When and what to learn in a changing world

César Barilla *

This version: July 3, 2025

Abstract

A decision-maker periodically acquires information about a changing state, controlling both the timing and content of updates. I characterize optimal policies using a decomposition of the dynamic problem into optimal stopping and static information acquisition problems. Eventually, information acquisition either stops or follows a simple cycle, with updates occurring at regular intervals and leading to consistent certainty levels; this enables precise characterizations of long run information acquisition across environments. In the limit as fixed costs vanish it is optimal to trade-off quality for frequency; surprisingly, this entails that both belief and action changes become lumpier. I highlight applications to portfolio diversification and asymmetries between safe and risky choices.

* Department of Economics, Columbia University; cesar.barilla@columbia.edu.

I am indebted to Yeon-Koo Che, Navin Kartik, Laura Doval, and Elliot Lipnowski for their guidance and support throughout this project and beyond. I am grateful to Hassan Afrouzi, Arslan Ali, Mark Dean, Alkis Georgiadis-Harris, Duarte Gonçalves, Jacopo Perego, Bernard Salanié, José Scheinkman, Ludvig Sinander, Akanksha Vardani, Nikhil Vellodi, Yu Fu Wong, Mike Woodford for their comments. I thank audiences at Columbia, Oxford, the 2024 Annual Meeting of the EAYE, the 2024 IO Theory Conference, the Paris IHP Game Theory Seminar, and the 2025 Canadian Economic Theory Conference for feedback.

1 Introduction

The world is constantly changing. Yet, most decision makers are not continuously upending their worldview. When making frequent decisions, it seems reasonable in the short run to act based on previously held beliefs or as if the relevant conditions are approximately fixed. However, past information eventually becomes outdated. Hence, periodically seeking to improve knowledge of current circumstances may be profitable, even if it is costly. This raises two natural questions: *when* should one decide to acquire new information and *what* should they learn?

Consider for instance the problem of an investor allocating resources between assets with uncertain returns. Market trends and the fundamentals that govern asset performance may change over time. How often and how thoroughly should the investor reconsider their current views? They could opt for infrequent but detailed research or frequent, less precise monitoring. The optimal balance between timing and quality of information acquisition depends on the stakes, information acquisition costs, and underlying volatility. In volatile markets, there is more to learn, but information becomes outdated faster. Higher certainty may be required for riskier investments, leading to quicker depreciation of information and higher costs. The investor's example is representative of a large class of problems: a government splitting budget between agencies with evolving needs, a producer choosing between available technologies, a retailer allocating inventory between locations with fluctuating demands, among others.

In this paper, I study a dynamic model of optimal information acquisition about a changing world, which provides a rich yet tractable way to capture the relation between the timing and content of infrequent information acquisition. A decision maker (DM) takes an action repeatedly at every instant; flow payoffs depend on their action choice and on an unobserved binary state of the world which changes over time. The DM sequentially chooses times at which they wish to acquire some information, which entails a fixed cost. At each such time they also flexibly decide what to learn, which entails a variable cost. The model is introduced in Section 2.

The first contribution of the paper is to rigorously solve the problem (Section 3). Well-known difficulties arise from the recursive nature of the value of information: the incentives to acquire information today simultaneously depend on all of the information the DM expects to acquire in the future and how the state changes; as a result, potentially complex learning dynamics induce nonlinear continuation values. However, the combination of the continuous time structure with infrequent information acquisition stemming from fixed costs makes this model tractable. I derive an appropriate Bellman equation that decomposes the DM's problem into a static optimal information acquisition problem and an optimal stopping problem. I show that the value function uniquely solves this equation even though the Bellman operator is not a contraction (Theorem 1). Optimal policies must consistently combine properties derived from optimal stopping and static information acquisition. This enables the precise characterization of optimal dynamic information acquisition (Theorem 2), which is described by two nested collections of belief intervals.

Second, I study the induced dynamics of information acquisition (Section 4). The main result is that optimal information acquisition must eventually either stop or settle into a simple repeating cycle (Theorem 3). In the cyclical case, information is acquired at regular intervals of time, when uncertainty reaches specific thresholds. Updates lead to two possible outcomes, captured by two "target posterior beliefs" reflecting endogenously chosen levels of relative confidence that one state is more likely than average. Each possible outcome leads to a waiting period of fixed length until the next update. In practice, this rules out more complex strategies with intermediary or irregular updates. This further simplifies the long run dynamics of optimal information acquisition: the problem can be reduced to choosing the *content* and the *frequency* of updates (Proposition 5); resulting expected payoffs have closed form expressions.

The convergence result enables the precise characterization and study of properties of long run optimal information acquisition. I show that optimal information acquisition may exhibit path dependency in the form of "learning traps" (Proposition 4): if initial beliefs belong to a "trap region" of sufficiently uninformed beliefs, then no learning ever occurs even if information would be regularly and perpetually acquired for higher initial levels of information. It also enables comparisons of information acquired under different policies or environments, which is a theoretically challenging question for general dynamic processes. I define three distinct types of "long-run informativeness" (Definition 3): (i) more informative experiments, (ii) lower uncertainty thresholds, and (iii) more frequent updates. I apply this definition to show that the world becoming more volatile has generically ambiguous effects on the frequency of information acquisition (Proposition 6). These results highlight that focusing on the long run neatly captures the dynamic incentives of repeated information acquisition, whereas short-run information acquisition obeys different incentives. I discuss intuition and examples which show that short-run information acquisition is less predictable and may exhibit counterintuitive patterns; for instance: optimally chosen information need not lead to a change in action with positive probability, unlike in static problems and in the long run.

Third, I study the limit of the model when fixed costs vanish, which prompts both methodological and qualitative contributions (Section 5). Technical issues arise as there is no unified formalism which explicitly allows for both discrete and continuous information acquisition (and the latter may become optimal without fixed costs). Nevertheless, I show that the definition of the problem can be naturally extended to allow for optimization over arbitrary belief processes, with and without fixed costs. This notably relies on defining a general cost function which extends and unifies costs from discrete information acquisition with existing cost specifications for continuous information acquisition. Solutions of the problem as fixed costs vanish are also shown to converge to solutions of the problems with no fixed cost even though the objective function is discontinuous at zero fixed costs. I derive an explicit characterization for optimal information acquisition in the limit as fixed costs vanish. Without fixed costs, it is optimal for the DM to either wait or acquire infinitesimal amounts of information to exactly confirm their current belief until rare news prompts a jump to a fixed alternative belief (Propostion 7). Optimal information must converge to a policy of this kind as fixed costs vanish (Proposition 8). In the long run, only two possible beliefs are ever held as the DM exactly prevents depreciation of information. Furthemore, the long run optimal belief process and the resulting value admit a closed form characterization which derives from the concavification of an appropriately defined "virtual net flow payoff" function (Theorem 5). The concentration on two beliefs delivers a new potential resolution of a tension between empirical observations and most models of dynamic learning (see e.g. Khaw et al. [2017]): decision makers do not continuously adjust their action, yet any changing world model in which beliefs are continuously changing and different beliefs imply different actions predicts continuous adjustment. While discrete action switches may be explained by frictions such as adjustment costs, the present model's prediction offers an alternative explanation based purely on optimal information acquisition. The ability to *flexibly and continuously* monitor the changing state (i.e without any discreteness either imposed or induced by fixed costs) allow the DM to optimally hold only two beliefs; *actions* in turn mirror the lumpy dynamics of belief.

Lastly, I provide some applications of the framework to concrete examples (Section 6). In a portfolio allocation problem between multiple assets with negatively correlated risk (Section 6.1), optimal behavior exhibits continuous rebalancing of the portfolio towards more diversification, punctuated by periodic shifts to a more extreme allocation. There may be information traps where initially uninformed investors are never able to acquire information and only ever buy a safe asset while informed ones retain better information and higher returns from risky assets. If fixed costs of information acquisition are negligible, there are no more information traps and there is no more continuous rebalancing: investors always hold risky portofolios and only adjust their allocation at discrete points in time. The frequency of adjustments is proportional to the underlying volatility of the environment. In a simple asymmetric problem with a safe and a risky action (Section 6.2), optimal information acquisition can generate distortions between "good" and "bad" news, typically leading to better quality but more frequent updating for information which suggests undertaking the risky action relative to a safer one – a behavior which may otherwise appear to qualitatively ressemble confirmation bias. This, as well as the non-monotonic dependence of frequency on the underlying volatility, can be connected to stylized facts from the literature on financial attention.

Related Literature

Explaining and studying the implications of imperfect adjustments to changing conditions has been a long standing theoretical and empirical agenda with important early contributions in macroeconomics and informational approaches gaining more attention in recent literature. Mankiw and Reis [2002], Reis [2006b,a] proposed an "inattentiveness" model where perfect observation of a changing state is subject to a fixed cost; further contributions expand on this framework and its quantitative implications, see e.g. Alvarez et al. [2011, 2016]. Similarly motivated by delays in macroeconomic adjustment processes, Sims [2003] proposed an alternative approach where agents flexibly acquire information subject to a limited processing capacity expressed in terms of reduction of entropy of beliefs. Subsequent literature on dynamic rational inattention (DRI) largely focuses either on environment with quadratic payoffs and gaussian states and information (see e.g. Maćkowiak and Wiederholt [2009], Maćkowiak et al. [2018], Afrouzi and Yang [2021a]), or on general implications of Shannon costs for induced random choice rules (Steiner et al. [2017]). Khaw et al. [2017] propose a discrete adjustment model which features quadratic payoffs, Shannon costs, and fixed costs to both information acquisition and adjusting actions, which they numerically solve and test against data from a laboratory experiment. The present paper combines a continuous time framework with fixed costs (which generates endogenously timed infrequent information acquisition, as in the first group of papers) with flexible information acquisition under a class of costs which generalizes entropy reduction. The limit case as fixed costs vanish relates to continuous time limits of DRI models; in this case, the endogenous timing of information acquisition generates some new predictions, notably in terms of action dynamics, relative to cases with an exogenously given discrete time grid (or exogenous constraints on action opportunities as in Afrouzi and Yang [2021b] or Davies [2024]).

Adjustment to a changing world also arises in the context of experimentation problems – see Whittle [1988] for a seminal reference and Che et al. [2024] for a recent contribution. Unlike with costly information acquisition, information in those problems is entangled with action choices. Although belief dynamics may bear some similarities, they have qualitatively different drivers and predictions since the agent is able to acquire information about alternative options without changing their action.[1] In the literature on social learning with a changing state (see e.g. Moscarini et al. [1998], Dasaratha et al. [2023]) learning occurs once and through observation of past actions; dynamics are driven by equilibrium forces rather than forward looking optimization. Within this literature, Lévy et al. [2022] features a similar environment as the present paper, as well as costly information acquisition and steady-state analysis.

The leading application of the model relates to information acquisition in finance. The interaction of strategic information acquisition and portfolio diversification is notably considered in a static setting by Van Nieuwerburgh and Veldkamp [2010]. The empirical and theoretical literature on "ostrich effects" documents more frequent monitoring after "good" news than after "bad" news (see for instance Karlsson et al. [2009], Galai and Sade [2006], Sicherman et al. [2016]); while this has natural interpretations as a behavioral bias, my model shows that costly information acquisition can generate similar patterns. Periodic inspections also arise in problems considering monitoring of strategic agents (e.g. Varas et al. [2020], Wong [2023], Ball and Knoepfle [2023]); however, in such contexts optimality is driven by incentive compatibility rather than informational motives.

Recent contribution on dynamic information acquisition with persistent states such as Che and Mierendorff [2019], Zhong [2022], Hébert and Woodford [2023], Georgiadis-Harris [2023] feature forms of similarly flexible but continuous information acquisition in the context of a one-shot decision. Since states are persistent and there is a single decision, incentives to acquire information over time in these models derive from convex costs (or budgets) over information flow. I consider costs linear in information flow, which entails that information is optimally acquired at most once in the persistent state limit; hence dynamics are fully driven by the changing state feature.

On a technical level, the present paper builds on the tools of static information acquisition and the related literature on communication and information design. Posterior separable costs are defined and studied in Caplin et al. [2022] and Denti [2022], and extend earlier work on Shannon entropy costs (Sims [2003]). The concavification method is developed in the literature on persuasion and communication; see Aumann et al. [1995] for an early use, Kamenica and Gentzkow [2011], Gentzkow and Kamenica [2014] for persuasion with and without costs, and Ely [2017] for dynamic persuasion integrating concavification and recursive analysis.

[1] See Lizzeri et al. [2024] for a recent paper analyzing differences between entangled and disentangled action and information choices with a persistent state.

2 Model

I begin by introducing the model, which formalizes the problem of an agent who makes decisions under uncertainty about an evolving state and chooses when and how to update their information.

Environment and decision problem Time is continuous and indexed by $t \geq 0$. There is a single decision maker (DM) who takes an action $a_t \in A$ at every instant in time; this generates flow payoffs which depend on the current action choice and the current value of a binary state of the world $\theta_t \in \Theta := \{0,1\}$. Denote by $\tilde{u} : A \times \Theta \to \mathbb{R}$ the utility function mapping actions and states to payoffs. Both the state and flow payoffs are unobserved.

The decision problem induces a convex *indirect utility function u*, which maps beliefs about the current state to expected payoff from the optimal action choice under those beliefs. Formally, denoting $\Delta(\Theta)$ the space of probability distributions over Θ, $u : \Delta(\Theta) \to \mathbb{R}$ is defined as:

$$u(p) := \max_{a \in A} \mathbb{E}_{\theta \sim p}\left[\tilde{u}(a,\theta)\right].$$

Assume an optimal action exists and u is continuous (e.g. A compact and $\tilde{u}$ continuous).

State transitions The state θ_t changes stochastically over time and follows Markovian dynamics: it jumps from 0 to 1 at rate $\lambda_0 > 0$ and from 1 to 0 at rate $\lambda_1 > 0$. Given that the state space is binary, the space of beliefs over Θ can be identified with the unit interval $[0,1]$, labeling beliefs in terms of the probability of the current state being 1. Markovian dynamics can be conveniently reparameterized in terms of the **total transition rate** $\lambda > 0$ and **invariant distribution** $\pi \in (0,1)$, which are formally defined as:

$$\lambda := \lambda_0 + \lambda_1, \text{ and } \pi := \frac{\lambda_0}{\lambda_0 + \lambda_1}.$$

Intuitively, π captures the long run average proportion of time that the state spends at 1; λ captures the total rate at which the state changes, which I will interpret as overall volatility.

Information acquisition and beliefs The DM chooses *when* to acquire information, and *what* information to acquire whenever they do. Formally, an information acquisition policy is described by sequences of (random) information acquisition times and information structures $\{\tau_i, F_i\}_{i \in \mathbb{N}}$ contingent on past information, where:

- $\{\tau_i\}_{i \in \mathbb{N}}$ are **information acquisition times**, i.e. $\tau_i \in \overline{\mathbb{R}}_+$ is the i-th time of information acquisition. The τ_i are a.s. increasing, strictly so when finite and $\tau_0 = 0$ by convention.

- $\{F_i\}_{i\in\mathbb{N}}$ are **information structures**, i.e. the content of signals being acquired at each τ_i. As is now standard in the information acquisition literature, each information structure is represented as a *probability distribution over posterior beliefs*: $F_i \in \Delta\Delta(\Theta)$ for all i.

The information acquisition policy $\{\tau_i, F_i\}_{i\in\mathbb{N}}$ induces the belief process $\{P_t\}_{t\geq 0}$. In between moments of information acquisition, beliefs about the current state drift towards the long run average π at an exponential rate controlled by λ: even in the absence of new information a Bayseian agent is aware that the hidden state might have changed. Fix some initial belief p and normalize the current time to 0; until the next update beliefs evolve according to:

$$dp_t = \lambda(\pi - p_t)dt, \quad \text{or equivalently: } p_t = e^{-\lambda t}p + \left(1 - e^{-\lambda t}\right)\pi.$$

Throughout the paper, I use lowercase p_t to denote the deterministic path of beliefs starting from $p_0 = p \in \Delta(\Theta)$ and reserve capital P_t for the overall belief process. In other words, if $P_{\tau_i} = p$ then $P_{\tau_i+t} = p_t$ for $t \in [0, \tau_{i+1} - \tau_i)$; at the next time of information acquisition, a new belief is drawn according to the corresponding experiment: $P_{\tau_{i+1}} \sim F_{i+1}$. Figure 1 illustrates possible belief dynamics.

The DM's information acquisition policy must be measurable with respect to the belief process and experiments must be Bayes plausible with respect to the current belief, i.e $F_i \in \mathscr{B}(P_{\tau_i^-})$ for all i, where:

$$\mathscr{B}(p) := \left\{F \in \Delta\Delta(\Theta) \,\middle|\, \int q\,dF(q) = p\right\}$$

is the set of feasible posterior distributions given current belief p. The rigorous construction of the belief process and the admissible class of controls is given in Appendix A.

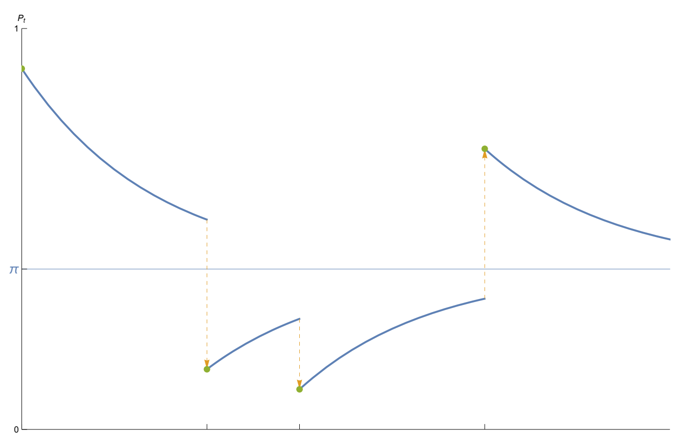


FIGURE 1: Possible belief dynamics
Beliefs continuously drift and periodically jump when information is acquired.

Information costs At any time when information is acquired, the DM incurs a fixed cost and a variable cost. Given an information acquisition policy $\{\tau_i, F_i\}$, at each time τ_i the DM pays a cost given by:

$$C(F_i) + \kappa$$

where $\kappa > 0$ is the fixed cost and the variable component $C : \Delta\Delta(\Theta) \to \overline{\mathbb{R}}_+$ maps information structures into (non-negative, potentially infinite) costs.

The variable component of information costs is **uniformly posterior separable** (UPS). Specifically, assume there exists a convex function $c : \Delta(\Theta) \to \overline{\mathbb{R}}_+$, finite and continuously differentiable over the interior of $\Delta(\Theta)$, such that for any $F \in \Delta\Delta(\Theta)$:

$$C(F) = \int_{\Delta(\Theta)} \big(c(q) - c(p)\big) dF(q) \text{ where: } p = \int q dF(q)$$

One natural interpretation of UPS costs is to see c as a "measure of certainty" at a given belief; hence $C(F)$ corresponds to the *expected increase in certainty* (reduction of uncertainty) induced by the chosen experiment relative to the current belief. See Frankel and Kamenica [2019] for a formalization of this interpretation. For more general references on UPS costs, their justification and relative merits or limits, see notably Caplin et al. [2022], Denti [2022], Denti et al. [2022]. Common choices for c include entropy (Shannon costs), negative variance, and the expected log-likelihood ratio (see notably Morris and Strack [2019], Pomatto et al. [2023]).

Optimal information acquisition problem The DM chooses an information acquisition policy so as to maximize total discounted expected utility under exponential discounting at rate $r > 0$. Hence, they solve the following **optimal information acquisition problem**:

$$v(p) := \sup_{\substack{\{\tau_i, F_i\}_{i \geq 0} \\ F_0 \in \mathscr{B}(p)}} \mathbb{E}\left[\int_0^\infty e^{-rt} u(P_t) dt - \sum_{i \geq 0} e^{-r\tau_i}\big(C(F_i) + \kappa\big)\right]. \tag{OIA}$$

The value function v captures the expected payoff from optimal information acquisition starting from initial belief p.

Remarks The assumption of binary states is made partly for expositional convenience, partly for tractability. The core of the approach as well as results about the structure of solution generalize to an arbitrary finite number of states; some qualitative properties of optimal dynamics only generalize in special cases. Section 7 discusses precisely the extent to which the main results generalize and qualitative departures.

Similarly, the recursive methodology generalizes beyond UPS costs (see Appendix B.1). The precise characterization of solutions, however, is dependent on this assumption. Because UPS costs are linear in the posterior distribution, they induce no intrinsic incentive to smooth information acquisition over time. Hence, they conveniently isolate the changing state as the sole source of dynamics in the model.[2] Section 7 discusses further the relative merits of UPS costs and alternatives.

3 Characterization of optimal policies

The characterization of solutions in the optimal information acquisition problem relies on a familiar dynamic programming approach, which suggests that the value function solves the recursive (Bellman) equation:

$$v(p) = \sup_{\tau \geq 0} \left[\int_0^\tau e^{-rt} u(p_t) dt + e^{-r\tau} \left(\sup_{F \in \mathscr{B}(p_\tau)} \int_{\Delta(\Theta)} v dF - C(F) - \kappa \right) \right]. \tag{$\star$}$$

Section 3.1 establishes this rigorously; Section 3.2 uses the implied decomposition between timing and content to derive the characterization of solutions.

3.1 Recursive equation and decomposition

Standard dynamic programming logic delivers a formal derivation as well as intution for ($\star$). To solve the problem from any starting belief, it suffices to focus on the next information acquisition time $\tau \geq 0$ and the corresponding information structure. Until τ, beliefs drift deterministically, and the DM accrues flow payoffs $u(p_t)$. At τ, the DM acquires information according to $F \in \mathscr{B}(p_\tau)$, incurring a cost $C(F) + \kappa$, and beliefs jump stochastically based on a draw from F. Given optimal behavior, the continuation value is the expected value function $\int v dF$, leading to the recursive equation ($\star$).

However, the Bellman operator in ($\star$) is not a contraction, hence one cannot appeal to standard methods to claim that v is its unique solution. To establish uniqueness, I decompose the Bellman equation into two operations and leverage the lattice structure of a suitably reduced domain to which the value function must belong.

Ex ante bounds on the value function Define the functions $\overline{v}$ and $\underline{v}$ as, respectively, the value from perfect costless observation of the state and from never getting any information

[2]In the persistent state limit, information would be acquired at most once.

about the true state, starting from an initial belief $p \in \Delta(\Theta)$ i.e.

$$\overline{v}(p) := \int_0^\infty e^{-rt}\mathbb{E}_{\theta\sim p_t}\left[\max_a u(a,\theta)\right]dt,$$

$$\underline{v}(p) := \int_0^\infty e^{-rt}u(p_t)dt.$$

Let $\mathbb{V}$ be the set of real-valued bounded measurable functions on $\Delta(\Theta)$ which are pointwise between $\underline{v}$ and $\overline{v}$. It can be directly verified that $v \in \mathbb{V}$. Hence I refer to $\mathbb{V}$ as the set of "candidate value functions" and restrict attention to this domain.

Decomposition and uniqueness The recursive equation can be decomposed into two parts: (i) the choice of an optimal information structure conditional on stopping, which reduces to an *"as-if-static" information acquisition problem* where the continuation values v itself plays the role of the indirect utility function; (ii) the choice of the optimal timing of information acquisition, which reduces to an *"as-if-one-off" deterministic optimal stopping problem* where the stopping payoff is given by value from instaneous information acquisition net of the fixed cost. The dynamic solution is uniquely characterized by consistency between solutions to both (sub)problems via the recursive structure: the value from information must incorporate future value from optimal timing; stopping payoffs must derive from future information acquisition. To formalize this logic, Definition 1 introduces corresponding functional operators.

Definition 1 (Recursive operators)**.** For any $w, g \in \mathbb{V}$ denote by

(i) $\mathcal{G}\, w$ the value from instantaneous information acquisition given continuation values w:

$$\mathcal{G}\, w(p) := \sup_{F\in\mathscr{B}(p)}\left[\int w\,dF - C(F)\right];$$

(ii) $\mathcal{W}\, g$ the value from optimal stopping given terminal payoffs g (under fixed cost κ):

$$\mathcal{W}\, g(p) := \sup_{\tau\in[0,\infty]}\left[\int_0^\tau e^{-rt}u(p_t)dt + e^{-r\tau}\big(g(p_\tau)-\kappa\big)\right];$$

(iii) Φ the composition operator capturing optimally timed one-shot information acquisition:

$$\Phi := \mathcal{W}\circ\mathcal{G}.$$

By definition v solves the recursive equation $(\star)$ if and only if it is a fixed point of Φ, i.e. $\Phi v = v$. The following result states that v is the unique such fixed point.

Theorem 1. *The value function v in the optimal information acquisition problem* (OIA) *is the unique solution to the recursive equation* ($\star$) *in* $\mathbb{V}$.

The formal proof is in Appendix B.1. As previously alluded to, Φ is not a contraction strictly speaking, hence one cannot apply the contraction mapping theorem. Instead, I rely on properties of the functional operators $\mathcal{G}$ and $\mathcal{W}$ to apply a Tarski-style fixed point theorem from Marinacci and Montrucchio [2019]. Namely, I show that $\mathcal{G}$ and $\mathcal{W}$ are monotone and order-convex operators over the lattice of real-valued bounded measurable functions over $\Delta(\Theta)$, which is a Riesz space, of which $\mathbb{V}$ is an order-interval. In addition, the value function is convex and differentiable; this partly carries over from static intuitions but involves some subtleties specific to the dynamic setting (see Appendix B.2).

Iterations of the fixed point operator Φ initalized at the lower bound $\underline{v}$ provide some economic intuition. Indeed, they correspond to the value in a *constrained* problem, where the DM is only allowed a finite number n of times of information acquisition, and converge to v as n goes to infinity. A symmetrical upper bound result holds, starting iterations from $\overline{v}$ instead, which can be interpreted as the solution of a *relaxed* problem: $\Phi^n\overline{v}$ represents the solution in a problem where the DM will be granted perfect observation of the state after the n^{th} time of information acquisition. Both results can be found in Appendix B.2.

3.2 Optimal policies

I now use the recursive equation ($\star$) to characterize optimal policies. Given conjectured continuation values w, the difference $\Phi w - w$ captures the *interim value from one-shot information acquisition*, i.e. the (signed) improvement in value from both optimal timing and content of the next update, relative to the direct continuation value w. It can be decomposed as:

$$\Phi w - w = \underbrace{\Phi w - \left(\mathcal{G}\, w - \kappa\right)}_{\text{stopping value}} + \underbrace{\mathcal{G}\, w - w}_{\text{gross static value of information}} - \kappa. \tag{1}$$

The first term captures the value from optimal stopping relative to "terminal" payoffs $\mathcal{G}\, w$. The second term captures the relative value from instantaneous information acquisition, gross of the fixed cost. Both are non-negative, which highlights that there is always weakly positive *gross* value in choosing timing and content of information acquisition.

The value function v in the optimal information acquisition problem is the unique candidate value that has exactly zero interim value of information acquisition everywhere. This is because v already incorporates all future value of information: hence v is unimprovable from one shot information acquisition ($\Phi v \geq v$) and is meanwhile an attainable continuation value ($v \geq \Phi v$).

At any p which triggers updating of information: the value must equal the payoff from stopping $v = \Phi v = \mathcal{G}\, v$, so the gross static value of information exactly equals the fixed cost $\mathcal{G}\, v - v = \kappa$.

These observations along with the decomposition previously outlined suggest the approach which delivers the characterization of solutions. First, establish independent properties of optimal stopping and static information acquisition. Second, combine the induced properties in the decomposition above. Consistency imposed by the recursive characterization in turn constrains values and policies. The general result gives geometric characterization in terms of the *net value function* $v-c$, and its concave envelope (denoted by $\mathrm{Cav}[v-c]$). This implicitly defines a simple class of policies.

