Information Acquisition and the Timing of Actions*

Alkis Georgiadis-Harris†

November 23, 2021

(CLICK FOR LATEST VERSION)

Abstract

This paper develops a dynamic model of information acquisition, in which the decision maker lacks control over the timing of their action. It characterizes the optimal dynamic experiment when the decision maker can flexibly choose all relevant aspects of the information they acquire. The cost of an experiment depends on the quantity of information it produces. At the optimum, the decision maker concentrates resources in generating a single piece of breakthrough news, *contradicting* their plan of action. In the absence of such news, the decision maker becomes more confident in their intentions. This leads them to sacrifice the frequency with which breakthroughs arrive in order to increase their impact on choice behaviour. These are in stark contrast with the case where the timing of actions is endogenous, in which breakthroughs confirm beliefs, and the resolution of the trade-off between frequency and precision is reversed.

1 Introduction

Information is an essential resource whose production, dissemination, and acquisition are fundamental economic activities. The modern information age presents decision makers with increasingly more sophisticated sources, from which they can learn about their environment in a granular way. Importantly, acquiring and processing information takes time, making them inherently dynamic activities. It is therefore critical to study how agents learn over time when they are allowed to choose freely all aspects of the information they acquire. Additionally, it is essential to understand how the nature of the problem they are facing shapes the information they collect. In many environments opportunities to act are scarce and uncertain, and decision makers are deprived of control over *when* they take actions. The aim of this paper is to shed light on learning behaviour in such cases.

*I am indebted to Balázs Szentes, Francesco Nava and Vasiliki Skreta for continual guidance and support. I am also grateful to Thomas Brzustowski, Carlo Cabrera, Angelos Diamantopoulos, Andrew Ellis, Maxi Guennewig, Kostas Kardaras, Matthew Levy, William Matcham, John Moore, Ronny Razin, and Christopher Sandmann for many helpful comments and discussions.

†Department of Economics, London School of Economics, London, UK. E-mail: a.georgiadis-harris@lse.ac.uk.

In particular, this paper considers a Bayesian decision maker whose preferences over actions depend on an unknown state, and for whom a single opportunity to irreversibly act arrives at some *exogenous* random time. The decision maker has access to a rich set of dynamic experiments which they may employ to learn about the state while waiting to act.

Many economic problems have this flavour. For instance, consider a pharmaceutical company researching a new vaccine technology over a known existing technology, where they can choose to market either. Before a pandemic occurs there is little demand for their product and because of the limited term of patents, they are not eager to market it early. Importantly, the onset of pandemics cannot be foreseen and will thus try to learn about the efficacy of the new technology in anticipation of the next outbreak. What kind of research agenda do these firms pursue? How do their research strategies and R&D expenditures vary with their estimates of the likelihood of success of new technologies? How do their estimates of the arrival of pandemics affect research output? As another example, consider an analyst offering expert forecasts on the state of the economy, but who cannot anticipate when a client will show up, and will collect information about the relevant variables while waiting. How do the learning strategies of such an analyst differ from one who actively prospects clients, or regularly publicizes reports? How do the payoff consequences of the analyst's reports affect their quality? As illustrated by these examples, understanding information acquisition in models where the timing of actions is exogenous is crucial; and yet, most of the literature following Wald (1947) and Arrow, Blackwell and Girshick (1949) has focused on models where decision makers choose when to act.

Furthermore, control over the timing of actions is an important theoretical dimension along which information acquisition problems should be distinguished. Choosing 'information' amounts to choosing lotteries over beliefs. Consequently, when the timing of actions is controlled a decision maker's preferences over time-lotteries take a central role in shaping their behaviour. As has been pointed out by DeJarnette et al. (2020), standard discounting generates risk-taking behaviour on the time dimension. In the present context, Zhong (2019) establishes the qualitative features of optimal information in an endogenous timing model. The author shows that the result is driven by the temporal-risk attitudes of the decision maker in the following way: the optimal experiment generates the maximal mean-preserving spread over decision times.[1] In contrast, the exogeneity of the decision time in this paper eliminates the role of preferences over time-lotteries, and we find that most qualitative aspects of the optimal experiment *reverse*.

Ultimately, in any decision problem there will be pieces of information whose timely arrival the decision maker values more. For instance, in the endogenous timing setting the decision maker would like to know whether they should *stop* and take an action as quickly as possible. When the timing of actions is exogenous, what is of value is whether, and how, they should *change* their intended action; and when the opportunity to act is likely to be soon, they want to learn this

[1] When there are linear delay costs, the decision maker is risk-neutral over time-lotteries, and the type of experiment becomes irrelevant.

promptly. An optimal dynamic experiment will try to deliver those relevant pieces of information as fast as feasibly possible.

Model outline.— We now outline the main ingredients of the model in more detail. An opportunity to act arrives at some *exogenous* random time, and the decision maker's preferences over actions depend one of two unknown states. The future is discounted at some constant rate. Time is continuous and the decision maker can flexibly design a dynamic experiment to learn prior to acting.

A dynamic experiment consists of a standard Blackwell experiment, and an information flow, which refines knowledge over time. Such experiments can be identified by the posterior process they induce, given a prior. Subject to some regularity assumptions, any dynamic experiment can be represented by a pair of characteristics. The first captures the speed at which 'incremental information' arrives. Such pieces of information generate erratic and local changes in the beliefs. The second characteristic governs the behaviour of 'breakthroughs'—that is, pieces of information which can generate abrupt and significant changes in beliefs. The decision maker can set up dynamic experiments delivering incremental information as well as breakthroughs of varying impact, arriving at different frequencies. Consequently, the set of experiments is rich enough to accommodate a large variety of dynamic learning patterns one may wish to explore.

To describe the information costs, we first define the total quantity of information in an experiment as a moment of the final posterior distribution it implements.[2] We can then specify a notion of information increment, or *flow-information*, generated by a dynamic experiment. As in Zhong (2019) and Hébert and Woodford (2021b), we specify a capacity constraint on the flow-information the experiments are allowed to generate.[3] Critically, such a constraint forces the decision maker to 'smooth' information acquisition over time; in the absence of such motive, the optimal experiment instantaneously implements a distribution of beliefs, and dynamics are irrelevant.

Preview of Results.—To describe the main results, we define the *information gain* as the difference between the decision maker's optimal payoff—their value, and their payoff from acting without any further information. The optimal experiment is as follows. The decision maker concentrates resources in generating a single breakthrough. That is, the optimal experiment is a *single-jump* process. It is characterised at each point in time by its *frequency*—the rate at which breakthrough news arrive; its *precision*—the significance of such news, captured by the magnitude of the change in beliefs they induce; and finally, its *direction*—whether breakthrough news confirm or contradict the decision maker's prior conviction. Close to the location of maximal information gain, the optimal experiment has high frequency and low precision. In the limit as the belief approaches the location

[2] In the tradition of the literature on rational inattention, we consider a uniformly posterior separable measure of information.

[3] Analyzing a problem with some smooth convex cost function applied to the information increment is similar to the constrained problem where the capacity is endogenized.

of maximal information gain, frequency explodes and precision implodes, producing purely incremental information. As the information gain decreases, the frequency of breakthroughs decreases in favour of more precision. Furthermore, the jumps in beliefs induced by breakthrough news are in the *direction of increasing information gain.* They are sufficiently precise so that they capture a substantial part of the potential gains from learning. That is, along a sequence of breakthroughs the information gain decreases.

Example.—To illustrate these results and facilitate a comparison with the endogenous timing models of Zhong (2019), and Hébert and Woodford (2021b), we consider the following example. An analyst learns about the state of the economy which can be either 'good' or 'bad.' An investor in need of advice arrives at some exogenous rate. The analyst is positively compensated if they correctly advise to invest in the good state, or if they correctly advise to avoid investing in the bad state. We will call the probability they assign to the state being good, their belief. Payoffs are given by the matrix:

	good	bad
invest	1	-1
not invest	-1	1

The left panel in Figure 1 below plots the value V (solid) and the expected payoff to the analyst U (dotted) if they were to quit acquisition forever. The information gain is the gap $V - U$ between the two, and it is maximized at 1/2.

To the right of 1/2 the information gain is decreasing. Consequently, optimal jumps are in the opposite direction than the prior—that is, they are *contradictory.* Moreover, they are sufficiently precise so that they land at a level where $V - U$ is lower than before the jump (grey areas). Importantly, as the analyst becomes increasingly more confident in the state, and they substitute resources away from frequency and into precision. The symmetric properties apply to the left of 1/2.

