Discussion Paper Series – CRC TR 224

Discussion Paper No. 352
Project B 03

Optimal Information Design of Online Marketplaces With Return Rights

Jonas von Wangenheim[1]

June 2022

[1] University of Bonn, Email: jwangenheim@uni-bonn.de

Funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) through CRC TR 224 is gratefully acknowledged.

Optimal Information Design of Online Marketplaces with Return Rights*

Jonas von Wangenheim†

May 30, 2022

Abstract

Customer data enables online marketplaces to identify buyers' preferences and provide individualized product information. Buyers learn their product value only after contracting when the product is delivered. I characterize the impact of such ex-ante information on buyer surplus and seller surplus, when the seller sets prices and refund conditions in response to the ex-ante information. I show that efficient trade and an arbitrary split of the surplus can be achieved. For the buyer-optimal signal low-valuation buyers remain partially uninformed. Such a signal induces sellers to sell at low prices without refund options, resulting in commonly observed practices of opaque sales.

JEL classification: D82, D86, D18,

Keywords: information disclosure, sequential screening, information design, strategic learning, Bayesian persuasion, mechanism design, platform economics, consumer protection

*I thank Andreas Asseyer, Dirk Bergemann, Helmut Bester, Giacomo Calzolari, Lucien Frys, Nicolas Schutz, Paul Heidhues, Johannes Johnen, Daniel Krähmer, Matthias Lang, Vincent Meisner, Thomas Schacherer, Roland Strausz, as well as the participants of the 2016 ASSET conference (Thessaloniki), the BIGSEM Doctoral Workshop (Bielefeld), the BGSE Micro Workshop (Bonn), the 2017 EARIE conference (Maastricht), and the 2021 CRC TRR 224 retreat for helpful comments and discussion. Funding by the German Research Foundation (DFG) through CRC TR 224 (Project B03) is gratefully acknowledged.

†University of Bonn, Institute for Microeconomics, Adenauer Allee 24-42, D-53113 Bonn (Germany). Email: jwangenheim@uni-bonn.de

1 Introduction

Over the recent decade, trade has increasingly shifted toward online marketplaces. Commonly in these markets, consumers do not know all the product characteristics at the time of purchase, but must rely on information and recommendations provided by the marketplace. User purchase histories and advances in data technologies increasingly enable marketplaces to identify customers' preferences, and generate customer-specific product information.

For instance, booking.com asks whether customers are business or leisure travelers and emphasizes that search results are fitted to individual needs. Amazon suggests products which have been bought by other customers with a similar purchase history. Amazon Video suggests various movie genres that are based on the personal watch history.

Moreover, marketplaces may design the precision of information. For instance, some marketplaces have customer rating systems which only allow for an overall rating in the form of a chosen number of stars. Others, like Airbnb, allow for individual comments or provide more detailed rating systems by splitting the ratings into different categories like Communication, Location, and Cleanliness.

Naturally, the provided information not only helps consumers to make better informed choices but it also affects the seller's optimal response. A seller may respond not only in setting prices but may also specify refund conditions under which the product can be returned after the consumer has accessed it and fully learned its match value.[1]

In this paper I analyze the interplay between the ex-ante information of a marketplace and the return rights chosen by a monopolistic seller in response. As the most natural benchmark I assume that the marketplace is unrestricted in the design of ex-ante information signals about buyers' match values (Kamenica and Gentzkow (2011)). On the other side, the seller is unrestricted in granting return rights and refund conditions.

The assumption that an information designer may only control ex-ante information, whereas the true value is eventually learned, markedly changes the information design problem compared to information design in standard static monopoly pricing, as in Roesler and Szentes (2017). In particular, the seller may

[1] Alternatively, one may assume that inspection only reveals some additional information, and interpret the buyer's valuation as his updated expected value. As buyers are risk neutral, this leaves all insights of the paper unchanged.

always grant full return rights, effectively restoring classical monopoly pricing under full information. Hence, different to Roesler and Szentes (2017), the standard static monopoly profit always constitutes a lower bound for seller profits.

In this paper I fully classify the possible effects of ex-ante information on efficiency and the division of surplus between the buyer and the seller in such a market with sequential information. As the main result, I show that the above lower bound on seller profit is the main constraint on achievable surplus division. More precisely, similar to Bergemann et al. (2015), I show that the only limits are imposed by the natural constraints that

1. buyer utility is non-negative,
2. the seller receives at least the static monopoly profit, and
3. aggregate surplus does not exceed the first-best gains from trade.

In particular, I illustrate how the marketplace can achieve any point on the Pareto frontier that provides at least the static monopoly profit to the seller. This insight implicitly solves the marketplace's objective to maximize any arbitrarily weighted combination of consumer surplus and producer surplus.

Due to the sequential structure of information, the seller faces a sequential screening problem as introduced in Courty and Li (2000). First, the marketplace designs a signal about the valuation for the good. The seller observes the signal distribution, but not its realization. This assumption expresses that the seller can observe what the buyer learns, but not how it translates into the buyer's valuation. Then, the seller offers a contract, before the buyer learns his true valuation. As studied in Courty and Li (2000), the seller optimally screens buyers by offering a menu of contracts, where contracts differ in prices and refund conditions.[2] Intuitively, buyers with higher valuation uncertainty are more attracted to contracts with a high refund flexibility. From a theoretic perspective, my paper extends the classical sequential screening framework by endogenizing the choice of ex-ante information. More specifically, a third party (the marketplace) chooses the ex-ante signal about the valuation of the good. My analysis allows for arbitrary objectives of the third party that take buyer surplus and seller surplus as

[2]The optimality of sequential screening also features, among others, in Baron and Besanko (1984), Battaglini (2005), Eső and Szentes (2007), Hoffmann and Inderst (2011), Krähmer and Strausz (2011), Nocke et al. (2011), and Pavan et al. (2014).

inputs.[3]

In modeling arbitrary signals my paper follows the vast literature on information design, initiated by Rayo and Segal (2010) and Kamenica and Gentzkow (2011). However, whereas in this literature typically the posterior mean of a signal realization is a sufficient statistic to describe receivers' optimal behavior, such an approach fails in my model. Since buyers as receivers eventually learn their true match value, the entire type distribution for a given signal realization matters when determining the receiver's optimal choice of contract.

For the optimal-design problem it is insightful to consider the benchmark of a marketplace whose interests are fully aligned with the buyer. On the simple intuition that more information cannot hurt the buyer, one might expect that consumer surplus can only be increasing in the amount of buyers' ex-ante private information. This is, however, not the case. Since the seller responds in her contract offer to the structure of buyers' private information, the choice of information exhibits a strategic effect on the subsequent contracting game.

I show that the buyer-optimal ex-ante information keeps the buyer to some extent uninformed about his valuation. The buyer-optimal signal pools low types in a specific way, while high types receive full information. As different low types obtain the same signal, the seller can sell to these types without providing any information rent. Consequently, the seller has an incentive to lower the price below the full information monopoly price, in order to increase participation. Lower prices increase efficiency as well as rents for high types. For the buyer-optimal signal the seller optimally chooses a simple take-it-or-leave it offer without refund at a low price. Trade is efficient, and the seller receives only full-information static monopoly profit.

