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]

April 2024
(First version: 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.

Collaborative Research Center Transregio 224 - www.crctr224.de
Rheinische Friedrich-Wilhelms-Universität Bonn - Universität Mannheim

Optimal Information Design of Online Marketplaces with Return Rights*

Jonas von Wangenheim†

March 11, 2024

Abstract

Consumer data increasingly enable online marketplaces to identify buyers' preferences and provide individualized product information. Buyers, however, fully 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 the seller to sell at low prices without refund options.

JEL classification: D18, D47, D82

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, Stephan Lauermann, Vincent Meisner, Benny Moldovanu, 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. I gratefully acknowledge financial support of the German Research Foundation through CRC TRR 224 and RTG 1659.

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

1 Introduction

Over the past decade, online marketplaces have become increasingly sophisticated in identifying customers' preferences and providing consumer-specific product information, prior to purchase. Platforms like Amazon, eBay, and AliBaba recommend products based on past purchase history, leading to substantial increases in revenue.[1] Moreover, marketplaces carefully choose what information to display. Some marketplaces have rating systems that only allow for an overall, coarse rating; others, like Amazon, eBbay, or Best Buy also allow for individual comments, and split the ratings into different categories, like value, quality, and ease of use.

While a platform may control the information prior to trade, buyers will typically learn their match value for the product after it has been delivered.[2] By offering return rights for a full or partial refund, a seller on the platform can effectively condition his offer on the information a buyer holds *after* he receives and inspects the product. In this paper I analyze the information design problem of a platform when the seller can set return rights in response to the designed information. I characterize all pairs of buyer surplus and seller surplus that can arise for different ex-ante information structures. I show that for any achievable surplus pair the platform can incentivize the seller to refrain from screening through refunds and post a single take-it-or-leave-it offer.

In my model, an information designer (platform) is unrestricted in designing an ex-ante information signal about a buyer's match value (Kamenica and Gentzkow (2011)). While the buyer learns the signal realization, the seller only observes the signal distribution. This assumption expresses that the seller can observe the information on the platform, but not how it translates into a buyer's match value. Then, the seller offers a contract before the buyer learns his true valuation. Due to the sequential structure of information, the seller faces a sequential screening problem, as studied in Courty and Li (2000). The seller can screen with respect to the ex-ante information by offering a menu of contracts, where each contract specifies a price and refund for returning the good.[3]

[1]For example, according to McKinsey, 35 % of purchases on Amazon stem from recommendations, see https://www.mckinsey.com/industries/retail/our-insights/how-retailers-can-keep-up-with-consumers.

[2]Alternatively, inspection after delivery may only reveal some additional information. We may then interpret the buyer's valuation as his updated expected value. This interpretation leaves all insights of the paper unchanged.

[3]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

To understand the impact of return rights, it is insightful to revisit the information design problem in the standard monopoly problem when return rights are unavailable. Suppose buyers' valuations θ are uniformly distributed on $[0,1]$, and the seller has zero cost. For a signal structure τ we identify each realization with its posterior mean, $\tau \equiv \mathbb{E}[\theta|\tau]$, and denote with $G(\tau)$ the respective distribution of the posterior means. Roesler and Szentes (2017) show that the designer can achieve efficiency through a signal structure τ that induces a distribution of posterior means[4]

$$G_q^B(\tau) = \begin{cases} 0 & \tau \in [0,q) \\ 1 - \frac{q}{\tau} & \tau \in [q,B) \\ 1 & \tau \in [B,1]. \end{cases}$$

Such a distribution generates the same seller profit q for every price on its support $[q,B]$. Hence, without return rights a price of q maximizes the seller's profit and induces efficient trade.

I now illustrate how a seller can improve her profits using return rights under this signal. Notice that the distribution G_q^B has a mass point of $\mu \equiv \frac{q}{B}$ on B. For the sake of illustration, suppose the mass point at B is formed by pooling all types $\theta \in \left[B - \frac{\mu}{2}, B + \frac{\mu}{2}\right]$, i.e., all types in that interval receive the same signal realization.[5] Then, the seller can improve her profits by offering two different contracts from which the buyer can choose. The first contract offers the good at a non-refundable price of B, the second contract offers the good at a fully refundable price of $B+\frac{\mu}{2}$. Types in $\left[B - \frac{\mu}{2}, B + \frac{\mu}{2}\right]$ do not benefit from the second contract since all types are weakly below the price. They take the first contract, ensuring the seller a profit of q. All other signal realizations have a posterior mean below B and would make a loss from the first contract. They choose the second contract and return the good whenever they learn upon delivery that their type satisfies $\theta < B+\frac{\mu}{2}$. This achieves an additional seller profit of $(1-(B+\frac{\mu}{2}))(B+\frac{\mu}{2})$

Strausz (2011), Nocke et al. (2011), and Pavan et al. (2014).

[4]The condition that the posterior distribution must be a mean-preserving contraction of the prior imposes a constraint on the parameters q and B. The buyer-optimal signal is given by the smallest q for which this constraint can be satisfied. For the uniform distribution this is the case for $q \approx 0.204$ and $B \approx 0.872$, see Roesler and Szentes (2017).

[5]Since the decision of whether to return a good depends on the refund and the exact type that realizes after delivery, we cannot simply work with the posterior mean of a signal realization, but we have to make additional assumptions about the underlying type distribution that leads to the posterior mean of B.

from the types that do not return the good.

