Disclosure and Pricing of Attributes*

Alex Smolin

July 8, 2020

Abstract

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

I characterize revenue-maximizing menus. All experiments belong to a class of linear disclosure policies. An optimal menu may be nondiscriminatory and qualitatively depends on the structure of buyer heterogeneity. The analysis informs on the benefits of partial disclosure in pricing settings.

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

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

*Toulouse School of Economics, 1 Esplanade de l'Université, 31080 Toulouse Cedex 06, France, alexey.v.smolin@gmail.com. I thank Dirk Bergemann and Daniel Krähmer for helpful conversations and Levent Celik for an excellent discussion session. I am very grateful to numerous participants of research seminars at Toulouse, Mannheim, Carlos III Madrid, Bonn, Aalto, Oxford, Emory, Essex, Nottingham, Bologna, Edinburgh, Warwick, Amsterdam, Royal Holloway, Boston University, HBS, Rice, Northwestern (Kellogg), Michigan State, Carnegie Mellon (Tepper), Utah (Eccles), Bielefeld, Valencia, Collegio Carlo Alberto, Einaudi, PSE as well as at the SAET 2017, ESSET Gerzensee 2018, Stony Brook 2018, EEA-ESEM 2018, EARIE 2018, Columbia/Duke/MIT/Northwestern IO Theory 2018, Warwick CRETA 2019, and HSE-Vienna 2019 conferences. I acknowledge funding from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

1 Introduction

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

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

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

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

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

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

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

In Section 4, I study the design of disclosure 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 buyer types prefer learning about different aspects of the object. Theorem 1 shows that an optimal way to combine these two functions is through a specific class of experiments—linear disclosures. A linear disclosure informs whether a linear combination of attributes is above or below a specified threshold. This disclosure guides the allocation and can be seen as informing the buyer about the valuation of a virtual type that accounts for incentive constraints.[2] Notably, this result requires no assumptions on the distributions of types or attributes and, as such, can be generalized to arbitrary valuation functions.

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

In Section 5, I study optimal pricing mechanisms. In Theorem 2, I establish that if all buyer types value the same, always positive, attribute, then no information is optimally provided and the seller posts a single price for the object. The intuition behind this result lies in the product structure of the buyer's valuation. When all types value the same attribute,

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

any disclosure realization simply scales their valuations and the corresponding demand curve. Even if the seller could condition the price on this realization, she would charge scaled prices, serve the same types, and obtain scaled revenue. By the martingale property of Bayesian expectations, the seller can obtain the same revenue by providing no information.[3]

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

I formalize this intuition beginning with the case in which each type values a distinct independent attribute so that type valuations are independently distributed (Section 5.2.1). I show that an optimal menu features free-of-charge partial disclosure and no price discrimination. Information is not priced as the payment can be backloaded into the object price. Price dispersion is not profitable, because it implies that the seller could extract more surplus on some items by simultaneously lowering their object price and changing their informational content. The optimal menu admits nondiscriminatory implementation—posting a single price for the object and informing the buyer whether the object is sufficiently good along each attribute.

In this optimal menu, the seller effectively persuades each type to purchase the object separately at a fixed price by informing him whether his valuation is sufficiently high. In Sections 5.2.2 and 5.3, I show that similar mechanisms are optimal in a broad range of settings as long as the types can be seen as belonging to distinct cohorts—with low valuation correlation across and high valuation correlation within cohorts.

In Section 6, I first discuss what product information needs to be priced and when. Second, I emphasize that attribute information can rotate the demand curve locally and can justify attribute shrouding. Finally, I indicate how a multiattribute framework complements the existing disclosure frameworks.

Related Literature This paper is about disclosure and pricing. One strand of the related 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 show that within a simple parameterized class, an optimal disclosure is extreme: either full or no disclosure. Bergemann and Pesendorfer (2007) further observe that if there is common knowledge of positive trade gains, then no disclosure dominates any other possible disclosure because it allows the seller to extract the full expected

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

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 the global rotation 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 partial disclosure can dominate both full and no disclosure, even if there is common knowledge of positive trade gains (Sections 4.4, 6.3).

At the same time, when the buyer has private information, it is natural to study discriminatory mechanisms and how they can be used to screen buyer types. In an influential paper, Eső and Szentes (2007) study settings in which the attribute and the buyer's type enter the valuation "additively." In such settings, under certain distributional assumptions, the seller optimally provides full disclosure.[5] However, Li and Shi (2017) show that the seller should withhold some information if the types represent private information about the object. Section 6.3 contains a detailed discussion related to these two papers.

All of the works described above operate in single-dimensional settings. Under complete object information, when comparing any two objects, all buyer types agree on the ranking. However, in practice, many products are multidimensional, with different attributes that appeal to different buyers. In this paper, I demonstrate that these settings can be successfully studied and lead to qualitatively different results. Despite the richness of the attribute space, optimal experiments belong to a tractable class of linear disclosures (Section 4.4). Optimal mechanisms feature partial disclosure but can be remarkably simple (Section 5.2).