The strategy that a marketplace deliberately conceals some product characteristics and charges a low but non-refundable price is a commonly observed practice, especially in the travel and hospitality industry. Platforms like priceline.com and hotwire.com typically use such opaque sales to sell residual capacities at heavily discounted prices.[4] Deals may guarantee specific features, such as the hotel star rating or the distance to the city center, but reveal the identity and other details of the hotel only after payment. In the recent academic literature such opaque

[3]Due to the third party's freedom to design information signals, the regularity conditions, imposed in Courty and Li (2000), may be violated. Thus, I cannot rely on their solution techniques to find the optimal contract.

[4]Green and Lomanno (2012) find that in 2010 about one quarter of all hotel bookings in online travel agencies involved opaque goods.

sales have been identified as a revenue-maximizing tool for a multi-product monopolistic seller to increase profits (Balestrieri et al. (2021) and Jiang (2007)) or to solve a mismatch between capacity and demand (Fay and Xie (2008)). My model puts a new perspective on the observed opaque sales. It shows that opaque sales and low prices may naturally arise on platforms that seek to attract buyers by granting high consumer surplus.

2 Related Literature

My paper contributes to the literature on dynamic mechanism design. Baron and Besanko (1984) were the first to study dynamic price discrimination in a two-period procurement model. My model builds on the framework of Courty and Li (2000), who analyze the optimal price discrimination of a monopolist in a two-period model. Battaglini (2005) and Pavan et al. (2014) provide general models for longer time horizons.

A recent branch of the literature, building on Lewis and Sappington (1994), studies sellers' strategic information revelation. In Bergemann and Pesendorfer (2007) a seller can choose the accuracy by which the buyers learn their private valuations in an auction. They identify a trade-off between allocation efficiency and information rents. Eső and Szentes (2007) show that the trade-off disappears when the information provision is part of the contractual relationship. Li and Shi (2017) show that this no longer holds when the seller can use discriminatory information disclosure. Hoffmann and Inderst (2011) characterize optimal contracts for the case where the buyer and seller's information are stochastically independent. Guo et al. (2022) take a complementary approach to mine and analyze seller-optimal information extensions in a sequential screening model for a fixed ex-ante signal. Terstiege and Wasser (2020) show that the buyer-optimal information structure of my model is robust toward additional seller information disclosure in a static environment.

My paper also relates to the literature on buyers' optimal information acquisition. If the marketplace aims to maximize consumer surplus, it provides consumers with costless information before the contractual relationship, to provide a strategic advantage in the contracting game. This timing is in contrast to the classical literature on buyer's information acquisition in principal-agent relationships, where the principal aims to contractually provide incentives for costly learning (e.g., Lewis and Sappington (1997), Crémer et al. (1998), Szalay (2009), Krähmer and Strausz (2011)).

My model is probably most closely related to Roesler and Szentes (2017), who characterize the buyer-optimal signal in a classical, static, one-unit trade environment. In contrast, I analyze internet markets where the buyer receives additional product information after delivery. As a result, the seller may combine the contract with refund options, which—different to Roesler and Szentes—induces a lower bound on seller profit and results in efficient trade for the buyer optimal signal, even for arbitrary production cost. Kessler (1998) was first to analyze the value of ignorance in a classical adverse selection model. She finds that even with costless learning a buyer favors a signal that is uninformative with some positive probability, in order to receive a more favorable contract.

The characterization result of feasible surplus divisions is reminiscent of Bergemann et al. (2015), who study surplus division in third-degree price discrimination environments. In their model the *seller* receives a signal, while the buyer is fully informed. In my model, the seller has to elicit information on the signal via an incentive-compatible mechanism.

The idea that one party can choose arbitrary information signals to influence another party's decision has initiated the literature on Bayesian persuasion (Rayo and Segal (2010), Kamenica and Gentzkow (2011)). My model setup relies on their formulation. However, unlike in the standard persuasion literature, in my model not only the posterior mean but the entire type distribution influences agents' behavior, since they learn their value eventually. This fact calls for different solution techniques.

3 The model

A risk-neutral seller has one unit of a nondivisible good for sale to a risk-neutral buyer. The valuation of the good for the buyer is drawn from a commonly known prior distribution $F(\theta)$ with positive support $[\underline{\theta}, \overline{\theta}]$ and positive, continuous density $f(\theta)$. The seller has a production cost (i.e., reservation value) of $c < \overline{\theta}$. Before contracting and learning the valuation, a third party chooses a signal about the buyer's valuation. The signal distribution is commonly known, the realization is private information to the buyer. I allow for any general signal structure in the form of a Borel-measurable signal space $T \subseteq \mathbb{R}$, together with a probability measure μ on the Borel σ-algebra of $[\underline{\theta}, \overline{\theta}] \times T$. The buyer observes a signal $\tau \in T$, which is distributed according to the signal distribution

$$G(\tau) = \int_{t \leq \tau} \int_{\theta \in [\underline{\theta}, \overline{\theta}]} \mathbb{1}(t, \theta) d\mu.$$

The only restriction on the signal is the "consistency" with the prior F in the sense that

$$\int_{T \times [\underline{\theta}, \theta]} \mathbb{1} d\mu = F(\theta)$$

for all $\theta \in [\underline{\theta}, \overline{\theta}]$.[5]

The setup includes the common examples of a finite signal space $T = \{\tau_1, ..., \tau_n\}$ with the restriction that

$$\sum_{i=1}^{n} F(\theta|\tau_i) Prob(\tau_i) = F(\theta),$$

as well as a continuous signal space $T = [\underline{\tau}, \overline{\tau}]$ with some distribution $G(\tau)$, and the restriction that

$$\int_{[\underline{\tau}, \overline{\tau}]} F(\theta|\tau) dG(\tau) = F(\theta).$$

The timing of the game is as follows:

1. The third party publicly chooses a signal structure.
2. The signal realization is privately observed by the buyer.
3. The seller offers a contract, the buyer accepts/rejects.
4. The buyer observes his type.
5. Transfers are made according to the rules of the contract.

For any signal structure that reveals at least some information to the buyer, the seller in Stage 3 faces a classical sequential screening problem, as described in Courty and Li (2000). They show that any optimal deterministic contract can be implemented as a menu of option contracts from which the buyer can choose at the contracting stage.[6] Hence, throughout this paper we restrict attention to

[5] We explicitly do not make common restrictions on the signal distribution, such as non-shifting support or an order by first-order stochastic dominance.

[6] This is an almost immediate consequence of the revelation principle.

option contracts. An option contract specifies an upfront payment a to the seller, and an option price p, for which the buyer can decide to buy, after he learns his true valuation. Equivalently, one can interpret such a contract as a buy price of $a + p$, together with the option to return the good for a refund of p.

