Information Acquisition and Time-Risk Preference

Daniel Chen*and Weijie Zhong†

September 7, 2024

Abstract

An agent acquires information dynamically until her belief about a binary state reaches an upper or lower threshold. She can choose any signal process subject to a constraint on the rate of entropy reduction. Strategies are ordered by "time risk"—the dispersion of the distribution of threshold-hitting times. We construct a strategy maximizing time risk (*Greedy Exploitation*) and one minimizing it (*Pure Accumulation*). Under either strategy, beliefs follow a compensated Poisson process. In the former, beliefs jump to the threshold that is closer in Bregman divergence. In the latter, beliefs jump to the unique point with the same entropy as the current belief.

1 Introduction

In this paper, we study information acquisition by a Bayesian agent about an unknown binary state that may be either zero or one. The agent wants to be reasonably certain about the state and is satisfied once her posterior belief that the state is one reaches either an upper or a lower threshold. She earns a unit payoff when this happens. She has great flexibility in how she can learn but has a resource constraint that limits her rate of learning. That is, she can choose any posterior belief process subject to a constraint on the rate of entropy reduction.

---

*dtchen@princeton.edu, Department of Economics, Princeton University

†wjzhong@stanford.edu, Graduate School of Business, Stanford University

Our simple model captures three important features of many economic settings: flexible learning, limited resources, and threshold decision rules. These features often appear in the contexts of research and development, clinical trials, digital marketing, user-experience testing, and others. For example, consider a data scientist at Facebook who must assess whether to introduce a new feature to the platform. The unknown state is whether adding the feature increases or decreases user engagement (and thus profits). The data scientist can learn about the state by conducting A/B tests. To provide incentives, her manager offers her a bonus when she has learned sufficiently precisely about the state (i.e. her A/B tests identify the state at some minimal level of statistical power[1]). She can freely design many aspects of the A/B tests—e.g., she can select the subpopulation of users that she gives access to the feature and she can adjust the length of time a user has access to the feature among other choices. However, there are limits to what she can do that constrain her speed of learning—e.g., her manager does not allow her to implement the feature for all users simultaneously as doing so could be disastrous for profits if the feature is disliked or she can only implement the feature for a given user for a short amount of time, etc.

Our main contribution is to show how, in such settings, the agent's optimal learning strategy depends on her *time-risk preferences*. That is, we allow for a rich set of preferences over threshold-hitting times beyond the standard case of exponential discounting. We say that the agent is time-risk loving (averse) whenever her utility over threshold-hitting times is convex (concave).[2] We derive a learning strategy that is optimal whenever the agent is time-risk loving and a strategy that is optimal whenever she is time-risk averse. Critically, the optimality of these strategies does not depend on the shape of the agent's utility function beyond its convexity or concavity.

In reality, there are many reasons why individuals may have time preferences that differ from the predominantly studied case of exponential discounting. This may be due to external factors such as explicit discounting, exogenous decision deadlines, and flow costs associated with foregone opportunities while learning. It may also be due to internal factors such as present bias resulting from hyperbolic discounting. We provide a simple framework that allows for these factors when studying optimal learning and

---

[1]In our binary state model, there is a one-to-one mapping between thresholds and the statistical power of a test (that is, the likelihood of Type II errors).

[2]In this paper, we model the agent as an expected-utility maximizer. However, it is easy to see that all of our results will go through as long as the agent has a preference relation over stopping times that is monotonic in the mean-preserving spread order.

derive strategies that are uniformly optimal up to the convexity or concavity of the utility function over threshold-hitting times.

We now briefly describe the two learning strategies and provide intuition for them. When the agent is time-risk loving, a *Greedy Exploitation* strategy is optimal. Under this strategy, the agent myopically maximizes the instantaneous rate that her beliefs jump to a threshold. She acquires a rare but decisive signal that, upon arrival, induces her belief to jump to the threshold that is closest in the Bregman divergence. By targeting the closer threshold, she can jump at a faster rate without violating her constraint on the rate of entropy reduction. In the absence of a signal arrival, her belief experiences compensating drift in the direction of the farther threshold. Eventually, her belief reaches a point that is equidistant in the Bregman divergence to the two thresholds. At this point, she acquires signals such that her beliefs may jump to either threshold but at rates set so that there is no net compensating drift so that her belief is stationary in the absence of a jump.

Intuitively, Greedy Exploitation is optimal because it produces a very risky distribution of threshold-hitting times. Because the strategy is greedy, it yields a high probability of an early hitting time. However, in the absence of a jump, since beliefs drift towards the farther threshold, the jump rate decreases so that the expected amount of time remaining until a threshold is reached *increases*. In this sense, the agent makes no "progress" in the absence of a jump. Thus there is a high probability of late threshold hitting times as well. We, in fact, show that among all strategies that exhaust the agent's resources (in that the constraint on the rate of entropy reduction is binding at all points in time), Greedy Exploitation yields a distribution of hitting times that is maximal in the mean-preserving spread order. In this sense, it maximizes time risk.

When the agent is time-risk averse, she instead seeks to minimize time risk. In this case, an optimal strategy is *Pure Accumulation.* Under this strategy, her beliefs reach a threshold at a *deterministic* time. Her beliefs follow a compensated Poisson process that jumps in the direction of the threshold that is farther away but to an interior belief that has the same entropy as her current belief. In the absence of a jump, her belief experiences a compensating drift toward the closer threshold. Pure accumulation is a continuous-time analog of the "suspense-maximal" policy in Ely et al. (2015). The strategy is, in effect, the "opposite" of Greedy Exploitation. Because jumps are always to beliefs with the same entropy as the current belief, in the

event of a jump, there is no progress: a jump does not reduce the expected amount of time until a threshold is hit. Instead, all progress is made through drift, which is why the threshold hitting time is deterministic. Thus, Pure Accumulation entails *no* time risk. It, therefore, produces a distribution of hitting times that is minimal in the mean-preserving spread order among all strategies that exhaust the agent's resources (in that the constraint on the rate of entropy reduction is binding at all points in time).

Our analysis of optimal learning through the lens of time-risk preferences has implications for both information acquisition in practice and economic modeling. Our model predicts that an agent who is time-risk-loving should use Greedy Exploitation, whereas an agent who is time-risk-averse should opt for Pure Accumulation, provided the agent has access to these learning strategies. Moreover, whenever the agent's space of available learning strategies is sufficiently rich, Brownian signals are generally suboptimal. Thus, when writing models where agents acquire information with parameterized signal structures, economists should be mindful of whether these signal structures are without loss of optimality given the time-risk preferences of the agents they seek to model.

