

Selling Supplemental Information ^{*}

Arlindo Skenderaj [†]

November 19, 2025

[Please click here for the most recent version.](#)

Abstract

I consider an environment in which a decision maker faces uncertainty and privately holds information in the form of a signal about the true state of the world. The decision maker purchases additional information from a data broker before receiving the signal realization. I characterize the data broker's optimal selling mechanism, which involves screening over all possible signals. I allow the space of all signals the data broker can sell to be arbitrarily correlated with the signal the decision maker owns. This plays a key role in designing the optimal menu. In the binary-action setting, the data broker extracts the efficient surplus by offering a distinct binary signal for each type. Moreover, this result holds even when the broker does not know the prior distribution over states. In more general environments, I provide conditions on the payoff structure and the decision maker's type space under which the data broker extracts the efficient surplus. I discuss scenarios in which efficient-surplus extraction is not possible.

Keywords: Data, Information Design, Mechanism Design, Multidimensional Screening, Selling information

^{*}I am indebted to Teddy Mekonnen for his continuous guidance, comments, and feedback. I also want to thank Jack Fanning for his comments and suggestions that improved this paper. I am grateful to Geoffroy De Clippel, Klajdi Hoxha, Bobby Pakzad-Hurson, Roberto Serrano, Rajiv Vohra, and Rakesh Vohra for their helpful comments. Finally, I thank the participants at the Brown University Theory Seminar and the Stony Brook Game Theory Conference. All errors are my own.

[†]Brown University, Department of Economics, arlindo_skenderaj@brown.edu.

1 Introduction

Firms increasingly rely on third-party data—such as demographic information, behavioral patterns, or market trends—to better understand what decisions they should make. While many decision makers collect first-party data through direct interactions or internal systems, these data are often incomplete, leading them to acquire supplemental information from data brokers such as Acxiom, Experian, or Nielsen. Many producers purchase additional consumer data to create more detailed customer profiles, which are then used for consumer segmentation, targeted advertising, and various forms of price discrimination. For example, firms may personalize prices, send tailored discount coupons, or steer consumers toward higher-priced alternatives based on predicted preferences.

As a result, investing in consumer data—both internal and external—has become a central part of many firms’ efforts to increase revenue through better targeting. According to a report by the U.S. Federal Trade Commission (2014), data brokers offer “data append” services that allow businesses to supplement their customer records with missing or additional details, such as household income, family composition, or shopping behavior. These tools help firms refine their marketing strategies and better predict individual-level demand. More generally, decision makers with limited information often rely on experts to make better-informed choices. Examples include a lender deciding whether to approve or reject loan applications, a farmer selecting what to plant on land with different soil qualities, or an employer evaluating whether to hire a particular applicant.

This paper studies a data broker who sells information to a decision maker. The decision maker faces uncertainty about the state of the world and, consequently, about which action to take. She privately holds partial information in the form of a signal à la [Gentzkow and Kamenica \(2017\)](#); this signal defines her type. The decision maker acquires additional information from the data broker to improve her decision. The general problem faced by the data broker consists of screening over all signals, which can be challenging for two main reasons. First, this is a multidimensional screening problem. Second, the type space is not totally ordered due to both horizontal and vertical differentiation across signals. The value of any given signal depends on the signal the decision maker already possesses; a particular signal might be highly valuable to one type but entirely uninformative to another. Much of the existing mechanism design literature relies on ranked type spaces where the single-crossing property holds. However, in this setting, it is unclear whether single-crossing conditions hold in general, as the decision maker’s type changes as she acquires additional data, and

her payoff is derived from a utility-maximization problem.

I first analyze the problem with binary actions and provide an explicit construction of the optimal menu. This menu allows the data broker to extract the efficient surplus from the decision maker. The decision maker attains full information, and the broker extracts all information rents. The key feature of the optimal menu is the *minimal complementary signal*: the least valuable (in payoff order) signal that, when combined with the given signal, fully reveals the state. In the binary-action case, for any given signal π , it is without loss of generality for its minimal complementary signal $\hat{\pi}$ to have two messages. These messages can be interpreted as “don’t switch” when the interim action induced by π is optimal, and “switch” otherwise. The first main result states that the decision maker’s benefit from following her recommendation is at least as high as any other type’s benefit from following it. The result is intuitive and can be illustrated when there are two types of decision maker; informed and uninformed. By following the recommendation, the informed type always takes the correct action. The uninformed type, however, may not: obeying can sometimes overturn an optimal default, yielding no benefit. If she disobeys, the outcome is equivalent to first switching the default and then obeying. Since switching the default imposes a loss—and the benefit from obeying cannot exceed that of the informed type—disobedience is no better. Moreover, in the binary-action setting, the broker can extract the efficient surplus even without knowledge of the state prior.

Nevertheless, the efficient surplus extraction result does not extend to settings with many actions and states. In Section 4, I show that when some decision maker type fully distinguishes some state, her recommendation (the minimal complementary signal) contains messages she ignores in that state but that other types can benefit from. This prevents the data broker from designing a menu that extracts the efficient surplus, as in the binary-action case. More generally, efficient surplus extraction still holds in a large class of cases. Proposition 2 states sufficient conditions for efficient-surplus extraction in the general setting. The optimal menu sends (i) a “correction” message that reveals the state whenever the interim action is suboptimal, and (ii) a “confirmation” message when the interim action is optimal.

The choice of modeling information as a signal à la [Gentzkow and Kamenica \(2017\)](#) is crucial for this environment. Since the decision maker’s private type corresponds to owning a signal and she acquires an additional signal, it is important to model information in a way that has a well-defined algebraic structure—one that allows combining signals and analyzing their joint informational content. Several studies on the optimal sale of information

represent information either through Blackwell experiments (e.g., [Bergemann et al. \(2018\)](#)) or market segmentation (e.g., [Yang \(2022\)](#)). However, these approaches do not provide an algebraic structure over the set of all information structures and are therefore not suitable for the environment considered in this paper. In contrast, representing information as a signal allows for a well-defined operation of combining two signals and captures how observations from different information sources are correlated.

The structure of the paper is as follows: In Section 2, I provide related literature. In Section 3, I introduce the model, and in Section 4, I conduct the main analysis. The conclusion is presented in Section 5. Proofs are provided in the Appendix.

2 Related Literature

This paper relates to the literature on information design, and in particular to the optimal sale of information. Methodologically, it contributes to the literature on multidimensional mechanism design. Seminal work includes [Admati and Pfleiderer \(1986\)](#), [Admati and Pfleiderer \(1990\)](#), and [Allen \(1990\)](#), which examine how a data broker can optimally design information structures for imperfectly informed decision makers. For an overview of markets for information, see [Bergemann and Bonatti \(2019\)](#). Recent work, such as [Babaioff et al. \(2012\)](#); [Bergemann et al. \(2018\)](#); [Liu et al. \(2021\)](#); [Yang \(2022\)](#); [Segura-Rodriguez \(2022\)](#); [Hoxha \(2024\)](#), study environments in which a data broker sells information to an agent facing a decision problem under uncertainty. In [Babaioff et al. \(2012\)](#), the data broker sells information to a decision maker who has a private payoff function. The information is sold after the data broker observes the realized state of the world. In [Liu et al. \(2021\)](#), the data broker sells information to a decision maker who takes passive and active action. In their case, the payoff depends on the state only under the active action, reducing the problem to one dimension. [Yang \(2022\)](#) studies the optimal selling mechanism of a data broker who sells a signal to a monopolist with a privately known constant marginal cost. [Segura-Rodriguez \(2022\)](#) studies a data broker who sells information about consumers to a monopolist who wants to predict one characteristic; which characteristic she wants to predict and how much she values the information are both private. [Hoxha \(2024\)](#) considers an environment in which the data broker sells information at the interim stage to a decision maker facing correlated states. This paper is closest to [Bergemann et al. \(2018\)](#), who analyze a setting in which the broker sells Blackwell experiments to a decision maker whose private information is an interim belief about the true state; the signal is used to improve decision choices across states. In contrast, I study an environment in which the decision maker’s private informa-

tion is the signal she owns, which represents a distribution over her interim beliefs. This captures situations where the decision maker must make frequent decisions and purchases information before the interim belief is realized. The key difference is that the monopolist's private information is itself a signal structure, and the analysis focuses on the ex-ante value of refining that information.

