Information Acquisit
and Time-Risk Prefer

Daniel Chen[*] 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 subject to a constraint on the rate of entropy reduc by "time risk"—the dispersion of the distribution We construct a strategy maximizing time risk (Greedy Exploitation) minimizing (Pure Accumulation). Under either strategy, belie compensated Poisson process. In the former, beli that is closer in Bregman divergence. In the latte point with the same entropy as the current belief.

1 Introduction

In this paper, we study information acquisition b binary state that may be either zero or one. The ag about the state and is satisfed once her posterio either an upper or a lower threshold. She earns a u has great fexibility in how she can learn but has her rate of learning. That is, she can choose any p constraint on the rate of entropy reduction.

[*]dtchen@princeton.edu, Department of Economics, Prince
[†]wjzhong@stanford.edu, Graduate School of Business, Sta
We thank Kayla Jiang for excellent research assistance.

Our simple model captures three important feat
flexible learning, limited resources, and thresh
appear in the contexts of research and developme
ing, user-experience testing, and others. For e
Facebook who must assess whether to introduce a n
unknown state is whether adding the feature incre
(and thus profits). The data scientist can learn a
tests. To provide incentives, her manager offers
sufficiently precisely about the state (i.e. her A
minimal level of statistical power).[1] She can flexibly design many asp
tests—e.g., she can select the subpopulation of u
ture and she can adjust the length of time a user
other choices. However, there are limits to what s
learning—e.g., her manager does not allow her to
simultaneously as doing so could be disastrous fo
she can only implement the feature for a given use

Our main contribution is to show how, in such set
ing strategy depends on her time-risk preferences. We allow for a
of preferences over threshold-hitting times bey
discounting. We say that the agent is time-risk
ity over threshold-hitting times is convex (concave).[2] We derive a
that is optimal whenever the agent is time-risk l
whenever she is time-risk averse. Critically, th
depend on the shape of the agent's utility functi

In reality, there are many reasons why individua
differ from the predominantly studied case of exp
due to external factors such as explicit discount
flow costs associated with foregone opportunities
internal factors such as present bias resulting f
a simple framework that allows for these factors w

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

[2] In this paper, we model the agent as an expected-utility
that all of our results will go through as long as the agent
times that is monotonic in the mean-preserving spread orde

derive strategies that are uniformly optimal up
utility function over threshold-hitting times.

We now briefy describe the two learning strategi
When the agent is time-risk loving, Greedy Exploitation is optimal.
this strategy, the agent myopically maximizes th
jump to a threshold. She acquires a rare but dec
induces her belief to jump to the threshold that i
By targeting the closer threshold, she can jump
her constraint on the rate of entropy reduction.
her belief experiences compensating drift in th
Eventually, her belief reaches a point that is eq
to the two thresholds. At this point, she acquire
jump to either threshold but at rates set so that
so that her belief is stationary in the absence of

Intuitively, Greedy Exploitation is optimal be
tribution of threshold-hitting times. Because t
probability of an early hitting time. However, i
liefs drift towards the farther threshold, the ju
amount of time remaining until a threshold is reached increases, the
agent makes no "progress" in the absence of a jump
of late threshold hitting times as well. We, in f
that exhaust the agent's resources (in that the c
duction is binding at all points in time), Greedy
hitting times that is maximal in the mean-preser
maximizes time risk.

When the agent is time-risk averse, she instead
this case, an optimal strategy is Pure Accumulation. Under this strategy, h
reach a threshold at a deterministic time. Her beliefs follow a comp
process that jumps in the direction of the thres
an interior belief that has the same entropy as h
of a jump, her belief experiences a compensating
Pure accumulation is a continuous-time analog of
Ely et al. (2015). The strategy is, in efect, the
Because jumps are always to beliefs with the same e

event of a jump, there is no progress: a jump does
of time until a threshold is hit. Instead, all pro
why the threshold hitting time is deterministic.
time risk. It, therefore, produces a distributio
mean-preserving spread order among all strategi
(in that the constraint on the rate of entropy re
time).

Our analysis of optimal learning through the le
plications for both information acquisition in p
model predicts that an agent who is time-risk-lo
tion, whereas an agent who is time-risk-averse s
provided the agent has access to these learning s
agent's space of available learning strategies i
generally suboptimal. Thus, when writing models
with parameterized signal structures, economist
signal structures are without loss of optimality
agents they seek to model.

