Information Selling under Prior Disagreement[*]

Ernesto Rivera Mora[†] Nicolas Rodriguez[‡]

December 15, 2025

Abstract

This paper studies monopolistic information selling under prior disagreement and limited intertemporal commitment. A Seller offers experiments to a Buyer across multiple periods. With a common prior, sequential information selling does not increase revenue. Under prior disagreement, revenue maximization requires sequential selling. Moreover, under generic prior beliefs, revenue strictly increases with the length of the interaction. In some environments, offering an initial free experiment maximizes revenue. Paternalistic policies that ban free samples of information can improve the Buyer's welfare. We apply our results to markets for attention, showing that mid-roll advertising extracts more revenue than pre-roll advertising.

Keywords: selling information; limited commitment; prior disagreement; free samples; sequential information design

---

[*] We are grateful to Andreas Blume, Tilman Börgers, Benjamin Brooks, Odilon Camara, Roberto Corrao, Inga Deimen, Amanda Friedenberg, Tan Gan, Marina Halac, Johannes Hörner, Doron Ravid, Stanley Reynolds, Ina Taneva, and Kun Zhang for valuable feedback and helpful comments. We thank the audiences at Stony Brook Game Theory Conference (2024), Durham Game Theory Conference (2024), and SAET Conference (2025) for insightful discussions.

[†] University of Colorado, Boulder, ernesto.riveramora@colorado.edu

[‡] Mississippi State University, nr749@msstate.edu

1 Introduction

In markets for professional advice, experts and clients often disagree about the value of information. Clients tend to undervalue recommendations from lawyers assessing legal risks, physicians diagnosing medical conditions, or financial advisors evaluating investment opportunities. To address this, experts often offer "complimentary consultations"—free initial sessions designed to increase clients' willingness to pay for further advice. This strategy may seem counterintuitive, as experts provide valuable information at no charge. This paper sheds light on the rationale behind this practice by showing that complementary consultations are often optimal and may strictly outperform selling schemes that bundle all sessions at a single price.

We study a sequential information-trading model with limited intertemporal commitment and prior disagreement. A (monopolistic) Seller offers experiments that reveal information about a payoff-relevant state to a Buyer over a fixed number of periods. However, in contrast with classic environments, the players *agree to disagree* about their beliefs about the state. Prior disagreement may arise from *overconfidence* (Grubb, 2009), differences in *opinions* (Che and Kartik, 2009), or simply distinct *views of the world* (Alonso and Câmara, 2016).[1]

Our main result, Theorem 1, characterizes the equilibrium payoffs for any Buyer's decision problem, prior beliefs, and number of trading periods. If the players share a common prior, then they also share the same value of information. In this case, sequential selling does not increase the Seller's revenue and bundling all information at a single price is optimal. In contrast, with prior disagreement, bundling all information in a single period is strictly suboptimal as sequential selling allows the Seller to extract higher surplus by exploiting belief differences. In other words, in environments with prior disagreement, it is optimal for the Seller to sell information slowly. The key insight is that different signal outcomes affect the Buyer's willingness to pay for future information. So, by tailoring experiments strategically, the Seller can ensure that outcomes she deems relatively more likely will increase the Buyer's

---

[1] While common priors are central to economic theory (Harsanyi, 1968; Aumann, 1976; Halpern, 2002), models without common priors remain rational under personalistic Bayesian perspectives. (Savage, 1972; Morris, 1995).

future demand for information.

Our main result also shows that, in multiple environments, the Seller's incentives to delay information repeat over time. More specifically, for a generic set of prior beliefs—in which the players reach belief agreement only once all uncertainty is resolved—the Seller's expected revenue strictly increases with the number of trading periods. In these environments, differences in beliefs can be exploited throughout the entire interaction, and each equilibrium selling strategy only reveals partial information in non-terminal periods.

This raises a question: Can the Seller extract unbounded (subjective) revenue by extending the interaction over an arbitrarily large number of periods? While one might expect that the Seller could design experiments to repeatedly shift the Buyer's beliefs and continually extract revenue, Theorem 2 shows that such infinite extraction is impossible. The Seller's expected revenue remains bounded, regardless of the number of periods in which they interact. Hence, as the number of periods grows, each additional period contributes only minimally to total revenue. We show that the revenue generated by sequential information selling is bounded by the revenue achievable in environments in which the Seller has a particular type of commitment: contracts with state-contingent payments subject to the Buyer's limited liability. Thus, Theorems 1 and 2 imply that—for a generic set of priors—commitment delivers a positive value for the Seller. Nevertheless, a positive value of commitment is not universal. Example 2 shows that—for a specific prior beliefs—commitment yields no additional value for the Seller.

Our results also show that, from the Seller's perspective, the Buyer ultimately pays more than he anticipates. Even when the Buyer places the highest possible value to learning the state, the Seller still extracts an expected payment that exceeds the Buyer's willingness to pay. (See Remark 1 in Section 2.) With this in mind, we examine policies that protect the Buyer's welfare. Consider a Regulator who understands that players have belief disagreement and whose prior is a convex combination of the players' priors. We assume the Regulator prioritizes consumer surplus and can choose the length of the interaction. Absent intervention, the Regulator anticipates that the Buyer may pay more than his maximum willingness to pay for information.

Proposition 1 shows that the Regulator's objective decreases with the number of trading periods, as fewer periods reduce the Buyer's expected payment without changing the amount of information revealed. As a result, shortening the interaction increases Buyer surplus. Policies that achieve this include banning sequential information sales, prohibiting the sale of partially informative experiments.

Remarkably, optimal regulation may require banning free information samples as a way to protect the Buyer. As we show in Section 2, in some environments, it is optimal for the Seller to provide free initial experiments. While such complementary information may appear beneficial to the Buyer, it allows the Seller to exploit prior disagreement by extracting higher expected payments in later periods. Hence, from the Regulator's perspective, introducing a paternalistic policy that bans free samples enhances the Buyer's welfare.

Section 8 applies our results to attention markets, in which the numeraire is not money but advertising time. Consider an expert who produces informative content—videos, articles, or podcasts—that offers advice to viewers facing a decision problem.[2] The expert monetizes content by bundling information with advertisements, so viewers "pay" by allocating attention to ads. Two common advertising strategies are used in attention markets. First, pre-roll advertising, which places all ads before content. This strategy corresponds to a static scheme that bundles all information at a single price. Second, mid-roll advertising, which places ads throughout the content. This strategy corresponds to a sequential scheme that enables sequential revelation. When the expert and viewer have belief disagreement, our results apply directly; mid-roll advertising allows the expert to extract higher revenue through sequential information disclosure. Moreover, we show that a regulator seeking to shift surplus to consumers can do so by mandating pre-roll advertising.

Literature Review This paper contributes to the literature on principals with limited commitment, which has largely focused on common-prior environments with

---

[2] We focus on informational content and abstract from artistic content. Moreover, we assume the expert has monopolistic control over the information, observes the state at no cost, the viewer values the content solely for informative purposes, and that dividing content into segments has not effect on its quality. See Kerkhof (2024) for empirical evidence on how the type of advertising impacts the endogenous creation of content in other settings.

private information (see Acharya and Ortner (2017); Bester and Strausz (2001); Krishna and Morgan (2008); Doval and Skreta (2022)).

This paper also contributes to the literature on information markets (Bergemann and Bonatti, 2019; Ichihashi, 2021; Ali, Haghpanah, Lin, and Siegel, 2022; Zhong, 2022; Bergemann, Bonatti, and Gan, 2022; Foerster and Närmann, 2025). Our paper contributes to this literature by developing a dynamic programming approach for settings with prior disagreement. By applying tools from information design (Kamenica and Gentzkow, 2011; Rayo and Segal, 2010) we reduce the sequential problem to a static Bayesian persuasion problem with prior disagreement (Alonso and Câmara, 2016).

However, our setting differs from static Bayesian persuasion with prior disagreement in two key ways. First, although persuasion can be valuable in one-shot environments with prior disagreement, it is not valuable in our model; it is the sequential structure that enables the Seller to exploit belief differences. Second, Application 2 in Alonso and Câmara (2016) illustrates that the value of persuasion depends on the cardinality of the state space: persuasion may only be beneficial when there are at least three states. By contrast, in our model, the Seller gains from sequential selling regardless of the number of states.

Three papers are closely related to our work. The first is Bergemann, Bonatti, and Smolin (2018), who analyze an environment in which a Seller offers information to a Buyer who observes a private signal correlated with the state. Since the Seller does not observe the private signal, the Buyer's belief about the state is private information. While the Buyer and the Seller have heterogeneous priors, their beliefs are not common knowledge. More importantly, conditional on the Buyer's private signal, both players agree about the likelihood of each path. As a result, the Seller cannot leverage disagreement about outcomes as in our setting.[3]

A second closely related paper is Hörner and Skrzypacz (2016), which shows that sequential information selling can arise in some equilibria in certain environments.

---

[3] Bergemann, Bonatti, and Smolin (2018) show that sequential mechanisms may improve the Seller's revenue by helping to screen the Buyer's types: the experiment in the first period provides information that helps identify which private signal the Buyer observed. They provide an example with three types in which a 2-step mechanism has more screening power than a 1-step mechanism.

However, there are important differences. While their analysis focuses on settings with binary states, Seller's private information, and a common prior, our model has arbitrary number of states, no private information, and prior disagreement. The results also differ. In their model, (i) the Seller is ex-ante indifferent between selling information gradually versus selling all information at once, and (ii) there are equilibria in which all information is sold in the first period. In our model, all equilibria under prior disagreement involve sequential information selling, and the Seller strictly increases her (subjective) ex-ante revenue by selling information sequentially.

Another related paper is Zheng and Chen (2021), which examines optimal advertising of information in two-period models in which the first experiment is required to be free. Our approach differs by allowing for any number of periods and arbitrary non-negative prices in each period. We show that, in a range of environments, free sampling of information emerges as the unique equilibrium strategy for the Seller. Thus, rather than imposing free sampling as a restriction, our model shows it may arise endogenously.

2 Example

A Manager (the information Buyer) faces a decision problem under uncertainty. There are two possible states: The firm's data is vulnerable ($\underline{\theta}$) or safe ($\overline{\theta}$). The Manager chooses whether to update the firewall ($\underline{a}$) or not ($\overline{a}$). Table 1 summarizes the Manager's payoff function $u : A \times \Theta \to \mathbb{R}$.

	$\underline{\theta}$	$\overline{\theta}$
$\underline{a}$	1	0
$\overline{a}$	0	1

Table 1 Manager payoffs.

So, updating the firewall is the right action when the data is vulnerable, and not updating is the right action when the data is safe. Let $\nu_B = \frac{9}{10}$ be the Manager's prior belief that the state is $\overline{\theta}$. Because learning the state perfectly guarantees a payoff of one, the Manager's willingness to pay for complete information given a belief $\mu_B \in (0,1)$ is

$$V(\mu_B) = 1 - \max_{a} \mathop{\mathbb{E}}_{\theta \sim \mu_B} [u(a,\theta)] = \min\{\mu_B, 1-\mu_B\}.$$

In particular, at the Buyer's prior $\nu_B = \frac{9}{10}$, his willingness to pay is $V\left(\frac{9}{10}\right) = \frac{1}{10}$. (See Figure 2.1.)

Before making a decision, the Manager may purchase information from a revenue-maximizing monopolistic Seller. Specifically, the Seller offers an experiment (i.e. firewall test) $\pi : \{\underline{\theta}, \overline{\theta}\} \to \Delta\{\underline{m}, \overline{m}\}$ for a price p. The Manager is free to accept or reject the offer. If the offer (π, p) is accepted, the Manager pays p and both players observe the outcome of the experiment. Otherwise, no payment is made and the experiment is not conducted.