On the literature of persuasion without transfers, [Kolotilin et al. \(2017\)](#), [Kolotilin \(2018\)](#), [Krähmer \(2020\)](#), [Candogan and Strack \(2021\)](#), [Yamashita and Zhu \(2021\)](#), [Ichihashi and Smolin \(2022\)](#), [Zhu \(2023\)](#) analyze optimal mechanisms without transfers for providing information to privately informed agents. Other studies have analyzed information transmission from experts to customers (see [Eső and Szentes \(2007b\)](#), [Milgrom and Roberts \(1986\)](#), [Pitchik and Schotter \(1987\)](#)). [Eső and Szentes \(2007a\)](#) analyze how a seller should reveal information to bidders in an auction and show that a handicap auction can implement the revenue-maximizing information disclosure policy. [Bergemann and Bonatti \(2015\)](#) studies how a data broker sells consumer tracking information (cookies) to advertisers, and how the sale of that information affects advertising strategies and market outcomes. In their paper, the price of the information was determined in a competitive market, and the data buyer did not have any private information. [Bergemann et al. \(2015\)](#) analyze the welfare outcomes that a designer can implement by offering any signal structure. They show that when the designer knows the monopolist's type, any division of the total surplus between the buyer and the seller can be achieved, provided that the total surplus does not exceed the efficient surplus, the monopolist's profit is at least as much as the profit without information, and the buyer's surplus is non-negative. [Krähmer \(2020\)](#) studies mechanism design problems with quasi-linear utility, where the principal can design and disclose additional information that affects agents' valuations. They show that the principal can design information and a mechanism to fully extract the efficient surplus in a large class of cases. However, in these papers the information receiver does not face a decision problem under uncertainty.

Finally, this paper relates to the literature on quality degradation (see [Mussa and Rosen \(1978\)](#) and [Maskin and Riley \(1984\)](#)). [Mussa and Rosen \(1978\)](#) analyze how a monopolist can use product quality differentiation to maximize profits by setting different qualities and prices for different consumer segments. [Maskin and Riley \(1984\)](#) study how a monopolist can design optimal incentive schemes under asymmetric information by offering quantity discounts tailored to consumers with private information about their preferences. In the current paper, the data seller's optimal mechanism may involve degrading the quality of information for some buyers to charge higher prices to others.

3 Model

A decision maker (*she*) faces uncertainty about the state of the world ω , which is drawn from a finite set $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$. The decision maker chooses an action $a \in A$, where $A = \{a_1, a_2, \dots, a_m\}$. The payoff function given the state and the action is

$$u : A \times \Omega \rightarrow \mathbb{R}.$$

Assume that in each state there is a unique optimal action. That is,

$$\forall \omega \in \Omega, \exists a^*(\omega) \in A, u(a^*(\omega), \omega) > u(a, \omega), \forall a \in A, a \neq a^*(\omega).$$

The payoff function is represented by the following matrix.

$u(a_i, \omega_j)$	ω_1	ω_2	$\dots$	ω_n
a_1	u_{11}	u_{12}	$\dots$	u_{1n}
a_2	u_{21}	u_{22}	$\dots$	u_{2n}
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
a_m	u_{m1}	u_{m2}	$\dots$	u_{mn}

Let $\mu_0 \in \Delta\Omega$ be the common prior on Ω . For any belief $\mu \in \Delta(\Omega)$, the decision maker takes an optimal action $a(\mu)$ that solves

$$\max_{a \in A} \sum_{\omega \in \Omega} \mu(\omega) u(a, \omega).$$

3.1 Information

Besides the prior, the decision maker privately holds additional information that enables her to make better decisions. I model this information as a signal π , following [Gentzkow and Kamenica \(2017\)](#): a *signal* is a finite partition of the extended state space $\Omega \times X$, where X is a random variable, independent of ω and uniformly distributed on $[0, 1]$. Let S denote the collection of nonempty Lebesgue-measurable subsets of $\Omega \times X$, and write $\pi \subset S$. An element $s \in \pi$ is called a *signal realization* (or *message*). Define

$$\mu(s \mid \omega) = \lambda\{x \mid (\omega, x) \in s\} \quad \text{and} \quad \mu(s) = \sum_{\omega \in \Omega} \mu(s \mid \omega) \mu(\omega),$$

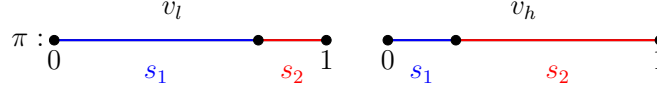


Figure 1: A signal.

where $\lambda(\cdot)$ denotes Lebesgue measure.

For any $s \in \pi$, $\mu(s \mid \omega)$ denotes the conditional probability of receiving s in state ω , and $\mu(s)$ denotes the unconditional probability of receiving s . Let Π denote the set of all signals. The definition is illustrated in Figure 1. Modeling information as in [Gentzkow and Kamenica \(2017\)](#) provides an algebraic structure over Π that allows us to “add” signals and analyze their joint information content.¹ It also captures how observations from one source correlate with observations from others.²

3.1.1 Value of supplemental information

Let $\pi \in \Pi$ and $s \in \pi$. Upon receiving a private message s , the decision maker chooses the action that maximizes expected payoff. Her *interim* payoff conditional on a message s is

$$u(s) = \max_{a \in A} \sum_{\omega \in \Omega} \mu(\omega \mid s) u(a, \omega).$$

Given any signal π , the decision maker takes the best action for each message. Her *ex-ante* payoff from π is

$$U(\pi) = \sum_{s \in \pi} \mu(s) u(s).$$

Denote by $\bar{\pi}$ the signal that fully reveals each state. The utility from the fully revealing signal is

$$\bar{U} = \sum_{\omega \in \Omega} \mu(\omega) u(a^*(\omega) \mid \omega).$$

Next, given any two signals $\pi, \pi' \in \Pi$, define their *join*, denoted by $\pi \vee \pi'$, by

$$\pi \vee \pi' = \{s \cap s' \mid s \in \pi, s' \in \pi'\}.$$

¹For more on the comparison of signals [Brooks et al. \(2024\)](#).

²Several studies on selling information represent information as Blackwell experiments [Bergemann et al. \(2018\)](#) or as market segmentations [Yang \(2022\)](#). These representations lack a convenient algebraic structure over the space of information structures and are therefore less suitable for the environment of this paper.