This paper builds on several existing frameworks. The multiattribute buyer's valuation follows the characteristic model of Lancaster (1966). An unrestricted search for an optimal disclosure policy is a defining feature of the Bayesian persuasion literature (Rayo and Segal (2010), Kamenica and Gentzkow (2011)).[6] The screening analysis builds on the mechanism design machinery of Myerson (1981, 1982) and Kolotilin, Mylovanov, Zapechelnyuk, and Li (2017). Finally, information design with screening and monetary transfers has already appeared in my previous, co-authored, work (Bergemann, Bonatti, and Smolin (2018)). There, the seller can price only information—the buyer's action is not contractable. By contrast, in this paper, the seller can price both services and as a result, in many settings, provides the information free of charge.

[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. Similarly, Dworczak (2020) shows that optimal disclosure can be partial in the presence of aftermarkets.

[5] Wei and Green (2019) investigate extensions in which the information must be provided free of charge.

[6] Almost all of this literature studies one-dimensional settings. Section 7.2 of Dworczak and Martini (2019) provides a recent elegant exception.

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* $x = (x_1, \ldots, x_J) \in X = \mathbb{R}^J$. The buyer's preferences towards each attribute constitute the buyer's *type* $\theta = (\theta_1, \ldots, \theta_J) \in \Theta \subseteq \mathbb{R}^J$. The ex post buyer's valuation for the object is:[7]

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

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

Prior Information Attributes are distributed over X according to a cumulative distribution function G. The buyer and seller are symmetrically informed about attributes. The type space Θ can be finite or infinite. The buyer's type is his privately known preferences, which are uncorrelated with attributes. From the seller's perspective, the types are distributed according to a cumulative distribution function F. Until Section 5, I do not impose any structural assumptions on the attribute and type distributions. The only technical 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.[8] That is, the experiment can provide no information, or no disclosure, $\underline{E} \triangleq (\underline{S}, \underline{\pi})$, with $\underline{S}$ being a singleton; it can fully reveal attributes, or provide full disclosure, $\overline{E} \triangleq \left(\overline{S}, \overline{\pi}\right)$, with $\overline{S} = X$ and $\overline{\pi}(x)$ placing probability 1 on $s = x$; or it can provide partial information.

Selling Mechanism To investigate the scope of screening, I allow the seller to price both the object and the information that she provides. The seller designs a menu of items, $i \in \mathcal{I}$:

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

The items consist of a collection of experiments $E(i)$ and tariff functions $r(i) \geq 0$, $p(i) \geq 0$. The first tariff captures the price of information—the upfront payment paid irrespective of a trade. The second tariff captures the price of the object, paid only if the trade occurs. Effec-

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

[8]Subject to a mild technical condition that requires that all conditional expectations exist.

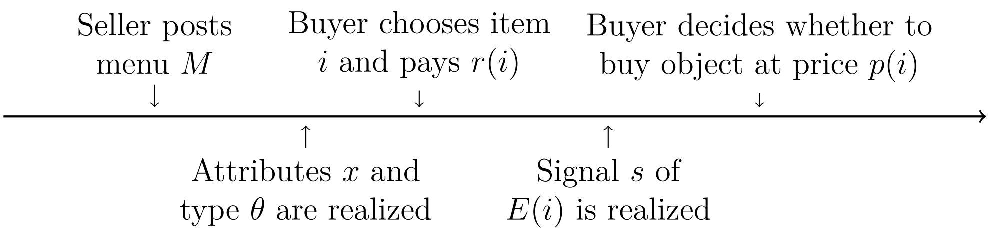

Figure 1: Timeline of the selling mechanism.

tively, the menu is a collection of call options that differ in monetary terms and information disclosure, which designed to screen different buyer types.

The timing is as follows.[9] The seller posts a menu M. The attribute vector x and the buyer's type θ are realized. Then, the buyer chooses an item $i \in \mathcal{I}$ and pays the corresponding price $r(i)$. Next, the buyer observes a signal s from the experiment $E(i)$ and decides whether to buy the object at the price $p(i)$. Finally, the payoffs are realized. The timing is illustrated in Figure 1.

The timing implies that the seller commits to a menu before the realization of the attributes x and the type θ. The attributes x and signals s are not contractible.[10] Sales are deterministic—a payment guarantees a transaction. Sequential interactions between players are excluded, so belief-elicitation schemes and scoring rules are not available.[11]

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

3 Illustrative Example

I begin by illustrating the workings of disclosure and pricing through a simple example. Let there be two attributes that are uniformly and independently distributed, $J = 2$, $x_1 \sim U[0,1]$, $x_2 \sim U[0,2]$. Let there be a continuum of buyer types split into two cohorts. Types $\theta_1 \in \Theta_1$ value only the first attribute, whereas types $\theta_2 \in \Theta_2$ value only the second attribute. Each cohort is equally likely, and within each cohort, the marginal type distributions are uniform over $[0,2]$ so that the average type is equal to 1.

