Engagement Maximization

Benjamin Hébert (Stanford and NBER) and Weijie Zhong (Stanford)*

March 25, 2026

Abstract

We investigate the management of information provision to maximize user engagement. A principal sequentially reveals signals to an agent who has a limited amount of information processing capacity and can choose to exit at any time. We identify a "dilution" strategy—sending rare but highly informative signals—that maximizes user engagement. The platform's engagement metric shapes the direction and magnitude of biases in provided information relative to a user-optimal benchmark. Even without intertemporal commitment, the platform replicates full-commitment revenue by inducing the user's belief to remain "as uncertain as" the prior until the rare, decisive signal arrives and induces stopping. We apply our results to two contexts: an ad-supported internet media platform and a teacher attempting to engage test-motivated students.

Key Words: Information Acquisition, Recommendation Algorithms, Polarization, Rational Inattention

JEL Codes: D83, D86

*The authors would like to thank Peter DeMarzo, Sebastian Di Tella, Darrell Duffie, Emir Kamenica, and Michael Woodford for helpful comments, and Ian Ball and Arjada Bardhi for excellent discussions. We would like to thank the editor (Juuso Toikka), several anonymous referees, and refine.ink for feedback that improved the paper. All remaining errors are our own. bhebert@stanford.edu and weijie.zhong@stanford.edu

1 Introduction

Free-to-use online platforms such as Facebook, Instagram, YouTube, and Pinterest are used by billions of people worldwide and earn substantial profits by displaying advertisements to their users. Their business models are powered by personalized recommendation algorithms that seek to maximize the "engagement" of each user by selectively displaying content (Lada et al. [2021], Sequoia [2018]). Large quantities of computing and other resources have been invested by these firms to develop algorithms that can predict billions of users' preferences in real time.

A key challenge for these algorithms is to manage users' incentives. Content is presented sequentially in "news feeds," "recommendations," or "timelines," and users can choose freely *what* to pay attention to and *when* to stop using the platform based on the entire history of content presented thus far. From the platform's perspective, providing the most useful content immediately is suboptimal, as a user might become satisfied and choose to stop using the platform, limiting the quantity of advertisements the platform can display to the user. On the other hand, if the user anticipates that the platform will never provide useful content, they will never begin to use the platform in the first place.

A second, subtler challenge is to manage the algorithm's own "incentives". As is suggested by the previous paragraph, to maximize users' *ex ante* engagement, platforms often design these algorithms with an initial commitment of providing reliable and wide-ranging content. Subsequent training of the algorithms based on user interactions, however, can induce algorithms to behave in a dynamically inconsistent manner. For instance, at a later time, the algorithm may "learn" to steer users into niche "rabbit holes" (WSJ [2021]) because doing so delivers short-term engagement gains. When the platform lacks long-term commitment power, users might question the credibility of the platform's stated design goals. It is therefore unclear ex ante whether concerns about dynamic inconsistency meaningfully constrain algorithm design.

Engagement maximization is relevant in many other contexts. We pair the example of an internet platform engaging a user with a second example, that of a teacher who seeks to maximize the engagement of a student who cares only about passing a test.[1] Our theoretical framework applies to this setting without modification.

We study the optimal design of sequential information presentation from a principal-agent perspective. The principal (the platform/teacher) provides the agent (the user/student)

[1] In this example, the teacher chooses the information flow, but the test itself is exogenous. We are grateful to Emir Kamenica for suggesting this alternative setting as an application of our model.

with information. The agent values this information because it affects the payoff she obtains when she eventually stops. Our model is agnostic about whether information is valuable for instrumental reasons (because it helps the agent make a better decision in some decision problem) or because the agent has an intrinsic motive for information acquisition. The agent chooses when to stop engaging with the provided information.[2] The rate at which the agent can process information is limited, and the agent experiences an opportunity cost of time spent on the platform. The principal's goal is to maximize the attention the agent allocates to the platform (which under some conditions is equivalent in equilibrium to maximizing the stopping time), and the principal accomplishes this by choosing the nature of the information provided to the agent. The key modeling assumptions we impose are that the principal knows perfectly the agent's preferences and that the principal can flexibly manipulate the entire information flow.

Two modeling features allow us to derive the rich implications in this environment. First, our model allows the measures of (i) engagement, (ii) information processing capacity, and (iii) the value of information to differ. This allows our model to capture a rich set of possible incentive misalignments between the principal and the agent. Second, we analyze our model both with and without intertemporal commitment, which allows us to characterize the implications of a requirement for dynamic consistency on the dynamics of information provision.

Our main result is the characterization of an optimal strategy for the principal. First, we completely characterize the optimal "overall" information structure (i.e. the signal-state joint distribution that describes the beliefs the agent will hold when choosing to stop). It is characterized by the solution to an augmented static rational inattention (RI) problem: the information structure maximizes a linear combination of the instrumental value, the engagement measure, and the informativeness measure of information with endogenous weights. Second, we identify one optimal sequential information structure for the principal: the principal sends a "dilution" of the overall information structure (a compound Poisson process such that a signal arrives at a Poisson rate and upon arrival, the signal is distributed according to the overall information structure).

Having described a single optimal strategy, we then show that the following features must hold for *any* optimal strategy:

- **Engagement-driven biases**: We compare the optimal information structure with the

[2] We show that it is without loss of generality to assume the agent does not process the information selectively; that is, the agent attends to the information the principal provides in equilibrium.

benchmark solution when the agent chooses both the information and the stopping time. We show that engaging with the platform leads to the agent being biased towards "more extreme" beliefs relative to the agent-optimal benchmark. More specifically, we show that the direction and magnitude of such biases are determined by the principal's engagement measure.

Leading Case: All Engagement is Profitable. Suppose that the principal's engagement measure is identical to the informativeness measure. In the internet platform context, we interpret this case as one in which the agent is exposed to ads in proportion to the content she receives, with no measurement of whether she attends to the ads. In this case, the optimal information structure is narrow but deep: the principal only provides information about the dimensions of interest to the agent. This information is provided in a rare but chunky way, leading the agent to spend additional time and acquire additional information relative to the agent-optimal benchmark. This result echoes the observation that engagement-maximizing recommendation algorithms often lead users down "rabbit holes."

Leading Case: Decision-Irrelevant Engagement is Profitable. Now suppose that the principal's revenue comes exclusively from engagement with information that is irrelevant for decision. In the internet platform context, we interpret this case as one in which the agent is exposed to decision-irrelevant ads, and the principal benefits only when the agent engages with these ads (e.g. if ad revenue comes from the click of an ad). In this case, we show that the optimal information structure is broad but noisy: the principal provides excess information (e.g. sponsored content) that the agent views as pure noise, intermixed with just enough decision-relevant information to keep the agent engaged. This result echoes the observation that engagement-maximizing recommendation algorithms often intersperse useful content with "clickbait" and "stealth marketing" material.

- **Credibility-driven dynamics**: To study policies that are credible under limited intertemporal commitment, we introduce a notion of subgame perfection to our model. First, we show that commitment has no value at all: the principal can achieve the same payoff from the optimal commitment policy even under limited commitment. However, limited commitment has strong implications for the dynamics of the policy: an optimal policy corresponds to a subgame-perfect equilibrium if (and nearly only if) the induced belief process of the agent jumps between a small set of very special beliefs, namely beliefs at which the agent is "as uncertain as" the prior belief. As an implication, the optimal and credible signal process must be a generalized "dilution": belief stays constant until a signal arrives at a Poisson rate. The signal brings the posterior belief either to the stopping region or to another such special interim belief. Consequently, the "dilution" policy we constructed earlier is indeed credible, and is uniquely credible in certain symmetric settings.

The intuition for the result is simple: we show that the principal-optimal policy leaves the agent with no surplus at the prior. If the agent were to become "more certain than" the prior at an interim stage, the surplus from future information would be negative, and the agent would stop engaging with the principal. On the other hand, if the agent were to become "less certain than" the prior at an interim stage, her surplus from future information would be positive. At this point, the principal would be tempted to adjust the information process and extract more surplus at the agent's expense. Thus, the only possible interim beliefs are those that are exactly "as certain as" the prior belief. Our main result in Section 4 is a formalization of this idea, along with a precise definition of what it means to be more certain, less certain, or as certain as the prior.

Our two leading cases are relevant to both the internet platform and teacher-student contexts. In the latter, they map to the degree to which the teacher values the student learning about test-relevant vs. test-irrelevant information. A teacher who values student learning but not test performance will provide just enough test-specific information to engage a student who cares only about the test, while providing additional content that is not part of the test. In contrast, a teacher who values all learning, regardless of whether it is covered on a test, will optimally provide only test-relevant information while inducing students to learn more test-relevant information than they would choose to learn on their own.

1.1 Related Literature

Our paper contributes to several strands of literature on the dynamic provision of information. Viewing our principal as a media company, our model is related to work on models of media bias (see Gentzkow et al. [2015] for a survey). We share with Kleinberg et al. [2022] an interest in explaining why the users of internet platforms would engage heavily with those platforms while perceiving themselves as gaining little from doing so. We derive this outcome as a result of strategic behavior by rational agents with conflicting incentives; those authors emphasize the time-inconsistency of user preferences. We share with Acemoglu et al. [2021] an emphasis on explaining what kind of information is available on internet platforms; our analysis focuses on content selection algorithms, whereas their analysis focuses on information sharing between users.

Closely related to our work is the literature on dynamic Bayesian persuasion (e.g. Ely [2017], Renault et al. [2017], Ely and Szydlowski [2020], Orlov et al. [2020], Che et al. [2020]) that build on the static model of Kamenica and Gentzkow [2011]. Our setting differs from most of these papers in that our principal's payoff solely depends on the engagement of the agent and not on her ultimate choice of action.[3] Most of these papers focus on the "Bayesian persuasion" settings where the principal's payoffs directly depend on the agent's action and the maximization of "engagement" is a side effect. Related papers without capacity constraints include Knoepfle [2020], Koh and Sanguanmoo [2022], Koh et al. [2024], and Saeedi et al. [2024].[4] These papers share several common predictions with our model, including Poisson dilution signals and the minimization of ex ante user welfare, indicating the robustness of these features across models. However, the absence of an informational constraint in these papers also drives stark differences: these papers predict full revelation of the state upon signal arrival, implying no bias from the agent's preferred posterior beliefs. Our analysis emphasizes the beliefs the agent will arrive at, and how these differ between the principal's optimal strategy and an agent-optimal benchmark, a comparison that is not possible absent information processing constraints. We elaborate further on the connection between our model and these models in Section 5. Our baseline analysis studies a forward-looking agent and a principal with full intertemporal commitment. Section 4.2 then asks which optimal policies remain credible without commitment, in contrast to the limited-commitment settings in Orlov et al. [2020], Che et al. [2020] or the myopic-agent settings in Ely [2017], Renault et al. [2017]. Interestingly, unlike the findings in these papers, neither commitment power nor forward-lookingness of the agent is necessary for sustaining our equilibrium strategy (see Section 4.2).

Our model predicts gradual information revelation over time, which is a feature shared by many of the dynamic information design models (e.g. Ely et al. [2015], Hörner and Skrzypacz [2016], Che et al. [2020], Orlov et al. [2020]). However, unlike these papers,

[3]The leading case of our model in which all engagement is profitable is equivalent to one where the principal's only goal is to ensure the agent continues to pay attention, as in Kawamura and Le Quement [2019]. However, our principal presents unbiased information, and hence is not engaged in cheap talk (Crawford and Sobel [1982], Cheng and Hsiaw [2022]).

[4]Knoepfle [2020]'s main focus is on the competition between multiple senders. The single sender case of Knoepfle [2020] is the closest to ours. Koh and Sanguanmoo [2022] and Koh et al. [2024] focus on more general implementability characterization. Saeedi et al. [2024] focuses on the case with asymmetric prior.

the gradual nature of belief evolution in our model arises from the agent's information processing constraint, as opposed to a desire to maximize suspense or address problems of limited commitment. In particular, Che et al. [2020] predicts that the agent's belief involves Poisson jumps and a drift, while our optimal strategy admits Poisson jumps without a drift.

Formally, our approach is a principal-agent version of Hébert and Woodford [2023]. Those authors consider a model in which a single decision maker chooses both what information to acquire and when to stop and act, whereas in our model the principal chooses the information and the agent chooses when to stop and act. We compare our model to a benchmark in which the agent chooses both the information and when to stop and act; this benchmark is characterized by results found in Hébert and Woodford [2023]. We follow Hébert and Woodford [2023] in assuming that the principal can choose any stochastic process for the agent's beliefs, subject only to the martingale requirement (which is imposed by Bayesian updating) and the upper bound on the agent's attention. We model this upper bound using a "uniformly posterior-separable" information cost, in the terminology of the rational inattention literature (Caplin et al. [2022]).[5]

The rest of the paper is organized as follows. We begin in Section 2 by describing the basic environment of our model. Section 3 characterizes optimal policy in our baseline model, with an emphasis on the beliefs the agent will hold when stopping. Section 4 discusses the dynamics of beliefs and the role of commitment. Section 5 discusses an extension of our baseline model, and Section 6 concludes.

2 The Environment

2.1 The Agent's Problem

We study the problem of a rational, Bayesian agent receiving signals about an underlying state for the purpose of taking an action. We model information acquisition as a continuous time process, building on results in Hébert and Woodford [2023] and Zhong [2022].

Let X be a finite set of possible states of nature. The state of nature is determined ex ante, does not change over time, and is not known to the agent. Let $q_t \in \mathscr{P}(X)$ denote the agent's beliefs at time $t \in [0,\infty)$, where $\mathscr{P}(X)$ is the probability simplex defined on X. We will represent q_t as vector in $\mathbb{R}_+^{|X|}$ whose elements sum to one, each of which corresponds to the likelihood of a particular element of X, and use the notation $q_{t,x}$ to denote the likelihood

[5] Examples of such information costs include mutual information, as applied in Sims [2010], as well as other proposed alternatives (Hébert and Woodford [2021], Bloedel and Zhong [2020]).

under the agent's beliefs at time t over the true state being $x \in X$. For each state x, we use $e_x \in \mathscr{P}(X)$ to denote the degenerate belief that state x occurs with probability 1.

At each time t, the agent can either stop and choose an action from a finite set A, or continue to acquire information. Let τ denote the time at which the agent stops and makes a decision, with $\tau = 0$ corresponding to making a decision without acquiring any information. The agent receives utility $u_{a,x}$ if she takes action a and the true state of the world is x, and pays a flow cost of delay per unit time, $\kappa > 0$, until an action is taken.[6]

Let $\widehat{u}(q')$ be the payoff (not including the cost of delay) of taking an optimal action under beliefs $q' \in \mathscr{P}(X)$:

$$\widehat{u}(q') = \max_{a \in A} \sum_{x \in X} q'_x u_{a,x}.$$

In what follows, the convex function $\widehat{u} : \mathscr{P}(X) \to \mathbb{R}$ will summarize the value the agent places on information. This function, as opposed to the utility function $u_{a,x}$, can be viewed as the primitive "value of information when acting" in our model; nothing in our analysis will depend on the nature of the action space A or on the predicted joint distribution of states and actions. As a result, our analysis extends without modification to the case in which the agent values information for its own sake (i.e. for non-instrumental reasons).

