Information Acquisition and Time-Ris

ByDaniel Chen and Weijie Zhong

An agent acquires information dynamically unt
a binary state reaches an upper or lower thresho
any signal process subject to a constraint on t
reduction. Strategies are ordered by “time risk
the distribution of threshold-hitting times. W
maximizing time risk (Greedy Exploitation) an
it (Pure Accumulation). Under either strategy
compensated Poisson process. In the former, be
threshold closer in Bregman divergence. In the
to the unique point with the same entropy as the

JEL: D80, D81, D83
Keywords: Information acquisition, stopping t
erence, rational inattention

Introduction

In this paper, we study information acquisition by
unknown binary state that may be either zero or one
reasonably certain about the state and is satisfed
the state is one reaches either an upper or a lower t
payof when this happens. She has great fexibility i
a resource constraint that limits her rate of learni
posterior belief process subject to a constraint or

Our simple model captures three important features
fexible learning, limited resources, and threshol
often appear in the contexts of research and develo
marketing, user-experience testing, and others.
scientist at Facebook who must assess whether to in
platform. The unknown state is whether adding the fe
user engagement (and thus profts). The data scienti
by conducting A/B tests. To provide incentives, her
when she has learned su ciently precisely about th
identify the state at some minimal level of certainty.[1] She can ... is
design many aspects of the A/B tests—e.g., she can s

Chen: dtchen@princeton.edu, Department of Economics, Princeton
jie.zhong@stanford.edu, Graduate School of Business, Stanford Univ
excellent research assistance.

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

users that she gives access to the feature and she ca
user has access to the feature among other choices.
what she can do that constrain her speed of learning-
allow her to implement the feature for all users sim
be disastrous for profts if the feature is disliked
feature for a given user for a short amount of time,

Our main contribution is to show how, in such sett
learning strategy depends on her time-risk preference. There, we allow for
a rich set of preferences over threshold-hitting t
of exponential discounting. We say that the agent
whenever her utility over threshold-hitting times[2] We derive
a learning strategy that is optimal whenever the ag
strategy that is optimal whenever she is time-risk a
of these strategies does not depend on the shape of
beyond its convexity or concavity.

In reality, there are many reasons why individuals
that difer from the predominantly studied case of e
may be due to external factors such as explicit dis
deadlines, and fow costs associated with foregone o
may also be due to internal factors such as present b
discounting. We provide a simple framework that al
studying optimal learning and derive strategies th
the convexity or concavity of the utility function

We now briefy describe the two learning strategie
them. When the agent is time-risk loving, Greedy Exploitation is
optimal. Under this strategy, the agent myopically
rate that her beliefs jump to a threshold. She acqui
that, upon arrival, induces her belief to jump to th
the Bregman divergence (under the entropy function
threshold, she can jump at a faster rate without vio
rate of entropy reduction. In the absence of a signa
compensating drift in the direction of the farthe
belief reaches a point that is equidistant in the B
thresholds. At this point, she acquires signals su
either threshold but at rates set so that there is no
her belief is stationary in the absence of a jump.

Intuitively, Greedy Exploitation is optimal beca
distribution of threshold-hitting times. Because
a high probability of an early hitting time. Howeve
since beliefs drift towards the farther threshold,

[2] In this paper, we model the agent as an expected-utility maximizer.
all of our results will go through as long as the agent has a preference
is monotonic in the mean-preserving spread order.

the expected amount of time remaining until a thres increases
In this sense, the agent makes no "progress" in the
there is a high probability of late threshold hitti
show that among all strategies that exhaust the age
constraint on the rate of entropy reduction is bin
Greedy Exploitation yields a distribution of hitti
mean-preserving spread order. In this sense, it max