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

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

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

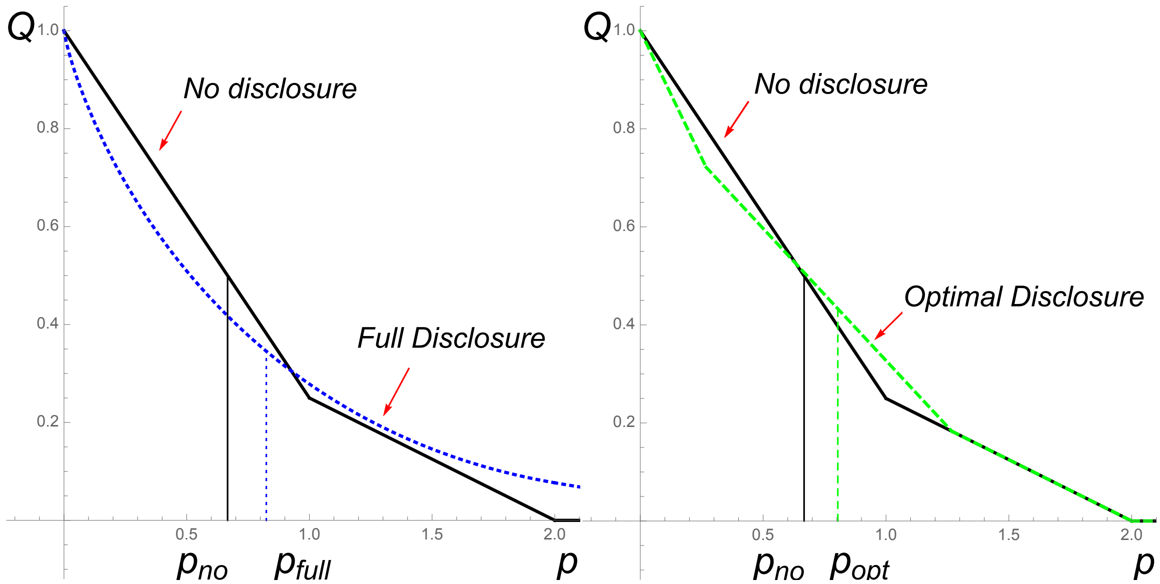

Figure 2: Attribute disclosure and demand transformation. Left: demand curves under no disclosure and full disclosure. Right: demand curves under no disclosure and optimal disclosure. Vertical lines indicate revenue-maximizing prices.

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

If the seller provides no disclosure, then she faces a piecewise-linear demand curve with a kink at price $p = 1$, as illustrated in Figure 2. Given this demand, the optimal no-disclosure price is $p_{no} = 2/3$, which results in revenue $1/3$. If the seller provides full disclosure, then the valuation of each type stochastically changes according to the realization of a relevant attribute; as a result, the type valuations spread. The demand decreases for lower prices, as intermediate types may learn that the attributes are low, and increases for higher prices, as the intermediate types may learn that the attributes are high. The overall effect is negative. The optimal full-disclosure price is $p_{full} \simeq 0.82$ with the corresponding revenue $0.28 < 1/3$.

However, the seller can increase revenue by providing partial disclosure. Consider an experiment that reveals whether the first attribute is above some threshold α_0 but provides no information about the second attribute. Under this disclosure, demand decreases for low prices, increases for medium prices, and remains the same for high prices. Intuitively, this disclosure targets types that ex ante value the object less and persuades them to buy at higher prices. If the disclosure threshold is chosen optimally, $\alpha_0^* \simeq 0.27$, then the overall effect is positive. The optimal price is $p_{opt} \simeq 0.80$, resulting in revenue $0.35 > 1/3$.

The effectiveness of partial disclosure raises a question of whether the seller can increase the revenue even further if she employs a discriminatory menu with upfront payments, offering a range of disclosures at different prices. Perhaps surprisingly, the answer is negative, as I show in Section 5.2. The simple "advertising" mechanism is optimal.

In this example, the attributes are independently distributed. Together with a particular type structure, it allows the optimal disclosure to be "one-dimensional"—informing about a single attribute. In a more general case, when attributes are correlated, one can expect optimal disclosure to be richer—informing about several attributes simultaneously.

4 Design of Disclosure Policies

I proceed with studying the revenue-maximizing menu design. I begin by 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 steps. First, the class of optimal disclosure policies can be identified without explicitly characterizing a pricing mechanism. I show this in the current section. Second, optimal pricing mechanisms can be derived in the leading settings, which is shown in Section 5.

4.1 Buyer's Problem

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

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

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. By 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 option i can be written as:

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

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

4.2 Responsive Menus

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

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

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

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

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

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

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

A focus on responsive menus enables the characterization of every experiment E in the menu by its *trade function*:

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

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

4.3 Seller's Problem

Proposition 1 enables each experiment to be associated with its trade function (8) and the seller's problem to be written 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 object 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 $k = 1, \ldots, 4$. These constraints ensure that each type prefers truth telling to all double-deviating strategies: misreporting and following the recommendations, "swapping" the buying decisions, always buying, or never buying. Deviations from θ to θ are included and ensure that the types are obedient on-path after truth telling.

