Disclosure and Pricing of Attributes*

Alex Smolin

January 19, 2019

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 disclose to the buyer arbitrary attribute information in the form of a statistical experiment. The seller decides how to price the object, what information to disclose, and how to price access to the information. To screen different types, the seller offers a menu of options that specify information prices, experiments, and object prices.

I characterize revenue-maximizing menus. If all types value the same attribute, then the seller cannot benefit from information disclosure and price discrimination. More generally, if each type values a single attribute and attributes are independent, then the seller can benefit from information disclosure but not from price discrimination. In other cases, a discriminatory menu can be profitable; however, optimal experiments always belong to a tractable class of linear disclosure policies. The analysis informs the operation of various intermediaries including business brokers and online recruiting platforms.

Keywords: attributes, call options, demand transformation, information design, intermediaries, linear disclosure, mechanism design, multidimensional screening, persuasion

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

*Institute for Microeconomics, University of Bonn, Lennéstraße 37, 53113, Bonn, Germany, alexey.v.smolin@gmail.com. I thank Alessandro Bonatti, Laura Doval, Piotr Dworczak, Sergei Izmalkov, Stephan Lauermann, Marta Troya Martinez, Xiaosheng Mu, Vasiliki Skreta, Andrzej Skrzypacz, Juuso Välimäki, and Andy Zapechelnyuk for helpful comments, and, especially, Dirk Bergemann and Daniel Krähmer for extended and productive conversations. I am grateful to the audiences of research seminars at Toulouse, Mannheim, Carlos III Madrid, Bonn, Aalto, Oxford, Emory, Essex, Nottingham, Bologna, Edinburgh, Warwick, Amsterdam, as well as at the SAET 2017, ESSET Gerzensee, Stony Brook, EEA-ESEM 2018, EARIE 2018, and Columbia/Duke/MIT/Northwestern IO Theory conferences.

1 Introduction

In many important markets, sellers have considerable control over information available to their buyers. Business brokers can control the extent of the firm investigation and documentation they supply, car retailers can limit the duration of test drives and amount of technical pre-sale support, and recruiting platforms can decide what parts of a job candidate's profile to reveal to employers. In all of these markets, the products (i.e., business, car, or meeting with a job candidate) 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 an operation of Ziprecruiter.com, a major online recruiting platform. The platform facilitates matching between job seekers and employers. Employers subscribe to the platform to advertise their open vacancies and obtain access to a large database of resumes. Ziprecruiter.com is actively innovating and experimenting with its algorithms and pricing. Just in October 2018, the platform raised $156 million investment to improve its matching technology, putting it at $1.5 billion valuation.[1] Currently, the platform employs a nonlinear pricing scheme for subscriptions, varying in the breadth of information provided and the ability to contact preferred candidates (Dubé and Misra (2017)).

The platform operates in the recruitment market that features substantial heterogeneity on both sides. The candidate profiles vary along many attributes, including work experience, education levels, technical skills, and standardized tests scores. The employers belong to distinct types, such as tech start-ups, chain stores, investment banks, or government agencies. Naturally, different types of employers are looking for different attributes in their candidates and, hence, differ in their willingness to pay to contact the same candidate.

Ziprecruiter.com has access to a large amount of data about the prospective candidates and facilitates employment matching by providing this data to employers. It can decide what information to provide and at what price. By programming its algorithms, the platform can deny access to some attributes in the data; alternatively, it can provide coarse statistics (e.g., instead of showing a full GPA, it can reveal only whether it surpasses a particular threshold). My goal is to study the trade-offs the platform is facing, to inform the revenue-maximizing design, and to evaluate allocation distortions introduced by the information control of the intermediary.

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

In this paper, I develop a framework to study information disclosure and pricing of multi-attribute 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 seller does not know what attributes the buyer likes or how much. These preferences are the buyer's private information and constitute the buyer's type. The seller controls pricing and, importantly, can disclose attribute information to the buyer. The players are Bayesian decision makers.

Both the object and the information about its attributes are valuable for a buyer, and I allow the seller to price them jointly. The seller offers a menu of options that differ in their informativeness. Each option consists of an information price paid upfront, an attribute information, and a strike 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. This menu mechanism provides a natural and practical framework for analyzing information disclosure and pricing together.

I study and characterize revenue-maximizing menus. The general revenue-maximization problem features information design and multidimensional screening with monetary transfers. As such, it entails two main methodological challenges. First, the class of all stochastic experiments is large. Not only can each experiment send many signals, but also the underlying uncertainty of the attribute vector generates a continuum of possible states, each having multiple dimensions. To understand distortions driven by the information design, it is important to pin down the structure of optimal experiments. Second, multidimensional screening problems are notoriously difficult. In the absence of a single-dimensional structure, it is unclear what incentive constraints are relevant for optimal design. This difficulty is further exacerbated by the presence of information disclosure, because different buyer types can respond differently to the same information.

I progress in both directions in turn. In Section 3, I study the design of disclosure policies. 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 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 of experiments—linear disclosures; there exists an optimal menu that contains only them. A linear disclosure informs whether a linear combination of attributes is above or below some threshold. Effectively, it splits the attribute space into two half-spaces and informs the buyer to which half-space the object belongs. A linear disclosure is nonstochastic almost everywhere and can be seen as informing the buyer about the valuation of a virtual type. It is a generalization of binary monotone partitions to multidimensional settings.[2] Notably, the result requires no assumptions on distribution of types or attributes.

In Section 4, 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. This result may look surprising. After all, the buyer has private information and the seller can separate the buyer types by designing complex menus. However, such menus cannot outperform a no-disclosure posted-price mechanism. 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 obtained a scaled revenue. By the martingale property of Bayesian expectations, the seller can obtain the same revenue by providing no information.[3]

The case of several attributes is qualitatively different because types can be differentiated not only vertically but also horizontally. As a result, an optimal allocation may depend on attribute realizations. Intuitively, the seller should allocate the object to types who value the realized attributes the most. To guide the allocation, she should provide some attribute information. As a result, an optimal menu can involve discriminatory pricing and disclosure.

In Section 4.3, I introduce and study the setting of a single-minded buyer. In this setting, attributes are independently distributed, and a buyer values only one attribute, but the seller does not know which one or how much. This setting allows for both horizontal and vertical heterogeneity but is sufficiently tractable. I start with a simpler case of orthogonal types, in which each type values a distinct attribute so that type valuations are independent. In Theorem 3, I show that an optimal menu features partial disclosure but no price discrimination. Moreover, the menu can be implemented by a nondiscriminatory mechanism—posting a single price for the object and informing the buyer whether the object is sufficiently good along each attribute. In Theorem 4, I generalize this finding to the case of a continuum of types valuing each attribute. I show that if the type distributions are log-concave, then the nondiscriminatory mechanism with partial disclosure remains optimal.

I conclude with discussion of my findings in Section 5. First, I illustrate how optimal disclosure transforms demand curves. I argue that by providing partial attribute information, the seller can target types within a specific range of valuations and hence rotate the demand

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

[3] This intuition points 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.

curve locally. It resonates with the analysis of global demand rotations by Johnson and Myatt (2006). Second, I show full disclosure is detrimental to the seller and, moreover, the seller may benefit from conditioning the price on the information disclosed. These results contrast with those of Eső and Szentes (2007) and highlight the qualitative difference between our frameworks.

Related Literature This paper contributes to the literature on private disclosure and pricing. One strand of this literature focuses on nondiscriminatory mechanisms—in which a seller provides a single disclosure. Lewis and Sappington (1994) introduce these mechanisms in a setting where a buyer has no prior information. They find optimal disclosure within a simple parameterized class and show that it is generally 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 because it allows the seller to extract the full expected surplus.[4] Johnson and Myatt (2006) extend the analysis to settings in which the buyer has prior information. They focus on disclosures that correspond to global rotations of a demand curve and show, once again, that extreme disclosures are optimal. My paper contributes to this literature by showing that if the product has several attributes, then a single multipartition disclosure can dominate both full and no disclosure, even if there is common knowledge of positive trade gains (Section 5.1).

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 the 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 these settings, the information disclosure can be seen as "valuation-rank" disclosure that corresponds to statements such as, "Your valuation is in your x-th percentile," with x being the same for all types. The authors show that in such settings, under certain distributional assumptions, the seller may optimally provide full disclosure and does not benefit from conditioning the price on the information disclosed.[5] However, Li and Shi (2017) show that these findings do not hold in common value settings in which the types represent private information about the object. In those settings, the information disclosure can be seen as "valuation-level" disclosure that corresponds to statements such as, "Your valuation is above x," with x being the same for all types. Li and Shi (2017) establish that in such settings, the seller should withhold some information but are

[4]See, however, Anderson and Renault (2006), who show that optimal disclosure is partial if the purchase is associated with search costs and the seller cannot commit to prices.

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

not able to identify optimal mechanisms.

All of this previous literature operates in single-dimensional settings. Under complete object information, when comparing any two objects, all buyer's types agree on their ranking. However, in practice, many products are multidimensional with different attributes appealing to different buyers. In this paper, I demonstrate that these settings can be successfully studied within a framework of attribute disclosure and lead to qualitatively different results. Despite the richness of the attribute space, optimal experiments belong to a tractable class of linear disclosures (Section 3.4). Optimal mechanisms feature partial disclosure but can be remarkably simple (Section 4.3). The seller can strictly benefit from conditioning the price on the information disclosed (Section 5.2).

Information design with screening and monetary transfers appears in my previous work (Bergemann, Bonatti, and Smolin (2018)). There, the seller offers a menu of information products to a buyer who seeks this information to resolve an exogenous decision problem; his action is not contractable. In contrast, in this paper, the seller's and buyer's problems are intertwined and situated within a multi-attribute framework. The seller can price both the information and the buyer's decision to buy the object. As a result, in many settings, the seller is willing to provide information free of charge (Section 4.3).

Finally, this paper builds on several existing frameworks. Multi-attribute buyer's valuation follows the characteristic model of Lancaster (1966). The mechanism timing is analogous to the sequential screening of Courty and Li (2000). An unrestricted search for a disclosure policy to optimally influence a single agent is a defining feature of Bayesian persuasion literature (Rayo and Segal (2010), Kamenica and Gentzkow (2011)). The screening analyses of single-attribute and single-minded-buyer settings build on the mechanism design machinery of Myerson (1981, 1982).

2 Model

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

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

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 with full support. The buyer and the seller are symmetrically informed about them. The type space Θ can be finite or infinite. The buyer's type is his privately known tastes, uncorrelated with attributes. From the seller's perspective, the types are distributed according to a cumulative distribution function F. Until Section 4, I do not impose any structural assumptions on the attribute and type distributions. The only technical 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:

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

The experiment can be arbitrarily informative about the attributes. It 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 probability 1 on $s = x$. Alternatively, it can provide partial information, for example, as illustrated in Figure 1. In this figure, upon observing a signal s_1, the buyer learns that attributes belong to the red area but does not know to which part of it.