We deliberately do not specify the objective function of the third party here, as we want to allow for various objectives. In the following section I analyze for a uniform prior the case where the third party's interest is fully aligned with the buyer. This case is interesting for multiple reasons. First, it provides an interesting theoretic benchmark, which illustrates that in our environment more consumer information can be detrimental to consumer surplus. Second, it may constitute a good approximation for heavily contested online platforms where consumers can switch to other platforms at virtually no cost. Third the buyer-optimal information structure provides important insights for consumer protection regulation.

In Section 5, I allow for arbitrary objectives with respect to buyer surplus and seller surplus. I show how the signal can be refined for arbitrary priors in order to induce mainly any arbitrary surplus pair.

4 The Buyer Optimal Signal – Uniform Case

It is instructive to analyze first the buyer-optimal signal for a uniform prior, as it catches the main economic intuitions. In Section 5, I show how the construction generalizes to arbitrary prior distributions.

Let the prior $F(\theta)$ be uniformly distributed on $[0, 1]$, and let $c = 0$. Consider, as a benchmark, that the buyer fully learns his type θ under signal τ. The seller will then charge the monopoly price of

$$p^M = \arg\max_p p(1 - F(p)) = 1/2.$$

She will therefore sell to the buyer if and only if the buyer's valuation exceeds $1/2$. The seller's profit is $\pi^M = 1/4$, while the buyer's expected surplus is $1/8$.

Note that the seller can always ignore the possibility to exploit the signal for ex-ante screening and just charge the monopoly price after the buyer learns the true valuation, i.e., $(a, p) = (0, p^M)$. Hence, the static monopoly profit of $\pi^M = 1/4$ defines a lower bound for the seller's utility.

Since for $c = 0$ trade is always efficient, the upper bound for buyer surplus is achieved, if trade always occurs, and the seller is left with her full-information monopoly profit π^M. The main insight of this section is that such a contract can be induced by the following signal:

$$\tau(\theta) = \begin{cases} 0 & \theta \leq \frac{1}{2} \\ \theta & \theta > \frac{1}{2}. \end{cases} \tag{1}$$

The buyer only learns his valuation if it is above 1/2. Buyers with a valuation below 1/2 are pooled in one signal of $\tau = 0$, which induces an expected valuation of $\mathbb{E}[\theta|\tau = 0] = 1/4$.

Suppose the seller offers a single contract $(a, p) = (1/4, 0)$, which means she offers the good at a price of 1/4 before the buyer learns θ with certainty. Since $\mathbb{E}[\theta|\tau] \geq 1/4$ for all τ, this offer will attract all buyers. I show in the appendix that, given this signal structure, there is no contract that generates a higher seller utility.

Proposition 1. *Given signal τ, there is no mechanism which generates a seller profit above $\frac{1}{4}$. In particular, the contract $(\frac{1}{4}, 0)$, which sells to all buyers ex ante at a price of $\frac{1}{4}$, is a seller-optimal trading mechanism.*

Since the seller is left with her lower bound utility of 1/4, and social surplus is maximized, the signal τ implements the upper bound of buyer utility. It is therefore a buyer-optimal signal.

Even though the above construction of the optimal signal is specific to the uniform distribution, the main intuitions from this example carry over to the general case. It is suboptimal for the buyer to be fully informed about his valuation. If buyers with relatively low valuations remain partly uninformed, then the seller has to provide less information rent to sell to these types. To include lower types in trade, the seller must set low prices for *all* buyers. While low types make zero profits in expectation, high types benefit from lower prices and buyer surplus increases. Since more types trade, efficiency increases as well.

Example: Opaque Sales in the Hotel Industry

When customers search for a hotel in a specific area, the webpage hotwire.com offers "opaque deals": a hotel booking is guaranteed at a certain fixed price in a predefined area. The specific name and all details of the hotel are only revealed after a non-refundable purchase. Prior to purchase, the webpage only offers coarse information, such as star rating and specific guaranteed amenities.

Star classification systems certify certain quality levels, typically represented by one star (basic) to five stars (luxury). While there exist various competing classification systems, the idea is commonly the same. Hotels gather points by providing various services and amenities. Hotels that want to certify a certain number of stars must achieve a respective number of points. Hence, whereas the number of stars may be a measure of overall quality, it leaves plenty of space for horizontal differentiation.[7]

Certainly, different customers have different needs for amenities and services. Suppose for the sake of this example that potential guests g are with equal probability either of type *business* ($g = B$), or of type *leisure* ($g = L$). Suppose that business and leisure travelers value hotel features in mostly different categories. Business travelers may value late check-in hours, good reception service, quiet rooms, and early breakfasts, whereas leisure travelers may value a large pool, sports and leisure facilities, and gastronomic services. Suppose hotels focus on one traveler type and provide high-quality service and amenities predominantly in the category suitable for their focus type. Additionally, within each of the two categories there is room for horizontal differentiation, as the hotels can choose which specific services to provide. We can think of this differentiation as two locations on two distinct circumferences of unit circles (Salop's circles). The position on the first circle represents the hotel's (horizontal) choice of which services to provide for business travellers, the position on the second circle represents the respective choice for leisure-traveller services. This again reflects the idea that the number of stars provides a measure for the total service quality, but the hotel is to a large degree free to choose which exact services to provide.

Hence, each traveler's preference, and each hotel's attribute can be represented as an element from the space $\{B, L\} \times S^1 \times S^1$, where S^1 describes the circumference of the unit circle. Suppose that hotels' and guests' locations on each circumference are independently and uniformly distributed. Now let the match value of a traveler of type (g, x_B, x_L) with hotel of type (h, y_B, y_L) be

[7] One of the most prominent classification systems is provided by the "Hotelstar Union" in Europe, which aims to harmonize the national standards of hotel certifications. By 2021 the system had been adapted by 19 European countries. Participating hotels gather points by providing features from a list of over 200 possible criteria. Besides some minimum requirements for each star, hotels are entirely free in how to achieve the number of points.

given by

$$\theta((g, x_B, x_L), (h, y_B, y_L), p) = 0.5 \cdot \mathbb{1}_{g=h} + (0.5 - d(x_g, y_g)),$$

where $d(\cdot,\cdot)$ describes the distance on the respective circle. In other words, a traveler receives a utility of 0.5 if the hotel has a focus suitable for his type, and an additional utility up to 0.5 if the hotel offers the most preferred features *within* the relevant focus categories. Note that without any information on the hotel, the match value for each traveler is ex ante uniformly distributed on $[0, 1]$, with values in $[0, 0.5]$ for types $g \neq h$ and values in $[0.5, 1]$ for types $g = h$.

Now, suppose that via the selection of displayed guaranteed amenities hotwire reveals the type (business/leisure) of the hotel, as well as the exact horizontal differentiation within that category only. This information structure corresponds exactly to the buyer-optimal information discussed above: types with a match value above 0.5 fully learn their match value, whereas types with a match value below 0.5 remain pooled. Hence, such an information structure achieves two things. First, the hotel agrees with hotwire's policy to sell the capacities at low rates without refund, since for the given information structure this is the profit-maximizing pricing strategy. Second, hotwire can offer highly competitive deals, since these offers maximize buyer surplus.