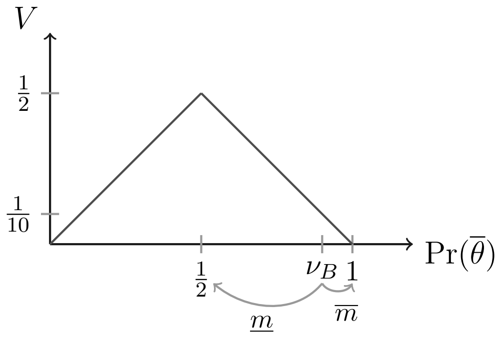


Figure 2.1 Manager's willingness to pay for observing the state.

Importantly, the Seller does not know the state and holds her own prior belief $\nu_S = \frac{1}{2}$ that the state is $\overline{\theta}$. We assume prior beliefs are common knowledge, both players are Bayesian, and they agree on the probabilities each experiment π specifies. In other words, although the players disagree about the prior probability of each state, they agree on how each experiment π maps states into signal distributions.

If the Seller is restricted to offering a single experiment, then her maximum expected revenue is $V\left(\frac{9}{10}\right) = \frac{1}{10}$. This is because the Seller's expected revenue equals the Manager's payment, and the Manager is unwilling to pay more than $V\left(\frac{9}{10}\right)$ for any experiment. If, however, the Seller is allowed to offer experiments sequentially, she can obtain a strictly higher expected subjective revenue.

To see this, consider the following sequential selling strategy. In the first period the Seller offers the following experiment for free. If the state is $\underline{\theta}$, the experiment yields outcome $\underline{m}$ with certainty. If the state is $\overline{\theta}$, it yields outcome $\underline{m}$ with probability $\frac{1}{9}$ and $\overline{m}$ with probability $\frac{8}{9}$. In the second period, the Seller offers a fully-revealing experiment at price $\frac{1}{2}$.

The Manager will accept the initial free experiment, as it provides information at no charge. As Figure 2.1 illustrates, if the outcome is $\overline{m}$, he becomes certain that the state is $\overline{\theta}$, and therefore rejects the second offer. If the outcome is $\underline{m}$, he updates his belief that the state is $\overline{\theta}$ to $\frac{1}{2}$. Hence, after such a signal, the Manager is willing to pay $V\left(\frac{1}{2}\right) = \frac{1}{2}$ for full information. Therefore, he accepts the second offer. The Seller thus receives a payment of $\frac{1}{2}$ if $\underline{m}$ occurs in the first experiment, and 0 otherwise. The key is that, from the Seller's perspective, the probability of $\underline{m}$ is $\frac{5}{9}$, so her expected revenue is $\frac{5}{18}$. This sequential selling strategy allows the Seller to secure almost 3 times her 1-period revenue, $\frac{1}{10}$.

This example illustrates that, under prior disagreement, sequential information selling strictly increases the Seller's expected revenue. Intuitively, given an experiment, prior disagreement leads the players to assign different probabilities to experimental outcomes. Consequently, they disagree on the likelihood of the Manager's future values of information induced by the initial experiment. Because the Seller places greater prior probability on $\underline{\theta}$, she believes that $\underline{m}$ is more likely than the Manager does. Therefore, from her perspective, the Manager underestimates the probability of having a high future value of information. The Seller thus increases revenue by exploiting this discrepancy.

Section 6 shows that, when the Seller and Manager interact over two periods, this sequential selling strategy is optimal. In particular, providing free initial information strictly dominates any strategy that does not involve free initial experiments. Moreover, our main results will show that revenue increases strictly with the length of the interaction. That is, selling information over $T+1$ periods yields strictly higher expected revenue than selling over only T periods.

Remark 1. In this example, the Seller believes that the initial value of information is high (her prior is $\frac{1}{2}$), while the Buyer believes that the value of information is low (his prior is $\frac{9}{10}$). So, the Seller aims to persuade the Buyer that assigning a low value to information is a mistake. Remarkably, our results show that sequential information revelation also arises in the reverse case. Suppose instead that the Buyer's prior is $\nu_B = \frac{1}{2}$, which assigns the highest possible ex ante value for information, while the Seller holds prior $\nu_S = \frac{9}{10}$ with a low value for information. Theorem 1 shows that—even though the Buyer already has the highest possible value for information—the Seller can still profit by disclosing information sequentially.

This highlights a fundamental difference from Alonso and Câmara (2016). In their model, a pessimistic Sender cannot benefit from disclosing information to an optimistic Receiver when there are only two states. In our model, by contrast, the Seller benefits from sequential revelation of information regardless of the number of states and prior beliefs, provided that the players agree to disagree.

3 Model

Throughout the paper, take the following conventions. Endow a compact metric space C with its Borel sigma-algebra. Denote by ΔC the set of probability measures on C and endow ΔC with the topology of weak convergence. Denote the relative interior of ΔC by $\operatorname{int} \Delta C$. For each $c \in C$ write $\delta_c \in \Delta C$ for the probability measure that assigns probability one to the singleton $\{c\}$.

Environment. There are two players, a Seller (she) and a Buyer (he). A payoff-relevant state θ is drawn from a finite set Θ, and is unknown to both players. Each player holds a belief $\nu_S, \nu_B \in \operatorname{int} \Delta\Theta$. While these beliefs may differ, they are common knowledge among the players. In this sense, the players may *agree to disagree.*

Experiments. Information about the state will be revealed through experiments. An experiment is a stochastic mapping $\pi : \Theta \to \Delta M$, in which M is a finite set of signals that satisfies $|M| \geq |\Theta|$. Importantly, while the players may disagree about their beliefs about the state, we assume that the players agree on the conditional probabilities specified by each experiment.

Interaction. The Buyer faces an exogenous decision problem, described by a compact set of actions A and a continuous utility function $u : A \times \Theta \to \mathbb{R}$ that has no weakly dominant action. The game proceeds over T periods of information selling, followed by the Buyer's decision problem (A, u). The number of periods T is exogenous. In each information selling period, the Seller offers an experiment π at a posted non-negative price p to the Buyer. Upon observing an offer (π, p), the Buyer either accepts or rejects. If he accepts, he pays p to the Seller and both players observe the realized signal. Otherwise, no payment is made and no player observes the signal. At the final stage, the Buyer selects an action $a \in A$.

Payoffs. Payoffs are as follows. If the Buyer selects action $a \in A$, the state is $\theta \in \Theta$, and the Buyer has accepted offers for a total sum of payments P, then the total payoff of the Buyer is $u(a, \theta) - P$. The Seller's payoff is the total sum of payments P.

Equilibrium. Because the game has complete and perfect information, we use subgame perfect equilibrium (SPE) as the solution concept. This reflects an environment in which the Seller has limited commitment: The Seller commits to honor current offers (charging the posted price and disclosing signals truthfully), but cannot commit to future transfers or offers contingent on the realized state or past outcomes.

4 Main Results

The main results of the paper characterize equilibrium payoffs in the game. Lemma A.5 in the Appendix establishes both the existence of SPE and the uniqueness of equilibrium payoffs for each T-period game. The argument proceeds by backward induction. In the final period, equilibrium existence and uniqueness of payoffs follow immediately from the fact that the final offer is an ultimatum game. For any earlier period, suppose equilibrium exists in all continuation games and that continuation payoffs are unique. The Seller's offer then reduces again to an ultimatum game: the Buyer accepts any offer yielding a payoff at least equal to his continuation value, and rejects otherwise. This pins down equilibrium behavior and payoffs uniquely.

Before presenting our main results, we introduce some notation. Write

$$U(\mu_B) := \max_{a \in A} \mathop{\mathbb{E}}_{\theta \sim \mu_B} [u(a, \theta)],$$

for the Buyer's payoff on the decision problem given belief μ_B, and

$$V^1(\mu_B) := \mathop{\mathbb{E}}_{\theta \sim \mu_B} [U(\delta_\theta)] - U(\mu_B),$$

for the Buyer's willingness to pay for fully learning the state given belief μ_B. Write $r(\theta) := \frac{\nu_S(\theta)}{\nu_B(\theta)}$ for the likelihood ratio of state θ under the players' prior beliefs.

Our main result describes equilibrium payoffs across environments that differ in the degree of prior disagreement. We distinguish three cases: *prior agreement (PA)* when $\nu_S = \nu_B$, *prior disagreement (PD)* when $\nu_S \neq \nu_B$, and *strong prior disagreement (SPD)* when $r(\theta) \neq r(\theta')$ for each pair of distinct states $\theta, \theta' \in \Theta$.[4] SPD and PD differ in the extent to which interim belief agreement can be achieved after some information is revealed. Under SPD, belief agreement occurs only when the state is fully revealed. Under PD, players may reach interim agreement even when uncertainty remains.[5]

Let $\mathcal{U}_T$ and $\mathcal{R}_T$ denote the Buyer's equilibrium expected utility and the Seller's equilibrium expected revenue, respectively, in the T-period game.

Theorem 1. *In each environment, $\mathcal{U}_T = U(\nu_B)$ for each $T \in \mathbb{N}$, $\mathcal{R}_T$ is weakly increasing in T, and $\mathcal{R}_1 = V^1(\nu_B) > 0$. Moreover,*

- *(i) Under PA, $\mathcal{R}_T = \mathcal{R}_1$ for each $T \in \mathbb{N}$.*
- *(ii) Under PD, $\mathcal{R}_2 > \mathcal{R}_1$.*
- *(iii) Under SPD, $\mathcal{R}_T$ strictly increases with T.*

Theorem 1 yields several insights. First, in all environments, the Buyer's equilibrium payoff remains constant in T and equals his payoff under no information. Second, the Seller's one-period revenue always matches the Buyer's willingness to pay for full information. Third, under PA, sequential information selling does not increase

---

[4] Proposition 1 in Alonso and Câmara (2016) implies that, given a Buyer's interim belief $\mu_B \in \Delta\Theta$, the players have interim belief agreement if and only if $r(\theta) = r(\theta')$ for all $\theta, \theta' \in \operatorname{Supp} \mu_B$.

[5] Specifically, agreement is achieved whenever the support of posterior beliefs includes only states with identical likelihood ratios.

revenue; offering a single fully-revealing experiment maximizes revenue. Fourth, under PD, sequential information selling strictly increases the Seller's revenue. Finally, under SPD, the Seller strictly benefits from longer interactions, as her equilibrium revenue $\mathcal{R}_T$ strictly increases with T.

The next example illustrates that, under PD (but not SPD), extending the interaction beyond two periods need not increase the Seller's revenue.

Example 2. Suppose the Buyer's decision problem is described by the following payoff table:

	θ_1	θ_2	θ_3
a_1	1	-1	0
a_2	-1	1	0

Table 2 Decision problem payoffs.

The players' prior beliefs are $\nu_S = (\frac{2}{5}, \frac{2}{5}, \frac{1}{5})$ and $\nu_B = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$. Since $\nu_S \neq \nu_B$ and $r(\theta_1) = r(\theta_2)$, this environment corresponds to PD but not to SPD. Notice, after any experiment, the players agree on which state in $\{\theta_1, \theta_2\}$ is more likely. They therefore always agree on which action is optimal, yet disagree about whether taking the optimal action is relevant (the state is in $\{\theta_1, \theta_2\}$) or not (the state is θ_3). Lemma A.9 in the Appendix shows that, in this example, $\mathcal{R}_T = \mathcal{R}_2 > \mathcal{R}_1$ for each $T \geq 2$. So, while the Seller increases her revenue from revealing information in two periods, she gains no additional revenue from extending the interaction beyond two periods. Intuitively, the Seller cannot improve upon showing the Buyer that his value for information is incorrect by first providing a free sample that reveals whether the action is relevant (the state is in $\{\theta_1, \theta_2\}$) or not (the state is θ_3).

