Disclosure and Pricing of Attributes*

Alex Smolin

August 8, 2022

Abstract

A monopolist sells an object characterized by multiple attributes. A buyer can be one of many types, differing in their willingness to pay for each attribute. The seller can provide arbitrary attribute information in the form of a statistical experiment. To screen different types, the seller offers a menu of options that specify information prices, experiments, and object prices.

I characterize revenue-maximizing menus. All experiments belong to a class of linear disclosure rules. An optimal menu may be nondiscriminatory and qualitatively depends on the structure of buyer heterogeneity. The analysis highlights the importance of demand microstructure and the benefits of information control in trade settings.

Keywords: advertising, attributes, bilateral trade, demand transformation, information design, intermediaries, mechanism design, multidimensional disclosure, persuasion

JEL Codes: D11, D42, D82, D83, L15

*Toulouse School of Economics, alexey.v.smolin@gmail.com. I thank Dirk Bergemann and Daniel Krähmer for helpful conversations as well as Heski Bar-Isaac and Levent Celik for excellent discussions. I am grateful to the participants in research seminars at numerous institutions. Finally, I thank the coeditor, Nicola Persico, and the anonymous referees for their many productive suggestions. I acknowledge funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

1 Introduction

In many important markets, sellers have considerable control over the information available to their buyers. Business brokers can control the extent of the firm investigation and documentation that they supply, recruiting platforms can decide what parts of a job candidate's profile to reveal to employers, and producers can decide what features of their consumer goods to advertise. In all of these markets, the products (i.e., a business, meeting with a job candidate, and consumer good) are characterized by multiple attributes that appeal to different types of buyers. To maximize revenue, sellers need to understand what attribute information to provide, how to price their products, and whether and how to price the information provided. These questions require unification of information and mechanism design paradigms to allow for joint control over information and monetary incentives.

As a concrete example, consider the operation of Ziprecruiter.com, a major online recruiting platform. The platform facilitates matching job seekers with employers: employers subscribe to the platform to advertise their open vacancies and obtain access to a large database of resumes.[1] The recruitment market features substantial heterogeneity on both sides. Candidate profiles vary in many attributes, including work experience, education levels, and standardized test scores. The employers belong to distinct types, such as tech start-ups, chain stores, and government agencies. Naturally, different types of employers are looking for different attributes in their candidates.

Ziprecruiter.com has access to a large amount of data about prospective candidates and, by programming its algorithms, it can commit to coarsening the data or denying access to some attributes. Moreover, the platform can price both information, through upfront fees, and the decision to contact a job candidate, through contact fees. Indeed, the platform currently employs a nonlinear pricing scheme for subscriptions, which varies in the breadth of information provided and the ability to contact preferred candidates (Dubé and Misra (2019)). My goal is to study the trade-offs that the platform faces and to evaluate the allocation distortions introduced by its information control.

In this paper, I develop a framework to study information disclosure and the pricing of multiattribute products. I consider a monopolist seller who has an indivisible object for sale to a single buyer and aims to maximize her revenue. The object has several attributes, and the buyer is uncertain about their values. The buyer's valuation for the object is linear in attributes. The strengths of the preferences are the buyer's private information and constitute the buyer's type. The seller controls the pricing and attribute information available to the buyer.

[1] "ZipRecruiter Is Valued at $1.5 Billion in a Bet on AI Hiring," (Carville (2018)).

Both the object and the information about its attributes are valuable for a buyer, and I allow the seller to price them separately. In particular, the seller offers a menu of options that differ in their informativeness. Each option consists of an information price, paid upfront, attribute information, and a price for the object. The attribute information is modeled as an arbitrary statistical experiment informative about attributes. Information control enables price discrimination. By varying the information price, the experiment, and the object price, the seller can screen buyer types. I illustrate the qualitative features of multiattribute disclosure and pricing in Section 3.

I emphasize that the model formulation implies the seller is not privately informed and cannot condition the price on the outcome of the experiment. In other words, the seller can provide the buyer access to attribute information about the product but cannot condition the price on the outcome of this information. This kind of information provision is referred to in the economic literature as "private disclosure" and has been widely used.[2] One common motivation for the privacy of disclosure is that it may be hard for the seller to assess the impact of information on the buyer. As such, the analysis is directly applicable to the settings of intermediaries who organize sales of goods and services and in which the buyer has more expertise in assessing the product information than the intermediary. However, the analysis also provides insights into the more traditional buyer-seller settings, such as advertising, as long as the seller does not condition the price on the outcome of the information provided.

The general revenue-maximization problem features information design and multidimensional screening. As such, it entails two methodological challenges. First, the class of stochastic experiments is large since the underlying uncertainty covers a continuum of possible states, each having multiple dimensions. To understand the distortions driven by the information design, it is important to determine the structure of optimal experiments. Second, in the absence of a single-dimensional structure, it is not clear which incentive constraints are relevant for optimal design. This difficulty is further exacerbated by the distinct feature of information according to which different buyer types can respond differently to the same signal. I progress in both directions in turn.

In Section 4, I study the design of disclosure rules.[3] Providing disclosure serves two functions. First, it swings the buyer's expectations and may persuade him to purchase the object at a higher price. Second, providing several disclosure options may facilitate screening because different buyer types prefer learning about different aspects of the object. Theorem 1 shows that an optimal way to combine these two functions is through a specific class

[2]See, for example, Anderson and Renault (2006), Eső and Szentes (2007), and Li and Shi (2017).

[3]This terminology follows Rayo and Segal (2010) and is not to be confused with the voluntary disclosure of Dye (1985), studied for example by Koessler and Renault (2012).

of experiments—linear disclosure rule. A linear disclosure rule informs whether a linear combination of attributes is above or below a specified threshold. This disclosure guides the allocation and can be seen as informing the buyer about the valuation of a virtual type that can differ from the demanding type to account for incentive constraints.[4] Notably, this result requires no assumptions about the distributions of types or attributes and, as such, can be generalized to arbitrary valuation functions.

In addition, I establish an analog of the "no distortion at the top" property, which is common in mechanism design settings: If gains from trade are commonly known to be positive, then some type is provided with no information and always purchases the object.

In Section 5, I study optimal pricing mechanisms. In Theorem 2, I establish that if all buyer types value the same, always positive, attribute, then no information is optimally provided and the seller posts a single price for the object. The intuition behind this result lies in the product structure of the buyer's valuation. When all types value the same attribute, any disclosure realization simply scales their valuations and the corresponding demand curve. Even if the seller could condition the price on this realization, she would charge scaled prices, serve the same types, and obtain scaled revenue. By the martingale property of Bayesian expectations, the seller can obtain the same revenue by providing no information.[5]

The case of several attributes is qualitatively different because types can be differentiated not only vertically but also horizontally. Therefore, an optimal allocation may depend on attribute realizations: the seller should aim to allocate the object to types who value the realization the most. To guide the allocation, she should provide some attribute information.

I formalize this intuition in the setting in which each type values one of many independent attributes (Section 5.2) but the seller does not know which attribute or the strength of the preference. I show that an optimal menu features free-of-charge partial disclosure and no price discrimination. Information is not priced since the payment can be backloaded into the object price. Price dispersion is not profitable, because it implies that the seller could extract more surplus on some items by simultaneously lowering their object prices and changing their informational content. The optimal menu admits nondiscriminatory implementation—posting a single price for the object and informing the buyer whether the object is sufficiently good along each attribute.

In the limit case, in which there is only one type per attribute, in the optimal menu the seller effectively persuades each type to purchase the object at a fixed price separately by

[4] Chakraborty and Harbaugh (2010) use linear disclosure rules to construct informative equilibria in a multidimensional cheap talk game.

[5] This intuition leads in the right direction but does not consider discriminatory menus and information pricing. I formally complete the argument and confirm the result by building on single-dimensional mechanism-design machinery.

informing him whether his valuation is sufficiently high. In Section 5.3, I show that similar mechanisms are optimal in a broad range of settings, as long as the types can be seen as belonging to distinct cohorts—with low valuation correlation across them.

I conclude the analysis with a discussion in Section 6. First, I discuss what product information must be priced and when. Second, I emphasize that attribute information can rotate the demand curve locally and can justify attribute shrouding. Finally, I indicate how a multiattribute framework complements the existing disclosure frameworks, and I emphasize the importance of explicitly modeling demand microstructure, i.e., how the consumer valuation is formed, in the settings with information control.

Related Literature This paper is about information provision and pricing. One strand of the related literature focuses on nondiscriminatory mechanisms—in which a seller commits to a single disclosure rule. Lewis and Sappington (1994) introduce these mechanisms in a setting where a buyer has no prior information. They show that within a simple parameterized class, an optimal disclosure rule is extreme: either full or no disclosure. Bergemann and Pesendorfer (2007) further observe that if there is common knowledge of positive trade gains, then no disclosure dominates any other possible disclosure rule because it allows the seller to extract the full expected surplus.[6] Johnson and Myatt (2006) extend the analysis to settings in which the buyer has prior information. They focus on disclosure rules that correspond to the global rotation of a demand curve and show, once again, that extreme disclosure rules are optimal. My paper contributes to this literature by showing that if the product has several attributes, then a partial disclosure can dominate both full and no disclosure, even if there is common knowledge of positive trade gains (Sections 4.4, 6.3).

At the same time, when the buyer has private information, it is natural to study discriminatory mechanisms and how they can be used to screen buyer types. In an influential paper, Eső and Szentes (2007) study settings in which the attribute and the buyer's type enter the valuation "additively." In such settings, under certain distributional assumptions, the seller optimally provides full disclosure. However, Li and Shi (2017) show that the seller should withhold some information if the types represent private information about the object.[7] Section 6.3 contains a detailed discussion related to these two papers.

All of the works described above operate in single-dimensional settings. Under complete object information, when comparing any two objects, all buyer types agree on the ranking.

[6] See, however, Anderson and Renault (2006), who show that optimal disclosure rule is partial if the purchase is associated with search costs and the seller cannot commit to prices. Similarly, Dworczak (2020) shows that optimal disclosure rule can be partial in the presence of aftermarkets. See also Bar-Isaac, Caruana, and Cuñat (2010).

[7] Wei and Green (2019) show that withholding information may also be optimal if the information must be provided free of charge.

However, in practice, many products are multidimensional, with different attributes that appeal to different buyers. In this paper, I demonstrate that these settings can be successfully studied and lead to qualitatively different results. Despite the richness of the attribute space, optimal experiments belong to a tractable class of linear disclosure rules (Section 4.4). Optimal mechanisms feature partial disclosure but can be remarkably simple (Section 5.2).

This paper builds on several existing frameworks. The multiattribute buyer's valuation follows the characteristic model of Lancaster (1966). An unrestricted search for an optimal disclosure rule is a defining feature of the Bayesian persuasion literature (Rayo and Segal (2010), Kamenica and Gentzkow (2011)).[8] The screening analysis builds on the mechanism design machinery of Myerson (1981, 1982) and Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017). Finally, information design with screening and monetary transfers has already appeared in my previous work (Bergemann, Bonatti, and Smolin (2018)). There, the seller can price only information—the buyer's action is not contractable. In contrast, in this paper, the seller can price both services and as a result, in many settings, provides the information free of charge.

2 Model

A buyer (he) decides whether to buy a single indivisible object from a seller (she). The object has a finite number J of characteristics or attributes. The attribute values constitute an *attribute vector* $x = (x_1, \ldots, x_J) \in X = \mathbb{R}^J$. The buyer's preferences toward each attribute constitute the buyer's *type* $\theta = (\theta_1, \ldots, \theta_J) \in \Theta \subseteq \mathbb{R}^J$. The ex post buyer's valuation for the object is:[9]

$$v(\theta, x) = \theta \cdot x = \sum_{j=1}^{J} \theta_j x_j. \tag{1}$$

The buyer's utility is quasilinear in transfers. The seller maximizes her revenue.

Prior Information Attributes are distributed over X according to a cumulative distribution function G. The buyer and seller are symmetrically informed about the attributes. The type space Θ can be finite or infinite. The buyer's type is his privately known preferences, which are uncorrelated with attributes. From the seller's perspective, the types are distributed according to a cumulative distribution function F. Until Section 5, I do not impose any structural assumptions on the attribute and type distributions. The only technical

[8] Almost all of this literature studies one-dimensional settings. The works of Rayo and Segal (2010) and Dworczak and Martini (2019) provide elegant exceptions.

[9] As I discuss in Sections 4.4 and 5.3, this formulation can be generalized.

requirement is that the ex ante expectations of all attributes are finite.

Information Disclosure The seller can disclose attribute information to the buyer. This information is modeled as a statistical experiment $E = (S, \pi)$ that consists of a signal set S and a likelihood function:[10]

$$\pi : X \to \Delta(S). \tag{2}$$

The experiment can be arbitrarily informative about the attributes. That is, the experiment can provide no information, or no disclosure, $\underline{E} \triangleq (\underline{S}, \underline{\pi})$, with $\underline{S}$ being a singleton; it can fully reveal attributes, or provide full disclosure, $\overline{E} \triangleq \left(\overline{S}, \overline{\pi}\right)$, with $\overline{S} = X$ and $\overline{\pi}(x)$ placing a probability of 1 on $s = x$; or it can provide partial information.

Selling Mechanism For the environments in which the designer can control players' private information, it is not generally known what class of mechanisms one can look at without loss of generality. To make progress while avoiding trivialities, I follow Eső and Szentes (2007) and Li and Shi (2017) and focus on a class of menu mechanisms, so that the seller designs a menu of items, $i \in \mathcal{I}$:

$$M = (r(i), E(i), p(i))_{i \in \mathcal{I}}. \tag{3}$$

Each item consists of an experiment $E(i)$ and two tariffs $r(i) \geq 0$ and $p(i) \geq 0$. The first tariff captures the price of information—the upfront payment made to observe the signal of experiment $E(i)$, irrespective of whether the buyer decides to purchase the object later. The second tariff captures the price of the object, paid only if the trade occurs. Effectively, the menu is a collection of call options that differ in monetary terms and information disclosure, designed to screen different buyer types.[11] This class of mechanisms provides a natural and rich framework to study how information disclosure and pricing interact in design problems.