When the agent is time-risk averse, she instead se
In this case, an optimal strategy is Pure Accumulation. Under this strategy, he
beliefs reach a threshold at a deterministic time. Her beliefs follow a comp
Poisson process that jumps in the direction of the t
but to an interior belief that has the same entropy
the absence of a jump, her belief experiences a com
closer threshold. Pure accumulation is a continuou
maximal" policy in Ely, Frankel and Kamenica (2015)
the "opposite" of Greedy Exploitation. Because jum
the same entropy as the current belief, in the event o
a jump does not reduce the expected amount of time u
Instead, all progress is made through drift, which
time is deterministic. Thus, not only is Pure Accumulation
produces a distribution of hitting times that is mi
spread order among all strategies that exhaust the
constraint on the rate of entropy reduction is bind

Our analysis of optimal learning through the lens
implications for both information acquisition in p
Our model predicts that an agent who is time-risk-l
ploitation, whereas an agent who is time-risk-aver
lation, provided the agent has access to these learn
ever the agent's space of available learning strate
signals are generally suboptimal. Thus, when writi
information with parameterized signal structures,
whether these signal structures are without loss of
preferences of the agents they seek to model.

I. Related Literature

Our paper contributes to a large literature on inf
Wald (1947) and Arrow, Blackwell and Girshick (194
sampling problem but allow the agent to fexibly de
in Zhong (2022), Hébert and Woodford (2017), Hébe
Steiner, Stewart and Matějka (2017), and Georgiad
most of these papers restrict attention to the stan
counting or a linear delay cost, we allow for more g
example, Zhong (2022) assumes exponential discoun

loving preferences whereas Hébert and Woodford (20
ford (2023) assume a linear delay cost which implies
Our results suggest that the assumed time-risk pref
features of the optimal strategies identifed in the[3]

Pure Accumulation is a continuous-time variant of
egy introduced by Ely, Frankel and Kamenica (2015)
with a deterministic and fnite time horizon. In Ely,
the suspense-maximal strategy is optimal in that it
that is increasing in the variances of future belie
fnds that a strategy similar to Pure Accumulation i
the stopping time is exogenous[4] and does not depend on the learning
The mechanisms behind the optimality of Pure Accum
papers are distinct from that of this paper where i
aversion.

In our analysis, the key summary statistic that d
a strategy is the distribution of the time that the
a threshold. This time is the hitting time, which is an object studi
in an emerging literature on time-risk preferences
and Chen (2013) show that the expected discounted u
preferences is that risk seeking over time lotteries (RSTL). DeJarnette et al
(2020) show that within a broad class of models RST
there is stochastic impatience. However, experim
subjects are risk averse over time lotteries (RATL) (Chesson and Viscus
(2003); Onay and Öncüler (2007)). Our model accomm
RATL and shows that optimal information acquisiti
under diferent time-risk preferences.

The optimal learning strategies that we identify
learning strategies that have been assumed in redu
in the literature. For example, Che and Mierendorf
and Nikandrova and Pancs (2018) adopt a framework t
Poisson signal processes in order to study optimal
formation. Poisson signals are also often assumed
experimentation (see a survey by Hörner and Skrzypa
Poisson learning has an optimization foundation u
ence.[5] The Pure Accumulation strategy is also related

[3]Hébert and Woodford (2023) allow for both discounting and a linear
consider the time-risk neutral limit for the majority of their analys
how diferent costs or constraints on information acquistion afect op
orthogonal to the objective of our paper. Zhong (2022) assumes a belie
does not have a threshold structure but shows that the optimal learnin
to Greedy Exploitation.

[4]In our paper, the stopping is determined by the exogenously threshold
chosen learning strategy.

