and the pure exploration policy is optimal.

[1] Zhong 2022 studies the case with exponential discounting. Hébert and Woodford 2023 studies both exponential discounting and fixed waiting cost. D. Chen and Zhong 2024 studies a one-dimensional setting with a fixed stopping threshold and general convex / concave time preference.

Our key technical innovation is a novel embedding theory: the characterization of all state-time distributions that can be implemented by stopping some martingale process with bounded variation. It could be viewed as extending the celebrated Skorokhod's embedding (Skorokhod 1982) to general martingale processes and the state-time product space. An extensive literature on stochastic analysis attempted to generalize Skorokhod's embedding to general stochastic processes (see, e.g., Obłój 2004 for a survey). Our approach differs from all these papers by embedding not only the distribution of states but the state-time joint distribution.

We show that strictly convex discount functions always lead to a Poisson exploration strategy where the state drifts deterministically or jumps to the stopping region, justifying the Poisson learning models adopted by papers on sequential sampling (see, e.g., Che and Mierendorff 2019, Mayskaya 2022 and Nikandrova and Pancs 2018). Strictly concave discount functions, on the other hand, always lead to pure exploration without stopping. The "pure accumulation" policy with deterministic stopping in D. Chen and Zhong 2024 is a special case of ours with binary state, threshold stopping rule and additively separable time preference. Pure exploration is a common model for studying the timing of innovation (see, e.g., Dasgupta and Stiglitz 1980, Lee and Wilde 1980 and Reinganum 1989). The result connects optimal exploration to the recent literature on the risk preference towards time lotteries (See, e.g., Chesson and Viscusi 2003, M. K. Chen 2013, Onay and Öncüler 2007 and DeJarnette et al. 2020).

Our second application studies the speed-accuracy tradeoff in dynamic exploration, which has been previously studied almost extensively using the drift-diffusion models (DDM) of binary choice problems (see, e.g., Fudenberg, Newey, et al. 2020; Fudenberg, Strack, and Strzalecki 2018; Ratcliff and Rouder 1998). Our exploration-stopping model provides a novel optimization foundation for the speed-accuracy complementarity and substitutability based on whether the DM's discount rate is increasing or decreasing over time.

Our third application is closely related to the literature on dynamic contests. Seel and Strack 2013, 2016 introduced the dynamic contest framework where contestants compete in the states of stopped Brownian motions. Several papers have extended this framework to allow for more general processes, prize structure, and preferences (see Feng and Hobson 2015, 2016a,b; Nutz and Zhang 2022). In another strand of literature, Anderson, L. Smith, and Park 2017; Park and L. Smith 2008 study the timing game where contestants compete in the stopping time. Our exploration contest merges the two approaches and provides a framework where the contest's prize depends on both the stopping state and the stopping time.

The rest of the paper is organized as follows. Section 2 addresses the question of attainable state-time distributions. Section 3 demonstrates the power of this result by solving the optimal exploration-stopping problem. Section 4 presents several applications.

2 Exploration & stoppinig : the embedding theory

In this section, we study the feasible outcomes in a problem of dynamic exploration. Closed subset T of $\mathbb{R}_+$ captures our timeline, which can be continuous or discrete. We assume that $0 \in T$ and that T contains at least two elements.

The state μ_t, $t \in T$, is a martingale with domain in a convex compact set $S \subset \mathbb{R}^n$. For example, the state could be interpreted as the belief of the DM about a binary "state of the world", in which case S is the interval $[0,1]$ Starting from point $\mu_0 \subset S$, the DM chooses the exploration strategy that determines the evolution of μ_t.

There is a rich set of exploration strategies, which lead to different laws of motion of the state μ_t. Any strategy is admissible as long as it satisfies the following restriction. Specifically, assume that there exists a strongly convex and continuous function $H : S \to \mathbb{R}$ and constant $\chi > 0$, such that cadlag martingale $\langle \mu_t \rangle_{t \in T}$ in S is *admissible* if and only if satisfies the *variation constraint*

$$\mathbb{E}\left[H(\mu_{t'}) - H(\mu_t) \,\middle|\, \mathcal{F}_t\right] \leq \chi(t' - t) \tag{1}$$

for all $t', t \in T$, $t' > t$. Condition (1) is a common constraint in information economics (See Zhong 2022, Hébert and Woodford 2023 and Georgiadis-Harris 2024). It captures the idea that there are many choices of how to explore - e.g., via Poisson or Brownian signals or a combination thereof - but there is a constraint on how quickly one can learn. Function H provides an appropriate measure of information received and χ, the rate of information arrival. The assumption nests familiar constraints like quadratic variation bound (when H is quadratic) and mutual information rate bound (when H is Shannon's entropy). We extend our analysis to allow for endogenous capacity constraint in Section 3.2.

Formally, denote by $(\Omega, \mathcal{F}, \mathcal{P})$ the underlying probability space.[2] The decision of when to stop exploration is captured by the stopping time τ w.r.t. the filtration $\langle \mathcal{F}_t \rangle_{t \in T}$. Let $\mathcal{M}$ denote the collection of all admissible pairs of $(\langle \mu_t \rangle, \tau)$. We are interested in joint probability measures over pairs (μ_τ, τ) attainable by some admissible state processes μ_t with some stopping time τ., i.e. the set

$$\mathbb{F} = \left\{ f \in \Delta(S \times T) \,\middle|\, \exists (\langle \mu_t \rangle, \tau) \in \mathcal{M} \text{ s.t. } f \sim (\mu_\tau, \tau) \right\}.$$

[2] Because exploration involves the choice of the type of information to receive, exploration strategy defines the probability space together with the process μ_t on it.

$\mathbb{F}$ is called the set of *embeddable* state-time distributions. We are now ready to present our first result: the characterization of embeddable distributions. In order to provide a clean expression for the necessary and sufficient condition, we normalize S to be a subset of the probability simplex of $\mathbb{R}^{n+1}$ and extend H homogeneously (of degree 1) from S to the convex cone $\{\alpha \cdot \mu \,|\, \alpha \in \mathbb{R}_+, \mu \in S\}$.[3]

Theorem 1 (Martingal Embeddings). *$f \in \mathbb{F} \iff \mathbb{E}_f[\mu] = \mu_0$ and $\forall t \in T$, f satisfies:*

$$\int_{\tau \le t} H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) + H\left(\int_{\tau > t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau) \right) - H(\mu_0) \le \chi \cdot \int \min\{t, \tau\} f(\mathrm{d}\mu, \mathrm{d}\tau) \tag{2}$$

Proof. See Appendix A. *Q.E.D.*

Let us interpret condition (2). The DM explores and then stops. The left-hand side captures the minimal information required to generate the portion of the distribution f over time interval $[0, t]$. The right-hand side captures the total information received over $[0, t]$, until stopping.

To see this in greater detail, define

$$\widehat{\mu_t} \equiv \begin{cases} \mu_\tau & \text{if } \tau \le t, \\ E[\mu_t \,|\, \tau > t] & \text{if } \tau > t. \end{cases}$$

Then, process $\widehat{\mu_t}$ contains weakly less information than process μ_t, because it does not refine the knowledge that μ_t contains in the event that $\tau > t$. Information contained in $\widehat{\mu_t}$ cannot exceed total information obtained until time t, hence we have the following necessary condition

$$E[H(\widehat{\mu_t})] - H(\mu_0) \le \chi E[\min(t, \tau)].$$

This inequality is equivalent to (2).

Theorem 1 states that condition (2) is not only necessary but sufficient, i.e., there exists a pair (μ_t, τ) that gives rise to distribution f. If equality in (2) held at all times, then $\widehat{\mu_t}$ would not only achieve distribution f, but also satisfy condition (1). If not, the proof of Theorem 1 constructs process μ_t that embeds maximal obtainable information at each time point t in such a way that we can target the desired joint distribution f at all times after t.

While the formal proof of sufficiency is relegated to the appendix, here we provide a sketch of the proof and a graphical illustration when the desired distribution f (which satisfies Equation (2)) is supported on finitely many points. Figure 1a depicts the support of one such distribution with 4 discrete periods. $(\langle \mu_t \rangle, \tau)$ is constructed backward in time.

[3] The first assumption is a normalization through shifting and scaling the space. The second assumption on H outside of S is immaterial. The two normalization assumptions are innocuous and help simplify the notations.

- *Step 1*. Take the mass that stops in the last period $f(\cdot,4)$ (the red dots in Figure 1b). First, let's find a continuous-time martingale that, when stopped at $t=4$, has the same distribution as $f(\cdot,4)$ and, importantly, keeps the inequality constraint (1) binding. One such martingale can be constructed using the two paths (the dashed curves) along which $H(\mu_t)$ increases at constant rate χ. Then, $\langle\mu_t\rangle$ is the compensated Poisson process that either drifts along the current path or jumps to the other path (illustrated by the dotted arrows). Construct the process backward in time until $t=5$. The two blue dots represent the distribution of μ_t at $t=5$. Note that by construction, Equation (1) is satisfied with equality for $t\in[3,4]$.

- *Step 2*. In Figure 1b, the two blue dots constitute a mean-preserving contraction of the two red dots. Now, consider the new joint distribution that replaces the mass represented by the red dots in period 4 with that represented by the blue dots in period 3. We claim that the new distribution still satisfies Equation (2). The reason is that when moving from $t=4$ to $t=3$, the reduction of accrued capacity (RHS of Equation (2)) is $\chi\cdot f(S,4)$. Meanwhile, the total variation (LHS of Equation (2)) also reduces by $\chi\cdot f(S,4)$ since our constructed process uses exactly χ unit of variation per unit of time.

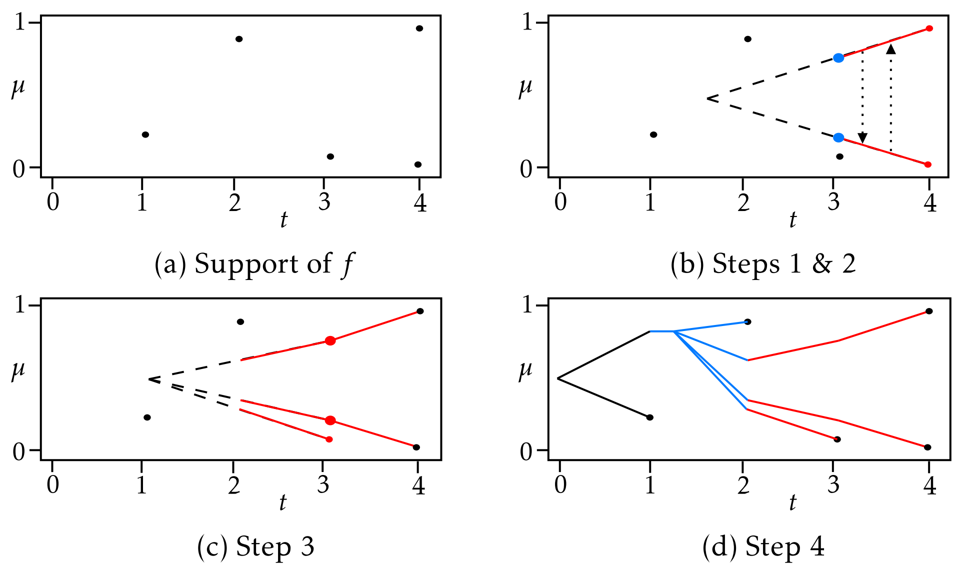
- *Step 3*. Then, we can recursively treat $t=3$ as the last period and construct a martingale that keeps the inequality constraint (1) binding for $t\in[2,3]$ and distributed according to the red dots when $t=3$. See Figure 1c.

- *Step 4*. We repeat steps 2 & 3 until $t=0$, depicted by Figure 1d. During the processes, there are two possible variants, which have been highlighted in blue and black. The first variant is period $[1,2]$, during which the constructed blue process becomes degenerate before $t=1$. In this case, we just keep it constant until $t=1$. Evidently, this means the reduction of total variation (LHS of Equation (2)) is strictly less than $\chi\cdot f(S,[2,4])$. Therefore, at $t=1$, Equation (2) holds with an even larger gap for the new distribution.

  The second variant is period $[0,1]$. Equation (2) at $t=1$ implies that the variation of the distribution is less than χ. Therefore, the process constructed backward must become degenerate before $t=0$, which guarantees that the entire process starts at μ_0 as required by $\mathcal{M}$.

2.1 Extensions of the embedding theory

The main embedding theory relies on several key features of the admissible processes: the martingale property, the inequality constraint, and the time-invariant capacity bound. In various economic applications, one or more of those features may be violated.



(a) Support of f (b) Steps 1 & 2

(c) Step 3 (d) Step 4

Figure 1: Graphical illustration of Theorem 1

In this section, we show that all these features can be relaxed via immediate corollaries of Theorem 1.

Equality constraint. If the admissible processes are defined by equality variation constraints: $\mathbb{E}\left[\frac{\mathrm{d}H(\mu_t)}{\mathrm{d}t}\middle|\mathcal{F}_t\right] = \chi$, then the embeddable state-time distributions are characterized by Equation (2) with one extra constraint:

$$\int_{S\times T} H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) - H(\mu_0) = \chi \int_{S\times T} \tau f(\mathrm{d}\mu, \mathrm{d}\tau). \tag{3}$$

The single equation (3) is sufficient to guarantee that all variation constraints are binding because it is effectively the aggregation of all the interim variation constraints.

Time-dependent variation bound. Let a bounded and strictly positive function $\chi_t : T \to \mathbb{R}_+$ be a time-dependent variation bound. We say cadlag martingale $\langle \mu_t \rangle$ is χ-admissible if and only if it satisfies

$$\mathbb{E}\left[H(\mu_{t'}) - H(\mu_t)\middle|\mathcal{F}_t\right] \le \int_t^{t'} \chi_s \mathrm{d}s \tag{4}$$

for all $t', t \in T$ and $t' > t$. Then, by transformation of the timeline via $t \to \varphi(t) = \int_{s\le t} \chi_s \mathrm{d}s$, Equation (4) is equivalent to $\mathbb{E}[H(\mu_{t'}) - H(\mu_t)] \le \varphi(t') - \varphi(t)$, i.e. $\langle \mu_{\phi^{-1}} \rangle$ is admissible with variation bound 1. Applying Theorem 1 to the transformed space immediately implies:

Corollary 1.1. *$f \in \Delta(S \times T)$ is attainable by χ-admissible process $\langle \mu_t \rangle$ and stopping time τ if and only if $\mathbb{E}_f[\mu] = \mu_0$ and $\forall t \in T$,*

$$\int_{\tau \le t} H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) + H\left(\int_{\tau > t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) - H(\mu_0) \le \int_{s \le t} \chi_s (1 - F(t)) \mathrm{d}s, \tag{5}$$

where $F(t) = \int_{\tau \le t} f(\mathrm{d}\mu, \mathrm{d}\tau)$.

Martingales with drift. Consider the collection of processes $\langle w_t \rangle$ that can be represented as some martingale process plus deterministic drift m: $w_t = \mu_t + m_t$ and satisfy Equation (1). Suppose H is a quadratic function, i.e., $H(w) = w^\top \cdot M \cdot w$ for some positive definite matrix M. Then,

$$\begin{aligned}
\mathbb{E}\left[H(w_{t'}) - H(w_t) \middle| \mathcal{F}_t\right] =& \mathbb{E}\left[w_{t'}^\top \cdot M \cdot w_{t'} - w_t^\top \cdot M \cdot w_t \middle| \mathcal{F}_t\right] \\
=& \mathbb{E}\left[\mu_{t'}^\top \cdot M \cdot \mu_{t'} - \mu_t^\top \cdot M \cdot \mu_t\right] + 2\mathbb{E}[\mu_{t'}^\top - \mu_t^\top \middle| \mathcal{F}_t] \cdot M \cdot m_t \\
=& \mathbb{E}\left[H(\mu_{t'}) - H(\mu_t) \middle| \mathcal{F}_t\right].
\end{aligned}$$

That is, $\langle w_t \rangle$ satisfies Equation (1) if and only if $\langle \mu_t \rangle$ also satisfies it. Applying Theorem 1 to the transformed space $(w, t) \to (\mu = w - m_t, t)$ immediately implies

Corollary 1.2. *Suppose H is a quadratic function. $f \in \Delta(\{S + m_t, t\}_{t \in T})$ is attainable by admissible process $\langle w_t \rangle$ and stopping time τ if and only if $\mathbb{E}_f[w - m_\tau] = w_0$ and $\forall t \in T$,*

$$\int_{\tau \le t} H(w - m_\tau) f(\mathrm{d}w, \mathrm{d}\tau) + H\left(\int_{\tau > t} (w - m_\tau) f(\mathrm{d}w, \mathrm{d}\tau)\right) - H(w_0) \le \chi \cdot \int \min\{t, \tau\} f(\mathrm{d}w, \mathrm{d}\tau) \tag{6}$$

3 The optimal exploration-stopping problem

In this section, we solve a dynamic exploration-stopping problem. When the martingale $\langle \mu_t \rangle$ is stopped at state μ in period t, the DM obtains a payoff of $U(\mu, t)$, where $U : S \times T \to \mathbb{R}_+$ is continuous and bounded. Then, given admissible strategy $(\langle \mu_t \rangle, \tau)$, the DM's expected payoff is $\mathbb{E}[U(\mu_\tau, \tau)]$. The DM solves the following optimization problem:

$$\sup_{(\langle \mu_t \rangle, \tau) \in \mathcal{M}} \mathbb{E}[U(\mu_\tau, \tau)]. \tag{C}$$

The nature of function $U(\mu, t)$ depends on the application - we provide several examples in Section 4. Here, we analyze problem (C) in its abstract form. Given Theorem 1, we can solve (C) by maximizing over state-time distributions rather than entire processes $\langle \mu_t \rangle$ and stopping time τ. Thus, problem (C) reduces to

$$\sup_{f \in \mathbb{F}} \int U(\mu, \tau) f(\mathrm{d}\mu, \mathrm{d}\tau), \tag{P}$$

where the set of distributions $\mathbb{F}$ is constrained by the information bound (2). We note that the solution of (P) exists under mild conditions:

Lemma 1. *Suppose* $\lim_{t\to\infty}\sup_{\mu\in S} U(\mu,t)=0$, (P) *has a solution.*

Proof. See Appendix B.1. *Q.E.D.*

Since Equation (2) is a concave constraint, (P) is a linear program and can be computed efficiently. To solve (P) analytically, we identify the shadow cost of information in the constraint (2). In the following example, we illustrate our model using the canonical real options problem and identify the shadow cost in a simple one-period setting.

Example 1 (Real options)**.** Real options play an important role in finance and economics (see Dixit and Pindyck 1994). We consider a DM who decides whether to take a risky investment. The investment gives a payoff of $\mu \in S$ net of the investment cost I. There is a safe outside option paying 0. Future payoffs are discounted with rate ρ. Let $S=[0,1]$ and $\mu_0 = 0.5$. The payoff function $U(\mu,\tau)=e^{-\rho\tau}\max\{\mu - I, 0\}$.

The martingale $\langle\mu_t\rangle$ captures the expected value of a potential investment. In the canonical real options problem, $\langle\mu_t\rangle$ is exogenously given (typically a Brownian motion). We consider the real options problem with *active exploration*, where the evolution of $\langle\mu_t\rangle$ is determined by the exploration strategy of the DM. For simplicity, $H(\mu_0)$ is normalized to 0.

The one-period case: We begin with the one-period problem, i.e., when $T=\{0,1\}$ and stopping can only occur at $t=1$.[4] In this case, (P) reduces to

$$\begin{aligned}&\sup_{f\in\Delta(S)} \mathbb{E}_f[U(\mu,1)]\\ &\quad\text{s.t. } \mathbb{E}_f[H(\mu)]\le\chi \text{ and } \mathbb{E}_f[\mu]=\mu_0.\end{aligned}$$

This special case is equivalent to the static "rational inattention" model of Caplin and Dean 2013. Here, we restate their analysis to illustrate the derivation of the shadow cost of information. Suppose at $t=1$, the agent could buy or sell information (measured by H) at the price of $\Lambda(1)$. Then, if the DM stops at μ, the utility net of the cost of information is $U(\mu,1)-\Lambda(1)H(\mu)$. Since the DM also needs to respect the constraint $\mathbb{E}_f[\mu]=\mu_0$, the solution is obtained by looking at points

$$(\mu, U(\mu,1)-\Lambda(1)H(\mu)),$$

[4] Such restriction is without loss when ρ is sufficiently small and the DM is sufficiently patient. Otherwise, the DM finds it optimal to stop at $t=0$ and the solution is trivial.

and taking a convex hull. Hence, it is optimal to stop at points where

$$U(\mu,1)-\Lambda(1)H(\mu)=a\cdot\mu$$

for some $a\in\mathbb{R}^2$ with inequality $U(\mu,1)\le a\cdot\mu+\Lambda(1)H(\mu)$ holding everywhere.[5] Evidently, the lower $\Lambda(1)$ is, the "wider" the distribution f becomes. The value of the multiplier $\Lambda(1)\ge 0$ is determined by a binding information constraint, i.e., [6]

$$\mathbb{E}_f[H(\mu)]=\chi. \tag{7}$$

Figure 2 illustrates the solution. The two red dots depict the points that U tangents $a\cdot\mu+\Lambda(1)H(\mu)$, i.e., the optimal stopping states. $\Lambda(1)$ is pinned down by leading to Equation (7).

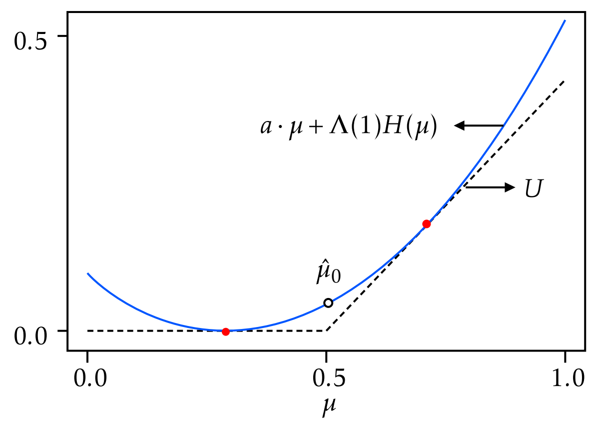


Figure 2: Real options (one period)
Computed with $H(\mu)=(\mu-\mu_0)^2$ and parameters $\chi=0.1$, $\rho=0.5$, $I=0.5$, $\mu_0=0.5$.

Dynamically, the shadow price of information is a function $\Lambda : T\to\mathbb{R}_+$ that sets the price for information in every period. It is weakly decreasing: earlier information is weakly more valuable as it gives the DM more opportunities to stop. Below, we set up the Lagrangian to understand the determinants of the dynamic shadow price of information for our problem.

3.1 Strong Duality & First-order Characterization

We make a few more definitions to set up the Lagrangian. Let $T^\circ := T\setminus\{0\}$. Define $G : T^\circ\to\mathbb{R}$ as the gap in the inequality constraint (2):

$$G(f)(t)=\frac{1}{t}\left(\chi\cdot\left(\int\min\{t,\tau\}f(\mathrm{d}\mu,\mathrm{d}\tau)\right)-H\left(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}\tau)\right)-\int_{\tau\le t}H(\mu)f(\mathrm{d}\mu,\mathrm{d}\tau)+H(\mu_0)\right).$$

[5] Note that our embedding of S in $\mathbb{R}^2$ converts $\mu\in S$ to a vector $(\mu,1-\mu)$. Hence, by a slight abuse of notation, $a\cdot\mu$ represents $(a_0-a_1)\mu+a_1$, which is an affine function of μ.

[6] For now, we ignore the corner case $\Lambda(1)=0$, i.e., the information constraint is slack.

Note that the capacity cumulates linearly in time; hence, we normalize the gap by a factor of $\frac{1}{t}$. The relevant space of state-time distributions is $\Delta_{\mu_0} := \{f \in \Delta(S \times T) | \mathbb{E}_f[\mu] = \mu_0,\ G(f) \in L^\infty(T^\circ)\}$. The shadow cost of information is a non-increasing function Λ on T°:

$$\Lambda(t) := \int_{s \geq t} \mathrm{d}\lambda(s),$$

for some Borel measure λ on T°. The relevant space of measures is $\mathbb{L} := \{\lambda \in \mathcal{B}(T^\circ) | \Lambda \in L^1(T^\circ)\}$, those for which the total shadow value of obtainable information is finite. We can write the Lagrangian for our problem as

$$\mathcal{L}(f,\lambda) := \int_{S \times T} U(\mu,\tau) f(\mathrm{d}\mu, \mathrm{d}\tau) + \int_{T^\circ} t \cdot G(f)(t) \mathrm{d}\lambda(t). \tag{8}$$

Then, the primal problem (P) is equivalently described by

$$\sup_{f \in \Delta_{\mu_0}} \inf_{\lambda \in \mathbb{L}} \mathcal{L}(f,\lambda). \tag{P}$$

The dual problem to (P) is given by

$$\inf_{\lambda \in \mathbb{L}} \sup_{f \in \Delta_{\mu_0}} \mathcal{L}(f,\lambda). \tag{D}$$

We show that under mild technical conditions, strong duality holds:

Lemma 2. *Suppose T is finite or a compact interval, then strong duality holds, i.e.* (P)=(D) *and there exists $\lambda \in \mathbb{L}$ that solves* (D).