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

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

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

Several challenges are involved in the problem. First, the seller maximizes over a large class of all functions from a multidimensional space X. Second, it is a priori not clear what kinds of deviations are binding and, hence, relevant for the design problem: the buyer's type has no single-dimensional structure, and there is an additional multiplicity of constraints caused by double deviations.[12]

The following observation is crucial to address the experimental complexity: only two

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

coarse statistics, not the entire trade function, matter for the revenue-maximizing problem. Namely, for a given responsive experiment E, the associated trade function q achieves the *attribute surplus* and the *(total) trade probability*:[13]

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

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

The formulations (9), (10), and (11) reveal that, due to the linearity of integration, these statistics are the only economically relevant parameters of the problem. A change in the trade function $q(\theta,\cdot)$ that does not affect the attribute surplus and the trade probability affects neither the buyer's incentives nor the seller's revenue. Accordingly, the seller can maximize directly over attribute surpluses and trade probabilities. In what follows, I will refer to $\mathcal{X}(q(\theta,\cdot))$ and $\mathcal{Q}(q(\theta,\cdot))$ as $\mathcal{X}(\theta)$ 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 trade never occurs, $q(\cdot)\equiv 0$, so the attribute surpluses must also be nil. At the other extreme, if the trade probability is 1, then the trade always occurs, $q(\cdot)\equiv 1$, so the attribute surplus is equal to its ex ante expectation $\mathbb{E}[x]$. Intermediate values of trade probability provide more freedom to choose attribute surpluses because the seller can select the regions in which the trade recommendations are sent. The corresponding feasibility set $\mathcal{F}\subseteq\mathbb{R}^{J+1}$ is:

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

The shape of $\mathcal{F}$ is determined by the attribute distribution G. It is typically strictly convex.

These observations reduce the search to the following problem:

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

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

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

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

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

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

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

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

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

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

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

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

4.4 Optimal Disclosure

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

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

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

Furthermore, Proposition 2 emphasizes the distinctive feature of the seller's problem that combines information and mechanism design. In a typical information design problem, the payoff structure is exogenously fixed and, unless the receiver's indirect utility is concave everywhere, disclosure appears in the optimal mechanism for some prior distributions. Indeed, if the prices were fixed and the buyer's preferences had sufficiently low intensity, then the seller would have to provide some information to persuade the buyer to buy the object. By contrast, when the seller has control over monetary incentives, she can compensate for the lack of information with lower prices and does find it optimal to do so.

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

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

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

A linear disclosure informs the buyer whether a linear combination of attributes is above or below a specified threshold. A linear disclosure is conditionally deterministic, assigning probability one to some some signal 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 price $p = \alpha_0$. If attributes are always positive and independently distributed, a linear disclosure admits additional interpretations. If elements of the coefficient vector α are positive, this disclosure can be viewed as a “level” disclosure. Observing a “trade” recommendation uniformly increases attribute expectation, whereas observing a “no-trade” recommendation uniformly decreases it. By contrast, if the elements of a coefficient vector α have different signs, then a linear disclosure can be viewed as a “comparative” disclosure between the attribute groups of different signs. A “trade” recommendation increases the attribute expectations in one group and decreases them in the other group.

Note that the likelihood function of a linear disclosure is not restricted on the defining hyperplane, $\{x \mid \alpha \cdot x = \alpha_0\}$. Furthermore, the hyperplane does not exist for $\alpha \equiv 0$ and α_0 being strictly positive or negative. Those linear disclosures correspond to never-trade and always-trade uninformative experiments.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Corollary 1. (General Payoffs)
Let X be an arbitrary attribute set, $v(\theta, x)$ be a general valuation function, and $|\Theta| < \infty$. Then, there exists an optimal menu in which every experiment has a linear form, i.e., for

any experiment, there exist $\alpha : \Theta \to \mathbb{R}$ *and* $\alpha_0 \in \mathbb{R}$*, not all zeros, such that:*

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

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

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

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

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

5 Design of Pricing Mechanisms

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

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

5.1 Single Attribute

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Remark 2. The no-disclosure mechanism may not be uniquely optimal. In fact, the analysis of Eső and Szentes (2007) can be applied to show that if the type distribution F has a monotone hazard rate property, then full disclosure is also optimal.[14] However, it must be accompanied by a complex structure of upfront payments and object prices.

5.2 Single-Minded Buyer

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, therefore, disclosure. In this section and introduce and study a tractable type structure which, nevertheless, allows to capture both vertical and horizontal heterogeneity of tastes.

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

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

Thus, single-minded types allow for a simpler notation. I can represent the types by J

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

attribute cohorts Θ_j such that all types within the same cohort value the same attribute. I slightly abuse the notation and let the type subscript identify the attribute cohort and the type value identify the valuation intensity, so that $\Theta_j \subseteq \mathbb{R}_+$ and

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

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

A buyer is *single-minded* if all types $\theta \in \Theta$ are single-minded and the attribute values are independently distributed such that $x_j \sim G_j$ and $G(x) = \times_j G_j(x_j)$.[15] The independence requirement is substantive. Without it, any buyer can be viewed as being single-minded by redefining the attributes as in the proof of Corollary 1.

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

The class of optimal experiments can be narrowed. Because attributes are independently distributed, a type $\theta_j \in \Theta_j$ values only information about attribute j. This observation suggests an optimal way to screen single-minded types: if the buyer reports type $\theta_j \in \Theta_j$, then the seller should provide information only about attribute j. Providing any other information would make misreporting more appealing without adding value for truth telling. At the same time, a linear disclosure informative only about attribute j is a binary monotone partition defined on this attribute.

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

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

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

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

[15]The name is inspired by "single-minded" bidders in combinatorial auctions who value specific bundles. See, for example, Lehmann et al. (2002). I thank Laura Doval for drawing my attention to this connection.