I highlight that attribute information affects valuation of different types differently according to the valuation function (1). For example, if attributes are independent, then the experiment informative about a subset of attributes is valuable only for those types who place non-zero weights on those attributes. Consequently, attribute information cannot be represented by an experiment that informs the buyer directly about his valuation.[6] Doing so would change the buyer incentives in his choice across experiments.

Selling Mechanism The seller has effectively two products valuable for a buyer—the object itself and the information about its attributes. To investigate the scope of screening, I allow the seller to price both the object and the information she provides.

The seller designs a menu of items indexed by $i \in \mathcal{I}$:

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

to be offered to a buyer. It consists of a collection of experiments $E(i)$ and tariff functions $r(i) \geq 0$, $p(i) \geq 0$. The first tariff captures a price of information—an upfront payment paid irrespectively of a trade. The second tariff captures a price of the object—a strike price

[6]This differs from the disclosure models of Eső and Szentes (2007) and Li and Shi (2017).

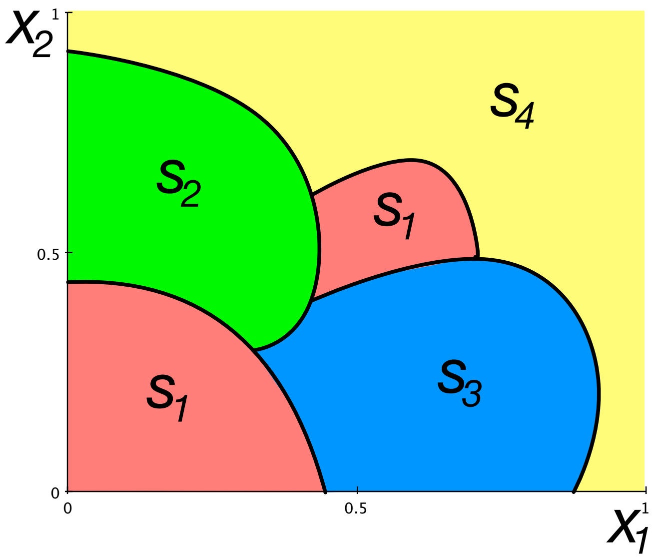

Figure 1: A partially informative experiment E with the attribute set $X = \mathbb{R}^2$ and the signal set $S = (s_1, s_2, s_3, s_4)$. Colors indicate regions in which the corresponding signals are sent.

paid only if the trade occurs.

Effectively, the menu is a collection of call options differing in monetary terms and information disclosure, designed to screen different buyer's types. The timing is as follows.[7] The seller posts a menu M. The attribute vector x and the buyer's type θ are realized. The buyer chooses an item $i \in \mathcal{I}$ and pays the corresponding price $r(i)$. He observes a signal s from the experiment $E(i)$ and decides whether to buy the object at the strike price $p(i)$. The payoffs are realized. The timing is illustrated in Figure 2.

The timing implies the seller commits to a menu before realization of the attributes x and the type θ. The true attributes x and experiment realization s are not contractible.[8] Sales are deterministic—the strike price once paid guarantees possession of the object. Sequential interactions between the players are excluded, so belief-elicitation schemes and scoring rules are not available.[9]

I highlight that no analog of a revelation principle is known in the environments in which the designer can privately provide additional information. As such, the posted-price menu mechanisms provide a natural and practical framework for studying how information disclosure and pricing interact in design problems. My goal is to characterize a revenue-

[7]The timing is analogous to that of Courty and Li (2000) and Li and Shi (2017).

[8]For instance, the buyer cannot claim a refund ex post. See Krähmer and Strausz (2015b), Heumann (2018) and Bergemann et al. (2017) for recent studies on ex-post incentive constraints.

[9]Krähmer (2017) investigates the usefulness of such schemes in screening problems with information design.

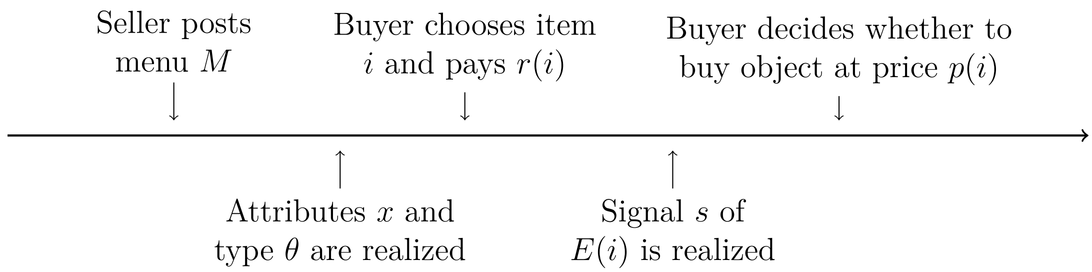

Figure 2: Timeline of the selling mechanism.

maximizing menu for the seller.

3 Design of Disclosure Policies

In this section, I proceed with studying the revenue-maximizing menu design. I begin with discussing the buyer's incentives and formalizing his choice in an arbitrary menu. I use this formalization to show that the design problem can be approached in two successive steps. First, it is possible to identify the class of optimal disclosure policies without explicitly characterizing a pricing mechanism. I show this in the current section. Second, one can build on this result and derive optimal pricing mechanisms in the leading settings. I do that in Section 4.

3.1 Buyer's Problem

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

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

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. Integrating over signal realizations, I can define the resulting *(total) trade probability* as:

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

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

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

The type θ chooses an option with the largest indirect utility. Naturally, types seek information that most fits their interests. For example, if a type values only one of many attributes, then that type places no information value on options that provide no information about that attribute. As a result, the types can disagree on the experiment ranking even if tariffs are the same. Faced with a menu, the types self-select different items. This gives the seller an opportunity to discriminate among them by carefully designing the menu.

3.2 Responsive Menus

The seller's problem lies at the intersection of 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, in the next couple of subsections, I show that an optimal class of experiments is simple and tractable.

I begin approaching this problem by binding the size of the optimal menus and signal sets. First, I appeal to the revelation principle and focus on *direct* menus:

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

with the size of the type space, which effectively ask the buyer his type and assign the experiment and the tariffs as functions of his report. Second, I follow the arguments of Bergemann, Bonatti, and Smolin (2018) to 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 $s = s^+$. Responsive experiments guide the buyer action. 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. □

The proof is analogous to the argument of the revelation principle of Myerson (1982). Intuitively, an experiment should provide information minimal to guiding the decision of a truth-telling type. 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 because the modified experiments are weakly less informative (Blackwell (1953)).

Proposition 1 puts the elementary structure on the exchange of information between the seller and the buyer. The buyer should inform the seller about his tastes. The seller should provide a recommendation whether to buy the object. The content of the recommendation should be chosen such that the buyer reports his tastes truthfully and obediently follows the recommendation.

Focus on responsive menus allows characterizing every experiment E in the menu 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 trade functions, one per each buyer's type. With a slight abuse of notation, I will refer to the trade function of type θ by $q(\theta, x)$.

3.3 Seller's Problem

Proposition 1 allows associating each experiment with its trade function (8) and writing the seller's problem in a standard mechanism design form. 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 strike price $p(\theta)$. The seller's problem is to maximize the total expected revenue over the tariff and trade functions:

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

subject to the incentive-compatibility constraints and individual rationality constraints.

The incentive-compatibility constraints require that for all $\theta, \theta' \in \Theta$:

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

where $\sigma(q(\theta',x),k)$ is a deviation function equal to $q(\theta',x)$, $1-q(\theta',x)$, 1, and 0 for $k = 1,\ldots,4$ respectively. These incentive-compatibility constraints ensure each type prefers truth telling to all double-deviating strategies: misreporting and following the recommendations, "swapping" the buying decisions, always buying, or never buying, respectively. 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 $\theta \in \Theta$:

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

There are several challenges involved in the seller's problem. First, the seller maximizes over a large class of stochastic experiments, captured by trade functions which are arbitrary functions from a multidimensional space X. Second, the buyer's type has no single-dimensional structure, and multidimensional screening problems are notoriously difficult. Third, the problem features an additional multiplicity of constraints caused by double deviations. It is a priori not clear what kinds of deviations are binding and, hence, relevant for the design problem.

The following observation is crucial to deal with the experiment complexity: not the whole trade function but only two coarse statistics matter for the revenue-maximizing problem. Namely, for a given responsive experiment E, say that the associated trade function q achieves the *attribute surplus*:

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

and the *(total) trade probability*:

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

It can be seen from the formulation (9), (10), (11) that these statistics are the only economically relevant parameters of the problem. A change in the trade function $q(\theta, x)$ that does not affect the attribute surplus and the total trade probability affects neither the buyer incentives nor the seller's revenue. Hence, instead of maximizing over the trade functions $q(\theta, x)$, the seller can maximize directly over attribute surpluses and trade probabilities, $\mathcal{X}(\theta)$ and $\mathcal{Q}(\theta)$.

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 attribute surpluses must be nil as well, because the trade never happens. At another extreme, if the trade probability is 1, then the attribute surplus is equal to its ex ante expectation $\mathbb{E}[x]$, because the trade always occurs. At the intermediate values of trade probabilities, there is more freedom of choosing attribute surpluses, because the seller can select at what regions the trade recommendation

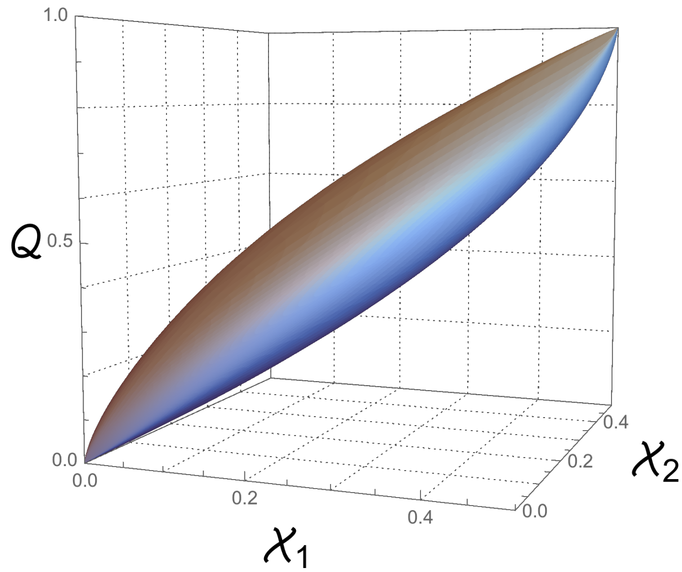

Figure 3: Feasibility set $\mathcal{F}$ of attribute surplus $\mathcal{X} = (\mathcal{X}_1, \mathcal{X}_2)$ and trade probability $\mathcal{Q}$ for the case of two attributes distributed uniformly over a unit square $X = [0,1]^2$.

is sent. Formally, define the feasibility set $\mathcal{F} \subseteq \mathbb{R}^{J+1}$ as:

$$\mathcal{F} \triangleq \{(\mathcal{X}, \mathcal{Q}) \mid \exists\, q : X \to [0,1] \text{ such that} \tag{14}$$
$$\mathcal{X} = \int_{x \in X} x q(x)\, dG(x) \text{ and } \mathcal{Q} = \int_{x \in X} q(x)\, dG(x) \Big\}.$$

The shape of the feasibility set is determined by the attribute distribution G. Figure 3 illustrates the feasibility set for the case of two uniformly and independently distributed attributes.