Proof. See Appendix B.2. *Q.E.D.*

We say that $\lambda \in \mathbb{L}$ gives the *shadow cost of information* if strong duality holds and λ solves (D). Then, given the shadow cost of information λ, we can find all solutions to (P) by maximizing the Lagrangian $\mathcal{L}(f,\lambda)$.

Next, we characterize candidate stopping points given shadow cost of information $\lambda \in \mathbb{L}$. We proceed somewhat informally to lead up to our next theorem. Consider point (μ,τ). The weight that measure f assigns to this point affects $\mathcal{L}$ linearly in three places, and nonlinearly through the term

$$H\left(\int_{\tau > t} \mu f(d\mu, d\tau)\right)$$

for all times $t \leq \tau$. Notice that, even though total measure f no longer integrates to 1 in this thought experiment, the Lagrangian is still well-defined.

The derivative of $\mathcal{L}$ with respect to mass f at (μ,τ) is

$$l_{f,\lambda}(\mu,\tau) := U(\mu,\tau) + \chi \int_{t\leq\tau} \Lambda(t)\mathrm{d}t - \int_{t\in(0,\tau)} \nabla H(\widehat{\mu_t})\mathrm{d}\lambda(t)\cdot\mu - \Lambda(\tau)H(\mu),$$

where $\widehat{\mu_t} := \int_{\tau>t} \mu f(\mathrm{d}\mu,\mathrm{d}\tau)$.[7]

Since we must also respect the constraints that $E_f[\mu] = \mu_0$ and total measure f must integrate to 1, it is optimal to stop only at points (μ,τ) where

$$l_{f,\lambda}(\mu,\tau) = a\cdot\mu$$

for appropriately chosen vector $a \in \mathbb{R}^{n+1}$, with inequality $l_{f,\lambda}(\mu,\tau) \leq a\cdot\mu$ holding at all other suboptimal points.

Theorem 2. *If for $\lambda \in \mathbb{L}$, $a \in \mathbb{R}^{n+1}$, $f \in \mathbb{F}$ and a selection of $\nabla H(0)$, for all $\mu \in S$,*

$$l_{f,\lambda}(\mu,\tau) \leq a\cdot\mu \tag{9}$$

holds with equality on the support of f, i.e.

$$\int (a\cdot\mu - l_{f,\lambda}(\mu,\tau)) f(d\mu,d\tau) = 0, \tag{10}$$

and the complementary slackness condition $\int G(f)(t)\mathrm{d}\lambda(t) = 0$ holds. Then, f solves problem (P), and λ gives the shadow cost of information.

Conversely, if λ gives the shadow cost of information, then for all f solving (P) with bounded $l_{f,\lambda}$ near $\tau = 0$, there exists $a \in \mathbb{R}^{n+1}$ such that for all $\mu \in S$, (9) and (10) hold for a selection of $\nabla H(0)$.

Proof. See Appendix B.3. *Q.E.D.*

We illustrate Theorem 2 by revisiting Example 1 in a dynamic setting.

Example 2 (Real options - two periods)**.** Consider the real options problem in Example 1 but with $T = \{0,1,2\}$. There are two possible times to stop $\tau = 1$ and $\tau = 2$. Let the shadow cost of information be $\Lambda(1)$ and $\Lambda(2)$ at $t = 1$ and $t = 2$, respectively.

In period 1, as we have already derived in Example 1, the DM's stopping utilities and continuation values at any stopping state μ_1 must be at level

$$a_1\cdot\mu + \Lambda(1)H(\mu) \tag{11}$$

[7] When $t \geq \bar{t} := \sup\operatorname{supp}(f)$, $\widehat{\mu_t} = 0$ and the subdifferential $\nabla H(0)$ is a set. A selection of $\nabla H(0)$ is needed to specify $l_{f,\lambda}$. For notational simplicity, we denote the selection of $\nabla H(0)$ also by $\nabla H(\widehat{\mu_t})$ when writing $l_{f,\lambda}$.

for some $a_1 \in \mathbb{R}^2$. In Figure 3, the top blue solid curve corresponds to (11) and contains two possible optimal stopping states μ_1 for our example — the red dot, which corresponds to stopping, and the black cross, which corresponds to belief $\mu_1 = \widehat{\mu}_1$ with which the DM would continue to the second period.

The period 2 problem is the same as the period 1, except that the "prior" is $\widehat{\mu}_1$. Analogously, in period 2, the DM's stopping utilities must be at level

$$a_2 \cdot \mu + \Lambda(2)H(\mu) \tag{12}$$

for an appropriate multiplier $a_2 \in \mathbb{R}^2$. The bottom blue solid curve in Figure 3 corresponds to (12) in our example. There are two possible optimal stopping states of μ_2 on this curve, depicted by the two red dots.

What is the relation between the two blue solid curves (11) and (12)? The portion of (12) between the two red dots is the DM's value function at time $t = 2$ when he can buy or sell information for the price of $\Lambda(2)$. Hence, the continuation value function at time 1 in case of not stopping is given by

$$a_2 \cdot \mu + \Lambda(2)H(\mu) + \Lambda(2)\chi, \tag{13}$$

since the DM can sell the χ amount of information acquired over time interval $(1,2]$ at price $\Lambda(2)$. The thin red curve in Figure 3 corresponds to (13). From optimality, the continuation value at $\widehat{\mu}_1$ (in the event of not stopping at time 1) must be on curve (11), and any suboptimal continuation value must be weakly below. It follows that the thin red curve (13) must lie weakly below the top blue curve (11), with smooth pasting at belief $\widehat{\mu}_1$, as illustrated in Figure 3.

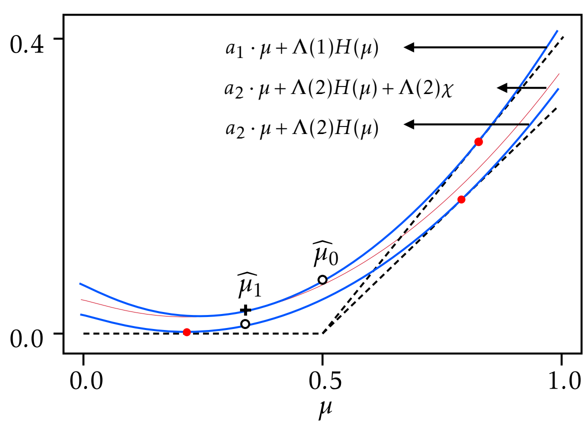


Figure 3: Real options (two periods)

Computed with $H(\mu) = (\mu - \mu_0)^2$ and parameters $\chi = 0.1$, $\rho = 0.5$, $I = 0.5$, $\mu_0 = 0.5$.

The smooth-pasting condition at $\widehat{\mu}_1$ implies

$$a_2 + \Lambda(2)\nabla H(\widehat{\mu}_1) + \chi\Lambda(2)\mathbf{1}^\top = a_1 + \Lambda(1)\nabla H(\widehat{\mu}_1),^{8} \tag{14}$$

or equivalently, $a_2 - a_1 = \nabla H(\widehat{\mu}_2)\lambda(2) - \chi\Lambda(2)\mathbf{1}^\top$, which pins down a_2 from a_1.

Extending the expressions (14) to the general setting implies that $\mathrm{d}a_t/\mathrm{d}t = \nabla H(\widehat{\mu}_t)\lambda(t) - \chi\Lambda(t)\mathbf{1}^\top$. Therefore, the DM's continuation value on the path of the optimal solution is of the form

$$a_t \cdot \mu + \Lambda(t)H(\mu)\,dt, \quad \text{where } a_t = a_0 + \int_{t\in(0,\tau)} \nabla H(\widehat{\mu}_t)\mathrm{d}\lambda(t) - \chi\int_0^\tau \Lambda(t)\mathbf{1}^\top,$$

which is exactly the FOC (9). This exercise delineates the economic implication of the FOC (9). The constrained optimization problem (P) is equivalent to an unconstrained problem with objective function

$$l_{f,\lambda}(\mu,\tau) - a\cdot\mu = \underbrace{U(\mu,\tau)}_{\text{Stopping payoff}} - \underbrace{\Lambda(\tau)H(\mu)}_{\substack{\text{Shadow cost of}\\ \text{“buying” information}}} + \underbrace{\chi\int_{t\le\tau}\Lambda(t)\mathrm{d}t}_{\substack{\text{Shadow benefit of}\\ \text{“selling” endowments}}} - \underbrace{\widehat{a}_t}_{\substack{\text{Shadow cost/benefit of}\\ \text{martingale constraint}}},$$

where $\widehat{a}_t = a_0 + \int_{t\in(0,\tau)} \nabla H(\widehat{\mu}_t)\mathrm{d}\lambda(t)$. In the unconstrained problem, the DM is endowed with χ unit of information per unit time. She “exploits” information to come to a stop and obtain a stopping payoff U. If there is either a surplus or a deficit of information, she can sell or buy information at price $\Lambda(t)$. The violation of the martingale constraint $\mathbb{E}_f[\mu] = \mu_0$ is punished at prices $\widehat{a}_t$.

With Theorem 2, solving the exploration-stopping problem boils down to solving the shadow prices $\Lambda(t)$. Let $\mu^*(t)$ denote the maximizer of $l_{f,\lambda}(\mu,\tau) - a\cdot\mu$ for every t. Then, the equality condition (10) defines an integral equation for $\Lambda(t)$ on the support of f:

$$U(\mu^*(t),t) + \chi\int_{s\le t}\Lambda(s)\mathrm{d}s - \Lambda(t)H(\mu^*(t)) - \widehat{a}_t\cdot\mu^*(t) = 0.$$

Therefore, we obtain a unified method for analytically solving the dynamic exploration-stopping problem. In Section 3.2 and Section 4, we leverage this method to derive general implications and analytical solutions in various economic applications. We revisit the first-order conditions (9) and (10) in Section 4.1, where we develop further understanding of the optimal exploration-stopping problem from the point of view of dynamic programming.

3.2 Implications and extensions

In this section, we derive several general implications of optimal exploration and stopping. Moreover, we derive an extension that generalizes our methodology to handle the case with endogenous capacity.

[8]Note that the smooth-pasting condition matches both the level and the slope of the value function in $\mathbb{R}^1$, which is equivalently represented by a slope condition in $\mathbb{R}^2$ due to our HD-1 normalization.

3.2.1 Coarse support

In Example 2, the solution exhibits a coarse support property, illustrated by Figure 3, where the stopping distribution involves at most two points in the support in each period. In Proposition 1 below, we prove that the coarse support property holds generally, with the size of support bounded by $n+2$.

Proposition 1. *Suppose time is discrete, i.e.* $T = \{t_0 = 0, t_1, \dots, t_k\}$ *and* $U \in C(S \times T)$, *there exists* f^* *solving problem* P *s.t.* $\forall t \in T$,

$$|\mathrm{supp}(f^*(\cdot, t))| \leq n+2.$$

Proof. See Appendix C.2. *Q.E.D.*

The numerical example below (depicted by Figure 4) shows that the bound $n+2$ is tight. In this example, we take the two period problem in Example 2, and add one extra possible option that pays in the low state. The stopping distribution of $\tau = 2$ involves three points, each corresponding to one possible option.

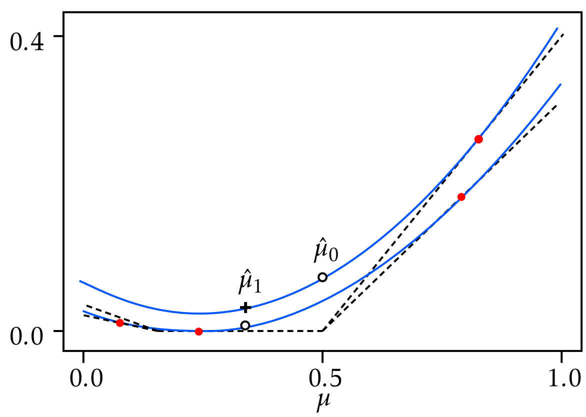


Figure 4: Illustration of Proposition 1

We state Proposition 1 in discrete time because the "support" of f at a single t is meaningless in continuous time. Nevertheless, Proposition 1 implies that the continuous-time problem has an approximate discrete solution that has coarse support. Technically, Proposition 1 is an extension of the coarse support property of the static information design problems (Doval and Skreta 2022; Kamenica and Gentzkow 2011; Zhong 2018) to the dynamic environment.

3.2.2 Time preferences

In the literature on dynamic information acquisition (see, e.g., Zhong 2022), a stark prediction is that the optimal exploration strategy is "Poisson", i.e., the martingale process either drifts along a deterministic path or jumps directly to the stopping region. In this section, we reveal that the optimality of such exploration strategies is closely related to the DM's time preference. For tractability, we consider the continuous time setting where $T = \mathbb{R}_+$ and make several technical assumptions on U: $\forall \mu \in S$, $U(S,\cdot) \in C^{(2)}(\mathbb{R})$ and $U_t'(\mu,t) < 0$. In what follows, we analyze three cases corresponding to convex, linear, and concave $U(\mu,\cdot)$ as a function of time.

Convex time preferences and Poisson process. Convex time preference is the most commonly adopted modeling assumption as it nests the canonical case of exponential discounting, where $U(\mu,t) = e^{-\rho t}u(\mu)$. More generally, it also covers settings with time-varying discount rate like $U(v,t) = e^{-\rho(t)}u(\mu)$, where $\rho'' \leq 0$ (e.g. hyperbolic discounting).

Proposition 2 (Convex time preference)**.** *Suppose* $\forall \mu, t \in S \times T$, $\frac{\partial^2}{\partial t^2}U(\mu,t) > 0$. *Then, if* f *solves* (P) *and* λ *gives the shadow cost of information and* $l_{f,\lambda}$ *is bounded, the* $(\langle \mu_t \rangle, \tau)$ *that implements* f *must satisfy*

$$\text{Prob}\left(\mu_t = \widehat{\mu_t} \middle| t < \tau\right) = 1,$$

where $\widehat{\mu_t} = \mathbb{E}[\mu_\tau | \tau > t]$.

Proof. See Appendix C.3. *Q.E.D.*

$\frac{\partial^2}{\partial t^2}U(\mu,t) > 0$ means that the stopping utility is convex in time, i.e., the DM wants to diversify the decision time. Proposition 2 predicts that the process that implements f must be degenerate conditional on continuation. In other words, the optimal martingale process must be a Poisson process that always jumps into the stopping region. This general result nests the models in Zhong 2022 and Hébert and Woodford 2023 that predict a Poisson belief process under exponential discounting. Moreover, it reveals that the feature of the Poisson learning process is the implication of the convexity exhibited by exponential discounting.

On a side note, Proposition 2 also predicts the uniqueness of the martingale process that embeds f. Since the stopping behavior is fully characterized by f, the multiplicity of the optimal process comes from the undetermined interim process $\mu_t | t < \tau$. In the environment described by Proposition 2, the interim process is uniquely pinned done by $\widehat{\mu_t}$. Therefore, the process that embeds f is essentially unique ($\langle \mu_{\min\{t,\tau\}} \rangle$ has unique distribution).

Linear time preferences and Brownian process. In contrast to the convex case, when U is linear in t, the optimal process exhibits multiplicity and can behave very differently. We begin with a result showing that Equations (C) and (P) reduce to a static problem.

Proposition 3 (Linear time preference[9]). *Suppose $U(\mu,t) = v(\mu) - \kappa t$, then, $(\langle\mu_t\rangle, \tau)$ solves (C) if and only if the distribution of μ_τ solves*

$$\sup_{\pi\in\Delta(S)} \mathbb{E}_\pi[v(\mu) - \kappa/\chi\,(H(\mu) - H(\mu_0))] \tag{15}$$

subject to $\mathbb{E}_\pi[\mu] = \mu_0$.

Proof. See Appendix C.4. *Q.E.D.*

Given $\pi \in \Delta(S)$ that solves Equation (15), one simple Poisson process that satisfies the information constraint and has the distribution of μ_τ given by π is the "dilution" of π: μ_t stays constant from μ_0 and jumps to a random location according to π at constant Poisson rate $\frac{\chi}{\mathbb{E}_\pi[H(\mu)-H(\mu_0)]}$. However, it is not the only such process. Another simple example is specified by

$$f(\mu,\tau) = \pi(\mu) \times \delta_{\tau = \frac{\mathbb{E}_\pi[H(\mu)-H(\mu_0)]}{\chi}},$$

namely, the stopping time is degenerate. Any such process solves (C) by Proposition 3.

In fact, Hébert and Woodford 2023 shows for a payoff of the form in Proposition 3, there exists a Brownian martingale that solves (C). Hébert and Woodford 2023 considers optimal learning for this class of utility functions, assuming that μ_t is restricted to Brownian martingale described by SDE

$$d\mu_t = \sigma_t\, dB_t \tag{16}$$

that satisfies our information constraint (1). Here, B_t is a Brownian motion of dimension n and σ_t is any vector whose entries add up to 0 to ensure that μ_t stays in the probability simplex. In this setting, the following result holds.

Proposition 4. *(Hébert and Woodford 2023) With constant waiting cost, the dynamic utility maximization problem under Brownian learning (16) subject to (1) is equivalent to Equation (15).*

Our general results regarding the convex and concave time preferences imply that the optimality of the Brownian learning process is non-generic: it is a knife-edge case that occurs only when the solution exhibits great multiplicity.

Concave time preferences and exploration. When the time preference is concave, we show that the optimal stopping time is contained in a window of bounded length, termed the "exploitation window". Hence, when the window is sufficiently short, the optimal

[9] Proposition 3 was proved in the working paper version of Zhong 2022.

exploration policy necessarily involves "pure exploration" at the beginning, i.e., acquiring information that will only be used later in the exploitation window. Define two functions:

$$\begin{cases} \bar{J}(t) = \max_{\mu \in S} U_t'(\mu, t); \\ \underline{J}(t) = \min_{\mu \in S} U_t'(\mu, t). \end{cases}$$

Evidently, $\underline{J}(t) \leq \bar{J}(t)$. When $U_t''(\mu, t) < 0$, both $\underline{J}(t)$ and $\bar{J}(t)$ are negative and strictly decreasing. Therefore, $\bar{J}^{-1} \circ \underline{J}$ defines a function satisfying $\bar{J}^{-1} \circ \underline{J}(t) \geq t$. The gap $\bar{J}^{-1} \circ \underline{J}(t) - t$ measures the variation of U_t' across different μ's. To state the result regarding concave time preference, we impose a technical regularity condition on the solution.

Definition 1. *Given f solving (P) and λ giving the shadow cost of information. Let $\bar{t} = \sup_t \mathrm{Supp}(f)$. (f, λ) are regular if (i) $\exists \mu^* = \lim_{t \to \bar{t}^-} \frac{\widehat{\mu_t}}{f(\tau > t)}$, (ii) $(\mu^*, \bar{t}) \in \mathrm{Supp}(f)$, and (iii) $\frac{\Lambda(\bar{t} - \Lambda(t))}{\bar{t} - t}$ is bounded for $t \to \bar{t}^-$.*

Definition 1 requires that the continuation belief converges to a belief that is in the support of the optimal f when t approaches the boundary of the support.

Proposition 5 (Concave time preference)**.** *Suppose $\frac{\partial^2}{\partial t^2} U(\mu, t) < 0$, regular f solves (P) and λ gives the shadow cost of information and $l_{f,\lambda}$ is bounded. Let $\underline{t} = \inf_{t \in T} \mathrm{Supp}(f)$, $\bar{t} = \sup_{t \in T} \mathrm{Supp}(f)$,*

$$\bar{t} \leq \bar{J}^{-1} \circ \underline{J}(\underline{t}).$$

Proof. See Appendix C.5. *Q.E.D.*

Proposition 5 states that when the time preference is concave, the stopping time must be contained within an interval, whose length is determined by the variation of U_t'. In the extreme case where $U_t'(\mu, t)$ does not vary with μ (e.g. U is additively separable), $\bar{J}$ and $\underline{J}$ coincide; hence, the optimal τ must be degenerate. More generally, fixing the variation of U_t' across v, the interval is narrower when U is more concave in time, i.e., when $U_t'(v, t)$ decreases faster.

The intuition for the result is exactly the opposite of the convex case. The concave utility in time means the DM wants concentrated decision time. An indirect implication of Proposition 5 is that concavity of the time preference incentivizes the DM to explore without exploitation for a period of time before stopping so that she can stop quickly within a short window of time.

We illustrate Proposition 5 using a numerical example, depicted by Figure 5. Figure 5 illustrates the distribution of f on the time dimension (the red histogram). $\bar{J}$ and $\underline{J}$ are

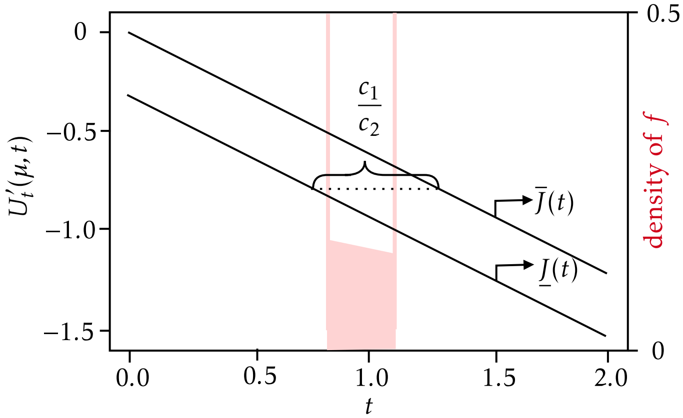


Figure 5: Concave time preference

Computed with $U(\mu,t) = \max\{\mu - 0.5, 0\} \cdot (1 - c_1 t) - c_2 t^2$, where $c_1 = 1/16$, $c_2 = 1/32$.

the two black lines. The dotted segment has length $\bar{J}^{-1} \circ \underline{J}(t) - t$, which equals $\frac{c_1}{c_2}$ for every t in this example.

The analysis in Section 3.2.2 nests a series of works on dynamic information acquisition and provides the near-complete characterization of the pattern of optimal information acquisition strategy. The Poisson learning is justified by convex time preferences (see Zhong 2022 and D. Chen and Zhong 2024). Brownian learning is justified by linear time preferences (see Hébert and Woodford 2023). The pure exploration is justified by concave time preferences (see D. Chen and Zhong 2024).[10] To further understand the connection between time preference and exploration, in Section 4.2, we provide the solution to an information acquisition problem under fully general time preference.

3.2.3 Endogenous capacity

In many applications, the DM may choose the rate of information arrival at cost. This section shows that our main theorems generalize to such a setting. The DM chooses martingale $\langle \mu_t \rangle$ in S, a bounded process $\langle \chi_t \rangle$ in $\mathbb{R}_+$ and a stopping time τ that are measurable w.r.t. the filtration $\langle \mathcal{F}_t \rangle$ generated by $\langle \mu_t \rangle$. The tuple $(\langle \mu_t \rangle, \langle \chi_t \rangle, \tau)$ is *admissible* if $\forall t, t' \in T, t' > t$

$$\mathbb{E}\left[H(\mu_{t'}) - H(\mu_t) | \mathcal{F}_t\right] \leq \mathbb{E}\left[\int_t^{t'} \chi_s \mathrm{d}s \middle| \mathcal{F}_t\right],$$

denoted by $(\langle \mu_t \rangle, \langle \chi_t \rangle, \tau) \in \mathcal{M}$. Given the learning rate χ_t at time t, the DM pays flow cost of $c_t(\chi_t)$, where $\forall t$, c_t is an increasing and convex function. Then, the DM's optimization

[10] Following a different approach, Georgiadis-Harris 2024 predicts pure exploration as an outcome of exogenous random stopping time that is not controlled by the DM.

problem is

$$\sup_{(\langle\mu_t\rangle,\langle\chi_t\rangle,\tau)\in\mathcal{M}} \mathbb{E}\left[U(\mu_\tau,\tau)-\int_{t\le\tau} c_t(\chi_t)\mathrm{d}t\right]. \tag{C1}$$

Analogous to the derivation of (P) from (C), a "relaxed" problem of (C1) is

$$\sup_{f\in\Delta(S\times T),\chi\in L^\infty(T)} \int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)-\int_{t\in T} c_t(\chi_t)(1-F(t))\mathrm{d}t \tag{P1}$$
$$\text{s.t.} \int_{\tau\le t} H(\mu)f(\mathrm{d}\mu,\mathrm{d}\tau)+H\left(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}\tau)\right)-H(\mu_0)\le\int_{s\le t}\chi_s(1-F(s))\mathrm{d}s,$$

where $F(t)=\int_{\tau<t} f(\mathrm{d}\mu,\mathrm{d}\tau)$.[11] Note that in (P1), we implicitly restrict the stochastic learning rate χ_t to be a deterministic function of time, while in (C1), $\langle\chi_t\rangle$ is a stochastic process. Nevertheless, we prove in Lemma 3 that such restriction is without loss.

Lemma 3. *(C1)=(P1) and for all (f,χ) feasible in (P1), there exists admissible strategy $(\langle\mu_t\rangle,\chi_t,\tau)$ of (C1) s.t. $(\mu_\tau,\tau)\sim f$.*

Proof. See Appendix C.6. *Q.E.D.*

Then, we can write the Lagrangian of Equation (P1) as

$$\mathcal{L}(f,\chi,\lambda):=\int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)-\int c_t(\chi_t)(1-F(t))\,\mathrm{d}t+\int\Bigg(\int_{s\le t}\chi_s(1-F(s))\,\mathrm{d}s$$
$$-H\left(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}\tau)\right)-\int_{\tau\le t}H(\mu)f(\mathrm{d}\mu,\mathrm{d}\tau)+H(\mu_0)\Bigg)\mathrm{d}\lambda(t)$$

The dual problem is

$$\inf_{\lambda\in\mathbb{L}}\sup_{f\in\Delta_{\mu_0},\chi\in L^\infty}\mathcal{L}(f,\chi,\lambda). \tag{D1}$$

The same technical conditions as in Lemma 2 guarantee strong duality.

