Tailoring Data for Profit*

Xueying Zhao †

March 19, 2025

- Click here for the latest version -

Abstract

A data buyer initially has *access* to a private information structure that provides partial information about a payoff-relevant state. The buyer's initial information determines their willingness to pay for additional data. A monopolistic data seller, capable of generating information structures that may correlate with the buyer's initial information, seeks to maximize revenue. The key findings demonstrate that the seller can extract the first-best surplus and achieve social efficiency by offering customized supplemental information products tailored to different buyer types. In particular, the seller can implement first-degree price discrimination by strategically leveraging data correlations.

Keywords: Mechanism design, information design, correlation, multidimensional screening, full surplus extraction, first-best implementation

JEL Codes: D42, D82, D86

*I thank especially Daniel Sgroi and Sinem Hidir for their guidance and support. I thank Pablo Beker, Costas Cavounidis, Daniele Condorelli, Rahul Deb, Miaomiao Dong, James Fenske, Alkis Georgiadis-Harris, Nima Haghpanah, Peter Hammond, Kevin He, Harry Pei, Kirill Pogorelskiy, Herakles Polemarchakis, Philip Reny, Jakub Steiner, Ao Wang, Yu Fu Wong, and participants at Stony Brook International Conference on Game Theory and Warwick MIWP Workshop for helpful comments and suggestions.

†Department of Economics, University of Warwick. Email: xueying.zhao@warwick.ac.uk

1 Introduction

In today's digital economy, firms collect vast amounts of data to make better-informed decisions. For instance, companies gather demographic information to understand consumer preferences and offer personalized prices, while banks analyze applicants' financial data to assess credit risk. However, the data firms initially obtain is often imperfect, creating a demand for additional information. This derives the emergence of data markets, where firms become buyers of information products. Firms' initial information is private and determines their willingness to pay for additional data. This information asymmetry presents challenges for data sellers who seek to maximize revenue. This paper explores the optimal design and pricing of information in such a setting.

A data buyer faces a decision problem under uncertainty, where his utility depends on his action and a payoff-relevant state. Initially, the buyer has *access* to an information structure that provides only partial information about the state. To further reduce uncertainty and make better-informed decisions, the buyer can purchase additional information from a data seller. The initial information structure is private to the buyer and determines his willingness to pay for any additional information. A monopolist data seller can generate new information structures that may correlate with the buyer's initial information and offer them for sale. We assume that only the information structure itself is contractible. The seller's objective is to design a revenue-maximizing menu of information structures with corresponding prices.

This setup presents a multidimensional screening problem in which the type space consists of various information structures. One interpretation of this model is that the data buyer already has a subscription to an initial information structure and will observe its realization only after deciding whether or not to purchase an additional information structure. Although the seller does not know the buyer's exact initial information structure, he knows the distribution of all possible information structures.

We analyze multiple information structures and allow for arbitrary correlations between them. Rather than modeling information structures as Blackwell experiments, we follow Green and Stokey (1978) by modeling information structures as partitions of an expanded state space. The formalism and graphical illustrations are based on Gentzkow and Kamenica (2017), which are useful for analyzing the joint informational content from multiple information structures and understanding correlations between them. The results build upon the concept of strong Blackwell order introduced by Brooks et al. (2024). Therefore, I adopt their terminology, referring to information structures as *signals* and to observations from these structures as *signal realizations.*

From the seller's perspective, the data buyer may be of different types, each characterized by its private signal. The realization of the private signal is observed **after** the buyer decides whether or not to purchase an additional signal. The value of this private signal is evaluated from an **ex-ante** perspective, representing the expected utility the buyer can achieve by making optimal decisions after observing each realization. The buyer's willingness to pay (WTP) for any additional signal depends on the incremental value it adds to his private signal. Consequently, the WTP varies across different buyer types for the same signal.

In the full-information benchmark, where the seller knows the buyer's private signal, the seller can offer a fully informative signal at a price equal to the buyer's WTP. The maximum possible revenue in this scenario is referred to as the first-best revenue. In the actual model, however, the seller designs a menu of signals that differ in informativeness and price, allowing the buyer to self-select based on his private signal. Selling a fully informative signal is not always necessary, as the buyer's private signal already provides information about the state. Instead, the seller can exploit the **correlation** between the buyer's private signal and the offered signals. By providing a supplemental signal that fills in gaps in the buyer's existing information, the seller ensures the buyer attains complete information about the state.

We characterize the properties of an optimal menu for the seller's revenue-maximizing problem in a setting with binary buyer types (Proposition 1). From the seller's perspective, a buyer who derives less value from his private signal is considered more valuable and is thus referred to as the high type, whereas a buyer with a higher-valued private signal is referred to as the low type. Two familiar properties emerge: (i) "Efficiency at the top," where the high type purchases a signal that supplements his private signal to achieve complete information, and (ii) "No rent at the bottom", where the seller extracts full surplus from the low type by offering the buyer his reservation utility and paying zero information rent.

The first-best implementation refers to achieving an optimal outcome under asymmetric information, which coincides with the outcome that would occur with full information. We introduce the WTP condition and show that the first-best implementation is achieved under this condition (Proposition 2). The key insight is that the seller can design a menu of differentiated signals that is both individual-rational and incentive-compatible, enabling the seller to sell a supplementary signal to each buyer type at a price equal to the buyer's WTP. This allows the seller to extract full surplus from both buyer types, resulting in a **socially efficient** outcome.

To avoid providing redundant information, we introduce the concept of a **minimal supplement**, which offers only the additional information necessary to complete the buyer's knowledge about the state. Using this concept, we construct two compound signals and derive an alternative interpretation of the WTP condition. A sufficient condition for the WTP condition to hold is established (Proposition 5), based on the Blackwell order between these compound signals. Given that the value of a signal is equal to the value of the experiment induced by the signal,[1] we focus on the Blackwell order between the experiments induced by the compound signals. We formalize a less commonly used definition of Blackwell order on experiments (Lemma 6) and demonstrate that the first-best implementation is surprisingly common through Example 1 and Example 2. For instance, in cases with binary states and private signals with limited realizations, the seller can achieve the first-best revenue (Proposition 6). Furthermore, when two Blackwell-ordered signals with two realizations each are available, the first-best implementation is achieved (Proposition 7).

1.1 Related Literature

This paper contributes to the recent literature on selling information to privately informed buyers. We explore the optimal design of **signals**, distinguished from existing studies that focus on the optimal design of **experiments** (Bergemann, Bonatti, and Smolin, 2018; Rodríguez Olivera, 2024). Formalizing an information source as a signal provides a

[1]A signal induces an experiment, as discussed in Section 2.1.

distinct approach compared to formalizing it as an experiment. Blackwell (1951) models an information source as an experiment, where the correlation between observations from that source and the state is specified. In contrast, Green and Stokey (1978) model an information source as a signal, which not only specifies the correlation between its observations and the state but also considers the correlation between its observations and those from other sources. While most prior work assumes conditional independence between information sources—making it sufficient to model them as experiments—we allow for arbitrary correlations, following the approach of Green and Stokey (1978), and model information sources as signals.

The most relevant paper is Bergemann, Bonatti, and Smolin (2018), and the novelty of this paper lies in the combination of two key features: (i) the type space consists of various information structures and (ii) correlations are allowed between the data buyer's initial information structure and the additional information structure offered by the data seller.

This paper also adds to the literature on multidimensional screening, a topic known for its inherent complexity and challenging tractability (Stole and Rochet, 2003). In our model, the buyer's private information is multidimensional, represented by a signal that induces a distribution of beliefs. We provide a characterization of an optimal mechanism that applies to other multidimensional screening problems, including those with type-dependent outside options.

Another strand of related literature examines the joint informational content of multiple information sources. Börgers, Hernando-Veciana, and Krähmer (2013) explore the substitutability and complementarity relations among signals. Gentzkow and Kamenica (2017) discuss Bayesian persuasion in a setting where multiple senders have access to a rich signal space, allowing for arbitrary correlation among the senders' signals. Brooks, Frankel, and Kamenica (2024) derive comparisons of information sources that remain robust to the potential presence of pre-existing information.

2 The Model

2.1 Model Setup

Consider two players: a single data buyer and a monopolist data seller. The buyer faces a decision problem under uncertainty and his utility function $u(a, \omega)$ depends on his action $a \in A$ and the state of the world $\omega \in \Omega$. The action space A is compact, and the state space Ω is finite. Assume that the state space consists of K elements:

$$\Omega \triangleq \{\omega_1, ..., \omega_k, ..., \omega_K\}.$$

A belief is a distribution over the state space, denoted by $\mu \in \Delta(\Omega)$.[2] Let μ_0 be the interior prior,[3] which is commonly known.

Information structures are modeled as signals rather than Blackwell experiments. A **signal** π is a finite partition of the expanded state space $\Omega \times [0, 1]$, with each element of this partition belonging to S, the set of non-empty Lebesgue-measurable subsets of $\Omega \times [0, 1]$ (Green and Stokey, 1978; Gentzkow and Kamenica, 2017).[4] An element $s \in S$ is a signal realization. This formalism distinguishes payoff-relevant states (Ω) from those that govern the realization of observations conditional on the state ($[0, 1]$). The interpretation is that a random variable x, drawn uniformly from $[0, 1]$, determines the signal realization conditional on the state. Specifically, the buyer with signal π will observe the realization $s \in \pi$ that contains $(\omega, x) \in \Omega \times [0, 1]$. Thus, the conditional probability of s given ω is

$$p(s \mid \omega) \triangleq \lambda(\{x \mid (\omega, x) \in s\}),$$

where λ denotes the Lebesgue measure. The unconditional probability of s is

$$p(s) \triangleq \sum_{\omega \in \Omega} \mu_0(\omega) p(s \mid \omega).$$

Upon observing the signal realization s, the posterior belief μ_s is formed via Bayes' rule

[2] $\Delta(X)$ denotes the set of all probability distributions over the set X.

[3] The prior probability of each state is strictly positive.

[4] Green and Stokey (1978, 2022) introduce the notion of signals as partitions of an expanded state space. The particular formalism used in this paper follows Gentzkow and Kamenica (2017).

(for $p(s) > 0$), where the posterior probability of ω given s is

$$\mu_s(\omega) \triangleq \frac{\mu_0(\omega)p(s \mid \omega)}{p(s)}.$$

This representation is useful for analyzing the joint informational content from multiple sources and understanding correlations between signal realizations. For a graphical illustration, see Figure 1. In this example, let $\Omega = \{\omega_1, \omega_2\}$ and $\pi = \{a, b\}$, where $a = (\omega_1, [0, 0.7]) \cup (\omega_2, [0.6, 1])$ and $b = (\omega_1, [0.7, 1]) \cup (\omega_2, [0, 0.6])$. The signal π is a finite partition of $\Omega \times [0, 1]$ with conditional probabilities $p(a \mid \omega_1) = 0.7$ and $p(b \mid \omega_2) = 0.6$.

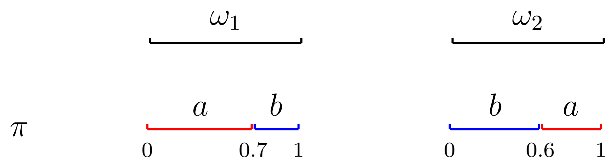

Figure 1: A signal π.

Let Π be the set of all signals.[5] The buyer has private information about the state, represented by a signal $\pi \in \Pi_0$, where $\Pi_0 \subset \Pi$ is the set of N possible private signals:

$$\Pi_0 \triangleq \{\pi_1, ..., \pi_n, ..., \pi_N\}.$$

The buyer's type is captured by his private signal, which induces a **distribution of beliefs**. Thus, the space of buyer types corresponds to the space of private signals. Specifically, a buyer who has private signal π_i is referred to as type π_i, where $i \in \{1, ..., N\}$.