Certainly, this example was stylized to match with the formerly discussed case of a uniformly distributed prior. Yet, the main insight generalizes: detailed information only on the strong points of a differentiated product may enable meaningful inference only for consumers with a high valuation of these points. Such information makes it hard for a seller to pool and extract information rent from high types. As low types remain pooled it is appealing for the seller to set low prices and serve high demand.

5 The Limits of Surplus Distribution

In the previous section, I derived a signal which maximizes buyer surplus for a uniform prior and no production costs. In this section, I fully characterize which combinations of buyer surplus and seller surplus are feasible for arbitrary signals and production costs $c \in [0, \overline{\theta}]$. Let us first characterize the natural constraints to this problem graphically.

First, by buyer's individual rationality, the expected buyer surplus must be

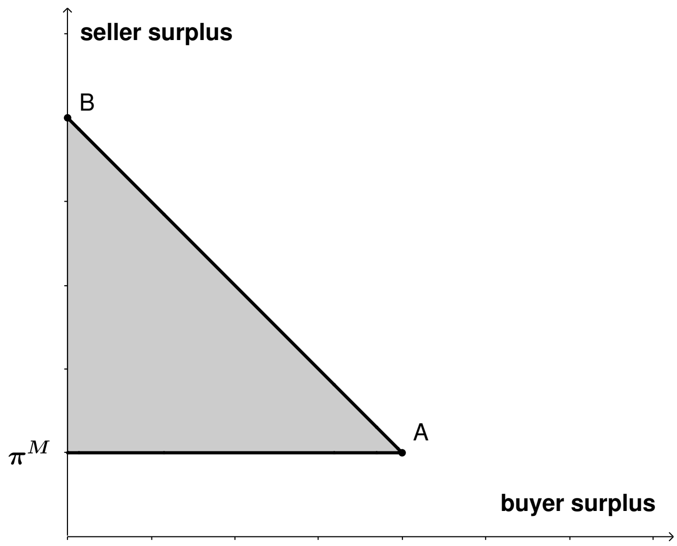

Figure 1: All potential pairs of surplus division

non-negative. Second, as argued in the previous section, the seller surplus can never fall below the static monopoly profit under full buyer information, since the seller can always use a full-refund mechanism. Finally, aggregate surplus cannot exceed first-best welfare, which is sketched as the diagonal Pareto frontier.

Consequently, any surplus pair must lie in the gray-shaded triangle. Point A corresponds to the buyer-optimal signal. Point B corresponds to the case where the buyer has no ex-ante information upon the prior distribution. In this case, the seller can extract the entire surplus by selling ex ante at a price of $\mathbb{E}[\theta]$.

The following theorem states that the above are the only constraints, and any arbitrary surplus pair in the triangle can be implemented.

Theorem 1. *There exists a signal and an optimal sequential selling mechanism with seller surplus u_S and buyer surplus u_B if and only if*

- $u_B \geq 0$,
- $u_S \geq \pi^M$, *and*
- $u_S + u_B \leq \int_c^{\overline{\theta}} (\theta - c) f(\theta) d\theta$,

where π^M is the standard static monopoly profit the seller can achieve, if the buyer has full information.

Any such surplus pair can be achieved by an optimal contract which specifies a take-it-or-leave it offer without refund.

A full proof can be found in the appendix. I will sketch the main steps of the construction here. Take an arbitrary surplus pair (u_B, u_S) which satisfies the above constraints. I will construct a corresponding signal that induces this surplus pair.

Define the threshold $x \geq c$ by

$$u_S + u_B = \int_x^{\overline{\theta}} (\theta - c) f(\theta) d\theta. \tag{2}$$

Note that welfare is indeed $u_S + u_B$, if we can construct a signal for which exactly all types above x buy.
Next, define the threshold $y \in [x, \overline{\theta}]$ by

$$u_S = (1 - F(x))\big(\mathbb{E}[\theta|\theta \in [x, y]] - c\big).$$

Furthermore, define $\overline{a} \in [x, y]$ by

$$\overline{a} := \mathbb{E}[\theta|\theta \in [x, y]].$$

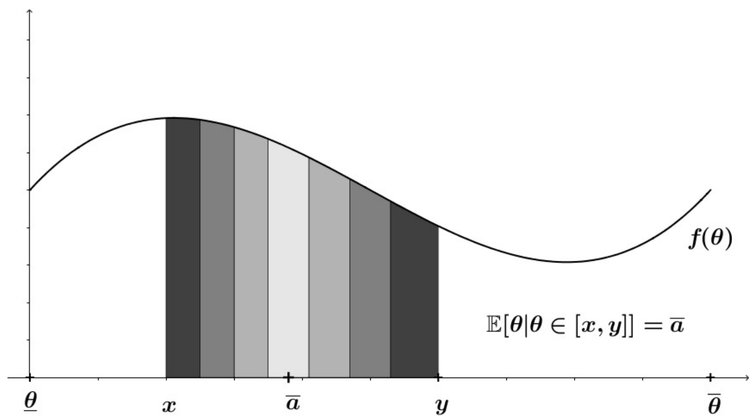

Figure 2: The signal to induce (u_B, u_S)

Note that seller surplus is indeed u_S if the seller successfully sells to all types $\theta \geq x$ at a price of $\overline{a}$.

Types outside $[x, y]$ fully learn their valuation, whereas types in $[x, y]$ learn that their type is in a certain pooling region, represented by the shade of gray assigned to their type, as depicted in Figure 2. The shaded areas are constructed in such a way that for any shade τ

$$\mathbb{E}[\theta|\tau] = \overline{a}.$$

Moreover, if τ_1 is darker than τ_0, then $F(\cdot|\tau_1)$ is a mean preserving spread of $F(\cdot|\tau_0)$.

If we let the number of different shades go to infinity, we obtain a continuum of shades. In the limit, each signal τ only pools two types $\{\theta_\tau^L, \theta_\tau^H\}$ with $\theta_\tau^L < \pi^M < \theta_\tau^H$. The signal structure can be represented by

$$\tau(\theta) = \begin{cases} \theta - \overline{\theta} & \theta < x \\ \int_\theta^{\overline{a}} f(s)(\overline{a} - s)ds & \theta \in [x, y] \\ \theta & \theta > y. \end{cases}$$

Compared to a signal with complete pooling in $[x, y]$ (which would be reminiscent of the construction for the uniform case), this refined signal structure achieves two things. First, it provides types in $[x, y]$ with maximal information by maintaining the condition that the ex-ante expected valuation remains $\overline{a}$. Note that under signal τ—in contrast to complete pooling on $[x, y]$—any contract (a, p) with $p \in [x, \overline{a}]$ and $a + p > \overline{a}$ would be rejected by all types with a lighter shade than type $\theta = p$. This insight enables me to prove that under τ no type $\theta \in [x, \overline{a}]$ will ever end up buying if $a + p > \overline{a}$.