Lemma 2-A. Suppose T is finite or a compact interval, then strong duality holds,i.e. (P1)=(D1) and there exists $\lambda^*\in\mathbb{L}$ that solves (D1).

Proof. See Appendix C.7 *Q.E.D.*

Lemma 2-A allows us to characterize the solution of Equation (P1) via first order conditions. Define the derivative of $\mathcal{L}$ with respect to f at (μ,τ) as:

$$l_{f,\chi,\lambda}(\mu,\tau):=U(\mu,\tau)-\int_{t\le\tau}c_t(\chi_t)\mathrm{d}t+\int_{t\le\tau}\Lambda(t)\chi_t\mathrm{d}t-\int_{t<\tau}\nabla H(\widehat{\mu}_t)\mathrm{d}\lambda(t)\cdot\mu-\Lambda(\tau)H(\mu).$$

[11] Whether t is included in the domain is inconsequential since it only affects $F(t)$ on a zero measure set.

The FOC w.r.t. f at (μ,τ) is

$$l_{f,\chi,\lambda}(\mu,\tau) \leq a \cdot \mu, \quad \text{with equality on } \mathrm{supp}(f). \tag{17}$$

The FOC w.r.t. χ at τ is

$$c_\tau'(\chi_\tau) = \Lambda(\tau). \tag{18}$$

Theorem 2-A. If there exists λ, f, χ, a and a selection of $\nabla H(0)$ such that (17),(18) and the complementary slackness condition hold, then (f,χ) solves (P1).

Conversely, if λ gives the shadow cost of information, for all (f,χ) solving (P1) with $l_{f,\chi,\lambda}$ bounded from above near $\tau = 0$, there exists $a \in \mathbb{R}^{n+1}$ such that (17),(18) hold for a selection of $\nabla H(0)$.

Proof. See Appendix C.8. *Q.E.D.*

The FOC (17) is the same as that with exogenous capacity constraint in our baseline model, except that the shadow price of information $\Lambda(t)$ must now coincide with the real marginal cost of information $c_t'(\chi_t)$ per the extra FOC (18).

4 Applications

4.1 Real Options

We have already analyzed the real options problem with one period and two periods in Examples 1 and 2, respectively. Recall from Example 1 that $U(\mu,\tau) = e^{-\rho\tau}\max\{\mu - I, 0\}$, representing a DM deciding whether to make a risky investment to obtain a stochastic payoff $\mu - I$. In what follows, we analyze the real options problem with active exploration in continuous time and connect it to the canonical problem with passive exploration. [12]

Passive exploration: Consider the canonical setting where the DM learns information about a potential investment *passively*, so the expected value of investment follows

$$d\mu_t = \sigma\, dZ_t, \tag{19}$$

were Z is a Brownian motion. For this problem, there exists a closed-form solution in which it is optimal to invest when μ_t reaches the critical threshold of

$$\mu^* = I + \frac{\sigma}{\sqrt{2\rho}}.$$

[12] The canonical real options problem does not limit the lower bound of the state. To make the analysis consistent, in this example, we let S have a sufficiently negative lower bound.

The value function is given in closed form by

$$V_P(\mu) = \begin{cases} (\mu^* - I)\exp\left(\frac{\sqrt{2\rho}}{\sigma}(\mu - \mu^*)\right) & \text{if } \mu \le \mu^* \\ \mu - I & \text{if} \mu > \mu^*. \end{cases}$$

Active exploration: How does this solution change if, instead of learning passively, the DM actively collects information subject to the constraint (1)? Specifically, assume that $H(\mu) = \mu^2$ and $\chi = \sigma^2$ so that the choice to learn Brownian information leads precisely to equation (19). This case is isomorphic to the information acquisition model of Zhong 2022 up to scaling and relabeling, which characterized the value function via the Hamilton-Jacobi-Bellman (HJB) equation

$$V(\mu) = \begin{cases} \max_\nu \frac{\chi}{\rho} \frac{\nu - I - V(\mu) - V'(\mu)(\nu - \mu)}{H(\nu) - H(\mu) - H'(\mu)(\nu - \mu)} & \text{if } \mu \le \bar{\mu}^* \\ \mu - I & \text{if } \mu > \bar{\mu}^*, \end{cases} \tag{20}$$

where $\bar{\mu}^* \ge \mu^*$ since the value function under active learning $V(\mu)$ must be (weakly) higher than $V_P(\mu)$. The left panel of Figure 6 illustrates these properties by comparing value functions $V(\mu)$ and $V_P(\mu)$.

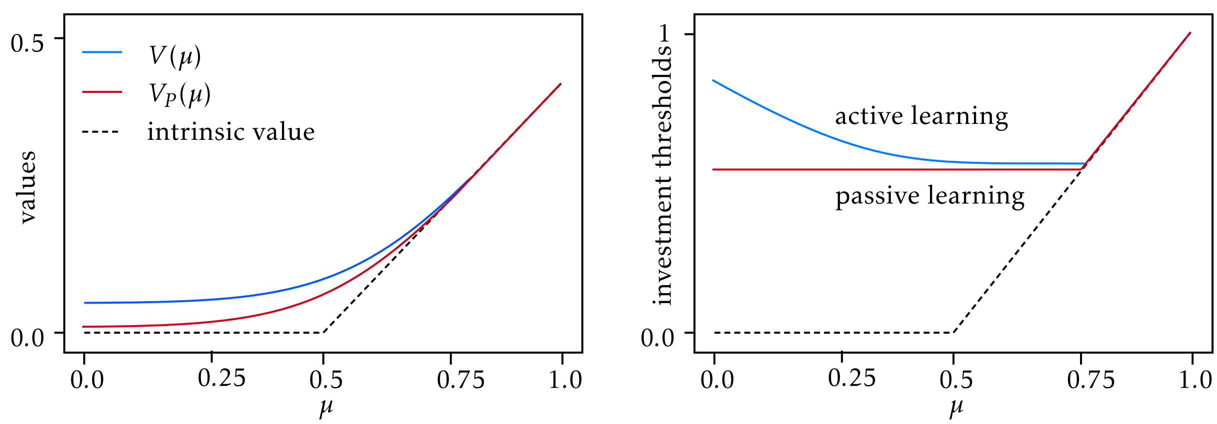


Figure 6: Active and passive learning

Computed with $H(\mu) = (\mu - \mu_0)^2$ and parameters $\chi = 0.1$, $\rho = 0.5$, $I = 0.5$, $\mu_0 = 0.5$

We also see that $V(\mu)$ is significantly higher than $V_P(\mu)$ when the option is deep out of the money. This is where active learning is different. While it takes an extremely long time for the option to get in the money due to volatility alone, active learning does much better by "shooting for the moon". Under active learning, optimal policy performs experiments that allow μ_t to jump up to $\nu(\mu_t)$ with Poisson intensity defined by the information constraint (1), with downward drift in the event of no jump. The right panel of Figure 6 illustrates that $\nu(\mu)$ starts from $\bar{\mu}^*$ and increases as μ drifts down. As the left panel indicates, experiments that reveal a high value of μ with a small probability can improve the option value significantly.

In what follows, we illustrate that our method nests the dynamic programming approach in the special case of exponential discounting by deriving Equation (20) from our first-order conditions. Let $e^{-\rho t}V$ and $e^{-\rho t}V'$ be the level and slope of $a_t\widehat{\mu}_t + \Lambda(t)H(\widehat{\mu}_t)$ at $\widehat{\mu}_t$. Then, the optimal stopping state μ^* is determined by Equation (10):

$$0 = \max_{\mu} e^{-\rho t}(\mu - I) - (a_t\mu + \Lambda(t)H(\mu)) \tag{21}$$

$$\iff \Lambda_t = e^{-\rho t}\max_{\mu}\frac{(\mu - I) - V - V'(\mu - \widehat{\mu}_t)}{H(\mu) - H(\widehat{\mu}_t) - H'(\widehat{\mu}_t)(\mu - \widehat{\mu}_t)}. \tag{22}$$

The equivalence of Equations (21) and (22) is illustrated in Figure 7, where the pink line represents $e^{-\rho t}V + e^{-\rho t}V'(\mu - \widehat{\mu}_t)$. Then, (21) corresponds to minimizing the segment α and (22) corresponds to maximizing the ratio γ/β, which are both attained at the red dot.

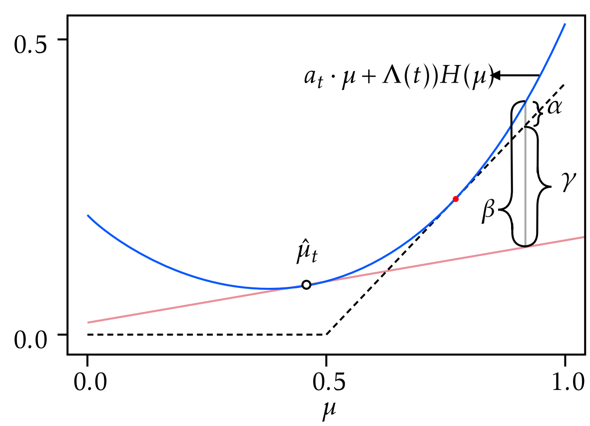


Figure 7: Derivation of the HJB equation

We have derived in Example 2 that the levels of the blue curves at $\widehat{\mu}_t$ in two periods dt apart differ by $\chi\Lambda_t$dt. Since the value function is stationary due to exponential discounting, this difference is $e^{-\rho t}V - e^{-\rho(t+\mathrm{d}t)}V$. Therefore, taking d$t \to 0$,

$$\chi\Lambda_t = \rho e^{-\rho t}V. \tag{23}$$

Combining Equations (22) and (23),

$$\frac{\rho}{\chi}V = \max_{\mu}\frac{\mu - I - V - V'(\mu - \widehat{\mu}_t)}{H(\mu) - H(\widehat{\mu}_t) - H'(\widehat{\mu}_t)(\nu - \widehat{\mu}_t)},$$

which is exactly Equation (20), where the value function is given by

$$V(\widehat{\mu}_t) = e^{\rho t}\left(a_t\widehat{\mu}_t + \Lambda(t)H(\widehat{\mu}_t)\right).$$

Of course, for a more general time preference that is not time-stationary, the HJB equation approach involves much more complicated partial differential equations, rendering the method underpowered. In the following sections, we illustrate the power of our method via two applications where the DM exhibits a general time preference.

4.2 Time preferences: exploration, exploitation, precision, and speed.

In this application, we apply our method to a canonical information acquisition problem with binary decision. There is an unknown payoff relevant state $x \in \{L, R\}$, with equal prior probability. There are two possible actions $a \in \{l, r\}$. The Bernoulli utility is $\frac{1}{2}\rho(t)$ if the action matches the state ($l|L$ or $r|R$) and $-\frac{1}{2}\rho(t)$ otherwise ($l|R$ or $r|L$). We assume that the discount function $\rho \in C^{(2)}\mathbb{R}_+$ is decreasing and $\lim_{t\to\infty}\rho(t) = 0$.

Let $S = \Delta(X) = [0,1]$. The DM chooses her belief process $\langle \mu_t \rangle$, the stopping time τ, and an action upon stopping. In this problem, we consider the variation constraint defined by $H(\mu) = |\mu - 0.5|^\alpha$, where $\alpha > 1$. The stopping utility is

$$U(\mu, t) = \rho(t) \cdot |\mu - 0.5|.$$

The FOC Equations (9) and (10) reduces to:

$$\rho(t)|\mu - 0.5| + \int_{s\leq t} \chi\Lambda(s)\mathrm{d}s - \Lambda(t)|\mu - 0.5|^\alpha \leq b, \tag{24}$$

with equality on the support of f.[13] It is easy to verify that $\mu^*(t) = (\rho(t)/\Lambda(t))^{\frac{1}{\alpha-1}}$ maximizes the LHS. Let $\xi(t)$ be the minimal gap in the inequality for every t. Equation (24) reduces to:

$$\frac{\alpha-1}{\alpha}\rho(t)\left(\frac{\rho(t)}{\alpha\Lambda(t)}\right)^{\frac{1}{\alpha-1}} + \chi\int_{s\leq t}\Lambda(s)\mathrm{d}s + \xi(t) = b, \tag{25}$$

with $\xi(t) = 0$ on the support of f.

4.2.1 The pattern of exploration

We first characterize the optimal pattern of exploration. We make the following assumption on the discount function ρ:

Assumption 1. $\left\{t\middle|\frac{\mathrm{d}}{\mathrm{d}t^2}\rho(t)^{\frac{\alpha}{\alpha-1}} > 0\right\}$ *and* $\left\{t\middle|\frac{\mathrm{d}}{\mathrm{d}t^2}\rho(t)^{\frac{\alpha}{\alpha-1}} < 0\right\}$ *each is constituted of finitely many intervals.*

[13] Observe that the problem has a solution that is symmetric around $\mu = 0.5$. Hence, $\widehat{\mu_t} \equiv (0.5, 0.5)$ and $\nabla H(\widehat{\mu_t}) \equiv 0$. This is a useful trick for simplifying the problem in any symmetric setting.

Assumption 1 states that the convexity of $\rho^{\frac{\alpha}{\alpha-1}}$ only switches finitely many times. It is a purely technical condition that ensures that the optimal exploration policy also switches pattern finitely many times.

Proposition 6. *Given Assumption 1, suppose f, λ and $\xi \in C\mathbb{R}_+$ solve Equation (25). Then, three finite collections of intervals form a partition of* $\mathrm{conv}(\mathrm{supp}(f))$*:*

1. *Region $\mathcal{A}$ where $\xi(t) > 0$ and $\lambda(t) \equiv 0$. Let (t', t'') be such an interval, the following holds:*
   (a) $\frac{\mathrm{d}}{\mathrm{d}t^2}\rho(t)^{\frac{\alpha}{\alpha-1}} \le 0$ *for* $t \to t''_-$*;*
   (b) *If $t' > 0$, then $\frac{\mathrm{d}}{\mathrm{d}t^2}\rho(t)^{\frac{\alpha}{\alpha-1}}$ switches sign at least once in (t', t'').*
   (c) $f(S, (t', t'')) = 0$.
2. *Region $\mathcal{E}$ where $\xi(t) \equiv 0$ and $\lambda(t) > 0$. In this region, the following holds:*
   (a) $\mathrm{supp}(f) = \left\{ \left(0.5 \pm \left(\frac{\rho(t)}{\alpha \Lambda(t)} \right)^{\frac{1}{\alpha-1}}, t \right) \right\}_{t \in \mathcal{E}}$.
3. *Region $\mathcal{R}$ where $\xi(t) = 0$ and $\frac{\mathrm{d}}{\mathrm{d}t^2}\rho(t)^{\frac{\alpha}{\alpha-1}} = 0$.*

Proof. See Appendix D.1. *Q.E.D.*

Proposition 6 shows that the optimal information acquisition strategy involves three distinct patterns that are dictated by the convexity of the (adjusted) discount function $\rho(t)^{\frac{\alpha}{\alpha-1}}$, or equivalently, the time-risk attitude. [14]

- *Pure exploration* region $\mathcal{A}$: in this region, there is no stopping probability (property 1.c). Therefore, information is "accumulated" for future use. A pure exploration period always involves a period of "convex then concave" $\rho(t)^{\frac{\alpha}{\alpha-1}}$ (property 1.a & 1.b). The key driving force behind pure exploration is that the concave part of the discount function implies time-risk aversion. Therefore, the DM would like to accumulate knowledge so that she can utilize the accumulated knowledge later to make decisions within a short period — the time risk is minimized.
- *Full exploitation* region $\mathcal{E}$: in this region, Equation (2) is binding, implying that the continuation belief of any implementing belief process must be degenerate and constant. Therefore, information is "exploited" at the maximal rate to reach immediate decisions. A full exploitation period typically involves "concave then convex"

[14] Since $\alpha > 1$, $\rho^{\frac{\alpha}{\alpha-1}}$ is "more convex" than ρ. So the convexity of $\rho(t)^{\frac{\alpha}{\alpha-1}}$ does not exactly match the convexity of ρ. The discount function is adjusted to accommodate the fact that achieving the same decision quality is easier later than earlier. Therefore, the later payoffs are discounted further by a factor of $\rho^{\frac{1}{\alpha-1}}$. In what follows, we refer to the "time-risk attitude" as defined by the convexity of the adjust discount function.

$\rho(t)^{\frac{\alpha}{\alpha-1}}$. The key driving force behind exploitation is that the convex discount function means time-risk loving. Therefore, the DM maximizes the time risk by inducing a dispersed decision time.

Note that exploration still continues during the period of full exploitation. However, all information explored is immediately exploited by stopping and making a decision.

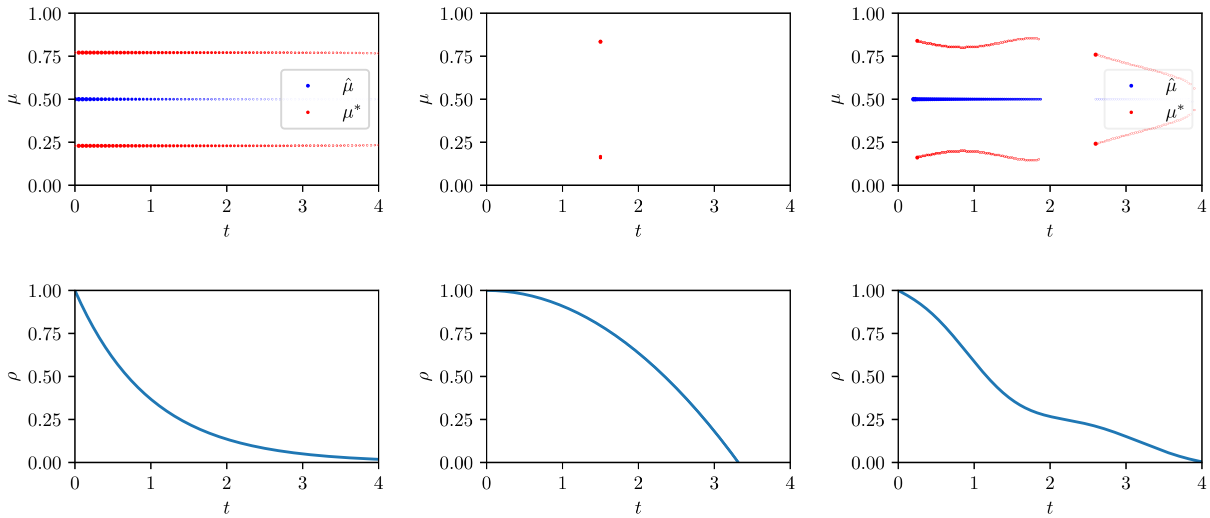
- *Time-risk neutral* region $\mathcal{R}$: in this region, the adjusted discount function $\rho^{\frac{\alpha}{\alpha-1}}$ is linear, implying that the DM is time-risk neutral. Unlike $\mathcal{A}$ and $\mathcal{E}$, there is no unique prediction of the optimal belief process since the DM is essentially indifferent between different distributions of the stopping time that have the same expectation.

It is worth pointing out that the switch between pure exploration and full exploitation strictly precedes the switch of the time-risk attitude. This is because the consequence of pure exploration is not instantaneous — it takes time to accumulate sufficient information to make a decision. Therefore, the DM will start to accumulate information while anticipating the time-risk aversion in the near future. Vice versa, anticipating time-risk loving in the sufficiently near future, the DM starts full exploitation right away. Proposition 6 immediately implies the following corollary.

Corollary 2.1. *The optimal policy involves pure exploration (full exploitation) when $\rho(t)^{\frac{\alpha}{\alpha-1}}$ is globally concave (convex).*

Figure 8 illustrates Proposition 6 and Corollary 2.1. From left to right, the first row of each column depicts the optimal information acquisition policy when the discount function ρ is given by the second row. In all figures, the red dots are the stopping beliefs, and the blue dots are continuation beliefs (plotted only when uniquely determined). The first two columns show the two corner cases in Corollary 2.1. In column 1, $\rho(t)$ is the standard exponential discounting function, which implies global time-risk-loving preference. The optimal belief stays at the prior until it jumps to one of the two constant stopping boundaries at a Poisson rate (pure exploitation). In column 2, the DM is globally time-risk averse. The optimal stopping time is degenerate (pure exploration). The belief process that implements the optimal f is not unique.

Column 3 illustrates a general case where the DM switches from time-risk averse to time-risk loving twice. The optimal belief process switches from pure exploration to exploitation exactly twice. As is predicted by Proposition 6, point (i), each exploration region (except the first one) covers at least one time-risk-loving region and ends in a time-risk-averse region. In other words, the switch between the two patterns strictly precedes the switch of time-risk attitude.



(a) Risk seeking (b) Risk averse (c) General preference

Figure 8: Information acquisition & time-risk preference

As a final remark, while the time-risk neutral region $\mathcal{R}$ might exist in general (where ρ is flat), there always exists an optimal exploration strategy that consists of only pure exploration and full exploitation.

Definition 2. *$f \in \Delta(S \times T)$ is a **pure strategy** if it alternates between only pure exploration and full exploitation, i.e. $\int G(f)(t) f(\mathrm{d}\mu, \mathrm{d}t) \equiv 0$.*

Proposition 7. *The information acquisition problem has a pure strategy solution f.*

Proof. See Appendix D.2 *Q.E.D.*

4.2.2 Speed v.s. accuracy

In this section, we study the speed-accuracy tradeoff in dynamic exploration. We focus on the full exploitation case ($\rho(t)^{\frac{\alpha}{\alpha-1}}$ is globally convex) where decision time has full support. The accuracy of decision is measured using parameter $\kappa(t) = (\mu^*(t) - 0.5)^{\alpha-1} = \frac{\rho(t)}{\alpha\Lambda}(t)$. Evidently, $\kappa(t)$ is isomorphic to the precision of the posterior belief upon stopping as well as the stopping payoff. Equation (25) reduces to the following ODE about κ:

$$-\frac{\mathrm{d}\log(\rho(t))}{\mathrm{d}t} = \frac{\kappa'(t) + \chi \cdot \kappa(t)^{\frac{-1}{\alpha-1}}}{(\alpha-1)\kappa(t)}. \tag{26}$$

Note that the LHS of Equation (26) is the discount rate (of an exponential discount function). Therefore, the sign of the rate of the LHS represents whether there is accelerating/decelerating discounting. Proposition 8 below shows that it is crucially related to the evolution of decision quality.

Proposition 8.

- Increasing accuracy: $\kappa'(t) > 0$ *&* $\kappa''(t) < 0 \implies$ *decreasing discount rate.*
- Decreasing accuracy: $\kappa'(t) < 0$ *&* $\kappa''(t) > 0 \implies$ *increasing discount rate.*
- Constant accuracy: $\kappa'(t) \equiv 0 \iff$ *constant discount rate.*

Proof. See Appendix D.3. *Q.E.D.*

Proposition 8 provides a foundation for the speed-accuracy/inaccuracy tradeoffs that are observed in binary choice experiments (see a survey by Ratcliff, P. L. Smith, et al. 2016). Instead of analyzing a parametric drift-diffusion model (DDM) (see, e.g., the DDM with optimal stopping studied in Fudenberg, Strack, and Strzalecki 2018), we fully endogenize the exploration process. The main focus of Proposition 8 is analogous to the study of time-varying stopping boundaries in the DDM models. Our model provides a closed-form characterization of the boundary and shows that its slope is closely related to the slope of the discount rate.

Proposition 8 predicts that the typical speed-accuracy tradeoff observed in the binary choice experiments is rationalized by a decreasing discount rate. This is intuitive — anticipating decelerating discounting in the future, the DM would take advantage of that and back-load the high-accuracy decisions. On the other hand, the speed-accuracy complementarity occurs under accelerating discounting, which fits decisions under time pressure. Constant accuracy occurs if and only if the discount rate is constant.

4.3 Continuous-time contest

In the second application, we apply our model to study a strategic contest. In particular, we are interested in the continuous-time contest setting. While the literature has studied competition in the dimension of the stopped state of a stochastic process (Seel and Strack 2013, Seel and Strack 2016) and the dimension of stopping time (Park and L. Smith 2008) separately, we study a novel setting where (i) both the stochastic process and stopping time are fully endogenized and (ii) contestants' payoffs depends on both the state and stopping time.