2 Related Literature

Our paper contributes to a large literature on information acquisition. As in Wald (1947) and Arrow et al. (1949) we study a sequential sampling problem but allow the agent to flexibly design the signal process as in Zhong (2022), Hébert & Woodford (2017), Hébert & Woodford (2023), Steiner et al. (2017), and Georgiadis-Harris (2023). Whereas most of these papers restrict attention to the standard case of exponential discounting or a linear delay cost, we allow for more general time preferences. For example, Zhong (2022) assumes exponential discounting which implies time-risk loving preferences whereas Hébert & Woodford (2017) and Hébert & Woodford (2023) assume a linear delay cost which implies time-risk neutral preferences. Our results suggest that the assumed time-risk preferences dictate the qualitative features of the optimal strategies identified in these papers.[3]

---

[3]Hébert & Woodford (2023) allow for both discounting and a linear delay cost. However, they consider the time-risk neutral limit for the majority of their analysis. Their main objective is to study how different costs or constraints on information acquistion affect optimal learning strategies which is orthogonal to the objective of our paper. Zhong (2022) assumes a belief-dependent payoff

Pure Accumulation is a continuous-time variant of the suspense-maximal strategy introduced by Ely et al. (2015) in a discrete-time setting with a deterministic and finite time horizon. In Ely et al. (2015), the suspense-maximal strategy is optimal in that it maximizes an expected utility that is increasing in the variances of future beliefs. Georgiadis-Harris (2023) also finds that a strategy similar to Pure Accumulation is optimal in a setting where the stopping time is *exogenous* and does not depend on the learning strategy.[4] The mechanisms behind the optimality of Pure Accumulation in both of these papers are distinct from that of this paper where it is shown to be time-risk aversion.

In our analysis, the key summary statistic that determines the payoff from a strategy is the distribution of the time that the agent's belief first reaches a threshold. This statistic defines a *time lottery*, which is an object studied in an emerging literature on time-risk preferences. Chesson & Viscusi (2003) and Chen (2013) show that the expected discounted utility framework implies preferences that are *risk seeking over time lotteries* (RSTL). DeJarnette et al. (2020) show that within a broad class of models RSTL can not be violated if there is stochastic impatience. However, experimental evidence suggests that subjects are often *risk averse over time lotteries* (RATL) (Chesson & Viscusi (2003); Onay & Öncüler (2007)). Our model accommodates both RSTL and RATL and shows that optimal information acquisition can differ dramatically under different time-risk preferences.

The optimal learning strategies that we identify are qualitatively similar to learning strategies that have been assumed in reduced form by many papers in the literature. For example, Che & Mierendorff (2019), Mayskaya (2019), and Nikandrova & Pancs (2018) adopt a framework that restricts attention to Poisson signal processes in order to study optimal stopping with endogenous information. Poisson signals are also often assumed in the literature on strategic experimentation (see a survey by Hörner & Skrzypacz (2017)). We show that Poisson learning has an optimization foundation under time-risk loving preferences.[5] The Pure Accumulation strategy is also related to classic models on the timing of innovation introduced by Dasgupta &

---

function that does not have a threshold structure but shows that the optimal learning strategy is nevertheless similar to Greedy Exploitation.

[4]In our paper, the stopping is determined by the exogenous thresholds as well as the *endogenously* chosen learning strategy.

[5]To be clear, we do not show that *all* Poisson learning strategies are optimal but rather that optimal strategies in our setting involve Poisson learning.

Stiglitz (1980) and Lee & Wilde (1980) (see a survey by Reinganum (1989)) which involve a deterministic time of innovation. The models in these papers assume a reduced-form learning process and are non-Bayesian. However, we show that the learning strategies in these papers can emerge endogenously in a Bayesian information acquisition framework when agents have time risk-averse preferences.

Our model also allows for Gaussian learning strategies. Gaussian signal processes are often assumed in reduced-form learning models (see for example Moscarini & Smith (2001); Ke & Villas-Boas (2019); Liang et al. (2019); Morris & Strack (2019)). Also, *drift-diffusion models* (DDM) of binary choice problems appear in Ratcliff & Rouder (1998) and Fudenberg et al. (2018). However, our results imply that Gaussian learning can not be justified by optimality except in the knife-edge case when agents have time-risk neutral preferences provided information can be acquired flexibly.

The optimality of a greedy strategy is also the main result in Liang et al. (2019). However, the mechanisms in our papers are very different. Liang et al. (2019)'s result crucially depends on the linear-Gaussian setup with exogenously given Gaussian information sources and holds for any time preferences. Our result allows for a flexible and endogenous choice of information sources, but crucially depends on time preferences. Also related is Gossner et al. (2021) which studies a model with learning but with exogenous information sources and derives a greedy learning strategy that leads to the shortest stopping time (when the belief hits an exogenously given threshold) in the sense of first-order stochastic dominance.

We model limits on the agent's learning resources via a constraint on the rate of entropy reduction. That is, the rate of resource depletion is determined by a *uniformly posterior separable* (UPS) function. The rational inattention literature also typically models information costs or constraints using a UPS function (Sims (2003); Matějka & McKay (2014); Steiner et al. (2017); Caplin et al. (2017)). Microfoundations for the UPS formulation can be found in Frankel & Kamenica (2019); Caplin et al. (2017); Zhong & Bloedel (2021); Morris & Strack (2019). In our paper, the UPS information constraint ensures that the expected threshold hitting time is equalized for all exhaustive strategies, which allows us to isolate the role of time *risk* in information acquisition. By Theorem 3 in Zhong (2022), a UPS information constraint is both necessary and sufficient for the expected learning time to be equalized for all exhaustive strategies.

3 Model

This section presents a simple model of an agent who wants to learn over time about an unknown state. The unknown state ω takes values in $\{0, 1\}$. At $t = 0$, the agent believes that ω is 1 with probability $\mu \in (0, 1)$. She receives a unit payoff when her posterior belief μ_t that ω is 1 reaches either an upper threshold $\overline{\mu} \in (\mu, 1)$ or a lower threshold $\underline{\mu} \in (0, \mu)$. However, she is impatient and her utility is a decreasing function $\rho : \mathbb{R}_+ \to \mathbb{R}$ of the threshold-hitting time. When ρ is convex, we say that the agent is *time-risk loving*. When ρ is concave we say that she is *time-risk averse*. A common case when ρ is convex arises when the agent discounts time at some exponential rate r: $\rho(t) = e^{-rt}$. On the other hand, a case when ρ is concave is if the agent incurs a flow cost $c(\cdot)$ of delay that increases with time: $\rho(t) = 1 - \int_0^t c(s)\, ds$.