2 Related Literature

Our paper contributes to a large literature on in
(1947) and Arrow et al. (1949) we study a sequent
the agent to fexibly design the signal process as
ford (2017), Hébert & Woodford (2023), Steiner e
(2023). Whereas most of these papers restrict att
nential discounting or a linear delay cost, we al
For example, Zhong (2022) assumes exponential di
loving preferences whereas Hébert & Woodford (20
assume a linear delay cost which implies time-ri
suggest that the assumed time-risk preferences d
optimal strategies identifed in these papers.[3]

[3] Hébert & Woodford (2023) allow for both discounting and
consider the time-risk neutral limit for the majority of t
study how diferent costs or constraints on information acq
which is orthogonal to the objective of our paper. Zhong (2

Pure Accumulation is a continuous-time variant
introduced by Ely et al. (2015) in a discrete-time
time horizon. In Ely et al. (2015), the suspense-
it maximizes an expected utility that is increas
Georgiadis-Harris (2023) also fnds that a strate
optimal in a setting where exogenous stopping time depends o
the learning strategy.[4] The mechanisms behind the optimality o
in both of these papers are distinct from that of
time-risk aversion.

In our analysis, the key summary statistic that
egy is the distribution of the time that the agen
This statistic, which we call the time of discovery, is an object studied in an
ture on time-risk preferences. Chesson & Viscusi
the expected discounted utility framework implies risk seeking i
over time (RSTL). DeJarnette et al. (2020) show th
of models RSTL can not be violated if there is sto
perimental evidence suggests that subjects are risk averse over time
(RATL) (Chesson & Viscusi (2003); Onay & Öncüler
modates both RSTL and RATL and shows that optimal
difer dramatically under diferent time-risk pre

The optimal learning strategies that we identif
ing strategies that have been assumed in reduced
ture. For example, Che & Mierendorf (2019), Mays
Pancs (2018) adopt a framework that restricts at
in order to study optimal stopping with endogenou
also often assumed in the literature on strategi
Hörner & Skrzypacz (2017)). We show that Poisson
foundation under time-risk preferences.[5] The Pure Accumulation
also related to classic models on the timing of in

function that does not have a threshold structure but show
nevertheless similar to Greedy Exploitation.

[V] In our paper, the stopping is determined by the endogenously
chosen learning strategy.

[5] To be clear, we do not show that Poisson learning strategies are opti
optimal strategies in our setting involve Poisson learnin

Stiglitz (1980) and Lee & Wilde (1980) (see a sur
involve a deterministic time of innovation. The
reduced-form learning process and are non-Bayes
learning strategies in these papers can emerge e
tion acquisition framework when agents have time

Our model also allows for Gaussian learning str
are often assumed in reduced-form learning mode
Smith (2001); Ke & Villas-Boas (2019); Liang et a
Also, drift-diffusion models (DDM) for binary choice problems app
Rouder (1998) and Fudenberg et al. (2018). Howeve
learning can not be justifed by optimality except
have time-risk neutral preferences provided inf

The optimality of a greedy strategy is also ther
However, the mechanisms in our papers are very dif
crucially depends on the linear-Gaussian setup w
formation sources and holds for any time preferen
and endogenous choice of information sources, bu
ences. Also related is Gossner et al. (2021) whic
with exogenous information sources and derives a
to the shortest stopping time (when the belief hi
in the sense of frst-order stochastic dominance.

We model limits on the agent's learning resourc
entropy reduction. That is, the rate of uni-resource
formly posterior separable (UPS) function. The rational inatt
typically models information costs or constrain
Matějka & McKay (2014); Steiner et al. (2017); Ca
tions for the UPS formulation can be found in Fran
et al. (2017); Zhong & Bloedel (2021); Morris & St
information constraint ensures that the expecte
for all exhaustive strategies, which is crucial for us to
mation acquisition. By Theorem 3 in Zhong (2022)
both necessary and sufcient for the expected lea
exhaustive strategies.