We assume that there are $n \geq 2$ contestants, each choosing privately a martingale process $\langle \mu_t^i \rangle$ in $S = [-M, M]$ starting at $\mu_0^i = 0$ and a stopping time τ^i. The payoff to contestant i given the profile of stopping time $\boldsymbol{t}$ and stopping state $\boldsymbol{\mu}$ is:

$$U^i(\mu^i, \boldsymbol{t}) = e^{-rt^i} \cdot |\mu^i| \cdot \frac{\mathbf{1}_{t^i = \min\{\boldsymbol{t}\}}}{\#\arg\min\{\boldsymbol{t}\}}.$$

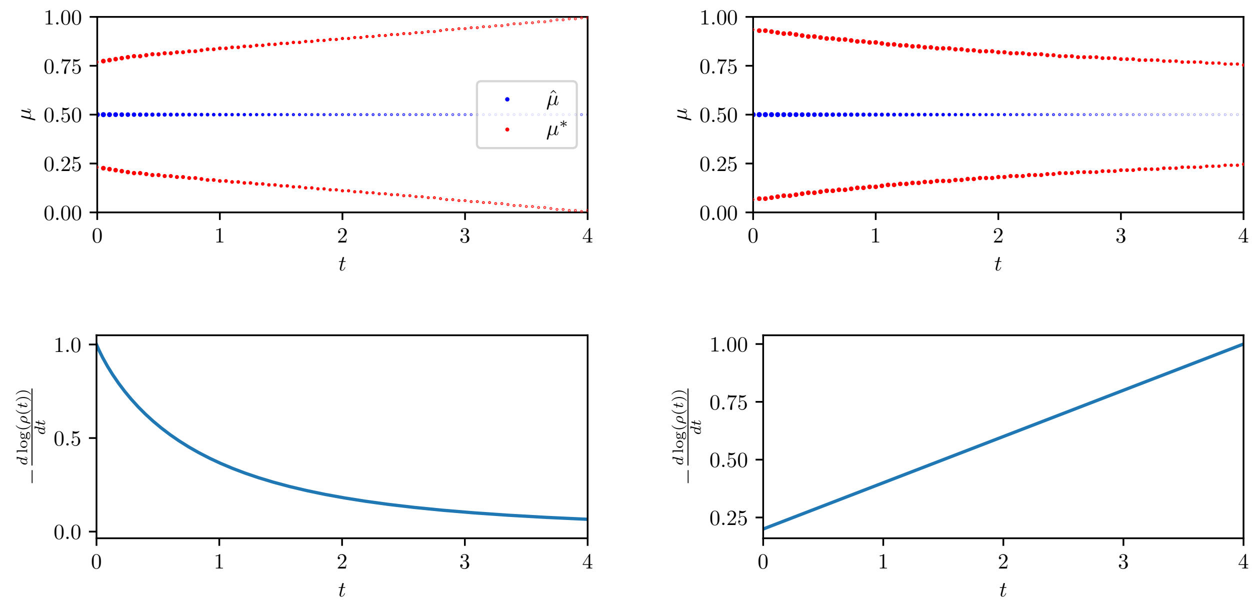


Figure 9: Decision accuracy & discount rate

The interpretation is that the contestants compete in conducting research on the same topic. Each contestant chooses privately how she explores, which affects the stochastic quality $\langle\mu_t^i\rangle$. The first contestant who stops (submits a paper/grant) receives a reward (publishing a paper or receiving a grant) proportional to the quality of the research $|\mu_t^i|$. For tractability, we assume that all contestants have the same variation bound specified by $H(\nu) = |\nu|^\alpha$, where $\alpha > 1$.

The equilibrium of the contest is specified by a collection of independent exploration-stopping strategies $(\langle\mu_t^{*i}\rangle, \tau^{*i})_{i=1}^n$ s.t. $\forall i$,

$$(\langle\mu_t^{*i}\rangle, \tau^{*i}) \in \arg\max_{(\langle\mu_t^i\rangle, \tau^i) \in \mathcal{M}^i} \mathbb{E}[U^i(\mu_{\tau^i}^i, \boldsymbol{\tau}^{*-i}, \tau^i)].$$

In an equilibrium, each contestant takes other contestants' strategies as given and best responds by choosing her own strategy. By defining the equilibrium this way, we implicitly assume that each contestant's research progress is private; hence, the strategy of player i does not depend on the realization of $(\langle\mu_t^{-i}\rangle, \tau^{-i})$. We are interested in equilibria with the following technical properties:

Definition 3. *Equilibrium $(\langle\mu_t^{*i}\rangle, \tau^{*i})_{i=1}^n$ is a* pure strategy equilibrium *if all $f^i \sim (\mu_{\tau^{*i}}^{*i}, \tau^{*i})$ are pure strategies.*

In other words, pure strategy equilibria are those in which each contestant alternates between pure exploration and exploitation as was described in Definition 2. The technical restrictions allow us to characterize all equilibria of the contest.

Assumption 2. $\frac{\chi}{(\alpha-1)r} < M^{\alpha}$.

Assumption 2 guarantees that the equilibria we identify will be interior. It is without loss of generality as M can be chosen arbitrarily large.

Proposition 9. *Suppose Assumption 2 holds. Let* $\zeta = \max\{1-(n-1)(\alpha-1), 0\}$. *In any pure strategy equilibrium of the game, all players adopt the identical strategy indexed by parameter* $\bar{t} \in [0, +\infty]$:

- *On the domain* $[0, \bar{t}]$, μ_t *starts at* 0 *until it jumps to* μ_t^* *or* $-\mu_t^*$ *at rate* $\frac{1}{2}\lambda_t^*$ *and* τ *is the first jump time of* μ_t, *where*[15]

$$\mu_t^* := \left(\frac{\chi}{r}\frac{\zeta(1-e^{(\alpha-1)r(t-\bar{t})})}{\alpha-1}\right)^{\frac{1}{\alpha}}$$
$$\lambda_t^* := \frac{(\alpha-1)r}{\zeta(1-e^{(\alpha-1)r(t-\bar{t})})}$$

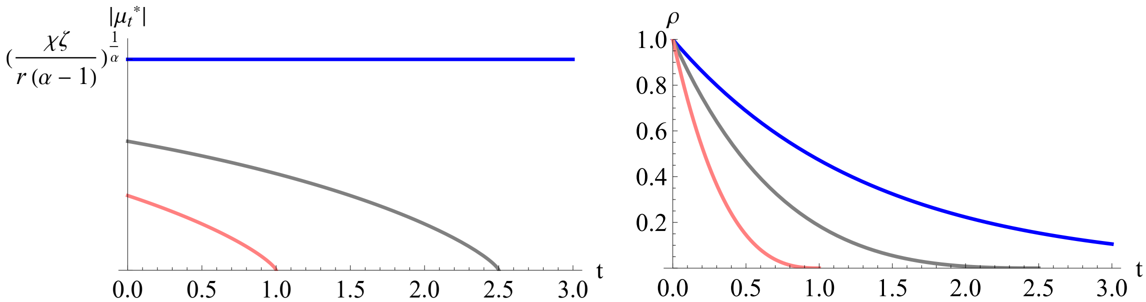
Proof. See Appendix D.4. *Q.E.D.*

Proposition 9 states that contestants use pure exploitation strategy in all pure strategy equilibria of the game. To illustrate the proposition, in figure Figure 10, we plot three possible equilibria of the game. In Figure 10(a), we plot the stopping quality $|\mu_t^*|$ as functions of t. Each color corresponds to one equilibrium in the game. In Figure 10(b), we plot the "effective discount factor", i.e., e^{-rt} scaled by the probability that at least one other contestant has stopped in the equilibrium.

As we have discussed in Section 4.2, Proposition 9 implies an endogenous time-risk loving preference among all contestants. As is illustrated by Figure 10(b), the effective discount factors are convex in time, justifying the Poisson exploration strategy. Importantly, Proposition 9 also predicts uniqueness: the endogenous effective discount factors are always convex. This is because, in the contest, each contestant solves the single-agent exploration-stopping problem, taking others' strategies as given. Propositions 6 and 7 implies that the only alternative exploration strategy that can occur is pure exploration. However, pure exploration leads to concentrated decision time (a point mass in stopping time). The point mass can never appear in equilibrium as other contestants can easily sacrifice quality a little bit and "undercut" by stopping a little earlier.

A key message of Proposition 9 is that contest rules have strong implications on the pattern of exploration. In our application, the "winner takes all" rule generates endogenous time-risk loving, leading to quite risky exploration policies. Contestants employ the full exploitation strategies that count on rare but significant breakthroughs.

[15] When $\bar{t} = 0$ or $\zeta = 0$, the strategy stops immediately at 0. When $\bar{t} = +\infty$, we define $e^{(\alpha-1)r(t-\bar{t})} \equiv 0$.



(a) Quality (b) Effective discount factor

Figure 10: Equilibrium strategies of the research contest game

4.3.1 Private v.s. public contest

So far, we focus on the case where each contestant's progress is private. A rather interesting observation is that the equilibria identified in Proposition 9 remain equilibria (up to minor modifications) even if each contestant's progress is revealed. The key observation is that under the learning strategy specified in Proposition 9, the event $\min\{\tau^{*-i}\} > t$ pins down a unique history for period t, that is, all μ_t^{-i} remain at 0.5. Then, the hazard rate for $s \geq t$

$$\omega_i(s) = r + (n-1)\lambda_s^*$$

remains the same conditional on this history. Therefore, conditional on any history where no contestant has stopped yet, strategy $(\langle \mu_t^{*i} \rangle, \tau^{*i})$ remains the best response for contestant i. Of course, conditional on the event $\min\{\tau^{*-i} \leq t\}$, i.e., someone has already stopped, no contestant has any incentive to learn anymore. Therefore, the equilibrium strategy involves immediate stopping.

5 Conclusion

In this paper, we characterize the possible outcomes of exploration and stopping and develop a general methodology for solving optimal exploration-stopping problems. By fully delineating the connection between time preference and the pattern of dynamic exploration, the current paper brought the theme of Zhong 2022 to completion. This methodology has the power to drive research in at least two distinct areas.

The first is contest design. In Section 4.3, we illustrate how to solve the equilibria of a contest given a specific reward structure. The same methodology can be used to solve a general multi-agent exploration game, including, for instance, a cooperative exploration setting. Ultimately, we hope that the current paper could be used to build a methodology for designing optimal contests to obtain general goals regarding the outcome and timing of exploration.

The second is dynamic persuasion/information design. A series of recent papers explore the optimal design of information provision to persuade an agent to engage with the principal for longer (see Knoepfle 2020, Hébert and Zhong 2022 and Koh and Sanguanmoo 2024). The existing papers focus on very special preference structures, leaving the general insight eluding. We hope that the current paper can be used to build a methodology that fully illuminates information provision in principal-agent settings.

References

Anderson, Axel, Lones Smith, and Andreas Park (2017). "Rushes in large timing games." In: *Econometrica* 85.3, pp. 871–913.

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.

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

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.

Che, Yeon-Koo and Konrad Mierendorff (2019). "Optimal Dynamic Allocation of Attention." In: *American Economic Review* 109.8, pp. 2993–3029.

Chen, Daniel and Weijie Zhong (2024). "Information Acquisition and Time-Risk Preference." In.

Chen, M Keith (2013). "The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets." In: *American Economic Review* 103.2, pp. 690–731.

Chesson, Harrell W and W Kip Viscusi (2003). "Commonalities in time and ambiguity aversion for long-term risks." In: *Theory and Decision* 54.1, pp. 57–71.

Dasgupta, Partha and Joseph Stiglitz (1980). "Uncertainty, industrial structure, and the speed of R&D." In: *The Bell Journal of Economics*, pp. 1–28.

DeJarnette, Patrick et al. (2020). "Time lotteries and stochastic impatience." In: *Econometrica* 88.2, pp. 619–656.

Dixit, Avinash K and Robert S Pindyck (1994). *Investment under uncertainty*. Princeton university press.

Doval, Laura and Vasiliki Skreta (2022). *Constrained Information Design*. arXiv: 1811.03588 [econ.TH].

Feng, Han and David Hobson (2015). "Gambling in contests modelled with diffusions." In: *Decisions in Economics and Finance* 38.1, pp. 21–37.

— (2016a). "Gambling in contests with random initial law." In: *The Annals of Applied Probability*, pp. 186–215.

— (2016b). "Gambling in contests with regret." In: *Mathematical Finance* 26.3, pp. 674–695.

Fudenberg, Drew, Whitney Newey, et al. (2020). "Testing the drift-diffusion model." In: *Proceedings of the National Academy of Sciences* 117.52, pp. 33141–33148.

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.

Georgiadis-Harris, Alkis (2024). *Preparing to Act*. Tech. rep. Working Paper.

Hébert, Benjamin and Michael Woodford (2023). "Rational inattention when decisions take time." In: *Journal of Economic Theory* 208, p. 105612.

Hébert, Benjamin and Weijie Zhong (2022). "Engagement maximization." In: *arXiv preprint arXiv:2207.00685*.

Jacod, Jean and Albert Shiryaev (2013). *Limit theorems for stochastic processes*. Vol. 288. Springer Science & Business Media.

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

Knoepfle, Jan (2020). *Dynamic Competition for Attention*. Tech. rep. University of Bonn and University of Mannheim, Germany.

Koh, Andrew and Sivakorn Sanguanmoo (2024). "Attention capture." In: *arXiv preprint arXiv:2209.05570*.

Lee, Tom and Louis L Wilde (1980). "Market structure and innovation: A reformulation." In: *The Quarterly Journal of Economics* 94.2, pp. 429–436.

Luenberger, David G (1997). *Optimization by vector space methods*. John Wiley & Sons.

Mayskaya, Tatiana (2022). "Dynamic choice of information sources." In: *California Institute of Technology Social Science Working Paper, ICEF Working Paper WP9/2019/05*.

Nikandrova, Arina and Romans Pancs (2018). "Dynamic project selection." In: *Theoretical Economics* 13.1, pp. 115–143.

Nutz, Marcel and Yuchong Zhang (2022). "Reward design in risk-taking contests." In: *SIAM Journal on Financial Mathematics* 13.1, pp. 129–146.

Obłój, Jan (Jan. 2004). "The Skorokhod embedding problem and its offspring." In: *Probability Surveys* 1.none. ISSN: 1549-5787. DOI: 10.1214/154957804100000060. URL: http://dx.doi.org/10.1214/154957804100000060.

Onay, Selçuk and Ayse Öncüler (2007). "Intertemporal choice under timing risk: An experimental approach." In: *Journal of Risk and Uncertainty* 34.2, pp. 99–121.

Park, Andreas and Lones Smith (2008). "Caller number five and related timing games." In: *Theoretical Economics* 3.2, pp. 231–256.

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

Ratcliff, Roger, Philip L Smith, et al. (2016). "Diffusion decision model: current issues and history." In: *Trends in cognitive sciences* 20.4, pp. 260–281.

Reinganum, Jennifer F (1989). "The timing of innovation: Research, development, and diffusion." In: *Handbook of industrial organization* 1, pp. 849–908.

Seel, Christian and Philipp Strack (2013). "Gambling in contests." In: *Journal of Economic Theory* 148.5, pp. 2033–2048.

— (2016). "Continuous time contests with private information." In: *Mathematics of Operations Research* 41.3, pp. 1093–1107.

Skorokhod, Anatoliy Vladimirovich (1982). *Studies in the theory of random processes*. Vol. 7021. Courier Dover Publications.

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

Zhong, Weijie (2018). *Information Design Possibility Set*. arXiv: 1804.05752 [math.OC].

— (2022). "Optimal dynamic information acquisition." In: *Econometrica* 90.4, pp. 1537–1582.

A Proof of Theorem 1

Necessity of Equation (2): For every admissible strategy $(\langle\mu_t\rangle, \tau) \in \mathcal{M}$, the variation constraint (1) implies that $H(\mu_t) - \chi \cdot t$ is a supermartingale. Then, applying the optional stopping theorem to $H(\mu_t) - \chi \cdot t$ w.r.t. stopping time $\min\{t, \tau\}$ yields an accounting inequality:

$$\mathbb{E}\left[H(\mu_{\min\{t,\tau\}}) - H(\mu_0)\middle|\mathcal{F}_0\right] \leq \chi \cdot \mathbb{E}[\min\{t,\tau\}]. \tag{27}$$

Let f be the joint probability measure of μ_τ and τ, Equation (27) implies

$$\begin{aligned}
&\mathbb{E}_{t<\tau}[H(\mu_t) - H(\mu_0)] + \mathbb{E}_{t\geq\tau}[H(\mu_\tau) - H(\mu_0)] \leq \chi \cdot \mathbb{E}[\min\{t,\tau\}] \\
\Longrightarrow\ & H(\mathbb{E}[\mu_t|t<\tau])\mathcal{P}(t<\tau) + \mathbb{E}_{t\geq\tau}[H(\mu_\tau)] - H(\mu_0) \leq \chi \cdot \mathbb{E}[\min\{t,\tau\}] \qquad (28) \\
\Longleftrightarrow\ & H\left(\frac{\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}\tau)}{\int_{\tau>t} f(\mathrm{d}\mu,\mathrm{d}\tau)}\right)\int_{\tau>t}\int_S f(\mathrm{d}\mu,\mathrm{d}\tau) + \int_{\tau\leq t}\int_S H(\mu)f(\mathrm{d}\mu,\mathrm{d}\tau) - H(\mu_0) \qquad (29) \\
&\leq \chi\cdot\int\min\{t,\tau\}f(\mathrm{d}\mu,\mathrm{d}\tau), \\
\Longleftrightarrow\ & H\left(\int_{s>t}\mu f(\mathrm{d}\mu,\mathrm{d}s)\right) + \int_{\tau\leq t}H(\mu)f(\mathrm{d}\mu,\mathrm{d}\tau) - H(\mu_0) \leq \chi\cdot\int\min\{t,\tau\}f(\mathrm{d}\mu,\mathrm{d}\tau).
\end{aligned}$$

The second inequality is from H being convex. The third inequality is from the optional stopping theorem and $\langle \mu_t \rangle$ being a martingale. The last inequality is from H being HD-1.

Sufficiency of Equation (2): We have shown that Equation (2) is a necessary condition for $f \in \mathbb{F}$. In what follows, we prove Theorem 1 by proving a slightly stronger sufficiency result. Let $\widetilde{\mathcal{M}}$ denote the collection of admissible processes and stopping times corresponding to $T = \mathbb{R}_+$. Define

$$\widetilde{\mathbb{F}} = \left\{ f \in \Delta(S \times T) \middle| \exists (\langle \mu_t, \tau \rangle) \in \widetilde{\mathcal{M}} \text{ s.t. } f \sim (\mu_\tau, \tau) \right\}.$$

$\widetilde{\mathcal{M}}$ extends the definition of the martingale and stopping time from T to $\mathbb{R}_+$. $\widetilde{\mathbb{F}}$ is the subset of embeddable distributions supported on $S \times T$. Note that $\widetilde{\mathbb{F}} \subset \mathbb{F}$ since any pair $(\langle \mu_t \rangle, \tau) \subset \widetilde{\mathcal{M}}$ such that $\operatorname{supp}(\tau) \subset T$ has projection $(\langle \mu_t \rangle_{t \in T}, \tau) \in \mathcal{M}$. In what follows, we prove Theorem 1 by showing Equation (2) is a sufficient condition for $f \in \widetilde{\mathbb{F}}$.[16]

For each n, discretize $\mathbb{R}_+$ to a finite grid $t \in T_n = \left\{ t_1 = 0, \dots, t_i^n, \dots, t_n^n \right\}$. The sequence of the grids (T_n) satisfies $\lim_{n\to\infty} \max_i \left\{ t_i^n - t_{i-1}^n \right\} \to 0$ and $\lim_{n\to\infty} t_n^n \to \infty$. Let $f_i^n = f(\tau \in (t_{i-1}^n, t_i^n])$ and $f_{n,i}(\nu) = f(\nu | \tau \in (t_{i-1}^n, t_i^n])$ for all $i < n$. Let $f_n^n = f(\tau \in (t_{n-1}^n, \infty)$ and

$$f_{n,n}(\nu) = \frac{f\left(\nu \middle| \tau \in (t_{n-1}^n, t_n^n]\right) + \delta_{\nu = \mathbb{E}_f[\nu | \tau > t_n^n]} f(\tau > t_n^n)}{f_n^n}.$$

In words, the discretized distribution f^n merges f within each interval $(t_{i-1}^n, t_i^n]$ and assigns the merged mass to the right end of the interval t_i^n. As a result, this operation only relaxes the constraints specified by Equation (2) for any $t \le t_{n-1}^n$. For the last interval $(t_{n-1}^n, t_n^n]$, it follows that

$$\begin{aligned}
\sum_{i=1}^n f_i^n \mathbb{E}_{f_{n,i}}[H(\nu)] - H(\mu_0) &= \int_0^{t_n^n} \int_S H(\nu) \mathrm{d} f(\nu, \tau) + f(\tau > t_n^n) H(\mathbb{E}_f[\nu | \tau > t_n^n]) - H(\mu_0) \\
&\le \chi \cdot \int \min\{\tau, t_n^n\} f(\mathrm{d}\mu, \mathrm{d}\tau) \\
&\le \chi \cdot \int \left(\sum_{i=1}^n \mathbf{1}_{\tau \in (t_{i-1}^n, t_i^n]} t_i^n + \mathbf{1}_{\tau \ge t_n^n} t_n^n \right) f(\mathrm{d}\mu, \mathrm{d}\tau) = \chi \cdot \sum_{i=1}^n t_i^n f_i^n
\end{aligned}$$

The first inequality is from Equation (2). The second inequality is from relaxing τ to the closest larger t_i. Then, f^n satisfies the conditions of Lemma 4; hence there exists a process $(\langle \mu_t^n \rangle, \tau^n) \in \widetilde{\mathcal{M}}$ that satisfies $(\mu_{\tau^n}^n, \tau^n) \sim f^n$.

Since f diminishes at infinity, the sequence of $(\langle \mu_t^n \rangle, \tau^n)$ satisfies the conditions of Lemma 6; hence, a limit point exists under the weak topology and $(\mu_{\tau^n}^n, \tau^n) \xrightarrow{d} (\mu_\tau, \tau)$. By construction, $(\mu_{\tau^n}^n, \tau^n) \sim f^n \xrightarrow{w} f$; hence, $(\mu_\tau, \tau) \sim f$. *Q.E.D.*

[16] Note that when $f \in \Delta(S \times T)$ satisfies Equation (2) for all $t \in T$, Equation (2) is automatically satisfied for all $t \in \mathbb{R}_+$, which will be taken as given in the subsequent analysis.

Lemma 4. *$f \in \Delta(S \times \mathbb{N})$ has discrete and finite support on the time dimension and f satisfies Equation (2) for all $t \in \mathbb{N}$; then, $f \in \widetilde{\mathbb{F}}$.*

Proof. Let the support of f on the time dimension be $\{t_i\}_{i=1}^n$. We prove by induction on n. When $n = 1$, the lemma is trivially true. Now, we assume by induction that the statement is proved for $n = k - 1$.

For notational simplicity, let $f^i = \mathbb{E}_f\left[\mathbf{1}_{t=t_i}\right]$ and $f_i(\nu) = \frac{f(\nu,t_i)}{f^i}$. Let $\delta t_i = t_i - t_{i-1}$. Equation (2) implies:

$$\sum_1^k f^i \mathbb{E}_{f_i}[H(\nu)] - H(\mu_0) \leq \sum_1^k \left(\sum_{j=i}^k f^j\right) I\delta t_i;$$

Now, apply Lemma 5 to construction a process that implements f_k and scale time by $\max\left\{\delta t_k, \frac{\mathbb{E}_\pi[H(\nu)-H(\mu)]}{I}\right\}$. Then, the process $\langle \mu_t \rangle$ is defined for $t \in [0, \max\left\{\delta t_k, \frac{\mathbb{E}_\pi[H(\nu)-H(\mu)]}{I}\right\}]$ and satisfies $\mathbb{E}\left[\frac{\mathrm{d}H(\mu_t)}{\mathrm{d}t}\middle|\mathcal{F}_t\right] = I$.

Case 1: $\delta t_k \leq \frac{\mathbb{E}_\pi[H(\nu)-H(\mu)]}{I}$, i.e., the time it takes to implement f_k from a degenerate starting state is longer than the k^{th} period. Then, by construction,

$$\mathbb{E}\left[\mathbb{E}\left[H(\mu_{\frac{\mathbb{E}_\pi[H(\nu)-H(\mu)]}{I}}) - H(\mu_{\frac{\mathbb{E}_\pi[H(\nu)-H(\mu)]}{I}-\delta t_k})\right]\middle|\mathcal{F}_{\frac{\mathbb{E}_\pi[H(\nu)-H(\mu)]}{I}-\delta t_k}\right] = I\delta t_k.$$

Let $\widetilde{\pi}$ be the distribution of $\mu_{\frac{\mathbb{E}_\pi[H(\nu)-H(\mu)]}{I}-\delta t_k}$. Therefore, we obtain a martingale process that starts with interim state distribution $\widetilde{\pi}$ and implements f_k within the k^{th} period.

Case 2: $\delta t_k > \frac{\mathbb{E}_\pi[H(\nu)-H(\mu)]}{I}$. Then, let $\widetilde{\pi} = \delta_{\mu=\mathbb{E}_{f_k}[\nu]}$.