Theorem 2. *Let v the unique fixed point of Φ. The information accquisition policy such that:*

1. *information is acquired whenever the gross value of information equals the fixed cost, which is described by the waiting time:*

$$\tau^*(p) := \inf\left\{t \geq 0 \,\middle|\, \mathcal{G}\, v(p_t) - v(p_t) = \kappa\right\},$$

2. *whenever information is acquired, the binary experiment supported over the two closest points at which there is no gross value of information ($\mathcal{G}\, v = v$) is chosen*

is optimal. Furthermore, the gross value of information is equal to the distance between the net value function $v - c$ *and its concave envelope* $\mathrm{Cav}[v - c]$. *Hence the optimal information acquisition policy is fully described by:*

1. *the region of the belief space where the net value function is strictly below its concave envelope, which is a countable collection of disjoint open intervals:*

$$\Gamma^* := \left\{p \in \Delta(\Theta) \,\middle|\, \mathrm{Cav}[v-c](p) > [v-c](p)\right\} = \bigcup_j (q_0^j, q_1^j),$$

2. *the* information acquisition region $\mathscr{I}^*$ *which is the set of beliefs where the net value function is at distance* exactly κ *from its concave envelope:*

$$\mathscr{I}^* := \left\{p \in \Delta(\Theta) \,\middle|\, \mathrm{Cav}[v-c](p) - [v-c](p) = \kappa\right\},$$

where endpoints of intervals in Γ^* describe optimal experiments and $\mathscr{I}^*$ the beliefs at which information acquisition occurs.

The statement of Theorem 2 is divided in two parts for clarity. The first part states the full optimal policy purely in terms of the *static* value of information induced by v. The second part gives an explicit form the value of information in terms of the net value function $v-c$, and uses this to reduce the description of optimal information acquisition to the choice of two regions.

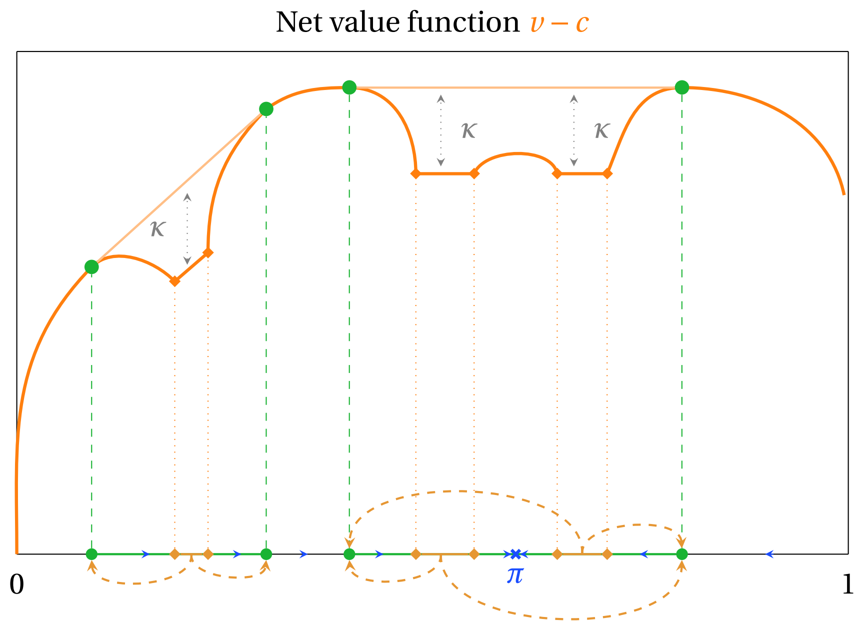


FIGURE 2: Geometrically solving for the optimal policy given the net value function $v-c$
The green intervals represent Γ (gross static value for information); the orange region represent $\mathscr{I}^*$ (information acquisition region); the green dots are the local "target" beliefs when information is acquired; blue arrowheads represent the direction of the drift and dashed orange arrows possible jumps.

This structure of optimal policies is best explained via a simple geometric visualization using the graph of $v-c$, which is illustrated in Figure 2. First, draw the *concave envelope* of $v-c$ and look for the region where it is strictly above $v-c$: this gives Γ^*, which is a collection of non-overlapping intervals. Each such interval defines a region where information *may* be acquired and the corresponding optimal binary experiment supported its endpoints. Within each interval, look for the subset of beliefs at which $v-c$ coincides with a shifted-down version of its own concave envelope, which in particular implies that $v-c$ is locally affine; note that this need not be an interval itself as it may have "holes".

The concave envelope property implies a somehow unusual restriction on $v-c$. In principle the shape $v-c$ is quite unrestricted (being the difference of two convex functions). For an arbitrary function w, $w-c$ can very well fall *strictly below* its own shifted down concave envelope. But since v must incorporate any profitable opportunity for information acquisition, $v-c$ must have flat truncated sections where it *exactly coincides* with its own shifted down concave envelope. In such sections, any strict convexity in v is coming purely from costs.

Important simplifications, both technical and intuitive, follow from the characterization of optimal experiments. The result combines three properties: (1) binary experiments are optimal, (2) within each interval where some experiment is profitable (ignoring the fixed cost), the same support is chosen and (3) the (convex hull of the) supports of optimal experiments do not overlap. The third property, i.e the decomposition of behavior into non-overlapping intervals, is most significant in ruling out certain complex dynamics information acquisition. The next section crucially relies on this fact to derive induced dynamics of beliefs, which in turn provides a clearer understanding of the structure of policies.

The proof of Theorem 2 relies on the combination of three results. First, Proposition 1 characterizes optimal experiments for arbitrary continuation values, which yields the sufficiency of binary experiments and a the interval decomposition, as well a general version of the concave envelope characterization. Second, Proposition 2 characterizes optimal timing of one shot information acquisition, taking as a given the value from optimal information acquisition. Putting it together is justified by a general verification result characterizing all optimal policies in terms of solutions in the Bellman equation (Proposition B.5 in Appendix B.3). I briefly state the first two results to give further intuition and because they provide more general statements that can be applied to arbitrary continuation values; further details are relegated to Appendix B.3.

Optimal experiments and the continuation value operator. Implications of uniformly posterior separable costs for *static* information acquisition are now well studied. The following proposition states a collection of key properties in the context of this model. These can be proven using, for instance, results in Caplin et al. [2022], Gentzkow and Kamenica [2014], Dworczak and Kolotilin [2024].

Proposition 1. *Consider an arbitrary continuous function w over $\Delta(\Theta)$.*

(i) ***Value of information via concave envelope:***

$$\mathcal{G}\, w(p) = \mathrm{Cav}[w - c](p) + c(p),$$

where Cav *denotes the concave envelope (smallest majorizing concave function).*

(ii) ***Geometric characterization of optimal experiments:*** *any optimal experiment at $p \in \Delta(\Theta)$ is supported over points where the supporting hyperplane of the convex hull of the subgraph of $w - c$ at p meets the graph of $w - c$ and conversely, any Bayes-Plausible experiment supported over those points is optimal.*

(iii) ***Sufficiency of binary experiments:*** *for any $p \in \Delta(\Theta)$ there must exist some optimal experiment which induces at most $|\Theta| = 2$ possible distinct posteriors (i.e. $|\mathrm{supp}(F_p)| \leq 2$).*

(iv) ***Optimality within intervals:*** *let F an optimal binary experiment at p, denote* $\text{supp}(F) =: \{q_0, q_1\}$*; then for any $q \in (q_0, q_1)$, the unique experiment in $\mathscr{B}(q)$ supported over* $\{q_0, q_1\}$ *is optimal at q.*

The concavification method for solving optimal information acquisition problems with uniformly posterior separable cost follows a geometric intuition. For a given belief p, the optimal experiment is derived by identifying the chord between two points of the graph of $w - c$ that attains the highest value at p. This follows from the problem's linear structure:

$$\mathcal{G}\, w(p) - c(p) = \sup_{F \in \mathscr{B}(p)} \int [w - c]\, dF.$$

The induced "target" beliefs determine a unique experiment via Bayes-plausibility (binary experiments can be identified by their support). If q_0 and q_1 are optimal from p, they remain optimal for any $q \in (q_0, q_1)$, thus partitioning the belief space into intervals where the optimal experiment is supported at such (local) endpoints, with the uninformative experiment optimal elsewhere.

Point (ii) implies that some informative experiment is conditionally optimal when $\text{Cav}[w - c](p) > [w - c](p)$, i.e. when the gross static value $\mathcal{G}\, w(p) - w(p)$ is strictly positive. This defines the region $\Gamma[w]$, which decomposes into disjoint intervals (q_0^j, q_1^j); these intervals also describe the induced optimal experiments. The information acquisition region $\mathscr{I}[w]$ is defined as the beliefs where the *net* static value $\mathcal{G}\, w(p) - w(p) - \kappa$ is non-negative; it is a subset of $\Gamma[w]$. Thus, the belief space is hierarchized into regions where information is either never acquired, potentially acquired, or actually acquired based on the net value being sufficiently high.

STOPPING VALUE AND DYNAMIC VALUE OF INFORMATION ACQUISITION. Turning to optimal timing, recall some classical facts about optimal stopping (see e.g. Peskir and Shiryaev [2006] for a textbook reference).

Proposition 2. *Let w a candidate value function. The value in the problem of optimal timing of one-shot information acquisition with continuation value w, which is given by $\Phi w = \boldsymbol{\mathcal{W}}(\mathcal{G}\, w)$, is the unique solution $\tilde{w}$ to:*

$$\min\Big\{ r\, \tilde{w}(p) - u(p) - \lambda(\pi - p)\, \tilde{w}'(p), \tilde{w}(p) - \mathcal{G}\, w(p) \Big\} = 0.$$

The stopping policy defined by:

$$\tau(p) := \inf\Big\{ t \geq 0 \;\Big|\; \Phi w(p_t) = \mathcal{G}\, w(p_t) \Big\}$$

is optimal.

The "stopping value" is the difference between the value from choosing the optimal stopping time and the stopping payoff, i.e., $\Phi w(p) - \mathcal{G}\, w(p)$ for any candidate value w and belief p. Define the stopping region $\mathscr{S}[w]$ as the set of beliefs where the stopping value is zero, i.e., where immediate stopping is optimal:

$$\mathscr{S}[w] := \left\{p \in \Delta(\Theta) \,\middle|\, \Phi w(p) = \mathcal{G}\, w(p)\right\}.$$

Note that $\mathscr{S}[w]$ need not equal $\mathscr{I}[w]$ for arbitrary w: it may be optimal to pay the fixed cost for the continuation value without acquiring information, or there could be net value for information without immediate stopping. For the true value function v, these regions must coincide: $\mathscr{I}[v] = \mathscr{S}[v]$, meaning stopping occurs when and only when acquiring information is beneficial.

Remark. In the remainder of the paper, optimal information acquisition refers to the optimal policy in Theorem 2. This policy is always well-defined, but cannot in general be guaranteed to be unique. In case of multiplicity, it corresponds to the selection of the earliest optimal stopping time and the least informative optimal experiment. The description via the interval decomposition defines a class of strategies which are relatively simple to describe and compute for the decision maker and the modeler. The next section establishes that their dynamic behavior sparsely captures the DM's incentives in periodic information acquisition.

4 Dynamics of information acquisition

The decomposition in Theorem 2 provides a simple way to visualize the evolution of beliefs and enables a precise description of long-run behavior. In this section, I first establish (Theorem 3 in Section 4.1) that learning must either stop in finite time or settle into a simple cyclical pattern. Then, I turn to characterizing conditions under which learning stops (Proposition 4 in Section 4.2) and study the existence of path dependent "learning traps". The results deliver explicit expressions for stationary payoffs and an equivalent simplified auxiliary problem, optimizing directly over long-run behavior (Section 4.3). This provides a way to compare informativeness accross environments (Section 4.4).

4.1 Convergence to cyclical information acquisition

Belief cycles. Under the optimal strategy, if at any point an experiment is chosen which leads to possible posterior beliefs q^0, q^1 on opposite sides of the long run average π ($q^0 < \pi < q^1$), then all future information acquisition leads to the same two posteriors. Indeed, beliefs drift towards π following information acquisition, hence they can only move inside of the interval (q^0, q^1). By Theorem 2, the same support must remain (conditionally) optimal for the next experiment (see Figure 2 for an illustration). This implies that the time between updates is simply the waiting time between either q^0, q^1 and the closest belief towards π which lies in the information acquisition region $\mathcal{I}$. In other words, if information acquisition is supported on beliefs which suggest that a different state is more likely than average, then beliefs must enter simple cyclical dynamics, where a fixed time between updates lead to restoring one of two possible levels of relative confidence that one state is more likely than average.

Definition 2 below formalizes this notion of "belief cycles", which is useful to state results concisely. It reduces description of periodic information acquisition to the fixed content of udpdates (a pair of target beliefs) and the thresholds beliefs for information acquisition – or, equivalently, conditional wait times. Figure 3 gives a representation of such cyclical dynamics.

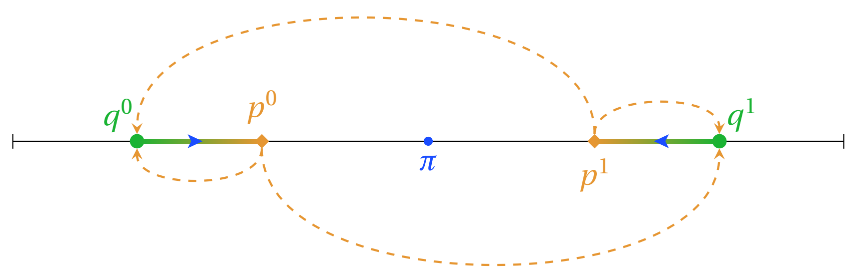


FIGURE 3: Representation of a belief cycle

Definition 2. A *belief cycle* Υ is a tuple $\Upsilon = \Big((q^0, q^1), (p^0, p^1), (\tau^0, \tau^1)\Big)$ composed of target beliefs (q^0, q^1), threshold beliefs (p^0, p^1), and waiting times (τ^0, τ^1) such that:

$$0 \leq q^0 \leq p^0 \leq \pi \leq p^1 \leq q^1 \leq 1$$
$$q^i_{\tau^i} = p^i \iff \tau_i = \frac{1}{\lambda}\log\left(\frac{\pi - q^i}{\pi - p^i}\right) \quad \text{for } i = 0, 1$$

It is called *non-degenerate* if $q^0 < p^0 < \pi < p^1 < q^1$ or, equivalently, $q^0 < \pi < q^1$ and $\tau^0, \tau^1 > 0$

A belief cycle parametrizes the law of motion of beliefs within the induced domain: information is acquired when beliefs reach p^0 or p^1, the resulting update triggers a jump to q^0 or q^1; the DM waits τ^0 or τ^1 respectively until the next update (see Figure 3). I refer to $[q^0, p^0] \cup [p^1, q^1]$ as the *domain* of the belief cycle.

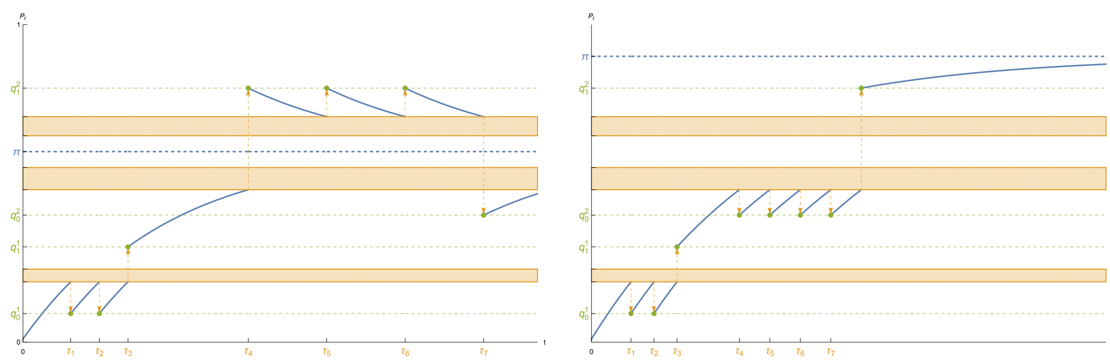
Observe that τ^θ (for $\theta \in \{0, 1\}$), which captures the amount of time that the DM waits after an "θ-news", is the time it takes to drift from q^θ to p^θ when acquiring no information. This description of the belief cycle therefore has some redundancy, but is convenient for completely describing long-run information acquisition behavior. It also implicitly encodes the dependence on parameters (λ, π), which allows for easier comparisons of optimal policies across environments (see Definition 3 and Section 4.4).

Convergence. It turns out that information acquisition must eventually either reach a belief cycle or stop altogether. Figure 2 and the examples of optimal belief dynamics in Figure 4 below provide intuition into the underlying logic: if both posteriors from a given experiment are on the same side of π, then if beliefs jump *towards* π they will only drift closer to π until they reach an experiment which triggers cyclical dynamics (if there is one).

Theorem 3. *Let $\{P_t\}$ the belief process deriving from optimal information acquisition. There exists an almost surely finite time $T \geq 0$ after which either:*

- **(A)** ***Learning stops:*** *no information is acquired, or*
- **(B)** ***Cyclical updates:*** *P_t follows dynamics described by a non-degenerate belief cycle.*

If learning stops, beliefs converge to their long-run average: $P_t \xrightarrow[t\to\infty]{a.s.} \pi$.



(A) Convergence to cyclical dynamics (B) Learning eventually stops

FIGURE 4: Examples of realized belief dynamics

The proof would be fairly intuitive if Γ^* could be guaranteed to have a finite numbers of disjoint intervals. However, there is no such guarantee: there could be infinitely many intervals, and these intervals could even accumulate at a point, so the fact that it must take a finite amount of time to cross each interval in Γ^* is not sufficient. To show convergence without further assumptions, I use Kolmogorov's 0-1 Law instead; details can be found in Appendix C.1.

The theorem implies that the DM must eventually settle on *periodic updates* with *fixed content*. This rules out many more complex yet perfectly reasonable behavior – for instance, any policy which would involve frequent "smaller" updates to check whether a more substantial (and infrequent) reassessment should occur. In some sense, optimal information acquisition is "bunched": the DM need only choose the unconditional *content* of information acquisition and its *frequency*.

The reduction of the problem to a choice over frequency and quality extends intuition from static information acquisition. In the static problem, it is optimal that the outcome of information acquisition optimally concentrates on two possible beliefs, each suggesting one state being relatively more likely. The long run solution extends this structure. It remains optimal to eventually concentrate all information acquisition on two possible outcomes, which must now suggest either state being more likely than in the long run average. The substantial difference is that the value of information stems from the periodic repetition of that experiment: the benefit from information acquisition is the sum of short-run improvements in decisions. Hence, it is necessary to specify not just content but also frequency. Furthermore, the choice of both objects is entangled: the periodicity determines not only the occurence of costs but also the time horizon (hence short-run value) from each update. This makes transparent the generally opaque recursive nature of the value of information acquisition, while providing a precise and tractable description of its determinants (see Section 4.3).

SHORT AND LONG RUN The convergence result in Theorem 3 illuminates the incentives behind short-run (i.e. non-cyclical) information acquisition. Suppose the DM has solved for the optimal stationary dynamics: they know that after entering some domain $[q^0, p^0] \cup [p^1, q^1]$, it will be optimal to remain in the corresponding cycle. Consider some starting belief p to the left of q^0, and assume the DM considers at most one experiment region outside of the stationary domain. The DM's goal is now to choose a stopping belief between p and q^0 and a local experiment supported by two posteriors, both to the left of q^0. If the leftmost posterior is realized, the DM repeats a short-run cycle; if the right posterior is realized, they anticipate drifting toward stationary dynamics and eventually receiving the known continuation value $v(q^0)$.

Paying the cost for this short-run experiment is essentially a gamble over the timing of convergence: it might extend time spent at higher certainty beliefs or jump faster to the steady state regime, which has lower certainty. This logic suggests an intuitive approach for decomposing optimal information acquisition: first solve for optimal stationary behavior, then solve for the nearest short-run acquisition region, and iterate outward, treating each inward continuation value as a fixed input from the previous problem.

This thought experiment highlights why short-run incentives differ from static intuition and may lead to counterintuitive patterns. For example, in static settings, information is only acquired if it leads to an action change—no experiment is optimal if it doesn't change the decision.

However, in dynamic settings, this isn't always true. Information may be acquired in the short run even without an immediate action switch because it alters the timing of action switches in the long run, which can be valuable. This phenomenon vanishes in the long run: periodicity implies that no information is acquired without action switches – otherwise, this would imply costs without benefits in the auxiliary problem, in line with the static logic.

ERGODIC DISTRIBUTION OF BELIEFS. The result on long run belief dynamics also enables the following characterization of the *ergodic distribution of beliefs* (formally derived in Appendix C.2).

Proposition 3. *Let μ the ergodic distribution of beliefs under optimal information acquisition, identified with its density. Assume information acquisition does not stop under the optimal policy and denote $[q^0, p^0] \cup [p^1, q^1]$ the support of the long run belief cycle. Then:*

$$\mu(p) = \begin{cases} \frac{1}{2}\frac{1}{p^0 - q^0} & \text{if } p \in [q^0, p^0] \\ \frac{1}{2}\frac{1}{q^1 - p^1} & \text{if } p \in [p^1, q^1] \end{cases}$$

This result expresses all objects of interest (spread of beliefs, average time to the next update,...) are in terms of the thresholds, which is useful for comparative statics. The ergodic distribution can be interpreted as the eventual distribution of beliefs within a population of identical decision-makers. This, in turn, enables using the model to study the effect of optimal information acquisition on population parameters, e.g. the spread and balance of beliefs within the population. Relating these outcomes to the model's primitives allows to address broader questions such as whether higher information costs lead to more or less disagreement.

4.2 Learning traps

When do optimal dynamics induce learning to stop in finite time? The following proposition gives simple conditions depending on whether or not there is gross value for information at π. It also highlights that the qualitative form of long run dynamics may depend on the initial belief.

Proposition 4. *Under optimal dynamics:*

(i) Learning stops in finite time only if π is not in the information acquisition region ($\pi \notin \mathscr{I}^*$).

(ii) If there no gross value for information at π (i.e. $\pi \notin \Gamma^*$), information is acquired at most a finite number of times.

(iii) If there is gross but not net value for information at π ($\pi \in \Gamma^* \setminus \mathscr{I}^*$), then either:

a. information acquisition is acquired at most a finite number of times for all priors, or

b. *there exists some open interval* $(\underline{p}, \overline{p})$ *such that no information is ever acquired for any prior* $p_0 \in (\underline{p}, \overline{p})$ *but any prior* not *in* $(\underline{p}, \overline{p})$ *leads to a belief cycle in the long run.*

Furthermore case (iii.b) occurs if and only if there exists some beliefs in the information acquisition region to the left and to the right of π *which both lead to the same conditionally optimal target beliefs as at* π*. Refer to the (possibly empty) interval* $(\underline{p}, \overline{p})$ *as the "trap region".*

I illustrate the logic behind Proposition 4 in Figure 5. Recall typical beliefs dynamics from Figures 2 and 4: whenever beliefs jump towards the long run average π, they can only get closer to π for ever after. Whether learning stops depend on whether the belief process drifts into points in the information acquisition region, coming from any side of π. If so, this triggers a cycle; otherwise, learning eventually stops. Proposition 4 follows this logic to delineate necessary and sufficient conditions. In particular problems, one can often determine which case obtains without fully solving the DM's problem. Indeed, one only needs to know the best attainable payoffs from *some* cycle initiated at or around π and its comparison with the fixed cost. For instance: if some cycle is profitable (taking into account the initial fixed cost), then it must be than $\pi \in \mathscr{I}^*$ and all priors lead to periodic information acquisition in the long run; if the same is true in gross but not net value, there is a hole and path dependency arise; etc.

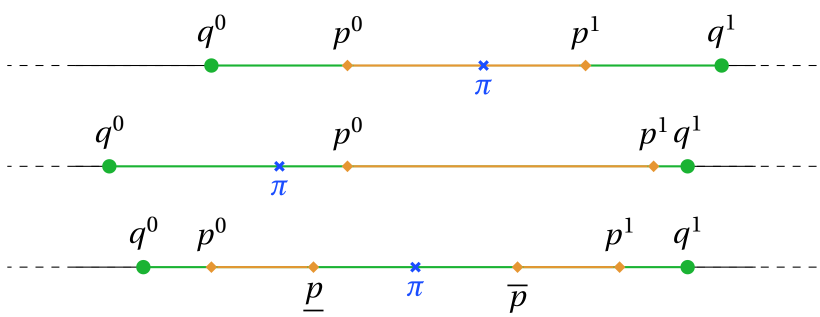


FIGURE 5: Possible cases from Proposition 4
The green interval represents Γ^* and the orange region is $\mathscr{I}^*$ (locally around π).
Case 1 leads to a belief cycle for all priors. Case 2 leads to learning stopping for all priors. Case 3 leads to a cycle for all priors outside of $(\underline{p}, \overline{p})$ and no information acquisition otherwise.

The most interesting case is the last one, (iii.b), in which path dependency arises. Whereas in all other cases, all beliefs lead to the same long run outcome (either a cycle or a finite number of updates), when there is a "hole" around π all initial beliefs lead to an eventual cycle except those close enough to π which lead to never acquiring information. Path dependency takes a stark form: no information is ever acquired because the DM started with uncertain enough beliefs. Even though there is gross value of information at π, it is not sufficiently high to warrant paying the initial cost of cyclical information acquisition. Even though this results from optimality, such path dependency might have welfare consequences in richer environments, e.g. in the presence of externalities from information acquisition or if inequality in how informed agents are is intrisically undesirable.

4.3 Stationary payoffs

Having established eventual convergence to cyclical dynamics, one can explicitly characterize stationary payoffs. This further enables a reduction of the problem to a choice over stationary dynamics, ignoring short-run behavior, as was previously suggested

STATIONARY PAYOFFS. Consider an arbitrary belief cycle $\Upsilon = \big((q^0, q^1), (p^0, p^1), (\tau^0, \tau^1)\big)$ and the associated stationary dynamics. Let v^0, v^1 the induced values for a DM placed in those stationary dynamics starting from q^0, q^1 respectively. From the recursive problem, they must verify:

$$\begin{aligned} v^0 &= \int_0^{\tau^0} e^{-rt} u(q_t^0) dt + e^{-r\tau^0} \left(\frac{q^1 - p^0}{q^1 - q^0} \big(v^0 - c(q^0)\big) + \frac{p^0 - q^0}{q^1 - q^0} \big(v^1 - c(q^1)\big) + c(p^0) - \kappa \right) \\ v^1 &= \int_0^{\tau^1} e^{-rt} u(q_t^1) dt + e^{-r\tau^1} \left(\frac{q^1 - p^1}{q^1 - q^0} \big(v^0 - c(q^0)\big) + \frac{p^1 - q^0}{q^1 - q^0} \big(v^1 - c(q^1)\big) + c(p^1) - \kappa \right) \end{aligned} \tag{CP}$$

The system above has a unique solution, which can be written out explicitly; its cumbersome expression is omitted here but can be found in Appendix C. This solution can be used to explicitly define the value from only engaging in cyclical information acquisition, $J(p, \Upsilon)$, which is formally derived from the following strategy:

- if $p \in [q^0, p^0] \cup [p^1, q^1]$, follow the belief cycle Υ starting from p
- if $p \in [0, q^0) \cup (q^1, 1]$, let beliefs drift until reaching q^0, q^1, then follow Υ
- if $p \in (p^0, p^1)$, immediately jump to q^0 or q^1, then follow Υ

This amounts to neutralizing short-run information acquisition: the DM is only allowed to acquire information as prescribed by a cycle, at most drifting first into the cycle.

AUXILIARY STATIONARY PROBLEM. The following result shows that the eventual stationary behavior from the full problem must coincide with the optimal belief cycle for J. In other words, it is without loss to ignore short-run information acquisition and solve the constrained problem which restricts all information acquisition to a single cycle. We can also use J to determine whether or not learning stops and the "trap" regions, if any.