3.4 Optimal Disclosure

I begin with observing a special feature of a responsive experiment that always recommends the buyer to buy, $q(x) \equiv 1$. 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, $\Theta \subseteq \mathbb{R}^J_{++}$, it means that this experiment is also a unique maximizer of the buyer surplus. Indeed, if a buyer always positively values the object, then his surplus is maximized if he always buys it.

Moreover, this always-trade experiment provides no information about attributes. As such, it also maximally limits the scope for deviations. If the buyer willingly chooses an uninformative option then he is determined to always buy the object.

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

Proposition 2 highlights the distinctive feature of the seller's problem that combines information and mechanism design. In a typical information design problem, the payoff structure is exogenously fixed and there is no guarantee that no disclosure would appear in an optimal mechanism, irrespectively of prior distributions. Indeed, for any given prices, if the buyer tastes are sufficiently bland, the seller has to provide some minimal 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.

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

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

$$q(x) = \begin{cases} 1, & \text{if } \alpha \cdot x > \alpha_0, \\ 0, & \text{if } \alpha \cdot x < \alpha_0. \end{cases} \tag{15}$$

A linear disclosure informs the buyer whether a linear combination of attributes is above or below some threshold. Equivalently, its trade function is an indicator function of an attribute half-space. A linear disclosure is conditionally nonstochastic almost everywhere. In the case of a single attribute, a linear disclosure corresponds to a binary monotone partition disclosure.

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 a 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 "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

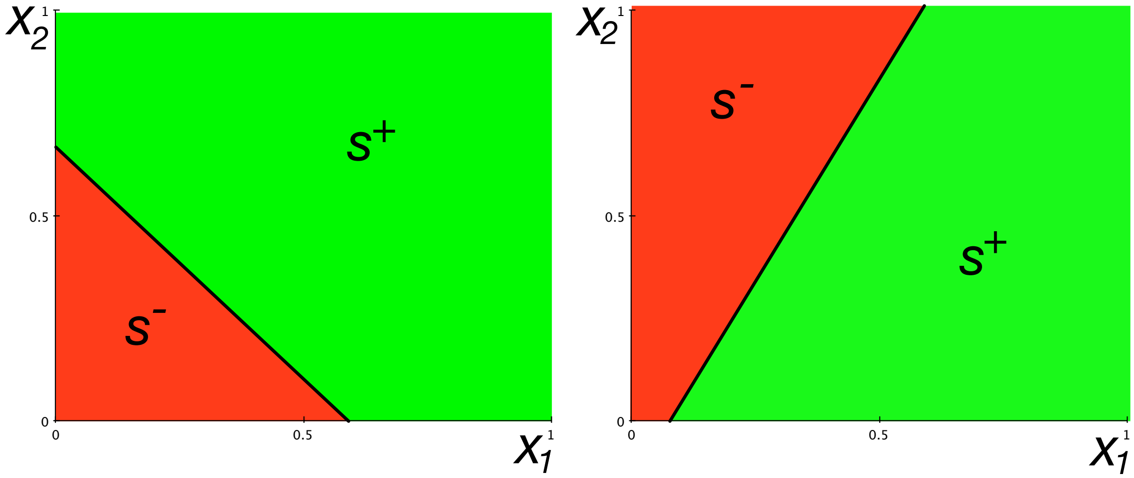

Figure 4: Linear disclosure in the case of two attributes, $J = 2$. Colors indicate regions in which the corresponding recommendations are sent. Left: level disclosure, $\alpha_1 > 0, \alpha_2 > 0$. Right: comparative disclosure, $\alpha_1 > 0, \alpha_2 < 0$.

them in the other group. In Figure 4, I illustrate these two types of linear disclosure in the case of two attributes.

Note that the likelihood function of a linear disclosure is not restricted on the defining hyperplane, $\{x \mid \alpha \cdot x = \alpha_0\}$. Any given parameters α, α_0 determine a class of linear disclosures that differ on the boundary. If attributes are continuously distributed, then this indeterminacy is irrelevant as it affects a zero probability event. However, if the attribute distribution G has a positive mass on the defining hyperplane, then the likelihood function should be additionally specified there. Furthermore, observe that the defining hyperplane does not exist for $\alpha \equiv 0$ and α_0 being strictly positive or negative. These linear disclosures correspond to never-trade and always-trade uninformative experiments, $q \equiv 0$ and $q \equiv 1$, respectively.

Given the standard topology on $\mathbb{R}^{J+1}$, denote the interior of the feasibility set by $\text{int}\,(\mathcal{F})$ and its boundary by $\partial\,(\mathcal{F})$.

Proposition 3. (Feasibility)
The feasibility set $\mathcal{F}$ is compact and convex. Its boundary $\partial\,(\mathcal{F})$ is spanned by linear disclosures. That is: (1) any linear disclosure achieves some boundary point of $\mathcal{F}$, and (2) any boundary point of $\mathcal{F}$ can be achieved by some linear disclosure.

The full proof of this central result is available in the Appendix. First, I show that the feasibility set is compact as a continuous image of a compact set. Second, I show that the feasibility set 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. The result follows.

Proposition 3 establishes importance of linear disclosures for attribute surpluses and trade probabilities. It also highlights the qualitative difference between the boundary and interior points of the feasibility set $\mathcal{F}$. Any boundary point is achieved by a linear disclosure. Any interior point is achieved by a stochastic combination of two linear disclosures. Furthermore, one can obtain an immediate corollary relating the trade expectation and the trade probability. To this end, for a given responsive experiment E, define the *trade expectation* as:

$$Y \triangleq \mathbb{E}\left[x \mid E, s^{+}\right] = \frac{\int_{x\in X} xq\left(x\right) \mathrm{d}G\left(x\right)}{\int_{x\in X} q\left(x\right) \mathrm{d}G\left(x\right)}. \tag{16}$$

Now, fix any $Y \in X$ and consider the class of responsive experiments $\mathcal{E}\left(Y\right)$ that induce Y as their trade expectation: $S = \{s^{+}, s^{-}\}$ and $\mathbb{E}\left[x \mid E, s^{+}\right] = Y$. This class is non-empty. For example, it contains an experiment that recommends the trade only when the attribute vector x is equal to a given expectation Y. However, under that experiment, the trade recommendation is sent with a zero probability if attributes are continuously distributed. It is possible to increase the likelihood of the trade recommendation by sending s^{+} from the progressively larger neighborhoods of Y. It turns out that the limit of this expansion is a linear disclosure.

Proposition 4. (Maximal Probability)
Consider any $Y \in X$ and the class of responsive experiments $\mathcal{E}\left(Y\right)$ that induce Y as its trade expectation. A linear disclosure maximizes the probability of a trade recommendation among all $E \in \mathcal{E}\left(Y\right)$.

Proposition 4 is straightforward in the case of a single attribute. The multidimensional extension immediately follows from Proposition 3. It suggests that a linear disclosure can be viewed as a natural multidimensional extension of a binary monotone partition disclosure.

In general multidimensional screening problems, one cannot be sure that an optimal bundle belongs to a boundary of a feasibility set. Indeed, optimal attribute surpluses $\mathcal{X}\left(\theta\right)$ might belong to an interior of their feasibility set. However, I show that the trade probability can always be minimized to bring the bundle $\left(\mathcal{X}\left(\theta\right), \mathcal{Q}\left(\theta\right)\right)$ to the boundary of $\mathcal{F}$.

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.

Proposition 5. (Implementability)
For any implementable allocation $(\mathcal{X}(\theta), \mathcal{Q}(\theta))_{\theta \in \Theta}$ *there exists an implementable allocation* $(\mathcal{X}(\theta), \mathcal{Q}'(\theta))_{\theta \in \Theta}$ *such that: (1) it delivers the same revenue and the same payoffs for all types and (2) for all* $\theta \in \Theta$, $(\mathcal{X}(\theta), \mathcal{Q}'(\theta)) \in \partial(\mathcal{F})$ *and* $\mathcal{Q}'(\theta) \leq \mathcal{Q}(\theta)$.

Intuitively, if $(\mathcal{X}(\theta), \mathcal{Q}(\theta))$ lies in the interior of $\mathcal{F}$, the seller can always reduce the total trade probability while keeping the attribute surplus the same. 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, the higher object price makes deviations to this type's item less appealing and, in fact, strictly so whenever the deviating types intend to always buy the object or swap their decisions with recommendations.

As an immediate corollary of Propositions 3 and 5, I obtain a tractable characterization of a class of optimal disclosures.

Theorem 1. (Linear Disclosure)
There exists an optimal responsive menu with every experiment in it being a linear disclosure.

Despite the complexity of the seller's problem, and in the absence of any assumptions on attribute and types distributions, all optimal experiments belong to a tractable class of linear disclosures. Note that Theorem 1 does not say which linear disclosures should be employed in an optimal menu; the optimal choice clearly depends on the problem at hand. However, the theorem identifies linear disclosures as an optimal way to screen buyer's types.

At this point, it is instructive to compare allocation distortions driven by the monopoly power in the cases of complete and incomplete information about the object. First, consider the situation with complete information so that the object's attributes are commonly known to be x_0. In this case, there is no scope for information control. The buyer's type θ matters only insofar as it affects the valuation $v(\theta) = \theta \cdot x_0$. If the seller could observe the type, she would engage in perfect price discrimination. She would allocate the object efficiently, selling it if and only if $v(\theta) \geq 0$, and extract full surplus. If the seller could not observe the type, she could try to screen different types by designing a menu of items varying in sale probabilities and prices. This screening problem was famously resolved by Myerson (1981). An optimal mechanism does not feature price discrimination. Each type θ is assigned a virtual valuation $\hat{v}(\theta) \leq v(\theta)$ that accounts for his private information. Under standard regularity conditions, an object is sold if and only if the virtual valuation is positive:

$$\hat{v}(\theta) \geq 0. \tag{17}$$

This allocation is typically inefficient. The seller does not extract all surplus generated by the transaction—the buyer obtains information rents.

Compare it to the current situation in which the object's attributes are uncertain and the seller can provide information about them. If the seller could observe the type θ, then she would profitably engage in information and price discrimination. By the argument of Bergemann and Pesendorfer (2007), the seller would inform type θ whether his valuation $v(\theta) = \theta \cdot x$ is positive and charge him a price $\mathbb{E}[v(\theta) \mid v(\theta) \geq 0]$. As in the case of complete information, the object would be allocated efficiently and the seller would extract 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 distortions would be remarkably similar to the case of complete information about the object. Each type θ is assigned a virtual type $\hat{\theta}(\theta)$ with the corresponding virtual valuation $\hat{v}(\theta) = \hat{\theta}(\theta) \cdot x$. An object is sold whenever the buyer virtual valuation is above some threshold, possibly with randomization on the boundary:

$$\hat{v}(\theta) \geq \alpha_0(\theta). \tag{18}$$

This allocation is also typically inefficient with two sources of inefficiencies. First, the virtual type $\hat{\theta}$ may differ from the true type θ. Second, the threshold α_0 may differ from 0. Again, the seller does not extract full generated surplus and the buyer obtains information rents.