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

timing of innovation introduced by Dasgupta and St
Wilde (1980) (see a survey by Reinganum (1989)) whi
time of innovation. The models in these papers assu
process and are non-Bayesian. However, we show tha
in these papers can emerge endogenously in a Bayesi
framework when agents have time risk-averse prefer

Our model also allows for Gaussian learning strat
cesses are often assumed in reduced-form learning mo
and Smith (2001); Ke and Villas-Boas (2019); Liang
Morris and Strack (2019)). Drift-diffusion (DDM) models in binary choice
problems appear in Ratclif and Rouder (1998) and Fud
lecki (2018). However, our results imply that Gauss
fed by optimality except in the knife-edge case when
preferences provided information can be acquired f

The optimality of a greedy strategy is also the mai
Syrgkanis (2019). However, the mechanisms in our
Liang, Mu and Syrgkanis (2019)'s result crucially d
setup with exogenously given Gaussian information s
preferences. Our result allows for a fexible and en
tion sources, but crucially depends on time prefere
Steiner and Stewart (2021) which studies a model wi
nous information sources and derives a greedy learn
shortest stopping time (when the belief hits an exo
the sense of frst-order stochastic dominance.

We model limits on the agent's learning resources
of entropy reduction. That is, the rate of resource
uniformly posterior-separable (UPS) function. The rational inatten
also typically models information costs or constra
(2003); Matějka and McKay (2014); Steiner, Stewart
Dean and Leahy (2017)). Microfoundations for the UP
in Frankel and Kamenica (2019); Caplin, Dean and Le
Bloedel (2021); Morris and Strack (2019). In our p
constraint ensures that the expected threshold hi
exhaustive strategies, which [illegible]
acquisition. By Theorem 3 in Zhong (2022), a UPS in
both necessary and su cient for the expected learn
all exhaustive strategies.

II. Model

This section presents a simple model of an agent wh
about an unknown state ω. The unknown state takes values in $\{0,1\}$. At $t=0$,
the agent believes that $\omega = 1$ with probability $\mu \in (0,1)$. She receives a unit
payof when her posterior that $\omega = 1$ (μ_t) reaches either an upper th

$\overline{\mu} \in (\mu, 1)$ or a lower threshold $\underline{\mu} \in (0, \mu)$. However, she is impatient an
utility is a decreasing function $\rho : \mathbb{R}_+ \to \mathbb{R}$ of the threshold-hitting time
ρ is convex, we say that the agent is risk-loving. When ρ is concave we
say that the agent is risk-averse. A common case where ρ is convex arises when
the agent discounts time at rate r: $\rho(t) = e^{-rt}$. On the other
hand, a case where ρ is concave is if the agent's flow cost of delay $c(\cdot)$ increases in time: $\rho(t) = 1 - \int_0^t c(s)\,ds$.

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

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

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 each J_t^i is a
Poisson point process that ticks at rate $\lambda^i(t,\mu_t)$: when J_t^i ticks, 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 $\boldsymbol{\mu} \in \mathcal{M}$ such
that for all t

$$(2)\qquad \mathcal{A}H(\mu_t) := \lim_{s \to 0} \mathbb{E}\left[\frac{H(\mu_{t+s}) - H(\mu_t)}{s} \middle| \mathcal{F}_t \right] \leq I,$$

where $\{\mathcal{F}_t\}$ is the natural filtration of $\boldsymbol{\mu}$, H is a strictly convex C^2 functio
defined on $[0,1]$, and $I \geq 0$ is a constant.[6] 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) corresponds
to a constraint on the well-known mutual informati
generality, $H(\overline{\mu}) = H(\underline{\mu}) = 0$ and $I = 1$ to simplify notations. T
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 a
beliefs are martingales (and thus, the drift of the

[6] Our restriction to jump-difusion belief processes is without loss o
of cádlág processes such that (2) is well defned. This follows from Pr
(2023).

is zero).

