Optimal Sequential Experimentation

Jorge D. Ramos-Mercado*

November 10, 2023

Abstract

An impatient decision-maker ($\mathcal{DM}$) learns about an unknown state by running a sequence of experiments. He does so by managing a jump-diffusion signal process with state-dependent dynamics. The $\mathcal{DM}$ controls the signal's precision and arrival rate of jumps, but faces flow costs convexly increasing in the signal's informativeness. Jumps describe precise, infrequently-arriving breakthroughs, while the diffusion models imprecise, frequently-arriving observations. If the $\mathcal{DM}$ could, instead, flexibly manage how he learns over time, then Zhong (2022) finds that only learning from breakthroughs is optimal. When the $\mathcal{DM}$ has to experiment as described above, however, it is without loss of generality to only consider experiments that never generate breakthroughs. Intuitively, the $\mathcal{DM}$ cannot separately manage the precision and composition of acquired information. Hence, the marginal experimentation costs equals the marginal benefits of producing both infrequent breakthroughs and frequent, noisy observations.

1 Introduction

Before making an ill-informed decision, a decision-maker ($\mathcal{DM}$) may experiment: acquire a sequence of noisy, payoff-relevant observations. The $\mathcal{DM}$ often controls data precision and elicits rare, highly informative observations i.e., breakthroughs. For example, regulators demand that drug manufacturers produce high volumes of precise data supporting

*Ramos-Mercado: ITESM, ramos237@umn.edu.

a new drug's efficacy. They also set up a stringent tests whose passing equivalently corroborates a drug's efficacy. In this paper, I study optimal experimentation and the type of information that such experiment generates. Should the $\mathcal{DM}$ learn gradually by managing his observation's precision or rare breakthroughs necessary? I find that an optimal experiment always exists in which the $\mathcal{DM}$ only learns gradually and never elicits breakthroughs.

Wald (1947) first studied sequential experimentation. He analyzed when a $\mathcal{DM}$ should stop acquiring noisy observations about an unknown, payoff-relevant state and make a decision. Moscharini and Smith (2000), *MS* from henceforth, extends this setting by allowing the $\mathcal{DM}$ to control data precision. Formally, the $\mathcal{DM}$ observes a diffusion with an exogenous, state-dependent drift that is obfuscated by an independent Brownian motion, whose precision he manages at convexly increasing costs. The drift models a signal to be detected and the diffusion represents measurement error. *MS* finds that precision increases in the expected payoff from making a decision.

This setting imposes that the decision-maker can only learn gradually. In other words, the $\mathcal{DM}$ never receives, rare very precise evidence that informs him about the state i.e., he cannot elicit breakthroughs. This is not always true in the design of real-world science and (mathematically) allowing the$\mathcal{DM}$ to elicit breakthroughs is the only meaningful way to extend the type of information that can be generated in continuous-time. For these reasons, is it with loss of generality that the $\mathcal{DM}$ can only experiment by acquiring noisy information? According to Zhong (2022), *Z22* from henceforth, this restriction matters.

Rather than extending *MS* , *Z22* finds a more tractable, reduced-form approach. He assumes that the $\mathcal{DM}$ directly picks a continuous-time process for his *beliefs* about the state. Meanwhile, the cost of experimenting increases in the amount of information acquired. This approach to modeling experimentation costs follows from the rational inattention literature e.g., Sims (2003), Hébert and Woodford (2021), Caplin et al (2022), Macowiak et al (2023), among others. It has garnered critique from Denti et al (2022) since this cost structure is inconsistent with a model of experiment specific costs. Nevertheless, this approach allows costs to be explicitly stated in terms beliefs.

This approach further yields a robust prediction: the optimal experiment should only seek breakthroughs confirming the most likely state. Generating said learning process via experiments, however, requires generating perfectly precise observations. The appropriateness of this assumption depends on contexts. For example, either a miner has or has not to found gold in a given site. Models positing this stark state (e.g., Keller et al 2005) find it sensible to only model experimentation as generating perfectly precise data.

Nevertheless, the miner in question cares more about the size of deposits when deciding to dig a site. In such context, the miner estimates gold reserves surveying i.e., generating noisy data. And in in this context, generating noise-less observations in infeasible. Moreover, when experimental data is noisy, managing its precision is a central concern. Data scientist spend roughly 60 percent of their time cleaning data (Press 2016), while said tasks takes clinical researchers 80 percent of their time (Rozario et al 2017).

In this paper, I extend the setting in *MS* by allowing the $\mathcal{DM}$ to elicit state-dependent breakthroughs. An admissible experiment is now a jump-diffusion where the diffusion is identical to *MS* , but the jumps have state dependent arrival rates also controlled by the $\mathcal{DM}$. The costs of said experiments is the same as in *Z22* in order to inherit said model's tractability. I then re-ask the following question: is gradual learning inherently inefficient when a decision-maker has to explicitly manage an experiment? I find that this is not the case i.e., there exists an payoff-maximizing experiment eliciting no breakthroughs.

The intuition goes as follows. The $\mathcal{DM}$ completes two, inseparable tasks: manage data precision and decide how much to learn from the diffusion and from jumps i.e., breakthroughs . As a consequence, if an optimal experiment generates breakthroughs, the $\mathcal{DM}$ equalizes the marginal costs of experimentation with the marginal benefit from information from both noisy data data or via breakthroughs. Hence, the marginal benefit of acquiring both types of information must be the same. The assumption that experimentation takes place in continuous time and that costs depend on the amount of information acquired, then allows me to construct an experiment that does not generate jumps and is payoff equivalent to any optimal experiment that $does$ generate breakthroughs.