5.2.1 Orthogonal Types

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

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

$$\theta_j = 1 \quad \forall\, j = 1, \ldots, J. \tag{36}$$

Hence, I will omit the dependence on the intensity 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 that is hence not informative about attribute j. Thus, the type has no reason to act on the experiment realization, and the tightest incentive-compatibility constraint is one in which he always buys. All others can be dropped. The seller's problem can be written as:

$$\max_{\{r_j, \mathcal{X}_j, p_j\}_{j=1}^J} \sum_{j=1}^J f(\theta_j)\left(r_j + \mathcal{Q}_j p_j\right) \tag{37}$$

$$\text{s.t.} \quad \mathcal{X}_j - p_j \mathcal{Q}_j - r_j \geq \mathbb{E}[x_j] - p_k - r_k,\ \forall\, j, k = 1, \ldots, J, \tag{38}$$

$$\mathcal{X}_j - p_j \mathcal{Q}_j - r_j \geq 0,\ \mathcal{X}_j \in [0, \mathbb{E}[x_j]],\ \mathcal{Q}_j = \mathcal{Q}_j(\mathcal{X}_j),\ \forall\, j = 1, \ldots, J. \tag{39}$$

This problem resembles a collection of one-dimensional mechanism design problems 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 attribute surpluses $\mathcal{X}_j$ are purely horizontal—they are valuable only to type θ_j. The object prices p_j are mixed as they are paid only if the type decides to trade. Second, the problem features nonlinear terms $p_j \mathcal{Q}_j$ and $\mathcal{Q}_j(\mathcal{X}_j)$.

Because of these differences, I cannot apply standard mechanism design techniques. Instead, I solve the problem via a sequence of simplifications. In the first step, I observe that using upfront payments is detrimental. For any $r_j > 0$, the seller can reduce the transfer and increase p_j while keeping the total expected transfer $r_j + \mathcal{Q}_j p_j$ the same. This change does not affect the 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 would always purchase the object.

In the second step, I show that object price discrimination is also not profitable. Because all experiments off the truth-telling path bring no information value, the optimal deviation

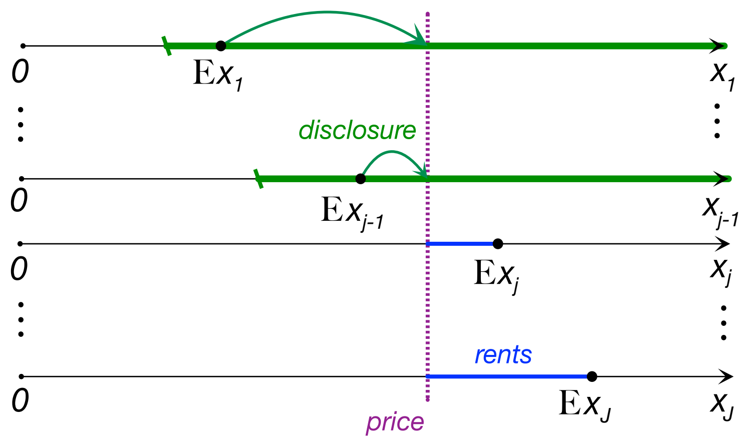

Figure 3: Optimal mechanism in the case of orthogonal types. Green indicates attribute regions in which a purchase recommendation is sent for the types with partial disclosure. Blue indicates the types' rent conditional on a trade, for the types with no disclosure. Attributes are ordered by increasing ex ante expectations.

is to the types associated with the lowest price $\underline{p}$. If any object price variation exists, 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. At the same time, for θ_j to be willing to pay a higher price, the item must contain partial disclosure, which 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.

The intuition is simple. If there was a single buyer type, then the seller could extract the full surplus. Similarly, with many types, the seller can extract additional surplus whenever some types get rents yet are irrelevant for the incentives of other types.

Once I establish that the optimal mechanism is nondiscriminatory, finding optimal experiments is straightforward. The seller should provide minimal information sufficient to convince the buyer to make a purchase. If the type θ_j is ex ante sufficiently optimistic, $\mathbb{E}[x_j] \geq p$, then the seller should provide no attribute information, $E_j = \underline{E}$. Otherwise, the seller should increase the type's expectation up to the object price.

Theorem 3. (Optimal Menu, Orthogonal Types)
If the buyer is single-minded and the types are orthogonal, then in an optimal responsive menu for all $j = 1, \ldots, J$: $r_j = 0$, $p_j = p$, and 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 3. I emphasize 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. 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. Because each type values only one attribute, the seller can provide a single combined disclosure followed by the optimal posted price.

5.2.2 Continuum of Types

The analysis can be extended to allow for vertical heterogeneity within the attribute cohorts (cf., Kolotilin et al. (2017)). 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 according to distribution function F_j with the monotone hazard rate property.

In this general case, for each attribute j, the seller needs to design the tariff and the attribute surplus functions, $r_j(\theta_j)$, $p_j(\theta_j)$, and $\mathcal{X}_j(\theta_j)$. This process requires incorporating additional incentive constraints that correspond to the type pretending to be another type within the same cohort and following the experiment recommendation.