Now consider the problem where the analyst can actively issue an investment recommendation, which investors use and based on which they are compensated as above, and we assume they discount the future. In this setting, they control the timing of their action. The value, V (solid), and the stopping payoff U (dotted), for that problem are depicted in the right panel of Figure 1. The optimal experiment has the following features. At the threshold belief 1/2, a signal arrives at some time, independent of the state, and implements a distribution supported at points within the stopping region. To the right of 1/2, a single-jump process is optimal, whose jumps are in the direction of increasing *value.* That is, they are always *confirmatory.* Moreover, they are sufficiently informative to warrant immediate action upon arrival. Finally, as the analyst becomes more confident in the state, they are acquiring increasingly more frequent signals, at the expense of precision. Again, the symmetric properties apply to the left of 1/2.

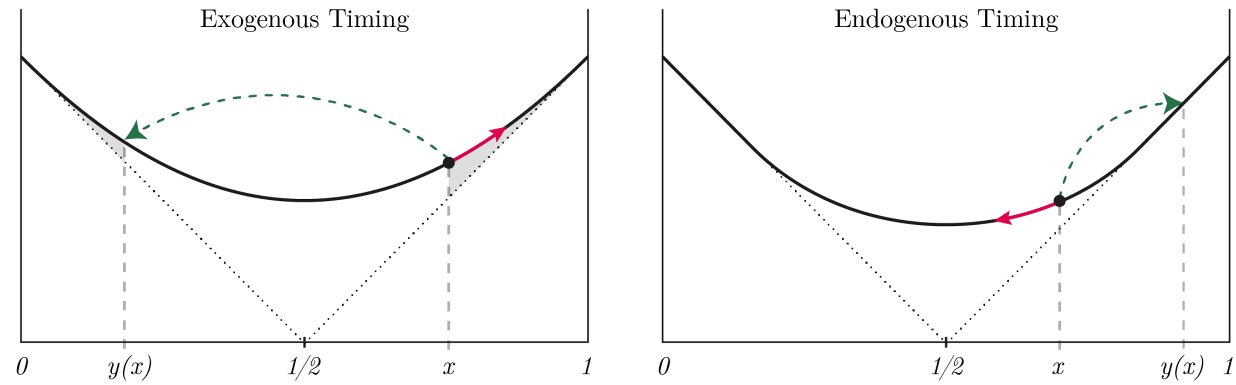

Figure 1: The horizontal axis measures the belief of the analyst. The left panel shows the exogenous timing example, and right panel shows the endogenous timing example. Green dashed arrows (success) correspond to jumps and red solid arrows (failure) indicate the direction of drift in the absence of arrival.

In both environments the optimal experiment is almost everywhere a single-jump process.[4] We see however that the qualitative properties are in stark contrast: the analyst always seeks contradictory evidence in the first case, and confirmatory evidence in the second. In addition, the preferences between frequency and precision as the belief varies are reversed: an analyst who controls the timing of their action, prefers frequency over precision when they are confident in the state; which is exactly when an analyst for whom the opportunity to act is exogenous, prefers precision over frequency.

To see why the optimal experiment has some directionality, notice that the analyst could set up an experiment which implements a zero-mean jump distribution upon arrival. This corresponds to an experiment whose 'arrival time' is uninformative about the state. When the analyst is willing to 'wait' for news, they might as well make the arrival time informative—and therefore learn while waiting. In the extreme case when they are not willing to wait at all, the signal is a diffusion.

Furthermore, this paper clarifies that single-jump processes are optimal since they can be viewed as extreme points of the set of processes. The objective can be defined to exhibit a form of linearity and hence the optimum is achieved at some such extreme point.

The other qualitative properties of the optimal experiment are more subtle. At the frontier of the feasible set, the analyst must trade off frequency for precision, in either jump direction (see Figure 2 below). Frequency and precision are substitutable, but as the analyst becomes *less* confident in the validity of their recommendation, frequency becomes *more* preferable to precision. Indeed, a more frequent signal is more likely to arrive before the opportunity to act, which the analyst values more when there is a high chance they are making a mistake. Consequently, the desire to substitute

[4]Zhong (2019) shows that this result is fairly robust, relying only on a continuity of the information measure. In a stopping problem with more general information measures, Hébert and Woodford (2021b) further advance the understanding of this result.

away from precision and into frequency is monotone in the potential gains from learning.

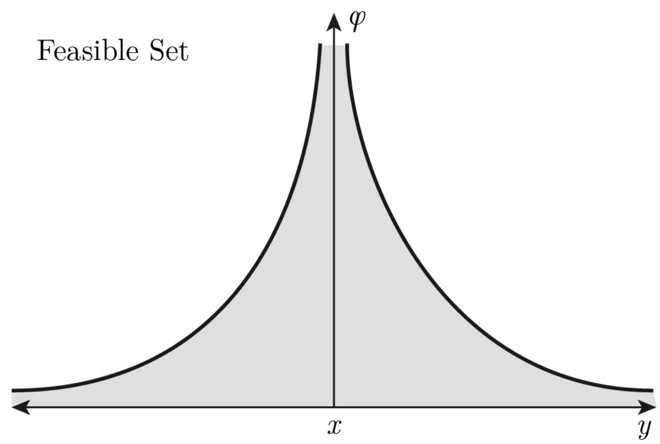

Figure 2: The shaded region indicates a typical set of feasible single-jump experiments for a prior $x \in (0,1)$. The horizontal axis corresponds to a jump location $y \neq x$, and the vertical axis to a frequency $\varphi > 0$.

Consider a confirmatory experiment. The absence of arrival makes the analyst less confident and hence more inclined to raise frequency. At the same time, precision must be increasing to make it possible for them to capture a sufficient portion of the information gain—which is also increasing along such a path. These competing forces create a tension which is resolved by a contradictory experiment: absence of arrival makes the analyst optimistic, wanting to decrease frequency and raise precision, which they can feasibly do.

Searching for contradictory evidence is risky since its arrival is less certain, and it is likely the analyst will not revise their recomendation quickly. However, the flexibility in designing the experiment allows them to compensate this risk by an ever more precise signal in the future, should the experiment be 'unsuccessful.' Anticipating this they become willing to acquire a contradictory experiment. Of course, they have to sacrifice frequency to implement such an increasing precision. Nevertheless, along a path of no arrival the analyst becomes more confident in the state which diminishes their desire for frequency.

It is worth noticing that under the optimal experiment, the information gain decreases over time, with probability 1. Indeed, the absence of a jump drifts beliefs in the direction of decreasing information gain, while the arrival of a jump directly implements such a reduction. This way the analyst can capture with certainty the benefits from information, along the entire path of beliefs they generate.

Related Research

The question of how to optimally gather information to decide on competing hypotheses about the world is probably as old as Statistics itself. In a series of seminal papers, Wald (1947) and Arrow, Blackwell and Girshick (1949) laid out the archetype sequential sampling model, in which a decision maker faces a dynamic problem of choosing between an irreversible decision whose consequences depend on an unknown state; or continuing to obtain information, in the form of independent draws from a distribution depending on the state. Costs are usually proportional to the number of observations. [5]

Although these early models recognize the importance of dynamics in statistical decision-making, they are limited in an important respect: the kinds of information sources available are beyond the decision maker's control.

Parallel to these models, the literature on 'experimentation' sought to merge information sources with payoff-generating activities: a decision maker chooses between competing options (bandits) of unknown quality whose payoffs are stochastic. The seminal papers in this literature are Robbins (1952), Weitzman (1979), and Gittins and Jones (1979).[6] Importantly, allocation of 'time' among these bandits produces dynamic information, and information costs arise as *opportunity costs* on the foregone payoffs of unchosen options. There too, the exogeneity of the structure of bandits limits the decision maker in the kind of information they can obtain.

The success of these dynamic models in capturing relevant economic phenomena spurred interest in moving away from their restrictive assumptions. Moscarini and Smith (2001) study a model of a decision maker who learns via observations of a Gaussian process whose 'precision' they control at a convex cost. By endogenizing this aspect they uncover a monotonicity between the precision of learning and the cost of delay. Che and Mierendorff (2019) consider a decision maker who learns via observing Poisson processes, but who faces a constraint in allocating their time between them. They are able to study the choice of 'direction' in the decision maker's information by considering choices between perfectly revealing Poisson signals, and uncover rich patterns of behaviour.[7] Ke and Villas-Boas (2019) study the problem of a decision maker who learns before choosing between two alternatives by sampling Gaussian processes, at a heterogeneous cost which is linear in the time spent with each source. Liang, Mu, and Syrgkanis (2021) consider a model where the decision maker allocates a fixed budget of time among Gaussian signals. In all of these models the decision maker *controls* the timing of their action.[8]

[5] An interesting generalization of this type of model is pursued in Morris and Strack (2017) who allow for belief-dependent sampling costs. Their goal is establishing equivalences between such sequential models, and static information acquisition models with appropriate cost functions.

[6] For a survey see Bergemann and Vallimaki (2006). Karatzas (1984), and Bolton and Harris (1999) treat continuous-time versions of the problem in a diffusion setting, while Keller, Rady, and Cripps (2005) treat the Poisson framework.

[7] See also Mayskaya (2020) for a related model.