In the last section, I study the effect of assuming that flow experimentation costs depend on the flow amount of information generated. I consider a numerical example. A $\mathcal{DM}$ runs a pure-diffusion experiment testing a hypothesis that is either true or false, but there are two cases. In the first case, the costs of running an experiment depends in the signal's precision—as in *MS* . The second case, assumes the same functional form but its input is the flow information generated and measured by entropy. My main result is that the optimal experiment in case 1 acquires more precise information and the $\mathcal{DM}$ learns faster than in the second case. In addition, the $\mathcal{DM}$ in case 1 only stops experimenting when his beliefs reach a higher threshold of certainty that the hypothesis is correct. This example clarifies that experimentation costs

The rest of the paper proceeds as follows. Section 2 presents the model. Section 3 then states the result. Next, section 4 presents a numerical example of note. Section 5 then discusses results and concludes.

2 Model

I now present the model. An impatient, Bayesian decision-maker ($\mathcal{DM}$) with Bernoulli preferences must pick from a finite set of alternatives $A, |A| > 1$. But his payoff depends on an unknown state $x \in \{x_i\}_{i=1}^n$ for $n \in \{2, 3, \ldots\}$ i.e., payoffs are $u : A \times \{x_i\}_{i=1}^n \to \mathrm{R} \gg 0$. The $\mathcal{DM}$ initial beliefs about the state are $\pi \equiv (\pi_i)_{i=1}^n \in \Delta^{n-1}$ where $\forall i, \pi_i \equiv$ Pr$(x = x_i) > 0$.

Next, the decision-maker does not have to make a decision from the outset. Instead, the $\mathcal{DM}$ can experiment until some time $T \in [0, \infty)$ and make his choice. The details of how he experiments are presented below, but (for now) it suffices to know that the $\mathcal{DM}$ derives beliefs π_T about the state and his decision problem is simply to pick an action $a \in A$ to solve

$$F(\pi_T) \equiv \max_{a \in A} \sum_{i=1}^n \pi_{iT} u(a, x_i). \qquad \text{(Terminal Payoffs)}$$

I assume that for each pair of states x_i, x_j, $\mathrm{argmax}_{a \in A}\, u(a, x_i) \cap \mathrm{argmax}_{a \in A}\, u(a, x_i) \neq \emptyset$ if and only if (iff) $x_i = x_j$. This ensures that learning about the state is payoff relevant. In the following section, I will go over the details of experimentation i.e., how the $\mathcal{DM}$ generates the π_T beliefs.

2.1 Information Acquisition Problem

Signals I now describe the information acquisition problem. Informally, the decision-maker generates a sequence of mostly-noisy observations coupled with rare ”breakthroughs” at a cost that depends on informativeness. The setting is further set in continuous-time as it affords me a significant tractability that does not exist in discrete-time. Formally, the $\mathcal{DM}$ picks a continuous-time signal process and a stopping time $T < \infty$. An admissible experiment $s \equiv (s_t)$ is a jump-diffusion process such that $s_0 = 0$ and at each time $t \in \mathrm{R}_+$

$$\mathrm{d}s_t = \mu(x)\ \mathrm{dt} + \frac{\mathrm{d}B_t}{\sqrt{h_t}} + \mathrm{d}N_t \tag{1}$$

where $B \equiv (B_t)$ is a Brownian motion with precision $h_t \gg 0$, $\mu : \{x_i\}_{i=1}^n \to \mathrm{R}$ is one-to-one; and $N \equiv (N_t)$ is a compensated jump process. Assume that $N \perp B$ and by time $t \geq 0$, (N_t) jumps by 1 at a rate of $\lambda_{it} \geq 0$ iff $x = x_i$. Note that $\mathcal{DM}$ controls the Parameter process $\phi_t \equiv (h_t, (\lambda_{it}))$, but that $(\phi_t, 1/\sqrt{h_t})$ satisfies the standard Lipschitz condition to ensure that (s_t) admits a weak solution (see for example Le Gall (2016), Jeanblanc et al (2009), Oksendal and Sulem (2019) among others). Intuitively, this assumption ensures that the experimentation problem is well-defined. Lastly, the stopping time T is s-adapted and the space of (s, T) pairs is E_{0t}.

Information and Costs I now model an experiment's flow costs. The costs of running an experiment increase in the flow amount on information generated. To do so, I first derive a measurement of the flow amount of information generated. This approach follows from the rational inattention literature.

Information is measured as follows. Let $H : \Delta^{n-1} \to \mathrm{R} \in C^2$ be a strictly concave function and $(\pi_t) \subset$ be the Bayes posterior, s-adapted beliefs. Then then flow amount of information generated by the signal at time t is $\mathrm{I}(\phi_t) \equiv -\mathcal{L}H(\pi_t)$ where $\mathcal{L}(\cdot)$ is the infinitesimal generator for (π_t). For example, $H(\cdot)$ can be entropy and $I(\cdot)$ is then a measure of how much uncertainty was reduced by unit of time.

I lastly define the flow cost of experimentation. Let $c : \mathrm{R}_+ \to \mathrm{R} \gg 0 \in C^2$ be a strictly convex function where $c'(0) = 0$ and $\lim_{x\to\infty} c'(x) = \infty$. The flow cost of signal (s_t) at time t is then $c[\mathrm{I}(\phi_t, \pi_t)]$. I conclude by noting that $c(\cdot) > 0$ ensures that the $\mathcal{DM}$

stops in finite time, while strict convexity prevents him from learning immediately.

Payoffs and the Experimentation Problem Now that flow costs have been described, I describe payoffs. If the $\mathcal{DM}$ selects (s, T), then at time t ($\leq T$) expected payoffs are

$$V_t(s,T) \equiv E_t\left[e^{-r(T-t)}F(\pi_T) - \int_t^T c[\mathbf{I}(\phi_t)]e^{-(\tau-t)}\mathrm{d}\tau\right] \qquad \text{(Payoffs)}$$

Hence, the $\mathcal{DM}$'s problem given initial beliefs π is

$$V_0(\pi) = \max_{(s,T)\in E_{00}} V_0(s,T). \qquad \text{(Unconstrained Problem)}$$

Alternatively, the $\mathcal{DM}$ may not be able to elicit breakthroughs (jumps) i.e., at each t, he picks from the set $E_{1t} \equiv \{(s,T) \in E_{0t} \mid \forall\tau \geq t, j, x, \lambda_{xjt} = 0 \text{ a.s.}\}$. The restricted problem given belief π_t is then

$$U_0(\pi) = \max_{(s,T)\in E_{10}} V_0(s,T). \qquad \text{(Constrained Problem)}$$

3 Results.

I now present my results. The proof has four parts. First, I explicitly derive beliefs. Next, I use standard techniques from stochastic calculus to derive a value function for the restricted and unrestricted problem. I then conduct a sequence of change of variables which simplifies the problems. Lastly, I show that necessary conditions for optimality imply that $U_0(\pi) = V_0(\pi)$.

