Persuasion and Optimal Stopping

Andrew Koh[*] Sivakorn Sanguanmoo[†] Weijie Zhong[‡]
MIT MIT Stanford GSB

This version: December 4, 2024

Abstract

We analyze the interplay between persuasion, timing, and commitment. A principal conducts a sequence of statistical experiments to persuade an agent to stop at the right time, in the right state, and choose the right action. We develop a *revelation principle* which delivers a first-order approach for solving the principal's problem under commitment, and an *anti-revelation principle* which establishes that commitment is unnecessary and transforms the solution via indirect recommendations to restore dynamic consistency. We further characterize how time and action preferences jointly shape optimal strategies featuring a *suspense-generation* stage which optimally concentrates the agent's stopping time, followed by an *action-targeting* stage which maximally correlates/anticorrelates persuasion and delay.

1 Introduction

At the heart of many economic decisions lies an optional stopping problem paired with a choice problem—a decision maker chooses *when* to stop gathering information as well as *what* irrevocable action to take. The decision maker's stopping time and action are, in turn, jointly determined by the flow of information over time, which makes information a powerful tool for shaping both timing and choice. For instance, a company might strategically reveal information about its financial health to steer investors to sell their shares at the right time in different states. Similarly, an advisor

[*]MIT Department of Economics; email: ajkoh@mit.edu
[†]MIT Department of Economics; email: sanguanm@mit.edu
[‡]Stanford GSB; email: weijie.zhong@stanford.edu

We are grateful to Daron Acemoglu, Ian Ball, Ben Brooks, Jeffery Ely, Drew Fudenberg, Stephen Morris, and Alex Wolitzky for insightful discussions. We also thank seminar participants at Harvard CS and SITE for comments.

might strategically reveal information about demand conditions in different markets to guide her clients to enter the right market at the right time.

In each of these examples, there are two distinct tensions. First, there is a tension between *persuasion* and *timing*: on the one hand, the information that—from the principal's point of view—best shapes the agent's actions is typically quite specific, and induces particular kinds of uncertainty; on the other hand, to manipulate the timing of the agent's action, the principal must release enough future information as a carrot to incentivize waiting. Second, there is a tension between *timing* and *commitment*: the principal would like to promise future information to shape the timing of the agent's action; yet, when the time to disclose this information arrives, the principal might be tempted to renege on her initial promise. We develop a unified analysis of how these tensions are traded off against each other.

In our model, there is an unknown, persistent payoff-relevant state. A principal chooses a sequence of history-dependent statistical experiments and an agent, observing the realizations of these experiments, makes an irreversible choice of when to stop and which action to take. The principal's and agent's preferences can depend on state, action, and time in arbitrary ways—in particular, they are not necessarily time-stationary.

Revelation principle and first-order characterization. When the principal has full intertemporal commitment, we establish a "revelation principle": it is without loss to use *simple recommendation* strategies that only send direct message "continue" or messages of the form "stop now and take a certain action". This converts the dynamic persuasion problem into a tractable semi-static linear program in which the principal directly controls the joint distribution of the agent's stopping time and stopping belief, subject to a series of interim obedience constraints. We establish strong duality and derive a sufficient and near necessary first-order characterization of the optimal policy (Theorem 1). The first-order condition makes precise how the agent's incentives are intertemporally linked: the optimal distribution of stopping time and belief "concavifies" a combination of the principal's and the agent's past and present utilities, where a time-dependent multiplier tracks the incentive value of information. Then, solving the dynamic persuasion problem boils down to solving a single-dimensional ordinary differential equation characterizing this multiplier.

The value of intertemporal commitment and an "anti-revelation principle". When the principal lacks intertemporal commitment, we recast the persuasion problem under limited commitment as a dynamic *experiment-selection game* where, in every period, the principal chooses a static statistical experiment and the agent chooses whether to stop to act based on her information so far, or continue. When the principal weakly enjoys delay, the value of intertemporal commitment is zero: the principal can achieve her preferred commitment *outcome*—the belief-stopping time distribution—in a subgame-perfect equilibrium of the experiment-selection game even without intertemporal commitment (Theorem 2 (i)).

How does the principal achieve this? We establish an "anti-revelation principle" (Theorem 2 (ii)) which pins down equilibrium strategies. Simple recommendations—which tell the agent to continue and nothing else—do not guarantee the principal her commitment payoff. This gap arises when the optimal simple recommendation instructs the agent to wait without offering any interim information. This leaves the agent extra continuation surplus from waiting which, in turn, creates incentives for the principal to renege on her promise to deliver information. Instead, in the optimal equilibrium strategy, the principal sends *indirect* interim messages which are maximally informative to the extent that they eliminate the agent's interim surplus. Because information is irreversible, this allows the principal to tie her own hands at future histories. Our results highlight the dual role of information as both a carrot to incentivize the agent and as a stick for the designer to discipline her future self.

Together, Therorems 1 and 2 yield a tractable and unified procedure for solving dynamic persuasion problems: First, one can appeal to the "revelation principle" to focus on simple recommendations and solve for the principal-optimal outcome via the first-order approach we develop (Theorem 1). Then, to recover the persuasion strategy that implements the same outcome in the absence of intertemporal commitment, one can appeal to the "anti-revelation principle" (Theorem 2) which pins down when and what kinds of extra interim information the principal should inject.

The form of optimal dynamic persuasion We use this methodology to shed light on how action and time preferences jointly shape the form of dynamic persuasion, especially when time preferences are non-stationary and the principal lacks commitment.

We begin by analyzing a dynamic version of the canonical persuasion model with binary states ($\theta = L, R$), binary actions ($a = \ell, r$), and continuous time. The impatient agent wishes to match her action to the state, while the principal strictly enjoys delay and obtains different payoffs from the two actions. Crucially, we allow the principal's gain (and agent's loss) from delay contingent on each action to be general: time preferences can be non-stationary, and the gain from either action can vary flexibly with time. We show that for a wide class of principal and agent preferences, optimal persuasion strategies consist of a *suspense-generation* stage, followed by an *action-targeting stage*. Suspense generation consists of inconclusive 'plot twists': the agent's beliefs jump stochastically between two interim paths, but the agent is never certain enough to stop. Action-targeting consists of a carefully chosen arrival rate of conclusive news about a particular state. The strategy concludes when the posterior belief reaches an endogenous threshold.

Simplicity along the state and action dimensions allow us to transparently convey the richness and subtleties involved in dynamic persuasion. This delivers a range of economic insights into how (i) time-risk preferences; (ii) persuasion-delay complementarity/substitutability; and (iii) persuasion gain shape optimal dynamic information:

1. *Time-risk preferences shape suspense window.* By generating suspense, the principal concentrates the agent's stopping time since the agent stops only in the action-targeting stage. When the principal is more time-risk averse (e.g., there is diminishing marginal gain from delay), the suspense-generation stage lengthens to concentrate the agent's stopping time around a late decision window (during action-targeting). Conversely, under time-risk-loving preferences, the principal necessarily generates no suspense at all to maximize the dispersion over stopping times. Indeed, we show that these qualitative insights on how time-risk preferences shape the optimal timing of information extend much more generally (Propositions 5 and 6).

   During the suspense window, by offering the arrival of inconclusive "plot twists", suspense generation pushes up the agents' interim outside option of stopping. This ensures that the surplus from continuation is zero which, per Theorem 2 (ii) is necessary and sufficient to ensure that the principal has no incentive to deviate—hence guaranteeing implementability even under limited commitment.

2. *Persuasion-delay complementarity/substitutability shape direction of targeting.* The principal initially prefers the action r to be taken and we call the principal's premium from action r over ℓ the "persuasion gain". When the principal's persuasion gain increases with time (persuasion-delay complementarity), we show that the second stage consists of random arrival of conclusive news that drives the agent to take the principal's dispreferred action ℓ; in the absence of such news, the agent eventually

takes the principal's preferred action. ℓ-targeting exploits the complementarity by *positively correlating* the events that the agent stops late, and the event that the agent takes the principal's preferred action.

Conversely, when the persuasion gain decreases with time (persuasion-delay substitutability), the action-targeting stage consists of conclusive news which drives the agent to take the principal's initially preferred action: this maximizes the chance that *either* the agent stops late, *or* the agent takes the principal's initially preferred action.

3. *Persuasion gain shape scope of persuasion.* When the initial gain from persuasion is large relative to the gain from delay, the principal reduces the scope of persuasion by shortening the window of both suspense-generation and action-targeting; the optimal distribution of stopping beliefs is then closer to that of the static benchmark (Kamenica and Gentzkow, 2011).

Related literature. We provide by far the most general treatment of persuasion in stopping problems, where we allow both the principal's and agent's time and action preferences to be fully general and accommodate both full and limited commitment. Our framework nests existing work in which the principal's preference is independent of the agent's action such as Ely and Szydlowski (2020), Orlov, Skrzypacz, and Zryumov (2020) (the commitment case),[1] Ely (2017) (the basic model), Knoepfle (2020) (the single sender case), Koh and Sanguanmoo (2024)[2] (the instrumental value case), and contemporaneous work by Saeedi et al. (2024) (the common prior case).[3] The novelty and analytical strength of our approach stem from our revelation and anti-revelation principles. These principles enable us to first address a tractable semi-static problem via a novel first-order approach that accommodates general preferences, then modify the solution to handle limited commitment. We use this to elucidate important qualitative features—such as non-stationarity, indirect messages, and imperfect revelations—which are fundamental to dynamic persuasion. By contrast, existing work has often relied on (i) specific payoff structures such as additive or multiplicative separability,

[1] Their setting involves an evolving public state, which is nested in an extension of our model in Online Appendix VI.

[2] The current paper succeeds the release of Koh and Sanguanmoo (2024) in 2022.

[3] Specifically, models that involve myopic agent, repeated actions (e.g. Renault et al. (2017), the general model of Ely (2017), Ball (2023), and Zhao et al. (2024)), informational frictions (e.g. Hébert and Zhong (2022) and Che et al. (2023)), or *both* stochastic states and limited commitment (Orlov, Skrzypacz, and Zryumov (2020)) are not nested by our framework. The asymmetric prior case of Saeedi et al. (2024) can be nested by our framework via a transformation technique they develop.

stationary time preferences and principal preference for revelation; (ii) Markovian restrictions; and (iii) principal commitment. While these assumptions have facilitated tractability, they typically led to the prediction of simple strategies that stochastically but fully reveals certain states.[4] Our results paint a richer and more complete picture.

Our analysis of the "revelation principle" in a multi-stage principal-agent setting in Section 3 is closely connected to the seminal work of Forges (1986); Myerson (1986); Sugaya and Wolitzky (2021), as well as the study of dynamic "mechanism-selection" games by Laffont and Tirole (1988); Bester and Strausz (2001); Skreta (2015); Doval and Skreta (2022). Our experiment-selection game can be viewed as the informational counterpart of mechanism-selection games. However, there is no nesting relation which makes direct comparisons difficult: our contribution is to highlight how providing information is qualitatively different from designing mechanisms. Indeed, in most mechanism-selection games there is a strict gap between the principal's payoff with and without commitment—considering direct mechanisms is not without loss. By contrast, we show that with information provision, there is no commitment gap and the revelation principle (i.e., in direct recommendations) holds for pinning down the *outcome* of such games. We then go on to develop an "anti-revelation principle" which makes precise the *strategy* the principal uses to close this commitment gap even in the absence of commitment.[5]

Our dynamic persuasion model is closely related to a series of papers seeking to implement the static Bayesian persuasion strategy in a dynamic setting with informational frictions (e.g. Henry and Ottaviani (2019), Escudé and Sinander (2023), Siegel and Strulovici (2020) and Che et al. (2023)). In these papers, the static strategy of Kamenica and Gentzkow (2011) is optimal but infeasible due to the frictions, and the dynamics are shaped by the information constraints/frictions. By contrast, our setting is frictionless and our principal can choose any dynamic persuasion strategy (including the static strategy). It is the principal's incentive to delay the agent's decision that endogenously leads to a non-degenerate dynamic strategy.[6] Indeed, the trade-off between

[4] Methodologically, previous work typically made progress by constructing a relaxed problem and guessing a candidate solution (Knoepfle, 2020; Ely and Szydlowski, 2020; Hébert and Zhong, 2022), dynamic programming (Orlov et al., 2020), or extreme-point techniques (Koh and Sanguanmoo, 2024).

[5] In mechanism-selection games, the agent is fully informed and the usage of indirect messages is to control the principal's own information. By contrast, our agent is uninformed, and, by carefully designing non-simple recommendations, the designer induce *interim agent learning* to tie her own hands and restore commitment.

[6] Intriguingly, Escudé and Sinander (2023) reports no commitment gap, while Che et al. (2023); Henry and Ottaviani (2019) report strict commitment gaps. This suggests that informational frictions have strong implications on the commitment gap. Less

persuasion and delay is intuitive, but to our knowledge has not been analyzed.[7] Our analysis of how the complementarity or substitutability between persuasion and delay shapes the direction of conclusive Poisson news is also novel, and offers an endogenous foundation for the "good news" and "bad news" models which are often assumed (see, e.g., Keller et al. (2005); Keller and Rady (2015); Che et al. (2023)) as well as delineates which one should be chosen.

The connection between time-risk attitudes induced by asymmetric discount rates and the timing of information revelation has been alluded to in Ely and Szydlowski (2020); Ball and Knoepfle (2023); Saeedi et al. (2024). Our analysis is the first that makes this connection explicit by allowing for general time-risk loving and time-risk averse preferences that are necessarily non-stationary. We further identify a novel connection between time-risk aversion, limited principal commitment, and the "suspense-generation" behavior of Ely et al. (2015).

The agent's optimal stopping problem is a classic statistical decision-making problem pioneered by Wald (1947) and Arrow, Blackwell, and Girshick (1949). Several recent papers seek to endogenize the information in the optimal stopping problem by giving the agent *no* control of information (e.g., Fudenberg, Strack, and Strzalecki (2018); Gonçalves (2024)), *some* control of information (e.g. Moscarini and Smith (2001), Che and Mierendorff (2019), Liang, Mu, and Syrgkanis (2022)), or *all* control of information (e.g. Hébert and Woodford (2023), Steiner, Stewart, and Matějka (2017), Zhong (2022), Sannikov and Zhong (2024)). Our paper complements this literature by studying the endogenous choice of information in stopping problems in a principal-agent setting.

The rest of the paper is organized as follows. Section 2 introduces the model and characterizes the solution under full commitment. Section 3 analyzes the model under limited commitment. Section 4 uses this methodology to shed light on the how time and action preferences shape optimal persuasion. Section 5 concludes.

2 The Framework

2.1 Model

Primitives. Θ is a finite set of payoff-relevant states. $\mu_0 \in \Delta(\Theta)$ is the common prior belief. A is a set of actions. The time space is a compact set $T \subset \mathbb{R}_+$ which can be

directly related is Mekonnen et al. (2023) who study persuasion in a different sequential search model where payoffs from stopping are iid. They show that intertemporal commitment is unnecessary, though for completely distinct reasons.

[7] A distinct force is at play in work on persuasion with rational inattention (Bloedel and Segal, 2020; Lipnowski et al., 2020).

finite or a continuum. There are two players: the principal (she) and the agent (he). The agent makes a one-time irreversible choice of action a at a stopping time t of his choosing. For any tuple of states, actions, and stopping times $(\theta, a, t) \in \Theta \times A \times T$, the agent obtains utility $u(\theta, a, t)$ and the principal obtains utility $v(\theta, a, t)$. We assume u and v are both bounded.

Information: At the start of the game (before $t = 0$) the principal *commits* to a persuasion strategy that resembles a sequential experimentation process. Formally, the principal's strategy is a cadlag martingale process $\langle \mu_t \rangle$ in $\Delta(\Theta)$ (accompanied by a suitable underlying probability space $(\Omega, \mathcal{F} = \langle \mathcal{F}_t \rangle, \mathcal{P})$) describing the common random posterior belief process induced by the realized outcome of the statistical experiments. The role of intertemporal commitment will be analyzed in detail in Section 3.

Stopping and action: The agent moves the second. His belief updates over time and makes a one-time choice (contingent on the history of the belief process) of the action. When the agent stops in period t with a belief μ, he can choose any action taken at a time no earlier than t. Therefore, we define the indirect utility with respect to stopping belief μ and time t: for all $\mu \in \Delta(\Theta)$ and $t \in T$,

$$U(\mu, t) := \max_{a \in A, s \geq t} \mathbb{E}_\mu[u(\theta, a, s)];$$

$$V(\mu, t) := \max_{a^*, s^* \in \underset{a \in A, s \geq t}{\operatorname{argmax}} \mathbb{E}_\mu[u(\theta, a, s)]} \mathbb{E}_\mu[v(\theta, a^*, s^*)].$$

By construction, U is convex in μ and non-increasing in t. (Indirect) utility functions U and V encode all the payoff relevant information although they abstract away from optimal actions.[8] Then, the agent's strategy is a stopping time τ with respect to $\langle \mathcal{F}_t \rangle$. Given $\langle \mu_t \rangle$, the agent solves an *optimal stopping problem*:

$$\max_\tau \mathbb{E}[U(\mu_\tau, \tau)].$$

Information design problem: The principal chooses the belief-stopping time pair $(\langle \mu_t \rangle, \tau)$, subject to the *obedience constraint* that the agent finds stopping at time τ optimal:

$$\sup_{\langle \mu_t \rangle, \tau} \mathbb{E}\left[V(\mu_\tau, \tau)\right] \tag{P}$$

$$\text{s.t. } \tau \in \arg\max_{\tau'} \mathbb{E}\left[U(\mu_{\tau'}, \tau')\right]. \tag{OC}$$

[8] By defining V, we implicitly assume that the agent breaks the tie in favor of the principal. The tie-breaking rule is inconsequential for our analysis as long as the resulting indirect utility is compatible with the technical assumptions we introduce later.

Remark 1. We model the information design problem (P) as the principal directly choosing the stochastic belief martingale, following the approach of Ely, Frankel, and Kamenica (2015). The definition of belief-based indirect utilities follows the approach of Kamenica and Gentzkow (2011), with a caveat that the agent chooses not only the action but also the action time given any posterior belief.[9]

2.2 Revelation principle and simplification

The logic of the celebrated revelation principle (Forges, 1986; Myerson, 1986) suggests that if the principal wants the agent to continue, she should pool all the messages that lead to continuation into a single message. While a formal revelation principle in continuous-time games is unknown to us—and establishing it is beyond the scope of the paper—we draw on its logic to derive "direct revelation mechanisms" in our framework and show that, with principal commitment, it is without loss to consider only such mechanisms.

We begin with an informal discussion. Take any obedient strategy $(\langle \mu_t \rangle, \tau)$, a "direct mechanism" should, at any $t < \tau$, send a pooled message "continue." to the agent to recommend continuation. Then, the pooled message conveys the information that $t < \tau$. Pooling can only improve the obedience constraint; hence,

$$\mathbb{E}\left[U\left(\mu_\tau, \tau\right) | \tau > t\right] \geq U\left(\mathbb{E}[\mu_t | \tau > t], t\right), \tag{1}$$

where the LHS is the agent's obedient payoff and the RHS is the deviation payoff from stopping. Since $\langle \mu_t \rangle$ is a martingale, the optional stopping theorem implies $\mathbb{E}[\mu_t | \tau > t] = \mathbb{E}[\mu_\tau | \tau > t]$. Then, let f be the distribution of (μ_τ, τ), (1) is equivalent to

$$\frac{\int_{y>t} U(\mu, y) f(\mathrm{d}\mu, \mathrm{d}y)}{\int_{y>t} f(\mathrm{d}\mu, \mathrm{d}y)} \geq U\left(\frac{\int_{y>t} \mu f(\mathrm{d}\mu, \mathrm{d}y)}{\int_{y>t} f(\mathrm{d}\mu, \mathrm{d}y)}, t\right) \tag{2}$$

The principal's payoff is $\int V(\mu, t) f(\mathrm{d}\mu, \mathrm{d}t)$. Therefore, f roughly characterizes the "direct mechanism" induced by $(\langle \mu_t \rangle, \tau)$ and (P) can thus be reduced to maximizing $\int V(\mu, t) f(\mathrm{d}\mu, \mathrm{d}t)$ by choosing f subject to (2). In what follows, we formalize this idea. We begin by formally mapping f to a class of "direct mechanisms" called simple recommendations.

Simple recommendations: Let $D = \Delta(\Theta) \times T$. We call elements of $\Delta(D)$ belief-time distributions. Let $\Delta_{\mu_0} = \{f \in \Delta(D) | \mathbb{E}_f[\mu] = \mu_0\}$ be all belief-time distributions consistent with the prior. Fixing $f \in \Delta_{\mu_0}$, let $\bar{t} = \sup\{t | (\mu, t) \in \mathrm{Supp}(f)\}$. For $t < \bar{t}$, define

[9] We emphasize that decoupling the stopping time and the action time is without loss given that both are fully controlled by the agent. We have found it expositionally convenient to state the model this way.

$\widehat{\mu}_t = \mathbb{E}_{f(\mu,\tau)}[\mu|\tau > t]$ as the continuation belief at time t conditional on not stopping. For all $\mu \in \Delta(\Theta)$, $t \leq \bar{t}$, let $x^{(\mu,t)}$ denote the belief path

$$x^{(\mu,t)}(s) = \begin{cases} \widehat{\mu}_s & s < t \\ \mu & s \geq t. \end{cases}$$

That is, we consider simple belief paths that jump from $\widehat{\mu}_t$ only once (which suggests a recommendation of stopping) and stay constant thereafter. Since each path $x^{\mu,t}$ leads to stopping at (μ, t), to achieve the belief-time distribution f, the persuasion strategy simply needs to randomize among all such paths according to distribution f.

Define $\Phi : (\mu, t) \mapsto (x^{\mu,t}, t)$ as the map which sends a belief-time pair to the path we constructed above and a jump time.[10]

Definition 1 (Simple Recommendation)**.** *For $f \in \Delta_{\mu_0}$, define the law $\mathcal{P}$ of the corresponding belief process $\langle \mu_t^f \rangle$ and stopping time τ^f as $\mathcal{P} \circ \Phi = f$, i.e., for all Borel sets $B_x \subset D_\infty$ and $B_t \subset T$:*

$$\mathcal{P}\left(\mu_t^f \in B_x, \tau^f \in B_t\right) = \int_{\Phi^{-1}(B_x \times B_t)} f(\mathrm{d}\mu, \mathrm{d}t).$$

A simple recommendation tells the agent to continue and nothing else, which induces the belief path $\widehat{\mu}_t$. The principal only releases information when the agent is meant to stop. The following lemma establishes the "revelation principle" within our framework.

Lemma 1. *$\forall$ feasible and obedient strategies $(\langle \mu_t \rangle, \tau)$, let $f \sim (\mu_\tau, \tau)$, then $f \in \Delta_{\mu_0}$ and f satisfies (2). Conversely, $\forall f \in \Delta_{\mu_0}$ that satisfy (2), $(\langle \mu_t^f \rangle, \tau^f)$ is a feasible and obedient strategy and $f \sim (\mu^f_{\tau^f}, \tau^f)$.*

Proof. See Online Appendix I.2. *Q.E.D.*

Lemma 1 states that the outcomes (belief-time distributions) of all feasible and obedient strategies are exactly the outcomes of all obedient simple recommendations. Therefore, (P) is equivalent to the following reduced problem:

$$\sup_{f \in \Delta_{\mu_0}} \int V(\mu, t) f(\mathrm{d}\mu, \mathrm{d}t) \tag{R}$$

[10] This is Borel-measurable; see Online Appendix I.1 for details.

$$\text{s.t.} \int_{y>t} U(\mu, y) f(\mathrm{d}\mu, \mathrm{d}y) \geq U\left(\int_{y>t} \mu f(\mathrm{d}\mu, \mathrm{d}y), t\right), \ \forall t \in T^\circ, \tag{OC-C}$$

where U is extended from $\Delta(\Theta)$ to $\mathbb{R}_+^{|\Theta|}$ homogeneously of degree 1 by defining $U(\mu, t) := \sum_\theta \mu(\theta) \cdot U\left(\frac{\mu}{\sum_\theta \mu(\theta)}, t\right)$ and $T^\circ := T \setminus \sup T$.[11]

2.3 Strong duality & first-order characterization

Since (R) is a canonical constrained optimization problem, we solve it using the method of Lagrange multipliers. Define the Lagrangian $\mathcal{L} : \Delta_{\mu_0} \times \mathcal{B}(T^\circ) \to \mathbb{R}$ as follows:[12]

$$\mathcal{L}(f, \Lambda) := \int V(\mu, \tau) f(\mathrm{d}\mu, \mathrm{d}\tau) + \int_{t \in T^\circ} \left(\int_{\tau > t} U(\mu, \tau) f(\mathrm{d}\mu, \mathrm{d}\tau) - U\left(\int_{\tau>t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau), t\right)\right) \mathrm{d}\Lambda(t).$$

Note that (R) is equivalent to $\sup_f \inf_\Lambda \mathcal{L}(f, \Lambda)$. We now make a series of regularity assumptions that are fairly straightforward and are satisfied in virtually all applications; we discuss them in Remark 2 at the end of this section.

Assumption 1 (Information is strictly valuable)**.** $\forall t \in T$, $\mathbb{E}_{\mu_0}[U(\delta_\theta, t)] > U(\mu_0, t)$.

Assumption 2 (Continuity)**.** *$U(\mu, t)$ is continuous in (μ, t), $\{V(\mu, t)\}_{\mu \in \Delta(\Theta)}$ is equicontinuous in t, and $V(\mu, t)$ is upper-semicontinuous in (μ, t).*

Lemma 2. *Given Assumptions 1 and 2, strong duality holds:*

$$\sup_f \min_\Lambda \mathcal{L}(f, \Lambda) = \min_\Lambda \sup_f \mathcal{L}(f, \Lambda). \tag{D}$$

Furthermore, the min is achieved by $\Lambda^* \in \mathcal{B}(T^\circ)$, the max is achieved by $f^* \in \Delta_{\mu_0}$.

Proof. See Online Appendix I.3. *Q.E.D.*

Strong duality ensures that solving the constrained optimization problem (R) is equivalent to solving the unconstrained optimization problem $\sup_f \mathcal{L}(f, \Lambda)$. With it, we are ready to characterize solutions to the dynamic information design problem (P). Define the "derivative" of the Lagrangian with respect to f at a specific (μ, t) pair as:

$$l_{f,\Lambda}(\mu, t) := V(\mu, t) + \Lambda(t) \cdot U(\mu, t) - \int_{\tau<t} \nabla_\mu U(\widehat{\mu}_\tau, \tau) \mathrm{d}\Lambda(\tau) \cdot \mu,$$

[11] The HD-1 extension of U is inconsequential, but serves as a useful modeling tool that eliminates the need for normalization when evaluating conditional distributions. A useful implication of the HD-1 extension is that $\forall \mu \in \Delta(\Theta)$, $k > 0$, $\nabla_\mu U(\mu, t) = \nabla_\mu U(k\mu, t)$, i.e. $\nabla_\mu U(\cdot, t)$ is HD-0.