The expected payoff from the join of any two signals is

$$U(\pi \vee \pi') = \sum_{s \in \pi} \sum_{s' \in \pi'} \mu(s \cap s') u(s \cap s').$$

Finally, the net value from adding π' to π is

$$V(\pi' | \pi) = U(\pi' \vee \pi) - U(\pi).$$

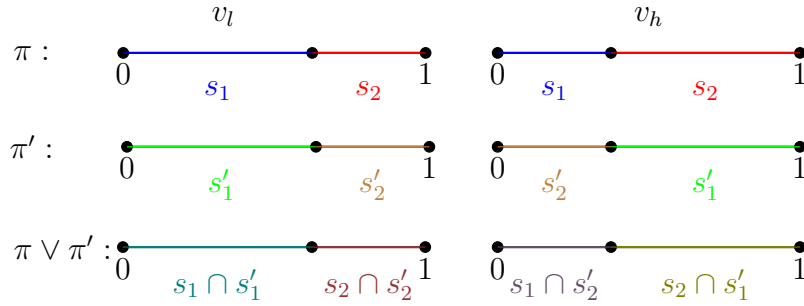


Figure 2: The join of two signals.

3.1.2 Minimal complementary signal.

Having defined the information, I will now describe how data broker induces full information to the decision maker with as little additional information as possible. For any signal π , I look for a minimal complementary signal π^c such that $\pi \vee \pi^c$ is fully revealing, and no strictly less valuable signal (in payoff order) has this property. Understanding the structure of the minimal complementary signal is essential: later I show that is the key to characterizing the optimal menu.

Definition 1. Given a signal π , the **minimal complementary signal** of π , denoted by π^c , is defined as the solution to

$$\min_{\pi'} U(\pi') \quad \text{subject to} \quad U(\pi \vee \pi') = \bar{U}.$$

Lemma 1. Let $|A| = 2$. Let $\pi \in \Pi$ be such that it does not perfectly reveal all the states. Then, the minimal complementary signal π^c is binary.

Proof (Appendix 6.1). In the two-action case, the minimal complementary signal has a simple interpretation: its messages can be interpreted as “don’t switch” when the interim action induced by π is optimal, and “switch” otherwise. Although the decision maker’s

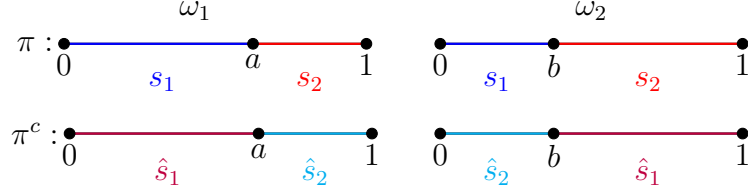


Figure 3: Minimal complementary signal

existing signal may contain many messages, with only two actions the decision maker either selects the correct action or the incorrect one after any message. The complementary signal therefore functions as a recommendation indicating whether to switch from the action chosen under the existing signal.

3.2 Data broker's problem

The data broker screens the decision maker based on the signal she already possesses. Let $\mathcal{T} \subseteq \Pi$ denote the decision maker's type space, with types distributed according to F on $\mathcal{T}$. By the revelation principle the data broker can restrict attention to direct mechanism which offers a signal $\sigma(\pi)$ at a price $t(\pi)$ for every $\pi \in \mathcal{T}$. Timing is as follows:

1. The data broker posts a menu $\mathcal{M} = \{(\sigma(\pi), t(\pi))\}_{\pi \in \mathcal{T}}$ of signal–price pairs before the true state is realized.
2. A decision maker of type π purchases a signal $\sigma(\pi)$ and pays $t(\pi)$, before her private message is realized.
3. The decision maker then holds $\pi \vee \sigma(\pi)$ and, after the true state and her private message are realized, takes an action that maximizes her expected payoff.

The data broker chooses a menu to maximize expected revenue

$$\int_{\mathcal{T}} t(\pi) dF(\pi)$$

subject to incentive compatibility constraints

$$V(\sigma(\pi) \mid \pi) - t(\pi) \geq V(\sigma(\pi') \mid \pi) - t(\pi'), \quad \forall \pi, \pi' \in \mathcal{T}, \quad (\text{IC})$$

and participation constraints

$$V(\sigma(\pi) \mid \pi) - t(\pi) \geq 0, \quad \forall \pi \in \mathcal{T}. \quad (\text{IR})$$

3.3 Illustrative examples

3.3.1 Monopolist buying consumer data.

A monopolist sells concert tickets to two consumer groups: students and professionals, with a willingness to pay of \$50 and \$100, respectively. The aggregate market consists of 30% students and 70% professionals. Suppose the monopolist owns a residential dataset that indicates whether a consumer lives in a house or an apartment. This information allows the monopolist to segment the market into two groups: one consisting mostly of professionals and one that includes a mix of both types. Specifically, 65% of the market lives in houses, of which 14% are students and 86% are professionals; the remaining 35% live in apartments, with 60% students and 40% professionals. Using this segmentation, the monopolist achieves a profit of

$$\frac{65}{100} \cdot 86 + \frac{35}{100} \cdot 50 = 73.4.$$

In contrast, if the monopolist had no access to any dataset, she would set a uniform price of \$100 and sell only to professionals, resulting in a profit of \$70. If the monopolist had full information, she would be able to perfectly price discriminate and extract the entire surplus from the market, yielding a profit of

$$50 \cdot 0.3 + 100 \cdot 0.7 = 85.$$

Now suppose a data broker can sell additional information to the monopolist. The broker does not know whether the monopolist already owns the residential dataset but believes she does with probability $\frac{1}{2}$. To maximize revenue, the broker offers a menu of datasets:

$$\mathcal{M} = \{(\text{fully-revealing data}, \$15), (\text{civil-status data}, \$11.6)\}.$$

Assume that only students who live in houses and professionals who live in apartments are married. Under this assumption, the civil-status dataset complements the residential dataset in such a way that their combination reveals the consumer type perfectly.

A monopolist with no prior data purchases the fully-revealing dataset and pays \$15, as she is indifferent between buying this option or acquiring partial information at a lower price. A monopolist who already owns the residential dataset purchases the civil-status data instead, because together they yield full segmentation, and her willingness to pay for this supplemental information is exactly \$11.6. In this way, the broker successfully screens both types and extracts the full information value from each.

3.3.2 Clinical diagnosis of a patient.

A physician faces three conditions $\Omega = \{\omega_1, \omega_2, \omega_3\}$ (bacterial pneumonia, viral pneumonia, lung cancer), each with prior probability $1/3$. There are three treatments $A = \{a_1, a_2, a_3\}$ (antibiotics, antivirals, chemotherapy). Payoffs are diagonal:

$$u(a_i, \omega_j) = \begin{cases} 10, & i = j, \\ 0, & i \neq j. \end{cases}$$

A bacterial culture test perfectly identifies bacterial pneumonia, whereas imaging does not distinguish viral pneumonia from early-stage lung cancer (both can produce similar opacities). The resulting partition is $\{\{\omega_1\}, \{\omega_2, \omega_3\}\}$. When the culture is positive, the physician prescribes a_1 (antibiotics). When imaging is ambiguous, she chooses one treatment—say a_2 (antivirals)—which is correct in ω_2 and incorrect in ω_3 . The resulting expected utility is $10/3 + 10/3 + 0 = 20/3$. Without any additional information, prescribing a_1 yields $10/3$.

Suppose a private laboratory can sell any diagnostic test but is uncertain whether the physician can run the culture test in house. The lab offers a menu of recommendation signals: conditional on the physician’s tentative treatment, the signal either (i) confirms it is correct or (ii) indicates it is incorrect and reveals the true state. Price the “informed-type” recommendation at $10/3$ and the “uninformed-type” recommendation at $20/3$.

Incentives are as follows. For the informed type (with partition $\{\{\omega_1\}, \{\omega_2, \omega_3\}\}$), the recommendation delivers full information and raises expected utility from $20/3$ to 10 , so the incremental value is $10/3$; paying $10/3$ leaves zero rent and imitation of the uninformed contract (priced at $20/3$) is unprofitable. For the uninformed type (who would otherwise prescribe a_1), the recommendation also yields 10 , i.e., an incremental value of $20/3$; paying $20/3$ leaves zero rent. Moreover, the uninformed type does not gain by imitating the informed contract: that signal corrects only the mistakes defined relative to the informed type’s baseline (it reveals in ω_3 but not in ω_2), yielding at most the same incremental value $10/3$, which is below the uninformed price.