The example illustrates how the seller may strictly benefit from using return rights to screen ex ante. Since trade is efficient in the example, the second contract induces an inefficiency due to returns. This distortion makes the second contract unattractive for buyers who receive signal B, which enables the seller to sell to these types efficiently at a high price.[6] Notice that for the example, the use of return rights substantially decreases efficiency and buyer surplus.[7]

When the designer wants to induce efficient trade she must ensure that the seller finds it optimal to refrain from screening via different refund schedules. Moreover, the seller has the option to grant return rights at a full refund. Under such a policy the buyer can make his decision under full information after delivery, which effectively restores classical monopoly pricing under full information. Hence—and in contrast to the static design problem in Roesler and Szentes (2017)—the monopoly profit with fully informed buyers always constitutes a lower bound for the seller's profits. The upper bound on buyer surplus is therefore achieved if trade is efficient, and if the seller only receives the full information monopoly profit. For the uniform distribution this goal can be achieved through a signal with a simple structure:

Suppose all types $\theta > \frac{1}{2}$ learn their types, whereas all types $\theta \in \left[0, \frac{1}{2}\right]$ receive the same signal, and remain pooled. This signal structure enables the seller to extract the entire surplus from types in the interval $\left[0, \frac{1}{2}\right]$ through a nonrefundable price of $\frac{1}{4}$. Such a price lets all buyers buy, induces efficiency, and leaves the seller with her lower bound profit of $\frac{1}{4}$. In Proposition 1, I show that for this signal the seller does not benefit from rationing types in $\left[0, \frac{1}{2}\right]$ at the benefit of higher prices.

For arbitrary distributions, the designer has to use more sophisticated information structures to dissuade the seller from using refunds as a screening device. Nevertheless, a similar strategy can be applied: By pooling a large mass of low-valuation buyers into signals with the same mean, the designer can incentivize the seller to offer a simple take-it-or-leave-it offer without refund. As the main

[6]This intuition is reminiscent of the general "no-distortion-at-the-top" result in Courty and Li (2000) when signals are ordered by first-order stochastic dominance: The highest signal receives an efficient allocation, whereas the allocation of lower signals is distorted in order to relax ex-ante incentive constraints, and increase the price for the high signal.

[7]For the signal that is buyer-optimal without return rights we have $B + \frac{\mu}{2} \approx 0.989$. For the described contracts only types above this threshold receive a rent, leading to a buyer surplus of about 0.00006.

result, I show that the described lower bound on seller profit is the main constraint on achievable surplus division. More precisely, similar as to Bergemann et al. (2015), I show that the only limits are imposed by the minimal 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 a platform's objective to maximize any arbitrarily weighted combination of consumer surplus and producer surplus.

Bergemann et al. (2015) obtain the same feasibility result for a model in which the *seller* observes the signal realization and the buyer knows his valuation. My result shows that, interestingly, the seller's necessity to screen with respect to the signal's information does not change the set of implementable outcomes as long as screening takes place before the buyer learns his valuation.

The result highlights that information design in marketplaces can remain a powerful tool even when the seller may effectively restore full information via return rights. Similar as in Roesler and Szentes (2017), coarse information may benefit the seller as well as the buyer. In particular, coarse information for buyers with low valuations may lead to lower prices and an increase in consumer surplus. Hence, 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) point out, this policy effectively destroys the ability of a seller to screen ex ante, and leaves the consumers with the same surplus as under full information. My results suggest that whether this surplus is above or below the surplus without mandatory return rights may depend on a platform's objective when designing ex-ante information. Contested online platforms where consumers can switch to other platforms at virtually no cost may be particularly concerned about consumer surplus. In this case mandatory return rights may backfire, and be detrimental for consumers.

My paper contributes to the vast literature on information design, initiated by Rayo and Segal (2010) and Kamenica and Gentzkow (2011). In contrast to most papers in that literature, I cannot reduce the design problem to implementing a distribution of posterior means, but rather I have to regard the entire type distribution for a signal, because the buyer's incentive to return the product depends on his exact type.

Lewis and Sappington (1994) were the first to study a sellers' strategic incentive to reveal information in a trade environment. The interaction between information supply and pricing schemes has since then been studied in a number of papers, see, e.g., Bergemann and Pesendorfer (2007), Eső and Szentes (2007), Hoffmann and Inderst (2011), and Li and Shi (2017).

In the aspect that buyers strategically prefer not to be fully informed, my paper is related to the literature on strategic ignorance. Kessler (1998) was first to notice this value in a classical adverse selection model, followed by Roesler and Szentes (2017) in the context of trade.

Complementary to my approach, Guo et al. (2022) analyze seller-optimal information extensions in a sequential screening model. 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.

In the context of the interaction between learning and return rights, Lyu (2022) analyzes how a seller can influence costly buyer learning through setting price and refund rules, when learning follows a poisson process.

2 The model

A risk-neutral seller has one unit of a non-divisible good for sale. There is a risk-neutral buyer, whose valuation of the good is drawn from a commonly known prior distribution $F(\theta)$ with positive support $[\underline{\theta}, \overline{\theta}]$ and positive density $f(\theta)$.[8] 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, while the realization is

[8]The implicit restriction to continuous and strictly increasing distributions F is innocuous and made only for mathematical convenience. None of the results or intuitions in this paper hinge on these assumptions.