The agent has great flexibility in how she can learn about ω but has limited resources and cannot learn infinitely fast. Let $\mathcal{M}$ denote the set of processes $\boldsymbol{\mu} = \{\mu_t; t \geq 0\}$ that take values in $[0, 1]$ and satisfy a stochastic differential equation (allowing for jumps) of the form

$$d\mu_t = \sum_{i=1}^{N} (\nu^i(t, \mu_t) - \mu_t) \left[dJ_t^i(\lambda^i(t, \mu_t)) - \lambda^i(t, \mu_t)\, dt \right] + \sum_{j=1}^{M} \sigma^j(t, \mu_t)\, dZ_t^j \quad (1)$$

with $\mu_0 = \mu$ for some positive integers N and M and functions $\{\nu^i\}_{i=1}^N$, $\{\lambda^i\}_{i=1}^N$, and $\{\sigma^j\}_{j=1}^M$. Above, each Z_t^j is a standard Brownian Motion and each J_t^i is a Poisson point process that ticks at rate $\lambda^i(t, \mu_t^i)$. If J_t^i ticks at time t, the belief μ_t jumps to the point $\nu^i(t, \mu_t)$. The number of distinct points that the belief can jump to at time t is the integer N and the number of distinct Brownian Motions is the integer M.

We assume that the agent can directly choose any belief process in $\mathcal{M}$ such that[6]

$$\mathbb{E}\left[H(\mu_s) - H(\mu_t) \Big| \mathcal{F}_t \right] \leq I(s - t) \quad (2)$$

for all t, s such that $s \geq t$ where $\{\mathcal{F}_t\}$ is the natural filtration of $\boldsymbol{\mu}$, H is a C^2 strictly convex function defined on $[0, 1]$, and $I \geq 0$ is a constant. $-H$ maps the agent's belief to its associated *entropy*. Thus, equation (2) is a constraint on the rate of entropy reduction. A special case is when $-H$ is Shannon's entropy so that (2) amounts to

---

[6]Our restriction to jump-diffusion belief processes is without loss of generality within the larger class of cádlág processes such that (2) is well defined. This follows from Proposition 1 in Georgiadis-Harris (2023).

a constraint on the well-known mutual information rate. Without loss of generality, we normalize $H(\overline{\mu}) = H(\underline{\mu}) = 0$ and $I = 1$. This can be done by redefining H to be

$$\frac{1}{I}\left[H(\mu) - \frac{H(\overline{\mu}) - H(\underline{\mu})}{\overline{\mu} - \underline{\mu}}(\mu - \underline{\mu})\right].$$

The same belief processes satisfy (2) before and after this redefinition because beliefs are martingales (and thus, the drift of the second term in brackets above is zero). The normalization is convenient because, provided (2) is binding at all times, the optional stopping theorem[7] implies that the expected time remaining until a threshold is reached is simply the current entropy:

$$-H(\mu_t) = \mathbb{E}\left[H(\mu_\tau) - H(\mu_t)|\mathcal{F}_t\right] = \mathbb{E}\left[\tau - t|\mathcal{F}_t\right].$$

To state the agent's learning problem, let $\tau_{\boldsymbol{\mu}}$ be the first time that her beliefs reach a threshold:

$$\tau_{\boldsymbol{\mu}} := \inf\{t|\mu_t \in [0, \underline{\mu}] \cup [\overline{\mu}, 1]\}.$$

Because $\tau_{\boldsymbol{\mu}}$ may be ∞ with positive probability for some belief processes (e.g., the agent can stop learning) we set $\rho(\infty) = -\infty$ to ensure that the agent never selects these processes.[8]

She solves

$$\max_{\boldsymbol{\mu} \in \mathcal{M}} \mathbb{E}[\rho(\tau_{\boldsymbol{\mu}})] \tag{3}$$

such that (2) holds.

Our simple model makes several assumptions in order to isolate the connection between optimal learning and time-risk preferences which is the main focus of our paper. For example, we assume that the agent experiences no costs from learning (as long as (2) is satisfied) and that she earns a common payoff regardless of the threshold that she ultimately hits. Though these assumptions are restrictive, we believe that our model aligns well with a number of economic applications as discussed in the Introduction. A key property of our model that allows for our main insights is that it accommodates general time-risk preferences. Indeed, our main contribution is to identify optimal learning strategies that apply across a range of time preferences

---

[7]See Theorem 3.22 of Karatzas & Shreve (1998).
[8]That is, we extend the domain and range of ρ.

and to highlight that it is the agent's attitude towards *time risk* that determines the qualitative properties of these strategies. Common examples of time-risk loving preferences beyond exponential discounting include: (i) hyperbolic discounting where $\rho(t) = (1+\alpha t)^{-\gamma/\alpha}$ (Loewenstein & Prelec, 1992); (ii) generalized expected discounted utility: $\rho(t) = \phi(e^{-rt})$ where ϕ is increasing and convex (DeJarnette et al., 2020); and linear delay cost up until a deadline (iii) $\rho^T(t) = \max\{T - t, 0\}$. Time-risk averse preferences are less commonly assumed in the literature though there is growing experimental evidence that agents may often be time-risk averse (Chesson & Viscusi, 2003; Onay & Öncüler, 2007). As discussed at the start of this section, time-risk aversion can arise when agents have flow costs of delay that increase with time.

4   Optimal Learning and Time-Risk Preferences

In this section, we present our main results: a strategy that is optimal whenever the agent is time-risk loving and a strategy that is optimal whenever she is time-risk averse. These results illustrate the connection between optimal learning and time-risk preferences.

4.1   Time-Risk Loving

We first consider the case when the agent is time-risk loving. A formal statement of her optimal learning strategy is given below in Definition 1 but we begin by describing it informally here.

An optimal strategy for the agent is *Greedy Exploitation* and is illustrated in Figure 1.

Let

$$d_H(\tilde{\mu}, \hat{\mu}) = H(\tilde{\mu}) - H(\hat{\mu}) - H'(\hat{\mu})(\tilde{\mu} - \hat{\mu})$$

denote the Bregman divergence between any two beliefs $\hat{\mu}$ and $\tilde{\mu}$ under the function H. In Figure 1, μ^* represents the unique belief that is equidistant in the Bregman divergence to the two thresholds: $d_H(\overline{\mu}, \mu^*) = d_H(\underline{\mu}, \mu^*)$. Initially, the agent's beliefs either jump to the threshold that is closer in Bregman divergence (in this case $\overline{\mu}$) or experience compensating drift towards the other threshold. By targeting the closer
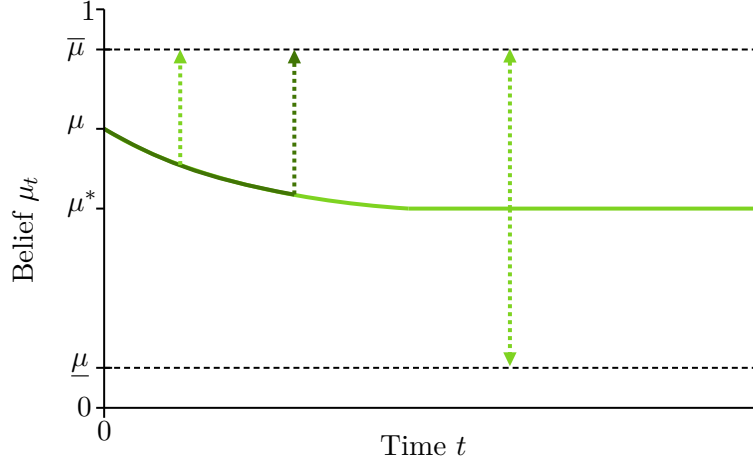
9

Figure 1: Greedy Exploitation

Notes: In dark green, we plot one possible realization of the belief path $\boldsymbol{\mu}^G$ under the Greedy Exploitation strategy. In light green, we plot other possible realizations of $\boldsymbol{\mu}^G$. The dashed lines with arrows represent jumps in the belief. The figure is computed for the case $H(\tilde{\mu}) = \tilde{\mu}^2$.

threshold, the agent greedily maximizes the "chance" that her beliefs reach a threshold in the "next instant." This is because the agent's beliefs can jump at a faster rate when she targets the closer threshold without violating her resource constraint (2). After some time, in the absence of a jump, her beliefs eventually drift to μ^*. At this point, her beliefs may jump to either threshold. The jump rates to the respective thresholds are such that there is no net compensating drift, and so, in the absence of a jump, her beliefs remain stationary.