The agent's beliefs, q_t, will evolve as a martingale. This property follows from "Bayes consistency." In a single-period model, Bayes consistency requires that the expectation of the posterior beliefs be equal to the prior beliefs. The continuous-time analog of this requirement is that the belief process is a martingale.

Formally, let Ω be the sample space. Let $q : \Omega \times \mathbb{R}_+ \to \mathscr{P}(X)$ be a canonical càdlàg stochastic process on $\mathscr{P}(X)$, let $\{\mathscr{F}_t\}$ be the natural filtration associated with this canonical process, and let $\mathscr{F} = \lim_{t \to \infty} \mathscr{F}_t$. Let $\mathscr{T}$ be the set of non-negative stopping times with respect to $\{\mathscr{F}_t\}$.

Given a probability measure P defined on $(\Omega, \mathscr{F})$, $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}, P)$ defines a probability space. The agent's problem, given this probability space, is to choose a stopping time to solve

$$V(P) = \sup_{\tau \in \mathscr{T}} \mathbb{E}^P[\widehat{u}(q_\tau) - \kappa\tau | \mathscr{F}_0].$$

[6]Note that we assume that the delay cost enters the agent's utility additively and the agent does not discount the utility from acting. Discounting, in our context, has the effect of changing the relative values of acting and never acting, and for this reason would complicate our analysis; see the more extensive discussions in Hébert and Woodford [2023] and Zhong [2022] on this point.

2.2 The Principal's Problem

The principal chooses the information the agent receives so as to maximize engagement. We begin with defining a measure for the agent's information processing using a continuous time version of what Caplin et al. [2022] call a "uniformly posterior-separable" cost function, as described in Hébert and Woodford [2023]. Uniformly posterior-separable cost functions are defined in terms of a "generalized entropy function," $H : \mathscr{P}(X) \to \mathbb{R}_+$. We assume that H is twice continuously-differentiable and strongly convex.[7] We assume the agent can process information at a rate no greater than $\chi > 0$. That is, the process $H(q_t) - \chi t$ must be an $\mathscr{F}_t$-supermartingale, requiring that for all t and $\delta > 0$,

$$\mathbb{E}^P[H(q_{t+\delta}) - \chi(t+\delta)|\mathscr{F}_t] \leq H(q_t) - \chi t, \tag{1}$$

where we adopt the convention that $q_t = q_\tau$ for $t > \tau$.[8]

The principal's goal is to maximize "engagement," which is not necessarily the same as maximizing the information acquired. Specifically, we assume that the principal earns profits in proportion to cumulative information acquisition as measured by the convex function $G : \mathscr{P}(X) \to \mathbb{R}_+$. The function G is not necessarily strictly or strongly convex, although $G = H$ will be one of our leading cases.

The principal's goal is to design the agent's belief process so as to maximize profits, taking into account the fact the agent will know the nature of the beliefs process and optimally choose when to stop paying attention. Let $\bar{q}_0 \in \mathscr{P}(X)$ be the agent's prior. The principal chooses his policies from the set $\mathscr{A}(\bar{q}_0)$, which is the set of probability measures on $(\Omega, \mathscr{F})$ such that q is martingale belief processes with $q_0 = \bar{q}_0$ and non-negative stopping times τ such that (1) is satisfied. Formally, the principal chooses the probability measure P, which is equivalent to choosing the law of the belief process q. In this sense, the principal can choose any càdlàg martingale belief process with $q_0 = \bar{q}_0$, subject to the constraint imposed by the agent's information processing capacity.[9]

Definition 1. The principal's problem given initial belief $\bar{q}_0 \in \mathscr{P}(X)$ is to maximize en-

[7] I.e. the Hessian matrix of H is strictly positive definite.

[8] This constraint is equivalent to the one studied in Hébert and Woodford [2023] within the class of uniformly posterior-separable costs, allowing us to invoke their results.

[9] An alternative interpretation is that the principal can send more information than the agent can process, in which case the agent chooses which information to attend to. Allowing the agent this choice cannot benefit the principal, and it is therefore without loss of generality to suppose the principal chooses a process that satisfies the agent's information processing constraint.

gagement,

$$J(\bar{q}_0) = \sup_{(P,\tau)\in\mathscr{A}(\bar{q}_0)} \mathbb{E}^P[G(q_\tau) - G(\bar{q}_0)|\mathscr{F}_0] \tag{P}$$

subject to the agent's stopping decision,

$$\mathbb{E}^P[\widehat{u}(q_\tau) - \kappa\tau|\mathscr{F}_0] \geq \sup_{\tau'\in\mathscr{T}} \mathbb{E}^P[\widehat{u}(q_{\tau'}) - \kappa\tau'|\mathscr{F}_0].$$

2.3 Discussion

The generalized entropy functions G and H and value of information $\hat{u}$ play conceptually distinct roles in our model. The function H governs the agent's information processing constraint. It determines which kinds of information are relatively easy or difficult for the agent to process. The function $\hat{u}$ describes the information the agent would like to receive, and the function G describes the information the principal would like the agent to receive.

To simplify the discussion that follows, we will assume that the state space has a product structure, $X = X_1 \times X_2$, and that the agent values only information about $x_1 \in X_1$ (i.e. $\hat{u}$ depends only on the marginal posterior over X_1). The agent benefits from learning about $x_2 \in X_2$ only if doing so makes it easier to learn about $x_1 \in X_1$. Whether or not this is the case is governed by the H function, and our general framework admits both possibilities.[10] We interpret x_1 as the useful content and x_2 as content that has no intrinsic value to the user. Throughout the paper, we assume that the user "passively" learns about both x_1 and x_2 whenever the principal supplies both types of content. This should be viewed as a tie-breaking convention in favor of the principal: when processing decision-irrelevant content lies within the agent's capacity, the agent is indifferent between ignoring it and processing it, and we select the equilibrium in which she processes it. In the internet platform context, this can be interpreted as reflecting "stealth marketing" or "clickbait" behavior that makes ads indistinguishable from relevant content.[11]

We focus our discussion on the extent to which G differs from H and $\hat{u}$. One leading case of interest is when $G = H$.[12] We interpret this case as reflecting a kind of indifference

[10]The specific property of the H function that governs whether or not learning about x_2 makes it cheaper to learn about x_1 is what Hébert and LaO [2023] call "R-monotonicity."

[11]A formal proof of this equilibrium-selection result appears in Proposition 8 of the working paper version of the paper [Hébert and Zhong, 2025].

[12]The case in which G is proportional to $\hat{u}$ is essentially identical to the $G = H$ case, and for this reason we will discuss only the latter.

on the part of the principal: he does not care what the agent learns about, only that she learn as much as possible. In equilibrium, the constraint (1) will bind; as a result, "learning as much as possible" is equivalent to "spending as much time learning as possible."[13] We will show in this case that the signals sent by the principal maximize the utility of a hypothetical agent with a lower cost of delay.

A second leading case is one in which G measures only information acquisition that is irrelevant to the agent's decision problem (i.e. it depends only on the marginal posterior of X_2). We interpret this case as capturing a multi-dimensional conflict between the principal and agent over both how much the agent should learn and what the agent should learn about. In this case, the signals sent by the principal are not equivalent to those that would be chosen by a hypothetical agent with a lower cost of delay; the principal also distorts the signals so that they are more informative about X_2. In the special case in which H is proportional to Shannon's entropy, we derive a sharp result: the principal provides the agent with information about $x_1 \in X_1$ that is identical to what the agent would optimally choose for herself, plus additional information about $x_2 \in X_2$ that the agent would never choose to receive.

There are also intermediate situations of interest. The information the principal would like the agent to acquire might be of some use to the agent without being neutral or aligned with what the agent would like to learn (i.e. G can depend in part of the posterior over X_1 without being a linear combination of $\hat{u}$ and H). Our two leading cases generate sharp predictions precisely because they are extreme.

3 Optimal Policy

We start by defining a relaxed version of the principal's problem. Any probability measure and stopping rule the principal can implement will induce a probability measure over beliefs the agent will hold when she chooses to stop (i.e. a law for q_τ). Define $\Pi(\bar{q}_0) = \{\pi \in \mathscr{P}(\mathscr{P}(X)) : \mathbb{E}^\pi[q] = \bar{q}_0\}$ as the set of probability measures over posterior beliefs that are consistent with the initial prior. Under any feasible policy (P, τ) chosen by the principal, the law of the stopped belief q_τ induced by this policy will be an element of $\Pi(\bar{q}_0)$, by the martingale property of beliefs.

The following lemma describes the date-zero participation constraint of the agent (i.e.,

[13]The authors would like to thank Emir Kamenica for suggesting this interpretation. Note that although time spent and information processed are equivalent when (1) binds, they are not equivalent more generally (e.g. if the principal sends no signals to waste time), which is why we describe the equivalence as "time spent learning."

incentive compatibility w.r.t. stopping at time 0) and an upper bound on the total engagement.

Lemma 1. *For all $(P,\tau) \in \mathscr{A}(\bar{q}_0)$ satisfying the agent's optimal stopping constraint in Definition 1, the law π of the stopped belief q_τ satisfies:*

1. $\mathbb{E}^\pi[\widehat{u}(q)] \geq \kappa \mathbb{E}^P[\tau|\mathscr{F}_0] + \widehat{u}(\bar{q}_0)$, *and*
2. $\mathbb{E}^\pi[H(q) - H(\bar{q}_0)] \leq \chi \mathbb{E}^P[\tau|\mathscr{F}_0]$.

Proof. See Appendix A.1. □

Lemma 1 presents a necessary condition for any admissible policy for the principal in the optimization problem (Definition 1). The first condition states that the agent's optimal stopping utility is weakly greater than the utility from stopping immediately. The second condition states that the expected cumulative information acquired by the agent is weakly less than $\chi\mathbb{E}[\tau|\mathscr{F}_0]$—the maximal attention permitted by the information constraint (1).

Combining these two constraints,

$$\mathbb{E}^\pi[\widehat{u}(q) - \widehat{u}(\bar{q}_0)] \geq \kappa \mathbb{E}^P[\tau|\mathscr{F}_0] \geq \frac{\kappa}{\chi}\mathbb{E}^\pi[H(q) - H(\bar{q}_0)].$$

Let us define the principal's relaxed optimization problem as a maximization over $\Pi(\bar{q}_0)$ incorporating only this combined constraint:

$$\begin{aligned} \bar{J}(\bar{q}_0) = \sup_{\pi \in \Pi(\bar{q}_0)} \ & \mathbb{E}^\pi[G(q) - G(\bar{q}_0)] \\ s.t.\ & \frac{\kappa}{\chi}\mathbb{E}^\pi[H(q) - H(\bar{q}_0)] \leq \mathbb{E}^\pi[\widehat{u}(q) - \widehat{u}(\bar{q}_0)]. \end{aligned} \tag{R}$$

Because this combined constraint must hold in the original principal's problem, we must have $\bar{J}(\bar{q}_0) \geq J(\bar{q}_0)$.

We assume that (1) it is possible for the principal to benefit from engagement and (2) it is prohibitively costly for the agent to learn enough information to achieve the principal's first best, to the point that the agent would prefer to learn nothing at all if confronted with only these two possibilities. We maintain Assumption 1 throughout unless stated otherwise.

Assumption 1. *The principal can benefit from engagement:* $\max_{\pi' \in \Pi(\bar{q}_0)} \mathbb{E}^{\pi'}[G(q)] > G(\bar{q}_0)$. *However, it is not possible for the principal to achieve the principal's first best: if* $\pi \in$

$\arg\max_{\pi' \in \Pi(\bar{q}_0)} \mathbb{E}^{\pi'}[G(q)]$, *then*

$$\widehat{u}(\bar{q}_0) - \frac{\kappa}{\chi} H(\bar{q}_0) > \mathbb{E}^{\pi}\left[\widehat{u}(q) - \frac{\kappa}{\chi} H(q)\right].$$

This rules out the possibility that the principal either has no desire for the agent to acquire information or the possibility of satiating this desire. It immediately implies that the constraint in (R) will bind.

Define the function $\mathscr{L}$ as the Lagrangian associated with the dual of (R),

$$\mathscr{L}(\pi, \lambda) = \mathbb{E}^{\pi}\left[\widehat{u}(q) - \frac{\kappa}{\chi} H(q) + \lambda G(q)\right].$$

The following proposition characterizes the solutions of (R).

Proposition 1. *A solution to* (R) *exists. Additionally,*

- *there exists a* $\lambda \geq 0$ *such that if* π^* *is a solution to* (R) *then*

$$\pi^* \in \arg\max_{\pi \in \Pi(\bar{q}_0)} \mathscr{L}(\pi, \lambda),$$

- *if* π^* *is a solution to* (R)*, then the constraint binds,*

$$\frac{\kappa}{\chi} \mathbb{E}^{\pi^*}[H(q) - H(\bar{q}_0)] = \mathbb{E}^{\pi^*}[\widehat{u}(q) - \widehat{u}(\bar{q}_0)],$$

- *if* π^* *is a solution to* (R)*, then starting* (R) *from any belief in the support of* π^**, it is not possible to induce engagement,* $\mathbb{E}^{\pi^*}[\bar{J}(q)] = 0$*, and*

- *at least one solution to* (R) *has finite support.*

Proof. See Appendix A.2. □

This proposition demonstrates that π^* is the solution to a static rational inattention problem, with a general UPS cost function (of the sort studied by Caplin et al. [2022]). Those authors show that a necessary condition for π^* is that it *concavifies* the function $\widehat{u} - \frac{\kappa}{\chi} H + \lambda G$.[14] Here, $\hat{u}$ captures the benefit of information to the agent, $\frac{\kappa}{\chi} H$ captures

[14] Related results appear in earlier working papers by those authors and in the Bayesian persuasion literature.

the costly delay required to acquire the information, and λG captures the benefit of the information acquisition to the principal, converted to the units of the agent's utility at the rate λ.

Engagement is infeasible ($\bar{J}(\bar{q}_0) = 0$) whenever the only π satisfying the constraint are ones with $\mathbb{E}^{\pi}[G(q) - G(\bar{q}_0)] = 0$, which is to say that it is not possible for the principal to induce information acquisition in the dimensions he cares about. If G is strictly convex (meaning that all information acquisition is at least somewhat beneficial to the principal, as in the $G = H$ case), this implies $\mathrm{Supp}(\pi) = \{\bar{q}_0\}$. When $\mathrm{Supp}(\pi) = \{\bar{q}_0\}$, we will say that π^* is degenerate, and otherwise say that π^* is non-degenerate.[15]

Let us take as given a solution π^* to this relaxed problem, and consider how it might be implemented in an incentive compatible way in the original principal's problem.