When can the data broker extract the efficient surplus? With two actions, efficient surplus extraction is feasible for any belief about the decision maker’s private information. This result does not generally extend to environments with multiple actions and states. For those settings, I provide sufficient conditions under which efficient surplus extraction is achievable.

4 Analysis

In the first part of this section, I analyze the binary-action case and show that the data broker extracts the efficient surplus. In the remainder, I analyze the general case with many actions and states, where the optimal menu does not achieve full surplus extraction in general. In this case, I provide sufficient conditions under which efficient surplus extraction is possible, and discuss the scenarios when this is not possible. The *efficient surplus extraction* is defined as follows.

Definition 2. *Data broker extracts the efficient surplus from the decision maker if there exists a feasible menu $\mathcal{M}^* = \{(\sigma^*(\pi), t^*(\pi))\}_{\pi \in \mathcal{T}}$ such that $\sigma^*(\pi) \vee \pi = \bar{\pi}$ and $t^*(\pi) = \bar{U} - U(\pi)$ for any $\pi \in \mathcal{T}$.*

4.1 Binary actions

I first analyze the case with two actions, i.e., $A = \{a_1, a_2\}$. The payoff function is

$u(a, \omega)$	ω_1	ω_2	$\cdots$	ω_n
a_1	u_{11}	u_{12}	$\cdots$	u_{1n}
a_2	u_{21}	u_{22}	$\cdots$	u_{2n}

For $\omega \in \Omega$, write $u_1(\omega) \equiv u(a_1, \omega)$ and $u_2(\omega) \equiv u(a_2, \omega)$. Partition the states into

$$\Omega_1 = \{\omega \in \Omega : u_1(\omega) \geq u_2(\omega)\} \quad \text{and} \quad \Omega_2 = \{\omega \in \Omega : u_2(\omega) > u_1(\omega)\}.$$

This aggregation reduces the problem to two regions in which each action is uniquely optimal, so the analysis follows the same logic as the binary state and binary action case.

Lemma 2. *Fix $\pi \in \Pi$ and let π^c be a minimal complementary signal to π (so $\pi \vee \pi^c$ is fully revealing and no strictly less valuable signal has this property). If $|A| = 2$, then for all $\pi' \in \Pi$,*

$$V(\pi^c \mid \pi) \geq V(\pi^c \mid \pi').$$

That is, the incremental value of π^c is weakly maximized at its base signal π .

The proof is provided in Appendix 6.2. The lemma implies a key property of binary-action environments: a minimal complementary signal is (weakly) most valuable to the type for which it is designed. Below, I will provide the key intuition for this result using the following example. In the next section, this need not hold when there are more than three states.

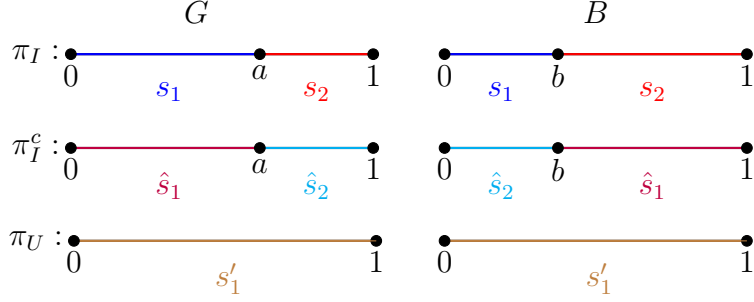


Figure 4

4.1.1 Intuition behind Lemma 6.2

Consider two states, good and bad, i.e., $\Omega = \{G, B\}$, and two actions $A = \{g, b\}$. Without loss of generality, normalize $u_{12} = u_{21} = 0$, $u_{22} = 1$, and $u_{11} = x > 0$. The payoffs are given by the following matrix

	g	b
G	x	0
B	0	1

Consider two types of decision makers:

1. *Uninformed (U)*. Only knows the prior $\mu_0 \in \Delta\Omega$.
2. *Partially informed (I)*. This type owns some signal π .

Assume the following

1. b is the default action, i.e., $\mu(B) \geq \mu(G)x$.
2. $\pi = \{s_1, s_2\}$ where s_1 induces action g , s_2 induces action b
3. $U(\pi) > \mu(B) \cdot 1$.

Let $\hat{\pi} = \{\hat{s}_1, \hat{s}_2\}$ denote the minimal complementary signal to π . Assume that $\hat{\pi}$ is such that $\hat{s}_2$ is realized whenever the message sent by π induces a suboptimal action, and $\hat{s}_1$ is realized whenever the message sent by π induces the optimal action. Interpret $\hat{s}_1$ as “keep the action induced by π ” and $\hat{s}_2$ as “switch the action.” The signals are illustrated in Figure 4.

The informed decision maker’s benefit from $\hat{\pi}$ is

$$V_I(\hat{\pi}) = \mu(G \cap s_2) \cdot x + \mu(B \cap s_1) \cdot 1.$$

The uninformed DM can imitate the informed DM by choosing to either obey or disobey to messages in $\hat{\pi}$. Her benefit from obeying $\hat{\pi}$.

$$V_U(\hat{\pi}) = \mu(G \cap s_2) x - \mu(B \cap s_1).$$

This payoff cannot exceed that of the informed decision maker. When receiving the “do not switch” message, neither type gains any benefit. When receiving the “switch” message, the informed type always improves her action, while the uninformed type takes the correct action with some probability and the wrong one with the remaining probability.

To complete the argument consider the benefit uninformed decision maker gets from disobeying the messages in $\hat{\pi}$.

$$\begin{aligned} V_U(\hat{\pi}) &= \mu(G \cap s_1) x - \mu(B \cap s_2) \\ &= [\mu(G) x - \mu(B)] + [\mu(B \cap s_1) - \mu(G \cap s_2) x] \\ &\leq \mu(B \cap s_1). \end{aligned}$$

Disobeying $\hat{\pi}$ is equivalent to first switching the default action from b to g and then following the messages of $\hat{\pi}$. The initial switch leads to a loss, since before receiving any information, taking action b yields a higher expected payoff. Moreover, as shown above, the uninformed decision maker’s gain from obeying $\hat{\pi}$ is no greater than that of the informed decision maker.

The following proposition presents the main result for binary states when the decision maker’s privately owned information is an arbitrary signal.

Proposition 1. *Assume $\Omega = \{\omega_1, \dots, \omega_n\}$, $A = \{a_1, a_2\}$, $\mathcal{T} = \Pi$ and $F \sim \mathcal{T}$. The data broker extracts the efficient surplus from the decision maker. In particular, the optimal menu is given by*

$$\mathcal{M} = \{(\pi^c, t(\pi))\}_{\pi \in \Pi},$$

where

$$t(\pi) = U(\pi \vee \pi^c) - U(\pi).$$

The proof is in Appendix 6.3. In the binary-action case the broker need not know the prior over states. The broker can elicit the action the decision maker would take under her private signal and then supply a minimal complementary signal instructing her to switch only when that action is suboptimal. However, as the example in the next section shows, this property fails with more than two actions or more than two states. Fully characterizing the multi-state, multi-action case is challenging; in the next section, I provide sufficient conditions on

the payoff structure and the type space under which the broker can still extract the efficient surplus.