The timing is analogous to that of Courty and Li (2000) and is as follows. The seller posts a menu M. The attribute vector x and the buyer's type θ are realized. If the buyer refuses to participate, then the players obtain zero payoffs. Otherwise, the buyer chooses an item $i \in \mathcal{I}$ and pays the corresponding price $r(i)$. Next, the buyer observes a signal s from the experiment $E(i)$ and decides whether to buy the object at the price $p(i)$. Finally, the payoffs are realized. The timing is illustrated in Figure 1.

The timing implies that the seller commits to a menu before the realization of the attributes x and the type θ. The attributes x and signals s are not contractible, corresponding to the setting of "private disclosure" (Li and Shi (2017)).[12] Sales are deterministic—a pay-

[10] S can be any Polish space. Throughout the paper, all introduced functions are (Borel) measurable.

[11] This setting is equivalent to one in which the seller provides information for free but can charge the buyer for opting out from the consecutive sale.

[12] For instance, the buyer cannot claim a refund ex post. See Krähmer and Strausz (2015b), Heumann (2020) and Bergemann et al. (2020) for recent studies of ex post incentive constraints.

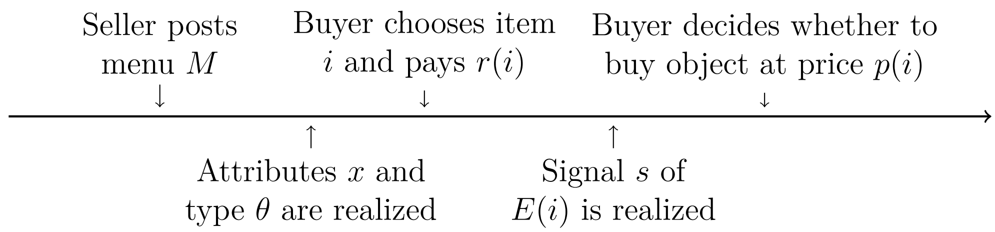

Figure 1: Timeline of the selling mechanism.

ment guarantees a transaction. Sequential interactions between players are excluded, so belief-elicitation schemes and scoring rules are not available.[13]

My goal is to characterize a revenue-maximizing menu, i.e., a menu upon which the seller cannot strictly improve by offering another menu.

3 Illustrative Example

I begin by illustrating the workings of disclosure and pricing through a simple example. There are two attributes that are uniformly and independently distributed, $J = 2$, $x_1 \sim U[0, 1]$, $x_2 \sim U[0, 2]$. There is a continuum of buyer types split into two cohorts. Types $\theta_1 \in \Theta_1$ value only the first attribute, i.e., each θ_1 is of the form $(\theta_{11}, 0)$, whereas types $\theta_2 \in \Theta_2$ value only the second attribute, i.e., each θ_2 is of the form $(0, \theta_{22})$. Each cohort is equally likely, and the marginal type distributions are uniform over $[0, 2]$. The sets of attributes and types are illustrated in Figure 2.

Perhaps the simplest way to think about the impact of attribute disclosure is in terms of the demand curves that the seller faces and how these curves are affected by the release of information. To this end, consider a simple class of mechanisms in which the seller provides some information free of charge and follows it with posting a single object price. These mechanisms can be viewed as marketing strategies that combine pricing with informative, persuasive advertising.

If the seller provides no disclosure, then the expectation of the first and second attributes stay at their prior values of $1/2$ and 1, respectively. As such, expected valuations of types in the first cohort are distributed uniformly over the interval $[0, 1]$ and those in the second cohort over the interval $[0, 2]$. These valuations translate into a piecewise-linear demand curve with a kink at price $p = 1$, as illustrated in Figure 3. Given this demand, the optimal no-disclosure price is $p_n = 2/3$ and it generates revenue of $1/3$.

[13] Krähmer (2020) and Doval and Ely (2020) emphasize the usefulness of such schemes in screening problems and in general games of incomplete information, respectively.

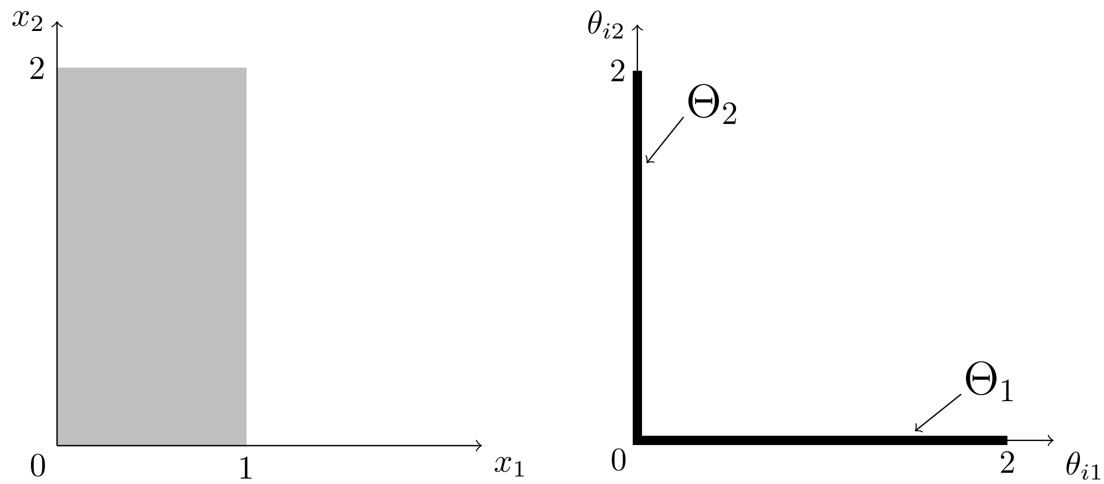

Figure 2: Illustrative example: Attributes are distributed uniformly over the gray rectangle (left); types are distributed uniformly over the L-shaped segment (right).

In contrast, if the seller provides full disclosure, then the valuation of each type stochastically changes according to the realization of a relevant attribute. As a result, a valuation of a first-cohort type with intensity θ_{11} is distributed uniformly over $[0, \theta_{11}]$, and a valuation of a second-cohort type with intensity θ_{22} is distributed uniformly over $[0, 2\theta_{22}]$. The demand curve aggregates these valuations across all types and is presented in Figure 3. Compared to the benchmark of no disclosure, the demand decreases at lower prices and increases at higher prices. The change is determined by the microstructure of the consumer types and the spread of their valuations. Roughly, one can think of this change as being driven by the types with intermediate ex ante valuations who, instead of remaining uninformed, learn that the attributes are too low, decreasing the demand at lower prices, or too high, increasing the demand at higher prices. The overall effect for the seller is negative: The optimal full-disclosure price is $p_f \simeq 0.82$ and it leads to revenue $0.28 < 1/3$.

To increase revenue, the seller should provide *partial* disclosure. Consider an experiment that reveals whether the first attribute is above some threshold α_0 but provides no information about the second attribute. This disclosure affects only valuations of first-cohort types θ_1 and can be viewed as transforming this cohort into two others. One new cohort captures valuations of types θ_1 who observed that the first attribute is below the threshold; these valuations are uniformly distributed over an interval of $[0, 2\mathbb{E}[x_1|x_1 < \alpha_0]]$. Another new cohort captures valuations of types θ_1 who observed that the first attribute is above the threshold; these valuations are uniformly distributed over an interval of $[0, 2\mathbb{E}[x_1|x_1 \geq \alpha_0]]$. As a result, the demand curve is piece-wise linear with two kinks. Relative to no disclosure,

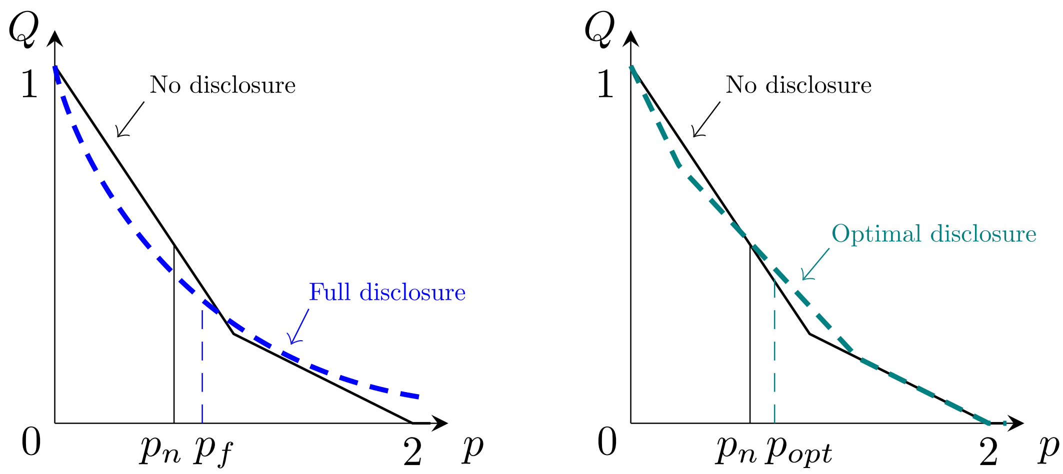

Figure 3: Attribute disclosure and demand transformation. Left: demand curves under no disclosure and full disclosure. Right: demand curves under no disclosure and optimal disclosure. Vertical lines indicate revenue-maximizing prices.

demand decreases at low prices, increases at medium prices, and remains the same at high prices. If the disclosure threshold is chosen optimally, $\alpha_0^* \simeq 0.27$, then this demand transformation benefits the seller. The optimal price is $p_{opt} \simeq 0.80$, resulting in an approximate revenue of $0.35 > 1/3$ (Figure 3). Intuitively, this disclosure targets types with lower ex ante valuations, i.e., those in the first cohort, and persuades them to buy at a higher price.

The effectiveness of partial disclosure raises the question of whether the seller can increase revenue even further if she employs a discriminatory menu with upfront payments, offering a range of experiments at different prices. After all, the cohort structure presents a clear opportunity to screen types from different cohorts by offering information about different attributes. Perhaps surprisingly, the answer is negative, as I show in Section 5.2. The simple partial-disclosure mechanism presented here is optimal.

In the example that we have just studied, the attributes were independently distributed and each type cared about a single attribute. As a result, an optimal disclosure rule provided information only about one attribute. In a more general case, one can expect an optimal disclosure rule to be richer—providing information about several attributes simultaneously. In the next section, I show how such a disclosure rule should optimally be constructed.

4 Design of Disclosure Rules

I proceed with studying the optimal menu design in general settings. I begin by discussing the buyer's incentives and formalizing his choice in an arbitrary menu. I use this formalization to approach the design problem in two consecutive steps. First, I identify the class of optimal disclosure rules without explicitly characterizing a pricing scheme; this step is presented in the current section. Second, I derive optimal pricing details and complete the characterization of optimal menus in several leading settings; that step is presented in Section 5.

4.1 Buyer's Problem

Consider the buyer's incentives when he chooses an item from a given menu. Let his type be θ. If he chooses option i, then he pays the upfront price $r(i)$. Then, a signal s is realized according to the likelihood function $\pi(E(i))$, leading to the interim valuation:

$$V(i, s, \theta) \triangleq \mathbb{E}[v(\theta, x) \mid E(i), s]. \tag{4}$$

Finally, the buyer decides whether to buy the object and does so optimally if and only if $V(i, s, \theta) - p(i)$ is greater than 0. By integrating over signal realizations, I can define the resulting *(total) trade probability* as:

$$Q(i, \theta) \triangleq \Pr(V(i, s, \theta) - p(i) \geq 0 \mid E(i)). \tag{5}$$

The corresponding indirect utility of choosing option i can be written as:

$$U(i, \theta) = -r(i) + \mathbb{E}[\max\{0, V(i, s, \theta) - p(i)\} \mid E(i)], \tag{6}$$

Type θ chooses an option with the largest indirect utility. Naturally, types seek information that most fits their interests. This feature gives the seller the opportunity to discriminate among the types by carefully designing the menu.

4.2 Responsive Menus

The seller's problem lies at the intersection of the mechanism and information design because the seller can both control the information available to the buyer and charge monetary transfers. In principle, she can offer complex experiments in an attempt to better discriminate among types; however, I show that an optimal class of experiments is simple and tractable.

I begin approaching the seller's problem by binding the size of the optimal menus and signal sets. First, there is no need to have more items than there are types, so I can focus on *direct* menus:

$$M = (r(\theta), E(\theta), p(\theta)), \tag{7}$$

which effectively ask the buyer his type and assign the experiment and the tariffs as functions of his report. Second, I can bound the size of the signal sets. For a given direct mechanism M, I call an experiment $E(\theta)$ *responsive* if $S(\theta) = \{s^+, s^-\}$, and type θ, when choosing this experiment, purchases the object if and only if $s = s^+$. I call the menu *responsive* if all of its experiments are responsive.

Proposition 1. (Responsive Menus)
The outcome of every menu can be replicated by a direct and responsive menu.

Proof. Detailed proofs of all formal statements can be found in the Appendix. □

Proposition 1 imposes an elementary structure on the exchange of information between the seller and the buyer. The buyer should inform the seller about his preferences, and the seller should provide a recommendation about whether to buy the object. The proof is analogous to the argument of the revelation principle of Myerson (1982). If the menu contains nonresponsive experiments, then the seller can replace them with responsive experiments that replicate the behavior of truth-telling types. After this modification, truth telling delivers the same payoff as before. Dishonesty, however, becomes weakly less appealing (Blackwell (1953)).

In a responsive menu, every experiment E is characterized by its *trade function*:

$$q(x) \triangleq \Pr\left(s^+ \mid E, x\right). \tag{8}$$

The function defines a probability of the trade recommendation for each attribute realization. The probability of the no-trade recommendation is then the complimentary $1 - q(x)$. A responsive menu features a collection of experiments, one per each buyer's type, that corresponds to a collection of trade functions. With slight abuse of notation, I refer to the trade function of type θ as $q(\theta, x)$.