3.1 Belief dynamics

I now characterize the $\mathcal{DM}$'s Bayes consistent beliefs over time derived from observing a given signal process.

Lemma 3.1. *Fix an admissible, experiment process $s = (s_t)$, whose parameters are $(\phi_t = \{h_t, (\lambda_{it})\})$, $(\pi_t = (\pi_{it} \equiv Pr_t(x = x_i)))$ be the e-adapted beliefs, and a function f :*

$\Delta^{n-1} \to \mathrm{R} \in C^2$. *Then the generator of beliefs* π_t *at time* t *when the parameters are* ϕ_t *equals to*

$$\mathcal{L}(f(\pi_t),\phi_t) = \overbrace{h_t \sum_{ij} \frac{\pi_{it}\pi_{jt}}{2}(\mu_i-\mu_t)(\mu_j-\mu_t)f_{ij}(\pi_t)}^{\text{Diffusion}} + \overbrace{\lambda_t[f(\nu_t)-f(\pi_t)-\nabla f(\pi_t)\cdot(\nu_t-\pi_t)]}^{\text{Jumps}} \quad (2)$$

such that $f_{ij}(\pi_t) = \partial_{\pi_i}\partial_{\pi_j} f(\pi_t)$, $\lambda_t \equiv \sum_i \pi_{it}\lambda_{it}$, $\mu_t \equiv \sum_i \pi_{it}\mu_i$, $\nu_t = \frac{1}{\lambda_{t^-}} \cdot (\pi_{it^-}\lambda_{it^-})_{i=1}^n$, *and for each* $i = 1, 2, \ldots, n$, *it holds that* $\pi_{it^-} = \lim_{dt\to 0} \pi_{it-dt}$.

I now provide an intuitive sketch of the proof, but delegate the derivation to section A.1 in the appendix. First, note that the jump and diffusion process are drawn independently from each other. This allows me to approximate each as separate binomials. For some time interval dt>0, the diffusion jumps up by $\pm\sqrt{h_t \mathrm{dt}}$ with probability $\approx (1 \pm \mu_i\sqrt{h_t \mathrm{dt}})/2$ when $x = x_i$. Likewise, the jump process jumps by 1 with probability $\approx \lambda_{it}$dt, but it otherwise remains equal to 0. This formulation then allows me to derive beliefs by considering two subsets of histories; those were a jump just occurred at time t and the rest. This is done by directly applying Bayes rule. Next, I use estimate the value of several first and second moments of the change in beliefs and use said results to approximate the generator of some arbitrary function $f \in C^2(\Delta^{n-1}, \mathrm{R})$.

Next, I make two observations. First, if at some time $t \geq 0$, $\forall x_i, \lambda_{x_i t} = 0$, then I can define $\nu_t = \pi_{t^-}$ since jumps are a 0-probability event. The second observation is that the dynamics when the $\mathcal{DM}$ cannot elicit jumps follow immediately and are stated below.

Corllary 3.2. *Let* $s = (s_t)$ *be a signal such that at each time* t *and* $x = 0, 1$, $\lambda_{xt} = 0$ *almost surely (a.s.), then for each twice continuously differentiable function* $f(\cdot)$, *it holds that*

$$\mathcal{L}(f(\pi_t),\phi_t) = \frac{h_t}{2}\sum_{ij} \pi_{it}\pi_{jt}(\mu_i-\mu_t)(\mu_j-\mu_t)f_{ij}(\pi_t) \quad (3)$$

3.2 Re-formulating the Experimentation Problems

Now that the beliefs dynamics are well-defined, I define the cost function, derive the Hamilton-Jacobi-Bellman (HJB) that the decision-maker's problem must solve, and then reformulate it in a more useful fashion. First, I characterize the costs function. Fix some signal $s = (s_t)$ with parameter process (ϕ_t), then at each time t the amount of information generate is

$$\mathrm{I}(\phi_t, \pi_t)) = \overbrace{h_t \Big[-\frac{1}{2} \sum_{ij} \pi_{it}\pi_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t) H_{ij}(\pi_t) \Big]}^{\text{Information derived from diffusion}} + \overbrace{\lambda_t [H(\pi_t) - H(\nu_t) - \nabla H(\pi_t) \cdot (\pi_t - \nu_t)]}^{\text{Information derived from jumps}}.$$

Now that I derived an expression for the amount of information generated, the flow costs are $c[\mathrm{I}(\phi_t, \pi_t))]$. The formula above further illustrates several points of note. First, the signal precision enters linearly into the total amount of information and separable from the information derived from jumps. This is a feature of the continuous-time modeling choice and plays a key role in the result below.