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

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

We extend the domain of ρ such that $\rho(\infty) = -\infty$ to ensure that the
agent never stops learning (i.e., she may choose $\tau_{\boldsymbol{\mu}} = \infty$ with belief p
positive probability). The agent 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 orde
between optimal learning and time-risk preferences
paper. For example, we assume that the agent experie
(as long as (2) is satisfed) and that she earns a comm
threshold that she ultimately hits. Though these as
believe that our model aligns well with a number of e
cussed in the Introduction. A key property of our mo
insights is that it accommodates general time-risk
contribution is to identify optimal learning strat
time preferences and to highlight that it is the age
that determines the qualitative properties of thes
of time-risk loving preferences beyond exponential
bolic discounting $\rho(t) = (1+\alpha t)^{-r/\alpha}$ where r (Loewenstein and Prelec, 199
generalized expected discounting $\rho(t) = \phi(e^{-rt})$ where ϕ is strictly increasing and
convex (DeJarnette et al., 2020); and linear delay
$\rho^T(t) = \max\{T - t, 0\}$. Time-risk averse preferences are less
in the literature though there is growing experimen
often be time-risk averse (Chesson and Viscusi, 200
As discussed at the start of this section, time-risk
have fow costs of delay that increase with time.

III. Optimal Learning and Time-Risk Preference

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

A. Time-Risk Loving

We frst consider the case when the agent is time-ris
of her optimal learning strategy is given below in D

describing it informally here.

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

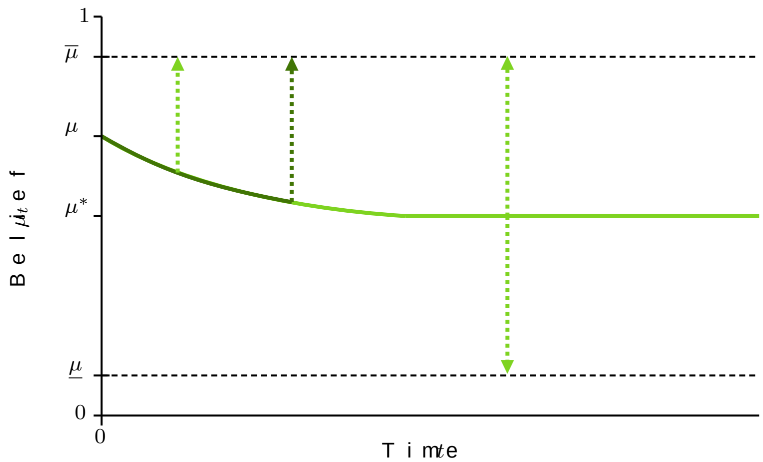

Figure 1. Greedy Exploitation

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

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 e Bregman divergence to the two thresholds: $d_H(\bar{\mu}, \mu^*) = d_H(\underline{\mu}, \mu^*)$. Initially, the agent's beliefs either jump to the threshold that i (in this case $\bar{\mu}$) or experience compensating drift towards t targeting the closer threshold, the agent greedily her beliefs reach a threshold in the "next instant. beliefs can jump at a faster rate when she targets t violating her resource constraint (2). After some t her beliefs eventually drift into μ^*. At this point, her beliefs may jum threshold. The jump rates to the respective thresh no net compensating drift, and so, in the absence o stationary.

DEFINITION 1: The Greedy Exploitation strategy μ^G is defined as follows. Le $\mu^* \in (0,1)$ be the unique belief such that $d_H(\bar{\mu}, \mu^*) = d_H(\underline{\mu}, \mu^*)$, 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 processe
indicated in parentheses.

THEOREM 1: If the agent is time-risk loving, then Gree
timal.

Before we sketch the proof of Theorem 1, we note that
tion is uniformly optimal for all μ, it must induce the riskiest di
of threshold hitting times among all strategies tha
satisfed at any point before stopping. To make this
following defnition.

DEFINITION 2: $\mathcal{T} = \{\tau_{\boldsymbol{\mu}} \mid \boldsymbol{\mu} \in \mathcal{M}$ such that (2) holds for almost all $t < \tau_{\boldsymbol{\mu}}$, a.s.$\}$.

The following Lemma 1 implies that the the expected
equal to the initial entropy for any exhaustive stra

LEMMA 1: For each $\boldsymbol{\mu} \in \mathcal{M}$, $\mathbb{E}[\tau_{\boldsymbol{\mu}}] \geq -H(\mu)$ with equality if and only if $\tau_{\boldsymbol{\mu}} \in \mathcal{T}$.

