Consumer-Optimal Information Design

Jonas von Wangenheim (Humboldt University Berlin)

Discussion Paper No. 53

November 2, 2017

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

Consumer-Optimal Information Design*

Jonas von Wangenheim†

November 2, 2017

Abstract

In many trade environments—such as online markets—buyers fully learn their valuation for goods only after contracting. I characterize the buyer-optimal ex-ante information in such environments. Employing a classical sequential screening framework, I find that buyers prefer to remain partially uninformed, since such an information structure induces the seller to set low prices. For the optimal information signal, trade is efficient, and the seller only extracts the static monopoly profit. Further, I fully characterize all possible surplus divisions that can arise in sequential screening for a given prior.

JEL classification: D82

Keywords: information disclosure, sequential screening, strategic learning, Bayesian persuasion, mechanism design

---

*I thank Andreas Asseyer, Helmut Bester, Giacomo Calzolari, Lucien Frys, Paul Heidhues, Daniel Krähmer, 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), and the 2017 EARIE conference (Maastricht) for helpful comments and discussion. I gratefully acknowledge financial support of the German Research Foundation through CRC TRR 190 and RTG 1659.

†Humboldt-Universität zu Berlin, Institute for Economic Theory 1, Spandauer Str. 1, D-10178 Berlin (Germany), jonas.wangenheim@fu-berlin.de

1 Introduction

Over the recent decade, trade has increasingly shifted towards online markets. Commonly in these markets, consumers do not observe all product characteristics at the time of purchase. For instance, if a consumer buys clothes he may only have a vague idea of the cut and the color. If he books a hotel online, he may learn some coarse information from the hotel's number of stars or the reviews of other customers, but many consumer-specific details will remain unclear until his actual arrival. Consequently, despite having some ex-ante information, the buyer will learn whether the good matches his private taste sufficiently only ex post, ie., after he contracts with the seller and gets access to the good.

A monopolist may exploit this partial uninformedness of the consumer, and offer contracts which leave only small information rents to consumers. This paper addresses the question to what extent regulation that obliges the monopolist to provide consumer-information, can protect consumers against an exploitation of the monopolist's market power. In particular, I derive the buyer-optimal information design in such markets.

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 sellers respond in their contract offers 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 information signal induces efficient trade and distributes all rents in excess of the classical static monopoly profit with fully informed buyers, to the buyer. Moreover, a seller-optimal contract for the buyer-optimal information always consists of a simple buy-now offer without refund. In a second step, I characterize all divisions of buyer surplus and seller surplus that can arise under different information signals for a given prior.

While a buyer may have incomplete information about his value for a product at the time of contracting, I assume that he learns his exact valuation ex post by inspecting the product after delivery.[1] Due to the sequential information structure,

[1] Alternatively, assume that inspection only reveals some additional information, and interpret

the seller faces a sequential screening problem. She optimally screens buyers by offering several contracts, which differ in price and refund conditions, as studied in Courty and Li (2000).[2] I extend their sequential screening framework, by allowing the buyer to decide how much he wants to learn about his valuation for the good before contracting. More specifically, the buyer, or the regulator on behalf of the buyer, first chooses 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.[3] Then, the seller offers a contract, before the buyer learns his true valuation.

The seller screens the buyer with respect to his signal realization. She optimally offers a menu of option contracts, each specifying a price, and the refund conditions. Intuitively, buyers with higher valuation uncertainty are more attracted by contracts with high refund flexibility.[4]

To better understand why a partial information revelation can be beneficial to the buyer, assume that trade is efficient for all buyer types, and consider the full information benchmark:

If the buyer learns his exact valuation by choosing a fully informative signal, there is no further learning. The seller charges the static monopoly price, leaving an information rent to the buyer. The rent, however, may come at the cost of trade inefficiencies, since the monopoly price in general induces only high types to buy.

This benchmark naturally lead to the question, whether there is an only partially informative signal that induces efficient trade, and distributes the additional rents of this more efficient allocation to the buyer. Indeed, I show that for any prior distribution there is a suitable signal structure, such that the seller chooses a contract for which

buyer's valuation as his updated value estimate. As buyers are risk neutral, this leaves all insights of the paper unchanged.

[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).

[3]I am only interested in product information that relates to subjective taste. By the unraveling argument of Viscusi (1978), every seller will disclose any information on quality, if he can credibly and costlessly do so.

[4]Due to the buyer's freedom to design information signals, the regularity conditions, imposed in Courty and Li (2000), may be violated. Thus, we cannot rely on their analysis to find the optimal contract.

1. trade is efficient, and

2. the seller only receives the static monopoly profit of fully informed buyers.

Note that this static monopoly profit is always a lower bound on the seller's profit, since she can always charge the static monopoly price and allow full refund, after the buyer learns his type. Hence, such a signal is buyer optimal in the sense that it maximizes the buyer surplus.

The optimal signal keeps low types partly uninformed, while high types have full information. Indeed, if different low types obtain the same signal, the seller can sell to these types by providing less information rent, since they have to break even only on average, rather than individually. Consequently, the seller has an incentive to lower the price below the static monopoly price, in order to increase participation. Lower prices increase efficiency as well as rents for high types.

Moreover, I fully characterize the possible combinations of buyer surplus and seller surplus that can arise in the sequential screening model for different signal distributions. Similar to Bergemann et al. (2015), I show that the only limits are imposed by the natural constraints that

1. buyer utility is nonnegative,

2. the seller receives at least the static monopoly profit, and

3. aggregate surplus does not exceed the first-best gains from trade.

The remainder of the paper is structured as follows: After discussing the relevant literature, I introduce the model in Section 3. Section 4 covers the case of a uniform distribution and provides an illustrative example. In Section 5, I construct the buyer-optimal signal. Section 6 characterizes all possible surplus division, before Section 7 concludes. All proofs are relegated to the Appendix.

