# Optimal Sequential Experimentation

Jorge D. Ramos-Mercado*

September 23, 2024

**Abstract**

A decision-maker (DM) must experiment when he is an early decider i.e., he decides before information is externally available. For example, when a scientist tests a new hypothesis or a business adopts an unproven technology. If DMs have ample resources, does being an early decides qualitatively changes the type of information acquires over time? The answer is yes. In the model, I extend the notion of an experiment to be a jump-diffusion whose drift and the arrival rate of jumps depends on a payoff-relevant state. The DM flexibly controls the diffusion's precision and the arrival rate of jumps, but faces costs convexly increasing in the flow amount of information generated. I find that an early decider's optimal experiment is a pure-diffusion i.e., it frequently generates imprecise information. In contrast, Zhong (2022) found that subsequent DMs only acquire infrequently-arriving but precise information.

A decision-maker (DM, he) often obtains payoff-relevant information prior to making a key decision. But if a DM is a early decider (i.e., must decide before relevant information is externally abundant), he must generate information by experimenting. That is, the DM must manage a noisy, data generating process. This constraint is salient, for example, when a firm is first to adopt a new technology, a scientists tests a new hypothesis, or a lawmaker is first to draw a new regulation. In such context, do early deciders qualitatively changes the type of information obtained over time? Even when a DM has ample resources, I find that being a early decider changes the type of information obtained before making

*Ramos-Mercado: EGADE Business School, Tecnológico de Monterrey, jdramosmercado@tec.mx.

a key decision. In particular, a early decider frequently generates imprecise information; meanwhile, subsequent DMs obtain infrequently arriving but precise information.

I extend Moscarini and Smith's (2001, MS from henceforth) framework. The DM learns about a payoff-relevant state by observing a continuous-time signal process: an experiment. In MS, an experiment is a diffusion whose drift depends on the state and the diffusion noise is flexibly controlled by the DM. The drift models the model's information which is obfuscated by noise (modeled by the diffusion) that the DM can reduce. Such noise the measurement inherent in all experimental, which cannot be fully avoided. I extend their notion of experiments by allowing experiments to be jump-diffusion processes. Both the drift and the arrival rate of jumps depends on the state and the DM flexibly controls the diffusion noise and the arrival rate of jumps.[1] Said jumps model rare, outlier observations that are highly informative of the state. For example, they model stringent sales benchmarks that a product or a startup generate; the rare detection of an outlier observation confirming a hypothesis; etc..

Experimentation is further costly. In Moscarini and Smith's model, flow costs increase convexly with precision i.e., decreasing measurement error is costly and costs increase at an increasing rate. Since the DM in my framework controls multiple parameters, it is not clear how one should extend the pre-existing cost structure. For tractability, I assume that costs increase convexly as one increases the flow amount of information generated i.e., extracting more information from the experiment is increasingly costly. This is an assumption which follows from the rational inattention literature e.g., Sims (2003), Steiner et al. (2017), Hébert et al. (2021), Zhong (2022).

These costs are useful, in the context of running an experiment, because it models the costs and burden of data processing—which is central in real-world applications. For example, chemists spend 60% of their time generating data (Princeton Review 2024), while data cleaning occupies 60% of data scientists' (Press 2016) and 80% of clinical researchers' time (Rozario et al. 2017). Moreover, issues related to poor-quality data cost the U.S. economy an estimated $3.1 trillion annually (Redman 2016). Thus, a model of

[1] If one assumes that experiments are Lévy processes and that are consistent with the classic notion of experiments, then the Lévy-Khintchine theorem implies that the set of experiments considered in this paper is general.

dynamic experimentation that does not account for these data processing consideration is unlikely to yield useful applications.

This cost assumption, however, has some drawbacks. In general, flow costs depend on the parameters chosen as well as current beliefs. Having costs of experimenting that depend on the DM's beliefs contradicts the assumption that experimental costs model the physical constraints associated with data cleaning and processing. Nevertheless, flow costs depending on the chosen parameters would require arbitrarily imposing conditions on the cross-derivative of said function. Such assumptions would also be contrived and only relevant for highly specific applications.

In spite of this limitation, the model yields a unique prediction, which clarifies how being a early decider affects the type of information that the DM acquires. If the DM was not a early decider and could flexibly acquire information from an external source, he would acquire a confirmatory, Poisson signal (Zhong 2022). That is, the DM acquires infrequent but precise information about the state. Crucially, Zhong (2022) assumes the came costs structure as adopted in the current model. In contrast, I find that the optimal experiment is a pure-diffusion i.e., he frequently generates imprecise information. Therefore, even when the DM faces the same type of source, forcing him to be a early decider (i.e., to experiment) qualitatively the type of information obtained.

The proof follows from making a simple but consequential observation. Rather than explicitly picking the experiment parameters, I reformulate his problem into a constrained information acquisition problem. He picks the flow amount of information generated by the jumps (Poisson information), his posterior beliefs conditional on a jump arriving, and the flow amount of information generated by the diffusion (Gaussian information). Such problem is further restricted, because the DM must always acquire some Gaussian information i.e., measurement error is never fully eliminated.

This change of variable immediately clarifies that Poisson and Gaussian information are perfect substitutes. Therefore, if an optimal experiment generates Poisson information, then there exists a payoff-equivalent information only generating Gaussian information i.e., it was feasible in MS. Additionally, I find that the DM cannot benefit from experimenting and acquiring Poisson information (with a positive probability) at the same time. Hence, with probability one, the optimal experiment is a pure-diffusion.

To characterize the optimal experiment, I modify the proof presented in Moscarini and Smith (1998,2001). The proof consists of reformulating the DM's problem in terms of the flow amount of information generated rather than the signal precision. Once one takes the appropriate steps to ensure that the set of feasible experiments remains well-defined, the rest of MS' proof follows without change.

I find that the optimal flow amount of information generated is a strictly increasing, Markov function of the value of experimenting. In fact, the policy functions (in terms of the value of experimenting) are the same in both settings. Nevertheless, the value of experimenting (as a function of the DM's beliefs) differs between both settings.

The rest of the paper proceeds as follows. The literature review is in section 1. Next, section 2 presents the model. Section 3 states results. Lastly, I conclude in section 5.

# 1 Literature Review

Sequential experimentation was first studied by Wald (1947), who analyzed the problem faced by a DM that sequentially acquires iid signals at a fixed, exogenously determined rate. Each signal represents an observation made by the DM, and the DM decides when to stop experimenting. While generalizing this problem in discrete time can be complex, Bolton and Harris (1999), El Karoui and Karatzas (1994), and others showed that the model becomes more tractable in continuous time. Specifically, they model the problem as a multi-armed bandit.

Moscarini and Smith (1998, 2001) extend this continuous-time approach as described in the introduction: they allowed the decision-maker to control each arm's precision. The key limitation with this approach is that the DM is only able to frequently generate imprecise information. What the DM controls is the data precision, but he has no way of acquiring any other type of information. In particular, a continuous-time signal process can also generate Poisson jumps i.e., infrequent but precise information.

To address this limitation, Zhong (2022) adopts an indirect approach. Instead of managing a signal process, the DM selects a Lévy martingale process for his beliefs. Lévy processes are a large class of continuous-time stochastic processes satisfying a very mild set of technical conditions and have been shown to be decomposable as the sum of two, independent processes: a diffusion and a compensated Poisson-process. His model assumes that the cost of acquiring information increases convexly with the flow of information generated, a standard assumption in the rational inattention literature and adopted in the current setting.

This approach has its own limitations. Many Lévy processes cannot easily be interpreted as experiments. For instance, a detector measuring the signature of a natural phenomenon is not well-represented by a Poisson process alone, as such detectors also capture "static" noise. This noise, while highly imprecise, depends on the phenomenon being studied and provides some information to the DM. Modeling the detector solely as a Poisson process overlooks the possibility that the DM can filter information from the statics noised while waiting for a precise signal to arrive. To account for this, my model restricts feasible beliefs to those that can be generated by noisy signal processes, while still allowing for Poisson signals.

# 2 Model

**Decision Problem** A Bayesian decision-maker (DM, he), with standard Bernoulli preferences, picks an alternative $a$ from a finite set of alternatives $A$ where $\#A > 1$. His payoff further depends on an unknown state $x$ that belongs to a finite set $X = \{x_i\}_{i=1}^n$ with $n = 2, 3, \dots$ and his initial beliefs are $p \equiv (p_i)_{i=1}^n \in \Delta^{n-1}$ where $p \gg 0$ and for each $x_i$, $p_i \equiv \Pr(x = x_i)$. The DM's payoff from picking alternative $a$ when the state is $x$ is $u(a, x) \gg 0$. Assuming that terminal payoffs are strictly positive ensures that the DM strictly benefits from eventually making a decision.