Next, I derive an expression for the HJB describing the $\mathcal{DM}$'s optimal experimentation problem. The $\mathcal{DM}$'s problem can be written as a function of his beliefs π_t at each time t. By the principle of optimality, the optimal experiment solves a value function $V(\cdot)$ as a function of beliefs that satisfies the following variational inequalities:

$$\max \Big\{ \overbrace{F(\pi_t) - V(\pi_t)}^{\text{Stopping value}}, \max_{\phi_t \in \mathrm{R}_{++} \times \mathrm{R}^n_+} \overbrace{-h_t \sum_{ij} \frac{\pi_{it}\pi_{jt}}{2} (\mu_i - \mu_t)(\mu_j - \mu_t) V_{ij}(\pi_t)}^{\Delta V(\pi_t) \text{ due to the diffusion}}$$
$$+ \underbrace{\lambda_t [V(\nu_t) - V(\pi_t) - \nabla V(\pi_t) \cdot (\nu_t - \pi_t)]}_{\Delta V(\pi_t) \text{ due to jumps}} - \underbrace{c[\mathrm{I}(\phi_t, \pi_t)]}_{\text{Costs}} - rV(\pi_t) \Big\} = 0. \tag{4}$$

Oksendal and Sulem (2019) establish that the HJB equation, at least, admits a viscosity solution. This is because the $\mathcal{DM}$'s problem reduces to picking a locally Lipschitz collection of parameter process. They even extend the existence proof to a much broader set of problems than the one studied in this paper.

This HJB equation is far too general to make any useful insights. Instead, I consider a change of variables that clarifies the structure of the value function. Assume that the $\mathcal{DM}$ picks γ_t, ν_t, and I_t such that $I_t \equiv \lambda_t[H(\pi_t) - H(\nu_t) - \nabla H(\pi_t) \cdot (\pi_t - \nu_t)]$ and $\gamma(\pi_t) \equiv -h_t \frac{1}{2} \sum_{ij} \pi_{it}\pi_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t) H_{ij}(\pi_t)$. Then the flow of acquiring information is

$$c(\gamma_t, \nu_t, I_t) = \underbrace{c(\gamma_t + I_t)}_{\text{Noisy + Precise info.}}$$

and the generator of any function $f : \Delta^{n-1} \to \mathrm{R} \in C^2$ becomes

$$\mathcal{L}(f(\pi_t), \bar{\phi}_t) = \gamma_t\Big[- \frac{\sum_{ij} \pi_{it}\pi_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t) f_{ij}(\pi_t)}{\sum_{ij} \pi_{it}\pi_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t) H_{ij}(\pi_t)}\Big] + I_t\Big[\frac{f(\nu_t) - f(\pi_t) - \nabla f(\pi_t)\cdot(\nu_t - \pi_t)}{H(\pi_t) - H(\nu_t) - \nabla H(\pi_t)\cdot(\pi_t - \nu_t)}\Big].$$

Lastly, define for each function f, belief π_t, and posterior ν_t, the function $G(f, \pi_t, \nu_t)$ as

$$G(f, \pi_t, \nu_t) \equiv \left[\frac{f(\nu_t) - f(\pi_t) - \nabla f(\pi_t)\cdot(\nu_t - \pi_t)}{H(\pi_t) - H(\nu_t) - \nabla H(\pi_t)\cdot(\pi_t - \nu_t)}\right];$$

meanwhile, the function $L(f, \pi_t)$ equals to

$$L(f, \pi_t) \equiv \left[- \frac{\sum_{ij} \pi_{it}\pi_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t) f_{ij}(\pi_t)}{\sum_{ij} \pi_{it}\pi_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t) H_{ij}(\pi_t)}\right]$$

I can now state a more useful reformulation of the value function as

$$0 = \max\Big\{F(\pi_t) - V(\pi_t), \max_{I_t, \nu_t \geq 0, h_t > 0, \nu_t \in \Delta^{n-1}} \gamma_t L(V, \pi_t) + I_t G(V, \pi_t, \nu_t) - c(\gamma_t + I_t) - rV(\pi_t)\Big\} \quad (5)$$

This formulation is useful since the problem separates the $\mathcal{DM}$'s problem into one of picking how much information to acquire from the diffusion (i.e., γ_t), how much information to acquire from the jumps (I_t), and the posterior belief conditional on the arrival of a jump (ν_t).

In a similar fashion, the restricted problem forces that for each $i = 1, 2, \ldots, n$ $\lambda_{it} = 0$ for certain. This implies that $I_t = 0$ and $\nu_t = \pi_t$. As a consequence, I can make the same derivations for the restricted problem and the value function $U(\cdot)$ satisfies that

$$0 = \max\left\{F(\pi_t) - U(\pi_t), \max_{I_t, \nu_t \geq 0, h_t > 0, \nu_t \in \Delta^{n-1}} \gamma_t L(U, \pi_t) - c(\gamma_t) - rU(\pi_t)\right\} \quad (6)$$

I now characterize the value function $U(\cdot)$ and its optimal control in the lemma below.

Lemma 3.3. *The HJB stated in equation 6 has a unique solution $U(\cdot)$ that is twice continuously differentiable and admits a unique, optimal control that is a Markov function of π_t i.e., $\bar{\sigma} : \Delta^{n-1} \to \mathrm{R}_+$.*

The lemma is an immediate application of Theorem 1 and 2 from Strulovici and Szydlowski (2015).

3.3 Main Result

This section presents my main result. I will first state the result and then provide the proof.

Theorem 3.4. *The $\mathcal{DM}$'s expected payoff in the constrained and unconstrained problems are the same:*

$$\forall_{\pi_t \in [0,1]} \ U(\pi_t) = V(\pi_t) \ \textit{a.s.} \tag{7}$$

This theorem implies that restricting the $\mathcal{DM}$ to picking a diffusion signal is without loss of generality. The proof goes as follows. First, it is immediate that for each belief $\pi_t \in \Delta^{n-1}$, it must be that $U(\pi_t) \leq V(\pi_t)$. This is because a control $\phi_U \equiv (\gamma, I, \nu) : \Delta^{n-1} \to \mathrm{R}_{++} \times \mathrm{R}_{+} \times \Delta^{n-1}$ such that $h(\cdot)$ is the same control maximizing the constrained problem, $I(\pi_t) = 0$ and $\nu(\pi_t) = \pi_t$ is an admissible control for the general problem. Hence, the payoff from using said control when the current belief is π_t can be defined as $V(\pi_t; \phi_U)$ and it holds that $V(\pi_t, \phi_U) = U(\pi_t)$ and $V(\pi_t, \phi_U) \leq V(\pi_t)$.