4.3 Seller's Problem

Proposition 1 enables the association of each experiment with its trade function (8). Vice versa, any trade function $q : X \to [0, 1]$ determines a responsive experiment that recommends trading with probability $q(x)$. As such, the seller's problem can be written in a standard mechanism design form as a maximization of the expected revenue over the tariff and trade functions:

$$\max_{(r(\theta),q(\theta,x),p(\theta))} \int_{\theta\in\Theta} \left(r\left(\theta\right) + p\left(\theta\right) \int_{x\in X} q\left(\theta, x\right) \mathrm{d}G\left(x\right) \right) \mathrm{d}F\left(\theta\right), \tag{9}$$

subject to the incentive-compatibility constraints and individual rationality constraints. The seller's revenue obtained from a particular type consists of the upfront payment $r(\theta)$ and, if the buyer decides to purchase the object, the object price $p(\theta)$. The incentive-compatibility constraints require that, for all $\theta, \theta' \in \Theta$:

$$\int_{x\in X} \left(\theta \cdot x - p\left(\theta\right)\right) q\left(\theta, x\right) \mathrm{d}G\left(x\right) - r\left(\theta\right) \geq \int_{x\in X} \left(\theta \cdot x - p\left(\theta'\right)\right) \sigma\left(q\left(\theta', x\right), k\right) \mathrm{d}G\left(x\right) - r\left(\theta'\right), \tag{10}$$

where $\sigma\left(q\left(\theta', x\right), k\right)$ is a deviation function equal to $q\left(\theta', x\right)$, $1 - q\left(\theta', x\right)$, 1, and 0 for $k = 1, \ldots, 4$. These constraints ensure that each type prefers truth telling over all double-deviating strategies: misreporting and following the recommendations, "swapping" the buying decisions, always buying, or never buying. Deviations from θ to θ are included and ensure that the types are obedient on-path after truth telling.

The individual-rationality constraints require that, for all $\theta \in \Theta$:

$$\int_{x\in X} \left(\theta \cdot x - p\left(\theta\right)\right) q\left(\theta, x\right) \mathrm{d}G\left(x\right) - r\left(\theta\right) \geq 0, \tag{11}$$

so that the seller cannot force the buyer to purchase an item from the menu.

Several challenges are involved in this problem. First, the seller maximizes over a large class of all functions from a multidimensional space X. Second, it is a priori not clear what kinds of deviations are binding and, hence, relevant for the design problem: the buyer's type has no single-dimensional structure, and there is an additional multiplicity of constraints caused by double deviations.[14]

The following observation is crucial to address the experimental complexity: only two coarse statistics, not the entire trade function, matter for the revenue-maximizing problem. Namely, for a given responsive experiment E, the associated trade function q achieves the

[14] Rochet and Choné (1998), Bergemann et al. (2012) and Daskalakis et al. (2017) highlight the difficulties associated with the multidimensional screening problems.

attribute surplus and the *(total) trade probability*:[15]

$$\mathcal{X}(q) \triangleq \int_{x\in X} xq(x)\, dG(x) \in \mathbb{R}^J, \tag{12}$$

$$\mathcal{Q}(q) \triangleq \int_{x\in X} q(x)\, dG(x) \in [0,1]. \tag{13}$$

The formulations (9), (10), and (11) reveal that, due to the linearity of integration, these statistics are the only economically relevant parameters of the problem. A change in the trade function $q(\theta,\cdot)$ that does not affect the attribute surplus and the trade probability affects neither the buyer's incentives nor the seller's revenue. Accordingly, the seller can maximize directly over attribute surpluses and trade probabilities. In what follows, I refer to $\mathcal{X}(q(\theta,\cdot))$ and $\mathcal{Q}(q(\theta,\cdot))$ as $\mathcal{X}(\theta)$ and $\mathcal{Q}(\theta)$, respectively.

Not all attribute surpluses and trade probabilities can be achieved by some trade function. At one extreme, if the trade probability is nil, then the trade never occurs, $q(\cdot)\equiv 0$, so the attribute surpluses must also be nil. At the other extreme, if the trade probability is 1, then the trade always occurs, $q(\cdot)\equiv 1$, so the attribute surplus is equal to its ex ante expectation $\mathbb{E}[x]$. Intermediate values of trade probability provide more freedom to choose attribute surpluses because the seller can select the regions in which the trade recommendations are sent. The corresponding feasibility set $\mathcal{F}\subseteq\mathbb{R}^{J+1}$, which is generated by all measurable trade functions, is:

$$\mathcal{F} \triangleq \{(\mathcal{X}(q), \mathcal{Q}(q)) \mid q : X \to [0,1]\}. \tag{14}$$

The set $\mathcal{F}$ is convex, and its shape is determined by the attribute distribution G.

These observations reduce the search to the following problem:

$$\max_{\{r(\theta),\mathcal{X}(\theta),\mathcal{Q}(\theta),p(\theta)\}} \int_{\theta\in\Theta} (r(\theta) + \mathcal{Q}(\theta)\, p(\theta))\, \mathrm{d}F(\theta) \tag{15}$$

subject to incentive-compatibility constraints: $\forall\, \theta, \theta' \in \Theta$,

$$\theta\cdot\mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq \theta\cdot\mathcal{X}(\theta') - \mathcal{Q}(\theta')\, p(\theta') - r(\theta'), \tag{16}$$

$$\theta\cdot\mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq \theta\cdot(\mathbb{E}[x] - \mathcal{X}(\theta')) - (1-\mathcal{Q}(\theta'))\, p(\theta') - r(\theta'), \tag{17}$$

$$\theta\cdot\mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq \theta\cdot\mathbb{E}[x] - p(\theta') - r(\theta'), \tag{18}$$

$$\theta\cdot\mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq -r(\theta'), \tag{19}$$

[15] The attribute surplus should not be confused with the trade surplus which depends on the match between the attribute surplus and the buyer type.

the individual-rationality constraints: $\forall\, \theta \in \Theta$,

$$\theta \cdot \mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq 0, \tag{20}$$

and the feasibility constraints: $\forall\, \theta \in \Theta$,

$$(\mathcal{X}(\theta), \mathcal{Q}(\theta)) \in \mathcal{F}. \tag{21}$$

Even though the seller sells a single object, information disclosure allows her to control the multidimensional attribute surpluses $\mathcal{X}(\theta)$ at the time of a purchase. Moreover, the surpluses directly affect the feasible trade probability in a nonlinear fashion.

4.4 Optimal Disclosure

I begin by observing a special feature of a responsive experiment that always recommends the buyer to buy and, as such, provides no information about attributes. If all attributes are strictly positive, $X \subseteq \mathbb{R}^J_{++}$, then this experiment is a unique maximizer of the attribute surplus along all dimensions. If all types are strictly positive, $\Theta \subseteq \mathbb{R}^J_{++}$, then this experiment is also a unique maximizer of the trade surplus.

Proposition 2. (No Disclosure)
If all attributes and types are strictly positive, $X \subseteq \mathbb{R}^J_{++}$ and $\Theta \subseteq \mathbb{R}^J_{++}$, and the number of types is finite, then in any optimal menu, some type buys the object with probability one. That is, no disclosure, $\underline{E}$, is part of any optimal responsive menu.

Proposition 2 is consistent with the "no distortion at the top" property, common in mechanism design problems: there is a type that is optimally served an efficient allocation. However, recall that a responsive experiment only recommends allocation, and the buyer always has an option to disobey. Hence, it is important that no disclosure also provides minimal information to the buyer and, thus, maximally limits the scope of deviation. No disclosure arises in an optimal mechanism because it maximizes efficiency and minimizes incentive costs simultaneously.

Furthermore, Proposition 2 emphasizes the distinctive feature of the seller's problem, which combines information and mechanism design. In a typical information design problem, the payoff structure is exogenously fixed and, unless the receiver's indirect utility is concave everywhere, disclosure appears in the optimal mechanism for some prior distributions. Indeed, if the prices were fixed and the buyer's preferences had sufficiently low intensity, then the seller would have to provide some information to persuade the buyer to buy the object.

In contrast, when the seller has control over monetary incentives, she can compensate for the lack of information with lower prices and does find it optimal to do so.

To provide a further understanding of optimal experiments, it is useful to understand the general properties of the feasibility set $\mathcal{F}$. To this end, I define a key class of experiments.

Definition 1. (Linear Disclosure)
A responsive experiment E is a linear disclosure if, for some coefficients $\alpha = (\alpha_1, \ldots, \alpha_J) \in \mathbb{R}^J$ and $\alpha_0 \in \mathbb{R}$ not all being equal to zero, its trade function is:

$$q(x) = \begin{cases} 1, & \text{if } \alpha \cdot x > \alpha_0, \\ 0, & \text{if } \alpha \cdot x < \alpha_0. \end{cases} \tag{22}$$

A linear disclosure informs the buyer whether a linear combination of attributes is above or below a specified threshold. A linear disclosure assigns probability one to some signal everywhere outside of the defining hyperplane, $\{x \mid \alpha \cdot x = \alpha_0\}$, on which is can possibly randomize. In the case of a single attribute, a linear disclosure corresponds to a binary monotone partition of the attribute space.

A linear disclosure can be viewed as a "reference" disclosure that informs the buyer whether some virtual type $\hat{\theta} = \alpha$ would like to buy the object at price $p = \alpha_0$. If attributes are always positive and independently distributed, a linear disclosure admits additional interpretations. If elements of the coefficient vector α are positive, this disclosure can be viewed as a “level” disclosure. Observing a “trade” recommendation uniformly increases attribute expectation, whereas observing a “no-trade” recommendation uniformly decreases it. In contrast, if the elements of a coefficient vector α have different signs, then a linear disclosure can be viewed as a “comparative” disclosure between the attribute groups of different signs. A “trade” recommendation increases the attribute expectations in one group and decreases them in the other group.

Note that the likelihood function of a linear disclosure is not restricted on the defining hyperplane. Furthermore, the hyperplane does not exist for $\alpha \equiv 0$ and α_0 being strictly positive or negative. Those disclosure rules correspond to never-trade and always-trade uninformative experiments.

Lemma 1. (Feasibility)
The feasibility set $\mathcal{F}$ is compact and convex. Any linear disclosure achieves some boundary point of $\mathcal{F}$. Any boundary point of $\mathcal{F}$ is achieved by some linear disclosure.

To prove this central result, I first show that $\mathcal{F}$ is compact as a continuous image of a compact set. Second, I show that $\mathcal{F}$ is convex because a convex combination of trade functions achieves a convex combination of attribute surpluses and trade probabilities. Then, I appeal to the supporting hyperplane theorem to show that a given trade function achieves a boundary point if and only if it maximizes a linear combination of attribute surpluses and trade functions. Any such trade function corresponds to a linear disclosure.

In general multidimensional screening problems, one cannot be sure that all optimal bundles can be found at the boundary of a feasibility set. However, the current problem is an exception. To this end, say that an allocation $\left(\mathcal{X}\left(\theta\right),\mathcal{Q}\left(\theta\right)\right)_{\theta\in\Theta}$ is *implementable* if there exist tariff functions $r\left(\theta\right)$, $p\left(\theta\right)$ such that each buyer's type $\theta\in\Theta$ reports his type truthfully.

Lemma 2. (Implementability)
For any implementable allocation $\left(\mathcal{X}\left(\theta\right),\mathcal{Q}\left(\theta\right)\right)_{\theta\in\Theta}$ there exists an allocation $\left(\mathcal{X}\left(\theta\right),\mathcal{Q}'\left(\theta\right)\right)_{\theta\in\Theta}$ such that (1) it can be implemented with the same revenue and the same payoffs for all types and (2) for all $\theta\in\Theta$, $\left(\mathcal{X}\left(\theta\right),\mathcal{Q}'\left(\theta\right)\right)$ is on the boundary of $\mathcal{F}$ and $\mathcal{Q}'\left(\theta\right)\leq\mathcal{Q}\left(\theta\right)$.

Lemma 2 clarifies that the buyer incentive structure leads the seller to minimize trade probability whenever it maintains a trade surplus. If $\left(\mathcal{X}\left(\theta\right),\mathcal{Q}\left(\theta\right)\right)$ lies in the interior of $\mathcal{F}$, then the seller can reduce the total trade probability while keeping the attribute surplus the same. Such change scales up the attribute expectation conditional on the trade recommendation without affecting the trade surplus. If the seller accompanies this change with a revenue-preserving increase in the object price, then the on-path payoff of type θ remains the same. However, a higher object price renders deviations to this type's item less appealing.

Theorem 1. (Optimal Disclosure)
There exists an optimal responsive menu in which every experiment is a linear disclosure.

The theorem is an immediate corollary of Lemmas 1 and 2 and does not require any assumptions about the attribute or type distributions. To appreciate this result, it is instructive to compare the allocation distortions driven by monopoly power in cases of complete and incomplete information about the object.

First, consider the situation in which the object's attributes are commonly known to be x_0 so that there is no scope for information control. If the seller could observe the type, she would allocate the object efficiently, selling it if and only if $v\left(\theta\right)\geq 0$, and would extract full surplus. If the seller could not observe the type, she could attempt to screen by designing a menu of items varying in sale probabilities and prices. This screening is not beneficial as famously resolved by Myerson (1981). Each type θ is assigned a virtual valuation $\hat{v}\left(\theta\right)$ and, under standard regularity conditions, an object is sold if and only if the virtual valuation is positive:

$$\hat{v}\left(\theta\right)\geq 0. \tag{23}$$

This allocation is typically inefficient since the virtual valuation differs from the true valuation.

Compare this scenario to the current situation in which the object's attributes are uncertain. If the seller could observe the type θ, then, according to Bergemann and Pesendorfer (2007), she would inform type θ whether his valuation is positive, charge a maximal acceptable price, and extract the full surplus. If the seller could not observe the type, she could design a menu varying in information content and prices. By Theorem 1, the optimal allocation distortion would be remarkably similar to the case of complete information about the object. Each type θ is assigned a virtual type $\hat{\theta}(\theta)$. An object is sold when the buyer's virtual valuation is above a specified threshold, possibly with randomization on the boundary:

$$\hat{v}(\theta) = \hat{\theta}(\theta) \cdot x \geq \alpha_0(\theta). \tag{24}$$

This allocation is also typically inefficient but is now with an additional distortion since the threshold may differ from 0.