2 Related Literature

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

A recent branch of the literature, building on the pioneer work of Lewis and Sappington (1994), studies sellers' strategic information revelation. Bergemann and Pesendorfer (2007) analyze auctions, where the seller can choose the accuracy by which the buyers learn their private valuations. 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, and argue that the seller should always disclose all relevant information. Li and Shi (2016) show that this no longer holds when the seller can use discriminatory information disclosure. If buyers have different ex ante types, and the provided information can depend to the reported types then partial information disclosure may be optimal. Hoffmann and Inderst (2011) characterize optimal contracts for the case where the buyer's and the seller's information are stochastically independent.

My paper conversely analyzes buyers' optimal information acquisition. The agent may acquire private information costlessly and observably *before* the contractual relationship, to obtain 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 principle aims to contractually provide incentives for costly learning.[5]

The model is probably closest related to Roesler and Szentes (2017), who characterize the buyer-optimal signal in a classical static one-unit trade environment. In contrast to their setup, I assume that after delivery the buyer receives additional information that affects his valuation. 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. I consider my framework to be more appropriate in the context such as internet markets, where consumer typically receive additional information upon the good's

[5]E.g., Lewis and Sappington (1997), Crémer et al. (1998), Szalay (2009), Krähmer and Strausz (2011).

delivery (see Krähmer and Strausz (2011)), while the context of Roesler and Szentes seems more appropriate in markets in which this learning effect is negligable.

Kessler (1998) analyzes the value of ignorance in a classical adverse selection model with two types. She finds that, even if the agent can learn his type costlessly, he will choose a signal that is uninformative with some positive probability, in order to receive a more favorable contract.

Bergemann et al. (2015) analyze trading contracts, where the seller has information beyond the prior distribution. In particular, they characterize the buyer-optimal seller information structure. In contrast to my 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 lately drawn a lot of attention, and produced a vast literature on Bayesian persuasion, based on the work of Kamenica and Gentzkow (2011). My setup is different from persuasion, since the buyer himself is uninformed. We can, however, make use of the tools from the framework of arbitrary signal choices. (One interpretation of my model is that the buyer tries to "persuade" himself, in the sense that he wants to manipulate his beliefs to obtain a strategic advantage towards the seller.)

3 The model

A seller can produce one unit of a good at zero cost. The valuation of the good for a buyer is drawn from a commonly known prior distribution $F(\theta)$ on some positive support $[\underline{\theta}, \overline{\theta}]$ with positive, continuous density $f(\theta)$.[6] Before contracting and learning the valuation, the buyer (or the regulator in the buyer's interest) can choose a signal structure to gain some information on the valuation. The signal distribution is commonly observed, the realization is private information to the buyer. I allow for any general signal structure in form of a Borel-measurable signal space $T \subseteq \mathbb{R}$, together with a probability measure μ on the Borel σ-algebra of

[6]The restriction to a positive support is only to keep the exposition transparent and tractable. It does not change the results. Indeed, if trade is inefficient for some buyer types, one can interpret an optimal learning process as a two step procedure. First the buyer learns whether $\theta > 0$, and then applies the optimal learning process described in this paper to the conditional distribution on θ being larger than 0.

$[\underline{\theta}, \overline{\theta}] \times T$. 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 $\theta \in [\underline{\theta}, \overline{\theta}]$.[7]

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

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

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 buyer 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. 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

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

true valuation.[8]

In the following section, I derive the buyer-optimal signal, which achieves the upper bound of buyer utility, for a uniform prior. In Section 5, I show how the construction generalizes to arbitrary prior distributions if we restrict the seller to the use of option contracts.

4 The Uniform Case

It is instructive to analyze first the case of a uniform prior, as it catches the main intuitions.

Let the prior $F(\theta)$ be the uniform distribution on $[0,1]$. 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 trade is always efficient, the upper bound for buyer surplus is achieved, if trade always occurs, and the seller is left with her monopoly profit π^M. The main result 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 valuation below 1/2 are pooled in one signal of $\tau = 0$, which induces an expected valuation of $\mathbb{E}[\theta|\tau = 0] = 1/4$.

[8] 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.

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. Using the tools of mechanism design, 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 utility 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' individual rationality constraints bind, and they make zero profits on average, high types benefit from lower prices and buyer surplus increases. Since more types trade, efficiency increases as well.

Applications and Discussion

There are numerous ways in which a regulator or intermediary can control the amount of product information exposed to consumers prior to trade.

One natural application for the use of information design are internet platforms. Especially in the hospitality and travelling industry it is common to offer car rentals, holiday packages, hotel stays, or airline tickets on internet platforms, such as online travel agencies.[9] By collecting personalized data, platforms can gather a profound understanding of consumers preferences. Further, they are able to discriminate product information with respect to individual consumers.

[9] According to Green and Lomanno (2012), in 2010 about 11 percent of all revenues in the US hotel industry were generated by online travel agencies like Expedia, Priceline, and Orbitz.

If the platform has to grant standard monopoly profits to hotels, but aims to maximize consumer surplus due to platform competition, it faces exactly the information design problem described in the model. As seen in Proposition 1 and generalized in Theorem 1, the platform optimally does not provide all product details to consumers, but leaves details of the deal somewhat opaque, and sells at low prices. Indeed, such "opaque deals", are common practice in online travel agencies such as priceline and hotwire: The platforms offer discounted deals, which guarantee specific features such as the number of hotel stars or location at the city center, but reveal the identity and other details of the hotel only after payment.[10]

Note that in contrast to Shapiro and Shi (2008) and Balestrieri et al. (2015), where opaque selling is a result of firm's profit maximizing behavior, in my model it appears as the natural tool to maximize consumer surplus, which provides a novel perspective for the use of opaque goods.[11]

Alternatively, a regulator may control the amount of product information by specific labelling requirements, certification standards, or—as Hoffmann et al. (2017) argue— by regulating the length of trial periods.

For instance, both the USA and the EU require sellers to label food ingredients on the package in descending order of predominance by weight, yet not by the exact amount.[12] The same EU regulation requires firms to label the nutrition value with a *Guideline Daily Amount* (GDA), however Grunert et al. (2010) find that only about 70 percent of customers have a conceptional understanding of its meaning. They find that the understanding is positively correlated with the interest in healthy eating, which suggests that the information design is particularly informative to consumers who have a high value for healthy food.