3.5 General Payoffs

Importantly, Theorem 1 places no structural assumptions on the attribute distribution. This allows careful definition of attributes and extension of the optimal disclosure characterization beyond the linear ex post valuation formulation (1). In particular, for a general valuation function $v(\theta, x)$, one can define auxiliary attributes to coincide with valuations of different buyer's types. In this auxiliary formulation, Theorem 1 can be applied to obtain the characterization of optimal disclosures as linear forms of type valuations.

Proposition 6. (General Payoffs)
Let the number of types be finite, $|\Theta| < \infty$, and the valuation function take a general form $v(\theta, x)$ for an arbitrary attribute set X. Then there exists an optimal menu with every experiment in it being a linear form, i.e., for every experiment in the menu, there exist $\alpha : \Theta \to \mathbb{R}$ and $\alpha_0 \in \mathbb{R}$, not all zeros, such that:

$$q(x) = \begin{cases} 1, & \text{if } \sum_{\theta \in \Theta} \alpha(\theta) v(\theta, x) > \alpha_0, \\ 0, & \text{if } \sum_{\theta \in \Theta} \alpha(\theta) v(\theta, x) < \alpha_0. \end{cases} \tag{19}$$

Proposition 6 allows characterization of the classes of optimal disclosures in alternative environments in which the buyer's type captures general preferences such as bliss points or degrees of risk aversion. To illustrate it, consider the following example.

Example 1. (Location Payoffs)

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

$$v(\theta, x) = v_0 - (x - \theta)^2 . \tag{20}$$

Let there be two types $\theta_1, \theta_2 \in \mathbb{R}^J$. Assume that X is bounded and v_0 is sufficiently high so that for all $x \in X$, the types' valuations are positive. In this case, optimal disclosures in the menu 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 Proposition 6, the other type is offered a linear form (19) that informs whether a linear combination of valuations $v(\theta_1, x)$ and $v(\theta_2, x)$ is above or below some threshold:

$$q(x) = \begin{cases} 1, & \text{if } \alpha_1 v(\theta_1, x) + \alpha_2 v(\theta_2, x) > \alpha_0, \\ 0, & \text{if } \alpha_1 v(\theta_1, x) + \alpha_2 v(\theta_2, x) < \alpha_0. \end{cases} \tag{21}$$

Plugging the location valuation function (20) into (21) provides a tractable characterization of optimal experiments. Generically, these experiments are *neighborhood disclosures* that inform whether the attribute vector lies in a neighborhood of a virtual type $\hat{\theta}$:

$$q(x) = \begin{cases} 1, & \text{if } -\left(x - \hat{\theta}\right)^2 \gtreqless \alpha_0', \\ 0, & \text{if } -\left(x - \hat{\theta}\right)^2 \lesseqgtr \alpha_0'. \end{cases} \tag{22}$$

where $\alpha_0' \in \mathbb{R}$ and the uncertain inequality sign allows the trade to happen in any of the disclosed regions. Moreover, the virtual type lies on the line connecting the types: $\hat{\theta} = \gamma\theta_1 + (1 - \gamma)\theta_2$ for some $\gamma \in \mathbb{R}$. That is, the class of optimal disclosures is rich but tractable. Coincidentally, as in the case of a linear ex post valuation (1), the disclosure informs about the valuation of a virtual type. ■

4 Design of Pricing Mechanisms

I proceed with studying revenue-maximizing mechanisms. In the previous section, I identified a class of optimal disclosures without placing any assumptions on the distributions of types or attributes. In this section, I similarly am able to identify a general class of optimal pricing mechanisms in the case of a single attribute. Obtaining the same level of generality with many attributes is problematic—the associated problem involves multidimensional screening, which is notoriously intractable. Nevertheless, I am able to identify key tradeoffs and characterize optimal mechanisms for specific classes of buyer types. First, I study the setting in which different types value different and independently distributed attributes. Second, I enrich the setting with vertical heterogeneity by allowing several types to value the same attribute with different intensities. Optimal mechanisms turn out to be remarkably simple and do not involve price discrimination.

From now on, I assume that all types and attributes are positive, $X \subseteq \mathbb{R}_+^J$, $\Theta \subseteq \mathbb{R}_+^J$. In this case, the buyer and the seller commonly know that there are positive gains from trade. It allows ignoring the efficiency role of disclosure and focusing solely on its screening effects.

4.1 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

$$v(\theta, x) = \theta x. \tag{23}$$

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. In this case, the revelation principle applies and implies that I can focus on direct mechanisms in which all payments are front loaded: the buyer reports his type θ, pays the upfront payment $r(\theta)$, and the trade happens with probability $q(x, \theta)$. The terms of trade determine the single-dimensional attribute surplus:

$$\mathcal{X}(\theta) = \int_{x \in X} x q(x, \theta) \, \mathrm{d}G(x) \tag{24}$$

that can be anywhere between 0 and $\mathbb{E}[x]$. I can then rewrite the seller's problem as maximizing the revenue directly over the payments $r(\theta)$ and the surpluses $\mathcal{X}(\theta)$ as:

$$\max_{r(\theta), 0 \leq \mathcal{X}(\theta) \leq \mathbb{E}[x]} \int_{\theta \in \Theta} r(\theta) \, \mathrm{d}F(\theta), \tag{25}$$

subject to incentive-compatibility and individual-rationality constraints:

$$\theta \mathcal{X}(\theta) - r(\theta) \geq \theta \mathcal{X}(\theta') - r(\theta'), \quad \forall \theta, \theta' \in \Theta, \tag{26}$$

$$\theta \mathcal{X}(\theta) - r(\theta) \geq 0, \quad \forall \theta \in \Theta. \tag{27}$$

This problem is analogous to a canonical mechanism design problem of Myerson (1981), with the attribute surplus taking the place of 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 corresponding optimal upfront payment $r(\theta)$ is equal to 0 for $\theta < \theta^*$ and equal to $r^* = \theta^* \mathbb{E}[x]$ for $\theta \geq \theta^*$.

The argument concludes by noting the optimal mechanism can be implemented by providing no disclosure and charging a strike price r^* for the object. This posted price mechanism can be implemented in the original, more restricted, problem and hence is optimal there as well.

Theorem 2. (Single Attribute)
If there is only one attribute, $J = 1$, and the buyer's valuation is always positive, $X \subseteq \mathbb{R}_+$, $\Theta \subseteq \mathbb{R}_+$, then an optimal menu is a posted price mechanism with no disclosure, i.e., it contains a single item with zero upfront payment, $r = 0$, and uninformative experiment, $E = \underline{E}$.

There is a simple intuition behind the optimality of no disclosure if the seller can only use a nondiscriminatory mechanism consisting of a single disclosure followed by a posted price. Consider an arbitrary disclosure. Any signal realization s scales the demand with the proportionality coefficient equal to the expected attribute $\mathbb{E}[x \mid s]$. If the seller could observe this realization, she would optimally charge a scaled price and obtained a scaled revenue. Importantly, the seller would sell the object to the same types irrespectively of the realization. Since any expectation under an information disclosure is a martingale, the seller would serve the same population and charge, on average, the same price. The seller can do just as well by using a posted price with no disclosure.

Figure 5 illustrates this argument. Consider the attribute and type distributions F, G such that under no additional information, the expected attribute value is $\mathbb{E}[x] = x_0$, the demand curve is $\mathcal{Q}_0(p)$, the optimal price is p_0, and the optimal trade probability is $\mathcal{Q}^*$. Consider an experiment that sends two signals s_1, s_2, inducing the posterior expectations $\mathbb{E}[x \mid s_1] = 1/2x_0$ and $\mathbb{E}[x \mid s_2] = 2x_0$. After signal s_1, expected valuations of all types are cut in half. As a result, the induced demand curve $\mathcal{Q}_1(p)$ is a scaled-down version of the original curve, $\mathcal{Q}_1(p) = \mathcal{Q}_0(2p)$. Hence, the new optimal price is twice as small as the original price, $p_1 = p_0/2$, inducing the same trade probability $\mathcal{Q}^*$. Similarly, after signal s_2, expected valuations of all types double. As a result, the demand curve scales up, $\mathcal{Q}_2(p) = \mathcal{Q}_0(p/2)$, and the induced optimal price is twice as large as the original price, $p_2 = 2p_0$, inducing, again, the same trade probability $\mathcal{Q}^*$. By the martingale property, an average posterior expectation is equal to the prior expectation, so $\Pr(s_1) = 6/10$ and $\Pr(s_2) = 4/10$. Thus, the average price is equal to the original price, $\Pr(s_1)\, p_1 + \Pr(s_2)\, p_2 = p_0$. Because the trade

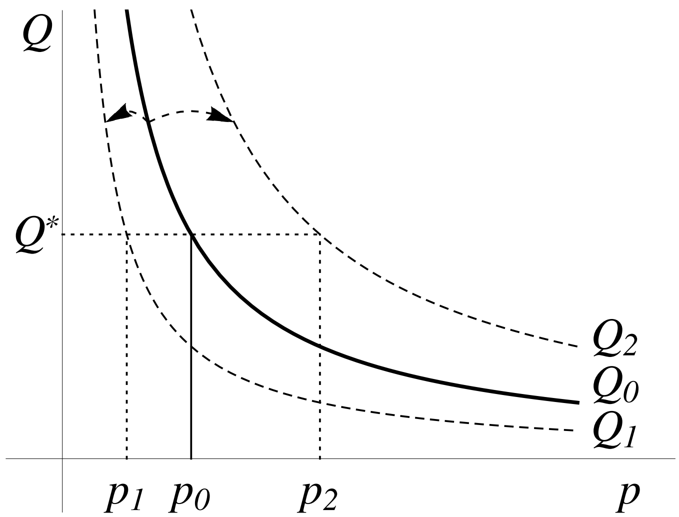

Figure 5: A stochastic split of a demand curve following an attribute disclosure. Optimal prices scale proportionally. Optimal trade probability remains the same.

probability remains constant, the expected revenue equals the no-disclosure revenue, $\mathcal{Q}^* p_0$. Disclosure, even if observed by the seller, does not benefit her.

Although intuitive, this argument does not account for discriminatory schemes with upfront payments. Theorem 2 confirms that no disclosure is optimal even if the principal can do that. I highlight that this result requires no assumptions on type distribution F or attribute distribution G beyond common knowledge of positive trade gains.

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

$$v(\theta, x) = \psi(\theta)\,\phi(x)\,, \tag{28}$$

for $\psi, \phi : \mathbb{R}^J \to \mathbb{R}_+$. For example, it applies, if all types belong to a ray, $\Theta = \{\beta\theta_0\}_{\beta\in\mathbb{R}_+}$ for some direction vector $\theta_0 \in \mathbb{R}^J_+$ (c.f., Armstrong (1996)). In this case, the indices can be defined as $\psi(\theta) = \beta(\theta)$, and $\phi(x) = \theta_0 \cdot x$.