3.1 Implementation

Take any non-degenerate $\pi \in \Pi(\bar{q}_0)$, and define the stochastic process q_t as:

$$q_t = \bar{q}_0 + \mathbf{1}_{N_\alpha(t) \geq 1} \cdot (Q - \bar{q}_0), \tag{2}$$

where $Q \in \mathscr{P}(X)$ is a random variable distributed according to π and $N_\alpha(t)$ is an independent Poisson counting process with parameter α. Here, q_t is a compound Poisson process that jumps according to π at rate α. We call such process q_t an $\alpha-$*dilution* of π.[16]

Proposition 2. *For all non-degenerate $\pi \in \Pi(\bar{q}_0)$ that satisfy the constraint in (R), let $\alpha = \frac{\chi}{\mathbb{E}^{\pi}[H(q) - H(\bar{q}_0)]}$ and let q_t be the α-dilution of π. Let P be the law of q_t and let $\tau = \inf\{t > 0 : q_t \neq \bar{q}_0\}$. Then (P, τ) is feasible in (P) and implements utility level* $\mathbb{E}^{\pi}[G(q) - G(\bar{q}_0)]$.

Proof. See Appendix A.3. □

Note that Proposition 2 immediately implies that $J = \bar{J}$; hence, we will not distinguish between the two functions. Combining Lemma 1, Proposition 1, and Proposition 2, we obtain the main characterization of the optimal policy:

[15] When G is not strictly convex, meaning that there are some dimensions of potential information acquisition the principal does not value, it is possible to have $J(\bar{q}_0) = 0$ with a non-degenerate π^*. However, the information acquired in this case generates no surplus for either the principal or the agent.

[16] Pomatto et al. [2018] first introduce the notion of dilution. They define the dilution of an information structure π as "producing π with probability α and uninformative signal with probability $1 - \alpha$." (the same notion appeared in Bloedel and Zhong [2020]). Our notion of α-dilution is essentially the repetition of a dilution in continuous time.

Theorem 1. *For all $\bar{q}_0 \in \mathscr{P}(X)$, there exists a $\pi^* \in \Pi(\bar{q}_0)$ with finite support solving* (R)*. If Supp$(\pi^*) = \{\bar{q}_0\}$, any feasible policy is optimal. Otherwise, let $\alpha^* = \frac{\chi}{\mathbb{E}^{\pi^*}[H(q)-H(\bar{q}_0)]}$, and let (P^*, τ^*) be the law and jumping time of the α^*-dilution of π^*. Then, (P^*, τ^*) solves the principal's problem* (P)*.*

There are generally many optimal policies in the principal's problem. First, there may be multiple π^* that solve the relaxed principal's problem (although uniqueness in static rational inattention problems is guaranteed under certain additional assumptions). Second, there are many stochastic processes q_t (equivalently, laws P) and stopping times τ that induce the same law for q_τ; provided that this law is equal to π^* and that incentive compatibility and the bounds on information acquisition are satisfied, all such policies are optimal. However, it is not the case that anything goes, as we show in section 4.

An immediate implication of Theorem 1 is that the agent's participation constraint binds. That is,

$$\mathbb{E}^{P^*}[\widehat{u}(q_{\tau^*}) - \kappa\tau^* | \mathscr{F}_0] = \widehat{u}(\bar{q}_0).$$

Strikingly, the agent is no better off receiving information from an engagement-maximizing principal than if she could not receive any information at all. The principal extracts the full surplus generated by the ability to produce information and apply it in the agent's decision problem, despite the agent's ability to choose when to stop and act. The principal, in maximizing the engagement of the agent, minimizes the agent's welfare subject to a participation constraint. Note, however, that the principal's ability to extract all of the surplus depends critically on the possibility of jumps in the belief process. We provide an example in Proposition 7 in which beliefs are required to be continuous and the agent-optimal outcome is attained.

3.2 The agent-optimal benchmark and extreme beliefs

We will compare our results to a benchmark in which the agent and not the principal chooses the probability space and martingale belief process (subject to (1)). This benchmark is a special case of the more general models described in Hébert and Woodford [2023] and Zhong [2022].

Those authors show that the optimal policies of this benchmark dynamic model are also equivalent to the solution to a static rational inattention problem. That is, under the agent-optimal policies (P_A, τ_A), given the initial belief $\bar{q}_0 \in \mathscr{P}(X)$,

$$\begin{aligned}\mathbb{E}^{P_A}[\widehat{u}(q_{\tau_A}) - \kappa \tau_A | \mathscr{F}_0] &= V^B(\bar{q}_0) \\ &= \max_{\pi \in \Pi(\bar{q}_0)} \mathbb{E}^{\pi}[\widehat{u}(q) - \frac{\kappa}{\chi}(H(q) - H(\bar{q}_0))]. \end{aligned} \tag{B}$$

This solution can be implemented, as above, by a compound Poisson process, but also by a diffusion (and many other processes). The following characterization is straightforward.

Lemma 2. *If the agent would choose to acquire information in the agent-optimal benchmark, the principal will induce engagement: for all $q \in \mathscr{P}(X)$, if $V^B(q) > \widehat{u}(q)$ then $\bar{J}(q) > 0$.*

Moreover, if π_A is a solution to (B), then starting (B) from any belief in the support of π_A, it is weakly optimal for the agent to stop, $\mathbb{E}^{\pi_A}[V^B(q) - \hat{u}(q)] = 0$.

Proof. See Appendix A.4. □

This lemma highlights two points. First, if the agent would choose to acquire information in the agent-optimal benchmark, the principal can profit by providing the agent with information (although that information will not in general be the information the agent would have chosen to acquire). Second, at any belief the agent might hold when stopping in the agent-optimal benchmark, the agent must have no benefit from continuing to acquire information.

The static rational inattention problem that characterizes stopping beliefs in the agent-optimal benchmark is essentially identical to the one that characterizes optimal policy in our principal-agent framework, except that the UPS information cost is $\frac{\kappa}{\chi}H(q)$ in the benchmark problem and $\frac{\kappa}{\chi}H(q) - \lambda G(q)$ in the principal agent case. The similarity in structure between these two problems will allow us to highlight the implications of engagement maximization.

Let $Q^i(\bar{q}_0) \subseteq \mathscr{P}(X)$ be the union of the support of all optimal stopping beliefs in the benchmark ($i = a$) and principal-agent ($i = p$) models. Let Conv $Q^i(\bar{q}_0)$ denote the convex hull of $Q^i(\bar{q}_0)$. The following proposition demonstrates that the beliefs the agent will hold after engaging with the principal are more extreme than the beliefs the agent would choose to acquire in the benchmark model. This result follows from the observation that in both models, stopping beliefs are characterized by the solution to a static rational inattention problem, with a lower information cost in the principal-agent case than in the benchmark case. There are three possible exceptions to this result. One is the case in which the principal can provide no strict incentive via information ($V^B(\bar{q}_0) = \widehat{u}(\bar{q}_0)$), in which case $Q^p(\bar{q}_0) \subset Q^a(\bar{q}_0)$. The second is the case in which the agent-optimal beliefs lie on the boundary of the simplex, in which case there may not be "room" for beliefs to become more extreme. The third is the possibility that the marginal benefit to the principal of additional information acquisition jumps downward exactly at the agent-optimal posterior beliefs, which can be ruled out by assuming G is differentiable.

Proposition 3. *Assume that G is continuously differentiable, $V^B(\bar{q}_0) > \widehat{u}(\bar{q}_0)$, and that $Q^a(\bar{q}_0)$ is a subset of the relative interior of $\mathscr{P}(X)$. Then for all $q \in Q^p(\bar{q}_0), q \notin \text{Conv}\, Q^a(\bar{q}_0)$.*

Proof. See the appendix, section A.5. □

That is, the stopping beliefs in the principal-agent problem will lie outside the convex hull of the stopping beliefs of the agent-optimal benchmark. When G is strictly convex, it is also possible to show that the stopping beliefs in the agent-optimal benchmark are in the relative interior of the convex hull of the stopping beliefs in the principal-agent problem. However, when the principal profits only by providing the agent with irrelevant information, G is not strictly convex, and the agent-optimal benchmark stopping beliefs will not necessarily lie in the relative interior. Figure 4 below provides an example of this possibility.

Extreme beliefs are a natural consequence of engagement maximization. By forcing the agent to ultimately acquire more information than she would choose for herself (the extreme beliefs), the principal can simultaneously delay the agent's stopping decision while providing information the agent is willing to attend to. Interpreted in the context of an internet platform and a user, our extreme beliefs result implies that the platform should provide in-depth content on a narrower set of topics than would be preferred by the user. We can interpret this loosely as encouraging the user to go down "rabbit holes" that lead to extreme beliefs. In the teacher-student context, extreme beliefs have a more benign interpretation: the teacher induces test-motivated students to learn more than they would choose to learn on their own.

3.2.1 Leading cases

Proposition 3 establishes that engagement maximization necessarily leads to more extreme stopping beliefs relative to the agent optimal benchmark. In what follows, we provide a more detailed analysis of the direction towards which engagement maximization biases the beliefs in two salient settings. To help interpretation, we assume that the state space has a product structure, $X = X_1 \times X_2$, and that the agent's decision problem u depends only on $x_1 \in X_1$. We further suppose that x_1 and x_2 are independent under the prior $\bar{q}_0$, so that an agent who receives signals that depend only on x_1 will not update about x_2. Thirdly, we assume throughout this subsection that H is the negative of Shannon's entropy H^S.

All Engagement is Profitable. When the principal benefits from all forms of engagement by the agent (our interpretation of the $G = H$ case), it is never worthwhile for the principal to provide the agent with decision-irrelevant information. Instead, the principal induces the agent to acquire more decision-relevant information than the agent would choose to acquire on her own.

The following result shows that this is what happens in the principal-agent problem–the principal will send the agent information about x_1 only. This occurs even though the principal would benefit from the agent learning about x_2 as well. Intuitively, signals about x_1 and x_2 that benefit the principal equally take an equal amount of time for the agent to process (because $G = H$), but only the former has a utility benefit for the agent. A principal who was sending signals about x_2 could switch to sending signals about x_1, relax the agent's participation constraint,[17] and then profit by adjusting the signals so that they induced even more extreme beliefs about x_1 for the agent.

Proposition 4. *Suppose $G = H$. Then,* (R) *is solved by π that reveals x_1 only, i.e. for all $q \in Supp(\pi)$ the marginals of q and $\bar{q}_0$ on X_2 are identical.*

Proof. See Appendix A.6. □

Decision-Irrelevant Engagement is Profitable. Let us now consider a different leading case, in which the information that benefits the principal is irrelevant to the agent's decision. Maintaining the assumption of a product structure for the state space, an independent prior, and that H is proportional to Shannon's entropy, we now suppose that G depends only on x_2 (recall that $\widehat{u}$ depends only on x_1).

Proposition 5. *Suppose G depends on x_2 only. Then,* (R) *is solved by $\pi = \pi_1^* \otimes \pi_2$, where $\pi_1^* \in \mathscr{P}(\mathscr{P}(X_1))$ is optimal for the agent and $\pi_2 \in \mathscr{P}(\mathscr{P}(X_2))$.*

Proof. See Appendix A.7. □

When the principal cares only about providing decision-irrelevant information, the principal will provide a mix of decision-relevant and decision-irrelevant information. The

[17] This argument depends on the assumption that sending a signal conditioned only on x_1 is least costly way to communicate information about x_1; that H is proportional to Shannon's entropy ensures this.

decision-relevant information will be exactly the information that the agent would have chosen for herself, but the agent will be unable to avoid learning the decision-irrelevant information that is "mixed in" with the decision-relevant information.[18]

3.3 Illustrative Examples

Example 1. We illustrate our results thus far in a simple example. An agent faces a choice between two actions, $A = \{l, r\}$. The payoffs from the actions are uncertain and depend on the state of the world $x \in X = \{L, R\}$. The agent assigns equal prior probability to both states ($\bar{q}_0$ is uniform). The agent gets utility one when the chosen action matches the state ($u_{l,L} = u_{r,R} = 1$) and utility negative one otherwise ($u_{l,R} = u_{r,L} = -1$). The agent is impatient and pays a constant cost of delay of two utils per unit of time ($\kappa = 2$). We assume that the information processing constraint H is the negative of Shannon's entropy H^S, and that the engagement measurement function G is also the negative of Shannon's entropy. We assume $\chi = 1$.

By Theorem 1, the principal's optimal strategy maximizes

$$\mathbb{E}^{\pi}[\widehat{u}(q) - (2-\lambda)(H^S(0.5) - H^S(q))]$$

for some $\lambda \in (0, 2)$. Meanwhile, the agent's optimal strategy maximizes

$$\mathbb{E}^{\pi}[\widehat{u}(q) - 2(H^S(0.5) - H^S(q))].$$

Both problems are static rational inattention problems. An application of Matêjka et al. [2015] implies that the agent's optimal stopping belief has support $\{\frac{e}{e+1}, \frac{1}{e+1}\}$ and the principal's optimal stopping belief has support $\left\{\frac{e^{2/(2-\lambda)}}{e^{2/(2-\lambda)}+1}, \frac{1}{e^{2/(2-\lambda)}+1}\right\}$. Both probability measures of stopping beliefs can be implemented by dilutions (compound Poisson processes). Since $\lambda > 0$, it is easy to see that the principal induces more extreme posterior beliefs and a longer wait compared to the agent-optimal benchmark.

Next, we vary the agent's prior belief $\bar{q}$, and focus our analysis on the agent's posteriors. Figure 1 illustrates the optimal q_τ for every prior belief: When $\bar{q}_0 \leq \underline{q}$ or $\bar{q}_0 \geq \bar{q}$, the agent stops immediately no matter what information she will receive. When $\bar{q}_0 \in (\underline{q}, \bar{q})$, the dashed lines illustrate the support of the agent-optimal policy. The lines are flat, meaning that the agent's stopping beliefs do not change with the prior belief, consistent with the

[18] We conjecture that the property required for the result is the R-mononticity property discussed in Hébert and LaO [2023].

prediction of the standard RI models.[19] The solid curves illustrate the support of the unique optimal π^*. They are closer to the boundaries zero and one, indicating that the posterior beliefs become more extreme under the principal's optimal policy.