Uniqueness Theorem 1 establishes the existence of an optimal responsive menu with each experiment being a linear disclosure. One might wonder whether there exist optimal responsive menus with non-linear disclosure rules. The proof argument does not preclude this possibility: although the adjustment to linear disclosure strictly relaxes constraints (17) and (18), it preserves the constraint (16); therefore, hypothetically, the relaxed constraints may be not exploited for additional revenue. It can be shown that with only two types, this relaxation *can* be exploited; thus, linear disclosure is uniquely optimal. With many types, the answer is less clear as the structure of incentive constraints is more complex.

At the same time, observe that the adjustment to linear disclosure strictly decreases the trade probability. Consequently, when the seller faces trading costs or, equivalently, attaches some value to the object, however small, linear disclosures are uniquely optimal.

General Payoffs The arguments behind Theorem 1 might seem to heavily rely on the linearity of the buyer's valuation function. However, Theorem 1 places no structural assumptions on the attribute distribution. This crucial feature allows for extending the optimal disclosure characterization beyond linear environments by carefully defining the relevant attributes. In particular, consider a general valuation function $v(\theta, x)$ and define auxiliary attributes to coincide with the valuations of different buyer types. In this auxiliary formulation, each type's valuation is linear in the relevant attribute, and Theorem 1 applies.

Corollary 1. (General Payoffs)

Let X be an arbitrary attribute set, $v(\theta, x)$ be a general valuation function, and $|\Theta| < \infty$.

Then, there exists an optimal menu in which every experiment has a linear form, i.e., for any experiment, there exist $\alpha : \Theta \to \mathbb{R}$ and $\alpha_0 \in \mathbb{R}$, not all zeros, such that:

$$q(x) = \begin{cases} 1, & \text{if } \sum_{\theta \in \Theta} \alpha(\theta) v(\theta, x) > \alpha_0, \\ 0, & \text{if } \sum_{\theta \in \Theta} \alpha(\theta) v(\theta, x) < \alpha_0. \end{cases} \tag{25}$$

Note the difference between the definitions of a linear form (25) and a linear disclosure (22). A linear disclosure operates in a space of attributes and can be specified independently of a buyer. A linear form, in contrast, operates in the space of valuations of different buyer types. The richer the buyer heterogeneity is, the more complex the linear form can be. However, in the case of linear payoffs the linear form always reduces to a linear disclosure.

Corollary 1 allows for a characterization of the classes of optimal disclosure rules in general environments with preferences that allow for bliss points or risk aversion. To illustrate, consider the case of *location payoffs* with the buyer's type capturing his bliss point in the attribute space $X \subseteq \mathbb{R}^J$, $\Theta \subseteq \mathbb{R}^J$, $v_0 > 0$, and:

$$v(\theta, x) = v_0 - (x - \theta)^2. \tag{26}$$

Let there be two types $\theta_1, \theta_2 \in \mathbb{R}^J$. Assume that X is bounded and for all $x \in X$, the types' valuations are positive. Optimal disclosure rules can be identified as follows. First, Proposition 2 can be applied to establish that one type is offered no disclosure and always buys. Second, by Corollary 1, the other type is offered a linear form (25) that informs whether a linear combination of valuations $v(\theta_1, x)$ and $v(\theta_2, x)$ is above or below a specified threshold. Generically, this experiment is a *neighborhood disclosure*: it informs whether the attribute vector is sufficiently close to a virtual type $\hat{\theta}$ located on the line that connects θ_1 and θ_2.

5 Design of Pricing Mechanisms

I proceed by studying pricing in the revenue-maximizing mechanisms. I identify a general class of optimal pricing mechanisms in the case of a single attribute. With many attributes, I am able to identify key trade-offs and characterize optimal mechanisms for specific classes of buyer types.

From now on, I assume that all types and attributes are positive, $X \subseteq \mathbb{R}^J_+$, $\Theta \subseteq \mathbb{R}^J_+$. In this scenario, it is commonly known that there are positive gains from trade. It makes it possible to ignore the efficiency role of disclosure and to focus solely on its screening effects.

5.1 Single Attribute

I begin with the basic case of a single attribute, $J = 1$, $X \subseteq \mathbb{R}_+$. The buyer's type is one dimensional, $\Theta \subseteq \mathbb{R}_+$, and the buyer's ex post valuation is

$$v(\theta, x) = \theta x. \tag{27}$$

This setting features only vertical type heterogeneity. I establish that providing no attribute information is optimal in this case. The argument starts by considering a more beneficial setting for the seller, in which she can condition payment and allocation directly on the attribute realization, as do Eső and Szentes (2007). In this case, the revelation principle applies, and I can focus on the direct mechanisms in which all payments are front loaded: the buyer reports his type θ, pays the upfront payment $r(\theta)$, and the trade occurs with probability $q(x, \theta)$. The relevant variable is the single-dimensional attribute surplus:

$$\mathcal{X}(\theta) = \int_{x \in X} xq(x, \theta)\,\mathrm{d}G(x) \tag{28}$$

which can be anywhere between 0 and $\mathbb{E}[x]$. I can then rewrite the seller's problem as:

$$\max_{r(\theta), 0 \leq \mathcal{X}(\theta) \leq \mathbb{E}[x]} \int_{\theta \in \Theta} r(\theta)\,\mathrm{d}F(\theta), \tag{29}$$

$$\text{s.t. } \theta\mathcal{X}(\theta) - r(\theta) \geq \theta\mathcal{X}(\theta') - r(\theta'), \quad \forall\, \theta, \theta' \in \Theta, \tag{30}$$

$$\theta\mathcal{X}(\theta) - r(\theta) \geq 0, \quad \forall\, \theta \in \Theta. \tag{31}$$

This problem is analogous to those of Myerson (1981) and Riley and Zeckhauser (1983), with the attribute surplus replacing the allocation probability. The optimal allocation $\mathcal{X}(\theta)$ is a step function equal to 0 for $\theta < \theta^*$ and to $\mathbb{E}[x]$ for $\theta \geq \theta^*$. The optimal upfront payment $r(\theta)$ is equal to 0 for $\theta < \theta^*$ and to $r^* = \theta^* \mathbb{E}[x]$ for $\theta \geq \theta^*$.

The argument concludes by noting that the optimal mechanism can be implemented by providing no disclosure and charging a price of r^* for the object. This posted price mechanism is feasible in the original problem with private disclosure and is therefore also optimal there.

Theorem 2. (Single Attribute)
If $J = 1$, $X \subseteq \mathbb{R}_+$, and $\Theta \subseteq \mathbb{R}_+$, then an optimal menu is a posted price mechanism with no disclosure, i.e., $r(\theta) \equiv 0$, $E(\theta) \equiv \underline{E}$.

Simple intuition underlies the optimality of no disclosure if the seller can only use a nondiscriminatory mechanism that consists of a single experiment followed by a posted price. Consider an arbitrary disclosure rule. Any signal realization s scales the demand proportionally to the attribute expectation $\mathbb{E}[x \mid s]$. If the seller could observe this realization, she would optimally charge a scaled price and obtain scaled revenue. Importantly, the induced allocation would not depend on the realization s. Since any expectation is a martingale, the seller would serve the same population at, on average, the same price. The seller can do equally well using a posted price with no disclosure.

Although intuitive, this argument does not consider discriminatory schemes with upfront payments. Theorem 2 confirms that no disclosure is optimal, even if the seller can use those schemes. Notably, this result requires no assumptions on the type or attribute distributions beyond the common knowledge of positive trade gains.

Remark 1. *(Attribute Index)* The same argument can be applied to the case of many attributes, $J > 1$, if the attributes and the types enter the valuation function through one-dimensional indices:

$$v(\theta, x) = \psi(\theta)\,\phi(x), \tag{32}$$

for $\psi, \phi : \mathbb{R}^J \to \mathbb{R}_+$. For example, the result applies if all types belong to a ray $\Theta = \{\beta\theta_0\}_{\beta\in\mathbb{R}_+}$ for some direction vector $\theta_0 \in \mathbb{R}^J_+$. In this case, the indices can be defined as $\psi(\theta) = \beta(\theta)$ and $\phi(x) = \theta_0 \cdot x$. □

Remark 2. *(Uniqueness)* The no-disclosure mechanism might not be uniquely optimal. In fact, the analysis of Eső and Szentes (2007) can be applied to show that, if the type distribution F has a monotone hazard rate property, then full disclosure is also optimal.[16] However, it must be accompanied by a complex structure of upfront payments and object prices. □

5.2 Single-Minded Buyer

I proceed with the case of multiple product attributes, $J \geq 2$. This case is qualitatively different because it may feature horizontal heterogeneity across buyer types: type θ may have a higher value for the object than type θ' for some attribute realization, yet a lower value than type θ' for another attribute realization. Since the seller should aim to allocate to object to the buyers who are willing to pay the most, she may need to provide attribute information through disclosure. To address such settings, I introduce and study a tractable type structure which allows to capture both vertical and horizontal heterogeneity of tastes.

I call a type *single minded* if he values only one attribute. For a generic single-minded type, the vector θ places a positive weight on only one dimension:

$$\theta = (0, \ldots, 0, \theta_j, 0, \ldots, 0). \tag{33}$$

[16] F has the monotone hazard rate property if $f(\theta)/(1 - F(\theta))$ is monotonically increasing.

Thus, single-minded types allow for a simpler notation. I can represent the types by J *attribute cohorts* Θ_j such that all types within the same cohort value the same attribute. I slightly abuse the notation and let the type subscript identify the attribute cohort and the type value identify the valuation intensity, so that $\Theta_j \subseteq \mathbb{R}_+$ and

$$v_j(\theta_j, x) = \theta_j x_j \quad \forall\, j, \theta_j \in \Theta_j. \tag{34}$$

I denote the frequency of a cohort Θ_j by $f(\Theta_j)$ and the cumulative type distribution within the cohort by $F_j(\theta_j)$.

A buyer is *single minded* if all types $\theta \in \Theta$ are single minded and the attribute values are independently distributed such that $x_j \sim G_j$ and $G(x) = \times_j G_j(x_j)$.[17] The independence requirement is substantive. Without it, any buyer can be viewed as being single minded by redefining the attributes as in the proof of Corollary 1.

If the buyer is single minded, then the seller knows that the buyer values only one of many independent attributes but does not know which one or the strength of the preference. Valuations of any two types are either perfectly correlated or independent.

Remark 3. *(Customizable Good)* The setting of a single-minded buyer resembles a unit-demand multi-good monopolist problem in which each buyer type values only one of the goods.[18] One difference between these problems is that I allow the seller to provide additional information to the buyer. Another difference is that in the multi-good monopolist problem a type who deviates across cohorts is guaranteed to obtain no value; therefore, the problem can effectively be separated into several single-good problems. In contrast, in my setting, when deviating across cohorts, the buyer can at least obtain the ex ante value of the relevant attribute, which links the problems together. At the same time, the setting of a single-minded buyer admits an interpretation of a sale of a customizable good. In this interpretation, the good that admits several possible configurations and the seller can inform the buyer about them. The buyer values only a single configuration and can set the configuration freely but only one time after the purchase. The question of optimal design can then be translated into what configuration information the seller provides and how this choice interacts with pricing. □

If the buyer is single minded, then the class of optimal experiments can be narrowed. Because attributes are independently distributed, a type $\theta_j \in \Theta_j$ values only information about attribute j. This observation suggests an optimal way to screen single-minded types: if the buyer reports type $\theta_j \in \Theta_j$, then the seller should provide information only about

[17] The name is inspired by “single-minded” bidders in combinatorial auctions who value specific bundles. See, for example, Lehmann et al. (2002).

[18] For example, see Section 8.4 of Hartline (2020).

attribute j. Providing any other information would make misreporting more appealing without adding value for truth telling. At the same time, a linear disclosure informative only about attribute j is a binary monotone partition defined on this attribute.

Proposition 3. (Directional Disclosure)
If the buyer is single minded, then there exists an optimal menu such that an experiment $E_j(\theta_j)$ *is a binary monotone partition of attribute* j.

It follows that an optimal experiment $E_j(\theta_j)$ can be characterized by its threshold $\alpha_{0j}(\theta_j)$ so that it informs the buyer whether attribute j is above or below this threshold. Incentive compatibility requires the buyer to purchase at higher attributes; thus, the attribute surplus can be written as:

$$\mathcal{X}_j(\theta_j) = \int_{\alpha_{0j}(\theta_j)}^{\infty} x_j \mathrm{d}G_j(x_j). \tag{35}$$

The attribute surplus can take any value between 0 and $\mathbb{E}[x_j]$. The corresponding total trade probability can be written as an increasing and convex function $\mathcal{Q}_j(\mathcal{X}_j)$.