Proposition 5. *For any p, the solution of the following problem:*

$$\max_{\Upsilon \text{ belief cycle}} J(p, \Upsilon)$$

coincides with the optimal belief cycle in the dynamic information acquisition problem. Furthermore, belef p is contained in the trap region $(\underline{p}, \overline{p})$ if and only if $\max_{\Upsilon} J(p, \Upsilon) < \underline{v}(p)$.

A convenient starting point in the above problem is π. Since beliefs do not drift from π, the choice is between getting payoffs $u(\pi)$ forever or jumping into the best feasible cycle:

$$v(\pi) = \max\left\{u(\pi),\ \max_{\Upsilon} \frac{\pi - q^0}{q^1 - q^0}\big(J(q^1, \Upsilon) - c(q^1)\big) + \frac{q^1 - \pi}{q^1 - q^0}\big(J(q^0, \Upsilon) - c(q^0)\big) + c(\pi) - \kappa\right\}$$

The main purpose of the auxiliary problems is to give a convenient non-recursive form to directly study properties of optimal long-run behavior.

4.4 Comparing informativeness across environments

INFORMATIVENESS COMPARISONS In general, there is no single or simple way to compare how much information is acquired via *dynamic* processes. Part of the difficulty lies in the comparibility of time dependent behavior. However, the structure of solutions in this setting suggests an intuitive approach. Informally, there are three natural criteria that capture "more information" being acquired: static informativeness of experiments, uncertainty thresholds triggering information acquisition, and frequency. The following definition formalizes these notions directly over long run belief cycles (and is illustrated in Figure 6 below).

Definition 3. Consider cycles $\Upsilon = ((q^0, q^1), (p^0, p^1), (\tau^0, \tau^1))$ and $\tilde{\Upsilon} = ((\tilde{q}^0, \tilde{q}^1), (\tilde{p}^0, \tilde{p}^1), (\tilde{\tau}^0, \tilde{\tau}^1))$ under possibly different respective environment $(\lambda, \pi), (\tilde{\lambda}, \tilde{\pi})$. Say that:

(i) Υ has more informative experiments than $\tilde{\Upsilon}$ if: $(\tilde{q}^0, \tilde{q}^1) \subset (q^0, q^1)$;

(ii) Υ has lower uncertainty thresholds for information acquisition than $\tilde{\Upsilon}$ if: $[\tilde{p}^0, \tilde{p}^1] \subset [p^0, p^1]$;

(iii) Υ has more frequent information acquisition than $\tilde{\Upsilon}$ if: $\tau^i \leq \tilde{\tau}^i$ for $i = 0, 1$.

The first notion captures that all experiments conducted in one belief cycle are Blackwell more informative than the ones in the other cycle.[3] These three notions induce distinct partial orders; they may not agree in ranking two policies. In particular, note that when varying λ or π the frequency comparison is not implied by the comparisons in terms of beliefs. This definition also illustrates how the cycle decomposition can be used to define behavioral features tailored

[3]The interval condition exactly states that the conditionally experiments under Υ are mean-preserving spreads (fixing a given belief) of those under $\tilde{\Upsilon}$.

to applications – for instance, asymmetric shifts may have natural interpretations in some contexts.

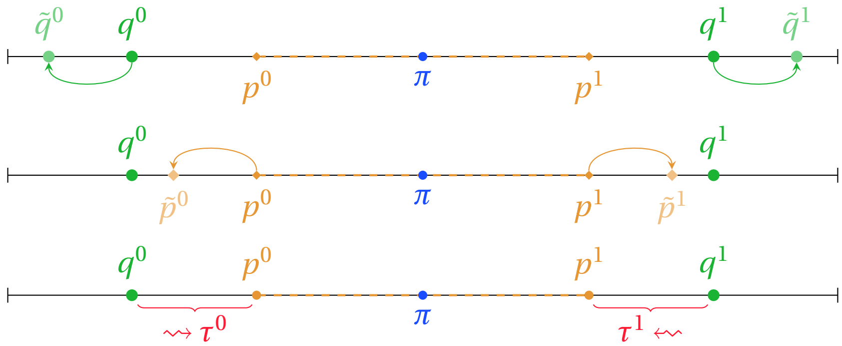


FIGURE 6: Visualizing the three notions of "more information acquisition in the long run"

SYMMETRIC PROBLEMS Previous characterizations drastically simplify in the class of problems which are invariant to a relabeling of the states, which I refer to as *symmetric problems*. Symmetric problems admit a useful dimensionality reduction: given the analysis of the previous section, the quality of information reduces to a one-dimensional object (distance of beliefs from the long-run average); this in turn enables an explicit representation of the problem in terms of frequency and quality. Symmetry is less restrictive than it might seem: natural cost functions (entropy, variance, log-likelihood ratio) are symmetric and normalizations may transform non-symmetric into equivalent symmetric problems. For instance, flow payoffs and costs only need to be symmetric up to adding an affine function – hence, for example, a choice between two actions is equivalent to some symmetric flow payoff, provided that the indifference belief is uniform.

Naturally, if the problem is symmetric so is the value function: $v(p) = v(1-p)$. This immediately entails that optimal information acquisition also has symmetry properties; in particular, when focusing on the long run, it must be that $q^0 = 1 - q^1$ and $p^0 = 1 - p^1$, hence $\tau^0 = \tau^1$. Abusing notations slightly, the description of the belief cycle can reduce to $\Upsilon = (q^*, p^*, \tau^*)$ with $1/2 \leq p^* \leq q^*$ and $q^*_{\tau^*} = p^*$. Here, p^* is the threshold level of certainty which triggers information acquisition, q^* is the target level of certainty that information acquisition aims to achieve, and τ^* the corresponding waiting time.

In symmetric problems, the characterization of cyclical payoffs values in (CP) is substantially simpler. Cyclical payoffs for $\Upsilon = (q^*, p^*, \tau^*)$, expressed at q^*, are given by:

$$J_S(\Upsilon) := \left(1 - e^{-r\tau^*}\right)^{-1} \left(\int_0^{\tau^*} e^{-rt} u(q^*_t) dt - e^{-r\tau^*} \left(c(q^*) - c(p^*) - \kappa \right) \right). \tag{SCP}$$

The above formula highlights the intertwined effects of quality and frequency of information acquisition. Over repeated periods of length τ^*, flow payoffs are accumulated corresponding to discounted indirect utility on the interval $[p,q]$. The information cost $c(q^*)-c(p^*)-\kappa$ is being paid at the *end* of the period to restore the certainty level q from its depreciated level q^*_τ. Each cycle is discounted over infinitely many discrete periods of length τ^*; however unlike with a fixed grid, the period length (hence discount factor $e^{-r\tau^*}$) is endogenous. The measure of "how much information" is acquired condenses to these three simple quantities (q^*, p^*, τ^*); hence each of the previous notions of informativeness each become a complete order.

VOLATILITY AND INFORMATIVENESS. Two counterveiling forces arise when the world becomes more volatile. There is more to learn, because the circumstances of decision making change faster; this suggests that the decision maker should pay more attention overall if they want to maintain a comparable level of accuracy of their decision. The benefits of information acquisition are more transient: since the state changes faster, information acquired at a given time becomes obsolete much faster; this makes information acquisition less profitable, which suggests the decision maker should acquire less information. Neither force ever dominates overall: the following proposition establishes that the frequency of information acquisition is non-monotonic in λ.

Proposition 6. *Consider a symmetric problem. Let $\tau(\lambda)$ the time between moments of information acquisition as a function $\lambda > 0$, fixing other parameters. Then:*

1. $\lim_{\lambda\to 0}\tau(\lambda)=\infty$*; furthemore for λ close enough to 0, τ is decreasing.*
2. *There exists $\overline{\lambda}$ such that for all $\lambda \geq \overline{\lambda}$, $\tau(\lambda)=\infty$; furthemore for λ smaller but close enough to $\overline{\lambda}$, τ is increasing.*

Low levels of volatility correspond to small departures from states being perfectly persistent towards a changing world; in this case the forces pushing for more frequent tracking dominate. At the other extreme, the cost of tracking a highly volatile state starts to become prohibitive: information degrades too fast. The DM eventually saturates their capacity to profitably track the state and starts gradually "giving up".

Precise comparative statics in λ are delicate, and regularity is not to be expected as volatility shapes payoffs through several distinct channels. As a thought experiment, fix a candidate belief cycle (p,q). Increasing λ decreases decreases payoffs along each cycle as faster paths from q to p spend more time towards lower values, and total discounted costs increase as each period gets shorter. Yet, the lifetime value of cycles also increases because they become more frequent.

Further, a quicker cycle lowers the "discrete" discount rate $(1-e^{-r\tau})$ which may improve payoffs if the previous effects are not too large.

In examples, being in a more volatile world seems to often lead to a decrease in the *quality* of information acquired and higher thresholds of uncertainty. This is not unintuitive: in a fast-changing world, seeking only small improvements in certainty over the long run average seems reasonable; yet it may be that it is optimal to acquire more unprecise information but compensate with more frequent verifications – leading to potentially no worse overall relative tracking of the state. This intuition suggests that in a more volatile world the overall cycle compresses around the long run average π, leading to *smaller* deviations of beliefs.

The only clear regularity, however, is that the change is frequency is always non-monotonic; more specifically, frequency decreases for small enough levels of volatility and eventually increases for high enough levels of volatility (getting close to the finite level such that information acquisition is no longer valuable). Even the previous conjecture would not pin down further properties – and for instance, τ^* is not in general single-peaked in λ. Indeed, recall that in the expression for the time between successive moments of information acquisition, it is subject to three distinct forces and results in behavior with few restrictions (even if p and q move co-monotonically). Examples can be produced where the frequency of information acquisition *oscillates* as the world becomes more volatile. Cyclicality gives a clear descriptive handle on information acquisition, yet there is enough richness in the model to produce wide variations in observed patterns even within well-behaved classes of cost and payoff functions.

5 Vanishing fixed costs

I now study the behavior of optimal information acquisition in the limit as fixed costs vanish, which raises some technical challenges but yields qualitative as well as methodological insights. This can be motivated either as an approximation to "small" fixed costs or to allow for possibly continuous information acquisition if fixed costs are literally absent.

I first provide (Section 5.1) informal intuition for the main result: in the vanishing fixed cost limit, after the initial instant it is always optimal to either *wait* or acquire information so as to precisely *confirm* the current belief until rare news prompts jump to a (fixed) new belief. This corresponds to a Poisson signal structure under which the absence of news exactly maintains the current belief (compensates the drift). The result is suggested by intuition that, as κ goes to zero, threshold and target beliefs should get "close to one another". But although this logic does give valuable insight into the mechanics of convergence, it cannot deliver rigorous arguments.

It nonetheless serves to illustrates the two kinds of difficulties that arise in tackling limits: first, directly characterizing the limit of dynamic information acquisition policies is challenging; second, it is a priori unclear how to define the "$\kappa = 0$ limit problem" and formalize approximation arguments.

A surprising solution to the former problem comes from solving the latter first: I begin with a preliminary step which consists in reformulating the problem in a way that naturally allows for extending to the case of no fixed cost. This is technical but involves interesting nuances and provides a unified tractable approach which subsumes existing specifications of costs for continuous information acquisition as well as the previous discrete cost, without loss. Furthermore I show that the limit of solutions as fixed costs vanish indeed converges to the solution of the limit problem despite discontinuities. These preliminary steps are summarized in Section 5.2 and detailed in Appendix D.

The new approach enables leveraging a limit characterization in the non-recursive formulation of the problem to obtain tight results, which are laid out in Section 5.3. After formally defining the class of wait-or-confirm policies, Theorem 4 characterizes an optimal such policy in terms of the net value function in the vanishing fixed cost problem, and establishes that the optimal policy converges in this class as κ goes to zero. Theorem 5 obtains an explicit expression for the long run optimal dynamics in terms of the concave envelope of an appropriately defined "virtual net flow payoff". I unpack some qualitative consequences of these results, in particular for induced action dynamics, in Section 5.4.

5.1 Informal intuition: infinitesimal information with vanishing fixed costs

Intuitive effects of vanishing fixed costs. When the fixed cost becomes negligible, the difference between the *gross* and *net* interim value of information collapses. The gap in these two values loosely captures the incentive to *wait*: in the regions where there is gross value of information, the DM has some profitable target beliefs that they would want to jump to, but *waits* until information has depreciated enough for the gains to warrant paying the fixed cost. This suggests that as the fixed cost vanishes so do incentives for waiting: if there are target beliefs that the DM could profitably jump to, they should do so immediately. In practice, this suggests that in the characterization of optimal policies from Theorem 2 the gap between target beliefs (the support of optimal experiments) and threshold beliefs (closest belief triggering information acquisition) would vanish.

For now, as an illustration and a thought experiment, assume that "target" and "threshold" beliefs simply get closer to one another when κ decreases to 0. For simplicity, focus on the long run interval (q^0, q^1). The previous logics suggests that (i) q^0 and p^0 should converge to the same point, (ii) similarly for q^1 and p^1 and (iii) the information acquisition region should have "no holes" in the limit (be given by the limit of $[p^0, p^1]$). What does this imply for the content of information acquisition and its dynamics? The change affects both the frequency of information acquisition and the probabilities in the chosen experiment. Consider the dynamics of the beliefs starting from q^0; recall the wait time until the next update is:

$$\tau^0 = \frac{1}{\lambda} \log \frac{\pi - q^0}{\pi - p^0}$$

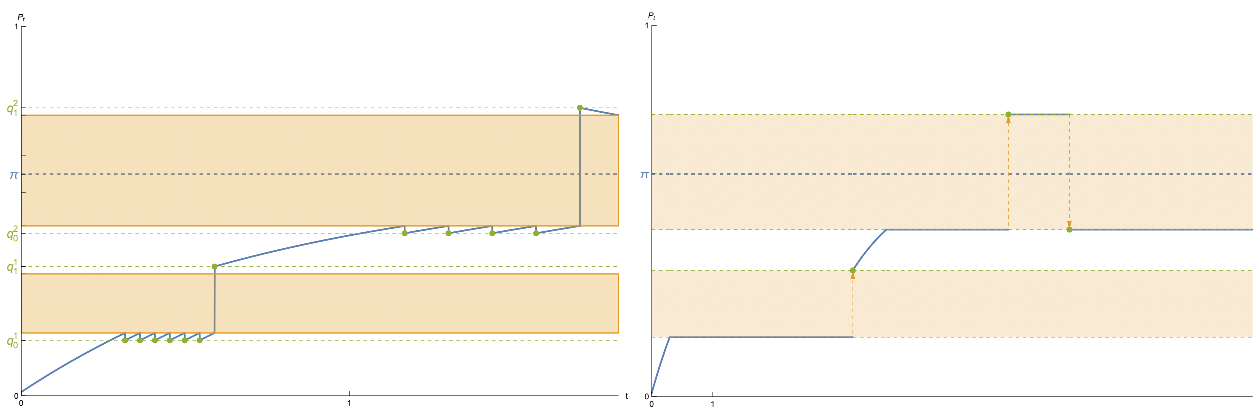
which converges to 0 if q^0 and p^0 converge to the same limit. This corresponds to the previous intuition on immediate information acquisition. What is more interesting is what happens to the jump probabilities; when acquiring information, the outcome is a jump:

$$\begin{cases} \text{to } q^0 \text{ with probability } \dfrac{q^1 - p^0}{q^1 - q^0} \\ \text{to } q^1 \text{ with probability } \dfrac{p^0 - q^0}{q^1 - q^0}. \end{cases}$$

Observe that if q^0 and p^0 converge to the same limit, the former probability goes to 1 and the latter to 0. This might seem counter-intuitive at first: the experiment looks as if it is uninformatively confirming belief q^0. But to understand what is really happening one needs to consider the effect of both limits (in frequency and content) happening simultaneously. At the limiting "target belief" there is no gross value of information, but after an infinitesimal amount of time the inward drift of the belief creates gross value of information; this triggers information acquisition which has an infinitesimally small probability of leading to a jump to the other target belief, and otherwise leads *immediately back* to the previous target belief. What this informally describes is an information technology where the DM *continuously checks* whether the current belief is still valid by acquiring incremental (infinitesimally informative) information. This prevents the belief from drifting inwards: either the current belief is confirmed, or a "rare signal" occurs which leads to a sudden jump to the other target belief.

In other words, the underlying information technology takes the familiar form of a "Poisson breakthrough" signal. The DM optimally chooses to acquire a signal structure where a breakthrough arrives at some constant rate, conditional on the true state. That rate is chosen so that (i) when a breakthrough arrives, it leads to the new belief which is exactly the other target belief and (ii) the inference from the lack of arrival of the breakthrough is such that it precisely confirms the current belief i.e. cancels out the unconditional drift. This random part of the belief process, excluding the deterministic drift, is a martingale (a compensated Poisson process) which captures the infinitesimal version of Bayes-plausibility.

The same logic applies to short-run information acquisition, except that following a jump the process drifts away from the information acquisition region. After a "Poisson breakthrough" in the non-cyclical regime, there will be a period of no information acquisition until a new information acquisition region is hit, at which points the DM starts acquiring information again in the same fashion. This extends the previous relationship between short- and long-run information acquisition, with the difference that information acquisition is continuous over successive time intervals. Whenever the belief process hits a short run information acquisition region it stays at the target belief, since the drift cancels out, until a jump randomly arrives and beliefs cross over the local information acquisition interval; the belief then proceeds to drift until it hits another region of information acquisition, and so on and so forth until it enters its stationary dynamics. Putting everything together, this means that in the limit the optimal policy is characterized by always either *waiting* (i.e no information acquisition, so that beliefs drift deterministically in some interval of time) or *confirming* the current belief, until a randomly timed jump to a fixed alternative belief in the direction of the steady state.



(A) Belief dynamics as fixed costs become small

(B) Belief dynamics in the limit with no fixed costs

FIGURE 7: Visualization of belief dynamics with vanishing fixed costs.

FORMALIZING LIMIT INFORMATION ACQUISITION. The nested intervals structure of solutions from Theorem 2 suggests it would be sufficient to track how thresholds and target beliefs change with κ, but this turns out not to be a tractable approach. First, incentives for information acquisition depends on the relative change in endogenous objects (specifically v and $\mathcal{G}\, v$), which is generally hard to characterize. Second, even the fact that the residual value of information collapses does not give information about how the *belief thresholds* behave. New multiplicity issues with indifference may also arise in the limit. Lastly, we lack a reliable way to "track" intervals: as κ changes, intervals in Γ and $\mathscr{I}$ could appear or disappear, merge or split in ways that are hard to discipline. Intuitively, this is driven by the complex interplay of payoffs and costs as we range along the belief space. When the fixed cost decreases, it may become optimal to merge two successive experiments on the path of convergence into one more frequent and more informative experiment, or on the contrary to split an experiment into successive experiments to exploit frequency but skip over certain regions of the belief space. It may also become profitable to run experiments in new regions of the belief space. As a result, the geometric characterization is not a suitable language: even though thresholds do effectively converge in the way that was described, this needs to be shown by other means. It is also useful to note that no claim can be made about *monotonic* convergence; indeed the behavior of the belief thresholds along the path of convergence can be fairly irregular.

A second, more subtle problem is highlighted by the previous illustration: how should we interpret precisely the "limit problem" with no fixed cost? Note that the suggested limit features continuous information acquisition, which the model does not allow for. Furthermore, even though it seems intuitive to assign an infinite cost to such a policy when $\kappa > 0$ (making our restriction without loss), our cost was only formally defined for punctual information acquisition. Even if we do extend the cost, it will naturally have some discontinuities: continuous information acquisition even with a very small fixed cost would create a spike to infinite costs. For these reasons, it is not straightforward to interpret the limit of solutions as κ goes to zero, in particular in regards to (approximate) optimality of the limit solution in some limit problem. Surprisingly, solving this problem holds a key to the previous problem as well.

5.2 Information acquisition with and without fixed costs

The essential preliminary step of the analysis consists in recasting the problem in terms of a choice directly over *belief processes*. The main difficulty is that, with no fixed cost, the policy space is not rich enough: it may be desirable to acquire information not just at discrete points but in continuous increments. When allowing for continuous information acquisition, one cannot simply substitute existing analogous cost specifications based on infinitesimal information flows, since those are not well defined for punctual information acquisition (hence do not include our initial specification) and vice versa. These two technologies can be reconciled using the following observations:

(i) every stochastic process (in the appropriate class corresponding to feasible belief processes) can be approximated arbitrarily well by a process where information is only acquired at countably many points in time;

(ii) for any such approximation, the objective function (in particular, the cost) has a unique well defined limit in the class of feasible belief processes.

Formally, let:

$$\boldsymbol{\mathcal{B}}(p) := \left\{(P_t)_{t\geq 0} \text{ càdlàg process in } [0,1] \;\middle|\; \mathbb{E}[P_0] = p \;\&\; \forall\, t, s \geq 0, \mathbb{E}[P_{t+s}|P_t] \overset{a.s.}{=} e^{-\lambda s} P_t + (1-e^{-\lambda s})\pi\right\}$$

and define $\boldsymbol{\mathcal{B}}_{\mathrm{d}}(p)$ to be the subset of such belief processes that can be generated by some discrete information acquisition policy $\{\tau_i, F_i\}$ with $F_0 \in \mathscr{B}(p)$. It is relatively standard to establish that $\boldsymbol{\mathcal{B}}(p)$ equipped with the Skorohod topology is a (compact subset of a) Polish space, of which $\boldsymbol{\mathcal{B}}_{\mathrm{d}}(p)$ is a dense subset – all of the detailed analysis for this section is relegated to Appendix D. It is also direct to define flow payoffs over $\boldsymbol{\mathcal{B}}(p)$ as:

$$\mathfrak{U}(P) := \mathbb{E}\left[\int_0^\infty e^{-rt} u(P_t)\,dt\right]$$

Costs, however, are only rigorously defined over $\boldsymbol{\mathcal{B}}_{\mathrm{d}}(p)$, although for $\kappa \geq 0$. An essential step consists in noticing that the UPS formulation allows to rewrite costs associated with a process $P \in \boldsymbol{\mathcal{B}}_{\mathrm{d}}(p)$ equivalently as:

$$\mathfrak{C}_\kappa(P) := \mathbb{E}\left[\sum_{t|P_t \neq P_{t^-}} e^{-rt}\Big(c(P_t) - c(P_{t^-}) + \kappa\Big)\right]$$

The technical result which enables the analysis is that $\mathfrak{C}_\kappa$ is uniformly continuous over $\boldsymbol{\mathcal{B}}_{\mathrm{d}}(p)$, hence it has a unique continuous extension over its closure $\boldsymbol{\mathcal{B}}(p)$. In the subcase when the process features only continuous information acquisition, the extended cost function over arbitrary belief processes coincides with previously introduced information technologies based on the infinitesimal generator – see for instance Zhong [2022], Hébert and Woodford [2023], Georgiadis-Harris [2023]. Hence, the result equivalently shows that there is a unique continuous extension of these costs to arbitrary belief processes that may not have a well-defined infinitesimal generators (like our punctual information acquisition processes), and that it coincides with the symmetrical extension of the "discrete" cost.

Note that, when the fixed cost is non-zero, continuous information acquisition will generate infinite costs, hence the initial formulation of the problem is indeed without loss. Density also entails that, for the case with no fixed costs ($\kappa = 0$) the definition of the value function in the exact same form as before (over punctual information acquisition processes) is still valid, albeit as a *supremum* (strictly speaking not an attained maximum), so that:

$$v(p) = \max_{P \in \boldsymbol{\mathcal{B}}(p)} \mathfrak{U}(P) - \mathfrak{C}_\kappa(P) = \sup_{P \in \boldsymbol{\mathcal{B}}_{\mathrm{d}}(p)} \mathfrak{U}(P) - \mathfrak{C}_\kappa(P).$$

The second supremum, corresponding to the original problem, is attained only if $\kappa > 0$.

Furthermore, I show that the limit (understood in the Skorohod topology, i.e. in distribution) of optimal belief processes for $\kappa > 0$ converges as the cost vanishes to a solution of the $\kappa = 0$

problem, in other words v_κ converges pointwise to v_0 and:

$$\text{if } P^\kappa \in \underset{P\in\mathcal{B}(p)}{\arg\max}\, \mathfrak{U}(P) - \mathfrak{C}_\kappa(P) \text{ and } P^\kappa \to P^0, \text{ then } P^0 \in \underset{P\in\mathcal{B}(p)}{\arg\max}\, \mathfrak{U}(P) - \mathfrak{C}_0(P)$$

This is despite the fact that $\mathfrak{C}_\kappa$ is *not* continuous in κ at 0; it is, however, *epi-continuous* in κ, which is the minimal notion which guarantees convergence of minimizers (see Appendix D). Convergence enables, modulo being careful about selection, intepreting optimal policies in the $\kappa = 0$ as natural behavioral approximation for the case with "small" fixed costs.[4]

Throughout the paper, the fixed cost is viewed as a substantial and economically meaningful assumption – both as an environmental feature and as a modelling device to capture lumpy information acquisition. Nonetheless, it has advantages even if one was primarily interested in the case with vanishing fixed cost. First, it provides compelling economic intuition for the properties of the optimal solution, via limit properties. Second, it provides a tractable structure and solution method that would not have been accessible if we had tried to write the problem directly over either the class of continuous information acquisition strategies or a richer class combining both types of strategies; it also bypasses some definitional issues that arise in that latter case e.g. for defining rigorously the "hybrid" class, or making formal sense of the resulting HJB which is generally singular. Approaching continuous information acquisition via its vanishing fixed cost limit amounts to taking limits in an *endogenously chosen time-discretization*. This contrasts with existing approaches (e.g. Zhong [2022]) where continuous time is obtained as the limit of an *exogenous uniform time-grid*.

5.3 Optimal information acquisition with vanishing fixed costs

Formalizing the intuition from Section 5.1 relies on an using the tools of the previous section to rewrite the problem as simple maximization of a "virtual flow payoff". This enables a precise analysis of the vanishing fixed cost limit of optimal information acquisition. From now on, denote v_κ the value function for any $\kappa \geq 0$, $w_\kappa := v_\kappa - c$ the corresponding net value function and P^κ the optimal belief process for $\kappa > 0$.

The virtual net flow payoff. The following reformulation is instrumental because it subsumes all costs into a "virtual net flow payoff".

[4]Note however, that even though the optimal belief process for $\kappa = 0$ is close in the distributional sense to the optimal belief process for small κ, the former is not technically approximately optimal because of the discontinuity: any continuous information acquisition generates infinite costs, even for small κ.

Lemma 1. *The net value function w_0 in the $\kappa = 0$ problem verifies:*

$$w_0(p) = \sup_{P \in \boldsymbol{\mathcal{B}}(p)} \mathbb{E}\left[\int_0^\infty e^{-rt} f(P_t)\,dt\right]$$

Where the "virtual net flow payoff" f at belief p is defined as:

$$f(p) := u(p) - rc(p) + \lambda(\pi - p)c'(p)$$

The resulting problem is only constrained by the compensated Bayes-plausibility constraint which defines feasible belief processes. An accompanying result allows us to state that the value function for $\kappa = 0$ is also the solution of the "limit HJB" for $\kappa > 0$, when we take κ (which is *not* the HJB for the general problem when $\kappa = 0$). This in turns allow for directly establishing several properties of the value function in the limit problem, e.g concavity of the net value function w_0. There also exists an analog of Lemma 1 for $\kappa > 0$, but it is less immediately useful (because the discounted fixed cost term is not an easy object to manipulate).

WAIT-OR-CONFIRM POLICIES. Informally, "wait-or-confirm" policies are such that the process drifts until it hits the (boundary of an interval) in the information acquisition region, then stays at that belief until it jumps over that interval. Since the rate of arrival of jumps is pinned down by those two beliefs and the compensated martingale condition, it is natural to parameterize the distribution of the whole process solely in terms of the open region where information is acquired immediately – which the belief process only possibly "jumps out of" at the initial time and never enters.