4.2 Many states and many actions

In this part assume that $|A| = |\Omega| = n \geq 3$. This section focuses on a class of payoff functions where $u(\omega, a)$ admits the following diagonal representation:

$u(a_i, \omega_j)$	ω_1	ω_2	$\cdots$	ω_n
a_1	u_1	0	$\cdots$	0
a_2	0	u_2	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
a_n	0	0	$\cdots$	u_n

In other words,

$$u(\omega_i, a_j) = \begin{cases} u_i & \text{if } j = i, \\ 0, & \text{if } j \neq i. \end{cases}$$

Let $u(\omega)$ to denote the payoff obtained from taking the optimal action in state ω i.e.,

$$u(\omega) = \arg \max_{a \in A} \sum_{i=1}^m u(a, \omega).$$

I focus on a type space in which each type corresponds to a simple signal.

Definition 3. A signal π is called simple if it can be represented as a partition of the state space Ω . Let $\tilde{\Pi}(\Omega)$ denote the set of all partitions of Ω .

A signal π is a partition of the state Ω , and each element $s \in \pi$ is an information set. For any $s \in \pi$, there exists ω_s such that

$$u(s) = \mu(\omega_s \mid s) u(\omega_s).$$

Lemma 3. Let $\pi \in \tilde{\Pi}(\Omega)$. The payoff from π is

$$U(\pi) = \sum_{s \in \pi} \mu(\omega_s) u(\omega_s).$$

Proof. Take any $s \in \pi$ and notice that

$$\mu(\omega \mid s) = \begin{cases} \frac{\mu(\omega)}{\mu(s)}, & \text{if } \omega \in s, \\ 0, & \text{otherwise.} \end{cases}$$

Let a_s be an optimal action conditional on receiving message s . That is, a_s solves

$$\max_{a \in A} \sum_{\omega \in \Omega} \mu(\omega \mid s) u(a, \omega).$$

Since the payoff function is the diagonal matrix, we have

$$u(s) = \mu(\omega_s \mid s) u(\omega_s) \quad \text{for some } \omega_s \in s.$$

The payoff from signal π is

$$U(\pi) = \sum_{s \in \pi} \mu(s) \mu(\omega_s \mid s) u(\omega_s) = \sum_{s \in \pi} \mu(\omega_s) u(\omega_s).$$

□

Lemma 4. Assume $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$, $A = \{a_1, a_2, \dots, a_n\}$, and $\mathcal{T} = \{\pi, \pi'\}$ for some $\pi, \pi' \in \tilde{\Pi}(\Omega)$. The data broker extracts the efficient surplus from the decision maker.

The formal proof is in Appendix 6.4. Below, I will provide the intuition through an example where the decision maker has two types.

4.2.1 Intuition behind Lemma 4

Consider two possible types for the decision maker:

1. *Uninformed (U)*: takes the default action a^* . Let ω^* be the state in which she takes the correct action, i.e.,

$$\mu(\omega^*)u(\omega^*) \geq \mu(\omega)u(\omega) \quad \forall \omega \in \Omega.$$

2. *Partly informed (I)*: owns a signal π . For any $s \in \pi$ and let a_s be an optimal action conditional on s . Moreover, let ω_s be such that $\mu(\omega_s \mid s) u(a_s, \omega_s) = u(s)$. Thus, for each $s \in \pi$, ω_s is the only state in s in which the decision maker takes the correct action. Let S be the collection of all states in which she takes the correct action. Since

the payoff from taking the incorrect action is always zero, the partly informed decision maker's willingness to pay for full information is

$$V_I = \sum_{\omega \in \Omega \setminus S} \mu(\omega) u(\omega).$$

Now consider selling the informed decision maker a signal $\hat{\pi}$ that fully reveals all states in which the informed decision maker is taking the wrong action and sends a common message, call it $m \equiv (\text{don't change})$, whenever the realized state is one of the ones in which the decision maker takes the right action. Note that, if the uninformed type mimics this signal, it must obey whenever the state is revealed. If $\omega^* \in S$, then she will obey message m^* to not change her default action a^* . This follows since a^* remains optimal in every signal that contains ω^* . In this case her benefit from $\hat{\pi}$ is

$$V_U = \sum_{\omega \in \Omega \setminus S} \mu(\omega) u(\omega).$$

If $\omega^* \in \Omega \setminus S$, then she will disobey message m^* and switch some other action a' , improving her action in some state $\omega' \in S$. Note that since a^* is the default action given the prior, it holds that

$$\mu(\omega^*) u(\omega^*) \geq \mu(\omega) u(\omega) \quad \forall \omega \in \Omega.$$

The uninformed type will improve her action after using $\hat{\pi}$ whenever the state is either $\omega' \in S$ or $\omega \in \Omega \setminus S \setminus \{\omega^*\}$. The uninformed type's benefit from $\hat{\pi}$ will therefore be

$$V_U = \sum_{\omega \in \Omega \setminus S} \mu(\omega) u(\omega) - \mu(\omega^*) u(\omega^*) + \mu(\omega') u(\omega') \leq \sum_{\omega \in \Omega \setminus S} \mu(\omega) u(\omega) \leq V_I.$$

Consider the following menu. First, offer the informed type the signal $\hat{\pi}$ at price V_I . By the argument above, the uninformed type would obtain at most V_I from mimicking this option, so she has no incentive to do so. Second, offer the uninformed type full information at a price equal to her willingness to pay for full information, which is strictly higher than V_I . Hence, the informed type will not mimic the uninformed option. Therefore, both incentive constraints hold, each type attains full information, and all rents are extracted.

In the following proposition, I show that the data broker can extract the efficient surplus from the decision maker when the type space is $\tilde{\Pi}(\Omega)$.

Proposition 2. *Assume $\Omega = \{\omega_1, \omega_2, \dots, \omega_n\}$ and $A = \{a_1, a_2, \dots, a_n\}$. Let the decision maker's type be distributed according to some F over $\tilde{\Pi}(\Omega)$. The data broker extracts the*

efficient surplus from the decision maker.

The proof is in appendix 6.5. The diagonal (state–action matching) is crucial for the full surplus extraction. With diagonal payoffs, any action–state mismatch yields the same baseline payoff. The marginal value of being corrected does not depend on the action the decision maker is taking; it depends only on the state. Under general payoffs, by contrast, the marginal value of being corrected typically depends on the action the decision maker would otherwise take. The example below shows that, in this case, efficient surplus extraction need not hold.

4.2.2 Example

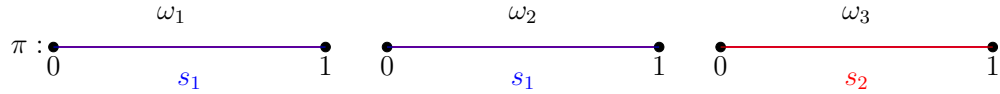
The following example shows that the data broker may not be able to extract the efficient surplus when there are more than two states. Consider the following payoff function represented by lower triangular matrix:

	ω_1	ω_2	ω_3
a_1	u_1	u_1	u_1
a_2	0	u_2	u_2
a_3	0	0	u_3

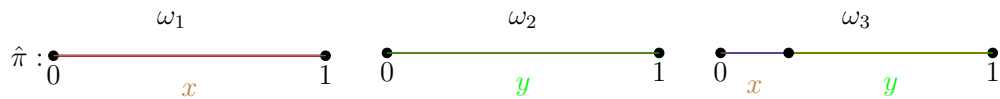
Assume that $u_1 < u_2 < u_3$ and that the prior $\mu = (\mu_1, \mu_2, \mu_3)$ is such that the decision maker is indifferent between taking any action $a \in A$. That is

$$u_1 = (\mu_2 + \mu_3) u_2 = \mu_3 u_3.$$

Moreover, assume $\mu_i > 0$ for $i = 1, 2, 3$. Consider two types for the decision maker i) fully uninformed, and ii) partially informed that owns a signal that only fully reveals when the state is ω_3 .