In the data market, the seller can generate any signals at no cost and offer them for sale to the buyer. Although the seller does not know the exact type of the buyer, he knows that the buyer is of type π_i with probability δ_i. To maximize revenue, the seller offers a menu of differentiated signals with associated prices. The buyer will observe a realization from his private signal only **after** deciding whether or not to purchase an additional signal. If the buyer opts out of the data market, he will observe one realization only from his private signal. However, if he chooses to buy an additional signal, he will observe two

[5]Arbitrary correlations between signal realizations across signals are allowed.

realizations: one from his private signal and another from the purchased signal, with a transfer made to the seller. After updating his beliefs about the state, the buyer selects an action. Our objective is to determine the revenue-maximizing menu of signals, along with their corresponding prices, for the seller.

Timing of Events

(i) The data seller offers the data buyer a menu M of information structures with associated prices.

(ii) Nature draws a state $\omega \in \Omega$ according to the prior belief μ_0. The buyer knows his initial information structure $\pi \in \Pi_0$.

(iii) The buyer decides whether or not to purchase an additional information structure. If he bought, the seller receives a transfer.

(iv) The buyer observes a realization from his initial information structure. If the buyer purchases an additional information structure, he will observe another realization from it.

(v) The buyer chooses an action $a \in A$, and his payoff is realized.

Signals vs. Experiments

Signals are distinct from experiments. An **experiment** consists of a set of possible outcomes and a family of conditional distributions over these outcomes given the state. While a signal induces an experiment, it also specifies the correlation with other signals. Two distinct signals can induce identical experiments. For instance, as illustrated in Figure 2, consider $\Omega = \{\omega_1, \omega_2\}$ and two signals, $\pi = \{c, d\}$ and $\pi' = \{e, f\}$. Both signals have identical conditional distributions: $p(c \mid \omega_1) = p(e \mid \omega_1) = 0.7$ and $p(d \mid \omega_2) = p(f \mid \omega_2) = 0.5$. Thus, they induce identical experiments.

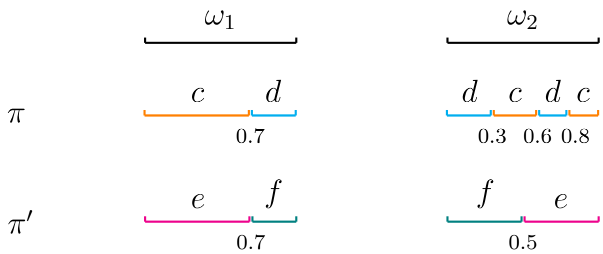

Figure 2: Two distinct signals π and π' induce identical experiments.

2.2 Value of Signals

Given a belief μ, the buyer selects an action $a \in A$ to maximize expected utility:

$$a(\mu) \in \arg\max_{a \in A} \mathbb{E}_{\omega \sim \mu}[u(a, \omega)].$$

Let $\hat{v}(\mu)$ denote the expected utility of the buyer from choosing the optimal action given belief μ, where

$$\hat{v}(\mu) \triangleq \max_{a \in A} \mathbb{E}_{\omega \sim \mu}[u(a, \omega)] = \max_{a \in A} \sum_{\omega \in \Omega} \mu(\omega) u(a, \omega).$$

Lemma 1. *The function $\hat{v}(\mu) : \Delta(\Omega) \to \mathbb{R}$ is convex.*

The proof of Lemma 1 is provided in the Appendix.

The ex-ante **value of signal** π, denoted by $v(\pi)$, is the expected utility that the buyer can achieve by acting optimally upon observing each realization of the signal π. It is given by

$$v(\pi) \triangleq \sum_{s \in \pi} p(s) \hat{v}(\mu_s) = \sum_{s \in \pi} \max_{a \in A} \sum_{\omega \in \Omega} u(a, \omega) p(s, \omega),$$

where $p(s, \omega) \triangleq \mu_0(\omega) p(s \mid \omega)$.

Remark 1. *The value of a signal is equal to the value of the experiment induced by that signal.*

To discuss multiple signals, it is useful to define the join of two signals.

Definition 1 (Refinement/Coarsening)**.** *A signal π is a **refinement** of π' (or equivalently, π' is a **coarsening** of π), if every element of π is a subset of one element of π'. Formally, for each $s \in \pi$, there is an $s' \in \pi'$ such that $s \subseteq s'$.*

Definition 2 (Join)**.** *The **join** of signals π and $\hat{\pi}$, denoted by $\pi \vee \hat{\pi}$, is the coarsest common refinement of π and $\hat{\pi}$. It is defined as the set of intersections formed by pairing each element of π with each element of $\hat{\pi}$:*

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

The signal $\pi \vee \hat{\pi}$ yields the same information as observing both signals π and $\hat{\pi}$. Figure 3 illustrates the join of two signals.

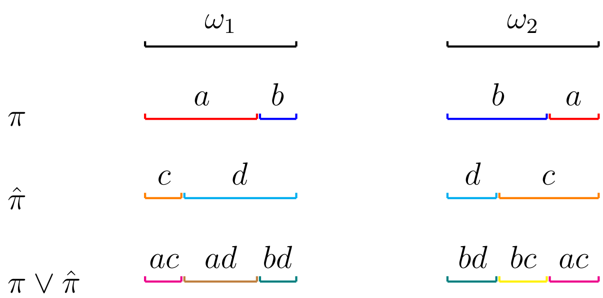

Figure 3: The signal $\pi \vee \hat{\pi}$ is the join of signals π and $\hat{\pi}$.

The value of the signal $\pi \vee \hat{\pi}$ is given by

$$v(\pi \vee \hat{\pi}) \triangleq \sum_{s' \in \pi \vee \hat{\pi}} p(s')\hat{v}(\mu_{s'}) = \sum_{s \in \pi} \sum_{\hat{s} \in \hat{\pi}} \max_{a \in A} \sum_{\omega \in \Omega} u(a, \omega) p(s \cap \hat{s}, \omega).$$

To establish the upper and lower bounds for the value of signals, we introduce two important signals: $\underline{\pi}$ and $\overline{\pi}$.

The signal $\underline{\pi}$ is the **coarsest uninformative signal**. An uninformative signal is one that induces a posterior identical to the prior.[6] The signal $\underline{\pi}$ is defined as:

$$\underline{\pi} = \{\Omega \times [0, 1]\}.$$

[6]Multiple uninformative signals can be generated by refining the signal $\underline{\pi}$.

Figure 4 illustrates this concept. In the case of a binary state space $\Omega = \{\omega_1, \omega_2\}$, $\underline{\pi} = \{a\}$, where $a = \Omega \times [0, 1]$.

Figure 4: The signal $\underline{\pi}$ in binary states.

The value of the signal $\underline{\pi}$, denoted by $\underline{v}$, is given by

$$\underline{v} \triangleq v(\underline{\pi}) = \hat{v}(\mu_0) = \max_{a \in A} \sum_{\omega \in \Omega} \mu_0(\omega) u(a, \omega).$$

The signal $\overline{\pi}$ is the **coarsest fully informative signal**. A fully informative signal provides complete information about the state by inducing a distribution of degenerate beliefs.[7] The signal $\overline{\pi}$ is defined as:

$$\overline{\pi} = \{(\omega_1, [0, 1]), ..., (\omega_k, [0, 1]), ..., (\omega_K, [0, 1])\}.$$

See Figure 5 for an illustration. For a binary state space $\Omega = \{\omega_1, \omega_2\}$, $\overline{\pi} = \{b, c\}$, where $b = (\omega_1, [0, 1])$, $c = (\omega_2, [0, 1])$.

Figure 5: The signal $\overline{\pi}$ in binary states.

The value of the signal $\overline{\pi}$, denoted by $\overline{v}$, is given by

$$\overline{v} \triangleq v(\overline{\pi}) = \sum_{\omega \in \Omega} \mu_0(\omega) \max_{a \in A} u(a, \omega).$$

[7] Multiple fully informative signals can be generated by refining the signal $\overline{\pi}$.

Lemma 2. *The value of signals is bounded such that:*

$$\underline{v} \leq v(\pi) \leq \overline{v}, \quad \forall \pi \in \Pi.$$

Refer to the Appendix for the proof of Lemma 2.

2.3 Seller' Problem

By the revelation principle, we can focus on direct mechanisms.[8] A direct mechanism M assigns a signal $\tilde{\pi}(\pi_i) : \Pi_0 \to \Pi$ and a price $t(\pi_i) : \Pi_0 \to \mathbb{R}$ to each type of buyer, where

$$M \triangleq \{(\tilde{\pi}(\pi_i), t(\pi_i))\}_{\pi_i \in \Pi_0}.$$

For simplicity, let $\tilde{\pi}_i$ represent $\tilde{\pi}(\pi_i)$ and t_i represent $t(\pi_i)$. Thus, any direct mechanism can be denoted by:

$$M = \{(\tilde{\pi}_i, t_i)\}_{i \in \{1,\ldots,N\}}.$$

In the data market, the seller offers a menu M of signals to the buyer. If the buyer of type π_i purchases signal $\tilde{\pi}_j$, where $j \in \{1, ..., N\}$, his payoff is

$$v(\pi_i \vee \tilde{\pi}_j) - t_j,$$

which is the value of having both the private signal π_i and the additional signal $\tilde{\pi}_j$, minus the transfer t_j made to the seller. Given that the buyer is of type π_i with probability δ_i, the expected revenue for the seller is

$$R \triangleq \mathbb{E}[t_i] = \sum_{i \in \{1,\ldots,N\}} \delta_i t_i.$$

The **seller's problem** is to choose a menu of signals with associated prices to maximize expected revenue:

$$\max_{\{(\tilde{\pi}_i, t_i)\}_{i \in \{1,\ldots,N\}}} \sum_{i \in \{1,\ldots,N\}} \delta_i t_i,$$

[8] See Lemma 1 in Myerson (1981).

subject to the individual-rationality constraints:

$$v(\pi_i \vee \tilde{\pi}_i) - t_i \geq v(\pi_i), \quad \forall i \in \{1, ..., N\},$$

and incentive-compatibility constraints:

$$v(\pi_i \vee \tilde{\pi}_i) - t_i \geq v(\pi_i \vee \tilde{\pi}_j) - t_j, \quad \forall i, j \in \{1, ..., N\}, \quad i \neq j.$$

To obtain the lower and upper bounds of revenue from selling data, it is useful to consider two benchmarks.

Definition 3 (Reservation Utility)**.** *The* ***reservation utility*** *of type π_i is defined as the value of the private signal π_i, denoted by $v(\pi_i)$.*

Assumption 1. *The reservation utility of each type of buyer* ***decreases*** *with the index n:*

$$v(\pi_1) \geq ... \geq v(\pi_n) \geq ... \geq v(\pi_N).$$

Definition 4 (Willingness to Pay)**.** *The* ***willingness to pay (WTP)*** *of type π_i for signal π is defined as the incremental value that signal π adds to type π_i. Formally, it is the difference between the value of signal $\pi_i \vee \pi$ and the value of π_i, denoted by*

$$v(\pi_i \vee \pi) - v(\pi_i).$$

The WTP varies across different buyer types for the same signal. For instance, the WTP of type π_i for a fully informative signal, such as $\overline{\pi}$, is given by

$$v(\pi_i \vee \overline{\pi}) - v(\pi_i) = v(\overline{\pi}) - v(\pi_i) = \overline{v} - v(\pi_i),$$

where the first equality comes from Remark 1.