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 that it must be Bayes-plausible with respect to the prior F, i.e.,

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

for all $\theta \in [\underline{\theta}, \overline{\theta}]$.[9]

After the buyer observes the realization of the signal, but before the buyer learns his valuation, the seller offers a menu of option contracts. An option contract specifies an upfront payment a to the seller and an option price p, at 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. The timing of the game is as follows:

1. The third party publicly chooses a signal structure.
2. The buyer privately observes the signal realization.
3. The seller offers a menu of option contracts, the buyer accepts one of the contracts or rejects all.
4. The buyer observes his value.
5. The buyer decides whether to exert the option to buy, and the contract is executed.

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 note that in such an environment any optimal deterministic contract can be implemented as a menu of option contracts.[10] Therefore, I restriction attention to menus of option contracts at the contracting stage.

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

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

As I want to allow for various objectives, I deliberately do not specify the objective function of the third party here. In the following section I analyze for case where the third party's interest is fully aligned with the buyer, and the prior is uniform. 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 4, 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 any arbitrary surplus pair that yields at least the full information static monopoly profit for the seller.

3 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.

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

$$p^M = \arg\max_p p(1-F(p)) = \frac{1}{2}.$$

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

Note that the seller can always ignore the possibility of exploiting the signal for ex-ante screening and just charge the monopoly price once the buyer has learned the true valuation, i.e., $(a,p) = (0,p^M)$. Hence, the static monopoly profit of $\pi^M = \frac{1}{4}$ 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 all types trade and if the seller is left with the static monopoly profit for fully informed buyers 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 $\frac{1}{2}$. Buyers with a valuation below $\frac{1}{2}$ are pooled at one signal of $\tau = 0$, which induces an expected valuation of $\mathbb{E}[\theta|\tau = 0] = \frac{1}{4}$.

Suppose the seller offers a single contract $(a, p) = \left(\frac{1}{4}, 0\right)$, which means she offers the good at a price of $\frac{1}{4}$ without a refund option. Since $\mathbb{E}[\theta|\tau] \geq \frac{1}{4}$ for all τ, this offer will attract all buyers. The seller is left with her lower bound utility of $\frac{1}{4}$, and social surplus is maximized.

The following proposition claims that, given this signal structure, there is no contract that generates a higher seller utility.

Proposition 1. *Given signal τ, there is no mechanism that generates a seller profit above $\frac{1}{4}$. In particular, the contract $\left(\frac{1}{4}, 0\right)$ is a seller-optimal trading mechanism.*

Proof. If the seller chooses a menu for which buyers with $\tau(\theta) = 0$ do not take any contract, only types with full information may engage in trade, and profits are bounded by the seller's lower bound $\pi^M = \frac{1}{4}$. Suppose therefore that types with $\tau(\theta) = 0$ pick some option contract (a, p). Since this contract is also available and profitable for all fully informed types with $\tau(\theta) \neq 0$, no type will ever pick a contract with a total price above $a + p$. Conversely, if the seller offered an additional contract with a total price below $a + p$, she would leave money on the table. Hence, we can assume without loss that the seller chooses the menu $\mathcal{M} = \{(a, p)\}$. Next, we determine for any $p < \frac{1}{2}$ the upfront fee $a(p)$ that lets the ex-ante participation constraint of types $\tau(\theta) = 0$ bind. Since exactly the types with $\theta \geq p$ will decide to buy, we have

$$a(p) = \frac{\int_p^{\frac{1}{2}} (\theta - p) d\theta}{\frac{1}{2}} = p^2 - p + \frac{1}{4}.$$

Hence, for any $p \in \left[0, \frac{1}{2}\right]$ the seller's profit from offering contract $(a(p), p)$ is

$$\pi((a(p), p) = a(p) + (1 - F(p))p = p^2 - p + \frac{1}{4} + (1 - p)p = \frac{1}{4}.$$

This concludes the argument that offering contract $(a(0), 0) = \left(\frac{1}{4}, 0\right)$ is a seller-optimal trading mechanism under signal τ. $\square$

By setting the upfront fee $a(p)$ optimally, the seller extracts all the surplus generated from the types in $\left[0, \frac{1}{2}\right]$. When choosing a price $p > 0$, the seller sacrifices the surplus from types $\theta \in [0, p]$ at the benefit of a higher total price $a(p) + p = p^2 + \frac{1}{4}$ from types in the interval $\left[\frac{1}{2}, 1\right]$. This displays the classic trade-off discussed in Courty and Li (2000) between the rationing of the low signals and information the rent given to high signals. For the uniform distribution these two effects exactly offset each other.

In general, selling to all consumers in the pooling area is optimal for the seller only if the distribution has a sufficiently thin left tail. Intuitively, if the prior distribution has a lot of mass around zero, the seller can exclude many types at the benefit of higher prices with only small effects on efficiency.[11] In the next section, I show how the signal structure needs to be modified for arbitrary distributions such that the seller still finds it optimal to post a single take-it-or-leave-it offer, and to induce efficient trade at a low price.

While the optimal signal structure for the general case is more complex, the main intuitions from this example carry over. 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 in order to sell to these types. To include lower types in trade, the seller must set a low price, which is then available 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.

4 The Limits of Surplus Distribution

In this section, I fully characterize which combinations of buyer surplus and seller surplus are feasible for arbitrary priors and production costs $c \in [0, \overline{\theta}]$.

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

[11] In Section 5.2, I discuss this in more detail and show that selling to all types in a pooling interval $[0, y]$ is optimal for the seller if it generates a profit of at least π^M and if the reverse hazard rate $\frac{f(\theta)}{F(\theta)}$ is above some bound.

the seller could always use a full-refund mechanism. Finally, aggregate surplus cannot exceed first-best welfare.