The π –minimal complementary signal is given by



Note that the minimal complementary signal should minimize $U(\hat{\pi})$ subject to $\hat{\pi} \vee \pi$ being the full informative signal. That is, the partition of $\omega_3, [0, 1]$ should be such that

$$\min_{a \in [0,1]} \left(\max\{(\mu_1 + a\mu_3)u_1, a\mu_3u_3\} + \max\{(\mu_2 + (1-a)\mu_3)u_2, (1-a)\mu_3u_3\} \right)$$

Note that having $(\mu_1 + a\mu_3)u_1 < a\mu_3u_3$ is not optimal. Increasing a by some $\varepsilon > 0$ will increase the first term by $\varepsilon\mu_3u_1$, but this will decrease the second term by at least $a\mu_3u_2$. So, $(\mu_1 + a\mu_3)u_1 \geq a\mu_3u_3$. On the other hand, it is (weakly) optimal to have $(\mu_1 + a\mu_3)u_1 = a\mu_3u_3$. If $(\mu_1 + a\mu_3)u_1 > a\mu_3u_3$, then decreasing a will decrease the first term by $a\mu_3u_3$ and increase the second term by at most $a\mu_3u_3$. So, one optimal solution is setting a such that $(\mu_1 + a\mu_3)u_1 = a\mu_3u_3$, i.e.,

$$a = \frac{\mu_1 u_1}{\mu_3 (u_3 - u_1)} \in (0, 1),$$

since $u_1(\mu_1 + \mu_3) < \mu_3u_3$. The DM is indifferent between taking actions a_1 or a_3 after receiving signal realization x . Low type's benefit from $\hat{\pi}$ is

$$V(\hat{\pi} \mid \pi) = \mu(\omega_2) (u_2 - u_1).$$

High type's benefit from $\hat{\pi}$ is

$$V(\hat{\pi} \mid \pi') = \mu(\omega_2) (u_2 - u_1) + \mu(\omega_3, y) (u_3 - u_1) > V(\hat{\pi} \mid \pi).$$

since $\mu(\omega_3, y) > 0$. Because $\hat{\pi}$ is the minimal signal that induces full information for the low type while yielding a higher benefit to the high type, the data broker cannot extract the efficient surplus. This happens because the high type already knows the state ω_3 and therefore ignores the messages that $\hat{\pi}$ provides in that state. Still, while the informed decision maker ignores signal y in state ω_3 , the uninformed type benefits from receiving it.

5 Conclusion

This paper studies how a data broker sells information, in the form of a signal, to a decision maker who privately owns a different signal. Unlike prior work on selling supplemental information, here the decision maker's private type is itself a signal structure—that is, a distribution over her interim beliefs. The analysis focuses on environments in which the decision maker purchases information before her private signal is realized, implying that the

data broker must screen over all signals. I allow the signals the data broker can sell to contain messages that are arbitrarily correlated with the ones the decision maker privately owns. Considering such a rich signal space plays a key role in the design of the optimal menu. It also aligns with practical settings, where a supplemental dataset is appended to an existing one.

This paper shows that, in a large class of cases, the data broker can perfectly induce full information for the decision maker and extract all information rents. Remarkably, this outcome can also be achieved even when the data broker does not know the prior distribution over the state of the world.

By modeling private information as a signal, the paper moves toward a general framework for selling information to informed agents. At the same time, screening over signal structures becomes substantially more complex as the number of actions and states grows. Characterizing optimal mechanisms in these richer environments remains an important extension.

References

- Admati, A. R. and P. Pfleiderer (1986). A monopolistic market for information. *Journal of Economic Theory* 39(2), 400–438.
- Admati, A. R. and P. Pfleiderer (1990). Direct and indirect sale of information. *Econometrica* 58(4), 901–928.
- Allen, F. (1990). The market for information and the origin of financial intermediation. *Journal of Financial Intermediation* 1(1), 3–30.
- Babaioff, M., R. Kleinberg, and R. P. Leme (2012). Optimal mechanisms for selling information. In *Proceedings of the 13th ACM Conference on Electronic Commerce*, New York, NY, USA, pp. 92–109. Association for Computing Machinery.
- Bergemann, D. and A. Bonatti (2015). Selling cookies. *American Economic Journal: Microeconomics* 7(3), 259–94.
- Bergemann, D. and A. Bonatti (2019). Markets for information: An introduction. *Annual Review of Economics* 11, 85–107.
- Bergemann, D., A. Bonatti, and A. Smolin (2018). The design and price of information. *American Economic Review* 108(1), 1–48.
- Bergemann, D., B. Brooks, and S. Morris (2015). The limits of price discrimination. *American Economic Review* 105(3), 921–57.
- Brooks, B., A. Frankel, and E. Kamenica (2024). Comparisons of signals. *American Economic Review* 114(9), 2981–3006.
- Candogan, O. and P. Strack (2021). Optimal disclosure of information to privately informed agents. Available at SSRN: <https://ssrn.com/abstract=3773326>.
- Eső, P. and B. Szentes (2007a). Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies* 74(3), 705–731.
- Eső, P. and B. Szentes (2007b). The price of advice. *The RAND Journal of Economics* 38(4), 853–870.
- Gentzkow, M. and E. Kamenica (2017). Bayesian persuasion with multiple senders and rich signal spaces. *Games and Economic Behavior* 104, 411–429.
- Hoxha, K. (2024). Selling correlated information products. *arXiv preprint arXiv:2405.11142*.

- Ichihashi, S. and A. Smolin (2022). Data provision to an informed seller. Available at SSRN: <https://ssrn.com/abstract=4085705>.
- Kolotilin, A. (2018). Optimal information disclosure: A linear programming approach. *Theoretical Economics* 13(2), 607–635.
- Kolotilin, A., T. Mylovanov, A. Zapechelnnyuk, and M. Li (2017). Persuasion of a privately informed receiver. *Econometrica* 85(6), 1949–1964.
- Krähmer, D. (2020). Information disclosure and full surplus extraction in mechanism design. *Journal of Economic Theory* 187, 105020.
- Liu, S., W. Shen, and H. Xu (2021). Optimal pricing of information. *arXiv preprint arXiv:2102.13289*.
- Maskin, E. and J. Riley (1984). Monopoly with incomplete information. *The RAND Journal of Economics* 15(2), 171–196.
- Milgrom, P. and J. Roberts (1986). Relying on the information of interested parties. *The RAND Journal of Economics* 17(1), 18–32.
- Mussa, M. L. and S. Rosen (1978). Monopoly and product quality. *Journal of Economic Theory* 18, 301–317.
- Pitchik, C. and A. Schotter (1987). Honesty in a model of strategic information transmission. *The American Economic Review* 77(5), 1032–1036.
- Segura-Rodriguez, C. (2022). Selling data. PIER Working Paper No. 19-006, Available at SSRN: <https://ssrn.com/abstract=3385500>.
- Yamashita, T. and S. Zhu (2021). Type-contingent information disclosure. Technical Report 21-1242, Toulouse School of Economics (TSE).
- Yang, K. H. (2022). Selling consumer data for profit: Optimal market-segmentation design and its consequences. *American Economic Review* 112(4), 1364–93.
- Zhu, S. (2023). Private disclosure with multiple agents. *Journal of Economic Theory* 212, 105705.