In what follows, I assume that that all attribute cohorts admit an upper bound, $\Theta_j = \left[0, \overline{\theta}_j\right]$, and θ_j are continuously distributed over Θ_j according to distribution function F_j with the monotone hazard rate property. The seller's problem is to design the tariff and the attribute surplus functions, $r_j(\theta_j)$, $p_j(\theta_j)$, and $\mathcal{X}_j(\theta_j)$ to maximize her objectives and can be written as follows (cf. Kolotilin et al. (2017)):

$$\max_{\{r_j(\theta_j), \mathcal{X}_j(\theta_j), p_j(\theta_j)\}} \sum_{j=1}^{J} f(\Theta_j) \int_{\theta_j \in \Theta_j} (r_j(\theta_j) + \mathcal{Q}_j(\theta_j) p_j(\theta_j)) \mathrm{d}F_j(\theta_j)$$

$$\text{s.t. } \theta_j \mathcal{X}_j(\theta_j) - p_j(\theta_j)\mathcal{Q}_j(\theta_j) - r_j(\theta_j) \geq \left(\theta_j \mathcal{X}_j\left(\theta_j'\right) - p_j\left(\theta_j'\right)\right)\mathcal{Q}_j\left(\theta_j'\right) - r_j\left(\theta_j'\right), \ \forall j, \theta_j, \theta_j' \in \Theta_j,$$

$$\theta_j \mathcal{X}_j(\theta_j) - p_j(\theta_j)\mathcal{Q}_j(\theta_j) - r_j(\theta_j) \geq \theta_j \mathbb{E}[x_j] - p_k(\theta_k) - r_k(\theta_k), \ \forall j, k, \theta_j \in \Theta_j, \theta_k \in \Theta_k,$$

$$\theta_j \mathcal{X}_j(\theta_j) - p_j(\theta_j)\mathcal{Q}_j(\theta_j) - r_j(\theta_j) \geq 0, \mathcal{X}_j(\theta_j) \geq \mathcal{Q}_j(\theta_j)\mathbb{E}[x_j], \forall j, \theta_j \in \Theta_j,$$

where $\mathcal{Q}_j(\theta_j) \equiv \mathcal{Q}_j(\mathcal{X}_j(\theta_j))\ \forall j, \theta_j \in \Theta_j$. This problem resembles a collection of one-dimensional mechanism design problems, one per attribute, with the following important differences. First, each item in this problem features both horizontal and vertical components. The upfront payments $r_j(\theta_j)$ are purely vertical—all types value them the same. In contrast, the attribute surpluses $\mathcal{X}_j(\theta_j)$ are only valuable to types from cohort Θ_j. The object prices p_j are mixed since they are paid only if the type decides to trade. Second, the problem features nonlinear terms $p_j(\theta_j)\mathcal{Q}_j(\mathcal{X}_j(\theta_j))$ and $\mathcal{Q}_j(\mathcal{X}_j)$.

Because of these differences, I cannot apply standard mechanism design techniques. Instead, I solve the problem via a sequence of simplifications. In the first step, I observe that using upfront payments is detrimental. For any $r_j(\theta_j) > 0$, the seller can reduce the transfer and increase $p_j(\theta_j)$ while keeping the total expected transfer $r_j(\theta_j) + \mathcal{Q}_j(\theta_j)\, p_j(\theta_j)$ the same. This change does not affect the utilities of truth-telling types or the seller's revenue; however, it renders misreporting less appealing. Intuitively, by shifting the expected transfer toward the object price, the seller better discriminates against the types who would always purchase the object.

In the second step, I use standard mechanism design arguments to show that incentive compatibility within the same cohort implies that $\mathcal{X}_j(\theta_j)$ is nondecreasing in θ_j. That is, higher types must trade with a higher probability but lower conditional expectations of the relevant attribute. Moreover, the expected transfers can be derived from the attribute surplus functions. This allows to rewrite the seller's problem solely as a choice over attribute surplus functions $\mathcal{X}_j(\theta_j)$.

Lemma 3. *The seller's problem can be written as*

$$\max_{\{\mathcal{X}_j(\theta_j)\}_{j=1}^J} \sum_{j=1}^{J} f(\Theta_j) \int_0^{\overline{\theta}_j} \left(\theta_j - \frac{1 - F_j(\theta_j)}{f_j(\theta_j)}\right) \mathcal{X}_j(\theta_j)\, \mathrm{d}F_j(\theta_j) \tag{36}$$

$$\text{s.t.} \quad \mathcal{X}_j(\theta_j)\, \text{is non} - \text{decreasing}, \;\; \mathcal{X}_j(\theta_j) \in [0, \mathbb{E}[x_j]], \tag{37}$$

$$\int_0^{\overline{\theta}_j} \mathcal{X}_j(\theta_j)\, d\theta_j \geq \overline{\theta}_j \mathbb{E}[x_j] - \underline{p}(\mathcal{X}_1(\cdot), \ldots, \mathcal{X}_J(\cdot)) \;\; \forall j = 1, \ldots, J. \tag{38}$$

The objective function and the monotonicity constraints capture the incentive-compatibility constraints within each attribute cohort. The integral constraints capture the incentive-compatibility constraints between different cohorts. In particular, these constraints require that the highest type within each cohort not want to purchase the object at the minimal price present in the menu $\underline{p}$.

In an optimal menu, the minimal price must be offered to the highest types. Toward a contradiction, assume that for some cohort Θ_j, a neighborhood of $\overline{\theta}_j$ is not offered the minimal price. Then, these types are not imposing externalities on other cohorts through the integral constraint. Moreover, to not go for the lowest price, these types should be offered some disclosure so $\mathcal{X}_j(\theta_j) < \mathbb{E}[x_j]$. It leads to a contradiction: the seller could marginally increase $\mathcal{X}_j(\theta_j)$ for these types, improving the revenue.

This observation allows for the stating of a relaxed problem in which the monotonicity and integral constraints are dropped but all high types are required to be offered a given price. If the type distributions have the monotone hazard rate property, then the solution to the problem is a collection of single-step functions. The solution corresponds to only one item per attribute cohort. It satisfies the original constraints and therefore solves the original problem.

Theorem 3. (Optimal Menu, Single-Minded Buyer)
If the buyer is single minded and the type distributions have the monotone hazard rate property, then in an optimal responsive menu for all $j = 1, \ldots, J$ *and* $\theta_j \in \Theta_j$*:* $r(\theta_j) = 0$, $p(\theta_j) = p$, *and* $E_j(\theta_j) = E_j$ *where* E_j *is a binary monotone partition of* x_j.

I emphasize the simplicity of both pricing and disclosure components of the optimal mechanism. With regard to pricing, the price for the object is the same for all types and the price of information is nil. With regard to disclosure, the mechanism admits a nondiscriminatory indirect implementation: the seller can simply inform the buyer whether each attribute is above the corresponding threshold and post a fixed price for the object. Because each type values only one attribute and the attributes are independent, each type will use only information about a relevant attribute when deciding whether to buy the object. This simplicity is not a consequence of some exogenous requirement but rather is a feature of a revenue-maximizing mechanism.

Furthermore, my analysis also provides a partial characterization in the case of general distributions F_j. The first statement of Theorem 3 remains the same—upfront payments are not used with a single-minded buyer. However, the second and the third statements must be modified since an optimal menu may feature limited price discrimination. In particular, the arguments of Samuelson (1984) can be applied to limit the number of optimal items to two per cohort. That is, the highest types are still offered the unique minimal lowest price but per each attribute cohort, there could be one more item that targets lower types.

5.3 Differentiated Types and Separate Persuasion

A notable limit case of the previous section is the case in which each attribute cohort is a singleton, $\Theta_j = \{\theta_j\}$. In this case, there is no vertical within-attribute heterogeneity, and the number of types equals the number of attributes $|\Theta| = J$. Any two different types $\theta, \theta' \in \Theta$ are orthogonal to one another as vectors in $\mathbb{R}^J$; accordingly, this case can be viewed as the setting of *orthogonal types*. Without loss of generality, all type intensities can be set equal to one, $\theta_j \equiv 1$ for all j.

By the arguments analogous to those in the previous section, an optimal mechanism is a free-of-charge disclosure followed by a single posted price. Finding optimal experiments is straightforward. The seller should provide minimal information sufficient to convince the buyer to make a purchase at the price posted. If the type θ_j is ex ante sufficiently optimistic, $\mathbb{E}[x_j] \geq p$, then the seller should provide no attribute information, $E_j = \underline{E}$. Otherwise, the seller should increase the type's expectation up to the object price.

The corresponding optimal mechanism is illustrated in Figure 4. The mechanism is simple

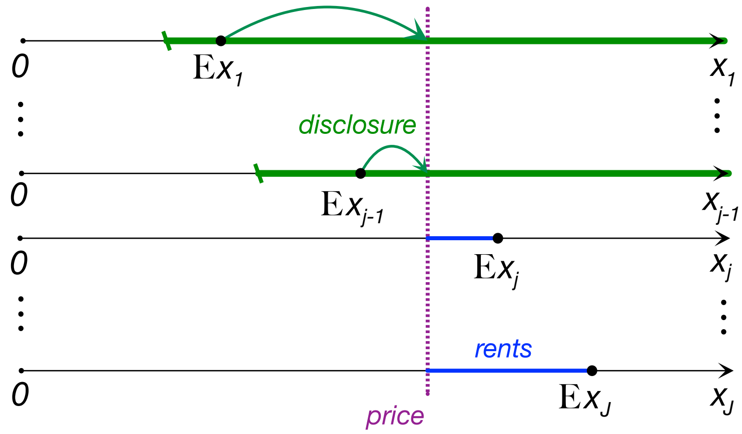

Figure 4: Optimal mechanism in the case of orthogonal types. Green indicates attribute regions in which a purchase recommendation is sent for the types with partial disclosure. Blue indicates the types' rent conditional on a trade, for the types with no disclosure. Attributes are ordered by increasing ex ante expectations.

and features a "no distortions at the top, no rents at the bottom" property: All types with the ex ante valuation above the optimal price always buy the object, whereas all other types are indifferent to participating in the mechanism. In this way, "the top" and "the bottom" are not single types, as is typical in mechanism design problems, but are two type classes that partition the type space.

The case of orthogonal types admits a clean illustration of an optimal mechanism and, importantly, suggests a generalization of the current analysis beyond a single-minded buyer. Toward this generalization, consider arbitrary attribute distribution $G(x)$ and valuation function $v(\theta, x)$. For a given price p, define a *separate persuasion mechanism* $M^{SP}(p)$ as a direct menu in which $r^{SP}(\theta) = 0$, $p^{SP}(\theta) = p$, and $E^{SP}(\theta)$ recommends to trade when $v(\theta, x)$ is above a threshold; the threshold is chosen so that $\mathbb{E}\left[v(\theta, x) \mid E^{SP}(\theta), s^+\right] = \max\{p, \mathbb{E}[v(\theta, x)]\}$. That is, the seller fixes an object price and for each type provides minimal valuation information to persuade him to make a purchase.

This mechanism is one with discriminatory disclosure that mimics the optimal menu for orthogonal types. If a type reports truthfully, then he is either left with no rents or is provided with no information; in both cases, the experiment brings no value to him. The mechanism may be incentive compatible or not—it depends on whether some types may benefit from information offered to other types. Denote with p^* a price that maximizes the revenue, if the buyer is assumed to report his type truthfully.

Theorem 4. (Separate Persuasion)
If the mechanism $M^{SP}(p^*)$ is incentive compatible, then it is an optimal mechanism.

This proposition is based on the observation that a separate persuasion mechanism solves a relaxation of the original problem, in which misreporting types are required to always purchase the object. If incentive compatible, the mechanism satisfies the relaxed constraints and also solves the original problem.[19]

Let me highlight the significance of Theorem 4. A priori, there is no reason to expect the separate persuasion mechanism to be optimal since it does not use all of the flexibility available to the seller: it does not price information and it does not vary the price of the object. Furthermore, the mechanism does not extract the full surplus because it generally induces inefficient allocation. Rather, the result builds on and generalizes the analysis of orthogonal types.

Theorem 4 can be used to find optimal mechanisms in environments in which the types are sufficiently differentiated. For the following statement, let the valuation function be linear (1) and the attributes be independently and continuously distributed.

Corollary 2. (Differentiated Types)
Fix the number of types J and their respective frequencies. Let p^* be a uniquely optimal price in a separate persuasion mechanism for some orthogonal types $\left(\hat{\theta}_1, \ldots, \hat{\theta}_J\right)$. If $p^* \neq \mathbb{E}\left[v\left(\hat{\theta}_j, x\right)\right]$ for all j, then there exists $\varepsilon > 0$ such that for any type profile $(\theta_1, \ldots, \theta_J)$ with $\left\|\theta_j - \hat{\theta}_j\right\| \leq \varepsilon$ for all j a separate persuasion mechanism is optimal.

According to the corollary, separate persuasion mechanisms are generically optimal even if the types value several attributes as long as they place most of the weight on distinct attributes. In these cases, the seller does not benefit from price discrimination but does benefit from information discrimination.

Note the structure of optimal disclosure when the types place weights on several attributes, however small. Even if the attributes are independent, providing information about each of them separately is *not* optimal. Instead, the disclosure leans towards the tastes of a reported type. At the same time, the informational content regarding a given attribute decreases as the type attaches less weight to it; thus, if the tastes are sufficiently concentrated, then providing information only about the leading attribute is approximately optimal.

[19]In any given setting, the incentive compatibility of this mechanism can be straightforwardly checked.

6 Discussion

6.1 Pricing of Product Information

In the setting of Section 5, the seller cannot benefit from pricing the information that she provides—upfront payments can always be set to zero. Similarly, in reality, it is uncommon to price product information. This fact raises a question regarding why it might be not beneficial to price product information separately from the product itself.

Proposition 4. (Payment Backloading)
In any given responsive menu, $r(\theta)$ can be decreased, and $p(\theta)$ can be increased without any loss of revenue when (i) $\mathcal{Q}(\theta) \leq 1/2$ or (ii) no type is willing to act contrary to both recommendations if being offered the experiment $E(\theta)$.

Proposition 4 holds for any valuation function. To understand the logic behind it, take an item of type θ in a responsive menu and consider the effect of *backloading* its payment, i.e., reducing the information price and increasing the object price to preserve the expected payment of type θ. By construction, this backloading preserves the incentives of the types who, if offered the item of type θ, would like to follow the recommendations. At the same time, it changes the incentives of the types who would like to disobey recommendations because the object price is paid only if the object is actually purchased. In particular, the backloading relaxes the incentives of the types who would purchase the object more frequently than θ when faced with $E(\theta)$. Hence, the incentives are clearly relaxed for the types who always purchase the object, because $1 \geq \mathcal{Q}(\theta)$; but also, whenever $\mathcal{Q}(\theta) \leq 1/2$, the incentives are relaxed for the types who would like to act contrary to recommendations because in that case $1 - \mathcal{Q}(\theta) \geq \mathcal{Q}(\theta)$. The incentives of the types who never purchase the object are captured by their individual rationality. As such, under the conditions of the proposition, the payments can be backloaded without any loss of revenue. In fact, this argument reveals that the second condition of the proposition need only hold for the types along the binding constraints.

By Proposition 4, a given experiment might need to be priced only if it induces a relatively frequent trade and some types would like to mismatch its recommendations. That is, a priced experiment should bring "bad news" by infrequently informing the buyer that the product is not worth its price. Moreover, this information should lead to a disagreement, with some type using it in the opposite fashion. In some cases, one can exclude the latter possibility.