Under Assumption 1, the WTP of type π_i for a fully informative signal **increases** with the index n, such that:

$$\overline{v} - v(\pi_1) \leq ... \leq \overline{v} - v(\pi_n) \leq ... \leq \overline{v} - v(\pi_N).$$

Single-item-menu Benchmark

The lower bound of revenue can be analyzed in a scenario where the seller is restricted to selling only a single item to the buyer. This benchmark provides a baseline for evaluating the seller's potential revenue compared to more complex mechanisms. The seller sets a uniform price to maximize expected revenue from selling a fully informative signal, such as $\overline{\pi}$, to the buyer. The buyer will purchase the signal if his WTP for $\overline{\pi}$ exceeds this price. The optimal price must equal the WTP of one type for the signal $\overline{\pi}$; otherwise, the seller could improve revenue by adjusting the price. The highest revenue achievable under this benchmark serves as the **lower bound of revenue**, denoted by:

$$\underline{R} \triangleq \max_{n\in\{1,\dots,N\}} (\overline{v} - v(\pi_n)) \sum_{i=n}^{N} \delta_i.$$

Full-information Benchmark

The upper bound of revenue is evaluated in the full-information benchmark, where the seller knows the type of buyer. In this case, the seller can sell a fully informative signal, such as $\overline{\pi}$, to each type of buyer at a price equal to that type's WTP for $\overline{\pi}$.

Definition 5 (First-best Revenue)**.** *The* ***first-best revenue*** *is the maximum possible revenue that the seller could achieve in the full-information benchmark, denoted by:*

$$\overline{R} \triangleq \sum_{i=1}^{N} \delta_i(\overline{v} - v(\pi_i)).$$

In the actual model, the seller does not know the type of buyer. Therefore, the goal is to design a menu of differentiated signals with corresponding prices to screen different buyer types, thereby extracting as much revenue as possible within the bounds established by the single-item-menu and full-information benchmarks.

3 Optimal Menu with Binary Types

In this section, we consider a scenario where the single data buyer has binary types: $\Pi_0 = \{\pi_1, \pi_2\}$, with $v(\pi_1) \geq v(\pi_2)$. The buyer is either of type π_1 or type π_2. Let

$\delta \in (0,1)$ be the probability that the buyer is of type π_2.

The seller's revenue-maximizing problem is given by

$$\begin{aligned}
\underset{\{(\tilde{\pi}_i, t_i)\}_{i\in\{1,2\}}}{\text{maximize}} \quad & (1-\delta)t_1 + \delta t_2 & \\
\text{subject to} \quad & v(\pi_1 \vee \tilde{\pi}_1) - t_1 \geq v(\pi_1) & (IR_1), \\
& v(\pi_2 \vee \tilde{\pi}_2) - t_2 \geq v(\pi_2) & (IR_2), \\
& v(\pi_1 \vee \tilde{\pi}_1) - t_1 \geq v(\pi_1 \vee \tilde{\pi}_2) - t_2 & (IC_1), \\
& v(\pi_2 \vee \tilde{\pi}_2) - t_2 \geq v(\pi_2 \vee \tilde{\pi}_1) - t_1 & (IC_2).
\end{aligned} \tag{1}$$

In the two benchmarks discussed previously, we established that the seller can sell a fully informative signal to the buyer. However, providing a fully informative signal is not always necessary, as the buyer already has access to a private signal that provides information about the state. To characterize the optimal menu, we first introduce the concept of a supplement of a signal, which is crucial for the subsequent analysis.

Two signals, π and π', are said to be **Blackwell equivalent**, denoted by $\pi \sim \pi'$, if the distribution of posteriors induced by π is identical to the distribution of posteriors induced by π'. For example, the signals π and π' shown in Figure 2 are Blackwell equivalent.

Definition 6 (Supplements)**.** *A* ***supplement*** *of a signal π, denoted by π^{su}, is a signal such that the join of π and π^{su} is Blackwell equivalent to a fully informative signal:*

$$(\pi \vee \pi^{su}) \sim \overline{\pi}.$$

A supplement π^{su} of a signal π provides complete information about the state when combined with π. To construct a supplement of π, consider partitioning each $s \in \pi$ into subsets $\{s_k\}_{k\in\{1,\ldots,K\}}$, where $s_k = \{(\omega_k, x) \mid (\omega_k, x) \in s\}$. The signal $\pi' = \cup_{s\in\pi, k\in\{1,\ldots,K\}} s_k$ is a supplement of π. By garbling π', multiple supplements of π can be generated.

For instance, in Figure 6, consider $\Omega = \{\omega_1, \omega_2\}$ and a signal $\pi = \{a, b\}$, where $a = (\omega_1, [0,1]) \cup (\omega_2, [0.7, 1])$ and $d = (\omega_2, [0, 0.7])$. The signal $\pi' = \{a_1, a_2, b_2\}$ is a supplement of π. However, for binary states, it is sufficient to consider signals with only two realizations. By garbling π', multiple supplements of π can be generated. For each $\gamma \in [0, 0.7)$, the signal $\pi_\gamma = \{c, d\}$ represents a supplement of π, where $c = (\omega_1, [0,1]) \cup (\omega_2, [0,\gamma])$ and $d = (\omega_2, [\gamma, 1])$. Similarly, the signal $\hat{\pi} = \{e, f\}$ is a supplement of π, where $c = (\omega_1, [0,1]) \cup (\omega_2, [0, 0.7])$ and $d = (\omega_2, [0.7, 1])$.

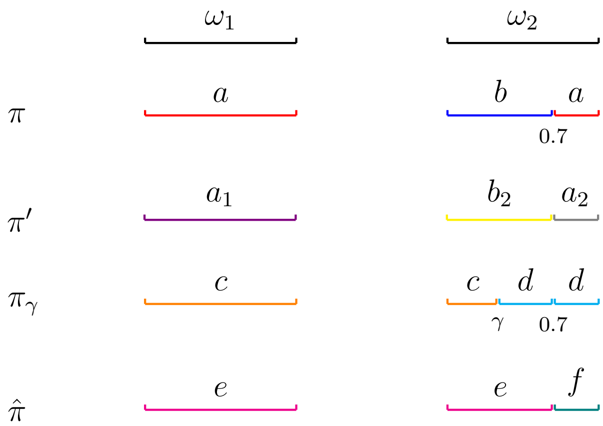

Figure 6: The signals π', π_γ, and $\hat{\pi}$ are supplements of π.

By Remark 1, the WTP of type π_i for the signal π_i^{su}, which is a supplement of the signal π_i, is given by

$$v(\pi_i \vee \pi_i^{su}) - v(\pi_i) = v(\overline{\pi}) - v(\pi_i) = \overline{v} - v(\pi_i).$$

We now characterize the properties of an optimal menu for the revenue-maximizing problem described in (1).

Proposition 1 (Menu Properties). *Consider $\Pi_0 = \{\pi_1, \pi_2\}$ with $v(\pi_1) \geq v(\pi_2)$. The following properties hold in an optimal menu:*

- *(i) Type π_2 pays a weakly higher price than type π_1: $t_2 \geq t_1$;*
- *(ii) Type π_2 purchases a signal $\tilde{\pi}_2$ that is a supplement of π_2: $(\pi_2 \vee \tilde{\pi}_2) \sim \overline{\pi}$;*
- *(iii) Type π_1 receives his reservation utility $v(\pi_1)$: IR_1 binds.*

Type π_2 has a lower reservation utility than type π_1, i.e., $v(\pi_2) \leq v(\pi_1)$. From the seller's perspective, type π_2 is more valuable and is considered as the "high type," while type π_1 is considered as the "low type." As indicated in Proposition 1, in an optimal menu, the high type pays a weakly higher price than the low type. Additionally, two familiar properties are observed: **"efficiency at the top,"** where the high type purchases a signal that supplements his private signal, thereby obtaining complete information about the state, which is as efficient as the full-information benchmark; and **"no rent at the bottom,"** where the low type receives a payoff equal to his reservation utility, implying that the seller pays zero information rent to the low type. The results can be established by contradiction. Details of proof can be found in the Appendix.

Claim 1 (Optimal Menu with A Single-item)**.** *If the two types have identical reservation utilities, i.e.,* $v(\pi_1) = v(\pi_2)$*, the seller can achieve the first-best revenue* $\overline{R}$ *with a single-item menu* $M = \{(\tilde{\pi}_i, t_i)\}_{i \in \{1,2\}}$*, which consists of a fully informative signal offered at a uniform price:*

$$\tilde{\pi}_i = \overline{\pi}, \quad t_i = \overline{v} - v(\pi_1), \ \ \forall i \in \{1, 2\}.$$

Proof. The first-best revenue is given by

$$\overline{R} = (1 - \delta)(\overline{v} - v(\pi_1)) + \delta(\overline{v} - v(\pi_2)) = \overline{v} - v(\pi_1).$$

By Remark 1, we know that $v(\pi_1 \vee \overline{\pi}) = \overline{v} = v(\pi_2 \vee \overline{\pi})$. Given the menu M described above, both types receive their reservation utility, and the seller obtains the first-best revenue $\overline{R}$. □

We now consider the case where $v(\pi_1) > v(\pi_2)$ for the set $\Pi_0 = \{\pi_1, \pi_2\}$. The characterization of an optimal menu will be completed by introducing the willingness to pay (WTP) condition.

Definition 7 (The WTP Condition)**.** *The WTP condition holds if, for the set* $\Pi_0 = \{\pi_1, \pi_2\}$ *with* $v(\pi_1) > v(\pi_2)$*, there exists a signal, denoted by* $\pi_1^* \in \Pi$*, such that*

(i) π_1^* *is a supplement of* π_1*:*

$$(\pi_1 \vee \pi_1^*) \sim \overline{\pi}; \tag{2}$$

(ii) The WTP of type π_1 *for* π_1^* *is weakly higher than the WTP of type* π_2 *for* π_1^**:*

$$v(\pi_1 \vee \pi_1^*) - v(\pi_1) \geq v(\pi_2 \vee \pi_1^*) - v(\pi_2). \tag{3}$$

The WTP condition ensures the existence of a signal that fully supplements the signal π_1, such that type π_1 has a stronger preference for it than type π_2.

Claim 2. *Under the WTP condition, the signal π_1^* cannot be $\overline{\pi}$.*

Proof. It is straightforward to see that the signal $\overline{\pi}$ is a supplement of any signal $\pi \in \Pi$. However, $\overline{\pi}$ violates formula (3) and thus cannot be π_1^* under the WTP condition. □

Proposition 2 (First-best Implementation)**.** *Consider $\Pi_0 = \{\pi_1, \pi_2\}$ with $v(\pi_1) > v(\pi_2)$. If the WTP condition holds, then in an optimal menu:*

(i) Type π_2 receives his reservation utility $v(\pi_2)$: IR_2 must bind;

(ii) The seller achieves the first-best revenue $\overline{R} = (1-\delta)(\overline{v} - v(\pi_1)) + \delta(\overline{v} - v(\pi_2))$.

According to Proposition 2, if the WTP condition holds, full surplus extraction is achievable because the seller pays zero information rent to the buyer, regardless of his type. Furthermore, this full surplus extraction is socially efficient, as the seller can obtain the first-best revenue. The intuition is that the seller can construct a menu of differentiated signals, tailored to each buyer type, priced at the buyer's willingness to pay, enabling efficient screening of the two buyer types.