This section presents a simple model of an agent w
an unknown state. The unknown state is $\omega \in \{0,1\}$. At $t = 0$, the agent
believes that $\omega = 1$ with probability $\mu \in (0,1)$. She receives a unit payof
posterior that $\omega = 1$, μ_t, reaches either an upper threshold $\bar{\mu} \in (\mu, 1)$ or a lower thresh
threshold $\underline{\mu} \in (0, \mu)$. However, she is impatient and her utili
$\rho : \mathbb{R}_+ \to \mathbb{R}$ of the threshold-hitting time. We say that the
time-risk loving. When ρ is concave we say the agent is time-risk averse. A common
case where ρ is convex arises when the agent discounts t
r: $\rho(t) = e^{-rt}$. On the other hand, ρ is concave, as is the case when if the agen
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. Let $\mathcal{M}$ denote the set of belief processes $\mu =$
$\{\mu_t; t \geq 0\}$ that take values in $[0,1]$ and satisfy a stochastic diffe
(allowing for jumps) of the form

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

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 Z_t^j is a standard Brownian motion and J_t^i is a Poisson
point process with arrival rate $\lambda^i(t, \mu_t)$. 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 b
t is then N and the number of distinct Brownian Moti M.

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

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

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

[6]Our restriction to jump-diffusion belief processes is wi
class of cádlág processes such that (2) is well defined. This
Harris (2023).

a constraint on the well-known mutual informatio

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
are martingales (and thus, the drift of the secon
The normalization is convenient because, provid
optional stopping theorem[7] implies that the expected time remai
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 the be
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 be
agent can stop learning, we set $\rho(\infty) = -\infty$ to ensure that the agent ne
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 o
between optimal learning and time-risk preferen
paper. For example, we assume that the agent expe
long as (2) is satisfied) and that she earns a commo
that she ultimately hits. Though these assumpti
our model aligns well with a number of economic a
Introduction. A key property of our model that al
it accommodates general time-risk preferences.
identify optimal learning strategies that apply

[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 attitudes towards time risk that determines the qualitative properties of these strategies. ... preferences beyond exponential discounting incl... $\rho(t) = (1+\alpha t)^{-\gamma/\alpha}$ (Loewenstein & Prelec, 1992); (ii) gene... utility $\rho(t) = \phi(e^{-rt})$ where ϕ is increasing and convex (DeJarnet... linear delay cost up to a deadline T: $\rho(t) = \max\{T-t, 0\}$. Time-risk avers... preferences are less commonly assumed in the li... experimental evidence that agents may often be ti... 2003; Onay & Öncüler, 2007). As discussed at the... aversion can arise when agents have fow costs of d...

4 Optimal Learning and Time-Risk

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

4.1 Time-Risk Loving

We first consider the case when the agent is time-r... her optimal learning strategy is given below in De... it informally here.

An optimal strategy for the agent is Greedy Exploitation, illustrated... 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 eq... divergence to the two thresholds: $d_H(\overline{\mu}, \mu^*) = d_H(\underline{\mu}, \mu^*)$. Initially, the agent... either jump to the threshold that is closer in Bre... ($\overline{\mu}$) or experience compensating drift towards the other...

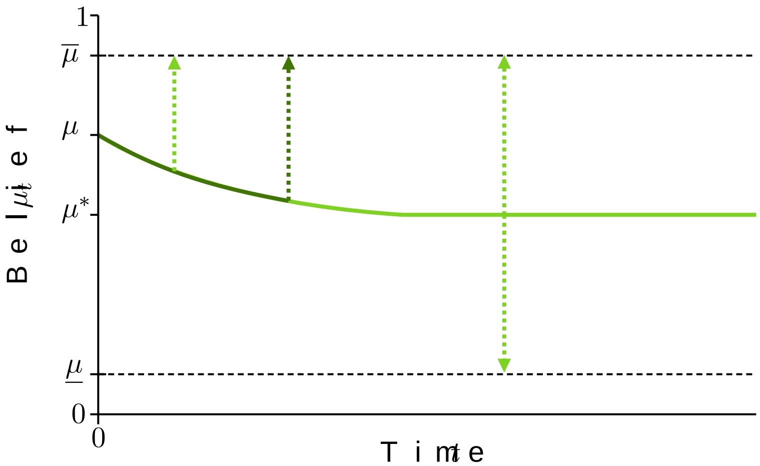

Figure 1: Greedy Exploitation

Note: In dark green, we plot one possible realization of the μ^G ... Greedy Exploitation strategy. In light green, we plot μ^G ... The dashed lines with arrows represent jumps in the beli... the case $H(\tilde{\mu}) = \tilde{\mu}^2$.

threshold, the agent greedily maximizes the "cha... in the "next instant." This is because the agent'... when she targets the closer threshold without vi... After some time, in the absence of a jump, her bel... μ^*. At this point, her beliefs may jump to either threshold. ... thresholds are such that there is no net compensa... a jump, her beliefs remain stationary.