Corollary 3. (No Information Pricing)
In an optimal menu, the price of an experiment E can be set to zero if (i) $J > 1$, attributes

are independent, $v(\theta, x) = \theta \cdot x$*,* $\Theta \subseteq \mathbb{R}_+^J$*, and* E *is a linear disclosure with* $\alpha \in \mathbb{R}_+^J$ *or (ii)* $J = 1$*,* $v(\theta, x)$ *is increasing in* x *for all* $\theta \in \Theta$ *and* E *is a binary monotone partition.*

In the cases of Corollary 3, the types agree on the ranking of their interim valuations after different signals of the experiment E. Therefore, no type likes to mismatch decisions with recommendations. The first case indicates that only comparative disclosures may need to be priced. The second case explains why binary monotone partitions, frequent in economic analyses, need not be priced.

At the same time, examples can be constructed in a general setting with sufficiently opposed types such that, in an optimal menu, the swapping constraint binds and the information should be priced to deter mismatching deviations.

6.2 Demand Transformation and Shrouded Attributes

As I discussed in Section 3, one way to think about attribute disclosure is in terms of its impact on the demand curve that the seller faces. This particular effect was studied by Johnson and Myatt (2006). They restrict their attention to disclosures that spread type valuations "uniformly." Such disclosures translate into *global* rotations of the demand curve. The authors show that in many settings, the optimal global rotations are extreme and correspond to either no disclosure or full disclosure: no disclosure is associated with a mass market characterized by a low price and high demand; full disclosure is associated with a niche market characterized by a high price and low demand.

In contrast, I show that attribute disclosure can rotate the demand curve *locally* (recall Figure 3). The local rotations correspond to partial disclosures that target specific types and, hence, affect the demand curve over a particular price segment. They can outperform full and no disclosure in both mass and niche markets. Multiple attributes are required for this result—recall that no disclosure remains optimal in a one-dimensional framework.

This idea of targeted disclosure provides additional justification for selective advertising and attribute shrouding, ubiquitous in practice (Gabaix and Laibson (2006)). Even if customers are perfectly rational, the seller may have incentives to suppress information about some attributes while providing information about the others. Intuitively, different attributes can appeal to different customer cohorts. At a given price, some cohorts may have to be persuaded to purchase the product while others may not have to be.

For example, the customer base of a smartphone company could consist of two main groups: high-value customers, who are primarily interested in reliability and the quality of customer service; and low-value customers, who are primarily interested in entertainment features such as the screen size and camera performance. A smartphone price will optimally balance the respective cohort valuations. In the absence of additional information, the price will be acceptable for the high-value customers but not for the low-value customers. The advertising campaign could thus optimally focus on the entertainment features to persuade the low-value cohort, while suppressing the information about reliability and services.

6.3 Alternative Disclosure Settings

The multiattribute disclosure setting complements the existing one-dimensional models of disclosure and pricing by Eső and Szentes (2007) and Li and Shi (2017). The main difference, however, lies not in the multidimensionality per se but rather in what kinds of information the seller can provide.

Eső and Szentes (2007) study discriminatory mechanisms in a valuation-rank framework, in which disclosure corresponds to statements such as "Your valuation is in your y-th percentile," with y being the same for all types. The authors obtain two main qualitative results: first, they show that full information disclosure is generally optimal; and second, they show that the seller cannot benefit from conditioning the price on the disclosure realization.[20]

Formally, their seller informs the buyer about an *orthogonal shock* $\xi(\theta)$, defined as the type's valuation percentile. By construction, these percentiles are uniformly distributed:

$$\xi(\theta) \sim U[0,1] \quad \forall \theta \in \Theta. \tag{39}$$

The implicit assumption of the valuation-rank framework is that these shocks are equal, i.e., $\xi(\theta) \equiv \xi(\theta')$ for all $\theta, \theta' \in \Theta$. However, despite having the same distribution, the shocks $\xi(\theta)$ are generally *different* random variables. This observation is crucial and is particularly evident in the multiattribute setting.

Consider the following example. Let there be two independently distributed attributes $J = 2$, $x_1 \sim U[0,1]$, $x_2 \sim U[0,2]$. Let there be two equally likely types: $\theta_1 = (1,0)$ and $\theta_2 = (0,1)$. The corresponding orthogonal shocks are $\xi(\theta_1) = x_1$ and $\xi(\theta_2) = x_2/2$. Both $\xi(\theta_1)$ and $\xi(\theta_2)$ are uniformly distributed on $[0,1]$. However, they depend on different attributes and are thus independent from each other.

Consequently, neither of the qualitative results of Eső and Szentes (2007) holds in this example. The optimal full-disclosure mechanism can be calculated to be $r_1 = r_2 = 1/2$, $p_1 = p_2 = 0$: anticipating information revelation, the seller would prefer to effectively sell the object in advance. The corresponding revenue is $1/2$. The seller can do strictly better by providing partial private disclosure. Building on the results of Section 5.2, it can be shown

[20] Eső and Szentes (2017) generalize the latter finding to dynamic environments. Krähmer and Strausz (2015a) discuss settings in which the full-disclosure distributional assumptions are violated.

that an optimal mechanism informs the buyer, free of charge, whether the first attribute is above or below 1/2 and then follows with a posted object price of 3/4. This mechanism obtains revenue $9/16 > 1/2$.

At the same time, the seller could further improve the revenue if she could condition the price directly on the disclosure realization. Consider the following mechanism. The seller provides full disclosure, observes the attributes, and chooses the price optimally given the realized valuation distribution. The corresponding revenue is:

$$\Pi = \int_0^1 \int_0^2 \frac{1}{2} \max \left\{ \min \{x_1, x_2\}, \frac{\max \{x_1, x_2\}}{2} \right\} \mathrm{d}x_1 \mathrm{d}x_2 = \frac{29}{48} > \frac{9}{16}. \tag{40}$$

This example also emphasizes the difference between the multiattribute setting and the setting of Li and Shi (2017). These authors study common value settings in which the types represent private information about the object. In their settings, information disclosure can be seen as a valuation-level disclosure that corresponds to statements such as "Your valuation is above x," with x being the same for all types. However, in general, attribute information affects the valuation of different types differently according to the valuation function. In the example above, any experiment informative only about attribute x_i affects the valuation of only type θ_i. Consequently, attribute information cannot be modeled as an experiment that informs the buyer directly about his valuation. This modeling would misrepresent the buyer incentives in terms of choice across experiments.

Overall, this discussion emphasizes the importance of explicitly modeling demand microstructure, i.e., how the consumer valuation is formed, in trade settings with information control.

7 Conclusion

I studied a monopolist who sells a multiattribute object to a privately informed buyer and showed that the seller can benefit from the disclosure of attribute information. The benefit comes through two channels. First, disclosure can be used as a screening device, leveraging the preferences of different buyer types for learning about different aspects of the object. Second, disclosure can lift the buyer's expectations and persuade him to buy the object at a higher price. Both channels are important. However, I show that in many settings screening is not beneficial and information should be disclosed partially and free of charge. In those settings, the choice of information content is more important than the choice of its pricing.

In this paper, I deliberately focused on the simplest model of pricing and information control. In practice, additional details might be important and should be considered. The seller could be restricted in the kinds of information that she may provide. The buyer might feature heterogeneity in his ability to process data. The market could involve imperfect competition. Each of these extensions can be approached within the multiattribute disclosure framework that I have outlined.

8 Appendix

Proof of Proposition 1. Consider any menu $M = (r(i), E(i), p(i))_{i \in \mathcal{I}}$. For any type θ, M induces the allocation distribution $\mu(\theta) : X \to \Delta(A)$, $A = \{\text{buy}, \text{not buy}\}$, the expected upfront payment $\hat{r}(\theta)$, and the expected object payment, conditional on a trade, $\hat{p}(\theta)$. Consider a direct responsive menu $M' = (r'(\theta), E'(\theta), p'(\theta))$ with $r'(\theta) = \hat{r}(\theta)$, $p'(\theta) = \hat{p}(\theta)$, and $E'(\theta) = (A, \mu(\theta))$. If all types are truthful and obedient, then M' results in the same allocation distribution and the same expected payments as M. At the same time, any deviation under M' is available to the buyer under M. Therefore, reporting truthfully and following the recommendations is incentive-compatible under M'.

Proof of Proposition 2. Toward a contradiction, assume that a responsive menu $M = (r(\theta), \mathcal{X}(\theta), \mathcal{Q}(\theta), p(\theta))_{\theta \in \Theta}$ is optimal, yet no type buys the object with probability one. Construct a new menu M' as follows. Select a type $\overline{\theta}$ with the highest expected payment $\overline{T} = r\left(\overline{\theta}\right) + p\left(\overline{\theta}\right) \mathcal{Q}\left(\overline{\theta}\right)$. Since Θ is finite, this type exists. Change this type's item to no disclosure followed by an object price as follows:

$$\left(r'\left(\overline{\theta}\right), \mathcal{X}'\left(\overline{\theta}\right), \mathcal{Q}'\left(\overline{\theta}\right), p'\left(\overline{\theta}\right)\right) = \left(0, \mathbb{E}[x], 1, \overline{T} + \overline{\theta} \cdot \left(\mathbb{E}[x] - \mathcal{X}\left(\overline{\theta}\right)\right)\right).$$

Keep all other items the same. In this menu, type $\overline{\theta}$ chooses the new item and always buys the object. This strategy gives him exactly the same payoff as that of the original menu:

$$\overline{\theta} \cdot \mathcal{X}'\left(\overline{\theta}\right) - p'\left(\overline{\theta}\right) = \overline{\theta} \cdot \mathcal{X}\left(\overline{\theta}\right) - p\left(\overline{\theta}\right) \mathcal{Q}\left(\overline{\theta}\right) - r\left(\overline{\theta}\right).$$

As $X \subseteq \mathbb{R}^J_{++}$, the uninformative experiment achieves a maximal attribute surplus, $\mathbb{E}[x] = \int_{x \in X} x dG > \mathcal{X}(\theta)$. As $\Theta \subseteq \mathbb{R}^J_{++}$, the new expected payment from type $\overline{\theta}$ is strictly higher than that in the original menu, $p'\left(\overline{\theta}\right) > \overline{T}$.

The new menu M' is not necessarily direct. The no disclosure item may be attractive to some types other than $\overline{\theta}$. However, the only profitable strategy under no disclosure is always buying. Such a deviation would only increase the seller's profit, as $\overline{T}$ was chosen to be the highest expected payment. Accordingly, menu M' brings strictly greater revenue than menu M, which is a contradiction.

Proof of Lemma 1. Set $\mathcal{F}$ is an image of a set of measures dominated by the prior distribution. That set of measures is compact in the weak* topology. Moreover, the corresponding map is continuous by the dominated convergence theorem, since the ex-ante attribute expectations exist. Therefore, $\mathcal{F}$ is compact.

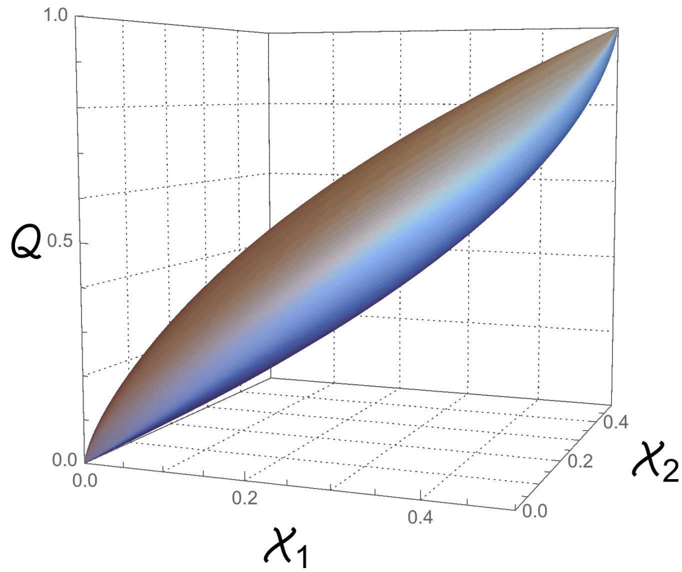

Figure 5: Feasibility set $\mathcal{F}$ of attribute surplus $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2)$ and trade probability $\mathcal{Q}$ for the case of two attributes distributed uniformly over a unit square $X = [0,1]^2$.

Set $\mathcal{F}$ is convex as a linear image of a convex set. Indeed, take any two points $(\mathcal{X}_1, \mathcal{Q}_1), (\mathcal{X}_1, \mathcal{Q}_2) \in \mathcal{F}$ and $\gamma \in [0,1]$. By construction, there exist trade functions $q_1, q_2 : X \to [0,1]$ that generate these two points. Then, the function $q_3 \triangleq \gamma q_1 + (1-\gamma) q_2$ is an admissible trade function that generates the attribute surplus:

$$\begin{aligned}\mathcal{X}_3(\theta) &= \int_{x \in X} x q_3(x)\, \mathrm{d}G(x) = \int_{x \in X} x\left(\gamma q_1(x) + (1-\gamma) q_2(x)\right) \mathrm{d}G(x) \\ &= \gamma \int_{x \in X} x q_1(x)\, \mathrm{d}G(x) + (1-\gamma)\int_{x \in X} x q_2(x)\, \mathrm{d}G(x) = \gamma \mathcal{X}_1(\theta) + (1-\gamma)\mathcal{X}_2(\theta).\end{aligned}$$

The same argument can be applied to the trade probability. Thus, $(\mathcal{X}_3, \mathcal{Q}_3)$ is a convex combination of $(\mathcal{X}_1, \mathcal{Q}_1)$ and $(\mathcal{X}_2, \mathcal{Q}_2)$, and it belongs to the feasibility set $\mathcal{F}$. Figure 5 illustrates the feasibility set for the case of two uniformly and independently distributed attributes.