Definition 1. The *Greedy Exploitation* strategy $\boldsymbol{\mu}^G$ is defined as follows. Let $\mu^* \in (0, 1)$ be the unique belief such that $d_H(\overline{\mu}, \mu^*) = d_H(\underline{\mu}, \mu^*)$,

- While $\mu_t^G > \mu^*$, her beliefs evolve according to

$$d\mu_t^G = (\overline{\mu} - \mu_t^G) \left[dJ_t^1(\lambda_t) - \lambda_t \, dt \right]$$

where $\lambda_t = I/d_H(\overline{\mu}, \mu_t^G)$.

10

- While $\mu_t^G = \mu^*$, her beliefs evolve according to

$$d\mu_t^G = (\overline{\mu} - \mu_t^G)\, dJ_t^2 \left(\frac{\mu_t^G - \underline{\mu}}{\overline{\mu} - \underline{\mu}} \lambda^* \right) + (\underline{\mu} - \mu_t^G)\, dJ_t^3 \left(\frac{\overline{\mu} - \mu_t^G}{\overline{\mu} - \underline{\mu}} \lambda^* \right)$$

where $\lambda^* = 1/d_H(\overline{\mu}, \mu^*)$.

- While $\mu_t^G < \mu^*$, her beliefs evolve according to

$$d\mu_t^G = (\underline{\mu} - \mu_t^G) \left[dJ_t^1(\lambda_t) - \lambda_t\, dt \right]$$

where $\lambda_t = 1/d_H(\underline{\mu}, \mu_t^G)$.

Above J_t^1, J_t^2, and J_t^3 are independent Poisson point processes with jump rates indicated in parentheses.

Theorem 1. *If the agent is time-risk loving, then Greedy Exploitation is optimal.*

Before we sketch the proof of Theorem 1, we note that because Greedy Exploitation is uniformly optimal for all convex ρ, it must induce the riskiest distribution of threshold hitting times among all strategies that are exhaustive in that (2) is satisfied at all points in time. To make this precise, we first state the following definition.

Definition 2. $\mathcal{T} = \left\{ \tau_{\boldsymbol{\mu}} \middle| \boldsymbol{\mu} \in \mathcal{M} \text{ such that (2) binds at all } t \right\}$.

The Greedy Exploitation strategy produces a threshold-hitting time that is maximal in the mean-preserving spread order among all threshold-hitting times in this set.

Corollary 1.1. *It holds that $\tau_{\boldsymbol{\mu}^G} \succeq_{\mathrm{mps}} \tau$ for each $\tau \in \mathcal{T}$.*

This result hinges on our assumption that the constraint on learning is of the form in (2). Because of this assumption, all exhaustive strategies have the same expected threshold hitting time, which is equal to the initial entropy $-H(\mu)$.

Proof of Theorem 1. The proof proceeds in seven steps.

Step 1. Set of Basis Discount Functions—We observe that any nonnegative convex ρ can be written as a conical combination of functions of the form

$$\rho_T(t) = \max\{T - t, 0\}$$

11

where $T \geq 0$. Thus, if Greedy Exploitation is optimal for each ρ_T, $T \geq 0$, then it must be optimal for any nonnegative convex ρ. In fact, it will necessarily be optimal for *any* convex ρ, including ρ that may take on negative values.

Lemma 1. *If Greedy Exploitation solves* (3) *for each ρ_T where $T \geq 0$, then it is optimal for any convex ρ.*

Proof of Lemma 1. See Theorem 3.6 in Müller (1996).[9] $\qquad\square$

Step 2. Candidate Value Function and Some Notation—Let $V(\mu, T) = \mathbb{E}\left[\rho_T(\tau_{\boldsymbol{\mu}^G})|\mu_0^G = \mu\right]$ denote the value function for Greedy Exploitation when $\rho = \rho_T$. That is, for each $T \geq 0$, let

$$V(\mu, T) = \begin{cases} \int_0^T (T-t)\lambda_t^G e^{-\int_0^t \lambda_z^G\, dz}\, dt, & \mu \in (\underline{\mu}, \overline{\mu}) \\ T, & \mu \in \{\underline{\mu}, \overline{\mu}\}. \end{cases} \tag{4}$$

In what follows, it is useful to observe that $\partial V(\mu, T)/\partial T = U(\mu, T)$ whenever $\mu \in (\underline{\mu}, \overline{\mu})$ where

$$U(\mu, T) = \int_0^T \lambda_t^G e^{-\int_0^t \lambda_z^G\, dz}\, dt$$

is the probability that a jump arrives by time T under Greedy Exploitation. It is easy to show that $\partial U(\mu, T)/\partial \mu > 0$ if $\mu \in [\mu^*, \overline{\mu})$ and that $\partial U(\mu, T)/\partial \mu < 0$ if $\mu \in (\underline{\mu}, \mu^*]$.

To ease the exposition, we adopt the following notation. We set $V_T(\mu) = V(\mu, T)$ and $U_T(\mu) = U(\mu, T)$. Also, given a function f and any two beliefs ν and μ, we let

$$d_f(\nu, \mu) = f(\nu) - f(\mu) - f'(\mu)(\nu - \mu)$$

whenever $f'(\mu)$ is well-defined. Note that if f is convex, then d_f is a Bregman divergence.

Step 3. Verification Lemma—To verify the optimality of Greedy Exploitation we use the following Lemma 2 which states that it suffices to prove that V satisfies the Hamilton-Jacobi-Bellman (HJB) equation (5).

---

[9]To apply Theorem 3.6 in Müller (1996) recall that because ρ is decreasing, any optimal strategy is exhaustive and thus has the expected threshold hitting time of $-H(\mu)$.

Lemma 2. *Given $T > 0$, if V in (4) satisfies*

$$U_t(\mu) = \max\left\{\max_\nu \frac{d_{V_t}(\nu, \mu)}{d_H(\nu, \mu)}, \frac{V_t''(\mu)}{H''(\mu)}\right\} \tag{5}$$

at each $(\mu, t) \in (\underline{\mu}, \overline{\mu}) \times [0, T]$ then $V_T(\mu)$ is equal to (3) when $\rho = \rho^T$.