[12] $\mathcal{B}(T^\circ)$ is the set of positive Borel measures on T°.

where $\widehat{\mu}_t = \mathbb{E}_{f(\mu,\tau)}[\mu|\tau > t]$ and $\Lambda(t) = \int_{\tau<t} \mathrm{d}\Lambda(\tau)$. Note that $\nabla_\mu U(\widehat{\mu}_t, t)$ is not a well-defined vector in two scenarios: $U(\cdot, t)$ has a kink at $\widehat{\mu}_t$ or $t \geq \bar{t}$. With a slight abuse of notation, we let $\nabla_\mu U(\widehat{\mu}_t, t)$ denote a selection of sub-gradients in $\nabla_\mu U(\widehat{\mu}_t, t)$ in the former case and a selection of sub-gradients in $\nabla_\mu U(0, t)$ in the latter case.

Theorem 1. *f solves (R) if (f, a, Λ) and a selection of sub-gradients satisfy (FOC):*

$$l_{f,\Lambda}(\mu, t) \leq a \cdot \mu, \text{ with equality on the support of } f, \tag{FOC}$$

(OC-C), and the complementary slackness condition $\mathcal{L}(f, \Lambda) = \mathbb{E}_f[V]$.

Conversely, if Assumptions 1 and 2 hold, then $\forall f$ solving (R), there exists Λ solving (D), $a \in \mathbb{R}^{|\Theta|}$ and a selection of sub-gradients such that (FOC) holds.

Proof. See Appendix A.1. *Q.E.D.*

Theorem 1 gives a sufficient and near-necessary first-order characterization of optimality. The sufficiency part is general; the necessity part relies on strong duality. Equation (FOC) can be thought of as a "concavification" condition. It states that the derivative of Lagrangian $l_{f,\Lambda}$ touches its upper supporting hyperplane $a \cdot \mu$ only on the support of the optimal distribution f. The function $l_{f,\Lambda}$ that is concavified is a combination of the principal's utility V, the agent's utility U and an aggregation of agent's past utilities. Conditioned on stopping at time t, the agent's static incentives are captured by the direct benefit V. Moreover, since the agent's incentives are intertemporally linked, there is a shadow benefit of $\Lambda \cdot U$ from relaxing the agent's constraints up to time t, as well as a shadow cost $\int_{\tau<t} \nabla_\mu U(\widehat{\mu}_\tau, \tau)\mathrm{d}\Lambda(\tau) \cdot \mu$ from affecting all past continuation beliefs. Thus, the principal's dynamic persuasion strategy is effectively a mean-preserving spread of the prior distribution onto both the belief and time dimensions, internalizing the agent's incentives into the principal's utility.

Theorem 1 provides a simple recipe for analytically solving (R): The key unknown variable to be solved is Λ, a one dimensional function. For a given Λ, one can solve the period-by-period optimization problem by choosing the stopping belief μ (as a function of Λ) to maximize $l_{f,\Lambda}$. Then, the concavification condition implies that $l_{f,\Lambda}$ is "flat" across periods at the optimal μ's, leading to a (differential) equation characterizing only Λ. The recipe is illustrated in detail in Sections 4.1 and 4.2, where we derive a closed-form solution of (FOC) in a dynamic persuasion model with binary states. Based on Theorem 1, we develop an efficient algorithm for solving (R) numerically in Online Appendix II. The algorithm reduces to $|\Theta \times T|$-dimensional gradient descent.

Remark 2. We briefly discuss Assumptions 1 and 2. Assumption 1 is a weak regularity condition that is fulfilled whenever information is valuable at μ_0 for the agent. Assumption 2 is a continuity assumption: the indirect utility U is generically continuous per the maximum theorem while V is not typically continuous. Hence, we only require time-continuity of V but permit discontinuity with respect to beliefs. The upper-semicontinuity assumption is standard for deriving the existence of a maximum. It is needed for Lemma 2 only for the existence of f^*.

Remark 3. We develop several special cases and extensions of our baseline model in Online Appendix VI. First, we offer a unified discussion of how optimal dynamic persuasion varies as the principal and agent's preferences are aligned/misaligned on the belief and time dimension. Second, we show how our framework can be extended to handle settings in which (i) the payoff relevant state is evolving; or (ii) public information about the state arrives over time so the principal is no longer the monopoly provider of information. Both settings can be readily nested in our model via a novel transformation technique that redefines the underlying space of feasible stopping belief and stopping time. Finally, we discuss how to incorporate informational frictions into our model.

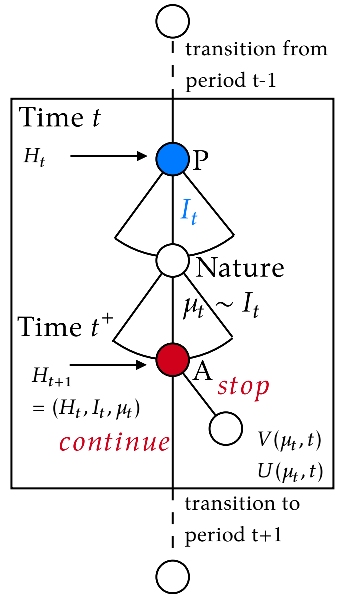
3 The Role of Intertemporal Commitment

Our framework assumes the principal chooses the agent's belief martingale. Implicit in this formulation is that the principal has full commitment power. We now relax this by recasting the dynamic persuasion problem (P) as an *experiment-selection game*: at each history, the principal chooses a one-period statistical experiment and, observing the principal's choice and the realization of the experiment, the agent decides whether to stop and act (game ends) or continue (game transitions to the next period). This will allow us to precisely pin down the role and value of intertemporal commitment in dynamic persuasion problems. We begin by defining the stochastic game.

Primitives Consider a finite time space $\mathcal{T} = 1, 2, 3, \dots \overline{T}$. The rest of the environment is as in Section 2: $U, V : \Delta(\Theta) \times \mathcal{T} \to \mathbb{R}$ are the agent's and principal's indirect utilities from stopping at belief μ at time T, respectively. $\mu_0 \in \Delta(\Theta)$ is the common prior belief. Let $\Gamma = (\mathcal{T}, U, V, \mu_0)$ denote the collection of primitives.

Extensive form A history H_t at time $t \in T$ is a profile of past belief-experiment pairs $H_t := (\mu_s, I_s)_{s<t}$, where belief $\mu_s \in \Delta(\Theta)$ and experiment $I_s \in \Delta^2_{\mu_{s-1}}(\Theta) := \{I \in \Delta^2(\Theta) | \mathbb{E}_I[\mu] = \mu_{s-1}\}$.[13] The null history is $H_1 = (\mu_0)$. Let $\mathcal{H}$ be the set of all possible histories. The timing of the game is as follows.

At each time $t < \overline{T}$ and history $H_t := (\mu_s, I_s)_{s<t}$, the period t is divided into two subperiods t and t^+.



(i) **Experiment selection.** At time t, the principal chooses an experiment, i.e., a random belief distribution $I_t \in \Delta^2_{\mu_{t-1}}(\Theta)$. Then, the history transitions stochastically to $H_{t+1} := (H_t, (\mu_t, I_t))$ where $\mu_t \sim I_t$.

(ii) **Stopping choice.** At time t^+ and history H_{t+1}, the agent chooses from $\{stop, continue\}$. If the agent stops, the game ends and the agent and principal's payoffs are $U(\mu_t, t)$ and $V(\mu_t, t)$. Otherwise, the game proceeds to period $t + 1$.

In the last period $t = \overline{T}$, the game ends when $\mu_{\overline{T}}$ is realized and payoffs are $U(\mu_{\overline{T}}, \overline{T})$ and $V(\mu_{\overline{T}}, \overline{T})$.

Our experiment-selection game is one of complete information where the state θ is observable to neither player and the principal's choice of statistical experiment as well as their realizations are commonly observable.[14] We have chosen this formulation to cleanly separate static commitment—the principal's inability to commit to a randomized message—and intertemporal commitment—the principal's inability to stick with a strategy prescribed in the past.[15]

Strategies A (behavioral) strategy for the principal is the map $\sigma_P : H_t \mapsto \Delta^2_{\mu_{t-1}}(\Theta)$. That is, it maps histories to a single static experiment. Restricting to pure strategies is

[13] $\Delta^2(\Theta)$ denotes the set of probability measures on the space $\Delta(\Theta)$.

[14] This builds on the observation that static persuasion (Kamenica and Gentzkow, 2011) can be decomposed into (i) commitment to acquire an experiment; and (ii) commitment to disclosure of the experiment, neither of which requires the sender to know the state. This is distinct from 'Wald persuasion' studied by Henry and Ottaviani (2019) (see also Morris and Strack (2019)) where the sender has no control over the flow information process but can control the stopping time. We discuss the timing of uncertainty resolution in Section 3.2.

[15] This is in the same spirit as Doval and Skreta (2022)'s 'mechanism selection game' which similarly abstracts away from the designer's ability to credibly run the current mechanism (McAdams and Schwarz, 2007; Akbarpour and Li, 2020) to focus on the role of limited intertemporal commitment.

without loss here as experiments themselves involve randomization. Let Σ_P be the set of all principal strategies. A strategy for the agent is the map $\sigma_A : H_t \mapsto \Delta\{stop, continue\}$. Let Σ_A be the set of all agent strategies. For all $\sigma \in \Sigma := \Sigma_P \times \Sigma_A$, let $(\langle \mu(\sigma)_t \rangle, \tau(\sigma))$ denote the induced stochastic belief process and stopping time. Define the players' payoffs under profile σ as

$$\mathcal{V}(\sigma) := \mathbb{E}^{\sigma}\Big[V(\mu(\sigma)_{\tau(\sigma)}, \tau(\sigma))\Big];$$
$$\mathcal{U}(\sigma) := \mathbb{E}^{\sigma}\Big[U(\mu(\sigma)_{\tau(\sigma)}, \tau(\sigma))\Big].$$

Let $\mathcal{V}(\sigma|H_t, t), \mathcal{U}(\sigma|H_t, t)$ denote the players' payoffs under profile σ in the subgame starting from history H_t in period t. Let $\mathcal{V}(\sigma|H_{t+1}, t^+), \mathcal{U}(\sigma|H_{t+1}, t^+)$ denote the players' payoffs under profile σ in the subgame starting from history H_{t+1} in subperiod t^+.

Equilibria Since the game is one of complete information, we focus on subgame perfection (*SPE*). Applying the one-shot deviation principle, a strategy profile $\sigma \in \Sigma$ is an *SPE* if for all times $t \in T$ and all histories $H_t, H_{t+1} \in \mathcal{H}$,

$$\mathcal{V}(\sigma|H_t, t) = \max_{I'} \mathbb{E}_{I'(\mu)}[\mathcal{V}(\sigma|(H_t, \mu, I'), t)];$$
$$\mathcal{U}(\sigma|H_{t+1}, t^+) = \max\big\{U(\mu_t, t), \mathcal{U}(\sigma|H_{t+1}, t+1)\big\};$$

Given the nature of our principal-agent setting, it is instructive to consider a special class of SPE that selects equilibria in favor of the principal.

Definition 2. *Given Γ, $\sigma \in \Sigma$ is a* ***principal-optimal SPE*** *(in shorthand, SPE^P), if for all histories $H_t \in \mathcal{H}$,*

$$\mathcal{V}(\sigma|H_t, t) = \max_{\sigma' \in SPE(\Gamma)} \mathcal{V}(\sigma'|H_t, t).$$

SPE^P requires a favorable selection of *SPE* not only at the null-history, but also at any history of the game. We view this as innocuous and simply ensures that the agent tiebreaks in favor of the principal.[16] The existence of SPE^P is not generally guaranteed in finite games—a favorable selection of *SPE* in the future might be suboptimal for the period 1 principal. Nevertheless, we will show that SPE^P exist and their outcomes correspond to that of the full-commitment solutions we derived in Section 2. Finally, we introduce an additional assumption on the principal's preference for delay.

[16] Note that in the principal's problem under commitment (Section 2), we also assumed the agent tiebreaks in favor of the principal.

Assumption 3. *For all beliefs μ, $V(\mu, t)$ (weakly) increases in t.*

Assumption 3 simply states that the principal weakly enjoys delay. It is satisfied in all applications of this paper and all special cases of our framework in the literature. Although this is natural and is fulfilled in many environments, it plays an important role which we illustrate in Section 3.2.

Theorem 2. *If $\mu_0 \in \Delta(\Theta)$, finite T, and U, V satisfying Assumptions 2 and 3 constitute Γ, then $SPE^P(\Gamma) \neq \emptyset$. Moreover,*

(i) ***Equilibrium outcome:*** *If finite support f solves (R), then $\exists \sigma \in SPE^P(\Gamma)$ such that $(\mu(\sigma)_{\tau(\sigma)}, \tau(\sigma)) \sim f$.*

Conversely, if $\sigma \in SPE^P(\Gamma)$, then $f \sim (\mu(\sigma)_{\tau(\sigma)}, \tau(\sigma))$ solves (R).

(ii) ***Equilibrium strategy:*** *If finite support $(\langle \mu_t \rangle, \tau)$ solves (P) and for any stopping time $\iota \leq \tau$, $\mathbb{E}[U(\mu_\tau, \tau)] = \mathbb{E}[U(\mu_\iota, \iota)]$, then $\exists \sigma \in SPE^P(\Gamma)$ s.t. $(\langle \mu_t \rangle, \tau) \stackrel{d}{=} (\mu(\sigma), \tau(\sigma))$.*

Conversely, if $V(\mu, t)$ strictly increases in t, then $\forall \sigma \in SPE^P(\Gamma)$ s.t. $\mu(\sigma)$ has finite support and $\max \operatorname{supp}(\tau(\sigma)) < \overline{T}$, *for any stopping time $\iota \leq \tau(\sigma)$, $\mathcal{U}(\sigma) = \mathbb{E}[U(\mu_\iota, \iota)]$.*

Proof. See Appendix A.2. *Q.E.D.*

Theorem 2 (i) establishes that there is no gap in *outcomes* whether or not the principal has intertemporal commitment. In particular, the principal's commitment payoff—in which she first commits to a dynamic persuasion strategy, then the agent best responds by optimally stopping—coincides with her payoff under every principal-optimal SPE. Thus, Theorem 2 (i) partially extends the revelation principle we developed in Section 2 to the setting with limited commitment: it is sufficient to consider only simple recommendations and solve (R) to characterize equilibrium outcomes of the SPE.

Theorem 2 (ii) establishes a crucial property of principal-optimal subgame perfect equilibria: an optimal dynamic persuasion strategy solving the original problem (P), which leaves the agent with zero surplus at any interim stage is sufficient and almost-necessary for it to correspond to the equilibrium path of an SPE^P. This suggests that simple recommendations, while helpful for reducing our original problem (P) to the more tractable problem (R), might not be implementable under limited commitment—indeed, they are often not. The intuition behind Theorem 2 (ii) is as follows. There are three types of potential deviations in the game:

1. Agent deviates from *continue* to *stop*. Such deviations have already been ruled out by the obedience constraint (OC) in the principal's problem (P) under commitment.

2. Principal deviates from the ex-ante optimal persuasion strategy. Such deviations cannot be profitable at an interim belief if the agent gets no surplus: if such deviation strictly benefits the principal and the agent responds by *continue*, i.e. it leaves the agent with non-negative surplus, then such a deviation is feasible and strictly profitable at the ex-ante stage. Conversely, if the agent gets any surplus at an interim belief, a principal that strictly enjoys delay would like to further delay the revelation of information to further extract this surplus.

3. Agent deviates from *stop* to *continue*, tempting the principal to restart the revelation of information. Such deviations are ruled out (only) when the principal and agent have a conflict of interest regarding delay: following the deviation, the principal responds by recommending the agent to *stop* in the next period while releasing no further information, making the patient principal better off and the agent worse off. Such a response is credible because if the principal can make herself even better off by releasing more information while the agent still finds it profitable to deviate in the first place, this joint deviation strictly improves the optimal full-commitment solution, which is impossible.

Theorem 2 (ii) can be interpreted as an "**anti-revelation principle**": simple recommendations that send minimally informative direct messages are often *not* consistent with subgame perfection. Instead, the key to restoring commitment is to strategically send indirect messages that are "unnecessary"—per the logic of the revelation principle—under full commitment. This delivery of unnecessary interim information exploits the simple but crucial observation that *information is irreversible*—once the agent learns, there is no way for the principal to systematically reverse this. This then pushes up the agents' outside option, which guards the principal against her own future deviations. Interestingly, there are general and economically interpretable conditions under which the "revelation principle" continues to hold in terms of the equilibrium strategy. There are also conditions under which the "anti-revelation principle" manifests. We provide such conditions in Sections 4.2 and 4.3.

In light of Theorem 2 (ii), we abuse terminology slightly and call an optimal strategy $(\langle \mu_t \rangle, \tau)$ "*dynamically consistent*" if it leaves no surplus to the agent at every positive-probability history:[17]

$$\text{for all stopping times } \iota \leq \tau, \ \mathbb{E}[U(\mu_\tau, \tau)] = \mathbb{E}[U(\mu_\iota, \iota)].$$

Practically, Theorem 2 delivers a straightforward process for constructing dynamically consistent strategies. First, derive the simple recommendation strategy which solves Equation (R). By Theorem 2 (i), the optimal joint distribution over beliefs and stopping times coincides with that in SPE^P. Then, we modify this simple recommendation by replacing direct recommendations with "indirect ones". This induces stochastic interim beliefs which pushes up the agents' endogenous "outside option" at interim histories.[18] The construction will be illustrated in Section 3.1, where we convert the dynamically inconsistent "moving the goalposts" strategy (Ely and Szydlowski (2020)) to a dynamically consistent strategy that replicates the same outcome.

3.1 Application: "Inching" or "Teleporting" the Goalposts?

We revisit the model introduced by Ely and Szydlowski (2020), where a principal dynamically persuades an agent to take on difficult tasks. We begin by recasting their model in our framework. The task has an unknown difficulty x that equals the time it takes to complete. The principal conducts a sequence of experiments that reveals the difficulty.[19] Let $\theta = x$, the effort threshold for success. c is the flow cost of effort. R is the reward to the agent upon completion. The agent and principal discount the future at rates r and r_p, respectively. Translated into our framework, the agent's indirect utility U is

$$U(\mu, t) = \sup_{s \geq t} \mathbb{E}\left[\mathbf{1}_{s \geq x} e^{-rs} R - c(1 - e^{-rs})\right].$$

Let $s(\mu, t)$ denote the largest maximizer; the principal's indirect utility is

$$V(\mu, t) = \mathbb{E}\left[1 - e^{-r_p s(\mu, t)}\right].$$

[17] This is a slight abuse of terminology because we will use this terminology also for settings where all other conditions of Theorem 2 hold but T is a continuum. In such cases, the extensive form game is not even well-defined.

[18] A special case of this result was obtained in previous work (Koh and Sanguanmoo, 2024) which focuses on the case where the agent is impatient. This excludes many environments of interest, such as that of "Moving the Goalposts" (Ely and Szydlowski, 2020). By contrast, Theorem 2 holds far more generally.

[19] Ely and Szydlowski (2020) assumes that the principal observes x at the beginning, which is different from our experimentation interpretation. Such difference is inconsequential in their setting with full dynamic commitment but is crucial when commitment is limited. We argue that our experiment-selection game fits the corporate succession/apprenticeships context of Ely and Szydlowski (2020): the skill level of a worker/apprentice is often not directly observed but can be tested by the manager/master.

In what follows, we study the binary task difficulty case, where $\Theta = \{x_l, x_h\}$, to illustrate the role of intertemporal commitment. We assume that (i) the agent is willing to complete the difficult task when he has already worked for x_l periods and (ii) the prior belief is pessimistic enough that, sans additional information, the agent stops immediately.[20] We now introduce two vastly different strategies that optimally persuade the agent when the principal has full commitment:

- **Teleporting the goalposts**: Ely and Szydlowski (2020), Proposition 3 shows that it is optimal to "move the goalposts":

  > *The principal provides disclosures at no more than two dates. At date zero, the disclosure is designed to maximize the probability that the agent begins working, and at date x_l, the disclosure is designed to maximize the probability that the agent continues to x_h.*

  The strategy is illustrated numerically in Figure 1.[21] The red (blue) bars in the top panel depict the probability of revealing x_h (x_l) at different points in time: with some probability, the strategy reveals x_h with certainty at $t = 0$, leading the agent to give up. Absent the initial revelation, the agent becomes just optimistic enough that the task is easy (with belief $\bar{x}$) to begin working until the true difficulty is fully revealed at $t = x_l$. We plot the agent's expected task difficulty $\mathbb{E}[x]$ (termed the "goalposts") on the bottom panel: it is initially set at $\tilde{x}$ and abruptly "teleported" to x_h at $t = x_l$ if the task is difficult. Call this policy *teleporting the goalposts*.

- **Inching the goalposts**: following the initial revelation at the beginning of the game, conclusive information that the task is easy is revealed at a Poisson rate to the agent such that it keeps him just willing to work. Then, the agent's posterior belief absent the revelation gradually reaches the true difficulty at $t^* = x_h - \bar{\tau}$, i.e., when she is indifferent between working to achieve the difficult task and stopping right away.

  In Figure 2, we numerically illustrate the solution that *inches the goalpost*, with the same color code as in Figure 1. For times $t \in [0, t^*]$, the fact that the task is easy (x_l) is revealed gradually at a decreasing Poisson rate with pdf depicted by the blue curve (not to scale). Hence, in the absence of evidence that the task is easy, the agent grows

[20] Formally, we maintain the parametric assumptions in Ely and Szydlowski (2020), Section IV that justifies "moving the goalposts": (i) $\bar{\tau}$ satisfies $e^{-r\bar{\tau}}R - c(1 - e^{-r\bar{\tau}}) = 0$ and $x_h - x_l \leq \bar{\tau}$ and (ii) $\mu_0(x_h) > \bar{\mu} = \frac{c}{R}\frac{1-e^{-rx_l}}{e^{-rx_l}}$.

[21] The parameters are set to $x_l = 1$, $x_h = 2$, $R = 2.5$, $c = 1$, $r_p = 0.8$, and $r = 1$. The prior belief is $(0.2, 0.8)$.

gradually more pessimistic that the task is hard. The corresponding goalpost $\mathbb{E}[x]$ increases gradually from $\bar{x}$ to x_h, keeping OC-C binding at all histories.

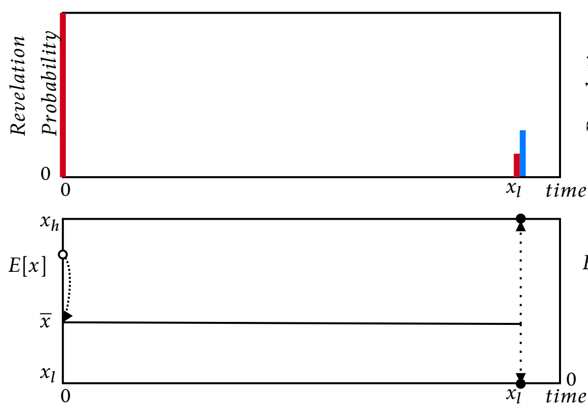


Figure 1: Teleporting the goalpost

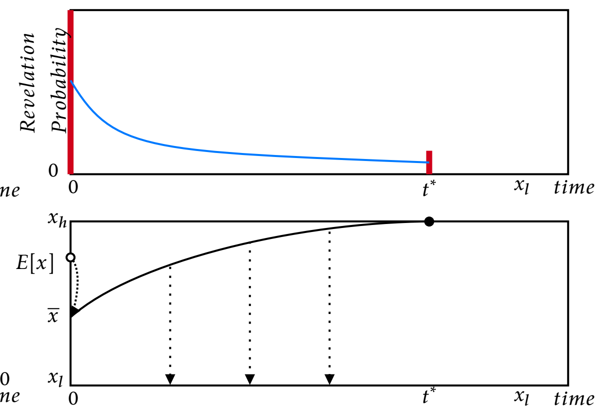


Figure 2: Inching the goalpost

The two policies in Figures 1 and 2 induce the same distribution over the agent's efforts and are thus outcome-equivalent.[22] However, we argue that in environments where intertemporal commitment is unnatural and cannot be taken for granted, how the goalpost is moved is crucial. As pointed out by Ely and Szydlowski (2020), the teleporting policy is dynamically inconsistent: right before x_l, the principal is tempted to renege on her original strategy of disclosing the true difficulty of the task to the agent since, by doing so, the agent would always work until x_h.[23] By contrast, the *inching* strategy is dynamically consistent: the agent is kept indifferent between stopping and continuing at any history, which does not allow the principal to extract excess effort.

Restoring dynamic consistency. We now show how the inching strategy can be systematically obtained via a constructive procedure. Consider a discrete-time version of the model where the time space is the finite set $T = \{t_i\}$.[24] Let μ denote the probability that the task is hard ($x = x_h$). First observe that *teleporting the goalposts* implements the

[22] This is because early revelation of x_l under the *inching* strategy still induces the agent to exert effort up to time $t = x_l$ to complete the simple task. The revelation of x_h at t^* also induces the agent to complete the difficult task.

[23] Indeed, Ely and Szydlowski (2020), discussing the relation to Orlov, Skrzypacz, and Zryumov (2020), write "The key difference is that in our setting, the principal has commitment power, while their sender cannot make use of promised disclosures. Consequently, we obtain entirely different policies." While Ely and Szydlowski (2020)'s model of observable x requires even stronger commitment to implement, note that the reneging behavior results from the lack of intertemporal commitment, which manifests in the experiment-selection game.

[24] We assume that T includes the several key times: $\{0, t^*, x_l, x_h\} \subset T$. As a result, the optimal belief-time distribution in the continuous time model remains optimal here.

same outcome by fully revealing the task difficulty at any time between t^* (where the hard task is just within reach) and x_l (where the agent stops without further information). Consider a policy that fully reveals the task difficulty at $t = t^*$. The outcome f is represented by the solid black dots in Figure 3-(a). Beliefs are depicted on the vertical axis, and time is depicted on the horizontal axis. We use $t_{-1} = t^*$ to denote the last period information is revealed. The teleporting strategy first splits the belief from μ_0 to μ_1 and 1 in the first period. Then, belief stays at μ_1 until t_{-2} and, at t^*, the task difficulty is fully revealed. This is depicted by the gray path in Figure 3 (a).

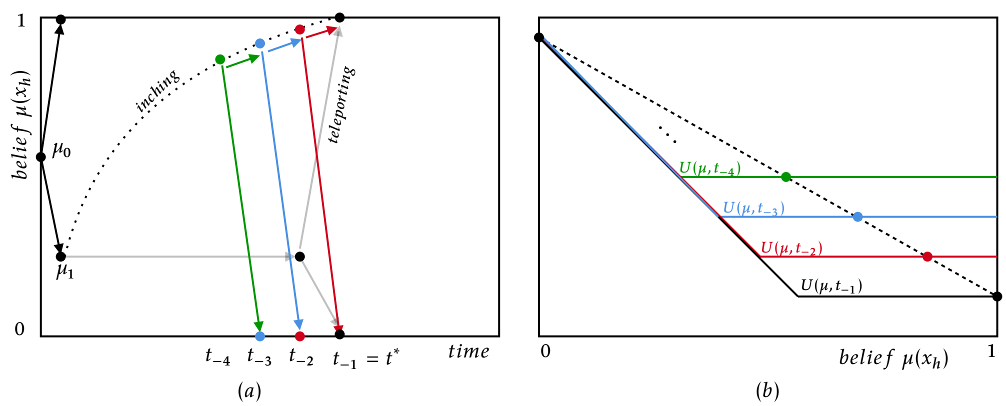


Figure 3: Construction of a dynamically consistent strategy.