Note that by the construction of $\widetilde{\pi}$:

$$\mathbb{E}_{f_k}[H(\nu)] - \mathbb{E}_{\widetilde{\pi}}[H(\nu)] \geq I\delta t_k \tag{30}$$

Let $\widetilde{f}^{k-1} = f^{k-1} + f^k$ and

$$\widetilde{f}_{k-1}(\nu) = \frac{f^{k-1} \cdot f_{k-1}(\nu) + f^k \cdot \widetilde{\pi}(\nu)}{f^{k-1} + f^k}.$$

In words, we redefine $\widetilde{f}_{k-1}$ as the total measure of f_{k-1} and "*f_k pushed back in time by t_k periods, following the trajectories specified by $\langle \mu_t \rangle$*". Let $\widetilde{f}$ be otherwise defined identically to f at other times. Then:

$$\begin{aligned}
\sum_1^{k-1} \widetilde{f}^i \mathbb{E}_{\widetilde{f}_i}[H(\nu)] - H(\mu_0) &= \sum_1^{k-1} f^i \mathbb{E}_{f_i}[H(\nu)] + f^k \mathbb{E}_{\widetilde{\pi}}[H(\nu)] - H(\mu_0) \\
&= \sum_1^k f^i \mathbb{E}_{f_i}[H(\nu)] - H(\mu_0) + f^k \mathbb{E}_{\widetilde{\pi}}[H(\nu)] - f^k \mathbb{E}_{f^k}[H(\nu)] \\
&\leq \sum_1^k \left(\sum_{j=i}^k f^j\right) I\delta t_i - f^k I\delta t_k = \sum_1^{k-1} \left(\sum_{j=i}^{k-1} \widetilde{f}^j\right) I\delta t_i.
\end{aligned}$$

The third equality is from Equations (2) and (30). Therefore, we verify Equation (2) for $\widetilde{f}$

for $k-1$. $\forall i < k-1$,

$$\begin{aligned}
&\sum_{j=1}^{i}\widetilde{f}^j\mathbb{E}_{\widetilde{f}_j}[H(\nu)] + H\left(\frac{\sum_{j=i+1}^{k-1}\widetilde{f}^j\mathbb{E}_{\widetilde{f}_j}[\nu]}{\sum_{j=i+1}^{k-1}\widetilde{f}^j}\right)\left(\sum_{j=i+1}^{k-1}\widetilde{f}^j\right) - H(\mu_0)\\
=&\sum_{j=1}^{i} f^j\mathbb{E}_{f_j}[H(\nu)] + H\left(\frac{\sum_{j=i+1}^{k} f^j\mathbb{E}_{f_j}[\nu]}{\sum_{j=i+1}^{k} f^j}\right)\left(\sum_{j=i+1}^{k} f^j\right) - H(\mu_0)\\
\le&\sum_{j=1}^{i}\left(\sum_{\ell=j}^{k} f^\ell\right) I\delta t_j = \sum_{j=1}^{i}\left(\sum_{\ell=j}^{k-1}\widetilde{f}^\ell\right) I\delta t_j.
\end{aligned}$$

The inequality is from Equation (2) applied at $i < k-1$.

By induction, $\widetilde{f}$ is implementable by appending $\langle mu_t\rangle$ to $\langle\mu_t^k\rangle$ on interval $[t_{k-1}, t_k]$ conditional on an event such that $\mu_{t_{k-1}}^k \sim f^k\widetilde{\pi}$. The stopping time τ^k is t_k on the event and τ^{k-1} otherwise. The filtration of $\langle\mu_t^k\rangle$ within $[t_{k-1}, t_k]$ is then expanded by the natural filtration of $\langle\mu_t\rangle$ and the (binary) stopping event. *Q.E.D.*

Lemma 5 (Zhong 2022, Lemma S.1)**.** *$\forall\pi\in\Delta(S)$, let $\mu = \mathbb{E}_\pi[\nu]$. $H \in C(S)$ is strictly convex. There exists a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and stochastic process $\langle\mu_t\rangle_{t\in[0,1]}$ s.t.*

1. *$\langle\mu_t\rangle$ is a martingale;*
2. *$\mu_0 = \mu$ and $\mu_1 \sim \pi$;*
3. *$\forall t_1 < t_2 \in [0,1]$, $\mathbb{E}\left[H(\mu_{t_2}) - H(\mu_{t_1})\middle|\mathcal{F}_{t_1}\right] = (t_2 - t_1)\mathbb{E}_\pi\left[H(\nu) - H(\mu)\right]$.*

Lemma 6. *$\{(\langle\mu_t^n\rangle, \tau^n)\} \subset \widetilde{\mathcal{M}}$ satisfies: $\forall\varepsilon > 0$, $\exists\bar{t} > 0$ s.t. $\forall n$, $\mathbb{P}(\tau^n \le \bar{t}) \ge 1-\varepsilon$. Then, there exists $(\langle\mu_t\rangle, \tau) \in \widetilde{\mathcal{M}}$ s.t. $(\langle\mu_t^n\rangle, \tau) \xrightarrow{w} (\langle\mu_t\rangle, \tau)$ and $(\mu_{\tau^n}^n, \tau^n) \xrightarrow{d} (\mu_\tau, \tau)$.*

Proof. $\forall\{(\langle\mu_t^n\rangle, \tau^n)\} \subset \widetilde{\mathcal{M}}$, it is wlog to assume that μ_t^n is constant for $t \ge \tau^n$ per the optional stopping theorem. $\forall\eta > 0$, define a delayed stopping time $\tau_\eta^n = \tau^n + \eta$. Define process $z_{\eta t}^n = 0$ if $\tau_\eta^n > t$ and $z_{\eta t}^n = \pm 1$, with equal probability if $\tau_\eta^n \le t$. Then, $\langle\mu_t^n, z_{\eta t}^n\rangle$ is a martingale process in the space $S\times[-1,1]$. Let P^n be the joint probability measure of $\langle\mu_t^n, z_{\eta t}^n\rangle$ on the Skorokhod space D_∞.

Next, we prove that the collection $\{P^n\}$ is tight. It is sufficient to check tightness for each marginal distribution. It's trivial that $(z_{\eta t}^n)$ is tight since the process is bounded and the cadlag modulus is zero. It is trivial that $(\langle\mu_t^n\rangle)$ is uniformly bounded in S. Next, we verify the Aldou's tightness criterion for $(\langle\mu_t^n\rangle)$. Since H is strictly convex, $\forall\varepsilon,\delta > 0$,

$$\mathbb{P}\left[|\mu_{t+\delta}^n - \mu_t^n| \ge \varepsilon\middle|\mathcal{F}_t\right] \le \mathbb{P}\left[H(\mu_{t+\delta}^n) - H(\mu_t^n) - \nabla H(\mu_t^n)(\mu_{t+\delta}^n - \mu_t^n) \ge \xi\middle|\mathcal{F}_t\right],$$

where $\xi = \min\limits_{\mu,\nu\in S, |\nu-\mu|\ge\varepsilon}(H(\nu) - H(\mu) - \nabla H(\mu)(\nu-\mu)) > 0$. The variation constraint implies:

$$\mathbb{P}\left[H(\mu_{t+\delta}^n) - H(\mu_t^n) - \nabla H(\mu_t^n)(\mu_{t+\delta}^n - \mu_t^n) \ge \xi\middle|\mathcal{F}_t\right]\cdot\xi + \mathbb{P}\left[H(\mu_{t+\delta}^n) - H(\mu_t^n) - \nabla H(\mu_t^n)(\mu_{t+\delta}^n - \mu_t^n) \le \xi\middle|\mathcal{F}_t\right]\cdot 0$$

$\leq \mathbb{E}\left[H(\mu^n_{t+\delta}) - H(\mu^n_t)\middle|\mathcal{F}_t\right] \leq \chi \cdot \delta.$
Therefore,

$$\mathbb{P}\left[|\mu^n_{t+\delta} - \mu^n_t| \geq \varepsilon\middle|\mathcal{F}_t\right] \leq \frac{\chi \cdot \delta}{\xi} \xrightarrow{\delta \to 0} 0.$$

Aldou's theorem implies that $\{\langle \mu^n_t \rangle\}$ is a tight collection of measures on D_∞ (Theorem 16.9 and 16.10 of Billingsley 2013). By Prokhorov's theorem, there exists weak limit of (P^n) when $n \to \infty$ under the weak topology, denoted by $\langle \mu_t, z_{\eta t} \rangle$. By Proposition IX.1.1 of Jacod and Shiryaev 2013, since $(\langle \mu^n_t \rangle, z^n_{\eta t})$ are uniformly bounded, $\langle \mu_t, z_{\eta t} \rangle$ is a cadlag martingale. Let $\tau_\eta = \inf\{t \| z_{\eta t}| = 1\}$. [17]

Next, we prove that $\langle \mu_t \rangle$ satisfies Equation (1). $\forall A \in \mathcal{F}$, define $d_t(A) = \sup_{\omega, \omega' \in A} |\mu_t(\omega) - \mu_t(\omega')|$. $\forall \epsilon > 0$, let δ be the continuity parameter of H. $\forall A \in \mathcal{F}_t$ s.t. $d_t(A) \leq \frac{1}{2}\delta$, $\forall \mu \in \{\mu_t(A)\}$, $\forall t' > t$, then:

$$\begin{aligned}\mathbb{E}[H(\mu_{t'})|A] \leq & \varliminf_{n \to \infty} \mathbb{E}\left[H(\mu^n_{t'})|A\right] \\ \leq & \varliminf_{n \to \infty} \left(\mathbb{E}[H(\mu^n_t)|A] + (t'-t)\chi\right) \\ \leq & \mathbb{E}[H(\mu_t)|A] + 2\varepsilon + (t'-t)\chi\end{aligned}$$

The first inequality is Fatou's lemma. The second inequality is Equation (1) applied to $\langle \mu^n_t \rangle$ at t. The last inequality is from the continuity of H and $\mathrm{d}_t(A) \leq \frac{1}{2}\delta$. Since ε can be chosen arbitrarily small, this implies $\mathbb{E}[H(\mu_{t'}) - H(\mu_t)|\mathcal{F}_t] \leq (t'-t)\chi$.

Next, by Skorokhod representation theorem, there exists probability space $(\Omega, \mathcal{F}, \mathcal{P})$ s.t. $(\langle \mu^n_t \rangle, \tau^n_\eta)$ converges a.s. to $(\langle \mu_t \rangle, \tau_\eta)$. $\forall \omega$ s.t. $(\mu^n_t(\omega), \tau^n_\eta(\omega)) \to (\mu_t(\omega), \tau_\eta(\omega))$. Pick $\varepsilon < \eta$. $\exists N$ s.t. $\forall n > N$, $|\tau^n_\eta(\omega) - \tau_\eta(\omega)| < \varepsilon \implies \forall \tau$ s.t. $|\tau - \tau_\eta(\omega)| < \varepsilon$, $\tau > \tau^n(\omega)$. Therefore, $\mu^n_t(\omega)$ are constant in $(\tau_\eta(\omega) - \varepsilon, \tau_\eta(\omega) + \varepsilon)$. Then $d(\mu^n_t(\omega), \mu_t(\omega)) \to 0$ implies $\mu^n_{\tau^n(\omega)}(\omega) = \mu^n_{\tau_\eta(\omega)}(\omega) \to \mu_{\tau_\eta(\omega)}(\omega)$. This suggests that: $\left(\mu^n_{\tau^n_\eta}, \tau^n_\eta\right) \xrightarrow{a.s.} \left(\mu_{\tau_\eta}, \tau_\eta\right)$.

Define $\tau = \tau_\eta - \eta$. Note that the analysis in last paragraph implies that with probability one, $\mu_t(\omega)$ is constant within $(\tau_\eta(\omega) - \eta, +\infty)$. Since $\langle \mu_t \rangle$ is cadlag, $\mu_t(\omega)$ is constant within $[\tau_\eta(\omega) - \eta, \infty) = [\tau(\omega), \infty)$. Therefore, $\langle \mu_t \rangle$ is a martingale w.r.t. the natural filtration of $(\langle \mu_t \rangle, \tau)$; hence, $(\langle \mu_t \rangle, \tau) \in \widetilde{\mathcal{M}}$. Since $\mu^n_{\tau^n} = \mu^n_{\tau^n_\eta}$ and $\mu_\tau = \mu_{\tau_\eta}$, $\left(\mu^n_{\tau^n}, \tau^n\right) \xrightarrow{a.s.} (\mu_\tau, \tau)$, *Q.E.D.*

[17] Note that the Skorokhod metric between $z^n_{\eta t}(\omega)$ and $z^n_{\eta t}(\omega')$ is equivalent to $\min\{1, |\tau^n_\eta(\omega) - \tau^n_\eta(\omega')|\}$. Then the weak convergence of $(z^n_{\eta t})$ is equivalent to the weak convergence of (τ^n_η). Therefore, since (τ^n_η) is a tight collection of measures on $\mathbb{R}_+$, the limit τ_η is supported on $\mathbb{R}_+$.

B Proofs in Section 3

B.1 Proof of Lemma 1

We note in Lemma 8 that $\mathbb{F}$ is closed.[18] If T is bounded, then the statement is trivial as $\Delta(S\times T)$ is tight. We focus on the case where T is unbounded. Then, there exists an increasing sequence $(t_n)\subset T$ such that $t_n\to\infty$. Define

$$\widehat{u_n}(\mu):=\sup_{\substack{h\in\Delta(S\times T\cap[t_n,\infty)),\\ \mathbb{E}_h[\nu]=\mu}}\mathbb{E}_h[U].$$

Claim. *$\widehat{u_n}$ is upper semicontinuous. Moreover, the maximum is attained.*

Proof. We begin with defining $u_n(\mu):=\sup_{t\geq t_n}U(\mu,t)$. Since $\lim_{t\to\infty}\sup_\mu U(\mu,t)=0$, $u_n(\mu)$ must be attained by finite t, denote this mapping by $t=\widehat{t}(\mu)$. Moreover, $\forall\mu_m\to\mu$ such that $u_n(\mu_m)$ converges, if $\widehat{t}(\mu_m)$ is unbounded, then $u_n(\mu_m)\to 0\leq u_n(\mu)$. If $t(\mu_m)$ is bounded, then wlog we can pick $\widehat{t}(\mu_m)$ to be converging. Then, $\lim u_n(\mu_m)=U(\lim\mu_m,\lim\widehat{t}(\mu_m))\leq u_n(\mu)$. Therefore, $u_n(\mu)$ is upper semicontinuous.

Observe that $\widehat{u_n}(\mu)$ is the upper concave envelope of $u_n(\mu)$. By Caratheodory's theorem, $\forall\mu$ there exists a finite support probability measure (p_i,μ_i) that has mean μ and attains $\widehat{u_n}(\mu)$. Therefore, $h=(p_i,\mu_i,\widehat{t}(\mu_i))$ attains $\widehat{u_n}(\mu)$. Denote $h_n(\mu)$ a mapping from μ to a maximizer that attains $\widehat{u_n}(\mu)$ (invoking the axoim of choice).

Next, we prove upper semicontinuity. Suppose for the purpose of contradiction that $\mu_m\to\mu$ but $\lim\widehat{u_n}(\mu_m)\geq\widehat{u_n}(\mu)+\epsilon$ for some $\epsilon>0$. Then, since $U(\mu,t)\xrightarrow[t\to\infty]{u}0$, there exists $\bar{t}$ s.t. $U(\mu,t)<\epsilon/2$ for $t>\bar{t}$. This implies that $\forall m$, $\exists h_m\in\Delta(S\times T\cap[t_n,\bar{t}])$ that attains $\widehat{u_n}(\mu_m)-\epsilon/2$. Note that the collection of h_m is tight, hence $\mathbb{E}_{\lim h_m}[U]\geq\lim\widehat{u_n}(\mu_m)-\epsilon/2>\widehat{u_n}(\mu)$. This contradicts the definition of $\widehat{u_n}(\mu)$. *Q.E.D.*

Next, define

$$\widehat{U}_n(\mu,t)=\begin{cases}U(\mu,t) & \text{if } t<t_n\\ \widehat{u_n}(\mu) & \text{if } t=t_n\end{cases}$$

for $\mu\in S$, $t\in T\&t\leq t_n$. Obviously, $\widehat{U}_n\geq U$. Since U is bounded and continuous, $\widehat{U}_n$ is bounded and upper semicontinous. This implies that $\int\widehat{U}_n(\mu,t)f(\mathrm{d}\mu,\mathrm{d}t)$ is upper semicontinuous. Therefore,

$$\sup_{f\in\mathbb{F}}\mathbb{E}_f[\widehat{U}_n]$$

has a solution $f_n\in\Delta(S\times T\cap[0,t_n])$.

[18] The proof is straightforward as $\mathbb{F}$ is defined by a collection of weak inequality constraints; hence, it is relegated to the online appendix.

Consider the collection of $\{f_n\}$. Suppose it is tight, then there exists $f \in \mathbb{F}$ s.t. $f_n \to f$ and $\mathbb{E}_f[U] = \lim \mathbb{E}_{f_n}[\widehat{U}_n] \geq$ Equation (P). Therefore, f solves Equation (P).

Now consider the remaining case that $\{f_n\}$ is not tight, i.e., $\exists \epsilon > 0$ s.t. $\forall t$, $\exists n$ s.t. $f_n(S \times T \cap [t, \infty)]) > \epsilon$. Since t is arbitrary, pick $t' = \frac{\sup H - \inf H}{\chi \epsilon}$. Then $f_n(S \times T \cap [t', \infty)]) > \epsilon$ implies $t_n \geq t'$. Define

$$f := \begin{cases} f_n & \text{if } t < t_n \\ h_n(\mathbb{E}_{f_n}[\mu | t \geq t_n]) & \text{if } t \geq t_n. \end{cases}$$

By definition, $\mathbb{E}_f[U] = \mathbb{E}_{f_n}[\widehat{U}_n | t < t_n] + \widehat{u}_n(\mathbb{E}_{f_n}[\mu | t \geq t_n]) \geq \mathbb{E}_{f_n}[\widehat{U}_n] \geq$Equation (P). We verify that $f \in \mathbb{F}$. Equation (2) is obviously satisfied for $t < t_n$. For $t \geq t_n$,

$$\int_{s \leq t} \chi(1 - F(s)) \mathrm{d}s \geq \chi \cdot t_n \cdot \epsilon \geq \sup H - \inf H.$$

The RHS is an obvious upper bound for the LHS of Equation (2). Therefore, f solves Equation (P). *Q.E.D.*

B.2 Proof of Lemma 2

Lemma 2 is trivially true when T is finite. We provide the proof when T is a compact interval. To prove Lemma 2, we invoke Theorem 1, chapter 8.6 of Luenberger 1997. Let $\Delta_{\mu_0}^C := \left\{ f \in \Delta_{\mu_0} | G(f)(t) \in UC(T^\circ) \right\}$, termed the *time-continuous* subset of Δ_{μ_0}.[19] We verify all conditions of the cited theorem, applied to $UC(T^\circ)$ and its dual space $\mathcal{B}(T^\circ)$. First, the objective functional $\int U(\mu, \tau) f(\mathrm{d}\mu, \mathrm{d}\tau)$ is a real-valued linear and continuous functional of f. $\Delta_{\mu_0}^C$ is a convex subset of the vector space of all probability measures on $(S \times T)$.

Next, we verify that G is a concave mapping of $\Delta_{\mu_0}^C$ into $UC(T^\circ)$. $\forall f_1, f_2 \in \Delta_{\mu_0}^C$, $\forall \alpha \in [0, 1]$, $\forall t \in T^\circ$,

$$\begin{aligned} -H\left(\int_{\tau > t} \mu(\alpha f_1 + (1 - \alpha) f_2)(\mathrm{d}\mu, \mathrm{d}\tau) \right) &= -H\left(\alpha \int_{\tau > t} \mu f_1(\mathrm{d}\mu, \mathrm{d}\tau) + (1 - \alpha) \int_{\tau > t} \mu f_2(\mathrm{d}\mu, \mathrm{d}\tau) \right) \\ &\geq -\alpha H\left(\int_{\tau > t} \mu f_1(\mathrm{d}\mu, \mathrm{d}\tau) \right) - (1 - \alpha) H\left(\int_{\tau > t} \mu f_2(\mathrm{d}\mu, \mathrm{d}\tau) \right). \end{aligned}$$

The inequality is from the convexity of H. This verifies the concavity of the only non-linear term in $G(\cdot)(t)$. Hence, G is concave.

Next, we verify that there exists $f \in \Delta_{\mu_0}^C$ s.t. $G(f)(\cdot)$ is an interior point of the positive cone. Let Let $f \sim \mathbf{1}_{\mu = \mu_0} \times U(T) \times \alpha + \mathbf{1}_{\mu = \mu_0, t = \sup T} \times (1 - \alpha)$, where $0 < \alpha < 1$ and $U(T)$ is the uniform distribution on T. In words, f stops uniformly on T at μ_0 with probability α and stops on $\sup T$ at μ_0 with probability $1 - \alpha$. By definition, $f \in \Delta_{\mu_0}^C$. $\forall t \in T^\circ$,

$$G(f)(t) \geq \frac{1}{t} (\chi \cdot t \cdot (1 - \alpha)) = (1 - \alpha) \chi.$$

[19] $UC(X)$ denotes all uniformly continuous fuctions on X. Note that $G(f)(t)$ is not defined at 0 and $\sup T$. We slightly abuse notation and define it as its continuous extension (which always exists).

Therefore, $\forall h \in UC(T)$ s.t. $\|h - G(f)\| < (1-\alpha)\chi$, $h \geq 0$; hence, $G(f)$ is an interior point. Then, the cited theorem implies

$$\sup_{f\in\Delta^C_{\mu_0}, G(f)\geq 0} \int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau) = \min_{\zeta\in\mathcal{B}(T^\circ)} \sup_{f\in\Delta^C_{\mu_0}} \left(\int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau) + \int G(f)(t)\mathrm{d}\zeta(t) \right), \quad (31)$$

where there exists $\zeta^* \in \mathcal{B}(T^\circ)$ achieving the RHS of (31). Note that $\zeta \in \mathcal{B}(T^\circ)$ equivalently defines $\lambda \in \mathbb{L}$ s.t. the radon-nikodym derivative $\frac{\mathrm{d}\zeta}{\mathrm{d}\lambda} = t$. Therefore, (31) is equivalent to

$$\sup_{f\in\Delta^C_{\mu_0}, G(f)\geq 0} \int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau) = \min_{\lambda\in\mathbb{L}} \sup_{f\in\Delta^C_{\mu_0}} \mathcal{L}(f,\lambda), \quad (32)$$

where there exists $\lambda^* \in \mathbb{L}$ achieving the minimum on the RHS. If f^* achieves the maximum on the LHS, then

$$\begin{cases} f^* \in \arg\max_{f\in\Delta^C_{\mu_0}} \mathcal{L}(f,\lambda^*); \\ \int_{t\in T^\circ} G(f^*)(t)\lambda^*(t)t\mathrm{d}t = 0. \end{cases} \quad (33)$$

$\forall f \in \Delta_{\mu_0}$, $\forall \epsilon > 0$, Lemma 7 implies that there exists $f' \in \Delta^C_{\mu_0}$ s.t. $\mathcal{L}(f',\lambda) \geq \mathcal{L}(f,\lambda) - \epsilon$ for all λ. Therefore, Equations (32) and (33) still hold when $\Delta^C_{\mu_0}$ is replaced by Δ_{μ_0}, which establishes strong duality. *Q.E.D.*

Lemma 7. *Suppose T is a compacty interval, $\forall f \in \Delta_{\mu_0}$, $\forall \epsilon > 0$, there exists $\widehat{f} \in \Delta^C_{\mu_0}$ s.t. $d_{lp}(\widehat{f}, f) \leq \epsilon$ and $G(\widehat{f}) \geq G(f) - \epsilon$.*[20]

Proof. $\forall f \in \Delta_{\mu_0}$, $G(f)(t)$ has bounded variation and only jumps down. Therefore, $G(f)(t)$ can be decomposed into $g(t)+h(t)$, where g is bounded and continuous and h is bounded and decreasing. Define the "delayed" measure

$$f^s(\mu,t) := \begin{cases} 0 & t < s \\ f(\mu, t-s) & t \in [s, \sup(T)) \\ f(\mu, [t-s, \sup(T)]) & t = \sup(T) \end{cases}$$

In words, f^s delay the distribution of f by s. By definition, $d_{lp}(f^s, f) \xrightarrow{s\to 0} 0$. Pick $\delta > 0$ s.t. $|U(\cdot,t) - U(\cdot,t-s)| < \epsilon$, $d_{lp}(f^s,f) < \epsilon$, and $|g(t) - g(t+s)| < \frac{1}{2}\epsilon$ when $s < \delta$. Then,

$$\begin{aligned} G(f^s)(t) \geq & G(f)(t-s) \\ = & g(t-s) + h(t-s) \\ \geq & g(t) + h(t)\epsilon \\ = & G(f)(t) - \epsilon. \end{aligned}$$

Let $\widehat{f}$ be the uniform randomization of f^s, for $s \in [\frac{1}{2}\delta, \delta]$. Then, $d_{lp}(\widehat{f}, f) \leq \epsilon$. Since G is a concave operator of f, $G(\widehat{f})(t) \geq G(f)(t) - \epsilon$.

[20] d_{lp} is the Levy-Prokhorov metric.