4.1 The Value of Commitment

This section extends the analysis by considering two distinct levels of commitment for the Seller, and analyzes how commitment influences equilibrium payoffs.

Suppose the Seller can commit to a contract $P \in \mathbb{R}^\Theta$ that specifies the Buyer's payment in each state in exchange for fully revealing the state. A contract P is individually rational (for the Buyer) if his value from learning the state, $V^1(\nu_B)$, exceeds his expected payment, $\mathbb{E}_{\theta\sim\nu_B}[P(\theta)]$. The Seller's problem is to choose an individually-rational contract P that maximizes her expected revenue, $\mathbb{E}_{\theta\sim\nu_S}[P(\theta)]$.

Full Commitment. Suppose the Seller can choose any contract in $\mathbb{R}^\Theta$, permitting both arbitrarily large positive and negative payments. Under PA, the Seller's expected revenue from any contract equals the Buyer's expected payment. Hence, the maximum revenue the Seller can obtain is $V^1(\nu_B)$. So, granting the Seller full commitment does not increase her revenue.

In contrast, under PD (and SPD), the Seller's revenue is unbounded. To see this, observe that when $\nu_S \neq \nu_B$, there exists a contract $P \in \mathbb{R}^\Theta$ such that

$$\underset{\theta\sim\nu_S}{\mathbb{E}}[P(\theta)] > 0 \quad \text{and} \quad \underset{\theta\sim\nu_B}{\mathbb{E}}[P(\theta)] \leq 0.$$

Such a contract can be constructed by assigning positive payments to states that the Seller considers relatively more likely and negative payments to states that the Buyer deems relatively more likely, ensuring the expectations differ in sign. This contract P is individually rational for the Buyer. Moreover, for any $\lambda > 1$, the scaled contract λP remains individually rational. Note, the Seller's expected revenue becomes arbitrarily large as λ increases. Therefore, the Seller's revenue is unbounded.

Commitment under Limited Liability. Suppose the Seller can only choose contracts in $\mathbb{R}_+^\Theta$. That is, the Seller has limited liability and can commit only to contracts with non-negative payments by the Buyer.

The following theorem establishes properties of optimal contracts under limited liability.

Theorem 2. *Assume the Seller has commitment under limited liability. In each environment, an optimal contract P^* exists. Moreover, for each $T \in \mathbb{N}$,*

$$\mathcal{R}_T \leq \underset{\theta\sim\nu_S}{\mathbb{E}}[P^*(\theta)] = V^1(\nu_B) \max_{\theta\in\Theta} r(\theta).$$

Theorem 2 yields three insights. First, it ensures the existence of an optimal contract under limited liability, and characterizes the Seller's revenue in this setting. Second, it provides an upper bound on equilibrium revenue with limited commitment. As a result, $\lim_{T\to\infty} \mathcal{R}_T$ exists and the marginal benefit of extending the number of periods vanishes. In particular, the Seller cannot manipulate the Buyer's beliefs in a way that extracts arbitrarily large expected revenue. Third, it clarifies the value of commitment under limited liability. Under PA, all likelihood ratios equal one, so commitment under limited liability has no value to the Seller. Under PD, commitment with limited liability yields strictly higher revenue than in the one-period case. Moreover, under SPD, the Seller always benefits from commitment under limited liability compared to limited commitment.

Section 6 shows that, in the example of Section 2, $\lim_{T\to\infty} \mathcal{R}_T < V^1(\nu_B) \max_{\theta\in\Theta} r(\theta)$, showing a strict gap in the limit. By contrast, in Example 2, $\mathcal{R}_T = V^1(\nu_B) \max_{\theta\in\Theta} r(\theta)$ for all $T \geq 2$, implying that commitment with limited liability provides no gain when there are two or more periods.

4.2 Policy

The results of the paper have policy implications. Consider an environment with prior disagreement in which a Regulator prioritizes the Buyer's surplus over the Seller's surplus. The Regulator holds an intermediate prior belief, $\nu_R = \lambda\nu_S + (1-\lambda)\nu_B$, for some $\lambda \in (0,1)$. The following result characterizes the interaction length that maximizes the Regulator's objective.

Proposition 1. *Under PD, the Regulator's objective is weakly decreasing in T. Moreover, under SPD, the objective is strictly decreasing in T.*

Proposition 1 implies that policies that reduce the number of trading periods improve the Buyer's welfare. Intuitively, from the Regulator's viewpoint, longer interactions give the Seller more opportunities to extract revenue through gradual selling, to the Buyer's detriment. Shorter interactions therefore limit the extent to which belief disagreement can be exploited.

The intuition behind the proof is straightforward. Because the Regulator places greater weight on the Buyer's surplus, his objective amounts to minimizing the expected payment from his own perspective. To characterize this expected payment, observe that players disagree only about the prior over states; so conditional on each state, they agree on the expected equilibrium payment. Moreover, since the Regulator's prior is a convex combination of the Buyer's and Seller's priors, and expectations are linear in beliefs, the expected payment from the Regulator's perspective is a convex combination of the Seller's expected revenue $\mathcal{R}_T$ and the Buyer's expected payment $\mathcal{R}_1$. Theorem 1 then implies that the Regulator's objective is constant in T under PA, weakly decreasing in T under PD, and strictly decreasing in T under SPD. In all cases, the Regulator's objective is maximized by choosing $T = 1$.

5 Dynamic Programming

To establish the main results of the paper, we develop a dynamic programming approach that characterizes (i) how the players' beliefs evolve after implementing any sequence of experiments, and (ii) how equilibrium payoffs and behavior recursively depend on such beliefs at each subgame.

Belief Dynamics We begin by characterizing how beliefs evolve after any sequence of experiments. Given priors ν_B and ν_S, with associated likelihood ratio $r(\theta)$ for each $\theta \in \Theta$, define

$$\Omega := \left\{ (\mu_S, \mu_B) \in \Delta\Theta \times \Delta\Theta : \mu_S(\theta) = \frac{r(\theta)}{\underset{\theta' \sim \mu_B}{\mathbb{E}}[r(\theta')]} \mu_B(\theta) \text{ for all } \theta \in \Theta \right\}.$$

Proposition 1 in Alonso and Câmara (2016) implies that a profile of posterior beliefs (μ_S, μ_B) can be generated by implementing a sequence of experiments if and only if $(\mu_S, \mu_B) \in \Omega$.[6] Thus, Ω is the set of belief profiles consistent with the priors (ν_S, ν_B). Moreover, the belief profile (μ_S, μ_B) is fully determined by the Buyer's belief μ_B alone: Given a Buyer's belief $\mu_B \in \Delta\Theta$, the Seller's belief μ_S is uniquely pinned down by the identity defining Ω. This observation allows us to track the evolution of the entire

---

[6] Note that any sequence of experiments is itself an experiment.

belief profile by following only the Buyer's belief.[7]

Fix a subgame in which the belief profile is $(\mu_S, \mu_B) \in \Omega$. Note that implementing an experiment π induces a distribution $\hat{\tau}_i \in \Delta\Omega$ over posterior profiles from the perspective of player i.[8] Since the belief profile is determined by the Buyer's belief, write $\tau_i \in \Delta(\Delta\Theta)$ for the marginal of $\hat{\tau}_i$ over the Buyer's belief. So, $\tau_i(\mu_B)$ is player i's belief that the Buyer's posterior after experiment π will be μ_B. Write $\mathrm{PS}[\mu] := \big\{\tau \in \Delta(\Delta\Theta) : \mathbb{E}_\tau[\mu'] = \mu\big\}$ for the set of (marginal) posterior spreads satisfying Bayes plausibility at belief μ.

Lemma 1. *Fix a belief profile $(\mu_S, \mu_B) \in \Omega$. There exists an experiment π inducing distributions $(\hat{\tau}_S, \hat{\tau}_B)$ over posterior profiles if and only if*

(i) $\operatorname{Supp} \hat{\tau}_S = \operatorname{Supp} \hat{\tau}_B \subseteq \Omega$,

(ii) $\tau_B \in \mathrm{PS}[\mu_B]$, *and*

(iii) For each $(\mu'_S, \mu'_B) \in \Omega$, it holds that

$$\hat{\tau}_S(\mu'_S, \mu'_B) = \rho(\mu'_B \mid \mu_B)\, \tau_B(\mu'_B) \qquad \textit{where} \qquad \rho(\mu'_B \mid \mu_B) := \frac{\mathbb{E}_{\theta\sim\mu'_B}[r(\theta)]}{\mathbb{E}_{\theta\sim\mu_B}[r(\theta)]}.$$

Lemma 1 characterizes how players may disagree about the likelihood of different posterior beliefs that they may hold after any experiment, given any interim belief profile $(\mu_S, \mu_B) \in \Omega$. To build intuition, suppose that players currently hold belief profile (μ_S, μ_B). An experiment π induces a distribution of posterior beliefs τ_B, which specifies the likelihood the Buyer assigns to each posterior belief μ'_B after observing the experiment's signal. With this in mind, define

$$H_{\rho>1}(\mu_B) := \{\mu'_B \in \Delta\Theta \mid \rho(\mu'_B \mid \mu_B) > 1\}$$

as the set of Buyer's posterior beliefs that the Seller deems relatively more likely. The key insight is that, for each interim belief $\mu_B \in \Delta\Theta$, the function $\rho(\mu'_B \mid \mu_B)$ is linear in μ'_B. Consequently, the set $H_{\rho>1}(\mu_B)$ forms an open half-space that contains μ_S, provided $\mu_S \neq \mu_B$.[9] Figure 5.1 depicts this open half-space as the shaded region

---

[7] The analysis is analogous if we track the Seller's belief instead.

[8] As the agents' beliefs disagree, $\hat{\tau}_S$ may be different than $\hat{\tau}_B$.

[9] Note, if $\mu_S \neq \mu_B$, then $\rho(\mu_S \mid \mu_B)\, \mathbb{E}_{\theta\sim\mu_B}[r(\theta)]^2 = \mathbb{E}_{\theta\sim\mu_B}[r^2(\theta)]$; Jensen's inequality ensures that $\rho(\mu_S \mid \mu_B) > 1$.

within the simplex.

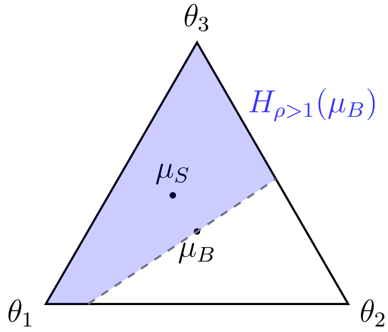


Figure 5.1 Illustration of the set $H_{\rho>1}(\mu_B)$.

Therefore, a Bayes plausible posterior spread $\tau_B \in \mathrm{PS}[\mu_B]$ can take one of two forms. First, consider the case in which $\operatorname{Supp} \tau_B \cap H_{\rho>1}(\mu_B) = \emptyset$, i.e., so, each posterior $\mu'_B \in \operatorname{Supp} \tau_B$ lies on the dotted line in Figure 5.1. In this case, each posterior satisfies $\rho(\mu'_B \mid \mu_B) = 1$, the Seller reveals only information that is orthogonal to the disagreement, and Bayes plausibility holds from both players' perspectives. Second, consider the case in which $\operatorname{Supp} \tau_B \cap H_{\rho>1}(\mu_B) \neq \emptyset$, i.e., some posteriors are relatively more likely from the Seller's perspective. In this case the Seller expects the Buyer's posterior μ'_B to drift towards μ_S and into $H_{\rho>1}(\mu_B)$. So, although the Buyer's posterior beliefs satisfy Bayes plausibility from his own perspective, they do not from Seller's. As we show below, the Seller can exploit this systematic disagreement to steer the Buyer's beliefs along paths that support higher payments.