What I need to show is that the inequality also holds in the opposite directions i.e., $U(\pi_t) \geq V(\pi_t)$. First observe that if at belief $\pi_t \in \Delta^{n-1}$, it holds that $V(\pi_t) = F(\pi_t)$, then $U(\pi_t) = F(\pi_t)$ since

$$\underbrace{F(\pi_t) \leq U(\pi_t)}_{\text{restricted } \mathcal{DM} \text{ can stop}} \leq \underbrace{V(\pi_t) = F(\pi_t)}_{\text{Unrestricted } \mathcal{DM} \text{ wants to stop}} .$$

This implies that it is without loss of generality to focus on the set of beliefs π_t for which the $\mathcal{DM}$ prefers to experiment when he is not restricted: $C \equiv \{\pi_t \in \Delta^{n-1} : V(\pi_t) > U(\pi_t)\}$. Next suppose that $\phi' = (\gamma', I', \nu')\Delta^{n-1} \to \mathrm{R}_{++} \times \mathrm{R}_{+} \times \Delta^{n-1}$ attains the maximum of the following problem below for each $\pi_t \in C$:

$$rV(\pi_t) = \max_{I_t, \nu_t \geq 0, h_t > 0, \nu_t \in \Delta^{n-1}} \gamma_t L(V, \pi_t) + I_t G(V, \pi_t, \nu_t) - c(\gamma_t + I_t)$$

At each belief $\pi_t \in C$, $\phi'(\pi_t)$ must satisfy two conditions. First, $\phi'(\pi_t)$ must satisfy the first order conditions. This implies that $\gamma_t \gg 0$ for $\pi_t \in C$ and satisfies $c'(\gamma_t + I_t) = L(V, \pi_t)$. Likewise, if $I'(\pi_t) > 0$, then it must satisfy that

$$\underbrace{c'(\gamma_t + I_t)}_{\text{Marg. cost of Info.}} = \underbrace{G(V, \pi_t, \nu_t)}_{\text{Marg. benefit of breakthroughs}} = \underbrace{L(V, \pi_t)}_{\text{Marg. benefit of noisy data}}$$

Secondly, it must satisfy the principle of optimality i.e., for each $\pi_t \in C$, it holds that

$$\begin{aligned} rV(\pi_t) &= \gamma'(\pi_t)L(V, \pi_t) + G(V, \pi_t, \nu'(\pi_t))I'(\pi_t) - c[\gamma_t(\pi_t) + I_t(\pi_t)] \\ &= \underbrace{[\gamma'(\pi_t) + I'(\pi_t)]}_{\text{total info.}} \underbrace{L(V, \pi_t)}_{\text{Benefit from experimentation}} - \underbrace{c[\gamma_t(\pi_t) + I_t(\pi_t)]}_{\text{Cost of info.}} \end{aligned}$$

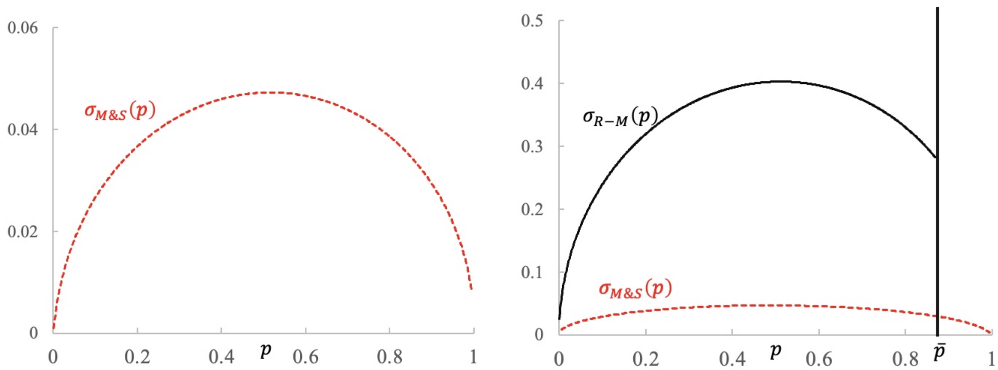
Notice that the second line follows from the observation that the marginal benefits both types of information must be equalized whenever the $\mathcal{DM}$ pick a strictly positive amount of both types of information. On the other hand, if $I'(\pi_t) = 0$, then the expression is unaffected by the $I'(\cdot)$ part of the expression. Alternatively, let $\bar{\phi} \equiv (\bar{\gamma}, \bar{I}, \bar{\nu}) : [0, 1] \to \mathrm{R}^2_+ \times [0, 1]$ be defined for each π_t as $\bar{\gamma}(\pi_t) = \gamma'(\pi_t) + I'(\pi_t), \bar{I}(\pi_t) = 0$, and $\nu(\pi_t) = \pi_t$. Then, by construction, $\bar{\sigma}(\pi_t)$ satisfies the first order conditions for the optimization problem and for each $\pi_t \in C$, it holds that

$$V(\pi_t) = \bar{\gamma}(\pi_t)L(V, \pi_t) - c(\bar{\gamma}_t(\pi_t))$$

Consequently, it holds that for each π_t $V(\pi_t; \bar{\phi}) = V(\pi_t)$. Likewise, $\bar{\phi}$ is admissible in the constrained problem. This implies that for each π_t, it holds that $V(\pi_t) = V(\pi_t; \phi) \leq U(\pi_t)$. This establishes the result.