PROOF:
Recall that we have normalized $H(\underline{\mu}) = H(\overline{\mu}) = 0$. Therefore,

$$\begin{aligned}
-H(\mu) &= \mathbb{E}[H(\mu_{\tau_{\boldsymbol{\mu}}}) - H(\mu)] \\
&= \mathbb{E}\left[\int_0^\infty \mathbf{1}_{\{t<\tau_{\boldsymbol{\mu}}\}}\mathcal{A}H(\mu_t)\,dt\right] \\
&\leq \mathbb{E}\left[\int_0^\infty \mathbf{1}_{\{t<\tau_{\boldsymbol{\mu}}\}}\,dt\right] \\
&= \mathbb{E}[\tau_{\boldsymbol{\mu}}]
\end{aligned}$$

where the second equality follows from Dynkin's for (2005)) which implies $H(\mu_t) - \int_0^t \mathcal{A}H(\rho_s)\,ds$ is a martingale and the opti stopping theorem (see Theorem 3.22 of Karatzas and equality follows from the constraint (2). The inequ if $\tau_\mu \in \mathcal{T}$.

Theorem 1 and Lemma 1 together imply that the Greedy produces a threshold-hitting time that is maximal i order among all threshold-hitting times in this set

COROLLARY 1: It holds that $\tau_\mu \succeq_{\text{mps}} \tau$ for each $\tau \in \mathcal{T}$.

PROOF OF THEOREM 1:

The proof proceeds in seven steps.

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

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

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

LEMMA 2: If Greedy Exploitation is optimal for (3) for each ρ_T where $T \geq 0$, then it is optimal for any convex ρ.

PROOF OF LEMMA 2:

See Theorem 3.6 in[7] Müller (1996).

Step 2. Candidate Value Function and Some Notation—Let $V(\mu, T) = \mathbb{E}[\rho_T(\tau_\mu^G) \mid \mu_0^G = \mu]$ denote the value function for $\rho = \rho_T$ under Greedy Exploitation, 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 $\partial V(\mu,T)/\partial T = U(\mu,T)$ for every $\mu \in (\underline{\mu},\overline{\mu})$ where

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

is the probability that T a jump occurs under Greedy Exploitation by time 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 $\mu \in (\underline{\mu},\mu^*]$.

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

To ease the exposition, we adopt the following no $V_T(\mu) =$ $V(\mu,T)$ and $U_T(\mu) = U(\mu,T)$. Also, given a function f and two beliefs μ, we let

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

whenever $f'(\mu)$ is well defined. Notice that when f is convex, d_f is the Bregman divergence.

Step 3. Verification Theorem—To confirm the optimality of Greedy E we use the following Lemma 3 which states that it suffices that V satisfies the Hamilton-Jacobi-Bellman (HJB) equation (5).

LEMMA 3: Given $T>0$, 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 3:

We first assert that condition (5) is equivalent to

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

where $\mathcal{A}^{\nu,\lambda,\sigma}$ is the operator defined for any 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 compe process (1). Because $\mathcal{A}^{\nu,\lambda,\sigma}$ is additively separable, it suces t jump point or volatility to achieve the max in (6). is chosen to maximize the "bang-for-the-buck"—tha V to the d H. Therefore, (5) and (6) must be equivalen

Next, suppose that (5) is satisfied. Consider an ar

$\{\nu^i\}$, $\{\lambda^i\}$, $\{\sigma^j\}$ with induced first hitting time τ. We then have

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

where the first equality uses Itô's formula for jump
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 martingales[8], the third equality uses
$\partial V/\partial T = U$ as noted in Step 2, the first inequality follow
inequality follows from the definition of V.