[8] A working paper version of Liang, Mu and Syrgkanis (2021), (Mu, Liang and Syrgkanis, 2017) also considers

The contribution of this paper vis à vis this line of research is two-fold. Firstly, it departs from the prevalent stopping-problem setting to consider other important environments. Namely: (i) settings where the timing of actions is exogenous; and (ii) settings with repeated actions and unobservable payoffs (see Footnote 11, after the presentation of the model, for a discussion of how this model can be viewed this way.) The former differ from the 'Wald-type' models above by shutting down any timing concerns. This introduces novel considerations orthogonal to the cost of delay in taking an action—which drives in large part the results in the aforementioned literature. The latter are distinct from experimentation models in that, although payoffs are generated by repeated actions, they are unobservable, and thus the decision maker must rely on additional information obtained elsewhere. This allows the study of information acquisition when decisions have incremental impact on payoffs, while retaining the flexibility in the design of experiments.

Schneider and Wolf (2020), also recognize the importance of departing from traditional timing assumptions. They develop a model where a decision maker experiments with exogenous exponential bandits until a fixed deadline. They show that 'time pressure' significantly alters the conclusions of the existing literature. In this paper, the 'deadline' is stochastic so the concept of time pressure is less pronounced.

Secondly, the present model allows for full freedom in what information the decision maker can generate, and thus is capable of capturing all subsets of the features of information studied in the literature. Methodologically, this paper is closely related to Zhong (2019) which studies flexible dynamic information acquisition in a 'Wald-type' model of optimal stopping. The recent literature on Bayesian Persuasion and Rational Inattention, and the techniques developed therein,[9] allow for flexibility in modelling information, which Zhong (2019) introduces to a dynamic model. In particular, Zhong (2019) associates to each signal a flow-information quantity, as the speed of reduction in an uncertainty measure. This is then used to specify costs of information, which are, importantly, super-additive in this flow-information. This forces the decision maker to smooth information acquisition over time generating non-trivial dynamics. Zhong (2019) characterizes the optimal signal as an (almost everywhere) single-jump Poisson process, with immediate stopping after arrival, and confirmatory direction. This paper adopts a similar specification in terms of costs of information, and the relationship between the results here and the stopping-problem in Zhong (2019) have been discussed in the Introduction.

Furthermore, this paper shares features with the literature on dynamic rational inattention. Indeed, the information acquisition problem can be thought of as the 'information-processing' problem of a rationally inattentive agent. Following Sims' (2003) influential work, a vast literature emerged on how agents process information in the presence of processing costs. See Máckowiak, Matějka and

---

a framework which accommodates an exogenous deadline. Their focus is in establishing when myopically optimal behaviour is equivalent to the dynamically optimal one. Importantly, their exogenous Gaussian framework cannot be reconciled in the present model, where optimal information is almost never achievable by observing Gaussian sources.

[9]See Kamenica (2019) for a survey of Bayesian Persuasion, and Bergemann and Morris (2019) for a survey of Information Design.

Wiederholt (2021) for an extensive survey. Steiner, Stewart and Matějka (2017) develop a dynamic rational inattention model where in each period the decision maker obtains arbitrary information about a non-persistent state, prior to choosing an action. The costs of information are linear in the reduction of Shannon entropy. However, their flow utilities depend on the entire history of actions and states. Miao and Xing (2020), also study a dynamic rational inattention model but with more general uniformly posterior-separable information measures. In the language of the present paper, in both these models costs are linear in flow-information and hence the DM has no smoothing motive.

Moreover, Hébert and Woodford (2021b) pursue an optimal stopping-problem with flexible information acquisition as in Zhong (2019), but generalize the functional forms delivering the informational constraints by considering general divergences. They explore which properties of these divergences generate diffusion experiments and which pure-jump experiments.

Finally, any model seeking to allow flexibility in information choices must specify a notion of information quantity or information costs at the required level of generality. Defining a 'sensible' quantity of information is a very subtle and deep question, with a long history. Here we will be content with representing the total quantity of information as a prior-*independent* moment of the final posterior distribution. As in, for example, Pomatto, Strack and Tamuz (2020), Mensch (2018), and Denti, Marinacci, and Rustichini (2021) we view the association between experiments and a numerical quantity of information as a completion of the Blackwell order, in that it generates a total order on final experiments, while respecting the Blackwell order. Although a representation as a prior-*dependent* moment is a consequence of a form of linearity of the information measure over Blackwell experiments,[10] the prior-independent representation pursued here, and in Zhong (2019), Hebért and Woodford (2021b), Steiner, Stewart and Matějka (2017), and Miao and Xing (2020), is harder to motivate—particularly as a 'physical' cost of information acquisition. This is the subject of recent debate in the literature on information costs and particularly the distinction between posterior-separable and *uniformly* posterior-separable costs. The specification in this paper falls in the latter class, which is nevertheless rich enough and includes widely used (or well-motivated) measures of information, such as Shannon entropy reduction of Sims (2003), Bayesian LLR costs of Pomatto, Strack and Tamuz (2020) and Bloedel and Zhong (2020), Neighbourhood-based Costs of Hébert and Woodford (2021a), and many others. Exploring a richer set of information measures is left for future work.

[10]Versions of this result appear in characterizations of posterior-separable information costs in Caplin, Dean and Leahy (forthcoming); of cardinal measures of information in Mensch (2018), and Azrieli and Lehrer (2008); and most generally in the characterization of monotone affine functionals of experiments in Torgersen (1991).

2 Model

The Decision Problem.— There is a set of unknown states $\Theta = \{\theta_0, \theta_1\}$, and a Bayesian decision maker (DM) with prior belief $\pi = \Pr(\theta = \theta_1) \in (0,1)$. A perishable opportunity to act arrives at some exogenous random time $\tau \sim \exp(\rho)$.[11]. Optimal learning behaviour for such a problem is covered by the solution characterized here. The DM has state-dependent preferences over their actions, described by some payoff function $u : \Theta \times A \to \mathbb{R}$, where $u(\theta, a)$ is the payoff from action $a \in A$ in state $\theta \in \Theta$. We assume A is compact and $u(\cdot, a)$ continuous in a, so that the maximum value function $U : [0,1] \to \mathbb{R}$, with:

$$U(\mu) = \max_{a \in A} \mathbb{E}_\mu[u(\theta, a)] = \max_{a \in A} \mu \cdot u(\theta_1, a) + (1-\mu) \cdot u(\theta_0, a)$$

is well-defined, and continuous. U is convex as the maximum of linear functions. The DM discounts the future are rate $r > 0$, so if they act at time t their payoffs are given by $e^{-rt} \cdot U(\mu_t)$.

Information Strategies.— While waiting for the opportunity to act, the DM may acquire information about the unknown state θ in the form of *dynamic* experiments, which we now describe. We identify a dynamic experiment by the posterior process $(\mu_t)_{t \geq 0}$ it induces, and in particular we will consider (time-homogeneous) Markov experiments so that the choice of an experiment corresponds to a choice of transition semigroup $(\mathcal{P}_t)_{t \geq 0}$. The interpretation of $\mathcal{P}_t$, is that for any $f \in C[0,1]$, and $x \in (0,1)$, $\mathcal{P}_t f(x) = \mathbb{E}^x[f(\mu_t)]$. Furthermore, to be able to apply dynamic programming techniques, we would like to have a way to translate the choice of the posterior process to a choice of some local characteristics of this process. For this reason we restrict attention to Feller processes.[12] Such processes can be described by their infinitesimal generators, $\mathcal{A} : D_{\mathcal{A}} \to C[0,1]$ where we write formally:

$$\mathcal{A}f = \lim_{h \to 0+} \frac{1}{h}(\mathcal{P}_h f - f)$$

The subset $D_{\mathcal{A}} \subseteq C[0,1]$, for which this limit exists and is continuous, is called the domain of $\mathcal{A}$. The operator $\mathcal{A}$ allows us to approximate changes in the expected value of functions of the posterior as: $\mathbb{E}^x[f(\mu_h) - f(x)] = \mathcal{A}f(x) \cdot h + o(h)$.

Under some regularity assumptions, the generator can be represented as:[13]

[11]Given the exponential assumption on τ, this model has an alternative interpretation. It captures the problem of a decision maker enjoying flow payoff from repeatedly taking an action, but who cannot learn from these payoffs. Such problems are economically meaningful. For example, consider a patient who has to take one of two life-saving treatments every day. In such a situation it is unlilkely they are able to observe the gradual effects of their choice, and will inevitably have to rely on other sources to learn about which treatment is best. This contrasts such an example to the 'experimentation' literature, where decision makers learn purely by choosing among different options yielding stochastic payoffs. See e.g. Gittins and Jones (1979), Bolton and Harris (1999) and Keller, Rady, and Cripps (2005).

[12]That is, for $f \in C[0,1]$: (i) $\mathcal{P}_t f \in C[0,1]$, and (ii) $\lim_{t \to 0} \mathcal{P}_t f(x) = f(x)$, for all $x \in (0,1)$. For further definitions see Applebaum (2009) Chapter 3.2.