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, $r_j(\theta_j) \equiv 0$. This change does not affect the revenue or incentive compatibility within each attribute cohort; however, it relaxes the 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. That is, higher types must trade with a higher probability but lower conditional expectations of the relevant attribute.

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

$$\max_{\{\mathcal{X}_j(\theta_j)\}_{j=1}^J} \sum_{j=1}^{J} f(\Theta_j) \int_0^{\overline{\theta}_j} \left(\theta_j - \frac{1 - F_j(\theta_j)}{f_j(\theta_j)}\right) \mathcal{X}_j(\theta_j) \, \mathrm{d}F_j(\theta_j) \tag{40}$$

$$\text{s.t.} \quad \mathcal{X}_j(\theta_j) \text{ is non} - \text{decreasing}, \;\; \mathcal{X}_j(\theta_j) \in [0, \mathbb{E}[x_j]], \tag{41}$$

$$\int_0^{\overline{\theta}_j} \mathcal{X}_j(\theta_j) \, d\theta_j \geq \overline{\theta}_j \mathbb{E}[x_j] - \underline{p}(\mathcal{X}_1(\cdot), \ldots, \mathcal{X}_J(\cdot)) \;\; \forall j = 1, \ldots, J. \tag{42}$$

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

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

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 a given price. If the type distributions have the monotone hazard rate property, then the solution to the problem is a collection of single-step functions. The solution corresponds to only one item per attribute cohort. It satisfies the original constraints and therefore solves the original problem.

Theorem 4. (Optimal Menu, Single-Minded Buyer)
If the buyer is single-minded and the type distributions have the monotone hazard rate property, then in an optimal responsive menu for all $j = 1, \ldots, J$ and $\theta_j \in \Theta_j$: $r(\theta_j) = 0$, $p(\theta_j) = p$, and $E_j(\theta_j) = E_j$ where E_j is a binary monotone partition of x_j.

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

5.3 Differentiated Types and Separate Persuasion

The findings of the previous section can be extended to general attribute distributions and valuation functions. For a given price p, define a *separate persuasion mechanism* $M^{SP}(p)$ as a direct menu in which $r^{SP}(\theta) = 0$, $p^{SP}(\theta) = p$, and $E^{SP}(\theta)$ recommends to trade when $v(\theta, x)$ is above a threshold; the threshold is chosen so that $\mathbb{E}\left[v(\theta, x) \mid E^{SP}(\theta), s^+\right] = \max\{p, \mathbb{E}[v(\theta, x)]\}$. That is, the seller fixes an object price and for each type provides minimal valuation information to persuade him to make a purchase.

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

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

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

Proposition 4 can be used to find optimal mechanisms in environments in which the types are sufficiently differentiated. Theorem 3 can be viewed as its corollary. More generally, fix a profile of orthogonal types $\hat{\Theta} = (\theta_1, \ldots, \theta_J) \in \mathbb{R}^{J \times J}$, the type frequencies, and the attribute distribution. Let the valuation function be linear (1) and the attributes be independently and continuously distributed.

Corollary 2. (Differentiated Types)
If p^* is a uniquely optimal price at an orthogonal type profile $\hat{\Theta}$ and $p^* \neq \mathbb{E}\left[v\left(\hat{\theta}, x\right)\right]$ for all $\hat{\theta} \in \hat{\Theta}$, then for any type profile in some Euclidean neighborhood of $\hat{\Theta}$, a separate persuasion mechanism is optimal.

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

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

6 Discussion

6.1 Pricing of Product Information

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

Proposition 5. (Payment Backloading)
In any given responsive menu, $r(\theta)$ can be decreased, and $p(\theta)$ can be increased without any loss of revenue when (i) $\mathcal{Q}(\theta) \leq 1/2$ or (ii) no type is willing to act contrary to both recommendations if being offered the experiment $E(\theta)$.

Proposition 5 holds for any valuation function. The logic is as follows. For any given item, the seller can backload the upfront transfer into the object's price, while maintaining the expected payment. This adjustment preserves the incentives of the types who, if being offered $E(\theta)$, would like to follow the recommendations. It relaxes the incentives of the types who would like to always purchase the object. The swapping constraint (17) tightens if and only if the trade probability is less than 1/2. In fact, this argument reveals that the second condition of the proposition needs to hold only for the types along the binding constraints.

By Proposition 5, a given experiment may need to be priced only if it induces a relatively frequent trade and some types would like to mismatch its recommendations. That is, a priced experiment should bring "bad news" by infrequently informing the buyer that the product is not worth its price. Moreover, this information should lead to a disagreement, with some type using it in the opposite fashion. In some cases, one can exclude the latter possibility.

Corollary 3. (No Information Pricing)
In an optimal menu, the price of an experiment E can be set to zero if (i) $J > 1$, attributes are independent, $v(\theta, x) = \theta \cdot x$, $\Theta \subseteq \mathbb{R}_{+}^{J}$, and E is a linear disclosure with $\alpha \in \mathbb{R}_{+}^{J}$ or (ii) $J = 1$, $v(\theta, x)$ is increasing in x for all $\theta \in \Theta$ and E is a binary monotone partition.