He is further able to experiment—described below—and pick a time $T \in [0, \infty)$ when to stop and make an irreversible decision. Lastly, if he stops at time $T$ when his beliefs are $p_T$, his payoff from making a decision is

$$F(p_T) \equiv \max_{a \in A} \sum_i p_{iT} u(a, x_i). \qquad \text{(Terminal Payoffs)}$$

Note that $F(\cdot) \gg 0$ since $u(\cdot, \cdot) \gg 0$.

## 2.1 Information Acquisition Problem

**Signals** I now describe the experimentation problem. The DM picks a pair $(s, T)$ where $T \geq 0$ is an $s$-adapted stopping time (i.e., rule for when to stop and decide)and $s = (s_t)$ is a signal process (n $experiment$) where $s_0 = 0$ and at each time $t \in [0, \infty)$

$$\mathrm{d}s_t = \mu(x)\ \mathrm{dt} + \frac{\mathrm{d}B_t}{\sqrt{h_t}} + \sum_k \Delta_k \mathrm{d}N_{kt}. \tag{1}$$

$(B_t)$ is a Brownian motion with precision process $(h_t) \gg 0$ and the drift is an injective function of the state i.e., for each distinct pair $x_i$ and $x_j$, $\mu(i) \neq \mu(j)$. [2]For clarity, I denote the drift when $x = i$ as $\mu_i$ for each $i \in X$. Next, for each $k = 1, 2, \ldots, K$, $(N_{kt})$ is a compensated, Poisson jump-process. Such process jumps by $1$ at time $t$ at a rate of $\lambda_{ikt} \geq 0$ if $x = i$. I further assume that for each distinct pair $k$ and $k'$, it holds that $\Delta_k \neq \Delta_{k'}$. This assumption implies that the DM can tell which process jumped.

Next, I assume that $\{(B_t), \{(N_{kt})\}_{k=1}^K\}$ are a collection of pairwise independent stochastic processes. This is an innocuous assumption due to the standard decomposition of Lévy processes. In addition, $\{h_t, \{\{\lambda_{ikt}\}_{i=1}^n\}_{k=1}^K\}$ satisfies the standard the standard Lipschitz condition ensuring that $(s_t)$ admits a weak solution (e.g., see Oksendal and Sulem 2019). In other words, I require that parameters that the DM chosen describe a unique and well-defined stochastic process.

**Information and Costs** I now describe how costs are modeled. Intuitively, I define a notion of flow information generated (consistent with Zhong 2022) and then assume that experiment costs convexly increase in the flow amount of information generated. Formally, let $H : \Delta(X) \to \mathrm{R}$ be a strictly convex and twice continuously differentiable function. Next, $(p_t) \subset \Delta(X)$ be the Bayes posterior belief process adapted to observing the signal paths i.e., $\{s_\tau \mid \tau \in [0, t)\}$. This means that $p_0 = p$ (i.e., the prior) and at times $t \geq 0$ beliefs evolve via Bayes rule from *only* observing the experiment's data.

The flow amount of information generated by $(s_t)$ at time $t$ is $I_t \equiv \frac{E_t \mathcal{L} H(p_t)}{\mathrm{dt}}$ where $\mathcal{L}(\cdot)$ is $(p_t)$'s infinitesimal generator.[3] For example, $-H(\cdot)$ could equal entropy and (in-

[2]Note that assuming that $(h_t) \gg 0$ simply ensures that the signal process is real-valued and well-defined.

[3]In particular, for each function $f \in C^2(\Delta(X), R)$ and belief $p_t \in \Delta(X)$, it holds that $\mathcal{L}f(p_t) \equiv$

tuitively) information is defined the expected amount by which uncertainty falls. I make further technical assumptions on $H(\cdot)$, but do so later since the context clarifies why such assumptions are useful. Next, the flow cost of experimenting at time $t$ is $c(I_t)$ where $c(0) \geq 0$ and $c(\cdot)$ is further strictly increasing, strictly convex, and continuously differentiable.[4] Since experimentation is a costly and the payoff from making a decision are strictly positive, then I ensure that the decision-maker strictly prefers eventually making a decision.

**Payoffs** I now describe payoffs. If the decision-maker picks a pair $(s, T)$, then his time $t \geq 0$ payoffs are

$$V_t(s,T) \equiv E_t\left[e^{-r(T-t)}F(p_T) - \int_t^T c(I_t)e^{-(\tau-t)}\mathrm{d}\tau\right] \tag{Payoffs}$$

for some discount rate $r > 0$. Intuitively, the DM's payoff equals the discounted value of making an informed decision minus the discounted flow costs of running an experiment. Next, given initial beliefs $p$, DM's optimization problem is

$$\mathcal{V}(p) \equiv \max_{s,T} V_0(s,T). \tag{Unconstrained Problem}$$

Alternatively, one may force him to pick an experiment generating pure-diffusion information i.e.,

$$U(p) = \max_{s,T} V_0(s,T) \text{ s.t. } \forall i,t, \lambda_{it} = 0 \text{ a.s.} \tag{Constrained Problem}$$

$\lim_{\Delta_t \to 0} \frac{f(p_t)-f(p_{t-\Delta_t})}{\Delta_t}$. Intuitively, the generator is the "derivative' of the expected change in $f$.

[4]It is worth noting that convexity ensures that the DM prefers learning over time rather than immediately at time $0$. Meanwhile, I assume differentiability of $c(\cdot)$ and $H(\cdot)$ to avoid technical issues that would play a limited and pathological role in the model.

# 3 Results

## 3.1 Overview of result and proof

Intuitively, my main result is that for each $p \in \Delta(p)$, $U(p) = \mathcal{V}(p)$ i.e., expanding the DM's set of experiments does not make him outright better off. I further find that an optimal, pure-diffusion experiment exists whose precision is a function of his current belief $p_t$ and that it is almost everywhere unique. The proof consists of several steps. First, I derive a closed-form expression for the DM's Bayes-consistent beliefs based on observing $(s_t)$. I then find that the value of experimenting can be re-written as a function of his current beliefs about $x$. Next, I present a simple change of variables and characterize the optimal experiment restricted to be a pure-diffusion i.e., belong to the set of experiments feasible in MS.

I then demonstrate that $\mathcal{V}(\cdot)$ and $U(\cdot)$ satisfy the same partial differential equation (PDE) and boundary conditions, proving they are identical. Thus, if an optimal experiment solves the constrained problem, it remains optimal in the unconstrained problem. Finally, I show that since $U(\cdot)$ is a continuously differentiable function, then no optimal experiment generates information via Poisson jumps with a positive probability.

## 3.2 Belief dynamics

I first characterize the DM's Bayes consistent belief conditional on observing an experiment $(s_t)$. Fix some experiment $(s_t)$ and let $(p_t) = ((p_{it})_{i=1}^n)$ be the resulting Bayes posterior beliefs conditional on observing signal $(s_t)$ and prior $p$. Define $\mu_t \equiv \sum_{i \in X} p_{it}\mu_i$ as the time $t \geq 0$ expected drift, where $\lambda_{kt} = \sum_{i \in X} p_{it}\lambda_{ikt}$ is the expected arrival rate of type $k = 1, 2, \ldots, K$ jump and $\nu_{kt} \equiv (\nu_{ikt} \equiv p_{it^-}\lambda_{ikt^-}/\lambda_{kt^-})$ is the Bayes posterior belief when such type of jump arrives and $p_{t^-} = \lim_{\Delta_t \to 0} p_{t-\Delta_t}$ i.e., the DM´s belief held right before the jump arrived. I can now characterize the evolution of beliefs.

**Lemma 3.1.** *Fix an experiment $s = (s_t)$, a function $f : \Delta(X) \to R \in C^2$, and let $(p_t)$ be the $s$-adapted Bayes consistent beliefs. Then $(p_t)$'s infinitesimal generator is*

$$\mathcal{L}f(p_t) = \frac{h_t}{2} \times \sum_{x_i x_j \in X} p_{it}p_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t)f_{ij}(p_t) + \sum_k \lambda_{kt}[f(\nu_{kt}) - f(p_t) - \nabla f(p_t)'(\nu_{kt} - p_t)]. \quad (2)$$

Intuitively, the infinitesimal generator associated with the belief process $(p_t)$ refers to the expected marginal change in a function that depends on beliefs. This is a useful result, because $(s_t)$ and $(p_t)'$s natural filtratios are the same.