Proof of Lemma 2. We first assert that condition (5) is equivalent to

$$U_t(\mu) = \max_{\{\nu^i\}, \{\lambda^i\}, \sigma} \mathcal{A}^{\nu, \lambda, \sigma} V_t(\mu) \tag{6}$$

$$\text{s.t } \mathcal{A}^{\nu, \lambda, \sigma} H(\mu_t) \leq 1$$

where $\mathcal{A}^{\nu, \lambda, \sigma}$ is the operator defined for functions $f \in C^2(\underline{\mu}, \overline{\mu})$ by

$$\mathcal{A}^{\nu, \lambda, \sigma} f(\mu) = \sum_i \lambda^i d_f(\nu^i, \mu) + \frac{1}{2} \sum_j (\sigma^j)^2 f''(\mu).$$

That is, $\mathcal{A}^{\nu, \lambda, \sigma}$ is the infinitesimal generator for the compensated jump diffusion process (1). Because $\mathcal{A}^{\nu, \lambda, \sigma}$ is additively separable, it suffices to select a single jump point or volatility to achieve the max in (6). The jump point or volatility is chosen to maximize the "bang-for-the-buck"—that is, the ratio of the drift of V to the drift of H. Therefore, (5) and (6) must be equivalent.

Next, suppose that (5) is satisfied. Consider an arbitrary admissible strategy $\{\nu^i\}$, $\{\lambda^i\}$, $\{\sigma^j\}$ with induced first threshold-hitting time τ. We have

$$
\begin{aligned}
V_T(\mu) =& \mathbb{E}\left[V_{T-\tau \wedge T}(\mu_{\tau \wedge T}) - \int_0^{\tau \wedge T}\left[-U_{T-t}(\mu_t) + \mathcal{A}^{\nu, \lambda, \sigma} V_{T-t}(\mu_t)\right] dt \right.\\
&+ \sum_j \int_0^{\tau \wedge T} \frac{\partial V_{T-t}(\mu_t)}{\partial \mu} \sigma_t^j \, dZ_t^j \\
&\left.+ \sum_i \int_0^{\tau \wedge T} \left[V_{T-t}(\nu_t^i) - V_{T-t}(\mu_t)\right] \left(dJ_t^i(\lambda_t^i) - \lambda_t^i \, dt\right)\right] \\
=& \mathbb{E}\left[V_{T-\tau \wedge T}(\mu_{\tau \wedge T}) - \int_0^{\tau \wedge T}\left[-U_{T-t}(\mu_t) + \mathcal{A}^{\nu, \lambda, \sigma} V_{T-t}(\mu_t)\right] dt\right] \\
\geq& \mathbb{E}\left[V_{T-\tau \wedge T}(\mu_{\tau \wedge T})\right] \\
\geq& \mathbb{E}\left[\rho_T(\tau)\right]
\end{aligned}
$$

where the first equality uses Itô's formula for jump diffusions and the fact that $\partial V / \partial T = U$ as noted in Step 2, the second equality follows from the fact that $\partial V_{T-t}(\mu_t)/\partial\mu$ and V_{T-t} are bounded which implies that the diffusion and jump terms are true martingales,[10] the first inequality follows from (6), and the last inequality follows from the definition of V. $\qquad\square$

Step 4. $\{V_{T-t}(\mu_t^G)\}$ *is a Martingale—* The remaining steps verify that V satisfies the conditions of Lemma 2. We begin with Lemma 3 which states that if the inner and outer max on the right-hand side of (5) is achieved by Greedy Exploitation then (5) is statisfied.

Lemma 3. *At each $t \in [0,\infty)$ the following hold:*

1. *If $\mu \geq \mu^*$, then*

$$U_t(\mu) = \frac{d_{V_t}(\overline{\mu}, \mu)}{d_H(\overline{\mu}, \mu)}.$$

2. *If $\mu \leq \mu^*$, then*

$$U_t(\mu) = \frac{d_{V_t}(\underline{\mu}, \mu)}{d_H(\underline{\mu}, \mu)}.$$

Proof of Lemma 3. Because $V_{T-t}(\mu_t^G) = \mathbb{E}\left[\rho_T(\tau_{\boldsymbol{\mu}^G})|\mu_t^G\right]$ and $\boldsymbol{\mu}^G$ is Markov it follows that $\{V_{T-t}(\mu_t^G)\}$ is a martingale for any given $T \geq 0$. By Itô's formula, the drift of $\{V_{T-t}(\mu_t^G)\}$ is zero if and only if conditions 1 and 2 of the lemma are satisfied. $\qquad\square$

*Step 5. Unimprovable by Poisson Learning—*The following Lemma 4 shows that Greedy Exploitation can not be improved on by any alternative Poisson learning strategy.

Lemma 4. *At each $(\mu, t) \in (\underline{\mu}, \overline{\mu}) \times [0, \infty)$ it holds that*

$$U_t(\mu) = \max_{\nu} \frac{d_{V_t}(\nu, \mu)}{d_H(\nu, \mu)}. \tag{7}$$

Proof of Lemma 4. We will prove the lemma when $\mu > \mu^*$. The proof when $\mu \leq \mu^*$ is analogous. By Lemma 3, it suffices to show that $\overline{\mu}$ achieves the max in (7). We split the proof into three cases.

---

[10]See Theorem 51 of Protter (2005).

- Case 1: $\nu \geq \mu$. We will show that $\nu = \overline{\mu}$ is the global maximizer of $d_{V_t}(\nu, \mu)/d_H(\nu, \mu)$ in the region $\nu \geq \mu$. To start, we observe that

$$\frac{d}{d\nu} \frac{d_{V_t}(\nu, \mu)}{d_H(\nu, \mu)} = \frac{V_t'(\nu) - V_t'(\mu)}{d_H(\nu, \mu)} - \frac{d_{V_t}(\nu, \mu)}{d_H(\nu, \mu)^2} \left[H'(\nu) - H'(\mu) \right].$$

This derivative is negative if and only if

$$\frac{V_t'(\nu) - V_t'(\mu)}{H'(\nu) - H'(\mu)} \geq \frac{d_{V_t}(\nu, \mu)}{d_H(\nu, \mu)}. \tag{8}$$

which is equivalent to

$$\frac{d_{V_t}(\overline{\mu}, \mu) - d_{V_t}(\overline{\mu}, \nu)}{d_H(\overline{\mu}, \mu) - d_H(\overline{\mu}, \nu)} \geq \frac{d_{V_t}(\nu, \mu)}{d_H(\nu, \mu)}. \tag{9}$$

Notice that (9) holds with equality at $\nu = \overline{\mu}$. We will show that $\overline{\mu}$ and in fact any local extremum of $d_{V_t}(\nu, \mu)/d_H(\nu, \mu)$ in the region $\nu \in (\mu, \overline{\mu}]$ is a local maximum. This immediately implies that $\overline{\mu}$ must actually be a global maximum in this region.