6 Appendix

6.1 Proof of Lemma 1

Proof. Let $\pi = s_1, s_2, \dots, s_k$ for some $k \in \mathbb{N}$. Define $\tilde{s}_1$ as the collection of all messages that induce action a_1 , and $\tilde{s}_2$ as the collection of all messages that induce action a_2 . Since π does not fully reveal the true state, assume without loss of generality that $\mu(\tilde{s}_2 \mid \omega_1) > 0$. That is, the decision maker takes the wrong action in state ω_1 with some positive probability.

First, the minimal complementary signal must contain at least two messages to identify the true state when the signal realization is in $\tilde{s}_2$. Let these messages be $\hat{s}_1$ and $\hat{s}_2$ with the following property: $\mu(\hat{s}_2 \mid \tilde{s}_2, \omega_1) = 1$ and $\mu(\hat{s}_1 \mid \tilde{s}_2, \omega_2) = 1$. This implies that $\mu(\omega_1 \mid \tilde{s}_2 \cap \hat{s}_1) = 1$ and $\mu(\omega_2 \mid \tilde{s}_2 \cap \hat{s}_2) = 1$. That is, $\hat{s}_1$ and $\hat{s}_2$ fully reveal the true state conditional on $\tilde{s}_2$. Hence, the minimal complementary signal must have at least two messages.

Second, it is without loss to restrict attention to minimal complementary signals with at most two messages. If $\mu(\tilde{s}_1 \mid \omega_2) > 0$, then to identify the true state conditional on $\tilde{s}_2$, two messages s'_1 and s'_2 with $\mu(s'_1 \mid \tilde{s}_1, \omega_1) = 1$ and $\mu(s'_2 \mid \tilde{s}_1, \omega_2) = 1$ are needed. Note that, it is weakly less informative to merge s'_1 with $\hat{s}_1$ as one signal, and similarly merge s'_2 with $\hat{s}_2$ as another signal.

□

6.2 Proof of Lemma 2

Proof. Take any $\pi, \pi' \in \Pi$ and consider the following collections of signals.

$$s_1 = \{s \in \pi \mid E[u(a_1, \omega) \mid s] \geq E[u(a_2, \omega) \mid s]\},$$

and

$$s_2 = \{s \in \pi \mid E[u(a_2, \omega) \mid s] > E[u(a_1, \omega) \mid s]\}.$$

Let $\hat{\pi} = \{\hat{s}_1, \hat{s}_2\}$ be the minimal complementary signal of π . Without loss of generality, assume that $\mu(\hat{s}_1 \mid \omega^1, s_1) = \mu(\hat{s}_2 \mid \omega^1, s_2) = 1$ whenever $\omega^1 \in \Omega_1$, and $\mu(\hat{s}_2 \mid \omega^2, s_1) = \mu(\hat{s}_1 \mid \omega^2, s_2) = 1$ whenever $\omega^2 \in \Omega_2$. That is, receiving $\hat{s}_1$ can be interpreted as receiving message "don't switch", whereas receiving $\hat{s}_1$ can be interpreted as receiving message "switch". The benefit from following this recommendation is i

$$V(\hat{\pi} \mid \pi) = \sum_{\omega \in \Omega_1} \mu(\hat{s}_2, \omega) (u_1(\omega) - u_2(\omega)) + \sum_{\omega \in \Omega_2} \mu(\hat{s}_2, \omega) (u_2(\omega) - u_1(\omega)).$$

Take any $s \in \pi'$. First let's show that $V(\hat{\pi} \mid s) \leq V(\hat{\pi} \mid \pi)$. If $\hat{s}_1$ and $\hat{s}_2$ induce the same action conditional on s' , then $V(\hat{\pi} \mid s') = 0$. Now, let's consider the other two cases separately:

First, assume that $\hat{\pi}$ is such that $\hat{s}_1$ induces action a_1 and $\hat{s}_2$ induces action a_2 . We have

$$\begin{aligned}
V(\hat{\pi} \mid s) &= \mu(\hat{s}_1 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_1, s) u_1(\omega) + \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_2, s) u_2(\omega) - \max_{a_1, a_2} \sum_{\omega \in \Omega} \mu(\omega \mid s) u(a, \omega) \\
&\leq \mu(\hat{s}_1 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_1, s) u_1(\omega) + \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_2, s) u_2(\omega) - \sum_{\omega \in \Omega} \mu(\omega \mid s) u_1(\omega) \\
&= \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_2, s) (u_2(\omega) - u_1(\omega)) \\
&= \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega_2} \mu(\omega \mid \hat{s}_2, s) (u_2(\omega) - u_1(\omega)) + \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega_1} \mu(\omega \mid \hat{s}_2, s) (u_2(\omega) - u_1(\omega)) \\
&\leq \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega_2} \mu(\omega \mid \hat{s}_2, s) (u_2(\omega) - u_1(\omega)).
\end{aligned}$$

The last inequality follows because $u_1(\omega) > u_2(\omega)$ for all $\omega \in \Omega_1$.

Second, assume that $\hat{\pi}$ is such that $\hat{s}_1$ induces action a_2 and $\hat{s}_2$ induces action a_1 . We have

$$\begin{aligned}
V(\hat{\pi} \mid s) &= \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_2, s) u_1(\omega) + \mu(\hat{s}_1 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_1, s) u_2(\omega) - \max_{a_1, a_2} \sum_{\omega \in \Omega} \mu(\omega \mid s) u(a, \omega) \\
&\leq \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_2, s) u_1(\omega) + \mu(\hat{s}_1 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_1, s) u_2(\omega) - \sum_{\omega \in \Omega} \mu(\omega \mid s) u_2(\omega) \\
&= \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_2, s) u_1(\omega) - \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_2, s) u_2(\omega) \\
&= \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega} \mu(\omega \mid \hat{s}_2, s) (u_1(\omega) - u_2(\omega)) \\
&= \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega_1} \mu(\omega \mid \hat{s}_2, s) (u_1(\omega) - u_2(\omega)) + \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega_2} \mu(\omega \mid \hat{s}_2, s) (u_1(\omega) - u_2(\omega)) \\
&\leq \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega_1} \mu(\omega \mid \hat{s}_2, s) (u_1(\omega) - u_2(\omega)).
\end{aligned}$$

The last inequality follows because $u_1(\omega) < u_2(\omega)$ for all $\omega \in \Omega_2$. So, overall we have

$$\begin{aligned} V(\hat{\pi} \mid s) &\leq \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega_2} \mu(\omega \mid \hat{s}_2, s) (u_2(\omega) - u_1(\omega)) + \mu(\hat{s}_2 \mid s) \sum_{\omega \in \Omega_1} \mu(\omega \mid \hat{s}_2, s) (u_1(\omega) - u_2(\omega)) \\ &= \mu(\hat{s}_2 \mid s) \left(\sum_{\omega \in \Omega_2} \mu(\omega \mid \hat{s}_2, s) (u_2(\omega) - u_1(\omega)) + \sum_{\omega \in \Omega_1} \mu(\omega \mid \hat{s}_2, s) (u_1(\omega) - u_2(\omega)) \right). \end{aligned}$$

This implies that

$$\begin{aligned} V(\hat{\pi} \mid \pi') &= \sum_{s \in \pi'} \mu(s) V(\hat{\pi} \mid s) \\ &\leq \sum_{s \in \pi'} \mu(s) \mu(\hat{s}_2 \mid s) \left[\sum_{\omega \in \Omega_2} \mu(\omega \mid \hat{s}_2, s) (u_2(\omega) - u_1(\omega)) + \sum_{\omega \in \Omega_1} \mu(\omega \mid \hat{s}_2, s) (u_1(\omega) - u_2(\omega)) \right]. \end{aligned}$$