The following theorem states that the above are the only constraints.

Theorem 1. *Denote by π^M the standard static monopoly profit the seller can achieve, if the buyer has full information. There exist 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*
- $u_S + u_B \leq \int_c^{\overline{\theta}} (\theta - c) f(\theta) d\theta$,

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

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

Define the threshold $x \geq c$ to satisfy

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

We construct a signal for which exactly all types above x buy, so that welfare is indeed $u_S + u_B$.
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

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

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

The boundaries x and y partition the type space. For the signal I construct, types in $[c, x]$ will not trade and will induce an efficiency loss. Hence, the location

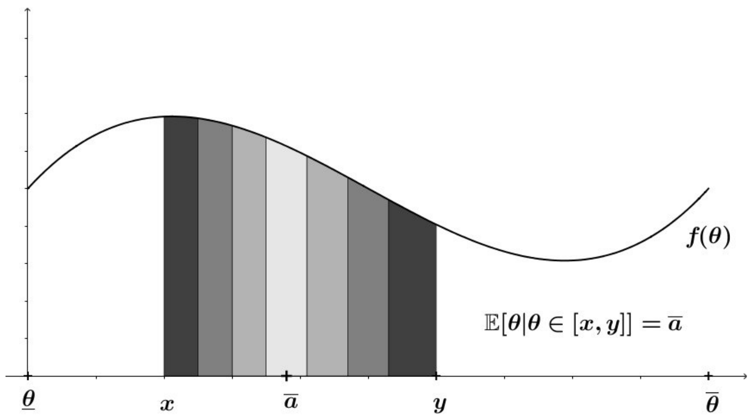

Figure 1: The signal to induce (u_B, u_S). Types of same shade are pooled together.

of x will determine welfare. All other types will trade at a non-refundable price of $\overline{a}$. Buyer types in $[x, y]$ will not receive any surplus in expectation. 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 Theorem 1.

To achieve this goal, consider the following signal construction: 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 $\{\theta^L_\tau, \theta^H_\tau\}$ with $\theta^L_\tau < \overline{a} < \theta^H_\tau$. The signal structure can be represented by

$$\tau(\theta) = \begin{cases} \theta - \overline{\theta} & \theta < x, \\ \int_\theta^{\overline{a}} f(s)(\overline{a} - s)ds & \theta \in [x, y], \\ \theta & \theta > y. \end{cases} \tag{3}$$

The signal $\tau(\theta)$, again, prescribes full learning for $\theta < x$ and $\theta > y$.[12] Notice

[12]The signal's range for types $\theta < x$ is shifted by $\overline{\theta}$ to ensure that it is disjoint to the range of

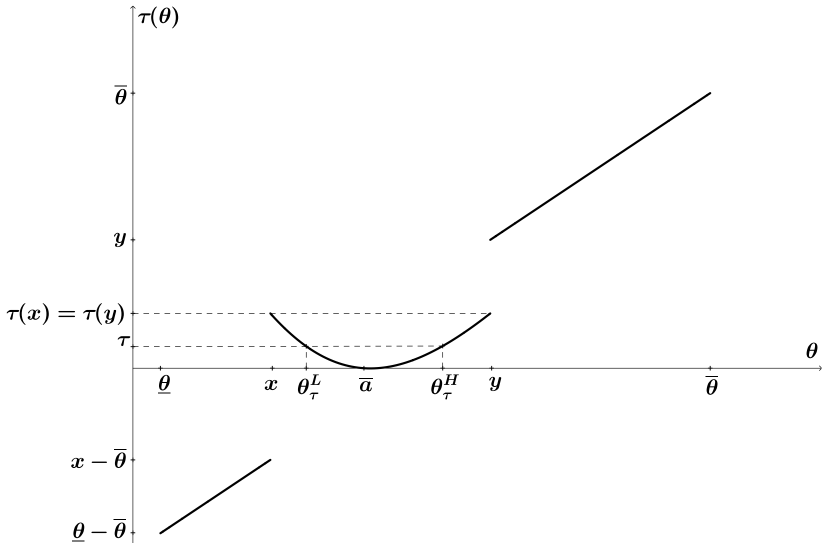

Figure 2: The signal structure $\tau(\theta)$.

that each signal realization τ in $[0, \tau(x)]$ indeed corresponds to two possible types $\theta^L_\tau, \theta^H_\tau$. Similar as in the discrete case, types are paired such that

$$\mathbb{E}[\theta|\theta \in [\theta^L_\tau, \theta^H_\tau]] = \overline{a}$$

for all respective pairs in $[x, y]$.[13] As I show in the proof of Theorem 1, this property implies that for each τ in the pairing region we have $\mathbb{E}[\theta|\tau] = \overline{a}$.

The fact that all types in $[x, y]$ receive a signal with mean $\overline{a}$ makes it attractive for the buyer to sell at a non-refundable price of $\overline{a}$. The pairwise pooling in the interval $[x, y]$ ensures that there is no better contract for the seller and that there is no scope for screening via a larger menu with different refund rules.