4 Numerical Example: Angel investor's problem

I now present a useful example. An investor is given the opportunity to invest in a startup. Said startup's long run profitability is $x = 0, 1$. Not investing nets him a payoff of 0. Otherwise, investing nets a payoff of $x - c$ for constant $c \in (0, 1)$. This implies that for



(a) ***Optimal Policy in case 1.*** *(b)* ***Optimal Policy in case 2.***

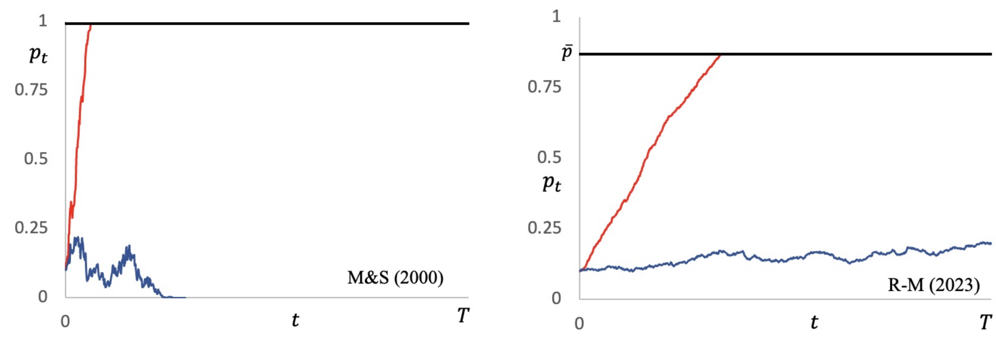
Figure 1: ***Optimal level of noise as a function of*** π_t***.***

each belief $\pi_t \equiv \text{Pr}_t(x = 1)$, terminal payoffs from making an investment decision at time t is $\Pi(\pi_t) = \max\{0, \pi_t - c\}$.

Before the investor decides, he can experiment and learn about x. He does so by managing a signal process $s \equiv (s_t)$ such that $s_0 = 0$ and $\text{d}s_t = \text{xdt+d}B_t/h_t$. The flow costs from running said experiments, however, can be one of two possible cases. In case 1, the flow costs are quadratic in h_t i.e., $h_t^2/2$ at time t. This is the type of cost function admissible in *MS*. In case 2, the signal generates a flow of information $I_t = -E_t[dH(\pi_t)]/\text{dt}$ where $H(\cdot)$ is entropy and the flow cost of the experimentation is quadratic on the flow $information$ generated i.e., $I_t^2/2$ (as in this paper).

I now discuss the optimal policy in each case. Figure 1's panel (a) illustrates the optimal level of experiment noise as a function of belief π_t (in red) when costs are given by case 1, said noise is single peaked in beliefs. Panel (b), for its part, adds the policy function for case 2—in black. Said policy function differs from the case 1 policy in 2 ways. First, the investors always acquires less precise information than in case 2. Secondly, the investor decides to invest when his beliefs reach a lower cutoff relative to the cutoff in case 1.

Next, I simulate the belief process (π_t) in each case when the true state of the world is $x = 0$ (in blue) or $x = 1$ (in red) in figure 2. Panel (a) describes how beliefs evolve in case 1, while panel (b) describes how beliefs evolve in case 2. Beliefs are simulated for the



(a) ***Case 1*** *(b)* ***Case 2.***

Figure 2: ***Simulation of*** π_t ***given the optimal experiments when*** $x = 0$ ***and when*** $x = 1$**.**

same time horizon and using the same randomization procedure. Said simulations show that beliefs evolve more gradually in case 2 relative to case 1. This observation, by itself, is not interesting. What is noteworthy is that experimentation takes considerably more in case 2 relative to case 1.

5 Discussion

This paper studies optimal, sequential experimentation when the decision-maker conducting the experiment flexibly picks his data generating process. My main result is that the $\mathcal{DM}$ loses nothing by not acquiring infrequently arriving breakthroughs that would decisively resolve most of the uncertainty faced. This is because rather than acquiring breakthroughs, the decision-maker can always generate additional information.

This result has 2 implications. First, suppose that one assumes that the variable flow costs of experimentation depend on the flow amount of information generated. Then the optimal salient decision in faced by a sequential experimenter is how much information is acquired. Secondly, the optimal way for a decision-maker to sequentially experiments is sensible to seemingly innocuous restrictions on the space of feasible signals and on the functional form of costs.

References

[1] Aliprantis, C., Border, K., 2006. Infinite Dimensional Analysis, second edition. Springer-Verlag.

[2] Bolton, Patrick, and Christopher Harris. 1999. "Strategic signal". Econometrica (ECTA) 67 (2), 349–374.

[3] Caplin, Andrew, Mark Dean, and John Leahy. 2022. "Rationally Inattentive Behavior: Characterizing and Generalizing Shannon Entropy." Journal of Political Economy 130 (6): 1676–1715.

[4] Bloedel, Alexander W., and Weijie Zhong. 2021. "The Cost of Optimally Acquired Information." Unpublished.

[5] Denti, Tommaso, Fabrizio Marinacci, and Aldo Rustichini. 2022. "Experimentation Cost of Information". American Economic Review (2022) 12(9): 3106–3123.

[6] Denti, Tommaso. Forthcoming. "Posterior Separable Cost of Information." American Economic Review.

[7] Denti, Tommaso, Massimo Marinacci, and Aldo Rustichini. 2022. "The experimental Order on Random Posteriors." Unpublished.

[8] Hébert, Benjamin, and Michael Woodford. 2021. "Rational Inattention When Decisions Take Time." Unpublished.

[9] Jeanblanc, Monique, Marc Yor, Marc Chesney. 2009. "Mathematical Methods for Financial Markets". Springer-Verlag Finance London.

[10] Mensch, Jeffrey. 2018. "Cardinal Representations of Information." Unpublished.

[11] Morris, Stephen, and Philipp Strack. 2019. "The Wald Problem and the Equivalence of Sequential Sampling and Static Information Costs." SSRN 2991567.

[12] Øksendal , Bernt and Agnès Sulem. 2019. "Applied Stochastic Control of Jump Diffusions". Springer-Verlag Switzerland.

[13] Platten, Eckhart (1982). ”An approximation method for a class of Ito processes with jump component”. Lient. Mat Rink 22(2): 124-136.

[14] Pomatto, Luciano, Philipp Strack, and Omer Tamuz. 2020. “The Cost of Information.” Unpublished.