To make this formal in a synthetic and convenient fashion, for any open set $\mathscr{I}$ divide its boundary $\partial\mathscr{I}$ into points from which drifting towards π leads either *in* or *out* of the set $\mathscr{I}$:

$$\partial_\pi^{\text{in}}\mathscr{I} := \left\{p \in \partial\mathscr{I} \;\middle|\; \exists \varepsilon > 0, b_\pi(p,\varepsilon) \subset \mathscr{I}\right\},$$
$$\partial_\pi^{\text{out}}\mathscr{I} := \partial O \setminus \partial_\pi^{\text{in}}\mathscr{I} = \left\{p \in \partial\mathscr{I} \;\middle|\; \forall \varepsilon > 0, \exists q \in b_\pi(p,\varepsilon) \setminus \mathscr{I}\right\};$$

where $b_\pi(p,\varepsilon)$ denotes the (open) "π-neighborhood" of size ε at p:

$$b_\pi(p,\varepsilon) := \begin{cases} (p, p+\varepsilon) & \text{if } p < \pi \\ (p-\varepsilon, p) & \text{if } p > \pi \end{cases}.$$

Definition 4 (Wait-or-confirm policies)**.** For any open set $\mathscr{I} \subset [0,1]$ and any $p \in [0,1]$, denote $\text{WoC}_p[\mathscr{I}]$ the distribution of the belief process $P \in \boldsymbol{\mathcal{B}}(p)$ such that:

- (Initial jump) If $p \in \mathscr{I}$, P is distributed according to the only binary experiment in $\mathscr{B}(p)$ supported over the two closest points from p not in $\mathscr{I}$, otherwise $P_0 = p$ a.s.

- (Waiting beliefs) At all $p \in \mathscr{I}^c \cup \partial_\pi^{\text{out}} \mathscr{I}$, P evolves according to no information acquisition, i.e it drifts deterministically with $dP_t = \lambda(\pi - P_t)dt$
- (Confirmation beliefs) At all $p \in \partial_\pi^{\text{in}} \mathscr{I}$, P stays at p (*confirming*) until some exponentially distributed time, at which it jumps to the closest belief $q(p)$ in the direction of π that is not in $\mathscr{I}$.

The belief process P is called a *wait-or-confirm (information acquisition) policy* with initial belief p, and $\mathscr{I}$ is called its (instantaneous) information acquisition region.

In simpler words, a wait-or-confirm belief process drifts until it hits the boundary of an interval in the (instantaneous) information acquisition region $\mathscr{I}$, at which points it confirms the current belief until some news prompts a jump over the interval. Note that at "confirmation" beliefs, the compensated martingale condition of the belief process pins down the rate of arrival of the exponential jumps to be $\lambda\frac{\pi-p}{q(p)-p}$. Further observe that the definition is well-posed since the open set $\mathscr{I}$ is decomposable into a countable collection of open intervals. The name "wait-or-confirm" implicitly ignores the instantaneous information acquisition region since those beliefs are only possibly relevant for Bayes-Plausibility at the initial time, and the belief process is almost surely in the union of the waiting set and the confirmation set at all times.

OPTIMALITY AND CONVERGENCE. The first main results in this section characterizes an optimal wait-or-confirm policy in terms of the value function w_0 and establishes convergence.

Theorem 4. *The optimal net value function w_0 in the information acquisition problem with $\kappa = 0$ is concave. Furthermore:*

1. *(optimality) The wait-or-confirm process $P \sim \text{WoC}_p\left[\text{int}\, L_0\right]$ is optimal, where:*

$$L_0 := \left\{p \in [0,1] \,\middle|\, \exists\varepsilon > 0, \exists q \in b_\pi(p,\varepsilon), w_0(q) = w_0(p) + d_\pi w_0(p)(p-q)\right\};$$

   and d_π denotes the directional derivative in the direction of π.

2. *(convergence) Assume $P^\kappa \to P$, then:*

$$P \sim \text{WoC}_p\left[\liminf_{\kappa\downarrow 0} \mathscr{I}_\kappa\right]$$

3. *(relation and local uniqueness) It is uniquely optimal to wait to acquire information at all beliefs such that w_0 is stricly concave is some π-neighborhood and* $\liminf_{\kappa\downarrow 0} \mathscr{I}_\kappa \subset$ *$intL_0$.*

The fact that w_0 is concave implies that when $\kappa = 0$ there is no residual interim value of information. This is intuitive, since if at any belief p we had $\text{Cav}[w_0](p) > w_0(p)$, then it would be optimal to immediately jump, implying $w_0(p) = \text{Cav}[w_0](p)$; hence $w_0 = \text{Cav}[w_0]$ everywhere and w_0 is concave. The concavity of w_0 underpins the decomposition of the belief space which is used to define an optimal wait-or-confirm policy in the result. An optimal wait-or-confirm policy is specified by the regions where w_0 is strictly concave or locally affine (in the direction of π). A more intuitive (though slightly informal) description of the optimal policy is as follows: in the region where w_0 is strictly concave, let belief drift deterministically until they hit an interval where w_0 is locally affine in the direction of π, at which point maintain the current belief so as to "skip over" the locally affine region to the next region where w_0 is strictly concave; if the initial belief is in a region where w_0 is locally affine, immediately acquire the experiment supported over the closest beliefs such that it is not.

The first part of Theorem 4 exhibits how to construct optimal policies but it is silent as to both (a) whether this specifies a unique one, and (b) whether the optimal policies for $\kappa > 0$ from the previous section converge to *one of these* optimal policies. It is easy to see, by taking an extreme counter-example, that neither uniqueness nor convergence to some arbitrarily selected wait-or-confirm policy can be guaranteed in general. Indeed, consider the case where f is affine over $[0,1]$: for $\kappa = 0$, every feasible belief process is optimal; for any $\kappa > 0$, it is uniquely optimal to never acquire information. Although this example is extreme, local versions of the same issue can exist in general. Note, however, that even in this case, the optimal policy when $\kappa > 0$ converges to *some* wait-or-confirm policy. This is what the second part of the result states formally.

Convergence reinforces that the simple class of wait-or-confirm policies is a natural one to consider because it is both without loss of optimality, consistent with continuity and limit considerations. The last point of Theorem 4 shows that the policy in the first point essentially breaks indifferences towards maximal information acquisition, whereas the limit optimal policy selects away from unecessary information acquisition. Robustness to the presence of a small fixed cost (along with selection of earliest stopping time and any selection criterion for optimal experiments) gives a natural selection of an optimal policy and that it is within the class of wait-or-confirm policies. Returning to the extreme illustration with f linear, it is quite intuitive that when the DM is indifferent between all information acquisition policies, infinitesimal perturbation in the form of adding a vanishing fixed cost would uniquely select the policy which consists in never acquiring information. In that sense, the limit of optimal policies as fixed costs vanish intuitively breaks indifferences (in the limit problem) in favor of waiting, which is arguably a desirable property.

Long run optimal policy. Theorem 4 gives an implicit characterization of optimal policies in terms of the (limit of the) value function, as was the case for our previous results when $\kappa > 0$. However, focusing on the long run in the case when $\kappa = 0$ allows a drastic improvement of the results, in the form of an *explicit* characterization of the optimal belief process in terms of the primitives of the problem. Special properties of long run behavior follow from the same logic as before: the process eventually converges to a stationary regime, which is easier to decribe and more regular. In the case with no fixed costs, this combines with the added tractability from the virtual flow payoff reformulation in Lemma 1 to give the next result.

Theorem 5. *In the information acquisition problem with no fixed costs:*

- *If* $\text{Cav}[f](\pi) = \pi$*, then it is optimal in the long run to not acquire any information, uniquely so if* $\text{Cav}[f] = f$ *in a neighborhood of* π *and* f *is locally strictly concave at* π*.*
- *If* $\text{Cav}[f](\pi) > \pi$*, i.e there exists some (generically unique) interval* (q_0^f, q_1^f) *containing* π *s.t. the concave envelope of* f *is everywhere above* f *inside, and* f *coincides with* $Cav[f]$ *at* q_0^f *and* q_1^f*, then:*

$$\forall p \in [q_0^f, q_1^f], \;\; w_0(p) = \int_0^\infty e^{-rt}\,\text{Cav}[f]\big(e^{-\lambda t}p + (1-e^{-\lambda t})\pi\big)dt$$

and the belief process which consists in:

- *Jumping immediately to* $\{q_0^f, q_1^f\}$ *from any* $p \in (q_0^f, q_1^f)$
- *Jumping from* q_0^f *to* q_1^f *at rate* $\rho_0 := \lambda \frac{\pi - q_0^f}{q_1^f - q_0^f}$
- *Jumping from* q_1^f *to* q_0^f *at rate* $\rho_1 := \lambda \frac{q_1^f - \pi}{q_1^f - q_0^f}$

is optimal.

The proof strategy relies on first establishing that the given expression of the value is an upper bound for w_0 (which follows from the concave envelope dominating the function pointwise, applying Jensen's inequality pathwise, and finally the compensated dynamic Bayes-plausibility constraint), then showing that the policy described is feasible and achieves this upper bound. It may be surprising that the payoffs are expressed as an integral over the *deterministic* drift path from p when the belief process is actually random; this captures the effect that the expected payoff from jumping between q_0^f and q_1^f is a linear combination of the payoffs at these two points and that the probabilities move towards their long run average on the line between $f(q_0^f)$ and $f(q_1^f)$. Because of the linearity of $\text{Cav}[f]$ (locally) and by definition of q_0^f and q_1^f, this is equivalent to moving along $\text{Cav}[f]$ and because of Bayes-plausibility expectations are equivalent to moving along the deterministic flow.

An immediate corollary of this result is that the limit of optimal long run belief cycles (and their associated belief process) for $\kappa > 0$ must converge to this policy as κ goes to 0. Let P^0 the optimal long run belief process from Theorem 5. Let $q_0^\kappa, p_0^\kappa, p_1^\kappa, q_1^\kappa$ describe the optimal belief cycle for any $\kappa > 0$ and P^κ the associated belief process. Then:

(i) P^κ converges in distribution to P^0 as κ converges to 0;

(ii) if $\text{Cav}[f](\pi) > \pi$, q_i^κ and p_i^κ converge to q_i^f as κ converges to 0, for $i = 0, 1$.

Although it is tempting to attempt extending the explicit characterization from Theorem 5 to short-run information acquisition as well, it generally does not hold. Outside of the long run regime, the upper bound given by $\int_0^\infty e^{-rt}\text{Cav}[f]\big(e^{-\lambda t}p + (1 - e^{-\lambda t})\pi\big)dt$ becomes strict. Because of transitory dynamics, it is no longer necessary that the intervals where $\text{Cav}[f] > f$ exactly correspond to regions where it is optimal to acquire Poisson signals (in other words, we cannot claim that intervals where w_0 is locally affine coincide with those where $\text{Cav}[f] > f$). One can still apply the idea of the recursive methodology from the previous section, solving for transitory optimal information acquisition within the Poisson class and "from the inside out" (constraining to one interval to the left of the stationary case, etc.),[5] but there seems to be no general explicit characterization outside of the stationary case.

5.4 Dynamics of beliefs and actions with vanishing fixed costs

Belief dynamics with vanishing fixed cost As fixed costs vanish, belief changes become lumpier. In the long run, information is gathered continuously within the relevant belief region, ensuring there are no delays in acquiring valuable signals. This derives from properties of UPS costs: there is no incentive to delay or break up information acquisition – although this does not imply that the DM never waits to acquire information, at least in the short run. Time intervals where beliefs are continuously changing are the ones in which no information is acquired, whereas learning compensates the depreciation of knowledge so that beliefs do not change until an eventual jump.

Return to the illustration of an investor checking their portfolio or market conditions: there may be a relatively small (but non-zero) cost to sitting down and opening up their account or a news platform. In that case optimal behavior can be interpreted as doing so frequently but spending little time processing the content: the investor does a coarse check for significant changes; if

[5] This approach generally requires some additional assumption(s) to be well-defined, namely to ensure that the induction is appropriate – this requires guaranteeing that that there is no accumulation point of intervals where $\text{Cav}[f] > f$. A simple way to ensure that would be that assume that both u and c are piecewise analytical, hence so is f. In that case there can only be finitely many intervals where information acquisition is profitable.

no news is salient, they assume that "nothing has changed" and revert to their previously held belief. They keep doing so until something unusual enough prompts a sudden reevaluation of their views. The frequency and precise coarseness of periodic checks are used to calibrate the target levels of confidence the investor aims to maintain.

The limit characterization of the belief process reveals a transition between two regimes: one where belief jumps are deterministically timed with random outcomes, and another where jumps are randomly timed with deterministic outcomes. In the standard model with significant fixed costs, information is acquired at predictable intervals, and the uncertainty lies in the outcome of the experiment. However, as fixed costs diminish, belief updates become rare but predictable, and the uncertainty shifts to the timing of belief changes. This shift implies that depending on the magnitude of fixed costs, an observer might perceive the agent as either frequently and significantly adjusting their beliefs or as holding a steady belief with rare, sudden shifts.

ERGODIC DISTRIBUTION The ergodic distribution in the limit is simply given by two point masses at q_0, q_1 with respective mass m_0 and m_1 pinned down by balance conditions:

$$m_0 = \frac{\rho_1}{\rho_0 + \rho_1} = \frac{\pi - q_0}{q_1 - q_0} \; ; \; m_1 = \frac{\rho_0}{\rho_0 + \rho_1} = \frac{q_1 - \pi}{q_1 - q_0}$$

The concentration to two points when fixed costs vanish is interesting within the population interpretation: it means that the population of agents in the long run is at any time split into two well-defined groups that hold exactly the same belief within each. The dynamics conditional on a given state similarly reduce to just a speed of transition between the two points: it only depends on which agents get the news early or late – this in particular (and intuitively) suggests that the overall population much more responsive to state changes in terms of beliefs, because reaction times are not impeded by exogenous waiting times.

ACTION SWITCHING DYNAMICS: LUMPY UPDATING VERSUS LUMPY ACTION Because the DM eventually only ever holds two beliefs, this also means they only ever take two actions – no matter how many actions we started from in the primitive space. This constitutes one possible resolution of an otherwise puzzling feature of dynamic settings with continuous information acquisition. If we start from a model with a continuum of action, such that a different action is optimal at each belief (say, in the context of a pricing problem), then by virtue of the belief continuously drifting the agent in our model has to be continuously adjusting their action. This can be seen as an unpalatable prediction in general: often, decision makers do not in practice adjust decision variables in real time but only lumpily. This is a well studied question in the macroeconomics literature. With non-zero fixed costs, unless the discreteness of action choice is *assumed* in the decision problem, the DM acquires information lumpily but might find it suitable to change their action continuously. However, as fixed costs vanish, the situation reverses: in the long run, information acquisition is continuous but the action changes lumpily.

This upshot is a novel source for discreteness in action choice: the flexibility of information acquisition without frictions. The mechanics of optimal information acquisition lead the DM to optimally only ever hold two beliefs and acquire information continuously so as to only jump between those two; hence because the belief change becomes lumpy (even though information acquisition is constant and incremental), the action choice becomes lumpy as well. Discreteness follows from the underlying discrete structure of optimal information structures and not from any frictions.

This opens up a potentially new avenue for interpretation and investigation, which should ultimately be confronted with empirical discipline. Can one differentiate between agents who change their action periodically because they bear a cost of switching actions, and agents who chose to acquire information so as to hold only finitely many beliefs, which in turn leads them to switching actions discretely? Naturally, realistic examples might feature a mixture of both explanations; the information-driven explanation for lumpy actions provides a complementary approach to understanding observed behavior.

6 Examples and applications

In this section, I examine two particular examples. Section 6.1 applies the result of the previous sections to a concrete problem of portfolio allocation, expanding on the specific interpretation and consequences of optimal information acquisition for diversification over time. Section 6.2 studies a canonical binary action example with a safe and a risky action. In particular, this allows for a simple exploration of the effects of asymmetry by considering one simple deviation from the symmetric case: moving the indifference point between the safe and risky actions, which amounts to making the risky action relatively riskier.

6.1 A portfolio allocation problem

Setup. An investor allocates a unit flow budget between three available assets over time. One asset is safe and yields fixed return $s \geq 0$. The other two assets are risky; they have the same expected return $m > s$ but different unobserved variances which may be low ($\sigma_L^2 > 0$) or high ($\sigma_H^2 > \sigma_L^2$). The variance of returns on the risky assets varies over time in a negatively correlated manner: if one asset has high variance, the other's is low; returns are conditionally independent given variances. Which asset is riskier changes at a given rate $\lambda > 0$. Denote x_0, x_1 the returns on the risky assets; the state of the world $\theta_t \in \{0,1\}$ labels which of the risky asset is riskier at time t. Denote γ the share of the portfolio invested in risky assets, and β the sub-share allocated to risky asset 0. Total return is given by:

$$y(\gamma,\beta) := (1-\gamma)s + \gamma\big(\beta x_0 + (1-\beta)x_1\big)$$

Lastly, assume that there is some fixed flow cost $z \geq 0$ to investing in the risky assets (e.g. a broker's fee), which is paid only when $\gamma > 0$. Fix the current belief p that $\theta_t = 1$. The investor is assumed to have mean-variance preferences so that flow utility from portfolio allocation (γ,β) is given by:

$$\mathbb{E}_p[y(\gamma,\beta)] - \frac{\alpha}{2}\mathbb{V}_p[y(\gamma,\beta)] - z\mathbb{1}_{\gamma>0} = (1-\gamma)s + \gamma m - \frac{\alpha}{2}\gamma^2\Big(\beta^2\mathbb{V}_p[x_0] + (1-\beta)^2\mathbb{V}_p[x_1]\Big) - z\mathbb{1}_{\gamma>0}.$$

The optimal subdvision β^* of the risky component of the portfolio is chosen purely so as to minimize risk exposure and depends only on the expected variances of the two assets:

$$\beta^*(p) := \frac{\mathbb{V}_p[x_1]}{\mathbb{V}_p[x_0] + \mathbb{V}_p[x_1]} = \frac{p\sigma_L + (1-p)\sigma_H}{\sigma_L + \sigma_H}$$

This gives the variance from optimal diversification of the risky component of the portfolio:

$$\sigma^*(p) := \frac{\big(p\sigma_H + (1-p)\sigma_L\big)\big(p\sigma_L + (1-p)\sigma_H\big)}{\sigma_H + \sigma_L}.$$

The optimal share of investment in risky assets is in turn given by:

$$\gamma^*(p) = \begin{cases} 1 & \text{if } \alpha\sigma^*(p) < m - s \text{ and } m - \dfrac{\alpha\sigma^*(p)}{2} \geq z \\ \dfrac{m-s}{\alpha\sigma^*(p)} & \text{if } \alpha\sigma^*(p) \geq m - s \text{ and } s + \dfrac{(m-s)^2}{2\alpha\sigma^*(p)} \geq z \\ 0 & \text{otherwise} \end{cases}$$

so that the resulting indirect utility, given beliefs p, is given by (see Figure 9):

$$u(p) = \begin{cases} m - \dfrac{\alpha}{2}\sigma^*(p) & \text{if } \sigma^*(p) < \dfrac{m}{\alpha} \text{ and } \sigma^*(p) \leq 2\dfrac{m-z}{\alpha} \\ \dfrac{m^2}{2\alpha\sigma^*(p)} & \text{if } \sigma^*(p) \geq \dfrac{m}{\alpha} \text{ and } z\sigma^*(p) \leq \dfrac{m^2}{2\alpha} \\ s & \text{otherwise} \end{cases}$$

Notice that $\sigma^*(p)$ is decreasing in $|p - 1/2|$: the less uncertain the DM is about the current state, the more they are able to select a portfolio with low expected variance but high payoff. This

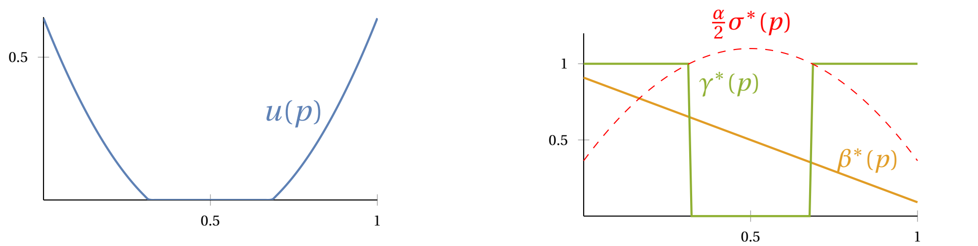


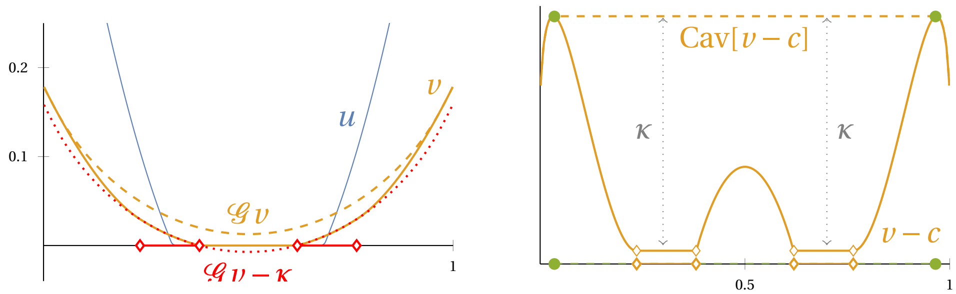
FIGURE 8: Flow payoffs and strategies in the portfolio choice problem.

means that, generally the DM chooses a less diversified allocation the more certain they are; for high enough certainty levels, they may buy none of the safe asset but generically they will always diversify to some extent between the two risky assets (even if $p = 0$ or 1). Conversely, the more uncertain they are, the more they prefer to hold a diversified portfolio; if they are *too* uncertain, they may stop holding risky assets altogether and only buy the safe asset (represented by the central flat section in Figure 9).

IMPLICATIONS OF OPTIMAL INFORMATION ACQUISITION. Results from the previous sections immediately imply a precise description of the investor's optimal long run information acquisition, provided it is optimal to do so. Optimal behavior features a repeating pattern: continuous rebalancing of the portfolio towards greater diversification as the investor becomes more uncertain about the current state, interrupted by periodic sudden restructuring towards holding a more extreme portfolio. This cycle may involve a phase where the investor temporarily exits the market and only holds the safe asset, especially if investment fixed costs are high or the market is intrinsically very volatile. Previous work in static contexts highlighted that information acquisition may lead to under-diversification (see e.g. Van Nieuwerburgh and Veldkamp [2010]). With costly periodic information acquisition in a changing environment, we see a similar effect taking place at times of information acquisition but unfolding over a cycle of endogenous length.

The solution may also feature path dependency, with some investors being effectively excluded from the market. If the information costs or the fixed cost of investing are high enough, there may exist a trap region for beliefs around 1/2. If that is the case, investors who start sufficiently uninformed will never acquire any information; their belief just drift to the no information average and they hold only the safe asset.[6] Meanwhile, investors who started with better information keep acquiring smaller amounts of information to maintain a cycle. As previously em-

[6]If there were no fixed cost of investing, they could still hold some of the risky assets in the limit, but converge towards the most diversified portfolio available.



(A) Value function

(B) Net value function

FIGURE 9: Solution in the portfolio choice problem.

phasized, any path dependency in the model comes from optimality – the trap exists because initial costs are not warranted by future benefits. Yet, if there were externalities from information acquisition or with a concern for inequality, path dependency in access to information would have welfare implications. Varying parameters also alters the domain of the trap region, which provides multiple possible explanations for differences across categories of investors.

The previous comparative statics also apply directly. In particular, as the fixed costs of information acquisition become small, portfolio choice concentrates over just two possible allocations, each favoring one of the risky assets. Reallocation towards a more diverse portfolio vanishes and we see only sporadic but relatively drastic rebalancing. This paradoxically suggests that an investor with easier access to information adjusts their allocation less frequently, and on the contrary sticks with a less extreme but stable investment strategy until a drastic change appears reasonable. By contrast, reallocation for investor with worse access to information comes from hedging against uncertainty because of the inability to continously monitor; this also leads them to holding a more extreme portolio when they do update, which displays more unstable holding patterns overall.

6.2 Asymmetry between a safe and a risky action

How do asymmetries in expected payoffs between states affect the quality and frequency of updating? To refine this question, I focus on simple environments where the DM has two actions available: a safe action which yields a deterministic known payoff, and a risky action whose payoff is determined by the state. The risky action action has better payoffs than the safe one if the state is good and worse if it is bad. Formally, indirect utility is given by:

$$u(p) = \max\left\{0, a(p - p^{\dagger})\right\}$$

where $p^\dagger \geq 1/2$ is the indifference threshold, and $a \geq 0$ corresponds to the difference between payoffs from the risky action in the good relative to the bad state (where I denote the good state by 1 and the bad state by 0).

Assume that all other primitives of the model are symmetric. The only asymmetry in this problem comes from shifting the indifference threshold *strictly* above 1/2. if $p^\dagger = 1/2$ the problem is effectively symmetric up to a normalization. This gives a simple structured way to capture a specific form of asymmetry in the form of the risky action being *relatively riskier*.

The main objective is to characterize properties (and in particular asymmetries) in the optimal long run information acquisition strategy which result from the payoff/state asymmetry. The first natural question is whether and how it is optimal to distort information acquisition: should the DM aim for a relatively high level of certainty when concluding in favor of the risky or the safe action? A second natural question is whether frequency responds asymmetrically to the content of news: do good/bad news lead to relatively slower or faster subsequent re-updating of information? If so, the content of a particular piece of news can impact how reactive the agent is to changes in the underlying state. For instance, some outcome might lead them to be less responsive and wait longer to reconsider their action choice.

Analytical properties are challenging to obtain but systematic examination of numerical simulations across a range of parameters and for natural cost functions suggest that:

(i) $q^1 - \pi > \pi - q^0$;

(ii) $p^1 - \pi > \pi - p^0$;

(iii) $\tau^1 < \tau^0$;

where q^θ, p^θ and τ^θ denote respectively the stationary target beliefs, thresholds, and times to update after a θ-news. My numerical analysis serves as suggestive evidence for a clear interpretable pattern. This suggests that: (i) the DM optimally chooses to target a higher level of confidence for the belief that suggest the risky action, (ii) the certainty threshold which triggers new information after a 1-news is correspondingly higher, and (iii) a 1-news leads to more rapid re-updating of information than a 0-news. These properties together look qualitatively like short-run confirmation bias: the DM keeps frequently acquiring information so as to maintain optimistic beliefs when they think taking the risky action is optimal, but if they switch to thinking the safe action is optimal, they hold relatively more pessimistic beliefs and wait a longer time before acquiring any new information. Nevertheless, this behavior is driven by purely rational motives: the structure of payoffs and the environments drives all asymmetries.