To understand why the pairwise pooling destroys the seller's incentive to screen, recall the seller's trade-off when the signal pools all types in $[x, y]$ together in one signal realization. As discussed in Section 3 for the example $[x, y] = \left[0, \frac{1}{2}\right]$, the seller may then offer a single contract (a, p) with $x < p < y$. Such a contract sacrifices the surplus from the types in $[x, p]$ at the benefit of selling to all higher

the types in $[x, y]$. Since this is a monotone transformation, the buyer can still unambiguously infer his type from the signal realization for any $\theta < x$.

[13]To see this formally, note that $0 = \frac{\tau(\theta^L_\tau) - \tau(\theta^H_\tau)}{F(\theta^H_\tau) - F(\theta^L_\tau)} = \frac{\int_{\theta^L_\tau}^{\theta^H_\tau} f(s)(\overline{a} - s)ds}{F(\theta^H_\tau) - F(\theta^L_\tau)} = \mathbb{E}[\overline{a} - \theta|\theta \in [\theta^L_\tau \theta^H_\tau]]$.

types at a higher total price $a + p$. This approach does not benefit the seller for the pairwise pooling signal: While a contract (a, p) still sacrifices the types in $[x, p]$ this sacrifice does not affect the willingness to pay of types in $[p, \overline{a}]$, who receive different signals.

The key point is that all types below $\overline{a}$ receive distinct signals. Suppose the seller wants to sell to *some* type $\hat{\theta} \in [x, \overline{a}]$. This type is the low type for the respective signal, hence the high type will buy as well. Ex-ante individual rationality then dictates that $a + p \leq \mathbb{E}[\theta|\tau(\hat{\theta})] = \overline{a}$. If such a cheap contract is available, no type $\theta > y$ from the full information region will ever pay more than a total price of $a + p = \overline{a}$. (Evidently, offering these types any contract at an even lower total price is not optimal for the seller, because she would leave money on the table.) Hence, whenever the seller decides to sell to *at least one* type in $[x, \overline{a}]$ this implies the necessity of offering some contract with $a + p = \overline{a}$, and the seller can readily sell to *all* types in $[x, y]$ at such a price.[14] This can be achieved by a simple take-it-or-leave-it offer at price $\overline{a}$, i.e., the contract $(\overline{a}, 0)$.[15]

I now explain why it is not optimal for the seller to sell only to types $\theta > \overline{a}$. Since the buyers in $[\overline{a}, y]$ receive different signals the seller may in principle use a large menu with different refund options to screen for different signal types. Here, the benefit of the specific anti-assortative pairing comes into play. The fact that for any signals τ_1, τ_2 we have $\theta^H_{\tau_1} > \theta^H_{\tau_2}$ whenever $\theta^L_{\tau_1} < \theta^L_{\tau_2}$ implies that the signals of types in $[\overline{a}, y]$ are ordered in the sense of mean preserving spreads. As I show in the proof of Lemma 1 in the appendix, this leads to an ordering on the ex-ante incentive constraints: Whenever some type $\hat{\theta} \in [\overline{a}, y]$ receives positive utility from a contract then so does any type $\theta > \hat{\theta}$. Hence, types $\theta > \hat{\theta}$ would only select a different contract than type $\hat{\theta}$ if it generated higher surplus to him, and hence less to the seller. This implies that the seller will not use more than one contract. The last step of the proof is to show that a single contract for which only types in $[\overline{a}, \overline{\theta}]$ buy cannot generate more profit than the full information monopoly profit, the seller's lower bound utility.[16]

[14]The formal arguments are displayed in Case 2 in the proof of Theorem 1.

[15]Effectively, the seller is indifferent between any contract $(\overline{a} - p, p)$ as long as $p \leq x$, but they all give rise to the same allocation. Any such contract induces a total price of $\overline{a}$ and no returns.

[16]The formal arguments are displayed in Case 3 in the proof of Theorem 1.

5 Discussion and Extensions

5.1 Return Rights, Shipping Costs, and Informational Frictions

In the previous section, I showed how a designer can implement efficiency via a signal that induces a remarkably simple menu with two features: First, the menu consists of only a single offer, and second, this offer consists of a simple price with no right of return. In this section, I discuss the role of return rights. I argue that providing efficiency through information rather than through returns is strictly optimal whenever returns impose some cost. I then discuss various frictions that may constrain the designer in practice to reveal all relevant information. I argue how return rights can help to increase efficiency in the face of these constraints.

The signal in Theorem 1 is able to induce efficiency without returns because it is fully revealing to all buyers with $\theta < c$. While it seems natural that an unconstrained information designer would reveal all inefficient types through the signal, this is in general not necessary to achieve efficiency. If the signal pools types above and below c into the same signal realization, the outcome can still be efficient if the signal structure again dissuades the seller from imposing distortions, but induces her to offer a contract (a, c). Such a contract can achieve efficiency through returns, because buyers return the good if and only if $\theta < c$. In reality, however, returns typically involve material costs in the form of sending back the good, as well as inconvenience costs for the buyer.

Denote the sum of these costs by γ and assume that it is the same for all buyer types. I now explain why we can assume without loss of generality that the seller bears the cost of returns. If the buyer bears the cost of returns, a contract (a, p) implies a price of $a + p$ if the buyer orders and keeps the good, and a price of $a + \gamma$ if the buyer first orders and then returns the good. This is equivalent to a contract $(a+\gamma, p-\gamma)$, together with the assumption that the seller bears the cost of returns.[17] Since the optimal contract for the signal in Theorem 1 generates no returns even when the return cost is zero, it is strictly optimal for the seller when she has to bear a positive cost for each return. Thus, under positive return costs, Theorem 1 still holds, and the feature that the given signal does not induce

[17] In principle, $p - \gamma$ could be negative. In this case, it is natural to assume free disposal, which implies that the buyer never returns the good. Such a contract is therefore equivalent to a take-it-or-leave-it offer at price of $a + \gamma$ with no refund.

returns becomes strictly optimal.