As suggested by Theorem 2 (ii), a necessary and sufficient condition to restore dynamic consistency is to raise the agent's stopping payoff at interim beliefs. To achieve this, we work recursively from the last period the agent receives information under the teleporting strategy. The agent's indirect utility in the last period t_{-1} is depicted by the black curve in Figure 3 (b). The sloped segment corresponds to the recommendation "work until $t = x_l = 1$" and the horizontal segment corresponds to "stop working".[25] Now consider any interim belief in the second-last period t_{-2}. First, observe that the expected payoff from revealing the state next period contingent on the period t_{-2} belief is given by the corresponding point on the black dotted line segment connecting $U(0, t_{-1})$ and $U(1, t_{-1})$. On the other hand, the expected payoff from stopping at t_{-2} is given by $U(\cdot, t_{-2})$, the red curve. Therefore, to eliminate the surplus from continuing, the (in this case, unique) interim belief is found by crossing the dashed line and the red curve, leading to the red dot that pins down beliefs at time t_{-2}.

[25] Recall that by construction, at $t_{-1} = t^*$ and with belief $\mu = 1$, the agent is indifferent between stopping immediately and working until $t = x_h = 2$.

Note that in this example, since the payoff from "work until 1" is time-invariant, the red curve coincides with the black curve for beliefs near 0. Therefore, we can also find a second red dot on the left which simply coincides with the black dot that equals the belief $\mu = 0$, i.e., the knowledge that the task is easy. We plot both red dots at time t_{-2} in Figure 3 (a). This fully pins down the information revealed in the last period: split the red dots into 0 and 1, as is depicted by the red arrows in Figure 3 (a).

Similarly, the interim belief in period t_{-3} is found by equating the segment connecting the dotted segment and the blue curve $U(\cdot, t_{-3})$, leading to the blue dots in Figure 3 (b) and the corresponding blue arrows in Figure 3 (a) depict the information structure. The same can be done for the green dots at time t_{-3} and we can recursively run this process backwards in time until we reach period t_1. The resulting strategy is exactly *inching the goalpost*. Then, Theorem 2 (ii) implies that this dynamically consistent strategy is the equilibrium path of an SPE^P of the corresponding experiment selection game. We relegate the full specification of the equilibrium to Online Appendix III.

We have illustrated how commitment can be restored in the canonical setting of (Ely and Szydlowski, 2020). Our construction applies to more general persuasion environments, but proceeds similarly: starting from the last time information is delivered under the optimal simple recommendation, we iteratively deliver interim information so that the agent's surplus is eliminated at all histories.

3.2 Discussion

Formulation of the extensive form game It is well-understood that in dynamic games of incomplete information, different extensive forms (corresponding to the same strategic form) have strong implications on the power of subgame perfection. We have chosen this particular timing to eliminate all information sets, i.e. our game is one of complete information so that every history corresponds to a proper subgame.[26]

No continuity gap Fix a continuous time space T. Consider a discretization of the time space $\widehat{T} = \{t_1, \dots, t_J\}$. Let $d_{\widehat{T}} = \max\{t_{j+1} - t_j\}$ denote the grid size of $\widehat{T}$. Let $\Delta_{\mu_0}(\widehat{T}) = \Delta_{\mu_0} \cap \Delta(\Delta(\Theta) \times \widehat{T})$ denote the set of feasible belief-time distributions supported on $\widehat{T}$. The following lemma shows that there is no continuity gap as the grid size converges

[26] An alternative formulation of the model is to have nature choosing a persistent θ first, and having the principal choosing statistical experiments about θ (observing past signal realizations but not θ). However, this approach leads to a game without any proper subgame and subgame perfection has no bite. While it is beyond the scope of this paper to fully investigate the implication of different extensive forms, we conjecture that the exact formulation of the extensive form game is inconsequential under the solution concept of sequential equilibrium.

to 0.

Lemma 3. *Given Assumptions 1 and 2, $\forall f \in \Delta_{\mu_0}$ that satisfies (OC-C), $\forall \epsilon > 0$, $\exists \delta > 0$ s.t. $\forall \widehat{T}$ with $d_{\widehat{T}} \leq \delta$, there exists $\widehat{f} \in \Delta_{\mu_0}(\widehat{T})$ that satisfies (OC-C) and $d_{lp}(f, \widehat{f}) < \epsilon$.*

Conversely, for any sequence of $(\widehat{T}^n)$ with $d_{\widehat{T}^n} \to 0$ and $\widehat{f}^n \in \Delta_{\mu_0}(\widehat{T}^n)$ that satisfies (OC-C), there exists $f \in \Delta_{\mu_0}$ that satisfies (OC-C) and a subsequence $(\widehat{f}^{n_k})$ s.t. $\widehat{f}^{n_k} \xrightarrow{w} f$.

Proof. See Online Appendix I.4. *Q.E.D.*

As a direct corollary of Lemma 3 and Theorem 2, the principal-optimal payoff in continuous-time can be approximated by SPE^P of the experiment-selection games when the grid size converges to 0, and conversely, the sequence of SPE^Ps of the experiment-selection games with converging grid size have an accumulation point that solves Equation (R) in continuous-time.

Necessity of a patient principal and an informational Coase conjecture Theorem 2 states that when the principal enjoys delay, the agent's deviation from *stop* to *continue* cannot induce the principal to reveal more information. We now give a simple sufficient condition under which the principal's impatience leads the SPE^P to consist of full-revelation at time $t = 1$, which is clearly suboptimal under full commitment.

Proposition 1. *Given $\mu_0 \in \Delta(\Theta)^\circ$, finite T and U, V satisfying Assumption 2. Suppose also:*

(i) $\forall \theta \in \Theta$, $U(\delta_\theta, t)$ is constant in t and $V(\delta_\theta, t)$ strictly decreases in t,
(ii) For every $\mu \in \Delta(\Theta)^\circ$, $\mathbb{E}_{\theta\sim\mu}[V(\delta_\theta, \overline{T})] > V(\mu, \overline{T})$.

Then, in the principal's preferred SPE, she fully reveals θ in $t = 1$.

Proof. See Online Appendix I.5. *Q.E.D.*

Condition (i) of Proposition 1 states that contingent on knowing the state, the agent is patient while the principal is strictly impatient. Condition (ii) states that the principal eventually prefers to reveal the state. Proposition 1 states that if the principal cannot commit to staying silent, her temptation to reveal the state in the future forces her to reveal the state right away. This might be viewed as a simple informational analog of the "Coase conjecture":[27] just as a monopolist who cannot commit to a time-path of prices must compete with her future self, so does our principal who cannot commit

[27] See, e.g., Coase (1972); Bulow (1982); Stokey (1979).

to staying silent. Indeed, for non-degenerate interim beliefs, the agent finds it optimal to deviate from *stop* to *continue* until $\overline{T}$, at which point the principal finds it optimal to reveal the state. But anticipating this, and because the principal dislikes delay, she fully reveals information immediately. We discuss Proposition 1 further in Online Appendix IV.

Decoupling different kinds of commitment Taken together, Theorem 2 and Proposition 1 offer a rich picture of dynamic persuasion games when the principal has the ability to commit to a static experiment, but lacks the ability to intertemporally commit. This complements a recent literature on strategic communication analyzing how the principal's payoff varies as her ability to commit to a static experiment diminishes (Lipnowski et al., 2022). Indeed, we view these as quite distinct kinds of commitment. Our contribution is to make precise when and how the principal can overcome her lack of intertemporal commitment to attain her commitment payoff (Theorem 2), and when she cannot (Proposition 1).

4 Persuasion v.s. Delay: The Trade-offs

We have developed a unified and tractable approach for solving dynamic persuasion problems. We now employ this to characterize the form of dynamic persuasion in environments where the principal would like to both *delay* the agent's decision, as well as influence it via *persuasion*.

There is, of course, a natural tension between these goals: delay entails delivering valuable future information, while persuasion necessarily garbles/degrades it. In this section we make these tradeoffs precise by developing a unified analysis of how (i) complementary vs substitutability of persuasion vs delay; and (ii) general properties of the principal and agent's time preference shape the optimal dynamic information policy.

In Sections 4.1 and 4.2 we first analyze a canonical binary-state, binary-action environment in which the principal has general preferences over the agent's action and stopping time. This offers a simple and transparent environment for us to analyze the key tradeoffs, contrast predictions against the well-known static solution (Kamenica and Gentzkow, 2011), and highlight distinctive implications of richer modelling along the time dimension. We will completely characterize optimal persuasion strategies there. Then, in Section 4.3, we establish that the main features from the binary environment hold more generally.

4.1 The binary environment

The state θ can be either L (left) or R (right). The common prior belief $\mu_0(R) \in (0, 0.5)$ and let $\mu_t := \mathbb{P}(\theta = R|\mathcal{F}_t)$ denote the agent's posterior belief at time t, conditional on observing the information provided by the principal. The agent takes a binary action ℓ or r. The agent wishes to match the state:

$$\text{Match: } u(L, \ell, t) = u(R, r, t) = 1 - c(t) \qquad \text{Mismatch: } u(L, r, t) = u(R, \ell, t) = -c(t)$$

Hence, the agent's indirect utility function is $U(\mu, t) = \max\{\mu, 1 - \mu\} - c(t)$. The principal's prefers that the agent stops later, but otherwise has arbitrary state-independent preferences over the agent's choice of action and stopping time:

$$\text{Choose } \ell\text{: } v_\ell + h_\ell(t) \qquad \text{Choose } r\text{: } v_r + h_r(t)$$

where h_l, h_r are strictly increasing and differentiable and we normalize $v_r \geq v_\ell$, $h_\ell(0) = h_r(0) = 0$, and $c(t) = t$ without loss.[28] Let $\Delta v := v_r - v_\ell$ and $\Delta h(t) := h_r(t) - h_\ell(t)$. Our framework is rich, and allows the (i) marginal delay gain i.e., h'_ℓ and h'_r; and (ii) marginal persuasion gain i.e., $h_r - h_\ell$ evolve in time.

We will analyze how these matter for the qualitative properties of optimal dynamic information. While our induced belief process in (P) can be arbitrary, two kinds of strategies will feature prominently:

1. **Suspense-generating**: The principal generates suspense by conducting a sequence of experiments that induces interim beliefs on one of two paths—the L path μ_t^L stays below 0.5 and the R path μ_t^R stays above 0.5. In each period, an experiment reveals a *plot twist*, i.e., belief jumps to the opposite path at an appropriate Poisson rate. The absence of the plot twist *confirms* the current belief, leading the agent's belief to drift along the current path. This is depicted in Figures 4 and 5 for times up to t_1.

2. **Action-targeting:** The principal conducts an experiment that reveals the L (resp., R) state at a Poisson rate. When L (resp., R) is revealed, the agent immediately stops and chooses ℓ (resp., r). The absence of further revelations induces the agent's belief to drift towards R (resp., L) gradually. This is depicted by Figure 4 (resp., 5) for times between t_1 and t_2.

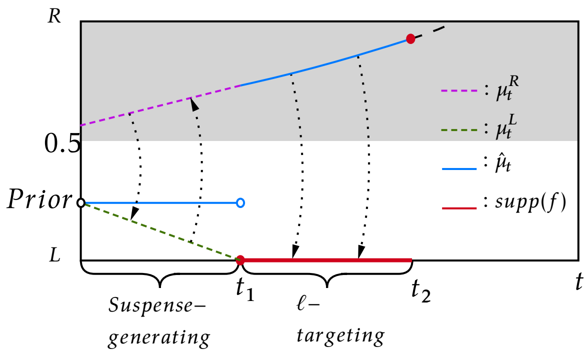


Figure 4: *Suspense-ℓ* strategy

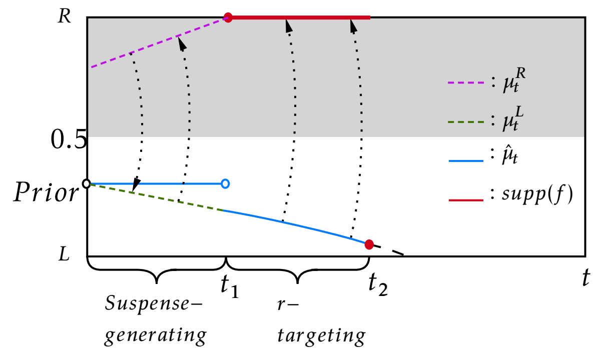


Figure 5: *Suspense-r* strategy

These strategies serve distinct but complementary roles; we will combine them to generate a rich class of dynamic persuasion strategies pinned down by t_1, $a \in \{\ell, r\}$, and t_2 as follows:

- **Suspense**^{t_1}-ℓ^{t_2} runs in two stages and is depicted in Figure 4. During the first stage $[0, t_1)$, the principal generates suspense with the two belief paths μ_t^L and μ_t^R.[29] The two paths are chosen such that the agent obtains no continuation surplus over the duration of suspense generation. At the end of the suspense stage $t = t_1$, the agent is either certain the state is L ($\mu_{t_1}^L = 0$) which induces stopping and choosing l, or $\mu_{t_1}^R = \frac{\mu_0}{2\mu_0 - t_1}$ which leads to the second stage.

  In the second stage $[t_1, t_2]$, the strategy reveals state L at a rate which is chosen such as to keep the agent's continuation incentive (OC-C) binding. This pins down a unique continuing belief path $\mu_{t_1}^*(t) := \mu_{t_1}^R e^{t - t_1}$. If the state L has not been revealed by time t_2, the agent stops and takes action r under belief $\mu_{t_1}^*(t_2) \in [0.5, 1]$.

- **Suspense**^{t_1}-r^{t_2}: runs in two stages and is depicted in Figure 5. As before, suspense is generated over $[0, t_1)$ with two belief paths chosen so that the agent has no continuation surplus.[30] At the end of the suspense stage $t = t_1$, the agent is either certain the state is R ($\mu_{t_1}^L = 1$) and stops and chooses r, or $\mu_{t_1}^L = \hat{\mu}_{t_1} = \frac{\mu_0 - t_1}{1 - t_1}$, which leads to the second stage. Then, over $[t_1, t_2]$ the strategy reveals state R, once again at a rate which is chosen such as to keep (OC-C) binding and pins done a unique

[28] Setting $v_r \geq v_\ell$ is wlog as otherwise we can flip the labels of the states. Setting $h_\ell(0) = h_r(0) = 0$ is wlog as otherwise we can adjust v_ℓ and v_r. $c(t) = t$ could be obtained by scaling the time space: $t \to c(t)$. Then, h_ℓ, h_r are converted to $h_\ell \circ c^{-1}, h_r \circ c^{-1}$.

[29] They are given by $\mu_t^L = \mu_0 \frac{t_1 - t}{t_1}$ and $\mu_t^R = \frac{\mu_0(1 - t_1 + t)}{2\mu_0 - t_1}$. Implicitly, we require $t_1 < \mu_0$ to guarantee that $\mu_{t_1}^R \in (0, 1)$.

[30] They are given by $\mu_t^L = \frac{(1 + t - t_1)\mu_0 - t}{1 - t_1}$ and $\mu_t^R = \frac{1 + t - (2 + t - t_1)\mu_0}{1 + t_1 - 2\mu_0}$ and depicted as the green and purple dotted lines in Figure 5.

continuing belief path $\mu^{\dagger}_{t_1}(t) := 1-(1-\mu^L_{t_1})e^{t-t_1}$. If the state R has not been revealed by time t_2, the agent stops and takes action l under belief $\mu^{\dagger}_{t_1}(t_2)$.

The first time t_1 represents the *duration* of suspense: as t_1 gets larger, the principal increasingly delays the beginning of revelation—leading the agent's stopping times to concentrate on the interval $[t_1, t_2]$ shifted further into the future. The action that is targeted represents the *direction* of persuasion, leading to different states being revealed over time. Finally, the second time t_2 represents the *scope* of persuasion: as t_2 gets larger, the posterior beliefs of the agent necessarily become more dispersed to compensate the agent for waiting, implying that more information is revealed throughout the process.

Instead of reducing the strategies to simple recommendations, we have leveraged our results from Section 3 to describe them in terms of indirect recommendations. This guarantees robust predictions irrespective of the level of intertemporal commitment.

4.2 Characterization of optimal strategy

Proposition 2. *Let* $\Psi(t,t_2) := \int_t^{t_2} e^{t-s}h'_\ell(s)ds + e^{t-t_2}h'_r(t_2)$ *and* $0 \le t_1 < t_2 < \mu^{*-1}_{t_1}(1)$. *If the* Suspense^{t_1}-ℓ^{t_2} *strategy is optimal, then the following conditions hold*

(a) ***FOC for*** t_2***:*** $\Delta v + \Delta h(t_2) = h'_r(t_2)$,

(b) ***FOC for*** t_1***:*** $\Psi(t_1,t_2) \ge h'_\ell(t_1)$ *with equality when* $t_1 > 0$,

(c) ***Local SOC:*** $h''_\ell(t_1)\cdot t_1 \le 0,\ h''_r(t_2) \le \Delta h'(t_2)$.

Conversely, if conditions (a)-(c) above hold and, in addition,

(d) ***Global SOC:*** $$\begin{cases} h''_\ell(t), h''_r(t) \le 0 & \forall t < t_1 \\ \max\{h''_r(t), 0\} \le \Psi_t(t,t_2) - h'_\ell(t) & \forall t \ge t_1, \end{cases}$$

then the Suspense^{t_1}-ℓ^{t_2} *strategy is optimal.*

Proof. See Online Appendix I.6. *Q.E.D.*

Proposition 2 provides a near complete characterization of the conditions that justify the suspense-ℓ policy. The FOCs in Proposition 2 pin down the three parameters t_1, t_2 and a and quantify three key trade-offs: (i) persuasion gain; (ii) time-risk preferences; and (iii) persuasion-delay complementarity vs substitutability. We discuss each in turn.

Persuasion gain determines the scope of persuasion. The scope of persuasion t_2 is pinned down by condition (a) of Proposition 2. In particular, the magnitude of Δv (the "time-0" persuasion gain) relative to $\Delta h(t)$ (additional persuasion gain at t) determines the location of t_2. If Δv is large relative to Δh, the principal would like to maximize the probability the agent chooses r, and so chooses an early terminal time (t_2 small) and a less dispersed terminal belief ($\mu^*_{t_1}(t_2)$ closer to 0.5). In the limit, this recovers the benchmark static persuasion policy which induces immediate stopping with posterior beliefs 0 and 0.5. Conversely, when the time-0 persuasion gain is small relative to the additional persuasion gain, the principal is willing to induce a later terminal time t_2 and compensate the agent for waiting by inducing a more dispersed terminal belief $\mu^*_{t_1}(t_2)$.

Time-risk attitudes determine the duration of suspense. The existence and duration of the suspense stage are tightly connected to the curvature of the principal's delay gain and is pinned down by condition (b) of Proposition 2. Conditional on having suspense, the duration is determined by $\Psi(t_1, t_2) = h'_\ell(t_1)$. $\Psi(t_1, t_2)$ can be viewed as the marginal loss from increasing the suspense window t_1 since the agent must be compensated by way of additional information over $[t_1, t_2]$. Comparative statics with respect to the curvature of h is straightforward, given our characterization. Fixing t_2 (which is pinned down by the persuasion gain) and $h'_\ell(t_1)$, suppose $h_\ell(t)$ gets more concave. Then, this implies that $h'_\ell(t)$ gets smaller for every $t > t_1$. This pushes down $\Psi(t_1, t_2)$ which implies the optimal suspense window t_1 becomes larger. In the extreme case where h_ℓ is strictly convex, the second-order condition (c) rules out any suspense.[31]

When and why does the principal prefer suspense? We have already seen that a time-risk loving principal with convex delay payoffs rules out suspense.[32] A simple sufficient condition for suspense to be optimal is concavity in delay payoffs h_ℓ, h_r so the principal is (relatively) time-risk averse.[33] Thus, the principal prefers a less dispersed stopping time which is precisely achieved by generating suspense: it keeps the agent engaged, but never certain enough to make a decision, which concentrates the agent's stopping times on the window $[t_1, t_2]$ shifted further into the future. This raises the

[31] We emphasize, however, that with h_ℓ strictly convex, ℓ-targeting with $t_1 = 0$ can still be optimal as long as the second condition of (d) is additionally satisfied.

[32] While the case of standard exponential discounting is not nested in our baseline model, the more general Proposition 5 rules out suspense generation under any time-risk loving preferences, including the case of exponential discounting.

[33] Recall that we normalized the agent's waiting cost to be linear in time so the time-risk attitude of the principal is relative to that of the agent.

question of why the principal does not simply eliminate all time risk by concentrating the agent's stopping on $t_1 = t_2$: this is indeed optimal when $h_\ell(t) = h_r(t)$ such the principal's time preferences do not change with the agent's action. But we will soon see, more generally, when the principal has differential delay preferences across actions, she sets a non-degenerate window $[t_1, t_2]$ to optimally trade off time-risk elimination for action prioritization.

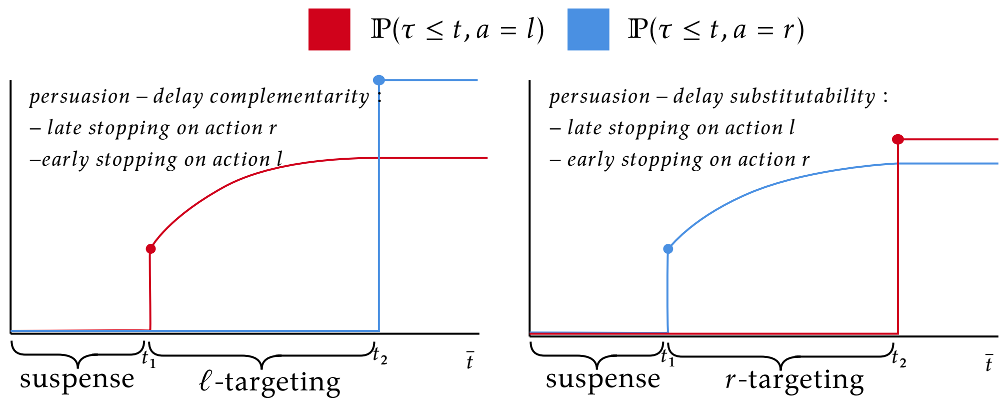
Note that the "plot twists" in the suspense-generating stage are not simple recommendations while the action-targeting stage is. Thus, our analysis also sheds light on how the time-risk preference interacts with intertemporal commitment: time-risk aversion necessitates frequent indirect messages to tie the principal's own hands under limited commitment. Meanwhile, under time-risk seeking preferences, simple recommendations / direct messages are sufficient with or without intertemporal commitment. Suspense generation is closely related to the "suspense-optimal" information of Ely, Frankel, and Kamenica (2015). However, there are key differences: in Ely, Frankel, and Kamenica (2015), the two belief paths are chosen so that the "residual variance" of the terminal belief conditional on the interim belief declines at a constant rate to maximally "smooth the consumption" of the variance reduction. In our model, however, the two belief paths are chosen that the continuation value of the agent declines at the rate that offsets his loss from delay so that the agent always obtains a zero surplus from continuation.[34]

Persuasion-delay complementarity/substitutability determine direction of targeting. The direction of targeting over the interval $[t_1, t_2]$ hinges on whether persuasion and delay are complements or substitutes. Suspense-ℓ is optimal (Proposition 2) when they are complements such that delay enhances the initial persuasion gain i.e., $\Delta h(t)$ is increasing in time. In this case, the principal benefits from positively correlating the event that the agent takes action r and the event that the agent stops late. ℓ-targeting facilitates this by offering the agent a specific kind of information to induce waiting: over the interval $[t_1, t_2)$, the agent might conclusively learn the state L, which induces immediate stopping and choosing ℓ; if this does not happen, the agent stops at t_2 and chooses r. This strategy maximizes the probability that *both* late stopping and action r obtains. The induced joint stopping time-action distribution is illustrated on the left

[34] The "suspense optimal" strategy as in Ely et al. (2015) is also optimal in frictional dynamic learning settings under time-risk aversion (See, e.g., Georgiadis-Harris (2024); Chen and Zhong (2024)). Although the time-risk attitude plays a similar role as in our model, the strategy is different from ours and is driven by a different force to smooth the acquisition of information.

in Figure 6.

Figure 6: Illustration of persuasion-delay complementarity vs substitutability



Now suppose, instead, that persuasion and delay are substitutes such that delay diminishes the initial persuasion gain i.e., $\Delta h(t) := h_r(t) - h_\ell(t)$ is decreasing in time—although, we emphasize, both h_ℓ and h_r remain increasing so the principal unambiguously enjoys delay. This can induce the principal to employ suspense-r; our next result pins down precise conditions under which this is optimal:[35]

Proposition 3. *Let* $\Psi(t_1, t_2) = \int_{t_1}^{t_2} e^{-s+t_1} h_r'(s) \mathrm{d}s + e^{-t_2+t_1} h_\ell'(t_2)$ *and* $0 \leq t_1 < t_2 < \mu_{t_1}^{+-1}(0)$*. If the* Suspense^{t_1}-r^{t_2} *strategy is optimal, then the following conditions hold*

(a) ***FOC for*** t_2***:*** $\Delta v + \Delta h(t_2) = -h_\ell'(t_2)$*,*

(b) ***FOC for*** t_1***:*** $\Psi(t_1, t_2) \geq h_r'(t_1)$*, with equality when* $t_1 > 0$*,*

(c) ***Local SOC:*** $h_r''(t_1) \cdot t_1 \leq 0, \;\; h_\ell''(t_2) \leq -\Delta h'(t_2)$*.*

Conversely, if conditions (a)-(c) above hold and, in addition,

(d) ***Global SOC:***
$$\begin{cases} h_\ell''(t), h_r''(t) \leq 0 & \forall t < t_1 \\ \max\{h_\ell''(t), 0\} \leq e^{t-t_1}\left(h_r'(t_1) - \int_{t_1}^{t} e^{-s+t_1} h_r'(s) \mathrm{d}s\right) - h_r'(t) & \forall t \geq t_1, \end{cases}$$

then the Suspense^{t_1}-r^{t_2} *strategy is optimal.*

The key difference across the conditions justifying suspense-r over suspense-l is condition (a): for suspense-r to be optimal, the principal's action preference must reverse: although the principal initially prefers action r over ℓ, she prefers ℓ over r at

[35] Proposition 3 is a direct corollary of Proposition 2 by switching the labels ℓ and r.

time t_2; that is, persuasion and delay are substitutes. In such cases, the principal benefits from negatively correlating the event that the agent takes action r and the event the agent stops late. r–targeting facilitates this by offering the agent Poisson arrivals of conclusive information that the state is R over the interval $[t_1, t_2)$. Upon arrival of such news, the agent is induced to stop immediately and choose r; if this does not happen, the agent stops at t_2 and chooses ℓ. This strategy maximizes the probability that *either* the agent chooses action r, or the agent stops late. The induced joint stopping time-action distribution is illustrated on the right in Figure 6.