Step 4: $\{V_{T-t}(\mu_t^G)\}$ is a Martingale — The remaining step is to verify that V sat-
isfies the conditions of Lemma 3. We begin with Lemm
the inner and outer max on the right-hand side of (5
Exploitation then (5) is satisfied.

LEMMA 4: 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 4:

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 follows that $\{V_{T-t}(\mu_t^G)\}$
is a martingale for $T \geq 0$. By Itô's formula $\{V_{T-t}(\mu_t^G)\}$ has drift of
zero if and only if conditions 1 and 2 of the lemma ar

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

Step 5. Unimprovable by Pure Learning—The following Lemma 5 shows th
Greedy Exploitation can not be improved on by any al
strategy.

LEMMA 5: 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 5:

We will prove the lemma when $\mu > \mu^*$. The proof for $\mu \le \mu^*$ is analogous.
By Lemma 4, it suffices to show that $\overline{\mu}$ attains the max in (7). We spl
proof into three cases.

- Case 1: $\nu \geq \mu$. We will show that $\nu = \overline{\mu}$ is the global maximizer o
  $d_{V_t}(\nu, \mu)/d_H(\nu, \mu)$ 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 $\overline{\mu}$ is actually the global maximum in this region.

  At any local extremum (9) holds with equality. We
  local extrema are necessarily local maxima simply
  tive of the left-hand side of (9) is negative. Thi
  the right-hand side is always zero at a local extr
  side expression is just 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 4 and that $U_t(\cdot)$ is increasing in ν as noted in Step 2.

- Case 2: $\nu \in (\mu^*, \mu)$. In this region, (following the same st
  it is easy to show that $d_{V_t}(\nu,\mu)/d_H(\nu,\mu)$ is nondecreasing if

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

This is the same condition as (9) except the inequ

As before, to determine whether a local extremum i
it suffices to check how the left-hand side changes as ν increases. In this cas
the left-hand side is increasing. This can be see

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

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

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

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

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

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

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), a (16), (17), and Lemma 4. The third line uses the fa $U_t(\nu)$ is decreasing for $\nu \in [\underline{\mu},\mu^*]$ as noted in Step 2. The last line uses Lemma

Step 6. Unimprovable by Diffusion.—The following Lemma 6 shows that Greedy Exploitation can not be improved on by d

LEMMA 6: 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 $U_t'(\mu) > 0$ whenever $\mu \in (\mu^*, \overline{\mu})$. 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 $\mu \in (\underline{\mu}, \mu^*)$. The weak inequality when $\mu = \mu^*$ follows from continuity.

Step 7. Putting it all together.—Lemmas 5 and 6 imply that (5) is sat
V for any $T \geq 0$. Lemma 3 then implies that Greedy Exploita
any discount function ρ_T, $T \geq 0$. Lemma 2 then implies optima
for all ρ convex. The proof of Theorem 1 is complete.

B. Time-Risk Averse

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

As discussed in Section I, the Pure Accumulation s
the suspense-maximal form Ely, Frankel and Kamenica (2015)
strategy, the agent's belief either jumps in the di
or experiences compensating drift. When her belie
with the same entropy as her current belief so that a
drift.

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

PROOF:

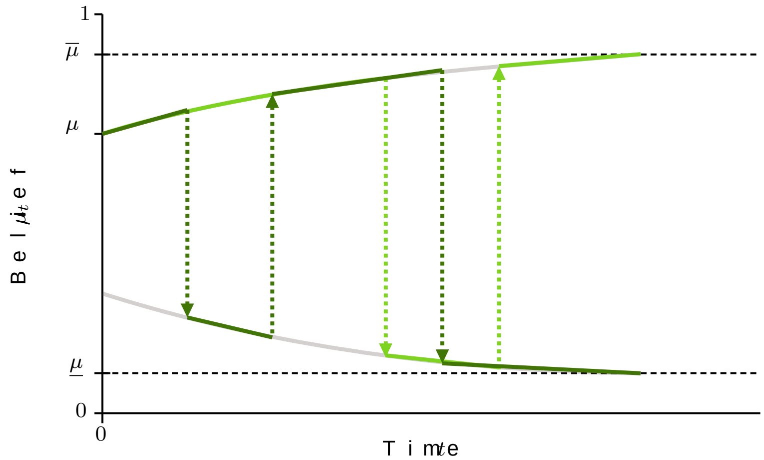

Figure 2. Pure Accumulation

Note: The dark green curve represents μ^P. The possible sample paths of represent jumps. The light green curve represents μ^P. another possibl figure is computed for the case $H(\tilde{\mu}) = \tilde{\mu}^2$.

Under Pure Accumulation, the agent is guaranteed deterministic time. Because Pure Accumulation is ex immediately from Lemma 1.

Because Pure Accumulation entails no time risk, we i ing result.

COROLLARY 2: It holds that $\tau \succeq_{\mathrm{mps}} \tau_{\mu}$ for each $\tau \in \mathcal{T}$.

IV. Concluding Discussion

In this paper, we have studied the relationship be and optimal information acquisition. We have shown for a time-risk loving agent is Greedy Exploitatio riskiest distribution over threshold hitting times On the other hand, an optimal strategy for a time- Accumulation. This strategy produces a determini and thus entails no time risk. Both of these strate up to the convexity or concavity of the utility fun impatient. Thus, they are immune to dynamic inconsi may have time preferences that difer from the well- discounting. Our analysis ofers some insight into acquire information and the kinds of signal struct using when modeling these agents.

In order to illustrate the connection between lear as sharply as possible we have made a number of spec sumptions of binary states, fxed stopping threshol on learning speed are critical because they ensure have the same expected threshold hitting times.[9] The key reason why Greedy Ex- ploitation and Pure Accumulation are optimal is bec the maximal and minimal threshold-hitting times in order among all exhaustive strategies. This allows risk preferences that determines their optimality. T depend only on threshold hitting times and not on wh us to derive closed-form solutions but is not criti

Indeed, we anticipate that many of the qualitativ ploitation or Pure Accumulation will persist under example, with other costs of learning, multiple st stopping thresholds (there are already examples in as reviewed in Section I). It is certainly possibl mal learning to accommodate these more general env charaterization.[10] The advantage of our special setup possible to solve for optimal strategies that are ex large classes of payofs and, moreover, allows us to optimal learning.

There are two promising avenues to explore in futu explore how our results may extend to the case when risk loving nor time-risk averse. For these more g the qualitative features of optimal information ac explore is to try to embed our model of information settings where there are multiple agents in order fexible information acquisition in games.

REFERENCES

Arrow, Kenneth J, David Blackwell, and M. A. Girsh 1949. "Bayes and minimax solutions of sequential decision pro Econometrica, Journal of the Econometric Society, 213–244.

Caplin, Andrew, Mark Dean, and John Leahy. 2017. "Rationally Inatte tive Behavior: Characterizing and Generalizing S Bureau of Economic Research Working Paper 23652.

[9] Specifcally, with a binary state and fxed thresholds, every learnin ability distribution over terminal beliefs by the martingale proper multiple states. That is, all learning strategies yield the same "ove straint then ensures that all learning strategies that result in the same expected threshold-hitting times.

[10] Our solutions were based on a guess and verify approach that was pos structure of our setup. In more general setups guessing the optimal st

Chen, M Keith. 2013. "The effect of language on economic dence from savings rates, health behaviors, and r American Economic Review, 103(2): 690–731.

Chesson, Harrell W, and W Kip Viscusi. 2003. "Commonalities in time ambiguity aversion for long-term risks." Theory and Decision, 54(1): 57–71.

Che, Yeon-Koo, and Konrad Mierendorff. 2019. "Optimal Dynamic All cation of Attention." American Economic Review, 109(8): 2993–3029.