[13]In particular we insist that $C^\infty[0,1] \subseteq D_{\mathcal{A}}$. This is the Courrége Representation Theorem (Applebaum, 2009, Theorem 3.5.3)

$$\mathcal{A}f(x) = \alpha(x) \cdot f''(x) + \int f(y) - f(x) - f'(x) \cdot (y-x) \mathrm{d}F(y|x)$$

for any $f \in C^\infty[0,1]$, where $\alpha(x) > 0$ is a positive diffusion coefficient, and $F(\cdot|x) \in \mathcal{M}(x)$ where:

$$\mathcal{M}(x) = \left\{ F \text{ positive measure} \,\middle|\, F(\{x\}) = 0, \text{ and } \int \min\{1, (y-x)^2\} \mathrm{d}F(y|x) < \infty \right\}$$

is a measure describing the jumps of the process. Given such $\mathcal{A}$, we refer to the mapping $x \mapsto (\alpha(x), F(\cdot|x))$ as the characteristics of $\mathcal{A}$, or of the corresponding posterior process.

Such posterior processes and the distributions over paths they induce are called *admissible information strategies*. We introduce a piece of notation to aid the exposition. For $f \in C^1[0,1]$, let $\mathcal{L}f(\cdot, x)$ be the deviation of f from its tangent line at x: $\mathcal{L}f(y,x) = f(y) - f(x) - f'(x) \cdot (y-x)$. With this notation we have for any $f \in C^\infty[0,1]$:

$$\mathcal{A}f(x) = \alpha(x) \cdot f''(x) + \int \mathcal{L}f(y,x) \mathrm{d}F(y|x)$$

Ultimately, the optimal experiment will be constructed by choosing a pair $(\alpha, F) \in \mathbb{R}^+ \times \mathcal{M}(x)$ optimally at each belief $x \in (0,1)$.

Information Costs.— We now describe the information costs which will depend on a notion of flow-information generated by an admissible information strategy. For this, we start from a uniformly posterior-separable measure of information, so that over a period of length h, and starting from belief x, the quantity of information generated by $\mathcal{A}$ is:

$$\mathbb{E}^x[G(\mu_h) - G(x)] = \mathcal{A}G(x) \cdot h + o(h)$$

for some convex function $G : [0,1] \to \mathbb{R}^+$. From the above, the term:

$$\mathcal{A}G(x) = \alpha(x) \cdot G''(x) + \int \mathcal{L}G(y,x) \mathrm{d}F(y|x)$$

can be interpreted as the information increment, or *flow-information*. Following Zhong (2019) and Hébert and Woodford (2021b), we use this quantity to formulate the flow-costs of information.

We make the following assumption:

Assumption 1. $G \in C^\infty[0,1]$ *with* $G''(x) > 0$ *for all* $x \in [0,1]$. *Moreover,* $\lim_{x \to 0} G'(x) = -\infty$, *and* $\lim_{x \to 1} G'(x) = +\infty$.

The first two conditions impose smoothness on the problem and guarantee that the increment $\mathcal{A}G$ is well-defined, while the third and fourth prevent the DM from choosing a strategy which jumps to the boundaries, 0 or 1.

For dynamics to matter in this problem, we need the flow-costs of information to exhibit strict super-additivity, which introduces an 'information-smoothing' motive. If flow-costs are sub-additive in the information increment, the DM acquires the optimal amount of information instantaneously.

To keep matters simple we consider a capacity constraint on the flow-information an experiment can generate. This allows for an analysis of the key trade-offs at play, in a relatively straightforward framework.[14] In particular, we define the set of *feasible information strategies* by specifying a set of feasible characteristics, $\mathcal{I}(x)$ at each $x \in (0,1)$, where:

$$\mathcal{I}(x) = \left\{ (\alpha, F) \in \mathbb{R}^+ \times \mathcal{M}(x) \;\middle|\; \alpha \cdot G''(x) + \int \mathcal{L}G(y,x)\mathrm{d}F(y) \le \kappa \right\}$$

for a capacity $\kappa > 0$. We will say that a process is *feasible* if the characteristics of its generator are feasible, and denote the set of feasible information strategies by:

$$\mathcal{I} = \left\{ \mathcal{A} \text{ generator} \;\middle|\; \big(\alpha(x), F(\cdot|x)\big) \in \mathcal{I}(x) \quad \forall x \in (0,1) \right\}$$

We will denote by $\Pi(\mathcal{I})$ the set of distributions over paths induced by feasible information strategies.

3 Optimization Problem

With all the ingredients at hand we are ready to state the optimization problem. Given a problem with data $(r, \tilde{\rho}, \tilde{U})$ we consider the value $\tilde{V} : [0,1] \to \mathbb{R}^+$ be defined by:

$$\tilde{V}(x) = \sup_{P \in \Pi(\mathcal{I})} \mathbb{E}^x_P \Big[e^{-r\tilde{\tau}} \tilde{U}(\mu_{\tilde{\tau}}) \Big]$$

where $\tilde{\tau} \sim \exp(\tilde{\rho})$, is the exogenous time at which the DM gets to act. The DM chooses an admissible law over belief paths subject to the information constraint, to maximize their expected discounted payoff at the time of decision. It is easy to see that this problem is equivalent to a problem with no discounting, where $\tau \sim \exp(\tilde{\rho} + r)$ and the decision payoff is given by $U = \frac{\tilde{\rho}}{\tilde{\rho}+r} \cdot \tilde{U}$.

Hence, in what follows we will dispense with discounting and solve the general problem with data $(0, \rho, U)$:

$$V(x) = \sup_{P \in \Pi(\mathcal{I})} \mathbb{E}^x_P \Big[U(\mu_\tau) \Big] = \sup_{P \in \Pi(\mathcal{I})} \mathbb{E}^x_P \left[\int_0^{+\infty} e^{-\rho t} \rho U(\mu_t) \, \mathrm{d}t \right] \tag{V}$$

where $\tau \sim \exp(\rho)$, is the exogenous time at which the DM gets to act.

The second equality above follows from the independence of τ and the belief process, as well as the exponential distribution assumption. It reveals an alternative interpretation of this model: this is the objective of a DM who repeatedly acts receiving flow payoff ρU, but who does not learn from these payoffs. The two models are behaviourally equivalent in terms of the optimal information

[14] Indeed, the case of smooth convex flow costs, can be analyzed in a similar manner by endogenizing the 'capacity constraint.'

strategies they generate.

Characterization

Before we give the characterization we collect some properties of the value function.

Lemma 1. *The value function V is convex, and continuously differentiable on* $(0,1)$.

Proof. Appendix A. □

Convexity of V is unsurprising since the DM strictly values any free experiment which generates a spread over their beliefs. Convexity also establishes the right- and left-differentiability of V, and to prove that it is continuously differentiable, it suffices to show that $V'_- = V'_+$. This is accomplished by constructing a simple dominated experiment and using the definition of the value.

We now proceed to characterize V and the optimal experiment.

Proposition 1. *The value function V is a viscosity solution to:*

$$\rho(V-U)(x) = \sup_{y\neq x} \frac{\mathcal{L}V(y,x)}{\mathcal{L}G(y,x)} \cdot \kappa \tag{HJB}$$

with $V(1)=U(1)$ *and* $V(0)=U(0)$.

Proof. Appendix B. □

Beyond giving an equation for the value function, this characterization shows that optimal experiments can be found among single-jump processes.[15] To see this we will argue heuristically, and derive the Hamilton-Jacobi-Bellman (HJB) equation as:[16]

$$\rho(V-U)(x) = \sup_{\mathcal{A}} \mathcal{A}V(x) \quad \text{subject to} \quad \mathcal{A}G(x) \leq \kappa \tag{I}$$

We now consider the optimization problem on the RHS, and construct the Langrangian:

$$\sup_{\mathcal{A}} \mathcal{A}V(x) - \lambda^*(x)\mathcal{A}G(x) + \lambda^*(x)\kappa$$

This objective is linear in the choice variable $\mathcal{A}$ so for this problem to have a finite solution (necessary to stay below the capacity constraint),

$$\mathcal{A}V(x) - \lambda^*(x)\mathcal{A}G(x) \leq 0, \;\; \forall \mathcal{A} \Rightarrow \lambda^*(x) \geq \sup_{\mathcal{A}} \frac{\mathcal{A}V(x)}{\mathcal{A}G(x)}$$

[15] This is so when the supremum is achieved; when the supremum is strict, the optimal signal is a diffusion.

[16] This is not justified in general without a priori knowledge about the smoothness of the value; this is why one needs to work within the viscosity solution framework.