At any local extremum (9) holds with equality. We can prove that all such local extrema are necessarily local maxima simply by proving that the derivative of the left-hand side of (9) is negative. This is because the derivative of the right-hand side is always zero at a local extremum since the right-hand side expression is the objective $d_{V_t}(\nu, \mu)/d_H(\nu, \mu)$. The left-hand side of (9) is decreasing in ν because

$$\frac{d}{d\nu} \frac{d_{V_t}(\overline{\mu}, \mu) - d_{V_t}(\overline{\mu}, \nu)}{d_H(\overline{\mu}, \mu) - d_H(\overline{\mu}, \nu)} = \frac{d}{d\nu} \frac{U_t(\mu)d_H(\overline{\mu}, \mu) - U_t(\nu)d_H(\overline{\mu}, \nu)}{d_H(\overline{\mu}, \mu) - d_H(\overline{\mu}, \nu)}$$

$$< \frac{d}{d\nu} \frac{U_t(\mu)d_H(\overline{\mu}, \mu) - U_t(\mu)d_H(\overline{\mu}, \nu)}{d_H(\overline{\mu}, \mu) - d_H(\overline{\mu}, \nu)}$$

$$= 0$$

where we have used the fact that $d_{V_t}(\overline{\mu}, \nu)/d_H(\overline{\mu}, \nu) = U_t(\nu)$ from Lemma 3 and that $U_t(\nu)$ is increasing in ν as noted in Step 2.

- Case 2: $\nu \in (\mu^*, \mu)$. In this region, (following the same steps as in Case 1) it is

easy to show that $d_{V_t}(\nu, \mu)/d_H(\nu, \mu)$ is nondecreasing if

$$\frac{d_{V_t}(\overline{\mu}, \mu) - d_{V_t}(\overline{\mu}, \nu)}{d_H(\overline{\mu}, \mu) - d_H(\overline{\mu}, \nu)} \leq \frac{d_{V_t}(\nu, \mu)}{d_H(\nu, \mu)}. \tag{10}$$

This is the same condition as (9) except the inequality has flipped.

As before, to determine whether a local extremum is a maximum or minimum it suffices to check how the left-hand side changes as ν increases. In this case, the left-hand side is increasing. This can be seen since

$$\begin{aligned}
\frac{d}{d\nu} \frac{d_{V_t}(\overline{\mu}, \mu) - d_{V_t}(\overline{\mu}, \nu)}{d_H(\overline{\mu}, \mu) - d_H(\overline{\mu}, \nu)} &= \frac{d}{d\nu} \frac{U_t(\mu)d_H(\overline{\mu}, \mu) - U_t(\nu)d_H(\overline{\mu}, \nu)}{d_H(\overline{\mu}, \mu) - d_H(\overline{\mu}, \nu)} \\
&> \frac{d}{d\nu} \frac{U_t(\mu)d_H(\overline{\mu}, \mu) - U_t(\mu)d_H(\overline{\mu}, \nu)}{d_H(\overline{\mu}, \mu) - d_H(\overline{\mu}, \nu)} \\
&= 0
\end{aligned}$$

where we have used the fact that the denominator is negative in this case. Thus, in this region, any local extremum must be a local minimum. Thus, no point $\nu \in (\mu^*, \mu)$ can achieve the max in (7).

- Case 3: $\nu \in [\underline{\mu}, \mu^*]$. Following analogous steps to those used to derive (9), we find that $d_{V_t}(\nu, \mu)/d_H(\nu, \mu)$ is decreasing in ν if and only if

$$\frac{d_{V_t}(\underline{\mu}, \mu) - d_{V_t}(\underline{\mu}, \nu)}{d_H(\underline{\mu}, \mu) - d_H(\underline{\mu}, \nu)} < \frac{d_{V_t}(\nu, \mu)}{d_H(\nu, \mu)}. \tag{11}$$

We will prove that the left-hand side of (11) is bounded above by $d_{V_t}(\overline{\mu}, \mu)/d_H(\overline{\mu}, \mu)$. Thus, there can not be a point $\nu \in [\underline{\mu}, \mu^*]$ that achieves a higher value than $d_{V_t}(\overline{\mu}, \mu)/d_H(\overline{\mu}, \mu)$, since if there was, at that point, $d_{V_t}(\nu, \mu)/d_H(\nu, \mu)$ would be decreasing in ν.

To show this, we first observe that

$$d_{V_t}(\underline{\mu}, \mu) = d_{V_t}(\underline{\mu}, \overline{\mu}) + d_{V_t}(\overline{\mu}, \mu) - (\underline{\mu} - \overline{\mu})\left(V_t'(\mu) - V_t'(\overline{\mu})\right), \tag{12}$$

and

$$d_H(\underline{\mu}, \mu) = d_H(\underline{\mu}, \overline{\mu}) + d_H(\overline{\mu}, \mu) - (\underline{\mu} - \overline{\mu})\left(H'(\mu) - H'(\overline{\mu})\right). \tag{13}$$

16

Define $f(\mu)$ and $g(\mu)$ as

$$f(\mu) = d_{V_t}(\overline{\mu}, \mu) - (\underline{\mu} - \overline{\mu})\left(V_t'(\mu) - V_t'(\overline{\mu})\right) \tag{14}$$

and

$$g(\mu) = d_H(\overline{\mu}, \mu) - (\underline{\mu} - \overline{\mu})\left(H'(\mu) - H'(\overline{\mu})\right). \tag{15}$$

Since (8) binds when $\nu = \overline{\mu}$, it follows that

$$\frac{f(\mu)}{g(\mu)} = \frac{d_{V_t}(\overline{\mu}, \mu) - (\underline{\mu} - \overline{\mu})\left(V_t'(\mu) - V_t'(\overline{\mu})\right)}{d_H(\overline{\mu}, \mu) - (\underline{\mu} - \overline{\mu})\left(H'(\mu) - H'(\overline{\mu})\right)} = \frac{d_{V_t}(\overline{\mu}, \mu)}{d_H(\overline{\mu}, \mu)}. \tag{16}$$

Also since $d_{V_t}(\overline{\mu}, \mu^*)/d_H(\overline{\mu}, \mu^*) = d_{V_t}(\underline{\mu}, \mu^*)/d_H(\underline{\mu}, \mu^*)$,

$$\frac{f(\mu^*)}{g(\mu^*)} = \frac{d_{V_t}(\underline{\mu}, \overline{\mu}) + f(\mu^*)}{d_H(\underline{\mu}, \overline{\mu}) + g(\mu^*)} \Rightarrow \frac{f(\mu^*)}{g(\mu^*)} = \frac{d_{V_t}(\underline{\mu}, \overline{\mu})}{d_H(\underline{\mu}, \overline{\mu})}. \tag{17}$$