Second, the anti-assortative pairing method in $[x, y]$ achieves an ordering on the participation constraints: If a contract (a,p) is profitable for some signal realization, it is a fortiori profitable for any signal realization of darker shade.[8] Consequently, any type $\theta \in [\overline{a}, y]$ who ends up buying, will buy for the same total price $a + p$. The inability to screen different types in $[\overline{a}, y]$, makes serving only buyers with $\theta > \overline{a}$ at a uniform price unattractive. Hence, selling to all buyers in $[x, \overline{\theta}]$ at a total price of $a + p = \overline{a}$ remains the best option.

Intuitively, the boundaries x and y partition the type space. Types in $[c, x]$ do not trade and induce an efficiency loss. Hence, the location of x determines welfare. Buyer types in $[x, y]$ don't receive any surplus. Hence, the location of y determines the distribution of surplus. By shifting the two boundaries one can realize any distribution of surplus that satisfies the natural constraints in

[8]I prove this formally in Lemma 1 in the appendix.

Theorem 1.

6 Conclusion

This paper emphasizes the important role of ex-ante information in a trade environment. Even if a buyer eventually learns his true valuation and the seller is free to guarantee return rights, there are almost no constraints to the division of buyer surplus and seller surplus that may result for different ex-ante information. One important insight is that more precise ex-ante information does not always benefit the buyer. Less information, in particular to buyers with low valuations, may lead to lower prices and an increase in consumer surplus. Under this light, any consumer protection policy for mandatory information disclosure should be regarded with care, as the overall effect on consumer utility may be ambiguous.

Similar considerations apply to mandatory return rights. Under Directive 2011/83/EU, the European Union grants any consumer the right to withdraw from online contracts within 14 days after delivery. As Krähmer and Strausz (2015) already point out, this policy effectively destroys the ability of a seller to screen ex ante, and leaves the consumers with the same information rent as under full information. Hence, such a policy may in particular restrict opaque sales, which, again, may be detrimental for consumer surplus.

With rapid advances in data analyses, large marketplaces will likely become even more sophisticated in the future in providing customers with targeted information. Their information design may have a significant impact on consumer welfare. Rather than regulating information or return rights, it may become a more effective consumer protection policy to find ways to align marketplace and consumer interests.

7 Appendix

Proof of Proposition 1. Evidently, the contract $(\frac{1}{4}, 0)$ generates a seller profit of $\frac{1}{4}$. Hence, it remains to show that an arbitrary menu $\mathcal{M} = \{(a_i, p_i)\}_{i \in I}$ of option contracts can generate a seller profit of at most $\frac{1}{4}$.

Take an arbitrary menu. If the low types with $\tau(\theta) = 0$ reject all contracts from the menu, then only types $\theta > \frac{1}{2}$ with full information may trade, and seller profit is bounded by the static full-information monopoly profit of $\pi^M = \frac{1}{4}$.

Suppose, contrary, that the low types with $\tau(\theta) = 0$ choose some contract (a_0, p_0) from the menu. Any buyer who picks this option contract chooses to buy ex post at the option price $p_0 \in \left[0, \frac{1}{2}\right]$ if and only if $\theta > p_0$. Since the expected utility of low types with $\tau(\theta) = 0$ must be non-negative this implies that

$$a_0 \leq \frac{\int_{p_0}^{\frac{1}{2}} (\theta - p_0) d\theta}{\frac{1}{2}} = 2\left[\frac{\theta^2}{2} - p_0\theta\right]_{p_0}^{\frac{1}{2}} = p_0^2 - p_0 + \frac{1}{4}.$$

Since the contract (a_0, p_0) is available to all buyers, no buyer type $\theta > \frac{1}{2}$ will pay more than a total price of $a_0 + p_0$. Hence, total profits are bounded by

$$a_0 + \int_{p_0}^{1} p_0 d\theta \leq \left(p_0^2 - p_0 + \frac{1}{4}\right) + \left(p_0 - p_0^2\right) = \frac{1}{4}.$$

□

Proof of Theorem 1. Take some arbitrary $u_S \geq \pi^M$ and $u_B \geq 0$, with $u_B + u_S \leq \int_c^{\overline{\theta}} (\theta - c) f(\theta) d\theta$. We need to construct a signal such that the seller's optimal mechanism induces seller utility u_S and buyer utility u_B.

Constructing the signal

Define $x \in [c, \overline{\theta}]$ implicitly by

$$u_S + u_B = \int_x^{\overline{\theta}} (\theta - c) f(\theta) d\theta = (1 - F(x))\mathbb{E}\left[(\theta - c)|\theta \in [x, \overline{\theta}]\right]. \tag{3}$$

Since f has full support, the right-hand side in (3) is strictly decreasing in x, from first-best surplus for $x = c$ to 0 for $x = \overline{\theta}$. Hence, there is indeed a unique $x \in [\underline{\theta}, \overline{\theta}]$, for which (3) is satisfied.[9] Define now y implicitly by

$$u_S = (1 - F(x))\mathbb{E}[(\theta - c)|\theta \in [x, y]]. \tag{4}$$

Note that the right-hand side in (4) is strictly increasing in y from

$$(1 - F(x))(x - c) \leq \pi^M \leq u_S$$

[9]The assumption that F is continuous and increasing is innocuous and only for mathematical convenience. If F has atoms, then $\tau(\theta)$ is not deterministic. If F is not increasing, we lose the uniqueness of x and y. None of the results or intuitions hinge on these assumptions.

for $y = x$ to

$$(1 - F(x))\mathbb{E}[(\theta - c)|\theta \in [x, \overline{\theta}]] = u_S + u_B \geq u_S$$

for $y = \overline{\theta}$. Hence, there is indeed a unique $y \in [x, \overline{\theta}]$ which satisfies (4). Further, define

$$\overline{a} := \mathbb{E}[\theta|\theta \in [x, y]].$$

Finally, define the following signal structure:

$$\tau(\theta) = \begin{cases} \theta - \overline{\theta} & \theta < x \\ \int_{\theta}^{\overline{a}} f(s)(\overline{a} - s)ds & \theta \in [x, y] \\ \theta & \theta > y. \end{cases}$$

As displayed in Figure 3, the signal prescribes full learning for $\theta < x$ and $\theta > y$. For $\theta \in [x, y]$ the function $\tau(\theta)$ is continuous and strictly decreasing on $[x, \overline{a}]$, and strictly increasing on $[\overline{a}, y]$, with

$$\begin{aligned} \tau(x) &= \int_x^{\overline{a}} f(s)(\overline{a} - s)ds \\ &= \int_x^{y} f(s)(\overline{a} - s)ds + \int_y^{\overline{a}} f(s)(\overline{a} - s)ds \\ &= (F(y) - F(x)) \underbrace{\left(\overline{a} - \frac{\int_x^y f(s)sds}{F(y) - F(x)}\right)}_{=0} + \int_y^{\overline{a}} f(s)(\overline{a} - s)ds \\ &= \tau(y). \end{aligned}$$