I now sketch the proof, but delegate the derivation to appendix B.1. Since each experiment $(s_t)$ is a jump-diffusion, then standard filtering theory implies that for each state $x_i$, the change in beliefs is

$$\mathrm{d}p_{it} = p_{it}(\mu_i - \mu_t)\mathrm{d}\bar{B}_t + \sum_k (\nu_{kit} - p_{it})d\bar{N}_{kt} \quad (3)$$

where $(\bar{B}_t)$ is a Brownian motion and for each $k = 1, \dots, K$ the process $(\bar{N}_{kt})$ is a compensate jump process with arrival rate $\lambda_{it} \geq 0$. The result simply applies Itô's lemma to the process $(p_t)$ given some twice-continuously differentiable function $f(\cdot)$. In the appendix, I provide a more direct proof in which $(s_t)$ is approximated in discrete-time as a the sum of $K + 1$ independent, binomial trees: one binomial tree approximates the diffusion and $K$ binomial trees approximate each of the compensated jump processes.

I conclude this subsection by making a clarification. Since the resulting dynamics for $(p_t)$ are standard, one is tempted (as in Zhong 2022) to start with $(p_t)$ as the primitive object. This is ill-advised, becuase taking $(p_t)$ as the primitive obfuscates the trivial fact that $(h_t) \gg 0$ in order for the signal process $(s_t)$ to be well-defined. In contrast, assuming that $\lambda_{ikt} = 0$ for some $i, k$, and $t$ does not impede $(s_t)$ from being well-defined. These observations clarify that the mere act of controlling a signal $(s_t)$ forces the DM to always learn from the diffusion but is not required to do the same with the jump process: an observation obfuscated by the belief approach to information design problems.

## 3.3 Cost Reformulation

Now that the beliefs dynamics are defined, I derive the Hamilton-Jacobi-Bellman (HJB) that the decision-maker's problem solves as well as several auxiliary results of note. First, I characterize costs. Fix some admissible experiment $(s_t)$ and let the Bayes consistent belief from observing such experiment be $(p_t)$. Then the flow amount of information generated at time $t$ is $I_t \equiv \mathcal{L}H(p_t)$:

$$\mathrm{I}_t \equiv \overbrace{\frac{h_t}{2} \times \sum_{ij} p_{it}p_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t)H_{ij}(p_t)}^{\text{Diffusion terms}} + \underbrace{\sum_k \lambda_{kt}[H(p_t) - H(\nu_{kt}) + \nabla H(p_t)'(\nu_{kt} - p_t)]}_{\text{Poisson terms}} \tag{4}$$

The equation above—alongside the fact that flow costs are just $c(I_t)$—makes the most important observation. It states that the diffusion and Poisson terms enter the total flow of information generated in an additively separable fashion. This happens to be one of the crucial steps in the main result.

## 3.4 DM's Value Function

I now derive an expression for the HJB describing the decision-maker's optimal experimentation problem. Since the filtration generated by $(p_t)$ has a natural one-to-one correspondence with the filtration generated by $(s_t)$ and $(p_t)$ satisfies the strong Markov property, then the value of experimentation can be written as just a function of the DM's current belief: $\mathcal{V}(p_t)$. Next, at each belief $p_t \in \Delta(X)$ such that the decision-maker strictly benefits from experimenting (i.e., $\mathcal{V}(p_t) > F(p_t)$), he picks parameter values $\phi_t = (h_t, (\lambda_{ikt}))$ to solve

$$r\mathcal{V}(p_t) = \max_{\phi_t} \mathcal{L}\mathcal{V}(p_t) - c(\mathrm{I}_t) \tag{5}$$

subject to $h_t > 0$ and for each $x_i$ and $k$, $\lambda_{ikt} \geq 0$.

For clarification, the DM can always stop experimenting at any time, so $\mathcal{V}(p_t) \geq F(p_t)$. This observation implies that the DM only experiments when $\mathcal{V}(p_t) > F(p_t)$ and stops as soon as $\mathcal{V}(p_t) = F(p_t)$. Such last condition is the boundary condition known as value matching. Lastly, Oksendal and Sulem (2019) establish that the HJB equation admits a viscosity solution, because DM's problem reduces to picking a locally Lipschitz collection of parameter process. This HJB equation, however, is far too general to allow for a tractable characterization.

## 3.5 Change of Variable

I consider a useful change of variables. Define, at each time $t$, the flow amount of Poisson information generated given a jump process $k = 1, \ldots, K$ (call it $j_{kt}$) as

$$j_{kt} \equiv \lambda_t[H(p_t) - H(\nu_{kt}) - \nabla H(p_t) \cdot (p_t - \nu_{kt})] \tag{6}$$

and the flow amount of Gaussian information as $\beta_t$ where

$$\beta_t \equiv \frac{h_t}{2} \times \sum_{ij} p_{it} p_{jt} (\mu_i - \mu_t)(\mu_j - \mu_t) H_{ij}(p_t). \tag{7}$$

The total amount of information is then $I_t = \beta_j + \sum_k j_{kt}$ and the generator for each function $f : \Delta(X) \to \mathrm{R} \in C^2$ becomes $\mathcal{L}f(p_t) = \beta_t L(f, p_t) + \sum_k j_{kt} G(f, p_t, \nu_t)$ where

$$L(f, p_t) \equiv \frac{\sum_{ij} f_{ij}(p_t) p_{it} p_{jt} (\mu_i - \mu_t)(\mu_j - \mu_t)}{\sum_{ij} H_{ij}(p_t) p_{it} p_{jt} (\mu_i - \mu_t)(\mu_j - \mu_t)}$$

and

$$G(f, p_t, \nu_{kt}) \equiv \frac{f(\nu_{kt}) - f(p_t) - \nabla f(p_t)'(\nu_{kt} - p_t)}{H(\nu_{kt}) - H(p_t) + \nabla H(p_t)'(p_t - \nu_{kt})}.$$

Intuitively, rather than picking $\phi_t = (h_t, (\lambda_{ikt}))$, the decision-maker picks $\hat{\phi} = (\beta_t, (j_{kt}, \nu_{kt}))$ i.e., the flow amount of information generated from the diffusion (i.e., $\beta_t$), the flow amount of information generated from the compensated jump process $k$ (i.e., $j_{kt}$), and his belief conditional on a jump. The only salient restriction that must be made is that $(\beta_t) \gg 0$, because $(h_t) \gg 0$ and strict convexity of $H(\cdot)$ implies that $\sum_{ij} p_{it}p_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t)H_{ij}(p_t) > 0$ for each $p_t$. Likewise, $j_{kt} \geq 0$ and $\nu_{kt} \in \Delta(X)$ due to the fact that $\lambda_{kit} \geq 0$ for each $x_i$ and $\nu_{kt}$ is a belief. I can now re-state the DM's HJB equation for beliefs $p_t$ such that $F(p_t) < \mathcal{V}(p_t)$ as

$$r\mathcal{V}(p_t) = \max_{(j_{kt},\nu_{kt}),\beta_t} \beta_t L(\mathcal{V}, p_t) + \sum_k j_{kt} G(\mathcal{V}, p_t, \nu_{kt}) - c\Big(\beta_t + \sum_k j_{kt}\Big) \tag{8}$$

s.t. $\beta_t > 0, \forall k, j_{kt} \geq 0, \nu_{kt} \in \Delta(X)$.

Notice that $\beta_t$, $j_{kt}$, and $j_{k't}$ are additively separable for each $k, k' = 1, 2, \ldots, K$. Thus, they are perfect substitutes, which is the key technical assumption needed to ensure the paper's main result.[5]

## 3.6 Payoff from acquiring Poisson Information

I now characterize the value for acquiring Poisson information. Note that $\nu_t$ only enters the problem in $G(V, p_t, \nu_{kt})$. Therefore, $\nu_{kt}$ need only solve

$$G(p_t) = \max_{\nu_{kt}\in\Delta(X)} G(V, p_t, \nu_{kt}) = \max_{\nu_{kt}\in\Delta(X)} \frac{\mathcal{V}(\nu_{kt}) - \mathcal{V}(p_t) - \nabla\mathcal{V}(p_t)'(\nu_{kt} - p_t)}{H(\nu_{kt}) - H(p_t) + \nabla H(p_t)'(p_t - \nu_{kt})} \tag{9}$$

Observe that if $\mathcal{V}(p_t)$ is continuously differentiable at some belief $p_t$, then the standard envelope condition holds and it holds that for each $x_i$,

$$\begin{aligned} G_i(p_t) = &\frac{-\mathcal{V}_i(p_t) + \mathcal{V}_i(p_t)}{H(\nu_{kt}) - H(p_t) + \nabla H(p_t)'(p_t - \nu_{kt})2} \\ &- [-H_i(p_t) + H_i(p_t)] \times \frac{\mathcal{V}(\nu_{kt}) - \mathcal{V}(p_t) - \nabla\mathcal{V}(p_t)'(\nu_{kt} - p_t)}{[H(\nu_{kt}) - H(p_t) + \nabla H(p_t)'(p_t - \nu_{kt})]^2} = 0 \end{aligned} \tag{10}$$

Heuristically, if the value function happens to be differentiable in an open, convex set, then the payoff from acquiring Poisson information (i.e., $G(p_t)$) is constant for every belief belonging to the set in question. I now summarize this result in the corollary below.