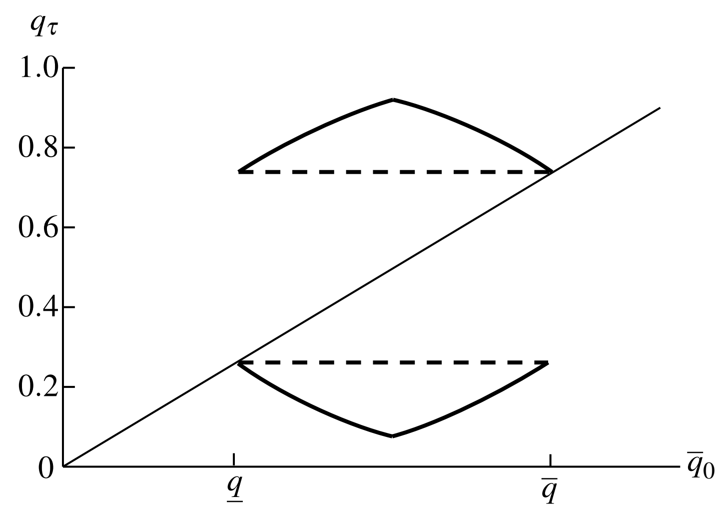

Figure 1: $\text{Supp}(q_\tau)$ as a correspondence of $\bar{q}_0$

Figure 2 illustrates the engagement level $\mathbb{E}^P[G(q_\tau) - G(\bar{q}_0)]$ for every prior belief. The dashed curves represent the agent-optimal policy and the solid curves represent the principal-optimal policy. The principal-optimal policy induces higher engagement level than the agent-optimal policy. However, the engagement level converges to zero when the prior belief goes to the boundary of the continuation region $(\underline{q}, \bar{q})$.

Example 2: Leading Cases. We next adapt example 1 to illustrate our two leading cases, in the context of engaging test-motivated students. Suppose the true state of the world T consists of both a test-relevant and test-irrelevant dimension, $T = T_1 \times T_2 = \{L, R\} \times \{0, 1\}$. The student and teacher know that the student will be asked a single question, $Q = \{Q_0\}$, whose responses are $A = \{l, r\}$, with $Q_0(L0) = Q_0(L1) = l$ and $Q_0(R0) = Q_0(R1) = r$. That is, the $\{L, R\}$ component of the true state is relevant, and the $\{0, 1\}$ component is irrelevant. The teacher chooses the information flow, not the test itself. The student (agent) again faces a binary choice $A = \{l, r\}$. The student's utility is one if she answers the question correctly and negative one otherwise, and therefore is identical (ignoring the test-irrelevant dimension) to that in example 1. The prior belief is uniform. We continue to assume the

[19]This property is known as the "locally invariant posteriors" (Caplin et al. [2022]).

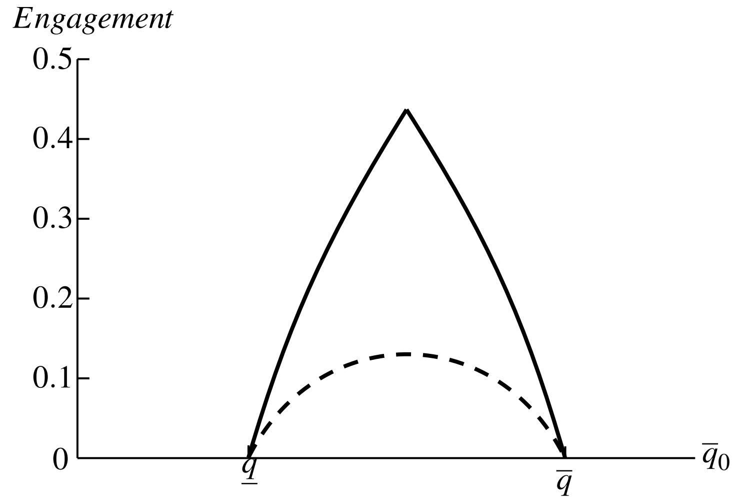

Figure 2: Engagement as a function of $\bar{q}_0$

agent's information processing capacity is defined by the negative of Shannon's entropy, and continue to use the parameters $\kappa = 2, \chi = 1$.

Leading Case: $H = G$. We first study the leading case of $H = G$, and illustrate Proposition 4 with a specific example. Like Example 1, both the principal's and the agent's optimization problems are equivalent to static rational inattention problems. By re-parameterizing the problems using the conditional probability of the chosen action, both problems can be rewritten as:

$$\sup_{P(a|t_1,t_2)} \sum_{a,t_1,t_2} \frac{1}{4} u(a,t_1) P(a|t_1,t_2) - C \cdot \Bigg(\sum_a \sum_{t_1,t_2} \frac{1}{4} P(a|t_1,t_2) \log(P(a|t_1,t_2)) - \sum_a \left(\sum_{t_1,t_2} \frac{1}{4} P(a|t_1,t_2) \right) \log \left(\sum_{t_1,t_2} \frac{1}{4} P(a|t_1,t_2) \right) \Bigg),$$

for $C = 2$ (the agent-optimal problem) and $C = 2 - \lambda$ (the principal-optimal problem). Note that replacing $P(a|t_1,t_2=0)$ and $P(a|t_1,t_2=1)$ with $\frac{P(a|t_1,0)+P(a|t_1,1)}{2}$ (denoted by $P(a|t_1)$) does not change the positive term while strictly reduces the negative term if $P(a|t_1,0)$ and $P(a|t_2,1)$ are not identical. Evidently, w.l.o.g., the optimization problem reduces to

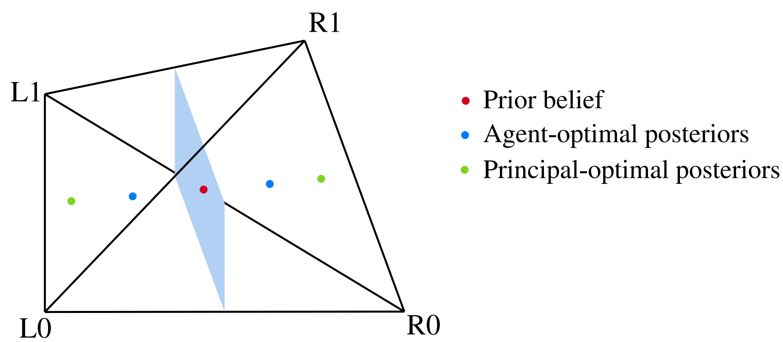

Figure 3: Leading case: $H = G$.

$$\sup_{P(a|t_1)} \sum_{a,t_1} \frac{1}{2} u(a,t_1) P(a|t_1) - C \cdot \left(\sum_{a,t_1} \frac{1}{2} P(a|t_1) \log(P(a|t_1)) \right.$$
$$\left. - \sum_{a,t_1} \frac{1}{2} P(a|t_1) \log\left(\sum_{t_1} \frac{1}{2} P(a|t_1) \right) \right),$$

which is equivalent to a static rational inattention problem with t_1 being the only unknown state.[20] As shown in Proposition 4, when there is a test-irrelevant state that is costly to learn about, no information about the state will be acquired either in the student-optimal benchmark or in the principal-agent problem, even though the student learning about state t_2 enters the teacher's payoff function.

Figure 3 illustrates the analysis above. The tetrahedron depicts the probability simplex that lies in $\mathbb{R}^3$. Each vertex of the tetrahedron is a degenerate belief (labeled by the state that occurs with probability one). The red dot is the uniform prior. The hyperplane represents the cutoff beliefs where the student is indifferent between actions l and r. To the right of the hyperplane, $q_R > q_L$ and r is optimal. Vice versa for the left of the hyperplane. The two blue dots are the student-optimal posterior beliefs. The two green dots are the solution to the principal-agent problem. Evidently, all the dots are on a line segment, which illustrates that the teacher provides only test-related information to the student. It is also evident that the teacher provides more information than the student would choose for herself.

Leading Case: Decision-Irrelevant G. Now, we turn to the second leading case, in which the teacher only cares about how much the student knows about the state $t_2 \in T_2$ (the test-

[20]This equivalence is special to mutual information and is the implication of a more general "compression invariance" property introduced by Caplin et al. [2022], Bloedel and Zhong [2020].

irrelevant dimension). We assume that the teacher's payoff is $2\mathbb{E}^{\pi}[|q_{t_2}-0.5|]$ if the measure of stopping belief is π, where q_{t_2} denotes the probability of $t_2 = 1$ under the student's stopping belief. Note that this payoff function is equivalent to the instrumental value of information in a hypothetical binary decision problem where the utility of matching (mismatching) the state y is 1 (-1). Theorem 1 shows that the teacher's optimal policy maximizes the following auxiliary problem (for some $\lambda > 0$):

$$\sup_{\pi} \mathbb{E}^{\pi}[2|q_{t_2}-0.5| + \lambda \widehat{u}(q) - 2\lambda(H(q)-H(\bar{q}_0))],$$

which is equivalent to a hypothetical static RI problem where the decision maker has four possible actions, whose utilities are

$u(a,t)$	$L0$	$R0$	$L1$	$R1$
a_1	$1+\lambda$	$1-\lambda$	$-1+\lambda$	$-1-\lambda$
a_2	$1-\lambda$	$1+\lambda$	$-1-\lambda$	$-1+\lambda$
a_3	$-1+\lambda$	$-1-\lambda$	$1+\lambda$	$1-\lambda$
a_4	$-1-\lambda$	$-1+\lambda$	$1-\lambda$	$1+\lambda$

.

Per Matêjka et al. [2015], the solution is logit with probabilities proportional to $\exp(u(a,t)/(2\lambda))$, because the Shannon information-cost coefficient in the auxiliary problem is 2λ. Along the test-relevant dimension, the relevant odds ratios are

$$\frac{P(a_1|L0)+P(a_3|L0)}{P(a_2|L0)+P(a_4|L0)} = e \qquad \text{and} \qquad \frac{P(a_1|R0)+P(a_3|R0)}{P(a_2|R0)+P(a_4|R0)} = \frac{1}{e},$$

and the same ratios hold in states $L1$ and $R1$. Therefore the induced conditional distribution over the test-relevant dimension is

$P(a\|t)$	$L0\&L1$	$R0\&R1$
$a_1\&a_3$	$\frac{e}{e+1}$	$\frac{1}{e+1}$
$a_2\&a_4$	$\frac{1}{e+1}$	$\frac{e}{e+1}$

,

which is identical to the student-optimal solution from Example 1. This is an illustration of the conclusion of Proposition 5. Along the test-irrelevant dimension, however,

$$\frac{P(a_1|L0)+P(a_2|L0)}{P(a_3|L0)+P(a_4|L0)} = e^{1/\lambda},$$

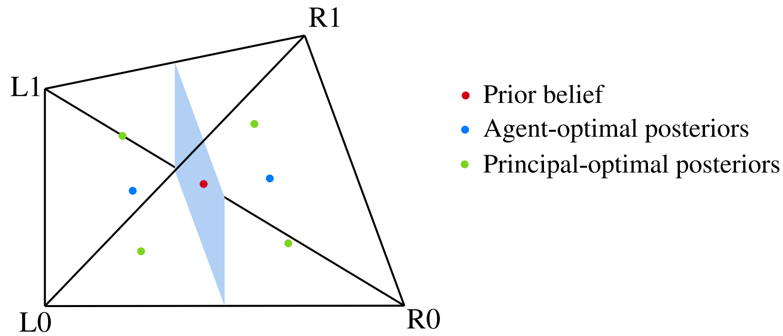

Figure 4: Leading Case: Decision-Irrelevant G.

and analogously in the other states, so the teacher induces additional learning about t_2. The unknown parameter λ can be pinned down by setting the student's IC condition binding. The analysis above suggests that when the teacher only cares about test-irrelevant knowledge, she gives the student exactly his preferred test-relevant information and extra knowledge such that the student gets barely enough welfare to be willing to participate.

Figure 4 illustrates the analysis above. Except the green dots, the figure is exactly the same as Figure 3. The green dots are the optimal posteriors of the principal-agent problem. For any pair of green dots on the same side of the hyperplane, they equal the blue dot in expectation. Their distance to the hyperplane is also the same as the blue dot, illustrating that the teacher provides the student his preferred test-relevant information together with extra test-irrelevant knowledge.

4 Characterization of dynamics

In Section 3, we showed that to solve for the optimal policy, it is sufficient to study a simple static rational inattention problem (R) and implement the dynamic policy by diluting the static solution. In this section, we seek to find necessary conditions that characterize the dynamics of the optimal policy. For the purpose of characterizing optimal polices, for this section, we focus on the non-degenerate case where the optimal policy is to acquire some information ($\bar{J}(\bar{q}_0) > 0$ in the context of Proposition 1). We show that two desiderata—interim incentive compatibility and credibility—jointly provide a tight characterization of the belief dynamics.

4.1 Implications of interim incentive compatibility

To begin, we define, given a measure $\pi \in \mathscr{P}(\mathscr{P}(X))$, the maximal value of information acquisition in a hypothetical restricted static rational inattention problem:

$$\overline{V}_R(q,\pi) = \max_{\pi' \in \mathscr{P}(\mathrm{Supp}(\pi)): E^{\pi'}[q']=q} \mathbb{E}^{\pi'}\left[\widehat{u}(q') - \widehat{u}(q) - \frac{\kappa}{\chi}(H(q') - H(q))\right].$$

This problem maximizes the agent's expected utility subject to the usual constraint of Bayes consistency and the additional constraint that the posteriors lie in the support of π. We use the notation $\overline{V}_R(q,\pi)$ to emphasize that the problem is restricted relative to the usual rational inattention problem. Note that the problem is feasible only for q that lie in in the convex hull of the support of π, which we denote by $\mathrm{Conv}(\mathrm{Supp}(\pi))$.

We interpret $\overline{V}_R(q,\pi)$ as describing the maximum possible value of information acquisition, in the sense that $\overline{V}_R(q,\pi)$ is the difference between the best possible utility the agent could achieve with and without information acquisition, under the restriction that the set of stopping beliefs lie in the support of π.[21] We define the set of "suspensive" beliefs given π as the set for which the value of information acquisition is weakly positive.

Definition 2. Given $\pi \in \mathscr{P}(\mathscr{P}(X))$, the belief $q \in \mathscr{P}(X)$ is *suspensive* if $q \in \mathrm{Conv}(\mathrm{Supp}(\pi)) \setminus \mathrm{Supp}(\pi)$ and satisfies $\overline{V}_R(q,\pi) \geq 0$; the belief $q \in \mathscr{P}(X)$ is *decisive* if $q \in \mathrm{Supp}(\pi)$.

Given a measure π that describes the stopping beliefs, it is intuitive that beliefs in $\mathrm{Supp}(\pi)$ are "decisive" because they directly lead to stopping. At the suspensive beliefs, the agent could benefit, relative to the prior consistent with π, from the signals that the principal is offering (which always result in some posterior in the support of π). Therefore, suspensive beliefs are those that the principal can possibly induce "suspense" before eventually leading the agent to stopping beliefs in $\mathrm{Supp}(\pi)$.

Suppose the optimal policy in (R), π^*, is unique. In this case, the stochastic process for beliefs q_t cannot leave the set of suspensive beliefs before $t = \tau$. If q_t left the convex hull of the support of π^*, then π^* would not describe the law of q_τ. If q_t was decisive or if $\overline{V}_R(q_t,\pi^*) < 0$, the agent would choose to stop. In the case of multiple optimal policies, this argument applies for some optimal π^*, which leads to the proposition below.