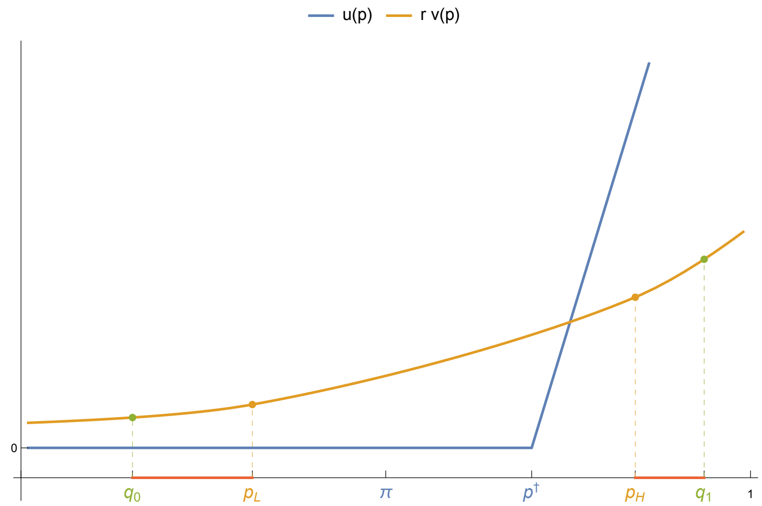


FIGURE 10: Asymmetric updating in choosing between a safe and a risky action

This brings up a potentially interesting conceptual point when linked with other possible (theoretical and empirical) patterns and mechanics of bias in dynamic information acquisition. In particular, the pattern described above is reminiscent of the "Ostrich effect", notably documented and analyzed in the finance literature (Sicherman et al. [2016], Galai and Sade [2006], Karlsson et al. [2009]), which consists in agents showing bias against information acquisition when expecting bad news. In particular, Sicherman et al. [2016] documents the frequency at which investors review the state of their portofolio, and two key stylized findings that are of interest in our context are that: (a) investors tend to check the status of their portofolio less frequently after poor performances than after good performances, and (b) investors tend to check less frequently when the market is more volatile. While (b) could potentially be tied to the kind of effects we studied in Section 4.4, (a) seems to be characterized by patterns similar to the ones derived here. Of course, my conclusion is not that this behavior is necessarily rational or necessarily derives from the incentives of optimal repeated information acquisition. Rather, my work develops analytical building blocks and highlights that asymmetries in the frequencies of information acquisition *can* arise from rationally optimal behavior in certain environments. I am hopeful that these tools can be useful for isolating and better understanding behavioral patterns which derive from various cognitive biases versus some form of adaptation to the environment. In that regard, note that the pattern of asymmetry in my example coincides with action switches – which might distinguish it from "ostrich"-type behavior (as the latter can feature changes in information seeking without any action switch). This line of inquiry could potentially help distinguish between similar behavior patterns, and eventually allow more precise quantitative studies of the multiple underlying mechanisms.

7 Discussion

Several technical assumptions in the model can be naturally relaxed for at least part of the results. More generally, the blueprint of analysis still remains valid even for substantially different assumptions (e.g. on information costs), although the actual content of the arguments would need to be substantially adapted. I also discuss alternative interpretations of the model and some more substantial extensions which are left for future work.

MORE THAN TWO VALUES. The recursive analysis remains largely unchanged if the state has more than two values, though the dynamics become more complex. With N states, transitions are governed by a time-homogeneous rate matrix Λ, which describes the rate of jumps between states. Beliefs evolve according to a differential equation that can be solved using the matrix exponential. While the recursive formulation and most of the analysis from the two-state case still hold, the richer belief dynamics mean that belief paths can now curve through a multi-dimensional space, rather than simply drifting along a line to π. This introduces greater complexity, as there are multiple ways for beliefs to converge toward π.

The core framework remains the same: information acquisition regions and costs are defined similarly, and optimal experiments are constructed using supporting hyperplanes. However, the dynamics are harder to characterize due to the multi-dimensional nature of belief paths, which can now drift in various directions after jumping to one of the extreme points of a convex polytope. Additionally, selecting optimal experiments becomes more challenging because there is generally no "least informative" experiment as in the two-state case.

In the long run, while the general intuition holds, dynamics may involve "super-cycles" rather than simple cycles, with beliefs drifting outwards before returning to regions of information acquisition. In some cases, particularly when state transitions occur uniformly, the simpler two-state cyclic dynamics may still apply. However, in most cases, the more complex dynamics introduce interesting possibilities for richer, multi-dimensional patterns of information acquisition. Although this makes analysis more difficult, this also gives the model potential to describe richer phenomena. In certain special cases, such as uniform transition rates across states, the analysis goes through almost as is, simplifying analysis and allowing us to extend our intuition to higher-dimensional settings.

INFORMATION COSTS. Generalizing the analysis to alternative information costs is challenging because the structure of optimal experiments and dynamics depends heavily on uniform posterior separability (UPS) costs. However, UPS costs cover a broad class, including many cases of interest that have garnered attention in the literature. They offer a simple belief-based framework that reduces complex dynamics to decisions over "certainty thresholds," which is useful for modeling agents with cognitive or resource constraints. Nonethelss, UPS costs have undesirable properties for certain applications, and are not well equipped to capture certain phenomena that require richer classes (see for example Denti [2022], Caplin et al. [2022], Denti et al. [2022], Bloedel and Zhong [2020], Hébert and Woodford [2021] for discussions).

For alternative costs, the overall framework remains valid. The recursive structure, fixed-point analysis, and general verification theorem extend to substantially weaker assumptions (see Appendix B); results like continuity, convexity, and existence of solutions require only minimal conditions such as continuity over feasible posterior distributions and concavity in the prior (see Denti et al. [2022]). Differences arise in the detailed characterization of optimal policies under different costs. Continuation values and the stopping problem must be adjusted for each specific cost structure, which will affect the recursive equations and value function properties. The general approach—linking optimal stopping and information acquisition—remains valid, but characterization may be less tractable. Future research could explore alternative cost classes, investigating how differences in static information costs translate into the dynamics of repeated information acquisition.

Relation to dynamic persuasion. Though the model's setup and analysis are motivated by dynamic information acquisition, it could be reinterpreted as capturing a situation of dynamic communication, where our "original DM" (designing the information acquisition policy) is a sender who periodically commits to sending information (at a fixed cost, but flexibly designed) about a changing state of the world to a strategic agent who then takes decisions accordingly. The main change would be that, if the interests of the sender and the receiver are not aligned, the indirect utility function u need not be convex, and the resulting value function need not be convex. This appears most clearly when rewriting the modified Bellman equation for the net value function:

$$w(p) = \sup_{\tau} \int_0^{\tau} e^{-rt} \hat{u}(p_t) dt + e^{-r\tau} \Big(Cav[w](p_\tau) - \kappa \Big)$$

where $w := v - c$ and $\hat{u}(p) := u(p) - rc(p) + \lambda(\pi - p)c'(p)$. If $\hat{u}$ is replaced with some other arbitrary continuous function, the recursive equation directly describes the sender's problem.

This equivalence mirrors results known in the static problem (see e.g. Gentzkow and Kamenica [2014]). The particular structure of the dynamic persuasion problem is particularly related formally and thematically to Ely [2017]. The recursive characterization mirrors the one in Theorem 1 in Ely [2017], with the notable difference of the fixed cost which allows to directly endogenize timing choice (versus a fixed time grid). My approach of studying the limit case as fixed costs vanish provides an alternative method to obtain continuous time solutions, which could complement the existing approach of taking limits in a fixed time grid's size – as in e.g. Ely [2017] or Zhong [2022].

Conclusion The question of how to optimally adapt to an evolving environment is fundamental and has immediate relevance in a wide range of contexts. Because attention is a finite resource and information is costly to obtain, it is natural to expect decision makers not to constantly seek new information, and instead periodically and imperfectly update their knowledge of current circumstances. As a result, how frequently and how precisely decision makers acquire information has important consequences. Yet, it is challenging to analyze the dynamic value of information, precisely because of the entanglement of its components across time. The model developped in this paper is a stepping stone, studying a tractable framework which both precisely captures the tradeoff between frequency and quality in general environments and delivers a solution method to study more detailed questions of interest. The general solution leads to further simplifications when focusing on specific environments. Developping precise applications in finance or policy and bringing the model to data is a promising avenue for future research. The tractability of the model and the characterization of solution also delineates promising paths for future theoretical work integrating optimal periodic information acquisition with a changing world in more complex settings like strategic environments.

References

Hassan Afrouzi and Choongryul Yang. Dynamic rational inattention and the phillips curve. 2021a. 4

Hassan Afrouzi and Choongryul Yang. Selection in information acquisition and monetary nonneutrality. Technical report, Mimeo, 2021b. 4

Fernando E Alvarez, Francesco Lippi, and Luigi Paciello. Optimal price setting with observation and menu costs. *The Quarterly journal of economics*, 126(4):1909–1960, 2011. 4

Fernando E Alvarez, Francesco Lippi, and Luigi Paciello. Monetary shocks in models with inattentive producers. *The Review of economic studies*, 83(2):421–459, 2016. 4

Robert J Aumann, Michael Maschler, and Richard E Stearns. *Repeated games with incomplete information*. MIT press, 1995. 5

Ian Ball and Jan Knoepfle. Should the timing of inspections be predictable? *SSRN Electronic Journal*, 2023. ISSN 1556-5068. doi: 10.2139/ssrn.4419716. URL http://dx.doi.org/10.2139/ssrn.4419716. 5

Patrick Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013. 69, 74

Alexander W Bloedel and Weijie Zhong. The cost of optimally-acquired information. *Unpublished Manuscript, November*, 2020. 46, 73

Andrew Caplin, Mark Dean, and John Leahy. Rationally inattentive behavior: Characterizing and generalizing shannon entropy. *Journal of Political Economy*, 130(6):1676–1715, 2022. 5, 8, 14, 46

Yeon-Koo Che and Konrad Mierendorff. Optimal dynamic allocation of attention. *American Economic Review*, 109(8):2993–3029, 2019. 5

Yeon-Koo Che, Jinwoo Kim, and Konrad Mierendorff. Predictive enforcement. *arXiv preprint arXiv:2405.04764*, 2024. 4

Krishna Dasaratha, Benjamin Golub, and Nir Hak. Learning from neighbours about a changing state. *Review of Economic Studies*, 90(5):2326–2369, 2023. 5

Benjamin Davies. Learning about a changing state. *arXiv preprint arXiv:2401.03607*, 2024. 4

Tommaso Denti. Posterior separable cost of information. *American Economic Review*, 112(10): 3215–3259, 2022. 5, 8, 46

Tommaso Denti, Massimo Marinacci, and Aldo Rustichini. Experimental cost of information. *American Economic Review*, 112(9):3106–3123, 2022. 8, 46, 54, 57

Piotr Dworczak and Anton Kolotilin. The persuasion duality. *Theoretical Economics*, 19(4):1701–1755, 2024. 14

Jeffrey C Ely. Beeps. *American Economic Review*, 107(1):31–53, 2017. 5, 46, 47

Stewart N Ethier and Thomas G Kurtz. *Markov processes: characterization and convergence*. John Wiley & Sons, 1986. 70, 72

Alexander Frankel and Emir Kamenica. Quantifying information and uncertainty. *American Economic Review*, 109(10):3650–3680, 2019. 8

Dan Galai and Orly Sade. The “ostrich effect” and the relationship between the liquidity and the yields of financial assets. *The Journal of Business*, 79(5):2741–2759, 2006. 5, 44

Matthew Gentzkow and Emir Kamenica. Costly persuasion. *American Economic Review*, 104 (5):457–462, 2014. 5, 14, 46

Alkis Georgiadis-Harris. Preparing to act: Information acquisition and the timing of actions. 2023. 5, 31, 73

Benjamin Hébert and Michael Woodford. Neighborhood-based information costs. *American Economic Review*, 111(10):3225–3255, 2021. 46

Benjamin Hébert and Michael Woodford. Rational inattention when decisions take time. *Journal of Economic Theory*, 208:105612, 2023. 5, 31, 73

Emir Kamenica and Matthew Gentzkow. Bayesian persuasion. *American Economic Review*, 101 (6):2590–2615, 2011. 5

Niklas Karlsson, George Loewenstein, and Duane Seppi. The ostrich effect: Selective attention to information. *Journal of Risk and uncertainty*, 38:95–115, 2009. 5, 44

Mel Win Khaw, Luminita Stevens, and Michael Woodford. Discrete adjustment to a changing environment: Experimental evidence. *Journal of Monetary Economics*, 91:88–103, 2017. 3, 4

Raphaël Lévy, Marcin Pęski, and Nicolas Vieille. Stationary social learning in a changing environment. *arXiv preprint arXiv:2201.02122*, 2022. 5

Alessandro Lizzeri, Eran Shmaya, and Leeat Yariv. Disentangling exploration from exploitation. Technical report, National Bureau of Economic Research, 2024. 5

Bartosz Maćkowiak and Mirko Wiederholt. Optimal sticky prices under rational inattention. *American Economic Review*, 99(3):769–803, 2009. 4

Bartosz Maćkowiak, Filip Matějka, and Mirko Wiederholt. Dynamic rational inattention: Analytical results. *Journal of Economic Theory*, 176:650–692, 2018. 4

N Gregory Mankiw and Ricardo Reis. Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *The Quarterly Journal of Economics*, 117(4):1295–1328, 2002. 4

Massimo Marinacci and Luigi Montrucchio. Unique tarski fixed points. *Mathematics of Operations Research*, 44(4):1174–1191, 2019. 11

Stephen Morris and Philipp Strack. The wald problem and the relation of sequential sampling and ex-ante information costs. *Available at SSRN 2991567*, 2019. 8

Giuseppe Moscarini, Marco Ottaviani, and Lones Smith. Social learning in a changing world. *Economic Theory*, 11:657–665, 1998. 5

Goran Peskir and Albert Shiryaev. *Optimal stopping and free-boundary problems*. Springer, 2006. 15

David Pollard. *Convergence of Stochastic Processes*. Springer New York, 1984. ISBN 9781461252542. doi: 10.1007/978-1-4612-5254-2. URL http://dx.doi.org/10.1007/978-1-4612-5254-2. 69

Luciano Pomatto, Philipp Strack, and Omer Tamuz. The cost of information: The case of constant marginal costs. *American Economic Review*, 113(5):1360–1393, 2023. 8

Ricardo Reis. Inattentive consumers. *Journal of monetary Economics*, 53(8):1761–1800, 2006a. 4

Ricardo Reis. Inattentive producers. *The Review of Economic Studies*, 73(3):793–821, 2006b. 4

Nachum Sicherman, George Loewenstein, Duane J Seppi, and Stephen P Utkus. Financial attention. *The Review of Financial Studies*, 29(4):863–897, 2016. 5, 44

Christopher A Sims. Implications of rational inattention. *Journal of monetary Economics*, 50(3): 665–690, 2003. 4, 5

Jakub Steiner, Colin Stewart, and Filip Matějka. Rational inattention dynamics: Inertia and delay in decision-making. *Econometrica*, 85(2):521–553, 2017. 4

Stijn Van Nieuwerburgh and Laura Veldkamp. Information acquisition and under-diversification. *The Review of Economic Studies*, 77(2):779–805, 2010. 5, 41

Felipe Varas, Iván Marinovic, and Andrzej Skrzypacz. Random inspections and periodic reviews: Optimal dynamic monitoring. *The Review of Economic Studies*, 87(6):2893–2937, 2020. 5

Peter Whittle. Restless bandits: Activity allocation in a changing world. *Journal of applied probability*, 25(A):287–298, 1988. 4

Yu Fu Wong. Dynamic monitoring design. *Available at SSRN 4466562*, 2023. 5

Weijie Zhong. Optimal dynamic information acquisition. *Econometrica*, 90(4):1537–1582, 2022. 5, 31, 32, 47, 73

APPENDIX

When and what to learn in a changing world

César Barilla

This version: July 3, 2025

A Preliminaries

A.1 Formal construction of the controlled belief process

It is worthwhile to give an explicit recursive construction of the belief process from a policy since (a) this gives intuitive content to the dynamics of beliefs, (b) this gives a rigorous definition for the class of controlled processes over which the DM optimizes and (c) this allows us to introduce explicitly some useful expressions for the law of motion of beliefs. Given an initial belief $p \in \Delta(\Theta)$, the belief process $\{P_t\}_{t\geq 0}$ is constructed as follows:

- First draw P_0 according to $F_0(p)$ (it is possible that the process jumps at the initial time; to allow for this case extend notations slightly and define the "true initial belief" p as the left-limit at 0 of the process $P_{0^-} := p$)
- For $i \geq 0$, iterate the following construction:
  - *Until the next time τ_{i+1} of information acquisition, the belief process is generated by the deterministic drift induced by the Markov chain*, i.e. set $P_t = p_t$ where p_t is the unique solution to :

    $$\text{for } t \in [\tau_i, \tau_{i+1}),\ dp_t = \big(\lambda_0(1-p_t) - \lambda_1 p_t\big)dt \text{ with initial condition } p_{\tau_i} = P_{\tau_i}$$

    where the equation for the flow is the usual (Kolmogorov) equation describing the law of motion of the unconditional distribution of the underlying Markov state, given initial probabilities; it has an explicit solution:

    $$p_t = e^{-\lambda t} p_{\tau_i} + (1-e^{-\lambda t})\pi$$

    If $\tau_{i+1} = \infty$, we can stop as the entire belief process is characterized.
  - Otherwise if $\tau_{i+1} < \infty$, *at τ_{i+1} the next time when information is acquired, a new belief is drawn according to the realization of the corresponding experiment*, i.e.:

    $$P_{\tau_{i+1}} \sim F_{i+1}$$

which captures the fact that new information arrives and the agent updates their knowledge about the current value of the state: the realization of the information generates an instantaneous jump in the process, which then proceeds to drifting as previously described until the next time of information acquisition.

If $\tau_i \to T$ for some finite T, we just extend the process by assuming that no information acquisition occurs after T:

$$P_t = e^{-\lambda t} P_T + (1 - e^{-\lambda t})\pi \quad \text{for } t \geq T$$

The law of motion of beliefs in between moments of information acquisition has a simple interpretation: given current belief p_t that the state is 1, in an infinitesimal interval of time dt, there is a probability $\lambda_0(1 - p_t)dt$ that the state was 0 but transitioned to 1 and a probability $\lambda_1 p_t dt$ that the state was 1 and transitioned to 0 – so the belief that the current state is 1 should increase by the former and decrease by the latter over any such small time interval.

Now, observe that this law of motion can be rewritten in a more convenient way by observing that for any $p_t \in \Delta(\Theta)$:

$$\lambda_0(1 - p_t) - \lambda_1 p_t = (\lambda_0 + \lambda_1)\Big(\frac{\lambda_0}{\lambda_0 + \lambda_1} - p_t\Big) = \lambda(\pi - p_t)$$

Hence until the next time of information acquisition and starting from some belief $p_0 = p \in \Delta(\Theta)$, beliefs solve:

$$dp_t = \lambda(\pi - p_t)dt \Longleftrightarrow p_t = e^{-\lambda t} p + (1 - e^{-\lambda t})\pi$$

i.e. beliefs *drift exponentially towards the invariant distribution* π at a rate which is controlled by the total volatility λ. The resulting local belief is a convex combination of the starting belief and the invariant belief π. As time grows longer without any information being acquired, beliefs converge to the long-run average probability that the state is 1, which is given by π. Note that this convergence is exponential, i.e. faster when far from π but slower as beliefs get closer to π – this captures the fact that information depreciates faster the further current beliefs are from the unconditional long run average.

A.2 The class of admissible policies

To close the loop, formally define the class of admissible information acquisition policies as the set of sequences of stopping times and experiments that verify the natural consistency requirements with respect to the induced belief process. Namely: the DM's choice at any time can be made contingent on information obtained up to that time and the experiments acquired at any given time are consistent with Bayes plausibility.

Definition A.1 (Information acquisition policies)**.** Define $\Xi(p)$ the set of **information acquisition policies** given initial belief p as the set of (random) sequences $\xi = \{\tau_i, F_i\}$ of information acquisition times and experiments, such that:

- The τ_i are progressively measurable with respect to the induced belief process $P_t^{p,\xi}$.
- Each experiment F_i is measurable with respect to the left-limit stopped process $\left\{\lim_{s\uparrow t} P_s^{p,\xi}\right\}_{t<\tau_i}$
- For all i, F_i is consistent with Bayes rule i.e. $F_i \in \mathscr{B}(p_{\tau_i^-})$ where:

$$\mathscr{B}(q) := \left\{F \in \Delta\Delta(\Theta) \,\middle|\, \int z dF(z) = q\right\}$$

  denotes the set of Bayes-plausible distribution over posteriors for a given belief $q \in \Delta(\Theta)$.

For any $\xi \in \Xi(p)$ we denote $\{\tau_i^\xi, F_i^\xi\}$ the corresponding full form. Let $\Xi = \cup_p \Xi(p)$.

The class of controls equivalently characterizes a class of belief processes (in distribution) and it is indifferent to consider the choice as being over either object.

B Characterization of solutions: general results, omitted proofs

In this section, we provide the basic results on the recursive characterization for the value function in the optimal information acquisition problem and policies. Whereas in the main text, we maintain the assumptions of binary states and UPS costs for the sake of homogeneity and readability, in this section we attempt to state and prove each result with more generality. This is done for the purpose of making it easier to adapt the analysis to different settings and highlight precisely the part that each assumption plays in each result. For the purpose of homogeneity, we give a set of more general assumptions on the state and costs, which nests the binary state-UPS cost setup and under which we can derive all the results in the main text. Note that these are not tight for *individual results* and natural extensions could be given – but for the sake of clarity, we only give one set of assumptions that generalize the ones in the main text while covering all the results. Specifically, we assume in this section that:

- the underlying state θ_t follows a continuous time Markov chain with a *finite* state space Θ and rate matrix Λ, so that the unconditional law of motion of beliefs in between moments of information acquisition is now given by (treating $p_t \in \Delta(\Theta)$ as a vector in the simplex) :

$$dp_t = \Lambda p_t dt \Longleftrightarrow p_t = e^{t\Lambda} p_0$$

where $e^{t\Lambda}$ denotes the matrix exponential.

- the cost function over experiments $C : \Delta\Delta(\Theta) \to \overline{\mathbb{R}}_+$ is:

  1. continuous over each Bayes-plausible subset, i.e. for any p, C is continuous over $\mathscr{B}(p)$ (where we consider $\Delta\Delta(\Theta)$ equipped with its weak-* topology, i.e. the topology of convergence in distribution); this is not needed for uniqueness of the fixed point but guarantees its convexity.

  2. "concave in the prior", following the definition in Denti et al. [2022], i.e. for any given belief p, the cost associated to the distribution of beliefs generated by the convex combination of any two Blackwell experiments via Bayes rule with prior p is weakly higher than the convex combination of the costs for the distribution of beliefs generated by the individual Blackwell experiments. As stated in Denti et al. [2022], this is a property that looks strong but is natural (it is, in particular, necessary for costs to be experimental). The main role of this property is to guarantee that the value function is convex.

B.1 Existence and uniqueness of a fixed point

In this section, I prove that the recursive operator Φ has a unique fixed point in $\mathbb{V}$ (which therefore must be the value function). Formally, the proof relies on the following general result, adapted from Marinacci and Montrucchio (2019) for the particular setting here:

Lemma B.1 (Theorem 1 in Marinacci and Montrucchio (2019), adapted)**.** *Let R a Riesz space (partially-ordered vector space with order structure $\leq$ making it a lattice) and $\Phi : A \to A$ a monotone and order-convex self-map defined on A an order-convex and chain-complete subset of R. Denote by $\partial^{\diamond} A$ the upper-perimeter of A (set of elements of A such that the segment with some other dominated element in A cannot be prolonged without exiting A). If $\Phi(x) \neq x$ for all $x \in \partial^{\diamond} A$, then Φ has a unique fixed point.*

Using the notations introduced in the main text, we want to prove that the Bellman operator Φ has a unique fixed point over the space of candidate value functions $\mathbb{V}$.

Proposition B.1 (FP Uniqueness)**.** *Φ has a unique fixed point in $\mathbb{V}$.*

Proof. The proof relies on applying Lemma B.1 to show that Φ has a unique fixed point in $[\underline{v}, \overline{v}]$. We just need to check the conditions of the theorem:

(i) **Space structure:**

The space of real-valued bounded measurable functions on $\Delta(\Theta)$ equipped with the pointwise order is a Riesz space; $\mathbb{V}$ is order-convex and chain-complete since on a Riesz space those sets are exactly the order intervals and $\mathbb{V} = [\underline{v}, \overline{v}]$.

(ii) **Φ is monotone:**
Recall $\Phi = \mathcal{W} \circ \mathcal{G}$ so it suffices to prove that both $\mathcal{W}$ and $\mathcal{G}$ are monotone, which is direct using pointwise domination inside of each supremum.

(iii) **Stability:**
We need to verify that $\Phi(I) \subseteq I$. Clearly for any $v \in I$ by definition of the supremum and feasibility of the policy which consists in never acquiring information $\Phi v \geq \underline{v}$. To show the other inequality observe that since $\overline{v}$ is linear:

$$\begin{aligned}
\mathcal{G}\,\overline{v}(p) &= \sup_{F \in \mathcal{B}(p)} \left\{ \mathbb{E}_{q \sim F}[\overline{v}(q)] - C(F) - \kappa \right\} \\
&= \sup_{F \in \mathcal{B}(p)} \left\{ \int_{\Delta(\Theta)} \int_0^\infty e^{-rt} \Big(\sum_{\theta \in \Theta} q_t(\theta) u(\delta_\theta) \Big) dt dF(q) - C(F) - \kappa \right\} \\
&= \sup_{F \in \mathcal{B}(p)} \left\{ \int_0^\infty e^{-rt} \Big(\int_{\Delta(\Theta)} \big(q e^{t\Lambda}\big) \cdot \vec{u} dF(q) \Big) dt - C(F) - \kappa \right\} \\
&= \sup_{F \in \mathcal{B}(p)} \left\{ \int_0^\infty e^{-rt} \big(p e^{t\Lambda}\big) \cdot \vec{u} dt - C(F) - \kappa \right\} \\
&= \overline{v}(p) - \kappa
\end{aligned}$$

Where the last line comes from observing that $F = \delta_p$ is optimal and that the integral term is just $\overline{v}(p)$. Since this establishes $\mathcal{G}\,\overline{v} = \overline{v} - \kappa$, we have:

$$\Phi \overline{v}(p) = \sup_\tau \int_0^\tau r e^{-rt} u(p_t) dt + e^{-r\tau} (\overline{v}(p_\tau) - \kappa) \leq \sup_\tau \int_0^\tau r e^{-rt} u(p_t) dt + e^{-r\tau} \overline{v}(p_\tau) = \overline{v}(p)$$

Where the first inequality is just by definition of the supremum and the second equality comes from the fact that $\overline{v}$ is the best achievable flow payoff starting from any belief. Using this and monotonicity of Φ gives for any $v \in A$, $\Phi v \leq \Phi \overline{v} < \overline{v}$. Putting the two together gives $\Phi(I) \subseteq I$.

(iv) **Φ is order-convex:**
It suffices to show that $\mathcal{W}$ and $\mathcal{G}$ are order-convex to get that $\Phi = \mathcal{W} \circ \mathcal{G}$ is order-convex (order-convexity of $\mathcal{G}$ trivially implies the same for $\mathcal{G}$). Let v, w in A; let $\beta \in [0,1]$. To show that $\mathcal{G}$ is order-convex observe that:

$$\begin{aligned}
\mathcal{G}\big[\beta v + (1-\beta) w\big](p) &= \sup_{F \in \mathcal{B}(p)} \left\{ \mathbb{E}_{q \sim F}[\beta v(q) + (1-\beta) w(q)] - C(F) \right\} \\
&= \sup_{F \in \mathcal{B}(p)} \left\{ \beta \mathbb{E}_{q \sim F}[v(q)] + (1-\beta) \mathbb{E}_{q \sim F}[w(q)] - C(F) \right\}
\end{aligned}$$

$$\leq \beta \sup_{F \in \mathscr{B}(p)} \Big\{ \beta \mathbb{E}_{q \sim F}[v(q)] - C(F) \Big\} + (1-\beta) \sup_{F \in \mathscr{B}(p)} \Big\{ \mathbb{E}_{q \sim F}[w(q)] - C(F) \Big\}$$
$$= \beta \mathcal{G}\, v(p) + (1-\beta) \mathcal{G}\, w(p)$$

It is similarly direct to prove that $\mathcal{W}$ is order-convex:

$$\mathcal{W}[\beta f + (1-\beta) g](p) = \sup_{\tau} \int_0^{\tau} e^{-rt} u(p_t) dt + e^{-r\tau} (\beta f + (1-\beta) g)(p_\tau)$$
$$\leq \beta \sup_{\tau} \Big\{ \int_0^{\tau} e^{-rt} u(p_t) dt + e^{-r\tau} f(p_\tau) \Big\} + (1-\beta) \sup_{\tau} \Big\{ \int_0^{\tau} e^{-rt} u(p_t) dt + e^{-r\tau} g(p_\tau) \Big\}$$
$$= \beta \mathcal{W} f(p) + (1-\beta) \mathcal{W} g(p)$$

(v) **Upper perimeter condition ($\Phi(v) \neq v$ for all $v \in \partial^{\diamond} A$) :**
First observe that the upper perimeter is given in this case by:

$$\partial^{\diamond} \mathbb{V} = \{ w \in \mathbb{V} \,|\, \inf_{p \in \Delta(\Theta)} \overline{v}(p) - w(p) = 0 \}$$

i.e. it is the set of functions that get arbitrarily close to the upper bound $\overline{v}$. This is obtained by definition of the upper perimeter and can also be seen from Proposition 4 in Marinacci and Montrucchio. Now take any $w \in \partial^{\diamond} I$ and any p such that $\overline{v}(p) - w(p) < \varepsilon$ for $\varepsilon < \kappa/2$. We can show a direct contradiction to $\Phi w = w$ by observing that $\Phi w(p) \leq \overline{v}(p) - \kappa$ (intuitively the RHS is an upper bound on the best possible outcome: the agent cannot do strictly better than perfect observation right now and forever after, which itself cannot be obtained without paying the fixed cost at least once, even if we ignore all other costs).