[5] It turns out that this observation is not enough to prove uniqueness when $n = 2$ among experiments with Markov controls.

**Corllary 3.2.** *Suppose that for some open, convex set $P \subset \Delta(X)$ it holds that for each $p \in P, \mathcal{V}(p) > F(p)$ and $\mathcal{V}(p)$ is continuously differentiable. Then $G(\cdot)$ is constant for each $p \in P$.*

It will turn out that $\mathcal{V}(\cdot)$ will end up being twice-continuously differentiable and that the result above results in no optimal experiment generating Poisson information in some set $P$ satisfying the conditions described above if $\#P > 1$. Consequently, the optimal experiment (if it exist) will be the unique experiment.

## 3.7 Technical Condition.

I now provide a technical condition ensuring that a solution beliefs are well-defined. Define for each $p \in \Delta(X)$ the function $K(p) \equiv \sum_{ij} p_i p_j (\mu_i - \mu(p))(\mu_j - \mu(p)) H_{ij}(p)$ where $\mu(p) \equiv \sum_i p_i \mu_i$, then the assumption goes as follows.

**Assumptions 3.3** (Boundedness from below)**.** *For each belief $p \in \Delta(X)$ such that $p \gg 0$, assume that there exist some $\epsilon > 0$ such that $\epsilon \leq K(p)$. In addition, for every sequence $(p_n) \in \Delta(X)$ such that $p_n \gg 0$ and $p_n \to p \in \partial\Delta(X)$, it holds that $\lim_{n\to\infty} 1/K(p_n) = 0$.*

I now clarify the role played by this assumption. Assume that the decision-maker manages a pure-diffusion experiment, then his time $t$ belief that $x = x_i$ is

$$\mathrm{d}p_{it} = p_{it}(\mu_i - \mu_t)\sqrt{h_t}\mathrm{d}\bar{B}_t \tag{11}$$

for $\mathrm{d}\bar{B}_t = \sqrt{h_t}[\mathrm{d}s_t - \mu_t \mathrm{dt}]$ is a Brownian motion. The equation above comes from theorem 9.1 in Lipster and Shiryaev (1977) and if one replaces $h_t$ for $\beta_t$ as the DM's control process, it holds that

$$\mathrm{d}p_{it} = p_{it}(\mu_i - \mu_t)\sqrt{\frac{2\beta_t}{K(p_t)}}\mathrm{d}\bar{B}_t. \tag{12}$$

Assumption 3.3 ensures that for any bounded, control process $(\beta_t)$, the resulting beliefs are well-defined.

## 3.8 Restricted Problem

I now characterize the optimal experiment when the decision-maker is limited to managing a pure-diffusion signal. This is a direct modification of Moscarini and Smith's (1998, 2002) optimal experiment, but the pure-diffusion experiment remains optimal even when the decision-maker can generate information via jumps.

This section focuses on the binary action and terminal action case, which simplifies the presentation of key arguments. It also allows me to computationally solve the restricted experiment and compare it to MS's solution.

**Binary action and state case** I begin by examining the simplest case of interest: $n = \#A = 2$. For each state $x_i$ , assume there is a unique optimal action $a_i$ such that for any pair of states $x_i$ and $x_j$ ($i \neq j$ ), it holds that $a_i \neq a_j$ . This ensures that learning directly affects the decision-maker's optimal choice.

For a fixed state $x_i$ , since the state can only take two values, the decision-maker's belief at time t can be represented as the probability that $x = x_i$ . The prior belief, denoted as $p \in (0, 1)$ , evolves over time $t \geq 0$ based on the observed signal path $\{s_\tau \mid \tau \in [0, t]\}$ , with the belief at time $t$ denoted by $p_t \in [0, 1]$.[6] Initially, $p_0 = p$ , and the belief evolves according to equation 12.

$$\mathrm{d}p_t = \mathrm{sgn}(\xi)\sqrt{\frac{2\beta_t}{H''(p_t)}}\mathrm{d}\bar{B}_t \tag{13}$$

where $\xi \equiv \mu_i - \mu_{-i}$, $\mathrm{sgn}(\xi) = 1$ if $\xi > 0$, and $\mathrm{sgn}(\xi) = -1$ if $\xi < 0$. This is a re-writing of equation 12 using the observation that $K(p_t) = H''(p_t)[\xi p_t(1 - p_t)]^2/2$. Further note that assumption 3.3 implies that as $p_t \to \{0, 1\}$, $dp_t \to 0$ for each bounded $(\beta_t)$ process.

Following the appendix B proof in Moscarini and Smith (1998), the current belief $p_t$ serves as a sufficient statistic for the entire signal path $\{s_\tau : \tau \in [0, t]\}$. Focusing on Markov controls based on current beliefs is without loss, because there exists a one-to-one

[6] The case when $p \in \{0, 1\}$ is trivial and uninteresting since the DM would be certain about the state and the experimentation problem would be moot.

correspondence between the natural filtration generated by $(s_t)$ and the filtration generated by $(p_t)$ . Therefore, it is equivalent to work with the belief process $(p_t)$ instead of the signal process $(s_t)$, and the change of variables does not affect the analysis.

In addition, the conditions imposed on $c(\cdot)$ and $H(\cdot)$ (i.e., $c(\cdot) \geq 0$ and assumption 3.3) ensure that the rest of the argument in Moscarini and Smith (1998)' appendix $B$ still follow. This implies that the value of experimenting and the optimal control can be written as a function of the current belief $p_t$ as $U(p_t)$ and $\beta(p_t)$, respectively. The control further solves

$$rU(p) = \max_{\beta > 0} \frac{U^{"}(p)}{H^{"}(p)}\beta - c(\beta). \tag{14}$$

A standard, interior first order condition (FOC) is a necessary condition and it states that $c'(\beta) = \frac{U^{"}(p)}{H^{"}(p)}$. By the principle of optimality (PO), it holds that for each belief $p \in [0,1]$ such that $U(p) > F(p)$ it must be the case that

$$rU(p_t) = \beta_t c'(\beta_t) - c(\beta_t) \tag{15}$$

Notice that $\frac{\mathrm{d}}{\mathrm{dx}}[xc'(x) - c(x)] = xc^{"}(x) > 0$, so $xc'(x) - c(x)$ admits an inverse function. Call such function $f(\cdot)$. As a function of the value of experimenting (i.e., $u = U(p)$), it thus holds that $\beta(u) = f(ru)$. Conversely, MS found that, as a function of the value of experimenting, $h(u) = f(ru)$.

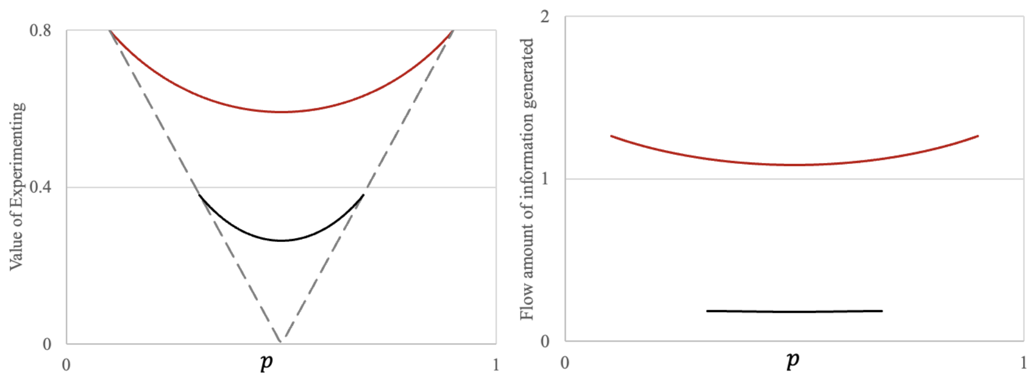
The main difference between both settings is that the value of experimenting differs between both settings. Let the value of experimenting given a belief $p \in [0,1]$ be $V(p_t)$. Then for each belief $p$ such that $V(p) > F(p)$, $V(\cdot)$ solves the ordinary differential equation (ODE)

$$rV(p_t) = V^{"}(p_t)\left[\frac{f[rV(p_t)][\xi p_t(1-p_t)]^2}{2}\right] - c[f[rV(p_t)]].$$

In contrast, I plug in the policy function $\beta(u) = f(ru)$ into the current setting's HJB equation and find that when for some current belief $p_t$ it holds that $U(p_t) > F(p_t)$, then

$$rU(p_t) = U"(p_t) \times \frac{f[U(p_t)]}{H"(p_t)} - c[f[U(p_t)]]. \tag{16}$$

In both cases, however, the boundary conditions are the same. Nevertheless, since these functions solve (in general) different ODEs, the value of experimenting differs between the two settings. Figure 1 right-hand panel plots $V(\cdot)$ (black), $U(\cdot)$ (red), and $F(\cdot)$ (dashed gray) as a function of belief $p$ for the case when $-H(\cdot)$ is entropy, $c(x) = x^2/2$, $r = 0.01$, $\xi = 1$, and the DM wishes to match the state i.e., $A = \{x_1, x_2\}$ and $u(a, x_i) = 1$ if $a = x_i$ and $-1$ is $a \neq x_i$. In contrast, the left-hand panel plots the flow amount of Gaussian information generated in each setting: black in the MS setting and red in the current setting. In the current setting, the DM stops experimenting when he is less certain about the state than in MS' case. Additionally, he generates less information in my setting in comparison to what he would have generated had costs depended of signal precision.