A Recursive Approach Lemma 1 reduces the dimensionality of the Seller's problem: After each experiment, the evolution of beliefs is fully determined by the Buyer's current interim belief μ_B and his associated marginal distribution τ_B. This allows us to recursively define value functions $V^t : \Delta\Theta \to \mathbb{R}$ based upon V^1. Formally, for each $t > 1$, write

$$V^t(\mu_B) := \sup_{\tau_B \in \mathrm{PS}[\mu_B]} \left(\mathop{\mathbb{E}}_{\mu'_B \sim \tau_B} \left[U(\mu'_B) \right] - U(\mu_B) + \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B} \left[V^{t-1}(\mu'_B)\, \rho(\mu'_B \mid \mu_B) \right] \right). \quad (1)$$

We claim that $V^t(\mu_B)$ is the Seller's maximum revenue provided the Buyer holds belief μ_B and there are t periods left.

To see this, fix an experiment π that induces a posterior spread $\tau_B \in \mathrm{PS}[\mu_B]$. The first component,

$$\mathop{\mathbb{E}}_{\mu_B' \sim \tau_B}\big[U(\mu_B')\big] - U(\mu_B),$$

captures the *Seller's current revenue*: The Buyer's willingness to pay for an experiment inducing τ_B. (The players anticipate that $U(\mu_B')$ is the Buyer's continuation value under belief μ_B'.)

The second component,

$$\mathop{\mathbb{E}}_{\mu_B' \sim \tau_B}\big[V^{t-1}(\mu_B')\,\rho(\mu_B' \mid \mu_B)\big] = \mathop{\mathbb{E}}_{\mu_B' \sim \tau_S}\big[V^{t-1}(\mu_B')\big],$$

captures the *Seller's future revenue*: The expected revenue from the remaining $t-1$ periods. Importantly, the Seller's and Buyer's assessments of the continuation revenue generally differ. From the Buyer's perspective, the continuation revenue is averaged using τ_B. From the Seller's perspective this expectation must be adjusted by the distortion factor $\rho(\mu_B' \mid \mu_B)$, which reflects the divergence between the Buyer's posterior distribution τ_B and the corresponding Seller's posterior distribution τ_S.

Lemma A.5 establishes equilibrium existence and uniqueness of equilibrium payoffs, showing that the Seller's expected revenue is $V^T(\nu_B)$ and the Buyer's expected utility is $U(\nu_B)$. The proof proceeds by induction, and follows the logic of ultimatum-style games. When $t = 1$ period remains, the Seller's revenue is maximized by a fully-revealing experiment. When $t > 1$ periods remain, the Seller takes the continuation payoffs $U(\cdot)$ and $V^{t-1}(\cdot)$ as given and offers an experiment that maximizes the sum of current and future revenues, as described by Equation (1). Notably, as in ultimatum-style games, the Seller secures the full surplus with probability one at every subgame in equilibrium.

6 Example Revisited

To illustrate the recursive approach, we revisit the example from Section 2. Consider the two-period game and priors $(\nu_B, \nu_S) = (0.9, 0.5)$. In the first period ($t = 2$), the Seller chooses an experiment that solves Equation (1). Because the term $U(\nu_B)$ is constant, the Seller's problem reduces to finding a posterior spread τ that solves

$$\max_{\tau \in \mathrm{PS}[\nu_B]} \mathop{\mathbb{E}}_{\mu_B \sim \tau} \big[\Lambda^2(\mu_B \mid \nu_B)\big],$$

where the objective function for $t = 2$ periods is defined as

$$\Lambda^2(\mu_B \mid \nu_B) := U(\mu_B) + V^1(\mu_B)\,\rho(\mu_B \mid \nu_B).$$

Standard results in information design show that the optimal experiment is induced by the concave envelope of $\Lambda^2(\cdot \mid \nu_B)$. This objective function includes the distortion factor $\rho(\cdot \mid \nu_B)$, which captures the difference between the Seller's and the Buyer's assessment of future continuation values. Specifically, $\rho(\mu_B \mid \nu_B) > 1$ for all beliefs that satisfy $\mu_B(\overline{\theta}) < 0.9$, reflecting the Seller's belief that lower posteriors for state $\overline{\theta}$ are more likely than the Buyer anticipates.

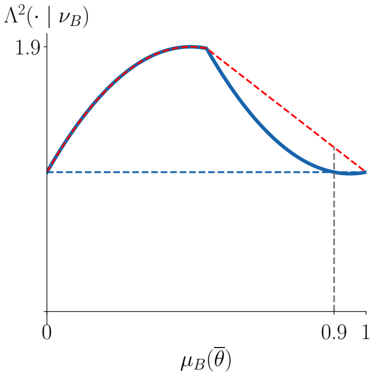


Figure 6.1 Objective $\Lambda^2(\cdot \mid \nu_B)$ and its concave envelope.

Figure 6.1 displays the objective $\Lambda^2(\cdot \mid \nu_B)$ (solid blue curve) and its concave envelope (red dashed line). The horizontal dashed blue line illustrates the surplus $U(\cdot) + V^1(\cdot)$ in the absence of belief disagreement. Importantly, without distortion, this sum is constant and equal to 1, meaning that the information sold in the first period would be irrelevant for total revenue, absent a distortion mapping.

The distortion ρ introduces a non-linearity in the objective function which leads the Seller to reveal some information in the first period. Geometrically, the distortion "inflates" the future value $V^1(\cdot)$ for the posterior beliefs that satisfy $\mu_B(\overline{\theta}) < \nu_B(\overline{\theta}) =$

0.9, creating a hump-shaped function for such beliefs. The concave envelope bridges the peak of this hump at $\mu_B(\overline{\theta}) = 0.5$ with the extreme belief $\mu_B(\overline{\theta}) = 1$, indicating that the optimal posterior spread τ_B^* randomizes between these two beliefs. Moreover, since the Buyer's expected utility U is linear over the interval $[0.5, 1]$, the Buyer's willingness to pay for this experiment is

$$\mathop{\mathbb{E}}_{\mu_B \sim \tau_B^*} \big[U(\mu_B)\big] - U(\nu_B) = U(\nu_B) - U(\nu_B) = 0.$$

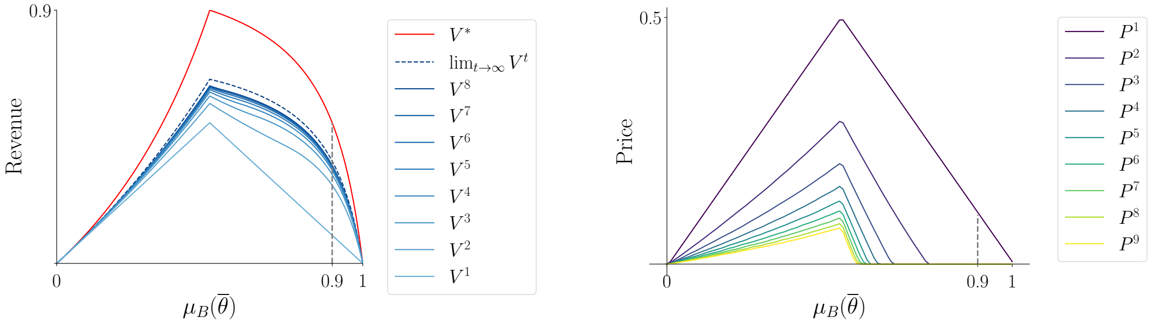
As a result, the free-sampling strategy identified in Section 2 is optimal, generating an expected revenue of $V^2(\nu_B) = 5/18$. Any strategy charging a positive price in the first period would strictly decrease expected revenue.

The analysis can be extended to a finite number of periods. The left panel of Figure 6.2 shows the value functions V^T for $T = 1, \ldots, 8$. The limit $\lim_{T\to\infty} V^T$ is shown by the blue dashed line, indicating decreasing marginal revenue as T grows large. The red mapping, $V^*(\cdot) = V^1(\cdot)\,\rho(\cdot \mid \nu_B)$, represents the Seller's revenue under commitment with limited liability.[10] The gap between V^* and the mappings V^T shows the value of commitment.

Notice, under the optimal contract, the Seller receives a positive payment only in state $\underline{\theta}$. This allocation allows the Seller to maximally exploit the belief disagreement by concentrating transfers in the state he considers relatively more likely. In contrast, in the sequential game without commitment, the Seller collects revenue by selling experiments at different interim beliefs. Thus, with positive probability, the Buyer makes payments even when the state is $\overline{\theta}$. This highlights the tension in the Seller's problem absent commitment: the Seller would prefer the Buyer to pay only in state $\underline{\theta}$. However, this is impossible in the sequential game, as the Buyer would never pay for information if she knew she would only have to pay in such state. So—in contrast with Example 2—selling information sequentially does not replicate the optimal commitment contract under limited liability, even as the number of periods diverges to infinity.

---

[10] The proof of Theorem 2 shows that (i) V^* is a fixed point of the Bellman operator, and (ii) evaluating V^* at the prior belief ν_B delivers the total revenue $V^*(\nu_B) = V^1(\nu_B)\max_{\theta\in\Theta} r(\theta)$.



(a)

(b)

Figure 6.2 Left panel: Value functions $V^1, \dots, V^8$, $\lim_{t\to\infty} V^t$, and V^*. Right panel: Price functions P^t for $t = 1, \dots, 8$. Each curve P^t shows the price of the optimal experiment as a function of the Buyer's belief for t remaining periods.

The right panel of Figure 6.2 describes the pricing of optimal experiments. For each period T, consider the price function

$$P^T(\mu_B) = \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B}\left[U(\mu'_B)\right] - U(\mu_B),$$

where τ_B is the solution to the Bellman equation (1) for T periods under belief μ_B. This function gives the price that the Seller charges for the optimal experiment when the Buyer's belief is μ_B and there are T periods remaining. These price functions are determined by the difference between the Buyer's expected value from observing the experiment and his current outside option $U(\mu_B)$. Observe that free sampling remains optimal at the prior for an arbitrary number of periods.

7 Proofs of Main Results

The proofs of our main results follow from identifying the geometric properties of the mappings $(V^t)_{t\in\mathbb{N}}$. With this in mind, let C^0_+ denote the space of non-negative continuous functions defined on $\Delta\Theta$, and write

$$\mathcal{F} := \left\{V \in C^0_+ : V(\delta_\theta) = 0 \text{ for all } \theta \in \Theta\right\}$$

for the set of candidate value functions—the functions that provide zero value at degenerate beliefs. Define the *Bellman operator* $\Phi : \mathcal{F} \to \mathcal{F}$ as follows:

$$\Phi(V)(\mu_B) := \sup_{\tau \in \mathrm{PS}[\mu_B]} \left\{ \mathop{\mathbb{E}}_{\mu'_B \sim \tau} \left[U(\mu'_B) + V(\mu'_B)\, \rho(\mu'_B \mid \mu_B) \right] \right\} - U(\mu_B).$$

Lemma A.6 in the Appendix ensures that Φ is well-defined, i.e., that $\Phi(V) \in \mathcal{F}$ for each $V \in \mathcal{F}$. Since $V^1 \in \mathcal{F}$, it follows by induction that $V^{t+1} = \Phi(V^t) \in \mathcal{F}$ for each $t \in \mathbb{N}$.[11]

As our next result shows, the properties of the mappings $(V^t)_{t \in \mathbb{N}}$ are determined by the properties of Φ and its associated distortion mapping ρ.