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

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

- 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 process
cated in parentheses.

Theorem 1. If the agent is time-risk loving, then Gre

Before we sketch the proof of Theorem 1, we note
tion is uniformly optimal for all convex ρ, it must induce the riskiest t
threshold hitting times among all strategies tha
at all points in time. To make this precise, we firs

Definition 2. $\mathcal{T} = \{\tau_{\boldsymbol{\mu}} \mid \boldsymbol{\mu} \in \mathcal{M}$ such that (2) binds at all $t\}$

The Greedy Exploitation strategy produces a th
imal in the mean-preserving spread order among a
set.

Corollary 1. It holds that $\tau_{\boldsymbol{\mu}^G} \succeq_{\text{mps}} \tau$ $\tau \in \mathcal{T}$

This result hinges on our assumption that the co
in (2). Because of this assumption, all exhausti
threshold hitting time, which $-H(\mu)$ is equal to the ini

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

Step 1. Set of Basis Functions—We observe that any nonnegat
ρ can be written as a conical combination of funct

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

where $T \ge 0$. Thus, if Greedy Exploitation is optimal for ρ_T, $T \ge 0$, then it must be optimal for any nonnegative ρ. No convexity is necessary for any convex ρ, including ones that may take on negative values.

Lemma 1 If Greedy Exploitation is optimal for each ρ_T, $T \ge 0$, then it is optimal for any $\rho \in \mathcal{I}^{+2}$.

Proof. See Theorem 3.6 in Müller (1996).[9] □

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

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

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 occurs by T under Greedy Exploitation. ... 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 $V_T(\mu) = V(\mu,T)$ and $U_T(\mu) = U(\mu,T)$. Also, given a function f and two points ν, 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, d_f is the Bregman divergence.

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

[9] To apply Theorem 3.6 in Müller (1996), ... ρ ... is exhaustive and thus has the expected threshold hitting ... $-H(\mu)$ ...

$T>0$ V M (4) satisfies

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

at each $(\mu,t) \in (\underline{\mu},\overline{\mu}) \times [0,T]$ i2 $V_T(\mu)$ 2[m i (3) when $\rho = \rho^T$

G We first assert that condition (5) is equ

$$U_t(\mu) = \max_{\{\nu^i\},\{\lambda^i\},\sigma} \mathcal{A}^{\nu,\lambda,\sigma} V_t(\mu) \qquad (6)$$
$$\text{s.t. } \mathcal{A}^{\nu,\lambda,\sigma} H(\mu_t) \le 1$$

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

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

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

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

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

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

Step 4. $\{V_{T-t}(\mu_t^G)\}$ $\mathcal{M}$ — The remaining step is to verify that V satisfies th
the conditions of Lemma 2. We begin with Lemma 3 w
and outer max on the right-hand side of (5) is achi
(5) is statisfed.

Lemma 3. At each $t \in [0, \infty)$ i2 ;

$\mu \geq \mu^*$ $\mathcal{M}$

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

$\mu \leq \mu^*$ $\mathcal{M}$

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

Because $V_{T-t}(\mu_t^G) = \mathbb{E}\left[\rho_T(\tau_{\boldsymbol{\mu}^G})|\mu_t^G\right]$ and μ^G is Markov it follo
that $\{V_{T-t}(\mu_t^G)\}$ is a martingale for any $t \geq 0$. By Itô's formula, the
$\{V_{T-t}(\mu_t^G)\}$ is zero if and only if conditions 1 □ and 2 of t

Step 5. Unimprovable — The following Lemma 4 show
Greedy Exploitation can not be improved on by an
strategy.

Lemma 4. At each $(\mu, t) \in (\underline{\mu}, \overline{\mu}) \times [0, \infty)$ i ii

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

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

  At any local extremum (9) holds with equality. W extrema are necessarily local maxima simply by the left-hand side of (9) is negative. This is b hand side is always zero at a local extremum sin is the objective $d_{V_t}(\nu,\mu)/d_H(\nu,\mu)$. The left-hand side of (9) is ν because

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

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

- Case 2: $\nu \in (\mu^*, \mu)$. In this region, (following the same 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 ine

As before, to determine whether a local extremu sufces to check how the left-hand side changes in ν. In this case left-hand side is increasing. This can be seen

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

where we have used the fact that the denominator in this region, any local extremum must be a lo $\nu \in (\mu^*,\mu)$ can achieve the max in (7).