*(a)* ***Value functions*** *(b)* ***Policy functions***

*Figure 1:* ***Optimal policy and value function in the MS (black) and current (red) setting.***

**Many states and actions case** Next, I characterize the case where $|A|, n \geq 2$ but both remain finite. The argument presented above remains intact with the exception that now beliefs follow the expression stated in equation 12. The function $U(\cdot)$ then satisfies for each belief $p \in \Delta(X)$ where $U(p) > F(p)$ the HJB equation

$$rU(p) = \max_{\beta>0} = L(U, p_t)\beta - c(\beta) \tag{17}$$

Plugging in the FOC into the HJB above still implies that $rU(p_t) = c'(\beta_t)\beta_t - c(\beta_t)$, so the optimal flow amount of Gaussian information as a function of the value of experimenting $u$ is still equal to $\beta(u) = f(ru)$. Once again, if for some belief $p$, it holds that $U(p_t) > F(p_t)$, then

$$rU(p) = f[rU(p_t)]L(U, p_t) - c[f(rU(p))] \tag{18}$$

This value of experimentation above and since $\beta_t = f[U(p_t)]$ this completes experiment's characterization.

# 4 Main Result

I now present the main result.

**Theorem 4.1.** $\forall p_t \in \Delta(X), \;\; U(p_t) = \mathcal{V}(p_t)$.

The proof goes as follows. First, the DM's optimal choice solves equation 8. The FOCs remain necesarry conditions and they state that $c'(I_t) = L(\mathcal{V}, p_t)$ where $I_t \equiv \beta_t + \sum_k j_{kt}$. And if at time $t$, the DM picks $j_{kt} > 0$, then it further satisfies the interior FOC stating that $c'(I_t) = G(\mathcal{V}, p_t, \nu_{kt}) = G(p_t)$. In other words, for each $k = 1, \ldots, K$ either $j_{kt} = 0$ or $j_{kt} > 0$ and $G(\mathcal{V}, p_t, \nu_{kt}) = L(\mathcal{V}, p_t)$. Hence, it is always the case that

$$\beta_t L(\mathcal{V}, p_t) + \sum_k j_{kt} G(\mathcal{V}, p_t, \nu_{kt}) = I_t L(\mathcal{V}, p_t) = I_t c'(I_t). \tag{19}$$

Next, if at some belief $p$, it holds that $\mathcal{V}(p) > F(p)$, then PO implies that $r\mathcal{V}(p_t) = I_t c'(I_t) - c(I_t)$. As a consequence, $I_t = I(\nu) \equiv f(r\nu)$. Lastly, if we plug in thefact that $I_t = f[r\mathcal{V}(p_t)]$, we conclude that

$$r\mathcal{V}(p_t) = f[r\mathcal{V}(p_t)]L(\mathcal{V}, p_t) - c[r\mathcal{V}(p_t)] \tag{20}$$

Note that equations 20 and 18 imply that $U(\cdot)$ and $\mathcal{V}(\cdot)$ solve the same partial differential equation (PDE). In both cases the boundary condition is the same i.e., value matching with the value of making a decision. Since the solution to a PDE is almost surely unique, then at each belief $p_t \in \Delta(X), U(p_t) = \mathcal{V}(p_t)$. This concludes the proof.

## 4.1 Uniqueness

Next, I claim that (with probability one) the previously characterized, pure-diffusion experiment is the unique, optimal experiment. The argument has two parts. First, let $C \equiv \{p \in \Delta(X) : \mathcal{V}(p) > F(p)\}$. I claim that, for each belief $p \in C$, $L(\mathcal{V}, p) > 0$. Suppose, for contradiction, that there exists some $p \in C$ such that $L(\mathcal{V}, p) \leq 0$. Then PO and equation 8 implies that $r\mathcal{V}(p) \leq L(\mathcal{V}, p)I(p)$ or that $\mathcal{V}(p) \leq 0$. This is a contradiction, because I assumed that $p \in C$ and hence $\mathcal{V}(p) > F(p)$ and $F(p) > 0$. I summarize this observation in a corollary below.

**Corllary 4.2.** *For every $p \in C$, $L(\mathcal{V}, p) > 0$.*

Next, I claim that there does not exists an optimal experiment for which the DM strictly prefers acquiring Poisson information with a strictly positive probability. Suppose, for contradiction, that there exists some optimal experiment such that the DM acquires Poisson information when his current belief $p$ lies in a non-empty, open set $P \subset C$. By corollary 3.2 and the fact that $\mathcal{V}(\cdot)$ is twice continuously differentiable, $G(p)$ is constant for each $p \in P$. The FOCs derived above and the fact that $I(p) = f[r\mathcal{V}(p)]$ implies that $G(p) = c'[f(r\mathcal{V}(p))]$. Since $c'[f(\cdot)]$ is one-to-one, then $r\mathcal{V}(p) = f^{-1}[c^{'-1}[G(p)]]$ i.e., $\mathcal{V}(\cdot)$ is constant on $P$. Therefore, $L(\mathcal{V}, p) = 0$ and this contradicts the assumption that $P$ is a subset of $C$. I now state the result below.

**Lemma 4.3.** *There does not exist an optimal experiment for which the DM obtains Poisson information if his current belief about the state belongs to some non-empty, open set $P \subset C$.*

This result implies that the probability that an optimal experiment generates a Poisson jump is $0$.

# 5 Conclusion

This paper builds on Moscarini and Smith's (2001) work on sequential experimentation by greatly extending the set of feasible experiments. I find that forcing the DM to experiment qualitatively changes the type of experiment that he chooses to obtain relative to the case when he can acquire it from an external source i.e., as in Zhong (2022). In fact, the type of information acquired precisely the opposite of each other.

This result serves as a caution against the belief-based approach in information design. The flexibility afforded by such approach, although elegant, may not account for trivial yet crucial restrictions based by real economic agents. Therefore, future work on information economics ought to consider whether its predictions are robust to, at the very least, basic data generation restrictions. Otherwise, the great recent advances in this literature may not find an application in future research.

# References

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

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

[3] Denti, Tommaso. 2022. "Posterior Separable Cost of Information." AER 112(10): 3215–59.

[4] N. El Karoui and I. Karatzas. 1994. Dynamic allocation problems in continuous time. *Annals of Applied Probability* 4(2):255–286.

[5] Moscarini, Giuseppe, and Lones Smith. 2001. "The Optimal Level of Experimentation". Econometrica 69(6): 1629-1644.

[6] ————-. 1998. "Wald Revisited: The Optimal Level of Experimentation". Cowles Foundation Discussion Papers.

[7] Princeton Review. 2024. *A Day in the Life of a Chemist*. Retrieved from: https://www.princetonreview.com/careers/34/chemist.

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

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

[10] Pomatto, Luciano, Philipp Strack, and Omer Tamuz. 2020. “The Cost of Information: The Case of Constant Marginal Costs.” American Economic Review 113(5): 1360–93.

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

[12] Redman, Thomas C. 2016. ”Bad Data costs the US $ 3 trillion per year”. Harvard Business Review retrieved from: $https : //dataladder.com/wp - content/uploads/2019/07/Bad - Data - Costs - the - U.S - 3 - Trillion - Per - Year.pdf$.

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

[14] Steiner, Jakub S., Colin Stewartm Filip Matéjka. 2017. ”Rational Inattention Dynamics: Inertia and Delay in Decision-making”. Econometrica 85(2): 521-553.

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

[16] Webster, Evan. 2023. *FP&A Professionals Only Spend 25% of Their Time on Analysis*. Retrieved from: https://www.venasolutions.com/blog/time-spent-on-analysis.