A different information design approach is taken by the Food Standard Agency (FSA) in the UK. In 2006 they introduced the traffic light rating system, under which nutrition values—such as sugar or saturated fat—are highlighted in red (high), amber (medium) or green (low). In their literature review on food labelling Hawley et al. (2013) conclude that traffic light ratings have "most consistently

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

[11] In Shapiro and Shi (2008), opaque selling arises as an equilibrium under competition with differentiated consumers, whereas Balestrieri et al. (2015) show that a monopolist's optimal selling strategy for substitutes may feature opaque options.

[12] USA: 21CFR §101.4, EU: Regulation 1169/2011

helped consumers to identify healthier products".

The following example depicts how certification regulation may provide individual consumers with the information signal that generates the consumer-optimal information structure.

Example: Hotel Certifications

This stylized example aims to illustrate that, in the terms of this paper, the current european system of hotel certification can be understood as providing consumers with only partial information. The example, however, also serves to show that, based on this paper's arguments, the information provision underlying the certification—designed by the hotel associations—is not consumer optimal. The consumer-optimal information regulation is derived.

In 2009 hotel associations from seven European countries founded the "Hotelstars Union" to harmonize the national standards of hotel certifications. By 2017 the system was adapted by 17 countries within the European Union, with only very slight differences between participating countries.

The grading system mainly consists of five different quality levels, represented by one to five stars[13]. Participating hotels gather points by providing features from a list of over 200 possible criteria, divided into categories as *reception, services, gastronomy,* and *leisure*. Hotels who want to certify a certain number of stars must achieve a respective number of points. Besides some minimum requirements for each star, hotels are entirely free in how to achieve the number of points.

While the number of stars may be a good measure for the overall hotel quality, it is quite uninformative about the match value with respect to private taste. While business travellers may be exceptionally concerned about reception opening hours and good Wi-Fi, leisure travellers may have a higher valuation for available sports equipment and wellness services.

For the following example, consider a hotel that has certified a certain number of stars. Guests who consider to book online can infer from the number of stars alone only how many total points the hotel achieved in the grading system. The provided features remain unknown to the guests until they arrive at the hotel and

[13]Sometimes there are intermediate grades denoted with the label "superior", in addition to the number of stars, for details on the national certification gradings see https://www.hotelstars.eu/

can inspect it.

Potential guests g are—with equal probability—either of type *business* ($g = B$), which only care about features in the categories *reception and services*, or of type *leisure* ($g = L$), which only care about features in the categories *gastronomy and leisure*. Moreover, any type $g \in \{B, L\}$ has private preferences over the specific features within her relevant categories, which can be represented by a location x_g on the circumference of a circle with perimeter one (Salop's circle). Assume for each type $g \in \{B, L\}$ that the location is uniformly distributed on the circumference. Guests are risk neutral and the realization of (g, x_g) is their private information.

The hotel h provides—with equal probability—either mainly features for the business type (focus $h = B$), or mainly for the leisure type (focus $h = L$). The exact features that the hotel offers for each, business and leisure type, may be represented by two locations (y_B, y_L) on the circumferences of two distinct circles, each with perimeter one. Assume, again, that y_B and y_L are each ex ante uniformly distributed on their circumference. Marginal costs for additional guests are zero. The values of (h, y_B, y_L) are private information of the hotel, but become observable after the guest arrives at the hotel.

Let the utility of a guest (g, x_g) who pays price p to stay in a hotel (h, y_B, y_L) be given by

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

where $d(x_g, y_g) \in [0, 0.5]$ describes the distance of x_g and y_g on the circumference of the circle. In other words, the guest 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 preferred features *within* the relevant categories. As neither the guest nor the hotel have ex ante any information on the match value, the value 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 $g = h$.

Since guests' valuation is ex ante uniform on $[0, 1]$ they have an expected value of 0.5. The hotel can set $p = 0.5$ without refund option, and all guests will accept the offer, leading to an efficient outcome where all rents are realized by the hotel.

Consider now the role of a regulator who is solely interested in consumer surplus and has the power to precisely regulate the information the hotel has to provide online. If the regulator forces the hotel to disclose its focus $h \in \{B, L\}$, as well as the available features only in its focal categories—thus the exact position of y_h— the guest receives a signal exactly as defined in (1): Guest types $g = h$ with a valuation above 0.5 learn their exact valuation as they learn the hotel's position y_g, while guest types $g \neq h$ don't learn the relevant position y_g and remain pooled with valuations in $[0, 0.5]$. As we have seen in Proposition 1, such a signal is guest-optimal, and implements the upper bound of the guests' utility.

The intuition of the example is simple: The regulator should exclusively allow information that is relevant to high types. High types receive high information rents, while the low types, who are relatively uninformed, induce the seller to set low prices.

5 The General Case

The main result of the paper is that the buyer can achieve his first-best for any arbitrary prior distribution:

Theorem 1. *Let π^M be the standard static monopoly profit the seller can achieve if the buyer privately learns his valuation before contracting. Then there exists an information signal such that for the seller's optimal menu of option contracts*

- *trade is efficient, and*
- *the seller receives π^M.*

Such a signal is buyer-optimal, since it maximizes aggregate surplus and leaves the seller with her lower bound of utility π^M.

The result follows immediately from the more general Theorem 2. I will, however, provide the intuition how to construct such a signal structure.

If a menu of contracts induces trade for all types θ, then each type will trade at the lowest available total payment $a + p$. This means that any signal which induces full trade and a seller profit of π^M must necessarily sell to all buyers at a uniform price of $a + p = \pi^M$.

We start by defining y as the type such that

$$\mathbb{E}[\theta | \theta \leq y] = \pi^M.$$

Consider first a signal structure similar that defined in Equation (1) of Section 4, where types above y learn their valuation, while types $\theta \leq y$ are pooled in one signal realization. By construction, the pooled buyers are ex ante willing to pay at most π^M for the good. Therefore, a uniform ex ante price of π^M would attract all buyers, and the seller would make monopoly profit π^M.