□

B.2 Properties of the fixed point

Continuity, convexity, differentiability. We first establish that the fixed point of Φ must be continuous.

Lemma B.2. *The unique fixed point of Φ must be continuous.*

Proof. Observe that both operators $\mathcal{G}, \mathcal{W}$ map continuous functions to continuous functions (this can be proven fai using e.g. Berge's theorem of the maximum). Therefore, we can redo the proof of existence and uniqueness as before over the subset of continuous functions in $\mathbb{V}$; since this yields a fixed point of Φ in a subset of $\mathbb{V}$, this fixed point must be the unique one over the whole set. □

Second, an expected result which carries over from static information acquisition is that the value function is *convex in the current belief.* The intuition is the same, even though the dynamics make its precise understanding more subtle: convex combinations of beliefs correspond to less information (garblings) and a rational DM is always better off with more information. Concavity in the prior of the cost guarantees that the value of information in any static information acquisition problem is convex. This is a well known result in the literature (see e.g. Proposition 5 in Denti et al. [2022]), which we merely restate in our context.

Lemma B.3. *For any continuous function w, $\mathcal{G}\, w$ is a convex function.*

Similarly, the "stopping value" operator also preserves convexity; the proof is also very standard: since u is convex in p, one can verify that $v(p)$ is a supremum of convex functions in p.

Lemma B.4. *If g is a convex continuous function, $\mathcal{W}\, g$ is convex.*

The combination of Lemma B.3 and Lemma B.4 immediately yields that a fixed point $v = \mathcal{W}\mathcal{G}\, v$ must be convex. We can also obtain results on the differentiability of v (proven here only for the binary state case, for simplicity).

Proposition B.2. *Assume that $\Theta = \{0,1\}$ and for any w, $\mathcal{G}\, w$ is differentiable inside the region where $\mathcal{G}\, w > w$. Then the value function v is differentiable everywhere.*

Proof. Let v the unique fixed point Φ; since v is convex, it is differentiable almost everywhere – hence it has left and right derivatives everywhere. From Proposition 1, v is is differentiable in the interior of the information acquisition region (since it is a collection of intervals and within each interval v is equal to an affine function plus c, which is differentiable). Fix any $p \in \Delta(\Theta) \setminus \text{int}\mathscr{I}^*$; denote $v'_-(p)$ and $v'_+(p)$ the left and right derivatives of v' at p. Next, consider two sequences $\{p_n^-\}$ and $\{p_n^+\}$ converging to p respectively from the left and from the right, such that v is differentiable at each p_n^-, p_n^+. If p is in the interior (but $p \neq \pi$ for simplicity) of the continuation region then for n large enough, using the differential characterization for the optimal stopping problem (in the viscosity sense in general, but here we only look at points of differentiability of the value function):

$$v'(p_n^-) = \frac{r\,v(p_n^-) - u(p_n^-)}{\lambda(\pi - p_n^-)} \xrightarrow[n\to\infty]{} \frac{r\,v(p) - u(p)}{\lambda(\pi - p)}$$
$$v'(p_n^+) = \frac{r\,v(p_n^+) - u(p_n^+)}{\lambda(\pi - p_n^+)} \xrightarrow[n\to\infty]{} \frac{r\,v(p) - u(p)}{\lambda(\pi - p)}$$

where the limit follows by continuity of u and v; this immediately implies $v'_-(p) = v'_+(p) = \frac{r\,v(p)-u(p)}{\lambda(\pi-p)}$, so v is differentiable at p. If $p = \pi$ is in the interior of the continuation region then v is differentiable at π if and only if u is differentiable at π.

Consider next p on the boundary of the information acquisition region $p \in \partial \mathscr{I}^*$ but *not on path except as an initial point*; if $p \geq \pi$ (resp. $p \leq \pi$) this means that point in an interval to the *right* (resp. *left*) of p are in the information acquisition region. Consider the latter case ($p \geq \pi$), the other one being symmetric. The same logic as before implies that $v'_-(p) = \frac{r\,v(p)-u(p)}{\lambda(\pi-p)}$; further we know that v is differentiable in an open neighborhood to the right of p and by optimality in the optimal stopping problem it must be that for all $q > p$ close enough p (i.e. such that $q \in \mathscr{I}^*$):

$$r\,v(q) \geq u(q) + \lambda(\pi - q)\,v'(q)$$

rearranging and taking the limit as q goes to p yields:

$$v'_+(p) \leq \frac{r\,v(p) - u(p)}{\lambda(\pi - p)} = v'_-(p)$$

but since we know $v'_- \leq v'_+$ by convexity, this means $v'_+(p) = v'_-(p)$ so v is differentiable at p.

In cases where $p \in \partial \mathscr{I}^*$ but p in on path, a standard smooth pasting argument delivers differentiability (assuming a kink yields a contradiction to the optimality of stopping at p). □

FIXED POINT ITERATIONS AND CONSTRAINED PROBLEMS.

Lemma B.5. *For any n, denote by $\underline{v}_n$ the n-th iteration of Φ starting from $\underline{v}$, i.e.:*

$$\underline{v}_n := \Phi^n \underline{v}$$

Then the sequence $\{\underline{v}_n\}$:

(i) is increasing:

$$\forall n \geq 0 \quad \underline{v}_{n+1} \geq \underline{v}_n$$

(ii) converges to the solution of the optimal information acquisition problem:

$$\underline{v}_n \longrightarrow v \quad \text{(pointwise)}$$

(iii) is equal for each n the solution of the constrained information acquisition problem where the DM is only allowed to acquire information at most n times.

Proof. Point (i) is direct given $\Phi w \geq \underline{v}$ for all w (it is always feasible not to acquire information) and using monotonicity of Φ to prove the clain by induction. Point (ii) is guaranteed by monotone convergence and the fixed point result. To prove point (iii), denote v^n the value function in the constrained problem with a "budget" of n times of information acquisition. Such values must verify the recursive equation:

$$v^n := \max\{\Phi v^{n-1}, v^{n-1}\}$$

where the maximum is understood pointwise. Furthermore by definition $v^0 = \underline{v}$. If $\underline{v}_n = v^n$ for $n = 0, \ldots, N$, then:

$$\Phi v^N = \Phi \underline{v}_N = \underline{v}_{N+1} \geq \underline{v}_N = v^N$$

which implies that:

$$v^{N+1} = \max\{\Phi v^N, v^N\} = \Phi v^N = \Phi \underline{v}_N = \underline{v}_{N+1}$$

hence this proves the claim by induction. □

A symmetrical proposition holds from the upper bound, with a similar proof.

Proposition B.3. *For any n, denote by $\overline{v}_n$ the n-th iteration of the fixed point operator Φ, starting from $\overline{v}$, i.e.:*

$$\overline{v}_n := \Phi^n \overline{v}$$

Then, the sequence $\{\overline{v}_n\}$:

(i) is decreasing:

$$\forall n \geq 0 \quad \overline{v}_{n+1} \leq \overline{v}_n$$

(ii) converges to the solution of the optimal information acquisition problem:

$$\overline{v}_n \longrightarrow v \quad \text{(pointwise)}$$

(iii) is equal for each n to the solution of the relaxed information acquisition problem where the DM is allowed to perfectly observe the state after n moments of information acquisition

Normalization. We briefly state a useful normalization result: starting from some arbitrary problem, the value function in the modified problem where an affine function is added to the indirect utility can itself be obtained as a (different but explicit) linear transformation of the original value function. This implies that the optimal policy in all problems obtained with such transformations is identical – which in practice allows us to avoid redundancy or normalize things in some convenient fashion to characterize solutions.

Proposition B.4 (Normalization by affine functions)**.** *Assume $rI - \Lambda$ is invertible. Let u some indirect utility function and v the corresponding value function; define some arbitrary linear function $L(p) = a \cdot p + b$, where $a \in \mathbb{R}^N$ and $\cdot$ denotes the scalar product; let the modified indirect utility $\tilde{u}$ be defined by:*

$$\tilde{u}(p) := u(p) + L(p)$$

Denote $\tilde{v}$ the value function corresponding to $\tilde{u}$. We have:

$$\tilde{v}(p) = v(p) + b + a \cdot \left((rI - \Lambda)^{-1} p\right)$$

Proof. Consider the original version of the problem (not in terms of beliefs) and define the random variable $X(\theta_t)$ which gives the "coordinate" of the current state in some arbitrary fixed ordering of the state space Θ: $X(\theta_t)$ is a vector of length $|\Theta|$ that has a 1 in the current value of θ_t and zeroes everywhere else. By definition $\mathbb{E}[X(\theta_t)] = p_t$. Define the "primary" utility function $\tilde{u}(\alpha_t, \theta_t) = u(\alpha_t, \theta_t) + a \cdot X(\theta_t) + b$ and the natural extension to indirect utility function $\tilde{u}(p_t) = u(p_t) + a \cdot p_t + b$. We have:

$$\begin{aligned}
\tilde{v}(p) &= \sup_{\alpha,\tau,F} \mathbb{E} \int_0^\infty e^{-rt}(u(\alpha_t;\theta_t) + a \cdot X(\theta_t) + b)dt - \sum e^{-r\tau_j}(C(F_j) - \kappa) \\
&= v(p) + \mathbb{E} \int_0^\infty e^{-rt}(a \cdot X(\theta_t) + b)dt \\
&= v(p) + \int_0^\infty e^{-rt}(a \cdot p_t + b)dt \\
&= v(p) + \int_0^\infty e^{-rt} a \cdot e^{t\Lambda} p\, dt + \frac{b}{r} \\
&= v(p) + a^\intercal \Big(\int_0^\infty e^{-rt} e^{t\Lambda} dt\Big) p + \frac{b}{r} \\
&= v(p) + a^\intercal (rI - \Lambda)^{-1} p + \frac{b}{r}
\end{aligned}$$

Where the fact that $\int_0^\infty e^{-rt} e^{t\Lambda} dt = (rI - \Lambda)^{-1}$ is a standard fact on the Laplace transform of matrix exponentials. □

Because the modified value function is a linear transformation, we can use the characterization of optimal policies via the "as-if-static" information acquisition problem to obtain that optimal policies are invariant to such transformations – indeed, by definition for any $\tilde{a}, \tilde{b} \in \mathbb{R}$ and any $w \in \mathbb{V}$, letting $\tilde{w}(p) := w(p) + \tilde{b} + \tilde{a}p$, we have $\mathcal{G}\, \tilde{w}(p) = \mathcal{G}\, w(p) + \tilde{b} + \tilde{a}p$. This observation directly implies the following corollary:

Corollary B.1. *Let u and $\tilde{u}$ two indirect utility functions such that $\tilde{u}$ can be obtained from u by adding an affine function. Any Markovian policy which is optimal in the information acquisition problem under u is optimal in the one under $\tilde{u}$ (all else being equal).*

For convenience, restate explicitly the one dimensional version of the previous result.

Corollary B.2 (Normalization by affine functions)**.** *Let u an indirect utility function and v the corresponding value function; define some arbitrary affine function $p \mapsto ap + b$, where $a, b \in \mathbb{R}$. Let the modified indirect utility $\tilde{u}$ be defined by:*

$$\tilde{u}(p) := u(p) + ap + b$$

Denote $\tilde{v}$ the value function corresponding to the information acquisition problem with $\tilde{u}$ as indirect utility function. We have:

$$\tilde{v}(p) = v(p) + \frac{b}{r} + a \cdot \left(\frac{r}{r+\lambda} p + \frac{\lambda}{r+\lambda} \pi \right)$$

This invariance to adding an affine function to the indirect utility is useful either directly to characterize optimal policies across classes of problems. Consider for instance two problems with binary actions $a \in \{0,1\}$: in one, the DM gets a flow payoff of 1 for choosing $a = \theta$; in the other, there is a safe action $u(a,\theta) = 0$ and a risky action which gives a payoff of 1 in state 1 and -1 in state in state 0. The former yields indirect utility function $\max\{p, 1-p\}$, the latter gives $\max\{0, 2p-1\}$. Since they can be obtained from each other by adding an affine function, they have the same optimal information acquisition policy. Another useful implication of this result for the purpose of some proofs is that, by adding a constant, we can assume without loss that the indirect utility function only takes non-negative values.

Other normalizations or rather equivalences between parameter transformations can be obtained rather easily in the problem but since they are of limited general usefulness we omit their explicit statement. It is useful, however, to mention that there is *no parameter reduction possible in general* – unlike what first intuition might suggest. In particular, the discount factor r and the volatility of the Markov chain λ, although they both control in some sense an exponential scaling of time, are *not* substitutable.

B.3 Verification, existence of solutions, sufficiency of Markov policies

First, state a preliminary result on existence of solutions in the subproblems which form the recursive equation.

Lemma B.6. *Let v the unique fixed point of Φ. For any $p \in \Delta(\Theta)$, let:*

$$S^*(p) := \underset{\tau \in [0,\infty]}{\arg\max} \int_0^\tau e^{-rt} u(p_t) dt + e^{-r\tau} \mathcal{G}\, v(p_\tau),$$
$$E^*(p) := \underset{F \in \mathscr{B}(p)}{\arg\max} \int v dF - C(F).$$

Both $S^*(p) \neq \emptyset$ and $E^*(p) \neq \emptyset$.

In general, there is no real existence problem for the stopping times, provided we allow for never stopping ($\tau = \infty$), as existence of a maximizer follows from continuity of the objective function over the extended real line. To argue that this generates a proper dynamic policy, however, requires verifying that the corresponding stopping times generate a well defined belief process as before. This is actually straightforward because the fixed cost $\kappa > 0$ guarantees that there is no incentive to continuously acquire information, so that the induced policy satisfies the requirements. Existence of optimal experiments is easily obtained with UPS costs since both $F \mapsto \int v dF$ and C are continuous over each of the $\mathscr{B}(p)$, which are compact in the weak-* topology over $\Delta\Delta(\Theta)$; note much weaker assumptions could be substituted here.

A GENERAL VERIFICATION THEOREM.

Proposition B.5 (Verification: optimal policies given value function)**.** *Let v be the unique fixed point of Φ. Any optimal strategy $\{\tau_i, F_i\} \in \Xi$ must verify a.s. for any $i \in \mathbb{N}$:*

$$\tau_i - \tau_{i-1} \in \underset{\substack{\tau \in [0,\infty] \\ p_0 = P_{\tau_{i-1}}}}{\arg\max} \int_0^\tau e^{-rt} u(p_t) dt + e^{-r\tau} \mathcal{G}\, v(p_\tau);$$

$$F_i \in \underset{F \in \mathscr{B}(P_{\tau_i^-})}{\arg\max} \int v dF - C(F).$$

Where both argmaxes are non-empty a.s. Conversely, any strategy which is almost surely induced in this way by iterated selections of measurable mappings is optimal.

C Dynamics and long run behavior: omitted proofs

C.1 Convergence

First, introduce some notation. Define the "long run domain" $D = [q^0, q^1]$ as either the (closure of the) interval in Γ^* that contains π, if there is one, or if not an arbitrarily chosen closed interval around π in which no information is acquired. Clearly, once the belief process enters D, it must either follow a cycle or information acquisition must stop. Hence to prove the claim it suffices to prove that the first entry time of the process in D is almost surely finite.

Consider an arbitrary initial belief p, assume without loss that $p < \pi$ (the proof is symmetric in the alternative case) and that p is in the waiting region (the initial jump makes no difference). Denote by $\{(q_n^0, p_n^0, p_n^1, q_n^1)\}_{n\in N}$ the collection of "effective on-path information acquisition intervals", i.e intervals in Γ^* such that $\Gamma^* \cap \mathscr{I}^* \neq \emptyset$ (where (q_n^0, q_n^1) denote the endpoints of the interval in Γ^* and (p_n^0, p_n^1) the minimum and maximum of $\Gamma^* \cap \mathscr{I}^*$ respectively) that are between p and π but not in D. Note N is countable but not necessarily finite – label intervals using the natural numbers in a natural ordered fashion from left to right:

$$p < p_0^0 \text{ and for all } n: q_n^0 < q_n^1 < q_0^{n+1}.$$

For any $\tilde{q} < \tilde{p} < \pi$, denote $\tau(\tilde{q}, \tilde{p})$ the time it takes for beliefs to deterministically drift from $\tilde{q}$ to $\tilde{p}$:

$$\tau(\tilde{q}, \tilde{p}) := \frac{1}{\lambda}\log\Big(\frac{\pi - \tilde{q}}{\pi - \tilde{p}}\Big)$$

Define the following sequence of independent random variables:

$$T_0 := \tau(p, p_0^0) + \tau(q_0^0, p_0^0) \times X_0 \text{ where } X_0 \sim \mathscr{G}\Big(\frac{p_0^0 - q_0^0}{q_0^1 - q_0^0}\Big)$$
$$\forall n \geq 1,\, T_n := \tau(q_1^{n-1}, p_{n-1}^0) + \tau(q_n^0, p_n^0) \times X_n \text{ where } X_n \sim \mathscr{G}\Big(\frac{p_n^0 - q_n^0}{q_n^1 - q_n^0}\Big)$$

where all the X_n are independent and defined on some probability space $(\Omega, \mathscr{F}, \mathbb{P})$, and $\mathscr{G}$ denotes the geometric distribution. Intuitively, T_n describes the amount of times it takes to "cross over" the n-th interval of information acquisition, after crossing the $n-1$-th. This decomposes into the deterministic time it takes to first reach an information acquisition moment after the previous cross ($\tau(q_1^{n-1}, p_{n-1}^0)$) and then repetitions of the path between q_n^0 and p_n^0. The length of each such spell is $\tau(q_n^0, p_n^0)$ and the number of repetitions is an exponential random variable with parameter corresponding to the probability of jumping to q_n^1 from p_n^0, i.e $\frac{p_n^0 - q_n^0}{q_n^1 - q_n^0}$.

The main object of interest is the *total* time it takes to cross over all effective on-path information acquisition intervals, which is:

$$\mathbf{T} := \sum_{n\in N} T_n$$

The event $\{\mathbf{T} = \infty\}$ is a tail event in the sense that it is in the terminal σ-algebra of the sequence of σ-algebras generated by the T_n, hence a classical application of Kolmogorov's 0-1 law entails that:

$$\mathbb{P}(\mathbf{T} = \infty) \in \{0,1\}$$

or equivalently $\mathbb{P}(\mathbf{T} < \infty) = 1 - \mathbb{P}(\mathbf{T} = \infty) \in \{0,1\}$. Denote $E := \{\omega \in \Omega | \forall n,\ X_n = 1\}$ the set of realizations such that the process jumps over each interval on the first information acquisition time. By definition of the X_n, $\mathbb{P}(E) > 0$ and by construction of the T_n, for any $\omega \in E$:

$$\mathbf{T}(\omega) \leq \tau(p, q^0) < \infty$$

Hence $\mathbb{P}(\mathbf{T} < \infty) > 0$, so it must be that $\mathbb{P}(\mathbf{T} < \infty) = 1$. Up to a constant, $\mathbf{T}$ is the first entry time of the belief process in D, so this completes the proof.

C.2 Ergodic distribution of beliefs

Assume that the long run domain of beliefs is non-empty $[q^0, p^0] \cup [p^1, q^1] \neq \emptyset$ with $p^0 < \pi < p^1$, and, to simplify exposition, assume that the initial distribution of beliefs has a density μ_0 supported over $[q^0, p^0] \cup [p^1, q^1]$. Denote by $\mu(t,p)$ the density of population with belief p at time t. Following standard logic, we can start by observing that mass should be preserved along the flow of the belief process: for any p in the interior of $[q^0, p^0] \cup [p^1, q^1]$ and any $dt > 0$ small enough, the mass that was at p at time t must be at $p_{dt} = e^{-\lambda dt} p + (1 - e^{-\lambda dt})\pi$ at $t + dt$, i.e.:

$$\mu(t + dt, e^{-\lambda dt} p + (1 - e^{-\lambda dt})\pi) = \mu(t,p) \text{ for all } p \in (q^0, p^0) \cup (p^1, q^1), dt \text{ small enough}$$

Dividing and taking the limit at dt goes to zero[7] yields that μ must solve the usual transport equation in the interior of the domain:

$$\frac{\partial}{\partial t}\mu(t,p) + \lambda(\pi - p)\frac{\partial}{\partial p}\mu(t,p) = 0 \text{ for all } p \in (q^0, p^0) \cup (p^1, q^1), t > 0$$

To complete the description, we need the initial condition $\mu(0,p) = \mu_0(p)$ and boundary conditions. We know that at every instant of time, each individual with a belief at say p^0 will acquire the Bayes-plausible experiment supported over $\{q^0, q^1\}$, so that the mass $\mu(t,p^0)$ will split according to the induced weights between q^0, q^1. This gives the boundary conditions, for any $t > 0$:

$$\mu(t,q^0) = \frac{q^1 - p^0}{q^1 - q^0}\mu(t,p^0) + \frac{q^1 - p^1}{q^1 - q^0}\mu(t,p^1)$$
$$\mu(t,q^1) = \frac{p^0 - q^0}{q^1 - q^0}\mu(t,p^0) + \frac{p^1 - q^0}{q^1 - q^0}\mu(t,p^1)$$

[7]We assume differentiability for ease of exposition; the differential characterization pins down dynamics and can be made rigorous even if μ is not classically differentiable, as is usual in the literature on partial differential equations.

Now, the ergodic distribution for that system must verify $\partial_t \mu = 0$ everywhere (i.e. be invariant over time), hence, from the transport equation it must verify in the interior of the domain:

$$\lambda(\pi - p)\frac{\partial}{\partial p}\mu(p) = 0$$

since $(\pi - p) \neq 0$, this entails that μ is piecewise constant on each interval $[q^0, p^0]$ and $[p^1, q^1]$, though not necessarily taking the same value; we denote m_0 and m_1 theses respective constant values. They are then pinned down by the boundary conditions and the preservation of mass $m_0(p^0 - q^0) + m_1(q^1 - p^1) = 1$, which yields the explicit expressions:

$$\begin{aligned} m_0 &= \frac{1}{2}\frac{1}{p^0 - q^0} \\ m_1 &= \frac{1}{2}\frac{1}{q^1 - p^1} \end{aligned}$$

C.3 Explicit expression of stationary payoffs

Fix some stationary thresholds p_L, p_H and target beliefs q_1, q_0 verifying:

$$0 < q_0 < p_L < \pi < p_H < q_1$$

Recall the stationary values from the induced belief cycle which we denote v_0, v_L, v_H, v_1 (respectively at q_0, p_L, p_H, q_1) must satisfy:

$$\begin{aligned} v_0 &= \int_0^{\tau(q_0,p_L)} e^{-rt} u(q_{0t}) dt + e^{-r\tau(q_0,p_L)} v_L \\ v_L &= \frac{q_1 - p_L}{q_1 - q_0}\big(v_0 - c(q_0)\big) + \frac{p_L - q_0}{q_1 - q_0}\big(v_1 - c(q_1)\big) + c(p_L) - \kappa \\ v_1 &= \int_0^{\tau(q_1,p_H)} e^{-rt} u(q_{1t}) dt + e^{-r\tau(q_1,p_H)} v_H \\ v_H &= \frac{q_1 - p_H}{q_1 - q_0}\big(v_0 - c(q_0)\big) + \frac{p_H - q_0}{q_1 - q_0}\big(v_1 - c(q_1)\big) + c(p_H) - \kappa \end{aligned}$$

To make notations more compact (and the eventual expressions readable), denote by F_L, F_H the induced experiments at p_L and p_H respectively, so that we can denote:

$$\begin{aligned} \mathbf{C}(F_L) &= \frac{q_1 - p_L}{q_1 - q_0}c(q_0) + \frac{p_L - q_0}{q_1 - q_0}c(q_1) - c(p_L) + \kappa \\ \mathbf{C}(F_H) &= \frac{q_1 - p_H}{q_1 - q_0}c(q_0) + \frac{p_H - q_0}{q_1 - q_0}c(q_1) - c(p_H) + \kappa \end{aligned}$$

Similarly, denote τ_0 and τ_1 the induced times to next update after a 0 and 1 news respectively:

$$\tau_0 := \tau(q_0, p_L) = \frac{1}{\lambda}\log\frac{\pi - q_0}{\pi - p_L}$$

$$\tau_1 := \tau(q_1, p_H) = \frac{1}{\lambda} \log \frac{q_1 - \pi}{p_H - \pi}$$

We also use the compact notations for flow payoffs:

$$U(q,p) := \int_0^{\tau(q,p)} e^{-rt} u(q_t) dt$$

Rewrite the four stationary values equations as two, in terms of v_0, v_1:

$$v_0 = U(q_0, p_L) + e^{-r\tau_0}\Big(\frac{q_1 - p_L}{q_1 - q_0} v_0 + \frac{p_L - q_0}{q_1 - q_0} v_1 - \mathbf{C}(F_L)\Big)$$
$$v_1 = U(q_1, p_H) + e^{-r\tau_1}\Big(\frac{q_1 - p_H}{q_1 - q_0} v_0 + \frac{p_H - q_0}{q_1 - q_0} v_1 - \mathbf{C}(F_H)\Big)$$

Rearrange the first equation to express v_0 in terms of v_1:

$$v_0 = \Big(1 - e^{-r\tau_0} \frac{q_1 - p_L}{q_1 - q_0}\Big)^{-1} \Big(U(q_0, p_L) + e^{-r\tau_0}\Big(\frac{p_L - q_0}{q_1 - q_0} v_1 - \mathbf{C}(F_L)\Big)\Big)$$

Plug it in the equation for v_1:

$$v_1 = U(q_1, p_H) + e^{-r\tau_1}\Bigg[\frac{q_1 - p_H}{q_1 - q_0} \quad \Big(1 - e^{-r\tau_0} \frac{q_1 - p_L}{q_1 - q_0}\Big)^{-1} \Big(U(q_0, p_L) + e^{-r\tau_0}\Big(\frac{p_L - q_0}{q_1 - q_0} v_1 - \mathbf{C}(F_L)\Big)\Big) + \frac{p_H - q_0}{q_1 - q_0} v_1 - \mathbf{C}(F_H)\Bigg]$$

which can be rearranged into:

$$v_1 = \frac{U(q_1, p_H) - e^{-r\tau_1}\mathbf{C}(F_H) + \frac{e^{-r\tau_1} \frac{q_1 - p_H}{q_1 - q_0}}{1 - e^{-r\tau_0} \frac{q_1 - p_L}{q_1 - q_0}} \Big(U(q_0, p_L) - e^{-r\tau_0}\mathbf{C}(F_L)\Big)}{1 - e^{-r\tau_1}\Big(\frac{p_H - q_0}{q_1 - q_0} + \frac{q_1 - p_H}{q_1 - q_0} \frac{e^{-r\tau_0}}{1 - e^{-r\tau_0} \frac{q_1 - p_L}{q_1 - q_0}} \frac{p_L - q_0}{q_1 - q_0}\Big)}$$