Complete comparative statics We have chosen to present Propositions 2 and 3 in terms of necessary and sufficient conditions for the optimality of suspense-ℓ/r strategies to explicitly quantify the three key tradeoffs. One might wonder how special these conditions that justify such strategies are. It turns out that such strategies are broadly optimal whenever the sender's delay gain is concave. With the formal statement relegated to Online Appendix V, we state an informal version of the proposition here.

Proposition 4 (Informal). *If h_ℓ, h_r are concave, a "generalized" Suspense-ℓ (Suspense-r) strategy is optimal if persuasion and delay are complements i.e., $\Delta h' > 0$ (substitutes i.e., $\Delta h' < 0$).*

The "generalized" Suspense-$\ell(r)$ strategy extends the Suspense-$\ell(r)$ strategy to cover several corner cases and close variants. Nonetheless, they are qualitatively similar, and a formal definition is in Online Appendix V where we derive complete comparative statics with respect to preferences to show that the insights we develop in Section 4.2 continue to hold.

Discussion of the binary environment Our setting captures a wide array of communication environments where the principal has incentives to both *extend* the length of the principal-agent relationship (by delaying the agent's stopping time), as well as *influence* the agent's decision. Such principal-agent incentives arise in the cases of prosecutor-jury (Kamenica and Gentzkow, 2011), advisor-politician (Morris, 2001), manager-board of directors, advisor-student (Fudenberg and Rayo, 2019), supervisor-worker (Ely, Georgiadis, and Rayo, 2023), consultant-client and so on.[36]

Importantly, marginal delay and persuasion gain richly vary across these communication environments. For instance, in financial contexts, the client's delay cost is

[36] See also Aghion and Tirole (1997); Armstrong and Vickers (2010); Che et al. (2013).

typically explicit time discounting, leading to time-risk loving preferences and, consequently, no suspense generation. Meanwhile, in legal contexts, the jury might suffer increasing marginal costs of delay, leading to time-risk aversion and the generation of suspense. Likewise, whether persuasion and delay are complements or substitutes differ: a prosecutor might wish to drag out a trial because of the publicity it generates, which enhances her payoffs from obtaining a conviction. Then, persuasion and delay are complements and Proposition 2 suggests ℓ-targeting i.e., searching for conclusive evidence of innocence is then optimal. Conversely, the prosecutor may benefit from either wearing down the defendant via delay which makes a plea deal more likely, or from obtaining a conviction. Then, persuasion and delay are substitutes: the prosecutor would thus like to maximize the chance that *either* event happens, which is achieved by r-targeting i.e., looking for conclusive evidence of guilt.

Propositions 2 and 3 relate directly to recent work on dynamic persuasion with binary states by Che, Kim, and Mierendorff (2023) and Escudé and Sinander (2023).[37] These papers exogenously constrain the set of available information structures. For instance, Escudé and Sinander (2023) imposes a constraint on the drift of signal process driven by Brownian noise (so the belief process is continuous). Che, Kim, and Mierendorff (2023) constrains the principal to allocate a limited capacity across conclusive L/R signals (which, if generated, are also costly). In both papers, the principal's gain is purely from persuasion. Different from these papers, persuasion in our environment is fully frictionless—the principal can choose any sequence of experiments—and delay arises as a by-product of differential time-risk preferences. Moreover, we allow the principal's marginal gain from delay vis-a-vis persuasion to vary flexibly with time. This serves as a benchmark case and unveils a richer picture of what matters for the form of optimal dynamic information.[38]

4.3 General time-risk preferences and the timing of persuasion

The binary environment allowed us to obtain a sharp characterization of optimal persuasion strategies. We will now show that the qualitative features underlying the persuasion vs delay tradeoff hold more broadly in general environments with time-risk loving / averse preferences. We will maintain an additional regularity assumption that $T = [0, \overline{T}]$ and U, V are twice differentiable with respect to t. When preferences are con-

[37] See also Henry and Ottaviani (2019) and McClellan (2022).

[38] In Online Appendix VI we also show how frictions like costly or constrained information generation for the principal can be readily embedded within our framework.

vex in time, the optimal simple recommendation strategy keeps the agent indifferent between continuing and stopping:[39]

Proposition 5 (Time-risk loving)**.** *Suppose ($V_t'' > 0$, $U_t'' \geq 0$) or ($V_t' > 0$ and $V_t'', U_t'' \geq 0$ (with at least one strict inequality)). If f, Λ solve* (D) *and* $\bar{t} = \sup \operatorname{supp}(f) < \overline{T}$*, then* (OC-C) *must be binding for all $t \leq \bar{t}$.*

Proof. See Online Appendix I.7. *Q.E.D.*

Proposition 5 states that when the principal and agent are both time-risk loving, the optimal simple recommendation strategy must keep the agent indifferent between stopping and continuing *at all times*.[40]

To see why this is a strong prediction, consider the case where the agent is strictly impatient and an interval (t, t') on which the agent does not stop. Then, the continuation payoff of the agent is constant on the interval but his stopping payoff strictly decreases over time; hence, (OC-C) must be strictly slack on the entire interval. Per Proposition 5, such strategy cannot be optimal. Therefore, Proposition 5 effectively rules out any strategy that creates "suspense" under time-risk loving preferences.

A further implication of Proposition 5 is that, since (OC-C) is binding for all times, Theorem 2 implies the optimal simple recommendation strategy is dynamically consistent. Thus, under time-risk loving preferences, the "revelation principle" holds even under limited commitment: considering simple recommendations is without loss both with and without intertemporal commitment.

To state the result regarding time-risk aversion, we impose a technical regularity condition on the solution.

Definition 3. *Suppose f and Λ solve* (D)*. Let $\bar{t} = \sup_t \operatorname{supp}(f)$. f, Λ are* regular *if (i) $\exists \mu^* = \lim_{t \to \bar{t}^-} \hat{\mu}_t$, (ii) $(\mu^*, \bar{t}) \in \operatorname{supp}(f)$, and (iii) $\frac{\Lambda(\bar{t}) - \Lambda(t)}{\bar{t} - t}$ is bounded for $t \to \bar{t}^-$.*

Definition 3 requires that the continuation belief $\hat{\mu}_t$, as it approaches the terminal time $\bar{t}$, converges to a belief-time pair in the support of the optimal distribution f. This

[39] This generalizes the Suspense^{t_1}-ℓ/r^{t_2} strategy in the special case in which $t_1 = 0$ i.e., no suspense-generating stage.

[40] The condition $\bar{t} < \overline{T}$ is innocuous: if it is not fulfilled, we can expand the domain T until $\overline{T}$ is not binding. There is a special case where f has infinite support when $\overline{T} = \infty$. In this case, the statement of Proposition 5 still holds. See the discussion in Footnote 46.

condition is satisfied in all our applications. Define

$$\begin{cases} \overline{J}_V(t) = \max_{\mu\in\Delta(\Theta)} V_t'(\mu,t) \\ \underline{J}_V(t) = \min_{\mu\in\Delta(\Theta)} V_t'(\mu,t) \end{cases} \quad \text{and} \quad \begin{cases} \overline{J}_U(t) = \max_{\mu\in\Delta(\Theta)} U_t'(\mu,t) \\ \underline{J}_U(t) = \min_{\mu\in\Delta(\Theta)} U_t'(\mu,t) \end{cases}.$$

In words, $\overline{J}_V$ and $\underline{J}_V$ describe the range of the time derivative V_t' across different beliefs. In the special case where V is separable in μ and t, $\overline{J}_V = \underline{J}_V$. The same applies to $\overline{J}_U$ and $\underline{J}_U$.

Proposition 6 (Binding the persuasion window)**.** *Suppose V_t'', $U_t'' < 0$, regular f and Λ solve* (D). *Let $\underline{t} = \inf_t \operatorname{supp}(f) > 0$. Then,*

$$\bar{t} \leq \max\left\{\overline{J}_V^{-1} \circ \underline{J}_V(\underline{t}), \overline{J}_U^{-1} \circ \underline{J}_U(\underline{t})\right\},$$

where the max is taken over the image of correspondences $\overline{J}_V^{-1}$ and $\overline{J}_U^{-1}$.

Proof. See Online Appendix I.8. *Q.E.D.*

Proposition 6 establishes a bound on the size of the "persuasion window", i.e., times at which the agent may stop and act. $\underline{J}_V(\underline{t})$ is the principal's minimal marginal delay gain at $\underline{t}$. Then, $\overline{J}_V^{-1} \circ \underline{J}_V(\underline{t})$ gives the latest time at which the principal's maximal marginal delay gain is as high as the minimal marginal delay gain at $\underline{t}$. Therefore, for all $t > \max\left\{\overline{J}_V^{-1} \circ \underline{J}_V(\underline{t}), \overline{J}_U^{-1} \circ \underline{J}_U(\underline{t})\right\}$, the marginal delay gain (for both the principal and the agent) must be lower than that at t for arbitrary belief. Roughly, $\max\left\{\overline{J}_V^{-1} \circ \underline{J}_V(t), \overline{J}_U^{-1} \circ \underline{J}_U(t)\right\} - t$ measures the time-concavity of V and U relative to their variation across different beliefs. The gap is smaller when V, U are more concave in t or there is less variation in V_t' and U_t'. Proposition 6 states that the more (relatively) time-concave U and V are, the smaller the persuasion window is. In the extreme case where U and V are separable in μ and t, the persuasion window must be degenerate. In contrast with the time-risk loving case we analyzed above, Proposition 6 implies that time-risk averse preferences necessarily leads to a period of time prior to any stopping.

Note that if the agent is strictly impatient, (OC-C) is necessarily slack for all $t < \underline{t}$, rendering the simple recommendation $(\langle \mu_t^f \rangle, \tau^f)$ dynamically inconsistent. Therefore, in time-risk averse settings with limited commitment, the "anti-revelation principle" applies: the principal must utilize a generalized suspense-generating strategy to maximally release interim information to the agent over $[0, \underline{t})$ up to the point that the agent gets no surplus from continuing.

Despite their importance, the implication of time-risk attitudes on dynamic persuasion has—in our view—been under-explored. In particular, the discounted utility model has been considered the standard model and provides technical convenience since it is amenable to canonical recursive techniques. However, as we have illustrated, this technical convenience comes at the cost of unintended consequences of dictating the timing of persuasion. This insight has been alluded to in a few recent papers that adopt the standard discounted utility model and show that asymmetric discount rates lead to non-trivial implications for the timing of revelation (See, e.g., Ely and Szydlowski (2020) and Saeedi et al. (2024)).[41] Two unique features of our framework allow us to fully delineate the implications of time-risk attitudes on dynamic persuasion. First, Lemma 1 and Theorem 1 allow us to handle richer preference structures that are not necessarily time-stationary—we brought this to bear in Propositions 2, 3, 5 and 6. Second, Theorem 2 allows us to better understand how time-risk attitudes and commitment (or lack thereof) jointly shape specific persuasion modes.

5 Discussion and Extensions

We have provided a unified analysis of optimal dynamic persuasion in stopping problems. Under full commitment, we established a "revelation principle" which converts the dynamic problem into a semi-static problem. We then established strong duality and showed how the problem can be solved with a first-order approach via a novel concavification technique (Theorem 1). Without intertemporal commitment, we established that equilibrium *outcomes* coincide: the principal obtains her commitment payoff even in the absence of commitment (Theorem 2 (i)). However, equilibrium*strategies* differ: direct recommendations no longer suffice; instead, the principal leverages maximally informative "indirect messages" to tie her hands. We made this precise via an "anti-revelation principle" (Theorem 2 (ii)).

We then analyzed how optimal dynamic information trades off persuasion against delay. Optimal information often utilizes both *suspense-generation* featuring "inconclusive plot twists" which concentrates the agent's stopping time, and *action-targeting* fea-

[41] Ball and Knoepfle (2023) studies a different setting of inspection. There, the asymmetric arrival rates of good / bad news lead to different time-risk attitudes, which determine the optimal timing of inspection.

turing stochastic arrivals of "good news" or "bad news" which correlates/anti-correlates persuasion and delay (Propositions 2 and 3).[42] Our analysis made precise how the principal's choice is jointly shaped by (i) time-risk preferences; (ii) persuasion vs delay gain; and (iii) persuasion-delay complementarity vs substitutability. We further showed that the impact of time-risk preferences on optimal information holds in more general environments (Propositions 5 and 6).

Finally, our results were obtained within a simple but canonical environment. This allowed us to transparently analyze the interplay between persuasion, timing, and commitment under general time-risk preferences. Nonetheless, our framework can be readily extended to handle more complicated environments such as those with costly information generation, or with a changing state (see Online Appendix VI).

References

Aghion, P. and J. Tirole (1997): "Formal and real authority in organizations," *Journal of political economy*, 105, 1–29.

Akbarpour, M. and S. Li (2020): "Credible auctions: A trilemma," *Econometrica*, 88, 425–467.

Armstrong, M. and J. Vickers (2010): "A model of delegated project choice," *Econometrica*, 78, 213–244.

Arrow, K. J., D. Blackwell, and M. A. Girshick (1949): "Bayes and minimax solutions of sequential decision problems," *Econometrica, Journal of the Econometric Society*, 213–244.

Ball, I. (2023): "Dynamic information provision: Rewarding the past and guiding the future," *Econometrica*, 91, 1363–1391.

Ball, I. and J. Knoepfle (2023): "Should the timing of inspections be predictable?" .

Bester, H. and R. Strausz (2001): "Contracting with imperfect commitment and the revelation principle: the single agent case," *Econometrica*, 69, 1077–1098.

Bizzotto, J., J. Rüdiger, and A. Vigier (2021): "Dynamic persuasion with outside information," *American Economic Journal: Microeconomics*, 13, 179–194.

Bloedel, A. W. and I. Segal (2020): "Persuading a rationally inattentive agent," *Available at SSRN*.

Bulow, J. I. (1982): "Durable-goods monopolists," *Journal of political Economy*, 90, 314–332.

Che, Y.-K., W. Dessein, and N. Kartik (2013): "Pandering to persuade," *American Economic Review*, 103, 47–79.

Che, Y.-K., K. Kim, and K. Mierendorff (2023): "Keeping the listener engaged: a dynamic model of bayesian persuasion," *Journal of Political Economy*, 131, 1797–1844.

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

[42] This also offered an optimality foundation for "bad news" and "good news" information structures often exogenously assumed in the extant literature.

Chen, D. and W. Zhong (2024): "Information Acquisition and Time-Risk Preference," *American Economic Review: Insights*, Forthcoming.

Coase, R. H. (1972): "Durability and monopoly," *The Journal of Law and Economics*, 15, 143–149.

Dixit, A. K. and R. S. Pindyck (1994): *Investment under uncertainty*, Princeton university press.

Doval, L. and V. Skreta (2022): "Mechanism design with limited commitment," *Econometrica*, 90, 1463–1500.

Ely, J., A. Frankel, and E. Kamenica (2015): "Suspense and surprise," *Journal of Political Economy*, 123, 215–260.

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

Ely, J. C., G. Georgiadis, and L. Rayo (2023): "Feedback Design in Dynamic Moral Hazard," .

Ely, J. C. and M. Szydlowski (2020): "Moving the goalposts," *Journal of Political Economy*, 128, 468–506.

Escudé, M. and L. Sinander (2023): "Slow persuasion," *Theoretical Economics*, 18, 129–162.

Forges, F. (1986): "An approach to communication equilibria," *Econometrica: Journal of the Econometric Society*, 1375–1385.

Fudenberg, D., D. Levine, and J. Tirole (1985): "Infinite Horizon Models of Bargaining with One-Sided Incomplete Information," *Game Theoretic Models of Bargaining, ed. by Alvin Roth*.

Fudenberg, D. and L. Rayo (2019): "Training and effort dynamics in apprenticeship," *American Economic Review*, 109, 3780–3812.

Fudenberg, D., P. Strack, and T. Strzalecki (2018): "Speed, accuracy, and the optimal timing of choices," *American Economic Review*, 108, 3651–3684.

Georgiadis-Harris, A. (2024): "Preparing to Act," Tech. rep., Working Paper.

Gonçalves, D. (2024): "Speed, Accuracy, and Complexity," *arXiv preprint arXiv:2403.11240*.

Hébert, B. and M. Woodford (2023): "Rational inattention when decisions take time," *Journal of Economic Theory*, 208, 105612.

Hébert, B. and W. Zhong (2022): "Engagement maximization," *arXiv preprint arXiv:2207.00685*.

Henry, E. and M. Ottaviani (2019): "Research and the approval process: The organization of persuasion," *American Economic Review*, 109, 911–955.

Kamenica, E. and M. Gentzkow (2011): "Bayesian persuasion," *American Economic Review*, 101, 2590–2615.

Keller, G. and S. Rady (2015): "Breakdowns," *Theoretical Economics*, 10, 175–202.

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

Knoepfle, J. (2020): "Dynamic Competition for Attention," Tech. rep.

Koh, A. and S. Sanguanmoo (2024): "Attention capture," *arXiv preprint arXiv:2209.05570*.

Laffont, J.-J. and J. Tirole (1988): "The dynamics of incentive contracts," *Econometrica: Journal of the Econometric Society*, 1153–1175.

Liang, A., X. Mu, and V. Syrgkanis (2022): "Dynamically Aggregating Diverse Information," *Econometrica*, 90.

Lipnowski, E., L. Mathevet, and D. Wei (2020): "Attention management," *American Economic*

Review: Insights, 2, 17–32.
Lipnowski, E., D. Ravid, and D. Shishkin (2022): "Persuasion via weak institutions," *Journal of Political Economy*, 130, 2705–2730.
Luenberger, D. G. (1997): *Optimization by vector space methods*, John Wiley & Sons.
McAdams, D. and M. Schwarz (2007): "Credible sales mechanisms and intermediaries," *American Economic Review*, 97, 260–276.
McClellan, A. (2022): "Experimentation and approval mechanisms," *Econometrica*, 90, 2215–2247.
Mekonnen, T., Z. Murra-Anton, and B. Pakzad-Hurson (2023): "Persuaded search," *arXiv preprint arXiv:2303.13409*.
Morris, S. (2001): "Political correctness," *Journal of political Economy*, 109, 231–265.
Morris, S. and P. Strack (2019): "The wald problem and the relation of sequential sampling and ex-ante information costs," *Available at SSRN 2991567*.
Moscarini, G. and L. Smith (2001): "The optimal level of experimentation," *Econometrica*, 69, 1629–1644.
Myerson, R. B. (1986): "Multistage games with communication," *Econometrica: Journal of the Econometric Society*, 323–358.
Orlov, D., A. Skrzypacz, and P. Zryumov (2020): "Persuading the principal to wait," *Journal of Political Economy*, 128, 2542–2578.
Renault, J., E. Solan, and N. Vieille (2017): "Optimal dynamic information provision," *Games and Economic Behavior*, 104, 329–349.
Saeedi, M., Y. Shen, and A. Shourideh (2024): "Getting the agent to wait," *arXiv preprint arXiv:2407.19127*.
Sannikov, Y. and W. Zhong (2024): "Martingale embeddings: theory and applications," *Working paper*.
Siegel, R. and B. Strulovici (2020): "The economic case for probability-based sentencing," Tech. rep., Working paper.[132].
Skreta, V. (2015): "Optimal auction design under non-commitment," *Journal of Economic Theory*, 159, 854–890.
Steiner, J., C. Stewart, and F. Matějka (2017): "Rational inattention dynamics: Inertia and delay in decision-making," *Econometrica*, 85, 521–553.
Stokey, N. L. (1979): "Intertemporal price discrimination," *The Quarterly Journal of Economics*, 93, 355–371.
Sugaya, T. and A. Wolitzky (2021): "The revelation principle in multistage games," *The Review of Economic Studies*, 88, 1503–1540.
Wald, A. (1947): "Foundations of a general theory of sequential decision functions," *Econometrica, Journal of the Econometric Society*, 279–313.
Zhao, W., C. Mezzetti, L. Renou, and T. Tomala (2024): "Contracting over persistent information," *Theoretical Economics*, 19, 917–974.
Zhong, W. (2018): "Information Design Possibility Set," .
——— (2022): "Optimal dynamic information acquisition," *Econometrica*, 90, 1537–1582.

A Appendix: Omitted Proofs

A.1 Proof of Theorem 1

Proof. *Sufficiency*: Suppose for the purpose of contradiction that (f, a, Λ) satisfies Equation (FOC),(OC-C), and the complementary slackness condition $\mathcal{L}(f, \Lambda) = \mathbb{E}_f[V]$ but f is suboptimal in (R). Then, there exists $f' \in \Delta_{\mu_0}$ s.t. $\mathcal{L}(f, \Lambda) < \mathcal{L}(f', \Lambda)$. Then, since $\mathcal{L}$ is concave, $\forall \alpha \in (0, 1)$,

$$
\begin{aligned}
&\frac{\mathcal{L}(\alpha f' + (1-\alpha) f, \Lambda) - \mathcal{L}(f, \Lambda)}{\alpha} \geq \mathcal{L}(f', \Lambda) - \mathcal{L}(f, \Lambda) > 0 \\
\Longrightarrow\ & \lim_{\alpha \to 0} \frac{\mathcal{L}(\alpha f' + (1-\alpha) f, \Lambda) - \mathcal{L}(f, \Lambda)}{\alpha} \geq \mathcal{L}(f', \Lambda) > 0 \\
\Longleftrightarrow\ & \int V(\mu, t)(f' - f)(\mathrm{d}\mu, \mathrm{d}t) + \int_{t \in T^\circ} \int_{\tau > t} U(\mu, t)(f' - f)(\mathrm{d}\mu, \mathrm{d}\tau) \mathrm{d}\Lambda(t) \\
& - \overline{\lim_{\alpha \to 0}} \int_{t \in T^\circ} \underbrace{\frac{U\left(\int_{\tau > t} \mu(\alpha f' + (1-\alpha) f)(\mathrm{d}\mu, \mathrm{d}\tau), t\right) - U\left(\int_{\tau > t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau), t\right)}{\alpha}}_{\geq \max_{\gamma \in \nabla_\mu U(\widehat{\mu_t})} \gamma \cdot \int_{\tau > t} \mu(f' - f)(\mathrm{d}\mu, \mathrm{d}\tau)} \mathrm{d}\Lambda(t) > 0 \\
\Longrightarrow\ & \int V(\mu, t)(f' - f)(\mathrm{d}\mu, \mathrm{d}t) + \int_{t \in T^\circ} \int_{\tau > t} U(\mu, t)(f' - f)(\mathrm{d}\mu, \mathrm{d}\tau) \mathrm{d}\Lambda(t) \\
& - \int_{t \in T^\circ} \nabla_\mu U\left(\int_{\tau > t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau), t\right) \cdot \int_{\tau > t} \mu(f' - f)(\mathrm{d}\mu, \mathrm{d}\tau) \mathrm{d}\Lambda(t) > 0 \\
\Longrightarrow\ & \int l_{f,\Lambda}(f' - f)(\mathrm{d}\mu, \mathrm{d}t) > 0 \\
\Longrightarrow\ & \mathbb{E}_{f'}[l_{f,\Lambda}] > a \cdot \mu_0.
\end{aligned}
$$

Note that in the fourth inequality, the selection of sub-gradients can be arbitrary. The last inequality violates Equation (FOC).

Necessity: Suppose f solves (R). Given strong duality, let Λ be the minimizer in (D). When defining $l_{f,\Lambda}(\mu, t)$, make the following selection: $\forall (\mu, t) \in D$, if $t < \bar{t}$, $\nabla_\mu U(\widehat{\mu_t}, t) = \arg\max_{\gamma \in \nabla_\mu U(\widehat{\mu_t}, t)} \gamma \cdot \mu$; if $t \geq \bar{t}$, $\nabla_\mu U(0, t)$ is chosen that $\nabla_\mu U(0, t) \cdot \mu = U(\mu, t)$. Define

$$\hat{l}(\mu) := \sup_{f' \in \Delta_\mu} \mathbb{E}_{f'}[l_{f,\Lambda}(\nu, t)].$$

Let $a \cdot \mu$ be the supporting hyperplane of $\hat{l}$ at μ_0. Then, $l_{f,\Lambda}(\mu, t) \leq a \cdot \mu$. Next, we prove that $\mathbb{E}_f[l_{f,\Lambda}(\mu, t)] = a \cdot \mu_0$. Suppose for the purpose of contradiction that $\mathbb{E}_f[l'_{f,\Lambda}(\mu, t)] \leq a \cdot \mu_0 - \epsilon$ for $\epsilon > 0$. There exists a finite support $f' \in \Delta_{\mu_0}$ s.t. $\mathbb{E}_{f'}[l'_{f,\Lambda}] > a \cdot \mu_0 - \frac{1}{2}\epsilon$. Then,

$$
\begin{aligned}
&\lim_{\alpha \to 0} \frac{\mathcal{L}(\alpha f' + (1-\alpha) f, \Lambda) - \mathcal{L}(f, \Lambda)}{\alpha} \\
=\ & \int V(\mu, t)(f' - f)(\mathrm{d}\mu, \mathrm{d}t) + \int_{t \in T^\circ} \int_{\tau > t} U(\mu, t)(f' - f)(\mathrm{d}\mu, \mathrm{d}\tau) \mathrm{d}\Lambda(t)
\end{aligned}
$$

$$
\begin{aligned}
&- \overline{\lim_{\alpha\to 0}} \int_{t\in T^\circ} \frac{U\left(\int_{\tau>t}\mu(\alpha f' + (1-\alpha)f)(\mathrm{d}\mu,\mathrm{d}\tau),t\right) - U\left(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}\tau),t\right)}{\alpha}\mathrm{d}\Lambda(t)\\
\geq & \int V(\mu,t)(f'-f)(\mathrm{d}\mu,\mathrm{d}t) + \int_{t\in T^\circ}\int_{\tau>t} U(\mu,t)(f'-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\Lambda(t)\\
&- \sup_{\gamma_t\in\nabla_\mu U(\widehat{\mu}_t,t)} \int_{t<\bar{t}} \gamma_t\cdot\int_{\tau>t}\mu(f'-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\Lambda(t)\\
&- \int_{t\geq\bar{t}} U\left(\int_{\tau>t}\mu f'(\mathrm{d}\mu,\mathrm{d}\tau),t\right)\mathrm{d}\Lambda(t)\\
\geq & \int V(\mu,t)(f'-f)(\mathrm{d}\mu,\mathrm{d}t) + \int_{t\in T^\circ}\int_{\tau>t} U(\mu,t)(f'-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\Lambda(t)\\
&- \int_{t<\bar{t}}\int_{\tau>t}\sup_{\gamma\in\nabla_\mu U(\widehat{\mu}_t,t)} \gamma\cdot\mu(f'-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\Lambda(t)\\
&- \int_{t\geq\bar{t}} U\left(\int_{\tau>t}\mu f'(\mathrm{d}\mu,\mathrm{d}\tau),t\right)\mathrm{d}\Lambda(t)\\
\geq & \int V(\mu,t)(f'-f)(\mathrm{d}\mu,\mathrm{d}t) + \int_{t\in T^\circ}\int_{\tau>t} U(\mu,t)(f'-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\Lambda(t)\\
&- \int_{t<\bar{t}}\int_{\tau>t}\sup_{\gamma\in\nabla_\mu U(\widehat{\mu}_t,t)} \gamma\cdot\mu(f'-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\Lambda(t)\\
&- \int_{t\geq\bar{t}}\int_{\tau>t} U(\mu,t)f'(\mathrm{d}\mu,\mathrm{d}t)\mathrm{d}\Lambda(t)\\
= & \int V(\mu,t)(f'-f)(\mathrm{d}\mu,\mathrm{d}t) + \int_{t\in T^\circ}\int_{\tau>t} U(\mu,t)(f'-f)(\mathrm{d}\mu,\mathrm{d}\tau)\mathrm{d}\Lambda(t)\\
&- \int_{t\in T^\circ}\left(\int_{\tau<t}\nabla_\mu U(\widehat{\mu}_\tau,\tau)\cdot\mu\mathrm{d}\Lambda(\tau)\right)(f'-f)(\mathrm{d}\mu,\mathrm{d}t)\\
=&\mathbb{E}_{f'-f}[l'_{f,\Lambda}(\mu,t)] > \frac{1}{2}\epsilon.
\end{aligned}
$$

Contradicts the optimality of f. *Q.E.D.*

A.2 Proof of Theorem 2

Proof. We prove Theorem 2 via an auxiliary problem. Given process $(\langle\mu_t\rangle,\tau)$, $\forall(\mu',t')\in D$, define $\Delta_{\mu',t'} = \{f\in\Delta(\Delta(\Theta)\times[t',\infty)\cap T)|\mathbb{E}_f[\mu]=\mu'\}$ and

$$
W(\mu',t') := \sup_{f\in\Delta_{\mu',t'}}\int V(\mu,t)f(\mathrm{d}\mu,\mathrm{d}t)
$$

$$
\text{s.t.} \int_{y>t} U(\mu,y)f(\mathrm{d}\mu,\mathrm{d}y) \geq U\left(\int_{y>t}\mu f(\mathrm{d}\mu,\mathrm{d}y),t\right),\ \forall t\geq t'.
$$

We consider pairs of $(\langle\mu_t\rangle,\tau)$ that satisfy the following condition:

$$
\forall(\mu',t')\in\mathrm{supp}\langle\mu_t|_{t<\tau}\rangle,\ \mathbb{E}[V(\mu_\tau,\tau)|\mathcal{F}_{t'},\mu_{t'}=\mu',\tau>t']\geq W(\mu',t'). \tag{3}
$$

Lemma 4. *If* $\sigma \in SPE(\Gamma)$*, then* $\mathbb{E}[V(\mu(\sigma)_{\tau(\sigma)}, \tau(\sigma))] \leq$ (R).