However, for many prior distributions F such a uniform price with full participation is suboptimal for the seller, given the signal. The seller may make higher profit, if she chooses to exclude types below some threshold $\hat{\theta} < y$ from trade. She could achieve this for example with a menu offer of $\mathcal{M} = \{(a, \hat{\theta})\}$, where a is chosen such that types $\theta \leq y$ in the pooling area receive a utility of zero in expectation. Such a contract comes at the cost of losing the participation of types $\theta < \hat{\theta}$, but at the gain of higher prices and thus more revenue from all types $\theta > y$.

One can modify the signal structure such that such contracts are never optimal. The idea is to give buyers more information by maintaining the property that $\mathbb{E}[\theta|\tau] = \pi^M$ for all $\theta \leq y$. Figure 1 illustrates how to achieve this goal.

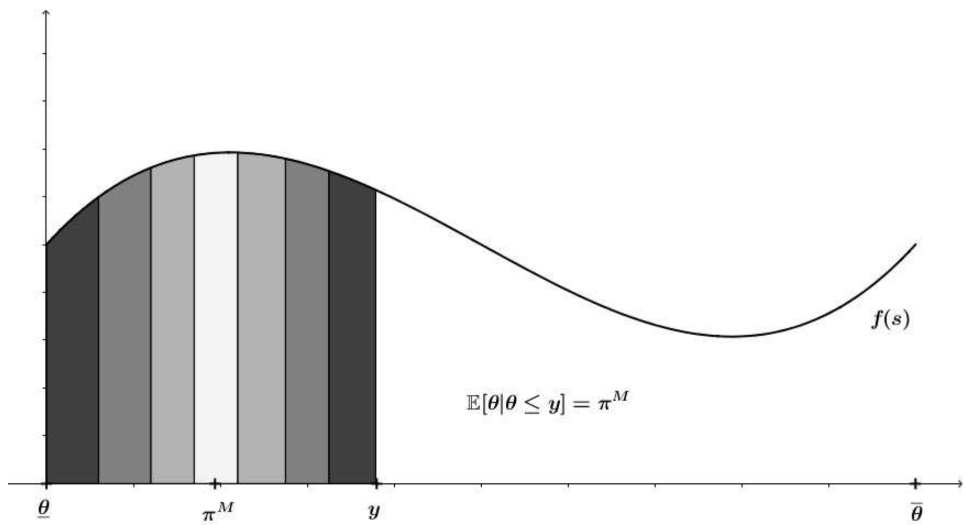

Figure 1: Types in intervals with the same color are pooled into the same signal.

Types above y fully learn their valuation, while types below y learn that their type is in a certain pooling region, represented by the shade of gray, assigned to their type in Figure 1. The shaded areas are constructed in such a way that for any shade signal τ

$$\mathbb{E}[\theta|\tau] = \pi^M.$$

Further, if τ_1 is darker than τ_0, then $F(\cdot|\tau_1)$ is a mean preserving spread of $F(\cdot|\tau_0)$. Now, we let the area of each shade signal shrink, while we let the number of different shades goes to infinity. In the limit, we obtain a continuum of shades, where each signal τ only pools two types $\{\theta_\tau^L, \theta_\tau^H\}$ with $\theta_\tau^L < \pi^M < \theta_\tau^H$, and $\mathbb{E}[\theta|\tau] = \pi^M$.

If the seller now aims to sell to any type $\hat{\theta} < \pi^M$, then—by ex ante individual rationality—she has to offer a contract that charges at most $a + p = \mathbb{E}[\theta|\tau(\theta) = \tau(\tilde{\theta})] = \pi^M$. Such a contract would attract all types, and generate a profit of π^M to the seller. It turns out, this is the best the seller can do: suppose that the seller aims for higher prices at the cost of participation. If it were optimal for the seller to use a menu for which the lowest type that buys satisfies $\hat{\theta} > \pi^M$, then one can show that the best the seller could do is to offer a contract $(0, \hat{\theta})$. Such a contract is equivalent to an ex post take-it-or-leave-it offer with price $\hat{\theta}$, which certainly cannot generate more profit to the seller than the optimal static monopoly price with fully informed buyers.

6 The Limits of Surplus Distribution

In the previous section, I have analyzed a signal which maximizes buyer surplus. In this section, I characterize which combinations of buyer surplus and seller surplus are feasible in a sequential screening framework. Similarly to Bergemann et al. (2015), let us first characterize the natural constraints to this problem graphically.

First of all, by buyer's individual rationality, expected buyer surplus will never be negative. Second, as argued in the previous section, seller surplus can never fall below the static monopoly profit, since the seller can always use a static mechanism after the buyer learned his true type. 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, as constructed in Section 5. Point B corresponds to the case, where the seller 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 $\mathbb{E}[\theta]$.

We will see that *any* arbitrary point C in the triangle can be implemented as the solution to the seller's problem for an appropriate signal distribution.

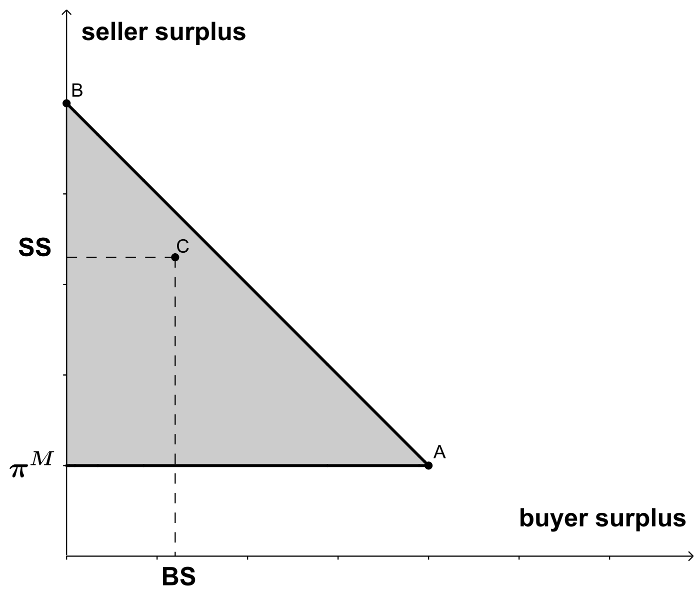

Figure 2: All potential pairs of surplus division

Theorem 2. *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*
- $u_S + u_B \leq \mathbb{E}[\theta]$,

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

A full proof can be found in the appendix. I will sketch the main steps here. Take an arbitrary surplus pair (BS, SS) which satisfies the above constraints. We will construct a corresponding signal, such that, indeed, buyer surplus and seller surplus are given by $(u_B, u_S) = (BS, SS)$.