Thus, for any τ with $0 < \tau \leq \tau(x)$ there are exactly two types $\theta_\tau^L, \theta_\tau^H$ with $\tau = \tau(\theta_\tau^L) = \tau(\theta_\tau^H)$, where without loss of generality $\theta_\tau^L < \overline{a} < \theta_\tau^H$. Let us call $\theta^L(\tau)$ the inverse function of $\tau(\theta)$ on $[x, \overline{a}]$, and $\theta^H(\tau)$ the inverse function of $\tau(\theta)$

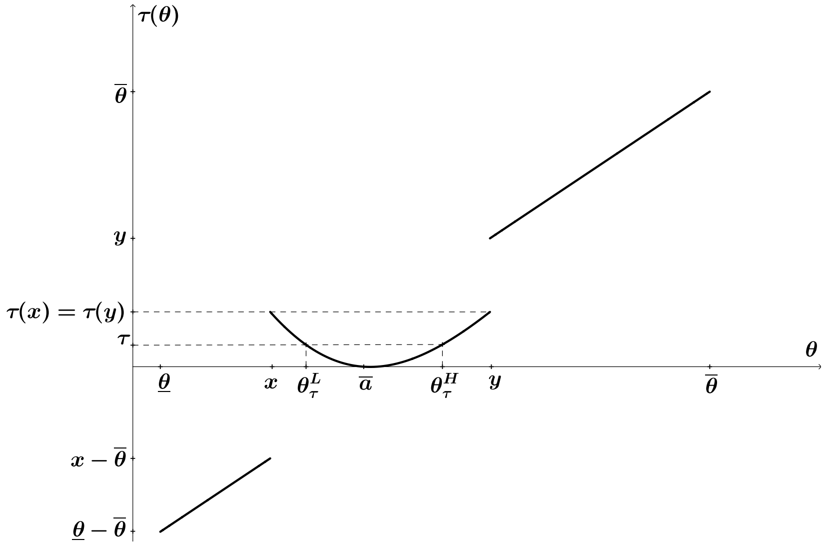

Figure 3: The signal structure $\tau(\theta)$

on $[\overline{a}, y]$. Then, for all $z \in (0, \tau(x)]$

$$
\begin{aligned}
0 &= \tau(\theta_z^L) - \tau(\theta_z^H) \\
&= \int_{\theta_z^L}^{\overline{a}} f(s)(\overline{a}-s)ds - \int_{\theta_z^H}^{\overline{a}} f(s)(\overline{a}-s)ds \\
&= \int_{\theta_z^L}^{\overline{a}} f(s)(\overline{a}-s)ds + \int_{\overline{a}}^{\theta_z^H} f(s)(\overline{a}-s)ds \\
&= -\int_{\theta_z^L}^{\theta_z^H} sf(s)ds + \overline{a}\cdot \mathrm{Prob}(\theta \in [\theta_z^L, \theta_z^H]).
\end{aligned}
$$

Using this result and the definition of the condition expectation we obtain for all $z \in (0, \tau(x)]$

$$
\int_{\{\tau \le z\}} \mathbb{E}[\theta|\tau] = \int_{\theta_z^L}^{\theta_z^H} sf(s)ds = \overline{a}\cdot \mathrm{Prob}(\theta \in [\theta_z^L, \theta_z^H]) = \int_{\{\tau \le z\}} \overline{a}.
$$

Since the intervals $(0, z]$ are generating the respective Borel-Algebra on $[0, \tau(x)]$, this implies for $\tau \in [0, \tau(x)]$

$$
\mathbb{E}[\theta|\tau] = \overline{a} \tag{5}
$$

almost surely.[10] Hence, for any $\tau_1, \tau_2 \in [0, \tau(x)]$ with $\tau_1 < \tau_2$ the distribution $F(\cdot|\tau_2)$ is a mean-preserving spread of $F(\cdot|\tau_1)$.[11]
For the resulting regular conditional probabilities we obtain

$$\mathbb{P}(\theta_\tau^H|\tau) = \frac{\overline{a} - \theta_\tau^L}{\theta_\tau^H - \theta_\tau^L} \qquad \text{and} \qquad \mathbb{P}(\theta_\tau^L|\tau) = \frac{\theta_\tau^H - \overline{a}}{\theta_\tau^H - \theta_\tau^L},$$

as these are the unique weights that simultaneously satisfy $\mathbb{P}(\theta_\tau^H|\tau) + \mathbb{P}(\theta_\tau^L|\tau) = 1$ and

$$\mathbb{E}[\theta|\tau] = \mathbb{P}(\theta_\tau^H|\tau)\theta_\tau^H + \mathbb{P}(\theta_\tau^L|\tau)\theta_\tau^L = \overline{a}.$$

The menu

We turn to the seller's decision problem to choose an optimal menu of option contracts, given signal τ. Consider the menu $\mathcal{M} = \{(\overline{a}, 0)\}$. All buyers with $\theta < x$ receive a fully informative signal $\tau < 0$, and know with certainty that their valuation satisfies $\theta < \overline{a}$, so they would reject the contract. Types $0 \leq \tau \leq \tau(x)$ satisfy $\mathbb{E}[\theta|\tau] = \overline{a}$, and types $\tau > \tau(x)$ satisfy $\mathbb{E}[\theta|\tau] = \tau > \overline{a}$, so they would both accept the contract $(\overline{a}, 0)$, which sells ex ante at a uniform price of $\overline{a}$. This means that under menu $\mathcal{M}$ we indeed have a seller utility of

$$(\overline{a} - c)(1 - F(x)) = u_S,$$

and buyer surplus

$$\int_x^{\overline{\theta}} (\theta - c) f(\theta) d\theta - u_S = (u_B + u_S) - u_S = u_B.$$

This shows that the menu $\mathcal{M}$ indeed implements the buyer and seller utility we want to construct. It remains to show, that $\mathcal{M}$ is an optimal menu for the seller for the given signal τ.

The optimality of the menu

[10]As usual, the conditional expectation and the following regular conditional probability are uniquely defined only almost surely. Since we are interested in the division of expected surplus, this restriction is irrelevant.

[11]Note, however, that the common assumption in Courty and Li (2000) of "non-shifting support" is violated. Thus, we cannot use their standard procedure to solve the seller's maximization problem.