Proof. Let $f \sim (\mu(\sigma)_{\tau(\sigma)}, \tau(\sigma))$. At any on-path history H_{t+1}, t^+, the agent may choose $\sigma'_A = \delta_{stop}$ to guarantee an interim payoff of $U(\mu_t, t) \leq \mathcal{U}(\sigma|H_{t+1}, t^+)$. Since U is convex in μ, $\mathbb{E}[U(\mu_t, t)|t < \tau(\sigma)] \geq U(\mathbb{E}[\mu_t|t < \tau(\sigma)], t)$. Combining the inequalities

$$\begin{aligned}\frac{\int_{\tau>t} U(\mu,\tau) f(\mathrm{d}\mu, \mathrm{d}\tau)}{\int_{\tau>t} f(\mathrm{d}\mu, \mathrm{d}\tau)} =& \mathbb{E}\left[\mathcal{U}(\sigma|H_{t+1}, t^+)|H_{t+1}, t^+\right] \\ \geq& \mathbb{E}[U(\mu_t, t)|t < \tau(\sigma)] \\ \geq& U(\mathbb{E}[\mu_t|t < \tau(\sigma)], t) \\ =& \frac{U\left(\int_{\tau>t} \mu f(\mathrm{d}\mu, \mathrm{d}\tau), t\right)}{\int_{\tau>t} f(\mathrm{d}\mu, \mathrm{d}\tau)}.\end{aligned}$$

On the RHS of the first line, the expectation is taken over all on-path histories in period t, which is identical to the event $t < \tau(\sigma)$ in period t. Therefore, f is feasible in (R). *Q.E.D.*

Lemma 5. *Suppose finite support* $(\langle \mu_t \rangle, \tau)$ *solves* (P)*,* $(\langle \mu_t \rangle, \tau)$ *satisfies condition* (3) *if for any stopping time* $\iota \leq \tau$*,* $\mathbb{E}[U(\mu_\tau, \tau)] = \mathbb{E}[U(\mu_\iota, \iota)]$.

Proof. Suppose condition (3) is violated at the history corresponding to filtration $\mathcal{F}_t$, conditional on μ_t and $\tau > t$. Let $f_1 \sim (\mu_\tau, \tau)|_{\mathcal{F}_t, \mu_t, \tau > t}$, $p_1 = \mathrm{Prob}(\mathcal{F}_t, \mu_t, \tau > t)$. Let $f_2 \sim (\mu_\tau, \tau)|_{\mathrm{not}(\mathcal{F}_t, \mu_t, \tau>t)}$ and $p_2 = 1 - p_1$. Then, $f = f_1 p_1 + f_2 p_2$. Since (3) is violated, let $f' \in \Delta_{\mu_t, t}$ be the deviation and $\hat{f} = f' p_1 + f_2 p_2$. Then since f' strictly improves on f_1,

$$\mathbb{E}_{\hat{f}}[V] = p_1 \mathbb{E}_{f'}[V] + p_2 \mathbb{E}_{f_2}[V] > \mathbb{E}_f[V].$$

Next, we verify that $\hat{f}$ is feasible under (P). For all times $t' \geq t$,

$$\begin{aligned}\int_{y>t'} U(\mu, y) \mathrm{d}\hat{f} =& p_1 \int_{y>t'} U(\mu, y) \mathrm{d}f' + p_2 \int_{y>t'} U(\mu, y) \mathrm{d}f_2 \\ \geq& p_1 U\left(\int_{y>t'} \mu \mathrm{d}f', t'\right) \\ =& U\left(p_1 \int_{y>t'} \mu \mathrm{d}f', t'\right) \\ =& U\left(\int_{y>t'} \mu \mathrm{d}\hat{f}, t'\right)\end{aligned}$$

The second line is from the feasibility condition in the deviation and f_2 having no mass when $y > t' \geq t$. The third line is because U is homogeneous of degree 1. The last line is from $\hat{f}$ being identical to $p_1 f'$ when $y > t$. For times $t' < t$,

$$\int_{y>t'} U(\mu, y) \mathrm{d}\hat{f} = p_1 \int_{y>t} U(\mu, y) \mathrm{d}f' + p_2 \int_{y>t'} U(\mu, y) \mathrm{d}f_2$$

$$
\begin{aligned}
&\geq p_1 U\left(\int_{y>t} \mu \mathrm{d} f', t\right) + p_2 \int_{y>t'} U(\mu, y) \mathrm{d} f_2 \\
&= p_1 U\left(\int_{y>t} \mu \mathrm{d} f_1, t\right) + p_2 \int_{y>t'} U(\mu, y) \mathrm{d} f_2 \\
&= p_1 \int_{y>t} U(\mu, y) \mathrm{d} f_1 + p_2 \int_{y>t'} U(\mu, y) \mathrm{d} f_2 \\
&= \int_{y>t'} U(\mu, y) \mathrm{d} f \\
&\geq U\left(\int_{y>t'} \mu \mathrm{d} f, t'\right) \\
&= U\left(\int_{y>t'} \mu \mathrm{d} \widehat{f}, t'\right)
\end{aligned}
$$

The second and third lines are from the feasibility condition in the deviation. The fourth line is from (3) since the agent is indifferent between stopping and continuing conditional on the history leading to distribution f_1. The last inequality is from the fact that $\mathbb{E}_{f_1}[\mu] = \mathbb{E}_{f'}[\mu]$. Therefore, $\widehat{f}$ is feasible and yields strictly higher payoff than f, a contradiction. *Q.E.D.*

Lemma 6. *There exists a finite support $(\langle \mu_t \rangle, \tau)$ that solves (P) and for any stopping time $\iota \leq \tau$, $\mathbb{E}[U(\mu_\tau, \tau)] = \mathbb{E}[U(\mu_\iota, \iota)]$.*

Proof. Step 1: We show that f solving (R) can be chosen to have finite support. Given an optimal f, $\forall t \in T$, solve

$$
\max_{f' \in \Delta(\Theta)} \mathbb{E}_{f'}[V(\mu, t)]
$$

$$
\text{s.t.} \begin{cases} \mathbb{E}_{f'}[U(\mu, t)] \geq \mathbb{E}_{f|_t}[U(\mu, t)] \\ \mathbb{E}_{f'}\mu = \mathbb{E}_{f|_t}\mu. \end{cases}
$$

By Theorem 1 of Zhong (2018), there exists a maximizer f' that has finite support. Then, replacing $f(\cdot, t)$ with $f'(\cdot) \times f(\Delta(\theta), t)$, weakly improves $\mathbb{E}_f[V]$ and $\mathbb{E}_f[U]$ without impacting $U(\widehat{\mu}_t, t)$. Therefore, the modified f is still optimal and satisfies (OC-C). By modifying f at each t, we obtain f solving (R) with a finite support.

Step 2: begin with the simple recommendation strategy $(\langle \mu_t^f \rangle, \tau)$. We show recursively that it can be modified to have the agent indifferent between continuing and stopping at every interim belief. Suppose that t is the last period of the finite support strategy $(\langle \mu_t \rangle, \tau)$ at which the agent strictly prefers continuing at interim belief $\mu = \widehat{\mu}_t$ at history $\mathcal{F}_t$. Suppose μ splits into $(\mu_i)_{i=1}^n$ in the next period, satisfying $\mu = \sum p_i \mu_i$. Then, for each i, pick $\lambda_i > 0$ s.t.

$$
\lambda_i U(\mu_i, t+1) + (1 - \lambda_i) \sum p_j U(\mu_j, t+1) = U(\lambda_i \mu_i + (1 - \lambda_i)\mu, t).
$$

Note that the LHS is weakly lower than the RHS if $\lambda_i = 1$ since U is weakly decreasing. The LHS is strictly higher than the RHS if $\lambda_i = 0$ since continuing is strictly better. Such λ_i exists due to the continuity of U. Let $\widehat{\mu}_i = \lambda_i \mu_i + (1-\lambda_i)\mu$. Then, modify $(\langle \mu_t \rangle, \tau)$ as follows. Conditional on $\widehat{\mu}_{t-1}$:

- For each $\mu_j \in \text{supp}(f(\cdot,t))$, $\text{Prob}(\mu_t = \mu_j, \tau = t | \widehat{\mu}_{t-1}) = f(\mu_j, t)/f(\Delta(\Theta), t)$.
- For each i, $\text{Prob}(\mu_t = \widehat{\mu}_i, \tau > t | \widehat{\mu}_{t-1}) = \frac{p_i/\lambda_i}{\sum p_j/\lambda_j} \frac{f(\mu,t)}{f(\Delta(\Theta),t)}$.

In words, the conditional stopping distribution is maintained. Conditional on continuing, μ is further split into $\widehat{\mu}_i$. It is easy to verify that $\mathbb{E}[\mu_t | \widehat{\mu}_{t-1}] = \widehat{\mu}_{t-1}$. Next, conditional on $\mu_t = \widehat{\mu}_i$:

- $\text{Prob}(\mu_{t+1} = \mu_j | \mu_i) = (1-\lambda_j)p_j + \mathbf{1}_{i=j}\lambda_i$.

The conditional stopping probability in period $t+1$ stays unchanged. In words, each $\widehat{\mu}_i$ is split into μ_j's in the next period. It is easy to verify that $\langle \mu \rangle_t$ remains a martingale. By the definition of λ_i, the agent is indifferent at each $\widehat{\mu}_i$.

Then, since the time space T is finite, we can iteratively apply the procedure above to obtain the finitely-supported strategy $(\langle \mu \rangle_t, \tau)$ where, by construction, the agent finds it optimal to stop at any interim belief in the support. *Q.E.D.*

Lemma 7. *If finite support $(\langle \mu_t \rangle, \tau)$ satisfies condition (3) and $V(\mu, t)$ strictly increases in t, then $\exists \sigma \in SPE^P(\Gamma)$ s.t. $(\langle \mu_t \rangle, \tau) \stackrel{d}{=} (\mu(\sigma), \tau(\sigma))$.*

Proof. We proceed by inducting on the size of the time space T. Consider the base case $T = \{1\}$. The history $H_1 = (\mu_0)$ is unique and condition (3) has no bite as there is no interim belief. Recall $I_1 \in \Delta^2(\Theta)$ be the distribution of μ_1. I_1 solves (R):

$$I_1 \in \arg\max_{I \in \Delta^2(\Theta)} \mathbb{E}_I[V(\mu, 1)] \quad \text{s.t.} \quad \mathbb{E}_I[\mu] = \mu_0.$$

$\sigma_P(H_1) = I_1$ constitutes an SPE^P with strategies $(\mu(\sigma), \tau(\sigma)) \stackrel{d}{=} (\mu_1, \delta_1)$.

Next, assume the induction hypothesis that the statement is true for times $\overline{T} = k-1 \geq 1$. Following the null history H_1, define $\sigma_P(H_1)$ as the distribution of μ_1 which specifies the principal's time 1 strategy. Since $\langle \mu_t \rangle$ is a martingale, $\sigma_P(H_1) \in \Delta^2_{\mu_0}(\Theta)$ hence this is feasible. Let $M = \text{supp}(\mu_1|_{\tau>1})$ and $M' = \text{supp}(\mu_1|_{\tau=1})$. Then, define the time $t = 1$ strategy for the agent as follows: $\sigma_A(\mu_0, \sigma_P(H_1), \mu_1 \in M) = \delta_{continue}$ and $\sigma_A(\mu_0, \sigma_P(H_1), \mu_1 \in M') = \delta_{stop}$. We define the continuation strategy after $t = 1$ by 'pasting' the equilibrium strategy from the induction hypothesis:

$$\sigma_P\big(H_t := (\mu_0, \sigma_P(H_1), \mu', I_s, \mu_s)_{2 \leq s < t}\big) = \sigma'_P\big(H_{t-1} := (\mu', I_s, \mu_s)_{1 \leq s < t-1}\big)$$
$$\sigma_A\big(H_{t+1} := (\mu_0, \sigma_P(H_1), \mu', I_s, \mu_s)_{2 \leq s \leq t}\big) = \sigma'_A\big(H_t := (\mu', I_s, \mu_s)_{1 \leq s \leq t-1}\big),$$

where (σ'_A, σ'_P) is chosen depending on three types of subgames as follows:

1. **Agent continues.** Following any history $H_2 = (\mu_0, \sigma_P(H_1), \mu')$ s.t. $\mu' \in M$, consider the (finite support) process $\nu_t = \mu_t | (\mathcal{F}_1, \mu_1 = \mu', 1 < \tau)$ and stopping time $\iota = \tau | (\mu_1 = \mu', 1 < \tau)$, i.e., $(\langle \nu_t \rangle, \iota)$ is a truncation of $(\langle \mu_t \rangle, \tau)$ following interim belief μ'.[43] Then, $(\langle \mu_t \rangle, \tau)$ satisfying condition (3) implies that $(\langle \nu_t \rangle, \tau)$ satisfies condition (3). From the induction hypothesis, we can choose an SPE^P (σ'_A, σ'_P) of the subgame with primitives $T = \{1, \dots k-1\}$, $U(\cdot, t+1)$, $V(\cdot, t+1)$ and prior μ' which satisfies $(\mu(\sigma'), \tau(\sigma')) \overset{d}{=} (\nu_\iota, \iota)$.

2. **Agent stops.** Following any history $H_2 = (\mu_0, \sigma_P(H_1), \mu')$ s.t. $\mu' \in M'$, $(\langle \nu_t \rangle, \iota) = (\mu', \delta_1)$ defines a finite support strategy satisfying condition (3). From the induction hypothesis, we can choose an SPE^P (σ'_A, σ'_P) of the subgame with primitives $T = \{1, \dots k-1\}$, $U(\cdot, t+1)$, $V(\cdot, t+1)$ and μ' which satisfies $(\mu(\sigma'), \tau(\sigma')) \overset{d}{=} (\delta_{\mu'}, \delta_1)$ i.e., the agent stops immediately at $t = 1$.

3. **Off-path.** Following any history $H_2 = (\mu_0, I_1, \mu_1)$ where $I_1 \neq \sigma_P(H_1)$ or $\mu_1 \notin M \cup M'$, Lemma 2 implies that there exists f solving (R) with primitives $T = \{1, \dots, k-1\}$, $U(\cdot, t+1)$, $V(\cdot, t+1)$ and $\mu_0 = \mu_1$. Then, Lemmas 5 and 6 (iii) implies that there exists a finite support strategy $(\langle \nu_t \rangle, \iota)$ satisfying condition (3) and $(\nu_\iota, \iota) \sim f$. From the induction hypothesis we can choose an SPE^P (σ'_A, σ'_P) of this subgame such that $(\mu(\sigma'), \tau(\sigma')) \overset{d}{=} f$. Additionally, define the $t = 1$ strategy for the agent as[44]

$$\sigma_A(H_1, I_1, \mu_1) = 1_{\mathcal{U}(\sigma') > U(\mu_1, 1)} \delta_{continue} + 1_{\mathcal{U}(\sigma') \leq U(\mu_1, 1)} \delta_{stop}$$

By construction, at any history at time $t > 1$, (σ_A, σ_P) specifies an SPE^P. To verify that (σ_A, σ_P) is an SPE, it remains to verify that there is no profitable deviation in period $t = 1$.

- Agent's deviation:
  - *continue* $\to$ *stop*. Following any history $H_2 = (\mu_0, \sigma_P(H_1), \mu')$ s.t. $\mu' \in M$, the continuation payoff is $\mathbb{E}[U(\mu_\tau, \tau) | \mathcal{F}_1, \mu_1 = \mu', 1 < \tau] \geq U(\mu', 1)$ which is guaranteed by the obedience condition. Therefore, deviating to *stop* is suboptimal.
  - *stop* $\to$ *continue*. Following any history $H_2 = (\mu_0, I_1, \mu_1)$ where $I_1 = \sigma_P(H_1)$ and $\mu_1 \in M'$, $\sigma_A(H_1) = \delta_{stop}$. Let σ' be the continuation SPE^P in period 2 specified above. Suppose for the purpose of contradiction that $\mathcal{U}(\sigma') > U(\mu_1, 1)$. Then, $\mathcal{V}(\sigma') \leq V(\mu_1, 2)$ must hold. If it did not, then since V is increasing in time, $\mathcal{V}(\sigma') > V(\mu_1, 2) \geq V(\mu_1, 1)$ but replacing $f(\mu_1, 1)$ with the distribution of $(\mu_{\tau(\sigma')}, \tau(\sigma'))$ strictly improves $\mathbb{E}_f[V]$ and weakly improves (OC-C), contradicting the optimality of f. But $\mathcal{V}(\sigma') \leq V(\mu_1, 2)$ implies that in the construction of σ' above, we set $(\mu(\sigma'), \tau(\sigma')) \overset{d}{=} (\delta_{\mu_1}, 2)$. This implies $\mathcal{U}(\sigma') = U(\mu_1, 2) \leq U(\mu_1, 1)$, contradiction. Since $\mathcal{U}(\sigma') \leq U(\mu_1, 1)$, it is suboptimal to deviate to *continue*.

[43] The truncation is well-defined since $\langle \mu_t \rangle$ has finite support; hence, it is conditioned on a positive measure event.

[44] Note that we have defined the time 1 strategy for all cases except this one.

– Following any other history, σ_A is optimal by definition.

- Principal's deviation: we verify that $\mathcal{V}(\sigma) = \mathbb{E}_f[V]$.

$$\begin{aligned}\mathcal{V}(\sigma) &= \int_{M'} V(\mu,1)\sigma_P(H_1)(\mathrm{d}\mu) + \int_M \mathcal{V}(\sigma'(\mu_0,\sigma_P(H_1),\mu))\sigma_P(H_1)(\mathrm{d}\mu)\\ &= \int V(\mu,1)f(\mathrm{d}\mu,1) + \int_M \mathbb{E}\left[V(\mu_\tau,\tau)|\mu_1=\mu, 1<\tau\right]\sigma_P(H_1)(\mathrm{d}\mu)\\ &= \int V(\mu,1)f(\mathrm{d}\mu,t) + \sum_{t>1}\int V(\mu,t)f(\mathrm{d}\mu,t)\\ &=\mathbb{E}_f[V].\end{aligned}$$

  Note that following any deviation of the principal, the agent may choose $\sigma_A = \delta_{stop}$ at any time to guarantee an interim payoff of at least $U(\widehat{\mu}_t,t)$. Hence, the belief-time distribution induced by the stategy profile satisfies (OC-C). Therefore, the deviation payoff must be no higher than $\mathbb{E}_f[V]$, which solves (R).

Finally, to verify that σ is an SPE^P it remains to verify $\mathcal{V}(\sigma) = \max_{\sigma'\in SPE(\Gamma)} \mathcal{V}(\sigma')$. Note that $\forall\sigma' \in SPE(\Gamma)$, Lemma 4 implies that $\mathcal{V}(\sigma')$ must be no higher than $\mathbb{E}_f[V]$ and the strategy σ we constructed achieves this upper-bound. *Q.E.D.*

Lemma 8. *If $V(\mu,t)$ strictly increases in t, then $\forall\sigma\in SPE^P(\Gamma)$ s.t. $\mu(\sigma)$ has finite support and* $\max\operatorname{supp}(\tau(\sigma)) < \overline{T}$, *for any stopping time $\iota\le\tau(\sigma)$, $\mathcal{U}(\sigma) = \mathbb{E}[U(\mu_\iota,\iota)]$.*

Proof. Suppose for the purpose of contradiction that there exists an on-path history H_{t+1} s.t. $\mathcal{U}(\sigma|H_{t+1},t^+) > U(\mu_t,t)$. Let σ' denote the strategy continuing from H_t i.e., $(\sigma|H_t,t)$. Let $f \sim (\mu(\sigma')_{\tau(\sigma')},\tau(\sigma'))$. Then, f satisfies (OC-C) as the agent may choose $\sigma_A = \delta_{stop}$ at any time $s^+ \ge t^+$ to guarantee an interim payoff of $U(\widehat{\mu}_s,s)$.

We claim that $\mathbb{E}_f[V] < W(\mu_t,t)$. Let t' be the earliest stopping time of f. Observe $t' < \overline{T}$ since we have assumed $\max\operatorname{supp}(\tau(\sigma)) < \overline{T}$. Define $f'(\mu,t) = f(\mu,t) - \xi\mathbf{1}_{t=t'}f(\mu,t') + \xi\mathbf{1}_{t=t'}f(\mu,t'+1)$ for $\xi\le 1$. Then, $\mathbb{E}_{f'}[U]$ is a continuous functional of ξ and for $\xi = 0$, it is strictly larger than $U(\mu_t,t)$. Since f' is continuous in ξ under the Prokhorov metric, there exists $\xi > 0$ s.t. $\mathbb{E}_{f'}[U] \ge U(\mu_t,t)$. Since V strictly increases in t, $\mathbb{E}_{f'}[V] - E_f[V] = \xi E_{f(\mu,t')}[V(\mu,t'+1) - V(\mu,t')] > 0$. Then, $W(\mu_t,t) \ge \mathbb{E}_{f'}[V] > \mathbb{E}_f[V]$.

However, Lemmas 5, 6 and 7 imply that $W(\mu_t,t) = \mathbb{E}_f[V]$ since $\mathbb{E}_f[V] = \mathcal{V}(\sigma') = \mathcal{V}(\sigma|H_t,t) = \max_{\sigma''\in SPE(\Gamma)} \mathcal{V}(\sigma''|H_t,t)$, contradiction. *Q.E.D.*

Putting the lemmas together, the sufficiency part of Theorem 2 (i) is an implication of Lemmas 4, 5, 6 and 7. The necessity part of Theorem 2 (i) is an implication of Lemma 4 plus the sufficiency part. The sufficiency part of Theorem 2 (ii) is an implication of Lemmas 5 and 7. The necessity part of Theorem 2 (ii) is Lemma 8. *Q.E.D.*

ONLINE APPENDIX TO 'PERSUASION AND OPTIMAL STOPPING'

ANDREW KOH SIVAKORN SANGUANMOO WEIJIE ZHONG

Online Appendix I: Omitted proofs

I.1 Simple recommendations are well-defined

Lemma 9. *Let $\left\{x^{(\mu,t)}\right\} \subset D_\infty$ be equipped with the Skorokhod topology. Mapping $\Phi : (\mu, t) \mapsto (x^{\mu,t}, t)$ is Borel measurable.*

Proof. Let σ be the Skorokhod metric. Then $\forall (\mu, t), (\mu', t')$ where $t < t'$:

$$\sigma(x^{(\mu,t)}, x^{(\mu',t')}) \leq \max\left\{|\mu - \mu'|, |t - t'|, \max_{s,s' \in [t,t')} \{|\widehat{\mu}_s - \widehat{\mu}_{s'}|\}\right\},$$

where the RHS is achieved via contracting time on $[t, t')$.

If $\widehat{\mu}_t$ is continuous at t. Then, $(\mu_n, t_n) \to (\mu, t)$ implies $\sigma(x^{(\mu_n,t_n)}, x^{(\mu,t)}) \to 0$. Hence, the mapping is Borel measurable (continuous) at (μ, t).

If $\widehat{\mu}_t$ is dis-continuous at t. Then, for all $t' > t$, $\sigma(x^{(\mu,t)}, x^{(\mu',t')}) \geq |\widehat{\mu}_t - \widehat{\mu}_{t+}| > 0$. Meanwhile, $(\mu_n, t_n) \to (\mu, t-)$ implies $\sigma(x^{(\mu_n,t_n)}, x^{(\mu,t)}) \to 0$. Then, for sufficiently small open ball around $x^{(\mu,t)}$, the inverse image is an open ball around (μ, t) truncated by $t' \leq t$; hence, a Borel set. Therefore, the mapping is Borel measurable at (μ, t). *Q.E.D.*

I.2 Proof of Lemma 1

Proof. The first statement: Suppose (2) is violated, i.e. $\exists t$ s.t.

$$\int_{y>t} U(\mu, y) f(\mathrm{d}\mu, \mathrm{d}y) < U\left(\frac{\int_{y>t} \mu f(\mathrm{d}\mu, \mathrm{d}y)}{\int_{y>t} f(\mathrm{d}\mu, \mathrm{d}y)}, t\right)$$

Then, define $\tau' = \min\{\tau, t\}$.

$$\begin{aligned}
\mathbb{E}\left[U(\mu_{\tau'}, \tau')\right] =& \text{Prob}(\tau \leq t)\mathbb{E}\left[U(\mu_\tau, \tau)|\tau \leq t\right] + \text{Prob}(\tau > t)\mathbb{E}\left[U(\mu_t, t)|\tau > t\right] \\
\geq& \text{Prob}(\tau \leq t)\mathbb{E}\left[U(\mu_\tau, \tau)|\tau \leq t\right] + \text{Prob}(\tau > t)U(\mathbb{E}\left[\mu_t|\tau > t\right], t) \\
& \text{(Jensen's inequality)} \\
=& \text{Prob}(\tau \leq t)\mathbb{E}\left[U(\mu_\tau, \tau)|\tau \leq t\right] + \text{Prob}(\tau > t)U(\mathbb{E}\left[\mu_\tau|\tau > t\right], t) \\
& \text{(Optional stopping theorem)} \\
>& \text{Prob}(\tau \leq t)\mathbb{E}\left[U(\mu_\tau, \tau)|\tau \leq t\right] + \text{Prob}(\tau > t)\mathbb{E}\left[U(\mu_\tau, \tau)|\tau > t\right] \\
=& \mathbb{E}[U(\mu_\tau, \tau)].
\end{aligned}$$

(OC) is violated. $f \in \Delta_{\mu_0}$ is a direct implication of the optional stopping theorem.