Proposition 6. *Given $\bar{q}_0 \in \mathscr{P}(X)$, let Π^* be the set of solutions to* (R). *Suppose (P,τ) solves* (P) *and $\sup_{\pi \in \Pi^*} \overline{V}_R(q,\pi)$ is a measurable function of q. Then, for every $t \in [0,\tau)$, q_t*

[21] The value of information acquisition is related to, but not the same as, the notion of uncertainty defined in Frankel and Kamenica [2019].

is almost surely suspensive:

$$\mathrm{Prob}^P\left(\sup_{\pi\in\Pi^*}\overline{V}_R(q_t,\pi)\geq 0\middle| t<\tau\right)=1.$$

Proof. See Appendix B.1. □

Solving (R) generically leads to the IC constraint binding, which suggests that the value of acquiring π^* is exactly zero at the prior for the agent. The set of suspensive beliefs thus defines beliefs that are "more uncertain" than the prior. Therefore, the interpretation of Proposition 6 is that two types of signals appear in any optimal policy almost surely. The first type are *decisive* signals that are so informative that the agent makes a decision immediately after receiving the signal. The second type are *suspensive* signals that causes the agent's beliefs to move to a posterior that is "more uncertain" than the prior.

Implication: The Necessity of Jumps in Beliefs. A direct implication of Proposition 6 is that the optimal policy cannot involve continuous sample paths for beliefs (we assume the initial prior lies in the continuation region). If sample paths were continuous, the belief process q_t would have to exit the set of suspensive beliefs, which would immediately induce the agent to stop. Consequently, beliefs must jump discontinuously from the set of suspensive beliefs (which lie in the continuation region of the benchmark model) to the set of decisive beliefs, which lie in the strict stopping region of the benchmark model (by Proposition 3).

To complement our discussion of the necessity of belief jumps, we analyze an extension of our model in which the belief process is required to be continuous. Formally, let $D^C\subset\mathbb{R}_+^{\mathscr{P}(X)}$ denote the set of *continuous* $\mathscr{P}(X)$-valued functions, i.e., all continuous paths. Let $\mathscr{A}^C(\bar{q}_0)$ denote the subset of $\mathscr{A}(\bar{q}_0)$ whose probability measures P are supported within D^C. Consider the optimization problem:

$$\begin{aligned} J^C(\bar{q}_0)=&\sup_{(P,\tau)\in\mathscr{A}^C(\bar{q}_0)}\mathbb{E}^P[G(q_\tau)-G(\bar{q}_0)|\mathscr{F}_0] \\ &\text{s.t.}\,\tau\in\arg\max_{\tau'\in\mathscr{T}}\mathbb{E}^P[\widehat{u}(q_{\tau'})-\kappa\tau'|\mathscr{F}_0]. \end{aligned}\tag{3}$$

Equation (3) is the same as Definition 1 except that the set of admissible strategies is restricted to have no jumps.

We begin our analysis of this restricted problem by observing that the agent-optimal benchmark is unchanged. This follows from results in Hébert and Woodford [2023], who show that the agent-optimal policy can be implemented by a pure diffusion process. Let $E^A = \{q \in \mathscr{P}(X) | V^B(q) > \widehat{u}(q)\}$ be the continuation region of the agent-optimal benchmark and $\bar{E}^A$ be the closure of E^A.

As we argued previously, the beliefs q_t must lie in the continuation region of the agent-optimal benchmark if $t < \tau$. Given that the $\mathscr{A}^C(\bar{q}_0)$ admits only continuous processes, it follows that the stopping beliefs must lie in $\bar{E}^A$.

Lemma 3. *If (P, τ) is admissible in Equation (3), then* $\mathrm{Supp}(q_\tau) \subset \bar{E}^A$.

Proof. See the appendix, section B.2. □

If in addition there are only two states ($|X| = 2$), the locally invariant posteriors property implies that the stopping beliefs will be identical to those the agent would choose in the agent-optimal problem; the only alternative the principal could choose would involve less information acquisition by the agent.[22]

Proposition 7. *If $|X| = 2$, then $J^C(\bar{q}_0)$ is achieved by an agent-optimal policy.*

Proof. See the appendix, section B.3. □

This proposition illustrates that the possibility of jumps in beliefs is required to guarantee the agent's participation constraint will bind. Without jumps in the belief process, the agent's ability to stop at any time allows the agent to extract rents that would otherwise accrue to the principal.

4.2 Implications of credibility

So far, our analysis assumes full intertemporal commitment on the part of the principal. However, in practice, commitment power might be limited, leading to a credibility concern. In this subsection, we tackle the issue of limited commitment by characterizing the necessary and sufficient conditions for a policy to be credible, i.e., immune to deviations in a subgame-perfection sense.

The definition of subgame perfection for a continuous-time game like ours is technically delicate. To avoid these complexities, we introduce the notion of a "strongly subgame-perfect" equilibrium. An equilibrium is strongly subgame-perfect if the principal is not

[22] When there are more than two states ($|X| > 2$), the principal's and agent's optimal policies need not coincide. Consider as an example the case of two actions ($|A| = 2$). The agent-optimal policy will in this case involve a diffusion on a line segment within the probability simplex (see Hébert and Woodford [2023]). The principal cannot induce the agent to follow this line segment beyond its endpoints, but can send the agent signals that cause the agent's beliefs to move orthogonal to this line segment.

willing to deviate, even if he can deviate to the full-commitment solution, and likewise the agent is willing to stop even if she can instead switch to the agent-optimal benchmark at that point.

In an environment in which commitment is valuable, such an equilibrium cannot exist. We will show in our environment that some (but not all) full-commitment solutions to the principal-agent problem are strongly subgame-perfect in this sense, which is to say that commitment has no value. Nevertheless, despite this lack of value, the principal's inability to commit will prevent some full-commitment solutions from being implemented as equilibria of the game without commitment.

Definition 3. For every $\bar{q}_0 \in \mathscr{P}(X)$, $(P, \tau) \in \mathscr{A}(\bar{q}_0)$ constitutes a *strongly subgame-perfect equilibrium* of the game if the agent's interim incentive compatibility constraints are satisfied and, in addition,

(i) the principal would not deviate to the commitment solution: $\forall t$, $\text{Prob}^P\Big(\mathbb{E}^P[G(q_\tau) - G(q_t) - J(q_t)|\mathscr{F}_t] \geq 0\Big) = 1$, and

(ii) the agent, when stopping, would not prefer to continue under the agent-optimal benchmark: $\text{Prob}^P\big(V^B(q_\tau) \leq \widehat{u}(q_\tau)\big) = 1$.

Let $\mathscr{V} : \mathscr{P}(X) \rightrightarrows \mathbb{R}^2_+$ denote the correspondence from prior to the set of strongly subgame-perfect equilibrium payoffs of the principal and agent.

This notion of strongly subgame-perfect equilibrium is "stronger" than subgame perfection in the sense that we overestimate deviation payoffs in conditions (i) and (ii): upon a deviation at belief, we award the principal with the full-commitment payoff and the agent with her agent-optimal payoff, respectively. If such an equilibrium exists, it is clearly subgame perfect in the usual sense as well.

In the analysis that follows, we will call a full-commitment optimal policy (P, τ) **credible** if (P, τ) is also a strongly subgame-perfect equilibrium in the game without commitment. The following theorem shows that the optimal dilution policy we construct in section 3 is indeed credible and hence constitutes a (strongly) subgame-perfect equilibrium. To avoid technicalities, what we will not discuss is whether there are other subgame-perfect equilibria that involve lower payoffs for the principal. Let $\mathscr{P}(X)^+ := \{q \in \mathscr{P}(X) | G(q) < \sum q_x G(e_x)\}$, namely, the set of beliefs that information has some value for the principal.

Theorem 2. *For all* $\bar{q}_0 \in \mathscr{P}(X)$*,* $J(\bar{q}_0)$ *can be achieved via a credible strategy, i.e.,* $(J(\bar{q}_0), 0) \in \mathscr{V}(\bar{q}_0)$*. Suppose* (P, τ) *solves* (P) *and* $\operatorname{supp}(q_\tau) \subset \mathscr{P}(X)^+$*.* (P, τ) *is credible if and only if the agent's interim surplus is almost surely* 0*:*

$$\textit{for all } t, \quad \operatorname{Prob}^P\left(\mathbb{E}^P[\widehat{u}(q_\tau) - \widehat{u}(q) - \kappa(\tau - t) | \mathscr{F}_t] = 0\right) = 1. \tag{4}$$

Proof. See Appendix B.4. □

Note that since $(J(\bar{q}_0), 0) \in \mathscr{V}(\bar{q}_0)$, $J(\bar{q}_0)$ is itself a strongly subgame-perfect equilibrium payoff. Condition (i) in Definition 3 must be satisfied for any weaker equilibrium notion that "selects SPNE in all subgames in favor of the principal." Therefore, (4) is a sufficient and necessary condition that guards the principal from deviations followed by a favorable selection of continuing equilibrium.

Using this condition, we can characterize a set of beliefs that cannot be reached in any credible solution to the principal-agent problem. To do this, we will define the set of "strictly suspensive" beliefs. In parallel to our definition of suspensive beliefs, we start by defining, given a measure $\pi \in \mathscr{P}(\mathscr{P}(X))$, the minimal value of information acquisition in a hypothetical restricted static rational inattention problem:

$$\underline{V}_R(q, \pi) = \min_{\pi' \in \mathscr{P}(\operatorname{Supp}(\pi)) : E^{\pi'}[q'] = q} \mathbb{E}^{\pi'}\left[\widehat{u}(q') - \widehat{u}(q) - \frac{\kappa}{\chi}(H(q') - H(q))\right].$$

We will say that a belief is strictly suspensive if this value is strictly positive.

Definition 4. Given $\pi \in \mathscr{P}(\mathscr{P}(X))$, the belief $q \in \mathscr{P}(X)$ is *strictly suspensive* if $q \in \operatorname{Conv}(\operatorname{Supp}(\pi)) \setminus \operatorname{Supp}(\pi)$ and satisfies $\underline{V}_R(q, \pi) > 0$.

Note the contrast between suspensive beliefs and strictly suspensive beliefs. A belief is suspensive given π if the *maximum* value of information is *weakly* positive, but is strictly suspensive only if the *minimum* value of information is *strictly* positive.

When $\pi' = \pi$ is the only measure on the support of π that satisfies Bayes consistency, there is no distinction between $\overline{V}_R$ and $\underline{V}_R$.[23] In this case, a belief is suspensive (resp. strictly suspensive) when this function is weakly (resp. strictly) positive.

The following proposition demonstrates that a solution to the principal-agent problem is credible only if interim beliefs are not strictly suspensive.

[23] This will be true generically when the number of actions is weakly lower than the number of states.

Proposition 8. *Given $\bar{q}_0 \in \mathscr{P}(X)$, let Π^* be the set of solutions to (R). Suppose (P,τ) solves (P),* $\mathrm{supp}(q_\tau) \subset \mathscr{P}(X)^+$*, and* $\inf_{\pi \in \Pi^*} \underline{V}_R(q,\pi)$ *is a measurable function of q. (P,τ) is credible only if, for every $t \in [0,\tau)$, q_t is almost never strictly suspensive:*

$$\mathrm{Prob}^P\left(\inf_{\pi \in \Pi^*} \underline{V}_R(q,\pi) \leq 0 \middle| t < \tau\right) = 1.$$

Proof. See Appendix B.5. □

Combining Propositions 6 and 8, in any solution that satisfies interim incentive compatibility and credibility, interim beliefs are suspensive but not strictly suspensive. They must be suspensive, as otherwise the agent would choose to stop, but cannot be strictly suspensive, as otherwise the principal would be tempted to deviate and extract more of the surplus.

Implication: Credible Beliefs Cannot Diffuse. When beliefs are suspensive but not strictly suspensive, the agent's value function is equal to $\widehat{u}(q_t)$ at all times. This implies that a pure diffusion process is infeasible almost everywhere in the continuation region. Because the agent recognizes that the principal will leave her with no surplus, only information that causes her to change her beliefs about the currently optimal action is valuable. If the belief q_t is such that one action is strictly optimal given those beliefs (which will be true generically), the principal must offer at least the possibility of jumping to a different region in which another action is optimal; otherwise, the agent will perceive no benefit from the information provided. In the internet platform context, we interpret this result as suggesting that the principal provides the agent with news articles containing extreme or sensational claims, which should cause the agent to either move her beliefs a lot or not at all, as opposed to providing more nuanced or qualified information. In the teacher-student context, to engage a test-motivated student, the teacher must provide signals that have some possibility of changing the student's belief about the correct answer on the test.

4.3 Illustrative Examples

Example 1 (continued). We illustrate Propositions 6 and 8 by continuing example 1. First, focus on the $\bar{q}_0 = \frac{1}{2}$ case. As is illustrated in Figure 1, the unique optimal π^* involves two posterior beliefs $\{q^1, q^2\}$, where $q^1 < \underline{q}$ and $q^2 > \bar{q}$. Per Proposition 2, the dilution of π^* implements an optimal policy. In this case, the dilution of π^* is the unique optimal policy. It is apparent from the symmetry of the problem that the value of information acquisition,

$\overline{V}_R(q,\pi^*)$, is maximized at $q = \frac{1}{2}$. As a result, the set of suspensive beliefs is a singleton, $\{\bar{q}_0\}$. By Proposition 6, beliefs will remain at $\bar{q}_0$ until they jump to either q^1 or q^2.

When $\bar{q}_0 < \frac{1}{2}$, the optimal policy is not unique. In Figure 5, we plot the simplest dilution policy, where belief stays at q_0 until it jumps to the two optimal posteriors q^1 and q^2. However, there can be other optimal policies. In this case, the set of suspensive beliefs is $[q_0, 1-q_0]$. Then, a full-commitment policy is optimal as long as beliefs remain in this interval $[q_0, 1-q_0]$ until eventually jumping to either q^1 or q^2. Specifically, Figure 6 plots one of such policies: when belief is at q_0 or $1-q_0$, it either jumps to the closer boundary, or drifts into the interior of $(q_0, 1-q_0)$. When belief is interior, it follows a diffusion. Here $\overline{V}_R(q,\pi^*) = \underline{V}_R(q,\pi^*)$ refers to the value of a restricted static redesign problem that keeps the support of π^* fixed; it need not coincide with the continuation payoff under the actual dynamic policy. Thus these objects can be positive on $(q_0, 1-q_0)$ even though the realized equilibrium surplus at the prior is zero.