Under the WTP condition, there exists a signal π_1^* that satisfies both (2) and (3). Note that π_1^* and $\overline{\pi}$ are distinct signals, as established in Claim 2. Therefore, the seller can use π_1^* and $\overline{\pi}$ to construct a menu $M' = \{(\tilde{\pi}_i', t_i')\}_{i\in\{1,2\}}$ of signals, defined as follows:

$$\begin{aligned} \tilde{\pi}_1' &\triangleq \pi_1^*, & t_1' &\triangleq \overline{v} - v(\pi_1), \\ \tilde{\pi}_2' &\triangleq \overline{\pi}, & t_2' &\triangleq \overline{v} - v(\pi_2). \end{aligned}$$

The feasibility and profitability of menu M' are discussed in the Appendix.

Proposition 3. *Consider $\Pi_0 = \{\pi_1, \pi_2\}$ with $v(\pi_1) > v(\pi_2)$. If the WTP condition fails, then IC_2 must bind in an optimal menu.*

For the detailed proof, see the Appendix.

3.1 First-best Implementation

The **first-best implementation** refers to a situation where the optimal outcome is achieved in a setting with asymmetric information, just as it would be if all participants had complete information. Under the WTP condition, the first-best implementation can be attained. We will now establish sufficient conditions for the WTP condition to hold.

Proposition 4 (Supplementary Private Signals)**.** *Consider* $\Pi_0 = \{\pi_1, \pi_2\}$ *with* $v(\pi_1) > v(\pi_2)$*. The WTP condition holds if* π_1 *and* π_2 *are supplements of each other:*

$$(\pi_1 \vee \pi_2) \sim \overline{\pi}.$$

Proof. The proof is straightforward. If $(\pi_1 \vee \pi_2) \sim \overline{\pi}$, then there exists a signal π_1^*, which is identical to π_2, such that π_1^* is a supplement of π_1. Additionally, the WTP of type π_1 for π_1^* is non-negative, while the WTP of type π_2 for π_1^* is zero. $\square$

The result in Proposition 4 relies on the supplementary relationship between signals π_1 and π_2. If π_2 is not a supplement of π_1, we must first identify supplements of π_1.

To ensure that the supplements of a signal contain only the necessary information, we define minimal supplements by using the concept of strong Blackwell order on signals, as introduced by Brooks, Frankel, and Kamenica (2024).

Definition 8 (Blackwell Dominance)**.** *Signal* π ***Blackwell dominates*** *signal* π' *if* π *has a weakly higher value than* π' *in any decision-making problem.*

The Blackwell order on signals is not a partial order, as it is not antisymmetric.[9] For example, in Figure 2, π Blackwell dominates π', and π' Blackwell dominates π, but they are not the same signals. The Blackwell order on signals is reflexive and transitive, making it a preorder.

Definition 9 (Strong Blackwell Dominance)**.** *Signal* π ***strongly Blackwell dominates*** *signal* π' *if, for any signal* $\hat{\pi} \in \Pi$, $\pi \vee \hat{\pi}$ *Blackwell dominates* $\pi' \vee \hat{\pi}$*.*

[9]This result is discussed in Brooks, Frankel, and Kamenica (2024).

The concept that π strongly Blackwell dominates π' is equivalent to the notion that π **reveals-or-refines** π', as established by Theorem 1 in Brooks, Frankel, and Kamenica (2024). This means that every signal realization of π either occurs in only one state (and thus "reveals" the state), or is a subset of one signal realization of π' (and thus "refines" π'). See Figure 7 for an illustration. In this example, consider $\Omega = \{\omega_1, \omega_2\}$ and two signals $\pi = \{c, d\}$ and $\pi' = \{e, f\}$. π strongly Blackwell dominates π' because c is a subset of e and d occurs only in state ω_2.

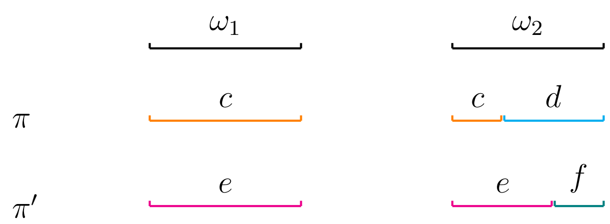

Figure 7: The signal π reveals-or-refines π'.

Definition 10 (Minimal Supplements)**.** *A* ***minimal supplement*** *of a signal π, denoted by π^{ms}, is a supplement of π such that there is no other supplement of π that is strongly Blackwell dominated by π^{ms}.*

Recall that in Figure 6, the signal π_γ is a supplement of π for any $\gamma \in [0, 0.7)$, and the signal $\hat{\pi}$ is also a supplement of π. However, π_γ cannot be considered as a minimal supplement of π because, for any $\gamma \in [0, 0.7)$, π_γ reveals-or-refines $\hat{\pi}$, or equivalently, $\hat{\pi}$ is strongly Blackwell dominated by π_γ. Therefore, $\hat{\pi}$ is a minimal supplement of π.

Lemma 3. *In a model with binary states, a signal with two realizations has a unique minimal supplement.*

Proof. Consider $\Omega = \{\omega_1, \omega_2\}$ and a signal $\pi = \{s_1, s_2\}$ that has two realizations. Let $s'_1 = \{(\omega, x) \mid (\omega_1, x) \in s_1, (\omega_2, x) \in s_2\}$ and $s'_2 = \{(\omega, x) \mid (\omega_2, x) \in s_1, (\omega_1, x) \in s_2\}$. By definition, the signal $\pi' = \{s'_1, s'_2\}$ is the unique minimal supplement of π. $\square$

A signal may have multiple minimal supplements. For example, in Figure 8, both $\breve{\pi}$ and $\hat{\pi}$ are minimal supplements of π.

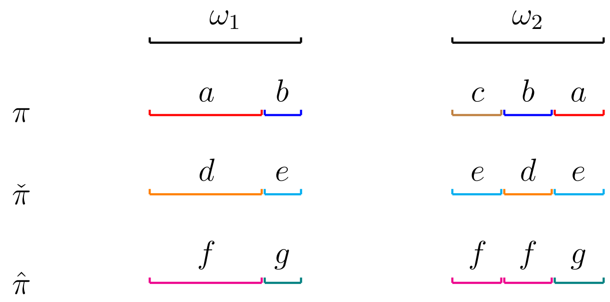

Figure 8: Both $\breve{\pi}$ and $\hat{\pi}$ are minimal supplements of π.

To find the sufficient condition for the WTP condition to hold, we first define two compound signals, π_A and π_B, generated by a random device. Consider the set $\Pi_0 = \{\pi_1, \pi_2\}$ with $v(\pi_1) > v(\pi_2)$, and a signal π_1^{ms} that is a minimal supplement of π_1. The construction of π_A and π_B is as follows.

The **signal** π_A consists of a realization from signal $\pi_1 \vee \pi_1^{ms}$ with probability $\frac{1}{2}$ and a realization from signal π_2 with probability $\frac{1}{2}$. The value of the signal π_A is given by

$$v(\pi_A) = \frac{1}{2}v(\pi_1 \vee \pi_1^{ms}) + \frac{1}{2}v(\pi_2).$$

The **signal** π_B consists of a realization from signal $\pi_2 \vee \pi_1^{ms}$ with probability $\frac{1}{2}$ and a realization from signal π_1 with probability $\frac{1}{2}$. The value of the signal π_B is given by

$$v(\pi_B) = \frac{1}{2}v(\pi_2 \vee \pi_1^{ms}) + \frac{1}{2}v(\pi_1).$$

Proposition 5. *Consider $\Pi_0 = \{\pi_1, \pi_2\}$ with $v(\pi_1) > v(\pi_2)$. The WTP condition holds if there exists a minimal supplement π_1^{ms} such that π_A Blackwell dominates π_B.*

Proof. Consider a minimal supplement π_1^{ms} of the signal π_1. The signals π_A and π_B are constructed from π_1, π_2, and π_1^{ms}, as defined. If π_A Blackwell dominates π_B, i.e., π_A has a weakly higher value than π_B in any decision-making problem:

$$\frac{1}{2}v(\pi_1 \vee \pi_1^{ms}) + \frac{1}{2}v(\pi_2) \geq \frac{1}{2}v(\pi_2 \vee \pi_1^{ms}) + \frac{1}{2}v(\pi_1),$$

which can be written as

$$v(\pi_1 \vee \pi_1^{ms}) - v(\pi_1) \geq v(\pi_2 \vee \pi_1^{ms}) - v(\pi_2),$$

then the WTP condition holds, as there exists a signal π_1^{ms} that satisfies both (2) and (3). $\square$

Lemma 4. *Signal π Blackwell dominates signal π' if and only if the experiment induced by π Blackwell dominates the experiment induced by π'.*

Proof. If the experiment induced by π Blackwell dominates the experiment induced by π', this means that the former experiment has a weakly higher value than the latter in any decision-making problem. According to Remark 1, this implies that signal π has a weakly higher value than signal π' in any decision-making problem, meaning that π Blackwell dominates π'. The converse is straightforward to verify. $\square$

By Lemma 4, to compare the Blackwell order on signals π_A and π_B, we can focus on the Blackwell order of the experiments induced by these signals.

Blackwell Order on Experiments

Given $\Omega = \{\omega_1, ..., \omega_K\}$, there are K possible states. An experiment, which has I possible outcomes, can be described by a $K \times I$ matrix $\mathbf{P} = \{P_{ki}\}$, where P_{ki} is the probability of outcome $i \in \{1, ..., I\}$ in state $k \in \{1, ..., K\}$. We have $P_{ki} \geq 0$ and $\sum_{i=1}^{I} P_{ki} = 1$ for each k, so that $\mathbf{P}$ is called a **Markov matrix**.

Let $\mathbf{P} = \{P_{ki}\}$ and $\mathbf{Q} = \{Q_{kj}\}$ be $K \times I$, $K \times J$ Markov matrices, i.e., any two experiments:

$$\mathbf{P} = \begin{bmatrix} P_{11} & P_{12} & \dots & P_{1I} \\ P_{21} & P_{22} & \dots & P_{2I} \\ \vdots & \vdots & \ddots & \vdots \\ P_{K1} & P_{K2} & \dots & P_{KI} \end{bmatrix}, \quad \mathbf{Q} = \begin{bmatrix} Q_{11} & Q_{12} & \dots & Q_{1J} \\ Q_{21} & Q_{22} & \dots & Q_{2J} \\ \vdots & \vdots & \ddots & \vdots \\ Q_{K1} & Q_{K2} & \dots & Q_{KJ} \end{bmatrix}.$$

We say that experiment $\mathbf{P}$ Blackwell dominates experiment $\mathbf{Q}$ if and only if $\mathbf{P}$ has a weakly higher value than $\mathbf{Q}$ in any decision-making problem. Lemma 5 presents a well-known definition of Blackwell order on experiments.

Lemma 5 (Blackwell and Girshick (1954)). *Experiment* $\mathbf{P}$ *Blackwell dominates experiment* $\mathbf{Q}$ *if and only if there exists an* $I \times J$ *Markov matrix* $\mathbf{D} = \{D_{ij}\}$ *such that*

$$\mathbf{PD} = \mathbf{Q}.$$

Theorem 12.2.2 in Blackwell and Girshick (1954) provides several equivalent definitions of Blackwell order on experiments. I formalize one of these definitions in Lemma 6, which, while less commonly used, is crucial for the subsequent analysis.