Remark 2. The no-disclosure mechanism may not be uniquely optimal. In fact, with a single attribute, informing the type-θ buyer about an attribute value x is equivalent to informing him about an attribute percentile $G(x)$, which is also equal to the valuation percentile for that type. It follows that the analysis of Eső and Szentes (2007) can be applied and, under some distributional assumptions, full disclosure is also optimal but must be accompanied by a complex structure of upfront payments and strike prices.

4.2 General Problem

The case of several attributes is qualitatively different because an optimal allocation may depend on attribute realization. Intuitively, the seller should tailor the allocation to the buyer types who value the realized attribute the most. This allocation adjustment requires attribute information and, hence, disclosure. Indeed, as I show in the following subsections, optimal mechanisms with multiple attributes generally provide some attribute disclosure.

In the previous section, I identified a class of optimal disclosures. The connection to linear disclosures simplifies the problem and allows optimizing the menu with respect to the defining parameters $\alpha(\theta)$, $\alpha_0(\theta)$ of (15). However, writing the seller's problem in terms of these parameters is cumbersome and not transparent. Instead, as I discussed in the previous section, the seller can optimize directly in terms of attribute surpluses $\mathcal{X}$. The optimal trade probability is then the minimal $\mathcal{Q}$ such that $(\mathcal{X}, \mathcal{Q})$ belong to the feasibility set $\mathcal{F}$. Because $\mathcal{F}$ is convex, this induced trade probability $\mathcal{Q}(\mathcal{X})$ is a convex function of attribute surpluses. The exact shape of $\mathcal{Q}(\mathcal{X})$ depends on the attribute distribution G and is generally nonlinear.

This observation reduces the search for an optimal menu to a concrete multidimensional screening problem, presented in full in the Appendix. Even though the seller sells a single object, information disclosure allows him to control the attribute surpluses $\mathcal{X}(\theta)$ at the time of a purchase. However, unlike in the bundling problem, the seller cannot choose the surpluses at will, they must belong to a convex feasibility set. Moreover, the surpluses directly affect the trade probability in a nonlinear fashion.

Even the most basic multidimensional screening problems are known to be prohibitively difficult.[10] The seller's problem is further complicated by the presence of multiple double-deviation constraints and the non-linearity of the trade probability. I make progress by focusing on a specific class of buyer's types.

4.3 Single-Minded Buyer

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

$$\theta = (0, \ldots, 0, \theta_j, 0, \ldots, 0). \tag{29}$$

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

[10] Bergemann et al. (2012) and Daskalakis et al. (2017) highlight the difficulties associated with the multiproduct monopolist problem.

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

$$v_j(\theta_j, x) = \theta_j x_j \quad \forall\, j, \theta_j \in \Theta_j. \tag{30}$$

I denote the frequency of a cohort Θ_j by $f(\Theta_j)$ and the marginal 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* attribute values are independently distributed, so that $x_j \sim G_j$ and $G(x) = \times_j G_j(x_j)$.[11] The independence requirement is substantive. Starting with the general case, one can always redefine attributes as valuations of the corresponding types as done in the proof of Proposition 6. In this formulation, each type naturally values only the attribute that arose from his original valuation. However, the so-defined attributes can be correlated with the correlation structure determined by the original attribute and type distributions.

If the buyer is single-minded, then the seller knows that the buyer values only one attribute but does not know which one or how much. This type structure allows further narrowing of the class of optimal experiments. Because attributes are independently distributed, a type $\theta_j \in \Theta_j$ values only information about attribute j. Information about other attributes does not change his ex-ante valuation and has no value for him. This observation suggests an optimal way to screen single-minded types in a direct menu: if the buyer reports type $\theta_j \in \Theta_j$, then the seller should provide information only about attribute j. Providing any other information would make misreporting more appealing without adding value for truth telling. The following proposition confirms this intuition.

Proposition 7. (Directional Disclosure)
If the buyer is single-minded, then there exists an optimal menu such that an experiment $E_j(\theta_j)$ is informative only about attribute j. That is, for all j, $\theta_j \in \Theta_j$, $q_j(\theta_j, x) = q_j(\theta_j, x')$ whenever $x_j = x'_j$.

Proposition 7 establishes that every experiment $E_j(\theta_j)$ provides information about a single attribute j. At the same time, by Theorem 1, $E_j(\theta_j)$ is a linear disclosure. However, any linear disclosure that is informative only about attribute j is effectively a binary monotone partition defined on this attribute.

Corollary 1. *If the buyer is single-minded, then an optimal experiment $E_j(\theta_j)$ is a binary monotone partition of attribute j.*

[11] The name is inspired by "single-minded" bidders studied in the literature on combinatorial auction design. A "single-minded" type values a specific attribute, whereas a "single-minded" bidder values a specific bundle. See, for example, Lehmann et al. (2002). I thank Laura Doval for drawing my attention to this helpful connection.

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. By construction, the buyer should purchase the object in only one element of the partition. Incentive compatibility requires this element correspond to attributes above the threshold, because higher attribute values are more attractive. The resulting trade function is:

$$q_j(\theta_j, x) = \begin{cases} 1, & \text{if } x_j > \alpha_{0j}(\theta_j), \\ 0, & \text{if } x_j < \alpha_{0j}(\theta_j). \end{cases} \tag{31}$$

As discussed in Section 4.2, posing the seller's problem in terms of $\alpha_{0j}(\theta_j)$ is not very convenient. Instead, I characterize the experiment by the attribute surplus:

$$\mathcal{X}_j(\theta_j) = \int_{\alpha_{0j}(\theta_j)}^{\infty} x_j \mathrm{d}G_j(x_j). \tag{32}$$

The total trade probability can be written as the function of the surplus $\mathcal{Q}_j(\mathcal{X}_j)$. The attribute surplus can take any value between 0 and $\mathbb{E}[x_j]$. As it increases, the corresponding disclosure threshold α_{0j} decreases and the total trade probability $\mathcal{Q}_j$ increases.

It follows that to characterize an optimal menu, I need to find the optimal attribute surplus functions $\mathcal{X}_j(\theta_j)$ and tariff functions $r_j(\theta_j)$, $p_j(\theta_j)$. I start with a simpler case that does not incorporate the valuation heterogeneity within each attribute. I then study a general case and show that an optimal mechanism remains qualitatively the same.

4.3.1 Orthogonal Types

I begin with the case in which each attribute cohort is a singleton, $\Theta_j = \{\theta_j\}$, so there is no vertical within-attribute heterogeneity and the number of types equals the number of attributes $|\Theta| = J$. Note that in this case, any two different types $\theta, \theta' \in \Theta$ are orthogonal to each other as vectors in $\mathbb{R}^J$, so I refer to this case as the setting of *orthogonal types*.

Without loss of generality, I can set all valuation intensities equal to one:

$$\theta_j \equiv 1 \quad \forall j = 1, \ldots, J. \tag{33}$$

Hence, I will omit the dependence on the intensity within each attribute and differentiate types by subscripts.

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 and hence not informative about attribute j. Thus, the type has no reason to act upon 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

$$\max_{\{r_j,\mathcal{X}_j,p_j\}_{j=1}^{J}} \sum_{j=1}^{J} f(\theta_j)(r_j + \mathcal{Q}_j p_j) \tag{34}$$

$$\text{s.t.} \quad \mathcal{X}_j - p_j \mathcal{Q}_j - r_j \geq \mathbb{E}[x_j] - p_k - r_k, \quad \forall\, j, k = 1, \ldots, J, \tag{35}$$

$$\mathcal{X}_j - p_j \mathcal{Q}_j - r_j \geq 0, \tag{36}$$

$$\mathcal{X}_j \in [0, \mathbb{E}[x_j]], \ \mathcal{Q}_j = \mathcal{Q}_j(\mathcal{X}_j). \tag{37}$$

This problem resembles a one-dimensional mechanism design problem with the following important differences. First, each item in this problem features both horizontal and vertical components. The upfront payments r_j are purely vertical—all types value them the same. The experiments E_j and the associated attribute surpluses $\mathcal{X}_j$ are purely horizontal—they are valuable only to the type θ_j. The object prices p_j are mixed as they are paid only if the trade occurs, probability of which depends on a type. Second, the problem is non-linear. Not only are there products between the object prices p_j and the trade probabilities $\mathcal{Q}_j$, but also the trade probability functions $\mathcal{Q}_j(\mathcal{X}_j)$ are generically nonlinear as well.

Because of these differences, I cannot apply standard mechanism design techniques. Instead, I solve the problem in a sequence of simplifications. In the first step, I observe that using upfront payments r_j is detrimental to the seller. For any strictly positive r_j the seller can reduce the transfer and increase the object price p_j to keep the total expected transfer $r_j + \mathcal{Q}_j p_j$ the same. This change does not affect utilities of truth-telling types or the seller's revenue. However, it makes misreporting less appealing. Intuitively, by shifting the expected transfer towards the object price, the seller better discriminates against the types who always purchase the object.

In the second step, I use the special structure of incentive-compatibility constraints to show object price discrimination is not profitable as well. Indeed, because all experiments off the truth-telling path bring no information value, an optimal deviation is to the types associated with the lowest price $\underline{p}$. If there is any object price variation, then there is a type θ_j with a price $p_j > \underline{p}$. This type's item is not attractive to any other type. Moreover, for the type to be willing to pay a higher price, the item must contain partial disclosure. This leads to a contradiction: the seller can simultaneously lower the price p_j and increase the attribute surplus $\mathcal{X}_j$ in such a way that the type's rents $\mathcal{X}_j - p_j \mathcal{Q}_j$ remain the same but the expected payment $\mathcal{Q}_j p_j$ increases. Intuitively, the seller should not lose surplus on types irrelevant for incentives of the others.

Once I establish that optimal mechanism is nondiscriminatory, finding optimal experiments is straightforward. The seller should maximize trade probability by providing minimal information sufficient to convince the buyer to make a purchase, attribute by attribute. 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 trade expectation up to the object price, $\mathcal{X}_j / \mathcal{Q}(\mathcal{X}_j) = p$. The following theorem summarizes the findings.

Theorem 3. (Optimal Menu, Orthogonal Types)
If the buyer is single-minded and the types are orthogonal, then an optimal responsive menu is characterized by the following properties:

1. *For any attribute j, $r_j = 0$,*
2. *For any attribute j, $p_j = p$, and*
3. *For any attribute j, E_j is a binary monotone partition of x_j such that $\mathbb{E}[x_j \mid E_j, s^+] = \max\{p, \mathbb{E}[x_j]\}$.*

The optimal menu is illustrated in Figure 6. I highlight its notable features. First, the pricing strategy is simple. The menu does not feature price discrimination and the disclosure is provided free of charge. Second, the menu has the standard "no distortions at the top, no rents at the bottom" property. Namely, 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, but two type classes that partition the type space. Third, the menu admits a nondiscriminatory indirect implementation. The seller can simply provide a single combined disclosure followed by the optimal posted price. Because each type values only one attribute, he will focus on the relevant attribute information. Third, for given attribute distributions $\{G_j\}$, the optimal mechanism can be found easily by a two-stage algorithm. In the first stage, for any fixed price p, the algorithm uses the third property of Theorem 3 to find optimal thresholds α_{0j} and define the corresponding trade probability $\mathcal{Q}$. The so-defined function $\mathcal{Q}(p)$ is effectively a "modified" demand curve that accounts for optimal disclosure. In the second stage, the algorithm uses the demand curve to find an optimal price.