Thus,

$$
\begin{aligned}
\frac{d_{V_t}(\underline{\mu}, \mu) - d_{V_t}(\underline{\mu}, \nu)}{d_H(\underline{\mu}, \mu) - d_H(\underline{\mu}, \nu)} &= \frac{d_{V_t}(\underline{\mu}, \overline{\mu}) + f(\mu) - d_{V_t}(\underline{\mu}, \nu)}{d_H(\underline{\mu}, \overline{\mu}) + g(\mu) - d_H(\underline{\mu}, \nu)} \\
&= \frac{U_t(\mu^*)d_H(\underline{\mu}, \overline{\mu}) + U_t(\mu)g(\mu) - U_t(\nu)d_H(\underline{\mu}, \nu)}{d_H(\underline{\mu}, \overline{\mu}) + g(\mu) - d_H(\underline{\mu}, \nu)} \\
&\leq \frac{U_t(\mu^*)d_H(\underline{\mu}, \overline{\mu}) + U_t(\mu)g(\mu) - U_t(\mu^*)d_H(\underline{\mu}, \nu)}{d_H(\underline{\mu}, \overline{\mu}) + g(\mu) - d_H(\underline{\mu}, \nu)} \\
&\leq U_t(\mu) \\
&= \frac{d_{V_t}(\overline{\mu}, \mu)}{d_H(\overline{\mu}, \mu)}.
\end{aligned}
$$

as desired. The first line uses (12), (13), (14), and (15). The second line uses (16), (17), and Lemma 3. The third line uses the fact that $U_t(\nu)$ is decreasing for $\nu \in [\underline{\mu}, \mu^*]$ as noted in Step 2. The last line uses Lemma 3.

$\square$

Step 6. Unimprovable by Diffusion Learning—The following Lemma 5 shows that Greedy Exploitation can not be improved on by diffusion learning.

Lemma 5. *At each $(\mu, t) \in (\underline{\mu}, \overline{\mu}) \times [0, \infty)$, it holds that*

$$U_t(\mu) \geq \frac{V_t''(\mu)}{H''(\mu)}.$$

Proof. Recall from Step 2 the observation that $U_t'(\mu) > 0$ when $\mu \in (\mu^*, \overline{\mu})$. Thus

$$U_t'(\mu) = \frac{d}{d\mu} \frac{d_{V_t}(\overline{\mu}, \mu)}{d_H(\overline{\mu}, \mu)} = \frac{-d_H(\overline{\mu}, \mu) V_t''(\mu)(\overline{\mu} - \mu) + d_{V_t}(\overline{\mu}, \mu) H''(\mu)(\overline{\mu} - \mu)}{d_H(\overline{\mu}, \mu)^2} > 0$$

which implies that

$$U_t(\mu) = \frac{d_{V_t}(\overline{\mu}, \mu)}{d_H(\overline{\mu}, \mu)} > \frac{V_t''(\mu)}{H''(\mu)}$$

as desired. An analogous argument applies when $\mu \in (\underline{\mu}, \mu^*)$. Weak inequality when $\mu = \mu^*$ follows from continuity.

$\square$

Step 7. Putting it All Together—Lemmas 4 and 5 imply that (5) is satisfied by V for any $T \geq 0$. Lemma 2 then implies that Greedy Exploitation is optimal for any discount function of the form ρ_T, $T \geq 0$. Lemma 1 then implies optimality for all convex ρ. The proof of Theorem 1 is complete.

$\square$

4.2 Time-Risk Averse

When the agent is time-risk averse, her optimal learning strategy is *Pure Accumulation*, illustrated graphically below in Figure 2.

As discussed in Section 2, the Pure Accumulation strategy is closely related to the *suspense-maximal strategy* from Ely et al. (2015). Under this strategy, the agent's belief either jumps in the direction of the farther threshold or experiences compensating drift. When her belief jumps, it jumps to a point with the same entropy as her current belief so that all progress is made through drift.

Definition 3. The *Pure Accumulation* strategy is defined as follows. Let $\mu^H : [0, 1] \setminus \{\mu^*\} \to [0, 1]$ denote the function that maps a belief $\hat{\mu}$ to the unique belief $\mu^H(\hat{\mu}) \neq \hat{\mu}$ such that $H(\mu^H(\hat{\mu})) = H(\hat{\mu})$. Under Pure accumulation, the agent's beliefs evolve
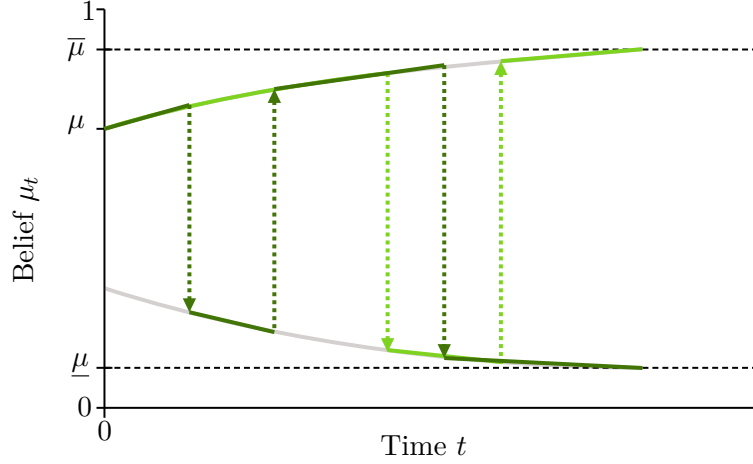
18

Figure 2: Pure Accumulation

Notes. The dark green curve represents one possible belief path of $\boldsymbol{\mu}^P$. The vertical segments represent jumps. The light green curve represents another possible belief path of $\boldsymbol{\mu}^P$. The figure is computed for the case $H(\tilde{\mu}) = \tilde{\mu}^2$.

according to

$$d\mu_t^P = \left[\mu^H(\mu_t^P) - \mu_t^P\right] dJ_t(\lambda_t) - \lambda_t \left[\mu^H(\mu_t^P) - \mu_t^P\right] dt$$

where J_t is a Poisson point process that ticks at rate $\lambda_t = 1/d_H(\mu^H(\mu_t^A), \mu_t^A)$.

Theorem 2. *If the agent is time-risk averse, then Pure Accumulation is optimal.*

Proof. Under Pure Accumulation, the agent is guaranteed to hit a threshold at the deterministic time $t = -H(\mu)$. $\qquad\square$

Because Pure Accumulation entails no time risk, we immediately have the following result.