Define

$$p_i^* \triangleq \sum_{k=1}^{K} P_{ki}, \qquad q_j^* \triangleq \sum_{k=1}^{K} Q_{kj},$$

$$\mathbf{p}^* \triangleq \begin{bmatrix} p_1^* & p_2^* & \dots & p_I^* \end{bmatrix}, \qquad \mathbf{q}^* \triangleq \begin{bmatrix} q_1^* & q_2^* & \dots & q_J^* \end{bmatrix},$$

$$\mathbf{P}^* \triangleq \begin{bmatrix} \frac{P_{11}}{p_1^*} & \frac{P_{12}}{p_2^*} & \dots & \frac{P_{1I}}{p_I^*} \\ \frac{P_{21}}{p_1^*} & \frac{P_{22}}{p_2^*} & \dots & \frac{P_{2I}}{p_I^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{P_{K1}}{p_1^*} & \frac{P_{K2}}{p_2^*} & \dots & \frac{P_{KI}}{p_I^*} \end{bmatrix}, \qquad \mathbf{Q}^* \triangleq \begin{bmatrix} \frac{Q_{11}}{q_1^*} & \frac{Q_{12}}{q_2^*} & \dots & \frac{Q_{1J}}{q_J^*} \\ \frac{Q_{21}}{q_1^*} & \frac{Q_{22}}{q_2^*} & \dots & \frac{Q_{2J}}{q_J^*} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{Q_{K1}}{q_1^*} & \frac{Q_{K2}}{q_2^*} & \dots & \frac{Q_{KJ}}{q_J^*} \end{bmatrix}.$$

Lemma 6 (Blackwell and Girshick (1954)). *Experiment* $\mathbf{P}$ *Blackwell dominates experiment* $\mathbf{Q}$ *if and only if there exists a* $J \times I$ *Markov matrix* $\mathbf{C} = \{C_{ji}\}$ *such that*

$$\mathbf{P}^*\mathbf{C}^T = \mathbf{Q}^*; \tag{4}$$

and

$$\mathbf{q}^*\mathbf{C} = \mathbf{p}^*. \tag{5}$$

In Lemma 6, equation (4) implies that each column of $\mathbf{Q}^*$ is a convex linear combination of the columns of $\mathbf{P}^*$, given that $\mathbf{C}$ is a Markov matrix.

3.2 Two Examples

Example 1. *Consider binary states* $\Omega = \{\omega_1, \omega_2\}$ *and two signals:* $\pi_1 = \{a, b\}$ *and* $\pi_2 = \{c\}$.

As illustrated in Figure 9, the signal realizations are as follows: $a = (\omega_1, [0, \alpha]) \cup (\omega_2, [\beta, 1])$*,* $b = (\omega_1, [\alpha, 1]) \cup (\omega_2, [0, \beta])$*, and* $c = (\omega_1, [0, 1]) \cup (\omega_2, [0, 1])$*, where* $\alpha, \beta \in [0, 1]$.

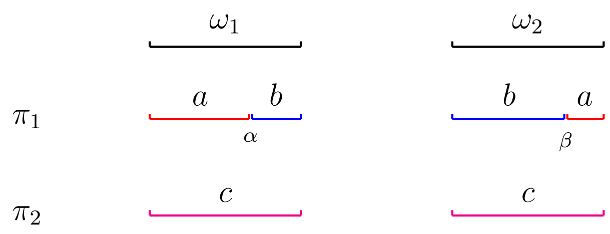

Figure 9: Example 1.

Proposition 6. *Given the set* $\Pi_0 = \{\pi_1, \pi_2\}$ *as defined in Example 1, there exists a menu that guarantees the first-best revenue* $\overline{R}$ *for the seller.*

According to Proposition 6, in a model with $\Pi_0 = \{\pi_1, \pi_2\}$ as defined in Example 1, the first-best implementation is achievable. This result relies on the fact that the WTP condition holds for the given set Π_0.

The signal $\pi_1^{ms} = \{e, f\}$, as shown in Figure 10, is a minimal supplement of π_1, as established in Lemma 3. Based on this, the signal π_A can be constructed from $\pi_1 \vee \pi_1^{ms}$ and π_2, while the signal π_B can be constructed from $\pi_2 \vee \pi_1^{ms}$ and π_1. We can prove that π_A Blackwell dominates π_B, which is a sufficient condition for the WTP condition to hold. Then, by Lemma 4, it is equivalent to prove that the experiment induced by π_A Blackwell dominates the experiment induced by π_B.

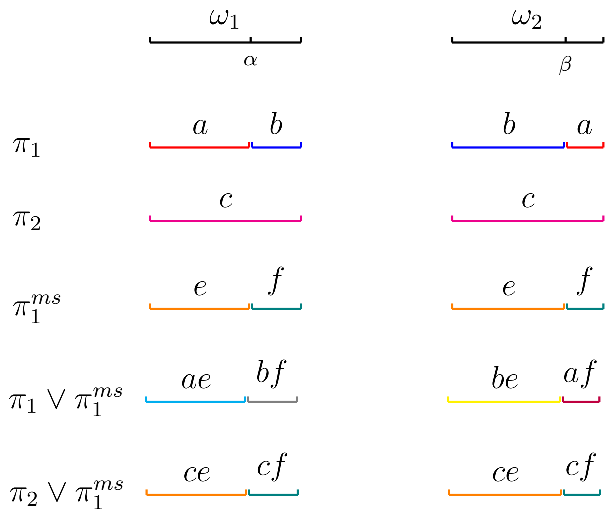

Figure 10: The minimal supplement in Example 1.

The experiment induced by π_A can be represented by a 2×5 Markov matrix $\mathbf{P}$, where

$$\mathbf{P} = \begin{array}{c} \\ \omega_1 \\ \omega_2 \end{array} \begin{array}{c} \begin{array}{ccccc} ae & bf & be & af & c \end{array} \\ \begin{bmatrix} \frac{\alpha}{2} & \frac{1-\alpha}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & \frac{\beta}{2} & \frac{1-\beta}{2} & \frac{1}{2} \end{bmatrix} \end{array}. \tag{6}$$

Similarly, the experiment induced by π_B can be represented by a 2×4 Markov matrix $\mathbf{Q}$, where

$$\mathbf{Q} = \begin{array}{c} \\ \omega_1 \\ \omega_2 \end{array} \begin{array}{c} \begin{array}{cccc} ce & cf & a & b \end{array} \\ \begin{bmatrix} \frac{\alpha}{2} & \frac{1-\alpha}{2} & \frac{\alpha}{2} & \frac{1-\alpha}{2} \\ \frac{\beta}{2} & \frac{1-\beta}{2} & \frac{1-\beta}{2} & \frac{\beta}{2} \end{bmatrix} \end{array}. \tag{7}$$

We can establish that experiment $\mathbf{P}$ in (6) Blackwell dominates experiment $\mathbf{Q}$ in (7) by constructing a 4×5 Markov matrix $\mathbf{C}$ that satisfies both (4) and (5), as demonstrated in the proof of Proposition 6.

Example 2. *Consider binary states* $\Omega = \{\omega_1, \omega_2\}$ *and two signals:* $\pi_1 = \{a, b\}$ *and* $\pi_2 = \{c, d\}$.

As illustrated in Figure 11, the signal realizations are as follows: $a = (\omega_1, [0, \alpha_1]) \cup (\omega_2, [\beta_1, 1])$, $b = (\omega_1, [\alpha_1, 1]) \cup (\omega_2, [0, \beta_1])$, $c = (\omega_1, [0, \alpha_2]) \cup (\omega_2, [\beta_2, 1])$, *and* $d = (\omega_1, [\alpha_2, 1]) \cup$

$(\omega_2, [0, \beta_2])$, *where* $\alpha_1, \alpha_2, \beta_1, \beta_2 \in [0, 1]$. *Assume* $\alpha_2 + \beta_2 \geq 1$, $\alpha_1 \geq \alpha_2$, *and* $\beta_1 \geq \beta_2$.

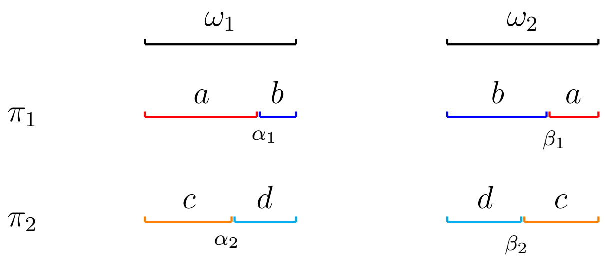

Figure 11: Example 2.

The assumptions $\alpha_2 + \beta_2 \geq 1$, $\alpha_1 \geq \alpha_2$, and $\beta_1 \geq \beta_2$ ensure that $v(\pi_1) \geq v(\pi_2)$, as established in Lemma 7.

Lemma 7. *In Example 2,* π_1 *Blackwell dominates* π_2*, which ensures that* $v(\pi_1) \geq v(\pi_2)$.

Proposition 7. *Given the set* $\Pi_0 = \{\pi_1, \pi_2\}$ *defined in Example 2, there exists a menu that guarantees the first-best revenue* $\overline{R}$ *for the seller.*

According to Proposition 7, in a model with $\Pi_0 = \{\pi_1, \pi_2\}$ as defined in Example 2, the first-best implementation is achievable. Detailed proofs can be found in the Appendix.

4 Conclusion

This paper discusses the optimal design and pricing of information products in a data market where a monopolist seller sells information to a privately informed data buyer. The buyer's initial information determines his willingness to pay for additional information, leading different buyer types to rank the same information product differently. This horizontal differentiation allows the seller to extract more surplus. By strategically leveraging data correlations, the seller can fully extract the first-best surplus by offering customized supplemental information products tailored to each buyer type. This full surplus extraction is socially efficient: (i) the data seller achieves first-degree price discrimination by charging the data buyer a price equal to his willingness to pay, and (ii) the data buyer attains complete information by combining his initial information and the purchased data, resolving uncertainty and making well-informed decisions in the downstream market.

These findings highlight that in data markets, sellers with significant market power can generate substantial revenue. While the seller extracts the full surplus, the outcome remains socially efficient. From a regulatory perspective, the key concern lies in the distribution of surplus between market participants, raising important considerations for policymakers evaluating market fairness.

APPENDIX

PROOF OF LEMMA 1.

Given the state space $\Omega = \{\omega_1, ..., \omega_K\}$, the set of beliefs is denoted by

$$\Delta(\Omega) = \{\mu \in \mathbb{R}_+^K \mid \sum_{k=1}^{K} \mu(\omega_k) = 1\}.$$

Note that $\Delta(\Omega)$ is convex, since for any $\mu, \mu' \in \Delta(\Omega)$ and $\theta \in [0, 1]$, we have $\theta\mu + (1-\theta)\mu' \in \Delta(\Omega)$.

By definition, for any $\mu \in \Delta(\Omega)$, we have

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

which implies that for any $a \in A$,

$$\sum_{\omega \in \Omega} \mu(\omega) u(a, \omega) \leq \hat{v}(\mu).$$

Similarly, for $\mu' \in \Delta(\Omega)$ and any $a \in A$, we have

$$\sum_{\omega \in \Omega} \mu'(\omega) u(a, \omega) \leq \hat{v}(\mu').$$

Thus, for any $a \in A$ and $\theta \in [0, 1]$, we obtain

$$\theta \sum_{\omega \in \Omega} \mu(\omega) u(a, \omega) + (1 - \theta) \sum_{\omega \in \Omega} \mu'(\omega) u(a, \omega) \leq \theta \hat{v}(\mu) + (1 - \theta)\hat{v}(\mu').$$

Since the above inequality holds for any $a \in A$, we have

$$\max_{a \in A} \left(\theta \sum_{\omega \in \Omega} \mu(\omega) u(a, \omega) + (1 - \theta) \sum_{\omega \in \Omega} \mu'(\omega) u(a, \omega) \right) \leq \theta \hat{v}(\mu) + (1 - \theta)\hat{v}(\mu'),$$

which can be rewritten as

$$\max_{a\in A}\left(\sum_{\omega\in\Omega}u(a,\omega)(\theta\mu(\omega)+(1-\theta)\mu'(\omega))\right)\leq\theta\hat{v}(\mu)+(1-\theta)\hat{v}(\mu').$$

Therefore, for any $\mu,\mu'\in\Delta(\Omega)$ and $\theta\in[0,1]$,

$$\hat{v}(\theta\mu+(1-\theta)\mu')\leq\theta\hat{v}(\mu)+(1-\theta)\hat{v}(\mu').$$

This completes the proof, demonstrating that $\hat{v}(\mu)$ is convex over the set $\Delta(\Omega)$. □

***PROOF OF LEMMA** 2.*

Let $F\in\Delta(\Delta(\Omega))$ and F' denote two distributions of beliefs with associated $\Delta(\Omega)$-based random variables Y and Y'.