4.3.2 Continuum of Types

I now extend the analysis to the general case of single-minded types and allow for the vertical heterogeneity within attribute cohorts. In particular, I assume that for all j, the attribute cohorts admit an upper bound, $\Theta_j = \left[0, \overline{\theta}_j\right]$ and θ_j are continuously distributed over Θ_j

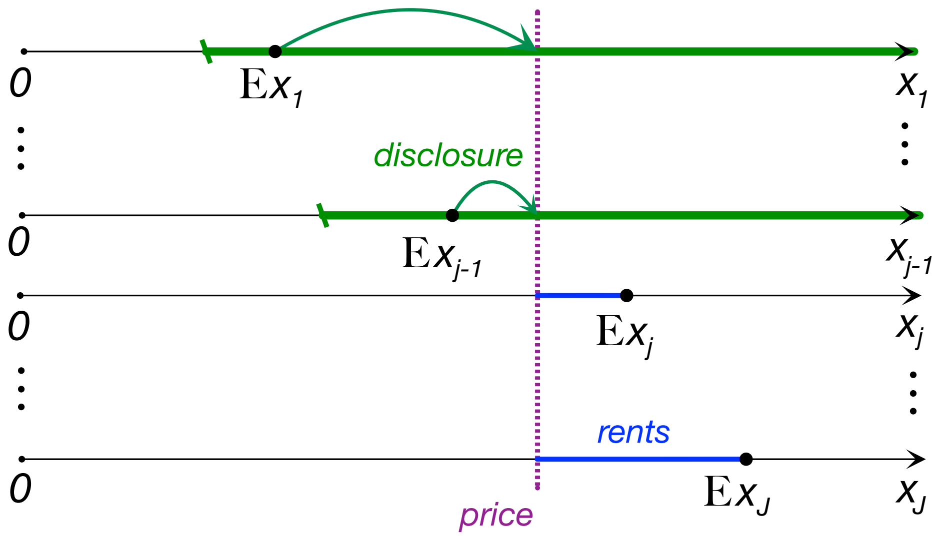

Figure 6: Optimal mechanism in the case of orthogonal types. Green color indicates attribute regions in which a purchase recommendation is sent, for types with partial disclosure. Blue color indicates the types' rent conditional on a trade, for types with no disclosure. Attributes are ordered by increasing ex-ante expectations.

according to a distribution function $F_j(\theta)$. For the main result of this section, I assume that each of the distributions $F_j(\theta)$ is log-concave.[12]

In this general case, for each attribute j, the seller needs to design the tariff and the attribute surplus functions, $r_j(\theta_j)$, $p_j(\theta_j)$, $\mathcal{X}_j(\theta_j)$. This requires incorporating more incentive constraints than in the case of orthogonal types, particularly, the constraints in which the type pretends to be another type within the same cohort and follows the experiment recommendation. Swapping the decision or never purchasing the object remains suboptimal, so those constraints may again be omitted.

As before, I approach this problem in a sequence of simplifying steps. First, I invoke the same argument as in the case of orthogonal types to establish that upfront payments are not used in an optimal menu, $r_j(\theta_j) \equiv 0$. Indeed, if some $r_j(\theta_j) > 0$, then the seller can simultaneously decrease $r_j(\theta_j)$ and increase $p_j(\theta_j)$ to keep the expected payment $r_j(\theta_j) + \mathcal{Q}_j(\theta_j)\, p_j(\theta_j)$ the same. This does not affect the revenue or incentive compatibility within each attribute cohort. However, it relaxes incentive constraints between different cohorts.

Second, by standard arguments, incentive compatibility within the same cohort implies that $\mathcal{X}_j(\theta_j)$ is nondecreasing in θ_j. It in turn implies that $\alpha_{0j}(\theta_j)$ is nonincreasing; hence, $\mathcal{Q}_j(\theta_j)$ is nondecreasing and $\mathcal{X}_j(\theta_j)/\mathcal{Q}_j(\theta_j)$ is nonincreasing in θ_j. That is, higher types

[12]The class of log-concave distributions includes normal, logistic, exponential, and uniform distributions, as well as their truncations.

must trade with higher probability but lower expectations.

The seller's problem can then be written solely in terms of the attribute surpluses.

Lemma 1. *The seller's problem can be written as:*

$$\max_{\{\mathcal{X}_j(\theta_j)\}_{j=1}^{J}} \sum_{j=1}^{J} f(\Theta_j) \int_0^{\overline{\theta}_j} \left(\theta_j - \frac{1-F_j(\theta_j)}{f_j(\theta_j)}\right) \mathcal{X}_j(\theta_j)\, \mathrm{d}F_j(\theta_j) \tag{38}$$

$$\text{s.t.} \quad \mathcal{X}_j(\theta_j) \text{ is non} - \text{decreasing}, \; \mathcal{X}_j(\theta_j) \in [0, \mathbb{E}[x_j]], \tag{39}$$

$$\int_0^{\overline{\theta}_j} \mathcal{X}_j(\theta_j)\, d\theta_j \geq \overline{\theta}_j \mathbb{E}[x_j] - \underline{p}(\mathcal{X}_1(\cdot), \ldots, \mathcal{X}_J(\cdot)) \;\; \forall j. \tag{40}$$

The objective function and the monotonicity constraints capture the incentive-compatibility constraints within each attribute cohort. They are derived by standard one-dimensional arguments. The integral constraints are novel and capture the incentive-compatibility constraints between different cohorts. In particular, they require that the highest type within each cohort does not want to purchase the object at the minimal price.

I now argue that the lowest price $\underline{p}$ is offered to the highest types $\overline{\theta}_j$. The argument and the result are analogous to that in the case of orthogonal types. Assume in an optimal mechanism some 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 but not affecting any other constraint.

This observation allows stating a relaxed problem in which the monotonicity and integral constraints are dropped but all high types are required to be offered the same minimal price. If type distributions are log-concave, then the solution to the relaxed problem is a single-step function. It satisfies the original constraints and hence solves the original problem. The solution corresponds to only one item per attribute cohort associated with the same object price. The following theorem summarizes the findings.

Theorem 4. (Optimal Menu, Single-Minded Buyer)
If the buyer is single minded and type distributions are log-concave, then an optimal menu is characterized by the following properties:

1. *For all j, $\theta_j \in \Theta_j$, $r(\theta_j) = 0$,*

2. *For all j, $\theta_j \in \Theta_j$, $p(\theta_j) = p$, and*

3. *For all j, $\theta_j \in \Theta_j$, $E_j(\theta_j) = E_j$ where E_j is a binary monotone partition of x_j.*

It is worth pointing out that my analysis provides a partial characterization in the case of general distributions $F_j(\theta_j)$ as well. The first statement remains the sam—upfront payments are not used with a single-minded buyer. However, the second and the third statements need to be modified to allow for limited price discrimination. In particular, the arguments of Samuelson (1984) and Bergemann, Bonatti, and Smolin (2018) can be applied to limit the number of optimal items to two per cohort. That is, the highest types are still offered the lowest price $\forall\, j,k,\ p\left(\overline{\theta}_j\right) = p\left(\overline{\theta}_k\right) = \underline{p}$, but per each attribute cohort there could be one more item that targets lower types.

5 Discussion

5.1 Demand Transformation

My analysis highlights that attribute disclosure can be profitably used to modify a demand curve. Consider the example in which there are two attributes $J = 2$, $x_1 \sim U[0,1]$, $x_2 \sim U[0,2]$ independently distributed and a continuum of single-minded types. Types $\theta_1 \in \Theta_1$ value only the first attribute and types $\theta_2 \in \Theta_2$ value only the second attribute. Each cohort is equally likely and within each cohort the types are uniformly distributed over $[0,2]$ so the average type is equal to 1.

If the seller provides no disclosure, then she faces a piecewise-linear demand curve. The optimal no-disclosure price is $p_{no} = 2/3$ with the corresponding revenue $1/3$. If the seller provides full disclosure, then the type valuations spread out. The demand decreases for lower prices and increases for higher prices. The overall effect is negative. The full-disclosure optimal price is $p_{full} \simeq 0.82$ with the corresponding revenue $0.28 < 1/3$. The left side of Figure 7 illustrates this case.

However, the seller can increase revenue by providing partial disclosure. The type distributions are log-concave so, by Theorem 4, the optimal mechanism can be implemented by a single multipartition disclosure followed by a posted object price. The optimal disclosure thresholds can be calculated numerically to be $\alpha_{01} \simeq 0.27$, $\alpha_{02} = 0$. The seller optimally reveals whether the first attribute is above 0.27, yet provides no information about the second attribute. This disclosure targets the types that ex-ante value the object less. As a result, the demand decreases for low prices, increases for medium prices, and remains the same for high prices. The overall effect is positive. The disclosure increases demand even at the optimal no-disclosure price p_{no}. The optimal-disclosure optimal price is $p_{opp} \simeq 0.80$ with the corresponding revenue $0.35 > 1/3$. The right side of Figure 7 illustrates.

These findings resonate with the analysis of Johnson and Myatt (2006), who also study

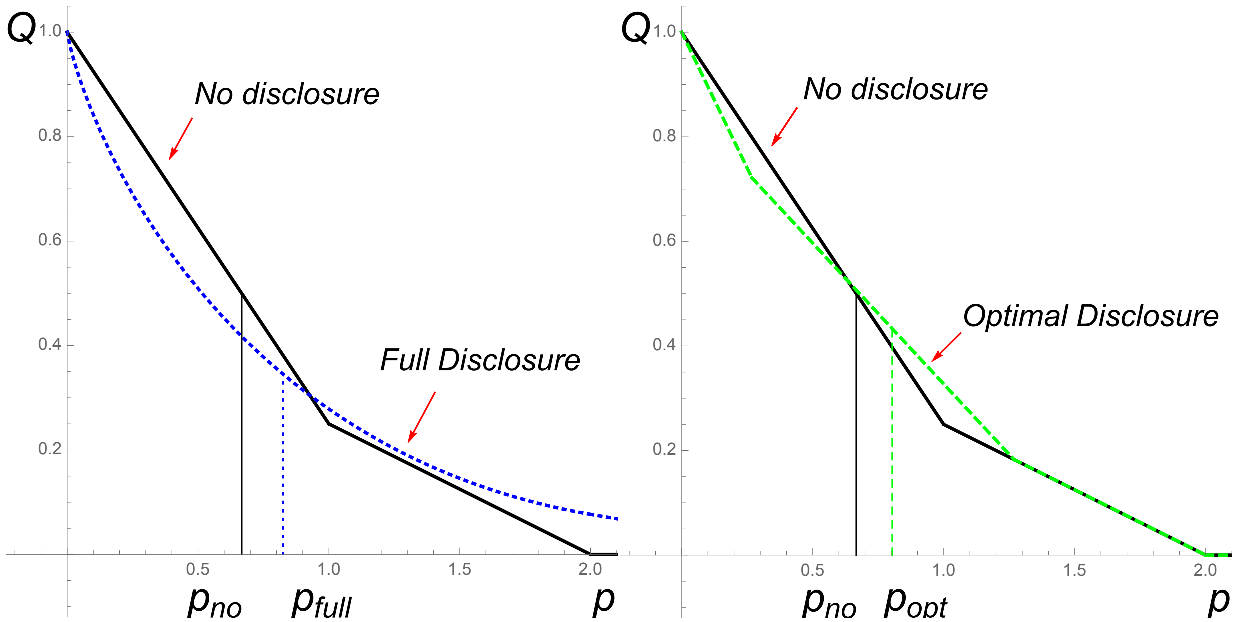

Figure 7: Attribute disclosure and demand transformation. Left: demand curves under no disclosure and full disclosure. Right: demand curves under no disclosure and optimal disclosure. Vertical lines indicate revenue-maximizing prices.

the impact of information disclosure on demand curves. They restrict attention to disclosures that spread type valuations uniformly. Such disclosures translate into *global* rotations of the demand curve. Johnson and Myatt (2006) 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 low price and high demand. Full disclosure is associated with a niche market characterized by high price and low demand.