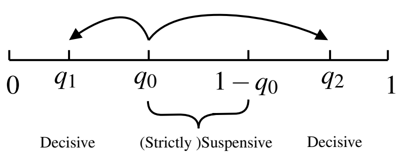

Figure 5: Dilution policy

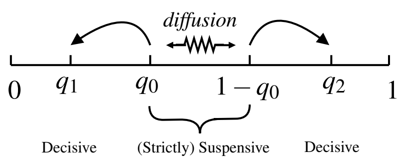

Figure 6: Jump/diffusion policy

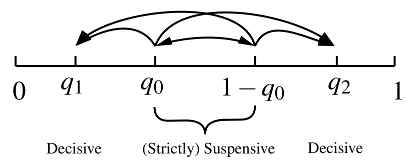

Figure 7: Credible policy

However, if the principal would like the policy to be credible, the interim belief cannot be strictly suspensive. In this example, $\bar{V}_R = \underline{V}_R > 0$ for all $q_t \in (q_0, 1-q_0)$, which is to say that these beliefs are strictly suspensive. It follows that a credible policy must satisfy $q_t \in \{q_0, 1-q_0\}$. Therefore, a credible belief process can only jump between q_0 and $1-q_0$ or jump to q^1 and q^2, as illustrated by Figure 7 (note that the dilution policy is also credible, but the jump/diffusion policy is not credible). Figure 8 illustrates the sample paths of one credible policy, in which beliefs jump between $\bar{q}_0$ and $1-\bar{q}_0$ before eventually jumping to q^1 or q^2.

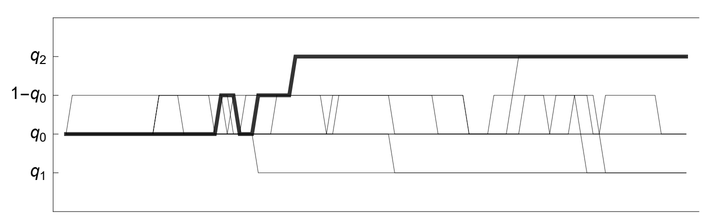

Figure 8: Sample paths of a credible policy

Example 2 (continued). We next illustrate the implications of Propositions 6 and 8 by continuing example 2, which discusses our two leading cases.

In the $H = G$ leading case, as discussed above, beliefs move on a line segment. As a result, even though the state space in example 2 is two-dimensional, our analysis for example 1 still applies. Because the prior is uniform, the only suspensive belief is the prior, and consequently the dilution policy is the unique optimal policy. This is illustrated below in Figure 9.

In the decision-irrelevant G leading case, in contrast, there are four possible optimal stopping beliefs under the unique π^* that solves (R). Because of the symmetry of the problem, we will still have $\bar{V}_R = \underline{V}_R$, and the set of beliefs that are suspensive but not strictly suspensive are characterized by the level set $\bar{V}_R(q, \pi^*) = 0$. These are illustrated by the curves shown in Figure 10.

Any credible, incentive-compatible policy must maintain interim beliefs on these curves. Because they are curves and not straight lines, this is not compatible with beliefs following a diffusion process. But it is compatible with a wide range of jump processes, which are the multi-dimensional analogs of the credible policies illustrated in Figures 7 and 8.

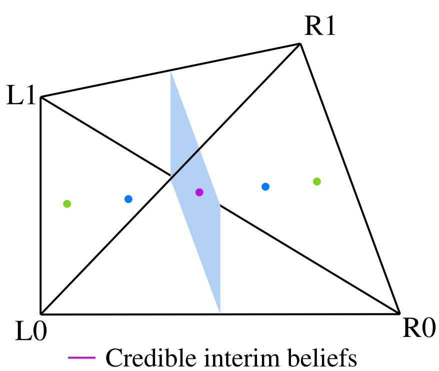

Figure 9: Credible interim beliefs when $H = G$

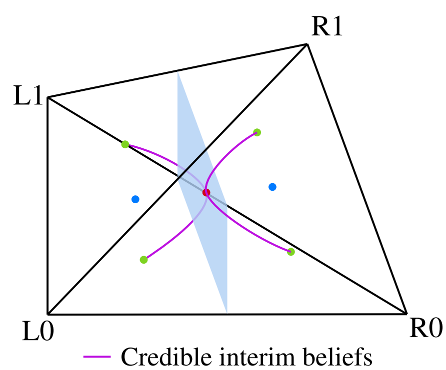

Figure 10: Credible interim beliefs with a decision-irrelevant G

5 Extension: Optimal Policy without Capacity Constraints

In this section, we consider a modified version of our model with a constant cost of delay in which the agent has an unlimited capacity to acquire information, and the principal's goal is to maximize the time spent by the agent on the platform. This modified model is the continuous-time analog of the discrete-time models studied by Knoepfle [2020] and Koh and Sanguanmoo [2022]. The purpose of this section is to illustrate the connection between these models and a special case of our more general framework. The agent-optimal policy in the modified model is for the agent to learn the optimal action with certainty immediately, as this avoids entirely the cost of delay.

The principal in this case chooses his policies from the set $\bar{\mathscr{A}}(\bar{q}_0)$, which is the set of probability measures on $(\Omega, \mathscr{F})$ such that q is martingale belief processes with $q_0 = \bar{q}_0$ and non-negative stopping times τ. Note that this set does not impose the constraint on the rate of information acquisition, (1), that was imposed in our main analysis. The principal solves

$$J(\bar{q}_0) = \sup_{(P,\tau)\in\bar{\mathscr{A}}(\bar{q}_0)} \mathbb{E}^P[\tau|\mathscr{F}_0]$$

subject to the same constraint with respect to the agent's stopping decision,

$$\tau \in \arg\max_{\tau'\in\mathscr{T}} \mathbb{E}^P[\widehat{u}(q_{\tau'}) - \int_0^{\tau'} \kappa dt|\mathscr{F}_0].$$

The first part of Lemma 1 remains applicable: if π is the law of q_τ, we must have $\mathbb{E}^\pi[\widehat{u}(q) - \widehat{u}(\bar{q}_0)] \geq \kappa\mathbb{E}^P[\tau|\mathscr{F}_0]$. Now observe that the expected utility under π is bounded above by the utility of fully learning the state. Let $\pi^{\max} \in \mathscr{P}(X)$ be the unique probability measure that places full support on the extreme points of $\mathscr{P}(X)$ (i.e. the e_x basis vectors), with $\pi^{max}(e_x) = \bar{q}_{0,x}$.

By the convexity of $\widehat{u}$, for all π such that $E^\pi[q] = \bar{q}_0$,

$$\mathbb{E}^{\pi^{max}}[\widehat{u}(q) - \widehat{u}(\bar{q}_0)] \geq \mathbb{E}^\pi[\widehat{u}(q) - \widehat{u}(\bar{q}_0)] \geq \kappa\mathbb{E}^P[\tau|\mathscr{F}_0] \geq \kappa J(\bar{q}_0).$$

It follows that $\kappa^{-1}\mathbb{E}^{\pi^{max}}[\widehat{u}(q) - \widehat{u}(\bar{q}_0)]$ is an upper bound on the utility achievable in this problem.

But now observe that the α-dilution of $\pi^{\max}$, as defined in (2), with intensity $\alpha = \kappa(\mathbb{E}^{\pi^{max}}[\widehat{u}(q) - \widehat{u}(\bar{q}_0)])^{-1}$, achieves this bound. Moreover, the policy is incentive-compatible: at each instant, the agent compares the utility benefit of the signal's arrival, $\alpha\mathbb{E}^{\pi^{max}}[\widehat{u}(q) - \widehat{u}(\bar{q}_0)]$, against the cost of delay, κ, and is willing to continue. It follows that this policy is optimal.

This policy is the optimal policy in a special case of our main model. Consider the case of our main model in which $\kappa = \chi$ and $G(q) = H(q) = \widehat{u}(q)$. In this special case, $\pi^{\max}$ is an optimal policy in our relaxed problem (R), because the constraint in that problem satisfied for any policy, and the process described in Proposition 2 is exactly the process above.

The intuition behind this equivalence is that in our main model, it is always optimal for the principal to exhaust the agent's information processing capacity. If exhausting this capacity necessarily involves satisfying the incentive compatibility constraint, then the principal is free to choose the process that simultaneously exhausts the agent's capacity and takes as long as possible to reach a decisive belief.

6 Conclusion

We have considered the problem of a principal who provides information to an agent so as to maximize the attention the agent pays to the principal's information (engagement). The agent values this information for instrumental purposes, is rational and Bayesian, and faces a constraint on the rate at which she can process information. Our main results are (i) that by maximizing engagement, the principal leaves the agent no better off than if she could not receive any information at all, and (ii) that the agent will end up holding extreme beliefs, relative to a benchmark in which the agent could choose the information for herself, and (iii) that optimal signals in the absence of commitment cannot either increase or decrease the value of future information prior to stopping (i.e. that they are suspensive but not strictly suspensive). Our results highlight the distortions created when information provision is designed to maximize engagement.

References

Daron Acemoglu, Asuman Ozdaglar, and James Siderius. Misinformation: Strategic sharing, homophily, and endogenous echo chambers. Technical report, National Bureau of Economic Research, 2021.

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

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

Yeon-Koo Che, Kyungmin Kim, and Konrad Mierendorff. Keeping the listener engaged: a dynamic model of bayesian persuasion. *arXiv preprint arXiv:2003.07338*, 2020.

Ing-Haw Cheng and Alice Hsiaw. Bayesian doublespeak. *Available at SSRN*, 2022.

Vincent P Crawford and Joel Sobel. Strategic information transmission. *Econometrica: Journal of the Econometric Society*, pages 1431–1451, 1982.

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

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

Jeffrey C Ely and Martin Szydlowski. Moving the goalposts. *Journal of Political Economy*, 128 (2):468–506, 2020.

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

Matthew Gentzkow, Jesse M Shapiro, and Daniel F Stone. Media bias in the marketplace: Theory. In *Handbook of media economics*, volume 1, pages 623–645. Elsevier, 2015.

Benjamin Hébert and Jennifer LaO. Information acquisition, efficiency, and nonfundamental volatility. *Journal of Political Economy*, 131(10):2666–2723, 2023.

Benjamin Hébert and Michael Woodford. Neighborhood-based information costs. *American Economic Review*, 111(10):3225–55, 2021.

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

Johannes Hörner and Andrzej Skrzypacz. Selling information. *Journal of Political Economy*, 124 (6):1515–1562, 2016.

Benjamin Hébert and Weijie Zhong. Engagement maximization, 2025. URL https://arxiv.org/abs/2207.00685v5.

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

Kohei Kawamura and Mark T Le Quement. News begets news: A model of endogenously repeated costly consultation. 2019.

Jon Kleinberg, Sendhil Mullainathan, and Manish Raghavan. The challenge of understanding what users want: Inconsistent preferences and engagement optimization. *arXiv preprint arXiv:2202.11776*, 2022.

Jan Knoepfle. Dynamic competition for attention. *Working paper*, 2020.

Andrew Koh and Sivakorn Sanguanmoo. Attention capture. *arXiv preprint arXiv:2209.05570*, 2022.

Andrew Koh, Sivakorn Sanguanmoo, and Weijie Zhong. Persuasion and optimal stopping. *arXiv preprint arXiv:2406.12278*, 2024.

Akos Lada, Meihong Wang, and Tak Yan. How machine learning powers facebook's news feed ranking algorithm, January 2021. URL https://engineering.fb.com/2021/01/26/ml-applications/news-feed-ranking/.

Filip Matêjka, Alisdair McKay, et al. Rational inattention to discrete choices: A new foundation for the multinomial logit model. *American Economic Review*, 105(1):272–98, 2015.

Dmitry Orlov, Andrzej Skrzypacz, and Pavel Zryumov. Persuading the principal to wait. *Journal of Political Economy*, 128(7):2542–2578, 2020.

Luciano Pomatto, Philipp Strack, and Omer Tamuz. The cost of information. *arXiv preprint arXiv:1812.04211*, 2018.

Jérôme Renault, Eilon Solan, and Nicolas Vieille. Optimal dynamic information provision. *Games and Economic Behavior*, 104:329–349, 2017.

Maryam Saeedi, Yikang Shen, and Ali Shourideh. Getting the agent to wait, 2024. URL https://arxiv.org/abs/2407.19127.

Sequoia. News feed ranking drives engagement, 2018. URL https://medium.com/sequoia-capital/engagement-part-iv-activity-feed-ranking-40d786b9d479.

Christopher A Sims. Rational inattention and monetary economics. *Handbook of Monetary Economics*, 3:155–181, 2010.

WSJ. How tiktok's algorithm figures you out, 2021.

Weijie Zhong. Information design possibility set. *arXiv preprint, arXiv: 1804.05752*, 2018.

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

A Omitted Proofs for Section 3

A.1 Proof of Lemma 1

Proof. Condition (1): let $\tau' \equiv 0$. Under this policy, the agent's utility from stopping is $\widehat{u}(\bar{q}_0)$. The optimality of τ and $\mathbb{E}^{\pi}[\hat{u}(q)] = \mathbb{E}^{P}[\hat{u}(q_\tau)]$ implies condition (1).

Condition (2): This follows from $\mathbb{E}^{\pi}[H(q)] = \mathbb{E}^{P}[H(q_\tau)]$, the optional stopping theorem, and the supermartingale property of $H(q_t) - \chi t$. The last two imply that $\mathbb{E}^{P}[H(q_\tau) - \chi\tau] \leq H(q_0)$. □

A.2 Proof of Proposition 1

Proof. Assumption 1 implies that the constraint in (R) must be binding, as otherwise, the optimal π^* will achieve the principal's first best, leaving the agent with negative surplus. Then, (R) is a convex optimization problem that satisfies the conditions in Theorem 4 of Zhong [2018], which immediately implies the existence of λ and of a finite support solution π^*.