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

# A Precision vs. Informational Cost comparison.

**Who experiments for longer?** The numerical examples above beg the following question: how does the optimal experiment changes when one shifts from precision to information-based costs? Fix a function $c(\cdot)$ and assume that $H(p) = p \ln p + (1-p)\ln(1-p)$ (i.e., - entropy). Assuming that information is measured as the expected reduction in entropy is standard in the rational inattention literature and happens to greatly simplify the subsequent result. I find a sufficient condition for when a decision-maker facing information-based costs acquires more information and makes a decision when he is better informed about the state than a peer facing precision-based costs.

I assume that all model factors remain the same except that in case **1** cost depend on the signal precision whereas in case **2** it depends on the flow amount of information generated. Define the sets $C_1 \equiv \{p \in [0,1] : V(p) > F(p)\}$ and $C_2 \equiv \{p \in [0,1] : U(p) > F(p)\}$ i.e., the set of beliefs for which the decision-maker benefits from experimenting in case **1** and **2**, respectively. In addition, define $h_j(\cdot)$, for $j = 1, 2$, as the optimal signal precision in each case $x_j$ as a function of current belief $p$. I can now state the result.

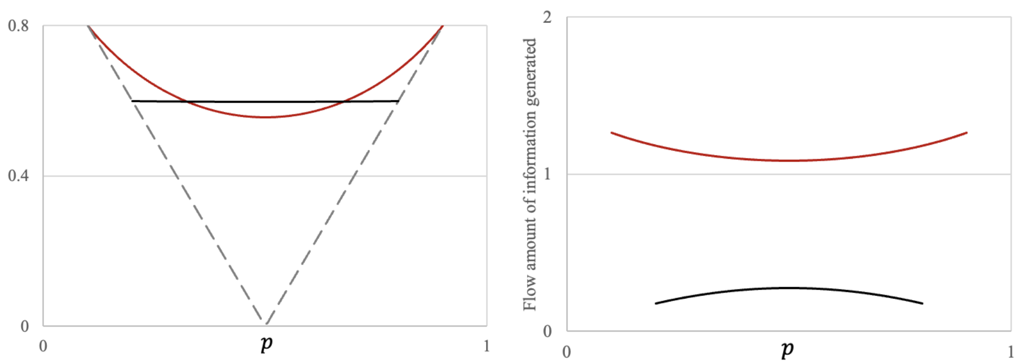
**Lemma A.1.** *Suppose that $|\xi| < \sqrt{2}$, then $C_1 \subset C_2$ and for each $p \in C_1, h_1(p) < h_2(p)$.*

I now provide the proof. First, I claim that both results follow from the observation that for each belief $p$, $U(p) \geq V(p)$. To see this, fix some belief $p$ such that $V(p) > F(p)$ i.e., $p \in C_1$. Since $V(p) \leq U(p)$, then $U(p) \geq V(p) > F(p)$, so $p \in C_2$. Therefore, $C_1 \subset C_2$. Next, at for each $p \in C_1$, it holds that $h_1(p) = f[V(p)] < f[U(p)] = h_2(p)\xi^2 p(1-p) < h_2(p)$. Observe that this argument uses the fact that $f(\cdot)$ is strictly increasing. Lastly, I provide the following lemma that concludes the proof above.

**Lemma A.2.** *Suppose that $|\xi| < \sqrt{2}$, then for each belief $p \in C_1$,$U(p) \geq V(p)$.*

The proofs intuition goes as follows. The assumption that information is measured as the reduction in entropy and $\xi^2 < 2$ implies that it is cheaper to pick any control $h(\cdot)$ when facing information rather than precision based costs. This implies that the decision-maker's maximum payoff from experimenting when he faces precision-based costs would be strictly higher if he faced information-based costs. But those controls would not be optimal even though they were feasible, so the maximum payoff from experimenting when the decision-maker faces information-based costs are higher than the maximum payoff when costs depend on signal precision.

I conclude this subsection by noting, in a numerical example, what happens when $\xi^2 > 2$. I consider the same setting as illustrated in figure 1 but now $\xi = 10$. Figure 2 plots the new value and policy functions. It is worthwhile noting that neither setting nets the decision-maker strictly higher value from experimenting than the other. In addition, the flow amount of information generated when costs depend on signal precision is now a single-peaked function of the current belief $p$. In contrast, when $\xi^2 < 2$, the aforementioned policy function was u-shaped.



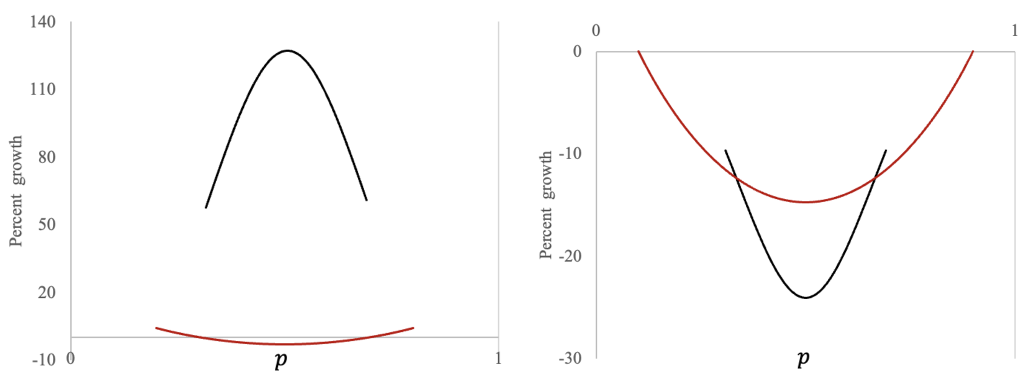

*(a)* ***Value functions*** *(b)* ***Policy functions***

*Figure 2:* ***Optimal policy and value function in the MS (black) and current (red) setting with matching the state terminal payoffs and*** $\xi^2 > 2$.

**Comparative, comparative statics** The numerical exercises and results presented above point out how shifting the cost structure qualitatively changes how the decision-maker chooses to experiment. I now conclude these numerical exercises by illustrating how the comparative statics qualitatively differ as one changes the cost structure.

Assume that the cost function takes on the functional cost $c(x) = \frac{x^{1+\frac{1}{\psi}}}{1+\frac{1}{\psi}}$ for some $\psi > 0$. The parameter $\psi$ models the curvature of the cost function. In the left-hand panel of figure 3, I plot the percent change in the policy function when one lowers $\psi$ from 1 to $1/4$. If costs depend on information precision, then lowering $\psi$ prompt the decision-maker to acquire more information for every belief. When costs depend on the flow amount of information generated, however, the magnitude and the direction of the change in the policy functions can be reversed.

Next, I consider lowering $r$ from $0.01$ to $0.005$ i.e., making the decision-maker more patient. The resulting change in policy functions are illustrated in the right-hand panel of figure 3. In both cases, making the decision-maker lowers the flow amount of information generated as a function of belief $p$. However, the decline (in percentage terms) is more significant when costs depend on signal-precision rather than on information generated. These comparative statics are important, because they clarify that changing the cost structure not only leads to different model predictions but also illustrate predictions that can be tested experimentally. This means that an experiment changing the $\psi$ or $|\psi|$ enable a researcher to discriminate between the two types of cost structures.




(a) **Lowering $\psi$.** (b) **Lowering $r$.**

*Figure 3:* ***Change in the optimal policy function as one lowers $\psi$ and $r$ when costs depend on flow information generated (red) or in the signal's precision (black).***

# B Proofs

## B.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 $\Delta_t > 0$, then an admissible signal $s = (s_t)$ can be approximated at times $t = 0, \Delta_t, \ldots$ as $s_0 = 0$ and $\mathrm{d}s_t \equiv s_{t+\Delta_t} - s_t$ $\mathrm{d}s_t = \mathrm{d}_{t\Delta_t} + \sum_k \mathrm{d}J_{kt\Delta_t}$ for $(\mathrm{d}_{t\Delta_t})_{t=0}^{\infty}$ is a sequence of independent random variables such that at time $t$, $d_{t\Delta_t} = \pm\sqrt{\Delta_t / h_{t-\Delta_t}}$ with probability $[1 \pm \mu_i \sqrt{h_{t-\Delta_t}\Delta_t}]/2$ iff $x = i$. Meanwhile, for each $k = 1, 2, \ldots, K$, $(J_{kt\Delta_t})_{t=0}^{\infty}$ is a sequence of independent random variables such that $\mathrm{d}J_{kt\Delta_t} = \Delta_k$ with probability $\lambda_{ikt-\Delta_t}\Delta_t$, $\mathrm{d}J_{kt\Delta_t} = 0$ otherwise, and the $\mathrm{d}J_{kt\Delta_t}$ and $\mathrm{d}J_{k't\Delta_t}$ are independent for each $t, k, k'$ . This means that the independent diffusion and jump processes are weakly approximated. Also, the decision-maker picks the parameters for realizations observed at time $t = 0\Delta_t, 2\Delta_t, \ldots$ at time $t - \Delta_t$.