Next, we prove the uniform continuity of $G(\widehat{f})$. Note that $\forall t < t' < t + \frac{1}{2}\delta$,

$$\begin{aligned}\widehat{f}((t,t']) =& \frac{2}{\delta}\int_{s=\frac{1}{2}\delta}^{\delta}(F(t'-s)-F(t-s))\mathrm{d}s \\ \leq& \frac{2}{\delta}\int_{(t-\frac{1}{2}\delta,t'-\frac{1}{2}\delta]\cup(t-\delta,t'-\delta]}F(s)\mathrm{d}s \\ \leq& \frac{4|t-t'|}{\delta}.\end{aligned}$$

When $t \in [0, \frac{1}{2}\delta]$, by construction, $G(\widehat{f})(t) \equiv \chi$. $\forall \epsilon > 0$, let γ be the continuity parameter of H corresponding to ϵ. When $t, t' > \frac{1}{2}\delta$ and $|t - t'| \leq \delta\gamma/4$,

$$|G(\widehat{f})(t)t - G(\widehat{f})(t')t'| \leq |t-t'|\chi + |t-t'|\max|H| + \epsilon$$

Therefore, $G(\widehat{f})(t)\cdot t$ is uniformly continuous for $t \geq \frac{1}{2}\delta$. Since $1/t$ is a uniformly continuous function when $t \geq \frac{1}{2}\delta$, so is $G(\widehat{f})(t)$. To sum up, $G(\widehat{f}) \in UC(T)$. *Q.E.D.*

B.3 Proof of Theorem 2

Sufficiency: Suppose for the purpose of contradiction that (f, λ, a) satisfies Equations (9) and (10) and the complimentary slackness condition, but f is suboptimal in (P). Then, there exists g s.t. $\mathcal{L}(f,\lambda) < \mathcal{L}(g,\lambda)$. Then, since $\mathcal{L}$ is concave, $\forall \alpha \in (0,1)$,

$$\begin{aligned}
& \frac{\mathcal{L}(\alpha g + (1-\alpha)f, \lambda) - \mathcal{L}(f,\lambda)}{\alpha} \geq \mathcal{L}(g,\lambda) - \mathcal{L}(f,\lambda) \\
\implies & \lim_{\alpha\to 0}\frac{\mathcal{L}(\alpha g + (1-\alpha)f, \lambda) - \mathcal{L}(f,\lambda)}{\alpha} > 0 \\
\iff & \int U(\mu,\tau)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) + \int_{t\in T^\circ}\left(\chi\cdot\int \min\{t,\tau\}(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) - \int_{\tau\leq t}H(\mu)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)\right)\mathrm{d}\lambda(t) \\
& -\overline{\lim_{\alpha\to 0}}\int_{t\in T^\circ}\underbrace{\frac{H\left(\int_{\tau>t}\mu(\alpha g+(1-\alpha)f)(\mathrm{d}\mu,\mathrm{d}t)\right) - H\left(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}t)\right)}{\alpha}}_{\geq y\cdot\int_{\tau>t}\mu(g-f)(\mathrm{d}\mu,\mathrm{d}\tau),\ \forall y\in\nabla H(\widehat{\mu}_t),\text{ by convexity of }H.}\mathrm{d}\lambda(t) > 0 \\
\implies & \int U(\mu,\tau)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) + \int_{t\in T^\circ}\left(\chi\cdot\int \min\{t,\tau\}(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) - \int_{\tau\leq t}H(\mu)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)\right)\mathrm{d}\lambda(t) \\
& -\int_{t\in T^\circ}\nabla H(\widehat{\mu}_t)\cdot\int_{\tau>t}\mu(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\lambda(t) > 0 \\
\iff & \int l_{f,\lambda}(\mu,\tau)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) > 0 \\
\implies & \int l_{f,\lambda}(\mu,\tau)g(\mathrm{d}\mu,\mathrm{d}\tau) > a\cdot\mu_0.
\end{aligned}$$

The last line contradicts $l_{f,\lambda}(\mu,\tau) \leq a\cdot\mu$. Note that the selection of $\nabla H(\cdot)$ is arbitrary.

Necessity: Suppose $f \in \arg\max_{f \in \Delta_{\mu_0}} \mathcal{L}(f, \lambda)$. $\forall \mu$, when $\tau \geq \bar{t}$, select $\nabla H(0)$ s.t. $\nabla H(0) \cdot \mu = H(\mu)$. Since $l_{f,\lambda}(\mu, \tau) \in L^\infty(S \times T^\circ)$,

$$\widehat{l}(\mu) = \sup_{\pi \in \Delta(S), \mathbb{E}_\pi[\mu] = \mu_0} \mathbb{E}_\pi[\sup_{\tau \in T^\circ} l_{f,\lambda}(\mu, \tau)]$$

is a well-defined real-valued concave function on S. Let $a \cdot \mu$ be the tangent hyperplane of $\widehat{l}$ at μ_0. Evidently, $l_{f,\lambda}(\mu, \tau) \leq a \cdot \mu$.

Next, we prove that $\int l_{f,\lambda}(\mu, \tau) f(\mathrm{d}\mu, \mathrm{d}\tau)) = a \cdot \mu_0$. Suppose for the purpose of contradiction that $\int l_{f,\lambda}(\mu, \tau) f(\mathrm{d}\mu, \mathrm{d}\tau) < a \cdot \mu_0$. Then, since $l_{f,\lambda}(\mu, \tau) \leq a \cdot \mu$, there exists an open set and $\epsilon > 0$ s.t. $\inf O > 0$ and $\int_O (a \cdot \mu - l_{f,\lambda}(\mu, \tau)) f(\mathrm{d}\mu, \mathrm{d}\tau) > \epsilon$. Let $(\mu, \tau) = \mathbb{E}_f[(\mu', \tau')|O]$.[21] Then, $\mathbb{E}_f[l_{f,\lambda}(\mu', \tau')|O] < a \cdot \mu - \epsilon$.

By the definition of $\widehat{l}$, there exists a finite support distribution $\pi \in \Delta(S)$ s.t. $\mathbb{E}_\pi[\sup_{\tau \in T^\circ} l_{f,\lambda}(\mu, \tau)] > \widehat{l}(\mu) - \varepsilon/4$. For each μ' in the support of π, there exists τ' s.t. $l_{f,\lambda}(\mu', \tau') > \sup_{\tau \in T^\circ} l_{f,\lambda}(\mu', \tau) - \frac{1}{4}\epsilon$. We slightly abuse notation and let π denote the distribution of such (μ', τ') pairs. Therefore, $\mathbb{E}_\pi[l_{f,\lambda}(\mu', \tau')] > a \cdot \mu - \frac{\epsilon}{2}$.

Define $g = f + f(O) \cdot (\pi - f|_O)$. Then,

$$\begin{aligned}
&\lim_{\alpha \to 0} \frac{\mathcal{L}(\alpha g + (1-\alpha) f, \lambda) - \mathcal{L}(f, \lambda)}{\alpha} \\
=& \int U(\mu, \tau)(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) + \int_{t \in T^\circ} \left(\chi \cdot \int \min\{t, \tau\}(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) - \int_{\tau \leq t} H(\mu)(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) \right) \mathrm{d}\lambda(t) \\
&- \overline{\lim_{\alpha \to 0}} \int_{t \in T^\circ} \frac{H\left(\int_{\tau > t} \mu(\alpha g + (1-\alpha) f)(\mathrm{d}\mu, \mathrm{d}\tau)\right) - H\left(\int_{\tau > t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right)}{\alpha} \mathrm{d}\lambda(t) \\
=& \int U(\mu, \tau)(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) + \int_{t \in T^\circ} \left(\chi \cdot \int \min\{t, \tau\}(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) - \int_{\tau \leq t} H(\mu)(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) \right) \mathrm{d}\lambda(t) \\
&- \overline{\lim_{\alpha \to 0}} \int_{t < \bar{t}} \frac{H\left(\int_{\tau > t} \mu(\alpha g + (1-\alpha) f)(\mathrm{d}\mu, \mathrm{d}t)\right) - H\left(\int_{\tau > t} \mu f(\mathrm{d}\mu, \mathrm{d}t)\right)}{\alpha} \mathrm{d}\lambda(t) \\
&- \overline{\lim_{\alpha \to 0}} \int_{t \geq \bar{t}} H\left(\int_{\tau > t} \mu g(\mathrm{d}\mu, \mathrm{d}\tau) \right) \mathrm{d}\lambda(t) \\
\geq& \int U(\mu, \tau)(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) + \int_{t \in T^\circ} \left(\chi \cdot \int \min\{t, \tau\}(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) - \int_{\tau \leq t} H(\mu)(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) \right) \mathrm{d}\lambda(t) \\
&- \int_{t < \bar{t}} \nabla H(\widehat{\mu}_t) \cdot \int_{\tau > t} \mu(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) \mathrm{d}\lambda(t) - \int_{t \geq \bar{t}} \int_{\tau > t} H(\mu) g(\mathrm{d}\mu, \mathrm{d}\tau) \mathrm{d}\lambda(t) \\
=& \int U(\mu, \tau)(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) + \int_{t \in T^\circ} \left(\chi \cdot \int \min\{t, \tau\}(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) - \int_{\tau \leq t} H(\mu)(g - f)(\mathrm{d}\mu, \mathrm{d}\tau) \right) \mathrm{d}\lambda(t) \\
&- \int \left(\int_{t < \min\{\tau, \bar{t}\}} \nabla H(\widehat{\mu}_t) \cdot \mu \mathrm{d}\lambda(t) + \int_{t \in [\bar{t}, \tau)} \nabla H(0) \cdot \mu \mathrm{d}\lambda(t) \right) (g - f)(\mathrm{d}\mu, \mathrm{d}\tau)
\end{aligned}$$

[21] Wlog, O can be chosen that $a \cdot \mu = \hat{l}(\mu)$ since O can be arbitrarily close to the whole domain.

$$
\begin{aligned}
&= f(O) \cdot \int l_{f,\lambda}(\mu,\tau)(\pi - f|_O)(\mathrm{d}\mu, \mathrm{d}\tau) \\
&> f(O) \cdot \left(a\mu - \frac{\epsilon}{2} - (a \cdot \mu - \epsilon) \right) > 0.
\end{aligned}
$$

The last line contradicts $f \in \arg\max_{f \in \Delta_{\mu_0}} \mathcal{L}(f, \lambda)$. *Q.E.D.*

Online Appendix to *Exploration and Stopping*

Yuliy Sannikov **Weijie Zhong**

C Proofs in Section 3

C.1 Lemmas for Lemma 1

Lemma 8. $\mathbb{F}$ *is closed under weak topology.*

Proof. We prove by showing that $\mathbb{F}^C$ is open. Since H is continuous on a compact set, it is bounded. WLOG, let H be non-negative. Define

$$\chi_{t_1,t_2}(t) = \begin{cases} 0 & \text{when } t > t_2 \\ \dfrac{t_2 - t}{t_2 - t_1} & \text{when } t \in [t_1, t_2] \\ 1 & \text{when } t < t_1 \end{cases}$$

for any $t_1 < t_2$. Note that $\chi_{t,t'}(\tau)$ is bounded and continuous and $\mathbf{1}_{\tau \le t} \le \chi_{t,t'}(\tau) \le \mathbf{1}_{\tau \le t'}$. $\forall f \notin \widehat{\mathbb{F}}$, there exists t s.t.

$$H\left(\int_{\tau>t}\int_S \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) + \int_{\tau \le t}\int_S H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) - H(\mu_0) > \chi \cdot \int_{s \le t} (1 - F(s))\mathrm{d}s.$$

Since F is right-continuous in t. Therefore, $\exists t' > t, \varepsilon > 0$ s.t.

$$H\left(\int_{\tau>t}\int_S \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) + \int_{\tau \le t}\int_S H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) - H(\mu_0) \ge \chi \cdot \int_{s \le t'} (1 - F(s))\mathrm{d}s + \varepsilon. \tag{34}$$

Now, consider

$$\begin{aligned}
&\int \chi_{t,t'}(\tau) H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) + H\left(\int (1 - \chi_{t,t'}(\tau)) \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) \\
=& \int \mathbf{1}_{\tau \le t} H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) + \int (\chi_{t,t'}(\tau) - \mathbf{1}_{\tau \le t}) H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) \\
&+ H\left(\int \mathbf{1}_{\tau > t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau) - \int (\chi_{t,t'}(\tau) - \mathbf{1}_{\tau \le t}) \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) \\
\ge& \int \mathbf{1}_{\tau \le t} H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) + \int (\chi_{t,t'}(\tau) - \mathbf{1}_{\tau \le t}) H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) \\
&+ H\left(\int \mathbf{1}_{\tau > t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) - H\left(\int (\chi_{t,t'}(\tau) - \mathbf{1}_{\tau \le t}) \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right)
\end{aligned}$$

$$\geq \int \mathbf{1}_{\tau \leq t} H(\mu) f(\mathrm{d}\mu, \mathrm{d}\tau) + H\left(\int \mathbf{1}_{\tau > t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right)$$
$$\geq \chi \cdot \int_{s \leq t'} (1 - F(s))\mathrm{d}s + H(\mu_0) + \varepsilon.$$

The first inequality is from H being convex and HD1. The second inequality is from the convexity of H. The last inequality is Equation (34). Since $\chi_{t,t'}(\tau)H(\mu)$ and $\chi_{t,t'}(\tau)\mu$ are both bounded and continuous functions of (μ, τ), there exists an open ball O s.t. $f \in O$ and $\forall f' \in O$,

$$\int \chi_{t,t'}(\tau) H(\mu) f'(\mathrm{d}\mu, \mathrm{d}\tau) + H\left(\int (1 - \chi_{t,t'}(\tau)) \mu f'(\mathrm{d}\mu, \mathrm{d}\tau)\right) \geq \chi \cdot \int_{s \leq t'} (1 - F'(s))\mathrm{d}s + H(\mu_0) + \frac{1}{2}\varepsilon. \tag{35}$$

Now, consider

$$\begin{aligned}
&\int \mathbf{1}_{\tau \leq t'} H(\mu) f'(\mathrm{d}\mu, \mathrm{d}\tau) + H\left(\int \mathbf{1}_{\tau > t'} \mu f'(\mathrm{d}\mu, \mathrm{d}\tau)\right) \\
=&\int \chi_{t,t'}(\tau) H(\mu) f'(\mathrm{d}\mu, \mathrm{d}\tau) + \int (\mathbf{1}_{\tau \leq t'} - \chi_{t,t'}(\tau)) H(\mu) f'(\mathrm{d}\mu, \mathrm{d}\tau) \\
&+ H\left(\int (1 - \chi_{t,t'}(\tau)) \mu f(\mathrm{d}\mu, \mathrm{d}\tau) - \int (\mathbf{1}_{\tau \leq t'} - \chi_{t,t'}(\tau)) \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) \\
\geq&\int \chi_{t,t'}(\tau) H(\mu) f'(\mathrm{d}\mu, \mathrm{d}\tau) + \int (\mathbf{1}_{\tau \leq t'} - \chi_{t,t'}(\tau)) H(\mu) f'(\mathrm{d}\mu, \mathrm{d}\tau) \\
&+ H\left(\int (1 - \chi_{t,t'}(\tau)) \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) - H\left(\int (\mathbf{1}_{\tau \leq t'} - \chi_{t,t'}(\tau)) \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) \\
\geq&\int \chi_{t,t'}(\tau) H(\mu) f'(\mathrm{d}\mu, \mathrm{d}\tau) + H\left(\int (1 - \chi_{t,t'}(\tau)) \mu f(\mathrm{d}\mu, \mathrm{d}\tau)\right) \\
\geq&\chi \cdot \int_{s \leq t'} (1 - F'(s))\mathrm{d}s + H(\mu_0) + \frac{1}{2}\varepsilon.
\end{aligned}$$

The first inequality is from H being convex and HD1. The second inequality is from the convexity of H. The last inequality is Equation (35). Therefore, $f' \in \mathbb{F}^C$; hence, $\mathbb{F}$ is a closed set. *Q.E.D.*

C.2 Proof of Proposition 1

We prove by showing that $\forall f \in \mathbb{F}$, there exists $f' \in \mathbb{F}$ s.t. $|\mathrm{supp}(f'(\cdot, t))| \leq n + 2$ and $\int U(\mu, \tau) f'(\mathrm{d}\mu, \mathrm{d}\tau) \geq \int U(\mu, \tau) f(\mathrm{d}\mu, \mathrm{d}\tau)$. $\forall t \in T$, consider the following optimization problem:

$$\sup_{f^t \in \Delta(S)} \int U(\mu, t) f^t(\mathrm{d}\mu) \tag{36}$$

$$\text{s.t.} \begin{cases} \mathbb{E}_{f^t}[\mu] = \dfrac{\int \mu f(\mathrm{d}\mu, t)}{f(S,t)}; \\ \mathbb{E}_{f^t}[H(\mu)] \leq \dfrac{\int H(\mu) f(\mathrm{d}\mu, t)}{f(S,t)}. \end{cases}$$

For a feasible f^t, modify f by replacing $f(\cdot,t)$ with $f^t \cdot f(S,t)$ and denote it by f'. This modification does not change any term in Equation (2) except that $\int_{\tau\leq t} H(\mu) f'(\mathrm{d}\mu, \mathrm{d}\tau)$ gets weakly lower. Hence, the modified probability measure f' is still in $\mathbb{F}$. Second, since $f(\cdot,t)/f(S,t)$ is a feasible probability measure of Equation (36), if f^t is the maximizer of Equation (36), then $\mathbb{E}_{f'}[U] \geq \mathbb{E}_f[U]$.

A direct application of corollary 3.1 of Doval and Skreta 2022 implies that there exists f^t solving Equation (36) with $|\mathrm{supp}(f^t)| \leq n+2$.[1] Therefore, by replacing each $f(\cdot,t)$ with the corresponding f^t, we obtain f' with $|\mathrm{supp} f(\cdot,t)| \leq n+2$ and $\mathbb{E}_{f'}[U] \geq \mathbb{E}_f[U]$.

Let f be the solution of Equation (P) (existence implied by Lemma 1), then the corresponding f' satisfies the statement of Proposition 1 *Q.E.D.*

C.3 Proof of Proposition 2

Let $\bar{t} = \sup \mathrm{Supp}(f)$. Suppose for the purpose of contradiction that $\Lambda(t)$ is not strictly decreasing for $t < \bar{t}$, then there exists an interval (t_1, t_2) s.t. $\Lambda(t) = \Lambda$ on the interval. $\forall t \in [0, \bar{t}]$, define $\xi(t) = \max_\mu (l_{f,\lambda}(\mu,t) - a\cdot\mu)$, the gap in the (9). Therefore, Theorem 2 implies that $\xi(t) \leq 0$ and $\xi(t_1) = \xi(t_2) = 0$. $\forall t \in (t_1, t_2)$, $\mu \in S$,

$$\frac{\partial^2}{\partial t^2}(l_{f,\lambda}(\mu,t) - a\cdot\mu) = U_t''(\mu,t) > 0.$$

Therefore, since $\xi(t)$ is the upper envelope of a collection of strictly convex functions, $\xi(t)$ is strictly convex on $[t_1, t_2]$; hence, $\xi(t) < 0$ on (t_1, t_2) and $\xi(t_2^-)' > 0$. $\xi(t) < 0$ on (t_1, t_2) implies that $f(\mu, (t_1, t_2)) = 0$; hence, (2) is strictly slack at t_2^-. If $t_2 < \bar{t}$, then (2) must be binding at t_2. If $t_2 = \bar{t}$, then the DM can strictly improve f by reduce t_2 if (2) is not binding at t_2. Therefore, since (2) is binding at t_2, f puts a point mass at t_2 that is different from $\widehat{\mu}_{t_2}$. Let $\mu^*_{t_2}$ be such a maximizer that attains $\xi(t_2)$. Consider

$$\begin{aligned} \xi(t_2 + \delta t) \geq & l_{f,\lambda}(\mu^*_{t_2}, t_2 + \delta t) - a\cdot \mu^*_{t_2} \\ =& \xi(t_2) + U(\mu^*_{t_2}, t_2 + \delta t) - U(\mu^*_{t_2}, t_2) + \chi \int_{\tau\in[t_2, t_2+\delta t)} \Lambda(\tau)\mathrm{d}\tau \\ &+ \int_{\tau\in[t_2,t_2+\delta t)} (H(\mu^*_{t_2}) - \nabla H(\widehat{\mu}_\tau)\cdot \mu^*_{t_2}) \mathrm{d}\lambda(\tau) \end{aligned}$$

[1] Doval and Skreta 2022 improved the bound derived by Zhong 2018 from $2(n+1)$ to a tight bound of n+2.

$$=\xi(t_2)+U_t'(\mu^*_{t_2},t_2)\delta t+\chi\Lambda(t_2^+)\delta t+(\Lambda-\Lambda(t_2^+))(H(\mu^*_{t_2})-\nabla H(\widehat{\mu}_{t_2})\cdot\mu^*_{t_2})+o(\delta t).$$

If $\Lambda>\Lambda(t_2^+)$, then since $\mu^*_{t_2}\neq\widehat{\mu}_{t_2}$ and H is strictly convex, $H(\mu^*_{t_2})-\nabla H(\widehat{\mu}_{t_2})>0$; hence, $\xi(t+\delta t)>0$ for sufficiently small $\delta t>0$. If $\Lambda=\Lambda(t_2^+)$, the last line above is weakly higher than

$$\begin{aligned}&\xi(t_2)+U_t'(\mu^*_{t_2},t_2)\delta t+\chi\Lambda\delta t+o(\delta t)\\ =&\xi(t_2^-)'\delta t+o(\delta t).\end{aligned}$$

Therefore, $\xi(t_2+\delta t)>0$ for sufficiently small $\delta t>0$. Both case contradict $\xi\leq 0$.

Since $\Lambda(t)$ is strictly decreasing, the complementary slackness condition implies that Equation (2) is binding all the time:

$$\begin{aligned}&\int_{\tau\leq t}H(\mu)f(\mathrm{d}\mu,\mathrm{d}\tau)+H\left(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}\tau)\right)-H(\mu_0)\equiv\chi\cdot\int\min\{t,\tau\}f(\mathrm{d}\mu,\mathrm{d}\tau)\\ \Longleftrightarrow\ &E[H(\widehat{\mu}_t)]-H(\mu_0)\equiv\chi E[\min(t,\tau)].\end{aligned}$$

Combine the equality with

$$\begin{aligned}&\mathbb{E}[H(\widehat{\mu}_t)]-H(\mu_0)\leq\mathbb{E}[H(\mu_{\min t,\tau})-H(\mu_0)]\leq\chi E[\min(t,\tau)]\\ \Longrightarrow\ &H(\widehat{\mu}_t)\equiv\mathbb{E}[H(\mu_t)|\tau>t].\end{aligned}$$

Since H is strictly convex, $\mu_t|\tau>t$ has to be degenerate and equal to $\widehat{\mu}_t$. *Q.E.D.*

C.4 Proof of Proposition 3

$\forall$ embeddable f, Equation (1) at $t\to\sup T$ implies that

$$\begin{aligned}&\mathbb{E}_f[H(\mu)-H(\mu_0)]\leq\chi\mathbb{E}_f[\tau]\\ \Longrightarrow\ &\mathbb{E}_f[v(\mu)]-\kappa\mathbb{E}_f[\tau]\leq\mathbb{E}_f[v(\mu)-\kappa/\chi(H(\mu)-H(\mu_0))]\leq(15).\end{aligned}$$

This proves sufficiency.