In contrast, I show that attribute disclosure can rotate the demand curve *locally*. The local rotations correspond to partial disclosures that target specific types. These disclosures can outperform full and no disclosure in both mass and niche markets. I highlight that multiple attributes are required for this result. As shown in Section 4.1, no disclosure remains optimal in a one-dimensional framework.

5.2 Full and Public Disclosure

In a closely related paper, Eső and Szentes (2007) study discriminatory mechanisms in a "valuation-rank" framework and obtain two main qualitative results. First, they show that full information disclosure is generally optimal. Second, they show that the seller cannot benefit from conditioning the price on the disclosure realization.

In this section, I illustrate that both of these results do not hold in the attribute setting.

Consider the example in which there are two attributes $J = 2$, $x_1 \sim U[0,1]$, $x_2 \sim U[0,2]$, independently distributed. Let there be two equally likely orthogonal types: θ_1 values only the first attribute and θ_2 values only the second attribute.

An optimal menu can be calculated by Theorem 3. The menu can be implemented by a single disclosure that reveals whether the first attribute is above or below 1/2 followed by a posted price of 3/4. This optimal mechanism obtains revenue 9/16.

The revenue is strictly higher compared to that under full disclosure. In that case, if the seller does not use upfront payments, then the optimal object price can be calculated to be 2/3, translating into revenue of 1/3. The seller can improve it by using upfront payments. The optimal mechanism can be easily calculated to be $r_1 = r_2 = 1/2$, $p_1 = p_2 = 0$. That is, anticipating full disclosure, the seller prefers to not discriminate between the types at all and sell the object in advance through an upfront fee. The corresponding maximal revenue under full disclosure is 1/2 falling short off optimal 9/16.

At the same time, the seller can strictly improve the revenue if she could condition the price on the disclosure realization. In fact, she can improve the revenue even under full disclosure. Consider the following mechanism. The seller provides full disclosure, observes the attributes, and chooses the price optimal for the realized valuation distribution. Under this scheme, the seller obtains the revenue:

$$\Pi = \int_0^1 \int_0^2 \frac{1}{2} \max\left\{\min\left\{x_1, x_2\right\}, \frac{\max\left\{x_1, x_2\right\}}{2}\right\} \mathrm{d}x_1 \mathrm{d}x_2 = \frac{29}{48} > \frac{9}{16}. \tag{41}$$

Naturally, observing the disclosure realization reduces the buyer informational advantage and helps the seller to better screen the types.

These observations highlight the distinction between the attribute setting and the "valuation-rank" setting of Eső and Szentes (2007). In their framework, the seller informs the buyer about an "orthogonal shock" $\xi(\theta)$, defined as the type's valuation percentile. By construction, these percentiles are uniformly distributed:

$$\xi(\theta) \sim U[0,1] \quad \forall\, \theta \in \Theta. \tag{42}$$

The implicit assumption of the valuation-rank framework is that these shocks in fact equal each other, $\xi(\theta) \equiv \xi$. This assumption is suitable for environments in which, conditional on the object's state, the buyer's type does not affect his valuation percentile. As Eső and Szentes (2007) discuss, these environments include the cases of additive valuations and Gaussian learning.

However, despite having the same distribution, the shocks $\xi(\theta)$ are generally different random variables. This distinction is particularly prominent in the case of a single-minded buyer with orthogonal types. Different shocks correspond to different attributes and thus are independently distributed. In the example above, these shocks are $\xi(\theta_1) = x_1$ and $\xi(\theta_2) = x_2/2$.

6 Conclusion

I studied a monopolist who sells a multi-attribute object to a privately informed buyer and showed that the seller can benefit from disclosure of attribute information. The benefit comes through two channels. First, disclosure can be used as a screening device, leveraging the fact that different buyer types prefer 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 if each type values a single attribute and attributes are independent, then screening is not beneficial and information should be disclosed partially and free of charge. That is, in such 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 may be important and should be accounted for. The seller may be restricted in what kinds of information she may provide. The buyer may feature heterogeneity in his ability to process data. The market may involve imperfect competition. Each of these extensions can be approached within the multi-attribute disclosure framework that I have outlined.

7 Appendix

Proof of Proposition 1. Consider an arbitrary menu $M = (r(i), E(i), p(i))_{i \in \mathcal{I}}$. For any type θ, the menu 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 report truthfully and follow the recommendations, then the menu M' results in the same allocation distribution and the same expected payments as the menu M. At the same time, any deviation under M' is available to the buyer under M. Hence, reporting truthfully and following the recommendations is incentive-compatible for all types under M'.

Proof of Proposition 2. Towards a contradiction, assume that a responsive menu $M = (r(\theta), \mathcal{X}(\theta), \mathcal{Q}(\theta), p(\theta))_{\theta \in \Theta}$ is optimal yet every type does not buy the object with some strictly positive probability. Construct a new menu M' as follows. Pick a type $\overline{\theta}$ with the highest expected payment $\overline{T} = r\left(\overline{\theta}\right) + p\left(\overline{\theta}\right) \mathcal{Q}\left(\overline{\theta}\right)$. As 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 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 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 deviation would only increase the seller's profit as $\overline{T}$ was chosen to be the highest expected payment. Hence, the menu M' brings strictly greater revenue than the menu M. Contradiction. The result follows.

Proof of Proposition 3. As the expected valuations of all types exist, the set $\mathcal{F}$ is a continuous image of a compact set. Hence, $\mathcal{F}$ is compact.

The set $\mathcal{F}$ is convex because for any trade functions q_1, q_2 their convex combination is a feasible trade function that achieves a convex combination of their attribute surpluses and trade probabilities. 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. Hence, $(\mathcal{X}_3, \mathcal{Q}_3)$ is a convex combination of $(\mathcal{X}_1, \mathcal{Q}_1)$ and $(\mathcal{X}_2, \mathcal{Q}_2)$ and belongs to the feasibility set $\mathcal{F}$.

The boundary of $\mathcal{F}$ is spanned by linear disclosures. Indeed, as $\mathcal{F}$ is a finite-dimensional closed set, the Supporting Hyperplane theorem (Rockafellar (1970), Theorem 11.6, Corollary 11.6.1, p. 100) can be applied.[13] In particular, a point $\left(\hat{\mathcal{X}}, \hat{\mathcal{Q}}\right)$ belongs to the boundary of $\mathcal{F}$ if only if there are coefficients (λ, λ_0) not all zero such that:

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

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

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

The integral is maximized pointwise. Its any maximizer is a linear disclosure (15) with coefficients $\alpha = \lambda$, $\alpha_0 = \lambda_0$.

[13]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. A sufficient condition for the existence of a supporting hyperplane is that $\mathcal{F}$ has a non-empty interior.

Proof of Proposition 4. For an arbitrary trade expectation $Y \in X$, the problem of finding the maximal probability experiment can be written as:

$$\max_{q:X\to[0,1]} \int q(x)\, dG(x)$$
$$\text{s.t. } \frac{\int xq(x)\, dG(x)}{\int q(x)\, dG(x)} = Y.$$

Equivalently, it can be written in terms of the attribute surpluses and trade probabilities as:

$$\max_{(\mathcal{X},\mathcal{Q})\in\mathcal{F}} \mathcal{Q}$$
$$\text{s.t. } \mathcal{X} - \mathcal{Q}Y = 0.$$

As $\mathcal{F}$ is compact, the solution to this problem exists and belongs to the boundary of $\mathcal{F}$. Thus, by Proposition 3, it can be achieved (only) by a linear disclosure.

Proof of Proposition 5. The seller's problem can be written in terms of attribute surpluses and trade probabilities as:

$$\max_{\{r(\theta),\mathcal{X}(\theta),\mathcal{Q}(\theta),p(\theta)\}} \int_{\theta\in\Theta} \left(r(\theta) + \mathcal{Q}(\theta)\, p(\theta)\right) \mathrm{d}F(\theta) \tag{43}$$

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

$$\theta \cdot \mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq \theta \cdot \mathcal{X}(\theta') - \mathcal{Q}(\theta')\, p(\theta') - r(\theta'), \tag{44}$$
$$\theta \cdot \mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq \theta \cdot (\mathbb{E}[x] - \mathcal{X}(\theta')) - (1 - \mathcal{Q}(\theta'))\, p(\theta') - r(\theta'), \tag{45}$$
$$\theta \cdot \mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq \theta \cdot \mathbb{E}[x] - p(\theta') - r(\theta'), \tag{46}$$
$$\theta \cdot \mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq -r(\theta'), \tag{47}$$

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

$$\theta \cdot \mathcal{X}(\theta) - \mathcal{Q}(\theta)\, p(\theta) - r(\theta) \geq 0, \tag{48}$$

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

$$(\mathcal{X}(\theta), \mathcal{Q}(\theta)) \in \mathcal{F}. \tag{49}$$

Consider any feasible profile of tariff functions, attribute surpluses, and trade probabilities $r(\theta)$, $\mathcal{X}(\theta)$, $\mathcal{Q}(\theta)$, $p(\theta)$ that satisfy constraints (44), (45), (46), (47), (48). Take any type

$\hat{\theta}$ such that $\left(\mathcal{X}\left(\hat{\theta}\right), \mathcal{Q}\left(\hat{\theta}\right)\right) \in \operatorname{int}(\mathcal{F})$. Consider the following perturbation: keeping $\mathcal{X}\left(\hat{\theta}\right)$ and $p\left(\hat{\theta}\right)\mathcal{Q}\left(\hat{\theta}\right)$ fixed, minimize $\mathcal{Q}\left(\hat{\theta}\right)$ within $\mathcal{F}$. By (3), $\mathcal{F}$ is compact so an optimum exists and belongs to the boundary. By construction, $\mathcal{Q}'\left(\hat{\theta}\right) \leq \mathcal{Q}\left(\hat{\theta}\right)$. The perturbation keeps the objective (43) and the constraints (44), (47), (48) intact. At the same time, it strictly increases $p\left(\hat{\theta}\right)$ and hence strictly relaxes the constraints (45) and (46) for all types deviating *to* type $\hat{\theta}$. Hence, the perturbed profile is implementable and delivers the same allocation as the original profile. As one can do this perturbation to all types $\theta \in \Theta$, the result follows.