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

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

6.2 Demand Transformation and Shrouded Attributes

As I discussed in Section 3, one way to think about attribute disclosure is in terms of its impact on the demand curve that the seller faces. This particular effect was studied by

Johnson and Myatt (2006). They restrict their attention to disclosures that spread type valuations "uniformly." Such disclosures translate into *global* rotations of the demand curve. The authors show that in many settings, the optimal global rotations are extreme and correspond to either no disclosure or full disclosure: no disclosure is associated with a mass market characterized by a low price and high demand; full disclosure is associated with a niche market characterized by a high price and low demand.

By contrast, I show that attribute disclosure can rotate the demand curve *locally* (see Figure 2). The local rotations correspond to partial disclosures that target specific types and, hence, affect the demand curve over a particular price segment. They can outperform full and no disclosure in both mass and niche markets. Multiple attributes are required for this result—recall that no disclosure remains optimal in a one-dimensional framework.

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

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

6.3 Alternative Disclosure Settings

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

Eső and Szentes (2007) study discriminatory mechanisms in a valuation-rank framework, in which disclosure corresponds to statements such as “Your valuation is in your y-th percentile,” with y being the same for all types. The authors obtain two main qualitative results: first, they show that full information disclosure is generally optimal; and second, they show that the seller cannot benefit from conditioning the price on the disclosure realization.[16]

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

$$\xi(\theta) \sim U[0,1] \quad \forall\, \theta \in \Theta. \tag{43}$$

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

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

Consequently, neither of the qualitative results of Eső and Szentes (2007) holds in this example. The optimal full-disclosure mechanism can be calculated to be $r_1 = r_2 = 1/2$, $p_1 = p_2 = 0$: anticipating information revelation, the seller would prefer to effectively sell the object in advance. The corresponding revenue is $1/2$. The seller can do strictly better by providing partial private disclosure. Building on the results of Section 5.2, it can be shown that an optimal mechanism informs the buyer, free of charge, whether the first attribute is above or below $1/2$ and then follows with a posted object price of $3/4$. This mechanism obtains revenue $9/16 > 1/2$.

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

$$\Pi = \int_0^1 \int_0^2 \frac{1}{2} \max\left\{\min\{x_1, x_2\}, \frac{\max\{x_1, x_2\}}{2}\right\} \mathrm{d}x_1 \mathrm{d}x_2 = \frac{29}{48} > \frac{9}{16}. \tag{44}$$

This example also emphasizes the difference between the multiattribute setting and the setting of Li and Shi (2017). These authors study common value settings in which the types represent private information about the object. In their settings, information disclosure can be seen as a valuation-level disclosure that corresponds to statements such as "Your valuation is above x," with x being the same for all types. However, in general, attribute information

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

affects the valuation of different types differently according to the valuation function. In the example, any experiment informative only about attribute x_i affects the valuation of only type θ_i. Consequently, attribute information cannot be modeled as an experiment that informs the buyer directly about his valuation. This would misrepresent the buyer incentives in terms of choice across experiments.

7 Conclusion

I studied a monopolist who sells a multiattribute object to a privately informed buyer and showed that the seller can benefit from the disclosure of attribute information. The benefit comes through two channels. First, disclosure can be used as a screening device, leveraging the 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 in many settings screening is not beneficial and information should be disclosed partially and free of charge. In those settings, the choice of information content is more important than the choice of its pricing.

In this paper, I deliberately focused on the simplest model of pricing and information control. In practice, additional details 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 multiattribute disclosure framework that I have outlined.

8 Appendix

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

Proof of Proposition 2. 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 no type buys the object with probability one. Construct a new menu M' as follows. Select a type $\overline{\theta}$ with the highest expected payment $\overline{T} = r\left(\overline{\theta}\right) + p\left(\overline{\theta}\right)\mathcal{Q}\left(\overline{\theta}\right)$. 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 that of the original menu:

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

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

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

Proof of Lemma 1. As the expected valuations of all types exist, set $\mathcal{F}$ is a continuous image of a compact set. Therefore, $\mathcal{F}$ is compact.

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

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

The same argument can be applied to the trade probability. Thus, $(\mathcal{X}_3, \mathcal{Q}_3)$ is a convex combination of $(\mathcal{X}_1, \mathcal{Q}_1)$ and $(\mathcal{X}_2, \mathcal{Q}_2)$ and belongs to the feasibility set $\mathcal{F}$.

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.[17] A point $\left(\hat{\mathcal{X}}, \hat{\mathcal{Q}}\right)$ belongs to the boundary of $\mathcal{F}$ if and only if there are coefficients (λ, λ_0), not all zero, such that:

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

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

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

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

Proof of Lemma 2. The seller's problem can be written in terms of attribute surpluses and trade probabilities (15). Consider any feasible profile $r(\theta)$, $\mathcal{X}(\theta)$, $\mathcal{Q}(\theta)$, $p(\theta)$ that satisfy constraints (16), (17), (18), (19), (20). 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 Lemma 1, $\mathcal{F}$ 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 (15) and the constraints (16), (19), (20) intact. At the same time, it strictly increases $p\left(\hat{\theta}\right)$ and hence strictly relaxes constraints (17) and (18) for all types that deviate to type $\hat{\theta}$. Accordingly, the perturbed profile is implementable and delivers the same allocation as the original profile. As one can apply this perturbation to all types $\theta \in \Theta$, the result follows.