**Approximating beliefs after a jump** I first consider the case when there are jumps. Suppose that heheld beliefs $p_{t-\Delta_t} = (p_{it-\Delta_t})$, then the Bayes posterior belief that $x = i$ given the jump is approximately equal to

$$\nu_{ikt} = \frac{p_{it-\Delta_t}\Delta_t\lambda_{ikt-\Delta_t}}{\sum_j p_{jt-\Delta_t}\Delta_t\lambda_{jkt-\Delta_t}} + o(\Delta_t) = \frac{p_{it-\Delta_t}\lambda_{ikt-\Delta_t}}{\sum_j p_{jkt-\Delta_t}\lambda_{jkt-\Delta_t}} + o(\Delta_t)$$

where the error term $o(\Delta_t)$ (such that $\lim_{\Delta_t \searrow 0} o(\Delta_t)/\Delta_t = 0$) follows from the observation that distribution of $d_{t\Delta_t}$ approximately gives equal weight to both outcomes as $\Delta_t$ goes to $0$. Further observe that as $\Delta_t$ goes to $0$, it holds that $\nu_{ikt} = \frac{p_{it-}\lambda_{it-}}{\lambda_{kt-}}$ where $\lambda_{kt-} \equiv \sum_j p_{jt-}\lambda_{jkt-}$. It is further the case that the discontinuous change in beliefs is $\mathrm{d}p_{ikt} \equiv \nu_{ikt} - p_{it-\Delta_t}$ and it converges to $\mathrm{d}p_{ikt} \equiv \nu_{ikt} - p_{it-}$ as $\Delta_t \searrow 0$.

**Approximating beliefs when there is no jump** Next, I characterize how beliefs change when $J_{t\Delta_t} = 0$. Suppose that the decision-maker observes $\mathrm{d}s_t = \pm\sqrt{\Delta_t/h_{t-\Delta_t}}$, then the probability of observing such signal realization conditional on $x = i$, for $i = 1, 2 \ldots, n$, is

$$\Pr_t(\mathrm{d}s_t = \pm\sqrt{\Delta_t / h_{t-\Delta_t}} | x = i) = \frac{1}{2}\left[1 \pm \mu_i \sqrt{h_{t-\Delta_t}\Delta_t} - \sum_k \lambda_{ikt-\Delta_t}\Delta_t\right] + o(\Delta_t)$$

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

$$\begin{aligned} p_{it} = & \frac{p_{it-\Delta_t}\left[1 \pm \mu_i \sqrt{h_{t-\Delta_t}\Delta_t} - \Delta_t \sum_k \lambda_{ikt-\Delta_t}\right]}{\sum_j p_{jt-\Delta_t}\left[1 \pm \mu_j \sqrt{h_{t-\Delta_t}\Delta_t} - \Delta_t \sum_k \lambda_{jkt-\Delta_t}\right]} + o(\Delta_t) \\ & = \frac{p_{it-\Delta_t}\left[1 \pm \mu_i \sqrt{h_{t-\Delta_t}\Delta_t} - \Delta_t \sum_k \lambda_{ikt-\Delta_t}\right]}{1 \pm \mu_{t-\Delta_t} \sqrt{h_{t-\Delta_t}\Delta_t} - \Delta_t \sum_k \lambda_{kt-\Delta_t}} + o(\Delta_t) \end{aligned}$$

where $\mu_t \equiv \sum_i p_{it}\mu_i$ and for each $k = 1, 2, \ldots, K$, $\lambda_{kt} \equiv \sum_{i=1}^K \lambda_{ikt}$. The change in beliefs $dp_{it} \equiv p_{it} - p_{it-\Delta_t}$ can be the approximated as

$$dp_{it} = \frac{p_{it-\Delta_t}\left[\pm(\mu_i - \mu_{t-\Delta_t})\sqrt{h_{t-\Delta_t}\Delta_t} - \Delta_t \sum_k (\lambda_{ikt-\Delta_t} - \lambda_{kt-\Delta_t})\right]}{1 \pm \mu_{t-\Delta_t}\sqrt{h_{t-\Delta_t}} - \Delta_t \sum_k \lambda_{kt-\Delta_t}} + o(\Delta_t)$$

and the probability that beliefs change by the amount described above occurs with a probability of approximately $(1 \pm \mu_{t-\Delta_t}\sqrt{h_{t-\Delta_t}\Delta_t} - \Delta_t \sum_k \lambda_{kt-\Delta_t})/2$.

**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

$$\frac{E[dp_{it}|\forall k, J_{kt\Delta_t}=0]}{\Delta_t} = \frac{p_{it-\Delta_t}}{2\Delta_t}\left[(\mu_i - \mu_{t-\Delta_t})\sqrt{h_{t-\Delta_t}\Delta_t} - \Delta_t\sum_k(\lambda_{ikt-\Delta_t} - \lambda_{kt-\Delta_t})\right]$$
$$+\frac{p_{it-\Delta_t}}{2\Delta_t}\left[-(\mu_i - \mu_{t-\Delta_t})\sqrt{h_{t-\Delta_t}\Delta_t} - \Delta_t\sum_k(\lambda_{ikt-\Delta_t} - \lambda_{kt-\Delta_t})\right] + \frac{o(\Delta_t)}{\Delta_t}$$
$$= -p_{it-\Delta_t}\sum_k(\lambda_{ikt-\Delta_t} - \lambda_{kt-\Delta_t}) + \frac{o(\Delta_t)}{\Delta_t}.$$

On the other hand, if jump $\Delta_k$ arrives, for some $k = 1, 2, \ldots, K$, then for each $i = 1, 2, \ldots, n$ it holds that $E[dp_{it}|J_{kt\Delta_t} = 1] = \nu_{ikt} - p_{it-\Delta_t}$. As a consequence, the unconditional change in beliefs just equals to the expectation over the conditional expectations i.e.,

$$\frac{E[dp_{it}]}{\Delta_t} = \sum_k \frac{\Delta_t\lambda_{kt-\Delta_t}}{\Delta_t}E[dp_{it}|J_{kt\Delta_t}=1] + \left(1-\Delta_t\sum_k\lambda_{kt-\Delta_t}\right)\frac{E[dp_{it}|\forall k, J_{kt\Delta_t}=0]}{\Delta_t}$$
$$= \sum_k(\nu_{ikt}-p_{it-\Delta_t})\lambda_{kt-\Delta_t} - \left(1-\Delta_t\sum_k\lambda_{kt-\Delta_t}\right)\sum_k(\lambda_{ikt-\Delta_t}p_{it-\Delta_t}-\lambda_{kt-\Delta_t}) + \frac{o(\Delta_t)}{\Delta_t}.$$
$$= \sum_k(\nu_{ikt}-p_{it-\Delta_t})\lambda_{kt-\Delta_t} - \left(1-\Delta_t\sum_k\lambda_{kt-\Delta_t}\right)\sum_k(\nu_{ikt-\Delta_t}-p_{it-\Delta_t})\lambda_{kt-\Delta_t} + \frac{o(\Delta_t)}{\Delta_t}$$
$$= -\Delta_t\sum_k\lambda_{kt-\Delta_t}\sum_k(\nu_{ikt-\Delta_t}-p_{it-\Delta_t})\lambda_{kt-\Delta_t} + \frac{o(\Delta_t)}{\Delta_t}.$$

The second first line is just the law of iterated expectations and the second line replaced the conditional expectations derived above and replaces them for the expressions derived above. The third line exploits the observation that $\nu_{ikt-\Delta_t} \equiv \lambda_{ikt-\Delta_t}p_{it-\Delta_t}/\lambda_{kt-\Delta_t}$ by Bayes rule. The last line simplifies the expression and yields that as$\Delta_t$ goes to $0$, the expected change in beliefs converges to $0$. Ergo, beliefs are a martingale as expected.