Call $AS = BS + SS$ the aggregate surplus we want to construct. Define the threshold x by

$$AS = \int_x^{\overline{\theta}} f(\theta)\theta d\theta = (1 - F(x))\mathbb{E}[\theta|\theta \in [x, \overline{\theta}]].$$

Note that the gains from trade $u_S + u_B$ are indeed given by AS, if we can construct a signal for which exactly all types above x trade.

Next, we define the threshold $y \leq \overline{\theta}$ by

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

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

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

We will use a similar construction as in Section 5 to build a signal for which the seller chooses to sell ex ante to all types $\theta \geq x$ at a uniform price of $\overline{a}$. In this case, seller surplus u_S is indeed given by SS, and buyer surplus is $u_B = AS - SS = BS$.

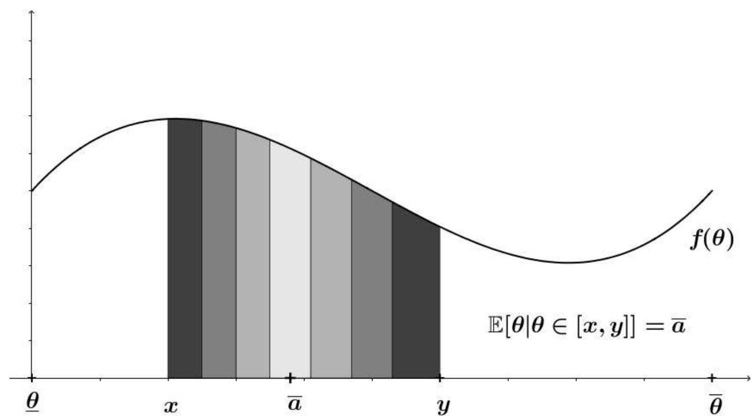

Figure 3: The signal to induce $(u_B, u_S) = (BS, SS)$

Types below x and above y fully learn their valuation. Types in the interval $[x, y]$ again learn their pooling region that is assigned to their type, illustrated by the corresponding shade of grey. The shaded areas are constructed in such a way, that for any shade signal τ we have

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

If we let the number of different shades go to infinity, we obtain a continuum of shades. In the limit, the signal can be represented by

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

While types $\theta \notin [x, y]$ learn their valuation, the signal $\tau(\theta)$ for each $\theta \in [x, y]$ corresponds to exactly two types $\{\theta_\tau^L, \theta_\tau^H\}$, which satisfy $\theta_\tau^L \leq \overline{a} \leq \theta_\tau^H$, and $\mathbb{E}[\theta|\tau] = \overline{a}$.

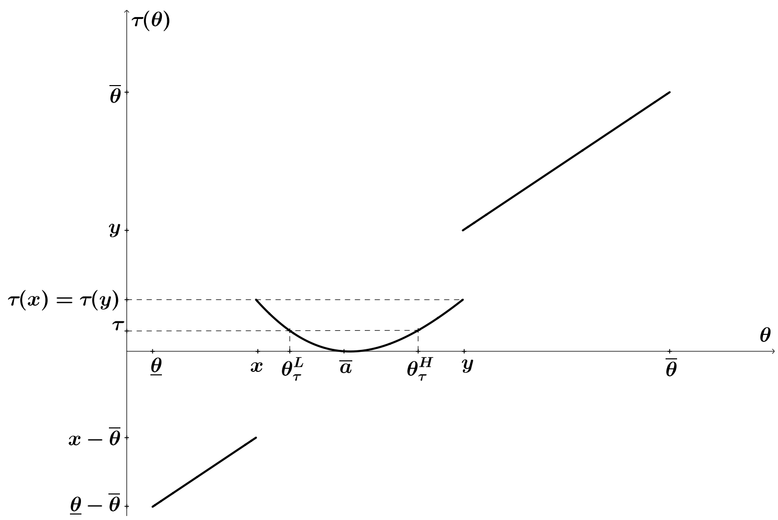

Figure 4: $\tau(\theta)$ for the distribution in Figure 3 (stylized).

Now, consider an optimal menu of contracts, the seller will offer to the buyer. Let $\hat{\theta}$ be the lowest type which buys given this menu. If $\hat{\theta} \notin [x, y]$, then the buyer learns $\hat{\theta}$ with certainty under τ, and the seller can charge at most $a+p = \hat{\theta}$ from this type. Since this contract is available to all types, expected profits cannot exceed

$$(1 - F(\hat{\theta}))\hat{\theta} \leq \max_p\{(1 - F(p))p\} = \pi^M \leq SS.$$

The seller can do (weakly) better if she decides to sell ex ante to all types in $[x, y]$. Since for all these types we have $\mathbb{E}[\theta|\tau] = \overline{a}$, a contract $(a, p) = (\overline{a}, 0)$ attracts exactly all types $\theta \geq x$, and seller's profit is $(1 - F(x))\overline{a} = SS$. One can show that this contract is indeed optimal from the seller's perspective.

The boundaries x and y define three partitions of types. In the optimal contract, types $\theta < x$ don't trade and therefore induce an efficiency loss. Types $\theta \in [x, y]$ trade, but make no surplus in expectation. Buyers $\theta > y$ receive an information rent and extract the entire buyer surplus. Intuitively, by moving the boundaries x and y one can realize any buyer and seller surplus that satisfies the natural constraints in Theorem 2.

7 Conclusion

This paper emphasizes the role of private information in sequential screening. It shows that there are almost no restrictions to the division of buyer and seller surplus, that can arise in sequential screening for different ex-ante information.

The buyer-optimal signal keeps the buyer to some extent uninformed about his valuation at the time of contracting. Thus, the ability of a monopolist to screen buyers sequentially may not necessarily harm consumers, but lead to lower prices, and increase efficiency.