The second statement: First, we show that $(\mu^f_{\tau^f}, \tau^f) \sim f$. $\forall$ Borel sets $B_\mu \subset \Delta(\Theta)$, $B_t \subset T$,

$$(\mu^f_{\tau^f}, \tau^f) \subset B_\mu \times B_t \iff t \in B_t \ \&\ x^{(\mu,t)}(t) \subset B_\mu \iff (\mu, t) \subset B_\mu \times B_t.$$

Therefore, $\mathcal{P}\left((\mu^f_{\tau^f}, \tau^f) \subset B_\mu \times B_t\right) = \int_{B_\mu \times B_t} f(\mathrm{d}\mu, \mathrm{d}t)$.

Second, we show that $\langle \mu_t^f \rangle$ is a martingale. Take any Borel subset $B \subset \mathcal{F}_t$ such that $\mathcal{P}(B) > 0$. B can be further divided into two sets of events (that are elements of $\mathcal{F}_t$ since τ is a stopping time).

- Case 1: $\tau^f \leq t$. $\forall t' > t$,

$$\begin{aligned}\mathbb{E}\left[\mu_{t'}^f | B\right] =& \mathbb{E}\left[\mu_{t'}^f | B, \tau^f \leq t\right] \\ =& \frac{1}{\mathcal{P}(B)} \int_{(x^{(\mu,s)},s)\in B} x^{(\mu,s)}(t') \mathrm{d}\mathcal{P} \\ =& \frac{1}{\mathcal{P}(B)} \int_{(x^{(\mu,s)},s)\in B} x^{(\mu,s)}(t) \mathrm{d}\mathcal{P} \\ =& \mathbb{E}\left[\mu_t^f | B\right]\end{aligned}$$

  The third line is from the definition of $x^{(\mu,s)}$ and $t' > t \geq s$.

- Case 2: $\tau^f > t$. In this case, $\forall \mu$ and $s > t$, $(x^{(\mu,s)}, s) \in B$ since all those $x^{(\mu,s)} = \widehat{\mu}_t$ up to period t. Therefore,

$$\begin{aligned}\mathbb{E}\left[\mu_{t'}^f | B\right] =& \mathbb{E}\left[\mu_{t'}^f | \tau^f > t\right] \\ =& \frac{1}{\mathcal{P}(B)} \int_{s>t} x^{\mu,s}(t') \mathrm{d}\mathcal{P} \\ =& \frac{1}{\mathcal{P}(B)} \left(\int_{s>t'} x^{\mu,s}(t') \mathrm{d}\mathcal{P} + \int_{t<s\leq t'} x^{\mu,s}(t') \mathrm{d}\mathcal{P} \right) \\ =& \frac{1}{\mathcal{P}(B)} \left(\int_{s>t'} \widehat{\mu}_{t'} \mathrm{d}\mathcal{P} + \int_{t<s\leq t'} \mu f(\mathrm{d}\mu, \mathrm{d}s) \right) \\ =& \frac{1}{\int_{s>t} f(\mathrm{d}\mu, \mathrm{d}s)} \left(\widehat{\mu} \cdot \int_{s>t'} f(\mathrm{d}\mu, \mathrm{d}s) + \int_{t<s\leq t'} \mu f(\mathrm{d}\mu, \mathrm{d}s) \right) \\ =& \widehat{\mu}_t \\ =& \mathbb{E}[\mu_t^f | \tau^f > t]\end{aligned}$$

Third, we verify (OC). By the definition of $\langle \mu_t^f \rangle$, conditional on the event $\tau^f \leq t$, μ_t^f is constant in t. Let $\widetilde{\tau} := \min\left\{\tau^f, \tau'\right\}$ and $\delta\tau := (\tau' - \tau^f)^+$. Then, $(\mu_{\widetilde{\tau}}, \tau')$ has the same distribution as $(\mu_{\tau'}, \tau')$.

$$\begin{aligned}\mathbb{E}\left[U(\mu_{\tau'}, \tau')\right] =& \mathcal{P}(\tau^f > \tau')\mathbb{E}\left[U(\mu_{\tau'}, \tau') | \tau^f > \tau'\right] + \mathcal{P}(\tau^f \leq \tau')\mathbb{E}\left[U(\mu_{\tau^f}, \tau^f + \delta\tau) | \tau^f \leq \tau'\right] \\ \leq& \mathcal{P}(\tau^f > \tau')\mathbb{E}\left[U(\mu_{\tau'}, \tau') | \tau^f > \tau'\right] + \mathcal{P}(\tau^f \leq \tau')\mathbb{E}\left[U(\mu_{\tau^f}, \tau^f) | \tau^f \leq \tau'\right] \\ & \text{(Because } \delta\tau \geq 0. \text{)} \\ =& \mathcal{P}(\tau^f > \tau')\mathbb{E}\left[U(\widehat{\mu}_{\tau'}, \tau') | \tau^f > \tau'\right] + \mathcal{P}(\tau^f \leq \tau')\mathbb{E}\left[U(\mu_{\tau^f}, \tau^f) | \tau^f \leq \tau'\right] \\ \leq& \mathcal{P}(\tau^f > \tau')\mathbb{E}\left[\mathbb{E}_f\left[U(\mu, t) | t > \tau'\right] | \tau^f > \tau'\right] + \mathcal{P}(\tau^f \leq \tau')\mathbb{E}\left[U(\mu_{\tau^f}, \tau^f) | \tau^f \leq \tau'\right] \\ & \text{(Implied by (2).)}\end{aligned}$$

$$=\mathcal{P}(\tau^f > \tau')\mathbb{E}\left[U(\mu_{\tau^f},\tau^f)|\tau^f > \tau'\right] + \mathcal{P}(\tau^f \le \tau')\mathbb{E}\left[U(\mu_{\tau^f},\tau^f)|\tau^f \le \tau'\right]$$
$$=\mathbb{E}\left[U(\mu_{\tau^f},\tau^f)\right].$$

Q.E.D.

I.3 Proof of Lemma 2

Proof of Lemma 2. The lemma is trivial when T is finite. We prove the case when T is a continuum. We first prove strong duality. Define the mapping G:

$$G(f)(t) = \int_{\tau>t} U(\mu,\tau)f(\mathrm{d}\mu,\mathrm{d}\tau) - U\left(\int_{\tau>t}\mu f(\mathrm{d}\mu,\mathrm{d}\tau),t\right).$$

Since U is convex, $G(\cdot,t)$ is concave. Since U is bounded, G maps Δ_{μ_0} into $L^\infty(T^\circ)$.

Next, we verify that there exists some $f\in\Delta_{\mu_0}$ s.t. $G(f)(\cdot)$ is an interior point (w.r.t. L^∞ norm) of the positive cone. Let $\widehat{U}(\mu,t)$ denote $\mathbb{E}_\mu[U(\delta_\theta,t)]$. Since T is compact and U is continuous, Assumption 1 implies that there exists $\epsilon>0$ s.t. $\widehat{U}(\mu_0,t)-U(\mu_0,t)\ge 3\epsilon$ for all $t\in T$. Since U is continuous, there exists a finite partition $\{t_i\}_{i=0}^I\subset T$ s.t. $t_0=0$ and $|U(\mu_0,t)-U(\mu_0,t_i)|\le\epsilon$ when $t\in T$ and t_i are adjacent. Therefore, $\forall t\in(t_i,t_{i+1}]\cap T$,

$$U(\mu_0,t)\le\widehat{U}(\mu_0,t_{i+1})-2\epsilon.$$

Let $M=\sup_t U(\mu_0,t)$. Wlog, we can pick $\epsilon<\frac{1}{3}M$. Next, we define f:

$$\text{for } i=1,\dots,I-1, \begin{cases} f(t_i)=\left(\frac{\epsilon}{M}\right)^{i-1}\left(1-\frac{\epsilon}{M}\right);\\ f(\mu|t_i)\text{ achieves }\widehat{U}(\mu_0,t_i).\end{cases}$$
$$\begin{cases} f(t_I)=\left(\frac{\epsilon}{M}\right)^{I};\\ f(\mu|t_I)\text{ achieves }\widehat{U}(\mu_0,t_I).\end{cases}$$

By definition of $\widehat{U}$, $\forall i$, $f(\mu|t_i)$ is well-defined since the upper concave hull can be achieved by a finite distribution per Caratheodory's theorem. Then, for any $i=0,\dots,|I|-1$ and $t\in(t_i,t_{i+1}]$,

$$G(f)(t)=\sum_{j=i+1}^{I-1}\left(\left(\frac{\epsilon}{M}\right)^{j-1}\left(1-\frac{\epsilon}{M}\right)\right)\cdot\widehat{U}(\mu_0,t_j)+\left(\frac{\epsilon}{M}\right)^{I}\widehat{U}(\mu_0,t_I)-U(\mu_0,t)\cdot\left(\sum_{j=i+1}^{I-1}\left(\left(\frac{\epsilon}{M}\right)^{j-1}\left(1-\frac{\epsilon}{M}\right)\right)+\left(\frac{\epsilon}{M}\right)^{I}\right)$$
$$\ge\left(\frac{\epsilon}{M}\right)^{i}\left(1-\frac{\epsilon}{M}\right)\widehat{U}(\mu_0,t_{i+1})-\left(\frac{\epsilon}{M}\right)^{i}\cdot U(\mu_0,t)$$
$$\ge\left(\frac{\epsilon}{M}\right)^{i}\left(1-\frac{\epsilon}{M}\right)(U(\mu_0,t_{i+1})+2\epsilon)-\left(\frac{\epsilon}{M}\right)^{i}\cdot U(\mu_0,t)$$
$$=\left(\frac{\epsilon}{M}\right)^{i}\left(2\epsilon-\frac{2\epsilon^2}{M}-\frac{\epsilon}{M}U(\mu_0,t)\right)$$
$$\ge\left(\frac{\epsilon}{M}\right)^{I}\cdot\frac{\epsilon}{3}.$$

Therefore, $G(f)(\cdot)$ is bounded away from 0 by at least $\left(\frac{\epsilon}{M}\right)^I \cdot \frac{\epsilon}{3}$ under the L^∞ norm.

The following lemma establishes that for any target joint distribution f, we can find a "nearby" continuous distribution which does not alter the value of G too much. Let $\Delta_{\mu_0}^C$ be the subset of Δ_{μ_0} such that $G(f)(t)$ is uniformly continuous on T°.

Lemma 10. *Given Assumption 2, for all $f \in \Delta_{\mu_0}$ and all ϵ, there exists $f' \in \Delta_{\mu_0}^C$ s.t. $\mathbb{E}_{f'}[V] \geq \mathbb{E}_f[V] - \epsilon$ and $G(f')(t) \geq G(f)(t) - \epsilon$.*

We now take stock of the conditions to apply Theorem 1, Chapter 8.6 of Luenberger (1997). $\int V(\mu,\tau) f(d\mu, d\tau)$ is a linear functional. G is a concave mapping, and from Lemma 10 there exists $f \in \Delta_{\mu_0}^C$ such that $G(f)(\cdot)$ is in the interior of the positive cone. Further, since $f \in \Delta_{\mu_0}^C$, $G(f)(\cdot)$ is uniformly continuous on T° hence $\mathcal{B}(T^\circ)$ is the appropriate dual space. We then have

$$\sup_{f \in \Delta_{\mu_0}^C} \inf_{\Lambda \in \mathcal{B}(T^\circ)} \mathcal{L}(f,\Lambda) = \min_{\Lambda \in \mathcal{B}(T^\circ)} \sup_{f \in \Delta_{\mu_0}^C} \mathcal{L}(f,\Lambda),$$

where the minimum is achieved by $\Lambda \in \mathcal{B}(T^\circ)$.

Then note that for all $f \in \Delta_{\mu_0}$, $\Lambda \in \mathcal{B}(T^\circ)$ and $\epsilon > 0$, Lemma 10 implies that there exists $f' \in \Delta_{\mu_0}^C$ such that $\mathcal{L}(f',\Lambda) \geq \mathcal{L}(f,\Lambda) - \epsilon$. Therefore,

$$\sup_{f \in \Delta_{\mu_0}^C} \inf_{\Lambda \in \mathcal{B}(T^\circ)} \mathcal{L}(f,\Lambda) \leq \sup_{f \in \Delta_{\mu_0}} \inf_{\Lambda \in \mathcal{B}(T^\circ)} \mathcal{L}(f,\Lambda) \leq \min_{\Lambda \in \mathcal{B}(T^\circ)} \sup_{f \in \Delta_{\mu_0}} \mathcal{L}(f,\Lambda) = \min_{\Lambda \in \mathcal{B}(T^\circ)} \sup_{f \in \Delta_{\mu_0}^C} \mathcal{L}(f,\Lambda).$$

The inner inequality must be equality since the smallest term is the same as the largest term.

Finally, we verify that there exists f achieving the maximum when V is upper semicontinuous. This implies that $\mathbb{E}_f[V]$ is upper semicontinuous. Then, it is sufficient to verify that the set of f s.t. $G(f) \geq 0$ is closed. Suppose not, then, there exists $f_n \to f$ s.t. $G(f_n) \geq 0$ but $G(f)(t) < 0$. Consider

$$\gamma_\epsilon(\tau) = \begin{cases} 0 & \tau \leq t \\ \frac{\tau - t}{\epsilon} & \tau \in (t, t+\epsilon) \\ 1 & \tau > t + \epsilon \end{cases}.$$

Then, $\int U \cdot \gamma_\epsilon \mathrm{d}f \to \int_{\tau > t} U \mathrm{d}f$ and $\int \mu \cdot \gamma_\epsilon \mathrm{d}t \to \int_{\tau > t} \mu \mathrm{d}f$ when $\epsilon \to 0$. For ϵ sufficiently small, $\int U \cdot \gamma_\epsilon \mathrm{d}f - U\left(\int \mu \cdot \gamma_\epsilon \mathrm{d}f, t\right) < 0$. However, since $U \cdot \gamma_\epsilon$ and $\mu \cdot \gamma_\epsilon$ are bounded and continuous functions,

$$\int U \cdot \gamma_\epsilon \mathrm{d}f - U\left(\int \mu \cdot \gamma_\epsilon \mathrm{d}f, t\right)$$
$$= \lim_{n \to \infty} \int U \cdot \gamma_\epsilon \mathrm{d}f_n - U\left(\int \mu \cdot \gamma_\epsilon \mathrm{d}f_n, t\right).$$

Then, there exists n s.t. $\int U \cdot \gamma_\epsilon \mathrm{d}f_n - U\left(\int \mu \cdot \gamma_\epsilon \mathrm{d}f_n, t\right) < 0$. Note that $f_n \cdot \gamma_\epsilon$ is a convex combination of $\mathbf{1}_{t>s} f_n$ for $s \in [t, t+\epsilon]$ and $G(\mathbf{1}_{t>s} f_n)(T) = G(f_n)(s) \geq 0$. Then, $\int U \cdot \gamma_\epsilon \mathrm{d}f_n -$

$U\left(\int \mu\cdot\gamma_\epsilon \mathrm{d}f_n, t\right) = G(f_n\cdot\gamma_\epsilon, T)\geq 0$ since G is concave. Contradiction. *Q.E.D.*

Proof of Lemma 10. $\forall f \in \Delta_{\mu_0}$, $G(f)(t)$ has bounded variation and only jumps down. Therefore, $G(f)(t)$ can be decomposed into $g(t)+h(t)$, where g is bounded and continuous and h is bounded and decreasing. Define the "delayed" measure

$$f^s(\mu,t) := \begin{cases} 0 & t<s \\ f(\mu,t-s) & t\in[s,\sup(T)) \\ f(\mu,[t-s,\sup(T)]) & t=\sup(T) \end{cases}$$

In words, f^s delays the distribution of f by s. Pick $\delta>0$ as a continuity parameter corresponding to $\frac{1}{2}\epsilon$ for U, V and g.[45] Then,

$$\begin{aligned} G(f^s)(t) &= \int_{\tau>t-s} U\mathrm{d}f - U\left(\int_{\tau>t-s}\mu\mathrm{d}f, t\right) \\ &\geq \int_{\tau>t-s} U\mathrm{d}f - U\left(\int_{\tau>t-s}\mu\mathrm{d}f, t-s\right) - \frac{1}{2}\epsilon \\ &= g(t-s)+h(t-s)-\frac{1}{2}\epsilon \\ &\geq g(t)+h(t)-\frac{1}{2}\epsilon \\ &= G(f)(t)-\frac{1}{2}\epsilon; \\ \mathbb{E}_{f^s}[V] &= \int_0^{\sup(T)-s} V(\mu,t+s)\mathrm{d}f + \int_{\sup(T)-s}^{\sup(T)} V(\mu,\sup(T))\mathrm{d}f \\ &\geq \mathbb{E}_f[V]-\frac{1}{2}\epsilon. \end{aligned}$$

Let $\widehat{f}$ be the uniform randomization of f^s, for $s\in[0,\delta]$. Then, since $\mathbb{E}_f[V]$ is linear operator of f, $\mathbb{E}_{\widehat{f}}[V]\geq \mathbb{E}_f[V]-\frac{1}{2}\epsilon$. Since G is a concave operator of f, $G(\widehat{f})(t)\geq G(f)(t)-\epsilon$.

Next, we prove the uniform continuity of $G(\widehat{f})$. Note that $\forall t<t'<t+\delta$,

$$\begin{aligned} &\widehat{f}((t,t']) \\ =&\frac{1}{\delta}\int_{s=0}^{\delta}(F(t'-s)-F(t-s))\mathrm{d}s \\ \leq&\frac{1}{\delta}\int_{(t,t']\cup(t-\delta,t'-\delta]} F(s)\mathrm{d}s \\ \leq&\frac{2|t-t'|}{\delta}. \end{aligned}$$

Therefore,

$$\left|\int_{\tau>t} U\mathrm{d}\widehat{f} - \int_{\tau>t} U\mathrm{d}\widehat{f}\right| = \left|\int_{(t,t']} U\mathrm{d}\widehat{f}\right|$$

[45] All the continuity are uniform since T and D are compact.

$$\leq |U| \hat{f}((t,t']) \leq |U| \cdot \frac{2|t-t'|}{\delta}.$$

Then, $\forall \epsilon$, let $|U| \cdot \frac{2|t-t'|}{\delta} < \frac{\epsilon}{2}$ and $2|t-t'|/\delta$ be less than than the continuity parameter of U for $\frac{\epsilon}{2}$. Then, $\left|U\left(\int_{\tau>t} \mu \mathrm{d}f, t\right) - U\left(\int_{\tau>t'} \mu \mathrm{d}f, t\right)\right| < \frac{\epsilon}{2}$. To sum up, $|G(\hat{f})(t) - G(\hat{f})(t')| < \epsilon$.

Q.E.D.

I.4 Proof of Lemma 3

Proof. Per the proof of Lemma 2, there exists $f^\circ \in \Delta_{\mu_0}$ and $1 > \xi > 0$ s.t. $G(f^\circ) \geq \xi$. Let δ be the continuity parameter of U given $\frac{1}{2}\xi\epsilon$. Define

$$\tilde{f} = \left(1 - \frac{1}{2}\epsilon\right) f + \frac{1}{2}\epsilon f^\circ.$$

Then, since G is concave, $G(\tilde{f}) \geq \frac{1}{2}\epsilon\xi$. Let $\hat{f}$ be the discretization of $\tilde{f}$ by pooling all mass in $(t_j, t_{j+1}]$ to t_j. By construction $d_{lp}(f, \hat{f}) \leq d_{lp}(f, \tilde{f}) + d_{lp}(\tilde{f}, \hat{f}) \leq \frac{1}{2}\epsilon + \frac{1}{2}\epsilon\xi < \epsilon$. $\forall j$,

$$\begin{aligned} G(\hat{f})(t_j) \geq & \int_{\tau > t_{j+1}} U \mathrm{d}\tilde{f} - U\left(\int_{\tau>t_{j+1}} \mu \mathrm{d}\tilde{f}, t_j\right) \\ \geq & \int_{\tau > t_{j+1}} U \mathrm{d}\tilde{f} - U\left(\int_{\tau>t_{j+1}} \mu \mathrm{d}\tilde{f}, t_{j+1}\right) - \frac{1}{2}\epsilon\xi \\ = & G(\tilde{f})(t_{j+1}) - \frac{1}{2}\epsilon\xi \geq 0. \end{aligned}$$

The converse direction is due to Δ_{μ_0} being a tight collection and the set of f s.t. $G(f) \geq 0$ being closed, as is established in the proof of Lemma 2. *Q.E.D.*

I.5 Proof of Proposition 1

Proof. First, we show that such an *SPE* exists. Let $\sigma_A(H_t)$ be *continue* whenever μ_t is interior and *stop* otherwise. Let $\sigma_P(H_t)$ be full revelation at all histories. It is trivial that the agent finds no profitable deviation as he receives his first best at any history. Given the agent's strategy, at any history, the principal has to eventually fully reveal the state. Otherwise, if she does not fully reveal the state so that the agent stops at $\overline{T}$, she would be better off by fully revealing at $\overline{T}$ by the condition (ii) of Proposition 1. Therefore, revealing it immediately is optimal as the principal is strictly impatient.

Now consider uniqueness. The agent can guarantee him the first best by always choosing *continue*. Therefore, the maximal principal's payoff that is consistent with the agent first best payoff is from full revelation at $t = 1$, which is an *SPE* as we have proved. *Q.E.D.*

I.6 Proof of Proposition 2

Proof for necessity. Under the Suspense^{t_1}-ℓ^{t_2} strategy, the corresponding joint distribution of stopping time and action is

$$\mathbb{P}(\tau \leq t, a = \ell) = 1 - (2\mu_0 - t_1)e^{t_1 - t}, \qquad \mathbb{P}(\tau = t_2, a = r) = (2\mu_0 - t_1)e^{t_1 - t_2}$$

for every $t \in [t_1, t_2]$. The principal's payoff under the Suspense^{t_1}-ℓ^{t_2} strategy is

$$V_\ell^*(t_1, t_2) := (1 - (2\mu_0 - t_1))(v_\ell + h_\ell(t_1)) + (2\mu_0 - t_1)\left(\int_{t_1}^{t_2} e^{-t+t_1}(v_\ell + h_\ell(t))\mathrm{d}t + e^{-t_2+t_1}(v_r + h_r(t_2))\right).$$

Optimality implies the first order condition w.r.t. t_2:

$$\frac{\mathrm{d}V_\ell^*(t_1, t_2)}{\mathrm{d}t_2} = (2\mu_0 - t_1)e^{-t_2+t_1}(\Delta v + \Delta h(t_2) - h_r'(t_2)) = 0,$$

which implies condition (a): $\Delta v + \Delta h(t_2) = h_r'(t_2)$ and the local SOC: $h_r''(t_2) \le \Delta h'(t_2)$ in condition (c). Let $\Delta = t_2 - t_2$, consider the FOC w.r.t. shifting t_1 and t_2 jointly:

$$\frac{\mathrm{d}V_\ell^*(t_1, t_1 + \Delta)}{\mathrm{d}t_1} = \underbrace{(2\mu_0 - t_1 - 1)}_{<0}\left(\Psi(t_1, t_2) - h_\ell'(t_1)\right) + e^{-t_2+t_1}\underbrace{(\Delta v + \Delta h(t_2) - h_r'(t_2))}_{=0 \text{ from (a)}},$$

which gives condition (b): $\Psi(t_1, t_2) \ge h_\ell'(t_1)$ with equality when $t_1 > 0$ and the local SOC: $h_\ell''(t_1) \cdot t_1 \le 0$ in condition (c), as desired. *Q.E.D.*

Proof for sufficiency. Note that by the construction of the strategy, (OC-C) is slack for $0 < t < t_1$. Therefore, the complementary slackness condition implies $\Lambda(t)$ must be constant prior to t_1, i.e., $\Lambda(0+) = \Lambda(t_1)$. The principle's FOC from choosing stopping belief ν at time $t > t_1$ is

$$l_{f,\Lambda}(\nu, t) = \begin{cases} v_0 + h_\ell(t) + \Lambda(t) - \Lambda(t)c(t) + \int_{t_1}^t c(s)\mathrm{d}\Lambda(s) - \Lambda(t_1) - 2\nu(\Lambda(t) - \Lambda(t_1)) & \nu < 0.5 \\ v_1 + h_r(t) - \Lambda(t)c(t) + \int_{t_1}^t c(s)\mathrm{d}\Lambda(s) - \Lambda(t_1) + 2\nu\Lambda(t_1) & \nu \ge 0.5 \end{cases}$$

Note that we select $\nabla U(0, t)$ to be $\nabla U(\mu_{t_1}^*(t^*), t)$. Define $l_{f,\Lambda}^*(\nu, t) = l_{f,\Lambda}(\nu, t) - 2\nu\Lambda(t_1)$ as an affine transformation of $l_{f,\Lambda}$. The key observation is $l_{f,\Lambda}^*$ as a function of ν takes a simple piecewise linear structure: it linearly decreases on $[0, 0.5)$ at rate $2\Lambda(t)$ and stays constant on $[0.5, 1]$, as is depicted by Figure 7. We show each part of Corollary 3.1 in turn.

The concavification method suggests that the following three conditions are sufficient for the optimality of f: (i) $l_{f,\Lambda}^*(0, t)$ is a constant function of t for $t \ge t_1$, (ii) $l_{f,\Lambda}^*(0, t) \le l_{f,\Lambda}(0, t_1)$ for $t < t_1$, and (iii) $l_{f,\Lambda}^*(\mu_{t_1}^*(t_2), t)$ is maximized at t_2. We construct Λ:

$$\Lambda(t) = \begin{cases} \Psi(t, t_2) := \int_t^{t_2} e^{t-s}h_\ell'(s)\mathrm{d}s + e^{t-t_2}h_r'(t_2), & \forall t \ge t_1 \\ \Psi(t_1, t_2), & \forall t < t_1 \end{cases},$$

where t_2 is given by $\Delta v + \Delta h(t_2) = h_r'(t_2)$. When $t_1 > 0$, condition (b) guarantees that $\Lambda'(t_1) = 0$, i.e. Λ is a smooth function. Note that Λ is monotonically increasing if $\forall t \ge t_1$,

$$\Lambda'(t) = \Psi_t'(t, t_2) \ge 0, \tag{4}$$

which is given by condition (d). Next, we verify the three conditions.