To prove necessity, it is sufficient to prove that Equation (15) is attainable. $\forall\pi$ that is feasible in Equation (15), define

$$f(\mu,\tau)=\pi(\mu)\cdot\delta_{\tau=\frac{\mathbb{E}_\pi[H(\mu)-H(\mu_0)]}{\chi}}.$$

It is straightforward that f satisfies Equation (1) and attains expected utility $\mathbb{E}_f[v(\mu)-\kappa\tau]=\mathbb{E}_\pi[v(\mu)-\kappa/\chi\,(H(\mu)-H(\mu_0))]$. *Q.E.D.*

C.5 Proof of Proposition 5

$\forall t \in [0,\bar{t}]$, define $\xi(t) = \max_\mu(l_{f,\lambda}(\mu,t) - a\cdot\mu)$, the gap in the (9). Therefore, Theorem 2 implies that $\xi(t) \leq 0$ and $\xi(\underline{t}) = \xi(\bar{t}) = 0$. The envelope theorem implies $\xi'(\underline{t}^-) = U'(\mu,\underline{t}^-) + \chi\Lambda(\underline{t}) \geq 0$, where μ is the maximizer that attains $\xi(\underline{t})$. Suppose for the purpose of contradiction that $\bar{t} > \overline{J}^{-1}\circ \underline{J}(\underline{t})$. Then, $\exists \epsilon > 0$ s.t. $\forall \mu \in S$, $\forall t > \bar{t} - \epsilon$,

$$U_t'(\mu,t) + \chi\Lambda(t) < -\epsilon.$$

Regularity implies that $\frac{\widehat{\mu_t}}{f(\tau>t)} \to \mu^*$ and $\frac{\Lambda(\bar{t})-\Lambda(t)}{\bar{t}-t}$ is bounded when $t \to \bar{t}^-$. Therefore, there exists $\delta \in (0,\epsilon)$ s.t. $\forall t \geq \bar{t} - \delta$,

$$\frac{\Lambda(\bar{t}-\delta)-\Lambda(\bar{t})}{\delta}(H(\mu^*) - \nabla H(\widehat{\mu_t})\cdot\mu^*) < \frac{1}{2}\epsilon.$$

Then, $\forall t_1, t_2 \in (\bar{t}-\delta, \bar{t})$ and $t_1 < t_2$,

$$\begin{aligned}
&l_{f,\lambda}(\mu^*,t_2) - a\cdot\mu^* \\
=&(l_{f,\lambda}(\mu^*,t_1) - a\cdot\mu^*) + \int_{t_1}^{t_2} U_t'(\mu^*,t) + \chi\Lambda(t)\mathrm{d}t + \int_{t_1}^{t_2}(H(\mu^*) - \nabla H(\widehat{\mu_t})\cdot\mu^*)\mathrm{d}\lambda(t) \\
<&-\epsilon(t_2-t_1) + \frac{\epsilon\delta}{2(\Lambda(\bar{t}-\delta)-\Lambda(\bar{t}))}\int_{t_1}^{t_2}\mathrm{d}\lambda(t) \\
&\to -\frac{\epsilon\delta}{2} \quad \text{when } t_2 \to \bar{t},\ t_1 = \bar{t}-\delta.
\end{aligned}$$

The analysis above implies that $l_{f,\lambda}(\mu^*,t)$ is bounded away from 0 when $t \to \bar{t}^-$, contradicting $(\mu^*,\bar{t}) \in \mathrm{Supp}(f)$. *Q.E.D.*

C.6 Proof of Lemma 3

$\forall$ admissible strategy $(\langle\mu_t\rangle, \langle\chi_t\rangle, \tau)$, let $\widehat{\chi_t} := \mathbb{E}[\chi_t | t < \tau]$. Then,

$$\begin{aligned}
\mathbb{E}\left[H(\mu_{\min\{\tau,t\}}) - H(\mu_0)\right] \leq &\mathbb{E}\left[\int_0^{\min\{t,\tau\}} \chi_s \mathrm{d}s\right] \\
=&\mathbb{E}\left[\int_0^{\min\{t,\tau\}} \mathbb{E}[\chi_s | s<\tau]\mathrm{d}s\right] \\
=&\mathbb{E}\left[\int_0^{\min\{t,\tau\}} \widehat{\chi_s}\mathrm{d}s\right].
\end{aligned}$$

Let $f \sim (\mu_\tau, \tau)$, the inequality is equivalent to

$$\int_{\tau\leq t} H(\mu) f(\mathrm{d}\mu,\mathrm{d}\tau) + H\left(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}\tau)\right) - H(\mu_0) \leq \int_{s\leq t}\widehat{\chi_s}(1-F(s))\mathrm{d}s.$$

Meanwhile,

$$\begin{aligned}\mathbb{E}\left[U(\mu_\tau,\tau)-\int_{t\le\tau}c_t(\chi_t)\mathrm{d}t\right]&=\mathbb{E}[U(\mu_\tau,\tau)]-\mathbb{E}\left[\int_{t\le\tau}\mathbb{E}[c_t(\chi_t)|t<\tau]\mathrm{d}t\right]\\&\le\mathbb{E}[U(\mu_\tau,\tau)]-\mathbb{E}\left[\int_{t\le\tau}c_t(\mathbb{E}[\chi_t|t<\tau])\mathrm{d}t\right]\\&=\int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)-\int c_t(\widehat{\chi}_t)(1-F(t))\mathrm{d}t.\end{aligned}$$

Therefore, (P1)≥(C1). As is established by Corollary 1.1, for any feasible (f,χ) in (P1), there exists an admissible strategy of (C1) implementing f and achiving the same payoff. Therefore, (P1)=(C1). *Q.E.D.*

C.7 Proof of Lemma 2-A

Defined mapping G:

$$G(f,\eta)(t)=\frac{1}{t}\left(\int_{s\le t}\eta_s\mathrm{d}s-H(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}\tau))-\int_{\tau\le t}H(\mu)f(\mathrm{d}\mu,\mathrm{d}\tau)+H(\mu_0)\right).$$

Rewrite the Lagrangian by replacing χ_t with $\frac{\eta_t}{1-F(t)}$:

$$\widetilde{\mathcal{L}}(f,\eta,\lambda):=\int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)-\int c_t\left(\frac{\eta_t}{1-F(t)}\right)(1-F(t))\,\mathrm{d}t+\int t\cdot G(f,\eta)(t)\mathrm{d}\lambda(t),$$

with the convention that $\frac{0}{0}=0$. Consider the space

$$\Omega=\left\{(f,\eta)\in\Delta_{\mu_0}\times L^\infty(T)\middle|\frac{\eta(t)}{1-F(t)}\in L^\infty(T)\right\}.$$

Ω is a subset of the vector space of $\Delta(S\times T)\times L^\infty(T)$, endowed with the product topology. Let Ω^C be the time-continuous subspace of Ω s.t. $G(f,\eta)(t)\in UC(T^\circ)$. We verify that Ω^C is convex: $\forall(f_1,\eta_1),(f_2,\eta_2)\in\Omega^C$, $\forall\alpha\in(0,1)$,

$$\left\|\frac{\alpha\eta_1+(1-\alpha)\eta_2}{1-\alpha F_1-(1-\alpha)F_2}\right\|\le\max\left\{\left\|\frac{\eta_1}{1-F_1}\right\|,\left\|\frac{\eta_2}{1-F_2}\right\|\right\}.$$

Therefore, $\alpha(f_1,\eta_1)+(1-\alpha)(f_2,\eta_2)\in\Omega^C$; hence, Ω^C is convex. By defintion, G is a concave mapping of Ω^C into $UC(T^\circ)$.

Next, we verify that there exists $(f,\eta)\in\Omega^C$ s.t. $G(f,\eta)(\cdot)$ is an interior point of the positive cone. Let $f=\delta_{\mu=\mu_0}\cdot U(T)$, where $U(T)$ is the uniform randomization and $\eta_t=\eta\cdot(1-F(t))$ for $\eta>0$. Therefore, $(f,\eta)\in\Omega^C$. $\forall t\in T^\circ$,

$$G(f,\eta)(t)=\frac{\eta}{t}\int_{s\le t}(1-F(s))\mathrm{d}s>0;$$

hence, $G(f,\eta)$ is an interior point.

Next, we verify that the objective

$$\int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)-\int c_t\left(\frac{\eta_t}{1-F(t)}\right)(1-F(t))\mathrm{d}t$$

is bounded , concave and continuous in (f,η). Since $\frac{\eta}{1-F}$ is bounded, the objective function is obviously bounded. Note that the function $c_t(x/y)/y$ has positive semi-definite Hessian matrix; hence, $\int c_t\left(\frac{\eta_t}{1-F(t)}\right)(1-F(t))$ is a convex functional. It is obvious that $\int U\mathrm{d}f$ is continuous in f. Since $\frac{\eta_t}{1-F(t)}$ is (uniformly) bounded, it is continuous in (f,η).

The cited theorem implies

$$\begin{aligned}
&\sup_{(f,\eta)\in\Omega^C, G(f,\eta)\geq 0}\int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)-\int c_t\left(\frac{\eta_t}{1-F(t)}\right)(1-F(t))\mathrm{d}t=\min_{\lambda\in\mathbb{L}}\sup_{(f,\eta)\in\Omega^C}\widetilde{\mathcal{L}}(f,\eta,\lambda)\\
\iff &\sup_{f\in\Delta^C_{\mu_0},\chi\in L^\infty, G(f,\chi/(1-F))\geq 0}\int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)-\int c_t(\chi_t)(1-F(t))\mathrm{d}t=\min_{\lambda\in\mathbb{L},f\in\Delta^C_{\mu_0},\chi\in L^\infty}\mathcal{L}(f,\chi,\lambda),
\end{aligned}$$

where there exists $\lambda^*\in\mathbb{L}$ achieving the minimum on the RHS. Note that the argument of Lemma 7 applies here as well; hence, it is wlog to replace $\Delta^C_{\mu_0}$ with Δ_{μ_0}:

$$\sup_{f\in\Delta_{\mu_0},\chi\in L^\infty, G(f,\chi/(1-F))\geq 0}\int U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)-\int c_t(\chi_t)(1-F(t))\mathrm{d}t=\min_{\lambda\in\mathbb{L},f\in\Delta_{\mu_0},\chi\in L^\infty}\mathcal{L}(f,\chi,\lambda)$$

The LHS is exactly Equation (P1). *Q.E.D.*

C.8 Proof of Theorem 2-A

Sufficiency: Suppose for the purpose of contradiction that (f,χ,a,λ) solves Equations (17) and (18) but (f,χ) does not solve Equation (P1). In other words, there exists admissible (g,ϕ) s.t. $\mathcal{L}(g,\phi,\lambda)>\mathcal{L}(f,\chi,\lambda)$. Since $\widetilde{\mathcal{L}}$ is concave, $\forall\alpha\in(0,1)$,

$$\begin{aligned}
&\lim_{\alpha\to 0}\frac{\mathcal{L}(\alpha g+(1-\alpha)f,\frac{\alpha\phi(1-G)+(1-\alpha)\chi(1-F)}{1-\alpha G-(1-\alpha)F},\lambda)-\mathcal{L}(f,\chi,\lambda)}{\alpha}>0\\
\implies &\int U(\mu,\tau)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)+\int_{t\in T^\circ}\left(\int_{\tau\leq t}(\phi_\tau(1-G(\tau))-\chi_\tau(1-F(\tau)))-H(\mu)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)\right)\mathrm{d}\lambda(t)\\
&-\overline{\lim_{\alpha\to 0}}\int_{t\in T}\frac{c_t\left(\frac{\alpha\phi_t(1-G(t))+(1-\alpha)\chi_t(1-F(t))}{1-\alpha G(t)-(1-\alpha)F(t)}\right)(1-\alpha G(t)-(1-\alpha)F(t))-c_t(\chi_t)(1-F_t)}{\alpha}\mathrm{d}t\\
&-\overline{\lim_{\alpha\to 0}}\int_{t\in T^\circ}\frac{H\left(\int_{\tau>t}\mu(\alpha g+(1-\alpha)f)(\mathrm{d}\mu,\mathrm{d}\tau)\right)-H\left(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}t)\right)}{\alpha}\mathrm{d}\lambda(t)>0\\
\implies &\int U(\mu,\tau)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)+\int_{t\in T^\circ}\left(\int_{\tau\leq t}(\phi_\tau(1-G(\tau))-\chi_\tau(1-F(\tau)))-H(\mu)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)\right)\mathrm{d}\lambda(t)
\end{aligned}$$

$$
\begin{aligned}
&-\int_{t\in T} c_t(\chi_t)(F(t)-G(t))+(1-G(t))(\phi_t-\chi_t)c_t'(\chi_t)\mathrm{d}t \\
&-\int_{t\in T^\circ}\nabla H(\widehat{\mu_t})\int_{\tau>t}\mu(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\lambda(t)>0 \\
\Longrightarrow &\int U(\mu,\tau)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)+\int_{t\in T^\circ}\left(\int_{\tau\le t}(\phi_\tau(1-G(\tau))-\chi_\tau(1-F(\tau)))-H(\mu)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)\right)\mathrm{d}\lambda(t) \\
&-\int_{\tau\in T}\int_{t\le\tau} c_t(\chi_t)\mathrm{d}t(f-g)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\tau-\int_{t\in T}(1-G(t))(\phi_t-\chi_t)\Lambda(t)\mathrm{d}t \\
&-\int_{t\in T^\circ}\nabla H(\widehat{\mu_t})\int_{\tau>t}\mu(g-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\lambda(t)>0 \\
\Longleftrightarrow &\int l_{f,\chi,\lambda}(\mu,\tau)(f-g)(\mathrm{d}\mu,\mathrm{d}\tau)>0 \\
\Longleftrightarrow &\int l_{f,\chi,\lambda}(\mu,\tau)g(\mathrm{d}\mu,\mathrm{d}\tau)>a\cdot\mu_0.
\end{aligned}
$$

Contradiction.

Necessity: We begin with Equation (18). Suppose $\int|c_t'(\chi_t)-\Lambda(t)|(1-F(t))\mathrm{d}t>0$, then consider an alternative path $\chi_t'=\chi_t-\mathrm{sgn}(c_t'(\chi_t)-\Lambda(t))\varepsilon$. Then,

$$
\begin{aligned}
\mathcal{L}(f,\chi',\lambda)-\mathcal{L}(f,\chi,\lambda)&=\int(c_t(\chi_t)-c_t(\chi_t'))(1-F(t))\mathrm{d}t+\int\left(\int_{s\le t}(\chi_s'-\chi_s)(1-F(s))\mathrm{d}s\right)\mathrm{d}\lambda(t). \\
&=\int_{c_t'(\chi_t)>\Lambda(t)}((c_t(\chi_t)-c_t(\chi_t'))+\Lambda(t)(\chi_t'-\chi_t))(1-F(t))\mathrm{d}t \\
&\quad+\int_{c_t'(\chi_t)<\Lambda(t)}((c_t(\chi_t)-c_t(\chi_t'))+\Lambda(t)(\chi_t'-\chi_t))(1-F(t))\mathrm{d}t \\
&=\int_{c_t'(\chi_t)>\Lambda(t)}((c_t(\chi_t)-c_t(\chi_t-\varepsilon))-\varepsilon\Lambda(t))(1-F(t))\mathrm{d}t \\
&\quad+\int_{c_t'(\chi_t)<\Lambda(t)}((c_t(\chi_t)-c_t(\chi_t+\varepsilon))+\varepsilon\Lambda(t))(1-F(t))\mathrm{d}t \\
\Longrightarrow \frac{\mathcal{L}(f,\chi',\lambda)-\mathcal{L}(f,\chi,\lambda)}{\varepsilon}&\to\int|c_t'(\chi_t)-\Lambda(t)|(1-F(t))\mathrm{d}t>0.
\end{aligned}
$$

This leads to a contradiction to (f,χ,λ) being a saddle point.

Next, we pin down a and ξ and verify Equation (17). $\forall\mu\in S$, for $\tau>\bar{t}$, select $\nabla H(0)$ s.t. $\nabla H(0)\cdot\mu=H(\mu)$ Since $l_{f,\chi,\lambda}$ is bounded,

$$
\widehat{l}(\mu)=\sup_{\pi\in\Delta(S),\mathbb{E}_\pi[\mu]=\mu_0}\mathbb{E}_\pi[\sup_{\tau\in T^\circ}l_{f,\chi,\lambda}(\mu,\tau)]
$$

is a well defined real valued concave function on S. Let $a\cdot\mu$ be the supporting hyperplane of $\widehat{l}$ at $\mu_0\mu_0$. Evidently, $l_{f,\chi,\lambda}(\mu,\tau)\le a\cdot\mu$.

Next, we prove that $\int l_{f,\chi,\lambda}(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)=a\cdot\mu_0$. Suppose for the purpose of contradiction that $\int l_{f,\chi,\lambda}(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau)<a\cdot\mu_0$. Then, since $l_{f,\chi,\lambda}(\mu,\tau)\le a\cdot\mu$, there exists an open

set and $\varepsilon > 0$ s.t. $\inf O > 0$ and $\int_O (a \cdot \mu - l_{f,\chi,\lambda}(\mu,\tau)) f(\mathrm{d}\mu,\mathrm{d}\tau) > \varepsilon$. Let $(\mu,\tau) = \mathbb{E}_f[(\mu',\tau')|O]$. Then, $\mathbb{E}_f[l_{f,\chi,\lambda}(\mu',\tau')|O] < a \cdot \mu - \varepsilon$.

Since $a\mu = \widehat{l}(\mu)$, there exists a finite support distribution $\pi \in \Delta(S)$ that attains $\widehat{l}(\mu) - \frac{1}{4}\varepsilon$. For each μ' in the support of π, there exists τ' s.t. $l_{f,\chi,\lambda}(\mu',\tau') > \sup_\tau l_{f,\chi,\lambda}(\mu',\tau) - \frac{1}{4}\varepsilon$. We slightly abuse notation and let π denote the distribution of (μ',τ') pairs. Therefore, $\mathbb{E}_\pi[l_{f,\chi,\lambda}(\mu',\tau')] > a \cdot \mu - \frac{\epsilon}{2}$.

Define $g = f - f(O) \cdot (f|_O - \pi)$. Then,

$$
\begin{aligned}
&\lim_{\alpha \to 0} \frac{\mathcal{L}(\alpha g + (1-\alpha) f, \chi, \lambda) - \mathcal{L}(f,\chi,\lambda)}{\alpha} \\
=& \int U(\mu,\tau)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) + \int_{t \in T^\circ} \left(\int_{s \le t} \chi_s (G(s) - F(s)) \mathrm{d}s - \int_{\tau \le t} H(\mu)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) \right) \mathrm{d}\lambda(t) \\
&- \overline{\lim_{\alpha \to 0}} \int_{t \in T^\circ} \frac{H\left(\int_{\tau > t} \mu(\alpha g + (1-\alpha) f)(\mathrm{d}\mu,\mathrm{d}\tau)\right) - H\left(\int_{\tau > t} \mu f(\mathrm{d}\mu,\mathrm{d}\tau)\right)}{\alpha} \mathrm{d}\lambda(t) \\
\ge& \int U(\mu,\tau)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) + \int_{t \in T^\circ} \left(\int_{s \le t} \chi_s (G(s) - F(s)) \mathrm{d}s - \int_{\tau \le t} H(\mu)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) \right) \mathrm{d}\lambda(t) \\
&- \int_{t < \bar{t}} \nabla H(\widehat{\mu}_t) \cdot \int_{\tau > t} \mu (g-f)(\mathrm{d}\mu,\mathrm{d}\tau) \mathrm{d}\lambda(t) - \int_{t \ge \bar{t}} \int_{\tau > t} H(\mu)(g-f)(\mathrm{d}\mu,\mathrm{d}\tau) \mathrm{d}\lambda(t) \\
=& f(O) \cdot \int l_{f,\chi,\lambda}(\mu,\tau)(\pi - f|_O)(\mathrm{d}\mu,\mathrm{d}\tau) \\
>& f(O) \cdot \left(a\mu - \frac{\epsilon}{2} - (a \cdot \mu - \epsilon) \right) > 0.
\end{aligned}
$$

This leads to a contradiction to (f,χ,λ) being a saddle point. *Q.E.D.*

D Proofs in Section 4

D.1 Proof of Proposition 6

Let $\widetilde{\rho}(t) = \rho(t)^{\frac{\alpha}{\alpha-1}}$, $\widetilde{l}(t) = \frac{\alpha-1}{\alpha} \rho(t) \left(\frac{\rho(t)}{\alpha \widehat{\Lambda}(t)} \right)^{\frac{1}{\alpha-1}} + \chi \int_{s \le t} \widehat{\Lambda}(s) \mathrm{d}s - b$. Since $\xi \in C\mathbb{R}_+$, the region where $\xi^* > 0$ is a countable collection of open intervals. Let (t',t'') be such an open interval. We first note that $\lambda(t) \equiv 0$ on (t',t''). This is because $(t',t'') \subset \mathrm{supp}(f)^C$, which implies Equation (3) being strictly slack. Then, the complementary slackness condition implies $\lambda(t) = 0$.

Suppose the first statement is not true. Then, for sufficiently small ϵ, on $(t''-\epsilon, t'')$, $\widetilde{l}''(t) > 0$. $\widetilde{l}(t'') = 0$ and $\widetilde{l}(t) \le 0$ for $t \ge t''$. Therefore, $\widetilde{l}$ has a strict downward kink at t''. On the other hand, $\widehat{\Lambda}$ is constant on $(t''-\epsilon, t'')$ and is decreasing when $t \ge t''$; hence, $\widehat{\Lambda}(t)^{\frac{1}{1-\alpha}}$ could only have an upward kink at t''. Contradiction.

Suppose $\widetilde{\rho}''$ does not switch sign in (t', t''). Since $\widetilde{l} < 0$ on (t', t'') and $t' > 0$, it has an interior minimizer $t^* \in (t', t'')$, which implies $\widetilde{l}''(t^*) > 0$. This necessarily leads to $\widetilde{l}$ having a strict downward kink at t''. Follow the same argument as before, there is a contradiction. Since $\frac{\mathrm{d}}{\mathrm{d}t^2}\rho(t)^{\frac{\alpha}{\alpha-1}}$ switches sign finitely many times, such intervals much be finite.

Since $\xi^*(t) > 0$ on (t', t''), the complementary slackness condition implies $f(S, (t', t'')) = 0$.

Next, consider the open region where $\xi^* = 0$ and $\widetilde{\rho}'' \neq 0$. We prove that $\lambda(t) > 0$. Since $\xi^*(t) \equiv 0$, Equation (25) implies $\widehat{\Lambda}$ being twice differentiable and

$$\frac{\alpha-1}{\alpha^{\frac{\alpha}{\alpha-1}}}\widetilde{\rho}''(t) + \frac{\alpha+1}{\alpha-1}\chi\widehat{\Lambda}(t)^{\frac{1}{\alpha-1}}\widehat{\Lambda}'(t) + \frac{1}{\alpha-1}\left(\chi\int_{s\le t}\widehat{\Lambda}(s)\mathrm{d}s - b\right)\left(\widehat{\Lambda}^{\frac{2-\alpha}{\alpha-1}}\widehat{\Lambda}''(t) + \frac{2-\alpha}{\alpha-1}\widehat{\Lambda}(t)^{\frac{3-2\alpha}{\alpha-1}}\widehat{\Lambda}'(t)^2\right) = 0$$

Suppose for the purpose of contradiction that $\lambda(t) = 0$. Since $\lambda \ge 0$, λ is locally minimized at t; hence, $\lambda'(t) = 0$. This implies $\widehat{\Lambda}'(t) = 0$ and $\widehat{\Lambda}''(t) = 0$. Therefore, $\widetilde{\rho}''(t) = 0$, leading to contradiction.

Since both the region where $\xi^* > 0$ and the region where $\xi^* = 0 \& \widetilde{\rho}'' \neq 0$ are finitely many open intervals, the remaining region ($\xi^* = 0 \& \widetilde{\rho}'' = 0$) constitutes finitely many closed intervals. *Q.E.D.*

D.2 Proof of Proposition 7

Let f be a solution to the information acquisition problem and λ be the corresponding multiplier. Let $\bar{t} = \sup \mathrm{Supp}(f)$. WLog, we assume that $\forall t < \bar{t}$, $\rho(t) > 0$ (otherwise f can be truncated at $\bar{t}$). Suppose Λ_t is not strictly increasing for $t < \bar{t}$, then the region where Λ_t is flat constitutes a countable collection of open intervals $\cup(l_i, r_i)$. We prove by induction that $\forall n$, there exists an optimal strategy f_n s.t. $\forall i \le n$, $\int_{[l_i, r_i]} G_{f_n}(t) f_n(\mathrm{d}\mu, \mathrm{d}t) = 0$. Easy to see that it is sufficient to prove the statement for $n = 1$.

Wlog, assume that $\int_{(l_1, r_1)} G_f(t) f(\mathrm{d}\mu, \mathrm{d}t) > 0$. Let μ^* solve

$$\mu^* \in \arg\max_{\mu \in S} l_{f,\lambda}(\mu, r_1),$$

i.e. $\mu^* = M(\frac{\rho(r_i)}{\Lambda(r_i)}) > 0.5$. Let

$$p^* = \max\left\{\frac{\chi\int_{t\in(l_i,r_i)}(r_i - l_i) f(\mathrm{d}\mu, \mathrm{d}t)}{H(\mu^*)}, \int_{t\ge r_i} f(\mathrm{d}\mu, \mathrm{d}\tau)\right\}.$$

Now, we claim that f_1 defined as

$$f_1(\mu,t) = \begin{cases} f(\mu,t) & t \le l_1 \\ 0 & t \in (l_1, r_1) \\ \frac{p^*}{2}(\delta_{\mu^*+0.5,r_1} + \delta_{0.5-\mu^*,r_1}) & t = r_1 \\ \left(1 - \frac{p^*}{\int_{t\ge r_1} f(\mathrm{d}\mu,\mathrm{d}\tau)}\right) f(\mu,t) & t > r_1 \end{cases}$$

solves the information acquisition problem and $\int_{(l_1,r_1)} G_{f_1}(t) f_1(\mathrm{d}\mu,\mathrm{d}t) = 0$. The latter is obvious. We first verify that f_1 is a feasible strategy:

$$G_{f_1}(t) \begin{cases} = G_f(t) \ge 0 & t \le r_1 \\ = \int_{\tau\in(l_1,t)} (t-l_1) f(\mathrm{d}\mu,\mathrm{d}\tau) > 0 & t \in (l_1,r_1) \\ = \int_{\tau\in(l_1,t)} (t-l_1) f(\mathrm{d}\mu,\mathrm{d}\tau) - p^* H(\mu^*) \ge 0 & t = r_1 \\ = \left(1 - \frac{p^*}{\int_{t\ge r_i} f(\mathrm{d}\mu,\mathrm{d}\tau)}\right) G_f(t) \ge 0 & t > r_1 (\text{ while well defined}) \end{cases}$$

Note that $\mathrm{Supp}(f_1) \subset \mathrm{Supp}(f) \cup \{(\mu^*, r_1), (1-\mu^*, r_1)\} \subset \arg\max l_{f,\lambda}$. Therefore, f_1 is optimal since it satisfies Equation (25) on $(0,\bar{t})$. By definition $\int_{[l_1,r_1)} G_{f_1}(t) f_1(\mathrm{d}\mu,\mathrm{d}t) = 0$. We only need to verify that $\int G_{f_1}(t) f_1(\mathrm{d}\mu, r_1) = 0$. Note that $G_{f_1}(r_1) > 0$ only if $p^* = \int_{t>r_1} f_1(\mathrm{d}\mu,\mathrm{d}\tau) = 0$. Then, r_1 is the last period. In this case, it is without of optimality to move $f(\cdot, r_1)$ earlier in time until G_{f_1} reaches zero at the mass point.