Since $\mathcal{F}$ is a finite-dimensional closed set, the supporting hyperplane theorem (Rockafellar (1970), Theorem 11.6, Corollary 11.6.1, p. 100) can be applied.[21] A point $\left(\hat{\mathcal{X}}, \hat{\mathcal{Q}}\right)$ belongs to the boundary of $\mathcal{F}$ if and only if there are coefficients (λ, λ_0), not all zero, such that:

$$\left(\hat{\mathcal{X}}, \hat{\mathcal{Q}}\right) \in \arg\max_{(\mathcal{X}, \mathcal{Q}) \in \mathcal{F}} \lambda \cdot \mathcal{X} + \lambda_0 \mathcal{Q}.$$

[21]This step might fail if there are infinitely many attributes, $|J| = \infty$. If $\mathcal{F}$ has infinite dimensions, then it might have some boundary points that cannot be supported by a hyperplane. In such cases, a sufficient condition for the existence of a supporting hyperplane is that $\mathcal{F}$ has a nonempty interior.

By the definition of $\mathcal{F}$, the trade function $\hat{q}$ that generates the point $\left(\hat{\mathcal{X}}, \hat{\mathcal{Q}}\right)$ is such that:

$$\begin{aligned}\hat{q}\left(x\right) &\in \arg\max_{q:X\to[0,1]} \lambda \cdot \int_{x\in X} xq\left(x\right)\mathrm{d}G\left(x\right) + \lambda_0 \int_{x\in X} q\left(x\right)\mathrm{d}G\left(x\right) = \\ &\in \arg\max_{q:X\to[0,1]} \int_{x\in X} \left(\lambda\cdot x + \lambda_0\right) q\left(x\right)\mathrm{d}G\left(x\right).\end{aligned}$$

The integral is maximized pointwise. Any maximizer of it is a linear disclosure (22) with coefficients $\alpha = \lambda$ and $\alpha_0 = \lambda_0$.

Proof of Lemma 2. The seller's problem can be written in terms of attribute surpluses and trade probabilities (15). Consider any profile $\left(r\left(\theta\right), \mathcal{X}\left(\theta\right), \mathcal{Q}\left(\theta\right), p\left(\theta\right)\right)_{\theta\in\Theta}$ that satisfies constraints (16), (17), (18), (19), (20). Define functions $\mathcal{Q}' : \Theta \to [0,1]$ and $p' : \Theta \to \mathbb{R}_+$ as:

$$\begin{aligned}\mathcal{Q}'\left(\theta\right) &= \min_{Q:(\mathcal{X}(\theta),Q)\in\mathcal{F}} Q, \\ p'\left(\theta\right) &= \begin{cases}\frac{\mathcal{Q}(\theta)p(\theta)}{\mathcal{Q}'(\theta)}, & \text{if } \mathcal{Q}'\left(\theta\right) > 0, \\ p(\theta), & \text{if } \mathcal{Q}'\left(\theta\right) = 0.\end{cases}\end{aligned}$$

Intuitively, these functions are perturbations of the original mechanism that minimize trade probability while keeping the expected revenue fixed. By Lemma 1, $\mathcal{F}$ is compact, so for each $\theta \in \Theta$, $\mathcal{Q}'\left(\theta\right)$ is well defined and $\left(\mathcal{X}\left(\theta\right), \mathcal{Q}'\left(\theta\right)\right)$ belongs to the boundary of $\mathcal{F}$. At the same time, by the measurable maximum theorem (Aliprantis and Border (2006), Thm. 18.19) $\mathcal{Q}'$, and hence p', is a measurable function. By construction, for any type $\theta \in \Theta$, $\mathcal{Q}'\left(\theta\right) \leq \mathcal{Q}\left(\theta\right)$, $p'\left(\theta\right) \geq p\left(\theta\right)$, and $\mathcal{Q}'\left(\theta\right)p'(\theta) = \mathcal{Q}\left(\theta\right)p(\theta)$. As such, allocation $\left(\mathcal{X}\left(\theta\right), \mathcal{Q}'\left(\theta\right)\right)_{\theta\in\Theta}$ is implementable via tariff functions $r\left(\theta\right), p'\left(\theta\right)$ and satisfies all conditions of the lemma. The result follows.

Proof of Theorem 1. The seller's problem (15) can be seen as a maximization of a continuous function over a compact set. Therefore, an optimal menu exists. By Lemma 2, there exists an optimal menu with all allocations located on the boundary of the feasibility set $\mathcal{F}$. By Lemma 1, such allocations are achieved by linear disclosures.

Proof of Corollary 1. Define an auxiliary attribute x'_θ as the valuation of a type θ, $x'_\theta \triangleq v\left(\theta, x\right)$. By construction, the valuation of each type can be defined as $v'\left(\theta, x'\right) = x'_\theta$. This instance is a special case of the formulation (1). Thus, Theorem 1 applies and there exists an optimal menu with every experiment in it being a linear disclosure of auxiliary attributes x':

$$q(x') = \begin{cases} 1, & \text{if } \sum_{\theta \in \Theta} \alpha_\theta x'_\theta > \alpha_0, \\ 0, & \text{if } \sum_{\theta \in \Theta} \alpha_\theta x'_\theta < \alpha_0, \end{cases}$$

for $\alpha \in \mathbb{R}^{|\Theta|}, \alpha_0 \in \mathbb{R}$, not all zeros. In the original formulation, these are linear forms.

Calculations behind Location Payoffs Example. Consider a linear form. If $\alpha_1 + \alpha_2 \neq 0$, then the sum can be normalized to equal 1. The linear form can be rewritten as:

$$q(x) = \begin{cases} 1, & \text{if } -(x - (\alpha_1\theta_1 + \alpha_2\theta_2))^2 \gtrless \alpha'_0, \\ 0, & \text{if } -(x - (\alpha_1\theta_1 + \alpha_2\theta_2))^2 \lessgtr \alpha'_0, \end{cases}$$

with $\alpha'_0 = -v_0 + \alpha_0 + \alpha_1\alpha_2(\theta_1 - \theta_2)^2$ and the inequality sign depending on the sign of the original $\alpha_1 + \alpha_2$. This is a neighborhood disclosure with $\hat{\theta} = \alpha_1\theta_1 + \alpha_2\theta_2$ and $\alpha_1 + \alpha_2 = 1$.

If $\alpha_1 = \alpha_2 = 0$, then the linear form provides no disclosure and, as X is bounded, is equivalent to a neighborhood disclosure for a sufficiently large $|\alpha_0|$.

If $\alpha_1 + \alpha_2 = 0$ and $\alpha_1 \neq 0$, then the linear form is a linear disclosure:

$$q(x) = \begin{cases} 1, & \text{if } (\theta_1 - \theta_2) \cdot x \gtrless \alpha'_0, \\ 0, & \text{if } (\theta_1 - \theta_2) \cdot x \lessgtr \alpha'_0, \end{cases}$$

with $\alpha'_0 = \alpha_0/(2\alpha_1) + (\theta_1^2 - \theta_2^2)/2$ and the inequality sign depending on the sign of α_1. However, the proof of Lemma 1 established that the attribute surplus and probability achieved by a linear form with parameters $(\alpha_1, \alpha_2, \alpha_0)$ correspond to a boundary point of $\mathcal{F}$ in the auxiliary attributes, supported by the hyperplane orthogonal to the vector $(\alpha_1, \alpha_2, \alpha_0)$. If $\theta_1 \neq \theta_2$, then $\mathcal{F}$ has a strict interior. Thus, the set of boundary points supported by hyperplanes with $\alpha_1 + \alpha_2 = 0$ has a measure of zero.

Proof of Theorem 2. The argument is given in the text. The only difference from the standard problems is that $\mathcal{X}$ can take values in $[0, \mathbb{E}[x]]$, not in $[0, 1]$. However, this difference does not affect the extreme nature of the solution.

Lemma 4. (Directional Decomposition)
Let $(x_1, \ldots, x_J)$ be J attributes independently distributed over $X \subseteq \mathbb{R}^J$ according to distributions $G_1, \ldots, G_J$. Let $E = (S, \pi)$ be an arbitrary experiment. Let $(\mu(s, E), \Pr(s, E))$ be the belief distribution induced by E so that $\mu(s, E)$ is a distribution over X conditional on s given E. Denote by $\mu_j(s, E)$ the jth marginal distribution of $\mu(s, E)$. Then, there exist ex-

periments $\{E_j\}_{j=1}^J$ *such that:* $E_j = (S, \pi_j)$ *induces a belief distribution* $(\mu(s, E_j), \Pr(s, E_j))$ *with* $\mu(s, E_j) = (\mu_j(s, E), G_{-j})$ *and* $\Pr(s, E_j) = \Pr(s, E)$ *for all* $s \in S$*.*

Proof. The proof is constructive. Introduce dummy variables $(x'_1, \ldots, x'_J)$ which are distributed as $(x_1, \ldots, x_J)$ but are drawn independently of them. For a given j, construct E_j as an experiment that provides information about the vector $\left(x_j, x'_{-j}\right)$ according to π. By construction, E_j induces the same marginal distribution of beliefs about attribute j. However, since $(x_1, x'_1, \ldots, x_J, x'_J)$ are independent, it provides no information about other attributes. $\square$

Proof of Proposition 3. Consider an arbitrary responsive experiment $E_j(\theta_j)$. By Lemma 4, there exists a linear disclosure $E'_j(\theta_j)$ such that $\mathcal{X}'_j(\theta_j) = \mathcal{X}_j(\theta_j)$, $\mathcal{Q}\left(E'_j(\theta_j)\right) = \mathcal{Q}(E_j(\theta_j))$, and $\mathcal{X}'_k(\theta_j) = \mathbb{E}[x_k]$ for all $k \neq j$. Replacing $E_j(\theta_j)$ with $E'_j(\theta_j)$ does not change the incentive compatibility within cohort Θ_j, but by Blackwell's Theorem, it relaxes the incentive compatibility of other cohorts. By Theorem 1, the result follows.

Proof of Lemma 3. Define the expected transfer function as $T_j(\theta_j) \triangleq \mathcal{Q}_j(\theta_j) p_j(\theta_j)$. I can use standard one-dimensional arguments within each cohort to establish the connection between the attribute surplus and the expected transfer function:

$$T_j(\theta_j) = \theta_j \mathcal{X}_j(\theta_j) - \int_0^{\overline{\theta}_j} \mathcal{X}_j(z)\, \mathrm{d}z.$$

Individual rationality and incentive compatibility within each cohort are satisfied by construction. However, deviations between different cohorts impose additional constraints,

$$\int_0^{\overline{\theta}_j} \mathcal{X}_j(\theta_j)\, d\theta_j \geq \overline{\theta}_j \mathbb{E}[x_j] - \underline{p},$$

where $\underline{p}$ is the minimal object price in the menu $\{\mathcal{X}_j\}_{j=1}^J$. The deviations from all other types $\theta_j \in \Theta$ follow because the indirect utility function is convex and grows slower than $\theta_j \mathbb{E}[x_j]$. Applying double integration to the objective function completes the derivation.

Proof of Theorem 3. The argument in the text establishes that all high types are offered the minimal price. The optimal mechanism should then solve the problem (36) with the additional constraints that all high types are offered the same fixed price $\underline{p}^*$ and are served the fixed attribute surplus $\mathcal{X}_j^*\left(\overline{\theta}_j\right)$. These constraints can be written as:

$$\int_0^{\overline{\theta}_j} \mathcal{X}_j\left(\theta_j\right) d\theta_j = \mathcal{X}_j\left(\overline{\theta}_j\right) - \underline{p}^* \mathcal{Q}_j\left(\mathcal{X}_j\left(\overline{\theta}_j\right)\right),$$
$$\mathcal{X}_j\left(\overline{\theta}_j\right) = \mathcal{X}_j^*\left(\overline{\theta}_j\right).$$

Consider a relaxed problem with the original integral constraints and the monotonicity constraints dropped. In this problem, by Luenberger (1969) (Chapter 8, Theorem 1), there exist Lagrange multipliers $\{\lambda_j\}$ such that the optimal $\mathcal{X}_j\left(\theta_j\right)$ maximize the Lagrange function:

$$\mathcal{L} \sim \sum_{j=1}^{J} f\left(\Theta_j\right) \int_0^{\overline{\theta}_j} \left(\theta_j - \frac{1-F_j\left(\theta_j\right)}{f_j\left(\theta_j\right)} - \lambda_j\right) \mathcal{X}_j\left(\theta_j\right) \mathrm{d}F_j\left(\theta_j\right),$$

over a domain $\mathcal{X}_j\left(\theta_j\right) \in \left[0, \mathcal{X}^*\left(\overline{\theta}_j\right)\right]$. If all type distributions have the monotone hazard rate property, then the integrands increase in θ_j. Therefore, the optimal $\mathcal{X}_j\left(\theta_j\right)$ are bang-bang: $\mathcal{X}_j\left(\theta_j\right) = 0$ for $\theta_j < \theta_j^*$, $\mathcal{X}_j\left(\theta_j\right) = \mathcal{X}^*\left(\overline{\theta}_j\right)$ for $\theta_j > \theta_j^*$. This solution corresponds to a single item per each attribute cohort and satisfies the relaxed constraints.

Proof of Theorem 4. I begin by characterizing the optimal mechanism in the case of orthogonal types, since formally it is not covered by 3. The class of orthogonal types features particularly tractable incentive constraints. If type θ_j misreports, then he is offered an experiment tailored to another orthogonal type that is hence not informative about attribute j. Thus, the type has no reason to act on the experiment realization, and the tightest incentive-compatibility constraint is one in which he always buys. All others can be dropped. The seller's problem can be written as:

$$\max_{\{r_j, \mathcal{X}_j, p_j\}_{j=1}^J} \sum_{j=1}^{J} f\left(\theta_j\right)\left(r_j + \mathcal{Q}_j p_j\right) \tag{41}$$

$$\text{s.t.} \quad \mathcal{X}_j - p_j \mathcal{Q}_j - r_j \geq \mathbb{E}\left[x_j\right] - p_k - r_k, \ \forall j, k = 1, \ldots, J, \tag{42}$$

$$\mathcal{X}_j - p_j \mathcal{Q}_j - r_j \geq 0, \ \mathcal{X}_j \in \left[0, \mathbb{E}\left[x_j\right]\right], \ \mathcal{Q}_j = \mathcal{Q}_j\left(\mathcal{X}_j\right), \ \forall j = 1, \ldots, J. \tag{43}$$

I show that the solution to this problem does not feature price discrimination. First, by the arguments of Theorem 3, upfront payments can be without loss of revenue set to zero, $r_j \equiv 0$. Second, consider an arbitrary solution to the seller's problem such that $r_j \equiv 0$ and define $\underline{p} = \min_j \{p_j\}$. Toward the contradiction, assume that $p_j > \underline{p}$ for some j.