[15] Mike Press. 2016. ”Cleaning Big Data: Most Time-Consuming, Least Enjoyable Data Science Task, Survey Says”. Forbes retrieved from: $https : //www.forbes.com/sites/gilpress/2016/03/23/data - preparation - most - time-consuming-least-enjoyable-data-science-task-survey-says/amp/$

[16] Sims, Christopher A. 2003. “Implications of Rational Inattention.” Journal of Monetary Economics 50(3): 665–90.

[17] Strulovici, Bruno and Martin Szydlowski. 2015. ”On the smoothness of value functions and the existence of optimal strategies in diffusion models”. Journal of Economic Theory 159:1016–1055.

[18] Wald, A. 1947. ”Sequential analysis”. John Wiley.

[19] Zhong, Weijie. 2022 ”Optimal Dynamic Information Acquisition”. Econometrica 90(4):1537-1582.

A Proofs

A.1 Deriving dynamics.

In this section, I present the characterization of beliefs. Informally, I approximate the signal process in discrete time and take limits.

Approximating signals in discrete-time Fix some small time interval dt> 0, then an admissible signal $s = (s_t)$ can be approximated at times $t = 0, \mathrm{dt}, \ldots$ as $s_0 = 0$ and $\mathrm{d}s_t \equiv s_{t+\mathrm{dt}} - s_t$

$$\mathrm{d}s_t = \mathrm{d}_t^{\mathrm{dt}} + J_t^{\mathrm{dt}} \tag{8}$$

for $(\mathrm{d}_t^{\mathrm{dt}})_{t=0}^{\infty}$ is a sequence of independent random variables such that at time t, $d_t^{\mathrm{dt}} = \pm\sqrt{h_t \mathrm{dt}}$ with probability $[1 - \mu_i\sqrt{h_t\mathrm{dt}}]/2$ iff $x = x_i$; meanwhile, $(J_t^{\mathrm{dt}})_{t=0}^{\infty}$ is a sequence of independent random variables such that $J_t^{\mathrm{dt}} = 1$ with probability $\lambda_{it}\mathrm{dt}$ and $J_t^{\mathrm{dt}} = 0$ with probability $1 - \lambda_{it}\mathrm{dt}$ iff $x = x_i$. This means that the independent diffusion and jump processes are weakly approximated to since the objective is to approximate a generator.

Approximating beliefs after a jump I first consider the case when there are jumps. Suppose that the $\mathcal{DM}$ held beliefs $\pi_{t-\mathrm{dt}} = (\pi_{it-\mathrm{dt}})$, then the Bayes posterior belief that $x = x_i$ given the jump is approximately equal to

$$\pi_{it} = \frac{\pi_{it-\mathrm{dt}}\mathrm{dt}\lambda_{it}}{\sum_j \pi_{jt-\mathrm{dt}}\mathrm{dt}\lambda_{jt}} + o(\mathrm{dt}) = \frac{\pi_{it-\mathrm{dt}}\lambda_{it}}{\sum_j \lambda_{jt}\pi_{jt-\mathrm{dt}}} + o(\mathrm{dt})$$

where the error term $o(\mathrm{dt})$ (such that $\lim_{\mathrm{dt}\searrow 0} o(\mathrm{dt})/\mathrm{dt} = 0$) follows from the observation that distribution of d_t^{dt} approximately gives equal weight to both outcomes as dt goes to 0. Further observe that as dt goes to 0, it holds that

$$\pi_{it} = \frac{\pi_{it-}\lambda_{it}}{\sum_j \pi_{jt-}\lambda_{jt}} = \frac{\pi_{it-}\lambda_{it}}{\lambda_t} = \nu_{it}.$$

It is further the case that the expected change in beliefs equals to $\mathrm{d}\pi_{it} \equiv \pi_{it} - \pi_{it-\mathrm{dt}}$ and satisfies that

$$\mathrm{d}\pi_{it} = \frac{\pi_{it-\mathrm{dt}}(\lambda_{it} - \lambda_t)}{\lambda_t}$$

Approximating beliefs when there is no jump Next, I characterize how beliefs change when $J_t^{\mathrm{dt}} = 0$. Suppose that the $\mathcal{DM}$ observes $\mathrm{d}s_t = \pm\sqrt{h_t\mathrm{dt}}$, then the probability of observing said signal realization conditional on $x = x_i$, for $i = 1, 2 \ldots, n$, is

$$\Pr_t(\mathrm{d}s_t = \pm\sqrt{h_t\mathrm{dt}}|x = x_i) = [1 - \lambda_{it}\mathrm{dt} \pm \mu_i\sqrt{h_t\mathrm{dt}}]/2 + o(\mathrm{dt})$$

Once again, if the prior belief is $\pi_{t-\mathrm{dt}} = (\pi_{it-\mathrm{dt}})$, then the Bayes posterior beliefs are

$$\pi_{it} = \frac{\pi_{it-\mathrm{dt}}[1-\lambda_{it}\mathrm{dt} \pm \mu_i\sqrt{h_t\mathrm{dt}}]}{\sum_j \pi_{jt-\mathrm{dt}}[1-\lambda_{jt}\mathrm{dt} \pm \mu_j\sqrt{h_t\mathrm{dt}}]} + o(\mathrm{dt}) = \frac{\pi_{it-\mathrm{dt}}[1-\lambda_{it}\mathrm{dt} \pm \mu_i\sqrt{h_t\mathrm{dt}}]}{1-\lambda_t\mathrm{dt} \pm \mu_t\sqrt{h_t\mathrm{dt}}} + o(\mathrm{dt}).$$

This equation implies that the change in beliefs $d\pi_{it} \equiv \pi_{it} - \pi_{it-\mathrm{dt}}$ can be approximated as

$$d\pi_{it} = \frac{\pi_{it-\mathrm{dt}}[\pm(\mu_i - \mu_t)\sqrt{h_t\mathrm{dt}} - (\lambda_{it} - \lambda_t)\mathrm{dt}]}{1-\lambda_t\mathrm{dt} \pm \mu_t\sqrt{h_t\mathrm{dt}}} + o(\mathrm{dt})$$

and the probability that beliefs change by the amount described above occurs with a probability of approximately $(1-\lambda_t\mathrm{dt} \pm \mu_t\sqrt{h_t\mathrm{dt}})/2$.