One can show that in fact the inequality holds as an equality and consequently, $\sup_{\mathcal{A}} \mathcal{A}V(x) - \lambda^*(x)\mathcal{A}G(x) = 0$, which we substitute in Equation (I) to get:

$$\rho(V-U)(x) = \lambda^*(x)\cdot\kappa = \sup_{\mathcal{A}} \frac{\mathcal{A}V(x)}{\mathcal{A}G(x)}\cdot\kappa \tag{II}$$

We now explain why it is enough to consider single-jump processes. As the ratio of two linear functionals of $\mathcal{A}$, the objective maximized on the RHS is quasi-convex. We re-write it as:

$$\sup_{(\alpha,F)} \frac{\alpha V''(x) + \int \mathcal{L}V(y,x)\mathrm{d}F(y)}{\alpha G''(x) + \int \mathcal{L}G(y,x)\mathrm{d}F(y)}\cdot\kappa = \sup_{(0,F)} \frac{\int \mathcal{L}V(y,x)\mathrm{d}F(y)}{\int \mathcal{L}G(y,x)\mathrm{d}F(y)}\cdot\kappa$$

where the second equality holds because the diffusion term can be approximated by jump measures.[17] Finally, the value of this ratio is independent of the total mass of F, and one can just consider probability measures on $\Delta(S^x)$ where $S^x = [0,x)\cup(x,1]$. Quasi-convexity then yields:

$$\sup_{P\in\Delta(S^x)} \frac{\int \mathcal{L}V(y,x)\mathrm{d}P(y)}{\int \mathcal{L}G(y,x)\mathrm{d}P(y)}\cdot\kappa = \sup_{P\in\mathrm{ext}\Delta(S^x)} \frac{\int \mathcal{L}V(y,x)\mathrm{d}P(y)}{\int \mathcal{L}G(y,x)\mathrm{d}P(y)}\cdot\kappa$$

The set of extreme points $\mathrm{ext}\Delta(S^x) = \{\delta_y : y \neq x\}$, are the point-masses at points other than x. Operators built from such point-masses correspond to single-jump processes.

Plugging in Equation (II), gives:

$$\rho(V-U)(x) = \sup_{y\neq x} \frac{\mathcal{L}V(y,x)}{\mathcal{L}G(y,x)}\cdot\kappa$$

which is the (HJB) equation we were after.

We can recognize in the LHS as the difference between the optimal payoff, V, and the payoff from taking the optimal action in the absence of further information, U. Their difference $V-U$ captures how much the DM gains from having the ability to learn optimally. We term this quantity the *information gain*.[18] It accrues at rate $\rho > 0$ which is the frequency at which the DM gets to take an action. The RHS proportional to the ratio between the expected change in the value and the flow-information, optimized over single-jump signals.[19] That ratio coincides with the *shadow-cost* of information. The HJB equation necessitates that, at the optimum, the information gain is

[17] Indeed, by setting $F^\varepsilon = \frac{\alpha}{\varepsilon^2}\cdot\delta_{x+\varepsilon}$, we have for any $f\in C^2[0,1]$,

$$\int \mathcal{L}f(y,x)\mathrm{d}F^\varepsilon(y) = \frac{\alpha}{\varepsilon^2}\cdot\big(f(x+\varepsilon) - f(x) - f'(x)\cdot\varepsilon\big) = \alpha f''(x) + \frac{o(\varepsilon^2)}{\varepsilon^2} \to \alpha f''(x) \quad \text{as } \varepsilon\to 0$$

[18] Interpreting the solution to the original problem with data $(r,\rho,\tilde{U})$ with discounting in terms of the solution to the transformed problem, amounts to simply recognizing that the payoff in the absence of further information is given by $U(\mu) = \mathbb{E}[e^{-r\tilde{\tau}}]\tilde{U}(\mu) = \frac{\rho}{\rho+r}\tilde{U}(\mu)$.

[19] Of course, when the supremum is strict no single-jump experiment is optimal in which case the signal collapses to a diffusion.

exactly proportional to the shadow-cost. This monotonicity between the information gain and the shadow-cost of information is a key driver of the results.

Finally, the local characteristics of the optimal experiment are identified as follows. If $\sup_{y\neq x}\frac{\mathcal{L}V(y,x)}{\mathcal{L}G(y,x)} > \frac{V''(x)}{G''(x)}$ then a pure-jump is optimal, and we recover it from:

$$y^*(x) \in \arg\max_{y\neq x} \frac{\mathcal{L}V(y,x)}{\mathcal{L}G(y,x)} \neq \emptyset$$

and specify its frequency by the binding information constraint:

$$\varphi^*(x) = \frac{\kappa}{\mathcal{L}G(y^*(x),x)}$$

If $\sup_{y\neq x}\frac{\mathcal{L}V(y,x)}{\mathcal{L}G(y,x)} = \frac{V''(x)}{G''(x)}$ then a diffusion is optimal, and is characterized by binding the information constraint: $\alpha^*(x) = \frac{\kappa}{G''(x)}$.

Properties

We will now examine the key features of the optimal experiment, in terms of the *frequency* of jumps, their *direction*, and their *location*.

Proposition 2 (diffusion signal at the max)**.** *Diffusion is optimal if and only if the information gain is maximal:*

$$\rho(V-U)(x) = \frac{V''(x)}{G''(x)}\cdot\kappa \iff \quad x \text{ is a local maximum of } V-U$$

Proof. Appendix C. □

This result establishes when a pure-jump process is not optimal, which occurs only at the local maxima of the information gain. At every other point, the optimal experiment generates a single jump.

We can illustrate this by considering a binary action setting. There, the information gain is maximal at the belief for which the DM is indifferent between the two actions. At that belief the DM is eager to rapidly break their indifference, and this desire leads them to substitute *all* impact for frequency. Such points are non-generic and only play a role if it so happens that the prior exactly coincides with such a point. If the DM has a prior different than a local maximizer of the information gain, their posterior will never visit any local maximizer along the optimal path.

Next, we establish two important properties of the optimal experiment.

Proposition 3 (direction)**.** *Any optimal jump is in the direction of increasing information gain:*

$$\rho(V-U)'(x) > 0 \iff y^*(x) - x > 0$$

Proof. Appendix D. □

This result states that any optimal jump is in the direction of increasing information gain. Stated differently, it necessitates that the information gain drifts downwards absent a breakthrough. In binary action examples, there will be a belief $\hat{x} \in (0, 1)$ where they are indifferent among actions and the information gain is maximal. This belief provides a payoff-adjusted threshold against which to measure the DM's 'conviction,' and we can say that the DM is more confident about state 1, if $\mu > \hat{x}$, and vice versa. For instance, in the introductory example where payoffs were symmetric across states, $\hat{x} = 1/2$, and hence it coincided with the most uncertain belief. There, the notion of conviction was what we would naturally think. In general, one needs to adjust for discrepancies in payoffs across states, since they generate a degree of 'bias' for the DM.

With this in mind, Proposition 3 states that optimal jumps are contradictory in the sense that they counter the DM's conviction: when to the right of $\hat{x}$ the DM is more confident about state 1, and yet seeks evidence which would contradict this; and vice versa.

Moreover, we have the following:

Proposition 4 (location)**.** *The information gain decreases along the sequence of optimal jumps:*

$$\rho(V - U)(y^*(x)) \leq \rho(V - U)(x)$$

Proof. Appendix E. □

This result establishes the location and, by extension, the impact of optimal jumps. In particular, it states that optimal jumps are sufficiently significant so that the information gain decreases after a jump. The idea is that the DM consolidates any potential intermediate jumps until sufficient precision is accumulated into one big breakthrough. More specifically, starting from x, one can show that if there was an intermediate jump $y^*(x)$ to a location of strictly larger information gain, $(V - U)(y^*(x)) > (V - U)(x)$, then the DM could profitably jump to the destination of the next jump, $x \to y^*(y^*(x))$.

To gain some intuition on the properties above we focus again on a binary action setting. There are two ways by which the DM can be induced to change their plan of action. They can seek evidence whose arrival reaffirms their intentions, and the absence of arrival slowly induces them to change their action. This is a 'confirmatory' information strategy. Alternatively, they can use a 'contradictory' strategy in which they seek evidence whose arrival generates a substantial decrease in confidence in their plan of action, and the absence of arrival slowly affirms their intentions. Notice that for a contradictory strategy to be valuable it must be possible for the DM to accumulate sufficient evidence against their plan, via potentially repeated jumps.

Proposition 3 states that the DM would never jump in a direction which makes them more convinced of the validity of their current intentions. That is, they find the contradictory information strategy optimal. Moreover, Proposition 4 states that they would never attempt to change their action through jumping repeatedly in the opposite direction. That is, they concentrate resources to generate a single *convincing* signal. Proposition 4 also quantifies *how* convincing this contradictory signal is: it causes the information gain to decrease.