Let $\tilde{\mathcal{M}} = \{(a_i, p_i)\}_{i \in I}$ be an arbitrary menu of option contracts. Denote with $\tilde{u}_B$ and $\tilde{u}_S$ the surplus pair resulting from $\tilde{M}$. We need to show that $\tilde{u}_S \leq u_S$, such that $\mathcal{M}$ is optimal.
Let $\hat{\theta}$ be the lowest type who purchases the good under $\tilde{\mathcal{M}}$, in the sense that he chooses some $(a, p) \in \tilde{\mathcal{M}}$ to pay the upfront fee a, and decides to buy the good at the price p, after he learns his type.
Case 1: $\hat{\theta} < x$ or $\hat{\theta} > y$
In this case $\hat{\theta}$ learns his type with certainty under τ. Since, by assumption, he accepts the contract (a, p), we can conclude that

$$a + p \leq \hat{\theta}.$$

Furthermore, any buyer's signal $\tau(\theta)$ reveals to the buyer with certainty whether his type satisfies $\theta > \hat{\theta}$. This means, that any buyer with $\theta > \hat{\theta}$ learns from his signal realization that he will receive a positive utility from contract (a, p). Consequently, no type $\theta > \hat{\theta}$ will accept a contract at total cost higher than $a+p$. Since $\hat{\theta}$ is by assumption the lowest type that buys, we can conclude that

$$\tilde{u}_S \leq (a+p-c)(1-F(\hat{\theta})) \leq (\hat{\theta}-c)(1-F(\hat{\theta}) \leq \max_p \{(1-F(p))(p-c)\} = \pi^M \leq u_S.$$

Case 2: $\hat{\theta} \in [x, \overline{a}]$
Then $\hat{\theta}$ is the low type for the respective signal realization, i.e., $\hat{\theta} = \theta^L_{\tau(\hat{\theta})} < \theta^H_{\tau(\hat{\theta})}$. Thus, since type $\theta^L_{\tau(\hat{\theta})}$ purchases the good under (a, p), so will type $\theta^H_{\tau(\hat{\theta})}$. Under the buyer's ex-ante individual rationality we have

$$a + p \leq \mathbb{E}[\theta | \tau(\hat{\theta})] = \overline{a}.$$

The contract (a, p) is therefore, in particular, also profitable to all types $\theta > y$, who learn their valuation ex ante with certainty. Hence, any of these types will also pay at most $a + p \leq \overline{a}$. Thus, even if the seller extracts all surplus from types $\theta \in [\hat{\theta}, y]$, her surplus is bounded by

$$\begin{aligned}
\tilde{u}_S &\leq \int_{\hat{\theta}}^{y} (\theta - c) dF(\theta) + (1 - F(y))(\overline{a} - c) \\
&\leq \int_{x}^{y} (\theta - c) dF(\theta) + (1 - F(y))(\overline{a} - c) \\
&= (F(y) - F(x))(\overline{a} - c) + (1 - F(y))(\overline{a} - c) \\
&= (1 - F(x))(\overline{a} - c) \\
&= u_S
\end{aligned}$$

Case 3: $\hat{\theta} \in [\overline{a}, y]$
Then $\hat{\theta}$ is the high type for the respective signal realization, i.e., $\hat{\theta} = \theta^H_{\tau(\hat{\theta})}$. Moreover, we have $\theta^H_{\tau(\hat{\theta})} \geq p > \theta^L_{\tau(\hat{\theta})}$, because otherwise $\theta^L_{\tau(\hat{\theta})}$ would purchase the good for p whenever $\theta^H_{\tau(\hat{\theta})}$ does, violating that $\theta^H_{\tau(\hat{\theta})}$ is the lowest type who purchases the good. Lemma 1 shows that since the ex-ante participation constraint is satisfied for $\tau(\hat{\theta})$, it cannot bind for any higher $\tau \in [\tau(\hat{\theta}), \tau(y)]$.

Lemma 1. *If for signal realizations $0 \leq \tau_1 < \tau_2 \leq \tau(y)$ and some contract (a, p) with $p > \theta^L_{\tau_1}$ we have*

$$-a + \mathbb{P}(\theta^H_{\tau_1} | \tau_1)(\theta^H_{\tau_1} - p) \geq 0, \tag{IR τ_1}$$

then

$$-a + \mathbb{P}(\theta^H_{\tau_2} | \tau_2)(\theta^H_{\tau_2} - p) > 0. \tag{IR τ_2}$$

Proof of Lemma 1. Call $\alpha_1 := \mathbb{P}(\theta^H_{\tau_1} | \tau_1)$ and $\alpha_2 := \mathbb{P}(\theta^H_{\tau_2} | \tau_2)$.
We thus need to show that

$$\alpha_1(\theta^H_{\tau_1} - p) < \alpha_2(\theta^H_{\tau_2} - p).$$

If $\alpha_2 > \alpha_1$ this is immediate, since $\theta^H_{\tau_2} > \theta^H_{\tau_1}$. Assume therefore in the following that $\alpha_2 \leq \alpha_1$.
Equation (5) can be rewritten as

$$(1 - \alpha_1)\theta^L_{\tau_1} + \alpha_1 \theta^H_{\tau_1} = \overline{a},$$

or respectively

$$(1 - \alpha_2)\theta^L_{\tau_2} + \alpha_2 \theta^H_{\tau_2} = \overline{a}.$$

It follows that

$$\alpha_1(\theta^H_{\tau_1} - \theta^L_{\tau_1}) = \overline{a} - \theta^L_{\tau_1} = (\overline{a} - \theta^L_{\tau_2}) + (\theta^L_{\tau_2} - \theta^L_{\tau_1}) = \alpha_2(\theta^H_{\tau_2} - \theta^L_{\tau_2}) + (\theta^L_{\tau_2} - \theta^L_{\tau_1}).$$

Now, since $\theta^L_{\tau_2} < \theta^L_{\tau_1} < p$ and $\alpha_2 \leq \alpha_1 < 1$, we have

$$\begin{aligned}
\alpha_1(\theta^H_{\tau_1} - p) &= \alpha_1(\theta^H_{\tau_1} - \theta^L_{\tau_1}) + \alpha_1(\theta^L_{\tau_1} - p) \\
&= \alpha_2(\theta^H_{\tau_2} - \theta^L_{\tau_2}) + (\theta^L_{\tau_2} - \theta^L_{\tau_1}) + \alpha_1(\theta^L_{\tau_1} - p) \\
&< \alpha_2(\theta^H_{\tau_2} - \theta^L_{\tau_2}) + \alpha_2(\theta^L_{\tau_2} - \theta^L_{\tau_1}) + \alpha_2(\theta^L_{\tau_1} - p) \\
&= \alpha_2(\theta^H_{\tau_2} - p).
\end{aligned}$$

□

Furthermore, any type $\theta > y$, who learns his type with certainty, obtains a utility of

$$-a + (\theta - p) > -a + (\hat{\theta} - p) > -a + \mathbb{P}(\hat{\theta}|\tau(\hat{\theta}))(\hat{\theta} - p) \geq 0$$

from contract (a, p). The contract thus generates a positive expected utility to all $\tau > \tau(\hat{\theta})$, and positive utility to all types $\theta > \hat{\theta}$. This means that the contract (a, p) alone induces all types $\theta \geq \hat{\theta}$ to purchase the good. Since, by assumption, $\hat{\theta}$ is the lowest type who purchases the good for menu $\tilde{\mathcal{M}}$, any additional contract in the menu does not increase trade efficiency. It could therefore only decrease seller utility, since a buyer would only take it if it yielded higher rents to him than the contract (a, p), and thus lower rents to the seller. Therefore, if $\tilde{\mathcal{M}}$ is an optimal menu, we can assume $\tilde{\mathcal{M}} = \{(a, p)\}$, and seller utility is given by

$$\begin{aligned}
\tilde{u}_S &= \text{Prob}(\tau > \tau(\hat{\theta}))a + (1 - F(\hat{\theta}))(p - c) \\
&= (1 - F(\hat{\theta}) + F(\theta^L_{\tau(\hat{\theta})}) - F(x))a + (1 - F(\hat{\theta}))(p - c).
\end{aligned}$$