Approximating the expected change in beliefs Given the approximations for the change in beliefs given above, I now take expectations. First, conditional on there being no jump, the expected change in beliefs equals to

$$\begin{aligned} E[d\pi_{it}|J_t^{\mathrm{dt}}=0] &= \pi_{it-\mathrm{dt}}[(\mu_i-\mu_t)\sqrt{h_t\mathrm{dt}} - (\lambda_{it}-\lambda_t)\mathrm{dt}]/2 - \pi_{it-\mathrm{dt}}[(\mu_i-\mu_t)\sqrt{h_t\mathrm{dt}} + (\lambda_{it}-\lambda_t)\mathrm{dt}]/2 + o(\mathrm{dt}) \\ &= -\pi_{it-\mathrm{dt}}(\lambda_{it}-\lambda_t)\mathrm{dt} + o(\mathrm{dt}) = -\lambda_t(\nu_{it} - \pi_{it-\mathrm{dt}})\mathrm{dt} + o(\mathrm{dt}). \end{aligned}$$

Note that the expected change in beliefs when there is no jump is

$$E[d\pi_t|J_t^{\mathrm{dt}}=1] = \nu_t - \pi_t.$$

As a consequence, the unconditional change in beliefs just equals to the expectation over the conditional expectations i.e.,

$$E[d\pi_{it}] = \lambda_t\mathrm{dt}\left(\frac{\pi_{it-\mathrm{dt}}(\lambda_{it}-\lambda_t)}{\lambda_t}\right) + (1-\lambda_t\mathrm{dt})E[d\pi_{it}|J_t^{\mathrm{dt}}=0] = o(\mathrm{dt}).$$

Observe that this term holds due to the Law of Iterated expectations. I will exploit this observation again when estimating the change of a variable. Further notice that the approximations imply that beliefs will form a martingale.

Approximating the co-movement of beliefs Next, conditional on no jumps to co-movement of the belief that $x = x_i$ or x_j (for $i, j = 1, 2, \ldots, n$) is

$$\mathrm{d}\pi_{it}\mathrm{d}\pi_{jt} = \left[\frac{h_t\pi_{it-\mathrm{dt}}\pi_{jt-\mathrm{dt}}(\mu_i-\mu_t)(\mu_j-\mu_t)}{1 \pm \mu_t^2 h_t\mathrm{dt}}\right]\mathrm{dt} + o(\mathrm{dt})$$

occurs with a probability of roughly $(1 \pm \mu_t^2 h_t \text{dt})/2$. The expected covariance in the change of said observations is then equal to

$$E[\text{d}\pi_{it}\text{d}\pi_{jt}|J_t^{\text{dt}} = 0] = h_t \pi_{it-\text{dt}} \pi_{jt-\text{dt}} (\mu_i - \mu_t)(\mu_j - \mu_t)\text{dt} + o(\text{dt}).$$

Approximating the generator Lastly, I approximate the generator. Let $f : \Delta^{n-1} \to \text{R}$ be a twice continuously differentiable function, then I need to calculate $E[df(\pi_t)]/\text{dt}$ as dt goes to 0. Note that $\text{d}f(\pi_t) = f(\pi_t) - f(\pi_{t-\text{dt}})$. I can partition the expectation by the law of iterated expectations. If there is a jump, then

$$E[df(\pi_t)|J_t^{\text{dt}} = 1] = f(\nu_t) - f(\pi_{t-\text{dt}}).$$

Alternatively, there may have been no jumps, then the change in beliefs can be approximated via a quadratic Taylor approximation as

$$\begin{aligned} E[df(\pi_t)|J_t^{\text{dt}} = 0] &= E[\nabla f(\pi_t) \cdot (d\pi_{it}) + (d\pi_{it}) H f(\pi_t) \cdot (d\pi_{it})/2 | J_t^{\text{dt}} = 0] + o(\text{dt}) \\ &= -\nabla f(\pi_t) \cdot (\nu_t - \pi_{t-\text{dt}}) \lambda_t \text{dt} + \frac{h_t}{2} \text{dt} \sum_{ij} \pi_{it-\text{dt}} \pi_{jt-\text{dt}} (\mu_i - \mu_t)(\mu_j - \mu_t) f_{ij}(\pi_{t-\text{dt}}) + o(\text{dt}) \end{aligned}$$

Since the probability of a jump is approximately $\lambda_t$dt, then the unconditional expectation equals to

$$\begin{aligned} &E_t[df(\pi_t)] = \lambda_t \text{dt}[f(\nu_t) - f(\pi_t)] + [1 - = \lambda_t \Delta] \times E[df(\pi_t)|J_t^{\text{dt}} = 0] \\ &= \frac{h_t}{2} \text{dt} \sum_{ij} \pi_{it-\text{dt}} \pi_{jt-\text{dt}} (\mu_i - \mu_t)(\mu_j - \mu_t) f_{ij}(\pi_{t-\text{dt}}) + [f(\nu_t) - f(\pi_{t-\text{dt}}) - \nabla f(\pi_t) \cdot (\nu_t - \pi_{t-\text{dt}})] \lambda_t \text{dt} + o(\text{dt}) \end{aligned}$$

Dividing both sides of the expression above by dt and taking the limit as dt goes to 0 it yields that

$$\mathcal{L}(f(\pi_t), \phi_t) = \frac{h_t}{2} \sum_{ij} \pi_{it} \pi_{jt} (\mu_i - \mu_t)(\mu_j - \mu_t) f_{ij}(\pi_t) + \lambda_t [f(\nu_t) - f(\pi_t) - \nabla f(\pi_t) \cdot (\nu_t - \pi_t)].$$

This concludes the proof.