Finally, we have the following:

Proposition 5 (frequency vs impact)**.** *Let $\varphi^*(x)$ be the optimal frequency. Then φ^* is increasing in the information gain:*

$$(V-U)(x_1) \leq (V-U)(x_2) \Rightarrow \varphi^*(x_1) \leq \varphi^*(x_2)$$

Proof. Appendix F. □

This result establishes that the frequency of the optimal experiment increases in the information gain. In binary action examples, this describes a key qualitative feature of the optimal experiment: the DM trades off frequency for impact as they become more confident in the state.

Conclusion

This paper characterizes the optimal dynamic experiment in an environment where the decision maker lacks control over the timing of actions. The decision maker can flexibly choose all aspects of the experiment but faces a capacity constraint on the flow-information it can generate. The key properties of learning in this environment are that resources are concentrated into generating single breakthroughs, which counter the decision maker's prior intentions. Consequently, absence of breakthroughs make the decision maker more confident in their action plan which leads them to sacrifice frequency in the arrival of breakthroughs, to increase their significance.

References

D. Applebaum. *Lévy processes and stochastic calculus.* Cambridge university press, 2009.

K. J. Arrow, D. Blackwell, and M. A. Girshick. Bayes and minimax solutions of sequential decision problems. *Econometrica*, 17(3/4):213–244, 1949. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1905525.

Y. Azrieli and E. Lehrer. The value of a stochastic information structure. *Games and Economic Behavior*, 63(2):679–693, 2008.

D. Bergemann and S. Morris. Information design: A unified perspective. *Journal of Economic Literature*, 57(1):44–95, 2019.

D. Bergemann and J. Välimäki. Bandit problems. 2006.

A. W. Bloedel and W. Zhong. The cost of optimally-acquired information. *Unpublished Manuscript, November*, 2020.

P. Bolton and C. Harris. Strategic experimentation. *Econometrica*, 67(2):349–374, 1999.

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

P. DeJarnette, D. Dillenberger, D. Gottlieb, and P. Ortoleva. Time lotteries and stochastic impatience. 88(2):619–656, 2020. doi: 10.3982/ecta16427. URL https://doi.org/10.3982%2Fecta16427.

T. Denti, M. Marinacci, A. Rustichini, et al. Experimental cost of information. Unpublished Manuscript, 2021.

W. H. Fleming and H. M. Soner. *Controlled Markov processes and viscosity solutions*, volume 25. Springer Science & Business Media, 2006.

J. C. Gittins and D. M. Jones. A dynamic allocation index for the discounted multiarmed bandit problem. *Biometrika*, 66(3):561–565, 12 1979. ISSN 0006-3444. doi: 10.1093/biomet/66.3.561. URL https://doi.org/10.1093/biomet/66.3.561.

B. M. Hébert and M. Woodford. Neighborhood-based information costs. Technical report, National Bureau of Economic Research, 2021a.

B. M. Hébert and M. Woodford. Rational inattention when decisions take time. Unpublished Manuscript, 2021b.

J. Jacod and A. Shiryaev. *Limit theorems for stochastic processes*, volume 288. Springer Science & Business Media, 2013.

E. Kamenica. Bayesian persuasion and information design. *Annual Review of Economics*, 11: 249–272, 2019.

I. Karatzas. Gittins indices in the dynamic allocation problem for diffusion processes. *The Annals of Probability*, 12(1):173–192, 1984.

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

G. Keller, S. Rady, and M. Cripps. Strategic experimentation with exponential bandits. *Econometrica*, 73(1):39–68, 2005.

A. Liang, X. Mu, V. Syrgkanis, et al. Dynamic information acquisition from multiple sources. *arXiv preprint arXiv:1703.06367*, 58, 2017.

A. Liang, X. Mu, and V. Syrgkanis. Dynamically aggregating diverse information. 2021.

D. G. Luenberger. *Optimization by vector space methods*. John Wiley & Sons, 1997.

B. Máckowiak, F. Matějka, and M. Wiederholt. Rational inattention: A review. Unpublished Manuscript, 2021.

T. Mayskaya. Dynamic choice of information sources. *California Institute of Technology Social Science Working Paper, ICEF Working Paper WP9/2019/05*, 2020.

J. Mensch. Cardinal representations of information. *Available at SSRN 3148954*, 2018.

J. Miao and H. Xing. Dynamic discrete choice under rational inattention. Technical report, Working Paper, 2020.

S. Morris and P. Strack. The wald problem and the equivalence of sequential sampling and static information costs. *Unpublished manuscript, June*, 2017.

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

L. Pomatto, P. Strack, and O. Tamuz. The cost of information. *arXiv preprint arXiv:1812.04211*, 2020.

H. Robbins. Some aspects of the sequential design of experiments. *Bulletin of the American Mathematical Society*, 58(5):527–535, 1952.

J. Schneider and C. Wolf. On risk and time pressure: When to think and when to do. *Available at SSRN 3490935*, 2020.

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

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

J. Stenger, F. Gamboa, and M. Keller. Optimization of quasi-convex function over product measure sets. *SIAM Journal on Optimization*, 31(1):425–447, 2021.

E. Torgersen. *Comparison of statistical experiments*, volume 36. Cambridge University Press, 1991.

A. Wald. Foundations of a general theory of sequential decision functions. *Econometrica*, 15(4): 279–313, 1947. ISSN 00129682, 14680262. URL http://www.jstor.org/stable/1905331.

M. L. Weitzman. Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, pages 641–654, 1979.

W. Zhong. Optimal dynamic information acquisition. unpublished manuscript, 2019.

Appendix

A Proof of Lemma 1

Convexity follows from the fact that in a hypothetical scenario where the DM observes some free information generating a spread $\lambda x_1 + (1-\lambda)x_2$ over their beliefs, the DM can still wait to observe which belief materializes and implement the optimal experiment then. Hence, $V\big(\lambda x_1 + (1-\lambda)x_2\big) \geq \lambda V(x_1) + (1-\lambda)V(x_2)$

We now prove differentiability. Fix some $x \in (0,1)$. Consider the following experiment: a signal arrives at some time σ independently of the state, and implements a 50:50 spread: $F = \frac{1}{2}\cdot\delta_{x+\varepsilon} + \frac{1}{2}\cdot\delta_{x-\varepsilon}$. The frequency of $\varphi(x,\varepsilon)$ of the arrival time σ is selected to bind the constraint, that is:

$$\varphi(x,\varepsilon) = \frac{\kappa}{\frac{1}{2}\Big[G(x+\varepsilon) - G(x)\Big] + \frac{1}{2}\Big[G(x-\varepsilon) - G(x)\Big]}$$

Call the resulting distribution of beliefs Q. For any time interval $h > 0$, we have:

$$V(x) \geq \mathbb{E}_Q\left[\int_0^h e^{-\rho t}\rho U(\mu_t)\,\mathrm{d}t + e^{-\rho h}V(\mu_h)\right]$$

since implementing this experiment until h time passes and then following the optimal experiment must be unprofitable. We re-arrange and expand to get for h small enough:

$$0 \geq \int_0^h e^{-\rho t}\rho U(x) - \rho V(x) + \varphi(x,\varepsilon)\cdot\left(\frac{1}{2}\cdot V(x+\varepsilon) + \frac{1}{2}\cdot V(x+\varepsilon) - V(x)\right)\mathrm{d}t$$

Now dividing by h and sending $h \to 0$, and substituting for $\varphi(x,\varepsilon)$ we have:

$$\frac{\rho}{\kappa}(V-U)(x)\cdot\Big[G(x+\varepsilon) - G(x) + G(x-\varepsilon) - G(x)\Big] \geq \Big[V(x+\varepsilon) - V(x) + V(x-\varepsilon) - V(x)\Big]$$

Dividing by $\varepsilon > 0$ we get:

$$\frac{\rho}{\kappa}(V-U)(x)\cdot\left[\frac{G(x+\varepsilon)-G(x)}{\varepsilon} + \frac{G(x-\varepsilon)-G(x)}{\varepsilon}\right] \geq \left[\frac{V(x+\varepsilon)-V(x)}{\varepsilon} + \frac{V(x-\varepsilon)-V(x)}{\varepsilon}\right]$$

We take limits $\varepsilon \to 0$ on both sides to get:

$$\frac{\rho}{\kappa}(V-U)(x)\cdot\Big[G'_+(x) - G'_-(x)\Big] \geq \Big[V'_+(x) - V'_-(x)\Big]$$

where we notice that the limits exist by convexity of V and G.

Since G is differentiable the LHS is zero and we have:

$$V'_-(x) \geq V'_+(x)$$

Consequently, V is differentiable at x. As a convex function we have that it is also continuously differentiable, which completes the proof.