Condition (i) is satisfied if $\forall t \ge t_1$, Λ satisfies

$$\frac{l_{f,\Lambda}^*(0, t)}{\partial t} = h_\ell'(t) + \Lambda'(t) - \Lambda(t) = 0,$$

which can be verified by direct calculation.

Condition (ii) has a bite only when $t_1 > 0$, in which case, it is satisfied if $\frac{\partial l^*_{f,\Lambda}(0,t)}{\partial t}|_{t=t_1-} \geq 0$ and $\frac{\partial^2 l^*_{f,\Lambda}(0,t)}{\partial t^2} \leq 0$ for $t < t_1$. The former condition is implied by condition (b). The latter condition is implied by (d).

Condition (iii): we verify the optimality of $\mu^*_{t_1}(t_2)$ and t_2 separately.

- $\mu^*_{t_1}(t_2) \in (0.5, 1)$ is optimal $\iff$ $l^*_{f,\Lambda}(\cdot, t_2)$ attains the same value at 0 and $\mu^*_{t_1}(t_2)$ $\iff$ $\Delta v + \Delta h(t_2) = \Lambda(t_2) = h'_r(t_2)$ (condition (a)). See the illustration in Figure 7, where the colored lines are $l^*_{f,\Lambda}(\nu, \hat{t})$ and the dashed line is the concave envelope.
- t_2 is optimal if $\frac{\partial l^*_{f,\Lambda}(1,t)}{\partial t}|_{t=t_2} = 0$ and $\frac{\partial^2 l^*_{f,\Lambda}(1,t)}{\partial t^2} \leq 0$. The former is equivalent to $h'_r(t_2) - \Lambda(t_2) = 0$, which is implied by the definition of Λ. The latter is implied by $h''_r(t) - \Lambda'(t) \leq 0$, i.e.
$$\begin{cases} h''_r(t) \leq 0, & \forall t < t_1 \\ \Psi'_t(t, t_2) - h''_r(t) \geq 0, & \forall t \geq t_1 \end{cases} \tag{5}$$
which are given by condition (d).

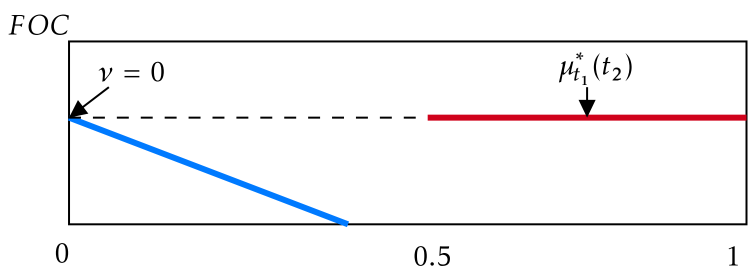


Figure 7: Concavification condition for *Suspense*^{t_1}-ℓ^{t_2} at $t = t_2$.

I.7 Proof of Proposition 5

Proof. Let Λ be the multiplier that solves (D). We claim that $\Lambda(t)$ must be strictly increasing on $[0, \bar{t}]$. Suppose for the purpose of contradiction that it is not true, i.e., there exists interval $[t_1, t_2]$ s.t. $\Lambda(t) = \kappa$ on the interval. Wlog, let $[t_1, t_2]$ be a maximal interval such that this is true. Since $\bar{t} < \overline{T}$, $t_2 < \overline{T}$. [46]

Define $\xi(t) = \max_\mu (l_{f,\Lambda}(\mu, t) - a \cdot \mu)$. Then, (FOC) implies that $\xi(t) \leq 0$ and $\xi(t_2) = 0$. Note that $l_{f,\Lambda}(\mu, t)$ is left continuous and only jumps up in t. The envelope theorem implies
$$\xi'(t_2^-) = V'(\mu, t_2) + \kappa U'(\mu, t_2),$$
where μ attains $\xi(t_2)$. Next, we calculate a lower bound for "$\xi'(t_2^+)$".
$$\begin{aligned} &\xi(t_2 + \delta) - \xi(t_2) \\ \geq &l_{f,\Lambda}(\mu, t_2 + \delta) - l_{f,\Lambda}(t_2) \\ = &V(\mu, t_2 + \delta) - V(\mu, t_2) + \Big(\Lambda(t_2 + \delta) - \kappa + \kappa\Big) U(\mu, t_2 + \delta) - \kappa U(\mu, t_2) - \int_{t_2}^{t_2+\delta} \nabla_\mu U(\widehat{\mu}_t, t) \cdot \mu \Lambda(\mathrm{d}t) \end{aligned}$$

[46] In the special case where $\overline{T} = \infty$, $t_2 = \infty$ is already ruled out becase this will imply $t_1 \geq \bar{t}$.

$$
\begin{aligned}
=&V(\mu,t_2+\delta)-V(\mu,t_2)+\kappa\Big(U(\mu,t_2+\delta)-U(\mu,t_2)\Big)\\
&+\Big(\Lambda(t_2+\delta)-\kappa\Big)U(\mu,t_2+\delta)-\int_{t_2}^{t_2+\delta}\nabla_\mu U(\widehat{\mu}_t,t)-\cdot\mu\Lambda(\mathrm{d}t)\\
\geq&V(\mu,t_2+\delta)-V(\mu,t_2)+\kappa\Big(U(\mu,t_2+\delta)-U(\mu,t_2)\Big)+\int_{t_2}^{t_2+\delta}\Big(U(\mu,t)-\nabla_\mu U(\widehat{\mu}_t,t)\cdot\mu\Big)\Lambda(\mathrm{d}t)\\
\geq&V(\mu,t_2+\delta)-V(\mu,t_2)+\kappa\Big(U(\mu,t_2+\delta)-U(\mu,t_2)\Big)\\
=&\xi'(t_2^-)\delta+o(\delta),
\end{aligned}
$$

where the second inequality is because $U(\mu,t_2+\delta)\leq U(\mu,t_2)$, and the third inequality is because U is convex in μ.

Suppose $V_t''>0$, $U_t''\geq 0$, then $V+\kappa U$ is strictly convex in t. Suppose $V_t'>0$, then we claim that $\kappa>0$. If not, $\xi(t_2+\delta)\geq V'(\mu,t_2)\delta+o(\delta)>0$, violating the (FOC) around t_2. Then, since $\kappa>0$, $V+\kappa U$ is strictly convex in t.

Therefore, $l_{f,\Lambda}(\mu,t)-a\cdot\mu$ is strictly convex in t on $[t_1,t_2]$ for fixed μ. Since $\xi(t):=\max_\mu(l_{f,\Lambda}(\mu,t)-a\cdot\mu)$ is the upper envelope of a collection of strictly convex functions, it is also strictly convex on $[t_1,t_2]$. Hence, $\xi'(t_2^-)>0$ which violates the (FOC) at t_2^+, leading to a contradiction.

Since $\Lambda(t)$ must be strictly increasing, the complementarity slackness condition implies that (OC-C) must be binding all the time. *Q.E.D.*

I.8 Proof of Proposition 6

Proof. Let $\xi(t)=\max_\mu(l_{f,\Lambda}(\mu,t)-a\cdot\mu)$. Then, (FOC) implies that $\xi(t)\leq 0$ and $\xi(\underline{t})=\xi(\overline{t})=0$. The envelope theorem implies $\xi'(\underline{t}^-)=V'(\mu,\underline{t})+\Lambda(0)U'(\mu,\underline{t})\geq 0$, where μ is the maximizer that attains $\xi(\underline{t})$. Suppose for the purpose of contradiction that $\overline{t}>\max\left\{\overline{J}_V^{-1}\circ\underline{J}_V^{-1}(\underline{t}),\overline{J}_U^{-1}\circ\underline{J}_U^{-1}(\underline{t})\right\}$. Then, $\exists\epsilon>0$ s.t. $\forall\mu\in\Delta(\Theta)$, $\forall t>\overline{t}-\epsilon$

$$
V_t'(\mu,t)+\Lambda(t)U_t'(\mu,t)<-\epsilon.
$$

Since $\widetilde{\mu}_t\to\mu^*$ and $\frac{\Lambda(\overline{t})-\Lambda(t)}{\overline{t}-t}$ is bounded when $t\to\overline{t}^+$, there exists $\delta>0$ s.t. $\delta<\epsilon$ and $\forall t\geq\overline{t}-\delta$

$$
\frac{\Lambda(\overline{t})-\Lambda(\overline{t}-\delta)}{\delta}(U(\mu^*,t)-\nabla_\mu U(\widehat{\mu}_t,t)\cdot\mu^*)<\frac{1}{2}\epsilon.
$$

Then, $\forall t_1,t_2$ s.t. $\overline{t}-\delta<t_1<t_2<\overline{t}$,

$$
\begin{aligned}
&(l_{f,\Lambda}(\mu^*,t_2)-a\cdot\mu^*)\\
=&(l_{f,\Lambda}(\mu^*,t_1)-a\cdot\mu^*)+\int_{t_1}^{t_2}V_t'(\mu^*,t)+\Lambda(t)U_t'(\mu^*,t)\mathrm{d}t+\int_{t_1}^{t_2}(U(\mu^*,t)-\nabla_\mu U(\widehat{\mu}_t,t)\cdot\mu^*)\Lambda(\mathrm{d}t)\\
<&-\epsilon(t_2-t_1)+\frac{\epsilon\delta}{2(\Lambda(\overline{t})-\Lambda(\overline{t}-\delta))}\int_{t_1}^{t_2}\Lambda(\mathrm{d}t)
\end{aligned}
$$

$$\to -\frac{\epsilon\delta}{2} \quad \text{when} t_2 \to \bar{t}, t_1 = \bar{t} - \delta.$$

The analysis above implies that $l_{f,\Lambda}(\mu^*, t)$ is bounded away from 0 when $t \to \bar{t}^-$, contradicting $(\mu^*, \bar{t}) \in \text{supp}(f)$. *Q.E.D.*

ONLINE APPENDIX II: AN EFFICIENT ALGORITHM FOR COMPUTING (R)

In this section, we propose an efficient algorithm for computing (R), which solves the shadow prices via $|T \times \Theta|$-dimensional gradient descent. Define function Φ:

$$\Phi(\Lambda, b, \nu, \mu, \tau) := V(\mu, \tau) + \Lambda(\tau)U(\mu, \tau) - \Lambda(0)U(\mu_0, 0) - b_\tau \mu + b_0\mu_0 + \int_{t\in(0,\tau)} (\beta_t \nu_t - U(\nu_t, t))\mathrm{d}\Lambda(t),$$

where $\Lambda \in \mathbb{L}$, $\nu \in \mathbb{D}(T)$ (the set of cadlag paths), $(\mu, \tau) \in D$ and $b_t = b_0 + \int_{s<t} \beta_s \mathrm{d}\Lambda(s)$ for $\beta \in L^1(T)$.

Theorem 3. *If (Λ^*, f, a^*) and a selection of ∇U satisfies (FOC). Let $\beta^*_t = \nabla U(\widehat{\mu}_t, t)$ and $b^*_0 = a^*$. Then,*

$$(\Lambda^*, b^*) \in \arg\min_{\Lambda,\alpha,a} \max_{\nu,\mu,\tau} \Phi(\Lambda, b, \nu, \mu, \tau).$$

Moreover, $\forall(\mu^*, \tau^*) \in \mathrm{supp}(f)$,

$$(\widehat{\mu}, \mu^*, \tau^*) \in \arg\max_{\nu,\mu,\tau} \Phi(\Lambda^*, b^*, \nu, \mu, \tau).$$

Specifically, in the special case where $T \subset \mathbb{N}$ is a discrete and finite set, Theorem 3 yields a simple algorithm for computing the shadow prices, given by

$$(\Lambda^*, b^*) \in \arg\min_{\Lambda,b} \max_{\nu,\mu,\tau} \Big(V(\mu, \tau) + \Lambda(\tau)U(\mu, \tau) - \Lambda(0)U(\mu_0, 0) - b_\tau\mu + b_0\mu_0 + \sum_{t=1}^{\tau-1} ((b_{t+1} - b_t) \cdot \nu_t - (\Lambda(t+1) - \Lambda(t)U(\nu_t, t)) \Big).$$

- *The inner problem*, for each given τ, can be solved by maximizing μ and ν_t period by period. Then, τ can be solved by enumerating finitely many values in T.
- *The outer problem*: Fixing any (ν, μ, τ), is an affine function of (Λ, b). Therefore, the objective is convex in (Λ, b) and it can be solved efficiently via gradient descent.

Proof of Theorem 3. $\forall f \in \Delta_{\mu_0}$, we consider two auxiliary objects.

- ϕ^f is a probability measure in $\Delta(\mathbb{D}(T) \times D)$ defined as follows. The marginal measures $\phi^f|_{\mathbb{D}(T)} = \delta_{\widehat{\mu}}$ and $\phi^f|_D = f$.
- $(\langle\mu^f_t\rangle, \tau^f)$ is the simple recommendation defined by f.

We first prove the claim that $(\Lambda^*, \alpha^*, a^*_0)$ and ϕ^f constitute a saddle point of

$$\min_{\Lambda,b} \max_{\phi\in\Delta(\mathbb{D}(T)\times D)} E_\phi[\Phi(\Lambda, b, \nu, \mu, \tau)].$$

- Fixing (Λ^*, b^*), we show the optimality of ϕ^f. It is equivalent to show that $\forall(\mu, \tau) \in \mathrm{supp}(f)$,

$$(\widehat{\mu}, \mu, \tau) \in \arg\max_{\nu,\mu,\tau} \Phi(\Lambda^*, b^*, \nu, \mu, \tau).$$

Observe that

$$\Phi(\Lambda^*, b*, \nu, \mu, \tau) = \underbrace{V(\mu,\tau) + \Lambda(\tau)U(\mu,\tau) - b_\tau\mu}_{\text{maximized by supp}(f)} - \Lambda(0)U(\mu_0, 0) + b_0\mu_0$$
$$+ \int_{t\in(0,\tau)} \underbrace{(\beta_t\nu_t - U(\nu_t,t))}_{\text{maximized if }\nabla U(\nu_t,t)=\beta_t} \mathrm{d}\Lambda(t).$$

The first maximizer is an implication of Theorem 1. The second maximizer is achieved by $\widehat{\mu}$.

- Fixing ϕ^f, we show that optimality of $\Lambda^*, b*$.

$$\mathbb{E}_{\phi^f}[\Phi(\Lambda, b, \nu, \mu, \tau)] = \mathbb{E}_f[\Phi(\Lambda, b, \widehat{\mu}, \mu, \tau)]$$
$$= \int \left[V(\mu,\tau) + \Lambda(\tau)U(\mu,\tau) - \Lambda(0)U(\mu_0,0) - b_\tau\mu + b_0\mu_0 + \int_{t\in(0,\tau)} (\beta_t\widehat{\mu}_t - U(\widehat{\mu}_t,t))\mathrm{d}\Lambda(t) \right] f(\mathrm{d}\mu, \mathrm{d}\tau)$$
$$= \mathbb{E}_f[V(\mu,\tau)] - \underbrace{\int \left(\int_{t\in(0,\tau)} b_t\mathrm{d}\widehat{\mu}_t + b_\tau(\mu - \widehat{\mu}_t) \right) f(\mathrm{d}\mu, \mathrm{d}\tau)}_{=\mathbb{E}_{\mathcal{P}}[\int_0^{\tau^f} b_t \mathrm{d}\mu_t^f]=0}$$
$$+ \int \left(\int_{t\in(0,\tau)} (U(\mu,0) - U(\widehat{\mu}_t,t))\Lambda(\mathrm{d}t) \right) f(\mathrm{d}\mu, \mathrm{d}\tau) + \Lambda(0)\mathbb{E}_f[U(\mu,\tau) - U(\mu_0,0)]$$
$$= \mathbb{E}_f[V(\mu,\tau)] + \underbrace{\int G(f)(t)\mathrm{d}\Lambda(t)}_{\text{minimized by }\Lambda^*}.$$

The second equality is by integration by parts. The first underbrace is from $\langle \mu_t^f \rangle$ being a martingale. The last equality is from switching the order of integration. The second underbrace is from the complementary slackness condition.

Therefore, we proved that

$$(\Lambda^*, b*) \in \arg\min_{\Lambda,b}\max_{\phi} \mathbb{E}_\phi[\Phi(\Lambda, b, \widehat{\mu}, \mu, \tau)]$$
$$= \arg\min_{\Lambda,b}\max_{\nu,\mu,\tau} \Phi(\Lambda, b, \widehat{\mu}, \mu, \tau).$$

Q.E.D.

Online Appendix III: Construction of SPE^P in Section 3.1

We illustrate Theorem 2 by constructing the full SPE of the corresponding discrete-time extensive form game. There are two specific continuation strategies that are important for our construction:

- The first one is the "inching the goalposts" policy, illustrated by the black curve and arrows in Figure 8. The agent's posterior expectation of task difficulty either jumps to x_l or drifts up gradually until it reaches x_h at t^*.
- The second policy is illustrated by the red curve and arrows in Figure 8. Importantly, the expected task difficulty does not reach x_h at t^*. Afterwards, the principal "inches the goalpost reversely" by revealing $x = x_h$ at a small probability so that the agent is indifferent between continuing and stopping. Absent the revelation, the expected task difficulty drifts down gradually until it reaches x_l at $t = x_h$.

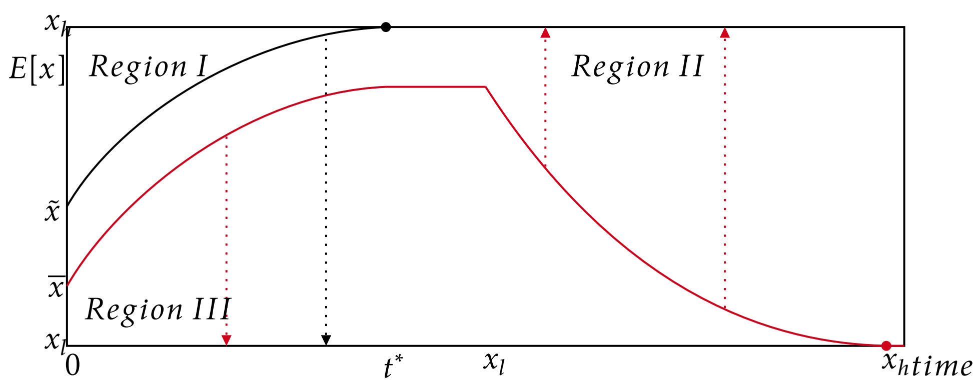


Figure 8: Construction of an SPE.

Figure 8 is depicted with smooth curves purely for illustration—the actual time space is discrete. Then, the SPE^P is constructed as follows:

- For any history that $(\mathbb{E}_{\mu_{t-1}}[x], t)$ is in region I, the principal's strategy induces $\mathbb{E}_{\mu_t}[x]$ either equals x_h or down to the black curve.
- For any history that $(\mathbb{E}_{\mu_{t-1}}[x], t)$ is in region II, the principal's strategy induces $\mathbb{E}_{\mu_t}[x]$ either up to the black curve or down to the red curve.
- For any history that $(\mathbb{E}_{\mu_{t-1}}[x], t)$ is in region III, the principal's strategy induces $\mathbb{E}_{\mu_t}[x]$ either up to the red curve or down to x_l.

In all three scenarios, the agent continues following an interior belief and stops following a boundary belief.

Online Appendix IV: Further discussion of Proposition 1

The sufficient conditions for immediate full-revelation in Proposition 1 were strong but transparent: condition (i) ensured that waiting is costly for the principal, while condition (ii) ensured that the principal is tempted to reveal additional information in the future. There are several natural environments which fulfill these conditions:

Investment advice Consider the classic environment where a patient firm (the agent) engages an advisor to decide which project to irreversibly invest in (see Dixit and Pindyck (1994)). In each state there is an optimal project a_θ which delivers a flow payoff which is constant over time. There is also a risk-free asset a_{rf} (e.g., the agent's outside option) which delivers a constant (but less good) flow payoff. Finally, the payoff from investing in a suboptimal project diminishes over time such that we have, for $\theta' \neq \theta$:

$$u(a_{\theta'}, \theta, \overline{T}) < u(a_{rf}, \theta, \overline{T}) < u(a_\theta, \theta, \overline{T})$$

The advisor (the principal) discounts the future, and obtains state-independent payoffs (e.g., commissions) by inducing the agent to adopt different investments, but least prefers the risk-free option.

Then, a simple chain of reasoning (two rounds of iterated deletion) ensures that, per Proposition 1, in the principal's preferred *SPE* she fully reveals the true state at time 1. First, the agent knows that if she waits until $\overline{T}$, the principal's dislike of the risk-free option would ensure full disclosure. The principal knows that the agent knows this, and so anything short of full disclosure in the interim would induce waiting until $\overline{T}$. But since the principal is impatient, the best she can do is immediate full disclosure.

Discussion of relationship to the Coase conjecture In the "Coase conjecture" (Coase, 1972), a monopolist sells a durable good to consumers of unknown types. This generates an endogenous incentive to reduce prices: on the equilibrium path of prices, after selling to *any* set of consumers, the remaining consumers must be of lower valuations (the so-called 'skimming property'; see e.g. Fudenberg, Levine, and Tirole (1985), Lemma 1). This endogeneously creates an incentive for the monopolist to lower its price to sell to the remaining types—and, anticipating this, consumers would rather wait.

In our informational environment, receiver types are known, and the principal (the monopolist provider of information) has a substantially richer action space. Proposition 1 is stated under the condition that the principal has an exogeneous incentive to deliver full information (price equal to marginal cost) as time goes on. This transparently conveys the possibility that "Coasian" dynamics might also arise in our environment. We expect that a more complete analysis (outside of the scope of this paper) with unknown receiver types will also generate endogeneous incentives for the principal to deliver more information as she learns about receiver types in the interim.

Online Appendix V: Complete comparative statics for Section 4.2

In order to state the complete comparative statics result, we generalize the definition of the *Suspense-ℓ/r* strategy to accommodate several corner cases and one close variant. We first define a variant of the *Suspense-ℓ/r* strategy:

- **Suspense^{t_1}-Inconclusive-ℓ^{t_2}** runs in two stages and is depicted in Figure 9. During the first stage $[0, t_1)$, the principal generates suspense with the two belief paths μ_t^L and μ_t^R.[47] The two paths are chosen such that the agent obtains no continuation surplus over the duration of suspense generation. At the end of the suspense stage $t = t_1$, either the agent is more confident that the state is L ($\mu_{t_1}^L < \mu_0$) and will not be provided further information which induces stopping and choosing ℓ, or $\mu_{t_1}^R = e^{-t_2+t_1}$ which leads to the second stage.[48]

  In the second stage $[t_1, t_2]$, the strategy reveals state L at a rate which is chosen such as to keep the agent's continuation incentive (OC-C) binding. This, in turn, pins down a unique continuing belief path $\mu_{t_1}^*(t) := \mu_{t_1}^R e^{t-t_1}$. By time t_2, The state L will be fully revealed. Thus, at time t_2 the agent is certain the state is R which induces stopping and choosing r.

- **Suspense^{t_1}-Inconclusive-r^{t_2}** is defined analogously. The strategy gradually reveals state R in the second stage $[t_1, t_2]$, which is depicted in Figure 10.

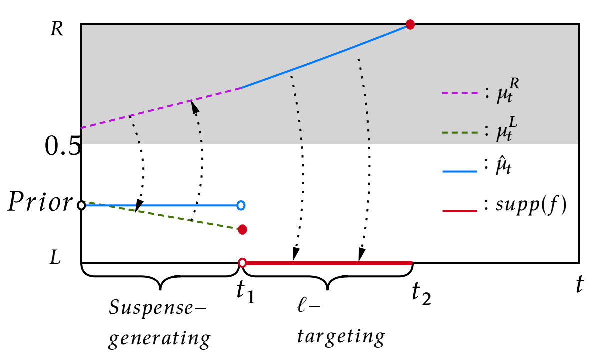


Figure 9: *Suspense-Inconclusive-ℓ* strategy

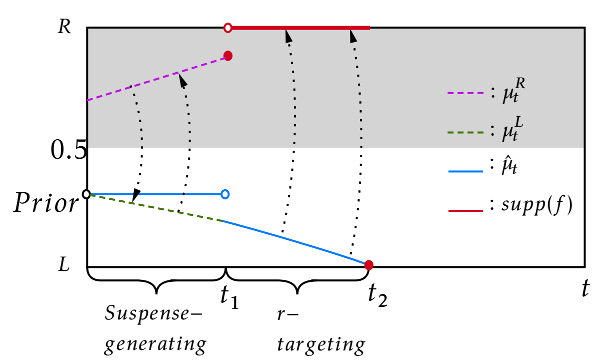


Figure 10: *Suspense-Inconclusive-r* strategy

The difference between Suspense^{t_1}-Inconclusive-ℓ^{t_2} and Suspense^{t_1}-ℓ^{t_2} is that they imperfectly revealing signal at the beginning and the end of the persuasion window, respectively. Second, we extend the definitions of Suspense^{t_1}-ℓ/r^{t_2} to accommodate the following corner cases:

[47] They are given by $\mu_t^L = \mu_{t_1}^L + \frac{t_1-t}{A}(\frac{1}{2} - \mu_{t_1}^L)$ and $\mu_t^R = \mu_{t_1}^R - \frac{t_1-t}{A}(\mu_{t_1}^R - \frac{1}{2})$, where $A = t_1\left(\frac{1/2-\mu_{t_1}^L}{\mu_0-\mu_{t_1}^L}\right)$. $\mu_{t_1}^L$ and $\mu_{t_1}^R$ will be determined shortly.

[48] $\mu_{t_1}^L$ is computed from the agent's indifference condition between stopping at time 0 and time t_1. We obtain $\mu_{t_1}^L = \frac{1}{2}(\frac{2\mu_0-t_1-p_+}{1-p_+})$, where $p_+ = \frac{t_1}{2e^{-t_2+t_1}-1}$. Implicitly, we require $t_1 < \mu_0$ to guarantee there exist t_2 such that $\mu_{t_1}^R \in (0,1)$

- $t_1 = t_2 = 0$: in this case, all strategies reduce to the static persuasion strategy of Kamenica and Gentzkow (2011), where beliefs $\{0, 0.5\}$ are induced at $t = 0$ if $\Delta v > 0$ and beliefs $\{0.5, 1\}$ are induced if $\Delta v < 0$.
- $t_2 = \mu_{t_1}^{*-1}(1)$ or $\mu_{t_1}^{\dagger-1}(1)$: in these cases, the terminal belief of the Suspense^{t_1}-ℓ^{t_2} upon no revelation is either 0 or 1. In these cases, the Suspense^{t_1}-ℓ/r^{t_2} strategy coincides with the Suspense^{t_1}-Inconclusive-ℓ/r^{t_2} strategy.

In short, we say a strategy is a *generalized Suspense-ℓ/r* strategy if it satisfies the definitions above for some t_1, t_2. The following corollary provides sufficient conditions of all generalized *Suspense-ℓ/r* strategies, including what we have in Proposition 2.