Dasgupta, Partha, and Joseph Stiglitz. 1980. "Uncertainty, industri ture, and the speed of R&D." The Bell Journal of Economics, 1–28.

DeJarnette, Patrick, David Dillenberger, Daniel G Ortoleva. 2020. "Time lotteries and stochastic impat Econometrica, 88(2): 619–656.

Ely, Jeffrey, Alexander Frankel, and Emir Kamenica. 2015. "Suspense and surprise." Journal of Political Economy, 123(1): 215–260.

Frankel, Alexander, and Emir Kamenica. 2019. "Quantifying informa and uncertainty." American Economic Review, 109(10): 3650–80.

Fudenberg, Drew, Philipp Strack, and Tomasz Strzal accuracy, and the optimal timing of choices." American Economic Review, 108(12): 3651–84.

Georgiadis-Harris, Alkis. 2023. "Preparing to Act." working pap

Gossner, Olivier, Jakub Steiner, and Colin Stewart. 2021. "Attention please." Econometrica, 89(4): 1717–1751.

Hébert, Benjamin, and Michael Woodford. 2017. "Rational inattentio sequential information sampling." National Burea

Hébert, Benjamin, and Michael Woodford. 2023. "Rational inattent when decisions take time." Journal of Economic Theory, 105612.

Hörner, Johannes, and Andrzej Skrzypacz. 2017. "Learning, Experime tation, and Information Design." Advances in Economics and Econom 1: 63–98.

Karatzas, Ioannis, and Steven E. Shreve. 1998. "Brownian motion." Brownian motion and stochastic calculus, 47–127.

Ke, T Tony, and J Miguel Villas-Boas. 2019. "Optimal learning be choices." Journal of Economic Theory, 180: 383–437.

Lee, Tom, and Louis L Wilde. 1980. "Market structure and innov reformulation." The Quarterly Journal of Economics, 94(2): 429–436.

Liang, Annie, Xiaosheng Mu, and Vasilis Syrgkanis. 2019. "Dynamically aggregating diverse information." arXiv:1910.07015.

Loewenstein, George, and Drazen Prelec. 1992. "Anomalies in intertemporal choice: Evidence and an interpretation." The Quarterly Journal of Economics, 107(2): 573–597.

Matějka, Filip, and Alisdair McKay. 2015. "Rational inattention to discrete choices: A new foundation for the multinomial logit model." American Economic Review, 105(1): 272–298.

Mayskaya, Tatiana. 2019. "Dynamic Choice of Information Sources." Working paper.

Morris, Stephen, and Philipp Strack. 2019. "The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs." Working Paper.

Moscarini, Giuseppe, and Lones Smith. 2001. "The optimal level of experimentation." Econometrica, 69(6): 1629–1644.

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

Nikandrova, Arina, and Romans Pancs. 2018. "Dynamic project selection." Theoretical Economics, 13(1): 115–143.

Onay, Selçuk, and Ayse Öncüler. 2007. "Intertemporal choice under timing risk: An experimental approach." Journal of Risk and Uncertainty, 34(2): 99–121.

Protter, Philip. 2005. Stochastic differential equations. Springer.

Ratcliff, Roger, and Jeffrey N. Rouder. 1998. "Modeling response times for two-choice decisions." Psychological Science, 9(5): 347–356.

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

Sims, Christopher A. 2003. "Implications of rational inattention." Journal of monetary Economics, 50(3): 665–690.

Steiner, Jakub, Colin Stewart, and Filip Matějka. 2017. "Rational Inattention Dynamics: Inertia and Delay in Decision-Making." Econometrica, 85(2): 521–553.

Wald, Abraham. 1947. "Foundations of a general theory of sequential decision functions." Econometrica, Journal of the Econometric Society, 279–316.

Zhong, Weijie. 2022. "Optimal dynamic information acquisition." Econometrica, 90(4): 1537–1582.

Zhong, Weijie, and Alexander Bloedel. 2021. "The cost of optimally acquired information." working paper.