Lemma 2.

(i) Under PA, $V^T = V^1$ *for all* $T \in \mathbb{N}$.

(ii) Under PD, $V^2(\mu_B) > V^1(\mu_B)$ *for all* $\mu_B \in \operatorname{int} \Delta\Theta$.

(iii) Under SPD, for each $T \in \mathbb{N}$, $V^{T+1}(\mu_B) > V^T(\mu_B)$ *if and only if* $V^1(\mu_B) > 0$.

Lemma 2 shows that the geometric properties observed in Section 6 extend to arbitrary finite state spaces and decision problems. Under PA, the likelihood ratio is constant, so $\rho(\cdot \mid \mu_B) = 1$. Therefore, the objective function $\Lambda^2(\cdot \mid \mu_B)$ becomes the sum of the Buyer's utility and the continuation value, which is linear in beliefs. Because a linear function coincides with its concave envelope, sequential selling cannot create additional revenue.

Under PD, the distortion $\rho(\cdot \mid \mu_B)$ creates a non-linearity in the objective $\Lambda^2(\cdot \mid \mu_B)$. As in our main example, the Seller assigns relatively higher likelihood to the posteriors that lie in the direction of her own belief. This distortion inflates the expected continuation value $V^1(\cdot)$ for such posteriors.[12] The resulting asymmetry creates a "hump" in the objective function, generating a gap between the objective and its concave envelope. As a result, revealing no information in the first period

---

[11] As there is no time discounting, the operator Φ is not a Banach contraction. As a result, asymptotic behavior of the value functions is not derived from standard tools.

[12] Importantly, the distortion does not affect the value of some posterior beliefs. Indeed, the posterior μ_B (at which there is no information revelation) and all the degenerate posteriors (at which there is no future revenue) are not affected by the distortion. This helps to establish a gap between the objective Λ^2 and its concave envelope.

becomes strictly suboptimal, yielding a value $V^2(\nu_B)$ strictly greater than the static benchmark $V^1(\nu_B)$.

Under SPD, this logic applies recursively. Whenever $V^1(\mu_B) > 0$, the optimal experiment induces at least one non-degenerate posterior with positive probability; by SPD, disagreement therefore persists. In each such induced subgame, the same distortion effect arises. Therefore, the Seller becomes able to continuously inflate the perceived value of future revenue. As a result, $V^{T+1}(\nu_B) > V^T(\nu_B)$ for all $T \in \mathbb{N}$.

We now use Lemma 2 to prove our main results.

Proof of Theorem 1. Fix $T \in \mathbb{N}$. Lemma A.5 ensures that $\mathcal{R}_T = V^T(\nu_B)$ and that $\mathcal{U}_T$ is constant in T and equal to $U(\nu_B)$. The result then follows directly from Lemma 2. ■

Proof of Theorem 2. Write $\mathcal{P} := \left\{P \in \mathbb{R}_+^\Theta : \mathbb{E}_{\theta\sim\nu_B}[P(\theta)] \leq V^1(\nu_B)\right\}$, for the set of individually-rational contracts. The optimal contract solves

$$\max_{P\in\mathcal{P}} \mathop{\mathbb{E}}_{\theta\sim\nu_S}[P(\theta)].$$

The objective function is linear, and the constraint set $\mathcal{P}$ is non-empty, compact, and convex. Therefore, an extreme point of $\mathcal{P}$ attains the maximum. The set $\mathcal{P}$ contains $|\Theta| + 1$ extreme points. The first is the contract P_0 that assigns a zero payment at all states. The remaining extreme points are characterized by each state $\theta' \in \Theta$: For state θ', define the contract $P_{\theta'}$ that assigns a payment of $\frac{V^1(\nu_B)}{\nu_B(\theta')}$ to state θ' (the maximum payment consistent with individual rationality) and zero to all other states. Notice that,

$$\mathop{\mathbb{E}}_{\theta\sim\nu_S}[P_{\theta'}(\theta)] = V^1(\nu_B)\, r(\theta').$$

Consequently, an optimal contract P^* satisfies

$$\mathop{\mathbb{E}}_{\theta\sim\nu_S}[P^*(\theta)] = V^1(\nu_B) \max_{\theta\in\Theta} r(\theta).$$

Now, we prove that $\mathcal{R}_T \leq \mathbb{E}_{\theta\sim\nu_S}[P^*(\theta)]$ for all $T \in \mathbb{N}$. Let $\theta^* \in \arg\max_{\theta\in\Theta} r(\theta)$

and define $V^* : \Delta\Theta \to \mathbb{R}$ by

$$V^*(\mu_B) := V^1(\mu_B)\,\rho(\delta_{\theta^*} \mid \mu_B).$$

Note that $V^*(\nu_B) = V^1(\nu_B) \max_{\theta \in \Theta} r(\theta)$. Moreover, Lemma A.5 ensures that $\mathcal{R}_T = V^T(\nu_B)$. Hence, it suffices to show that $V^T \leq V^*$ for all $T \in \mathbb{N}$.

We proceed by induction. First, observe that $\rho(\delta_{\theta^*} \mid \mu_B) \geq 1$ for all $\mu_B \in \Delta\Theta$. This implies that $V^1 \leq V^*$. Now, assume $V^T \leq V^*$ for some $T \geq 1$. Lemma A.7 establishes that the operator Φ is weakly increasing, and Lemma A.8 establishes that V^* is a fixed point of Φ. Consequently, $V^T \leq V^*$, which implies that

$$V^{T+1} = \Phi(V^T) \leq \Phi(V^*) = V^*,$$

as claimed. ■

Proof of Proposition 1. Fix an SPE in the T-period game. Notice that players agree on the distribution over signals that each experiment specifies. As a result, conditional on each state θ, they also agree on the distribution over all paths of play and, therefore, on the total expected payment along those paths. Thus, the expected total payment conditional on θ does not depend on the players' prior beliefs. We can therefore denote this common quantity by $\mathcal{P}_T(\theta)$: the expected total payment in the T-period game conditional on the state θ.

Given that the Regulator's prior is $\nu_R = \lambda\nu_S + (1-\lambda)\nu_B$, his expected payment satisfies:

$$\begin{aligned}\mathop{\mathbb{E}}_{\theta\sim\nu_R}[\mathcal{P}_T(\theta)] &= \lambda \mathop{\mathbb{E}}_{\theta\sim\nu_S}[\mathcal{P}_T(\theta)] + (1-\lambda)\mathop{\mathbb{E}}_{\theta\sim\nu_B}[\mathcal{P}_T(\theta)] \\ &= \lambda\mathcal{R}_T + (1-\lambda)\mathcal{R}_1,\end{aligned}$$

where the second equality follows from Lemma A.5 and Theorem 1. Therefore, the Regulator's subjective expected payment is increasing in T. Consequently, consumer surplus from the Regulator's perspective, decreases (strictly decreases under SPD) as T increases. Consumer surplus is therefore maximized when $T = 1$. ■

8 Application: Markets for Attention

While our model assumes monetary transfers, the analysis extends to attention-based markets, in which the "price" of information is the time the Buyer spends watching advertisements. Consider an Expert (the Seller) who produces informative online content, such as videos, articles, newsletters, or podcasts, that advises a Viewer (the Buyer) on a particular decision problem (e.g. whether to buy or sell certain stocks.). The Expert monetizes by embedding advertisements in the content, treating time as the numeraire. Thus, the Viewer pays with attention, and the Expert's revenue is proportional to ad time consumed.

We focus on environments with belief disagreement. For instance, settings in which the Expert is more optimistic about the value of the content (holding an intermediate belief) than an overconfident Viewer (holding an extreme belief) or vice versa. Note that this setup is consistent with the assumptions of limited liability and limited commitment. Because the Viewer cannot engage in negative advertising time, the players cannot bet on the realized state, so full-commitment infinite revenue cannot be achieved. Moreover, since attention is voluntary, the Expert cannot enforce a contract that specifies state-contingent advertising time.

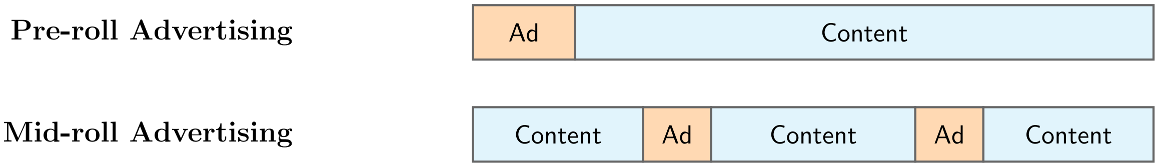


Figure 8.1 Advertising Strategies.

We focus on two common monetization strategies: pre-roll advertising, in which ads are shown before the content begins, and mid-roll advertising, in which ads are inserted in blocks throughout the content. (See Kerkhof (2024); Du, Luo, and Hu (2025)). Figure 8.1 compares pre-roll advertising with mid-roll advertising using three informative periods, including an initial segment of content that is free of ads.[13] We assume the Expert chooses both the advertising strategy and the information

---

[13] The need to preserve content coherence limits the maximum number of segments; this defines the maximum number of periods T of the interaction.

structure.[14] Pre-roll advertising represents a static scheme: all information is bundled and provided in exchange for a single upfront advertising exposure. In contrast, mid-roll advertising enables a sequential scheme with sequential information revelation and multiple ad breaks. Viewers can consume an initial free segment and then decide at each ad break whether to continue. This structure allows the Expert to strategically release information over time.

Our results show that, under a common prior, both strategies yield the same expected revenue, leaving no strategic advantage of mid-roll advertising. Nevertheless, under prior disagreement, mid-roll advertising strictly dominates pre-roll advertising. The Expert can tailor the information structure in a way that—on average—increases the Viewer's perceived value of later content. This encourages the Viewer to continue watching and tolerate additional ad breaks. Moreover, the analysis yields a simple and enforceable policy implication. A Regulator seeking to limit surplus extraction from consumers could mandate pre-roll advertising, thereby protecting consumers of informative content.

9 Discussion

Several extensions merit discussion. Our model abstracts from time discounting, costly experiments, and risk aversion—relaxing these assumptions would impact the Seller's incentives to delay information. We briefly discuss such extensions below.

Discounting the Future Our main finding reveals that the Seller benefits from longer interactions. Introducing a discount factor $\delta \in (0,1)$ makes delaying payments less attractive for the Seller. As a result, the Seller may forgo exploiting trading opportunities in the long run and instead disclose more information in early stages. At one extreme, when the Seller is highly impatient ($\delta \approx 0$), the optimal strategy

---

[14] For instance, on online platforms such as Twitch, mid-roll ads are commonly triggered by the streamer during live broadcasts. Other examples include platforms like Buzzsprout, Libsyn, or Substack, in which creators typically control ad placement. This assumption is without loss in environments in which the Platform chooses the advertising strategy but shares the Expert's objective (maximizing advertising time). In such cases, the Platform and the Expert can be regarded as a single decision maker.

is to disclose all information in a single period. At the other extreme, with a sufficiently patient Seller ($\delta \approx 1$), a continuity argument implies that our characterization provides a close approximation of equilibrium behavior and payoffs.

Costly Experiments This paper assumes that the Seller faces no cost for executing experiments. This benchmark reflects many economic interactions in which the marginal cost of an experiment is negligible. For example, software firms face zero marginal costs when running antivirus tests.

Introducing costs would affect the equilibrium design of experiments. With a small fixed cost per experiment, the Seller would reveal more information than in the no-cost case, aiming to reduce the expected number of experiments performed. If the cost per experiment is sufficiently high, the Seller would instead choose to offer a fully-revealing experiment in the first period.