Corollary 2.1. *It holds that $\tau \succeq_{\mathrm{mps}} \tau_{\boldsymbol{\mu}^A}$ for each $\tau \in \mathcal{T}$.*

5 Concluding Discussion

In this paper, we have studied the relationship between time-risk preferences and optimal information acquisition. We have shown that an optimal strategy for a time-risk

loving agent is Greedy Exploitation. This strategy produces the riskiest distribution over threshold hitting times among all exhaustive strategies. On the other hand, an optimal strategy for a time-risk averse agent is Pure Accumulation. This strategy produces a deterministic threshold hitting time and thus entails no time risk. Both of these strategies are uniformly optimal up to the convexity or concavity of the utility function, provided the agent is impatient. Thus, they are immune to dynamic inconsistency. In practice, agents may have time preferences that differ from the well-studied case of exponential discounting. Our analysis offers some insight into how these agents may seek to acquire information and the kinds of signal structures economists may consider using when modeling these agents.

In order to illustrate the connection between learning and time-risk preferences as sharply as possible we have made a number of special assumptions. The assumptions of binary states, fixed stopping thresholds, and a capacity constraint on learning speed are critical because they ensure that all exhaustive strategies have the same expected threshold-hitting times.[11] The key reason why Greedy Exploitation and Pure Accumulation are optimal is because they respectively yield the maximal and minimal threshold-hitting times in the mean-preserving spread order among all exhaustive strategies. This allows us to emphasize that it is time *risk* preferences that determines their optimality. The assumption that payoffs depend only on threshold hitting times and not on which threshold is hit allows us to derive closed-form solutions but is not critical for the economic insights.

Indeed, we anticipate that many of the qualitative properties of Greedy Exploitation or Pure Accumulation will persist under other model formulations, for example, with other costs of learning, multiple states, and endogenously chosen stopping thresholds (there are already examples in the literature where this is so as reviewed in Section 2). It is certainly possible to extend our model of optimal learning to accommodate these more general environments though explicit charaterizations will be rare.[12] The advantage of our special setup is that it is possible to solve for optimal strategies that are explicit and uniformly optimal for large classes of payoffs and,

---

[11]Specifically, with a binary state and fixed thresholds, every learning strategy yields the same probability distribution over terminal beliefs by the martingale property which is not necessarily so with multiple states. That is, all learning strategies yield the same "overall experiment." The capacity constraint then ensures that all learning strategies that result in the same "overall experiment" have the same expected threshold-hitting times.

[12]Our solutions were based on a guess and verify approach that was possible because of the special structure of our setup. In more general setups guessing the optimal strategy seems more difficult.

moreover, allows us to isolate the role of time risk for optimal learning.

There are two promising avenues to explore in future work. The first is to explore how our results may extend to the case when the agent is neither time-risk loving nor time-risk averse. For these more general preferences, what are the qualitative features of optimal information acquisition? A second avenue to explore is to try to embed our model of information acquisition into strategic settings where there are multiple agents in order to study the implications of flexible information acquisition in games.

References

Arrow, K. J., Blackwell, D., & Girshick, M. A. (1949). Bayes and minimax solutions of sequential decision problems. *Econometrica, Journal of the Econometric Society*, (pp. 213–244).

Caplin, A., Dean, M., & Leahy, J. (2017). *Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy.* Working Paper 23652, National Bureau of Economic Research.

Che, Y.-K. & Mierendorff, K. (2019). Optimal dynamic allocation of attention. *American Economic Review*, 109(8), 2993–3029.

Chen, M. K. (2013). The effect of language on economic behavior: Evidence from savings rates, health behaviors, and retirement assets. *American Economic Review*, 103(2), 690–731.

Chesson, H. W. & Viscusi, W. K. (2003). Commonalities in time and ambiguity aversion for long-term risks. *Theory and Decision*, 54(1), 57–71.

Dasgupta, P. & Stiglitz, J. (1980). Uncertainty, industrial structure, and the speed of r&d. *The Bell Journal of Economics*, (pp. 1–28).

DeJarnette, P., Dillenberger, D., Gottlieb, D., & Ortoleva, P. (2020). Time lotteries and stochastic impatience. *Econometrica*, 88(2), 619–656.

Ely, J., Frankel, A., & Kamenica, E. (2015). Suspense and surprise. *Journal of Political Economy*, 123(1), 215–260.

Frankel, A. & Kamenica, E. (2019). Quantifying information and uncertainty. *American Economic Review*, 109(10), 3650–80.

Fudenberg, D., Strack, P., & Strzalecki, T. (2018). Speed, accuracy, and the optimal timing of choices. *American Economic Review*, 108(12), 3651–84.

Georgiadis-Harris, A. (2023). Information acquisition and the timing of actions.

Gossner, O., Steiner, J., & Stewart, C. (2021). Attention please! *Econometrica*, 89(4), 1717–1751.

Hébert, B. & Woodford, M. (2017). *Rational inattention and sequential information sampling*. Technical report, National Bureau of Economic Research.

Hébert, B. & Woodford, M. (2023). Rational inattention when decisions take time. *Journal of Economic Theory*, (pp. 105612).

Hörner, J. & Skrzypacz, A. (2017). Learning, experimentation, and information design. *Advances in Economics and Econometrics*, 1, 63–98.

Karatzas, I. & Shreve, S. E. (1998). Brownian motion. *Brownian motion and stochastic calculus*, (pp. 47–127).

Ke, T. T. & Villas-Boas, J. M. (2019). Optimal learning before choice. *Journal of Economic Theory*, 180, 383–437.

Lee, T. & Wilde, L. L. (1980). Market structure and innovation: A reformulation. *The Quarterly Journal of Economics*, 94(2), 429–436.

Liang, A., Mu, X., & Syrgkanis, V. (2019). Dynamically aggregating diverse information. *arXiv preprint arXiv:1910.07015*.

Loewenstein, G. & Prelec, D. (1992). Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2), 573–597.

Matějka, F. & McKay, A. (2014). Rational inattention to discrete choices: A new foundation for the multinomial logit model. *The American Economic Review*, 105(1), 272–298.

Mayskaya, T. (2019). Dynamic choice of information sources. unpublished manuscript.

Morris, S. & Strack, P. (2019). *The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs.* Working paper.

Moscarini, G. & Smith, L. (2001). The optimal level of experimentation. *Econometrica*, 69(6), 1629–1644.

Müller, A. (1996). Orderings of risks: A comparative study via stop-loss transforms. *Insurance: Mathematics and Economics*, 17(3), 215–222.

Nikandrova, A. & Pancs, R. (2018). Dynamic project selection. *Theoretical Economics*, 13(1), 115–143.

Onay, S. & Öncüler, A. (2007). Intertemporal choice under timing risk: An experimental approach. *Journal of Risk and Uncertainty*, 34(2), 99–121.

Protter, P. E. (2005). *Stochastic differential equations.* Springer.

Ratcliff, R. & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, 9(5), 347–356.

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

Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3), 665–690.

Steiner, J., Stewart, C., & Matějka, F. (2017). Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica*, 85(2), 521–553.

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

Zhong, W. (2022). Optimal dynamic information acquisition. *Econometrica*, 90(4), 1537–1582.

Zhong, W. & Bloedel, A. (2021). The cost of optimally acquired information.