- Case 3: $\nu \in [\underline{\mu},\mu^*]$. Following analogous steps to those fnd 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 b $d_{V_t}(\overline{\mu},\mu)/d_H(\overline{\mu},\mu)$. Thus, there can not be a point $\nu \in [\underline{\mu},\mu^*]$ that achieves a higher va $d_{V_t}(\overline{\mu},\mu)/d_H(\overline{\mu},\mu)$, since if there was, at that point $d_{V_t}(\nu,\mu)/d_H(\nu,\mu)$ would not 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

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

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 with $\nu = \overline{\mu}$, 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

(16), (17), and Lemma 3. The third line uses that $U_t(\nu)$ is decreasing

for $\nu \in [\underline{\mu}, \mu^*]$ as noted in Step 2. The last line uses Lem

□

Step 6. Unimprovable by Diffusion—The following Lemma 5 show

Greedy Exploitation can not be improved on by dif

$(\mu, t) \in (\underline{\mu}, \overline{\mu}) \times [0, \infty)$ i ii

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

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 $\mu \in (\underline{\mu}, \mu^*)$. We make ... the case when $\mu = \mu^*$ follows from continuity.

□

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

□

4.2 Time-Risk Averse

When the agent is time-risk averse, Pure Accumulation ... tion, illustrated graphically below in Figure 2.

As discussed in Section 2, the Pure Accumulation ... the suspense-maximal strategy from Ely et al. (2015). Under this s... belief either jumps in the direction of the farth... sating drift. When her belief jumps, it jumps to ... her current belief so that all progress is made th...

Definition 3. The Pure Accumulation strategy is defined as follows. ... $\mu^H : [0,1] \setminus \{\mu^*\} \to [0,1]$ denote the function that maps each belief $\hat{\mu}$ to the belief $\mu^H(\hat{\mu}) \neq \hat{\mu}$ such that $H(\mu^H(\hat{\mu})) = H(\hat{\mu})$. Under Pure accumulation, the age...

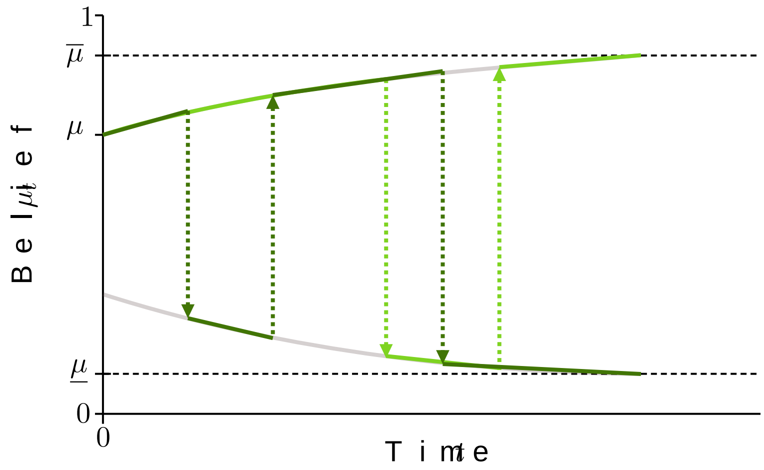

Figure 2: Pure Accumulation

Note: The dark green curve represents μ_t^P. The possible belie
segments represent jumps. The light green curve represe
of μ^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 Pur

Proof. Under Pure Accumulation, the agent is guarant
determin is $t = -H(\mu)$. time □

Because Pure Accumulation entails no time risk
ing result.

Corollary 2. It holds that $t \succeq_{\text{mps}} \tau_{\mu^A}$ $\tau \in \mathcal{T}$

$\mathcal{MM}$ $\mathcal{M}$

In this paper, we have studied the relationship be
timal information acquisition. We have shown tha

loving agent is Greedy Exploitation. This strate over threshold hitting times among all exhaustiv optimal strategy for a time-risk averse agent is produces a deterministic threshold hitting time of these strategies are uniformly optimal up to t ity function, provided the agent is impatient. T inconsistency. In practice, agents may have tim well-studied case of exponential discounting. O how these agents may seek to acquire information economists may consider using when modeling thes

In order to illustrate the connection between as sharply as possible we have made a number of spe tions of binary states, fxed stopping thresholds speed are critical because they ensure that all ex pected threshold-hitting times.[11] The key reason why Greedy Exploit Accumulation are optimal is because they respect mal threshold-hitting times in the mean-preserv strategies. This allows us to emphasize its key features that de their optimality. The assumption that payofs dep and not on which threshold is hit allows us to deri critical for the economic insights.