Risk Aversion Our analysis assumes a risk-neutral Seller who maximizes expected revenue. A risk-averse Seller, however, would face a different trade-off. Under the optimal mechanism, the Seller's revenue depends on the realized path of signals—some paths yield higher payments than others. A sufficiently risk-averse Seller may prefer to secure a more certain payment rather than gamble on Nature selecting a favorable path. Risk aversion therefore pushes the Seller toward earlier information disclosure. By revealing more information upfront, the Seller reduces the variance of revenue at the expense of expected revenue. In the extreme case, a highly risk-averse Seller would fully disclose in the first period, eliminating all randomness in payments.

A Appendix

A.1 Subgame Perfect Equilibrium

As described in Section 3, the players in this game have no private information. Moreover, since prior beliefs are transparent and all signals are public, the players' posterior beliefs remain transparent after any sequence of experiments.

By using subgame perfect equilibrium, we implicitly impose the restriction that players "cannot signal what they do not know." That is, the actions of the co-player do not convey information about the state θ. At each history, the players' beliefs about the state depend solely on the signals selected by chance. More specifically, given a stream of experiments $\boldsymbol{\pi} = (\pi^t)_{t\in\mathcal{T}}$ purchased by the Buyer at periods $\mathcal{T} \subseteq \{1, 2, \dots, T\}$ the conditional distribution of the stream of signals $\boldsymbol{m} = (m^t)_{t\in\mathcal{T}}$ is given by

$$\mathbb{P}_{\boldsymbol{\pi}}[\boldsymbol{m} \mid \theta] := \prod_{t\in\mathcal{T}} \pi^t(m^t \mid \theta).$$

Therefore, since i has prior $\nu_i \in \operatorname{int}\Delta\Theta$, i's posterior beliefs at period $\hat{t} < \min(\mathcal{T})$ are given by

$$\mu_i^{\hat{t}}(\theta) = \frac{\nu_i(\theta)\,\mathbb{P}_{\boldsymbol{\pi}}[\boldsymbol{m} \mid \theta]}{\displaystyle\sum_{\theta'\in\Theta} \nu_i(\theta')\,\mathbb{P}_{\boldsymbol{\pi}}[\boldsymbol{m} \mid \theta']}. \tag{2}$$

Absent any experiment, the belief of each agent i remains fixed at the prior ν_i, even after a deviation of the co-player.

In this game, the players' posterior beliefs are transparent after any history. Hence, subgame perfect equilibrium is equivalent to strong perfect Bayesian equilibrium under the following requirements: First, at each history that implements some experiments, the players' beliefs are described by Equation (2). Second, at each history in which the Buyer does not purchase any experiment, the belief of agent i equals its prior ν_i.

A.2 Omitted Proofs

Proof of Lemma 1. Assume there exists an experiment π inducing distributions $(\hat{\tau}_S, \hat{\tau}_B)$ over posterior profiles. Alonso and Câmara (2016) shows $\operatorname{Supp}\hat{\tau}_S \subset \Omega$ and (ii). We will show (i) and (iii):

(i) The profile (μ'_S, μ'_B) is in $\operatorname{Supp}\hat{\tau}_S$ if there is a signal m inducing the posterior profile (μ'_S, μ'_B) such that $\mathbb{E}_{\theta'\sim\mu_S}[\pi(m \mid \theta')] > 0$. Hence, there exists θ such that $\pi(m \mid \theta)\mu_S(\theta) > 0$. Because $(\mu_S, \mu_B) \in \Omega$ implies that

$\mu_S(\theta) = \frac{r(\theta)}{\mathbb{E}_{\theta' \sim \mu_B}[r(\theta')]}\, \mu_B(\theta)$. Since $\mu_S(\theta) > 0$ implies $\mu_B(\theta) > 0$, it follows that $\mathbb{E}_{\theta' \sim \mu_B}[\pi(m \mid \theta')] > 0$. Hence, (μ'_S, μ'_B) is in $\operatorname{Supp} \hat{\tau}_B$. An analogous argument shows the converse.

(iii) Fix a belief profile (μ'_S, μ'_B). Note, item (i) ensures that $\operatorname{Supp} \hat{\tau}_S = \operatorname{Supp} \hat{\tau}_B$, the equality of interests holds trivially if $(\mu'_S, \mu'_B) \notin \operatorname{Supp} \hat{\tau}_S$. Now, consider the case $(\mu'_S, \mu'_B) \in \operatorname{Supp} \hat{\tau}_S$. Observe that item (i) implies that the messages inducing the posterior profile (μ'_S, μ'_B) from the Seller's and Buyer's perspectives are the same. Write $M(\mu'_S, \mu'_B) \subseteq M$ for the set of messages that induce the posterior profile (μ'_S, μ'_B). Therefore, for each player i,

$$\hat{\tau}_i(\mu'_S, \mu'_B) = \sum_{m \in M(\mu'_S, \mu'_B)} \mathop{\mathbb{E}}_{\theta \sim \mu_i} \big[\pi(m \mid \theta)\big].$$

Notice, Bayes rule implies that, for each $m \in M(\mu'_S, \mu'_B)$ and each $\theta \in \operatorname{Supp} \mu_i$,

$$\pi(m \mid \theta) = \tfrac{\mu'_i(\theta)}{\mu_i(\theta)} \mathop{\mathbb{E}}_{\theta' \sim \mu_i} \big[\pi(m \mid \theta')\big]. \tag{3}$$

In addition, since $\operatorname{Supp} \mu_S = \operatorname{Supp} \mu_B$ and $\mu_S(\theta) = \frac{r(\theta)}{\mathbb{E}_{\theta' \sim \mu_B}[r(\theta')]} \mu_B(\theta)$ for each $\theta \in \Theta$, it follows that

$$\begin{aligned}
\mathop{\mathbb{E}}_{\theta \sim \mu_S} \left[\tfrac{\mu'_B(\theta)}{\mu_B(\theta)}\right] &= \sum_{\theta \in \operatorname{Supp} \mu_B} \frac{\mu_S(\theta)}{\mu_B(\theta)} \mu'_B(\theta) \\
&= \frac{1}{\mathbb{E}_{\theta' \sim \mu_B}[r(\theta')]} \sum_{\theta \in \operatorname{Supp} \mu_B} r(\theta) \mu'_B(\theta) \\
&= \tfrac{\mathbb{E}_{\theta \sim \mu'_B}[r(\theta)]}{\mathbb{E}_{\theta \sim \mu_B}[r(\theta)]}.
\end{aligned} \tag{4}$$

Therefore,

$$\begin{aligned}
\hat{\tau}_S(\mu'_S,\mu'_B) &= \sum_{m\in M(\mu'_S,\mu'_B)} \mathop{\mathbb{E}}_{\theta\sim\mu_S}[\pi(m\mid\theta)] \\
&= \sum_{m\in M(\mu'_S,\mu'_B)} \mathop{\mathbb{E}}_{\theta\sim\mu_S}\left[\frac{\mu'_B(\theta)}{\mu_B(\theta)} \mathop{\mathbb{E}}_{\theta'\sim\mu_B}[\pi(m\mid\theta')]\right] \\
&= \mathop{\mathbb{E}}_{\theta\sim\mu_S}\left[\frac{\mu'_B(\theta)}{\mu_B(\theta)} \sum_{m\in M(\mu'_S,\mu'_B)} \mathop{\mathbb{E}}_{\theta'\sim\mu_B}[\pi(m\mid\theta')]\right] \\
&= \mathop{\mathbb{E}}_{\theta\sim\mu_S}\left[\frac{\mu'_B(\theta)}{\mu_B(\theta)}\right] \hat{\tau}_B(\mu'_S,\mu'_B) \\
&= \frac{\mathbb{E}_{\theta\sim\mu'_B}[r(\theta)]}{\mathbb{E}_{\theta\sim\mu_B}[r(\theta)]}\,\hat{\tau}_B(\mu'_S,\mu'_B),
\end{aligned}$$

where the second and fifth equalities follow from Equations (3) and (4), respectively.

Now, suppose that $(\hat{\tau}_S,\hat{\tau}_B)$ satisfy conditions (i)-(iii). Assume without loss of generality that $M = \operatorname{Supp}\hat{\tau}_S$. Consider the experiment $\pi:\Theta\to\Delta M$ given by

$$\pi\big((\mu'_S,\mu'_B)\mid\theta\big) = \frac{\mu'_S(\theta)}{\mu_S(\theta)}\,\hat{\tau}_S(\mu'_S,\mu'_B).$$

Simple computations show that π induces $(\hat{\tau}_S,\hat{\tau}_B)$. ■

Lemma A.3. *The mapping $U:\Delta\Theta\to\mathbb{R}$ is continuous and convex.*

Proof. For each $a\in A$, the mapping $u_a:\Delta\Theta\to\mathbb{R}$ defined by $u_a(\mu):=\mathbb{E}_{\theta\sim\mu}[u(a,\theta)]$ is linear. So, $U(\mu)=\max_{a\in A}u_a(\mu)$ is convex and continuous as it is the maximum of a set of linear functions. ■

Lemma A.4. *Fix a mapping $\Lambda:\Delta\Theta\times\Delta\Theta\to\mathbb{R}$ and let $F:\Delta\Theta\to\mathbb{R}$ be defined by*

$$F(\mu) := \sup_{\tau\in\mathrm{PS}[\mu]} \mathop{\mathbb{E}}_{\mu'\sim\tau}\big[\Lambda(\mu',\mu)\big].$$

If Λ is continuous, then F is continuous.

Proof. Fix $\mu\in\Delta\Theta$ and a sequence $(\mu_k)_{k\in\mathbb{N}}$ with $\mu_k\to\mu$. We show $\lim F(\mu_k)=F(\mu)$ by proving $\limsup_{k\to\infty}F(\mu_k)\le F(\mu)$ and $\liminf_{k\to\infty}F(\mu_k)\ge F(\mu)$. By Kamenica and Gentzkow (2011), there exists $\tau \in \mathrm{PS}[\mu]$ achieving $F(\mu)$ and an experiment $\pi : \Theta \to \Delta M$ inducing τ. There exists an affine mapping $L : \Delta\Theta \to \mathbb{R}$ such that $L(\mu) = F(\mu)$ and $L(\mu') \geq \Lambda(\mu', \mu)$ for all $\mu' \in \Delta\Theta$.

Step 1. Since $\Delta\Theta$ is compact and Λ and L are continuous, they are uniformly continuous. Fix $\varepsilon > 0$. There exists $\delta > 0$ such that $\|\mu_k - \mu\|_\infty < \delta$ implies $|L(\mu_k) - L(\mu)| < \varepsilon/2$ and $|\Lambda(\mu', \mu_k) - \Lambda(\mu', \mu)| < \varepsilon/2$ for all μ'. Thus,

$$F(\mu_k) \leq \sup_{\tau' \in \mathrm{PS}[\mu_k]} \mathop{\mathbb{E}}_{\mu' \sim \tau'} \left[L(\mu') + \tfrac{\varepsilon}{2}\right] = L(\mu_k) + \tfrac{\varepsilon}{2} < L(\mu) + \varepsilon = F(\mu) + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary, $\limsup_{k\to\infty} F(\mu_k) \leq F(\mu)$.