Note that since

$$\mu(s) \mu(\hat{s}_2 \mid s) \mu(\omega \mid \hat{s}_2, s) = \mu(\omega, \hat{s}_2, s).$$

we have

$$\begin{aligned} V(\hat{\pi} \mid \pi') &\leq \sum_{s \in \pi'} \left[\sum_{\omega \in \Omega_2} \mu(\omega, \hat{s}_2, s) (u_2(\omega) - u_1(\omega)) + \sum_{\omega \in \Omega_1} \mu(\omega, \hat{s}_2, s) (u_1(\omega) - u_2(\omega)) \right] \\ &= \sum_{\omega \in \Omega_2} \sum_{s \in \pi'} \mu(\omega, \hat{s}_2, s) (u_2(\omega) - u_1(\omega)) + \sum_{\omega \in \Omega_1} \sum_{s \in \pi'} \mu(\omega, \hat{s}_2, s) (u_1(\omega) - u_2(\omega)). \end{aligned}$$

Now, since π' is the full support of s we have

$$\sum_{s \in \pi'} \mu(\omega, \hat{s}_2, s) = \mu(\omega, \hat{s}_2).$$

That is,

$$\begin{aligned} V(\hat{\pi} \mid \pi') &\leq \sum_{\omega \in \Omega_2} \mu(\omega, \hat{s}_2) (u_2(\omega) - u_1(\omega)) + \sum_{\omega \in \Omega_1} \mu(\omega, \hat{s}_2) (u_1(\omega) - u_2(\omega)) \\ &= V(\hat{\pi} \mid \pi). \end{aligned}$$

□

6.3 Proof of Proposition 1

Proof. Consider the following menu that consists of the collections of all minimal complementary signals

$$\mathcal{M} = \{(\pi^c, t(\pi))\}_{\pi \in \Pi}, \quad \text{where} \quad t(\pi) = U(\pi \vee \pi^c) - U(\pi).$$

Note that this is the maximum possible revenue the data broker can achieve, since each type ends up with full information and is charged its maximum willingness to pay for that information. Lemma 2 shows that this optimal menu is incentive compatible, as no type benefits from mimicking and purchasing another type's minimal complementary signal. $\square$

6.4 Proof of Lemma 4

Proof. Without loss of generality assume π, π' are such that $U(\pi) > U(\pi')$. Henceforth, I refer to the decision maker with signal π as the *low type* and the one with signal π' as the *high type*, ordered by their willingness to pay for the fully revealing signal. Since π is a finite partition of the state space Ω , for any $s \in \pi$ we have $s \subseteq \Omega$. From lemma 3, the expected payoff from π is

$$U(\pi) = \sum_{s \in \pi} \mu(\omega_s) u(\omega_s).$$

Define the set $S = \bigcup_{s \in \pi} \{\omega_s\}$, which consists of all the set of states in which the chosen action is optimal. Note that

$$U(\pi) = \sum_{\omega \in S} \mu(\omega) u(\omega).$$

Consider the following signal

$$\hat{\pi} = \{\{\omega\} \mid \omega \in \Omega \setminus S\} \cup \{S\}.$$

It is immediate that $\pi \vee \hat{\pi}$ is fully revealing, since $\hat{\pi}$ reveals precisely the states in which the action induced by π is not optimal. The signal $\hat{\pi}$ is interpreted as follows: when the action induced by π is optimal, it sends a “do not switch” message; when it is not, it fully reveals the state. Low type DM's benefit from $\hat{\pi}$ is

$$V(\hat{\pi} \mid \pi) = \sum_{\omega \in \Omega \setminus S} \mu(\omega) u(\omega).$$

Now, compute the benefit to the high type from mimicking the low type, i.e., evaluate $\pi' \vee \hat{\pi}$.

$$\begin{aligned}\hat{\pi} \vee \pi' &= \{s' \cap \hat{s} \mid s' \in \pi', \hat{s} \in \hat{\pi}\} \\ &= \{s' \cap S \mid s' \in \pi', s' \cap S \neq \emptyset\} \cup \{\{\omega\} \mid \omega \in \Omega \setminus S\}.\end{aligned}$$

If $s' \cap S = \emptyset$, define $\omega_\phi = \omega_{s'}$. The payoff from $\hat{\pi} \vee \pi'$ is

$$U(\hat{\pi} \vee \pi') = \sum_{s' \in \pi'} \mu(\omega_{s' \cap S}) u(\omega_{s' \cap S}) + \sum_{\omega \in \Omega \setminus S} \mu(\omega) u(\omega).$$

Observe that $\mu(\emptyset) = 0$. The payoff from π' is

$$U(\pi') = \sum_{s' \in \pi'} \mu(\omega_{s'}) u(\omega_{s'}).$$

Thus, the high type's benefit from $\hat{\pi}$ is

$$\begin{aligned}V(\hat{\pi} \mid \pi') &= U(\hat{\pi} \vee \pi') - U(\pi') \\ &= \sum_{s' \in \pi'} \mu(\omega_{s' \cap S}) u(\omega_{s' \cap S}) + \sum_{\omega \in \Omega \setminus S} \mu(\omega) u(\omega) - \sum_{s' \in \pi'} \mu(\omega_{s'}) u(\omega_{s'}) \\ &= \sum_{\omega \in \Omega \setminus S} \mu(\omega) u(\omega) - \sum_{s' \in \pi'} [\mu(\omega_{s'}) u(\omega_{s'}) - \mu(\omega_{s' \cap S}) u(\omega_{s' \cap S})] \\ &\leq \sum_{\omega \in \Omega \setminus S} \mu(\omega) u(\omega) \\ &= V(\hat{\pi} \mid \pi).\end{aligned}$$

The last inequality follows because $\omega_{s' \cap S} \in s' \cap S \subseteq s'$ and, by definition of $\omega_{s'}$,

$$\mu(\omega_{s'}) u(\omega_{s'}) \geq \mu(\omega) u(\omega) \quad \forall \omega \in s'.$$

To complete the proof, consider the optimal menu

$$\mathcal{M} = \{(\bar{\pi}, t'), (\hat{\pi}, t)\},$$

where $\bar{\pi}$ is the fully revealing signal, $t' = V(\bar{\pi} \mid \pi')$, and $t = V(\hat{\pi} \mid \pi)$.

□

6.5 Proof of Proposition 2

Proof. Pick any $\pi \in \tilde{\Pi}(\Omega)$. Define

$$\hat{\pi} = \{\{\omega\} \mid \omega \in \Omega \setminus S\} \cup \{S\} \quad \text{where} \quad S = \bigcup_{s \in \pi} \{\omega_s\}.$$

Observe that $\pi \vee \hat{\pi}$ fully reveals all states, i.e., $U(\hat{\pi} \vee \pi) = \overline{U}$. Consider the following menu of signal–price pairs

$$\mathcal{M} = \{(\hat{\pi}, t_\pi) \mid \pi \in \tilde{\Pi}(\Omega)\} \quad \text{where} \quad t_\pi = U(\hat{\pi} \vee \pi) - U(\pi).$$

This menu extracts all information rents from the decision maker. It is now left to check that this menu satisfies the IC and IR constraints. It is clear IR constraints hold. To verify that the IC holds, take any $\pi' \in \tilde{\Pi}(\Omega)$. If $U(\pi) > U(\pi')$, then type π will not mimic type π' because $t_\pi < t_{\pi'}$ and $V(\hat{\pi} \mid \pi) \geq V(\hat{\pi}' \mid \pi)$. If $U(\pi) < U(\pi')$, then from Lemma 4, type π DM will not benefit from mimicking type π' . $\square$