Since all the quantities in the right hand side have an explicit solution, this gives the solution of the system (as we can then express v_0 as a function of v_1 and v_L, v_H as a function of v_0, v_1). This expression essenitally amounts to the same simple intuition that appears more clearly in the symmetric case: payoffs derive from simple repetition of the cycle. Indeed notice that the terms:

$$U(q_1, p_H) - e^{-r\tau_1}\mathbf{C}(F_H) \text{ and } U(q_0, p_L) - e^{-r\tau_0}\mathbf{C}(F_L)$$

capture the total accumulated net payoffs over each cycle, starting from q_1 and q_0 respectively. The overall denominator captures the discounting coming from average expected period length, conditional on starting from q_1, while the term in front of $U(q_0, p_L) - e^{-r\tau_0}\mathbf{C}(F_L)$ captures the relative expected prevalence of 0-cycles, given we are starting from a 1 cycle. Note that few further simplifications are possible that are obviously generally useful, but many manipulations can be done in particular problems; also observe that we can express the exponential terms explicitly in terms of beliefs as:

$$e^{-r\tau_1} = e^{-\frac{r}{\lambda}\log\frac{\pi-q_0}{\pi-p_L}} = \left(\frac{\pi-q_0}{\pi-p_L}\right)^{-\frac{r}{\lambda}}$$

$$e^{-r\tau_0} = e^{-\frac{r}{\lambda}\log\frac{q_1-\pi}{p_H-\pi}} = \left(\frac{q_1-\pi}{p_H-\pi}\right)^{-\frac{r}{\lambda}}$$

We can give a full expression for the objective function in the π-initialized problem, which highlights that the continuation value are a convex combination of the cycle payoffs; keeping the same notations as before denote v_π the value from starting at π and immediately jumping at q_0, q_1 (with the appropriate probabilities and at the corresponding cost):

$$v_\pi := \frac{\pi-q_0}{q_1-q_0}v_1 + \frac{q_1-\pi}{q_1-q_0}v_0 - \mathbf{C}(F_0)$$

where F_0 is the initial jump distribution i.e.:

$$\mathbf{C}(F_0) = \frac{\pi-q_0}{q_1-q_0}\big(c(q_1)-c(\pi)\big) + \frac{q_1-\pi}{q_1-q_0}\big(c(q_1)-c(\pi)\big) + \kappa$$

Making this expression explicit yields:

$$\begin{aligned}
v_\pi = {} & \frac{\pi-q_0}{q_1-q_0}\frac{U(q_1,p_H) - e^{-r\tau_1}\mathbf{C}(F_H) + \frac{e^{-r\tau_1}\frac{q_1-p_H}{q_1-q_0}}{1-e^{-r\tau_0}\frac{q_1-p_L}{q_1-q_0}}\Big(U(q_0,p_L)-e^{-r\tau_0}\mathbf{C}(F_L)\Big)}{1-e^{-r\tau_1}\left(\frac{p_H-q_0}{q_1-q_0}+\frac{q_1-p_H}{q_1-q_0}\frac{e^{-r\tau_0}}{1-e^{-r\tau_0}\frac{q_1-p_L}{q_1-q_0}}\frac{p_L-q_0}{q_1-q_0}\right)} \\
& + \frac{q_1-\pi}{q_1-q_0}\frac{U(q_0,p_L) - e^{-r\tau_0}\mathbf{C}(F_L) + \frac{e^{-r\tau_0}\frac{p_L-q_0}{q_1-q_0}}{1-e^{-r\tau_1}\frac{p_H-q_0}{q_1-q_0}}\Big(U(q_1,p_H)-e^{-r\tau_1}\mathbf{C}(F_H)\Big)}{1-e^{-r\tau_0}\left(\frac{q_1-p_L}{q_1-q_0}+\frac{p_L-q_0}{q_1-q_0}\frac{e^{-r\tau_1}}{1-e^{-r\tau_1}\frac{p_H-q_0}{q_1-q_0}}\frac{q_1-p_H}{q_1-q_0}\right)} \quad - \mathbf{C}(F_0) \\
=: {} & \gamma_1 \times \big(U(q_1,p_H) - e^{-r\tau_1}\mathbf{C}(F_H)\big) + \gamma_0 \times \big(U(q_0,p_L) - e^{-r\tau_0}\mathbf{C}(F_L)\big) - \mathbf{C}(F_0)
\end{aligned}$$

where:

$$\gamma_1 := \frac{\frac{\pi-q_0}{q_1-q_0}}{1-e^{-r\tau_1}\left(\frac{p_H-q_0}{q_1-q_0}+\frac{q_1-p_H}{q_1-q_0}\frac{e^{-r\tau_0}}{1-e^{-r\tau_0}\frac{q_1-p_L}{q_1-q_0}}\frac{p_L-q_0}{q_1-q_0}\right)} + \frac{\frac{q_1-\pi}{q_1-q_0}\frac{e^{-r\tau_0}\frac{p_L-q_0}{q_1-q_0}}{1-e^{-r\tau_1}\frac{p_H-q_0}{q_1-q_0}}}{1-e^{-r\tau_0}\left(\frac{q_1-p_L}{q_1-q_0}+\frac{p_L-q_0}{q_1-q_0}\frac{e^{-r\tau_1}}{1-e^{-r\tau_1}\frac{p_H-q_0}{q_1-q_0}}\frac{q_1-p_H}{q_1-q_0}\right)}$$

$$\gamma_0 := \frac{\frac{\pi-q_0}{q_1-q_0}\frac{e^{-r\tau_1}\frac{q_1-p_H}{q_1-q_0}}{1-e^{-r\tau_0}\frac{q_1-p_L}{q_1-q_0}}}{1-e^{-r\tau_1}\left(\frac{p_H-q_0}{q_1-q_0}+\frac{q_1-p_H}{q_1-q_0}\frac{e^{-r\tau_0}}{1-e^{-r\tau_0}\frac{q_1-p_L}{q_1-q_0}}\frac{p_L-q_0}{q_1-q_0}\right)} + \frac{\frac{q_1-\pi}{q_1-q_0}}{1-e^{-r\tau_0}\left(\frac{q_1-p_L}{q_1-q_0}+\frac{p_L-q_0}{q_1-q_0}\frac{e^{-r\tau_1}}{1-e^{-r\tau_1}\frac{p_H-q_0}{q_1-q_0}}\frac{q_1-p_H}{q_1-q_0}\right)}$$

D Information acquisition as optimization over path measures with and without fixed cost

In this section, we introduce the formalism that allows us to express the optimal information acquisition problem as an optimization over a measure space, specifically a space of measures over *belief paths*. This has several purposes:

1. it provides a convenient abstract formalism which highlights in particular that the optimal information acquisition problem is *a linear problem* over the appropriate space; this formalism is useful down the line for some proofs and of general methodological interest

2. it allows us to give a rigorous content to the "extension" of the problem in two directions:

   (a) allowing for non-discrete information acquisition

   (b) allowing for the case $\kappa = 0$

the technical contents of the first point consists mainly in careful definitions and topological considerations; the main difficulty for the second point is that our primitive cost is per se well-defined only over belief processes that feature only punctual information acquisition – hence we need a way to extend it over the whole space of belief processes. Fortunately, the topological arguments from the first point allow us to establish that $\mathfrak{C}$ (defined for any $\kappa \geq 0$) is uniformly continuous over the subset of belief processes which derive from punctual information acquisition, which is a dense subset of the whole space of belief processes: therefore it has a unique continuous extension. This has several consequences:

1. the unique extension of the cost over the whole space of belief processes coincides with existing costs in the literature (based on the infinitesimal generator) for processes that feature continuous information acquisition, but allows for a more general representation that explicitly allows for discontinuous information acquisition

2. this allows us to state rigorously that when $\kappa > 0$, punctual information acquisition is without loss (because any continuous information acquisition, even if allowed, generates infinite costs)

3. we also obtain natural approximation results: even when $\kappa = 0$ (so that continuous information acquisition *might* be optimal), punctual information acquisition can approximate optimal solutions over the whole class of belief processes.

D.1 Preliminaries: the space(s) of belief processes

We first need to formally define and give some topological considerations over the general space of belief processes. The key here is a classical methodological point from probability theory: we need to systematically view "belief processes" as random variables over the appropriate metric space – or equivalently as measures in the corresponding space; this is turn allows us to endow this space with the usual topology of convergence in distribution (the weak-* topology of the equivalent space of measures). In this subsection, we formally define all the appropriate objects – both for the general space of belief processes, and the subset which is generated by punctual information acquisition.

Belief paths. The first building block is to define a metric structure for the space of belief paths (the metric space over which our random belief processes take values). Throughout, we let D the space of $\Delta(\Theta)$-valued càdlàg functions over $[0,\infty)$; since the domain will always be time we call elements of D "belief paths". We equip D with the usual Skorohod metric which we denote d; we refer to Billingsley [2013] for the formal definition (or Pollard [1984] for an alternative version). Recall that D equipped with (a well-chosen version of) the Skorohod metric is a (compact subset of a) complete separable metric space. We further equip D with the Borel σ-algebra induced by its metric topology and denote $\Delta(D)$ the set of probability measures over this measurable space.

Belief processes. Let $(\Omega, \mathscr{F}, \mathbb{Q})$ a probability space. Formally, a belief process can be viewed as a random variable over D, i.e. a measurable function from $(\Omega, \mathscr{F}, \mathbb{Q})$ to D equipped with the Borel σ-algebra generated by its metric topology induced by d. Not every such random variables is a reasonable candidate to be a belief process: we need to impose consistency requirements (which follows from Bayes rule); this is done naturally by defining the space of belief processes as follows:

$$\boldsymbol{\mathcal{B}} := \left\{ P : \Omega \to D \text{ measurable} \;\middle|\; \forall\, t, s \geq 0,\;\; \mathbb{E}[P_{t+s} | P_t] = e^{s\Lambda} P_t \right\}$$

Where the condition over expectations is the equivalent of the martingale property of beliefs with a changing state: over any time interval, the expected evolution of beliefs conditional on the initial value should follow the unconditional law of motion prescribed by the underlying Markov chain (equivalently: the component of the change in belief due to information should be a martingale). Since we eventually only care about expectations induced by belief processes, we will identify elements in $\boldsymbol{\mathcal{B}}$ with their distributions and slightly abuse notations to interpret

$\mathcal{B}$ either as a space of random variables or as a space of measures, where for a measure $\mu \in \Delta(D)$, "$\mu \in \mathcal{B}$" means that there exists a random variable $P \in \mathcal{B}$ such that μ is the distribution of P.

Belief processes generated by punctual information acquisition. We now rephrase the definition given in Section 2 for the class of belief processes which is generated by punctual information acquisition. Recall the notation that for any information acquisition policy $\xi \in \Xi$, P^{ξ} denotes the corresponding belief process (where we omit the dependence on the initial point). Denote by $\mathcal{B}_{\text{d}}$ (where c stands for "countable") the subset of $\mathcal{B}$ such that information is only acquired at countably many moments in time:

$$\mathcal{B}_{\text{d}} = \left\{ Q \in \mathcal{B} \,\middle|\, \exists \xi \in \Xi,\ \ Q = P^{\xi} \text{ a.s.} \right\}$$

It is straightforward to verify that $\mathcal{B}_{\text{d}}$ is a dense class in $\mathcal{B}$ – indeed, it contains in particular any discrete-time approximation on a fixed countable grid.[8] For any belief process in $P \in \mathcal{B}_{\text{d}}$, we define $\{\tau_i^P\}_{i \in \mathbb{N}}$ the ordered random times of the discontinuities of P and for any i let F_i^P the conditional distribution of $P_{\tau_i} | P_{\tau_i^-}$. Straightforwardly $\{\tau_i^P, F_i^P\} \in \Xi$ and this pins down a unique (in distribution) element of Ξ so that there is a one to one mapping between $\mathcal{B}_{\text{d}}$ and Ξ (where, again, we identify random variables and their distributions).

D.2 Payoffs and costs over belief processes

Expected utility over belief processes. I tackle first the (easier) definition of the expected discounted utility in terms of some arbitrary belief process. First, observe that it is intuitive to define expected discounted utility over *paths*; for some arbitrary càdlàg belief path $p \in D$, let:

$$\begin{aligned} \mathscr{U} : D &\longrightarrow \mathbb{R} \\ p &\longmapsto \int_0^{\infty} e^{-rt} u(p_t) dt \end{aligned}$$

An important observation is that U is a continuous function over the space of paths: for any sequence of paths $p_n \in D$, $p_n \xrightarrow{d} p$ implies that for some sequence of increasing continuous bijective function $\lambda_n : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\lambda_n(t) \to t$ for all t as $n \to \infty$, $p_n(\lambda_n(t)) \to p(t)$ for almost every t (since every càdlàg path has countably many discontinuities, see Lemma 5.1. in Ethier and Kurtz [1986]); using convergence almost everywhere and continuity of u with a dom-

[8]For any $P \in \mathcal{B}$ and any $h > 0$ we can construct the approximate process $P^h \in \mathcal{B}_{\text{d}}$ by letting $\tau_0 = 0$, $\tau_{i+1} = \tau_i + h$, and $P^h_{\tau_i + t} = e^{t\Lambda} P_{\tau_i}$ for all $t \in [\tau_i, \tau_{i+1})$ (this is the càdlàg process which follows the deterministic drift in between h-spaced updates at which it just takes on the current value of P). By construction of $P \in \mathcal{B}$, $\mathbb{E}[P^h_t - P^h_{t^-} | P^h_{t^-}] = 0$ so that the distribution of the jumps of P^h indeed verifies Bayes Plausibility and P^h is a well defined element of $\mathcal{B}_{\text{d}}$. Clearly $P^h \xrightarrow[h \to 0]{a.s.} P$ hence in distribution as well.

inated convergence argument over the bound $|u(p_n(t)) - u(p(t))| \leq |u(p_n(t)) - u(p_n(\lambda_n t))| + |u(p_n(\lambda_n t)) - u(p(t))|$, we get that $U(p_n) \to U(p)$. Utility over belief processes is simply expected utility, where the expectation is over random paths:

$$\begin{aligned} \mathfrak{U} : \boldsymbol{\mathcal{B}} &\longrightarrow \mathbb{R} \\ P &\longmapsto \mathbb{E}\Big[\mathscr{U}(P)\Big] \end{aligned}$$

which is always well-defined since u is bounded and continuous. $\mathfrak{U}$ is defined using notations in terms of random variables but it is insightful and convenient to rewrite it equivalently in terms of the *distribution* of the process instead of the process itself. Abusing notations as previously in interchanging the measure and the process, we write:

$$\mathfrak{U}(\mu) := \int_D \mathscr{U}\, d\mu$$

This highlights that, because we are endowing $\boldsymbol{\mathcal{B}}$ with its weak-* topology and since U is continuous in D, $\mathfrak{U}$ is linear and continuous by construction. Naturally, this is also true within the dense subset $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$, and nothing in this construction needs to be made specific.

INFORMATION COSTS WITH PUNCTUAL INFORMATION ACQUISITION. Costs pose a different challenge: a priori, the construction that we are given for information costs is specific to the class of belief processes with punctual information acquisition. The key here will be to use the density of $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$ in $\boldsymbol{\mathcal{B}}$ to show that there exists a unique way to continuously extend the cost (rigorously defined over the $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$) on $\boldsymbol{\mathcal{B}}$. First, we recall the construction of the cost over $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$, recast in our notation over belief processes.

Another difference between payoffs and costs is that, whereas payoffs obviously depend only on the realized path of beliefs, costs generally do not – they depend on the *distribution* of the information being acquired and not just on its realization. Nevertheless, the assumption of UPS costs on primitive experiments precisely allows us to still write the cost as an expectation over paths. Let us start with the intuitive definition of the cost from the induced experiments:

$$\begin{aligned} \mathfrak{C} : \boldsymbol{\mathcal{B}}_{\mathrm{d}} &\longrightarrow \overline{\mathbb{R}}_+ \\ P &\longmapsto \mathbb{E}\Bigg[\sum_{i \in \mathbb{N}} e^{-r\tau_i^P} \mathbf{C}(F_i^P)\Bigg] \end{aligned}$$

Where recall that our assumption on the cost is:

$$\mathbf{C}(F) = C(F) + \kappa = \int c\, dF - c(p) + \kappa \text{ where } p = \int q\, dF(q)$$

Where importantly we now let $\kappa \geq 0$ (we do not impose a strictly positive fixed cost anymore as we will show in the next subsection that the problem is still well defined even when $\kappa = 0$). By construction, this then means for any $P \in \mathcal{B}_\text{d}$, we can rewrite:

$$C(F_i^P) = \mathbb{E}\left[c(P_{\tau_i}) - c(P_{\tau_i^-}) \mid P_{\tau_i^-}\right]$$

This immediately highlights that $\mathfrak{C}$ can be written as an expectation over paths. Define:

$$\begin{aligned} \mathscr{C} : D &\longrightarrow \overline{\mathbb{R}}_+ \\ p &\longmapsto \sum_{\{t \mid p_t \neq p_{t^-}\}} e^{-rt}\left(c(p_t) - c(p_{t^-}) + \kappa\right) \end{aligned}$$

this is well defined over D since any càdlàg path must have countably many discontinuities (Lemma 5.1. in Ethier and Kurtz [1986]). Then it is direct to observe that by construction, for any $\mu \in \mathcal{B}_\text{d}$:

$$\mathfrak{C}(\mu) = \int \mathscr{C}\, d\mu$$

The key observation is the following: $\mathscr{C}$ is continuous over D (this is fairly direct by definition and using continuity of c; it follows that $\mathfrak{C}$ is linear and uniformly continuous over $\mathcal{B}_\text{d}$ equipped with its induced weak-* topology.

Information costs from arbitrary belief processes. We now have a well-defined defined $\mathfrak{C}$ over $\mathcal{B}_\text{d}$ and we know that it is uniformly continuous. It is a classical result that given a uniformly continuous mapping from a subset of a metric space (to another metric space), there exists a unique continuous extension of this mapping to the closure of the original domain (which in the case of a dense subspace is the whole space). In other words, we can define for any $\mu \in \mathcal{B}$:

$$\mathfrak{C}(\mu) := \lim_{n\to\infty} \mathfrak{C}(\mu_n) \text{ where } \mu_n \in \mathcal{B}_\text{d} \text{ is an arbitrary sequence such that } \mu_n \to \mu$$

where the latter limit is naturally understood in the weak-* topology on $\mathcal{B}$. Observe that, crucially, uniform continuity is what guarantees that this is well defined and the choice of the approximating sequence is irrelevant – since it implies that for any μ_n, ν_n such that $\rho(\mu_n, \nu_n) \to 0$ (where ρ is a distance which metrizes the weak-* topology on $\mathcal{B}$), $|\mathfrak{C}(\mu_n) - \mathfrak{C}(\nu_n)| \to 0$, this guarantess that if any two sequences of belief processes converge to the same limit, the sequences of cost also do.

Several remarks on this extended cost function are in order. First, and crucially, it can very well take the value $+\infty$. Actually, if we assume $\kappa > 0$ any belief process which features continuous information acquisition on some interval of time a.s. will generate infinite cost – which will warrant looking for optimal processes in $\mathcal{B}_\mathrm{d}$. Second, if we let $\kappa = 0$, then the "extended" cost function considered over processes that feature "continuous information acquisition everywhere" coincides with a natural extension to the changing state environment of existing cost functions based on the infinitesimal generator – see e.g. Zhong [2022], Bloedel and Zhong [2020], Hébert and Woodford [2023], Georgiadis-Harris [2023]. Indeed, assume P is a belief process; let M_t be the *martingale component* of P_t, i.e. the "information acquisition" part of the change in beliefs, formally defined as:

$$M_t := P_t - \int_0^t \Lambda P_t dt$$

Further assume that M_t has a well-defined infinitesimal generator $\mathscr{A}_t$ defined as:

$$\mathscr{A}_t f(x) := \lim_{h \to 0} \frac{1}{h} \mathbb{E}\left[f(M_{t+h}) - f(M_t) \,\middle|\, M_t = x \right]$$

over some appropriate domain of functions $\mathscr{D}(\mathscr{A}_t)$, then:

$$\mathfrak{C}(P) = \mathbb{E}\left[\int_0^\infty e^{-rt} \mathscr{A}_t c(M_t) dt \right]$$

The overarching point here is that this class of processes forms another class where we can give a simple definition for the cost, but *those two classes do not overlap and do not cover the whole space of belief processes* – indeed, belief processes in $\mathcal{B}_\mathrm{d}$ do not have a well defined infinitesimal generator (though they have a useful and well-defined *extended* generator) so we cannot simply apply to them the cost as defined for processes that do. In either case, this poses a similar challenge of extending the cost function over the entire space of belief processes: we have just shown that there is a unique consistent (continuous) way to do so, through approximating any belief process arbitrarily well with processes in $\mathcal{B}_\mathrm{d}$. The properties we have proven also pave the way to arguing further in the next subsection that the problem written over $\mathcal{B}_\mathrm{d}$ is essentially without loss in terms of characterizing solutions as well.

D.3 Optimization over belief processes: equivalence and approximation

With all the proper definitions, consistent general notations, and useful properties at hand, we can now finally turn to the optimal information acquisition problem. First, we can rewrite in terms of optimization over belief processes the problem which is constrained to countably many times of information acquisition – which is the one we have been considering throughout the paper, up to the difference that we now explicitly allow for $\kappa = 0$. Let:

$$v_c(p) := \sup_{P \in \mathcal{B}_\mathrm{d}(p)} \mathfrak{U}(P) - \mathfrak{C}(P)$$

where $\boldsymbol{\mathcal{B}}_{\mathrm{d}}(p)$ is the subset of $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$ with fixed initial value $p \in \Delta(\Theta)$ for the process. The first important observation is that this is a well-defined problem for all $\kappa \geq 0$; however we have only proven in the main text that the sup is attained for $\kappa > 0$ – and indeed it is generally not the case for $\kappa = 0$. Straightforwardly thanks to the definition of the previous subsection, can define the same problem over the *entire* space of belief processes:

$$v(p) := \sup_{P \in \boldsymbol{\mathcal{B}}(p)} \mathfrak{U}(P) - \mathfrak{C}(P)$$

The main result follows straightforwardly by combining density of $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$ and continuity of $\mathfrak{U}$ and $\mathfrak{C}$: for any $p \in \Delta(\Theta)$ we have

$$v_c(p) = v(p),$$

furthermore any maximizing sequence in the unconstrained problem (i.e. any sequence $\{P^l\}$ in $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$ such that $\mathfrak{U}(P^l) - \mathfrak{C}(P^l) \to v(p)$ as $l \to \infty$) if it converges in $\boldsymbol{\mathcal{B}}$, must converge to a maximizer in the right-hand problem. In other words, even when the supremum is not attained any approximate maximizer in the class $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$ is an approximate maximizer in $\boldsymbol{\mathcal{B}}$ and the limits of approximate solutions in $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$ is a solution in $\boldsymbol{\mathcal{B}}$. To get existence of solutions in the unconstrained problem, it is relatively straightforward to verify that, using the fact that $\Delta(\Theta)$ is compact, we can apply Theorem 12.3 in Billingsley [2013] to verify that $\boldsymbol{\mathcal{B}}(p)$ is a (weak-*) compact subset of $\Delta(D)$.

We briefly highlight an issue that requires a little bit of care: when considering the approximate problem (optimization over $\boldsymbol{\mathcal{B}}_{\mathrm{d}}$ for the case $\kappa = 0$, the recursive characterization that we have used throughout the paper *does not directly extend*. To make this precise, denote v^κ the value function in the optimal information acquisition problem and Φ^κ the fixed point operator to highlight the dependence in κ, i.e.:

$$\Phi^\kappa w(p) := \sup_\tau \int_0^\tau e^{-rt} u(p_t) + e^{-r\tau} \sup_{F \in \mathscr{B}(p_\tau)} \left\{ \int w dF - C(F) - \kappa \right\}.$$

Recall that we have established before that for $\kappa > 0$, v_κ is the unique fixed point of Φ^κ. For $\kappa = 0$, we can however a priori only claim that v_0 is *a fixed point of* Φ^0 but Φ^0 does not have a unique fixed point. The problem appears worse at first, but it is not as severe as it seems: Φ^0 generically has infinitely many fixed points, but this is because it has many "unrealistic" fixed points. From observing the structure of Φ^0, one can observe that if w is better than any stream of payoffs which can be generated by flow payoffs from u, in the "as-if-one-off" stopping problem there will immediate stopping at every belief, leading to $\Phi^0 w = w$ but such a fixed point could never be the value of the optimal information acquisition problem. Technically, what fails in the proof of uniqueness in the upper perimeter condition.

D.4 Limit of solutions and solution of the limit problem.

In general, convergence of functionals is not equivalent to convergence of the maximizers; to claim convergence to solutions as $\kappa \downarrow 0$ to a solution of the problem when $\kappa = 0$, we need a different notion of convergence – canonically, the two equivalent notions of Γ-convergence and epi-convergence are used. [9] Here, we cannot use a stronger result like Berge's Theorem of the maximum because the cost function $\mathfrak{C}_\kappa$ is not continuous in the parameter κ at 0. For our purposes, epi-convergence is relatively straightforward to use given the construction for $\mathfrak{C}$ (we detail why below). We first recall its definition.

Definition D.1. Let (X,d) a metric space and a sequence of functionals $f_n : X \to \overline{\mathbb{R}}$. We say that f_n epi-converges to $f : X \to \overline{\mathbb{R}}$ if for every $x \in X$

(i) For any x_n s.t. $x_n \to x$, $f(x) \leq \liminf_n f_n(x_n)$

(ii) There exists $x_n \to x$ such that $f(x) \geq \limsup_n f_n(x_n)$

We will denote $f_n \xrightarrow[n\to\infty]{\text{epi}} f$.

Note that the condition in (ii), given (i), could equivalently be replaced by $f(x) = \lim_n f_n(x_n)$ by definition. The following result justifies the use of epi-convergence:

Proposition D.1. *If $f_n \xrightarrow[n\to\infty]{epi} f$ then for any sequence $x_n \in \arg\min f_n$:*

$$x_n \to x \Longrightarrow x \in \arg\min f$$

Proof. Let $x_n \in \arg\min f_n$, by definition of epi-convergence:

$$f(x) \leq \liminf_n f_n(x_n) = \liminf_n \{\inf f_n\} \leq \limsup_n \{\inf f_n\} \leq \inf f$$

Hence $x \in \arg\min f$. □

To apply these results, write $\mathfrak{C}_\kappa, \mathscr{C}_\kappa, \mathbf{C}_\kappa$ for the costs as previously introduced, highlighting the dependence in κ. As previously alluded to, epi-convergence of $\mathfrak{C}_\kappa$ to $\mathfrak{C}_0$ as $\kappa \downarrow 0$ is not only the minimal notion that we need to ensure convergence of the solutions, but also in some sense the best we can hope for: in general for a given μ in $\boldsymbol{\mathcal{B}}$, it is *not true* that $\mathfrak{C}_\kappa(\mu) \to \mathfrak{C}_0(\mu)$. The problem

[9] Γ-convergence is technically more general since it is defined over topological spaces; the two notions coincide over first countable spaces, in particular metric spaces.

here is essentially one of the order of limits: if we take "first" the limit in an approximating sequence μ_n and it so happens that for a *fixed* κ, $\mathfrak{C}_\kappa(\mu) = +\infty$, then taking the limit in κ "second" will only give $+\infty$; but it need not be that if we taken the limit over κ first instead, the cost would still blow up.