Step 2. Let $\tau_k \in \mathrm{PS}[\mu_k]$ be induced by π under μ_k. For each $m \in M$, let $\mathbb{P}_\pi[m \mid \mu]$ be the probability of m given μ, and $\mathcal{P}_m(\mu)$ the posterior after m. The maps $\mathbb{P}_\pi[m \mid \cdot]$ and $\mathcal{P}_m(\cdot)$ are continuous. Since Λ is continuous,

$$\lim_{k\to\infty} \mathbb{P}_\pi[m \mid \mu_k]\, \Lambda\big(\mathcal{P}_m(\mu_k), \mu_k\big) = \mathbb{P}_\pi[m \mid \mu]\, \Lambda\big(\mathcal{P}_m(\mu), \mu\big).$$

Summing over $m \in M$,

$$\lim_{k\to\infty} \mathop{\mathbb{E}}_{\mu' \sim \tau_k} \left[\Lambda(\mu', \mu_k)\right] = \mathop{\mathbb{E}}_{\mu' \sim \tau} \left[\Lambda(\mu', \mu)\right] = F(\mu).$$

Since $F(\mu_k) \geq \mathbb{E}_{\mu' \sim \tau_k}\left[\Lambda(\mu', \mu_k)\right]$, it follows that $\liminf_{k\to\infty} F(\mu_k) \geq F(\mu)$. ■

Lemma A.5. *Fix* $t \in \mathbb{N}$ *with* $0 \leq t \leq T$, *and* $\mu_B \in \Delta\Theta$. *In any subgame with* t *remaining information-trading periods in which the Buyer's belief is* μ_B, *the following is true:*

- *(i) An SPE exists and SPE payoffs are unique.*
- *(ii) The Buyer's equilibrium payoff is* $U(\mu_B)$.
- *(iii) For each* $t \geq 1$, *the Seller's expected revenue is* $V^t(\mu_B)$.

Proof. The proof proceeds in two steps. Step 1 shows that if an SPE exists, then SPE payoffs are unique and equal to $U(\mu_B)$ and $V^t(\mu_B)$. Step 2 establishes SPE existence.

Step 1. Assume an SPE exists. We first compute the Buyer's SPE payoffs by induction. Note that in any SPE when $t = 0$, the Buyer's expected payoff is

$$\max_{a \in A} \mathop{\mathbb{E}}_{\theta \sim \mu_B} [u(a, \theta)] = U(\mu_B).$$

Now, suppose the statement holds for t with $0 \leq t < T$. Let $\bar{u}$ be a Buyer's equilibrium payoff for $t + 1$. Fix an SPE inducing Buyer's payoff $\bar{u}$. Because it is not profitable for the Buyer to deviate at the first period by rejecting all experiments and then sticking back to the SPE, it must be that $\bar{u} \geq U(\mu_B)$.

Now, assume that $\bar{u} > U(\mu_B)$. Notice, this implies that the Buyer accepts the Seller's offer (π, p) at the first period. Otherwise, we would have that $U(\mu_B) \geq \bar{u}$, a contradiction. Therefore,

$$\bar{u} = \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B} [U(\mu'_B)] - p > U(\mu_B),$$

where τ_B is the Buyer's distribution over Buyer's posteriors that π induces. Note, the Buyer will accept π at a slightly higher price at the first period, which implies that the Seller has incentives to deviate in such period. In conclusion, $\bar{u} \leq U(\mu_B)$.

We now compute the Seller's SPE revenue by induction. Let υ be a Seller's SPE revenue for $t = 1$. Since, for any $\varepsilon > 0$, the Buyer will accept a fully-revealing experiment at a price of $V^1(\mu_B) - \varepsilon$, it follows that $\upsilon \geq V^1(\mu_B)$. In addition, the Buyer will reject any experiment for a price higher than $V^1(\mu_B)$. Hence, $\upsilon \leq V^1(\mu_B)$.

Now, assume that the statement holds for t with $0 \leq t < T$. Write π^* for a solution to the maximization problem defining $V^{t+1}(\mu_B)$, τ_S^* and τ_B^* for the Buyer's and Seller's distribution over Buyer's posteriors μ'_B that π^* induces, and define $p^* = \mathbb{E}_{\mu'_B \sim \tau_B^*}[U(\mu'_B)] - U(\mu_B)$.

Let υ be a Seller's equilibrium revenue for $t + 1$. Pick an SPE inducing υ, let (π, p) be the first experiment it prescribes, and let τ_B be the Buyer's distribution over Buyer's posteriors that π induces. Recall, the corresponding Seller's distribution τ_S over Buyer's posteriors is $\tau_B(\cdot)\rho(\cdot \mid \mu_B)$. (See Lemma 1.) If the Buyer rejects (π, p), then $\upsilon = V^t(\mu_B) \leq V^{t+1}(\mu_B)$. If the Buyer accepts (π, p), then $\mathbb{E}_{\mu'_B \sim \tau_B}[U(\mu'_B)] - p \geq$

$U(\mu_B)$. Consequently,

$$\begin{aligned}
\upsilon &= p + \mathop{\mathbb{E}}_{\mu'_B \sim \tau_S}\left[V^t(\mu'_B)\right] \\
&= p + \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B}\left[V^t(\mu'_B)\,\rho(\mu'_B \mid \mu_B)\right] \\
&\leq \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B}\left[U(\mu'_B) - U(\mu_B) + V^t(\mu'_B)\,\rho(\mu'_B \mid \mu_B)\right] \\
&\leq V^{t+1}(\mu_B)
\end{aligned}$$

In any case, $\upsilon \leq V^{t+1}(\mu_B)$.

Suppose now that $\upsilon < V^{t+1}(\mu_B)$. Notice, the Buyer will accept experiment π^* for a price $p' = p^* - \varepsilon$ for each $\varepsilon > 0$ small enough, because

$$\mathop{\mathbb{E}}_{\mu'_B \sim \tau^*_B}[U(\mu'_B)] - p' > \mathop{\mathbb{E}}_{\mu'_B \sim \tau^*_B}[U(\mu'_B)] - p^* = U(\mu_B).$$

Therefore, the Seller has incentives to deviate from the priced experiment (π, p) to (π^*, p') and then conform back to the SPE. Indeed, deviating yields a payoff of

$$\begin{aligned}
p' + \mathop{\mathbb{E}}_{\mu'_B \sim \tau^*_S}[V^t(\mu'_B)] &= p' + \mathop{\mathbb{E}}_{\mu'_B \sim \tau^*_B}[V^t(\mu'_B)\rho(\mu'_B \mid \mu_B)] \\
&= p^* + \mathop{\mathbb{E}}_{\mu'_B \sim \tau^*_B}[V^t(\mu'_B)\rho(\mu'_B \mid \mu_B)] \\
&= V^{t+1}(\mu_B) - \varepsilon.
\end{aligned}$$

Since $\varepsilon > 0$ is small, this payoff is strictly higher than υ, contradicting that υ is a SP payoff. In conclusion, $\upsilon \geq V^{t+1}(\mu_B)$.

Step 2. Consider the following strategy profile:

> **Buyer's strategy:** In any Buyer's history with $t = 0$ and Buyer's belief μ_B, choose any action in $\arg\max_{a \in A} \mathbb{E}_{\theta \sim \mu_B}[u(a, \theta)]$. In addition, at each Buyer's history with $t > 0$ remaining periods, Buyer's belief μ_B, and priced experiment (π, p), the Buyer accepts the offer if and only if $\mathbb{E}_{\mu'_B \sim \tau_B}[U(\mu'_B)] - p \geq U(\mu_B)$, where τ_B is the Buyer's distribution over Buyer's posteriors that π induces.

Seller's Strategy: In any Seller's history with $t \geq 1$, offer any experiment π that solves the maximization problem defining $V^t(\mu_B)$ at a price

$$p = \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B}[U(\mu'_B)] - U(\mu_B),$$

where τ_B is the Buyer's distribution over Buyer's posteriors that π induces.

Using Step 1, the one-shot deviation principle ensures that this strategy profile constitutes an SPE. ■

Lemma A.6. *If $V \in \mathcal{F}$, then $\Phi(V) \in \mathcal{F}$.*

Proof. Fix $V \in \mathcal{F}$. Note, the mapping $\Lambda(\mu', \mu) := U(\mu') + V(\mu')\rho(\mu' \mid \mu)$ is continuous. Lemma A.4 guarantees that $\Phi(V)$ is continuous. Fix $\theta \in \Theta$. Observe that $\tau \in \mathrm{PS}[\delta_\theta]$ if and only if $\mathrm{Supp}\,(\tau) = \{\delta_\theta\}$. Thus,

$$\Phi(V)(\delta_\theta) = U(\delta_\theta) + V(\delta_\theta) - U(\delta_\theta) = 0.$$

Therefore, $\Phi(V) \in \mathcal{F}$. ■

Lemma A.7. *The Bellman operator Φ is increasing.*

Proof. Fix $\mu \in \Delta\Theta$, and $V, W \in \mathcal{F}$. Assume $V \geq W$. Then,

$$\begin{aligned}
\Phi(W)(\mu) &= \sup_{\tau' \in \mathrm{PS}[\mu]} \mathop{\mathbb{E}}_{\mu' \sim \tau'}\big[U(\mu') + W(\mu')\rho(\mu' \mid \mu)\big] - U(\mu) \\
&\geq \sup_{\tau' \in \mathrm{PS}[\mu]} \mathop{\mathbb{E}}_{\mu' \sim \tau'}\big[U(\mu') + V(\mu')\rho(\mu' \mid \mu)\big] - U(\mu) \\
&= \Phi(V)(\mu).
\end{aligned}$$

■

Lemma A.8. *Fix $\theta^* \in \arg\max_{\theta \in \Theta} r(\theta)$ and define $V^*(\mu_B) := V^1(\mu_B)\ \rho(\delta_{\theta^*} \mid \mu_B)$. Then V^* is a fixed point of Φ.*

Proof. Fix $\mu_B, \mu'_B \in \Delta\Theta$. Since $\rho(\mu \mid \mu')\ \rho(\mu' \mid \mu'') = \rho(\mu \mid \mu'')$ for all $\mu, \mu', \mu'' \in \Delta\Theta$, then

$$\begin{aligned}
\Lambda(\mu'_B, \mu_B) &:= U(\mu'_B) + V^*(\mu'_B)\ \rho(\mu'_B \mid \mu_B) \\
&= U(\mu'_B) + V^1(\mu'_B)\ \rho(\delta_{\theta^*} \mid \mu_B) \\
&= \big(1 - \rho(\delta_{\theta^*} \mid \mu_B)\big)\, U(\mu'_B) + \rho(\delta_{\theta^*} \mid \mu_B) \mathop{\mathbb{E}}_{\theta \sim \mu'_B}[U(\delta_\theta)]
\end{aligned}$$

Since $\rho(\delta_{\theta^*} \mid \mu_B) = \frac{r(\theta^*)}{\mathbb{E}_{\theta\sim\mu_B}[r(\theta)]} \geq 1$ and U is convex (see Lemma A.3), then Λ is concave in its first argument. In addition, note that $\Lambda(\mu_B, \mu_B) = U(\mu_B) + V^*(\mu_B)$. Therefore,

$$\begin{aligned}
\Phi(V^*)(\mu_B) &= \sup_{\tau \in \mathrm{PS}[\mu_B]} \mathop{\mathbb{E}}_{\mu'_B \sim \tau}[\Lambda(\mu'_B, \mu_B)] - U(\mu_B) \\
&= \mathop{\mathbb{E}}_{\mu'_B \sim \delta_{\mu_B}}[\Lambda(\mu'_B, \mu_B)] - U(\mu_B) \\
&= \Lambda(\mu_B, \mu_B) - U(\mu_B) \\
&= V^*(\mu_B).
\end{aligned}$$

■

Proof of Lemma 2.

(i) Assume PA. So, $\rho(\cdot \mid \cdot) = 1$. Since $U + V^1$ is linear, it follows that $V^2 = V^1$. By using an induction argument, $V^T = V^1$ for all $T \in \mathbb{N}$.