Theorem 1 provides an optimal signal for the natural benchmark that the designer fully controls all relevant information and is unconstrained in her ability to design any signal. In practice, there may be different kinds of constraints. Most obviously, the designer may be constrained in her ability to identify buyers' preferences or to design targeted information for buyers. For example, if the signal contains "mistakes" and sends high signals to buyers with valuations of $\theta < c$ with some positive probability, return rights may help to increase efficiency as well as buyers' ex-ante willingness to pay.

Moreover, even if a platform perfectly controls all relevant product information, there may still be uncertainty about the quality of the seller. By offering return rights, a seller may signal that she is trustworthy and separate herself from fraudulent sellers on a platform. Alternatively, on platforms for used or refurbished products, a seller may hold private information about the quality of a particular item. Return rights may then help to overcome adverse selection problems in the sense of Akerlof (1970). Interestingly, in these two cases the mere provision of a right of return may help to overcome the problems. As pointed out in Footnote 15, for the efficient signal in Theorem 1 the contract $(\overline{a} - c, c)$ induces the same allocation as the contract $(\overline{a}, 0)$: All types $\theta > c$ trade at a total price of $\overline{a}$, and no one returns the good. Thus, return rights may not even be exercised in equilibrium, but they can still be an effective tool to increase trust in the transaction.

5.2 Full Pooling versus Anti-Assortative Pairing

In Section 4, I showed how for arbitrary distributions a pairwise pooling construction on some interval $[c, y]$ can achieve efficiency. In Section 3, I demonstrated that for a uniform prior, $c = 0$, and $y = \frac{1}{2}$ full pooling on $[c, y]$ was sufficient to implement efficiency. This raises the natural question of under which conditions the more complicated pairwise pooling is necessary to implement efficiency.

To save notation, normalize $[\underline{\theta}, \overline{\theta}] = [0, 1]$ and $c = 0$ for this section.[18] We analyze under which conditions the signal

[18] This normalization implies that we focus on the case of efficient trade. However, for the implementability of efficiency this is without loss of generality. If there are types who do not trade under efficiency, we can easily extend the signal construction to be fully revealing for these types. This implies that they never buy and all results in this section remain to hold.

$$\tau(\theta) = \begin{cases} 0 & \theta \leq y \\ \theta & \theta > y \end{cases} \tag{4}$$

implements efficiency.

If the seller offers some contract (a, p) to types $\tau = 0$ we can again assume without loss that this is the unique contract in the menu: Since this contract is also available and profitable to all higher types, no type will ever pick a contract with a higher total price than $a + p$, whereas offering another contract at a lower price would not be profit maximizing. Moreover, an optimal contract lets the ex-ante participation constraint of type $\tau = 0$ bind, hence for any optimal contract choice (a, p) with $p \in [0, y]$ we have

$$a(p) = \frac{\int_p^y (\theta - p) f(\theta) d\theta}{F(y)}.$$

Efficiency requires that $p = 0$, since otherwise types in $[0, p]$ do not trade. Hence, we need to analyze under which conditions the contract $(a(0), 0)$ maximizes the seller's utility.

A necessary condition is that selling to all consumers at a price of $a(0) = \mathbb{E}[\theta|\theta < y]$ generates at least the profit from selling to all types under full information, hence

$$\mathbb{E}[\theta|\theta < y] \geq \pi^M, \tag{5}$$

where π^M is the static full information monopoly profit. Suppose (5) is satisfied. Then signal τ induces efficiency if and only if $p = 0$ maximizes

$$\begin{aligned} \pi(a(p), p) &= \frac{\int_p^y (\theta - p) f(\theta) d\theta}{F(y)} + (1 - F(p))p \\ &= \frac{1}{F(y)} \left[\int_p^y \theta f(\theta) d\theta + \big(- F(y) + F(p) + F(y) - F(p)F(y) \big) p \right] \\ &= \frac{1}{F(y)} \left[\int_p^y \theta f(\theta) d\theta + (1 - F(y)) F(p) p \right] \end{aligned}$$

among all $p \in [0, y]$.

A sufficient condition is that for all $p \in [0, y]$

$$\frac{\partial \pi(a(p), p)}{\partial p} = -f(p)p + (1 - F(y))\frac{F(p)}{F(y)} \leq 0. \tag{6}$$

Equation (6) illustrates the trade-off between the marginal and infra-marginal effect of a price increase. Recall that the seller extracts all surplus from types in $[p, y]$. The first summand stems from losing the surplus of the marginal consumer, the second summand captures the infra-marginal effect of a higher total price $a + p$ paid by buyers in $[y, 1]$. Notice that the second (positive) effect of an increase in p only depends on the mass of consumers $F(p)$ that are excluded from trade, whereas the first (negative) effect depends on p directly. Intuitively, if there is a lot of mass around $\theta = 0$, then excluding these types is not very costly in terms of lost efficiency, whereas the benefit from the price increase is independent of the valuation of the excluded consumers and only depends on *how many* consumers are excluded. Hence, whenever there is too much mass in the left tail of the distribution, efficient trade is not feasible under full pooling in $[0, y]$.[19] By rearranging the terms in (6), we can immediately relate this intuition to a bound on the reversed hazard rate $\frac{f(\theta)}{F(\theta)}$, and have shown the following result:

Corollary 1. *Let $c = 0$. Consider some $y > 0$ with $\mathbb{E}[\theta|\theta \in [0, y]] \geq \pi^M$. If the reversed hazard rate of the type distribution satisfies*

$$\frac{f(\theta)}{F(\theta)} \geq \frac{1 - F(y)}{\theta F(y)}$$

for all $\theta \leq y$ then the signal which pools all types in $[0, y]$ and fully reveals to all other types induces efficient trade.