Next, we prove $\mathbb{E}^{\pi^*}[\bar{J}(q)] = 0$. Suppose for the purpose of contradiction that $\mathbb{E}^{\pi^*}[J(q)] > 0$. For each $q \in \text{Supp}(\pi^*)$, let π^q be a probability measure that implements $\bar{J}(q)$. Define $\pi' \in \Pi(\bar{q}_0)$ by $d\pi'(q') = \mathbb{E}^{\pi^*(q)}[d\pi^q(q')]$, observing that

$$\mathbb{E}^{\pi'(q')}[q'] = \mathbb{E}^{\pi^*(q)}[\mathbb{E}^{\pi^q(q')}[q']] = \mathbb{E}^{\pi^*(q)}[q] = \bar{q}_0.$$

Then,

$$\begin{aligned}
\mathbb{E}^{\pi'(q')}[G(q') - G(\bar{q}_0)] =& \mathbb{E}^{\pi^*(q)}[\mathbb{E}^{\pi^q(q')}[G(q')] - G(q) + G(q) - G(\bar{q}_0)] \\
=& \mathbb{E}^{\pi^*(q)}[\mathbb{E}^{\pi^q(q')}[G(q') - G(q)]] + \bar{J}(\bar{q}_0) \\
=& \mathbb{E}^{\pi^*(q)}[\bar{J}(q)] + \bar{J}(\bar{q}_0) \\
>& \bar{J}(\bar{q}_0).
\end{aligned}$$

Meanwhile,

$$\begin{aligned}\mathbb{E}^{\pi'(q')}\left[\widehat{u}(q')-\frac{\kappa}{\chi}H(q')-\widehat{u}(\bar{q}_0)+\frac{\kappa}{\chi}H(\bar{q}_0)\right]=&\mathbb{E}^{\pi^*(q)}\left[\mathbb{E}^{\pi^q(q')}\left[\widehat{u}(q')-\frac{\kappa}{\chi}H(q')\right]-\widehat{u}(\bar{q}_0)+\frac{\kappa}{\chi}H(\bar{q}_0)\right]\\ =&\mathbb{E}^{\pi^*(q)}\left[\mathbb{E}^{\pi^q(q')}\left[\widehat{u}(q')-\frac{\kappa}{\chi}H(q')-\widehat{u}(q)+\frac{\kappa}{\chi}H(q)\right]\right]\\ &+\mathbb{E}^{\pi^*(q)}\left[\widehat{u}(q)-\frac{\kappa}{\chi}H(q)-\widehat{u}(\bar{q}_0)+\frac{\kappa}{\chi}H(\bar{q}_0)\right]\\ \geq&0.\end{aligned}$$

Therefore, π' is feasible in (R) and strictly improves on π^*, contradicting the optimality of π^*. It follows by the non-negativity of $J(q)$ that $\mathbb{E}^{\pi^*}[J(q)]=0$. □

A.3 Proof of Proposition 2

Proof. Under the α-dilution policy, if $t<\tau$,

$$\mathbb{E}^P[H(q_{t+\delta})-\chi(t+\delta)|\mathscr{F}_t,\tau>t]=(1-e^{-\alpha\delta})\mathbb{E}^\pi[H(q)-H(\bar{q}_0)]+H(\bar{q}_0)-\chi(t+\delta).$$

Using $\alpha=\chi(\mathbb{E}^\pi[H(q)-H(\bar{q}_0)])^{-1}$ and $x\geq 1-e^{-x}$,

$$\mathbb{E}^P[H(q_{t+\delta})-\chi(t+\delta)|\mathscr{F}_t,\tau>t]\leq H(\bar{q}_0)-\chi(t).$$

as required by (1). That constraint holds by assumption if $t\geq\tau$.

By $q_\tau\sim\pi$, the principal's utility is $\mathbb{E}^\pi[G(q)-G(\bar{q}_0)]$. What remains to be verified is the agent's optimality condition.

Take any admissible stopping time τ',

$$\begin{aligned}\mathbb{E}^P[\widehat{u}(q_{\tau'})-\kappa\tau'|\mathscr{F}_0]=&\mathrm{Prob}(\tau'<\tau|\mathscr{F}_0)\left(\widehat{u}(\bar{q}_0)-\kappa\mathbb{E}^P\left[\tau'|\tau'<\tau,\mathscr{F}_0\right]\right)\\ &+\mathrm{Prob}(\tau'\geq\tau|\mathscr{F}_0)\left(\mathbb{E}^P\left[\widehat{u}(q_\tau)-\kappa\tau'|\tau'\geq\tau,\mathscr{F}_0\right]\right)\\ \leq&\mathrm{Prob}(\tau'<\tau|\mathscr{F}_0)\left(\mathbb{E}^\pi[\widehat{u}(q)]-\kappa\mathbb{E}^P[\tau|\mathscr{F}_0]-\kappa\mathbb{E}^P\left[\tau'|\tau'<\tau,\mathscr{F}_0\right]\right)\\ &+\mathrm{Prob}(\tau'\geq\tau|\mathscr{F}_0)\left(\mathbb{E}^\pi[\widehat{u}(q)]-\kappa\mathbb{E}^P\left[\tau|\tau'\geq\tau,\mathscr{F}_0\right]\right)\\ =&\mathrm{Prob}(\tau'<\tau|\mathscr{F}_0)\left(\mathbb{E}^\pi[\widehat{u}(q)]-\kappa\mathbb{E}^P\left[\tau|\tau'<\tau,\mathscr{F}_0\right]\right)\\ &+\mathrm{Prob}(\tau'\geq\tau|\mathscr{F}_0)\left(\mathbb{E}^\pi[\widehat{u}(q)]-\kappa\mathbb{E}^P\left[\tau|\tau'\geq\tau,\mathscr{F}_0\right]\right)\\ =&\mathbb{E}^P[\widehat{u}(q_\tau)-\kappa\tau|\mathscr{F}_0].\end{aligned}$$

The first equality is from the definition of conditional expectations, the process defining q_t, equation (2), and the convention that $q_t=q_\tau$ for $t>\tau$. The first inequality is from the constraint $\mathbb{E}^\pi[\widehat{u}(q)-\widehat{u}(\bar{q}_0)]\geq\kappa\mathbb{E}^P[\tau|\mathscr{F}_0]$ in the relaxed problem (R) and $E[\tau|\tau'\geq\tau]\leq E[\tau'|\tau'\geq\tau]$. The second equality is from the memorylessness property of τ,[24] and the last from the definition of conditional expectations. Note that if the constraint in (R) is slack,

[24] Specifically, $\mathbb{E}^P[\tau|\tau'<\tau,\mathscr{F}_0]=\mathbb{E}^P[\tau-\tau'|\tau'<\tau,\mathscr{F}_0]+\mathbb{E}^P[\tau'|\tau'<\tau,\mathscr{F}_0]=\mathbb{E}^P[\tau|\mathscr{F}_0]+\mathbb{E}^P[\tau'|\tau'<\tau,\mathscr{F}_0]$.

the first inequality is strict if τ' is not equal to τ P-a.e. In this case, the α-dilution of π strongly implements principal's utility level $\mathbb{E}^{\pi}[G(q)-G(\bar{q}_0)]$. □

A.4 Proof of Lemma 2

Proof. If $V^B(\bar{q}_0) > \widehat{u}(\bar{q}_0)$, then let π implement $V^B(\bar{q}_0)$. By the assumption that full information is strictly beneficial, let π' be the distribution of posterior beliefs corresponding to full information and note by Assumption $\mathbb{E}^{\pi'}[G(q')-G(\bar{q}_0)] > 0$. Then, the dilution of a convex combination of π and π' is feasible, incentive compatible for sufficiently small weight on π' and strictly profitable for the principal. Therefore, $J(\bar{q}_0) > 0$.

Next, suppose for the purpose of contradiction that $\mathbb{E}^{\pi_A}[V^B(q)-\widehat{u}(q)] > 0$. For each $q \in \mathrm{Supp}(\pi_A)$, let π^q be the probability measure that implements $V^B(q)$. Define $\pi' \in \Pi(\bar{q}_0)$ by $d\pi'(q') = \mathbb{E}^{\pi_A(q)}[d\pi^q(q')]$. Then

$$
\begin{aligned}
\mathbb{E}^{\pi'(q')}\left[\widehat{u}(q')-\frac{\kappa}{\chi}H(q')-\widehat{u}(\bar{q}_0)+\frac{\kappa}{\chi}H(\bar{q}_0)\right] =& \mathbb{E}^{\pi_A(q)}\left[\mathbb{E}^{\pi^q(q')}\left[\widehat{u}(q')-\frac{\kappa}{\chi}H(q')\right]-\widehat{u}(\bar{q}_0)+\frac{\kappa}{\chi}H(\bar{q}_0)\right] \\
=& \mathbb{E}^{\pi_A(q)}\left[\mathbb{E}^{\pi^q(q')}\left[\widehat{u}(q')-\frac{\kappa}{\chi}H(q')-\widehat{u}(q)+\frac{\kappa}{\chi}H(q)\right]\right] \\
&+\mathbb{E}^{\pi_A(q)}\left[\widehat{u}(q)-\frac{\kappa}{\chi}H(q)-\widehat{u}(\bar{q}_0)+\frac{\kappa}{\chi}H(\bar{q}_0)\right] \\
=& \mathbb{E}^{\pi_A(q)}\left[V^B(q)-\hat{u}(q)\right]+V^B(\bar{q}_0)-\widehat{u}(\bar{q}_0) \\
>& V^B(\bar{q}_0)-\widehat{u}(\bar{q}_0).
\end{aligned}
$$

This contradicts the optimality of π_A, and by $V^B(q) \geq \hat{u}(q)$ we must have $\mathbb{E}^{\pi_A}[V^B(q) - \hat{u}(q)] = 0$. □

A.5 Proof of Proposition 3

Proof. Let

$$U^a(q) = \widehat{u}(q) - \frac{\kappa}{\chi}H(q) \qquad \text{and} \qquad U^p(q) = U^a(q) + \lambda G(q),$$

where λ is the multiplier from Proposition 1.

We first note that $\lambda > 0$. Indeed,

$$V^B(\bar{q}_0) = \frac{\kappa}{\chi}H(\bar{q}_0) + \max_{\pi:\, E^{\pi}[q]=\bar{q}_0} E^{\pi}[U^a(q)].$$

Hence the assumption $V^B(\bar{q}_0) > \widehat{u}(\bar{q}_0)$ implies

$$\max_{\pi:\, E^{\pi}[q]=\bar{q}_0} E^{\pi}[U^a(q)] > U^a(\bar{q}_0).$$

If $\lambda = 0$, then Proposition 1 would imply that any principal-optimal distribution maximizes $E^{\pi}[U^a(q)]$, while the binding participation constraint in Proposition 1 gives

$$E^{\pi^*}[U^a(q)] = U^a(\bar{q}_0),$$

a contradiction. Therefore $\lambda > 0$.

Now fix $q^* \in Q^p(\bar{q}_0)$, and let π^p be a principal-optimal distribution with $q^* \in \mathrm{Supp}(\pi^p)$. By the usual concavification/Lagrangian lemma for Bayes-plausible problems, there exists an affine function L^p such that

$$U^p(q) - L^p(q) \leq 0 \qquad \forall q \in \mathscr{P}(X),$$

with equality for every $q \in \mathrm{Supp}(\pi^p)$.

We first show that $q^* \notin Q^a(\bar{q}_0)$. Suppose instead that $q^* \in Q^a(\bar{q}_0)$. Let π^a be a benchmark-optimal distribution with $q^* \in \mathrm{Supp}(\pi^a)$. Again by the Lagrangian lemma, there exists an affine function L^a such that

$$U^a(q) - L^a(q) \leq 0 \qquad \forall q \in \mathscr{P}(X),$$

with equality for every $q \in \mathrm{Supp}(\pi^a)$.

Because $q^* \in \mathrm{relint}(\mathscr{P}(X))$, the existence of these touching affine majorants implies that U^a and U^p are differentiable at q^*. The only possible nondifferentiability comes from the piecewise-linear convex function $\widehat{u}$, and an interior convex kink cannot admit a global affine majorant touching from above. Hence

$$\nabla L^a = \nabla U^a(q^*) \qquad \text{and} \qquad \nabla L^p = \nabla U^p(q^*) = \nabla U^a(q^*) + \lambda \nabla G(q^*).$$

Therefore, for every $q \in \mathscr{P}(X)$,

$$U^p(q) - U^p(q^*) \leq \nabla U^p(q^*) \cdot (q - q^*),$$

and

$$U^a(q) - U^a(q^*) \leq \nabla U^a(q^*) \cdot (q - q^*).$$

Subtracting yields

$$\lambda \big(G(q) - G(q^*)\big) \leq \lambda \nabla G(q^*) \cdot (q - q^*) \qquad \forall q \in \mathscr{P}(X).$$

Since G is convex and differentiable,

$$G(q) - G(q^*) \geq \nabla G(q^*) \cdot (q - q^*) \qquad \forall q \in \mathscr{P}(X).$$

Because $\lambda > 0$, both inequalities must bind for every q. Thus G is affine on $\mathscr{P}(X)$. But then for every Bayes-plausible π,

$$E^{\pi}[G(q)] = G(E^{\pi}[q]) = G(\bar{q}_0),$$

so $J(\bar{q}_0) = 0$, contradicting Lemma 2. We conclude that

$$Q^p(\bar{q}_0) \cap Q^a(\bar{q}_0) = \varnothing.$$

Now suppose $q^* \in Q^p(\bar{q}_0) \cap \mathrm{Conv} Q^a(\bar{q}_0)$. By Carathéodory's theorem, there is some $\pi' \in \mathscr{P}(Q^a(\bar{q}_0))$ such that $\mathbb{E}^{\pi'}[q] = q^*$, and by $q^* \notin Q^a(\bar{q}_0)$,

$$U^a(q^*) < \mathbb{E}^{\pi'}[U^a(q)].$$

It follows by the convexity of G and $\lambda > 0$ that

$$U^p(q^*) = U^a(q^*) + \lambda G(q^*) < \mathbb{E}^{\pi'}[U^a(q) + \lambda G(q^*)] \leq \mathbb{E}^{\pi'}[U^p(q)],$$

which contradicts $q^* \in Q^p(\bar{q}_0)$. We conclude that $q^* \in Q^p(\bar{q}_0)$ implies $q^* \notin \mathrm{Conv}Q^a(\bar{q}_0)$ □

A.6 Proof of Proposition 4

Proof. Per Proposition 1, the optimal π necessarily leads to $J(\bar{q}_0) = \mathbb{E}^\pi[\widehat{u}(q) - \widehat{u}(\bar{q}_0)]$. Then, (R) reduces to:

$$\sup_{\pi \in \mathscr{P}(\mathscr{P}(X)):\mathbb{E}^\pi[q]=q_0} \mathbb{E}^\pi[\widehat{u}(q) - \widehat{u}(\bar{q}_0)]$$
$$s.t.\ \frac{\kappa}{\chi}\mathbb{E}^\pi[H(q) - H(\bar{q}_0)] \leq \mathbb{E}^\pi[\widehat{u}(q) - \widehat{u}(\bar{q}_0)].$$

Proposition 1 implies that π could be picked with finite support. Let s be the signal that induces π per the Bayes rule. Let $p(s|x_1,x_2)$ denote the conditional distribution of s. Let s' be another signal that has conditional distribution $p(s' = z|x_1,x_2) = \mathbb{E}^{\bar{q}_0}[p(s = z|x_1,x_2)|x_1]$, i.e., s' contains s's information about x_1 only. Let π' be the distribution of posterior induced by s'. Evidently, since u depends on x_1 only,

$$\mathbb{E}^\pi[\widehat{u}(q) - \widehat{u}(\bar{q}_0)] = \mathbb{E}^{\pi_1}[\widehat{u}(q, \bar{q}_0|_{X_2}) - \widehat{u}(\bar{q}_0)] = \mathbb{E}^{\pi'}[\widehat{u}(q) - \widehat{u}(\bar{q}_0)].$$

Therefore, π' is feasible. Next, we show that π' weakly improves π. Note that

$$\begin{aligned}
&\mathbb{E}^\pi[H(q) - H(\bar{q}_0)] = I(s;x_1,x_2) = \mathbb{E}^{\bar{q}_0}[D_{KL}(p(s|x_1,x_2))||p(s)];\\
&\mathbb{E}^{\pi'}[H(q) - H(\bar{q}_0)] = I(s';x_1,x_2) = \mathbb{E}^{\bar{q}_0}[D_{KL}(p(s'|x_1))||p(s')]\\
\implies &\mathbb{E}^\pi[H(q) - H(\bar{q}_0)] - \mathbb{E}^{\pi'}[H(q) - H(\bar{q}_0)]\\
&= \mathbb{E}^{\bar{q}_0}\left[\log\frac{p(s|x_1,x_2)}{p(s)} - \log\frac{p(s|x_1)}{p(s)}\right]\\
&= \mathbb{E}^{\bar{q}_0}[D_{KL}(p(s|x_1,x_2)||p(s|x_1))] \geq 0,
\end{aligned}$$

strictly if $p(s|x_1,x_2) \neq p(s|x_1)$ for some (x_1,x_2) in the support of $\bar{q}_0$.