It is worth noting that the European Union has taken a clear stand on the issue of consumer rights in online markets. By Directive 2011/83/EU, any consumer is granted 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 monopolist to screen ex ante, granting the consumers the same information rent, as under full information.[14,15] This paper shows, that regulators, who care about consumer utility, can improve on this regulation, if they have sufficent control over information provided to individual buyers. The regulation of information may be more flexible and powerful than the regulation of contracts, and therefore deserves further study.

[14] Krähmer and Strausz (2015) already find that the welfare effect of such a policy is ambiguous, and depends on consumers' ex ante private information.

[15] The buyer can be forced to bear the shipping cost of returning the good, so there may be scope for ex-ante screening to some very limited extent.

8 Appendix

Proof of Proposition 1. By the revelation principle for dynamic games (e.g., Myerson (1986)), we can restrict attention to direct, incentive compatible mechanisms:

The buyer reports his private information sequentially. After learning τ, he reports its realization to the seller. If $\tau \in (0.5, 1]$ then $\tau = \theta$, thus a truthful report of τ reveals θ already. If the buyer reports $\tau = 0$ then the seller asks for a report of $\theta \in [0, 0.5]$ after the buyer observes its realization.[16]

A direct mechanism specifies the trading rules as a function of the buyer's reports. Formally, the allocation rule

$$q : (\{0\} \times [0, 0.5]) \cup (0.5, 1] \to [0, 1],$$

assigns to each complete report a probability of receiving the good. The transfer rule

$$t : (\{0\} \times [0, 0.5]) \cup (0.5, 1] \to \mathbb{R},$$

assigns to each complete report a monetary transfer from the agent to the principal.

Note that since τ defines a partition on θ, each feasible report corresponds exactly to one claim of being some type $\theta \in [0, 1]$. Identifying the report space with the type space, let

$$u(\hat{\theta}|\theta) = \theta q(\hat{\theta}) - t(\hat{\theta})$$

be the utility of buyer of type θ reporting as if being type $\hat{\theta}$. Let us further simplify notation by $u(\theta) := u(\theta|\theta)$. The incentive constraints, which guarantee truthful reporting, read

[16] In Myerson (1986), the agent's report space is the entire support of his private information at each stage. That is, if the agent lies about the value of τ, he may still report θ truthfully and inconsistent with the report of τ. Since in our case τ defines a partition on all types θ, however, the seller can immediately detect and punish any untruthful report (τ, θ) with $\tau(\theta) \neq \theta$, such that the buyer would never choose such a report. It is therefore without loss of generality to consider only direct mechanisms, which restrict the reports of θ to values that are admissible for the reported τ.

$$\forall \theta \in (0.5,1], \hat{\theta} \in [0,1] \qquad u(\theta) \geq u(\hat{\theta}|\theta) \qquad \text{(IC } \tau \neq 0\text{)}$$

$$\forall \hat{\theta} \in [0.5,1] \qquad \mathbb{E}[u(\theta)|\tau = 0] \geq \mathbb{E}[u(\hat{\theta}|\theta)|\tau = 0] \qquad \text{(IC } \tau = 0\text{)}$$

$$\forall \theta \in [0,0.5], \hat{\theta} \in [0,0.5] \qquad u(\theta) \geq u(\hat{\theta}|\theta). \qquad \text{(IC } \theta\text{)}$$

First period individual rationality reads

$$\forall \theta \in [0.5,1] \qquad u(\theta) \geq 0 \qquad \text{(IR } \tau \neq 0\text{)}$$

$$\mathbb{E}[u(\theta)|\tau = 0] \geq 0 \qquad \text{(IR } \tau = 0\text{)}$$

Since the seller's utility equals social surplus minus buyer's utility, her program is

$$\begin{aligned} \mathcal{P}: \qquad & \max_{(q,t)} \int_0^1 (\theta q(\theta) - u(\theta))d\theta \\ s.t. \qquad & \text{(IC } \tau \neq 0\text{),(IC } \tau = 0\text{),(IC } \theta\text{),(IR } \tau \neq 0\text{),(IR } \tau = 0\text{).} \end{aligned}$$

We will derive the optimum for a so called "relaxed" problem $\mathcal{P}'$ with less constraints, and verify ex post that the remaining constraints are satisfied for the derived solution. The constraint (IC $\tau \neq 0$) directly implies the weaker condition

$$\forall \theta \in (0.5,1], \hat{\theta} \in [0.5,1] \qquad u(\theta) \geq u(\hat{\theta}|\theta). \qquad \text{(IC' } \tau \neq 0\text{)}$$

We now define

$$\begin{aligned} \mathcal{P}': \qquad & \max_{(q,t)} \int_0^1 (\theta q(\theta) - u(\theta))d\theta \\ s.t. \qquad & \text{(IC' } \tau \neq 0\text{),(IC } \theta\text{),(IR } \tau = 0\text{).} \end{aligned}$$

The solution to program $\mathcal{P}'$ must implement weakly higher seller surplus than program $\mathcal{P}$, as it faces less constraints.

By Revenue Equivalence (e.g., Myerson (1981)), (IC θ) is equivalent to

1. $q(\theta)$ is increasing on $[0, 0.5]$, and

2. $u(\theta) = u(0) + \int_0^\theta q(s)ds$ for all $\theta \in [0, 0.5]$.

From 2. it follows that

$$u(\theta) = u(0.5) - \int_\theta^{0.5} q(s)ds.$$

Further, in any optimal solution, (IR $\tau = 0$) must bind, because otherwise the seller could uniformly raise the transfer for all types. Using integration by parts we obtain

$$\begin{aligned}
0 &= \int_0^{0.5} u(\theta)d\theta \\
&= \int_0^{0.5} \left(u(0.5) - \int_\theta^{0.5} q(s)ds \right) d\theta \\
&= 0.5u(0.5) - \left[\theta \int_\theta^{0.5} q(s)ds \right]_0^{0.5} + \int_0^{0.5} -\theta q(\theta)d\theta \\
&= 0.5u(0.5) - \int_0^{0.5} \theta q(\theta)d\theta,
\end{aligned}$$