Indeed, we anticipate that many of the qualitat tion or Pure Accumulation will persist under oth ple, with other costs of learning, multiple stat thresholds (there are already examples in the li in Section 2). It is certainly possible to exten accommodate these more general environments tho be rather.[12] The advantage of our special setup is that it strategies that are explicit and uniformly opti

[11] Specifcally, with a binary state and fxed thresholds, ev probability distribution over terminal beliefs by the mar with multiple states. That is, all learning strategies yie constraint then ensures that all learning strategies that the same expected threshold-hitting times.

[12] Our solutions were based on a guess and verify approach th structure of our setup. In more general setups guessing the

moreover, allows us to isolate the role of time ri

There are two promising avenues to explore in fu how our results may extend to the case when the age time-risk averse. For these more general prefere of optimal information acquisition? A second av our model of information acquisition into strate agents in order to study the implications of flexil

References

Arrow, K. J., Blackwell, D., & Girshick, M. A. (19 of sequential decision problems. Econometrica, Journal of the Eco (pp. 213–244).

Caplin, A., Dean, M., & Leahy, J. (2017). Rationally Inattentive Beh acterizing and Generalizing Shannon Entropy. Working Paper 23652, Na Bureau of Economic Research.

Che, Y.-K. & Mierendorff, K. (2019). Optimal dynam American Economic Review, 109(8), 2993–3029.

Chen, M. K. (2013). The effect of language on econ savings rates, health behaviors, and retirement American Economic Review, 103(2), 690–731.

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

Dasgupta, P. & Stiglitz, J. (1980). Uncertainty of r&d. The Bell Journal of Economics (pp. 1–28).

DeJarnette, P., Dillenberger, D., Gottlieb, D., and stochastic impatience. Econometrica, 88(2), 619–656.

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

Frankel, A. & Kamenica, E. (2019). Quantifying in
ican Economic Review, 109(10), 3650–80.

Fudenberg, D., Strack, P., & Strzalecki, T. (2018
timing of choices. American Economic Review, 108(12), 3651–84.

Georgiadis-Harris, A. (2023). Information acqu

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

Hébert, B. & Woodford, M. (2017). Rational inattention and seque
sampling. Technical report, National Bureau of Econ

Hébert, B. & Woodford, M. (2023). Rational inatt
Journal of Economic Theory, (pp. 105612).

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

Karatzas, I. & Shreve, S. E. (1998). Brownian motion. In Brownian motion and stoc
tic calculus (pp. 47–127).

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

Lee, T. & Wilde, L. L. (1980). Market structure a
The Quarterly Journal of Economics, 94(2), 429–436.

Liang, A., Mu, X., & Syrgkanis, V. (2019). Dynam
mation. arXiv preprint arXiv:1910.07015.

Loewenstein, G. & Prelec, D. (1992). Anomalies i
and an interpretation. The Quarterly Journal of Economics, 107(2), 573–597.

Matějka, F. & McKay, A. (2014). Rational inatten
foundation for the multinomial logit model. The American Economic Re
105(1), 272–298.

Mayskaya, T. (2019). Dynamic choice of inform
manuscript.

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

Moscarini, G. & Smith, L. (2001). The optimal lev
Econometrica, 69(6), 1629–1644.

Müller, A. (1996). Orderings of risks: A compara
Insurance: Mathematics and Economics, 17(3), 215–222.

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

Onay, S. & Öncüler, A. (2007). Intertemporal cho
mental approach. Journal of Risk and Uncertainty, 34(2), 99–121.

Protter, P. E. (2005). Stochastic diferential equations. Springer.

Ratclif, R. & Rouder, J. N. (1998). Modeling respo
Psychological Science, 9(5), 347–356.

Reinganum, J. F. (1989). The timing of innovatio
difusion. Handbook of industrial organization, 1, 849–908.

Sims, C. A. (2003). Implications of rational ina
Journal of Monetary Economics, 50(3), 665–690.

Steiner, J., Stewart, C., & Matějka, F. (2017). Ra
and delay in decision-making. Econometrica, 85(2), 521–553.

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

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

Zhong, W. & Bloedel, A. (2021). The cost of optima