B Proof of Proposition 1

Before we begin we give the definition of a viscosity solution to:

$$\rho(v-U)(x) - \sup_{\mathcal{A}\in\mathcal{I}(x)} \mathcal{A}v(x) = 0 \quad \text{and} \quad V = U \quad \text{on } \{0,1\}$$

A continuous function $V : [0,1] \to \mathbb{R}^+$ is a viscosity solution to the above if:[20]

- For all $v \in C^\infty[0,1]$, such that $V \geq v$ and $V - v$ has a global minimum of 0, at $\hat{x} \in (0,1)$:

$$\rho(v-U)(\hat{x}) - \sup_{\mathcal{A}\in\mathcal{I}(x)} \mathcal{A}v(\hat{x}) \geq 0 \tag{super}$$

- For all $v \in C^\infty[0,1]$, such that $V \leq v$ and $V - v$ has a global maximum of 0, at $\hat{x} \in (0,1)$:

$$\rho(v-U)(\hat{x}) - \sup_{\mathcal{A}\in\mathcal{I}(x)} \mathcal{A}v(\hat{x}) \leq 0 \tag{sub}$$

Firstly, we prove the following:

Lemma 2. *For any test-function $v \in C^\infty[0,1]$,*

$$\sup_{\mathcal{A}\in\mathcal{I}(x)} \mathcal{A}v(x) = \sup_{y\neq x} \frac{\mathcal{L}v(y,x)}{\mathcal{L}G(y,x)} \cdot \kappa$$

Proof. To prove this we first appeal to the saddle-point characterization of the optimal solution to this problem. First, $\mathcal{A} \to \mathcal{A}v(x)$ and $\mathcal{A} \to \mathcal{A}G(x)$ are linear and the set $\mathcal{E}(x)$ is convex. Moreover, there is clearly an element $\mathcal{A} \in \mathcal{E}(x)$ for which $\mathcal{A}G(x) < \kappa$, and the cone in $\mathbb{R}$ is closed with non-empty interior. By the Langrange Multiplier Theorems in Luenberger (1997, p.219 and 221),

$$\sup_{\mathcal{A}\in\mathcal{I}(x)} \mathcal{A}v(x) = L(\mathcal{A}^*, \lambda^*, x)$$

where $(\mathcal{A}^*, \lambda^*)$ forms a saddle-point of the functional over the domain $\mathcal{E}(x) \times \mathbb{R}^+$:

$$L(\mathcal{A}, \lambda, x) = \mathcal{A}v(x) - \lambda\mathcal{A}G(x) + \lambda \cdot \kappa$$

That is,

$$L(\mathcal{A}, \lambda^*, x) \leq L(\mathcal{A}^*, \lambda^*, x) \leq L(\mathcal{A}^*, \lambda, x) \qquad \text{for all } \mathcal{A} \in \mathcal{E}(x), \lambda \geq 0$$

Next, we provide such a saddle-point. Firstly, define $\lambda^*(x) = \sup_{\mathcal{A}\in\mathcal{E}(x)} \frac{\mathcal{A}v(x)}{\mathcal{A}G(x)}$. We will show that:

$$\lambda^*(x) = \sup_{y\neq x} \frac{\mathcal{L}v(y,x)}{\mathcal{L}G(y,x)}$$

that is, the supremum can be taken over single-jump processes. Firstly, because we can always

[20] See Fleming and Soner (2006), Chapter II.4.

approximate a diffusion by an appropriate choice of jump measures, we have:

$$\lambda^*(x) = \sup_{(0,F)\in\mathcal{E}(x)} \frac{\int \mathcal{L}v(y,x)\mathrm{d}F(y)}{\int \mathcal{L}G(y,x)\mathrm{d}F(y)} = \sup_{(0,P)\in\mathcal{E}(x)} \frac{\int \mathcal{L}v(y,x)\mathrm{d}P(y)}{\int \mathcal{L}G(y,x)\mathrm{d}P(y)}$$

where P are probability measures. The second equality holds because the value of the ratio is independent of the total mass of the measures F.

Next, let $S^x = [0,x)\cup(x,1]$, and $\Delta(S^x)$ the set of Borel probability measures on S^x. The space S^x is Polish as a G_δ subset of a Polish space. The extreme points are given by $\mathrm{ext}\Delta(S^x) = \{\delta_y : y \neq x\}$.[21] Since the ratio above is a quasiconvex, lower semicontinuous functional of $P \in \Delta(S^x)$, we have by Theorem 3.2 in Stenger, Gamboa, and Keller (2021):

$$\sup_{P\in\Delta(S^x)} \frac{\int \mathcal{L}v(y,x)\mathrm{d}P(y)}{\int \mathcal{L}G(y,x)\mathrm{d}P(y)} = \sup_{P\in\mathrm{ext}\Delta(S^x)} \frac{\int \mathcal{L}v(y,x)\mathrm{d}P(y)}{\int \mathcal{L}G(y,x)\mathrm{d}P(y)} = \sup_{y\neq x} \frac{\mathcal{L}v(y,x)}{\mathcal{L}G(y,x)}$$

which proves the claim.

Next, define:

$$\mathcal{A}^* = \begin{cases} \varphi^*\cdot\delta^*_y \ \text{ with } \ y^* \in \arg\max_{y\neq x} \frac{\mathcal{L}v(y,x)}{\mathcal{L}G(y,x)} \ \text{ and } \ \varphi^* = \frac{\kappa}{\mathcal{L}G(y^*,x)}, & \text{if } \lambda^*(x) > \frac{v''(x)}{G''(x)} \\ \alpha^* = \frac{\kappa}{G''(x)}, & \text{if } \lambda^*(x) = \frac{v''(x)}{G''(x)} \end{cases}$$

Notice that $\mathcal{A}^*G(x) = \kappa$ and $\mathcal{A}^*v(x) - \lambda^*(x)\mathcal{A}^*G(x) = 0$.[22] Therefore, $L(\mathcal{A}^*, \lambda^*, x) = \lambda^*(x)\cdot\kappa$. We argue that $(\mathcal{A}^*, \lambda^*)$ is a saddle-point.

For arbitrary $\mathcal{A}$, from the definition of λ^* (we omit the dependence on x for clarity):

$$L(\mathcal{A}, \lambda^*) = \mathcal{A}v - \lambda^*\cdot\mathcal{A}G + \lambda^*\kappa \leq \lambda^*\cdot\kappa = L(\mathcal{A}^*, \lambda^*, x)$$

Moreover, for arbitrary λ:

$$\begin{aligned} L(\mathcal{A}^*, \lambda) = \mathcal{A}^*v - \lambda\mathcal{A}G + \lambda\cdot\kappa &= \mathcal{A}^*v - \lambda^*\mathcal{A}G + \lambda\cdot(\kappa - \mathcal{A}^*G) + \lambda^*\mathcal{A}^*G \\ &= \lambda\cdot(\kappa - \mathcal{A}^*G) + \lambda^*\mathcal{A}^*G \\ &= \lambda^*\kappa \end{aligned}$$

Consequently, the pair $(\mathcal{A}^*, \lambda^*)$ defined above is a saddle-point for L.

□

We now prove that the value V is a viscosity solution to:

$$\rho(V-U)(x) = \sup_{\mathcal{A}\in\mathcal{I}(x)} \mathcal{A}V(x)$$

The claim then follows since the sub- and super-solution properties are tested on test-functions $v \in C^2[0,1]$, and we can apply Lemma 2.

[21] See Stenger, Gamboa, and Keller (2021), Theorem 2.1. Here we are taking the trivial moment class.

[22] Note that the constructed $\mathcal{A}^*$ achieves the supremum defining λ^*.

We check the conditions for Theorem 5.1 in Fleming and Soner (2006). In fact, we only need to check the continuity in x of $\sup_{\mathcal{A}\in\mathcal{I}(x)} \mathcal{A}v(x)$ for $v \in C^\infty[0,1]$. This follows from the compactness of $\mathcal{I}(x)$ which is implied by the strict convexity $G'' > 0$.

Lemma 3. *For $v \in C^\infty[0,1]$, the mapping $x \mapsto \sup_{\mathcal{A}\in\mathcal{I}(x)} \mathcal{A}v(x)$ is continuous.*

Proof. Online Appendix. □

C Proof of Proposition 2

($\Leftarrow$): Suppose $(V-U)(x)$ is maximal but an optimal jump exists; that is the supremum in (HJB) is achieved. The envelope theorem applies and we have:

$$0 = \rho(V-U)'(x) = \varphi^*(x) \cdot \Big(\lambda^*(x) \cdot G''(x) - V''(x)\Big) \cdot (y^*(x) - x)$$

where $\lambda^*(x) = \sup_{y\neq x} \frac{\mathcal{L}V(y,x)}{\mathcal{L}G(y,x)}$. The first equality holds since $V-U$ has a local maximum at x, and the maximizer is interior. By the hypothesis that an optimal jump exists, $\lambda^*(x) > \frac{V''(x)}{G''(x)}$, and the above implies $y^*(x) - x = 0$, a contradiction.