Since, according to ex-ante IR we have $a \leq \mathbb{P}(\hat{\theta}|\tau(\hat{\theta}))(\hat{\theta} - p)$, it follows that

$$\tilde{u}_S \leq \big(1 - F(\hat{\theta}) + F(\theta^L_{\tau(\hat{\theta})}) - F(x)\big)\mathbb{P}(\hat{\theta}|\tau(\hat{\theta}))(\hat{\theta} - p) + (1 - F(\hat{\theta}))(p - c).$$

Recall that $0 \leq \theta^L_{\tau(\hat{\theta})} < p \leq \hat{\theta}$, since $\hat{\theta}$ is the lowest type who buys. If

$$\big(1 - F(\hat{\theta}) + F(\theta^L_{\tau(\hat{\theta})}) - F(x)\big)\mathbb{P}(\hat{\theta}|\tau(\hat{\theta})) > 1 - F(\hat{\theta}),$$

then

$$\begin{aligned}
\tilde{u}_S &\leq \big(1 - F(\hat{\theta}) + F(\theta^L_{\tau(\hat{\theta})}) - F(x)\big)\mathbb{P}(\hat{\theta}|\tau(\hat{\theta}))(\hat{\theta} - c) \\
&\leq (1 - F(x))\mathbb{P}(\hat{\theta}|\tau(\hat{\theta}))(\hat{\theta} - c) \\
&\leq (1 - F(x))\big(\mathbb{P}(\theta^H_{\tau(\hat{\theta})}|\tau(\hat{\theta}))(\hat{\theta} - c) + \mathbb{P}(\theta^L_{\tau(\hat{\theta})}|\tau(\hat{\theta}))(\theta^L_{\tau(\hat{\theta})} - c)\big) \\
&= (1 - F(x))(\overline{a} - c) \\
&= u_S.
\end{aligned}$$

Alternatively, if

$$\big(1 - F(\hat{\theta}) + F(\theta^L_{\tau(\hat{\theta})}) - F(x)\big)\mathbb{P}(\hat{\theta}|\tau(\hat{\theta})) \leq 1 - F(\hat{\theta}),$$

then

$$\begin{aligned}
\tilde{u}_S &\leq \big(1 - F(\hat{\theta}) + F(\theta^L_{\tau(\hat{\theta})}) - F(x)\big)\mathbb{P}(\hat{\theta}|\tau(\hat{\theta}))(\hat{\theta} - p) + (1 - F(\hat{\theta}))(p - c) \\
&= (1 - F(\hat{\theta}))(\hat{\theta} - p) + (1 - F(\hat{\theta}))(p - c) \\
&= (1 - F(\hat{\theta}))(\hat{\theta} - c) \\
&\leq \max_p (1 - F(p))(p - c) \\
&= \pi^M \\
&\leq u_S
\end{aligned}$$

This concludes the proof that there is no menu $\tilde{\mathcal{M}}$ which yields the seller a surplus above u_S. Consequently, $\mathcal{M}$ is a seller-optimal contract. $\square$

References

Balestrieri, F., Izmalkov, S., and Leao, J. (2021). The market for surprises: selling substitute goods through lotteries. *Journal of the European Economic Association*, 19(1):509–535.

Baron, D. P. and Besanko, D. (1984). Regulation and information in a continuing relationship. *Information Economics and policy*, 1(3):267–302.

Battaglini, M. (2005). Long-term contracting with markovian consumers. *The American Economic Review*, 95(3):637–658.

Bergemann, D., Brooks, B., and Morris, S. (2015). The limits of price discrimination. *The American Economic Review*, 105(3):921–957.

Bergemann, D. and Pesendorfer, M. (2007). Information structures in optimal auctions. *Journal of Economic Theory*, 137(1):580–609.

Courty, P. and Li, H. (2000). Sequential screening. *The Review of Economic Studies*, 67(4):697–717.

Crémer, J., Khalil, F., and Rochet, J.-C. (1998). Contracts and productive information gathering. *Games and Economic Behavior*, 25(2):174–193.

Eső, P. and Szentes, B. (2007). Optimal information disclosure in auctions and the handicap auction. *The Review of Economic Studies*, 74(3):705–731.

Fay, S. and Xie, J. (2008). Probabilistic goods: A creative way of selling products and services. *Marketing Science*, 27(4):674–690.

Green, C. E. and Lomanno, M. V. (2012). *Distribution channel analysis: A guide for hotels*. HSMAI Foundation.

Guo, Y., Li, H., and Shi, X. (2022). Optimal discriminatory disclosure. Technical report, Discussion paper.

Hoffmann, F. and Inderst, R. (2011). Pre-sale information. *Journal of Economic Theory*, 146(6):2333–2355.

Jiang, Y. (2007). Price discrimination with opaque products. *Journal of Revenue and Pricing Management*, 6(2):118–134.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *The American Economic Review*, 101(6):2590–2615.

Kessler, A. S. (1998). The value of ignorance. *The Rand Journal of Economics*, 29(2):339–354.

Krähmer, D. and Strausz, R. (2011). Optimal procurement contracts with pre-project planning. *The Review of Economic Studies*, 78(3):1015–1041.

Krähmer, D. and Strausz, R. (2015). Optimal sales contracts with withdrawal rights. *The Review of Economic Studies*, 82(2):762–790.

Lewis, T. R. and Sappington, D. E. (1994). Supplying information to facilitate price discrimination. *International Economic Review*, pages 309–327.

Lewis, T. R. and Sappington, D. E. (1997). Information management in incentive problems. *Journal of political Economy*, 105(4):796–821.

Li, H. and Shi, X. (2017). Discriminatory information disclosure. *American Economic Review*, 107(11):3363–85.

Nocke, V., Peitz, M., and Rosar, F. (2011). Advance-purchase discounts as a price discrimination device. *Journal of Economic Theory*, 146(1):141–162.

Pavan, A., Segal, I., and Toikka, J. (2014). Dynamic mechanism design: A myersonian approach. *Econometrica*, 82(2):601–653.

Rayo, L. and Segal, I. (2010). Optimal information disclosure. *Journal of Political Economy*, 118(5):949–987.

Roesler, A.-K. and Szentes, B. (2017). Buyer-optimal learning and monopoly pricing. *The American Economic Review*, 107(7):2072–80.

Szalay, D. (2009). Contracts with endogenous information. *Games and Economic Behavior*, 65(2):586–625.

Terstiege, S. and Wasser, C. (2020). Buyer-optimal extensionproof information. *Journal of Economic Theory*, 188:105070.