**Approximating the co-movement of beliefs** The next step is to calculate the co-movement of beliefs conditional on there being no jumps. This step is important, because the main result of this proof is to approximate the change of a twice continuously differentiable function over an instant i..e., the generator for the belief process induced from observing a given, admissible signal process. Pick some pair of state realizations $x_i$ and $x_j$, then

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

occurs with a probability of roughly $(1 \pm \mu_t^2 h_t \Delta_t)/2$. The equation above holds, because the additional terms get multiplied by either a term $\epsilon_1(\Delta_t) = \Delta_t^{\frac{3}{2}}$ or $\epsilon_2(\Delta_t) = \Delta_t^2$. There terms are of the magnitude $o(\Delta_t)$ (i.e., $\lim_{\Delta_t \searrow 0} \frac{\epsilon_j(\Delta_t)}{\Delta_t} = 0$ for each $j = 1, 2$), so they will not add anything to the final approximation. Taking expectations of the expected co-movement then reveals that

$$\frac{E[\mathrm{d}p_{it}\mathrm{d}p_{jt}|J_{t\Delta_t} = 0]}{\Delta_t} = h_t p_{it-\Delta_t} p_{jt-\Delta_t}(\mu_i - \mu_t)(\mu_j - \mu_t) + \frac{o(\Delta_t)}{\Delta_t}.$$

**Approximating the generator** Lastly, I approximate the generator. Let $f : \Delta(X) \to \mathrm{R}$ be a twice continuously differentiable function, then I need to calculate $\mathcal{L}f(p_t) \equiv E[df(p_t)]/\Delta_t$ as $\Delta_t$ goes to $0$ for $df(p_t) \equiv f(p_t) - f(p_{t-\Delta_t})$. I can partition the expectation by the law of iterated expectations. If there is a jump of magnitude $\Delta_k$ for some $k = 1, 2, \dots, K$, then

$$E[df(p_t)|\mathrm{d}J_{kt\Delta_t} = \Delta_k] = f(\nu_{kt}) - f(p_{t-\Delta_t}).$$

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

$$\begin{aligned}\frac{E[df(p_t)|\forall k, J_{kt\Delta_t} = 0]}{\Delta_t} &= \frac{1}{\Delta_t}E\left[\nabla f(p_t)'\mathrm{d}p_t + \frac{1}{2}\mathrm{d}p_{it} H f(p_t)'\mathrm{d}p_t|\forall k, J_{kt\Delta_t} = 0\right] + \frac{o(\Delta_t)}{\Delta_t}\\ &= \frac{1}{\Delta_t}E\left[\sum_i f_i(p_t)\mathrm{d}p_{it} + \frac{1}{2}\sum_{ij} f_{ij}(p_t)\mathrm{d}p_{it}\mathrm{d}p_{jt}|\forall k, J_{kt\Delta_t} = 0\right] + \frac{o(\Delta_t)}{\Delta_t}\end{aligned}$$

$$= \sum_i f_i(p_t) \frac{E\big[\mathrm{d}p_{it} | \forall k, J_{kt\Delta_t} = 0\big]}{\Delta_t} + \frac{1}{2} \sum_{ij} f_{ij}(p_t) \frac{E\big[\mathrm{d}p_{it}\mathrm{d}p_{jt} | \forall k, J_{kt\Delta_t} = 0\big]}{\Delta_t} + \frac{o(\Delta_t)}{\Delta_t}$$

$$= -\sum_i f_i(p_t) p_{it-\Delta_t} \sum_k (\lambda_{ikt-\Delta_t} - \lambda_{kt-\Delta_t})$$

$$+ \frac{1}{2} \sum_{ij} f_{ij}(p_t) h_t p_{it-\Delta_t} p_{jt-\Delta_t} (\mu_i - \mu_{t-\Delta_t})(\mu_j - \mu_{t-\Delta_t}) + \frac{o(\Delta_t)}{\Delta_t}$$

$$= -\sum_i \sum_k f_i(p_t)(\lambda_{ikt-\Delta_t} - \lambda_{kt-\Delta_t}) p_{it-\Delta_t}$$

$$+ \frac{1}{2} \sum_{ij} f_{ij}(p_t) h_{t-\Delta_t} p_{it-\Delta_t} p_{jt-\Delta_t} (\mu_i - \mu_{t-\Delta_t})(\mu_j - \mu_{t-\Delta_t}) + \frac{o(\Delta_t)}{\Delta_t}$$

$$= -\sum_k \lambda_{kt-\Delta_t} \nabla f(p_t)' (\nu_{kt} - p_{t-\Delta_t})$$

$$+ \frac{1}{2} \sum_{ij} f_{ij}(p_t) h_{t-\Delta_t} p_{it-\Delta_t} p_{jt-\Delta_t} (\mu_i - \mu_{t-\Delta_t})(\mu_j - \mu_{t-\Delta_t}) + \frac{o(\Delta_t)}{\Delta_t}$$

The first line states the Taylor approximation to a second degree and the second step re-states the argument in summation form. Next, the third line exploits the fact that expectations are linear operators, while the fourth line imports the approximations for these changes in value from the previous sections of the proof. The next equality regroups terms, while the last re-writes the equation above taking into account that above and replaces them for the expressions derived above. The third line exploits the observation that $\nu_{ikt-\Delta_t} \equiv \lambda_{ikt-\Delta_t} p_{it-\Delta_t} / \lambda_{kt-\Delta_t}$ by Bayes rule.

Next, I calculate $E[\mathrm{d}f(p_t)]/\Delta_t$. Since the probability of a jump of size $\Delta_k$ is approximately $\lambda_{kt-\Delta_t}\Delta_t$, then the unconditional expectation equals to

$$\frac{E[\mathrm{d}f(p_t)]}{\Delta_t} = \sum_k \frac{\lambda_{kt}\Delta_t}{\Delta_t} E[\mathrm{d}f(p_t) \mid J_{kt\Delta_t} = \Delta_k] + \left[1 - \Delta_t \sum_k \lambda_{kt}\right] \times \frac{E[df(p_t) | \forall k, J_{kt\Delta_t} = 0]}{\Delta_t}$$

$$= \sum_k \lambda_{kt} [f(\nu_{kt}) - f(p_{t-\Delta_t})] + \left[1 - \Delta_t \sum_k \lambda_{kt}\right]$$

$$\times \left[\frac{1}{2} \sum_{ij} f_{ij}(p_t) h_{t-\Delta_t} p_{it-\Delta_t} p_{jt-\Delta_t} (\mu_i - \mu_{t-\Delta_t})(\mu_j - \mu_{t-\Delta_t}) - \sum_k \lambda_{kt-\Delta_t} \nabla f(p_{t-\Delta_t})' (\nu_{kt} - p_{t-\Delta_t})\right] + \frac{o(\Delta_t)}{\Delta_t}$$

$$= \sum_k \lambda_{kt}[f(\nu_{kt}) - f(p_{t-\Delta_t}) - \nabla f(p_{t-\Delta_t})'(\nu_{kt} - p_{t-\Delta_t})]$$
$$+ \frac{1}{2}\sum_{ij} f_{ij}(p_t) h_{t-\Delta_t} p_{it-\Delta_t} p_{jt-\Delta_t}(\mu_i - \mu_{t-\Delta_t})(\mu_j - \mu_{t-\Delta_t}) + \frac{o(\Delta_t)}{\Delta_t}$$

The first equality just applies the Law of iterated expectations and the second equality replaces the conditional expectations with the approximations that were derived above. The last equality just collects terms. Lastly, take the limit as $\Delta_t$ goes to $0$ it yields that

$$\mathcal{L}f(p_t) = \sum_k \lambda_{kt}[f(\nu_{kt}) - f(p_t) - \nabla f(p_t)'(\nu_{kt} - p_t)] + \frac{h_t}{2}\sum_{ij} f_{ij}(p_t) p_{it} p_{jt}(\mu_i - \mu_t)(\mu_j - \mu_t)$$

This concludes the proof.

## B.2 Proof of Lemma A.2.

Fix some belief $p \in C_1$. Then by the PO, it holds that at each time $t \geq 0$

$$V(p) = E_t\left[e^{-(T_1 - t)}F(p_{T_1}) - \int_t^{T_1} c[h_1(p_s)]\text{ds} \mid p_t = p\right] \tag{21}$$

where at time $t$, $T_1 = \inf\{s \geq t : p_t \notin C_1\}$ i.e., the first time when the payoff from making a decision is weakly higher than the value of experimenting. Since $\xi^2 < \sqrt{2}$ and $H(p) = p\ln p + (1-p)\ln(1-p)$, then the flow amount of information generated at each belief $p \in C_1$ is $\beta_1(p) = h_1(p)\xi^2[p(1-p)]^2 H''(p) = h_1(p)\xi^2[p(1-p)] < h_1(p)$ since $H''(p) = 1/[p(1-p)]$. As a consequence, the decision-maker's payoff if he faced information based costs would be strictly higher i.e.,

$$V(p) < E_t\left[e^{-(T_1 - t)}F(p_{T_1}) - \int_t^{T_1} c[\beta_1(p_s)]\text{ds} \mid p_t = p\right] \leq U(p) \tag{22}$$

The second lines uses the fact that the set of experiments is the same in both settings, so the optimal experiment when facing precision-based costs cannot yield the decision-maker strictly higher payoffs than the maximum payoffs that a decision-maker faces precision-based costs. This concludes the proof.