Suppose F is a mean-preserving spread of F', i.e. $\mathbb{E}[Y\mid Y']=Y'$.

Then for any convex function $h:\Delta(\Omega)\to\mathbb{R}$, we have

$$\mathbb{E}[h(Y)]=\mathbb{E}[\mathbb{E}[h(Y)\mid Y']]\geq\mathbb{E}[h(\mathbb{E}[Y\mid Y'])]=\mathbb{E}[h(Y')],$$

where the first equality comes from the law of iterated expectations, the inequality comes from Jensen's inequality, and the last equality follows from F being a mean-preserving spread of F'.

Now, consider an arbitrary signal $\pi\in\Pi$. The value of the signal π is given by

$$v(\pi)=\mathbb{E}_{\mu\sim\pi}[\hat{v}(\mu)],$$

where $\hat{v}:\Delta(\Omega)\to\mathbb{R}$ is convex, as established in Lemma 1.

Since the distribution of beliefs induced by π is a mean-preserving spread of the prior μ_0, we have

$$v(\pi)=\mathbb{E}_{\mu\sim\pi}[\hat{v}(\mu)]\geq\hat{v}(\mu_0)=\underline{v}.$$

Similarly, since the distribution of degenerate beliefs induced by $\overline{\pi}$ is a mean-preserving spread of the distribution of beliefs induced by π, we have

$$\overline{v} = v(\overline{\pi}) = \mathbb{E}_{\mu\sim\overline{\pi}}[\hat{v}(\mu)] \geq \mathbb{E}_{\mu\sim\pi}[\hat{v}(\mu)] = v(\pi).$$

Therefore, for any $\pi \in \Pi$, we have $\underline{v} \leq v(\pi) \leq \overline{v}$. $\square$

PROOF OF PROPOSITION 1.

(i) Prove by contradiction. Suppose $t_2 < t_1$ in an optimal menu $M = \{(\tilde{\pi}_i, t_i)\}_{i\in\{1,2\}}$. Then, the data seller's expected revenue, given by $(1-\delta)t_1 + \delta t_2$, would be strictly less than $\overline{v} - v(\pi_1)$, since t_1 is at most $\overline{v} - v(\pi_1)$. However, since the revenue is bounded from the bottom, the data seller could increase revenue by selling a fully informative signal to both types at a price of $\overline{v} - v(\pi_1)$, which contradicts the optimality of menu M.

(ii) Suppose that $\tilde{\pi}_2$ is not a supplement of π_2 in an optimal menu $M = \{(\tilde{\pi}_i, t_i)\}_{i\in\{1,2\}}$. Then, we have $v(\pi_2 \vee \tilde{\pi}_2) < v(\overline{\pi}) = \overline{v}$. We will show that the data seller could increase revenue by choosing an alternative menu M', which is both individually rational and incentive-compatible.

The menu M' is constructed by replacing $\tilde{\pi}_2$ with π', and charging a strictly higher price of $t_2 + \epsilon$ for it, where π' is a supplement of signal π_2: $(\pi_2 \vee \pi') \sim \overline{\pi}$ and $\epsilon = \overline{v} - v(\pi_2 \vee \tilde{\pi}_2) > 0$. If type π_1 strictly prefers to purchase signal π' rather than $\tilde{\pi}_1$, then we can complete the construction of menu M' by replacing $\tilde{\pi}_1$ with π' and charging a price of $t_2 + \epsilon$, which will be higher than t_1 because $t_2 \geq t_1$ from (i) of Proposition 1. If, however, type π_1 does not prefer to purchase π' rather than $\tilde{\pi}_1$, then we finalize menu M' by keeping $\tilde{\pi}_1$ and t_1 unchanged.

(iii) Suppose that IR_1 is slack in an optimal menu $M = \{(\tilde{\pi}_i, t_i)\}_{i\in\{1,2\}}$. Then,

$$v(\pi_1 \vee \tilde{\pi}_1) - t_1 > v(\pi_1).$$

IR_2 must bind; otherwise, the data seller could increase revenue by increasing both t_1 and t_2 while keeping $t_1 - t_2$ constant.

Additionally, IC_1 must bind; otherwise, the data seller could increase revenue by increasing t_1. Since $(\pi_2 \vee \tilde{\pi}_2) \sim \overline{\pi}$ from (ii) of Proposition 1, IR_2 and IC_1 determine the optimal transfers, where $t_2 = \overline{v} - v(\pi_2)$ and $t_1 = v(\pi_1 \vee \tilde{\pi}_1) - v(\pi_1 \vee \tilde{\pi}_2) + \overline{v} - v(\pi_2)$.

Hence, we have

$$v(\pi_1 \vee \tilde{\pi}_1) - t_1 = v(\pi_1 \vee \tilde{\pi}_2) - \overline{v} + v(\pi_2).$$

Since $v(\pi_1 \vee \tilde{\pi}_2) \leq \overline{v}$, it follows that:

$$v(\pi_1 \vee \tilde{\pi}_1) - t_1 \leq v(\pi_2) \leq v(\pi_1),$$

which leads to a contradiction.

□

PROOF OF PROPOSITION 2.

Given $\Pi_0 = \{\pi_1, \pi_2\}$ with $v(\pi_1) > v(\pi_2)$. Assume that the WTP condition holds, then there exists a signal π_1^* satisfying (2) and (3).

(i) Proof by contradiction. Consider an optimal menu $M = \{(\tilde{\pi}_i, t_i)\}_{i \in \{1,2\}}$ and suppose that IR_2 is slack:

$$t_2 < v(\pi_2 \vee \tilde{\pi}_2) - v(\pi_2). \tag{A1}$$

Then, IC_2 must bind:

$$t_2 = v(\pi_2 \vee \tilde{\pi}_2) - v(\pi_2 \vee \tilde{\pi}_1) + t_1; \tag{A2}$$

otherwise, the data seller could improve revenue by increasing t_2.

From (iii) in Proposition 1,

$$t_1 = v(\pi_1 \vee \tilde{\pi}_1) - v(\pi_1). \tag{A3}$$

Then rewrite (A2) into

$$t_2 = v(\pi_2 \vee \tilde{\pi}_2) - v(\pi_2 \vee \tilde{\pi}_1) + v(\pi_1 \vee \tilde{\pi}_1) - v(\pi_1). \tag{A4}$$

Combining (A1) and (A4), we have

$$v(\pi_1 \vee \tilde{\pi}_1) - v(\pi_1) < v(\pi_2 \vee \tilde{\pi}_1) - v(\pi_2). \tag{A5}$$

The revenue from menu M is $(1-\delta)t_1 + \delta t_2$, with t_1 and t_2 in (A3) and (A4), respectively. We will now demonstrate that the data seller can improve revenue by choosing an alternative menu $M' = \{(\tilde{\pi}_i', t_i')\}_{i\in\{1,2\}}$ of differentiated signals, where

$$\begin{aligned} \tilde{\pi}_1' &\triangleq \pi_1^*, & t_1' &\triangleq \overline{v} - v(\pi_1), \\ \tilde{\pi}_2' &\triangleq \overline{\pi}, & t_2' &\triangleq \overline{v} - v(\pi_2). \end{aligned}$$

To verify the feasibility of menu M', first check the incentive-compatibility constraints. Both signal π_1^* and $\overline{\pi}$ provide the same additional value to type π_1, as $v(\pi_1 \vee \pi_1^*) = v(\overline{\pi}) = v(\pi_1 \vee \overline{\pi})$. However, since $v(\pi_1) > v(\pi_2)$, π_1^* is strictly cheaper than $\overline{\pi}$: $t_1' < t_2'$. Therefore,

$$v(\pi_1 \vee \pi_1^*) - t_1' > v(\pi_1 \vee \overline{\pi}) - t_2',$$

which implies that type π_1 strictly prefers to purchase signal π_1^* rather than $\overline{\pi}$.

From (2) and (3), we know

$$v(\pi_2 \vee \overline{\pi}) - t_2' \geq v(\pi_2 \vee \pi_1^*) - t_1',$$

which implies that type π_2 prefers to purchase signal $\overline{\pi}$ rather than π_1^*. It is straightforward to verify that menu M' satisfies the individual-rationality constraints, as both types receive their reservation utilities.

We now discuss the profitability of menu M'. The value of signals is bounded from the above: $v(\pi_1 \vee \tilde{\pi}_1) \leq \overline{v}$. Thus,

$$t_1 \leq t_1'.$$

From (ii) in Proposition 1, $v(\pi_2 \vee \tilde{\pi}_2) = v(\overline{\pi}) = \overline{v}$. Then reduce (A4) into

$$t_2 = \overline{v} - v(\pi_2 \vee \tilde{\pi}_1) + v(\pi_1 \vee \tilde{\pi}_1) - v(\pi_1),$$

which, combined with (A5), implies

$$t_2 < t_2'.$$

Thus, for any $\delta \in (0,1)$, we have

$$(1-\delta)t_1 + \delta t_2 < (1-\delta)t_1' + \delta t_2',$$

which implies that the expected revenue from menu M is strictly lower than that from menu M'. This leads to a contradiction.

(ii) Consider the menu M' constructed in the proof of (i) of Proposition 2. It is easy to verify that this menu guarantees the first-best revenue $\overline{R}$ for the data seller, where $\overline{R} = (1-\delta)(\overline{v} - v(\pi_1)) + \delta(\overline{v} - v(\pi_2))$.

□

PROOF OF PROPOSITION 3.

Given $\Pi_0 = \{\pi_1, \pi_2\}$ with $v(\pi_1) > v(\pi_2)$. Assume that the WTP condition fails, then for any supplement of signal π_1, denoted by π_1^{su}, we must have

$$v(\pi_1 \vee \pi_1^{su}) - v(\pi_1) < v(\pi_2 \vee \pi_1^{su}) - v(\pi_2). \tag{A6}$$

Consider an optimal menu $M = \{(\tilde{\pi}_i, t_i)\}_{i\in\{1,2\}}$ and suppose that IC_2 is slack:

$$t_2 < v(\pi_2 \vee \tilde{\pi}_2) - v(\pi_2 \vee \tilde{\pi}_1) + t_1. \tag{A7}$$

From (ii) in Proposition 1, we have $v(\pi_2 \vee \tilde{\pi}_2) = v(\overline{\pi}) = \overline{v}$.

From (iii) in Proposition 1, we have $t_1 = v(\pi_1 \vee \tilde{\pi}_1) - v(\pi_1)$. Then rewrite (A7) into

$$t_2 < \overline{v} - v(\pi_2 \vee \tilde{\pi}_1) + v(\pi_1 \vee \tilde{\pi}_1) - v(\pi_1). \tag{A8}$$

IR_2 must bind:

$$t_2 = v(\pi_2 \vee \tilde{\pi}_2) - v(\pi_2) = \overline{v} - v(\pi_2); \tag{A9}$$

otherwise, the data seller could improve revenue by increasing t_2.