Now that we have a collection $f_i \subset \mathcal{F}$ s.t. $\forall i \le n$, $\int_{[l_i,r_i]} G_{f_n}(t) f_n(\mathrm{d}\mu,\mathrm{d}t) = 0$. Since $\mathcal{F}$ is compact and U is bounded and continuous, there exists a limit point $f^* \in \mathcal{F}$ and f^* is optimal as it achieves the same expected utility. Note that $f_n \xrightarrow{w} f^*$ implies $G_{f^*} \le \underline{\lim}\, G_{f_n}$ and $\forall i$, $\int_{t\in[l_i,r_i)} f^*(\mathrm{d}\mu,\mathrm{d}t) \le \underline{\lim} \int_{t\in[l_i,r_i)} f_n(\mathrm{d}\mu,\mathrm{d}t) = 0$. Therefore, $\int G_{f^*}(t) f^*(\mathrm{d}\mu,\mathrm{d}t) = 0$. *Q.E.D.*

D.3 Proof of Proposition 8

Let $r(t) = -\frac{\mathrm{d}\log(\rho(t))}{\mathrm{d}t}$. Differentiate the RHS of Equation (26) w.r.t. t:

$$\frac{\mathrm{d}r(t)}{\mathrm{d}t} = \frac{(\alpha-1)\kappa(t)\kappa''(t) - \alpha\chi\kappa(t)^{\frac{1}{1-\alpha}}\kappa'(t) - (\alpha-1)\kappa'(t)^2}{(\alpha-1)^2\kappa(t)^2}.$$

When $\kappa'(t) > 0$ and $\kappa''(t) < 0$, $r'(t) < 0$, which proves the first point of Proposition 8.

When $\kappa'(t) < 0$ and $\kappa''(t) > 0$, $r(t) > 0$ implies $-\kappa'(t) < \chi \cdot \kappa(t)^{\frac{1}{1-\alpha}}$. Therefore,

$$r'(t) > \frac{(\alpha-1)\kappa(t)\kappa''(t) - \alpha\chi\kappa(t)^{\frac{1}{1-\alpha}}\kappa'(t) + (\alpha-1)\chi\kappa(t)^{\frac{1}{1-\alpha}}\kappa'(t)}{(\alpha-1)^2\kappa(t)^2} > 0,$$

which proves the second point of Proposition 8.

When $\kappa'(t) \equiv 0$, $r(t)$ is constant, which implies

$$\kappa(t) = \left(-\frac{\chi\left(e^{C\cdot\alpha r + \alpha r t} - 1\right)}{(\alpha-1)r}\right)^{-\frac{1-\alpha}{\alpha}}.$$

Note that when $t \to \infty$, $-\frac{\chi\left(e^{C\cdot\alpha r+\alpha r t}-1\right)}{(\alpha-1)r} \to -\infty$ for any $C \in \mathbb{R}$. Therefore, the only possible case where $\kappa(t)$ is well-defined is $C = -\infty$ and $\kappa(t) \equiv \left(\frac{\chi}{(\alpha-1)r}\right)^{\frac{\alpha-1}{\alpha}}$. *Q.E.D.*

D.4 Proof of Proposition 9

Step 1. We claim that in the equilibrium, the stopping time could not involve any point mass. For the purpose of contradiction, suppose that player i's stopping time involves a point mass at $t > 0$. Then, for $j \neq i$, the effective discount factor

$$\rho_j(t) = \mathbb{E}_{\min\{\tau^{*-j} \geq t\}}\left[\frac{1}{\#(\tau^{-j} \leq t) + 1}\right] \cdot e^{-rt}$$

jumps down at t. Let f^* be the optimal strategy of player j. Let Borel measure $f_t^{*\varepsilon} = \int_{s\in[t,t+\varepsilon]} f^*(\mu, \mathrm{d}s)$. Define $f^\varepsilon = f^* - f^* \cdot \delta_{\tau\in[t,t+\varepsilon]} + f_t^{*\varepsilon}\delta_{\tau=t-\varepsilon}$ for $\varepsilon \in (0,t)$. We claim that $\exists \varepsilon > 0$ s.t. $f_t^{*\varepsilon}(S) = 0$. If not, let λ_j be the multiplier from the dual problem:

$$\begin{aligned}
&\mathcal{L}_j(f^\varepsilon, \lambda_j) - \mathcal{L}_j(f^*, \lambda_j) \\
&= \int_{s\in[t,t+\varepsilon]} |\mu|(\rho_j(t-\varepsilon) - \rho_j(s)) f^*(\mathrm{d}\mu, \mathrm{d}s) \\
&\quad + \int_{s\in(t-\varepsilon,t+\varepsilon]} \left(-\chi(F^*(t) - F^\varepsilon(s)) - \mathbb{E}_{f^\varepsilon}[H(\widehat{\mu}_s) + H(\mu_s)] - \mathbb{E}_{f^*}[H(\widehat{\mu}_s) + H(\mu_s)]\right) \mathrm{d}\lambda_j(s) \\
&\geq \int_{s\in[t,t+\varepsilon]} |\mu|(\rho_j(t-\varepsilon) - \rho_j(s)) f^*(\mathrm{d}\mu, \mathrm{d}s) \\
&\quad - (\Lambda_j(t-\varepsilon) - \Lambda_j(t+\varepsilon)) f_t^{*\varepsilon}(S)(\chi + \sup H - \inf H) \\
\implies &\lim_{\varepsilon\to 0} \frac{\mathcal{L}_j(f^\varepsilon, \lambda_j) - \mathcal{L}_j(f^*, \lambda_j)}{f_t^{*\varepsilon}(S)} \geq (\rho_j(t-) - \rho_j(t)) - (\Lambda(t-) - \Lambda(t))(\chi + \sup H - \inf H) > 0
\end{aligned}$$

In the last inequality, we use the fact that $t\lambda_j(t)$ is L^1 on $[t-\varepsilon, t]$. This contradicts the fact that f^* maximizes $\mathcal{L}_j(f^*, \lambda_j)$.

Therefore, $\exists \varepsilon > 0$ s.t. $f_{-i}(S, [t, t+\varepsilon]) = 0$. This implies

$$\rho_i(s) \begin{cases} \geq \rho_i(t) e^{-r(s-t)} & s < t \\ = \rho_i(t) e^{-r(s-t)} & s \in [t, t+\varepsilon] \end{cases}$$

Since f_i^* involves a point mass at t, this implies $\lambda_i(s) = 0$ on $(t-\delta, t)$ for some $\delta > 0$. On $(t-\delta, t)$:

$$\begin{aligned} &\frac{\alpha-1}{\alpha} \frac{\rho_i(s)^{\frac{\alpha}{\alpha-1}}}{(\alpha \Lambda_i(t)^{\frac{1}{\alpha-1}})} + \chi \int_{\tau \leq t-\delta} \Lambda_i(\tau) \mathrm{d}\tau + \chi(s - t + \delta) \Lambda_i(t) < b \\ \Longrightarrow\ &\frac{\alpha-1}{\alpha} \frac{\rho_i(t) e^{-\frac{\alpha r}{\alpha-1}(s-t)}}{(\alpha \Lambda_i(t)^{\frac{1}{\alpha-1}})} + \chi \int_{\tau \leq t-\delta} \Lambda_i(\tau) \mathrm{d}\tau + \chi(s - t + \delta) \Lambda_i(t) < b, \end{aligned}$$

with equality holds at b. Note that LHS is strictly convex, hence strictly increasing when $s \to t-$. On the other hand, inequality

$$\frac{\alpha-1}{\alpha} \frac{\rho_i(t) e^{-\frac{\alpha r}{\alpha-1}(s-t)}}{(\alpha \Lambda_i(s)^{\frac{1}{\alpha-1}})} + \chi \int_{\tau \leq s} \Lambda_i(\tau) \mathrm{d}\tau \leq b$$

holds for $s \in [t, t+\varepsilon)$. This means LHS is decreasing when $s \to t+$, requiring $\Lambda_i(s)^{\frac{1}{1-\alpha}}$ to be strictly decreasing when $s \to t+$. However, $\Lambda_i(s)^{\frac{1}{1-\alpha}}$ could only have an upward kink at t. Contradiction.

Therefore, $\forall i$, $\xi_i \equiv 0$. The equilibrium is characterized by Λ_i's.

Step 2. We rule out any equilibrium that involves corner solutions, i.e. $f(\{-M, M\} \times T) > 0$. Suppose it is optimal to stop at M at t, let $z(t) = \Lambda_i(t)/\rho_i(t)$, Equation (9) implies

$$\begin{aligned} &\rho_i(s) M + \chi \int_{\tau \leq s} \Lambda_i(\tau) \mathrm{d}\tau - M^\alpha \Lambda_i(s) \leq b \text{ with equality at } t \\ \Longrightarrow\ &\rho_i'(t) M + \chi \rho_i(t) z(t) - M^\alpha (\rho_i(t) z'(t) + \rho_i'(t) z(t)) = 0 \\ \Longleftrightarrow\ &\rho_i'(t)(M - M^\alpha z(t)) + \chi \rho_i(t) z(t) - M^\alpha \rho_i(t) z'(t) = 0 \\ \Longrightarrow\ &(z(t)(\chi + r M^\alpha) - r M) - M^\alpha z'(t) \geq 0 \end{aligned}$$

Note that whenever $z(t) \leq \frac{1}{\alpha} M^{1-\alpha}$, Assumption 2 together with the inequality above implies $z'(t) < 0$. This means, for any $t' > t$, it is optimal to stop at $\pm M$.

Next, we prove that for any ρ_i, stopping at only $\pm M$ from t is dominated by stopping at $\pm \left(\frac{\chi}{(\alpha-1) r} \right)^{\frac{1}{\alpha}}$. Since ρ_i is arbitrary, we normalize t to 0. Suppose for contradiction that stopping at $\pm M$ is optimal, then

$$M \in \arg\max_{\mu \leq M} \mu \cdot \int_{\tau \geq 0} \rho_i(\tau) \left(e^{-\frac{\chi}{\mu^\alpha} \tau} \frac{\chi}{\mu^\alpha} \right) \mathrm{d}\tau$$

$$\implies \int_{\tau\geq 0} \rho_i(\tau) e^{-\lambda\tau}\left(\frac{\alpha-1}{\alpha} - \lambda\tau\right)\mathrm{d}\tau \leq 0,$$

where $\lambda = \frac{\chi}{M^\alpha}$. Let $\rho_i(t) = e^{-\int \omega_i(s)\mathrm{d}s}$, where $\omega_i > r$, then $\forall s \geq 0$,

$$\frac{\mathrm{d}}{\mathrm{d}s}\int_{\tau\geq 0} \rho_i(\tau) e^{-\lambda\tau}\left(\frac{\alpha-1}{\alpha} - \lambda\tau\right)\mathrm{d}\tau = -\int_{\tau\geq s} \rho_i(\tau)\left(\frac{\alpha-1}{\alpha} - \lambda\tau\right)\mathrm{d}\tau \geq 0,$$

where inequality is strict when $s > 0$. This implies that

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{1}{\lambda^{\frac{1}{\alpha}}}\int_{\tau\geq 0} e^{-(r+\lambda)\tau}\lambda \mathrm{d}\tau = \frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{\lambda^{1-\frac{1}{\alpha}}}{r+\lambda} < 0$$
$$\implies (\alpha-1)r - \lambda = (\alpha-1)r - \frac{\chi}{M^\alpha} < 0$$

However, the last inequality violates Assumption 2. Therefore, we focus on only interior solutions.

Step 3. We derive an ODE system characterizing the equilibrium. The sufficient and necessary FOCs for an interior equilibrium define $(\mu_i^*, \rho_i, \Lambda_i)$ solving:

$$\begin{cases} -\dfrac{\mathrm{d}\log(\rho_i(t))}{\mathrm{d}t} = r + \displaystyle\sum_{j\neq i}\frac{\chi}{\mu_j^*(t)^\alpha} \\ \rho_i(t) = \alpha\cdot\mu_i^*(t)^{\alpha-1}\Lambda_i(t) \\ \dfrac{\alpha-1}{\alpha}\rho_i(t)\left(\dfrac{\rho_i(t)}{\alpha\Lambda_i(t)}\right)^{\frac{1}{\alpha-1}} + \chi\displaystyle\int_{s\leq t}\Lambda_i(s)\mathrm{d}s = b_i \end{cases} \tag{37}$$

Define $\omega_i(t) := -\frac{\mathrm{d}\log(\rho_i(t))}{\mathrm{d}t}$, then Equation (37) is equivalent (with additional initial conditions $\rho_i(0) = 1$) to an ODE system for ω_i's:

$$\omega_i'(t) = \sum_{j\neq i}\left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\omega_j(t)-r)\right)\left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\alpha\omega_j(t)-r)\right), \tag{38}$$

where $\overline{\omega} = \frac{\sum\omega_i}{n}$ and $\frac{n}{n-1}(\overline{\omega}(t)-r) - (\omega_j(t)-r) > 0$.

Figure 11 illustrates the phase diagram of Equation (38) for the $n = 2$ case. In the phase diagram, there is a unique interior steady state (the red point). There are more stable points on the boundary (the black points). We would like to argue that a path of $\boldsymbol{\omega}$ constitutes an equilibrium if and only if it starts from the red line.

Step 4. We verify the proposed strategies (conditional on $\zeta > 0$) constitute all symmetric regular equilibria of the game with $\omega_i(0) \geq \frac{r}{\zeta}$. The proposed strategies defines

$$\begin{aligned}\omega_i(t) =& r + (n-1)\lambda^*(t) \\ =& \frac{r}{1 - \frac{1-\zeta}{1-\zeta e^{(\alpha-1)r(t-\bar{t})}}}.\end{aligned}$$

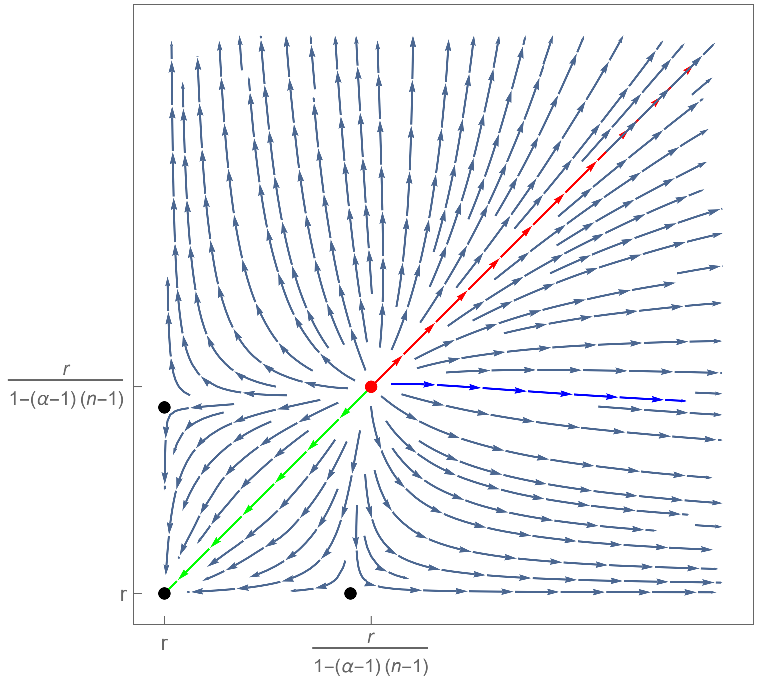


Figure 11: Phase diagram of Equation (38).

It is easy to verify that such ω_i's correspond to all symmetric solutions of Equation (38):

$$\begin{aligned}\omega_i'(t) =&(n-1)\left(\frac{n}{n-1}(\omega_i(t)-r)-(\omega_i(t)-r)\right)\left(\frac{n}{n-1}(\omega_i(t)-r)-(\alpha\omega_i(t)-r)\right)\\ =&(\omega_i-r)\left(\frac{\omega_i-r}{n-1}-(\alpha-1)\omega_i\right)\end{aligned}$$

with initial value no less than $\frac{r}{\zeta}$.

Step 5. We rule out any asymmetric equilibrium where $\overline{\omega}(t)$ is <u>ever</u> weakly higher than $\frac{r}{\zeta}$. This corresponds to the blue curve in Figure 11. Equation (38) implies

$$\begin{aligned}\overline{\omega}'(t) =&\frac{n-1}{n}\sum_i\left(\frac{n}{n-1}(\overline{\omega}(t)-r)-(\omega_i(t)-r)\right)\left(\frac{n}{n-1}(\overline{\omega}(t)-r)-(\alpha\omega_i(t)-r)\right)\\ =&\frac{n-1}{n}\left(n\frac{n^2}{(n-1)^2}(\overline{\omega}-r)^2-\frac{n}{n-1}(\overline{\omega}(t)-r)n((\alpha+1)\overline{\omega}(t)-2r)+\sum_i(\omega_i(t)-r)(\alpha\omega_i(t)-r)\right).\\ \geq&\left(\frac{n^2}{n-1}(\overline{\omega}(t)-r)^2-n(\overline{\omega}(t)-r)((\alpha+1)\overline{\omega}(t)-2r)+(n-1)(\overline{\omega}(t)-r)(\alpha\overline{\omega}(t)-r)\right).\\ =&(\overline{\omega}(t)-r)\left(\frac{\overline{\omega}-r}{n-1}-(\alpha-1)\overline{\omega}\right).\end{aligned}$$

The inequality is the Jensen's inequality (strict if the equilibrium is asymmetric). This implies that $\overline{\omega}$ is always higher than $\widetilde{\omega}$, the solution of

$$\widetilde{\omega}'(t)=(\widetilde{\omega}(t)-r)\left(\frac{\widetilde{\omega}-r}{n-1}-(\alpha-1)\widetilde{\omega}\right)$$

when $\widetilde{\omega}$ and $\overline{\omega}$ have the same initial condition. $\widetilde{\omega}(t)$ has an explicit general solution:

$$\widetilde{\omega}(t) = \frac{r}{1 - \frac{(n-1)(\alpha-1)}{1+Ce^{(\alpha-1)rt}}}.$$

Suppose $\overline{\omega}(t) \geq \frac{r}{1-(n-1)(\alpha-1)}$ and ω_i's are asymmetric, then $\overline{\omega}'(t) > 0$. Therefore, there exists $t' > t$ s.t. $\overline{\omega}(t') > \frac{r}{1-(n-1)(\alpha-1)}$. Then, $\overline{\omega}(t)$ converges to ∞ in finite time. This implies that at least one $\mu_i^*(t)$ converges to 0 in finite time. As a result, at least $n-1$ $\omega_i(t)$'s diverges to ∞ in finite time (at the same time $\bar{t}$ where the first $\mu_i^*(t)$ converges to 0). For notational simplicity, denote these $n-1$ indices $2, \ldots, n$.

Next, we argue that $\omega_2 = \cdots = \omega_n$. Suppose for the purpose of contradiction that $\omega_2(t) < \omega_3(t)$ for some t, where ω_3 is the largest among all ω_i. Wlog, we pick t that for all $t' > t$, $\omega_2'(t'), \omega_3'(t') > 0$. Then,

$$\begin{aligned}
\omega_3'(t) - \omega_2'(t) =& \left(\frac{n}{n-1}(\overline{\omega}(t) - r) - (\omega_2(t) - r)\right)\left(\frac{n}{n-1}(\overline{\omega}(t) - r) - (\alpha\omega_2(t) - r)\right) \\
&- \left(\frac{n}{n-1}(\overline{\omega}(t) - r) - (\omega_3(t) - r)\right)\left(\frac{n}{n-1}(\overline{\omega}(t) - r) - (\alpha\omega_3(t) - r)\right) \\
\geq& \left(\frac{2n}{n-1}(\overline{\omega}(t) - r) - ((\alpha+1)\omega_3(t) - 2r)\right) \cdot (\omega_3(t) - \omega_2(t)) \\
\implies \frac{\mathrm{d}\log(\omega_3(t) - \omega_2(t))}{\mathrm{d}t} \geq& \left(\frac{2n}{n-1}(\overline{\omega}(t) - r) - ((\alpha+1)\omega_3(t) - 2r)\right) \\
\geq& \frac{n}{n-1}(\overline{\omega}(t) - r) - \frac{1}{\alpha}\left(\frac{n}{n-1}(\overline{\omega} - r) + r\right) + r \\
>& \frac{n(\alpha-1)}{\alpha(n-1)}(\overline{\omega}(t) - r) \\
\implies (\omega_3 - \omega_2)(t+s) \geq& (\omega_3 - \omega_2)(t) \cdot e^{\int_t^s \frac{n(\alpha-1)}{\alpha(n-1)}(\overline{\omega}(y) - r)\mathrm{d}y}.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
\omega_3'(t) \leq& \frac{n^2}{n-1}(\overline{\omega}(t) - r)^2 \\
\implies \omega_3(t+s) \leq& \omega_3(t) + \int_t^{t+s} \frac{n^2}{n-1}(\overline{\omega}(y) - r)^2 \mathrm{d}y
\end{aligned}$$

Note that $\omega_3 - \omega_2$ is growing in exponential rate while ω_3 is growing in polynomial rate when $\overline{\omega} \to \infty$. Therefore, $\omega_2 \to -\infty$, which contradicts $\omega_2(t) \to +\infty$. As a result, $\omega_2 = \omega_3$, i.e. $\omega_2 = \ldots \omega_n$. Note that if $\omega_1 \to \infty$, then ω_1 is also identical to all other ω_i's. So we focus on the case $\omega_1 < K < \infty$.

Next, we argue that $\omega_1(t) \to r$. Suppose not, i.e. $\omega_1 - r \geq \varepsilon > 0$, Equation (38) reduces to

$$\omega_1'(t) = (\omega_1(t) - r)\left(\frac{\omega_1(t) - r}{n-1} - (\alpha - 1)\omega_2(t)\right)$$

$$\begin{aligned}\implies \omega_1'(t) \leq & \frac{K^2}{n-1} - \varepsilon(\alpha-1)\omega_2(t)\\ \leq & \frac{K^2}{n-1} - \varepsilon(\alpha-1)\widetilde{\omega}(t)\end{aligned}$$

It can be easily verified that the RHS integrates to $-\infty$ when $t \to \bar{t}$. However, $\omega_1(t) \to \infty$ implies $\mu_i^*(t) \to \infty$ for $i \geq 2$. This implies contestant $2, \ldots, n$ stops with probability strictly less than one at $\bar{t}$, which is clearly suboptimal.

However, $\omega_1(t) \to r$ implies that the strategies is not interior; hence, we rule out this possibility.

Step 6. We rule out any equilibrium where $\overline{\omega}(t)$ is always strictly lower than $\frac{r}{\zeta}$. This corresponds to the green curve in Figure 11. Note that whenever $\omega_i = \omega_j$, $\omega_i' = \omega_j'$. Therefore, the order of ω_i's does not change. Let ω_i be the largest (with possible ties). Choose an arbitrary t, Equation (37) implies:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\omega_i(t) - \frac{n}{n-1}\overline{\omega}(t)\right) = \left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\omega_i(t)-r)\right)\left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\alpha\omega_i(t)-r)\right),$$

where

$$\begin{aligned}\left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\omega_i(t)-r)\right) &\geq \left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\overline{\omega}(t)-r)\right) > 0;\\ \left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\alpha\omega_i(t)-r)\right) &\leq \left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\alpha\overline{\omega}(t)-r)\right) < 0\end{aligned}$$

Therefore:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\omega_i(t) - \frac{n}{n-1}\overline{\omega}(t)\right) \leq \left(\frac{n}{n-1}(\overline{\omega}(t)-r) - ((\overline{\omega}(t)-r)\right)\left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\alpha\overline{\omega}(t)-r)\right) < 0.$$

$\frac{n}{n-1}(\overline{\omega}(t)-r)-(\omega_i(t)-r)$ is strictly increasing with an upper bound. Therefore, it converges. Suppose $\lim_{t\to\infty}\frac{n}{n-1}(\overline{\omega}(t)-r)-(\omega_i(t)-r) = \eta > 0$, then

$$\begin{aligned}&\overline{\lim_{t\to\infty}}\frac{\mathrm{d}}{\mathrm{d}t}\left(\omega_i(t) - \frac{n}{n-1}\overline{\omega}(t)\right)\\ \leq &\overline{\lim_{t\to\infty}}\,\eta\left(\frac{n}{n-1}(\overline{\omega}(t)-r) - (\alpha\overline{\omega}(t)-r) + \alpha(\overline{\omega}(t)-\omega_i(t))\right).\end{aligned}$$

Suppose $\overline{\omega}$ is bounded away from $\frac{r}{1-(n-1)(\alpha-1)}$ or ω_i is bounded away from $\overline{\omega}$, the RHS is bounded away from 0. But this contradicts the existence of limit. Therefore, there exists a sequence of t_ℓ s.t. $\lim\overline{\omega} = \lim\omega_i = \frac{r}{1-(n-1)(\alpha-1)}$; then, $\eta = \frac{(\alpha-1)r}{1-(n-1)(\alpha-1)}$. Since ω_i is the largest, this implies $\lim\omega_j = \frac{r}{1-(n-1)(\alpha-1)}$. However, in this limit:

$$\frac{\mathrm{d}\omega(t)}{\mathrm{d}t} = \eta\alpha\left(\omega(t) - \mathbf{1}\cdot\frac{r}{1-(n-1)(\alpha-1)}\right) + O\left(\left\|\omega(t) - \mathbf{1}\cdot\frac{r}{1-(n-1)(\alpha-1)}\right\|^2\right)$$

i.e. ω diverges from the limit. Hence, this is not possible that $\eta > 0$.

Next, we rule out the possibility that $\eta = 0$, i.e. $\omega_i(t) \to r$. However, this contradicts the strategies being interior. *Q.E.D.*