or equivalently

$$0.5u(0.5) = \int_0^{0.5} \theta q(\theta)d\theta. \tag{2}$$

Again by Revenue Equivalence, (IC' $\tau \neq 0$) implies that on any closed interval $[\tilde{\theta}, 1] \subset (0.5, 1]$, the allocation $q(\theta)$ is weakly increasing, and further for any $\theta \in [\tilde{\theta}, 1]$

$$u(\theta) = u(\tilde{\theta}) + \int_{\tilde{\theta}}^{\theta} q(s)ds.$$

Therefore, $u(\theta)$ is continuous on $(0.5, 1]$, and, because $q(\theta)$ is weakly positive and bounded, $\lim_{\theta \searrow 0.5} u(\theta)$ exists. Since by (IC' $\tau \neq 0$) for all $\theta \in (0.5, 1]$

$$u(\theta) \geq u(0.5, \theta) = (\theta - 0.5)q(0.5) + u(0.5) \geq u(0.5),$$

we have necessarily

$$\lim_{\theta \searrow 0.5} u(\theta) \geq u(0.5).$$

Moreover, if we had $\lim_{\theta \searrow 0.5} u(\theta) = u(0.5) + \varepsilon$, for some $\varepsilon > 0$, the seller could increase all transfers of types $\theta \in (0.5, 1]$ uniformly by ε and still satisfy all constraints of $\mathcal{P}'$. A mechanism with $\lim_{\theta \searrow 0.5} u(\theta) > u(0.5)$ can therefore not be optimal. It follows that any solution to $\mathcal{P}'$ must satisfy

$$u(\theta) = u(0.5) + \int_{0.5}^{\theta} q(s)ds$$

for all $\theta \in [0.5, 1]$.

For the seller's objective function in $\mathcal{P}'$ we obtain

$$\begin{aligned}
\int_0^1 (\theta q(\theta) - u(\theta))d\theta &= \int_0^1 \theta q(\theta)d\theta - \underbrace{\int_0^{0.5} u(\theta)d\theta}_{=0} - \int_{0.5}^1 u(\theta)d\theta \\
&= \int_0^1 \theta q(\theta)d\theta - \int_{0.5}^1 \left(u(0.5) + \int_{0.5}^{\theta} q(s)ds \right) d\theta \\
&= \int_0^1 \theta q(\theta)d\theta - 0.5u(0.5) - \int_{0.5}^1 \int_{0.5}^{\theta} q(s)dsd\theta \\
&\overset{(2)}{=} \int_0^1 \theta q(\theta)d\theta - \int_0^{0.5} \theta q(\theta)d\theta - \int_{0.5}^1 \int_{0.5}^{\theta} q(s)dsd\theta \\
&= \int_{0.5}^1 \theta q(\theta)d\theta - \left[\theta \int_{0.5}^{\theta} q(s)ds \right]_{0.5}^1 + \int_{0.5}^1 \theta q(\theta)d\theta \\
&= \int_{0.5}^1 \theta q(\theta)d\theta - \int_{0.5}^1 q(s)ds + \int_{0.5}^1 \theta q(\theta)d\theta \\
&= \int_{0.5}^1 (2\theta - 1)q(\theta)d\theta.
\end{aligned}$$

Note that the seller's utility is independent of the allocation for types $\theta \leq 0.5$. Indeed, any attempt to increase surplus from these types equally increases the information rent the seller has to provide to types $\theta \in [0.5, 1]$.

Since $(2\theta - 1) > 0$ for $\theta > 0.5$, the seller maximizes her utility by setting $q(\theta) = 1$ for any $\theta > 0.5$. The seller's maximal utility under $\mathcal{P}'$ therefore is

$$\int_{0.5}^1 (2\theta - 1)d\theta = [\theta^2 - \theta]_{0.5}^1 = 0.25.$$

If the seller chooses to set $q(\theta) = 1$ for all $\theta \leq 0.5$ then the direct mechanism takes the form

$$q(\theta) \equiv 1,$$

and

$$t(\theta) \equiv 0.25,$$

which corresponds exactly to the offer to sell the product ex ante at a uniform price of 0.25. Since all buyer types obtain the same offer in this contract, it satisfies all incentive constraints of $\mathcal{P}$. Moreover, the contract yields positive profit for all $\theta \geq 0$, therefore it satisfies the constraint (IR $\tau \neq 0$) of program $\mathcal{P}$ as well. $\square$

Proof of Theorem 2. Take some arbitrary $SS \geq \pi^M$ and $BS \geq 0$, with $BS + SS \leq \mathbb{E}[\theta]$. We need to construct a signal such that the seller's optimal mechanism induces seller utility $u_S = SS$ and buyer utility $u_B = BS$.

Constructing the signal

Define x implicitly by

$$BS + SS = \int_x^{\overline{\theta}} \theta dF(\theta) = (1 - F(x))\mathbb{E}[\theta|\theta \in [x, \overline{\theta}]]. \tag{3}$$

Since f has full support, the right hand side is strictly decreasing in x for $x \in [\underline{\theta}, \overline{\theta}]$, with $\int_{\underline{\theta}}^{\overline{\theta}} \theta dF(\theta) = \mathbb{E}[\theta]$, and $\int_{\overline{\theta}}^{\overline{\theta}} \theta dF(\theta) = 0$. Since

$$0 < BS + SS \leq \mathbb{E}[\theta],$$

there is exactly one $x \in [\underline{\theta}, \overline{\theta}]$, for which (3) is satisfied.[17]

Define now y implicitly by

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

The right hand side is strictly increasing in y and since

$$(1-F(x))\mathbb{E}[\theta|\theta \in [x, x]] = (1-F(x))x \leq \pi^M \leq SS \leq BS+SS = (1-F(x))\mathbb{E}[\theta|\theta \in [x, \overline{\theta}]],$$

there is exactly one $y \in [x, \overline{\theta}]$, which satisfies (4). Further, we call

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

[17] 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 loose uniqueness of x and y. None of the results or intuitions hinge on these assumptions.