6 Conclusion

In this paper, I analyzed the role of ex-ante information when buyers eventually learn their valuation and the seller is free to guarantee return rights. I fully characterized all pairs of buyer surplus and seller surplus that an information designer can generate, and identified a lower bound on seller profit as the only constraint. Moreover, each surplus pair can be induced with a remarkably simple optimal contract, namely a simple posted price offer without refund.

[19] This holds in particular whenever there is any mass point in the distribution sufficiently close to (but not at) zero.

For the buyer-optimal outcome, buyers receive incomplete ex-ante information. Pooling low-value consumers in a specific way induces the seller to refrain from screening and to sell efficiently at a single low price.

With rapid advances in data analytics, large marketplaces are likely to become even more sophisticated at providing targeted information to customers in the future. Their information design may have a significant impact on consumer welfare. This insight naturally raises the question of effective consumer protection on these platforms. Traditional consumer protections, such as mandatory information disclosures and right of return, aim to increase the information available to the buyer when making a purchase decision. As argued in this paper, the effect of such interventions on consumer surplus can be ambiguous. The question of how to incentivize marketplaces to use targeted information for the benefit of consumers remains an important topic for future research.

7 Appendix: 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

$$u_S + u_B = \int_x^{\overline{\theta}} (\theta - c) f(\theta) d\theta = (1 - F(x)) \mathbb{E}\big[(\theta - c) | \theta \in [x, \overline{\theta}]\big]. \tag{7}$$

Since f has full support, the right-hand side in (7) 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 (7) is satisfied. Define now y implicitly by

$$u_S = (1 - F(x)) \mathbb{E}[(\theta - c) | \theta \in [x, y]]. \tag{8}$$

Note that the right-hand side in (8) is strictly increasing in y. For $y = x$ it takes a value

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

whereas for $y = \overline{\theta}$ it takes a value

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

Hence, by the intermediate value theorem, there is indeed a unique $y \in [x, \overline{\theta}]$ which satisfies (8). Furthermore, define

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

Finally, define the following signal structure:

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

Since all types $\theta < x$ and $\theta > y$ lead to different signal realizations, the signal prescribes full learning for these types. Next, we show that any realization τ from types in $[x, \overline{a}) \cup (\overline{a}, y]$ corresponds to exactly two types $\theta^L_\tau, \theta^H_\tau$, where $\theta^L_\tau < \overline{a} < \theta^H_\tau$.[20]

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

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

Thus, for any τ with $0 < \tau \leq \tau(x)$ there are exactly two types $\theta^L_\tau, \theta^H_\tau$ with $\tau = \tau(\theta^L_\tau) = \tau(\theta^H_\tau)$, where, without loss of generality, $\theta^L_\tau < \overline{a} < \theta^H_\tau$.

Next, we establish that $\mathbb{E}[\theta|\tau] = \overline{a}$ for all signal realizations from types in the

[20] See Figure 3 for a graphic illustration.

pairing region $[x, y]$.[21] Let us call $\theta^L(\tau)$ the inverse function of $\tau(\theta)$ on $[x, \overline{a}]$, and $\theta^H(\tau)$ the inverse function of $\tau(\theta)$ on $[\overline{a}, y]$. Then, for all $z \in (0, \tau(x)]$

$$\begin{aligned} 0 &= \tau(\theta_z^L) - \tau(\theta_z^H) \qquad (9) \\ &= \int_{\theta_z^L}^{\overline{a}} f(s)(\overline{a} - s)ds - \int_{\theta_z^H}^{\overline{a}} f(s)(\overline{a} - s)ds \\ &= \int_{\theta_z^L}^{\overline{a}} f(s)(\overline{a} - s)ds + \int_{\overline{a}}^{\theta_z^H} f(s)(\overline{a} - s)ds \\ &= -\int_{\theta_z^L}^{\theta_z^H} sf(s)ds + \overline{a} \cdot \text{Prob}(\theta \in [\theta_z^L, \theta_z^H]). \end{aligned}$$

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

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

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

$$\mathbb{E}[\theta|\tau] = \overline{a} \qquad (10)$$

almost surely.[22] 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 $F(\cdot|\tau_1)$.

Finally, we derive probability weights for the paired types, given their signal realization. For the resulting regular conditional probabilities we obtain that

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

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

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

The menu

[21] Intuitively, the key property we employ is the fact that for any pooled pair, $\theta_\tau^L, \theta_\tau^H$, we have $\mathbb{E}[\theta|\theta \in [\theta_\tau^L, \theta_\tau^H]] = \overline{a}$, see (9), or Footnote 13. The fact that this holds for a nested interval structure that is a generator for the Borel sets in $[x, y]$ yields the result.

[22] 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.

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 have a seller utility of

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

and a buyer surplus of

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

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

The optimality of the menu

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 $\tilde{\mathcal{M}}$. To show that $\mathcal{M}$ is optimal we need to show that $\tilde{u}_S \leq u_S$.