The intuition of the solution to this apparent problem brings together the definition of epi-convergence and our construction for $\mathfrak{C}_\kappa$. This is exactly why epi-convergence is broken down into conditions (i) and (ii); (i) is consistent with the behavior we just described: if we take limits in the "wrong order", we still have the right ordering of the "limit" value with the liminf; (ii) weakens the actual limit condition into existence of some well chosen sequence. Our construction for $\mathfrak{C}$ suggests exactly how to do that: for $\kappa > 0$, only consider approximations in $\boldsymbol{\mathcal{B}}_\mathrm{d}$ (this is always possible by density and ensures that we don't do the "wrong limit first"), and choose an approximation so that the discounted sum over all information acquisition times converges to ∞ slower than κ goes to 0.

Let us now state and prove the convergence result formally – this will be sufficient to conclude convergence of the maximizers.

Lemma D.1. *Let κ_n a sequence of strictly positive real numbers such that $\kappa_n \to 0$ as $n \to \infty$. Then:*

$$\mathfrak{C}_{\kappa_n} \xrightarrow[n\to\infty]{epi} \mathfrak{C}_0$$

Proof. Throughout, $\mathfrak{C}_\kappa$ is understood as a function the metric space $\boldsymbol{\mathcal{B}}$ understood as the space of measures and equipped with its weak-* topology. Let κ_n a sequence of strictly positive real numbers such that $\kappa_n \to 0$ as $n \to \infty$; to alleviate notation (and mirror the notations above) denote $\mathfrak{C}_n := \mathfrak{C}_{\kappa_n}$. Further denote:

$$\begin{aligned}\mathscr{C}_0 : D &\longrightarrow \overline{\mathbb{R}}_+\\ p &\longmapsto \sum_{\{t|p_t \neq p_{t^-}\}} e^{-rt}\big(c(p_t) - c(p_{t^-})\big)\end{aligned}$$

$$\begin{aligned}f_n : D &\longrightarrow \overline{\mathbb{R}}_+\\ p &\longmapsto \sum_{\{t|p_t \neq p_{t^-}\}} e^{-rt}\kappa_n\end{aligned}$$

so that we can break down the cost as $\mathscr{C}_n(p) := \mathscr{C}_0(p) + h_n(p)$ ($\mathscr{C}_0$ is the variable cost part and f_n is the fixed cost part) and we can write the costs over processes as:

$$\mathfrak{C}_n(\mu) = \lim_{m\to\infty} \int \mathscr{C}_n(p)\, d\mu_m(p) \text{ for any } \mu_m \xrightarrow[m\to\infty]{} \mu$$

$$= \lim_{m\to\infty} \int \big(\mathscr{C}_0(p) + f_n(p)\big) d\mu_m(p)$$

Now consider some arbitrary $\mu \in \boldsymbol{\mathcal{B}}$:

(i) Let any $\mu_n \to \mu$. Consider sequences $\mu_{n,m} \in \boldsymbol{\mathcal{B}}_{\mathrm{d}}$ such that for any n, $\mu_{n,m} \xrightarrow[m\to\infty]{} \mu_n$. We can write:

$$\begin{aligned}
\mathfrak{C}_0(\mu) &= \lim_{n\to\infty} \mathfrak{C}_0(\mu_n) \\
&= \liminf_{n\to\infty} \lim_{m\to\infty} \int \mathscr{C}_0(p) d\mu_{n,m}(p) \\
&\leq \liminf_{n\to\infty} \lim_{m\to\infty} \int \big(\mathscr{C}_0(p) + f_n(p)\big) d\mu_{n,m}(p) \\
&= \liminf_{n\to\infty} \mathfrak{C}_n(\mu_n)
\end{aligned}$$

Where the first equality is by continuity of $\mathfrak{C}_0$, the second by definition of $\mathfrak{C}_0$ (and replacing the limit with a lim inf without loss); the inequality is direct since we are adding a positive term and the last equality is by definition.

(ii) Construct the sequence μ_n as follows: let $P \sim \mu$; for each n, μ_n is the distribution of the process P^n which approximates P on a time grid with some fixed step size h_n – formally construct the grid by setting $\tau_0 = 0$, $\tau_j = \tau_{j-1} + h_n$, let P^n follow the unconditional drift for each $t \in [\tau_j, \tau_{j+1})$ and "update" the process to P at each τ_j: $P^n_{\tau_j} = P_{\tau_j}$ for all j. The key step then is to choose wisely the grid steps h_n for each successive approximation; we let h_n be such that:

$$h_n \xrightarrow[n\to\infty]{} 0 \text{ and } \frac{\kappa_n}{1 - e^{-rh_n}} \xrightarrow[n\to\infty]{} 0$$

it is fairly transparent that there are many ways to construct such an h_n sequence – we just need to choose it so that it converges to 0 "slowly enough" (relative to κ_n). The first condition naturally guarantees that $\mu_n \to \mu$ (since $P^n \xrightarrow{a.s.} P$ hence in distribution as well). The purpose of the second condition should become clear when writing out the cost; indeed because $\mu_n \in \boldsymbol{\mathcal{B}}_{\mathrm{d}}$ for all n, we have:

$$\mathfrak{C}_n(\mu_n) = \int \big(\mathscr{C}_0(p) + f_n(p)\big) d\mu_n(p)$$

by construction of μ_n the discontinuities of any $p \in \mathrm{supp}(\mu_n)$ are at the fixed h_n-spaced grid points, so that:

$$f_n(p) = \sum_{j\in\mathbb{N}} \kappa_n e^{-rh_n j} = \frac{\kappa_n}{1 - e^{-rh_n}}$$

the convergence $f_n \to 0$ is obviously uniform on compact sets, so modulo using Prokhorov's theorem to reduce attention to a compact set, we eventually obtain:

$$\int_n f_n d\mu_n \to 0$$

hence:

$$\lim_{n\to\infty} \mathfrak{C}_n(\mu_n) = \lim_n \int \mathscr{C}_0(p) d\mu_n(p) = \mathfrak{C}_0(\mu)$$

where the last equality is by definition.

Hence we can conclude that $\mathfrak{C}_n$ epi-converges to $\mathfrak{C}_0$. □

Naturally, this does not guarantee that *every* solution of the limit problem can be found as the limit of solutions when $\kappa \downarrow 0$; this only ensures that every limit of solutions of the problem with a fixed cost goes to a solution of the problems with no fixed cost as κ vanishes. Nonetheless, even if and when there is a strict selection being made from considering only solutions of the $\kappa = 0$ problem that are limits of solutions as $\kappa \downarrow 0$, one might be tempted to argue that such a selection is very natural: it picks the solutions of the problem that are not knife-edge in the sense that introducing a small fixed cost would lead a small (measured in the distance over path measures) change in the solution.

Proposition D.2. *Let v_κ the value function corresponding to any $\kappa \geq 0$ and P^κ the optimal belief process. Then for any $\overline{\kappa} \geq 0$:*

(i) v_κ converges pointwise to $v_{\overline{\kappa}}$ as $\kappa \to \overline{\kappa}$.

(ii) if P^κ converges in distribution to $P^{\overline{\kappa}}$ as $\kappa \to \overline{\kappa}$, then $P^{\overline{\kappa}}$ is an optimal belief process in the problem with fixed cost $\overline{\kappa}$

D.5 Optimal policies with vanishing fixed costs

Proof of Lemma 1. Rewriting the $\kappa = 0$ problem in the virtual flow firm first relies on rewriting the objective for $\kappa > 0$. Consider an arbitrary belief process $P \in \mathcal{B}_\mathrm{d}$ with discontinues at times $(\tau_i)_{i\geq 0}$, and to make notations more transparent introduce a notation for the deterministic flow operator:

$$\zeta_s q := e^{-\lambda s} q + (1 - e^{-\lambda s})\pi \text{ for any } q \in [0,1]$$

So that for each $i \geq 1$, by definition:

$$P_{\tau_i^-} = \zeta_{\tau_i - \tau_{i-1}} P_{\tau_{i-1}} \text{ a.s.}$$

The realized payoff from P can be rewritten as (where all equalities are understood pathwise):

$$\begin{aligned}
&\int_0^\infty e^{-rt}u(P_t) - \sum_i e^{-r\tau_i}\Big(c(P_{\tau_i}) - c(P_{\tau_i^-}) + \kappa\Big) \\
&\quad = \int_0^\infty e^{-rt}u(P_t) + \sum_i e^{-r\tau_i}\Big(c(P_{\tau_i^-}) - c(P_{\tau_{i-1}})) + \kappa\Big) + c(p) \\
&\quad = \int_0^\infty e^{-rt}u(P_t) + \sum_i e^{-r\tau_{i-1}}\Big(e^{-r(\tau_i-\tau_{i-1})}c(P_{\tau_i^-}) - c(P_{\tau_{i-1}}))\Big) + c(p) - \sum_i e^{-r\tau_i}\kappa \\
&\quad = \int_0^\infty e^{-rt}u(P_t) + \sum_i e^{-r\tau_{i-1}}\Big(e^{-r(\tau_i-\tau_{i-1})}c(\zeta_{\tau_i-\tau_{i-1}}P_{\tau_{i-1}}) - c(P_{\tau_{i-1}}))\Big) + c(p) - \sum_i e^{-r\tau_i}\kappa \\
&\quad = \int_0^\infty e^{-rt}u(P_t) + \sum_i \int_{\tau_{i-1}}^{\tau_i} \frac{\partial}{\partial t}\Big(e^{-rt}c(P_t)\Big)dt + c(p) - \sum_i e^{-r\tau_i}\kappa \\
&\quad = \int_0^\infty e^{-rt}\underbrace{\Big(u(P_t) - rc(P_t) + \lambda(\pi - P_t)c'(P_t)\Big)}_{=f(P_t)}dt + c(p) - \sum_i e^{-r\tau_i}\kappa.
\end{aligned}$$

Hence:

$$v_\kappa(p) - c(p) = \sup_{P\in\boldsymbol{\mathcal{B}}_{\mathrm{d}}(p)} \mathbb{E}\Big[\int_0^\infty e^{-rt}f(P_t)dt - \sum_i e^{-r\tau_i}\kappa\Big].$$

Since the discounted fixed cost term epi-converges to zero by the same arguments as previously, and by density:

$$w_\kappa(p) = \sup_{P\in\boldsymbol{\mathcal{B}}_{\mathrm{d}}(p)} \mathbb{E}\Big[\int_0^\infty e^{-rt}f(P_t)dt\Big] = \sup_{P\in\boldsymbol{\mathcal{B}}(p)} \mathbb{E}\Big[\int_0^\infty e^{-rt}f(P_t)dt\Big]$$

Proof of Theorem 4. The proof of Theorem 4 is separated into the following two results, which are proven below.

Proposition 7. *The optimal net value function w_0 in the information acquisition problem with $\kappa = 0$ is concave and:*

(i) for every belief p such that w_0 is strictly concave in a $\pi-$neighborhood p, it is uniquely optimal to not acquire information at p;

(ii) for every belief p such that w_0 is locally affine at p, it is optimal to immediately acquire information at p;

(iii) for every belief p such that neither previous condition hold, it is optimal to acquire information so as to confirm *p until some exponentially distributed time, at which beliefs jump to some prescribed belief $q(p)$ in the direction of π.*

Proof. If w is strictly concave locally in the direction of π at p, then it must be that the optimal process has $P_0 = p$ a.s. since otherwise we would have $\mathbb{E}[w(P_0)] < w(p)$, as any feasible belief process must put positive probability in the direction of π. If this is true for any $\tilde{p}$ in some π-neighborhood of p, then the only possible belief process is one that follows the deterministic drift $dP_t = \lambda(\pi - P_t)dt$ in that π-neighborhood. If instead w is locally affine in some neighborhood $[q_0, q_1]$ of p, then clearly:

$$\frac{p - q_0}{q_1 - q_0} w(q_1) + \frac{q_1 - p}{q_1 - q_0} w(q_0) = w(p),$$

so it is optimal to immediately acquire information so as to jump to $\{q_0, q_1\}$. Now consider any p such that w is not locally affine at p but w is also not strictly concave in any π-neighborhood of p. Denote by $w'_\pi(p)$ the directional derivative of w at p in the direction of π (which exists by Alexandrov's theorem). Concavity implies that for all q in a π-neighborhood of p, $w(q) \leq w(p) + w'_\pi(p)(q - p)$ but since w is strictly concave in no π-neighborhood of p we must be able to find a q such that this holds with equality. Fix such a q and now consider the belief process which stays at p until a random time when it jumps to q, and that random time is given by:

$$T \sim \mathscr{E}\Big(\lambda \frac{\pi - p}{q - p}\Big)$$

Denote $\rho := \lambda \frac{\pi - p}{q - p}$. It is direct to verify that this is a feasible belief process, assuming any arbitrary consistent distribution following the jump to q. First decompose the expectation of P_t conditionally on the jump time, replace with explicit expression depending on whether or not the jump has occured at t, then rearrange and compute integrals to obtain:

$$\begin{aligned}
\mathbb{E}[P_t] &= \int_0^\infty \rho e^{-\rho s} \mathbb{E}[P_t | T = s] ds \\
&= \int_0^t \rho e^{-\rho s} \big(q + (1 - e^{-\lambda(t-s)})(\pi - q)\big) ds + \int_t^\infty \rho e^{-\rho s} p ds \\
&= p + \int_0^t \rho e^{-\rho s} \big(q - p + (1 - e^{-\lambda(t-s)})(\pi - q)\big) ds \\
&= p + \big(1 - e^{-\rho t}\big)\big(q - p\big) + \big(\pi - q\big)\Big(\big(1 - e^{-\rho t}\big) - \rho e^{\lambda t} \frac{1 - e^{(-\rho+\lambda)t}}{\rho - \lambda}\Big) \\
&= p + \big(1 - e^{-\rho t}\big)\big(\pi - p\big) - \big(\pi - q\big) \rho e^{\lambda t} \frac{1 - e^{(-\rho+\lambda)t}}{\rho - \lambda} \\
&= p + \big(1 - e^{-\rho t}\big)\big(\pi - p\big) - \big(\pi - q\big) \lambda \frac{\pi - p}{q - p} e^{\lambda t} \frac{1 - e^{(-\rho+\lambda)t}}{\lambda \frac{\pi - p}{q - p} - \lambda} \\
&= p + \big(1 - e^{-\rho t}\big)\big(\pi - p\big) - \big(\pi - p\big) e^{\lambda t} \big(1 - e^{(-\rho+\lambda)t}\big) \\
&= p + \big(1 - e^{-\lambda t}\big)\big(\pi - p\big).
\end{aligned}$$

This is sufficient to establish that P is feasible. It only remains to establish that P is optimal – again, under a local understanding, fixing any optimal continuation policy after the jump to

q. To do so, it suffices to establish that, by construction, the expected payoff at p attains $w(p)$. First write this payoff explicitly:

$$\begin{aligned}\mathbb{E}\left[\int_0^T e^{-rt} f(p)dt + e^{-rT} w(q)\right] &= \mathbb{E}\left[1-e^{-rT}\right]\frac{f(p)}{r} + \mathbb{E}\left[e^{-rT}\right] w(q)\\ &= \left(1-\frac{\rho}{r+\rho}\right)\frac{f(q)}{r} + \frac{\rho}{r+\rho} w(q).\end{aligned}$$

Now observe that by optimality since waiting is always feasible at any point it must be that $rw(p) \geq f(p) + \lambda(\pi - p)w'_\pi(p)$. Given the assumptions on p and the previous arguments, it must be optimal to wait arbitrary close to p in the opposite direction from π, hence this actually must hold with equality at p. Furthermore recall that q has been chosen so that $w(q) = w(p) + w'_\pi(p)(q-p)$. Replacing in the previous expression and rearranging gives that the payoff at p is equal to:

$$\begin{aligned}&\frac{r}{r+\rho}\left(w(p) - \frac{\lambda}{r}(\pi-p)w'_\pi(p)\right) + \frac{\rho}{r+\rho}\left(w(p) + w'_\pi(p)(q-p)\right)\\ &= w(p) + w'_\pi(p)\left(\frac{\rho}{r+\rho}(q-p) - \frac{\lambda}{r}\frac{r}{r+\rho}(\pi-p)\right)\\ &= w(p) + w'_\pi(p)\left(\frac{\lambda\frac{\pi-p}{q-p}}{r+\rho}(q-p) - \frac{\lambda}{r}\frac{r}{r+\rho}(\pi-p)\right)\\ &= w(p) + w'_\pi(p)\left(\frac{\lambda(\pi-p)}{r+\rho} - \frac{\lambda(\pi-p)}{r+\rho}\right)\\ &= w(p)\end{aligned}$$

Which concludes the proof. □

Proposition 8. *Let P^κ the optimal belief process for $\kappa > 0$ and let $\mathscr{I}_\kappa$ the corresponding information acquisition region. If P^κ converges in distribution to P, then P is a wait-or-confirm belief process with:*

- *Waiting beliefs region* $\liminf_{\kappa\downarrow 0}\mathscr{I}_\kappa$
- *Initial jump beliefs region* $\mathrm{int}_\pi\left(\liminf_{\kappa\downarrow 0}\mathscr{I}_\kappa^c\right)$
- *Confirmation beliefs region* $[0,1]\setminus\left(\mathrm{int}_\pi\left(\liminf_{\kappa\downarrow 0}\mathscr{I}_\kappa^c\right)\cup\liminf_{\kappa\downarrow 0}\mathscr{I}_\kappa\right)$

Proof. Consider the optimal belief process P^κ for $\kappa > 0$ and denote its information acquisition region $\mathscr{I}_\kappa$. Assume P^κ converges to P in distribution (i.e in $\boldsymbol{\mathcal{B}}$ equipped with its weak-* topology induced by the Skorohod metric) as κ goes to zero. First consider a belief $p \in \liminf_{\kappa\to 0}\mathscr{I}_\kappa$, i.e p is evenutally in all information acquisition regions for κ small enough. Because convergence in distribution implies convergence in distribution at all continuity points and P is càdlàg, hence in particular continuous at the initial time, this must imply that P involves immediate information acquisition at p. (In other words, $P_0 = p$ would involve a contradiction, so the initial distribution of P_0 must involve immediate information acquisition.)

Now consider instead a belief $p \in \text{int}_\pi \liminf_{\kappa\to 0} \mathscr{I}_\kappa^c$ i.e there exist a π-neighborhood $b_\pi(p,\varepsilon)$ of p such that all points in $b_\pi(p,\varepsilon)$ are eventually not in all information acquisition regions for κ small enough. Again using convergence in distribution at all continuity points, it must be that the limit process involves no information acquisition arbitrarily close to p in the direction of π, which in light of the arguments in the proof of Proposition 7 establishes that waiting must be uniquely optimal in a π-neighborhood of p – hence by optimality of the limit P must involve waiting at p.

Lastly, consider p is neither of the previous sets. Without loss assume $p < \pi$ (the other case is symmetrical). From the previous points, this must mean that p is arbitrarily close in the direction of π to a point where immediate information acquisition is optimal in the limit problem, and arbitrarily close in the opposition direction from a point where waiting is uniquely optimal in the limit problem. By definition since $p \notin \liminf_{\kappa\to 0} \mathscr{I}_\kappa$ this means we can find a sequence κ_n converging to zero such that for all n, $p \in \mathscr{I}_{\kappa_n}^c$ i.e no information acquisition occurs at p under κ_n. Let z_n the closest point to p in the direction of π which is in $\mathscr{I}_{\kappa_n}$. Again by definition, it must be that z_n gets arbitrarily close to p as n goes to infinity otherwise this would contradict $p \notin \text{int}_\pi \liminf_{\kappa\to 0} \mathscr{I}_\kappa^c$, so $z_n \to p$. Denote $\{q_n^0, q_n^1\}$ the support of the optimal experiment at z_n for any n. By construction $p \in [q_n^0, q_n^1]$ for all n. Up to a subsequence, denote q_∞^0, q_∞^1 the limits of q_n^0, q_n^1 respectively; clearly $p \in [q_\infty^0, q_\infty^1]$. First consider the possibility that $q_\infty^0 < q_\infty^1$ and $p \in (q_\infty^0, q_\infty^1)$. Then observe that for all n, for any $q \in [q_\infty^0, q_\infty^1]$:

$$\text{Cav}[w_{\kappa_n}](q) = \frac{q - q_n^0}{q_n^1 - q_n^0} w_{\kappa_n}(q_n^1) + \frac{q_n^1 - q}{q_n^1 - q_n^0} w_{\kappa_n}(q_n^0)$$

Continuity of $w_\kappa(q)$ in q and pointwise convergence in κ, along with the assumption that $q_\infty^0 < q_\infty^1$ implies that w_0 is locally affine at p, which contradicts the fact that p is arbitrarily close in the opposition direction from π to a point where waiting is uniquely optimal in the limit problem. Now consider instead the possibilty that $p = q_\infty^1$: again this implies a contradiction because by assumption on p, w_0 must be strictly concave in some π-neighborhood of p. Hence the only remaining possibility is $p = q_\infty^0$ and $q_\infty^0 < q_\infty^1$.

Having established that both q_0^n and z_n converge to p, and that $p < q_\infty^1$, it remains to establish that the distribution of the belief process at p converges to the "confirmation" process which stays at p until an exponentially distributed jump time to q_∞^1. To do so, it suffices to consider the distribution of the time T_n that it takes for the process to reach q_n^1 from p. This can be expressed as:

$$T_n = \tau(p, z_n) + \tau(q_n^0, z_n) X_n \text{ where } X_n \sim \mathcal{G}\Big(\frac{z_n - q_n^0}{q_n^1 - q_0^n}\Big)$$

Where $\mathcal{G}$ denotes the geometric distribution as before. From the previous arguments, the first term $\tau(p, z_n)$ goes to zero as n converges to infinity, hence it is enough to prove that:

$$\tau(q_n^0, z_n) \times X_n \xrightarrow[n\to\infty]{d} \mathcal{E}\Big(\lambda \frac{\pi - p}{q_\infty^1 - p}\Big).$$

Which is established by a direct but cumbersome computation, working directly with the CDFs of X_n. □

Proof of Theorem 5. The proof relies first on establishing the upper bound for w_0, then on exhibiting a policy which achieves it in some region around π. The upper bound is straightforwardly derived from:

$$\begin{aligned}
w_0(p) &= \sup_{P \in \boldsymbol{\mathcal{B}}(p)} \mathbb{E}\Big[\int_0^\infty e^{-rt} f(P_t) dt\Big] && \because \text{ definition of } w_0 \\
&\leq \sup_{P \in \boldsymbol{\mathcal{B}}(p)} \mathbb{E}\Big[\int_0^\infty e^{-rt} \text{Cav}[f](P_t) dt\Big] && \because \text{ definition of Cav} \\
&\leq \sup_{P \in \boldsymbol{\mathcal{B}}(p)} \int_0^\infty e^{-rt} \text{Cav}[f]\big(\mathbb{E}[P_t]\big) dt && \because \text{ Jensen (pointwise and pathwise)} \\
&= \int_0^\infty e^{-rt} \text{Cav}[f]\big(p_t\big) dt && \because \text{ definition of } \boldsymbol{\mathcal{B}}(p).
\end{aligned}$$

Now, consider first the case where $\text{Cav}[f] = f$ in a neighborhood of π. In that case, not acquiring information at any p in this neighborhood clearly achieves the upper bound, uniquely so if f is locally strictly concave (by another application of Jensen's inequality). This must mean that no information is acquired in the long run under any optimal belief process in the $\kappa = 0$ problem. In the case where $\text{Cav}[f](\pi) = \pi$ but we do not have $\text{Cav}[f] = f$ in a neighborhood around π, some straightforward but tedious casework shows that either (i) jumping to π and then not acquiring information or (ii) eventually stopping information acquisition must be optimal in a neighborhood of p. Essentially, case (i) happens if the belief processes an appropriate neighborhood around π from a side where $\text{Cav}[f]$ is linear and (ii) when it approaches from a side where $\text{Cav}[f]$ is strictly concave. In either case, uniqueness cannot be guaranteed because of knife-edge indifferences if f is locally affine.

The main part of the proof consists of establishing optimality of the "confirmatory" policy when $\text{Cav}[f](\pi) > \pi$. In that case, denote (q_0, q_1) an interval such that $\pi \in (q_0, q_1)$, $\text{Cav}[f] > f$ in

(q_0, q_1) and $\text{Cav}[f] = f$ at q_0 and q_1. Fix any initial belief $p \in [q_0, q_1]$ and consider the belief process which:

- Immediately jumps to $\{q_0, q_1\}$ if $p \in (q_0, q_1)$
- From q_0 jumps to q_1 at rate $\rho_0 := \lambda \frac{\pi - q_0}{q_1 - q_0}$
- From q_1 jumps to q_0 at rate $\rho_1 := \lambda \frac{q_1 - \pi}{q_1 - q_0}$.

This means that the belief process effectively follows a continuous Markov chain between q_0 and q_1 with rate matrix:

$$\Psi := \begin{pmatrix} -\rho_0 & \rho_0 \\ \rho_1 & -\rho_1 \end{pmatrix},$$

and initial distribution given by the unique Bayes-plausible experiment in $\mathscr{B}(p)$ supported over $\{q_0, q_1\}$. Denote Q_t the resulting belief process and $M(t)$ the matrix of conditional probabilities:

$$M(t) := \begin{pmatrix} \mathbb{P}(Q_t = 0 | Q_0 = q_0) & \mathbb{P}(Q_t = 1 | Q_0 = q_0) \\ \mathbb{P}(Q_t = 0 | Q_0 = q_1) & \mathbb{P}(Q_t = 1 | Q_0 = q_1) \end{pmatrix},$$

Which solves the Kolmogorov equation $M'(t) = M(t)\Psi$ i.e $M(t) = e^{t\Psi}$ which in this case simplifies to an explicit expression:

$$M(t) := \begin{pmatrix} \frac{q_1 - \pi}{q_1 - q_0} + \frac{\pi - q_0}{q_1 - q_0} e^{-\lambda t} & \frac{\pi - q_0}{q_1 - q_0} - \frac{\pi - q_0}{q_1 - q_0} e^{-\lambda t} \\ \frac{q_1 - \pi}{q_1 - q_0} - \frac{q_1 - \pi}{q_1 - q_0} e^{-\lambda t} & \frac{\pi - q_0}{q_1 - q_0} + \frac{q_1 - \pi}{q_1 - q_0} e^{-\lambda t} \end{pmatrix}.$$

This, in particular, allows to verify that this belief process is feasible since $\mathbb{E}[Q_0] = p$ by construction and we can directly compute from the explicit expression of $M(t)$, skipping algebraic simplifications:

$$\mathbb{E}[Q_t | Q_0 = q_0] = \pi - (\pi - q_0)e^{-\lambda t} \text{ and } \mathbb{E}[Q_t | Q_0 = q_1] = \pi + (q_1 - \pi)e^{-\lambda t}.$$

To compute the induced expected value observe that for anyt t, since $Q_t \in \{q_0, q_1\}$ a.s. then f and $\text{Cav}[f]$ coincide over the support of Q_t. Hence:

$$\mathbb{E}[f(Q_t)] = \mathbb{E}\big[\text{Cav}[f](Q_t)\big].$$

Furthermore since $\text{Cav}[f]$ is affine over $[q_0, q_1]$:

$$\mathbb{E}\big[\text{Cav}[f](Q_t)\big] = \text{Cav}[f]\big(\mathbb{E}[Q_t]\big) = \text{Cav}[f](p_t),$$

where the last equality is just from the compensated martingale constraint. This immediately means that:

$$\mathbb{E}\left[\int_0^{\infty} e^{-rt} f(Q_t)dt\right] = \int_0^{\infty} e^{-rt}\,\text{Cav}[f](p_t)dt.$$

Which proves the desired result.