Finally, we define the following signal structure:

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

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 $\tau(\theta)$ on $[\overline{a}, y]$. This means that the distribution of τ is given by

$$G(\tau) = F(\theta^H(\tau)) - F(\theta^L(\tau)).$$

It follows[18] that for any $\tau \in (0, \tau(x)]$

[18] We denote by $\mathbb{P}(A|\tau)$ the regular conditional probability for A given τ. This notion extends the concept of conditional probabilities to the case where one conditions on events of probability zero. The regular conditional probability is defined by the condition that for any measurable sets A, B the equality $\mathbb{P}(\theta \in A, \tau \in B) = \int_B \mathbb{P}(A|\tau)dG(\tau)$ holds. It is unique almost surely. Since we are interested in expectations only, this restriction is innocuous. For formal details see for example §7 on regular conditional distributions in Shiryaev (1996).

$$
\begin{aligned}
\mathbb{P}(\theta^H_\tau|\tau) &= \mathbb{P}(\theta > \overline{a}|\tau) \\
&= \lim_{\varepsilon\to 0} \mathbb{P}(\theta > \overline{a}|\tau(\theta) \in [\tau, \tau+\varepsilon]) \\
&= \lim_{\varepsilon\to 0} \frac{F(\theta^H(\tau+\varepsilon)) - F(\theta^H(\tau))}{F(\theta^H(\tau+\varepsilon)) - F(\theta^H(\tau)) + F(\theta^L(\tau)) - F(\theta^L(\tau+\varepsilon))} \\
&= \frac{f(\theta^H_\tau)\theta^{H\prime}(\tau)}{f(\theta^H_\tau)\theta^{H\prime}(\tau) - f(\theta^L_\tau)\theta^{L\prime}(\tau)} \\
&= \frac{f(\theta^H_\tau)/\tau'(\theta^H_\tau)}{f(\theta^H_\tau)/\tau'(\theta^H_\tau) - f(\theta^L_\tau)/\tau'(\theta^L_\tau)} \\
&= \frac{1/(\theta^H_\tau - \overline{a})}{1/(\theta^H_\tau - \overline{a}) + 1/(\overline{a} - \theta^L_\tau)} \\
&= \frac{\overline{a} - \theta^L_\tau}{\theta^H_\tau - \theta^L_\tau}.
\end{aligned}
$$

Similarly, we have

$$\mathbb{P}(\theta^L_\tau|\tau) = \frac{\theta^H_\tau - \overline{a}}{\theta^H_\tau - \theta^L_\tau}.$$

It follows that

$$\mathbb{E}[\theta|\tau] = \frac{\theta^H_\tau - \overline{a}}{\theta^H_\tau - \theta^L_\tau}\theta^L_\tau + \frac{\overline{a} - \theta^L_\tau}{\theta^H_\tau - \theta^L_\tau}\theta^H_\tau = \overline{a}. \tag{5}$$

This means that 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 $F(\cdot|\tau_1)$.[19]

The menu

We turn to the seller's decision problem to choose an optimal menu of option contracts, given 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 contract $\mathcal{M}$ we have

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

[19]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.

and

$$u_B = \int_x^{\overline{\theta}} \theta dF(\theta) - u_S = (BS + SS) - SS = BS.$$

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

Let $\tilde{\mathcal{M}} = \{(a_i, p_i)\}_{i \in I}$ be an arbitrary menu of option contracts. We need to show that it does not generate higher seller utility than SS.

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

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

Further, 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 receives positive utility from contract (a, p). Consequently no type $\theta > \hat{\theta}$ will accept a contract at higher total cost than $a + p$. Since $\hat{\theta}$ is by assumption the lowest type that buys, we can conclude that

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

Case 2: $\hat{\theta} \in [x, \overline{a}]$

Then $\hat{\theta}$ is the low type for the respective signal realization, ie. $\hat{\theta} = \theta^L_{\tau(\hat{\theta})} < \theta^H_{\tau(\hat{\theta})}$. Thus, since type $\theta^L_{\tau(\hat{\theta})}$ purchases the good under (a, p), so will type $\theta^H_{\tau(\hat{\theta})}$. By 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 as well 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

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

Case 3: $\hat{\theta} \in [\overline{a}, y]$

Then $\hat{\theta}$ is the high type for the respective signal realization, ie. $\hat{\theta} = \theta^H_{\tau(\hat{\theta})}$. 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 can't bind for any higher $\tau \in [\tau(\hat{\theta}), \tau(y)]$.

Lemma 1. *If for signal types $0 \leq \tau_1 < \tau_2 \leq \tau(y)$ 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 we have

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

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 $\alpha_2 < 1$, we have

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

□

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

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

from contract (a, p). The contract thus generates 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

$$u_S = (1 - G(\tau(\hat{\theta})))a + (1 - F(\hat{\theta}))p \quad = (1 - F(\hat{\theta}) + F(\theta^L_{\tau(\hat{\theta})}) - F(x))a + (1 - F(\hat{\theta}))p.$$

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

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

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

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

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

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

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

References

Balestrieri, F., Izmalkov, S., and Leao, J. (2015). The market for surprises: selling substitute goods through lotteries.

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.

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

Grunert, K. G., Fernández-Celemín, L., Wills, J. M., genannt Bonsmann, S. S., and Nureeva, L. (2010). Use and understanding of nutrition information on food labels in six european countries. *Journal of Public Health*, 18(3):261–277.

Hawley, K. L., Roberto, C. A., Bragg, M. A., Liu, P. J., Schwartz, M. B., and Brownell, K. D. (2013). The science on front-of-package food labels. *Public health nutrition*, 16(3):430–439.

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

Hoffmann, F., Inderst, R., and Turlo, S. (2017). Regulating cancellation rights with consumer experimentation.

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. (2016). Discriminatory information disclosure. *Discussion paper*.

Myerson, R. B. (1981). Optimal auction design. *Mathematics of operations research*, 6(1):58–73.

Myerson, R. B. (1986). Multistage games with communication. *Econometrica*, 54(2):323–358.

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.

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

Shapiro, D. and Shi, X. (2008). Market segmentation: The role of opaque travel agencies. *Journal of Economics & Management Strategy*, 17(4):803–837.

Shiryaev, A. N. (1996). *Probability, volume 95 of Graduate texts in mathematics*. Springer-Verlag, New York.

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

Viscusi, W. K. (1978). A note on "lemons" markets with quality certification. *The Bell Journal of Economics*, 9(1):277–279.