Corollary 3.1. *Suppose $t_1, t_2 \in [0, \mu_{t_1}^{*-1}(1)]$ such that $t_1 < t_2$. Let conditions (b)-(d) from Proposition 2 hold. Then, the following statements are true:*

1. *If $t_1 < t_2 < \mu_{t_1}^{*-1}(1)$ and $\Delta v + \Delta h(t_2) = h_r'(t_2)$, then* Suspense^{t_1}-$\ell^{t_2}$ *is optimal.*
2. *If $t_1 < t_2 = \mu_{t_1}^{*-1}(1)$ and $\Delta v + \Delta h(t_2) \in [h_r'(t_2) - 2\Psi(t_1, t_2), h_r'(t_2)]$, then* Suspense^{t_1}-$\ell^{t_2}$ *is optimal.*
3. *If $t_2 < \mu_{t_1}^{*-1}(1)$ and $\Delta v + \Delta h(t_2) = h_r'(t_2) - 2\Psi(t_1, t_2)$, then* Suspense^{t_1}*-Inconclusive-*ℓ^{t_2} *is optimal.*

Moreover, if h_ℓ and h_r are concave and $|\Delta v| \geq \max\{h_r'(0), h_\ell'(0)\}$, then the Suspense^0-$\ell^0$ strategy is optimal.

Proof of Corollary 3.1. Note that part 1 is the same as Proposition 2.

Proof of part 2. We use the same Λ used in the proof of Proposition 2. It remains to verify that $\max_t l_{f,\Lambda}^*(\mu, t)$ is concavified at 0 and 1. If $\nu \geq 0.5$, $\max_t l_{f,\Lambda}^*(\nu, t) = \max_t l_{f,\Lambda}^*(1, t) = l_{f,\Lambda}^*(1, t_2)$. If $\nu < 0.5$,

$$\frac{\partial l_{f,\Lambda}^*(\nu, t)}{\partial t}|_{t=t_1-} = h_\ell'(t_1) - \Lambda(t_1) \geq 0$$

$$\frac{\partial l_{f,\Lambda}^*(\nu, t)}{\partial t}|_{t=t_1+} = h_\ell'(t_1) + (1 - 2\nu)\Lambda'(t_1) - \Lambda(t_1) \leq h_\ell'(t_1) + \Lambda'(t_1) - \Lambda(t_1) = 0.$$

Since $l_{f,\Lambda}^*(\nu, \cdot)$ is concave, we have $\max_t l_{f,\Lambda}^*(\nu, t) = l_{f,\Lambda}^*(\nu, t_1)$ if $\nu < 0.5$. Thus, the statement $\max_t l_{f,\Lambda}^*(\mu, t)$ is concavified at 0 and 1 is equivalent to

$$l_{f,\Lambda}^*(0, t_2) - 2\Lambda(t_1) \leq l_{f,\Lambda}^*(1, t_2) \leq l_{f,\Lambda}^*(0, t_2).$$

The above inequality is equivalent to $\Delta v + \Delta h(t_2) \in [h_r'(t_2) - 2\Psi(t_1, t_2), h_r'(t_2)]$, as given in the condition of Proposition 6.(2)

Proof of part 3. We use the same Λ as before. We showed earlier that $\max_t l_{f,\Lambda}^*(\mu^*(t_1), t) = l_{f,\Lambda}^*(\mu^*(t_1), t_1)$. It remains to verify that $\max_t l_{f,\Lambda}^*(\mu, t)$ is concavified at 0, 1, and $\mu^*(t_1)$. Since

$\mu^*(t_1) < 0.5$, the concavification condition is equivalent to

$$l^*_{f,\Lambda}(\mu^*(t_1), t_1) = \mu^*(t_1) l^*_{f,\Lambda}(0, t_1) + (1 - \mu^*(t_1)) l^*_{f,\Lambda}(1, t_2)$$
$$\iff \Delta v + \Delta h(t_2) = h_r'(t_2) - 2\Psi(t_1, t_2),$$

as given in the condition of part 3.

Proof of optimality of the Suspense^0-ℓ^0 strategy. Under Suspense^0-ℓ^0 strategy with $\Lambda(t) = |\Delta v|$, the principle's FOC from choosing stopping belief ν at time t is

$$l_{f,\Lambda}(\nu, t) = \begin{cases} v_0 + h_\ell(t) - |\Delta v| t, & \nu < 0.5 \\ v_1 + h_r(t) + |\Delta v|(2\nu - 1) - |\Delta v| t, & \nu > 0.5 \end{cases}$$

Since h_l and h_r are concave and $|\Delta v| \geq \max\{h_r'(0), h_l'(0)\}$, $l_{f,\Lambda}(\nu, t) \leq l_{f,\Lambda}(\nu, 0)$ for every $\nu \in [0,1]$. It is easy to check that $l^*_{f,\Lambda}(\cdot, 0)$ is concavified at 0, 0.5, and 1, as desired. *Q.E.D.*

We finally provide a full characterization of optimal information structures when the principal's value is concave and either complementary or substitutable with the agent's action.

Proposition 4. *If h_ℓ and h_r are concave and $\Delta v > 0$, then*

1. *(Complements) If $\Delta h' > 0$, then there exist t_1 and t_2 such that Suspense^{t_1}-ℓ^{t_2} strategy is optimal. Moreover, t_1, t_2 are decreasing in Δv.*

2. *(Substitutes) If $\Delta h' < 0$, then there exist t_1 and t_2 such that either Suspense^{t_1}-r^{t_2} strategy or Suspense^{t_1}-Inconclusive-r^{t_2} strategy is optimal.*

Proof of Proposition 4. We prove the following equivalent (by flipping labels) statement: if h_ℓ and h_r are concave and $\Delta h' > 0$, then

1. (Complements) If $\Delta v > 0$, then there exist t_1 and t_2 such that Suspense^{t_1}-ℓ^{t_2} strategy is optimal.

2. (Substitutes) If $\Delta v < 0$, then there exist t_1 and t_2 such that either Suspense^{t_1}-ℓ^{t_2} strategy or Suspense^{t_1}-Inconclusive-ℓ^{t_2} is optimal.

Since h_ℓ and h_r are concave and $\Delta h' > 0$, conditions (c) and (d) in Proposition 2 always hold. Let $F(t_1, t_2) = \Psi(t_1, t_2) - h_\ell'(t_1)$. Define $t^*_{1,\ell}(t_2) := \min\{t_1 \geq 0 : F(t_1, t_2) \geq 0\}$. This is well-defined because $F(t_2, t_2) = \Delta h'(t_2) > 0$. Consider that

$$e^{-t_1} F(t_1, t_2) = \int_{t_1}^{t_2} e^{-s} h_\ell''(s) \mathrm{d}s + e^{-t_2} \Delta h'(t_2)$$

is strictly increasing and continuous in t_1. This implies 1) $F(t_1, t_2) \geq 0$ if and only if $t_1 \geq t^*_{1,\ell}(t_2)$ and 2) $t^*_{1,\ell}(t_2)$ is continuous.

- $t^*_{1,\ell}(t_2)$ is increasing in t_2. To see this, $\frac{\partial F}{\partial t_2} = e^{t_1-t_2}(-\Delta h'(t_2) + h''_r(t_2)) < 0$. Consider any $t_2 < t'_2$. If $t_1 > t^*_{1,\ell}(t'_2)$, then $F(t_1, t_2) > F(t_1, t'_2) \geq 0$ since F is decreasing in t_2. This implies $t^*_{1,\ell}(t'_2) > t^*_{1,\ell}(t_2)$, as desired.

Define $\hat{t}_2 := \max\{t_2 \geq 0 : t_2 \leq \mu^{*-1}_{t^*_{1,\ell}(t_2)}(1)\}$. This is well-defined because $\mu^{*-1}_t(1)$ is bounded above. Let $\hat{t}_1 := t^*_{1,\ell}(t_2)$. From the definition of $\hat{t}_2$ we must have $\hat{t}_2 = \mu^{*-1}_{\hat{t}_1}(1)$. Note that if $t_2 \leq \hat{t}_2$, then $t_2 \leq \mu^{*-1}_{t^*_{1,\ell}(t_2)}(1)$. This is because 1) $t^*_{1,\ell}(t_2)$ is increasing in t_2 and 2) $\mu^{*-1}_{t_1}(1) = t_1 + \log\left(\frac{2\mu_0 - t_1}{\mu_0}\right)$ is decreasing in t_1.

We consider the following three cases of Δv:

- Case 1: $\Delta v \in [h'_r(\hat{t}_2) - 2\Psi(\hat{t}_1, \hat{t}_2) - \Delta h(\hat{t}_2), h'_r(\hat{t}_2) - \Delta h(\hat{t}_2)]$. Corollary 3.1 part 2 implies Suspense$^{\hat{t}_1}$-$\ell^{\hat{t}_2}$ strategy is optimal. Note that Suspense$^{\hat{t}_1}$-$\ell^{\hat{t}_2}$ are exactly the same as Suspense$^{\hat{t}_1}$-Inconclusive-$\ell^{\hat{t}_2}$ because $\hat{t}_2 = \mu^{*-1}_{\hat{t}_1}(1)$.
- Case 2: $\Delta v > h'_r(\hat{t}_2) - \Delta h(\hat{t}_2)$. We have the following subcases.
  - Case 2.1: $\Delta v \geq h'_r(0)$. the last part of Corollary 3.1 implies Suspense^0-ℓ^0 strategy is optimal.
  - Case 2.2: $\Delta v < h'_r(0)$. This implies there exists $t^*_2 \in (0, \hat{t}_2)$ such that $\Delta v = h'_r(t^*_2) - \Delta h(t^*_2)$. Set $t^*_1 = t^*_{1,\ell}(t^*_2)$. Corollary 3.1 part 1 implies Suspense$^{t^*_1}$-$\ell^{t^*_2}$ strategy is optimal.
- Case 3: $\Delta v < h'_r(\hat{t}_2) - 2\Psi(\hat{t}_1, \hat{t}_2) - \Delta h(\hat{t}_2)$. This implies $\Delta v < 0$.[49] We have the following subcases:
  - Case 3.1: $\Delta v \leq -h'_r(0)$. the last part of Corollary 3.1 implies Suspense^0-Inconclusive-ℓ^0 strategy is optimal.
  - Case 3.2: $\Delta v > -h'_r(0)$. Note that $h'_r(0) - 2\Psi(0,0) - \Delta h(0) = -h'_r(0)$. This implies there exists $t^*_2 \in (0, \hat{t}_2)$ such that $\Delta v = h'_r(t^*_2) - \Psi(t^*_1, t^*_2) - \Delta h(t^*_2)$, where $t^*_1 = t^*_{1,\ell}(t^*_2)$ Corollary 3.1 part 3 implies the Suspense^{t_1}-Inconclusive-$\ell^{t^*_2}$ strategy is optimal.

Q.E.D.

Another way of thinking about Proposition 4 is to fix h_r, h_ℓ such that $\Delta h' > 0$ and vary Δv, the time-0 gain from the agent taking action r over ℓ; this illustrated in Figure 11. For the case in which $\Delta v > 0$, we are in case 1 of Proposition 4 so the suspense-ℓ strategy is optimal.[50] Now consider the case in which $\Delta v < 0$; this corresponds to case 2. of Proposition 4 with the labels flipped so either Suspense^{t_1}-Inconclusive-ℓ^{t_2} or Suspense^{t_1}-ℓ^{t_2} is optimal.[51]

[49] This is because $h'_r(t_2) - 2\Psi(t_1, t_2) \leq (1 - 2e^{t_1-t_2})h'_r(t_2) \leq 0$ since $t_2 < \mu^{*-1}_{t_1}(1) = t_1 - \log \mu^R_{t_1} < t_1 + \log 2$.

[50] Note that time-0 static persuasion to maximize the probability the agent chooses r is a special case.

[51] Even though $\Delta v < 0$, Suspense^{t_1}-ℓ^{t_2} can be optimal when $h'_r(\hat{t}_2) - \Delta h(\hat{t}_2) < 0$

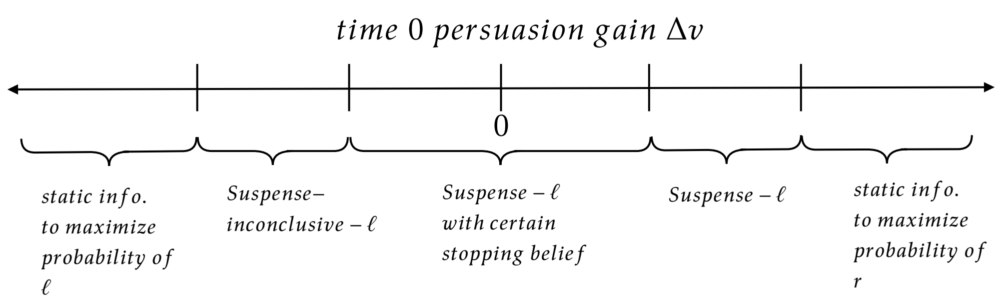


Figure 11: Strategies as Δv varies (fix $\Delta h' > 0$)

Online Appendix VI: Further Implications and Extensions of The Main Model

VI.1 Simple special cases

Preference for revelation. Suppose the principal exhibits "preference for revelation", i.e. $V(\cdot,t)$ is weakly convex for every t. In this case, the optimal persuasion strategy is perfectly revealing. This can be seen from the following modification: splitting a stopping belief μ to degenerate beliefs does not change continuation beliefs $\widehat{\mu}_t$ but weakly improves both the principal's and agent's stopping utilities. Formally, $\forall f \in \Delta_{\mu_0}$ and Borel set $B \subset T$, define $\pi(\theta,B) := \int_{t\in B}\mu_\theta f(\mathrm{d}\mu,\mathrm{d}t)$. Then, $\int\sum U(\delta_\theta,t)\pi(\theta,\mathrm{d}t) = \int\sum U(\delta_\theta,t)\mu_\theta f(\mathrm{d}\mu,\mathrm{d}t) \geq \int U(\mu,t)f(\mathrm{d}\mu,\mathrm{d}t)$. The same inequality holds for V when V is convex in μ. Then, the principal's strategy can be directly modeled by the conditional stopping probability in every state $\pi \in \Delta(\Theta\times T)$ and (R) reduces to:

$$\sup_{\pi}\sum_{\theta}\int V(\delta_\theta,t)\pi(\theta,\mathrm{d}t)$$

$$\text{s.t.}\begin{cases}\displaystyle\int\pi(\theta,\mathrm{d}t) = \mu_0(\theta),\ \forall\theta\in\Theta\\ \displaystyle\sum_{\theta}\left(\int_{\tau>t}U(\delta_\theta,\tau)\pi(\theta,\mathrm{d}\tau)\right) - U\left(\left(\int_{\tau>t}\pi(\theta,\mathrm{d}\tau)\right)_{\theta\in\Theta},t\right)\geq 0,\ \forall t\in T^\circ\end{cases}$$

with the simplified FOC given by:

$$l_{\pi,\Lambda}(\theta,t) := V(\delta_\theta,t) + \Lambda(t)\cdot U(\delta_\theta,t) - \int_{\tau<t}\nabla_\mu U(\widehat{\mu}_\tau,\tau)\mathrm{d}\Lambda(\tau)\cdot\delta_\theta \leq a_\theta,$$

with equality on the support of π. This simplification further reduces the problem's dimensionality from $\Delta(\Theta)\times T$ to $\Theta\times T$.

More broadly, in many of the dynamic persuasion problems studied in the literature, the principal's utility only depends on time but not the action and state, which is a special case of preference for revelation. In such cases, the reduced problem provides an even simpler solution to such problems: the optimal persuasion strategy involves full revelation of the state with stochastic delay, which is a common feature shared by the optimal strategies identified in the literature.[52]

Impatient principal. Suppose $V(\mu,\cdot)$ is weakly decreasing for every μ.[53] In this case, it is straightforward that for all obedient strategy $(\langle\mu_t\rangle,\tau)$, $\mathbb{E}[V(\mu_\tau,\tau)] \leq \mathbb{E}[V(\mu_\tau,0)]$, i.e. it is without loss of optimality to only release information at time-0. In this special case, the optimal solution of (P) reduces to that of the static Bayesian persuasion problem in

[52] Specifically, Ely and Szydlowski (2020); Koh and Sanguanmoo (2024), the single agent case of Knoepfle (2020) and the symmetric prior case of Saeedi et al. (2024) are special cases of the reduced model. In all these papers, the optimal persuasion strategy features *full* revelation with stochastic delay.

[53] $V(\mu,\cdot)$ being decreasing does not necessarily require the principal to be impatient, i.e., $v(\theta,a,\cdot)$ is decreasing. For example, $V(\mu,\cdot)$ is decreasing if (i) $v(\theta,a,\cdot)$ and $u(\theta,a,\cdot)$ are quasi-concave and (ii) the agent's utility always peaks later than the principal's. Therefore, we interpret decreasing $V(\mu,\cdot)$ more broadly as the principal being "more impatient than the agent".

Kamenica and Gentzkow (2011) under indirect utility $V(\mu, 0)$:

$$(\text{P}) = \mathbb{E}_{\mu_0}[V(\delta_\theta, 0)].$$

This reduction, together with Proposition 1, have important implications for the analysis of static persuasion games. Such games are often thought to well-approximate environments in which both the principal and agent dislike delay. Under this view, information transmission and agent's choice both take place at the start of the game, rendering dynamics irrelevant. But, in the richer extensive-form game, this prediction hinges on whether the principal can commit to staying silent.

These two reductions, combined with our analysis in the previous sections, deliver a unified picture how optimal persuasion strategies are shaped by the alignment of the principal's and agents' preferences on the dimensions of information and time (Table 1).

Preference for Info. \ Delay	Aligned ($V(\mu,\cdot)\downarrow$)	Misaligned ($V(\mu,\cdot)\uparrow$)
Aligned	Full revelation at $t=0$ No commitment gap	Full revelation with stochastic delay No commitment gap
Misaligned	Partial revelation at $t=0$ (reduces to KG2011) **Commitment gap**	Persuasion v.s. Delay trade-off No commitment gap

Table 1: Optimal persuasion strategies

VI.2 Evolving state/knowledge.

While our model assumes a persistent state that the principal may fully reveal at the beginning, the cases with an evolving state or evolving principal's knowledge are readily nested by our model. Formally, consider the case where there are finitely many state variables $\{\theta_{t_i}\}$, where the principal can design experiments about each θ_{t_i} at or after period t_i. This directly models the case where the principal learns about the state gradually. An alternative interpretation is that each θ_{t_i} represents an increment of a stochastic process, and observing θ_{t_i}'s means monitoring the path of the stochastic process, representing an evolving state $X_t = \sum_{t_i \le t} \theta_{t_i}$.

Let $\widehat{\theta} = (\theta_{t_i})$ be the full state of the world that contains all relevant information, and Θ be the state space of $\widehat{\theta}$. Let $\nu_t(\widehat{\theta})$ denote the posterior belief of $\widehat{\theta}$ conditional on the realization of the θ_{t_i}'s for $t_i \le t$.

Unobservable state In this setting, $\theta_{t_i}'s$ are unobservable to the principal and the agent. The informational constraint restricts the belief process to fall in the following set $\overline{D}$:

$$\overline{D} := \left\{ (\mu, t) \in D \middle| \mu \in \operatorname{conv}\left(\bigcup_{\theta} \{\nu_t(\theta)\} \right) \right\},$$

where conv(·) denotes the convex closure of a set. In other words, $\overline{D}$ contains all the beliefs that may be induced by only revealing information about states that are observable up to now. Then, the dynamic persuasion problem is equivalent to (R), with the domain D being replaced by $\overline{D}$.[54] Then, Lemma 1 and Theorem 1 apply in a straightforward way. We show in Example 1 that this technique allows us to embed the basic model of Ely (2017) in our setting.

Public information Another variant of interest is the case where all state variables except θ_0 are *public information*, observed by both the principal and the agent (See, e.g., Orlov et al. (2020); Bizzotto et al. (2021)). In this case, the informational constraint restricts the belief process to fall in the following set $\underline{D}$:

$$\underline{D} := \left\{ (\mu, t) \in D \middle| \mu \in \bigcup_{\theta_{t_i}} \operatorname{conv}\Big(\bigcup_{\theta_0} \{\nu_t(\theta)\} \Big) \right\}.$$

$\underline{D}$ differs from $\overline{D}$ in that the convex closure is taken over only $\nu_t(\theta)$ conditional on θ_0, as opposed to all θ_{t_i}. In other words, $\underline{D}$ contains all the beliefs that may be induced by only revealing information about θ_0 while fully revealing all the observable states up to now. Similar to the previous case, the dynamic persuasion problem is equivalent to (R) with the domain $\underline{D}$. We show in Example 2 that this technique allows us to embed the model of Orlov et al. (2020) in our setting.

As has been pointed out in Ely (2017) and Orlov et al. (2020), full intertemporal commitment is necessary for the implementation of the optimal policies in Examples 1 and 2. This is consistent with our analysis as Theorem 2 fails to extend to the case with constrained D: to restore dynamic consistency, the principal needs to release information early to raise the agent's outside option (analogous to the "pipetting" strategy in Orlov et al. (2020)). However, such information might not be available at the time when it is needed due to the gradual arrival of information.

Examples of evolving state and public information

Example 1 (Beeps)**.** For each t_i, $\theta_{t_i} \in \{0, 1\}$. The prior belief is defined as follows. conditional on $\theta_{t_{i-1}} = 0$ (or if $i = 0$), the probability that $\theta_{t_i} = 1$ is $1 - e^{-\gamma(t_i - t_{i-1})}$. Conditional on $\theta_{t_{i-1}} = 1$, $\theta_{t_i} = 1$ with certainty.

[54] It is easy to show that $\overline{D}(t) := \{\mu \in \Delta(\Theta) | (\mu, t) \in D\}$ increases in set inclusion order. This is sufficient for the simple recommendation strategy to be feasible and the proof of Theorem 1 to hold without modification.

In this example, θ_{t_i} describes whether an email has arrived in the mailbox at time t_i (with Poisson rate γ), observed by the principal only when $t \geq t_i$. The principal chooses a revelation process of the *past* states. The agent chooses when to stop working and check the email, with indirect utility $U(\mu,t) = e^{-rt}\text{Prob}_\mu(\sum_{t_i \leq t} \theta_{t_i} > 0)$: the agent gets a unit of utility if he stops after the arrival of email and zero otherwise. The utility is discounted by rate r. The principal's indirect utility is $V(\mu,t) = t$, i.e., the principal only cares about inducing the agent to work for longer. This example resembles a version of the basic model of Ely (2017) but with a forward-looking agent.[55]

Example 2 (Persuading the agent to wait)**.** Let $T = \{0, 1, \ldots, I\}$. $\theta_0 \in \{\theta_L, \theta_H\}$, representing the quality of a project (consumer's willingness to pay). For $t \geq 1$, θ_t are independently distributed with normal distribution $N(0,\sigma)$. Let $X_t = (m + \theta_t)X_{t-1}$ be defined recursively, where m is the trend growth of the process. X_t describes the market size of the project.

The agent decides whether to exercise a real option upon stopping. If the option is exercised, the agent pays a cost of I_A and gets payoff $\theta_0 \cdot X_t$. The agent's indirect utility is

$$U(\mu,t) = \mathbb{E}_\mu[\max_{a\in\{0,1\}} a \cdot e^{-rt}(\theta_0 X_t - I_A)]$$

The principal's indirect utility is

$$V(\mu,t) = \mathbb{E}_\mu[a(\mu,t)e^{-rt}(\theta_0 X_t - I_P)],$$

where $a(\mu,t)$ is the optimal choice of the agent. This example resembles the main model of Orlov et al. (2020).

VI.3 Costly or constrained information generation.

Consider the extension where the principal bears an additive and separable flow cost of experimentation, defined by

$$c_t := \frac{\mathrm{d}\mathbb{E}[H(\mu_t)|\mathcal{F}_t]}{\mathrm{d}t},$$

where H is a convex function on $\Delta(\Theta)$ (e.g., $-H$ is the Shannon's entropy). Then, for a strategy profile $(\langle\mu\rangle_t, \tau)$, the principal's payoff is $\mathbb{E}\big[V(\mu_\tau,\tau) - \int_0^\tau c_t \mathrm{d}t\big]$. As is well known in the literature on rational inattention, the experimentation cost features *uniform posterior separability*. Then, the "chain-rule" implies $\mathbb{E}\big[\int_0^\tau c - t\mathrm{d}t\big] = \mathbb{E}[H(\mu_\tau) - H(\mu_0)]$.[56] Therefore, by redefining the indirect utility: $\widehat{V}(\mu,t) = V(\mu,t) - H_P(\mu)$, (P) fully nests the setting with costly experimentation.

The more complicated extension is when the flow of information is constrained. Consider the setting where the principal's choice of belief process is subject to an extra infor-

[55] While the spirit is the same, the agent is myopic in Ely (2017). Our framework does not nest the general model with repeated action choice in Ely (2017).

[56] This observation has been made in Steiner et al. (2017); Zhong (2022); Hébert and Zhong (2022).

mation capacity constraint (ICC):

$$\frac{\mathrm{d}\mathbb{E}[H(\mu_t)|\mathcal{F}_t]}{\mathrm{d}t} \leq \chi. \tag{ICC}$$

(ICC) is a common "flow information" constraint in information economics (see Zhong (2022), Hébert and Woodford (2023) and Georgiadis-Harris (2024)). (ICC) imposes a nontrivial constraint on implementation because simple recommendation strategys, which are typically not smooth in the flow of information, are no longer feasible. Therefore, the relaxed semi-static problem is not equivalent to the original problem. A partial resolution is provided by Sannikov and Zhong (2024), which characterizes the belief-time distributions that can be implemented by stopping a martingale subject to (ICC). As a direct corollary of Theorem 1 of Sannikov and Zhong (2024), a relaxed problem of (P) subject to (ICC) is

$$\sup_{f\in\Delta_{\mu_0}} \int V(\mu,t) f(\mathrm{d}\mu,\mathrm{d}t) \tag{R}$$

$$\text{s.t.} \int_{y>t} U(\mu,y) f(\mathrm{d}\mu,\mathrm{d}y) \geq U\left(\int_{y>t} \mu f(\mathrm{d}\mu,\mathrm{d}y), t\right),\ \forall t\in T^\circ, \tag{OC-C}$$

$$\int_{y\leq t} H(\mu) f(\mathrm{d}\mu,\mathrm{d}y) + H\left(\int_{y>t} \mu f(\mathrm{d}\mu,\mathrm{d}y)\right) - H(\mu_0) \leq \chi \int \min\{t,y\} f(\mathrm{d}\mu,\mathrm{d}y),\ \forall t\in T. \tag{ICC-C}$$

Note that (ICC-C) is a convex constraint. Hence, solving (R) subject to (OC-C) and (ICC-C) remains a simple linear program. As is observed by Sannikov and Zhong (2024), if the solution f *keeps* (ICC-C) *binding for every* t, then the simple recommendation strategy $(\langle\mu_t^f\rangle, \tau^f)$ achieves the same payoff and is feasible in the original problem (P) subject to (ICC). This technique allows our model to nest settings with costly information generation like Hébert and Zhong (2022).