Substituting (A9) into (A8), we get

$$v(\pi_1 \vee \tilde{\pi}_1) - v(\pi_1) > v(\pi_2 \vee \tilde{\pi}_1) - v(\pi_2).$$

Then from (A6), signal $\tilde{\pi}_1$ cannot be a supplement of π_1, which implies that $v(\pi_1 \vee \tilde{\pi}_1) < \overline{v}$.

Given that $\tilde{\pi}_1$ is not a supplement of π_1 and IC_2 is slack, the data seller can always improve revenue by charging a higher price for adding even a small piece of information to $\tilde{\pi}_1$, which is valuable to the decision maker. This leads to a contradiction. $\square$

PROOF OF PROPOSITION 6.

Consider the experiment $\mathbf{P}$ in (6), induced by signal π_A, and the experiment $\mathbf{Q}$ in (7), induced by signal π_B. We will prove that $\mathbf{P}$ Blackwell dominates $\mathbf{Q}$.

First, we focus on the cases where $0 < \alpha < 1$ and $0 < \beta < 1$. which ensure that all elements in matrices $\mathbf{P}$ and $\mathbf{Q}$ are positive.

By definition,

$$\begin{aligned}
\mathbf{p}^* &= \begin{bmatrix} \frac{\alpha}{2} & \frac{1-\alpha}{2} & \frac{\beta}{2} & \frac{1-\beta}{2} & 1 \end{bmatrix}, \\
\mathbf{q}^* &= \begin{bmatrix} \frac{\alpha+\beta}{2} & \frac{2-\alpha-\beta}{2} & \frac{\alpha+1-\beta}{2} & \frac{1-\alpha+\beta}{2} \end{bmatrix}, \\
\mathbf{P}^* &= \begin{bmatrix} 1 & 1 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 1 & \frac{1}{2} \end{bmatrix}, \\
\mathbf{Q}^* &= \begin{bmatrix} \frac{\alpha}{\alpha+\beta} & \frac{1-\alpha}{2-\alpha-\beta} & \frac{\alpha}{\alpha+1-\beta} & \frac{1-\alpha}{1-\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{1-\beta}{2-\alpha-\beta} & \frac{1-\beta}{\alpha+1-\beta} & \frac{\beta}{1-\alpha+\beta} \end{bmatrix}.
\end{aligned}$$

There exists a 4×5 Markov matrix $\mathbf{C}$ such that $\mathbf{P}^*\mathbf{C}^T = \mathbf{Q}^*$ and $\mathbf{q}^*\mathbf{C} = \mathbf{p}^*$.

If $\alpha \geq \beta$, let

$$\mathbf{C} = \begin{bmatrix} \frac{(\alpha-\beta)\alpha}{\alpha+\beta} & \frac{(\alpha-\beta)(1-\alpha)}{\alpha+\beta} & 0 & 0 & \frac{2\beta}{\alpha+\beta} \\ 0 & 0 & \frac{(\alpha-\beta)\beta}{2-\alpha-\beta} & \frac{(\alpha-\beta)(1-\beta)}{2-\alpha-\beta} & \frac{2(1-\alpha)}{2-\alpha-\beta} \\ \frac{\alpha\beta}{\alpha+1-\beta} & \frac{\beta(1-\alpha)}{\alpha+1-\beta} & \frac{(1-\alpha)\beta}{\alpha+1-\beta} & \frac{(1-\alpha)(1-\beta)}{\alpha+1-\beta} & \frac{2(\alpha-\beta)}{\alpha+1-\beta} \\ \frac{(1-\alpha)\alpha}{1-\alpha+\beta} & \frac{(1-\alpha)^2}{1-\alpha+\beta} & \frac{\beta^2}{1-\alpha+\beta} & \frac{\beta(1-\beta)}{1-\alpha+\beta} & 0 \end{bmatrix};$$

and if $\beta > \alpha$, let

$$\mathbf{C} = \begin{bmatrix} 0 & 0 & \frac{(\beta-\alpha)\beta}{\alpha+\beta} & \frac{(\beta-\alpha)(1-\beta)}{\alpha+\beta} & \frac{2\alpha}{\alpha+\beta} \\ \frac{(\beta-\alpha)\alpha}{2-\alpha-\beta} & \frac{(\beta-\alpha)(1-\alpha)}{2-\alpha-\beta} & 0 & 0 & \frac{2(1-\beta)}{2-\alpha-\beta} \\ \frac{\alpha^2}{\alpha+1-\beta} & \frac{\alpha(1-\alpha)}{\alpha+1-\beta} & \frac{(1-\beta)\beta}{\alpha+1-\beta} & \frac{(1-\beta)^2}{\alpha+1-\beta} & 0 \\ \frac{(1-\beta)\alpha}{1-\alpha+\beta} & \frac{(1-\beta)(1-\alpha)}{1-\alpha+\beta} & \frac{\alpha\beta}{1-\alpha+\beta} & \frac{\alpha(1-\beta)}{1-\alpha+\beta} & \frac{2(\beta-\alpha)}{1-\alpha+\beta} \end{bmatrix}.$$

Since the sum of each row in matrix $\mathbf{C}$ is 1, to verify that $\mathbf{C}$ is a Markov matrix, we only need to ensure that all of its elements are non-negative. The non-negativity condition holds in both cases: when $\alpha \geq \beta$ and when $\beta > \alpha$.

According to Lemma 6, $\mathbf{P}$ Blackwell dominates $\mathbf{Q}$. Consequently, by Lemma 4, π_A Blackwell dominates π_B. Therefore, by Proposition 5, the WTP condition must hold.

Given this, Proposition 2 guarantees that the data seller can achieve the first-best revenue $\overline{R}$ with the following menu $M = \{(\tilde{\pi}_i, t_i)\}_{i\in\{1,2\}}$, where

$$\tilde{\pi}_1 \triangleq \pi_1^{ms}, \qquad t_1 \triangleq \overline{v} - v(\pi_1),$$

$$\tilde{\pi}_2 \triangleq \overline{\pi}, \qquad t_2 \triangleq \overline{v} - v(\pi_2).$$

In the cases where $\alpha = 0, 0 < \beta < 1$; $\alpha = 1, 0 < \beta < 1$; $0 < \alpha < 1, \beta = 0$; or $0 < \alpha < 1, \beta = 1$, the matrix $\mathbf{P}$ reduces to a 2×4 matrix, and the proof follows analogously from the previous discussion.

In the cases where $\alpha = \beta = 0$ or $\alpha = \beta = 1$, the signal π_1 becomes $\overline{\pi}$. Thus, type π_1 already has full information about the state: $v(\pi_1) = \overline{v}$. The first-best revenue in this case is given by $\overline{R} = \delta(\overline{v} - v(\pi_2))$. The data seller can achieve $\overline{R}$ by selling a fully informative signal to type π_2 at a price of $\overline{v} - v(\pi_2)$.

In the cases where $\alpha = 0, \beta = 1$ or $\alpha = 1, \beta = 0$, the signals π_1 and π_2 are uninformative about the state: $v(\pi_1) = \underline{v} = v(\pi_2)$. Then, the data seller can achieve the first-best revenue, as established in Claim 1. $\square$

PROOF OF LEMMA 7.

Under assumptions $\alpha_2 + \beta_2 \geq 1$, $\alpha_1 \geq \alpha_2$, and $\beta_1 \geq \beta_2$, we know that $\alpha_1 + \beta_1 \geq 1$.

If $\alpha_1 + \beta_1 = 1$, it must be the case that $\alpha_2 + \beta_2 = 1$. In this scenario, the only possibility is $\alpha_1 = \alpha_2$ and $\beta_1 = \beta_2$, implying that signals π_1 and π_2 are identical. Then the value of π_1 is equal to the value of π_2 in any decision-making problem: $v(\pi_1) = v(\pi_2)$.

We will now consider the cases where $\alpha_1 + \beta_1 > 1$.

The experiment induced by π_1 can be represented by a 2×2 Markov matrix $\mathbf{P}$, where

$$\mathbf{P} = \begin{bmatrix} \alpha_1 & 1-\alpha_1 \\ 1-\beta_1 & \beta_1 \end{bmatrix}.$$

Similarly, the experiment induced by π_2 can be represented by a 2×2 Markov matrix $\mathbf{Q}$, where

$$\mathbf{Q} = \begin{bmatrix} \alpha_2 & 1-\alpha_2 \\ 1-\beta_2 & \beta_2 \end{bmatrix}.$$

There exists a 2×2 Markov matrix $\mathbf{D}$ such that $\mathbf{PD} = \mathbf{Q}$, where

$$\mathbf{D} = \begin{bmatrix} \dfrac{\alpha_2\beta_1 - (1-\alpha_1)(1-\beta_2)}{\alpha_1+\beta_1-1} & \dfrac{(1-\alpha_2)\beta_1 - (1-\alpha_1)\beta_2}{\alpha_1+\beta_1-1} \\ \dfrac{\alpha_1(1-\beta_2) - \alpha_2(1-\beta_1)}{\alpha_1+\beta_1-1} & \dfrac{\alpha_1\beta_2 - (1-\alpha_2)(1-\beta_1)}{\alpha_1+\beta_1-1} \end{bmatrix}.$$

To verify that $\mathbf{D}$ is a Markov matrix, note that the sum of each row in $\mathbf{D}$ is equal to 1.

Additionally, given that $\alpha_2+\beta_2 \geq 1$, $0 \leq \alpha_2 \leq \alpha_1 \leq 1$, and $0 \leq \beta_2 \leq \beta_1 \leq 1$, we have

$$\alpha_2\beta_1 \geq \alpha_2\beta_2 \geq (1-\alpha_2)(1-\beta_2) \geq (1-\alpha_1)(1-\beta_2),$$

$$\alpha_1\beta_2 \geq \alpha_2\beta_2 \geq (1-\alpha_2)(1-\beta_2) \geq (1-\alpha_2)(1-\beta_1),$$

$$(1-\alpha_2)\beta_1 \geq (1-\alpha_1)\beta_2,$$

$$\alpha_1(1-\beta_2) \geq \alpha_2(1-\beta_1),$$

which imply that all elements in $\mathbf{D}$ is non-negative.

Thus, by Lemma 5, the experiment $\mathbf{P}$ Blackwell dominates the experiment $\mathbf{Q}$. Then according to Lemma 4, we know that the signal π_1 Blackwell dominates the signal π_2, which ensures that $v(\pi_1) \geq v(\pi_2)$. □

PROOF OF PROPOSITION 7.

Given the set $\Pi_0 = \{\pi_1, \pi_2\}$ of signals as defined in Example 2. By Lemma 7, we have $v(\pi_1) \geq v(\pi_2)$. If $v(\pi_1) = v(\pi_2)$, then the data seller can achieve the first-best revenue, as established in Claim 1.

In the following cases, the signal π_2 becomes $\underline{\pi}$, and the data seller can obtain the first-best revenue, as established in Proposition 6:

(i) If $\alpha_2 = 0$, we must have $\beta_2 = 1$ since $\alpha_2 + \beta_2 \geq 1$;

(ii) If $\beta_2 = 0$, we must have $\alpha_2 = 1$ since $\alpha_2 + \beta_2 \geq 1$;

(iii) If $\alpha_1 = 0$, we must have $\alpha_2 = 0$ since $\alpha_2 \leq \alpha_1$;

(iv) If $\beta_1 = 0$, we must have $\beta_2 = 0$ since $\beta_2 \leq \beta_1$.

We will now consider the remaining cases where $\alpha_2 + 1 - \beta_2 > 0$, $1 - \alpha_2 + \beta_2 > 0$, $\alpha_2 + \beta_1 - \beta_2 > 0$, $\alpha_1 - \alpha_2 + \beta_2 > 0$, $\alpha_1 + 1 - \beta_1 > 0$, and $1 - \alpha_1 + \beta_1 > 0$. These conditions ensure that the denominator of each element in the matrices $\mathbf{P}^*$ and $\mathbf{Q}^*$, defined below, is positive.