D Proof of Proposition 3

Suppose an optimal jump $y^*(x)$ exists at x, and let $\lambda^*(x) = \sup_{y\neq x} \frac{\mathcal{L}V(y,x)}{\mathcal{L}G(y,x)}$. Then, this supremum is achieved by $y^*(x)$. Consequently, the envelope theorem applies and from (HJB) we have:

$$\rho(V-U)'(x) = \frac{\kappa}{\mathcal{L}G(y^*(x),x)} \cdot \Big(\lambda^*(x) \cdot G''(x) - V''(x)\Big) \cdot (y^*(x) - x)$$

The first two terms in the RHS are strictly positive, and thus $\rho(V-U)'(x) > 0 \iff y^*(x) - x > 0$.

E Proof of Proposition 4

We prove this by contradiction. Suppose this is not the case, so that for some $x \in (0,1)$:

$$\sup_{y\neq y^*(x)} \frac{\mathcal{L}V(y,y^*(x))}{\mathcal{L}G(y,y^*(x))} = \lambda^*\big(y^*(x)\big) > \lambda^*(x) = \sup_{y\neq x} \frac{\mathcal{L}V(y,x)}{\mathcal{L}G(y,x)}$$

We will show that the original jump to $y^*(x)$ was not optimal. First, there exists some $\hat{y} \neq y^*(x)$ such that:

$$\frac{\mathcal{L}V(\hat{y},y^*(x))}{\mathcal{L}G(\hat{y},y^*(x))} > \lambda^*\big(y^*(x)\big) - \varepsilon > \lambda^*(x) \tag{*}$$

We will show that $\hat{y}$ constitutes a profitable deviation to $y^*(x)$ starting from x.

Indeed, for any $f \in C^2[0,1]$:

$$\mathcal{L}f(\hat{y},x) = \mathcal{L}f(\hat{y},y^*(x)) + \mathcal{L}f(y^*(x),x) + \big(f'(y^*(x)) - f'(x)\big) \cdot (\hat{y} - y^*(x))$$

and therefore:

$$\frac{\mathcal{L}V(\hat{y},x)}{\mathcal{L}G(\hat{y},x)} = \frac{\mathcal{L}V(\hat{y},y^*(x)) + \mathcal{L}V(y^*(x),x) + \big(V'(y^*(x)) - V'(x)\big)\cdot(\hat{y}-y^*(x))}{\mathcal{L}G(\hat{y},y^*(x)) + \mathcal{L}G(y^*(x),x) + \big(G'(y^*(x)) - G'(x)\big)\cdot(\hat{y}-y^*(x))} \tag{**}$$

Next, we consider optimizing:

$$\max_{y\neq x} \frac{\mathcal{L}V(y,x)}{\mathcal{L}G(y,x)}$$

The FOC necessitates that the maximizer $y^*(x)$ satisfies:

$$\big(V'(y^*(x)) - V'(x)\big) - \big(G'(y^*(x)) - G'(x)\big)\cdot\frac{\mathcal{L}V(y^*(x),x)}{\mathcal{L}G(y^*(x),x)} = 0 \Rightarrow$$

$$\big(V'(y^*(x)) - V'(x)\big) = \lambda^*(x)\cdot\big(G'(y^*(x)) - G'(x)\big)$$

Consequently, we can re-write (**) as:

$$\begin{aligned}\frac{\mathcal{L}V(\hat{y},x)}{\mathcal{L}G(\hat{y},x)} &= \frac{\frac{\mathcal{L}V(\hat{y},y^*(x))}{\mathcal{L}G(\hat{y},y^*(x))}\cdot\mathcal{L}G(\hat{y},y^*(x)) + \lambda_*(x)\cdot\mathcal{L}G(y^*(x),x) + \lambda_*(x)\big(G'(y^*(x)) - G'(x)\big)\cdot(\hat{y}-y^*(x))}{\mathcal{L}G(\hat{y},y^*(x)) + \mathcal{L}G(y^*(x),x) + \big(G'(y^*(x)) - G'(x)\big)\cdot(\hat{y}-y^*(x))} \\ &> \frac{\lambda_*(x)\cdot\mathcal{L}G(\hat{y},y^*(x)) + \lambda_*(x)\cdot\mathcal{L}G(y^*(x),x) + \lambda_*(x)\big(G'(y^*(x)) - G'(x)\big)\cdot(\hat{y}-y^*(x))}{\mathcal{L}G(\hat{y},y^*(x)) + \mathcal{L}G(y^*(x),x) + \big(G'(y^*(x)) - G'(x)\big)\cdot(\hat{y}-y^*(x))} \\ &= \lambda_*(x)\end{aligned}$$

where the inequality follows from (*). This contradicts the optimality of $y^*(x)$.

F Proof of Proposition 5

In each region where an optimal jump exists, the FOC holds and we get:

$$\big(V'(y^*(x)) - V'(x)\big) = \lambda^*(x)\cdot\big(G'(y^*(x)) - G'(x)\big)$$

and using the Implicit Function Theorem we differentiate to obtain:

$$V''(y^*(x))\cdot\frac{dy^*(x)}{dx} - V''(x) = \lambda^*(x)\cdot\left(G''(y^*(x))\frac{dy^*(x)}{dx} - G''(x)\right) + (\lambda^*)'(x)\cdot\big(G'(y^*(x)) - G'(x)\big)$$

which we re-arrange to get:

$$\frac{dy^*(x)}{dx}\cdot\Big(V''(y^*(x)) - \lambda^*(x)\cdot G''(y^*(x))\Big) = V''(x) - \lambda^*(x)G''(x) + (\lambda^*)'(x)\cdot\big(G'(y^*(x)) - G'(x)\big) \tag{*}$$

Moreover, from the envelope theorem:

$$(\lambda^*)'(x) = \frac{1}{\mathcal{L}G(y^*(x),x)}\cdot\Big(\lambda^*(x)\cdot G''(x) - V''(x)\Big)\cdot(y^*(x) - x) \Rightarrow$$

$$V''(x) - \lambda^*(x) \cdot G''(x) = -(\lambda^*)'(x) \cdot \frac{\mathcal{L}G(y^*(x), x)}{y^*(x) - x}$$

Substituting into (*) yields:

$$\frac{dy^*(x)}{dx} \cdot \Big(V''(y^*(x)) - \lambda^*(x) \cdot G''(y^*(x))\Big) = (\lambda^*)'(x) \cdot \left(G'(y^*(x)) - G'(x) - \frac{\mathcal{L}G(y^*(x), x)}{y^*(x) - x}\right) \Rightarrow$$

$$\frac{dy^*(x)}{dx} \cdot \Big(V''(y^*(x)) - \lambda^*(x) \cdot G''(y^*(x))\Big) \cdot (y^*(x) - x) = (\lambda^*)'(x) \cdot \Big(G'(x) - G'(y^*(x)) - G'(y^*(x)) \cdot \big(x - y^*(x)\big)\Big)$$

Re-arranging, and using the expression of λ^* in terms of the information gain we get:

$$-\frac{dy^*(x)}{dx} \cdot \big(y^*(x) - x\big) = \frac{\rho}{\kappa}(V - U)'(x) \cdot \mathcal{L}G\big(x, y^*(x)\big) \cdot \Big(\lambda^*(x) \cdot G''(y^*(x)) - V''(y^*(x))\Big)^{-1}$$

The term $\mathcal{L}G\big(x, y^*(x)\big) > 0$ by convexity of G. Additionally, the SOC necessitates:

$$V''(y^*(x)) - \lambda^*(x) \cdot G''(y^*(x)) \leq 0$$

Hence, we express the derivative as:

$$-\frac{dy^*(x)}{dx} = \frac{(V - U)'(x)}{y^*(x) - x} \cdot K(x, y^*(x)) \qquad (**)$$

with $K(x, y^*(x)) > 0$.

From Proposition 3, $(V - U)'(x)$ has the same sign as $y^*(x) - x$. Therefore, the RHS in (**) is always positive, which implies that:

$$\frac{dy^*(x)}{dx} < 0$$

We use this to show that precision $|y^*(x) - x|$ is increasing in the distance from the maximal information gain. Let $\hat{x}$ be a local maximum of $(V - U)$. Consider, $x_1 < x_2 < \hat{x}$. We have $y^*(x_1) > y^*(x_2)$, and:

$$|y^*(x_1) - x_1| = y^*(x_1) - x_1 > y^*(x_2) - x_1 > y^*(x_2) - x_2 = |y^*(x_2) - x_2|$$

where the first equality follows from Proposition 3. Similarly, for $\hat{x} < x_2 < x_1$, $y^*(x_2) > y^*(x_1)$, and

$$|y^*(x_1) - x_1| = x_1 - y^*(x_1) > x_1 - y^*(x_2) > x_2 - y^*(x_2) = |y^*(x_2) - x_2|$$

This proves the assertion. Since frequency and impact are inversely related at the optimum we have that:

$$(V - U)(x_1) < (V - U)(x_2) \Rightarrow |\hat{x} - x_1| > |\hat{x} - x_2| \Rightarrow \varphi^*(x_1) < \varphi^*(x_2)$$

which completes the proof.