(ii) Assume PD. Fix a belief $\mu_B \in \operatorname{int}\Delta\Theta$ and let $\theta^* \in \arg\max_{\theta\in\Theta} r(\theta)$. Observe that

$$\rho(\delta_{\theta^*} \mid \mu_B) = \frac{r(\theta^*)}{\mathop{\mathbb{E}}_{\theta\sim\mu_B}[r(\theta)]} > 1, \quad \text{while} \quad \rho(\mu_B \mid \mu_B) = 1.$$

Since $\rho(\cdot \mid \mu_B)$ is linear and μ_B has full support, there is some full-support belief $\hat{\mu}_B$ in a neighborhood of μ_B that satisfies $\rho(\hat{\mu}_B \mid \mu_B) > 1$. Since $\{\delta_\theta : \theta \in \Theta\}$ are the extreme points of $\Delta\Theta$, there is some $\tau_B \in \mathrm{PS}[\mu_B]$ with support $\{\hat{\mu}_B\} \cup \{\delta_\theta : \theta \in \Theta\}$. Observe, since $V^1(\hat{\mu}_B) > 0$, $\rho(\hat{\mu}_B \mid \mu_B) > 1$, and $V^1(\delta_\theta) = 0$ for each θ, we obtain

$$\begin{aligned}
V^2(\mu_B) &= \sup_{\tau'_B \in \mathrm{PS}[\mu_B]} \mathop{\mathbb{E}}_{\mu'_B \sim \tau'_B} \big[U(\mu'_B) + V^1(\mu'_B)\, \rho(\mu'_B \mid \mu_B)\big] - U(\mu_B) \\
&\geq \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B} \big[U(\mu'_B) + V^1(\mu'_B)\, \rho(\mu'_B \mid \mu_B)\big] - U(\mu_B) \\
&> \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B} \big[U(\mu'_B) + V^1(\mu'_B)\big] - U(\mu_B) \\
&= \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B} \big[\mathop{\mathbb{E}}_{\theta \sim \mu'_B} [U(\delta_\theta)]\big] - U(\mu_B) \\
&= \mathop{\mathbb{E}}_{\theta \sim \mu_B} [U(\delta_\theta)] - U(\mu_B) \\
&= V^1(\mu_B),
\end{aligned}$$

where the second equality follows from the definition of V^1 and the third follows from the fact that $\tau_B \in \mathrm{PS}[\mu_B]$.

(iii) Assume SPD. We show the result by using induction over T.

Base case. We show the result holds for $T = 1$. Suppose that $V^1(\mu_B) = 0$. So, each $\mu'_B \in \Delta\Theta$ with $\operatorname{Supp} \mu'_B \subseteq \operatorname{Supp} \mu_B$, satisfies $V^1(\mu'_B) = 0$. Then,

$$\begin{aligned}
V^2(\mu_B) &= \sup_{\tau \in \mathrm{PS}[\mu_B]} \mathop{\mathbb{E}}_{\mu'_B \sim \tau} [U(\mu'_B) + V^1(\mu'_B)\rho(\mu'_B \mid \mu_B)] - U(\mu_B) \\
&= \sup_{\tau \in \mathrm{PS}[\mu_B]} \mathop{\mathbb{E}}_{\mu'_B \sim \tau} [U(\mu'_B)] - U(\mu_B) \\
&= V^1(\mu_B) \\
&= 0.
\end{aligned}$$

Now, assume that $V^1(\mu_B) > 0$. Observe that $|\operatorname{Supp} \mu_B| > 1$. Let $\hat{\theta} = \arg\max_{\theta \in \operatorname{Supp} \mu_B} r(\theta)$. Note, SPD implies that

$$\rho(\delta_{\hat{\theta}} \mid \mu_B) = \frac{r(\hat{\theta})}{\mathbb{E}_{\theta \sim \mu_B}[r(\theta)]} > 1, \quad \text{while} \quad \rho(\mu_B \mid \mu_B) = 1.$$

Therefore, there exists some $\hat{\mu}_B \in \Delta\Theta$ such that $\operatorname{Supp} \hat{\mu}_B = \operatorname{Supp} \mu_B$ and $\rho(\hat{\mu}_B \mid \mu_B) > 1$. Moreover, $\operatorname{Supp} \hat{\mu}_B = \operatorname{Supp} \mu_B$ implies that $V^1(\hat{\mu}_B) > 0$. So, in a similar way as in (ii), there is some $\tau_B \in \mathrm{PS}[\mu_B]$ with support $\{\hat{\mu}_B\} \cup \{\delta_\theta : \theta \in \operatorname{Supp} \mu_B\}$. Hence,

$$\begin{aligned} V^2(\mu_B) &\geq \mathop{\mathbb{E}}_{\mu'_B \sim \tau_B} \left[U(\mu'_B) + V^1(\mu'_B)\, \rho(\mu'_B \mid \mu_B) \right] - U(\mu_B) \\ &> V^1(\mu_B). \end{aligned}$$

Inductive step. Fix $T \geq 1$ and assume that the statement holds for all t with $1 \leq t \leq T$. First, suppose by contrapositive that $V^1(\mu_B) = 0$. Then, for each $\mu'_B \in \Delta\Theta$ such that $\operatorname{Supp} \mu'_B \subseteq \operatorname{Supp} \mu_B$, it follows that $V^1(\mu'_B) = 0$. By the inductive hypothesis, we obtain $V^{T+1}(\mu'_B) = 0$ for all $\mu'_B \in \Delta\Theta$ with $\operatorname{Supp} \mu'_B \subseteq \operatorname{Supp} \mu_B$. Consequently,

$$\begin{aligned} V^{T+2}(\mu_B) &= \sup_{\tau \in \mathrm{PS}[\mu_B]} \mathop{\mathbb{E}}_{\mu'_B \sim \tau} \left[U(\mu'_B) + V^{T+1}(\mu'_B)\, \rho(\mu'_B \mid \mu_B) \right] - U(\mu_B) \\ &= \sup_{\tau \in \mathrm{PS}[\mu_B]} \mathop{\mathbb{E}}_{\mu'_B \sim \tau} \left[U(\mu'_B) \right] - U(\mu_B) \\ &= V^1(\mu_B) \\ &= 0. \end{aligned}$$

Consequently, $V^{T+2}(\mu_B) = V^{T+1}(\mu_B)$.

Next, suppose by contrapositive that $V^{T+2}(\mu_B) = V^{T+1}(\mu_B)$. Let $\tau' \in \mathrm{PS}[\mu_B]$ be the posterior distribution that attains $V^{T+1}(\mu_B)$. Then,

$$V^{T+2}(\mu_B) = \mathop{\mathbb{E}}_{\mu'_B \sim \tau'} \left[U(\mu'_B) + V^T(\mu'_B)\, \rho(\mu'_B \mid \mu_B) \right] - U(\mu_B).$$

Therefore, for each $\mu'_B \in \operatorname{Supp} \tau'$, we must have $V^{T+1}(\mu'_B) = V^T(\mu'_B)$. By the inductive hypothesis, this implies $V^T(\mu'_B) = V^1(\mu'_B) = 0$ for each $\mu'_B \in \operatorname{Supp} \tau'$. Consequently, $V^{T+1}(\mu_B) = 0$. Since $V^{T+1} \geq V^1$, we conclude that $V^1(\mu_B) = 0$.

■

Lemma A.9. *In Example 2, $\mathcal{R}_T = \mathcal{R}_2 > \mathcal{R}_1$ for each $T \geq 2$.*

Proof. Consider the following two-period selling strategy. First, the Seller offers a free experiment that reveals whether the state is θ_3 or not. Second, the Seller offers a fully-revealing experiment for a price of 1. Observe that the Buyer accepts the initial offer as it provides free information. If the state is revealed to be θ_3, the Buyer rejects the second offer. Otherwise, the Buyer updates his belief to $\mu'_B = (\frac{1}{2}, \frac{1}{2}, 0)$ and is willing to pay $V^1(\mu'_B) = 1$ for full information. Hence, conditional on the first experiment discarding state θ_3, he accepts the second offer.

The Seller thus receives a payment of 1 if θ_3 is not the true state, and 0 otherwise. Observe, from the Seller's perspective, this sequential selling strategy yields an expected revenue of $\frac{4}{5}$. Consequently, $\mathcal{R}_2 \geq \frac{4}{5}$. Moreover, since

$$V^*(\nu_B) = V^1(\nu_B) \max_{\theta \in \Theta} r(\theta) = \tfrac{4}{5},$$

Theorem 2 implies that $\mathcal{R}_T = \frac{4}{5}$ for each $T \geq 2$. $\square$

References

Avidit Acharya and Juan Ortner. Progressive learning. *Econometrica*, 85(6):1965–1990, 2017.

S Nageeb Ali, Nima Haghpanah, Xiao Lin, and Ron Siegel. How to sell hard information. *The Quarterly Journal of Economics*, 137(1):619–678, 2022.

Ricardo Alonso and Odilon Câmara. Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory*, 165:672–706, 2016.

Robert J Aumann. Agreeing to disagree. *The Annals of Statistics*, pages 1236–1239, 1976.

Dirk Bergemann and Alessandro Bonatti. Markets for information: An introduction. *Annual Review of Economics*, 11:85–107, 2019.

Dirk Bergemann, Alessandro Bonatti, and Alex Smolin. The design and price of information. *American Economic Review*, 108(1):1–48, 2018.

Dirk Bergemann, Alessandro Bonatti, and Tan Gan. The economics of social data. *The RAND Journal of Economics*, 53(2):263–296, 2022.

Helmut Bester and Roland Strausz. Contracting with imperfect commitment and the revelation principle: the single agent case. *Econometrica*, 69(4):1077–1098, 2001.

Yeon-Koo Che and Navin Kartik. Opinions as incentives. *Journal of Political Economy*, 117(5):815–860, 2009.

Laura Doval and Vasiliki Skreta. Mechanism design with limited commitment. *Econometrica*, 90(4):1463–1500, 2022.

Shaofu Du, Zaichen Luo, and Li Hu. A game-theoretic analysis for creative mid-roll ads on a content platform. *International Journal of Production Research*, 63(11): 4217–4237, 2025.

Manuel Foerster and Fynn Närmann. Sequential information selling: Perfect price discrimination and the role of encryption. *Working paper*, 2025.

Michael D. Grubb. Selling to overconfident consumers. *American Economic Review*, 99(5):1770–1807, 2009.

Joseph Y Halpern. Characterizing the common prior assumption. *Journal of Economic Theory*, 106(2):316–355, 2002.

John C Harsanyi. Games with incomplete information played by “bayesian” players part ii. bayesian equilibrium points. *Management Science*, 14(5):320–334, 1968.

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

Shota Ichihashi. The economics of data externalities. *Journal of Economic Theory*, 196:105316, 2021.

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

Anna Kerkhof. Advertising and content differentiation: Evidence from youtube. *The Economic Journal*, 134(663):2912–2950, 2024.

Vijay Krishna and John Morgan. Contracting for information under imperfect commitment. *The RAND Journal of Economics*, 39(4):905–925, 2008.

Stephen Morris. The common prior assumption in economic theory. *Economics & Philosophy*, 11(2):227–253, 1995.

Luis Rayo and Ilya Segal. Optimal information disclosure. *Journal of political Economy*, 118(5):949–987, 2010.

Leonard J Savage. *The foundations of statistics.* Courier Corporation, 1972.

Shuran Zheng and Yiling Chen. Optimal advertising for information products. In *Proceedings of the 22nd ACM Conference on Economics and Computation*, pages 888–906, 2021.

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