Let $\hat{\theta}$ be the lowest type who purchases the good under $\tilde{\mathcal{M}}$, in the sense that he chooses some $(a, p) \in \tilde{\mathcal{M}}$, and then decides to buy the good at the price p, after he learns his type $\hat{\theta}$.

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 a 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 $\theta^H_{\tau(\hat{\theta})}$. Hence, both— that is, *all*—types corresponding to signal $\tau(\hat{\theta})$ purchase the good.[23] 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

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

where in the transition from second to third line we exploited the fact that, by construction, $\mathbb{E}[\theta|\theta \in [x, y]] = \overline{a}$.

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

[23]This is the key step, where we exploit that all types in $[x, \overline{a}]$ receive different signals. If we want to sell to one of them, we are immediately selling to all types in the respective signal realization, and a buyer with that signal realization is ex ante not willing to pay more than $\overline{a}$, his expected type. The remainder of the calculations of Case 2 simply show that with such a cheap contract available the seller cannot make a profit above u_S. Note that this logic is in contrast to the uniform example, where types below $\overline{a} = \frac{1}{4}$ were pooled. In that case, excluding some of the low types in $\left[0, \frac{1}{4}\right]$ lead to a total price $a + p$ above $\overline{a} = \frac{1}{4}$ for the remaining types.

good for p whenever $\theta^H_{\tau(\hat{\theta})}$ does, violating that $\theta^H_{\tau(\hat{\theta})}$ is the lowest type to purchase the good. The following technical Lemma establishes an order on the ex-ante participation constraints, that stems from the fact that types are paired in an anti-assortative way. Due to this pairing, the type distributions conditional on the signal realizations for types in $[x, y]$ are ordered in the sense of mean-preserving spreads. This ordering leads to the property that if the ex-ante participation constraint is satisfied for $\tau(\hat{\theta})$, it is a fortiori satisfied for any higher $\tau \in [\tau(\hat{\theta}), \tau(y)]$.[24]

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 (10) 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

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

[24] Courty and Li (2000) show a similar result for the case where signals are ordered in the sense of mean preserving spreads and have non-shifting support. The latter is violated in our case.

Now, since $\theta^L_{\tau_2} < \theta^L_{\tau_1} < p$ and $\alpha_2 \leq \alpha_1 < 1$, we have[25]

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

□

Next, we argue that this order on the ex-ante incentive constraints implies that it is unprofitable for the seller to use more than one contract. Lemma 1 implies that if signal type $\tau(\hat{\theta})$ receives weakly positive utility from contract (a, p) then so does any signal type $\tau(\theta)$ with $\theta \in (\hat{\theta}, y]$. Moreover, 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). This means that such a contract (a, p) induces all types $\theta \geq \hat{\theta}$ to purchase the good. Since, by assumption, $\hat{\theta}$ is the lowest type to purchase the good for menu $\tilde{\mathcal{M}}$, any additional contract in the menu does not increase trade efficiency. Moreover, a buyer would only take another contract if it yielded higher rents to him than the contract (a, p), and thus lower rents to the seller. Therefore, the seller can only lose from separating any types in $[\hat{\theta}, \overline{\theta}]$ with more contracts in the menu. Hence, we can assume $\tilde{\mathcal{M}} = \{(a, p)\}$.

We have established that in Case 3 it is optimal for the seller to offer a unique contract (a, p), and all types $\theta \geq \hat{\theta}$ end up buying under this contract. Seller utility is given by

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

When choosing a contract (a, p) under which type $\hat{\theta}$ buys, it is optimal for the seller to make this type's ex-ante IR bind, i.e.,

$$a = \mathbb{P}(\hat{\theta}|\tau(\hat{\theta}))(\hat{\theta} - p).$$

[25]Notice that here we exploit the anti-assortative pairing, since only under this pairing does $\theta^H_{\tau_2} > \theta^H_{\tau_1}$ imply that $\theta^L_{\tau_2} < \theta^L_{\tau_1}$.

By plugging the IR constraint into the seller's profit function we obtain the following profit function

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

This equation displays a trade-off between a high ex-ante fee a and high price p. The ex-ante fee is paid by all types in $[x, \theta^L_{\tau(\hat{\theta})}] \cup [\hat{\theta}, \overline{\theta}]$, whereas the price p is only paid by types in $[\hat{\theta}, \overline{\theta}]$. However, by the ex-ante IR, a unit decrease in the price can only be translated into an increase in the upfront fee of $\mathbb{P}(\hat{\theta}|\tau(\hat{\theta}))$.

In the following we do not analyze how the seller solves the trade-off between a and p optimally but rather show that in any case her profit is bounded by u_s. First, consider the case

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

which corresponds to the case where a high upfront fee is optimal. The following calculations show that in this case the profit is lower than the one that can be achieved by $(a, p) = (\overline{a}, 0)$.

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

Conversely, consider

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

which corresponds to the case where a low upfront fee is optimal. The following calculation shows that in this case the profit is lower than the one that can be achieved by $(a,p) = (0,\hat{\theta})$.

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

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

References

Akerlof, G. A. (1970). The Market for "Lemons": Quality Uncertainty and the Market Mechanism. *The Quarterly Journal of Economics*, 84(3):488–500.

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.

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

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.

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.

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

Lyu, Q. (2022). Optimal refund mechanism. Technical report, Discussion paper.

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.

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