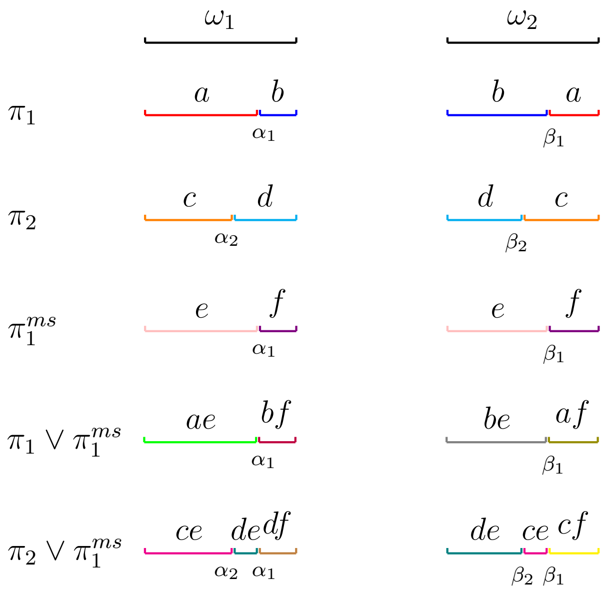

Figure 12: The minimal supplement in Example 2.

As illustrated in Figure 12, the signal π_1 has a unique minimal supplement $\pi_1^{ms} = \{e, f\}$. We will prove that the WTP condition holds for the set $\Pi_0 = \{\pi_1, \pi_2\}$ of signals and thus the data seller can receive the first-best revenue from the following menu $M = \{(\tilde{\pi}_i, t_i)\}_{i\in\{1,2\}}$, where

$$\tilde{\pi}_1 \triangleq \pi_1^{ms}, \qquad t_1 \triangleq \overline{v} - v(\pi_1),$$
$$\tilde{\pi}_2 \triangleq \overline{\pi}, \qquad t_2 \triangleq \overline{v} - v(\pi_2).$$

First, construct the signals π_A and π_B as defined.

The experiment induced by signal π_A can be represented by a 2×6 Markov matrix $\mathbf{P}$, where

$$\mathbf{P} = \begin{array}{c} \\ \omega_1 \\ \omega_2 \end{array} \begin{array}{c} \begin{array}{cccccc} ae & bf & be & af & c & d \end{array} \\ \begin{bmatrix} \frac{\alpha_1}{2} & \frac{1-\alpha_1}{2} & 0 & 0 & \frac{\alpha_2}{2} & \frac{1-\alpha_2}{2} \\ 0 & 0 & \frac{\beta_1}{2} & \frac{1-\beta_1}{2} & \frac{1-\beta_2}{2} & \frac{\beta_2}{2} \end{bmatrix} \end{array}.$$

Similarly, the experiment induced by signal π_B can be represented by a 2×6 Markov matrix $\mathbf{P}$, where

$$\mathbf{Q} = \begin{array}{c} \\ \omega_1 \\ \omega_2 \end{array} \begin{array}{c} \begin{array}{cccccc} ce & de & df & cf & a & b \end{array} \\ \begin{bmatrix} \frac{\alpha_2}{2} & \frac{\alpha_1-\alpha_2}{2} & \frac{1-\alpha_1}{2} & 0 & \frac{\alpha_1}{2} & \frac{1-\alpha_1}{2} \\ \frac{\beta_1-\beta_2}{2} & \frac{\beta_2}{2} & 0 & \frac{1-\beta_1}{2} & \frac{1-\beta_1}{2} & \frac{\beta_1}{2} \end{bmatrix} \end{array}.$$

By definition,

$$\mathbf{p}^* = \begin{bmatrix} \frac{\alpha_1}{2} & \frac{1-\alpha_1}{2} & \frac{\beta_1}{2} & \frac{1-\beta_1}{2} & \frac{\alpha_2+1-\beta_2}{2} & \frac{1-\alpha_2+\beta_2}{2} \end{bmatrix},$$

$$\mathbf{q}^* = \begin{bmatrix} \frac{\alpha_2+\beta_1-\beta_2}{2} & \frac{\alpha_1-\alpha_2+\beta_2}{2} & \frac{1-\alpha_1}{2} & \frac{1-\beta_1}{2} & \frac{\alpha_1+1-\beta_1}{2} & \frac{1-\alpha_1+\beta_1}{2} \end{bmatrix},$$

$$\mathbf{P}^* = \begin{bmatrix} 1 & 1 & 0 & 0 & \frac{\alpha_2}{\alpha_2+1-\beta_2} & \frac{1-\alpha_2}{1-\alpha_2+\beta_2} \\ 0 & 0 & 1 & 1 & \frac{1-\beta_2}{\alpha_2+1-\beta_2} & \frac{\beta_2}{1-\alpha_2+\beta_2} \end{bmatrix},$$

$$\mathbf{Q}^* = \begin{bmatrix} \frac{\alpha_2}{\alpha_2+\beta_1-\beta_2} & \frac{\alpha_1-\alpha_2}{\alpha_1-\alpha_2+\beta_2} & 1 & 0 & \frac{\alpha_1}{\alpha_1+1-\beta_1} & \frac{1-\alpha_1}{1-\alpha_1+\beta_1} \\ \frac{\beta_1-\beta_2}{\alpha_2+\beta_1-\beta_2} & \frac{\beta_2}{\alpha_1-\alpha_2+\beta_2} & 0 & 1 & \frac{1-\beta_1}{\alpha_1+1-\beta_1} & \frac{\beta_1}{1-\alpha_1+\beta_1} \end{bmatrix}.$$

There exists a 6×6 Markov matrix $\mathbf{C}$ such that $\mathbf{P}^*\mathbf{C}^T = \mathbf{Q}^*$ and $\mathbf{q}^*\mathbf{C} = \mathbf{p}^*$, where

$$\mathbf{C} = \begin{bmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \mathbf{c}_3 \\ \mathbf{c}_4 \\ \mathbf{c}_5 \\ \mathbf{c}_6 \end{bmatrix},$$

with

$$
\begin{aligned}
\mathbf{c}_1 &= \begin{bmatrix} \frac{(1-\beta_1)\alpha_2\alpha_1}{(\alpha_2+\beta_1-\beta_2)(1-\beta_2)} & \frac{(1-\beta_1)\alpha_2(1-\alpha_1)}{(\alpha_2+\beta_1-\beta_2)(1-\beta_2)} & 0 & 0 & \frac{(\beta_1-\beta_2)(\alpha_2+1-\beta_2)}{(\alpha_2+\beta_1-\beta_2)(1-\beta_2)} & 0 \end{bmatrix}, \\
\mathbf{c}_2 &= \begin{bmatrix} 0 & 0 & \frac{(1-\alpha_1)\beta_2\beta_1}{(\alpha_1-\alpha_2+\beta_2)(1-\alpha_2)} & \frac{(1-\alpha_1)\beta_2(1-\beta_1)}{(\alpha_1-\alpha_2+\beta_2)(1-\alpha_2)} & 0 & \frac{(\alpha_1-\alpha_2)(1-\alpha_2+\beta_2)}{(\alpha_1-\alpha_2+\beta_2)(1-\alpha_2)} \end{bmatrix}, \\
\mathbf{c}_3 &= \begin{bmatrix} \alpha_1 & 1-\alpha_1 & 0 & 0 & 0 & 0 \end{bmatrix}, \\
\mathbf{c}_4 &= \begin{bmatrix} 0 & 0 & \beta_1 & 1-\beta_1 & 0 & 0 \end{bmatrix}, \\
\mathbf{c}_5 &= \begin{bmatrix} \frac{((1-\beta_2)\alpha_1-(1-\beta_1)\alpha_2)\alpha_1}{(\alpha_1+1-\beta_1)(1-\beta_2)} & \frac{((1-\beta_2)\alpha_1-(1-\beta_1)\alpha_2)(1-\alpha_1)}{(\alpha_1+1-\beta_1)(1-\beta_2)} & 0 & 0 & \frac{(1-\beta_1)(\alpha_2+1-\beta_2)}{(\alpha_1+1-\beta_1)(1-\beta_2)} & 0 \end{bmatrix}, \\
\mathbf{c}_6 &= \begin{bmatrix} 0 & 0 & \frac{((1-\alpha_2)\beta_1-(1-\alpha_1)\beta_2)\beta_1}{(1-\alpha_1+\beta_1)(1-\alpha_2)} & \frac{((1-\alpha_2)\beta_1-(1-\alpha_1)\beta_2)(1-\beta_1)}{(1-\alpha_1+\beta_1)(1-\alpha_2)} & 0 & \frac{(1-\alpha_1)(1-\alpha_2+\beta_2)}{(1-\alpha_1+\beta_1)(1-\alpha_2)} \end{bmatrix}.
\end{aligned}
$$

Since the sum of each row in matrix $\mathbf{C}$ is 1, to verify that $\mathbf{C}$ is a Markov matrix, we only need to ensure that all of its elements are non-negative. Note that when $\alpha_2 = 1$, it implies $\alpha_1 = 1$, and similarly, when $\beta_2 = 1$, it implies $\beta_1 = 1$. In both cases, the corresponding matrix element becomes 0. Furthermore, under the assumptions $0 \leq \alpha_2 \leq \alpha_1 \leq 1$ and $0 \leq \beta_2 \leq \beta_1 \leq 1$, we establish that the following inequalities hold: $(1-\beta_2)\alpha_1 - (1-\beta_1)\alpha_2 \geq 0$ and $(1-\alpha_2)\beta_1 - (1-\alpha_1)\beta_2 \geq 0$. Given that the denominators of each element in the matrices $\mathbf{P}^*$ and $\mathbf{Q}^*$ are positive, we conclude that all elements of $\mathbf{C}$ are non-negative. Thus, matrix $\mathbf{C}$ satisfies the requirements of a Markov matrix.

Thus, by Lemma 6, $\mathbf{P}$ Blackwell dominates $\mathbf{Q}$. This implies that π_A Blackwell dominates π_B according to Lemma 4. Therefore, by Proposition 5, the WTP condition holds. □

References

Bergemann, D., A. Bonatti, and A. Smolin (2018). The design and price of information. *American Economic Review 108*(1), 1–48.

Blackwell, D. (1951). Comparison of experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Mtatistics and Probability*, Volume 2, pp. 93–103. University of California Press.

Blackwell, D. and M. A. Girshick (1954). *Theory of Games and Statistical Decisions*. Wiley.

Börgers, T., A. Hernando-Veciana, and D. Krähmer (2013). When are signals complements or substitutes? *Journal of Economic Theory 148*(1), 165–195.

Brooks, B., A. Frankel, and E. Kamenica (2024). Comparisons of signals. *American Economic Review 114*(9), 2981–3006.

Gentzkow, M. and E. Kamenica (2017). Bayesian persuasion with multiple senders and rich signal spaces. *Games and Economic Behavior 104*, 411–429.

Green, J. R. and N. L. Stokey (1978). Two representations of information structures and their comparisons. *Working paper*.

Green, J. R. and N. L. Stokey (2022). Two representations of information structures and their comparisons. *Decisions in Economics and Finance 45*(2), 541–547.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of operations research 6*(1), 58–73.

Rodríguez Olivera, R. (2024). Strategic incentives and the optimal sale of information. *American Economic Journal: Microeconomics 16*(2), 296–353.

Stole, L. and J.-C. Rochet (2003). The economics of multidimensional screening. Advances in Economic Theory.