Proof of Theorem 1. The seller's problem (43) can be seen as a maximization of a continuous function over a compact set. Hence, an optimal menu exists. By Proposition 5, there exists an optimal menu with all allocations located on the boundary of the feasibility set $\mathcal{F}$. By Proposition 3, such allocations are achieved by linear disclosures. The result follows.

Proof of Proposition 6. Introduce auxiliary attributes, as many as there are types. Define an auxiliary attribute x'_θ as the valuation of a type θ, $x'_\theta \triangleq v(\theta, x)$. The new attribute vector, x', is distributed over a set $X' \subseteq \mathbb{R}^{|\Theta|}$, according to the distribution of original attributes G and the valuation function $v(\theta, x)$. By construction, the valuation of each type can be defined as the corresponding auxiliary attribute, $v'(\theta, x') = x'_\theta$. This 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. The result follows.

Calculations behind Example 1. Consider the linear form (21). If $\alpha_1 + \alpha_2 \neq 0$, then, by rescaling the term α_0, the sum can be normalized to equal 1. By rearranging the terms, the linear form can be rewritten as:

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

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 coefficient sum. 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 $|\alpha_0|$.

Finally, if $\alpha_1 + \alpha_2 = 0$ and $\alpha_1 \neq 0$, then the linear form (21) is a linear disclosure that informs about the direction of types' difference:

$$q(x) = \begin{cases} 1, & \text{if } (\theta_1 - \theta_2) \cdot x \gtreqless \alpha_0', \\ 0, & \text{if } (\theta_1 - \theta_2) \cdot x \lesseqgtr \alpha_0', \end{cases}$$

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 Proposition 3 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 the feasibility set $\mathcal{F}$ in the auxiliary attributes, supported by the hyperplane orthogonal to the vector $(\alpha_1, \alpha_2, \alpha_0)$. If $\theta_1 \neq \theta_2$, the feasibility set has a strict interior. Hence, the set of boundary points supported by hyperplanes with $\alpha_1 + \alpha_2 = 0$ has a measure zero and, hence, generically does not matter for an optimal mechanism.

Proof of Theorem 2. The argument is given in the text. The only difference from Myerson (1981)'s problem is that $\mathcal{X}$ can take values in $[0, \mathbb{E}[x]]$, not in $[0,1]$. However, it does not affect the extremal nature of the solution.

Lemma 2. (Directional Decomposition)
Let $(x_1, x_2, \ldots, x_J)$ be J attributes distributed independently over $X \subseteq \mathbb{R}^J$ according to prior distributions $G_1, \ldots, G_J$. Let $E = (S, \pi)$, $\pi : X \to \Delta(S)$ 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 exists a collection of experiments $\{E_j\}_{j=1}^J$ such that experiment E_j provides the same information as E about attribute j and provides no information about other attributes: $E_j = (S, \pi_j)$ 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 $(x_1', x_2', \ldots, x_J')$ that have the same prior distributions as $(x_1, x_2, \ldots, x_J)$ but drawn independently of them. For a given j, construct E_j as an experiment that informs about the vector $\left(x_j, x_{-j}'\right)$ according to the likelihood function of E. By construction, E_j induces the same marginal distribution of beliefs about attribute j. However, as $(x_1, x_1', x_2, x_2' \ldots, x_J, x_J')$ are independent, it provides no information about other attributes. The result follows. $\square$

Proof of Proposition 7. Consider an arbitrary responsive experiment $E_j(\theta_j)$. By Lemma 2 there exists a linear disclosure $E'_j(\theta_j)$ such that the relevant attribute surplus and trade probability remain the same $\mathcal{X}'_j(\theta_j) = \mathcal{X}_j(\theta_j)$, $\mathcal{Q}\left(E'_j(\theta_j)\right) = \mathcal{Q}(E_j(\theta_j))$, and all other surpluses stay at the ex-ante expectations, $\mathcal{X}'_k(\theta_j) = \mathbb{E}[x_k]$ for all $k \neq j$. Hence, replacing $E_j(\theta_j)$ with $E'_j(\theta_j)$ does not change incentive compatibility within cohort Θ_j but, by Blackwell's Theorem, relaxes the incentive-compatibility constraints of other cohorts.

Proof of Theorem 3. Consider the seller's problem (34) and its arbitrary solution. Define $\underline{p} = \min_\theta \{p(\theta)\}$. This is the relevant price of the deviations across types. Towards the contradiction, assume $p(\theta) > \underline{p}$ for some type θ.

If $\mathbb{E}[x_{j(\theta)}] \geq \underline{p}$, then the incentive compatibility constraint is binding. Hence, $\mathcal{Q}(\theta)\, p(\theta) = \mathcal{X}(\theta) - \mathbb{E}[x_{j(\theta)}] + \underline{p}$. For small $\varepsilon > 0$ consider a modified mechanism with $\mathcal{X}'(\theta) = \mathcal{X}(\theta) + \varepsilon$, $\mathcal{Q}'(\theta)\, p'(\theta) = \mathcal{X}'(\theta) - \mathbb{E}[x_{j(\theta)}] + \underline{p}$. Because $\mathcal{Q}(\mathcal{X})$ is continuous, the mechanism remains incentive compatible yet brings higher revenue. Contradiction.

If $\mathbb{E}[x_{j(\theta)}] < \underline{p}$, then the individual-rationality constraint is binding. Hence, $\mathcal{Q}(\theta)\, p(\theta) = \mathcal{X}(\theta)$. For small $\varepsilon > 0$, consider the modified mechanism with $p'(\theta) = p(\theta) - \varepsilon$, $\mathcal{X}'(\theta)/\mathcal{Q}'(\theta) = \mathcal{X}(\theta)/\mathcal{Q}(\theta) - \varepsilon$. The mechanism remains incentive compatible yet brings higher revenue. Contradiction.

Now, consider optimal disclosure for a given object price. By feasibility and individual rationality, $\mathcal{X}(\theta)/\mathcal{Q}(\theta) \geq \max\left\{p, \mathbb{E}[x_{j(\theta)}]\right\}$. If $\mathcal{X}(\theta)/\mathcal{Q}(\theta) > \max\left\{p, \mathbb{E}[x_{j(\theta)}]\right\}$, then for small $\varepsilon > 0$ the mechanism with $\mathcal{X}'(\theta)/\mathcal{Q}'(\theta) = \mathcal{X}(\theta)/\mathcal{Q}(\theta) - \varepsilon$ is incentive compatible and increases trade probability, $\mathcal{Q}'(\theta) > \mathcal{Q}(\theta)$, and consequently, revenue. Contradiction.

Proof of Lemma 1. The seller's problem can be written as:

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

$$\begin{aligned}
\text{s.t.}\quad & \theta_j \mathcal{X}_j(\theta_j) - p_j(\theta_j)\, \mathcal{Q}_j(\theta_j) - r_j(\theta_j) \geq \left(\theta_j \mathcal{X}_j\left(\theta'_j\right) - p_j\left(\theta'_j\right)\right) \mathcal{Q}_j\left(\theta'_j\right) - r_j\left(\theta'_j\right), \quad \forall\, j, \theta_j, \theta'_j \in \Theta_j, \\
& \theta_j \mathcal{X}_j(\theta_j) - p_j(\theta_j)\, \mathcal{Q}_j(\theta_j) - r_j(\theta_j) \geq \theta_j \mathbb{E}[x_j] - p_k(\theta_k) - r_k(\theta_k), \quad \forall\, j, k, \theta_j \in \Theta_j, \theta_k \in \Theta_k, \\
& \theta_j \mathcal{X}_j(\theta_j) - p_j(\theta_j)\, \mathcal{Q}_j(\theta_j) - r_j(\theta_j) \geq 0, \\
& \mathcal{X}_j(\theta_j) \geq \mathcal{Q}_j(\theta_j)\, \mathbb{E}[x_j], \;\; \mathcal{Q}_j(\theta_j) = \mathcal{Q}_j(\mathcal{X}_j(\theta_j)), \quad \forall j, \theta_j \in \Theta_j.
\end{aligned}$$

Define the expected transfer function:

$$T_j(\theta_j) \triangleq \mathcal{Q}_j(\theta_j)\, p_j(\theta_j).$$

I can use standard one-dimension arguments within each cohort to establish the connection between the attribute surplus and the expected transfer function. Incentive compatibility requires the slope of the indirect utility function be equal to $\mathcal{X}_j(\theta)$ almost everywhere. Hence, the indirect utility function is convex and, by the Envelope Theorem, the optimal transfers can be recovered to be

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

Individual rationality and incentive compatibility within each cohort are satisfied by construction. However, incentive compatibility between different cohorts imposes one additional constraint,

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

where $\underline{p}$ is the minimal object price in the menu determined by $\{\mathcal{X}_j\}_{j=1}^J$. 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 4. The argument in the text establishes that all high types are offered the minimal price. The optimal mechanism should then solve the problem (38) 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:

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

I can then consider a relaxed problem with the original integral constraints and the monotonicity constraints dropped. In this relaxed problem, by Luenberger (1969), there exist the Lagrange multipliers $\{\lambda_j\}$ such that optimal $\mathcal{X}_j(\theta_j)$ maximize the Lagrange function

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

over a domain $\mathcal{X}_j(\theta_j) \in \left[0, \mathcal{X}^*\left(\overline{\theta}_j\right)\right]$. If all type distributions are log-concave, then the integrand is increasing in θ_j. Hence, the optimal $\mathcal{X}_j(\theta_j)$ are bang-bang: $\mathcal{X}_j(\theta_j) = 0$ for $\theta_j < \theta_j^*$, $\mathcal{X}_j(\theta_j) = \mathcal{X}^*\left(\overline{\theta}_j\right)$ for $\theta_j > \theta_j^*$. This relaxed solution corresponds to a single item per each attribute cohort. Hence, the relaxed constraints are satisfied and the relaxed solution solves the original problem as well.

References

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

Armstrong, M. (1996): "Multiproduct Nonlinear Pricing," *Econometrica*, 64, 51–75.

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. Y. Weintraub (2017): "The Scope of Sequential Screening with Ex-Post Participation Constraints," *Discussion paper*.

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.

Dubé, J.-P. and S. Misra (2017): "Scalable Price Targeting," *NBER Working Paper*.

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.

Heumann, T. (2018): "Optimal Sequential Screening with Ex Post Participation Constraint," *Discussion paper*.

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.

Krähmer, D. (2017): "Information Disclosure and Monopolistic Screening," *Discussion paper*.

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.

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

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