□

A.7 Proof of Proposition 5

Proof. We slightly abuse notation and let x_1, x_2 denote the random variables that determines the state realization. $\forall\pi$, let $\pi_i = \pi|_{X_i}$. Let s_1 be the random variable that is independent to x_2 and induces posterior belief π_1 for x_1. Let s_2 be the random variable that induces posterior belief π. Let I denote mutual information. Then,

$$\begin{aligned}\mathbb{E}^{\pi}[H(q)-H(\bar{q}_0)] =& I(s_1,s_2;x_1,x_2)\\ =& I(s_1,s_2;x_1)+I(s_1,s_2;x_2|x_1)\\ \geq& I(s_1;x_2)+I(s_2;x_2|x_1)\\ =& I(s_1;x_2)+I(s_2,x_1;x_2)-\underbrace{I(x_1;x_2)}_{=0}\\ \geq& I(s_1;x_2)+I(s_2;x_2)\\ =& \mathbb{E}^{\pi_1}[H(q,\bar{q}_0|_{X_2})-H(\bar{q}_0)]+\mathbb{E}^{\pi_2}[H(\bar{q}_0|_{X_1},q)-H(\bar{q}_0)].\end{aligned}$$

In the derivation, $(q,\bar{q}_0|_{x_2})$ denotes the belief vector where the probability of state x_1 is given by q and the probability of state x_2 is given by $\bar{q}_0$ restricted to x_2, respectively. Belief vector $(\bar{q}_0|_{X_1},q)$ is defined analogously. Therefore, let $\pi'=\pi_1\otimes\pi_2$,

$\mathbb{E}^{\pi}[H(q)-H(\bar{q}_0)]\geq\mathbb{E}^{\pi_1}[H(q,\bar{q}_0|_{X_2})-H(\bar{q}_0)]+\mathbb{E}^{\pi_2}[H(\bar{q}_0|_{X_1},q)-H(\bar{q}_0)]=\mathbb{E}^{\pi'}[H(q)-H(\bar{q}_0)]$;

$\mathbb{E}^{\pi}[\widehat{u}(q)-\widehat{u}(\bar{q}_0)]=\mathbb{E}^{\pi_1}[\widehat{u}(q,\bar{q}_0|_{X_2})-\widehat{u}(\bar{q}_0)]=\mathbb{E}^{\pi'}[\widehat{u}(q)-\widehat{u}(\bar{q}_0)]$;

$\mathbb{E}^{\pi}[G(q)-G(\bar{q}_0)]=\mathbb{E}^{\pi_2}[G(\bar{q}_0|_{X_1},q)-G(\bar{q}_0)]=\mathbb{E}^{\pi'}[G(q)-G(\bar{q}_0)]$.

Then, π' is feasible and weakly improves π. Proposition 1 implies that $J(\bar{q}_0)$ can be solved by maximizing

$$\begin{aligned}&\mathbb{E}^{\pi_1\otimes\pi_2}\left[\widehat{u}(q)-\frac{\kappa}{\chi}H(q)+\lambda G(q)\right]\\ =&\underbrace{\mathbb{E}^{\pi_1}\left[\widehat{u}(q,\bar{q}_0|_{X_2})-\frac{\kappa}{\chi}H(q,\bar{q}_0|_{X_2})\right]}_{\text{maximized by }\pi_1^*\text{ solving }V^B}+\mathbb{E}^{\pi_2}\left[-\frac{\kappa}{\chi}H(\bar{q}_0|_{X_1},q)+\lambda G(\bar{q}_0|_{X_1},q)\right].\end{aligned}$$

□

B Omitted Proofs for Section 4

B.1 Proof of Proposition 6

Proof. By Theorem 1, there is some $\pi\in\Pi^*$ that is the law of q_τ under the optimal policy. By the martingale property of the belief process, $q_t=E^P[q_\tau|\mathscr{F}_t,t<\tau]$ for all $t\in[0,\tau)$, and thus q_t must lie in the convex hull of the support of π; hence, $\overline{V}_R(q_t,\pi)$ is well defined.

For the purpose of contradiction, suppose that the inequality in Proposition 6 does not hold. Then, there exists $\varepsilon>0$ and a positive probability event $\Omega'\in\mathscr{F}_t|_{t<\tau}$ s.t. $\forall\omega\in\Omega'$, $\overline{V}_R(q_t(\omega),\pi)<-\varepsilon$. By the definition of $\overline{V}$,

$$\mathbb{E}\left[\widehat{u}(q_\tau)-\widehat{u}(q_t)-\frac{\kappa}{\chi}(H(q_\tau)-H(q_t))\Big|t<\tau,\omega\in\Omega'\right]\leq-\varepsilon.$$

However, this violates the incentive compatibility condition of the agent. By deviating to stopping immediately at event Ω', the agent's expected utility improves by $\varepsilon \cdot \text{Prob}^P(\Omega')$. Contradiction. $\square$

B.2 Proof of Lemma 3

Proof. Define:

$$\tau' = \tau \wedge q_t \text{ first leaves } E^A.$$

By definition $\tau' \leq \tau$. Let $\bar{E}^A$ denote the closure of E^A. Since $\text{Supp}(\mathrm{q}_{\tau'-}) \subset E^A$, $\text{Supp}(q_{\tau'}) \subset \bar{E}^A$. We prove by contradiction that $\tau' = \tau$. Suppose $\tau' < \tau$ on a positive measure, on which

$$\begin{aligned}\mathbb{E}^P\left[\widehat{u}(q_\tau) - (\tau - \tau') \cdot \kappa \middle| \tau' < \tau\right] =& \mathbb{E}^P\left[\mathbb{E}^P\left[\widehat{u}(q_\tau) - (\tau - \tau') \cdot \kappa \middle| \tau'\right] \middle| \tau' < \tau\right] \\ <& \mathbb{E}^P\left[\widehat{u}(q_{\tau'}) \middle| \tau' < \tau\right].\end{aligned}$$

The inequality is from the fact that $\tau' < \tau \implies q_{\tau'} \notin E^A$. Therefore, τ' strictly improves upon τ; hence, τ is not incentive compatible. The contradiction implies that $\text{Supp}(q_\tau) \subset \bar{E}^A$. $\square$

B.3 Proof of Proposition 7

Proof. As is discussed in Section 3, the agent-optimal policy can be solved by concavifying $\widehat{u}(q) - \frac{\kappa}{\chi} H(q)$. Therefore, there exists a linear function $L(q)$ that is weakly higher than $\widehat{u}(q) - \frac{\kappa}{\chi} H(q)$ and tangents it at two beliefs $q^1 < \bar{q}_0 < q^2$. WLOG, let q^1 and q^2 be the smaller and largest such beliefs, respectively. Since $\widehat{u} - \frac{\kappa}{\chi} H$ is piece-wise strictly concave, the interval $[q^1, q^2]$ is bounded away from the rest of E^A.

Lemma 3 implies that any admissible principal's strategy has $\text{Supp}(q_\tau) \subset \bar{E}^A$. Moreover, any continuous path that starts from $\bar{q}_0$ and ends outside of $[q^1, q^2]$ leaves E^A; hence, it is not admissible. Therefore, $\text{Supp}(q_\tau) \subset [q^1, q^2]$.

Therefore, $J^C(\bar{q}_0)$ is bounded above by the following relaxed problem:

$$\sup_{\pi \in \mathscr{P}(\mathscr{P}([q^1, q^2]))} \mathbb{E}^\pi[G(q) - G(\bar{q}_0)].$$

Suppose G is affine on $[q^1, q^2]$, then the principal is completely indifferent between any strategy; hence, the agent optimal strategy is optimal for the principal as well. If G is not affine on $[q^1, q^2]$, the relaxed problem is solved by π^* with support $\{q^1, q^2\}$ (such π^* is unique). By Hébert and Woodford [2023], there exists a Gaussian process that implements π^* and satisfies the information constraint. Note that this Gaussian process also implements the agent-maximal continuation payoff (the upper concave hull of $\widehat{u}(q) - \frac{\kappa}{\chi}(H(q) - H(q_t))$) for every interim belief; hence, it is incentive compatible. $\square$

B.4 Proof of Theorem 2

Proof. **Sufficiency**: condition (ii) of Definition 3 is a direct implication of Propositions 1 and 2. Next, we verify condition (i). Suppose for the purpose of contradiction that condition (i) is violated, i.e. there exist some t, a positive probability Borel set $B \in \mathscr{F}_t$ and $\varepsilon > 0$ s.t. $\forall B' \subset B$ and $B' \in \mathscr{F}_t$,

$$\mathbb{E}^P[G(q_\tau) - G(q_t)|\omega \in B'] \leq \mathbb{E}^P[J(q_t)|\omega \in B'] - \varepsilon;$$

Let π'_q be the policy that implements $J(q)$ for each q. Let π be the distribution of q_τ. Then, the inequality above implies

$$\mathbb{E}^P[\mathbb{E}^{\pi'_{q_t(\omega)}}[G(q)|\omega \in B]P(B) + \mathbb{E}^P[G(q_\tau)|\omega \notin B](1-P(B)) > \mathbb{E}^\pi[G(q)].$$

Meanwhile, (4) implies that

$$\mathbb{E}^P[\widehat{u}(q_\tau) - \widehat{u}(q_t) - \kappa(\tau - t)|\omega \in B'] = 0.$$

Since each π_q is incentive compatible itself,

$$\begin{aligned}
&\mathbb{E}^P[\mathbb{E}^{\pi'_{q_t(\omega)}}[\widehat{u}(q) - \widehat{u}(q_t(\omega)) - \frac{\kappa}{\chi}(H(q) - H(q_t(\omega))]|\omega \in B] \\
&\geq\ 0\ = \mathbb{E}^P[\widehat{u}(q_\tau) - \widehat{u}(q_t) - \frac{\kappa}{\chi}(H(q_\tau) - H(q_t))|\omega \in B] \\
\implies\ &\mathbb{E}^P[\mathbb{E}^{\pi'_{q_t(\omega)}}[\widehat{u}(q) - \frac{\kappa}{\chi}H(q)]|\omega \in B] \geq \mathbb{E}^P[\widehat{u}(q_\tau) - \frac{\kappa}{\chi}H(q_\tau)|\omega \in B]
\end{aligned}$$

The two inequalities above implies that the probability measure $\pi'' = \mathbb{E}^P[\pi'_{q_t(\omega)}|\omega \in B]P(B) + \mathbb{E}^P[\delta_{q_\tau}|\omega \notin B](1-P(B))$ is feasible in (R) and strictly improves upon π, leading to a contradiction.

Due to the sufficiency result, the dilution policy defined in Theorem 1 specifies a strong equilibrium.

Necessity: Suppose for the purpose of contradiction that the equality doesn't hold. Then, there exist some t, a positive probability Borel set $B \in \mathscr{F}_t$ and $\varepsilon > 0$ s.t. $\forall B' \subset B$ and $B' \in \mathscr{F}_t$,

$$\begin{aligned}
&\mathbb{E}^P[\widehat{u}(q_\tau) - \widehat{u}(q_t) - \kappa(\tau - t)|\omega \in B'] \geq \varepsilon \\
\implies &\mathbb{E}^P\left[\widehat{u}(q_\tau) - \widehat{u}(q_t) - \frac{\kappa}{\chi}(H(q_\tau) - H(q_t))|\omega \in B'\right] \geq \varepsilon.
\end{aligned}$$

Let $\bar{I} = \max_{q \in \mathscr{P}(X)}(\sum q_x H(e_x) - H(q))$. Let $\delta = \bar{I}/(\bar{I} + \varepsilon)$. Let π'_q be the distribution of $q_\tau|_{q_t = t}$ for $q \in \text{supp} q_t$. Let $\pi''_q = \delta\pi'_q + (1-\delta)\sum q_x e_x$. Then, by construction, $\forall B' \subset B$,

$$\mathbb{E}^P\left[\mathbb{E}^{\pi''_{q_t(\omega)}}\left[\widehat{u}(q) - \widehat{u}(q_t) - \frac{\kappa}{\chi}(H(q) - H(q_t))\right] \middle| \omega \in B'\right] \geq 0.$$

Since π''_q leaves the agent with non-negative utility, it is feasible in (R); hence, $\mathbb{E}^{\pi''_{q_t(\omega)}}[G(q) - G(q_t(\omega))] \leq J(q_t(\omega))$. By assumption, $\text{supp}(q_\tau) \subset \mathscr{P}(X)^+$; hence, $\text{supp}(q_t) \subset \mathscr{P}(X)^+$

due to the martingale property. Therefore, $\forall B' \subset B$, $\mathbb{E}^P[G(q_\tau)|\omega \in B'] < \mathbb{E}^P[\sum q_x(\omega) G(e_x)|\omega \in B']$. This implies,

$\mathbb{E}^P[G(q_\tau) - G(q_t(\omega))|\omega \in B'] < \mathbb{E}^P\left[\mathbb{E}^{\pi''_{q_t(\omega)}}[G(q) - G(q_t(\omega))]\,|\omega \in B'\right] \leq \mathbb{E}^P[J(q_t)|\omega \in B']$.

This violates condition (i), which states

$$\mathrm{Prob}^P\left(\mathbb{E}^P[G(q_\tau) - G(q_t) - J(q_t)|\mathscr{F}_t] \geq 0\right) = 1.$$

□

B.5 Proof of Proposition 8

Proof. By Theorem 1, there is some $\pi \in \Pi^*$ that is the law of q_τ under the optimal policies. By the martingale property of the belief process, $q_t = E^P[q_\tau|\mathscr{F}_t, t < \tau]$ for all $t \in [0, \tau)$, and thus q_t must lie in the convex hull of the support of π.

For the sake of contradiction, suppose Proposition 8 does not hold. Then, there exists positive probability event $B' \in \mathscr{F}_t|_{t<\tau}$ and $\varepsilon > 0$ such that $\forall \omega \in B'$,

$$\inf_{\pi \in \Pi^*} \underline{V}_R(q_t(\omega), \pi) \geq \varepsilon.$$

Then, $\mathbb{E}^P\left[\widehat{u}(q_\tau) - \widehat{u}(q_t) - \frac{\kappa}{\chi}(H(q_\tau) - H(q_t))\Big|\omega \in B'\right] \geq \varepsilon$, i.e., the agent obtains positive interim surplus on positive probability event B', which violates Theorem 2. □