[17] 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 nonempty interior.

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

Proof of Corollary 1. Define an auxiliary attribute x'_θ as the valuation of a type θ, $x'_\theta \triangleq v(\theta, x)$. By construction, the valuation of each type can be defined as $v'(\theta, x') = x'_\theta$. 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.

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

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

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 $\alpha_1 + \alpha_2$. This is a neighborhood disclosure with $\hat{\theta} = \alpha_1\theta_1 + \alpha_2\theta_2$ and $\alpha_1 + \alpha_2 = 1$.

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

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

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

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

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

Lemma 4. (Directional Decomposition)
Let $(x_1, \ldots, x_J)$ be J attributes independently distributed over $X \subseteq \mathbb{R}^J$ according to distributions $G_1, \ldots, G_J$. Let $E = (S, \pi)$ be an arbitrary experiment. Let $(\mu(s, E), \Pr(s, E))$ be the belief distribution induced by E so that $\mu(s, E)$ is a distribution over X conditional on s given E. Denote by $\mu_j(s, E)$ the jth marginal distribution of $\mu(s, E)$. Then, there exist experiments $\{E_j\}_{j=1}^J$ such that: $E_j = (S, \pi_j)$ induces a belief distribution $(\mu(s, E_j), \Pr(s, E_j))$ with $\mu(s, E_j) = (\mu_j(s, E), G_{-j})$ and $\Pr(s, E_j) = \Pr(s, E)$ for all $s \in S$.

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

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

Proof of Theorem 3. Consider an arbitrary solution to the seller's problem (37). Define $\underline{p} = \min_\theta \{p(\theta)\}$. Towards the contradiction, assume that $p(\theta) > \underline{p}$ 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, which is a contradiction. Moreover, note that as $\mathcal{X}'(\theta) > \mathcal{X}(\theta)$, $\alpha_{0j(\theta)}' < \alpha_{0j(\theta)}$; hence, $\mathcal{X}'(\theta)/\mathcal{Q}'(\theta) = \mathbb{E}\left[x_{j(\theta)} \mid x_{j(\theta)} \geq \alpha_{0j(\theta)}'\right] < \mathbb{E}\left[x_{j(\theta)} \mid x_{j(\theta)} \geq \alpha_{0j(\theta)}\right] = \mathcal{X}(\theta)/\mathcal{Q}(\theta)$ and $\mathcal{Q}'(\theta) > \mathcal{Q}(\theta)$. Thus, $p'(\theta) < p(\theta)$.

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, which is a contradiction.

Now, consider the optimal disclosure for a given object price. According to 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. This is a contradiction.

Proof of Lemma 3. The seller's problem can be written as:

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

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

where $\mathcal{Q}_j(\theta_j) \equiv \mathcal{Q}_j(\mathcal{X}_j(\theta_j))\ \forall j, \theta_j \in \Theta_j$. Define the expected transfer function as $T_j(\theta_j) \triangleq \mathcal{Q}_j(\theta_j)\, p_j(\theta_j)$. I can use standard one-dimensional arguments within each cohort to establish the connection between the attribute surplus and the expected transfer function:

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

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

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

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

Proof of Theorem 4. The argument in the text establishes that all high types are offered the minimal price. The optimal mechanism should then solve the problem (40) 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:

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

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

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

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

Proof of Proposition 4. Introduce auxiliary attributes as in the proof of Corollary 1. By the proof of Theorem 3, $M^{SP}(p^*)$ solves a relaxed problem in which the constraints (16) and (17) are dropped. If $M^{SP}(p^*)$ is incentive compatible, then these relaxed constraints are satisfied and, thus, $M^{SP}(p^*)$ solves an original problem.

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

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

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

Now, replace orthogonal $\hat{\Theta}$ by a generic Θ. As long as $\Theta \geq 0$, the mismatching strategies remain irrelevant. As attributes are continuously distributed, in some neighborhood of $\hat{\Theta}$ both sides of the inequalities are continuous and the constraints remain satisfied. Hence, the buyer cannot benefit from misreporting, and the result follows from Proposition 4.

Proof of Proposition 5. Consider any optimal menu and a type θ with $r(\theta) > 0$. Let $r'(\theta) = r(\theta) - \varepsilon$, $p'(\theta) = r(\theta) + \varepsilon / \mathcal{Q}(\theta)$ for a small $\varepsilon > 0$. If incentive compatible, this modification preserves the seller's revenue and the buyer's payoff. The constraint (16) remains the same. The constraint (18) is relaxed and is strictly so when $\mathcal{Q}(\theta) < 1$. The constraint (19) is satisfied by the individual rationality. For the constraint (17):

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

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

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

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

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

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

References

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

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

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

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

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

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

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

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

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

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

Doval, L. and J. C. Ely (forthcoming): "Sequential Information Design," *Econometrica*.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

LI, H. AND X. SHI (2017): "Discriminatory Information Disclosure," *American Economic Review*, 107, 3363–85.

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

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

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

RAYO, L. AND I. SEGAL (2010): "Optimal Information Disclosure," *Journal of Political Economy*, 118, 949–987.

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

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

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

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

WEI, D. AND B. S. GREEN (2019): "(Reverse) Price Discrimination with Information Design," *Discussion paper*.