If $\mathbb{E}[x_j] \geq \underline{p}$, then the incentive-compatibility constraint is binding. Hence, $\mathcal{Q}_j p_j = \mathcal{X}_j - \mathbb{E}[x_j] + \underline{p}$. For small $\varepsilon > 0$, consider a modified mechanism with $\mathcal{X}_j' = \mathcal{X}_j + \varepsilon$, $\mathcal{Q}_j' p_j' =$

$\mathcal{X}_j' - \mathbb{E}[x_j] + \underline{p}$. Because $\mathcal{Q}_j(\mathcal{X}_j)$ is continuous, the mechanism remains incentive compatible yet brings higher revenue, which is a contradiction. Moreover, note that as $\mathcal{X}_j' > \mathcal{X}_j$, $\alpha_{0j}' < \alpha_{0j}$; hence, $\mathcal{X}_j'/\mathcal{Q}_j' = \mathbb{E}\left[x_j \mid x_j \geq \alpha_{0j}'\right] < \mathbb{E}\left[x_j \mid x_j \geq \alpha_{0j}\right] = \mathcal{X}_j/\mathcal{Q}_j$ and $\mathcal{Q}_j' > \mathcal{Q}_j$. Thus, $p_j' < p_j$.

If $\mathbb{E}[x_j] < \underline{p}$, then the individual-rationality constraint is binding. Hence, $\mathcal{Q}_j p_j = \mathcal{X}_j$. For small $\varepsilon > 0$, consider the modified mechanism with $p_j' = p_j - \varepsilon$, $\mathcal{X}_j'/\mathcal{Q}_j' = \mathcal{X}_j/\mathcal{Q}_j - \varepsilon$. The mechanism remains incentive compatible yet brings higher revenue, which is a contradiction.

Now, consider the optimal disclosure for a given object price. According to feasibility and individual rationality, $\mathcal{X}_j/\mathcal{Q}_j \geq \max\{p, \mathbb{E}[x_j]\}$. If $\mathcal{X}_j/\mathcal{Q}_j > \max\{p, \mathbb{E}[x_j]\}$, then for small $\varepsilon > 0$, the mechanism with $\mathcal{X}_j'/\mathcal{Q}_j' = \mathcal{X}_j/\mathcal{Q}_j - \varepsilon$ is incentive compatible and increases trade probability, $\mathcal{Q}_j' > \mathcal{Q}_j$, and consequently, revenue. This is a contradiction.

These arguments establish that in the case of orthogonal types an optimal mechanism sets $r_j = 0$, $p_j = p$, and E_j is a binary monotone partition of x_j such that $\mathbb{E}\left[x_j \mid E_j, s^+\right] = \max\{p, \mathbb{E}[x_j]\}$ for all $j = 1, \ldots, J$.

Now, consider a general setting. Introduce auxiliary attributes as in the proof of Corollary 1. By the arguments above, $M^{SP}(p^*)$ solves a relaxed problem in which the constraints (16) and (17) are dropped. If $M^{SP}(p^*)$ is incentive compatible, then these relaxed constraints are satisfied and, thus, $M^{SP}(p^*)$ solves an original problem.

Proof of Corollary 2. For a profile $(\theta_1, \ldots, \theta_J)$, denote by $p^*(\theta_1, \ldots, \theta_J)$ an optimal price in a separate persuasion mechanism, by $\alpha_{0j}(\theta_1, \ldots, \theta_J)$ optimal persuasion thresholds, and by $\Pi^{SP}(p, \theta_1, \ldots, \theta_J)$ the revenue function. As attributes are continuously distributed, $\Pi^{SP}(\cdot)$ is continuous. If $p^*(\theta_1, \ldots, \theta_J)$ is a singleton, then, by the Maximum Theorem, $p^*(\cdot)$ and $\alpha_0(\cdot)$ are continuous functions in a neighborhood of $(\theta_1, \ldots, \theta_J)$.

Fix an orthogonal type profile $\hat{\Theta} = \left(\hat{\theta}_1, \ldots, \hat{\theta}_J\right)$. If $\mathbb{E}\left[\hat{\theta} \cdot x\right] \neq p\left(\hat{\theta}_1, \ldots, \hat{\theta}_J\right)$ for all $\hat{\theta} \in \hat{\Theta}$, then the types, when deviating, strictly prefer to not act contrary to their no-information action. As $\hat{\Theta} \geq 0$, the mismatching strategies are irrelevant; moreover, for all $\hat{\theta}_j, \hat{\theta}_k \in \hat{\Theta}$, $k \neq j$:

$$\mathbb{E}\left[\hat{\theta}_j \cdot x \mid \hat{\theta}_k \cdot x > \alpha_{0k}\left(\hat{\theta}_1, \ldots, \hat{\theta}_J\right)\right] < p^*\left(\hat{\theta}_1, \ldots, \hat{\theta}_J\right), \text{ if } \mathbb{E}\left[\hat{\theta}_j \cdot x\right] < p,$$
$$\mathbb{E}\left[\hat{\theta}_j \cdot x \mid \hat{\theta}_k \cdot x < \alpha_{0k}\left(\hat{\theta}_1, \ldots, \hat{\theta}_J\right)\right] > p^*\left(\hat{\theta}_1, \ldots, \hat{\theta}_J\right), \text{ if } \mathbb{E}\left[\hat{\theta}_j \cdot x\right] > p.$$

Now, replace orthogonal $\hat{\Theta}$ by a generic Θ. As long as all types in Θ are positive, the mismatching strategies remain irrelevant. As attributes are continuously distributed, in some neighborhood of $\hat{\Theta}$ both sides of the inequalities are continuous and the constraints remain satisfied. Hence, the buyer cannot benefit from misreporting, and the result follows from Theorem 4.

Proof of Proposition 4. Consider any optimal menu and a type θ with $r(\theta) > 0$. Let $r'(\theta) = r(\theta) - \varepsilon$, $p'(\theta) = p(\theta) + \varepsilon / \mathcal{Q}(\theta)$ for a small $\varepsilon > 0$. If incentive compatible, this modification preserves the seller's revenue and the buyer's payoff. The constraint (16) remains the same. The constraint (18) is relaxed and is strictly so when $\mathcal{Q}(\theta) < 1$. The constraint (19) is satisfied by the individual rationality. For the constraint (17):

$$r'(\theta) + (1 - \mathcal{Q}(\theta)) p'(\theta) = r(\theta) + (1 - \mathcal{Q}(\theta)) p(\theta) + \varepsilon \frac{1 - 2\mathcal{Q}(\theta)}{\mathcal{Q}(\theta)}.$$

For the modification to violate incentive constraints, it must be that $1 - 2\mathcal{Q}(\theta) < 0$ and the constraint (17) binds for some type θ'.

Proof of Corollary 3. I present the arguments for the case when the higher signal is sent with probability one at the threshold. The cases with randomization are analogous.
i.) Whenever $\alpha_j \geq 0$:

$$\mathbb{E}[x_j \mid \alpha x \geq \alpha_0] = \mathbb{E}\left[x_j \mid \alpha_j x_j \geq \alpha_0 - \sum_{k \neq j} \alpha_k x_k\right] \geq \mathbb{E}\left[x_j \mid \alpha_j x_j < \alpha_0 - \sum_{k \neq j} \alpha_k x_k\right].$$

$$\mathbb{E}[v(\theta, x) \mid \alpha x \geq \alpha_0] = \sum_{j=1}^{J} \theta_j \mathbb{E}[x_j \mid \alpha x \geq \alpha_0] \geq \sum_{j=1}^{J} \theta_j \mathbb{E}[x_j \mid \alpha x < \alpha_0] = \mathbb{E}[v(\theta, x) \mid \alpha x < \alpha_0].$$

ii.) For any increasing function $v(\theta, \cdot)$, and $x_0 \in \mathbb{R}$, $\mathbb{E}[v(\theta, x) \mid x \geq x_0] \geq \mathbb{E}[v(\theta, x) \mid x < x_0]$.

References

Aliprantis, C. and K. Border (2006): "Infinite Dimensional Analysis," .

Anderson, S. P. and R. Renault (2006): "Advertising Content," *American Economic Review*, 96, 93–113.

Bar-Isaac, H., G. Caruana, and V. Cuñat (2010): "Information Gathering and Marketing," *Journal of Economics & Management Strategy*, 19, 375–401.

Bergemann, D., A. Bonatti, and A. Smolin (2018): "The Design and Price of Information," *American Economic Review*, 108, 1–48.

Bergemann, D., F. Castro, and G. Weintraub (2020): "The Scope of Sequential Screening with Ex-Post Participation Constraints," *Journal of Economic Theory*, 105055.

Bergemann, D. and M. Pesendorfer (2007): "Information Structures in Optimal Auctions," *Journal of Economic Theory*, 137, 580–609.

Bergemann, D., J. Shen, Y. Xu, and E. Yeh (2012): "Multi-Dimensional Mechanism Design with Limited Information," in *Proceedings of the 13th ACM Conference on Electronic Commerce*, ACM, 162–178.

Blackwell, D. (1953): "Equivalent Comparisons of Experiments," *Annals of Mathematical Statistics*, 24, 265–272.

Carville, O. (2018): "ZipRecruiter Is Valued at $1.5 Billion in a Bet on AI Hiring," *Bloomberg.com*, accessed at https://www.bloomberg.com/news/articles/2018-10-04/ziprecruiter-is-valued-at-1-5-billion-in-a-bet-on-ai-hiring.

Chakraborty, A. and R. Harbaugh (2010): "Persuasion by Cheap Talk," *American Economic Review*, 100, 2361–82.

Courty, P. and H. Li (2000): "Sequential Screening," *Review of Economic Studies*, 67, 697–717.

Daskalakis, C., A. Deckelbaum, and C. Tzamos (2017): "Strong Duality for a Multiple-Good Monopolist," *Econometrica*, 85, 735–767.

Doval, L. and J. C. Ely (2020): "Sequential information design," *Econometrica*, 88, 2575–2608.

Dubé, J.-P. and S. Misra (2019): "Personalized Pricing and Customer Welfare," *Discussion Paper*.

Dworczak, P. (2020): "Mechanism Design with Aftermarkets: Cutoff Mechanisms," *Discussion paper*.

Dworczak, P. and G. Martini (2019): "The Simple Economics of Optimal Persuasion," *Journal of Political Economy*, 127.

Dye, R. A. (1985): "Disclosure of Nonproprietary Information," *Journal of Accounting Research*, 123–145.

Eső, P. and B. Szentes (2007): "Optimal Information Disclosure in Auctions and the Handicap Auction," *Review of Economic Studies*, 74, 705–731.

——— (2017): "Dynamic Contracting: An Irrelevance Theorem," *Theoretical Economics*, 12, 109–139.

Gabaix, X. and D. Laibson (2006): "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets," *Quarterly Journal of Economics*, 121, 505–540.

Hartline, J. (2020): "Mechanism Design and Approximation," Accessed at http://jasonhartline.com/MDnA/.

Heumann, T. (2020): "Information Design and Sequential Screening with Ex Post Participation Constraint," *Theoretical Economics*, 15, 319–359.

Johnson, J. P. and D. P. Myatt (2006): "On the Simple Economics of Advertising, Marketing, and Product Design," *American Economic Review*, 96, 756–784.

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

Koessler, F. and R. Renault (2012): "When Does a Firm Disclose Product Information?" *The RAND Journal of Economics*, 43, 630–649.

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a Privately Informed Receiver," *Econometrica*, 85, 1949–1964.

Krähmer, D. (2020): "Information Disclosure and Full Surplus Extraction in Mechanism Design," *Journal of Economic Theory*, 105020.

Krähmer, D. and R. Strausz (2015a): "Ex Post Information Rents in Sequential Screening," *Games and Economic Behavior*, 90, 257–273.

——— (2015b): "Optimal Sales Contracts with Withdrawal Rights," *Review of Economic Studies*, 82, 762–790.

Lancaster, K. J. (1966): "A New Approach to Consumer Theory," *Journal of Political Economy*, 74, 132–157.

Lehmann, D., L. I. Oćallaghan, and Y. Shoham (2002): "Truth Revelation in Approximately Efficient Combinatorial Auctions," *Journal of the ACM*, 49, 577–602.

Lewis, T. R. and D. E. Sappington (1994): "Supplying Information to Facilitate Price Discrimination," *International Economic Review*, 309–327.

Li, H. and X. Shi (2017): "Discriminatory Information Disclosure," *American Economic Review*, 107, 3363–85.

Luenberger, D. G. (1969): *Optimization by Vector Space Methods*, John Wiley & Sons.

Myerson, R. B. (1981): "Optimal Auction Design," *Mathematics of Operations Research*, 6, 58–73.

——— (1982): "Optimal Coordination Mechanisms in Generalized Principal-Agent Problems," *Journal of Mathematical Economics*, 10, 67–81.

Rayo, L. and I. Segal (2010): "Optimal Information Disclosure," *Journal of Political Economy*, 118, 949–987.

Riley, J. and R. Zeckhauser (1983): "Optimal Selling Strategies: When to Haggle, When to Hold Firm," *Quarterly Journal of Economics*, 98, 267–290.

Rochet, J.-C. and P. Choné (1998): "Ironing, Sweeping and Multidimensional Screening," *Econometrica*, 66, 783–826.

Rockafellar, R. T. (1970): *Convex Analysis*, Princeton University Press.

Samuelson, W. (1984): "Bargaining under Asymmetric Information," *Econometrica*, 995–1005.

Wei, D. and B. S. Green (2019): "(Reverse) Price Discrimination with Information Design," *Discussion paper*.